Title

What conditions control the incidence of convergent evolution? A study of fossorial reptiles.

ISBN

9798189271397

Author

Krone, Isaac Weston

Publication Date

2026-05-01

Peer reviewed|Thesis/dissertation

What conditions control the incidence of convergent evolution? A study of fossorial reptiles.

By
Isaac Krone

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jimmy A. McGuire, Chair

Professor Ian J. Wang

Professor Charles R. Marshall

Summer 2026

What conditions control the incidence of convergent evolution? A study of fossorial reptiles.

By

Isaac Krone

Abstract

What conditions control the incidence of convergent evolution? A study of fossorial reptiles.

by

Isaac Krone

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Professor Jimmy A McGuire, Chair

In this thesis I use limb-reduced reptiles as a focal group to explore several largely ignored areas of inquiry in paleontology, herpetology, comparative anatomy, and macroecology. The convergence on limb-reduced or limbless body plans adapted for a burrowing lifestyle is a major theme in the evolution of squamate reptiles, but its implications for the anatomy, evolution, biogeography, and long-term success of clades are far from fully understood. Here, I develop novel techniques for understanding fossilization potential of extant species (Chapter 1), and understanding the evolution of surface patterns (Chapter 3), and apply these to limbless squamates to understand what information about their evolution might actually be available via the fossil record, and how scalation patterns may reflect their evolution and ecology. I address a major shortfall in the current understanding of the diversity of limbless squamates by constructing the first densely-sampled phylogeny of dibamid lizards (Chapter 2), and analyzing their diversity and biogeographic history. Finally, I use a microcosm of evolution, the island of Madagascar, as a test case to develop a comprehensive understanding of niche evolution and the role of specialization for fossoriality in shaping the biogeographic and diversification histories of fossorial reptiles (Chapter 4), plotting a path towards a comprehensive, global understanding of where, when, how, and why this body plan evolves, and what is possible for lineages after evolving this specialization.

Dedication

to the unnoticed, unknown, and underfoot.

Table of Contents

Table of Figures

Table of Figures

Index of Tables

Index of Tables

Acknowledgements

I have made it this far only with a great deal of support and a healthy helping of luck. I owe my success in science first and foremost to my parents, Carla and Gary, who always encouraged my interest in paleontology and biodiversity. They went to great lengths to make sure that I got the best education possible at home and throughout my K-12 schools, which, in Southeastern Wisconsin around the turn of the millennium, were an incredible achievement in public instruction. My parents have supported my science ever since, as has my ever-creative sister Kenna, a biologist of the microscopic, pathogenic, and generally bothersome realms.

Drs. Thomas Carr and Meaghan Seitz were the first scientists to train and support me, and were very influential in setting me up for success in vertebrate paleontology. They gave me the chance to work closely in the lab and the field with a wealth of fossil material and a diverse group of colleagues, and those early experiences in research–especially field work–were the spark of my scientific career. Discussions in the field on extinction, fossilization and stratigraphy with Dr. Carr were a key influence in my development of the first chapter of this thesis.

I owe thanks to many, many people at the University of Chicago. Chief among them are Dr. Eric Larsen, who started my more-than-decade-long love affair with the Sonoran desert, which turned my interests more and more towards the lizard world, and Dr. Ken Angielczyk, who gave me a great dataset and let me run with it for a senior thesis project. That project is just as indebted to Dr. Christian Kammerer, who had collected most of those data in the first place, and gave me just as much good advice on what to do with it. Dr. Mike Coates always met me with good humor and trained me very well in comparative anatomy and the wonders of cartilaginous fish, even if my work in his lab fizzled. After graduation, Dr. Mark Westneat kept my career going with a brief stint in fish biomechanics and morphometrics, keeping me in a good place to learn and prepare for the next steps of my career.

My long obsession with the Sonoran desert would not have gone on this long without the training and camaraderie of the now departed Dr. Emma Grieg, who taught me how to be a field biologist and a mentor to new ones. The world is much, much poorer without her in it.

The greatest source of blame for my herpetological and neontological turn is Dr. David Blackburn. It was through working with him that I learned just how diverse reptiles and amphibians

are, and realized how poorly they are known. Dave's influence is also manifested in my first chapter, particularly in its focus on the unique challenges in amphibian paloentology.

But after those dalliances in dinosaurs, sharks, frogs, synapsids, fish, and birds, I knew what I liked: lizards. I also knew that I wanted to describe new species, and to learn to produce the phylogenies that I relied on to compare them. I could not have found a better person to work with than Dr. Jim McGuire. Jim's extensive work underpins all of my taxonomic and phylogenetic projects, only a portion of which are reflected in this document, and his high standards and genuine curiosity for the methodologies he employs have served as an example for me in the back half of my backwards PhD process. Moreover, my non-phylogenetics chapters would never have existed if not for that curiosity; whenever I came up with a new idea that was more interesting than what I was working on, Jim always encouraged me to pursue it, even if it wasn't in his wheelhouse; more than anything else, it's Jim's willingness to see what I come up with that has allowed me to produce such a diverse dissertation. Jim's offer to take me into the field in 2023 began a new chapter in my scientific career, full of more biodiversity than I'll probably ever have time to deal with.

In joining Jim's lab, I came into contact with the most influential presence of all; the Museum of Vertebrate Zoology and Integrative Biology communities. The MVZ has been my intellectual home for a long time now, and I suspect it always will be; my interests and perspectives have been profoundly shaped by spending eight years here. The intellectual and financial support of the MVZ have been indispensable to my work. That kind of support came from the broader IB community as well; not only am I leaving with a wealth of connections to an increasingly far-flung and influential cohort of scientists, but I'm also leaving in a secure financial position. The vast majority of my colleagues were willing and determined to challenge the University of California and fight in the 2021 strike, and without the victories gained in that strike, I'd have had a much tougher time living in Berkeley. The IB graduate student community demonstrated so many things that scientists can and should be; collaborative, caring, and deeply committed.

I will thank just a few of the MVZ members who have been so important to my time here. First, Dr. Carol Spencer, without whom I would have probably made permanent enemies of US customs and would understand very little of what has to happen to make a museum work, and Lydia Smith, who has patiently guided me through far more wet lab work than I was ever prepared for, and without whom I would have probably written off the whole endeavor. Thanks to Drs. Rauri Bowie and Alan Shabel,

who have greatly influenced my teaching and supported my desire to continue that teaching. In the McGuire Lab, thanks to Drs. Ammon Corl, Simon Scarpetta, and Umilaela Arifin, who have been incredible models of what you can do after a PhD, and have always been sources of good advice and encouragement, and to Dr. Jefferey Frederick, Dr. Ben Karin, and soon-to-be-Dr. Sina Amini, for sharing so much knowledge and accumulating laundry lists of accomplishments that encourage me to strive to keep up with them. Beyond the McGuire Lab, I will always recommend Majo Navarette, Kannon Pearson, and John Myles as some of the best possible people to go into the field with.

I have had the pleasure of working closely in the lab and in the field with a number of undergraduates, all of whom have taught me a lot about biology and a lot about mentorship. In particular, I owe a lot of any future success I have in mentoring to my experience working with Meg Scudder, Prakrit Jain, Edward Pandji, Natalie Ng, Kinsi Petersen, and Elisha Pador.

My academic committees have been universally helpful and supportive of whatever strange directions I pivot to. Dr. Marvalee Wake has always been a superbly insightful sounding board for new ideas and a source of great technical and taxonomic expertise. Dr. Rauri Bowie, again, introduced me to a bevy of influential literature and always knows a paper that I should read. Dr. Ian Wang made me think spatially and encouraged me to meticulously develop the methods in chapter 3. And Dr. Charles Marshall always challenges me to think more deeply, consider ever-broader scales, and avoid the pitfalls of a neontology-first mindset.

Finally, I have to thank the people outside of my science with whom I have shared my PhD years. My Chicago friends, near and far and many now also “Dr.”, made starting a PhD in Berkeley feel like coming home; Johnny Guy, Benno Giammarinaro, Will Craft, Hazel Court, Lawrence Wang, and Brandon Dixon made sure that my first years here were good ones. Thomas Wester, Alex Kapadia, Ted Litchfield, Johnathan Gerstell, Aidan Lilienfeld, and Avery Robinson always kept in touch with me and with reality from around the globe, and made sure I had a place to crash and friends to spend time with when I was away from home.

Last but not least, the back half of my PhD owes most of its happiest moments to soon-to-be-Dr. fern Wildtruth, who cajoles and inspires me to take action, take trips, take photos, take vows, and pay attention to plants. I would be half the person that I am now without her.

Chapter 1 - All the Earth will not remember: how geographic gaps structure the record of diversity and extinction

Introduction

A central problem in the study of macroevolutionary processes is the difficulty of comparing the modern biota to the biota of the past. Without a trustworthy record of the past, we are unable to connect patterns seen in modern organisms and ecosystems to processes that span timescales longer than the existence of the human species, let alone scientific inquiry. Our knowledge of the past is generally informed by data from fossils, from which we can learn exceedingly little about ecology, behavior, biogeography, and detailed anatomy compared to modern organisms, from which a wealth of such data can be obtained.

In order to reconcile patterns in modern and extinct biodiversity, we must somehow transform our data to make the realms comparable. Increasing the data available to us on extinct organisms is fundamentally unrealistic; the fossil record cannot provide what we need. Therefore, the reconciliation must come from decreasing the data available on modern organisms to a level resembling that of the fossil record.

Here, I develop a first-order approximation of what the fossil record of modern tetrapods might look like in several million years' time, using a purely geographic filtering process. I present here my published work (Krone et al. 2024), completed in collaboration with Dr. Ryan Yohler and Dr. Kat Magoulick. In the transition section, I use our results to analyze patterns of fossilization probability within squamates with different substrate preferences (aquatic, arboreal, fossorial, or terrestrial).

Abstract

We know the fossil record is incomplete, but just how much biodiversity does it miss? We produce the first geographically controlled estimate by comparing the geographic ranges of 34,266 modern tetrapods with a map of the world's sedimentary basins. By modeling which tetrapods live within sedimentary basins, we produce a first-order estimate of what might be found in the fossil record

of the future. In this record, nearly 30% of tetrapod species have almost no chance of fossilizing, and more stringent criteria for fossilization exclude far more diversity. This geographically structured fossil record preserves disparate patterns of taxonomic and phylogenetic diversity in different tetrapod groups and underpreserves projected extinctions. For the globally threatened amphibians, the magnitude of the extinction of all endangered species would be underestimated by 66–98% in our future record. These results raise profound questions about the structure of the fossil record. Is it capable of recording major origination and extinction events on land? Have swaths of terrestrial diversity gone unrecorded based on geography alone? There are chapters of Earth history that paleontologists can never hope to know, but what is missing, and why?

Introduction

The fossil record is a profound gift of biological knowledge, allowing us to glimpse the past and prompting us to imagine the future. Fossils are central to understanding the processes and scope of evolution and have been since humans first contemplated evolution. Yet fossils do not record all of life history; there is a contingent structure to the record itself. The primary control over which organisms are preserved in the fossil record is the accumulation of sediment and subsequent lithification to form sedimentary rocks (Nyberg and Howell 2015), because sedimentary rocks preserve and store fossilized remains. This results in vast spatial heterogeneity in fossilization potential. Although paleontologists have long recognized the preservation potential of wet and lowland environments over dry and upland ones (Newell 1959; Knoll and Niklas 1987; Holland 2022), the question of which ancient organisms evaded fossilization is fundamentally moot; the rocks do not exist.

Yet understanding this spatial structure is crucial to our ability to contextualize the dwindling biota of the modern diversity crisis within Earth history. The effects of anthropogenic habitat destruction and climate change on geology (Spalding and Hull 2021), ecology (Palkovacs et al. 2012; Stuart-Smith et al. 2021), and biodiversity (Wake 2008; Barnosky et al. 2011; Plotnick et al. 2016; Crooks et al. 2017) are an ongoing and time-sensitive area of research. Much research has focused on the possibility of a “sixth mass extinction” (Barnosky et al. 2011), comparable to the five mass extinctions recognized by paleontologists (Raup and Sepkoski 1982; Sepkoski 1984), which coincide with large shifts in climate (Song et al. 2021) and elevated extinction rates (Raup and Sepkoski 1982). Measuring extinction rates in the fossil record presents many difficulties (Foote 2000), and comparing

those rates to those observed over the short timescale of recorded history is particularly fraught (McCallum 2007; Barnosky et al. 2011; Ceballos et al. 2015; Spalding and Hull 2021). But these comparisons could be improved by incorporating a structural model of the fossil record itself, one based on the first-order control over its production: sedimentation (Holland 2016). Organisms that die outside of an area of net sediment accumulation (active sedimentary basins) cannot enter the fossil record, as there are no sedimentary rocks being formed to hold them. This is a profound structuring mechanism; many currently accumulating sediments are geologically “doomed” (Holland 2016) and will be lost to erosion quickly after they are deposited; just 16% of the Earth's terrestrial surface is currently accumulating sediments likely to persist more than 1 million years (Nyberg and Howell 2015; Fig. 1).

We produce a first-order quantification of the potential fossil record of modern tetrapods by measuring the extent to which their geographic ranges overlap with areas of active sedimentation, producing a predicted fossil geographic range (FGR) for 34,266 extant tetrapod species. In doing so, we place bounds on the scope of that record and begin to explore the implications of this geographic structure. Because this is a rough approximation of fossilization potential, we consider a range of geographically structured “fossil records” that selectively include species based on the sizes of their FGRs. Species with smaller FGRs are more likely to have their fossils destroyed by future erosion or to be overlooked by future paleontologists than species with larger FGRs. At our most permissive, we include any species with an FGR area greater than 1 km^2 in the fossil record. We then consider successively stricter minimum areas for inclusion: 10 km^2, 100 kkm^2 , 1000 km^2, 10,000 km2, and 100,000 km^2.

Using this species-level dataset of FGR areas, we examine how taxonomic differences in fossilization potential contribute to the phylogenetic fidelity of the fossil record by measuring the loss of phylogenetic diversity (PD) when comparing the extant fauna with our predicted fossil records. We find marked heterogeneity in the fossil record's ability to accurately record the evolutionary history of different tetrapod groups.

We also measure the effect of this geographic structuring on the record of extinction. We compare modern tetrapod diversity with the diversity of a depauperate tetrapod fauna resulting from a simulated extinction event that removes species listed by the International Union for Conservation of Nature (IUCN) as Endangered (EN) and Critically Endangered (CR), Extinct (EX) and Extinct in the

Wild (EW), or Data Deficient (DD). We then compare this loss in diversity to the loss in diversity that a geographically structured fossil record could preserve, demonstrating its profound inability to accurately record the rate and magnitude of such an extinction event, with a particular focus on the heavily threatened amphibians (Wake 2008; Alroy 2015).

Materials and Methods

Geographic Methods

We collected all available polygonal geographic range maps for modern tetrapods from the IUCN Red List database (for mammals [accessed 3 September 2021]) and amphibians [accessed 7 March 2022]; IUCN 2021), BirdLife's species range maps (for birds [accessed 18 November 2021]; BirdLife International 2021), and the Global Assessment of Reptile Distributions (GARD) range map database (for crocodilians, turtles, squamates, and the tuatara [accessed 23 June 2020]; Roll et al. Reference Roll, Feldman, Novosolov, Allison, Bauer, Bernard and Böhm2017). We then created a single maximally inclusive polygonal range map for each tetrapod species by using the dissolve function in ArcGIS Pro (ESRI 2022) to combine all polygons associated with a species (which originally were labeled with different IDs) into a single polygonal feature. We performed a pairwise intersect operation between these range polygons and the polygonal sedimentary basin maps from Nyberg and Howell (Nyberg and Howell 2015), including both terrestrial and shallow-marine basins, again dissolving all output layers corresponding to one species into a single polygonal feature and repairing geometry as necessary to avoid self-intersects. The resulting polygons estimate where on the Earth's surface that species has a chance to enter the fossil record. We hereafter refer to these areas as fossil geographic ranges (FGRs). The areas (km2) of these FGRs are the basis for our subsequent analyses. All geographic operations were performed in ArcGIS Pro (ESRI 2022) and resulting data were projected using the WGS 1984 Cylindrical Equal Area projection (https://pro.arcgis.com/en/pro-app/3.1/help/mapping/properties/cylindrical-equal-area.htm).

We include both terrestrial and marine sedimentary basins in our analysis in order to include marine species and to account for inexact matching between the borders of species range

polygons and sedimentary basin polygons. Our methods produce a maximally inclusive FGR given the geographic ranges and data available.

Criteria for Inclusion in the Fossil Record

For most of our tests, we consider inclusion in the fossil record as a binary trait; a species is either included in the record or absent (excluded). Because the fossil record is sampled by paleontologists, who cannot sift through of all the Earth's sedimentary layers (whether exposed or buried), species with very small FGRs are less likely to be recognized as part of the fossil record than species with large FGRs. To account for this undersampling, we analyze our data under six geographically structured scenarios, representing more or less complete sampling of the fossil record. In the first scenario, all species with an FGR of greater than 1 km^2 are included in the fossil record. This is nearly identical to including all species with an FGR greater than zero (less than 2% difference in taxonomic completeness). In the second scenario, all species with an FGR greater than 10 km^2 are included in the record. In the third, all species with an FGR greater than 100 km2 are included, and so on for minimum FGR sizes of 1000 km^2, 10,000 km^2, and 100,000 km^2. Because the inclusion scenarios are hierarchical (i.e., the species “included” in the record at the 10 km2 scenario are a subset of those included in the tree in the 1 km^2 scenario), higher minimum FGR sizes result in progressively smaller samples of total species diversity. We analyze taxonomic completion and PD loss under these six geographically structured scenarios.

Taxonomic Completion

Because our phylogenetic and geographic data sources do not share consistent taxonomies, we first used synonymies from the Integrated Taxonomic Information System (ITIS) database accessed through the R package *taxize* (Chamberlain and Szöcs 2013) to match taxa in trees to taxa in our geographic data, using only unambiguous cases of synonymy. Then, for each genus in our geographic data represented by five or more species not found in our phylogenetic datasets, we manually checked for synonyms in the IUCN Red List and BirdLife International databases. Using these techniques, we matched more than 90% of species in each geographic database to taxa in a phylogeny. Using the inclusion scenarios described above, we tabulate the number of species whose FGRs meet each

minimum FGR cutoff value. To measure genus- and family-level completeness, we simply tabulate the number of genera and families represented by species in the fossil record (Table 1).

We checked our data against the Paleobiology Database (Uhen et al. 2023) in order to find false negatives: extant taxa that have a known fossil record, but that we do not predict will enter the fossil record. We found no clear instances of false negatives. We compared our dataset against all PBDB taxonomic records for Reptilia, Aves, Amphibia, and Mammalia (accessed 8 November 2022).

Phylogenetic Signal

We performed all statistical and phylogenetic analyses in R (R v. 4.1.2; R Core Team 2021). Our code is available at https://doi.org/10.5281/zenodo.8417641. Our phylogenetic tests use subsets of 1000 trees downloaded from the full phylogeny datasets provided by VertLife (Jetz et al. 2012; Tonini et al. 2016, Jetz and Pyron 2018*a*; Upham et al. 2019). For our investigation of phylogenetic signal, we built a maximum-clade-credibility tree from each of these datasets using the *maxCladeCred()* function in the R package *phangorn* (Schliep 2011) and then used the *phylosig()* function from the *phytools* package (Revell 2024) to measure Pagel's λ for geographic range area and FGR area in these clades and selected subgroups (Table 2). Pagel's λ measures the degree to which continuous data distributed across the tips of a tree fit the expectation of evolution under Brownian motion and varies from 0 (data of sister tips cannot predict each other) to 1 (the data fit a Brownian motion model), rarely exceeding 1 (Freckleton et al. 2002).

Phylogenetic Diversity

We tested each major taxonomic group independently, trimming each initial tree set to include only those taxa which occur both in that tree and our geographic datasets, which resulted in trees with 9755 species for Squamata (100% complete), 6636 species for Amphibia (91.7% complete), 5564 species for Mammalia (91.7% complete), and 9391 species for Aves (94.0% complete). As the predicted fossil records of crocodilians and turtles are quite complete, we do not report PD (Faith 1992) results for these groups. To understand the phylogenetic signature of the potential fossil record of modern tetrapods, we compare the PD of the maximally inclusive trees described above with the PD of trees representing only those species found in a projected fossil record. Because PD for an entire tree is defined as the sum of all branch lengths, our methodology works by “pruning” branches from trees to

remove species we predict will not enter the fossil record, obtaining a subtree containing only the remaining “fossilized” species.

We first measure the total PD for a tree containing all the extant species, then prune the tree to exclude the branches of species whose FGRs are smaller than some cutoff value given by our inclusion scenarios. We then measure the PD of the pruned tree, expressed as a proportion of the total PD. We repeat this process to measure PD loss for that tree in each of the six inclusion scenarios. We then repeat this for each of the 1000 trees in the phylogeny subset to generate a distribution of 1000 PD loss values corresponding to 1000 alternate tree topologies under each inclusion scenario.

We also tracked the age of the internal nodes of the tree that are lost as taxa are removed. A preservation regime erases phylogenetic history more quickly by removing long branches from the tree at a higher rate than short ones, but removing many short branches can have the same effect as removing a few long branches. By observing the ages of nodes that are removed from the tree, we can understand whether a fossilization regime tends to favor short or long branches and whether branching events are likely to go unrecorded at particular times. For each pruned tree, we count how many internal nodes fall within log-scale time bins from 10−2 to 102.25 Ma, in exponential increments of 0.25.

We contrast these data on the diversity lost by a geographically structured fossil record with a null model for how the fossil record should record diversity. For our null model, we model the fossil record as a random preserver of species. We achieve this via Monte Carlo simulation, generating randomized datasets in which observed FGR areas are randomly shuffled between species. In this approach, the chance of fossilization for any species becomes independent of its phylogenetic position and geographic range, but distribution of FGR areas is identical to the geographically structured dataset, and therefore we can easily follow the same PD measurement process as with our geographically structured data. We generate 1000 of these “random” fossil records to compare against our geographically structured data. For each of these randomized datasets, we follow all the steps taken to generate PD loss and node age distributions for our geographically structured data, using the same 1000 phylogenetic trees and the same cutoff values for minimum FGR. We then compare the PD distributions of our geographically structured and random fossil records at each FGR size cutoff level using Cohen's D to determine the effect size of nonrandom fossilization at these different range-size cutoff values. To understand the pattern of node loss versus node age, we first compute the average

difference in node number between geographically structured and random records within each node age bin. We then multiply that difference by the node age of the bin to estimate the difference in PD (in Ma) recovered by geographically structured and random fossil records.

Extinction Rates and Extinction Magnitudes

To measure the fossil record's ability to capture the rate and magnitude of large extinction events, we compared our full dataset with a subset excluding species listed by the IUCN as DD, EX, EW, CE, CR, or EN. The full dataset represents current tetrapod diversity, while the reduced dataset represents diversity in a subsequent geologic interval in which the excluded species are all extinct. Plotnick et al. (Plotnick et al. 2016) used a similar approach to investigate potential mass extinction in mammals, though here we include DD species among those present in the modern interval but absent in the post-extinction interval, slightly increasing the magnitude of extinction. We have calculated extinction rates following Barnosky et al. (Barnosky et al. 2011) over a 100 year time interval. Extinction magnitudes are those recovered by the potential fossil record, comparing the diversity at a pre-extinction interval and diversity at a post-extinction interval at the same FGR cutoff value.

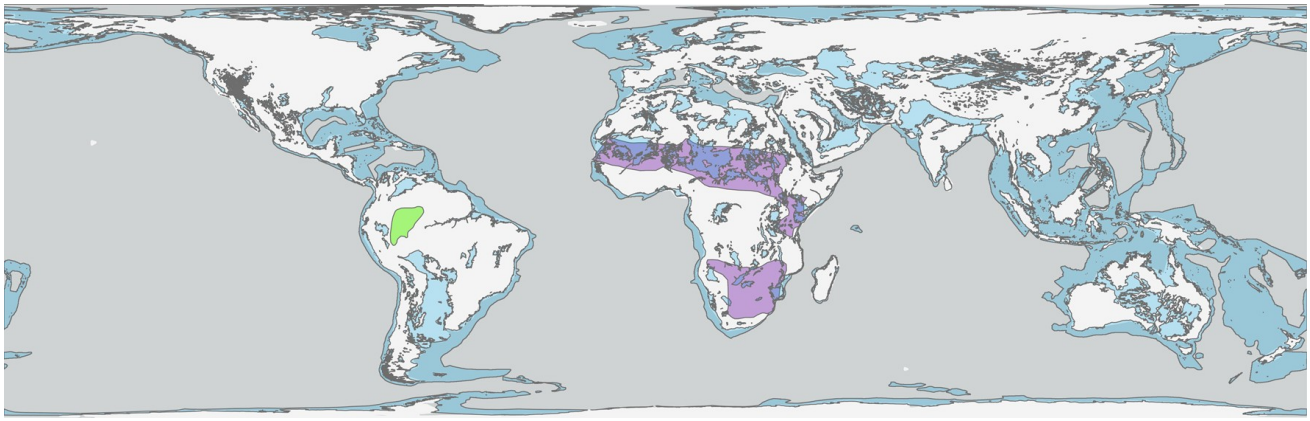

Figure 1: The world's active sedimentary basins (from Nyberg and Howell 2015), and the ranges of the Catequero bullfrog (Tomopterna cryptotis) and the sapito confuso (confused toad, Leptodactylus diedrus). Because the extant geographic range of T. cryptotis (purple) overlaps with sedimentary basins (blue), it may enter the fossil record, producing a fossil geographic range (FGR) of 2,993,359 km2. Leptodactylus diedrus (green) does not range over any sedimentary basins, excluding it from the fossil record. Basemap from ESRI (2017).

Results

Composition of a Geographically Structured Record

Just over 73% of tetrapod species at least partially overlap with areas of long-term deposition, giving them some chance of fossilization. However, there are striking differences in potential fossilization between the four major tetrapod clades at the species (Fig. 2, Supplementary Fig. S1), genus, and family levels. Amphibians are poorly represented, with just 41.9% of species having any chance of entering the fossil record, compared with 72.9% of reptilian species, 81.3% of mammalian species, and 86.0% of bird species (Table 1). None of the taxa that we predict will have an FGR size of zero are represented in the Paleobiology Database (Supplementary Table S1).

Table 1: Proportions of extant tetrapod taxa recovered in a fossil record under six minimum fossil geographic range (FGR) areas for inclusion in the record.

	>1 km^2	>10 km^2	>100 km^2	>1000 km^2	>10,000 km^2	>100,000 km^2
Amphibian species	0.42	0.39	0.33	0.25	0.15	0.07
Amphibian genera	0.79	0.77	0.71	0.63	0.45	0.27
Amphibian families	0.95	0.95	0.89	0.83	0.71	0.56
Bird species	0.86	0.85	0.8	0.71	0.59	0.41
Bird genera	0.93	0.93	0.91	0.86	0.78	0.65
Bird families	0.98	0.97	0.96	0.93	0.87	0.82
Mammalian species	0.81	0.79	0.75	0.64	0.5	0.29
Mammalian genera	0.92	0.92	0.89	0.83	0.72	0.55
Mammalian families	0.99	0.99	0.98	0.94	0.89	0.79
Reptile species	0.73	0.68	0.62	0.51	0.34	0.17
Reptile genera	0.93	0.91	0.88	0.82	0.7	0.5
Reptile families	0.96	0.96	0.94	0.93	0.9	0.75

Phylogenetic Patterns

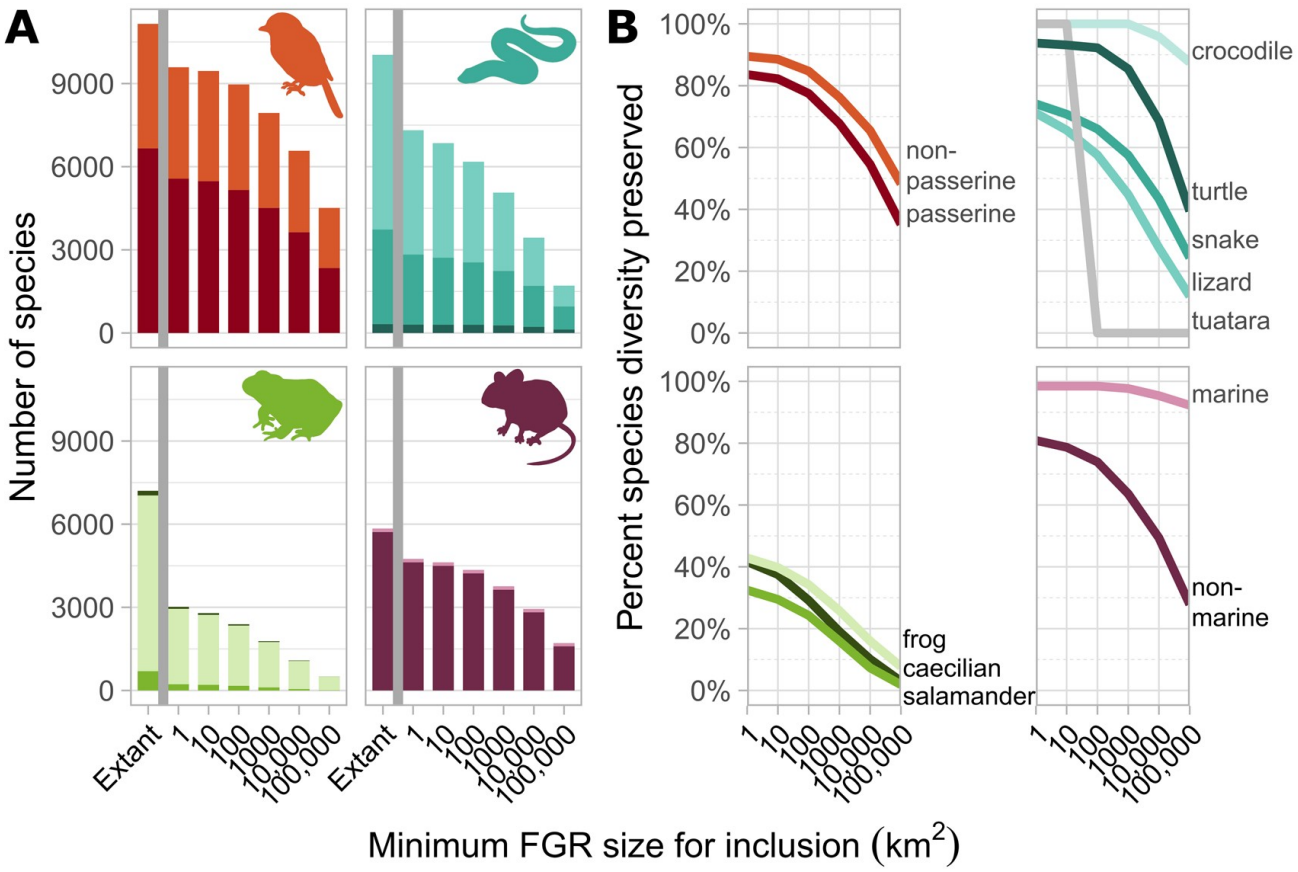

Figure 2: Predicted diversity of the fossil record of modern tetrapods. A, Species diversity in each group (clockwise from upper left: birds, reptiles, mammals, and amphibians) given a minimum fossil geographic range (FGR) size for inclusion. Smaller minimum sizes for inclusion produce more complete records. B, The proportion of extant diversity preserved in the fossil record given a minimum FGR size for inclusion. Colors are the same as in Fig. 4.

Phylogenetic signal (Pagel's λ) of both extant and fossil geographic range areas varies widely between tetrapod groups (Table 2). In squamates and amphibians, extant and FGR areas have a similar level of phylogenetic signal, with FGR signal being slightly stronger, because FGRs can only be equal to or smaller than extant geographic ranges, thereby increasing the overall similarity of tip data. The same is true for mammals, although phylogenetic signal is much higher in general in their geographic

ranges. The phylogenetic signal difference is more pronounced between fossil and extant geographic range areas in birds. Salamanders notably depart from the pattern of increased FGR area signal with a Pagel's λ = 0.76 for extant geographic range areas, but only 0.28 for FGR areas.

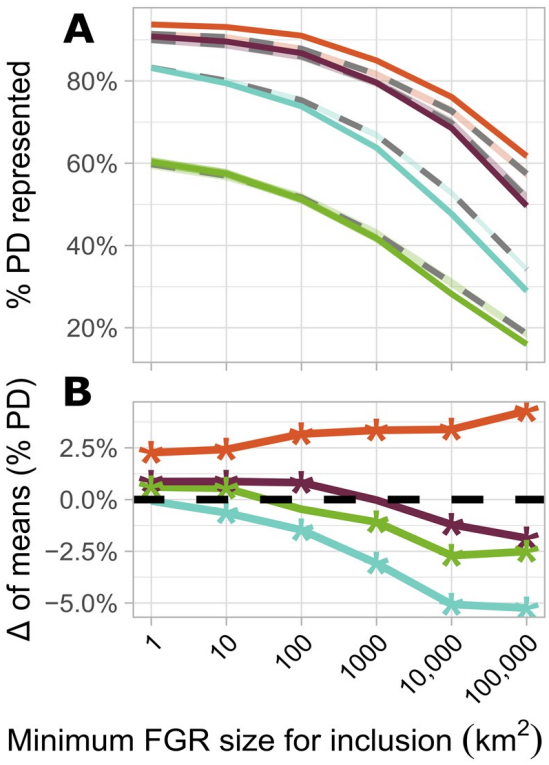

Figure 3: Phylogenetic completeness of the fossil record in squamates (blue), birds (orange), amphibians (green), and mammals (purple). A, Percent phylogenetic diversity (PD) retained by the record at different fossil geographic range (FGR) inclusion cutoff values. Dark intervals with solid lines represent a geographically structured record; light intervals with dashed lines represent simulated data where inclusion in the fossil record is random. B, Difference in PD preserved by geographically structured vs. random fossil records. Positive values indicate geographically structured PD is greater than PD in a random record. Points marked with an asterisk correspond to a "large" Cohen's D (>0.8).

The PD of a fossil record made up of randomly selected species is very different from the PD of a geographically structured record. Surprisingly, geographic structuring does not have a consistent effect on PD across taxa. PD is consistently underpreserved (compared with a randomly selected fossil record) in reptiles, with diminishing preservation with progressively less complete records (higher FGR cutoff values). In mammals and amphibians, more complete fossil records slightly overpreserve PD, while less complete fossil records underpreserve it. In birds, we find the opposite pattern: PD is consistently overrepresented, and overrepresentation is higher in less complete records (Fig. 3).

The pattern of phylogenetic node loss is likewise dissimilar across the four tetrapod trees (Fig. 4). In birds, a geographically structured fossil record preferentially preserves phylogenetic nodes across the entirety of crown-avian history, regardless of the inclusion scenario. Squamates, amphibians, and mammals exhibit a different pattern; younger nodes

are preferentially preserved by a geographically structured fossil record, while older nodes are preferentially lost. However, the timing of the pattern is different. In squamates, late Neogene nodes are favored by geographic preservation, while early Neogene and Paleogene nodes are preferentially lost. In Amphibians, nodes are favored across the Neogene but preferentially lost in the Paleogene. In Mammals, nodes in the late Neogene and early Quaternary are preferentially preserved, while nodes in the middle Neogene and, to a small extent, the early Neogene are preferentially lost. The magnitude of preferential recovery or loss generally increases as the fossil record becomes more incomplete (under higher FGR cutoff scenarios). Under more complete geographically structured records (FGR cutoff 1–100 km2), nodes 1 Ma old or younger are preferentially lost in mammals and squamates.

Table 2: Phylogenetic signal (Pagel's λ) of extant geographic range area and fossil geographic range (FGR) area of extant tetrapod groups.

	Geographic range Pagel's λ	FGR Pagel's λ	Difference in Pagel's λ	Geographic range log likelihood	FGR log likelihood
Bird	0.34	0.45	0.11	-175463	-153095
Passerine	0.07	0.14	0.71	-102191	-88329.8
Nonpasserine	0.3	0.14	0.08	-72429.8	-63495.4
Squamate	0.34	0.42	0.02	-149823	-135232
Lizard	0.4	0.36	-0.04	-94910.4	-86534.5
Snake	0.35	0.37	0.02	-54018.1	-48530
Amphibian	0.43	0.45	0.02	-100716	-88514.8
Frog	0.34	0.35	0.01	-88907.5	-78362.1
Salamander	0.76	0.28	-0.48	-9321.04	-7460.47
Caecilian	0.14	0.32	0.18	-2439.62	-2034.17
Mammal	0.75	0.77	0.03	-99372.9	-88942.2
Terrestrial mammal	0.37	0.41	0.04	-89235.3	-79570.2
Marine Mammal	0.84	0.91	0.08	-2427.9	-2194.92

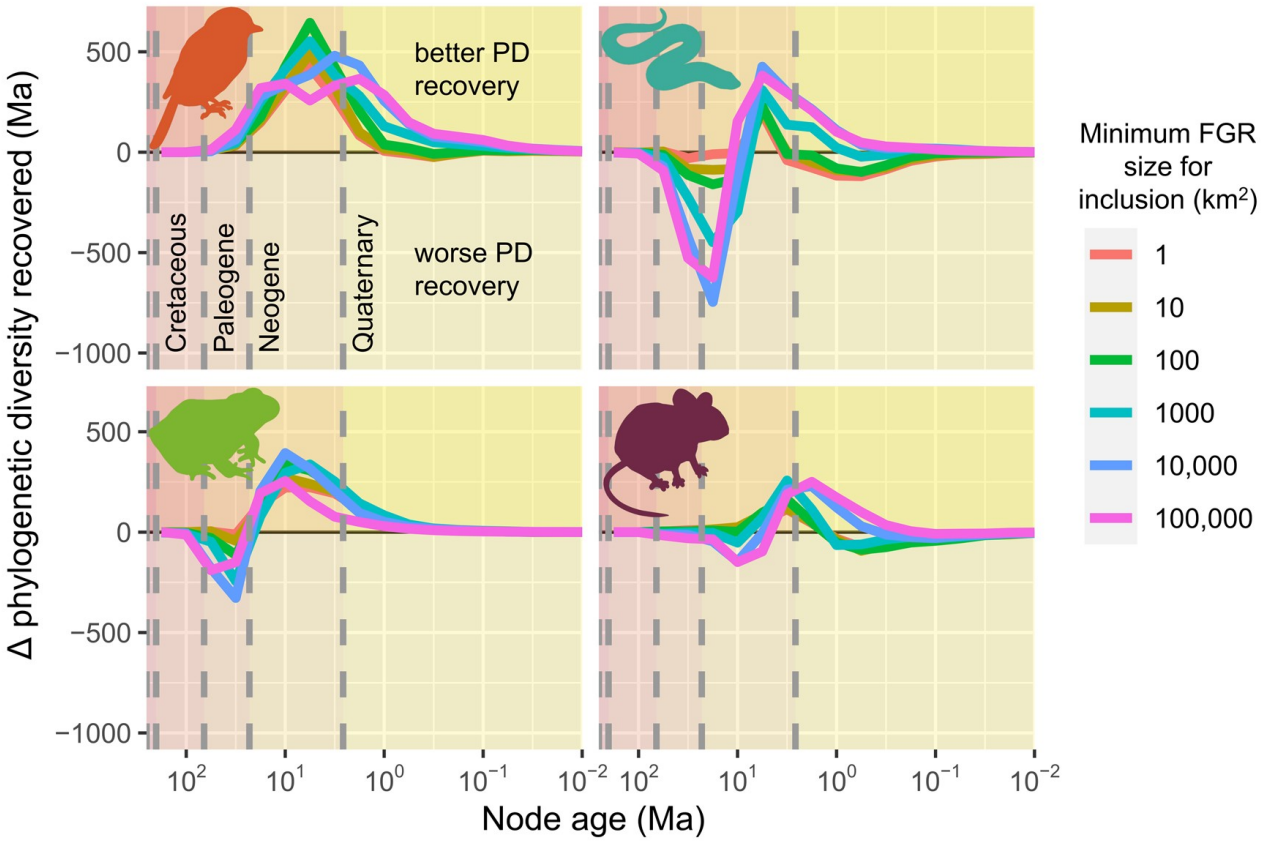

Figure 4: Comparison of temporal patterns of phylogenetic diversity (PD) lost in geographically structured vs. random inclusion of species in the fossil record of birds, squamates, amphibians, and mammals; icons as in Fig. 2. The x-axis represents the ages of nodes in a group's phylogenetic tree, while y-axis values represent the difference in PD recovered by a geographically structured vs. random fossil record for a given age of node. Positive y- axis values indicate that the geographically structured record preserves more PD than a random fossil record.

Extinction

The fossil record would severely underrepresent the magnitude of a large extinction event in amphibians. Our simulated extinction produces a loss of 40% of amphibian species and 13% of amphibian genera. The magnitude of this extinction would be severely underrepresented in even our most complete fossil record. When the minimum FGR area for fossilization is 1 km2, the fossil record registers a a loss of 13.5% of amphibian species and 4% of amphibian genera. In worse fossil records, the magnitude of species diversity loss quickly drops below 5% (Fig. 5B).

This translates to drastically underrepresented extinction rates. For amphibians, the actual magnitude of our simulated extinction is dramatically high (4025 extinctions per million species-years [E/MSY]) (Fig. 5A). This drops to 600 E/MSY in the most inclusive fossil record and is just 7 E/MSY at an FGR cutoff of 100,000 km2: a 99.83% reduction in the apparent extinction rate. We observe this underrepresentation in all tetrapods, although the most severe underrepresentation of extinction rate is in amphibians (Fig. 5A).

Table 3: Extinction magnitude (EM) of tetrapod taxa in a simulated extinction of all tetrapod species listed by the IUCN as Endangered, Critically Endangered, Extinct, Extinct in the Wild, or Data Deficient under different scenarios for inclusion in the fossil record. “Modern diversity” is a count of extant taxa. “Moden EM” is the EM (proportion of extant taxa that will be extinct in the next interval). Subsequent columns report the EM recovered by a fossil record that only preserves species with fossil geographic range (FGR) areas greater than a given value.

	Modern Diversity	Moden EM	>1 km^2	>10 km^2	>100 km^2	>1000 km^2	>10,000 km^2	>100,000 km^2
Amphibian species	7204	0.4	0.14	0.12	0.08	0.05	0.02	0.01
Amphibian genera	549	0.13	0.04	0.04	0.03	0.02	0.02	0
Amphibian families	75	0.05	0.01	0.01	0.03	0.02	0.06	0
Bird species	10064	0.19	0.19	0.2	0.19	0.19	0.18	0.17
Bird genera	1158	0.07	0.07	0.08	0.08	0.07	0.08	0.1
Bird families	88	0.02	0.02	0.02	0.01	0.01	0.03	0.08
Mammalian species	5844	0.28	0.2	0.19	0.17	0.14	0.11	0.06
Mammalian genera	1277	0.13	0.1	0.1	0.09	0.08	0.16	0.04
Mammalian families	157	0.05	0.05	0.06	0.06	0.06	0.05	0.07
Reptile species	11154	0.08	0.05	0.05	0.04	0.04	0.03	0.03
Reptile genera	2359	0.05	0.03	0.03	0.03	0.03	0.02	0.02
Reptile families	244	0.03	0.02	0.02	0.02	0.02	0.01	0.01

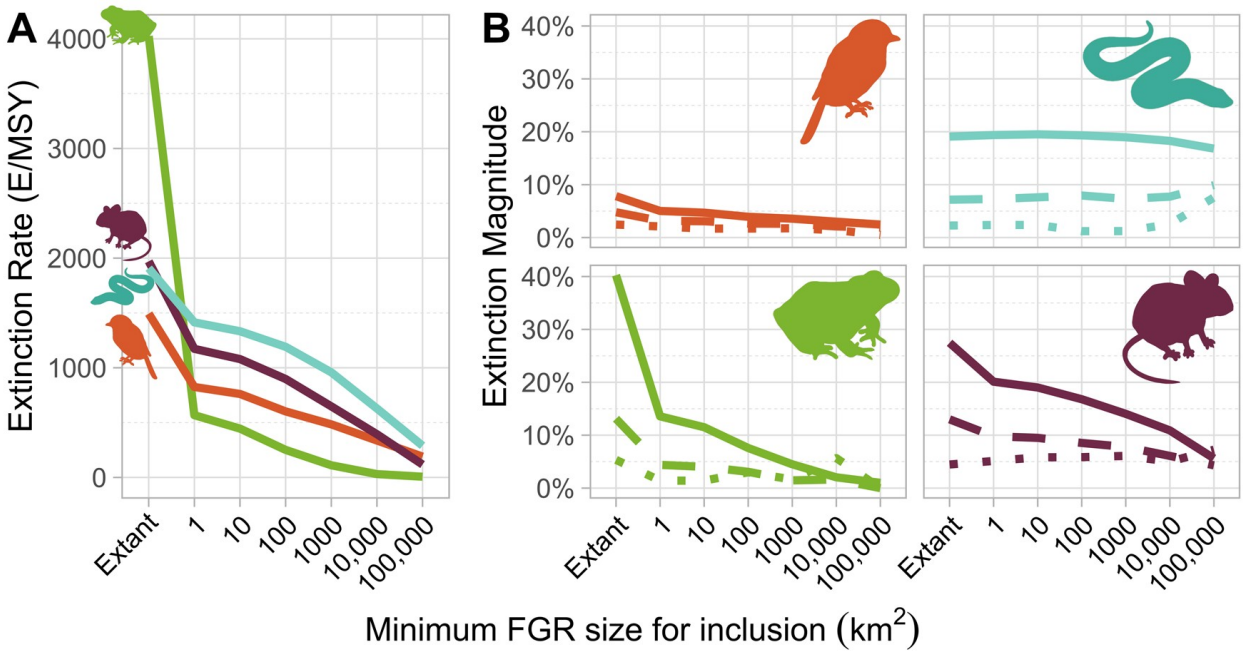

Figure 5: Appearance of the extinction of endangered species in the fossil record. A, Extinction rates (in extinctions per million species-years [E/MSY]) of tetrapods, calculated over a 100 year time interval (following Barnosky et al. Reference Barnosky, Matzke, Tomiya, Wogan, Swartz, Quental and Marshall2011) given a minimum fossil geographic range (FGR) size for inclusion. B, Magnitude of extinction recovered at the species (solid lines), genus (dashed lines), and family (dotted lines) levels given a minimum FGR size for inclusion. Colors as in Fig. 3.

Yet a poor record of extinction rates does not necessarily translate to an underestimate of extinction magnitude. As the minimum FGR size required for inclusion in the fossil record increases, the extinction magnitude that it preserves decreases at the species and genus levels in birds, amphibians, and mammals. However, in reptiles, the species and genus extinction magnitudes recovered by the fossil record hold approximately constant across FGR cutoff scenarios, with a slight increase in approximate genus-level extinction magnitude under the strictest scenarios. At the family level in all four groups, extinction magnitude is quite low under most FGR cutoff scenarios (Fig. 5B, Table 3).

Discussion

Geographic Controls on Taxonomic Diversity

Geography is a major driver of the taxonomic structure of the fossil record, but its effects vary widely between tetrapod clades and different taxonomic levels of analysis. These results imply that ectothermic tetrapods are poorly represented in the fossil record compared with endotherms, with amphibians particularly underrepresented. In part, this may be attributable to their smaller ranges, but the relationship between modern geographic range size and predicted FGR size is not simple; both modern and fossil geographic ranges reflect complex and contingent biogeographic patterns (Fig. 2).

These results reflect fundamental differences in the biogeography of amphibians and amniotes. Compared with biodiversity of other tetrapods, amphibian biodiversity is more constrained by temperature and water availability (Buckley and Jetz 2007), more concentrated in upland environments (Rahbek et al. 2019), and more influenced by topographic complexity (Brown et al. 2014) and elevational heterogeneity (Murali et al. 2021). The proportion of amphibian species we find to be excluded from the fossil record (65%) is remarkably consistent with the 62% of amphibians whose geographic range is primarily upland (Rahbek et al. 2019). This may help to explain the low family stage–level completeness of the lissamphibian fossil record compared with that of other tetrapods (Benton 1989).

Preservation and Extinction

Any gap in fossil preservation hampers our ability to record changes in diversity through time, and we identified potentially large gaps in tetrapod preservation. Thus, extinction rates recovered in the terrestrial fossil record may not reflect actual extinction rates; in our study, they are all much lower. In our simulated modern extinction, extinction rates from even relatively complete fossil records (a minimum FGR size for inclusion of 1000 km^2) are at least halved compared with the “true” rates (Fig. 5A). Extinction magnitudes can vary to a similar degree. Plotnick et al. (2016), following a procedure similar to ours, found that an ~16% magnitude extinction event in extant mammals would appear as an ~8% magnitude extinction at the species level due to the small number of extant mammals found in fossil databases. This recovered magnitude is about half of the true magnitude. At the genus level, the recovered extinction magnitude (6.5%) was slightly more than half of the true extinction magnitude

(10.7%). This approximate halving of extinction magnitude corresponds again to our mammal dataset when the minimum FGR area for inclusion is between 1000 and 10,000 km^2.

Amphibian extinction is particularly obscured; poor fossil records underpreserve extinction rates by orders of magnitude (Fig. 4), because endangered amphibians have smaller FGRs than non-endangered ones (Supplementary Fig. S2). It therefore seems unlikely that future paleontologists would recover any indication of the modern amphibian crisis. The fossil record is simply geographically incapable of representing the type of extinction that appears imminent, and the low stage-level completeness of amphibians (Benton 1989) and low family- and genus-level diversity of Lissamphibia in the fossil record (Uhen et al. 2023) suggest that the record of modern amphibians would be severely restricted by further factors of ecology, anatomy, taphonomy, sedimentation, worker effort, and so on.

This suggests a number of interesting possibilities. First, perhaps amphibian crisis is a common phenomenon; amphibian extinction rates are always high (or extinction pulses frequent), but origination rates are fast enough to maintain diversity. Both phylogenetic (Roelants et al. 2007, Jetz and Pyron 2018*b*) and paleontological (Tietje and Rödel 2017) studies have suggested a high turnover rate in amphibians, which bolsters this idea. Yet it is inconceivable that origination could keep pace with this modern crisis; were modern extinction rates to continue, amphibians would be wiped out within a few tens of thousands of years (Buckley and Jetz 2007; McCallum 2007).

Second, perhaps amphibian evolution is a largely “offstage” phenomenon. A sizable amount of modern amphibian diversity is attributable to the effects of topographic heterogeneity, refugia, and the “montane species pump” (Brown et al. 2014), all of which produce or maintain diversity outside the reach of sedimentation. Even in a relatively complete fossil record, this mode of evolution would produce long ghost lineages as clades move out of and back into areas of sedimentation, skirting fossilization just as they evolve new traits that would allow paleontologists to understand their evolutionary history. Given the difficulty of resolving lissamphibian origins and the long ghost lineages involved (Ruta and Coates 2007), this may be a major and intractable problem if lissamphibian ancestors also preferred upland habitats.

Third, if other extinctions follow the geographic structure of our simulated extinction, their real toll on the terrestrial biota will be consistently underestimated. Although we may have a good understanding of the proportions of marine taxa lost in mass extinctions (in some cases, the extinction proportion can hardly increase), our results highlight a mechanism by which a large proportion of

terrestrial taxa skirt fossilization. These un-fossilizable species may represent major innovations in the history of life, yet we cannot know them directly. These ephemeral taxa lived and died beyond the scope of history.

Phylogenetic Diversity in the Fossil Record

The fossil record is not a random sampler of PD on land, instead preserving each major tetrapod group in its own idiosyncratic regime. This is not surprising, given that phylogenetic endemism in tetrapods is highest in areas of low seasonality and, crucially, high topographic relief (Murali et al. 2021). Yet the heterogeneous patterns of PD preservation in the fossil record are unexpected.

Birds' large geographic ranges seem to translate to high predicted preservation rates (Table 1, Supplementary Fig. S1), similar to those of mammals, in our models. Yet bird PD is far better sampled by our geographically structured fossil record than a random fossil record (Fig. 4). Unlike all other tetrapods, both recent and deep nodes in the avian tree are preferentially preserved (Fig. 5). This pattern is consistent across all FGR cutoff sizes, including those where the rate of bird preservation is similar to rates of mammalian and reptilian preservation at which older nodes are underpreserved. For instance, when the mammalian preservation rate is 50.4%, nodes around 10 Ma old are underpreserved, but even when the bird preservation rate dips to 40.5%, phylogenetic nodes are all preferentially overpreserved. Because of this, the consistent overpreservation of birds cannot be due to mere abundance in the record. Instead, the preferential retention of deep avian nodes must be due to the particular interaction between the shape of the avian tree and the geographic distribution of modern bird species. While passerine birds inhabit diverse environments worldwide, the older-diverging birds of Aequorlitornithes (Prum et al. 2015) are often closely associated with marine, freshwater, or wetland environments—exactly where deposition occurs. The PD of this group, and of the wide-ranging waterfowl, should therefore be very comprehensively sampled by sedimentary basins and represents a larger number of long branches with few tip taxa than the "bushier" passerine tree. This is consistent with a higher degree of phylogenetic signal in FGR size among nonpasserines than among passerines (Pagel's $\lambda = 0.3$ vs. $\lambda = 0.14$; Table 2).

The striking difference in phylogenetic signal between salamander extant geographic range and FGR areas ($\lambda = 0.76$ vs. $\lambda = 0.28$; Table 2) likely stems from a strong tendency for elevational speciation in salamanders. In Plethodontidae, the family containing more than 60% of salamander species, high- and low-elevation species have repeatedly evolved from middle-elevation origins (Wiens

2007; Kozak and Wiens 2010). As species repeatedly invade high-elevation habitats, in which little deposition is likely to occur, their FGR areas become zero, while their middle-elevation sister species may have appreciable FGRs (although likely smaller than their extant geographic ranges). Likewise, species that invade low-elevation environments are likely to increase the sizes of their FGRs, approaching the sizes of their extant geographic ranges. Through elevational speciation from this middle-elevation origin, sister species FGRs become less similar, and the phylogenetic signal of FGR area drastically drops.

We do not recover any signal similar to that seen in salamanders in passerine birds (Table 2). Despite the group's remarkable diversity and elevational diversity gradient, passerines' high proportion of downslope to upslope dispersal (van Els et al. 2021) likely brings high- and middle-elevation passerine lineages regularly back into sedimentary basins over long timescales. Alternatively, it may be that signals of salamander-like patterns of elevational diversification in some groups of passerines are simply swamped by other evolutionary dynamics in this incredibly diverse group.

The relatively high levels of amniote PD sampled by our hypothetical record implies that, at least in principle, amniote evolutionary history can be well preserved. Fossils of early members of important modern clades are crucial to understanding the clades' histories and rates of diversification, and it seems that sedimentation itself is not an insurmountable barrier to the discovery of such fossils. Worker effort in finding and interpreting these fossils is likely a greater impediment to understanding amniote history than initial preservation.

Understanding the Structure of the Terrestrial Fossil Record

How good is the terrestrial fossil record, really? Foote's (1997) and Tietje and Rodel's (2017) estimates of the completeness of the mammalian and amphibian fossil record, though quite disparate in scope, correspond with our predicted record when the minimum FGR size for inclusion is 100–1000 km2. In this record, 25–33% of amphibians, 71–80% of birds, 64–75% of mammals, and 51–62% of reptiles might fossilize (Table 1). But our "completeness" is quite different from the stratigraphic "completeness" estimated in these sources. Here, we compare a global living assemblage to a predicted global death assemblage, asking what fraction of "species that lived" might fossilize, rather than what fraction of "species that lived and are detectable" in the fossil record appear between intervals (Toivonen et al. 2022), as Foote (1997) and Tietje and Rodel (2017) do. Using this definition, Žliobaitė

and Fortelius (2022) estimate that perhaps 4–10% of Miocene mammalian species left fossil remains. This is a smaller proportion than the ~15% of modern mammalian species represented by fossils in Plotnick et al.'s (2016) analysis, perhaps resulting from a better fossil record for the Pleistocene than the Miocene—a strong “Pull of the Recent.” But what accounts for this tiny proportion of preservation?

Moving hierarchically through the structuring processes of the fossil record, we can begin to estimate the relative contributions of different processes to the completeness (percent of total living fauna preserved) of the record. Here, we estimate that at least 19% of mammalian species can be eliminated from the record based on geography alone (Table 1). To reach Plotnick et al.'s (2016) 15% completeness estimate, another 66% of diversity must be lost due to (1) ecological, anatomical, taphonomic, and sedimentary processes that prevent taxa from entering the sedimentary record; 2) sedimentary, tectonic, and environmental processes that destroy sedimentary rocks or make them unavailable; (3) some combination of worker effort and sampling imperfection that has failed to uncover and correctly identify all of the fossilized species; and (4) taxonomic uncertainty, as some species are not diagnosable based on fossilizable traits.

We have designed our study to capture maximally inclusive FGRs based on the data available, making our estimates of diversity lost to geography quite conservative. Geographic ranges are labile and currently in flux for many species spreading into new habitats, retreating from inhospitable environments, or experiencing population declines (Lenoir and Svenning Reference Lenoir and Svenning2015), and these changes may nudge some species toward better fossilization potential and others away from it.

Organisms’ anatomical and ecological characteristics matter a great deal to their probability of preservation. Alongside geographic range size (Wagner and Marcot 2013; Plotnick et al. 2016), body size (Behrensmeyer et al. 2000; Wagner and Marcot 2013; Plotnick et al. 2016; Mannion et al. 2019) and species abundance (Wagner and Marcot 2013; Plotnick et al. 2016) are known to correlate with preservation, and their effects could be considered in more comprehensive models of the potential fossil record. Moreover, particular traits of clades can influence how well they are represented by fossils. For instance, modern crocodiles are ideal fossilizers: large-bodied, robust-boned animals living in and around low-energy aquatic environments. We predict their inclusion in the future fossil record is secure, continuing the already fairly complete crocodilian record (Markwick 1998; Mannion et al. 2019). The skeletons of birds are made up of much smaller elements, usually quite fragile, and are

more prone to taphonomic degradation than the skeletons of other tetrapods, probably contributing to their poorer fossil record (Benton 1989). (Although it is still somewhat phylogenetically complete (Ksepka and Boyd 2012).) Amphibians' small, often weakly ossified skeletons likewise seem to preserve poorly, again likely contributing to their remarkably spotty fossil record (Benton 1989; Tietje and Rödel 2017). In contrast, mammals' tough, complex, and highly diagnostic teeth are ideal for both preservation and later worker recognition.

We have considered all sedimentary basins equivalent and static, but variation in sediment sources, sediment transport mechanisms, sedimentation rates, fluvial regimens, accommodation rates and basin architecture (e.g., difference between an active foreland basin vs. a passive margin), climate, and long-term changes within a given basin will leave some sedimentary basins with far better preservation potential than others. Ecological gradients, which are inherently coupled to elevation gradients and climate variables, also act as a control on the structure of the fossil record (Holland and Loughney 2021; Holland 2022) in a way not considered in these models. However, they are in part expressed in the tetrapod ranges themselves. The climate and environment within a sedimentary basin may also affect taxonomic fidelity in its fossil record, with cool, dry climates producing the highest preservation potential (Behrensmeyer et al. 2000). Furthermore, large shifts in geomorphology (e.g., the retreat of the Laurentide Ice Sheet) can cause massive reorganization of the erosional and depositional makeup of sedimentary basins (Wickert 2016), which can lead to the erosion of previously stable sediments. Human impacts on sedimentary rock production are already apparent, producing increases in both erosion (Reusser et al. 2015) and sedimentation (Jenny et al. 2019). These effects will further confound our assumption of preservational homogeneity, perhaps by producing higher sedimentation rates and a higher-fidelity fossil record in more biologically depauperate areas. However, these human impacts on the sedimentary record are likely to be insignificant in the whole of a basin's stratigraphy.

Our model also fails to account for atypical modes of preservation. Cave and fissure-fill deposits can preserve species otherwise absent from the fossil record (Lundelius 2006; Jass and George 2010), although these deposits are generally short-lived, and very few are known that represent deep-time upland environments (Lundelius 2006). Preservation of terrestrial species in marine environments, however, is not uncommon (Schultze 1995; Butler and Barrett 2008). The extent to which postmortem

transport and burial add otherwise unsampled species to the fossil record is unclear, and a potential source of fossil diversity for which our geographic approximations cannot account.

Moreover, a thorough comparison of modern and fossil faunal diversity must account for exceptional spatial heterogeneity in paleontological effort; western Europe, the United States, and Canada are blanketed with known fossil localities, while the Global South—including areas of extremely high modern tetrapod diversity and endangerment—is poorly represented (Raja et al. 2021).

Beyond collecting, incomplete worker effort in identifying and publishing specimens, as well as keeping up with taxonomic changes, contributes to a mismatch between our accounting of the modern fauna and their recent fossil record. Neontologists are increasingly describing cryptic species based on ecological, genetic, or otherwise un-fossilizable characteristics (Bickford et al. 2007; Barrowclough et al. 2016; Burgin et al. 2018), making species an increasingly difficult unit of comparison for modern and fossil data.

The Future Fossil Record

It is difficult to consider our results outside the context of global warming and human impacts on the planet. Some researchers refer to this time of heightened human impact as the Anthropocene (Lewis and Maslin 2015). While not recognized as a subdivision of geologic time by the International Commission on Stratigraphy (ICS) or the International Union of Geologic Sciences (IUGS), the stratigraphic record will undoubtedly retain evidence of human activity. Climate change will globally disrupt ecosystems, changing ecological gradients just as sea-level rise changes the sediments accumulating along them (for a detailed discussion, see Holland 2022).

Our results indicate a possible existential problem: Will the future fossil record capture the current diversity crisis? How will range shifts appear in the future fossil record, with organisms migrating into and out of sedimentary basins in the midst of species extinction, evolution, and turnover? And ultimately, how many diversity crises are missing from the geologic record? These questions may be ultimately untestable, but we hope that they prompt studies that broaden the taxonomic and phenomenological scope of this study.

Conclusions

Our results are at once reassuring and deeply troubling. We confirm that mammals, long used as a model for terrestrial vertebrate biodiversity, are the most reliable recorders of tetrapod history, having the highest proportions of potential fossilization and eventual recovery (Fig. 2). Yet amphibians, representing a similar proportion of extant vertebrate diversity, leave behind far less; a clear understanding of amphibian history may be fundamentally impossible to recover.

Modern biodiversity is the outcome of contingent geologic, climatic, and evolutionary processes, and as such is not a perfect model for biodiversity in past or future points in Earth history. However, large-scale patterns in tetrapod biogeography are not mere recent developments; they are a momentary state in a long history structured by profound biological differences between major clades, differences that produce a contingent and idiosyncratic fossil record through the interaction of climate, geography, evolution, extinction, geology, and taphonomy. The scale of these differences demonstrates a need to carefully consider the histories of these groups and understand their unique macroevolutionary dynamics. Future work comparing modern and fossil biodiversity will greatly benefit from this geographic perspective, as well as from a broader taxonomic scope. What other evolutionary stories—of insects, conifers, or pulmonates—might the fossil record pass over?

That nearly 20% of modern mammalian species have almost no chance of fossilizing should caution any paleontologist against ignoring the geographic structure of the fossil record. Much past biodiversity may be unknowable, but by understanding the processes that failed to record these species, we can better understand the structure and scope of history itself.

Transition

The spatial heterogeneity of fossilization strongly influences the taxonomic makeup of the fossil biota – but what of their functional groups? Specifically, are taxa with different substrate preferences, for instance, fossoriality, more or less likely to fossilize than others?

For 7986 squamate species with assessed FGRs and whose substrate preference could be categorized as “arboreal” (listed as at least somewhat arboreal), “fossorial” (listed as at least somewhat fossorial), “aquatic” (listed as marine or at least somewhat aquatic), or none of the above (assumed to be purely terrestrial) according to a recent database (Meiri 2024). I measured percent fossil record completeness for species in each substrate use category at the FGR cutoffs used in the original study. I generated null distributions for expected completeness in these categories via monte carlo simulation, shuffling FGR sizes between species 10,000 times, and generated 95% quantile intervals from those null distributions to visualize the difference between random preservation and geographically structured preservation.

Substrate use can be highly influential on fossilization potential. Unsurprisingly, aquatic species (n = 344) are over-represented in the fossil record, and to a large degree: by 10% or more across all FGR sizes. At low minimum FGR cutoffs, fossorial species are slightly under-represented, by 2-3% at FGR cutoffs from 0-10 km^2, but their under-representation is within the 95% interval of random preservation at higher FGR cutoffs. Arboreal species are under-represented at FGR cutoffs from 1 to 10,000 km^2, by a maximum of about 5% for all species with FGRs greater than 1000km^2. The proportional under-representation of arboreal (n = 2517) and fossorial (n = 1565) species drives an over-representation of unspecialized terrestrial species (n = 3560) (Figure 6). While fossorial species tend to be under-preserved, they are only significantly under-preserved in very high-quality fossil records, which are unlikely to be reflective of real recovered records.

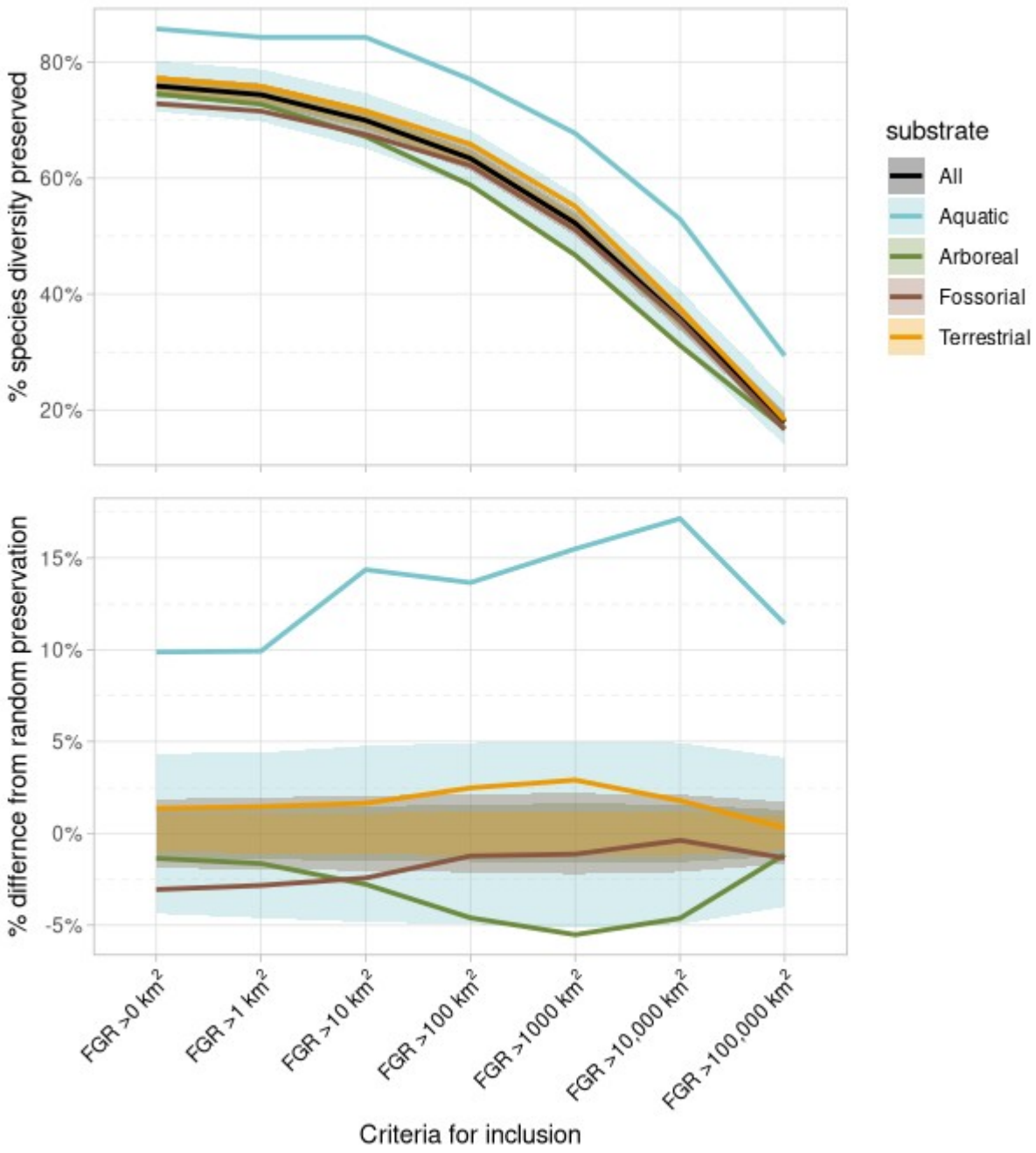

Figure 6: Percent fossilizaiton potential for squamates with different substrate preferences across seven minimum FGR cutoff sizes. Colored bands represent 95% confidence intervals based on simulations randomizing FGR size 10,000 times.

Conclusion

Our work is the first attempt to investigate the geographic bottleneck on the preservation of biodiversity at a global scale and across a large and globally important group of organisms. Its most valuable findings are two insights onto the heterogeneity of fossil preservation.

The first important finding is that the amphibian fossil record should be inherently poor based solely on the global biogeography of the group, and that the reduction in the fidelity of amphibians purely due to geography is quite extreme compared to other tetrapods. This result was unexpected and has major implications for our understanding of macro-evolutionary processes in the group. Amphibians are overwhelmingly threatened by human-caused extinction (CITE) and their declines are the subject of much research and debate, especially as they may relate to the extinction risk of tetrapods in general. Yet our results imply that an accounting of global historical amphibian diversity via the fossil record is simply impossible. The fact that a significant extinction event in a major tetrapod group could pass almost unnoticed through the fossil record should cause paleontologists to consider very carefully how apparent patterns in fossil taxa may be shaped by purely geographic forces; specifically, the concentration of diversity among upland environments. It also throws weight behind paleontological models that associate rarer taxa with more upland environments (e.g. Brinkman et al. 1998).

The second important finding is that, for most groups, the fossil record has a propensity to under-preserve phylogenetic diversity. It is yet unclear exactly why this might be, and several possible explanations have interesting implications for the structure of the fossil record.

First, this may reflect the propensity for mountain ranges to function as “museums” (Rahbek et al. 2019), where heterogeneous conditions allow lineages to hold on in small pockets of suitable habitat long after more homogeneous lowland habitats have been rendered inhospitable by changing climate regimes. In this scenario, we would expect the fossil record to generally record “early” extinctions, as lineages may persist outside of the geographic scope of fossilization long after their last fossils are buried, and in particular, we should expect this effect to be associated with periods of warming, in which habitable zones shift upslope. Likewise, we would expect the reappearance of Lazarus taxa in cool intervals, when conditions usually present only in higher altitudes spread to the lower basins in which fossils form.

The second possibility is that traits that favor diversification either are or tend to co-occur with traits that favor fossilization. This seems somewhat paradoxical, as mountain ranges are not simply museums for biodiversity but prolific generators of it (Marder et al. 2025). Therefore we would expect fossilizability of clades to be to some degree inversely correlated with their diversity. In birds, it does appear that the pattern of phylogenetic diversity over-preservation is trait-linked, but the increasing over-preservation of phylogenetic diversity in poor fossil records suggests that the overall topology of the avian tree is the important driver. Older groups tend to be both aquatic and wide-ranging. We see the same inverse diversification-fossilizaiton pattern in the investigation of squamate substrate use. Arboreality has a positive effect on squamate net diversification rates (Bars Closel et al. 2017; Li and Wiens 2022), but a negative effect on fossilizability.

Chapter 2 - The phylogenomic and biogeographic history of *Dibamus* (Squamata: Dibamidae)

Introduction

Among the many groups of reptiles that have evolved a snakelike body plan, few are more mysterious than the burrowing lizards of the family Dibamidae. These poorly-studied fossorial lizards are a phylogenetic and biogeographic enigma; the two genera, *Dibamus* and *Anelytropsis*, are separated by both the Pacific Ocean and tens of millions of years of evolutionary time, yet are undoubtedly each others' closest relatives (Greer 1985; Townsend et al. 2011; Simões and Pyron 2021). Both *Anelytropsis*, endemic to Mexico's Sierra Madre Oriental, and *Dibamus*, found from Southern China through Indochina and Wallacea to New Guinea, are small burrowers with heavily reduced eyes, covered by a broad ocular scale, no pectoral girdle, and tiny, flap-like hindlimbs present only in males. Their relationships to the rest of squamata have been the subject of much revision (see Simões and Pyron 2021 for discussion), and even now are unclear; Dibamidae appears to be either sister to Gekkonidae or sister to all other Squamates (Burbrink et al. 2020; Simões and Pyron 2021; Singhal et al. 2021).

The biology of dibamids is almost totally unstudied. In part, this is because they are rarely encountered on the surface. *Dibamus* are most often found in soil, leaf litter, or under cover (Kliukin et al. 2026), though they may be forced out of these low-lying microhabitats during rainstorms. Dibamids are presumed to be insectivorous, though their diets are almost entirely unknown (Krone 2024), and are sometimes associated with termites (Kliukin et al. 2023, 2024*a*). More than half of *Dibamus* species are known only from their type localities (Meiri et al. 2018), and many from only one or a handful of type specimens. *Anelytropsis papillosus* is known from perhaps forty specimens in collections, and scientific documentation of the species is restricted largely to comments on its geographic range (Axtell 1958; Valdez-Villaviceno et al. 2016; Montiel-Veranza et al. 2022).

Given this neglect, Dibamidae presents a ripe opportunity for both the discovery of new species and the study of the persistence and diversification of clades with highly reduced body plans. Here, we use a large phylogenomic dataset and morphological data to study the biogeographic history

of *Dibamus*, with special respect to its diversification in Wallacea, and identify several candidate species within the Wallacean region.

Taxonomic history

The majority of scientific work on Dibamidae has been taxonomic, largely over the past twenty-five years. Duméril & Bibron erected the genus *Dibamus* in 1839 with the description of *Dibamus novaeguineae* based on specimens from the Museum in Leyden, identified as *Acontias subcaecus* (Duméril et al. 1834). These specimens were likely collected by Salomon Müller, who notes *Acontias* among the reptiles collected by the voyage of the Triton in 1828 from Uru Languru (Triton Bay) (Haan, W. de et al. 1839). Cope (1885) erected a family, Anelytropidae, in his description of *Anelytropsis papillosus*, not recognizing the similarity of *Anelytropsis* to *Dibamus* likely due to a lack of familiarity with the latter genus. Boulenger (1884) followed Cope's opinion, yet to be published, of Anelytropidae as a unique family in his synopsis of lizard families, which established the family Dibamidae. By the time of Greer's (1985) monographic study uniting the two genera in the family Dibamidae, several more species of *Dibamus* had been named, of which Greer recognized six; *D. alfredi* (Taylor 1962) from Thailand, *D. bourreti* (Angel 1935) from Northern Vietnam, *D. celebensis* (Schlegel 1858) from Sulawesi, *D. leucurus* (Bleeker 1860) from Agam, Sumatra, with *D. argenteus* (Taylor 1915) from Augusan province in Mindanao (Philippines), recognized as its junior synonym, *D. montanus* (Smith 1921) from Central Vietnam, and *D. novaeguineae* (Duméril et al. 1834). Greer then described three additional species; *D. seramensis* from Seram island, in the Moluccas East of Sulawesi, *D. smithi* from Central Vietnam, and *D. taylori* from the Lesser Sunda Islands in southern Wallacea. Four more *Dibamus* species, *D. bogadeki* and *D. greeri* (Darevsky 1992), *D. somsaki* (Honda et al. 1997), and *D. deharvengi* (Ineich 1999) were described by the year 2000.

Since 2000, the number of recognized *Dibamus* species has ballooned to 29, based both on reappraisals of existing specimens and active collecting throughout Indochina and Indonesia. Almost all of these species have been described using morphology alone, largely based on details of scalation and vertebral counts, though recent publications have involved more detailed morphometric analysis (Prasetyo et al. 2025), descriptions of skeletal morphology from Computed-Tomography scans (Kliukin et al. 2023, 2024*b*, *a*, 2025, 2026), and analysis of DNA sequence data to generate phylogenetic hypotheses (Kliukin et al. 2024*a*, 2025, 2026), building a framework for more integrative taxonomic work on *Dibamus* and a deeper understanding of the phylogenetic history of Dibamidae.

Dibamus are known from throughout Wallacea, and have crossed both Wallace's and Lydekker's lines. Currently, six species of *Dibamus* are recognized from Wallacea and the Philippines, though the geographic ranges and boundaries of these species are unclear. Prasetyo et al. (2025) present the most recent and best-considered accounting of this diversity. They consider *D. novaeguineae* as a species complex inhabiting the Philippines and Eastern Indonesia on both sides of Lydekker's line, with representative specimens from western New Guinea, the Moluccas, North Sulawesi, the Lesser Sundas, and the Philippines. A broad definition of *D. celebensis* includes all other *Dibamus* from mainland Sulawesi, while *D. taylori* remains the primary species in and restricted to the Lesser Sundas. Three insular endemic species are described from around Sulawesi; *D. seramensis* on Seram, in the Moluccas (Greer 1985), *D. manadotuaensis* Koppetsch et al. 2019 from Manado Tua, a small volcano off of the Northern Peninsula of Sulawesi, and *D. oetamai* Prasetyo et al. 2025 from Buton island off of the Southeast Peninsula of Sulawesi. Also reportedly present in the Philippines is *D. leucurus* Bleeker 1860 (Greer 1985).

Dibamid evolution

Dibamid phylogeny is not well-understood, with only four published phylogenetic works on the interrelationships of the family. Townsend et al. (Townsend et al. 2011) used six nuclear protein-coding genes (BDNF, CMOS, DNAH3, NKTR, RAG1 and R35) and one mitochondrial protein-coding gene (ND2) to investigate the divergence time between *Dibamus* and *Anelytropsis*, finding *Anelytropsis* nested between two geographically concordant *Dibamus* clades. The "mainland" clade, was represented in their study by *D. booliati* and *D. greeri* from Laos and Northern and Central Vietnam (with one specimen previously referred to *D. montanus* subsequently re-identified as *D. greeri* (Kliukin et al. 2023)). The "Peninsular/Island" clade is made up of specimens from Peninsular Malaysia (*Dibamus tiomanensis*, from Tioman Island), Sulawesi (*D. celebensis*), Seram (*D. seramensis*), and *D. novaeguineae* (locality unknown). The placement of *Anelytropsis* within this tree was unclear; a maximum likelihood analysis including *Sphenodon* and 63 squamate taxa placed *Anelytropsis* as a distant sister to the Peninsular/Island clade with low support (bootstrap value of 0.71), while a Bayesian analysis including only the eight dibamid specimens recovered a robustly-supported tree with *A. papillosus* as an outgroup to the "mainland" clade. Divergence dating using a Bayesian relaxed-clock method over the full dataset with external fossil calibrations (at the time, no fossil dibamids were

known) recovered the same *Anelytropsis* + mainland topology, and dated the split between this clade and the Peninsular/Island clade at 69 Ma, with a 95% confidence interval spanning 49-90 Ma.

In the first more comprehensive treatment of Dibamid phylogeny, (Kliukin et al. 2024*a*) also recovered *Dibamus* as paraphyletic using a 224 bp fragment of the Mitochondrial 16s gene, though their sampling somewhat clarifies the geographic extent of the Mainland and Peninsular/Island clades. In this study, *Dibamus bogadeki*, endemic to Hong Kong, and *Dibamus bourreti*, found in Northern Vietnam, form a clade sister to a moderately well-supported clade putting *Anelytropsis* sister to Cambodian and island *Dibamus.* The Cambodian *D. dalaiensis* and *D. elephantinus* form a clade sister to the Indonesian/Papuan *Dibamus* novaeguineae, with moderately high branch support, Indicating that the "Peninsular/Island" clade extends West of the Malay Peninsula into far Southern Cambodia, contra the range indicated by Townsend et al. (2011).

More recent and more comprehensive phylogenies of the group based on mitochondrial ND2, 16S rRNA, and COI genes (Kliukin et al. 2025, 2026) again find *Anelytropsis* within *Dibamus*, though allied with *D. greeri, D. bogadeki* and *D. bourreti*. Outside of this clade containing *Anelytropsis* and the northern *Dibamus*, Kliukin et al. (2025) recover a "mainland" *Dibamus* clade sister to a Peninsular and Island clade, while a more recent assessment including more taxa (Kliukin et al. 2026) finds two *Dibamus* clades as outgroups to the Peninsular and Island clade; a largely Cambodian/Southern Vietnam group containing *D. dalaiensis*, *D. elephantinus*, *D. montanus,* and *D. annae*, and a second clade containig *D. irregularis*, *D. tropcentr,* and *D. deimontis*.

Recent phylogenomic analyses have failed to resolve the relationships of Dibamidae within Squamata, though it is clear that Dibamidae branches quite early in squamate phylogeny, likely in the Early Jurassic (Streicher and Wiens 2017; Burbrink et al. 2020; Singhal et al. 2021). Three primary hypotheses have been supported in recent work on squamate phylogenomics. First, and least broadly supported, is the hypothesis that Dibamidae is sister to all non-Gekkotan squamates (clade Unidentata) – for simplicity's sake, the D+U hypothesis. Second, with broad support, is the hypothesis that Dibamidae is sister to Gekkota; the D+G hypothesis. The third hypothesis places Dibamidae as a sister to all other squamates, the D+S hypothesis, and is also broadly supported. Streicher & Wiens (2017) used ultra-conserved elements (UCEs) to produce a phylogenetic hypothesis for squamata as a whole, but were unable to confidently place Dibamidae (*Anelytropsis* + *D. novaeguineae*); species tree methods favored D+S, while concatenated ML analysis favored D+G. Capture efficiency for

Anelytropsis papillosus in this study was notably poor; just 90 UCEs were recovered for this species, only 3.3% of the 2738.7 UCE loci recovered on average for 37 taxa. Singhal et al. (2021), in a study with a broader taxonomic scope using the same Dibamid data and including a different marker set for other squamates (Singhal et al. 2017), recovered two likely positions for Dibamidae; D+S or D+U, strongly favoring the position as sister to squamata (D+S). In contrast, Burbrink et. al (2020), using a set of 394 anchored phylogenomic loci, but only a single dibamid, *D. novaeguineae,* weakly favored a Gekkota+Dibamidae clade (D+G), according to both ML concatenated and species-tree estimates, with D+S the next most likely hypothesis, weakly favored among loci that did not support the D+G hypothesis. Title et al. (2024), favored D+G, though their approach to constructing a phylogeny appears to have excluded dibamids from the initial backbone creation (Title et al. Table S1), and their placement of dibamids is based on 118 samples covering 50 loci sampled from *Anelytropsis papillosus* and 8 species of *Dibamus*.

A shortfall of these phylogenomic studies is the lack of genomic-scale representation of the "mainland" clade of *Dibamus*. Given the ancient (~70 Ma) divergence time estimated for the dibamid crown (Townsend et al. 2011), the inclusion of only a single branch of *Dibamus* misses an opportunity to more uniformly sample genomic diversity within Dibamidae and reduce problems of long branch attraction that may erroneously ally Dibamidae with Gekkota.

The poor resolution of dibamid intra-relationships and squamate-dibamid inter-relationships may in part stem from the same cause; fairly rapid sets of ancient divergence events. A dataset predisposed to resolve such events is therefore a logical next step in addressing this difficult node. Rapidly-Evolving Long Exon Capture (RELEC) loci could provide such a dataset; this set of long exons are up to ten times more phylogenetically informative over long (50+ Ma) timescales (Karin et al. 2020).

In this study, we sequence a set of 324 molecular markers including both RELEC and other loci to infer the phylogenetic relationships of *Dibamus*, building a dataset that will be necessary to resolve the outstanding issues in dibamid phylogenetics. We present the most comprehensive phylogeny of *Dibamus*, including multiple representative samples from each of the previously identified clades and a particularly large sample of specimens from Wallacea. We also present a small morphological dataset of measurements of 42 Wallacean *Dibamus* specimens and comment on unrecognized diversity within the clade.

Materials & methods

Material examined

We obtained samples of flash-frozen or alcohol-preserved liver tissue from 73 *Dibamus* specimens from seven institutions, nominally representing 23 currently recognized species, approximately 78% of *Dibamus* diversity. A partially overlapping set of 36 Wallacean *Dibamus* specimens were used for our morphological analysis. Institutional/ collector abbreviations are as follows: ABV - Anna B Vassilieva field series; AMNH – American Museum of Natural History (herpetology collection); FMNH – Field Museum of Natural History (herpetology collection); JAM – Jimmy A. McGuire field series; LSU – Louisiana State University (herpetology collection); KUH – University of Kansas Herpetology collection; MCZ – Harvard Museum of Comparative Zoology; MZB – Museum Zoologicum Bogoriense; MVZ – Museum of Vertebrate Zoology (herpetology collection); NAP – Nikolay A. Poyarkov field series; NCSM – North Carolina State Museum (herpetology collection); RMB – Rafe M. Brown field series; ROM – Royal Ontario Museum; TNHC – Texas Natural History Collection (herpetology collection); UF – Florida Museum of Natural History (herpetology collection); ZMFK - Zoologisches rschungsmuseum Alexander Koenig (herpetology collection); ZMMU – Zoological Museum of the University of Moscow (herpetology collection). Specimen Metadata are available in table S1.

DNA Extraction

All DNA extraction and library prep steps were performed in the MVZ Evolutionary Genetics Laboratory (EGL). We subsampled tissues and allowed the tissue fragments to dissolve in a solution of lysis buffer + proteinase K for 3-12 hours at 65C before extracting DNA using an Omega Mag-Bind magnetic bead extraction kit. Post-extraction, we assessed DNA concentration using a qBit spectrofluorometer and ran 1-2uL of each sample on a 1% agarose gel to assess DNA fragment length.

Library preparation

After assessing DNA quality, we sonicated samples in a Qsonica Q800 sonicator for either two rounds of 2 minutes and 30 seconds (for samples with high DNA concentration & a high molecular weight band) or two rounds of one minute and thirty seconds (for samples with low DNA concentration

or lacking high molecular weight DNA), in either case pulsing for 15 seconds on at 40% amplitude followed by 15 second rest intervals and spinning down the sonication tubes between the rounds.

After plating, library preparation was carried out based on the KAPA Hyper Prep Kit protocol with enzymatic steps quartered to reduce costs. We used a double-sided SPRI bead cleanup process to select for DNA fragments between 300 and 500 bp long, first using a 0.5x ratio of SPRI bead solution to select for fragments smaller than 500 bp, then using a 0.65x ratio to select for fragments from that pool larger than 300 bp, retaining the rejected DNA fragments of any future use. Following size selection, we performed an end repair and A-tailing reaction following the KAPA recommended program, holding 15 µL of each sample in a thermocycler at 20°C for 30 minutes, 65°C for 30 minutes, and then holding the mixture at 4°C before immediately proceeding to adapter ligation, allowing the ligation reaction to proceed overnight at 4°C.

Post-ligation, we performed a bead cleanup using a 0.8x ratio, and then amplified libraries using 8-15 rounds of indexing PCR, depending on the initial DNA volume. We then measured DNA concentration using a Biotium AccuClear plate reader and analyzed fragment length via electrophoresis on a 2% agarose gel. Based on concentration, ibraries were further enriched for 2-7 rounds of PCR in order to normalize DNA concentration between them, then bead-cleaned again using a 0.55x ratio.

Pooling & Capture

Samples were grouped by DNA quality into pools for capture in batches of 8-12. We used a modified version of Arbor Biosciences myBaits Hybridization Capture for Targeted NGS Version 5.02 protocol for capture. To each pool, we added 15 uL of chicken Hybloc (Applied Genetics Laboratories) and 5 uL of universal blocking oligomers, then used a vacuum centrifuge to concentrate the pools from their initial volumes to approximately 8uL, spiking water in to any pools below 8uL to increase their volume. We then added 5uL of Arbor's recommended blockers mix to 7uL of each library, and incubated this library mixture at 95°C for five minutes. We incubated our library mix and hybridization mix for another 5 minutes at 65°C, the temperature recommended for target-bait divergence of less than 10%. We then added the hybridization mix to the library mix and incubated this capture mixture at 95°C for 16 hours.

We followed Arbor's recommended protocol for binding the hybridized libraries to biotinylated magnetic beads, transferring each capture reaction to bead aliquots heated to 65°C, incubating the

mixture for 2.5 minutes, gently agitating the mixture, and incubating again for 2.5 minutes. This was followed by four washes with wash buffer kept at 65°C, again incubating the mixture for 2.5 minutes, agitating, incubating again for 2.5 minutes, and then removing the buffer. After washing, we resuspended the washed beads in 60uL of Arbor's Buffer E and split this mixture into three sets of tubes, each filled with 20uL of the pooled library mixture.

Library amplification proceeded immediately following the wash, using one set of 20uL pooled mixtures to determine the appropriate number of PCR cycles necessary to recover enough DNA for sequencing. Our PCR master mix followed Arbor's recommendations except for the substitution of 5 uL of 10x Illumina primer cocktail for arbor's recommended P5 & P7 library primers. Depending on the initial DNA mass present in the library pools before capture, we performed either 13 or 14 cycles of PCR using arbor's recommended thermal program, then performed a 0.6x bead cleaning using low-ratio SPRI beads, resuspended the libraries in 20 uL EB, and assessed DNA concentration via qBit. Based on the qBit concentrations, we then performed 13-15 cycles of PCR on the second and third sets of tubes, then pooled these two tube sets together, performing another 0.6x SPRI bead cleanup and re-suspending in 10uL EB. This cleaned library was combined with the first set of cleaned libraries to produce ~29 uL final captured library, and we assessed the DNA concentration of these libraries via qBit before performing fragment length and concentration analysis via an Aligent Technologies Bioanalyzer capillary electrophoresis machine.

We sent 15uL of each pooled capture to the California institute for Quantitative biosciences (QB3) for sequencing using the Illumina platform, sequencing 20Gbp per sample pool.

Probe design

We designed a 324-gene probeset using 283 RELEC loci (Karin et al. 2020) and 41 legacy markers chosen for overlap with previous studies of squamate phylogeny (Townsend et al. 2011; Zheng and Wiens 2016), with a total target size of 930,020 nucleotides and an average GC content of 45.8%. Bait design was carried out by Arbor Scientific; after finding no repetitive sequences within the targets and substituting T for any strings of Ns 1-10 nucleotides long and removing any ambiguities in bases, they produced a set of 22,009 baits using 80 nucleotide probes and 2x tiling, using one probe every ~40 nucleotides. A total of 19,592 baits passed stringent BLAST filtration, having a GC content of 30-55% with no BLAST hits to a *Dibamus bourreti* mitogenome (Wang et al. 2021).

Data Processing

We processed, assembled, aligned, and trimmed raw sequence data using the FrogCap R pipeline (Hutter et al. 2022). The pipeline removes low-quality reads and adapter contamination using fastp (Chen et al. 2018), then removes bacterial contamination and merges reads using bbsplit and bbmerge (Bushnell et al. 2017). It uses dipSPAdes (Safonova et al. 2015) to create de novo contigs and uses BLAST (Altschul et al. 1990) to match contigs to target genes, which are then aligned using MAFFT (Katoh et al. 2002) and trimmed using trimAl (Capella-Gutiérrez et al. 2009). Sequencing depth and coverage form deduped reads was assessed using MosDepth (Pedersen and Quinlan 2018).

Phylogenetic Inference

We performed a Maximum Likelihood (ML) phylogenetic analysis of a concatenated data set including all markers in iqTree (Minh et al. 2020) with automated nucleotide model selection for each partition. Data were partitioned by target loci. Additionally, we produced a summary multispecies coalescent (MSC) species tree using ASTRAL-PRO (Sayyari and Mirarab 2016), with ML gene tree estimates for 313 loci obtained using RAxML (Stamatakis 2006).

Morphology

I measured 27 morphological characters for 36 Wallacean *Dibamus* specimens, augmenting the dataset with measurements from five *D. oetamai* specimens reported by Prasetyo et al. (2025). Snout-vent length (SVL) was taken with a string to the nearest millimeter. All other linear measurements were taken with Mitutoyo digital calipers to the nearest hundredth of a millimeter. Linear measurements taken are as follows; tail length from vent to tip (TL); width at midbody (BW); width of base of tail (TW), length of hindlimb (measured from the hip) (HLL), head depth (HD), head length from occiput to tip of snout (HL2), head length from back of posteriormost supralabial scale to snout (HL2), width of head at jaw angle (HW), distance from anterior point of eye to back of nare (EN), distance from anterior point of eye to snout (ES), Distance between nares (IN), distance between orbits measured across thinnest portion of the skull (IOD), length of frontonasal scale (FNSL), width of frontonasal scale (FNSW), length of frontal scale (FSL), width of frontal scale (FSW), width of interparietal scale (IPW), and width of nuchal scale (NSW). I took 6 scale counts; number of scale rows encircling the body just behind the head (ATSR); number of scale rows encircling the body at midbody (MBSR);

number of scales encircling the body just before the vent (PTSR); number of ventral scales from the mental to the vent (VEN); number of subcaudal scales (SC), and number of postorbital scales (POS).

For the *D. seramensis* holotype (MCZ 7623), measurements were taken based on photographs of the holotype specimen using imageJ. For 4 *Dibamus* specimens from Gunung Klabat, I used SVL data from the field catalog and counted ventral and subcaudal scales based on photos. These measurements were not included in statistical analyses.

We analyzed the resulting 42-specimen dataset (Supplementary table S2) using Multi-Factor Analysis (MFA), a robust extension of Principal Components Analysis useful for incorporating and weighting different data types, following Grismer (2025). I imputed missing data using the R missMDA package (Husson and Josse 2026) and performed size-correction of the measurements following Grismer et al. (2021), using a single regression coefficient for the entire dataset rather than calculating regression coefficients for a priori populations. I specified six groups of variables for the multi-factor analysis; body and tail measures (TL, BW, TW, HLL), head shape (HD, HL, HL2, HW, EN, ES, IN, IOD), cranial scales (FNSL, FNSW, FSL, FSW, IPW, NSW), body scale rows (ATBSR, MTBSR, PTBSR), body scale counts (SVL, TL), and number of postoculars (POS).

Results

Capture efficiency

At least half of the target sequences were sequenced in all but one *Dibamus* sample (*D. bourreti*, NAP 10923), with a median of 304 target sequences per sample. The aligned data matrix included 313 genes and 270,307 variable sites. For most samples, 90% of targeted base pairs were sequenced between 100-500x.

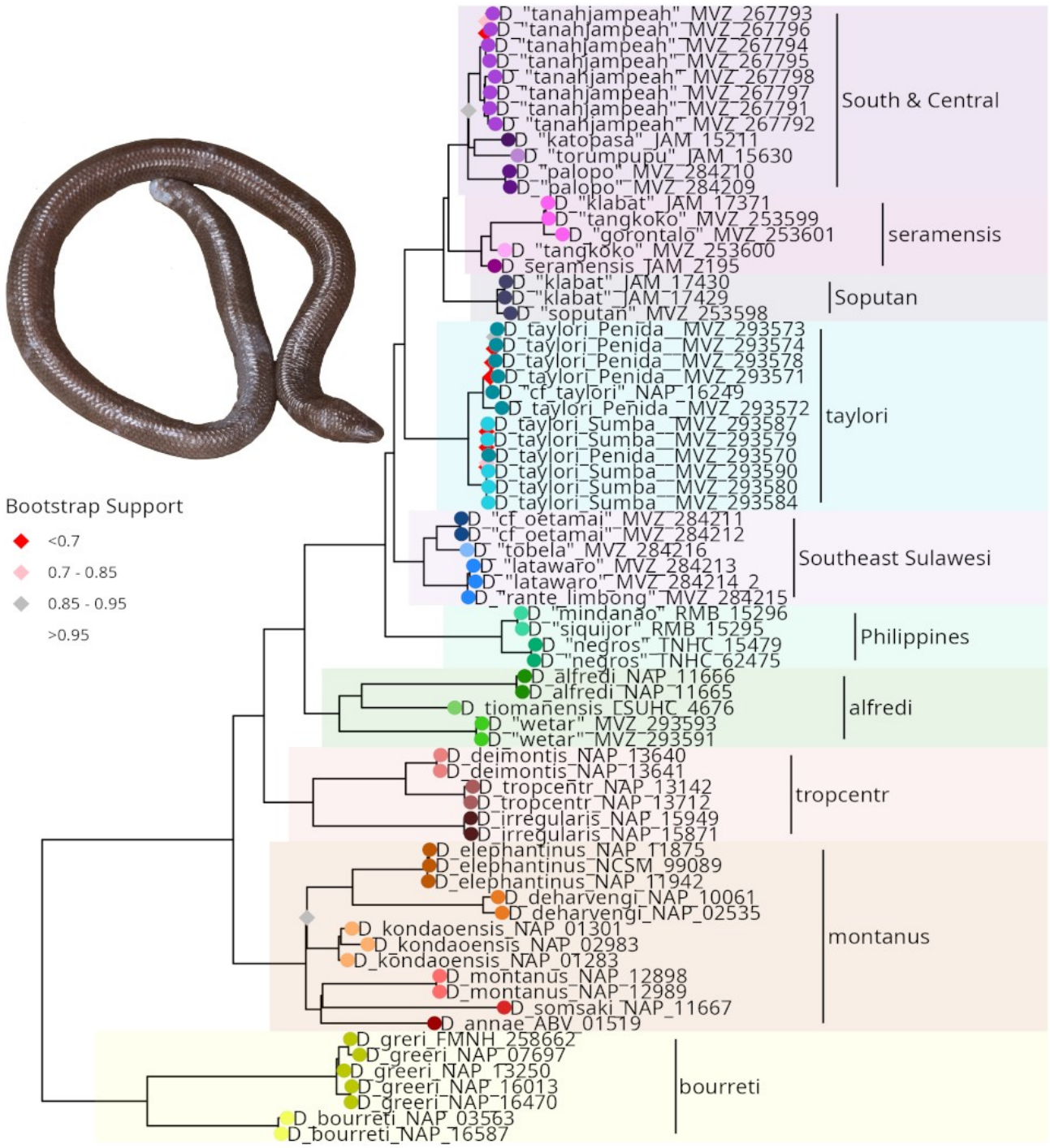

Figure 1: ASTRAL PRO phylogenetic tree for 72 Dibamus *specimens, with major clades labeled and nodes colored by binned bootstrap confidence level. Tip colors reflect those in subsequent figures.*

Phylogenetic Inference

Our dataset produced a well-resolved ASTRAL species tree with bootstrap support values greater than 95% for all but two nodes representing species-level or higher clades, and resolves all currently recognized non-Wallacean *Dibamus* species as monophyletic clades. Only two nodes subtending named species-level clades were resolved with less than 95% confidence; the node subtending the *D. elephantinus + D. deharvengi + D. kondaoensis* clade (89.7% support), and the node subtending the *D. tanahjampeahensis + D. "katopasa" + D "torumpupu"* clade (90.3% support). Nodes with less than 70% bootstrap support occur only within populations of *D. taylori* and *D. "tanahjampeahensis,"* in most cases between syntopic specimens.

Phylogenetic relationships within our phylogeny largely agree with recent work based on smaller genetic datasets (Kliukin et al. 2024*a*, 2025, 2026). We resolve four distinct clades of primarily mainland Southeast Asian *Dibamus* that form successive outgroups to a well-structured but biogeographically complex Wallacean clade. The northernmost *Dibamus* species, *D. bourreti* and *D. greeri*, are deeply diverged from each other but form the "bourreti" clade sister to all other *Dibamus*. The second clade of *Dibamus,* hereafter the "montanus" clade, contains the majority of other species found on the Indochinese peninsula, including a coastal subclade comprising *D. deharvengi* and *D. elephantinus* as sister species closely related to the Island endemic *D. kondaoensis*. This subclade is sister to a clade including the upland *D. annae*, sister to the pair of *D. somsaki* from low-elevation Eastern Thailand and the upland *D. montanus* (Figure 1).

The third branch within *Dibamus* is the "tropcentr" clade, containing species from South-Central Vietnam, the recently described *D. irregularis* and the sister species *D. deimontis* and *D. tropcentr* from Nui Chua Mountain in Ninh thuan Province, which present a probable example of elevational speciation within *Dibamus.*

The fourth *Dibamus* subclade, hereafter the "alfredi" group, contains species from Wallacea and the Malay Peninsula. The first branch in the "alfredi" clade is an unnamed species, originally identified as *D. taylori*, collected from the far eastern end of the lesser Sunda Islands on Wetar Island, representing a remarkable early dispersal of *Dibamus* from the Sunda shelf to nearly the edge of the Sahul Shelf, apparently without any intervening populations remaining extant. Sister to this vagrant

species is a clade containing *D. alfredi* from the central Malay Peninsula and another island endemic, *D. tiomanensis* from Pulau Tioman in the coastal waters on the east of the Malay Peninsula.

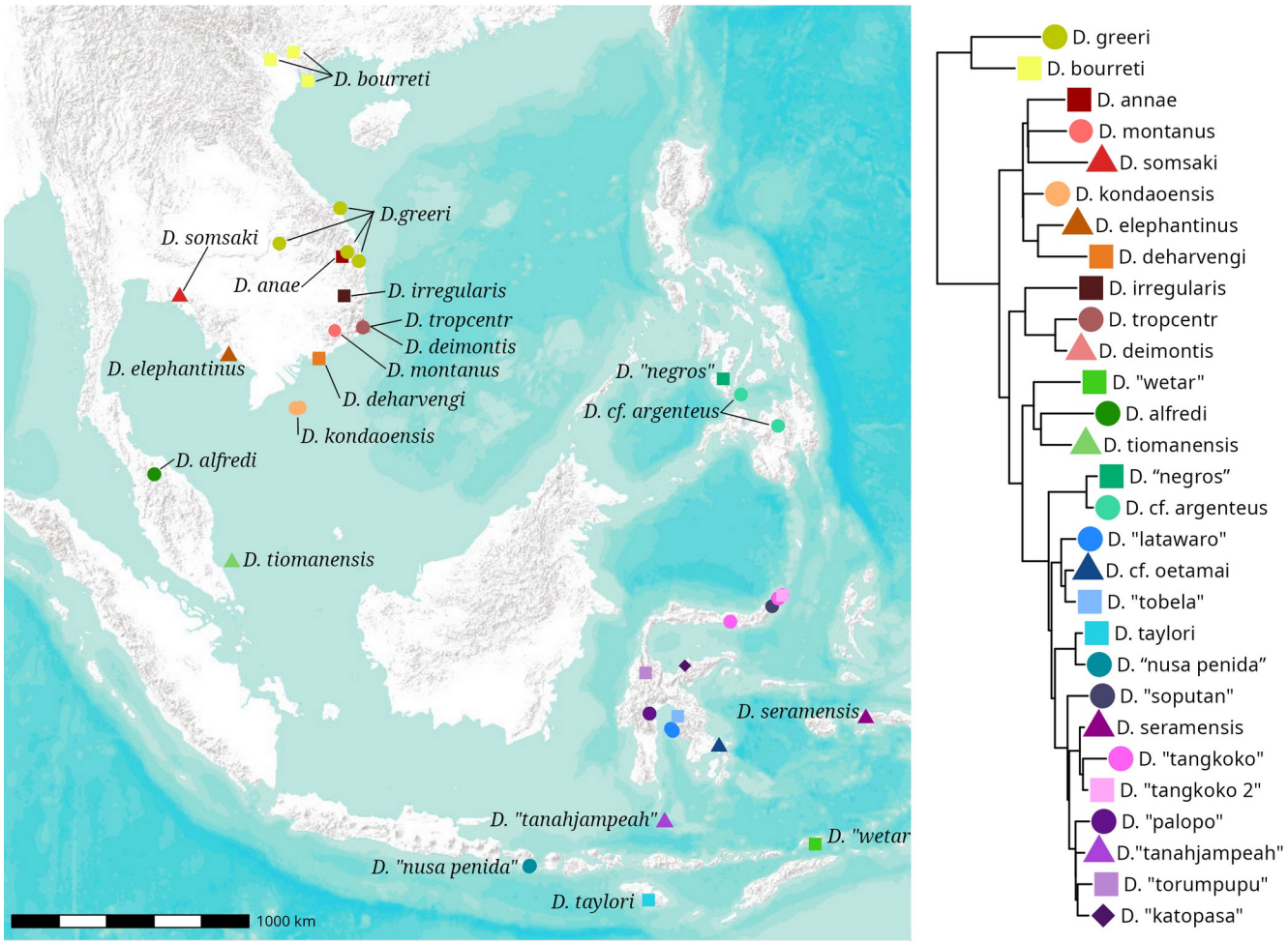

Figure 2: Topographic relief map of Southeast Asia and Wallacea with localities for all Dibamus specimens used in this study. Simplified phylogeny from Figure 1 shown for color reference. Background map data from ESRI.

The major Philippine and Wallacean radiation of *Dibamus* contains specimens assigned to *D. celebensis*, *D. leucurus*, D. *novaeguineae*, *D. oetamai*, *D. taylori*, and *D. seramensis*, but only positively recovers *D. taylori* as a clade. The first branch in this radiation contains all four specimens from the Philippines; one specimen referred to *D. leucurus* collected on Mindanao (KUH 334071 from Misamis Oriental Province), one from Siquijor (KUH 33127), and two colected on Negros (TNHC 64275 and TNHC-GDC 15479). The latter three specimens were originally referred to *D. novaeguineae*. This clade is geographically structured, with the two specimens from Negros recovered

as sisters, and the Mindanao and Siquijor specimens, possibly referable with Taylor's *D. argenteum*, recovered as sisters. These specimens comprise the "Philippines" clade.

Beyond the Philippines, the radiation of Wallacean *Dibamus* is centered on Sulawesi, and structured around its four primary peninsulas. The Southeast Sulawesi clade, branching first, contains two clades from the Southeast Peninsula. One, containing the southernmost specimens as well as one from Tobela on the northwest side of the peninsula, we tentatively associate with the recently described *D. oetamai* from Buton Island. The second clade represents an undescribed species collected at two localities near the Village of Latawaro.

The Lesser Sundas *Dibamus*, *D. taylori*, is the next branch on the phylogeny after the Southeast clade. We recover two geographically concordant clades of *D. taylori* based on our sampling from Sumba (a single population near the type locality for the species) and Nusa Penida.

Two small clades of *Dibamus* associated with Sulawesi's Northern Peninsula constitute the next two branches of the *Dibamus* phylogeny. The "Soputan" clade contains specimens from the low slopes of two mountains near the tip of the Northern Peninsula; one specimen from Gunung Soputan and two from the nearby (more slightly more northerly) Gunung Klabat. The "seramensis" clade, containing five specimens, is geographically puzzling. *Dibamus seramensis*, from Seram Island to the east of Sulawesi, is the most basal branch in this group, followed by a specimen from Tangkoko on the eastern tip of the Northern Peninsula. These two appear somewhat distantly related to a clade including three specimens, from Gorontalo, near the middle of the Northern Peninsula, as well as from Gunung Klabat and Tangkoko National Park.

Finally, a "South & Central" clade of *Dibamus* contains specimens from four localities; Palopo, near the eastern coast of Sulawesi's Central Core, Gunung Katopasa, at the proximal end of the Eastern Peninsula, Gunung Torumpupu in the northwestern portion of the Central Core, and Tanah Jampeah Island, about midway between the southern edge of the Southwest Peninsula of Sulawesi and Flores Island in the Lesser Sundas.

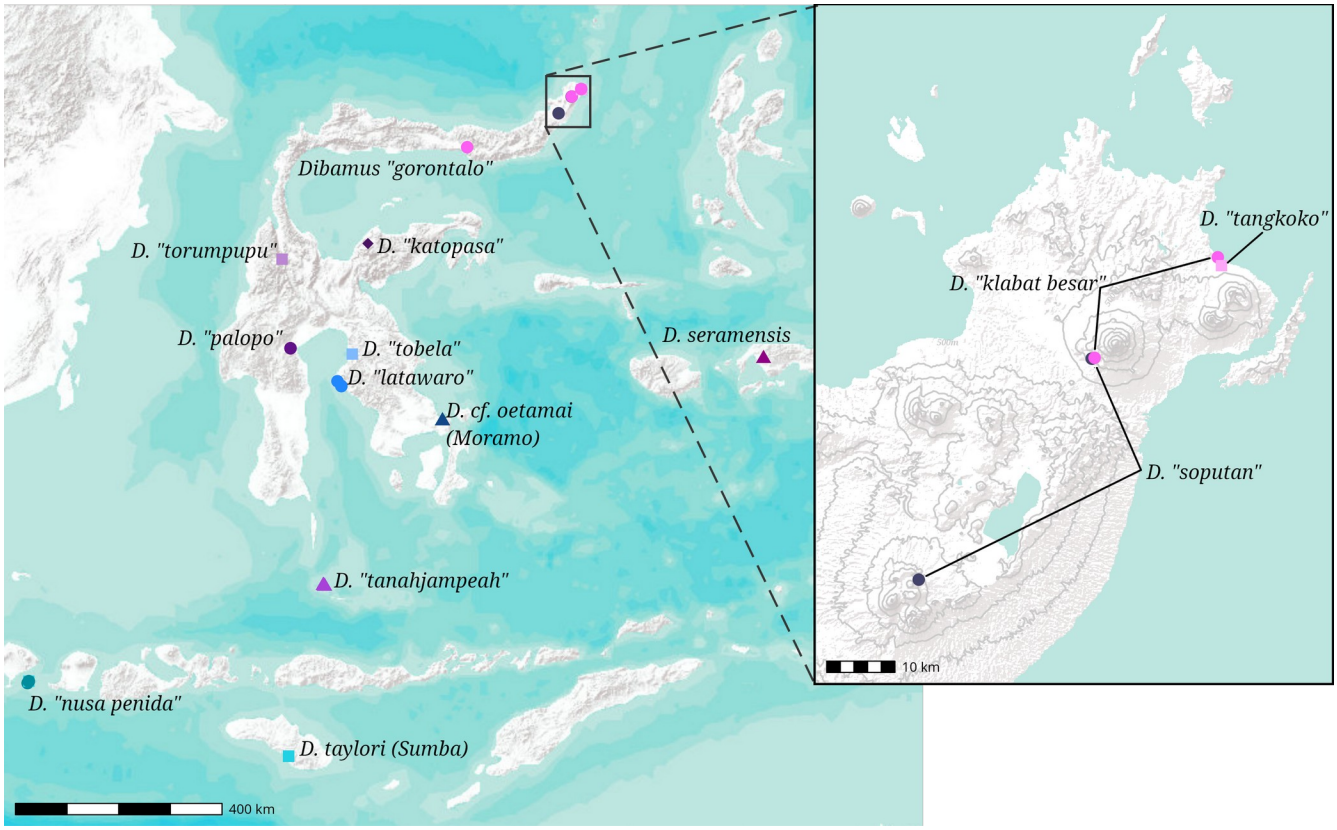

Figure 3: Topographic relief map of Sulawesi and surrounding islands showing locations of specimens within the Sulawesi Dibamus *radiation. Inset shows detail of the Eastern end of Sulawesi's northern Peninsula, visualizing distinct lineages in sympatry at Tangkoko and Gunung Klabat. Contour lines show 250m intervals. Colors follow Figures 1 & 2. Background map data from ESRI.*

Multi-factor Analysis

Multi-Factor analysis results partitioned morphological variation in Wallacean *Dibamus* specimens primarily along three axes. Dimension 1, accounting for 29.72% of variance, correlates strongly and positively with head depth, head length, head width, and body width. *Dibamus seramensis* specimens sort towards the positive extreme in this axis, along with the Gunung Soputan specimen (MVZ 253598) and one specimen from central Sulawesi (AMNH 142976) collected at 4500 ft elevation. At the opposite end of dimension 1, *Dibamus oetamai* specimens group with small specimens from Tanah Jampeah island.

Dimension 2, accounting for 24.2% of variance, correlates very strongly and positively with the

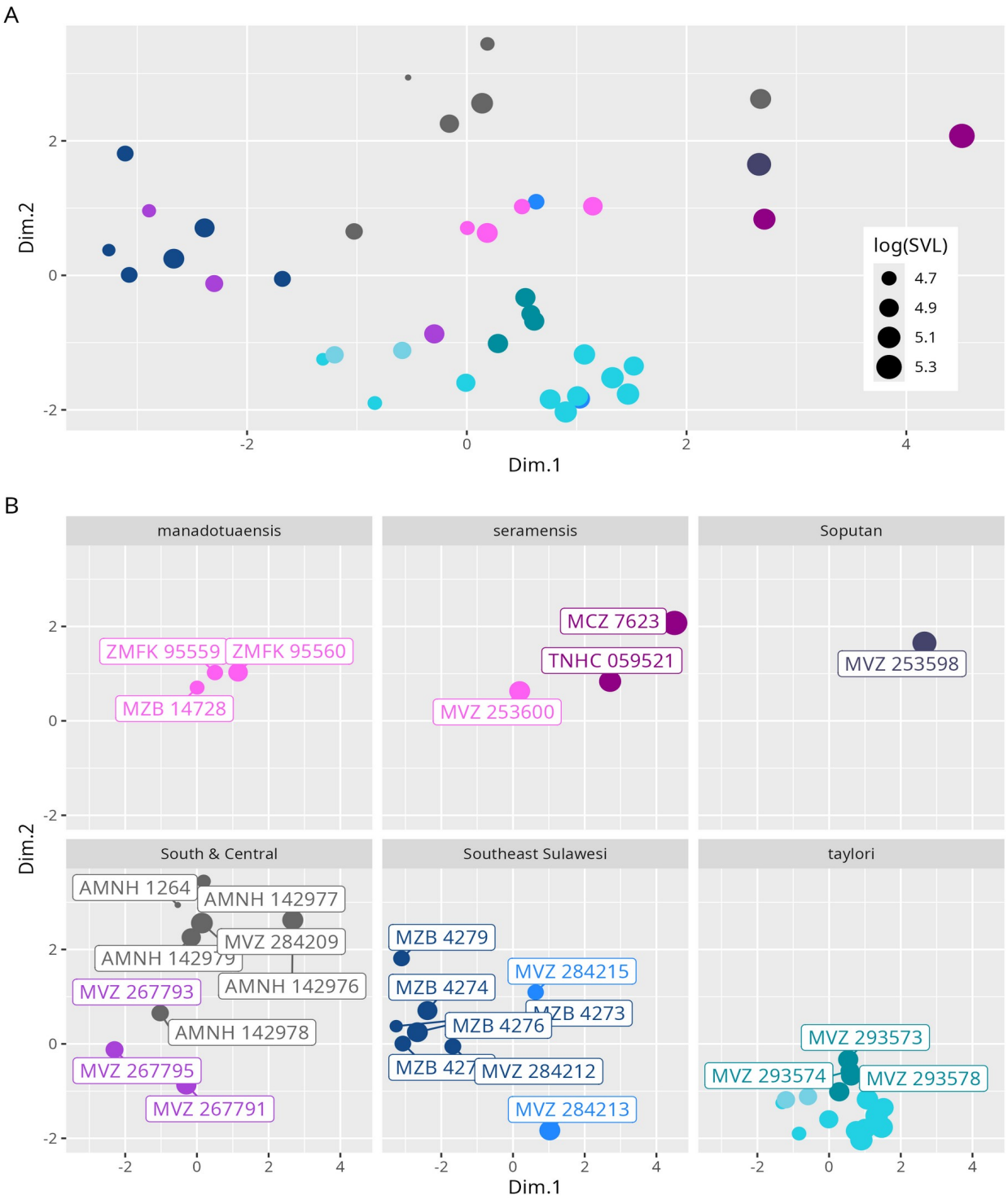

Figure 4: Multi-factor analysis of Wallacean Dibamus *morphological dataset. A) all Wallacean specimens >90mm SVL plotted by MFA Dimensions 1 (29.72% of variance) & 2 (24.2% of variance). Point size corresponds to the natural logarithm of SVL. B) the same data as above, faceted by species group (putative) and with specimen numbers plotted. Colors follow figures 1-3.*

This is the first phylogenetic study to comprehensively sample the family Dibamidae, and includes 6 species not included in previously published studies; (*D. deharvengi*, *D. kondaoensis*, *D. deimontis*, *D. tropcentr*, *D. alfredi*, and *D. taylori*). Our phylogenetic hypothesis for *Dibamus* recovers some key features found in previous studies (Figure 1). First, we find a deep split between the "bourreti" clade and other *Dibamus* species, as recovered by Townsend et al. (2011) and Kliukin et al. (2025) who note Townsend et al's (2011) mis-identification of MVZ 258662 as *D. montanus*. Likewise, we do recover a clade containing *D. tiomanensis* and the Wallacean *Dibamus*, as found by Townsend et al. (2011) and Kliukin et al. (2025), though we find *D. tiomanensis* as part of a clade sister to the greater Wallacean clade, rather than sister to a Philippine specimen as do Kliukin et al. (2025; LSUMZH 9546 from Negros). Among the mainland *Dibamus*, only relationships between species from our "bourreti" and "montanus" clades have been previously investigated. Kliukin et al. (2025) recover *D. montanus* as an outgroup to a clade containing *D. annae* and *D. elephantinus*, and *D. dalaiensis* as sister species. We recover a different branching arrangement, with *D. elephantinus* sister to a clade containing distantly-related *D. annae* and *D. montanus*, but cannot place *D. dalaiensis* in our study due to a lack of available material. As material from more *Dibamus* species becomes available, we hope to expand this hypothesis to include the many species not yet sampled, or likely not yet sampled, by this study.

From an origin in continental Asia, crown dibamids have crossed Wallace's Line and spread into the far east of Wallacea not once, but twice, and at least once also crossed Lydekker's line from Wallacea onto the Sahul shelf. At the same time, several ancient lineages have persisted and diversified in mainland SE Asia and on the Sunda shelf. The most speciose of these, the "montanus" group, likely contains seven species ranging from Thailand to Southern Vietnam, and from near sea level to 1340 m elevation, and the neighboring "tropcentr" clade, while confined to central Vietnam, spans a similar elevational range. The same wide elevational range is likely occupied by the Wallacean *Dibamus*; specimens referred to *D. celebensis* have been reported as high as 1372 m (4500 ft; AMNH R-142976, collected by mammalogist Guy Musser). Clearly, *Dibamus* lizards are quite capable of geographic spread and persistence in varied environments.

The capacity of *Dibamus* species to spread geographically appears to be driven primarily by their proximity to water. The Wallacean *Dibamus* radiation is likely quite young compared to the mainland clades, and yet contains clades ranging from Sulawesi to Seram, and from the Philippines to

the Lesser Sundas. Unsurprisingly, the island of Sulawesi appears to be central to this radiation, though given its geological history, we might more accurately say that the Sulawesi paleo-archipelago has been central to the diversification and dispersal of *Dibamus* throughout Wallacea.

This capacity for long-distance overwater dispersal contrasts with apparently low vagility of individual animals in their natural environment. Somewhat well-resolved structure within populations of *Dibamus* on Tanah Jampeah, Nusa Penida, and Sumba suggests that *Dibamus* populations do not freely interbreed even at very fine spatial scales, as seen in other fossorial ectotherms (Martinez-Solano 2006)

Given its position nested within a clade of specimens from Northern Sulawesi, it seems likely that *D. seramensis* represents a lineage that has dispersed southeast from Sulawesi's Northern Peninsula. While we lack specimens from the Banggai and Sula Islands, we might expect closely related *Dibamus* from these archipelagos, conforming to a stepping-stone model of dispersal. If this is the case, even the Eastern Peninsula of Sulawesi may be home to "seramensis" group *Dibamus*. Alternatively, it may be that *D. seramensis* is the result of a single long-distance dispersal event, likely borne on a mass of soil and vegetation swept from Sulawesi to Seram across the Maluku sea current.

Dibamus seems to have taken a unique path from the Philippines through Wallacea, dominated by North-to-South dispersals that coincide with the Indonesian throughflow (Nugraha and Hall 2018; Lubis et al. 2025). Four North-to-South dispersals are implied by the branching order in our phylogeny. First, from the Philippines to Sulawesi as may be the case in *Draco* (Mcguire et al. 2023) and has occurred in *Limnonectes* (Evans et al. 2003); second, from Sulawesi to the Lesser Sundas, a jump known to have also occurred in *Lamprolepis* (Linkem et al. 2013; Reilly et al. 2025) and several times in *Cyrtodactylus* geckos (Reilly et al. 2023) and soil-dwelling *Trigonopterus* weevils (Tänzler et al. 2014), third, from North Sulawesi to Seram, though incomplete sampling in our study may neglect a connection between Seram and Sulawesi's Eastern or Southeastern Peninsulas as seen in *Draco* (Mcguire et al. 2023), and fourth, from Sulawesi to Tanah Jampeah, as seen in *Cyrtodactylus* (Reilly et al. 2023) and *Varanus* (Arida 2017). Moreover, the branching order of even the mainland *Dibamus* clades follows something of a North-to-South trajectory, matching the consistent Southerly surface currents along the Indochinese Coast in both Northeast and Southwest Monsoons. Though here, the presence of *Dibamus* deep within the interior of the Indochinese Peninsula attests to their ability to spread across land; the broad range of *D. greeri* exemplifies this; there is surprisingly little genetic

diversity within our sample, and specimens from Northern Vietnam and Southern Cambodia are not highly diverged.

Taxonomy

The phylogenetic positions of the Indochinese *Dibamus* specimens used in this study overwhelmingly confirm the validity of recent taxonomic work by Kliukin and colleagues (Kliukin et al. 2023, 2024*b*, *a*, 2025, 2026) indicating that the current understanding of *Dibamus* diversity in this region is largely complete. Conversely, our sampling is woefully incomplete with regards to more southerly species on the Sunda shelf. We have yet to incorporate data from *D. booliati*, and *D. floweri* from the Malay Peninsula, *D. tebal*, *D. dezwaani* and *D. leucurus* from Sumatra and surrounding islands, *D. nicobaricum* from the Nicobar islands, and *D. ingeri* and *D. vorisi* from Borneo. The most parsimonious hypothesis for their affinities lies with the "alfredi" group, but the matter is far from settled.

The situation among the Wallacean *Dibamus* is just the opposite. The concept of *D. taylori* as a single species found throughout the lesser Sundas is untenable, as specimens from Wetar belong to a different clade altogether, and the Nusa Penida *Dibamus* appear to be a sister species to *D. taylori*. Pending morphological examination of other members of the "seramensis" group, the concept of *D. seramensis* may need to be expanded to include *Dibamus* specimens from mainland Sulawesi. Beyond these two recognized species, the taxonomic status of Wallacean *Dibamus* are here either untenable or unaddressed. The positions of *D. taylori* and *D. seramensis* within a complex radiation of *Dibamus* on Sulawesi renders the concept of a Sulawesi-wide *D. celebensis* polyphyletic, and the absence of a clear type locality for *D. celebensis* (the type locality is "Celebes"; Greer 1985) means that assigning *D. celebensis* to any Sulawesi lineage included in our study is not possible at present. Yet another North Sulawesi lineage, *D. manadotuaensis*, is not yet clearly referable to any clade on our tree because of the presence of both "seramensis"-group and "soputan"-group *Dibamus* in the immediate vicinity. A lack of material from *D. oetamai* in our study means that we cannot clearly refer *D. oetamai* to the "Southeast" clade, though we suspect that *D. oetamai* is closely related to our two specimens from the nearby Moramo locality.

Finally, the placement of *D. novaeguineae* is here unaddressed. Specimens referred to *D. novaeguineae* have been collected throughout Wallacea, New Guinea, and the Philippines, and many of

these may have been referred to *D. novaeguineae* based on characters that are known to be variable within populations or overlap between species, such as the number of postocular scales. In practice, a set of characters that unambiguously distinguishes *D. celebensis* from *D. novaeguineae* does not exist, so the identifications of Wallacean *Dibamus* should be treated with some skepticism. While the similarity between *D. novaeguineae* and *D. celebensis* my indicate a close relationship, the presence of a "alfredi"-group *Dibamus* on Wetar suggests that easterly *Dibamus* are not necessarily referable to the main Wallacean radiation.

Candidate Species

Among the taxonomic and phylogenetic uncertainty evident for Wallacean dibamids, we have recovered several lineages that may warrant recognition as species. Disregarding the attribution of our lineages to known species, there appear to be from 9 to 16 Wallacean species-level lineages in our phylogeny, several of which we know to be morphologically distinctive.

From the Southeast Sulawesi group, we have inspected one female from the "Rante Limbong" population from Southeast Sulawesi (MVZ 284215), which possesses a scalation character unique among dibamids: its frontonasal scale is fused to its frontal scale, creating a single scale shaped roughly like a figure "8". However, this character is not shared with MVZ 284213, a closely related male from Latawaro. The two specimens are well separated in MFA space, which may be due to a combination of changes related to size and sex. These Rante Limbong and Latawaro specimens are fairly well-diverged from specimens found further South at Air Tejun Moramo, tentatively referred to *D. oetamai* (MVZ 284211 and MVZ 284212), which they are well-separated from morphologically (Figure 2B). A specimen from Desa Tobela, North of Latawaro and Rante Limbong, is sister to these *D. cf. oetamai* specimens, but has not been morphologically examined (Figure 2). It seems that there are at least two and perhaps three *Dibamus* species on the Southeast Peninsula: (1) *D. cf. Oetamai,* (2) an unnamed species from Rante Limbong and Latawaro, and (3) a putatively separate species from Tobela (Table 1).

It seems likely that populations of *Dibamus taylori* from Sumba and Nusa Penida represent distinct species. In addition to a complete genetic divergence between the populations, Nusa Penida specimens score higher along MFA dimension 2 than Sumba, Flores and Komodo specimens (Figure 4B). As *D. taylori*'s type locality is Sumba, the Nusa Penida population may warrant recognition as a new species (Table 1).

The two lineages of *Dibamus* inhabiting Gunung Klabat are certainly different species, though none of the specimens have been fully measured. Based on photos taken in the field, the Seramensis-group specimen (JAM 17371) appears morphologically referable to *D. seramensis*, with 267 ventral scales, light overall coloration, and parafrontal scales separating the frontal from the oculars (these make up the anteriormost of the four scales bordering the posterior edge of the ocular). The other three specimens are all darker in color, have 214-225 ventral scales, and have a wide frontal scale that directly contacts the oculars, which are bordered posteriorly by two or three scales. These specimens seem morphologically similar to Greer's updated description of *D. celebensis*, though they lack a clear lateral labial suture. The absence of parafrontal scales separates them from *D. manadotuaensis*.

Interestingly, the presence of a parafrontal scale, larger body size, and higher ventral scale count are also features of these specimens' closest relative, MVZ 253598 from Gunung Soputan, just West of Gunung Klabat. In MFA space, this Gunung Soputan specimen is conspicuously similar to *D. seramensis.* This raises the possibility that there has been trait displacement between the two species found on Gunung Klabat. The larger, seramensis-group species perhaps displacing the Soputan-group species into a smaller, celebensis-like morphotype. These smaller Soputan-group Klabat specimens were found both in secondary forest and in cleared agricultural areas, while the single, large seramensis-group specimen was found in primary forest, only 40m away from the collecting site of a Soputan-group specimen.

Despite their morphological differences, Gunung Soputan and Gunung Klabat specimens are quite closely related, and we tentatively refer them to a single species (Table 1). Character displacement between sympatric fossorial lizards has been well-documented in other taxa (Huey 1974), and may be an unrecognized phenomenon in *Dibamus* due to their rarity in collections.

Within the "seramensis" clade, there are likely three species; *D. seramensis* proper, the Gunung Klabat "seramensis" species, which also appears at Tangkoko, just east of Gunung Klabat, and a second Tangkoko species, represented by a single specimen, MVZ 253600 (Table 1). This second Tangkoko specimen is strikingly similar in measurements to *D. manadotuaensis*, suggesting that *D. manadotuaensis* could belong to this "seramensis" clade (Figure 4). The single *Dibamus* specimen from Gorontalo, MVZ 253601, is sister to the Gunung Klabat "seramensis" species, though on a somewhat long branch. In the absence of physical specimens to examine, we tentatively consider it conspecific with the Gunung Klabat *D. "seramensis*."

Within the South & Central Sulawesi clade, four geographically distant lineages are separated by fairly deep divergences and are likely each a separate species (Table 1). Only two of these have been examined morphologically. One specimen from Palopo (MVZ 284209) groups morphologically with some other Central Sulawesi specimens not included in our phylogeny, AMNH 142977 from Sungai Sadaunta near Lore Lindu and AMNH 142979 from Sungai Tolewonu near the North coast of the Central Core (Figure 4B). Both of these specimens are from at least 170km north of Palopo, much closer to Gunung Torompupu, home to a second putative species of *Dibamus* for which only tissue is available. The third putative species from Gunung Katopasa is accessioned at the MZB and has not yet been examined.

All specimens from Tanah Jampeah Island appear to be part of a single species, collected both at near sea level close to the village of Maminasa (MVZ 267791 & 267792) and at 501 m on the Telkom Tower road (all other specimens). These appear polymorphic, perhaps because of allometric changes in scalation. Large specimens of *Dibamus* from Tanah Jampeah (e.g. MVZ 267791) possess a unique preocular scale that separates the ocular from the prefrontal. In one of these specimens (MVZ 267795), these scales appear to be developing beneath a large frontonasal that spans from ocular to ocular, which is the condition seen in smaller specimens from Tanah Jampeah. MVZ 267791 and 267795 are from different localities on the island (Maminasa and Telkom Tower road, respectively) and come from two different weakly divergent clades within this Tanah Jampeah lineage, suggesting that they represent different size classes of a single species found on the island.

Specimens from Central Sulawesi not included in our phylogeny, may represent several species. AMNH 142978, 142979, and 142980 are all from a lowland locality (Sungai Tolewonu) and the large specimens (142978, 142979) have longer tails and fewer midbody scale rows than AMNH 142977 (Sungai Sadaunta, 2500ft), and AMNH 142976 (Sungai Tokararu) from highland localities. At both localities, male and female specimens are well-separated in morphospace.

Table 1. Preliminary assessment of species diversity within Wallacean and Philippines *Dibamus*, including 15 candidate species.

Species group	**Species**	**Referred specimens**	**Comments**
"alfredi"	*Dibamus "wetar"*	MVZ 293593 MVZ 293591	
Philippines	*Dibamus "negros"*	TNHC 15479 TNHC 62475	These specimens may currently be referred to *D. leucurus*, but it is unlikely that they are in fact conspecific with any Sumatran *Dibamus*.
	Dibamus cf. argenteus	RMB 15295 RMB 15296	*D. argenteus* is currently recognized as a junior synonym of *D. lecurus*.
Southeast Sulawesi	*Dibamus cf. oetamai*	MVZ 284212 MVZ 284211 MVZ 284216	MVZ 284216 may represent a separate but closely related species.
	Dibamus "latawaro"	MVZ 284213 MVZ 284214 MVZ 284215	
"taylori"	*Dibamus taylori* Greer 1985	MVZ 293570 MVZ 293579 MVZ 293580 MVZ 293484 MVS 293587 MVZ 293590	Specimens referred here are from the same island (Sumba) as the *D. talyori* type.
	Dibamus "nusa penida"	MVZ 293571 MVZ293572 MVZ 293573 MVZ 293574 MVZ 293578 NAP 16249	
"Soputan"	*Dibamus "soputan"*	MVZ 253598 JAM 17370 JAM 17429 JAM 17430	Treated as one polymorphic species, possibly two species.
"seramensis"	*Dibamus seramensis*	JAM 2195	
	Dibamus "tangkoko"	MVZ 253600	

	Dibamus "klabat besar"	MVZ 253599 MVZ 253601 JAM 17371	
"South & Central"	*Dibamus "palopo"*	MVZ 284209 MVZ 284210	
	Dibamus "torumpupu"	JAM 15630	No physical specimen
	Dibamus "katopasa"	JAM 15211	
	Dibamus "tanahjampeah"	MVZ 267791 MVZ 267792 MVZ 267793 MVZ 267794 MVZ 267795 MVZ 267796 MVZ 267797 MVZ 267798	

Conclusions

Here, we resolve a robust phylogeny for *Dibamus*, including definite representatives of sixteen named species and many specimens of uncertain affinity that may support the recognition of several new species. The phylogenetic and biogeographic histories of *Dibamus* are more complex than previously appreciated, with four ancient, low-diversity clades inhabiting continental Asia and two distantly related lineages having crossed into Wallacea. The diversity within these continental *Dibamus* clades is now well-understood, but the species diversity of *Dibamus* in Wallacea is currently underestimated, with a large number of potentially new species having recently diverged within a single clade that originated in the Philippines and spread southward to the Sulawesi paleo-archipelago, lesser Sundas, and Seram, and likely to New Guinea as well. This Wallacean *Dibamus* radiation is morphologically conservative, having few clearly useful distinguishing characters between clades or species, of which there may be more than a dozen.

Outstanding questions in the evolutionary history of Dibamidae will require better spatial and taxonomic sampling, which can be achieved by incorporating morphological and genetic data from formalin-fixed museum specimens. The position of *Anelytropsis*, timing of divergences within Dibamidae, and phylogenetic placement of the group within squamata are remain unclear.

Chapter 3: Scalation Morphometrics

Abstract

From the heads of sunflowers to insect wings to the scales on the heads of reptiles, discrete, two-dimensional patterns abound in the natural world. While some of these patterns, like plant phyllotaxis and the regular scalation of snake bodies, are created by repetition following a simple algorithm, others are far more complex and idiosyncratic, containing irregular elements that nonetheless have clear homology relationships between species. Despite the importance of these patterns in species delimitation and taxonomy researchers have not developed holistic methods for comparing them.

Here, I present such a method, developed to systematically compare reptile scalation patterns (often referred to as "pholidosis,") but applicable in general to surface patterns constructed from units with clear homology relationships. Rather than treating scale numbers and scale adjacency relationships as discrete characters, I draw inspiration from the techniques of geometric morphometrics and anatomical network analysis to analyze the pattern of pholidosis as a network spread across the body of the animal. I describe a simple method for producing these networks and then describe and implement an algorithm that provides an edit-distance measure between networks, allowing for comparative analysis of their topologies–a "topometric" tool. Using a group of fossorial lizards (family Dibamidae) as a study system, I demonstrate the utility of these techniques in understanding the evolution of scalation patterns in a comparative context, investigating the relationships between pholidosis patterns, phylogeny, morphology, ecology and biogeography in this poorly-understood clade. These techniques are implemented in an R package, "pholidosis."

Introduction

Patterns in the bodies of organisms arise through structured development, and the characters that we recognize as differing between species differ because of changes to that developmental pattern. Many characters differ continuously (e.g. length, mass, color), while others differ discretely (count, presence/absence). Some patterns of discrete characters, though, differ between species in complex

ways. The scintillating patterns of vein-bordered domains in insect wings and the patterns of scales on the heads of snakes and lizards are made up of discrete units, which in some cases have clear homology relationships between species. These discrete units produce a complex pattern via their presences, absences, and adjacency relationships.

Though these pattens can be distinctive, taxonomically and biomechanically important (Miralles et al. 2011; Salcedo et al. 2019),workers lack a framework by which to compare these topologies at a fine scale. Herpetologists rarely analyze suites of scalation characters, and even studies of insect wing patterning that incorporate network models do so largely to measure coarse network characteristics rather than to investigate fine-scale topological similarity (Hoffmann et al. 2018; Salcedo et al. 2019). To move beyond description and coarse analysis of these patterns into the study of the tempo and mode of their evolution, researchers require tools to directly and consistently compare their topologies.

Pholidosis

Squamates are defined by their scales; these keratinous structures sheath their bodies and serve as their main point of contact with their environment. Scales are diverse in color, shape, size, texture, and of course, function (Chang et al. 2009). Studies of the ontogeny (Alibardi 1997; Tzika et al. 2023), histology (Alibardi 1997; Dujsebayeva et al. 2021), and microstructure (Alibardi 1997; Comanns et al. 2011; Dujsebayeva et al. 2021) of squamate scales demonstrate how they have been adapted to serve the diverse lifestyles of the squamate radiation, but comparatively little attention has been paid to the inter-relationships of scales on the body surface. This is despite the importance of those inter-relationships to our concepts of squamate species; a discussion of scalation patterns, often termed called “squamation” (e.g. Greer 1985; Das and Yaakob 2003), “scutellation” (Leviton et al. 2011) or “pholidosis” (Scherz et al. 2017; Barros-Filho et al. 2019) is a crucial component in descriptions of most new squamate species (e.g. Da Silva and Ávila-Pires 2013; Liu et al. 2020, Kliukin et al. 2024b).

Many squamate species are diagnosed based on scale counts, sizes, or spatial relationships, but little effort has gone into studying the evolution of these characters; an important oversight given that homoplasy in such relied-on characters could lead to much taxonomic confusion.

As a key characteristic of squamate reptiles, keratinous scales are ubiquitous in lizards and snakes, and are often considered in two separate categories; head scales and body scales. In some taxa,

including the majority of snakes, the numbers and spatial arrangements of body scales follow a clear and consistent pattern, organized by the interaction of reaction-diffusion systems and dorsal-ventral positional information (Tzika et al. 2023). Consequently, herpetologists describe patterns of snake body scalation via just a few characters. In some cases, only two characters are necessary; the number of scale rows at mid-body, and the number of ventral scales between the head and the cloaca (Liu et al. 2020). Conversely, some lizard taxa sport a dazzling array of scale morphologies and interrelationships, with different types scales specialized to their backs, bellies, legs, and tails (Da Silva and Ávila-Pires 2013). Head scalation is rarely simple, and the ontogenetic patterning of these scales appears to follow a different logic than body scales (Tzika et al. 2023).

Though heavily relied on for identification, pholidosis patterns are not necessarily fixed for species, and may vary between populations (Gans 1971; Barros-Filho et al. 2019), between individuals (Gans 1971; Greer 1985; Darevsky 1992; Barros-Filho et al. 2019), or even over the life of an individual (Tomović et al. 2008). In many cases, pholidosis patterns involve both large, placoid scales and small granular scales (Da Silva and Ávila-Pires 2013), and the arrangement and homologies of small, granular scales may be impossible to determine. For this study, I focus on large, placoid scales in the study of pholidosis, though the incorporation of fields, patches, or series of scales into a pholidosis network is possible with the methods introduced herein.

Properties of surface networks

Biological surface networks represent patterns of adjacent units on some surface of an organism. In pholidosis networks, scales are the units, represented in a network by vertices which are connected to adjacent vertices (scales) by edges. Networks of pholidosis are not necessarily unique from other surface networks, but for the sake of simplicity they will be the focus of this paper.

Surface networks have two critical properties:

First, because the units themselves exist on a two-dimensional body surface, surface networks are necessarily planar, meaning that the network can be embedded in a plane such that its edges do not cross (Diestel 2017). Planarity introduces a constraint into the measurement of edit distances between networks, because it is not possible to arbitrarily link two vertices in a planar network while maintaining its planarity. For instance, the right and left ocular scales of a snake cannot be arbitrarily

linked with an edge, because this implies that the right and left eyes have been made adjacent through some untenable folding of the head.

Second, vertices can be considered homologous between networks. In pholidosis networks, this is based on scale homology. Though scale homologies are not universally implied by scale nomenclature in squamates (e.g. Greer 1985), the assumption of homology between scales or groups of scales in closely-related species of squamates is implied by the consistent use of these characters in taxonomic literature. Just as in other morphological studies, the analyses of surface networks are only as trustworthy as the homology assumptions that underlie them.

Beyond these necessary properties, the networks studied herein are constructed with two additional topological constraints that support the ease of computation of edit distances between networks. First, the networks are undirected, meaning that their edges do not have any directionality; edges simply link pairs of units rather than implying an ordered relationship between them. Second, the networks are fully triangulated (better referred to as maximally planar); that is, that the addition of any edge to the graph would result in a nonplanar graph. Because of this second constraint, the networks can be best thought of as existing on a surface of a sphere rather than on a flat plane, as the networks cannot have any "perimeter." The network therefore defines the edges and vertices of a polyhedron with all triangular faces. For a regular polyhedron, the number of edges, E, is equal to plus the number of faces, F, Plus the number of vertices, V, plus 2 ($E = F+V+2$). In a triangulated polyhedron, there are tree faces for every two edges ($E = 3/2\ F$) (Bowen & Fisk, 1967). Therefore, the networks described herein have a simple relationship between the number of edges and the number of faces: $E = 3V – 6$.

Finally, there are some inter-network relationships that the methodology described herein cannot account for. In the method of measuring network edit distance used herein, large changes in the position of some subset of scales relative to another subset are equivalent to the complete separation and re-attachment of these sub-networks (Supplementary figure S1). Step-wise movement of the subsets relative to each other is not explicitly modeled via the algorithm presented here, as no such movement is required to calculate edit distances among the exemplar dataset of dibamid lizards.

The properties of squamate scales can be embedded in these networks as attributes of edges or vertices. In this study, I demonstrate the use of edge weight to denote fusion or near-fusion of scales, but in principle scalation networks can embed many more types of data; for edges, the length of the shared scale perimeter or the presence and directionality of overlap; for scales, size, shape, color,

surface texture, ossification, innervation, special sensory or exocrine function, etc. The challenge of analysis of any character lies in implementing a reasonable model of character state transition. In the methods described and implemented herein, only topological characters (presence, adjacency, and fusions) are considered.

Dibamid lizards (family Dibamidae)

Lizards of the family Dibamidae are among the most poorly understood reptiles on Earth. These wormlike lizards have highly reduced eyes covered by an ocular scale, no forelimbs, and flap-like hindlimbs only in males. They are presumed to be fossorial insectivores, though precious little has been published on their diet or behavior (see Darevsky 1992). Though dibamids resemble amphisbaenians and fossorial skinks, they are not closely related to either group; recent phylogenomic studies place Dibamidae as sister either to Gekkota (Burbrink et al. 2020) or to all other squamates (Singhal et al. 2021). Twenty-seven species of *Dibamus* live in the tropics of Southeast Asia, Indonesia, the Philippines, and New Guinea (Kliukin et al. 2024*a*), while *Anelytropsis papillosus* is found in Mexico. Despite this major geographic rift, *Anelytropsis* appears to be most closely related to a clade of *Dibamus* residing in mainland Southeast Asia. This clade is sister to a "Peninsular-Island" clade of *Dibamus* inhabiting Peninsular Malaysia, Indonesia, the Philippines, and New Guinea (Townsend et al. 2011).

Dibamids are quite conservative in their external morphology (Greer 1985; Quah et al. 2017), differing primarily in scalation patterns, body size, and tail length, and no morphological characters are known that distinguish the *Dibamus* clades recovered by Townsend et al. (2011). Most dibamids exhibit a "rostral shield," of scales fused into a single keratinous structure covering the snout, though the structure is quite variable; different species exhibit many different phenotypes of scale fusion.

From the available data, dibamid pholidosis characters are largely stable at the species level (Greer 1985; Quah et al. 2017), though a lack of known specimens may contribute to an underestimation of variability. There is known intraspecific variation in the number of post-ocular and post-infralabial scales in some species, including the apparently quite variable and widespread *D. taylori (Greer 1985)*, and variability in the degree of fusion between scales of the rostral shield in some species (see D. bourreti, Greer 1985; Darevsky 1992). As this study is primarily a demonstration of

pholidosis network methodologies, rather than a particularly in-depth study of dibamid evolution, I represent each dibamid species using only a single scale configuration.

Methods – Pholidosis Networks

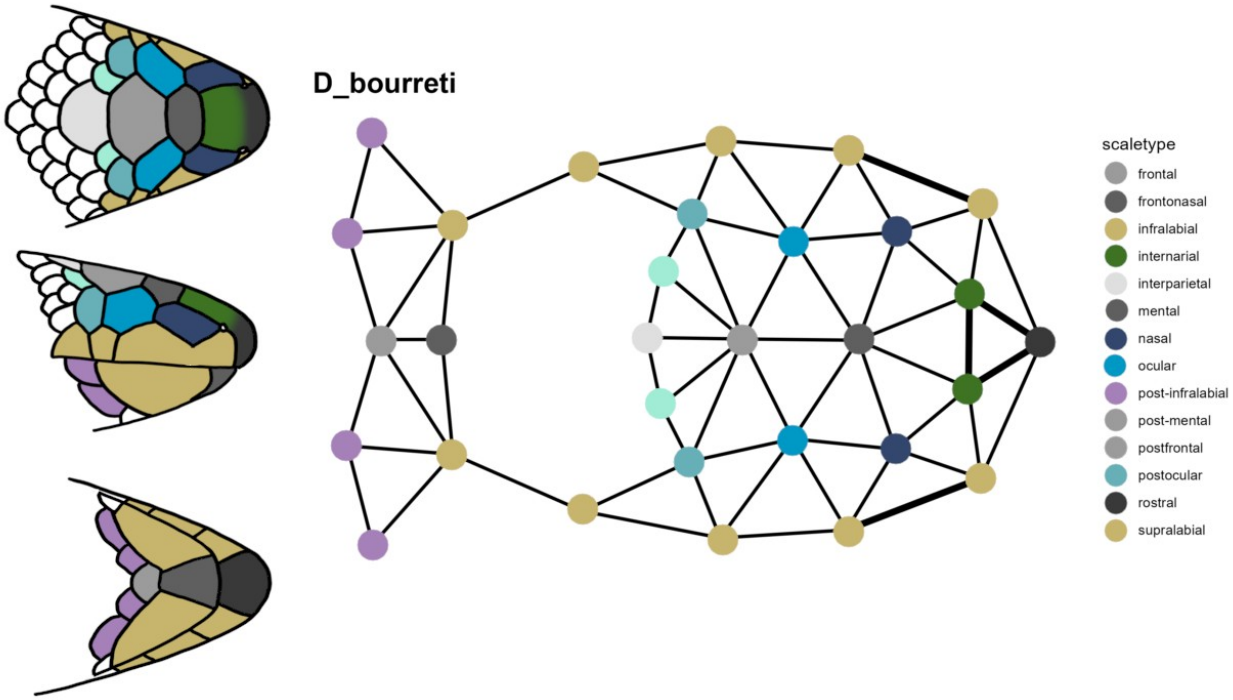

Figure 1: Pholidosis network of Dibamus bourreti. A. Dorsal, lateral, and ventral views of Dibamus bourreti (redrawn after Darevsky (1992) but with the fused supralabials reported by Greer (1985)) with scales colored according to their scale type. B. The resulting pholidosis network; thick lines indicate fusions between scales.

All of the methods described herein are implemented in an R package, *pholidosis*. A brief guide to using the package can be found in an accompanying publication (Krone, 2025). The package can be downloaded at github.com/dibamus/pholidosis.

Encoding the network using an adjacency matrix

The first step in quantifying surface pattern is to encode it in a machine-readable format. I follow (Esteve-Altava and Rasskin-Gutman 2014) in beginning with an adjacency matrix, encoded as follows:

Let *S* be the set of scales involved in a pholidosis pattern. *S* contains *n* scales (s_1, s_2, … s_n). *M* is a square matrix of dimensions n*n, where each scale in S occupies a row and column in this matrix. If a scale s_i borders a scale s_j, the matrix cell corresponding to $[s_i,s_j]$ is filled, producing an edge between those vertices. If the scales are adjacent and fully separated, I denote this with the number 1. Scales that are partially fused together are denoted by the number 2, while scales that are fully fused are denoted by the number 3. On importation into R, this number encodes an "edge weight" in the resulting network, *G*.

I find the best practice when creating these matrices is to include in *S* all possible scales within the pholidosis patterns of all species involved, as well as any structures that these scales abut, such as the eyes, mouth, and remaining body surface. If some pholidosis patterns have two pretemporal scales ($s_{pretemporal1}$, $s_{pretemporal2}$), while others have only one ($s_{pretemporal1}$), the same blank matrix *M*, containing rows and columns for both supratemporal scales, can be filled in to describe the whole set of patterns; matrices corresponding to species lacking the second pretemporal scale will simply have a corresponding blank row and column. This blank row and column produces an "isolated" vertex in the network, which can be removed for analysis.

The *pholidosis* package uses *igraph* (Csárdi and Nepusz 2006) and the *igraph* R package (Csárdi et al. 2025) to represent and perform operations on networks.

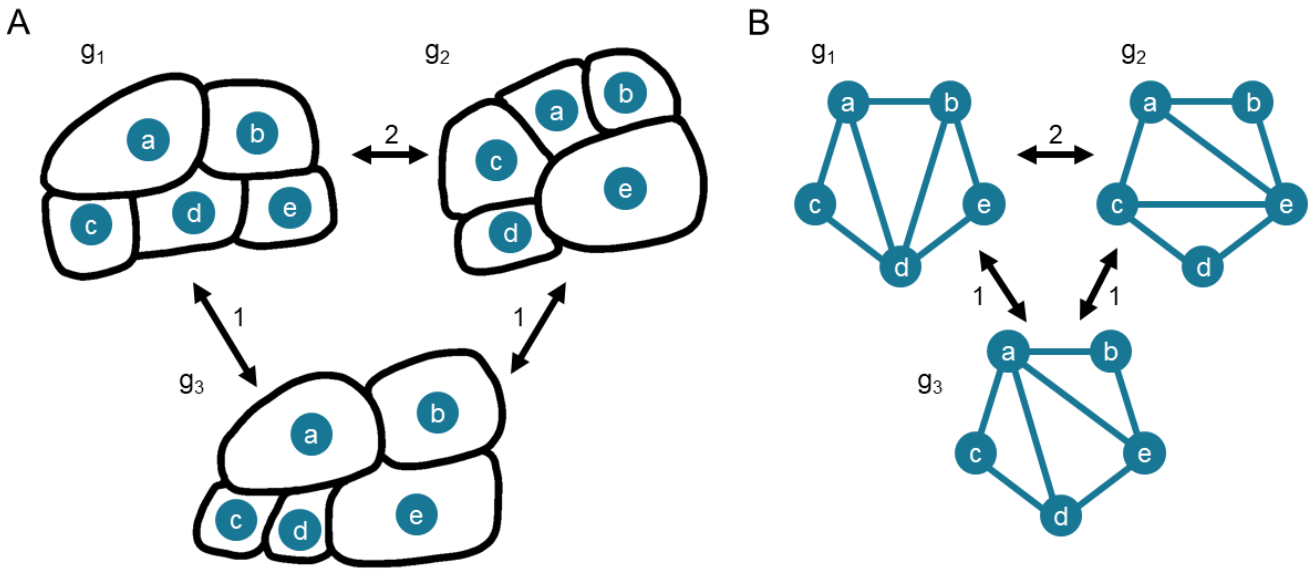

Figure 2: Edit distance accounting for topological changes. A. Three pholidosis networks (g1, g2, and g3) involving five scales, a, b, c, d, & e B. The same pholidosis networks represented as networks rather than physical arrangements of scales, with adjacency relationships between the scales represented by edges. The five vertices, a, b, c, d, & e, form a cycle with n-3 = 2 internal edges. Note that the full graphs also contain sub-cycles of four vertices with a single internal edge. Arrows between the networks track how many changes it takes to transform one network into the other (the edit distance). Though the networks g1, g2, and g3 all contain the same scales/nodes, their topologies are different. To transform g1 into g3, the edge (b, d) is removed and the edge (a, e) added. We can see that this topological change is the result of scale e growing in size such that it contacts scale a, therefore necessitating the loss of a contact between b and d. Therefore, though this change involves the loss of one edge and the gain of one edge, it is counted as a single change, since the gain and loss cannot occur separately in this arrangement, as the intermediate state would violate the need for the network to be fully triangulated. By a similar process, g3 can be transformed into g2; this is the "edge swap." Notice that from g1 to g2, there is a similar loss of two edges (a, d) and (b, d) and a gain of two edges (a, e) and (c, e). The transformation of g1 into g2 may occur through an intermediate scale arrangement (g3), but this intermediate step is not modeled, as we can assume that two changes have occurred between g1 and g2, since the two 5-vertex cycles of nodes surround unmatched edges, of which there are necessarily 2.

Network edit distance

Surface networks can be compared by a measurement of edit distance, the number of changes needed to change one network into another. Any network can be transformed into any other network by the addition and subtraction of edges (and therefore vertices), but simply counting the number of different edges between two networks can produce an overestimate of the number of biological changes needed to change one network into another. For instance, the addition or subtraction of a scale from a pholidosis network should be considered a single biological change, but this can involve the addition or subtraction of any number of edges, depending on what scales the new scale is adjacent to. Likewise, some shifts in the adjacency relationships of scales shared by both networks involve both the addition and subtraction of edges (Figure 2), though a biological interpretation implies a simpler process, in which a single change, for instance, the enlargement of a scale, necessitates cascading effects.

The following method is sufficient to account for these cases in the dibamid dataset, producing an estimate for the minimum number of changes in vertices, edges, and edge weights to transform one scalation network into another. It is implemented in the *pholidosis* package by the *pnet_distance* and *graph_edit_distance* functions.

The *graph_edit_distance* function measures the number of topological changes needed to transform one network, *G*, into a second network, *H*.

In order to tabulate and keep track of the necessary changes to transform one graph into another, we must create descriptive tables of edge properties for the graphs *G* and *H*. For each edge in *G*, we assess the following properties: 1) the names of the vertices that the edge connects; 2) matched edges: whether an edge in *H* connects two vertices with the same names; 3) alternate edge paths: if no edge in *H* connects vertices of the same names, the shortest path (if any exists) between two vertices with the same names in *H* via edges unique to *H*; 4) whether, in those alternate edge paths, all intermediate vertices are unique to *H*; and 5), whether the edge is “resolved” (can be linked to an identical or alternative edge in *H*). We produce the same table for *H* with respect to *G*. The *pholidosis graph_edit_distance* function produces this table as an R data frame, and relies on the *igraph* function *shortest_paths* to find alternative paths between vertices.

If an edge in *G* 1) does not contain unique vertices, 2) is not matched by an edge in *H*, and 3) its vertices cannot be connected by a path in *H* that contains only vertices unique to *H*, it is considered

“unresolved”, and marked as such in the descriptive table. These unresolved edges represent topological differences between *G* and *H* that involve only vertices found in both networks (Figure 2). In order to determine the smallest number of changes required to transform *G* into *H*, we must determine if these unresolved edges represent cases of edge “swapping,” in which sub-networks of *G* and *H* share a ring of homologous vertices surrounding different internal edges (Figure 2B). In these “swapping” cases, simply comparing the number of unshared edges between *G* and *H* would double-count simple changes in scalation (Figure 2).

When descriptive tables for the edges of *G* and *H* have been produced, we compare their unresolved edges to identify these systems of “swapping” and count the number of edges in *G* involved in the systems. This is achieved in a few steps. First, for all unresolved edges, we identify the two triangles (3-cycles) that the missing edge was a part of; since each graph is fully triangulated, every edge of the graph is part of exactly two triangles. One can visualize this as a square with a diagonal edge from corner to corner; that diagonal is the unresolved edge, and we can call this 4-vertex subgraph the unresolved edge’s “square graph”. In order to define the two 3-cycles, we must know the names of the vertices in the square graph that are not linked by the unresolved edge. Once these two vertices of the square graph are found for each unresolved edge, we construct an adjacency matrix of unresolved edges; unresolved edges are adjacent if they both appear in each others’ square graphs.

In order to identify corresponding systems of unresolved edges in *G* and *H*, we must identify systems of unresolved edges linked by this adjacency (e.g., internal edges in Figure 2B). From this adjacency matrix, we construct a summary graph wherein the identified unresolved edges are vertices, and the adjacent vertices are linked by edges. We then identify connected vertices, here corresponding to distinct groups of “problem edges” linked by adjacency. We then assign the problem edges to these groups. These groups are then identified by the set of vertices in the original graph that are part of the unresolved edges and their square graphs; this set of vertices forms a cycle around the unresolved edges of length equal to the number of unique vertices of the unresolved edges plus two vertices from the square graphs that were not involved in adjacent edges. The names of vertices in this cycle are then added to the descriptive table. In the *pholidosis* function *graph_edit_distance*, identification of the unresolved edge groups is achieved via a depth-first search of the summary graph using the *igraph* function *dfs*.

We then compare these cycles between graphs G and H. If both the cycles and number of unresolved edges in each cycle are identical between the graphs, all of these unresolved edges are considered "resolved." In this case, the number of edge-swaps to turn *G* into *H* is equal to the number of unresolved edges in *G* or *H*. If there is any difference between the recovered cycles and number of edges in the cycles between the graphs, then our methodology has failed; the differences between *G* and *H* are too great for this method to resolve an unambiguous edit distance (see Supplementary Figure S1). In this case, the *graph_edit_distance* function returns an NA value for the number of edge differences.

With descriptive tables for *G* and *H*, we can count changes in the number of vertives and changes in the structure of edges in *G* and *H* (node changes). We call these the f (vertex) and e (edge) distances. In graphs with differing edge weights denoting fusion or partial fusion of vertex elements, we can also count the number of changes in edge weight differences between the networks, the w distance. Since the fusion of two scales should be evolutionarily equivalent to the deletion of one scale, full fusions are counted as 1 change, and semi-fusions as 0.5 changes.

The *pholidosis* function *pnet_distance* produces an edit distance between two networks using *graph_edit_distance* and counts these weight differences, returning three distance measures; f, the number of vertex differences between the networks; e, the number of edge differences between the networks, and w, the number of edge weight differences between the networks. If desired, it will also return the descriptive tables for *G* and *H* as data frames produced by *graph_edit_distance*.

In some cases, the algorithm outlined here is insufficient to resolve the edit distance. Some of these involve real large shifts in the position of some graph elements relative to others; an example of this is given in supplementary figure S1. In other cases, insufficiency may be due to an incorrect homology assumption that produces a drastic shift in the network. For instance, in several *Dibamus* species such as *D. seramensis*, a scale located behind the ocular scale separates the ocular and postfrontal scales. In the case where this scale is labelled "postocular 1," and the species has a series of 4 postoculars, the *graph_edit_distance* function fails to calculate a distance between species with and without this scale arrangement. However, when we consider this a new type of scale, a "parafrontal," our algorithm can calculate distances between animals with and without this scale arrangement.

Distance and character matrices for pholidosis networks

The *pholidosis* package contains functions to build both distance and character matrices from collections of pholidosis networks. Given a list of pholidosis networks, the function *net_dist_mat* calculates distance measures f, e, and w for each possible pair of networks in the list using *graph_edit_distance*, and returns a 3*n*n array, containing matrices for f, e, and w distances between n networks. This distance matrix is a measure of distances between pholidosis networks in topospace, analogous to the morphospace in a geometric morphometric analysis.

Pholidosis networks can also be broken down into sets of atomized traits. The *pholidosis* function *morpho_matrix* produces a morphological character matrix that treats edges in a network as individual characters. The matrix rows correspond to the pholidosis networks, and the matrix contains a column for each unique edge in the list of networks. If a given edge is present in a given network, the corresponding cell in the matrix is scored with that edge's weight.

Network edit distance as a metric

In order for the network edit distance to be a useful measure, it should define a metric space. A metric space is a space in which four conditions hold; 1) that the distance from a point to itself is zero; 2) that the distance between any two distinct points is positive; 3) that the distance from point *a* to point *b* is the same as point *b* to point *a*; and 4) that the triangle inequality holds true. The triangle inequality states that the distance between points *a* and *b* must be less than or equal to the sum of the distances from *a* to *c* and *b* to *c*.

The first two of these conditions are necessarily fulfilled by the network edit distance calculated herein; all differences between networks are accounted for in the course of measurement by the descriptive tables and each difference is assigned some positive value. Since networks are compared pairwise and neither is given any preference by the edit distance algorithm, the third condition holds.

Investigating the condition of the triangle inequality is more complex. Since the f distance is based on counts of unshared vertices, this intuitively must be true. All differences are measured in positive single units, so any vertices unique to graph C will increase its distance from graphs A and B. For the e distance, we assume axiomatically that the smallest change between graphs is 1 edge swap. This type of change, a swap of the central "chord" edge in a 4-cycle (figure 2B, g_1 to g_3), provides the

basis of the metric. Figure 2B demonstrates the equivalency of the swapping of two chords in a 5-cycle to two steps of swapping in 4-cycles. This logic can be extended to larger cycles with more chords. One can find a series of changes such that each edge in A can be swapped for an edge in B to resolve one square graph at a time in this larger cycle. This produces the result that the number of swaps required to change A into B is equal to the number of edges in A that are not found in B, so long as all of those edges are chords of homologous cycles of A and B. As before, edges unique to graph C will only increase its distance from graphs A and B.

We can check whether the implementation of this graph edit distance algorithm in *pholidosis* produces any triads of networks in the dataset that violate the triangle inequality. For the dibamid dataset here, the inequality holds.

Methods – Pholidosis networks of dibamid lizards (Dibamidae)

I referenced diagrams and photographs from literature sources to code scale adjacency matrices for all published dibamid species, basing the matrices on the holotypes whenever possible (Supplementary Materials: Appendix). I am grateful for the help of Dr. S.R. Chandramouli who sent me high-resolution photos of *Dibamus nicobaricum*, without which I could not have fully resolved its pholidosis pattern (personal communication). The scalation patterns in *D. montanus* and *A. papillosus* elucidated which scales fused together to make the rostral shield in other species. These include the rostral scale, the first three supralabial scales, right and left nasal scales, and right and left internarial scales. In the adjacency matrices, adjacent scales were coded as adjacent (1), partially fused (2), or fully fused (3), but these values are changed to adjacent (1), partially fused (1.5), or fully fused (2) on importation via a call to a setup function, by default, *lizard_setup()*, by the *scale_network()* function.

To calculate a distance matrix for the set of 27 dibamid species (all except the recently published *Dibamus elephantinus*), I used the *net_dist_mat()* function implemented in *pholidosis*. The three components of the distance matrix (the *w*, *e*, and *f* distances) were added together to obtain a single distance measurement between each species pair. I then performed Principal Coordinate Analysis on the summed dibamid distance matrix to map dibamid scale network diversity onto two major axes of

variation using the base R function *cmdscale()*. I used the *WDBdisc()* function in the R package *WeDiBaDis* (Irigoien et al. 2017) to perform a unweighted discriminant function analysis based on a reduced version of this distance matrix (excluding *Anelytropsis papillosus*) to attempt to distinguish between Mainland and Peninsular/island *Dibamus* groups.

Maximum snout-vent length (SVL) and maximum proportional tail length measures used in this paper were compiled by (Quah et al. 2017) and supplemented with data from Koppetsch, Böhme & Koch, (2019), and (Kliukin et al. 2023, 2024*b*).

Since no comprehensive phylogenetic hypothesis exists for Dibamidae, dibamid scalation cannot be tested within a phylogenetic context. To account for some degree of possible phylogenetic signal in scalation patterns, I categorized species into three geographic groups (Mainland Southeast Asia, Peninsular & Island Southeast Asia, and Mexico) following the clades identified by (Townsend et al. 2011).

From the scale networks, I calculated the number of scales, number of fusion events, network strength, and composite scale count (scale count in which fully fused scales are counted as a single scale). I built linear models with the base R *lm()* function using these summary statistics, with and without location as a second term, to understand the basis of the principal axes identified by the principal coordinates analyses. I then compared these eight models for both axes in an AIC framework using the *aictab()* function of the *AICcmodavg* package(Mazerolle 2023).

Using the same approach, I investigated covariation between head scalation and body shape (SVL and maximum proportional tail length). I built linear models using the network summary statistics and Principal Coordinates as predictors, with and without location as a second term. Again, I compared these twelve models for each body shape variable within an AIC framework.

I used distance- and parsimony-based methods to generate phylogenetic hypotheses for Dibamidae based on pholidosis network data alone. From the pholidosis network distance matrix, I produced a distance-based tree following the algorithm by Saitou and Nei (1987) using the *ape* package's neighbor-joining function *nj (Paradis and Schliep 2019)*. For the parsimony tree, the scalation networks were decomposed into discrete characters (one for each edge) using the pholidosis morpho_matrix function. I then performed a maximum parsimony analysis of this matrix, with edge weight coded as an ordered state, using the *phangorn* function *pratchet()* (Schliep 2011).

Results

Dibamus networks

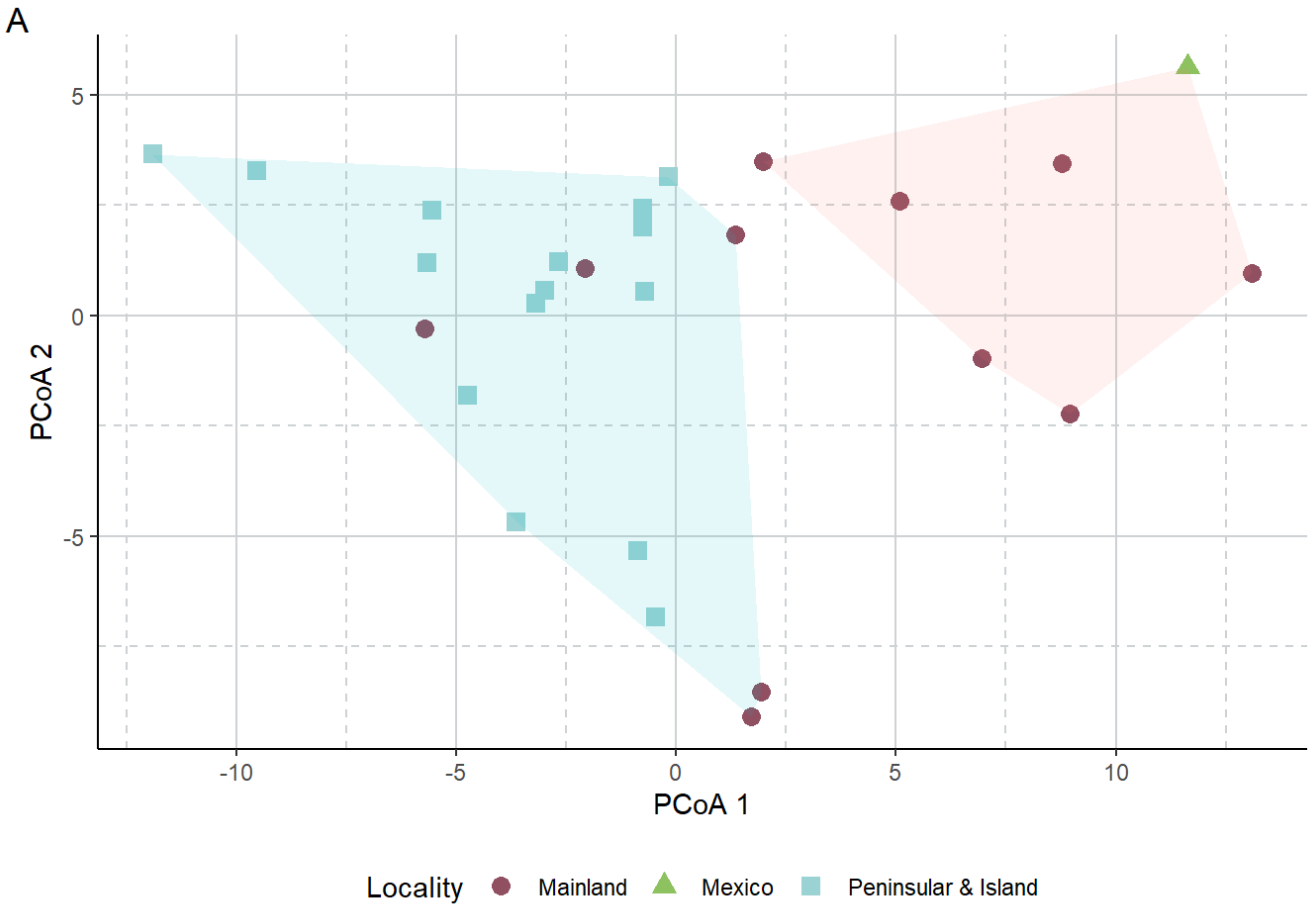

Figure 3: Principal coordinates of Dibamid scalation topospace. This principal coordinates plot reduces dibamid head scalation network topospace to two dimensions. Blue squares represent Peninsular/island species while purple circles represent Mainland species. Anelytropsis papillosus is represented by a green triangle. Shaded areas identify clusters identified with K-means clustering (k=2).

Edit distances between dibamid species (summing differences in scale count, scale fusion, and topological differences between scale configurations) are normally distributed (Shapiro-Wilks W = 0.99, p = 0.003) Edit distances range from 1 (*D. tropcentr* to *D. smithi*) to 26 (*D. bourreti* to *D.*

manadotuaensis), with a median of 11 and a mean of 10.93. Across the distance matrix, edge weight differences accounted for a total of 1727 steps of edit distance, edge topology changes 102 steps, and additions and subtractions of vertices for 2008 steps.

Dimensionality reduction via principal coordinates analysis of the network distance matrix into two dimensions fails to separate dibamids geographically (Figure 3). K-means clustering of principal coordinate data with k=2 and k = 3 groups does not identify geographically coherent sets of taxa; “Mainland” and “Peninsular-Island” groups occupy overlapping and approximately equal areas of morphospace. With k=2 groups, the species are partitioned primarily along PCoA 1; first a small group with PCoA 1 values of five and above containing *Anelytropsis papillosus*, *D. bogadeki*, *D. bourreti*, *D. deharvengi*, *D. greeri* and *D. somsaki*, plus *D. montanus* with PCoA 1 value of just over 2; and second, a larger group at PCoA 1 values below 2.5 containing all other Dibamids. Correspondingly, distance-based discriminant function analysis could not fully distinguish between Mainland and Peninsular/island groups of Dibamus, finding a classifier function that grouped two mainland *Dibamus* species with the Peninsular/island group.

The first principal coordinate axis is best explained by network strength, while PCoA 2 is best predicted by the number of fusion events and locality (Figure 4). The number of fusion events and locality jointly explain 82% of the variation in this axis (Supplementary Table S1).

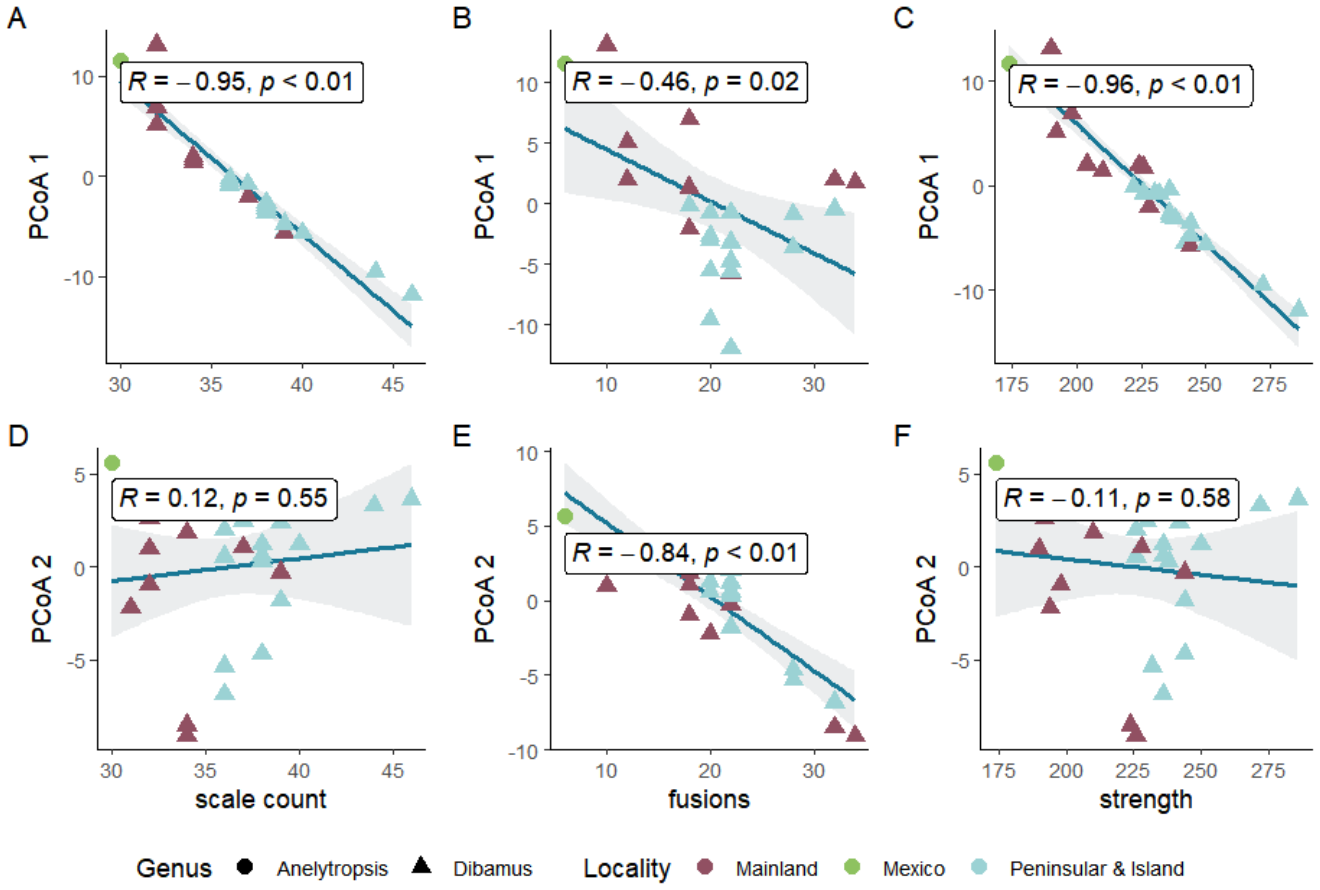

Figure 4: Correlations of principal Coordinate axes (PCoA) and graph properties. A-C show significant negative relationships between scale count, fusions, and strength and PCoA 1. D-E show relationship between these variables and PCoA 2.

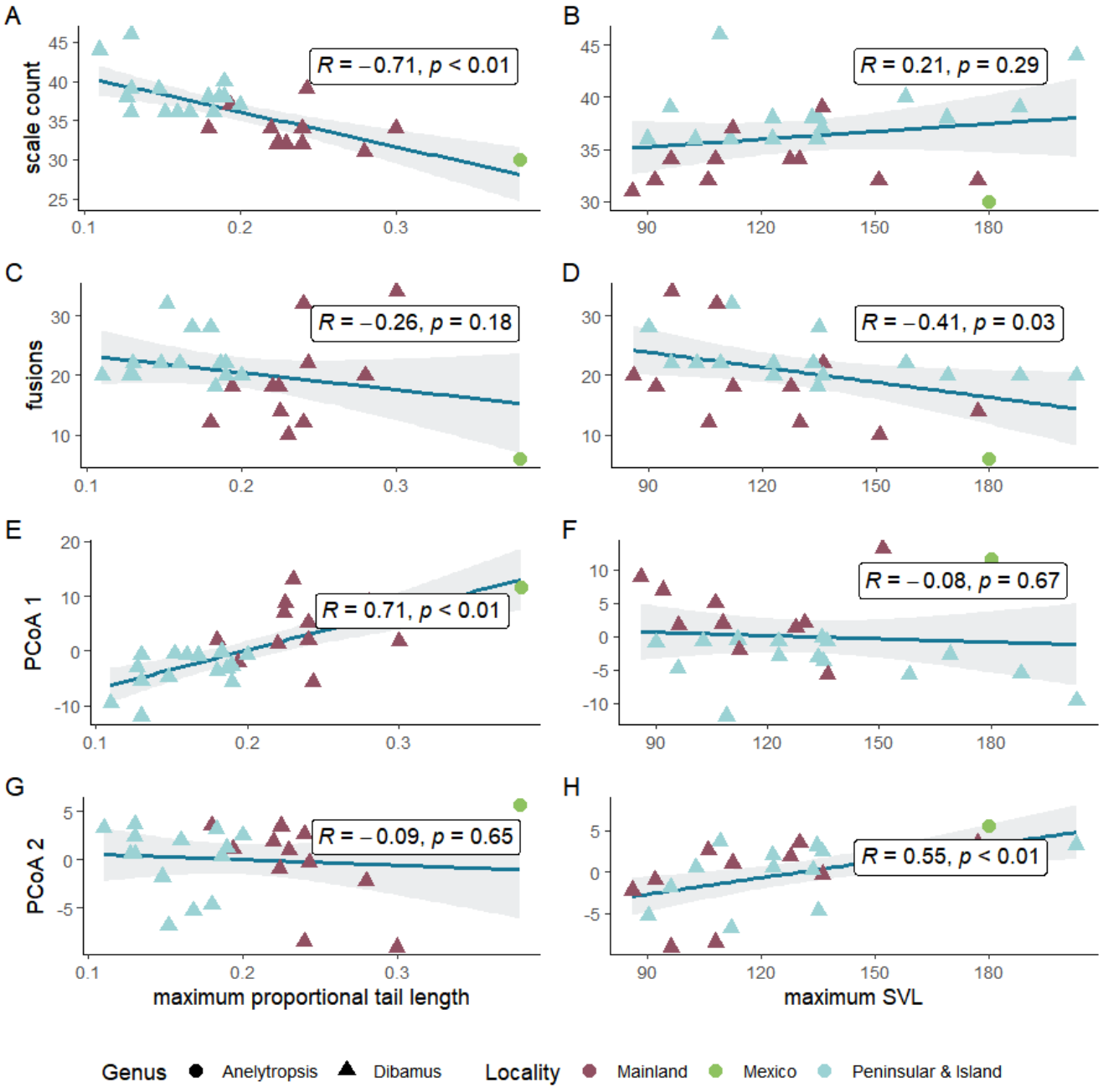

Figure 5: Regressions of dibamid head scalation network properties (vertical axes) on tail length and snout-vent length (SVL) (horizontal axes).

Dibamid pholidosis networks are influenced by body size (SVL). Maximum SVL is correlated with PCoA 2 (R^2=0.55, p< 0.01) and with the number of fusion events, with larger species having fewer fusions (R^2=-0.41, p = 0.03) (Figure 5D).

Surprisingly, pholidosis patterns of dibamid heads are strongly connected to tail length. Tail length is a probable proxy for fossoriality (Wiens et al. 2006; Lawing et al. 2012; Binfield et al. 2026) that is not correlated with SVL in dibamids (R = -0.03, p = 0.84). Figure 5 (A,C,E) shows that scale count and therefore PCoA 1 are tightly correlated with maximum proportional tail length (both R^2>0.73, all p<0.01). However, Mainland and Peninsular/island *Dibamus* groups have very different average tail lengths; the mainland species tend to have much longer tails (Welch's t = -5.9, df = 19.4, p < 0.001) and correspondingly weaker/smaller scalation networks (Welch's t = 4.5, df = 21.0, p < 0.001). Models including locality and either PCoA 2 values or composite scale count best explained differences in tail length in dibamids, with locality the major predictor (Supplementary Table S2).

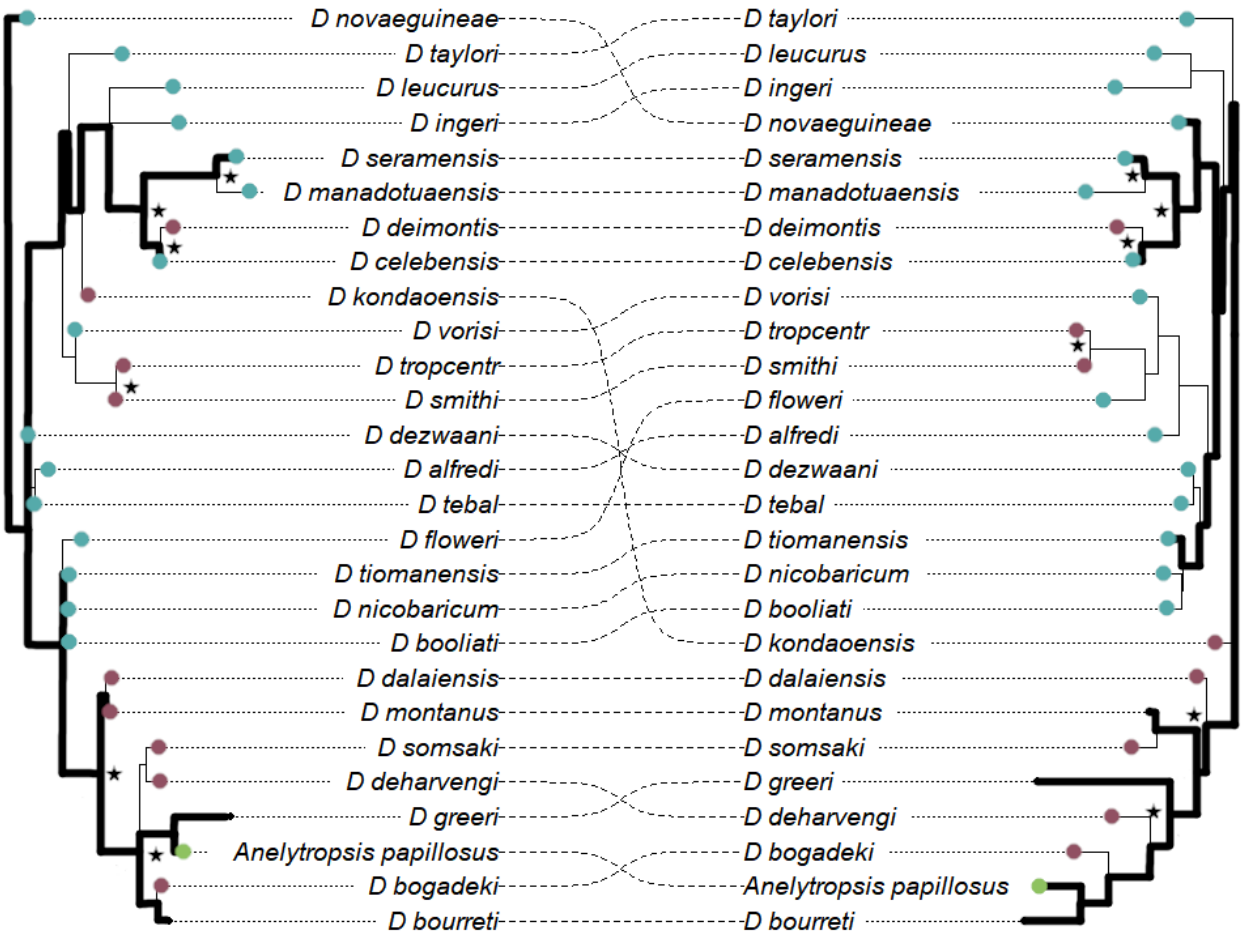

Figure 6: Phylogenies for Dibamidae based on scalation networks. The left tree is a consensus maximum parsimony tree based on a matrix of edge states for all dibamid pholidosis networks. The right tree is a neighbor-joining tree based on graph edit distances between those networks. Mainland species are indicated with purple, peninsular/island species are indicated with blue. Bold edges connect species found in Townsend et al. (2011).Clades found in both trees are marked with a star.

Scalation characters do not resolve particularly robust phylogenies for Dibamidae. The consensus parsimony tree includes several polytomies and has a consistency index of 0.52 and retention index of 0.72, indicating high homoplasy. The neighbor-joining tree

Both distance- and parsimony-based phylogenetic hypotheses agree partially with the Bayesian phylogeny of (Townsend et al. 2011) (Figure 6). Most saliently, they embed *Anelyropsis papillosus* within a clade containing the “mainland” species, rendering that “mainland” clade paraphyletic. Both trees also recover a primarily Sulawesi-adjacent clade (*D. celebensis*, *D. manadotuaensis*, and *D.*

seramensis) rendered paraphyletic by *D. deimontis* from Vietnam. The neighbor-joining tree recovers the “Peninsular/Island” clade with the same branching order as the molecular tree; if re-rooted such that the “Mainland” clade would be sister to all other dibamids, the parsimony tree would recover the same.

However, the phylogenies have rather low agreement; few clades near the tips of the trees are shared between the two, and quartet agreement between the two trees is 52% (9165 agreeing quartets of 17550 possible quartets).

Discussion

Using surface networks

By connecting a wealth of otherwise atomized character states into a biologically representative mathematical structure, the methods described herein allow for rigorous comparative analysis of surface patterns. Here, they have produced insight into possible allometric and ecological patterns in a secretive and poorly-understood group of lizards. The distance measurement method described here produces biologically intuitive results when comparing networks using principal coordinates analysis, discriminant function analysis, and tree estimation techniques.

Dibamid evolution and ecology

The evolution of dibamid pholidosis networks appears to be slow and piece-wise, yet still related to the animals’ ecology. Peninsular/island and mainland *Dibamus* species do not appear to be differentiable based on any single character despite a putative crown age of 54–95 million years for Dibamidae (Townsend et al. 2011). Moreover, though scalation networks produce poorly-resolved phylogenetic hypotheses, they still do appear to carry some amount of phylogenetic information. This suggests an extremely slow rate of evolution in scale characters, as a high rate of character state changes among the small number of characters involved would likely overprint this signal over such a long timespan.

It is surprising that scalation patterns seem capable of resolving biogeographic structure within *Dibamus*. While the phylogenetic hypotheses produced by scalation data are far from robust, they do suggest a more complex phylogenetic structure than a simple Mainland/ Peninsular & Island split. Despite living in Vietnam, *Dibamus kondaoensis*, *D. deimontis*, *D. smithi*, and *D. tropcentr* all appear

outside of the “Mainland clade” of *Dibamus*. Nevertheless, the differentiability of Mainland and Peninsular/Island Dibamus based on tail length suggests that Townsend et al.’s clades are ecologically distinct, with the short-tailed Peninsular/Island *Dibamus* species being more obligately endogeic.

Pholidosis patterns may be adaptive in Dibamidae. However, if shorter tails are indicative of more fossorial lifestyles, the number of fusion events (and therefore the degree of development of the “rostral shield”) does not necessarily carry the same signal. Given the apparent morphological split between Mainland and Peninsular/Island *Dibamus*, it may be that the development of the “rostral shield” is merely overprinted in dibamids by a strong phylogenetic signal. Alternatively, rostral shield development may be a consequence of miniaturization more than adaptation, as body size and number of fusion events in dibamid pholidosis networks are inversely correlated (Figure 5). Even with a comprehensive dibamid phylogeny in hand, the adaptive significance of these scalation patterns will remain unclear until these obscure lizards’ ecology is much better known.

Intraspecific variability in pholidosis may complicate these conclusions. The differences between some very similar *Dibamus* species are attributable to a single change in scale count or single incomplete fusion event, and intraspecific variation is known to exceed this in many species (Kliukin et al. 2024*b*). Future taxonomic work on dibamids informed by molecular phylogenetics may re-partition some of this intraspecific variation into interspecific variation. Until more dibamid specimens are carefully examined and correctly assigned to species, the degree to which scalation patterns are inter- and intraspecifically variable will be unclear. But the relationship between body size and scalation on the interspecific level, combined with the known variability within populations, suggests that there may be an ontogenetic component to dibamid pholidosis, which my examinations of some series of *Dibamus* appear to support.

Future directions in topometrics

I hope that herpetologists will use the tools presented here to rigorously investigate pholidosis, a long-overlooked aspect of squamate evolution. These preliminary results are encouraging; pholidosis patterns appear to retain some phylogenetic information and appear to be influenced by allometry and ecology. But dibamids are likely a poor model of general pholidosis evolution, just as they would be a poor model for general squamate morphology or ecology. Expansion of these methods to diverse

groups of reptiles will certainly provide a better understanding of the evolution of pholidosis patterns in squamates and likely result in better-informed methods of species identification.

By providing comparative tools for the study of biological surface networks, I hope to encourage others to take an interest in these patterns and attempt their comparative analysis. This methodology is ripe for further refinement and expansion. The distance metric implemented in this study will likely be sufficient to analyze a large number of surface patterns, but could be expanded to process large-scale shifts in network structure that might occur in other organisms, such as between families or orders of insects (Salcedo et al. 2019). Further incorporation of vertex- and edge- specific characters and models of their transition could prove useful in studying aspects of surface networks beyond bare topology, including surface properties such as texture, shape, and color. A fully developed methodology of topometric comparison and vertex characters, including edge shape and surface color could, for instance, be used to completely describe the appearance of a butterfly wing, and to quantify the interplay between methods of reaction-diffusion and positional information in producing differing color patterns between different species of butterflies.

Chapter 4: Substrate specialization in the evolution of squamate communities: a case study in Madagascar.

Abstract

The repeated evolution of limbless, fossorial squamates has captured much scientific attention. This extreme, and likely irreversible, change in body plan has evolved dozens of times within squamates, and current explanations for this phenomenon have centered on geographic isolation as a major driver of convergence, allowing the trait to evolve in situ more rapidly than fossorial clades can spread geographically.

Here, I use an extensive ecological and geographic dataset on Malagasy squamates to investigate the tempo of substrate preference evolution and the consequences of substrate specialization. I find that substrate specialization is rarely reversed, but that different specializations have vastly different macroevolutionary consequences; in Madagascar, fossoriality depresses diversification, but not geographic range size, and arboreality has no clear effect on diversification but does decrease geographic range size. Fossorial taxa tend to overlap in diet, size, and activity patterns, but not in climate or geographic space, while arboreal taxa exhibit the opposite pattern. Within clades with morphological pre-adaptations for substrate specialization, specialization evolves at similar rates, suggesting that the repeated evolution of fossoriality is not unusually frequent. These results suggest that fossorial specialization primarily inhibits ecological opportunity but not geographic spread, suppressing diversification through an ecological rather than biogeographic mechanism.

Introduction

Squamate reptiles comprise an exceptionally species-rich and diverse clade that is a major component of terrestrial vertebrate communities almost worldwide. Their diversity in body size, morphology, physiology, life history traits, and ecology makes the clade an ideal group in which to study how these traits influence evolutionary dynamics of diversification and biogeographic spread, as well as how the traits influence each other. Though key morphological characters (e.g., Burress and Muñoz 2022), ecological opportunity (e.g., Pereira and Schrago 2017) and elevational preference

(Pepper et al. 2011; Skeels et al. 2023) are important drivers of diversification in some clades, few traits have been recognized as globally important drivers of diversification in squamates. Among organismal traits (phenotypic traits expressed in individuals), substrate use has the clearest and strongest effect on diversification (Bars Closel et al. 2017; Li and Wiens 2022). Clades containing a higher proportion of arboreal species have significantly higher speciation rates, while clades containing a high proportion of fossorial or aquatic species tend to have lower speciation rates (Bars Closel et al. 2017).

The mechanisms behind substrate-use mediated diversification are not entirely understood. The complexity of arboreal environments creates opportunities for fine-scale spatial partitioning of the environment, as seen repeatedly in the *Anolis* radiation (Losos 1992; Moreno-Arias and Calderón-Espinosa 2016; Huie et al. 2021) which should lead to an increase in overall diversity. The mechanism underlying decreased diversity among fossorial squamates is less clear; it may be that the converse is true, and that the homogeneity of available underground habitat does not admit specialization, at least to a degree comparable with other habitats. However, the lower diversification may also be attributable to inhibited rates of geographic spread, as the sandy soils necessary for some fossorial squamates tend to occur in isolated patches. These isolated substrate islands may be difficult for these generally small and presumably not very vagile organisms struggle to reach and subsequently speciate in (Wiens et al. 2006; Li and Wiens 2022). Distinguishing between these explanations requires more detailed study of the geography and ecology of fossorial groups.

A detailed study of the geography and ecology of fossorial groups at a global scale will require more data on the geographic distributions and ecologies of squamates than are currently available. Therefore, I here use the island of Madagascar as a microcosm. Madagascar is a center of global endemism in which the course of terrestrial evolution has been largely uncoupled from the course of evolution on other major landmasses since their separation in the Cretaceous. While the island retains various biotic ties to other Gondwanan landmasses (India, South America, the Seychelles, and Australia), it is estimated that about one third of its modern biota have their closest relatives in Africa, and that the bulk of this diversity is made up of lineages that have dispersed from Africa after the separation of these landmasses (Yoder and Nowak 2006).

Despite the island's separation from the continents, its fauna suffered the same end-Cretaceous mass extinction as did the mainland, which appears to have largely wiped out the island's native

vertebrate fauna (Crottini et al. 2012; Michielsen et al. 2023). The modern fauna of the island therefore largely represents the descendants of a few lineages that have been able to cross the Mozambique channel from a continental origin and radiate on Madagascar. Within Squamata, only two extant lineages appear likely to have been present before the K-Pg extinction; the monotypic, endemic blind snake family Xenotyphlopidae (Miralles et al. 2018), and perhaps the island's endemic boids, *Acrantophis a*nd *Sanzinia* (Noonan and Chippindale 2006).

The majority of Cenozoic invasions of Madagascar have been by African fauna, especially among squamates (Crottini et al. 2012). Two clades of Chameleons, the miniaturized leaf-litter specialist *Brookesia*, and a clade of large, arboreal chameleons (*Calumma* and *Furcifer*) independently colonized Madagascar from Africa, with *Brookesia* likely arriving very close temporally to the end-Cretaceous mass extinction and the larger-bodied clade in the Mid-Eocene ca. 47 Ma (Tolley et al. 2013). The endemic typhlopid genus *Madatyphlops* may also have first arrived on the island in the Eocene or Paleocene, though its African origin is not entirely certain (Miralles et al. 2018). Madagascar's main skink radiation is derived from a clade of long-bodied, primarily African skinks including *Chalcides* and *Scelotes*, that arrived on the island in the Eocene at least ca. 50 Ma (Miralles et al. 2026). Another colonization from Africa, of *Trachylepis* skinks, occurred in the Early Oligocene ca. 30 Ma (Weinell et al. 2019). Snake-eyed skinks (*Cryptoblepharus*), are a recent arrival (ca. 4 Ma) from an origin at the Eastern margin of Indian Ocean and almost entirely restricted to coastal environments (Blom et al. 2019). The Malagasy gerrhosaurids, consisting of their own subfamily, Zonosaurinae, colonized the island from Africa in the Early Miocene, ca. 18 Ma (Blair et al. 2015). Pseudoxyrhophiine snakes have diversified extensively in Madagascar, producing the bulk of the subfamily's diversity since their arrival on the island from Africa in the early Miocene, ca. 22 Ma, while the closely related lamprophiid *Mimophis* arrived from Africa in the Middle to Late Miocene and is far less diverse, having only two species (Burbrink et al. 2020).

The deep biogeographic history of Madagascar's geckos has not been studied comprehensively, but current evidence points to all major extant lineages originating soon after the K-Pg extinction (Crottini et al. 2012). The clade containing most of Madagascar's geckos contains many other genera from Africa, India, and the Seychelles, suggesting a complex biogeographic history for this group immediately following the K/Pg boundary and making the origins of the Malagasy taxa unclear (Pyron et al. 2013). Nevertheless, Madagascar's geckos are certainly not newcomers to the island, and an

African origin for at least some of the genera, if not the entire radiation seems quite likely. The only recent arrival is Madagascar's single native *Hemidactylus* gecko, *Hemidactylus mercatorius*, arriving within the last 5 million years from the African continent (Crottini et al. 2012).

Madagascar has not only accumulated biodiversity from elsewhere, but produced many lineages that have subsequently colonized the [islands in the Indian Ocean and Eastern Africa. A major radiation of diurnal geckos including the partially frugivorous day geckos (*Phelsuma*), found throughout the Indian Ocean as far East as the Andaman islands (Rocha et al. 2007), and the smaller *Lygodactylus* geckos, found throughout Madagascar, Africa, and with a few species in South America (Gippner et al. 2021), is likely to have originated in Madagascar. The elusive, small-bodied *Ebenavia* geckos have colonized Mauritius, the Comoros, and Pemba island off the coast of Tanzania (Hawlitschek et al. 2018), and the Southern African gecko genus *Homopholis* appears to be nested within a clade otherwise consisting of Malagasy fish-scaled geckos *Geckolepis*, and velvet geckos (*Blaesodactylus*, sister to *Homopholis*) (Greenbaum et al. 2007).

Madagascar's large size and heterogeneous environments provided ample space and opportunity for lineages to diversify and specialize after their arrival, producing an extremely diverse squamate fauna that is nearly entirely endemic to the island (Glaw and Vences 2007; Brown et al. 2014). This makes the island an excellent testing ground for hypotheses on the evolution of diverse biotic communities (Razafindratsima et al. 2013; Brown et al. 2014; Casquet et al. 2015; Hackel et al. 2018; Burbrink et al. 2019). In particular, the repeated evolution of limblessness and fossoriality in the island's skinks, which have evolved limblessness or extreme limb reduction independently at least five times (Miralles et al. 2026), and the island's numerous radiations of arboreal geckos and chameleons make it a compelling microcosm in which to study the consequences of substrate specialization.

Here, I investigate the evolution of substrate use specialization on Madagascar, attempting to understand how squamate substrate specialists have originated on the island and how their particular substrate preferences have structured their diversification. I investigate three main questions in the evolutionary history of these specialists; first, how frequently have substrate specialists evolved, and whether they have evolved in situ or evolved elsewhere and subsequently colonized the island; second, whether different substrate specialists have different rates of net diversification and biogeographic spread; and third, along which niche axes different substrate specialists have diversified, and what consequences specialization has on the evolution of other traits.

To investigate these questions, I have assembled diverse geographic and trait data on all 54 genera of native Malagasy squamates, and constructed a phylogenetic tree containing 353 malagasy species belonging to these genera and 31 species representing their non-Malagasy sister taxa in order to map the evolution of traits and measure their divergence and overlap within a phylogenetic context. Using these data, I estimate per lineage evolutionary rates of substrate specialization, map relationships between ecological niches and diversification and biogeographic range, and measure niche overlap across several axes to understand how specialization for arboreal and fossorial lifestyles influences niche evolution and diversity.

Materials & Methods

In this paper, I employ four basic types of analysis:

1) Detection of the phylogenetic signal of various traits, using a recently published unified framework for phylogenetic signal, *M*, (Yao and Yuan 2025), which varies between 0 (phylogeny has no predictive power for traits) and 1 (phylogeny perfectly predicts traits), and can be measured for continuous, categorical, and multivariate traits.

2) Estimation of rates of character change using Markov models (Lewis 2001) in order to measure differences in rates of character evolution and build statistical support for putative shifts in those traits.

3) Linear modeling of relationships between traits within an evolutionary framework using phylogenetic generalized least squares (PGLS).

4) Partial Mantel tests to correlate between-genus niche overlap values, controlling for phylogeny using a phylogenetic distance matrix.

I perform these analyses at the species level where appropriate, otherwise at the genus level.

Through these analyses, I establish relationships between substrate use, body size, and diversification and geographic range size, and study how substrate use specialization relates to specialization in other niche axes.

Phylogenetic Tree

I produced a phylogenetic tree of all Malagasy squamates for which published phylogenetic data were available. This tree was produced via manually grafting taxa together based on family- or genus-level phylogenies, using published data files when possible and manually constructing tree files based on figures when published data files were not available. These lower-level phylogenies were grafted onto a dated backbone phylogeny based on two recent large-scale phylogenies of Squamata (Tonini et al. 2016; Zheng and Wiens 2016), with non-malagasy immediate outgroup taxa included. The resulting tree contains 384 species, 353 of which are malagasy taxa, and 31 of which represent sister groups to malagasy taxa. I produced a genus-level version of this tree by pruning all tips from the tree except a single tip representing each genus using the *drop.tip()* function from *phytools* (Revell 2024)(Table S1).

Environmental Data

I use the R package geodata (Hijmans et al. 2025) to download 19 Bioclimatic variables (Hijmans et al. 2005) and a global DEM (Jarvis et al. 2026) at 5 arc-minute resolution, as well as global physical soil characteristics from the SoilGrids database (Poggio et al. 2021) at a 30-arcsecond resolution. SoilGrid layers included bulk density, clay content, coarse fragments, sand, and silt at 0-5, 5-15, 15-30, and 30-60cm soil depths.

I down-sampled the SoilGrids layers to a 5 arc-minute resolution (a factor of 100) to match the bioclimatic and elevational data resolution and performed a PCA on the layers to reduce dimensionality using the *raster.pca* function from the *ENMTools* package (Warren and Dinnage 2025). Four principal components accounted for >96% of the variance in the 16 soil layers, so I produced rasters for these four components to use as global soil variables.

I then combined the bioclimatic, elevation, and soil PC variables into a single dataset and performed an iterative variance inflation reduction procedure to remove collinearity from the dataset using the *vifstep()* function of the “usdm” package (Naimi 2023). Ten environmental layers were rejected due to multicollinearity, leaving fourteen non-collinear layers, including nine bioclimatic variables, elevation, and the four soil principal components. Details of the layers retained can be found in Table S2.

Occurrence Data

I downloaded occurrence data from GBIF for all native Malagasy squamate genera using the rgbif package (Chamberlain et al. 2026) and filtered points using a series of R scripts. I used the coordinateCleaner package to flag observations with inappropriate coordinates (country and city centroids, coordinates in the ocean, and coordinates that represented outliers within species) and manually confirmed the validity of flagged points using GBIF metadata (for instance, whether a verbatim locality was listed and whether the point coordinates occurred at the verbatim locality listed) and IUCN range maps. I produced point occurrence maps for each genus (Supplementary materials) to be used as inputs for maximum entropy niche equivalency modeling.

Polygonal Range Maps

I merged polygonal range maps for reptile species from the GARD database (Roll et al. 2017)

to construct genus-level polygonal range maps using the aggregate function in the *terra* R package (Hijmans et al. 2026) after adjusting GARD taxonomy to match the taxonomy used in my cleaned occurrence data. I checked the accuracy of these maps by plotting them overlain with point occurrence data, and made adjustments to several maps that poorly matched point distributional data by substituting inaccurate or artefactual range polygons with range polygons from IUCN RedList data. Final range maps were produced by aggregating the cleaned IUCN range aggregate polygons with polygons produced via adding a 10 km buffer around all point occurrence data, and intersecting this aggregate with a high-resolution polygonal map of Madagascar to ensure that species ranges do not extend into surrounding ocean. I calculated the geographic area and centroid coordinates for each final generic range map.

Geographic Occupancy

To build a dataset of geographic co-occupation, I quantized polygonal range maps into raster data, producing raster maps of occupancy for all genera at the same 5 arc-minute resolution as the climatic variables using the *terra* function *rasterize()* (Hijmans et al. 2026). I transformed these data into a matrix of occupancy in each grid square for each genus.

Environmental Niche Equivalency Modeling

Occurrence datasets for each genus were thinned using the *GeoThinneR* package (Mestre Tomás 2025) to restrict the datasets to one occurrence per 5 arc-minute by 5 arc-minute grid square, matching point resolution to the spatial resolution of the environmental data. Following point thinning, I measured niche equivalency for each pair of genera using the dismo *nicheEquivalency()* function (Warren and Dinnage 2025) , producing 10 randomizations as a null dataset to determine significance of niche non-equivalency. I chose this small distribution because the degree of niche equivalency, rather than its significance, is the key comparative factor between generic pairs. I measured niche equivalency by both *I* (Warren et al. 2008) and Schoener's *D*.

Niche equivalency *I* and *D* values were highly correlated (Spearman's rank correlation $\rho = 0.981$, $p < 0.001$), so I proceeded with analysis using Schoener's D, as it was the most feasible method for measuring other axes of niche overlap. I constructed a distance matrix by subtracting all *D* values from 1 and set the distance to zero for any pairs where the *D* value was found to be non-significant (meaning that the null hypothesis of niche equivalency could not be rejected). I identified genera with particularly unique environmental niches by calculating average *D* values across the rows of the *D* distance matrix and finding genera with particularly low average *D* values.

Traits

I used the Meiri (2024) trait database to assemble data on maximum snout-vent length (SVL), maximum tail length (TL), maximum mass, limb development, activity time, substrate use, and detailed diet information ("diet metadata") on all native Malagasy reptiles. I supplemented these with data from various taxonomic papers (e.g., Sakata and Hikida 2003; Miralles et al. 2026).

Diet

I re-categorized diet data according to a diet hierarchy (see Supplementary Materials) and added data on snake diets from my own datasets (chapter 2, & iNaturalist). Species listed as "omnivorous" but having small body mass (<5g) were coded as eating invertebrates and herbivorous.

To construct a table of diet item-use for each genus, I first constructed a table counting the number of times species in the genus were asserted to use each terminal-level diet category (e.g. "insects" or "honeydew"). For animals whose diet data included higher-level diet categories (e.g.

“eggs” or “omnivorous”), I added partial observations to this table based on the existing distribution of diet items normalized by the number of times the higher-level diet category was asserted for the genus. For instance, if two *Phelsuma* species were said to be “omnivorous,” and the distribution of terminal-level diet items form *Phelsuma* was 20 nectar, 14 insects, 1 honeydew (35 total observations), the “omnivorous” diets of these two species were coded as 1.142 nectar (20/35)* 2 species, 0.8 insects (14/35) * 2 species, and 0.057 honeydew. This approach may bias species towards smaller niches but avoids interpreting overlapping niches based on vague data while using all available data to inform dietary niche.

Substrate Use and Activity Pattern

Substrate data from Meiri (2024) were adjusted to classify “Semi-Aquatic” species as “Aquatic & Terrestrial” and “Cryptic” (n = 2: *Madascincus polleni* and *M. pyrurus*) as “Terrestrial.” I used imputed substrate preference for *Madatyphlops* (Fossorial, however, see Fleck et al. (2026) for discussion of occasional arboreal behavior in the genus) and designated *Elapotinus* as fossorial due to its small eyes and proportionally short tail.

I re-coded activity pattern data from Meiri (2024) to designate species either as nocturnal, diurnal, or both nocturnal and diurnal.

Body Size

Because of the radical differences in body shape between snakes and lizards, the comparison of linear measurements between the two groups is highly problematic. Following Feldman et al. (2016), I used log-transformed body mass estimates (from Meiri (2024)). I recorded maximum and minimum body sizes of genera, and, to allow for measurement of body size overlap, I binned species into 10 discrete body size categories.

Niche Overlap

I calculated niche overlap for each niche axis (diet, substrate use, activity pattern, body size, geographic space) using genera as the unit of measurement. For each genus, I counted the number of attributed uses of each resource (e.g. “fruit”, “terrestrial”, “or “nocturnal”) and then calculated the proportion of use of each resource along its respective axis. For each pair of genera, I calculated Schoener’s D along each niche axis (diet, substrate use, activity pattern, body size). Following Geange

et al. (2011), I calculated unified niche overlap (*NO*) as the mean of the Schoener's D values across climate/soil, diet, substrate use, activity pattern, and body size for each pair of genera.

Dimensionality Reduction

I performed principal components analysis on substrate proportional use patterns using the R *princomp()* function. For diet data, I used the *vegan* function *metaMDS()* (Oksanen et al. 2022) to perform non-metric dimensional scaling (NMDS) on a reduced diet item matrix, excluding all genera for which no diet data were available and rounding diet counts to the nearest whole number. Because of very high between-genus dissimilarity and data completeness, I opted to use the Cao distance (Cao et al. 1997) as a metric, following recommendations of the *vegan* R package (Oksanen et al. 2022). I used the *dimcheck_out()* function from the *goeveg* package (Lampe and Jenny 2026) to measure MDS ordination stress across k=1 through k=15 nMds dimensions, and selected k=6 dimensions to keep for later analyses based on an apparent inflection point in a plot of stress versus change in stress.

I visualized the positions of genera in the first two dimensions of PCA and MDS space through scatterplots, assigning colors to scatterplot points according to their position along the first three dimensions using the *colors3D* R package, mapping each dimension to color channel (red, green, or blue), and selecting for combinations of variable mappings that maximized visual distinctness of niches and niche usage within them (Kling 2023).

Species-level comparisons

I modeled species-level rates of substrate use evolution using binary substrate use data, using whether species are or are not listed as being particular substrate users (terrestrial, fossorial, arboreal, & saxicolous) for all species for which both a phylogenetic placement and substrate use data were available. I used the *phylosignalDB* R package (Yao and Yuan 2025) to measure M to quantify phylogenetic signal of these binary traits, and followed Revell (2024) in fitting equal rates, all rates different, gain-only and loss-only Mk character models (Lewis 2001) to each binary substrate use character using the *fitMk()* function in phytools. For these models, I constrained the common ancestor of the squamate assemblage to be terrestrial, non-fossorial, non-arboreal, non-saxicolous, and four-limbed. I used ANOVA tests to compare model fit via Akaike Informaiton Criterion (AIC).

For each substrate use character, I fit a rate shift Mk model to the data, specifying a priori one or two lineages in which evolutionary rates might differ from the rest of the tree based on the observed

pattern of character states at the tips (Revell et al. 2024). I then compared the AIC-selected single rate-matrix model to the rate shift model via a likelihood ratio test using the R package *lmtest* (Hothorn et al. 2022).

I tested for correlation between leglessness and fossoriality using Pagel's test of correlated trait evolution (Pagel 1994), implemented in the phytools package (Revell 2024), again constraining the root node to be four-legged and non-fossorial.

Intergeneric Comparisons

As in species-level comparisons, I tested for phylogenetic signal in all non-environmental traits using Yao & Yuan's *M*. I treated substrate use proportions as a single complex trait, and used coordinates from the first 3 NMDS axes of diet space and climate space as two complex "diet" and "climate" traits.

To understand relationships between continuous traits, I used phylogenetic generalized least-squares regression on continuous variables scaled to unit variance, so that differences in effect size (regression slope) could be easily compared between variable types with vastly different ranges. I performed PGLS using correlation matrices calculated by the *corBrownian()* function in the *ape* package (Paradis and Schliep 2019) and computed Nagelkerke's pseudo-R^2 using the *nagelkerke()* function from the *rcompanion* package (Mangiafico 2026).

I first analyzed the relationship between three markers of evolutionary success (species richness, net diversification rate, and geographic range size) and three common predictors of these variables; clade age, minimum body size, and maximum body size. Though clade age is generally not an important predictor of diversity (Rabosky et al. 2012), it provides an important null hypothesis for patterns of diversity (diversity is a function of constant diversification rates with little to no extinction). Body size and geographic range size have previously been found to have positive correlations with species diversity in squamates (Feldman et al. 2016; Li and Wiens 2022).

After determining that some of these common predictor variables are important predictors for different markers of evolutionary success, I ran two-predictor PGLS regressions using a chosen common predictor and each of the first three principal components of substrate use as predictors of the three markers of evolutionary success. For predicting clade diversity and net diversification rate, I used generic minimum body size, as small body size is often related to fossoriality. For predicting log

generic range size, I used maximum body size, as larger-bodied species tend to have larger geographic ranges, while smaller-bodied species can have large or small ranges (Camacho et al. 2017).

I conducted partial mantel tests of correlation between Schoener's D values for axes of niche divergence, using a phylogenetic distance matrix as a covariate to account for between-genus phylogenetic similarity. In the first set of tests, I compared Schoener's D matrices for geographic, climatic, dietary, substrate use, diel activity pattern, and body size similarity in a pairwise pattern, to measure general correlation between axes of niche overlap. In the second set of tests, I constructed binary similarity matrices to designate groups of substrate specialists (genera with >=50% utilization of a particular substrate category) and correlated these with the set of Schoener's D matrices in order to measure whether commonality in a particular substrate specialization was predictive of niche overlap.

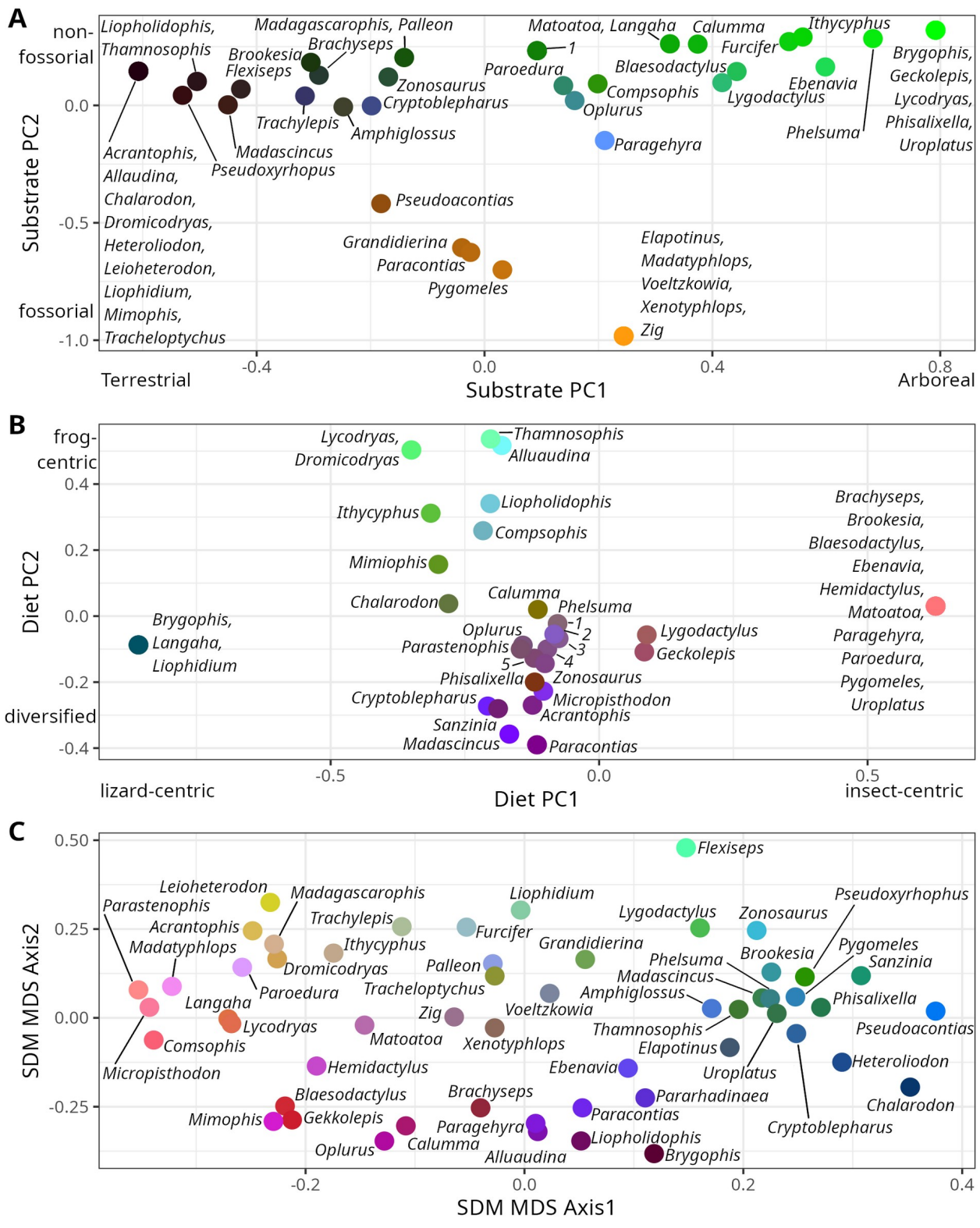

Figure 1: Niche diversity in Malagasy squamates. A) Substrate use diversity, 1-Sanzinia, Parastenophis, Micropisthodon, Hemidactylus. B) Dietary diversity, 1-Phelsuma 2-Pseudoxyrhopus, 3-Leioheterodon, 4-Furcifer, 5-Madagascarophis. C) Climatic and soil niche diversity

Results

Niche diversity

The malagasy herpetofauna exhibits varied patterns of diversification in substrate use, diet, and climatic and soil preferences. Substrate use space is filled with a bifurcating pattern; most taxa fall along a line or relative terrestriality versus arboreality, while fossorial taxa depart from this pattern; fossoriality and arboreality do not co-occur within malagasy squamate genera (Figure 1A). Dietary NMDS space is dominated by a primary axis of opposing hyper-specialization on either lizard or insect prey. The second axis of diet space separates frog specialists from more generalized or diverse diets (Figure 1B). Further axes of diet NMDS space separate harder-bodied prey from softer-bodied prey, and frugivory from the eating of endotherms. SDM NMDS space is more difficult to interpret; Axes 1 and 2 cluster some largely sympatric groups close to more widespread or allopatric groups (e.g, *Parastenophis*, *Micropisthodon,* & *Compsophis* near *Madatyphlops*) (Figure 1C.

Species-level comparisons

In Malagasy squamates, substrate use evolves from unspecialized to specialized, rarely in reverse. Best-fit models for the evolution of terrestriality (Table S4), arboreality, and fossoriality (Table 1) indicate that the evolution of specialized substrate preference (arboreality or fossoriality) occurs 5-6x more frequently than any reversals to terrestriality within clades that possess pre-adaptations for fossoriality (body elongation in skinks) or arboreality (adhesive toe pads in geckos or zygodactyl feet in large-bodied chameleons). Pre-adaptation is extremely important in driving these patterns; models designating rate shifts in pre-adapted clades fit the observed data far better than models assuming unified rates across the tree (Table 1). Fossoriality appears to be a prerequisite for limb reduction according to Pagel's test of correlated trait evolution, with a fossoriality-dependent model of limb reduction far outperforming independent models of trait evolution ($p << 0.001$).

Intergeneric comparisons

Among Malagasy squamate genera, substrate specialization is associated with reduced diversification and geographic range size, with different specializations having different consequences. While individual traits (activity pattern, substrate use, diet, and body size) all show significant phylogenetic signal, clade-level traits (climatic/soil niche preference, net diversification rate, and

geographic range size) are not phylogenetically structured in this assemblage (Table S5). PGLS regressions found that clade age does not correlate with diversification or geographic range size, but that body size does influence diversification and geographic range size (Table S6). Fossoriality significantly suppresses genus-level diversity and diversification rate, while arboreality suppresses geographic range size (Table 2).

Table 1: Dual rate-matrix models of character evolution for substrate use and limb development, including all-rates-different models (ARD) and all-rates-different models with a single rate shift (ARD+S). Models favored by AIC are in bold. Log likelihoods of favored tests are marked with asterisks to denote levels of significance. ** = p <0.001, ***= p <0.0001.

Trait	model	shift point	loglik	d.f	rate gain 1	rate loss 1	rate gain 2	rate loss 2
fossorial	ARD		-59.41	2	0	0.01	-	-
fossorial	**ARD+S**	**skinks**	**-51.08*** **	**4**	**0**	**0**	**0.02**	**0**
saxicolous	ARD		-110.44	2	0	0.02	-	-
saxicolous	**ARD+S**	**toxicoferans**	**-104.96****	**4**	**0.01**	**0.02**	**0**	**0.01**
four-limbed	ARD		-38.93	2	0	0	-	-
four-limbed	**ARD+S**	**skinks**	**-26.495*** **	**4**	**0**	**0**	**0**	**0.02**
arboreal	ARD		-142.71	2	0	0.02	-	-
arboreal	**ARD +S**	**geckos, *Calumma* + *Furcifer* clade**	**-133.93*** **	**4**	**0**	**0.02**	**0.02**	**0**

These differing outcomes are reflected in the type and degree of niche overlap seen between genera with similar substrate preference. Fossorial genera converge in individual traits, while arboreal genera converge in geographic and environmental preference. Fossorial genera tend not to co-occur, and are more similar in size and diet than arboreal genera, which do tend to co-occur, but are diverse in diet, body size, and activity pattern (Figure 2, Table 3). Co-occurrence among arboreal genera is high in the Eastern rain forest (Figure 3).

Table 2: Results of PGLS correlation of markers of evolutionary success with substrate specialization (PC scores) and body size within 52 genera of Malagasy squamates. Significant relationships ($p < 0.05$) are in bold.

correlation	substrate value	size value	substrate p-value	size p-value	pseudo-R^2	AIC
div. ~ substrate pc1 (arboreality) + min. body size	0.02	-0.32	0.83	0.02	0.13	131.52
div. ~ substrate pc2 (fossoriality) + min. body size	**-0.23**	**-0.38**	**0.04**	**0**	**0.21**	**126.93**
div. ~ substrate pc3 (saxicoly) + min. body size	0.11	-0.34	0.4	0.01	0.14	130.32
net. div. rate ~ substrate pc1 (arboreality) + min. body size	0.1	-0.46	0.47	0.01	0.13	166.82
net. div. rate ~ substrate pc2 (fossoriality) + min. body size	**-0.42**	**-0.56**	**0.01**	**0**	**0.25**	**159.86**
net. div. rate ~ substrate pc3 (saxicoly) + min. body size	0.16	-0.48	0.4	0.01	0.13	166.1
log range size ~ substrate pc1 (arboreality) + max. body size	**-0.28**	**0.8**	**0.03**	**0**	**0.39**	**155.64**
log range size ~ substrate pc2 (fossoriality) + max. body size	-0.13	0.73	0.45	0	0.34	159.34
log range size ~ substrate pc3 (saxicoly) + max. body size	0.04	0.8	0.82	0	0.33	159.86

Table 3: Correlation values from partial mantel tests of substrate specialization (more than 50% of substrate usage within a genus) versus niche overlap within 52 genera of Malagasy squamates. Only significant values ($p < 0.05$) shown.

	arboreal genera (n = 17)	fossorial genera (n = 9)	Terrestrial genera (n = 27)	Saxicolous genera (n = 2)
geograph	0.13		0.13	
climate	0.13		0.05	
diet		0.18		0.01
activity		0.13		0
size		0.01		

Discussion

Niche evolution within the Malagasy squamate fauna occurs at dramatically different rates along different niche axes and within different clades. Diet and especially substrate use are highly phylogenetically conserved (Tables S3 & S5). In particular, rates of substrate use evolution can differ dramatically from the background rate in specialist clades; within skinks, the evolution of fossoriality occurs at a rate 251x faster than in other Malagasy squamate groups, and in geckos and large-bodied chameleons, rates of gain and loss of arboreality are completely inverted relative to the background (Table 1). One can easily link both of these cases of rate shifts with key morphological characters; in skinks, an elongate body covered in smooth scales, in geckos, adhesive toe pads (though these have been lost multiple times), and in chameleons, zygodactylus feet and a prehensile tail that aid in climbing.

Of these two specializations, only fossoriality seems to produce long-term diversification consequences in the Malagasy squamates. My analyses show that the evolution of fossoriality is largely irreversible (Table 1), and clades tend to have either no fossorial species or a very high proportion of fossorial species (Figures 1 & 2). Paths to other strategies of substrate specialization are closed for fossorial reptiles, and paths of diversification along other niche axes are restricted (Table 3), likely contributing to slightly depressed rates of diversification (Table 2).

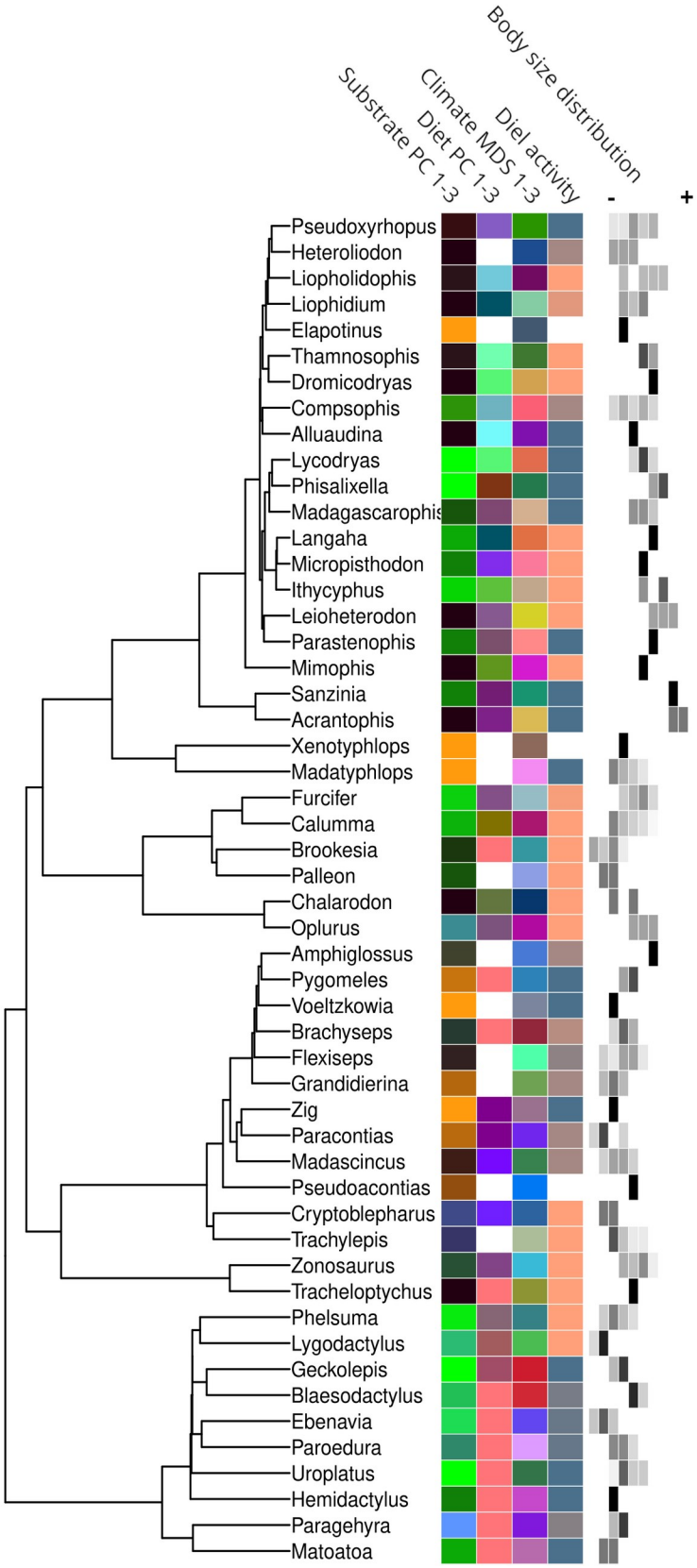

Figure 2: Phylogenetic tree of 52 genera of squamates on Madagascar, showing several ecological traits. Note that, among sister genera, there tends to be clear differentiation in traits along at least 1 niche axis. Colors follow those in Figure 1. Trait column 1 is colored by substrate preference according to the first 3 principal components of substrate use; black represents terrestriality, green arboreality, orange fossoriality, and blue saxicoly. Trait column 2 is colored by the first 3 principal components of diet; pink represents higher insectivory, green represents higher frog-eating, browns represent a higher intake of hard-bodied versus soft-bodied prey. Trait column 3 is colored according to MDS coordinates of climate similarity between genera and its colors do not map clearly onto input climate variables. Trait column 4 is colored according to diel activity pattern; pink corresponds to diurnality, while blue corresponds to nocturnality. The final trait column visualizes the distribution of log-transformed body mass into ten discrete categories, shaded according to the proportion of species within a genus whose maximum body mass falls into a size bin.

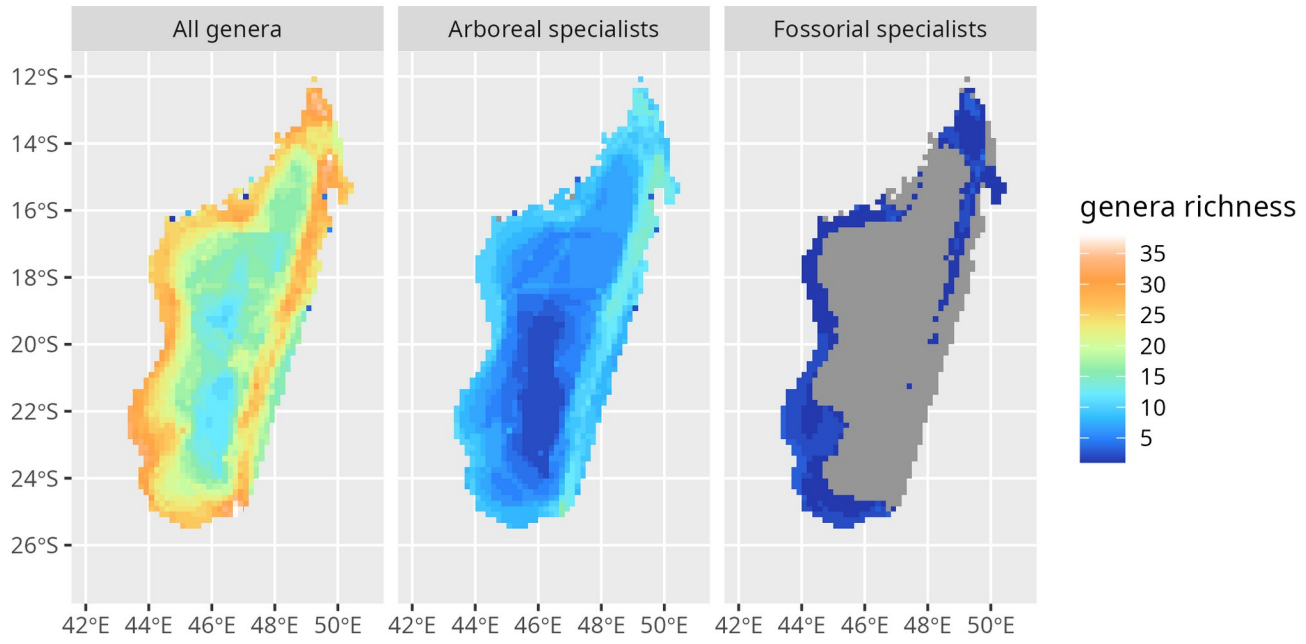

Figure 3: Maps of generic richness on Madagascar for 54 squamate genera, only arboreal specialists, and only fossorial specialists.

Yet fossoriality continues to evolve again and again. My results are consistent with the hypothesis that fossoriality tends to evolve in specific habitats that are geographically isolated from one another, preventing the geographic spread of fossorial lineages (Wiens et al. 2006; Bars Closel et al. 2017; Li and Wiens 2022). Highly fossorial genera do not tend to overlap geographically or climatically, but do tend to overlap in diet, body size, and diel activity pattern (Table 3), consistent with the hypothesis that niche space is restricted for fossorial lineages. However, the fact that there are many sites in Madagascar at which two (or rarely three) highly fossorial skink genera coexist, and that many potentially important white sands habitats across the island have never been sampled for these lizards (Miralles et al. 2026), suggest that the current understanding of fossorial skink diversity and biogeography on Madagascar is quite incomplete. The lack of a relationship between geographic range size and fossoriality suggests that niche restriction plays a larger role in depressing diversification than geographic isolation.

Niche diversity within fossorial squamates is underappreciated, but recent studies have begun to investigate sub-specialization within the limbless, fossorial ecomorph. Limbless, fossorial squamates

tend to fall into two ecomorphological paradigms; humicolous denizens of moist soils and leaf litter, and psammicolous sand-swimmers (Anelli et al. 2024; Miralles et al. 2026). Miralles et al. (2026) suggest that the pattern of biogeographic isolation within the psammophilous members of the *Zig + Paracontias* clade is a better fit for a model of evolution in which the clade's ancestral niche was humicolous, and that psammophily (those limited to sandy environments) has evolved several times in parallel within the clade during the aridification and cooling of Madagascar in the Mid-Miocene ~14 Ma during which the three other psammophilous Malagasy skink genera (*Pygomeles*, *Grandidierina*, and *Voeltzkowia*) originated.

The evolutionary dynamics found in Madagascar largely recapitulate patterns found on a global scale. The low phylogenetic signal in climatic niche overlap (Tables S5 & S7) suggests that climate is not a strong driver of overall patterns of species richness (Bars Closel et al. 2017; Li and Wiens 2022). Diversity is positively correlated with body size (Table S6), which is a key predictor of geographic range size (Feldman et al. 2016; Li and Wiens 2022). However, the finding that arboreality has no significant relationship to genus level diversity and net diversification rates is contrary to the family-level pattern across squamates (Bars Closel et al. 2017; Li and Wiens 2022). Without taking phylogeny into account, there is a significant positive relationship ($p = 0.035$) between percent arboreality and diversity within Malagasy squamate genera, but the concentration of highly arboreal genera into a few clades with high diversity (Gekkonidae, *Furcifer + Calumma*) likely means that this relationship cannot be distinguished from the general phylogenetic signal of diversity. Similarly, the finding of slightly depressed diversity and rates of net diversification in fossorial genera on Madagascar recapitulates the global trend (Bars Closel et al. 2017). In both Madagascar and at a global scale, substrate use is a likely a driver of differences in diversity, though this is likely clade-dependent (e.g., Burbrink et al. 2019).

Within clades pre-adapted to either fossoriality or arboreality, estimated rates of gain and loss of these specialist traits are nearly identical, with specialization evolving at a rate 4-6x faster than loss of specialization (Table 1). This suggests that the differential diversification in highly arboreal and highly fossorial clades is not due to differences in the ability of these groups to lose potentially constraining specializations, but is instead driven by fundamental constraints or opportunities of their particular niches. Specialization alone does not dictate evolutionary success; the dynamics of fossoriality and arboreality differ. The distribution of fossoriality tends to be bimodal, with clades either completely

non-fossorial or specializing highly for fossoriality; conversely, the distribution of arboreality is not so skewed, with many genera exhibiting intermediate levels of arboreality (Figure 2), consistent with the higher general rate of evolution. At large scales, arboreality may be indicative of general evolutionary lability in substrate use, which may in part contribute to its association with diversification rates. My finding that arboreality is significantly associated with decreased geographic range sizes is counterintuitive, but may be reflective of the particular geography of Madagascar, in which rain forests occur in a narrow band along the Eastern margin of the island.

Conclusions

Within a fauna largely insulated from exchange with continents, the evolution of squamate fossoriality has occurred largely along predicted paths; it has evolved several times in elongate-bodied clades, and resulted in extreme convergence. Fossoriality is a trait with unique evolutionary consequences; fossoriality limits the evolution of diet, body size, body shape, and activity pattern, and likely depresses diversification. Specialization for arboreality imposes none of these limits, despite having very similar evolutionary rates within pre-adapted clades. The largely non-overlapping geographic occurrence of fossorial genera is consistent with a hypothesis of multiple geographic origins of fossoriality, though the lack of an association between fossoriality and geographic range size, and the overlap of fossorial genera at many sites in Madagascar, suggests that fossoriality does not limit geographic spread over long timescales (the ~14 Ma since many of these fossorial clades originated).

Bibliography

Alibardi, L. 1997: Morphogenesis of the digital pad lamellae in the embryo of the lizard Anolis lineatopus. Journal of Zoology 243:47–55.

Alroy, J. 2015: Current extinction rates of reptiles and amphibians. Proceedings of the National Academy of Sciences 112:13003–13008.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990: Basic local alignment search tool. Journal of Molecular Biology 215:403–410.

Anelli, V., M. Bars-Closel, A. Herrel, and T. Kohlsdorf. 2024: Different selection regimes explain morphological evolution in fossorial lizards. Functional Ecology n/a.

Angel, F. 1935: Un Lézard nouveau de la famille des Dibamidés. Bulletin du Muséum national d'histoire naturelle 2:354–356.

Arida, E. 2017: Genetic Divergence and Phylogenetic Relationships Among Indonesian Species of Monitor Lizards of the Genus *Varanus* Based on Cytochrome Oxidase I Sequences. HAYATI Journal of Biosciences 24:41–45.

Axtell, R. W. 1958: A Northward Range Extension for the Lizard Anelytropsis papillosis, with Notes on the Distribution and Habits of Several Other Mexican Lizards. Herpetologica 14:189–191.

Barnosky, A. D., N. Matzke, S. Tomiya, G. O. U. Wogan, B. Swartz, T. B. Quental, C. Marshall, J. L. McGuire, E. L. Lindsey, K. C. Maguire, B. Mersey, and E. A. Ferrer. 2011: Has the Earth's sixth mass extinction already arrived? Nature 471:51–57.

Barros-Filho, J. D. de, M. A. de Freitas, T. F. S. Silva, M. F. de C. Loguercio, and M. C. C. Valverde. 2019: Redescription of Leposternon octostegum (Duméril, 1851), with an identification key for Brazilian Leposternon species, remarks on meristic methodology, and a proposal for pholidosis nomenclature (Squamata: Amphisbaenidae). Journal of Threatened Taxa 11:13058–13086.

Barrowclough, G. F., J. Cracraft, J. Klicka, and R. M. Zink. 2016: How Many Kinds of Birds Are There and Why Does It Matter? PLOS ONE 11:e0166307.

Bars Closel, M., T. Kohlsdorf, D. Moen, and J. Wiens. 2017: Diversification rates are more strongly related to microhabitat than climate in squamate reptiles (lizards and snakes). Evolution 71.

Behrensmeyer, A. K., S. M. Kidwell, and R. A. Gastaldo. 2000: Taphonomy and Paleobiology. Paleobiology 26:103–147.

Benton, M. J. 1989: Mass extinctions among tetrapods and the quality of the fossil record. Philosophical Transactions of the Royal Society of London. B, Biological Sciences 325:369–386.

Bickford, D., D. J. Lohman, N. S. Sodhi, P. K. L. Ng, R. Meier, K. Winker, K. K. Ingram, and I. Das. 2007: Cryptic species as a window on diversity and conservation. Trends in Ecology & Evolution 22:148–155.

Binfield, O., M. Camaiti, and L. Roberts. 2026: Drivers of tail evolution in squamates and their implications for the fossorial origin of snakes. The Anatomical Record 309.

Blair, C., B. P. Noonan, J. L. Brown, A. P. Raselimanana, M. Vences, and A. D. Yoder. 2015: Multilocus phylogenetic and geospatial analyses illuminate diversification patterns and the biogeographic history of Malagasy endemic plated lizards (Gerrhosauridae: Zonosaurinae). Journal of Evolutionary Biology 28:481–492.

Bleeker, P. 1860: Reptilien van Agam. Natuurkundig tijdschrift voor Nederlandsch Indië 20:325–389.

Blom, M. P. K., N. J. Matzke, J. G. Bragg, E. Arida, C. C. Austin, A. R. Backlin, M. A. Carretero, R. N. Fisher, F. Glaw, S. A. Hathaway, D. T. Iskandar, J. A. McGuire, B. R. Karin, S. B. Reilly, E. N.

Rittmeyer, S. Rocha, M. Sanchez, A. L. Stubbs, M. Vences, and C. Moritz. 2019: Habitat preference modulates trans-oceanic dispersal in a terrestrial vertebrate. Proceedings of the Royal Society B: Biological Sciences 286:20182575.
Boulenger, G. A. 1884: XVI.—Synopsis of the families of existing Lacertilia. Annals and Magazine of Natural History 14:117–122.
Brinkman, D. B., M. J. Ryan, and D. A. Eberth. 1998: The paleogeographic and stratigraphic distribution of ceratopsids (Ornithischia) in the upper Judith River Group of Western Canada. PALAIOS 13:160–169.
Brown, J. L., A. Cameron, A. D. Yoder, and M. Vences. 2014: A necessarily complex model to explain the biogeography of the amphibians and reptiles of Madagascar. Nature Communications 5:5046.
Buckley, L. B., and W. Jetz. 2007: Environmental and historical constraints on global patterns of amphibian richness. Proceedings of the Royal Society B: Biological Sciences 274:1167–1173.
Burbrink, F. T., S. Ruane, A. Kuhn, N. Rabibisoa, B. Randriamahatantsoa, A. P. Raselimanana, M. S. M. Andrianarimalala, J. E. Cadle, A. R. Lemmon, E. M. Lemmon, R. A. Nussbaum, L. N. Jones, R. Pearson, and C. J. Raxworthy. 2019: The Origins and Diversification of the Exceptionally Rich Gemsnakes (Colubroidea: Lamprophiidae: Pseudoxyrhophiinae) in Madagascar. Systematic Biology 68:918–936.
Burbrink, F. T., F. G. Grazziotin, R. A. Pyron, D. Cundall, S. Donnellan, F. Irish, J. S. Keogh, F. Kraus, R. W. Murphy, B. Noonan, C. J. Raxworthy, S. Ruane, A. R. Lemmon, E. M. Lemmon, and H. Zaher. 2020: Interrogating Genomic-Scale Data for Squamata (Lizards, Snakes, and Amphisbaenians) Shows no Support for Key Traditional Morphological Relationships. Systematic Biology 69:502–520.
Burgin, C. J., J. P. Colella, P. L. Kahn, and N. S. Upham. 2018: How many species of mammals are there? Journal of Mammalogy 99:1–14.
Burress, E. D., and M. M. Muñoz. 2022: Ecological Opportunity from Innovation, not Islands, Drove the Anole Lizard Adaptive Radiation. Systematic Biology 71:93–104.
Bushnell, B., J. Rood, and E. Singer. 2017: BBMerge – Accurate paired shotgun read merging via overlap. PLOS ONE 12:e0185056.
Butler, R. J., and P. M. Barrett. 2008: Palaeoenvironmental controls on the distribution of Cretaceous herbivorous dinosaurs. Naturwissenschaften 95:1027–1032.
Camacho, A., R. Recoder, M. Teixeira Jr, T. Kohlsdorf, M. T. Rodrigues, and M. S. Y. Lee. 2017: Overcoming phylogenetic and geographic uncertainties to test for correlates of range size evolution in gymnophthalmid lizards. Ecography 40:764–773.
Cao, Y., W. P. Williams, and A. W. Bark. 1997: Similarity Measure Bias in River Benthic Aufwuchs Community Analysis. Water Environment Research 69:95–106.
Capella-Gutiérrez, S., J. M. Silla-Martínez, and T. Gabaldón. 2009: trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973.
Casquet, J., Y. X. C. Bourgeois, C. Cruaud, F. Gavory, R. G. Gillespie, and C. Thébaud. 2015: Community assembly on remote islands: a comparison of Hawaiian and Mascarene spiders. Journal of Biogeography 42:39–50.
Ceballos, G., P. R. Ehrlich, A. D. Barnosky, A. García, R. M. Pringle, and T. M. Palmer. 2015: Accelerated modern human–induced species losses: Entering the sixth mass extinction. Science Advances 1:e1400253.
Chamberlain, S., D. Oldoni, V. Barve, P. Desmet, L. Geffert, D. Mcglinn, K. Ram, and J. Waller (0000-0002-7302-5976). 2026: rgbif: Interface to the Global Biodiversity Information Facility API. .

Chamberlain, S. A., and E. Szöcs. 2013: taxize: taxonomic search and retrieval in R. F1000Research 2:191.
Chang, C., P. Wu, R. E. Baker, P. K. Maini, L. Alibardi, and C.-M. Chuong. 2009: Reptile scale paradigm: Evo-Devo, pattern formation and regeneration. The International journal of developmental biology 53:813–826.
Chen, S., Y. Zhou, Y. Chen, and J. Gu. 2018: fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890.
Comanns, P., C. Effertz, F. Hischen, K. Staudt, W. Böhme, and W. Baumgartner. 2011: Moisture harvesting and water transport through specialized micro-structures on the integument of lizards. Beilstein Journal of Nanotechnology 2:204–214.
Cope, E. D. 1885: A Contribution to the Herpetology of Mexico. Proceedings of the American Philosophical Society 22:379–404.
Crooks, K. R., C. L. Burdett, D. M. Theobald, S. R. B. King, M. Di Marco, C. Rondinini, and L. Boitani. 2017: Quantification of habitat fragmentation reveals extinction risk in terrestrial mammals. Proceedings of the National Academy of Sciences of the United States of America 114:7635–7640.
Crottini, A., O. Madsen, C. Poux, A. Strauß, D. R. Vieites, and M. Vences. 2012: Vertebrate time-tree elucidates the biogeographic pattern of a major biotic change around the K–T boundary in Madagascar. Proceedings of the National Academy of Sciences 109:5358–5363.
Crottini, A., D. J. Harris, A. Miralles, F. Glaw, R. K. B. Jenkins, J. C. Randrianantoandro, A. M. Bauer, and M. Vences. 2015: Morphology and molecules reveal two new species of the poorly studied gecko genus Paragehyra (Squamata: Gekkonidae) from Madagascar. Organisms Diversity & Evolution 15:175–198.
Csárdi, G., and T. Nepusz. 2006: The igraph Software Package for Complex Network Research. InterJournal Complex Systems:1695.
Csárdi, G., T. Nepusz, K. Müller, S. Horvát, V. Traag, F. Zanini, and D. Noom. 2025: igraph for R: R interface of the igraph library for graph theory and network analysis. .
Da Silva, M. B., and T. C. S. Ávila-Pires. 2013: The genus *Cnemidophorus* (Squamata: Teiidae) in State of Piauí, northeastern Brazil, with description of a new species. Zootaxa 3681.
Darevsky, I. S. 1992: Two new species of the worm-like lizard Dibamus (Sauria: Dibamidae) with remarks of the distribution and ecology of Dibamus in Vietnam. Asiatic Herpetological Research 4:1–12.
Das, I., and N. S. Yaakob. 2003: A New Species of Dibamus (Squamata: Dibamidae) from Peninsular Malaysia. The Raffles Bulletin of Zoology 51:143–147.
Diestel, R. 2017: Planar Graphs. Pp.89–118 *in* R. Diestel, ed. Graph Theory. Springer, Berlin, Heidelberg.
Dujsebayeva, T. N., N. Ananjeva, and A. M. Bauer. 2021: Scale Microstructures of Pygopodid Lizards (Gekkota: Pygopodidae): Phylogenetic Stability and Ecological Plasticity. Russian Journal of Herpetology 28:291–308.
Duméril, C., G. Bibron, and A. H. A. Duméril. 1834: Erpétologie générale, ou, Histoire naturelle complète des reptiles. Roret, Paris :, p.
van Els, P., L. Herrera-Alsina, A. L. Pigot, and R. S. Etienne. 2021: Evolutionary dynamics of the elevational diversity gradient in passerine birds. Nature Ecology & Evolution 5:1259–1265.
ESRI. 2022: ArcGIS Desktop. .
Esteve-Altava, B., and D. Rasskin-Gutman. 2014: Theoretical morphology of tetrapod skull networks. Comptes Rendus Palevol 13:41–50.

Evans, B. J., R. M. Brown, J. A. McGuire, J. Supriatna, N. Andayani, A. Diesmos, D. Iskandar, D. J. Melnick, and D. C. Cannatella. 2003: Phylogenetics of Fanged Frogs: Testing Biogeographical Hypotheses at the Interface of the Asian and Australian Faunal Zones. Systematic Biology 52:794–819.

Faith, D. P. 1992: Conservation evaluation and phylogenetic diversity. Biological Conservation 61:1–10.

Feldman, A., N. Sabath, R. A. Pyron, I. Mayrose, and S. Meiri. 2016: Body sizes and diversification rates of lizards, snakes, amphisbaenians and the tuatara. Global Ecology and Biogeography 25:187–197.

Fleck, N., T. Glaw, J. Rafanoharana, A. René de Roland, H. Maheritafika, M. Vences, and F. Glaw. 2026: Arboreality in the typhlopid Madatyphlops ocularis from north-eastern Madagascar, with a review of climbing behaviour in scolecophidian snakes. Australian Journal of Zoology:ZO25060.

Foote, M. 1997: Estimating taxonomic durations and preservation probability. Paleobiology 23:278–300.

———. 2000: Origination and extinction components of taxonomic diversity: General problems. Paleobiology 26:74–102.

Freckleton, R. P., P. H. Harvey, and M. Pagel. 2002: Phylogenetic Analysis and Comparative Data: A Test and Review of Evidence. The American Naturalist 160:712–726.

Gans, C. 1971: Studies on amphisbaenians (Amphisbaenia, Reptilia). 4, A review of the amphisbaenid genus Leposternon. Bulletin of the AMNH ; v. 144, article 6. .

Geange, S. W., S. Pledger, K. C. Burns, and J. S. Shima. 2011: A unified analysis of niche overlap incorporating data of different types. Methods in Ecology and Evolution 2:175–184.

Gippner, S., S. L. Travers, M. D. Scherz, T. J. Colston, M. L. Lyra, A. V. Mohan, M. Multzsch, S. V. Nielsen, L. Rancilhac, F. Glaw, A. M. Bauer, and M. Vences. 2021: A comprehensive phylogeny of dwarf geckos of the genus *Lygodactylus*, with insights into their systematics and morphological variation. Molecular Phylogenetics and Evolution 165:107311.

Glaw, F., and M. Vences. 2007: A Field Guide to the Amphibians and Reptiles of Madagascar. Vences & Glaw, p.

Greenbaum, E., A. M. Bauer, and T. R. Jackman. 2007: Homopholis and blaesodactylus (Squamata: Gekkonidae) revisited: New insights from a molecular phylogeny. African Journal of Herpetology 56:101–114.

Greer, A. E. 1985: The Relationships of the Lizard Genera Anelytropsis and Dibamus. Journal of Herpetology 19:116–156.

Grismer, L., P. Geissler, N. Thy, T. Hartmann, P. Wagner, and N. Poyarkov. 2021: Molecular phylogenetics, PCA, and MFA recover a new species of Cyrtodactylus (Squamata: Gekkonidae) from an isolated sandstone massif in northwestern Cambodia. Zootaxa 4949:261–288.

Grismer, L. L. 2025: Introducing multiple factor analysis (MFA) as a diagnostic taxonomic tool complementing principal component analysis (PCA). ZooKeys 1248:93–109.

Haan, W. de, W. de Haan, P. W. Korthals, S. Müller, H. Schlegel, C. J. Temminck, and N. C. in Oost-Indië (Netherlands). 1839: Verhandelingen over de natuurlijke geschiedenis der Nederlandsche overzeesche bezittingen. [v.1] Land-en Volkenkunde. In commissie bij. S. en J. Luchtmans en C.C. van der Hoek, Leiden, 1–669 p.

Hackel, J., M. S. Vorontsova, O. P. Nanjarisoa, R. C. Hall, J. Razanatsoa, P. Malakasi, and G. Besnard. 2018: Grass diversification in Madagascar: In situ radiation of two large C3 shade clades and

support for a Miocene to Pliocene origin of C4 grassy biomes. Journal of Biogeography 45:750–761.
Hawlitschek, O., M. D. Scherz, B. Ruthensteiner, A. Crottini, and F. Glaw. 2018: Computational molecular species delimitation and taxonomic revision of the gecko genus Ebenavia Boettger, 1878. The Science of Nature 105:49.
Hijmans, R. J., M. Barbosa, A. Ghosh, and A. Mandel. 2025: geodata. .
Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005: Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965–1978.
Hijmans, R. J., M. Barbosa, R. Bivand, A. Brown, M. Chirico, E. Cordano, K. Dyba, E. Pebesma, B. Rowlingson, and M. D. Sumner. 2026: terra: Spatial Data Analysis. .
Hoffmann, J., S. Donoughe, K. Li, M. K. Salcedo, and C. H. Rycroft. 2018: A simple developmental model recapitulates complex insect wing venation patterns. Proceedings of the National Academy of Sciences 115:9905–9910.
Holland, S., and K. M. Loughney. 2021: The Stratigraphic Paleobiology of Nonmarine Systems. *in Elements of Paleontology*. Cambridge University Press, p.
Holland, S. M. 2016: The non-uniformity of fossil preservation. Philosophical Transactions of the Royal Society B: Biological Sciences 371:20150130.
———. 2022: The structure of the nonmarine fossil record: predictions from a coupled stratigraphic–paleoecological model of a coastal basin. Paleobiology 48:372–396.
Honda, M., J. Nabhitabhata, H. Ota, and T. Hikida. 1997: A New Species of Dibamus (Squamata: Dibamidae) from Thailand. The Raffles Bulletin of Zoology 45:275–279.
Hothorn, T., A. Zeileis, R. W. Farebrother (pan.f), C. Cummins (pan.f), G. Millo, and D. Mitchell. 2022: lmtest: Testing Linear Regression Models. .
Huey, R. B. 1974: Ecological Character Displacement in a Lizard. Integrative and Comparative Biology 14:1127–1136.
Huie, J. M., I. Prates, R. C. Bell, and K. de Queiroz. 2021: Convergent patterns of adaptive radiation between island and mainland Anolis lizards. Biological Journal of the Linnean Society 134:85–110.
Husson, F., and J. Josse. 2026: missMDA: Handling Missing Values with Multivariate Data Analysis. .
Hutter, C. R., K. A. Cobb, D. M. Portik, S. L. Travers, P. L. Wood Jr., and R. M. Brown. 2022: FrogCap: A modular sequence capture probe-set for phylogenomics and population genetics for all frogs, assessed across multiple phylogenetic scales. Molecular Ecology Resources 22:1100–1119.
Ineich, I. 1999: A New Species of Dibamus (Reptilia, Squamata, Dibamidae) from Vietnam. Bulletin de la Societe Zoologique de France 124:279–286.
Irigoien, I., F. Mestres, and C. Arenas. 2017: Weighted Distance Based Discriminant Analysis: The R Package WeDiBaDis. The R Journal 8:434–450.
Jarvis, A., H. I. Reuter, A. Nelson, and E. Guevara. 2026: Hole-filled seamless SRTM data V4. .
Jass, C. N., and C. O. George. 2010: An assessment of the contribution of fossil cave deposits to the Quaternary paleontological record. Quaternary International 217:105–116.
Jenny, J.-P., S. Koirala, I. Gregory-Eaves, P. Francus, C. Niemann, B. Ahrens, V. Brovkin, A. Baud, A. E. K. Ojala, A. Normandeau, B. Zolitschka, and N. Carvalhais. 2019: Human and climate global-scale imprint on sediment transfer during the Holocene. Proceedings of the National Academy of Sciences 116:22972–22976.

Jetz, W., and R. A. Pyron. 2018a: The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. Nature Ecology & Evolution 2:850–858.
———. 2018b: The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. Nature Ecology & Evolution 2:850–858.
Jetz, W., G. H. Thomas, J. B. Joy, K. Hartmann, and A. O. Mooers. 2012: The global diversity of birds in space and time. Nature 491:444–448.
Karin, B. R., T. Gamble, and T. R. Jackman. 2020: Optimizing Phylogenomics with Rapidly Evolving Long Exons: Comparison with Anchored Hybrid Enrichment and Ultraconserved Elements. Molecular Biology and Evolution 37:904–922.
Katoh, K., K. Misawa, K. Kuma, and T. Miyata. 2002: MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research 30:3059–3066.
Kling, M. 2023: colors3d: Generate 2D and 3D Color Palettes. .
Kliukin, N., A. Bragin, T. Nguyen, D. VU, L. Xuan Son, V. Gorin, I. Krone, and N. Poyarkov. 2026: A new species of blind skink (Squamata: Dibamidae: Dibamus) from Chu Yang Sin National Park, southern Vietnam, with the first description of a Dibamus clutch. Zootaxa 5760:501–528.
Kliukin, N. S., T. V. Nguyen, S. X. Le, A. M. Bragin, T. T. V. Tran, V. A. Gorin, and N. A. Poyarkov. 2023: A new species of the genus Dibamus Duméril & Bibron, 1839 (Squamata: Dibamidae) from the driest and hottest place of Vietnam. Zootaxa 5380:301–320.
Kliukin, N. S., A. M. Bragin, T. Neang, V. A. Gorin, A. A. Lisenkova, V. Y. Bogatyreva, and N. A. Poyarkov. 2024a: A new species of blind skink of the genus Dibamus Duméril ET Bibron, 1839 (Squamata, Dibamidae) from Cambodia. Zoologičeskij žurnal 103:97–120.
Kliukin, N. S., A. M. Bragin, T. V. Nguyen, S. X. Le, T. T. V. Tran, V. A. Gorin, and N. A. Poyarkov. 2024b: Another new species of Dibamus Duméril & Bibron, 1839 (Squamata: Dibamidae) from Nui Chua National Park, Ninh Thuan Province, Vietnam. Zootaxa 5406:87–104.
Kliukin, N. S., T. V. Nguyen, P. Pawangkhanant, H. M. Pham, S. X. Le, V. A. Gorin, C. Bos, I. W. Krone, and N. A. Poyarkov. 2025: A new species of Dibamus from the Central Highlands of Vietnam with redescription of Dibamus montanus Smith, 1921 (Squamata: Dibamidae). Zootaxa 5693:1–31.
Knoll, A. H., and K. J. Niklas. 1987: Adaptation, plant evolution, and the fossil record. Review of Palaeobotany and Palynology 50:127–149.
Koppetsch, T., W. Böhme, and A. Koch. 2019: A new species of Dibamus Duméril & Bibron, 1839 (Squamata: Dibamidae) from Pulau Manado Tua, Northern Sulawesi, Indonesia. Zootaxa 4555:331.
Kozak, K. H., and J. J. Wiens. 2010: Niche Conservatism Drives Elevational Diversity Patterns in Appalachian Salamanders. The American Naturalist 176:40–54.
Krone, I. W., K. M. Magoulick, and R. M. Yohler. 2024: All the Earth will not remember: how geographic gaps structure the record of diversity and extinction. Paleobiology 50:214–225.
Ksepka, D. T., and C. A. Boyd. 2012: Quantifying historical trends in the completeness of the fossil record and the contributing factors: an example using Aves. Paleobiology 38:112–125.
Lampe, F. von, and S. Jenny. 2026: goeveg: Functions for Community Data and Ordinations. .
Lawing, A. M., J. J. Head, and P. D. Polly. 2012: The Ecology of Morphology: The Ecometrics of Locomotion and Macroenvironment in North American Snakes. Pp.117–146 *in* J. Louys, ed. Paleontology in Ecology and Conservation. Springer, Berlin, Heidelberg.
Leviton, A. E., R. M. Brown, and C. D. Siler. 2011: The Dangerously Venomous Snakes of the Philippine Archipelago. Pp.473–530 *in* The Coral Triangle: The 2011 Hearst Biodiversity Philippine Expedition. California Academy of Sciences.

Lewis, P. O. 2001: A Likelihood Approach to Estimating Phylogeny from Discrete Morphological Character Data. Systematic Biology 50:913–925.
Lewis, S. L., and M. A. Maslin. 2015: Defining the Anthropocene. Nature 519:171–180.
Li, P., and J. J. Wiens. 2022: What drives diversification? Range expansion tops climate, life history, habitat and size in lizards and snakes. Journal of Biogeography 49:237–247.
Linkem, C. W., R. M. Brown, C. D. Siler, B. J. Evans, C. C. Austin, D. T. Iskandar, A. C. Diesmos, J. Supriatna, N. Andayani, and J. A. McGuire. 2013: Stochastic faunal exchanges drive diversification in widespread Wallacean and Pacific island lizards (Squamata: Scincidae: Lamprolepis smaragdina). Journal of Biogeography 40:507–520.
Liu, Q., J.-W. Yan, S.-B. Hou, P. Wang, S. N. Nguyen, R. W. Murphy, J. Che, and P. Guo. 2020: A new species of the genus Sinomicrurus (Serpentes: Elapidae) from China and Vietnam. Zoological Research 41:194–198.
Losos, J. B. 1992: The Evolution of Convergent Structure in Caribbean Anolis Communities. Systematic Biology 41:403.
Lubis, M. Z., S. Hidayat, and S. Hu. 2025: Indonesian Throughflow in the Halmahera Sea: A Review. Journal of Marine Science and Engineering 13:1974.
Lundelius, E. L. 2006: Cave site contributions to vertebrate history. Alcheringa: An Australasian Journal of Palaeontology 30:195–210.
Mangiafico, S. 2026: rcompanion: Functions to Support Extension Education Program Evaluation. .
Mannion, P. D., A. A. Chiarenza, P. L. Godoy, and Y. N. Cheah. 2019: Spatiotemporal sampling patterns in the 230 million year fossil record of terrestrial crocodylomorphs and their impact on diversity. Palaeontology 62:615–637.
Marder, E., T. M. Smiley, B. J. Yanites, and K. Kravitz. 2025: Direct effects of mountain uplift and topography on biodiversity. Science 387:1287–1291.
Markwick, P. J. 1998: Crocodilian diversity in space and time: the role of climate in paleoecology and its implication for understanding K/T extinctions. Paleobiology 24:470–497.
Martinez-Solano, I., J. Teixeira, D. Buckley, and M. Garcia-Paris. 2006: Mitochondrial DNA phylogeography of Lissotriton boscai (Caudata, Salamandridae): evidence for old, multiple refugia in an Iberian endemic. Molecular Ecology 15(11): 3375-3388.
Mazerolle, M. J. 2023: AICcmodavg: Model Selection and Multimodel Inference Based on (Q)AIC(c). .
McCallum, M. L. 2007: Amphibian Decline or Extinction? Current Declines Dwarf Background Extinction Rate. Journal of Herpetology 41:483–491.
Mcguire, J. A., X. Huang, S. B. Reilly, D. T. Iskandar, C. Y. Wang-Claypool, S. Werning, R. A. Chong, S. Z. S. Lawalata, A. L. Stubbs, J. H. Frederick, R. M. Brown, B. J. Evans, U. Arifin, A. Riyanto, A. Hamidy, E. Arida, M. S. Koo, J. Supriatna, N. Andayani, and R. Hall. 2023: Species Delimitation, Phylogenomics, and Biogeography of Sulawesi Flying Lizards: A Diversification History Complicated by Ancient Hybridization, Cryptic Species, and Arrested Speciation. Systematic Biology 72:885–911.
Meiri, S. 2024: SquamBase—A database of squamate (Reptilia: Squamata) traits. Global Ecology and Biogeography 33:e13812.
Meiri, S., A. M. Bauer, A. Allison, F. Castro-Herrera, L. Chirio, G. Colli, I. Das, T. M. Doan, F. Glaw, L. L. Grismer, M. Hoogmoed, F. Kraus, M. LeBreton, D. Meirte, Z. T. Nagy, C. de C. Nogueira, P. Oliver, O. S. G. Pauwels, D. Pincheira-Donoso, G. Shea, R. Sindaco, O. J. S. Tallowin, O. Torres-Carvajal, J.-F. Trape, P. Uetz, P. Wagner, Y. Wang, T. Ziegler, and U. Roll. 2018: Extinct,

obscure or imaginary: The lizard species with the smallest ranges. Diversity and Distributions 24:262–273.
Mestre Tomás, J. 2025: Supplementary material for the preprint “GeoThinneR: An R Package for Efficient Spatial Thinning of Species Occurrences and Point Data.” .
Michielsen, N. M., S. M. Goodman, V. Soarimalala, A. A. E. van der Geer, L. M. Dávalos, G. I. Saville, N. Upham, and L. Valente. 2023: The macroevolutionary impact of recent and imminent mammal extinctions on Madagascar. Nature Communications 14:14.
Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. 2020: IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution 37:1530–1534.
Miralles, A., J. Köhler, D. R. Vieites, F. Glaw, and M. Vences. 2011: Hypotheses on rostral shield evolution in fossorial lizards derived from the phylogenetic position of a new species of Paracontias (Squamata, Scincidae). Organisms Diversity & Evolution 11:135–150.
Miralles, A., J. Marin, D. Markus, A. Herrel, S. B. Hedges, and N. Vidal. 2018: Molecular evidence for the paraphyly of Scolecophidia and its evolutionary implications. Journal of Evolutionary Biology 31:1782–1793.
Miralles, A., R. Schmidt, F. Belluardo, N. A. Rahagalala, E. Monvoisin, F. M. Ratsoavina, J. Köhler, F. Glaw, and M. Vences. 2026: Zigzags in the White Sand Belt: A new, highly divergent lineage of sand-swimmer skink from Madagascar (Squamata: Scincidae). Megataxa 19:176–212.
Montiel-Veranza, J., V. Vásquez Cruz, and I. Morales-Leal. 2022: Nuevo Registro De Anelytropsis Papillosus Cope, 1885 (Squamata: Dibamidae) Para El Municipio De Comapa, En El Centro-Oeste De Veracruz, México. Revista Latinoamericana de Herpetología 5:139–141.
Moreno-Arias, R. A., and M. L. Calderón-Espinosa. 2016: Patterns of morphological diversification of mainland Anolis lizards from northwestern South America. Zoological Journal of the Linnean Society 176:632–647.
Murali, G., R. Gumbs, S. Meiri, and U. Roll. 2021: Global determinants and conservation of evolutionary and geographic rarity in land vertebrates. Science Advances 7:eabe5582.
Naimi, B. 2023: usdm: Uncertainty Analysis for Species Distribution Models. .
Newell, N. D. 1959: Adequacy of the Fossil Record. Journal of Paleontology 33:488–499.
Noonan, B. P., and P. T. Chippindale. 2006: Dispersal and vicariance: The complex evolutionary history of boid snakes. Molecular Phylogenetics and Evolution 40:347–358.
Nugraha, A. M. S., and R. Hall. 2018: Late Cenozoic palaeogeography of Sulawesi, Indonesia. Palaeogeography, Palaeoclimatology, Palaeoecology 490:191–209.
Nyberg, B., and J. A. Howell. 2015: Is the present the key to the past? A global characterization of modern sedimentary basins. Geology 43:643–646.
Oksanen, J., G. L. Simpson, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O’Hara, P. Solymos, M. H. H. Stevens, E. Szoecs, H. Wagner, M. Barbour, M. Bedward, B. Bolker, D. Borcard, G. Carvalho, M. Chirico, M. D. Caceres, S. Durand, H. B. A. Evangelista, R. FitzJohn, M. Friendly, B. Furneaux, G. Hannigan, M. O. Hill, L. Lahti, D. McGlinn, M.-H. Ouellette, E. R. Cunha, T. Smith, A. Stier, C. J. F. T. Braak, and J. Weedon. 2022: vegan: Community Ecology Package. .
Pagel, M. 1994: Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. Proceedings of the Royal Society B: Biological Sciences 255:37–45.

Palkovacs, E. P., M. T. Kinnison, C. Correa, C. M. Dalton, and A. P. Hendry. 2012: Fates beyond traits: ecological consequences of human-induced trait change: Fates beyond traits. Evolutionary Applications 5:183–191.
Paradis, E., and K. Schliep. 2019: ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528.
Pedersen, B. S., and A. R. Quinlan. 2018: Mosdepth: quick coverage calculation for genomes and exomes. Bioinformatics 34:867–868.
Pepper, M., S. Y. W. Ho, M. K. Fujita, and J. Scott Keogh. 2011: The genetic legacy of aridification: Climate cycling fostered lizard diversification in Australian montane refugia and left low-lying deserts genetically depauperate. Molecular Phylogenetics and Evolution 61:750–759.
Pereira, A. G., and C. G. Schrago. 2017: Arrival and diversification of mabuyine skinks (Squamata: Scincidae) in the Neotropics based on a fossil-calibrated timetree. PeerJ 5:e3194.
Plotnick, R. E., F. A. Smith, and S. K. Lyons. 2016: The fossil record of the sixth extinction. Ecology Letters 19:546–553.
Poggio, L., L. M. de Sousa, N. H. Batjes, G. B. M. Heuvelink, B. Kempen, E. Ribeiro, and D. Rossiter. 2021: SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 7:217–240.
Prasetyo, M. D., D. S. Yudha, A. A. T. Amarasinghe, I. Ineich, G. R. Gillespie, and A. Riyanto. 2025: A New Blind Skink (Reptilia: Dibamidae: Dibamus) from Buton Island, Indonesia. Taprobanica 14.
Prum, R. O., J. S. Berv, A. Dornburg, D. J. Field, J. P. Townsend, E. M. Lemmon, and A. R. Lemmon. 2015: A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526:569–573.
Pyron, R. A., F. T. Burbrink, and J. J. Wiens. 2013: A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. BMC Evolutionary Biology 13:93.
Quah, E. S. H., A. M. Shahrul, L. L. Grismer, and R. Grassby-Lewis. 2017: A new species of Dibamus Duméril & Bibron 1839 (Squamata: Dibamidae) from a hill station in Peninsular Malaysia. Raffles Bulletin of Zoology 65:10.
Rabosky, D. L., G. J. Slater, and M. E. Alfaro. 2012: Clade Age and Species Richness Are Decoupled Across the Eukaryotic Tree of Life. PLOS Biology 10:e1001381.
Rahbek, C., M. K. Borregaard, A. Antonelli, R. K. Colwell, B. G. Holt, D. Nogues-Bravo, C. M. Ø. Rasmussen, K. Richardson, M. T. Rosing, R. J. Whittaker, and J. Fjeldså. 2019: Building mountain biodiversity: Geological and evolutionary processes. Science 365:1114–1119.
Raja, N. B., E. M. Dunne, A. Matiwane, T. M. Khan, P. S. Nätscher, A. M. Ghilardi, and D. Chattopadhyay. 2021: Colonial history and global economics distort our understanding of deep-time biodiversity. Nature Ecology & Evolution 6:145–154.
Raup, D. M., and J. J. Sepkoski. 1982: Mass Extinctions in the Marine Fossil Record. Science 215:1501–1503.
Razafindratsima, O. H., S. Mehtani, and A. E. Dunham. 2013: Extinctions, traits and phylogenetic community structure: insights from primate assemblages in Madagascar. Ecography 36:47–56.
Reilly, S. B., H. Kaiser, B. R. Karin, U. Arifin, A. L. Stubbs, E. Arida, A. Hamidy, D. T. Iskandar, and J. A. McGuire. 2025: Rampant Dispersal Without Gene-Flow: Reproductively and Geographically Isolated Lineages of the Supertramp Lizard Lamprolepis smaragdina Permeate the Lesser Sunda Archipelago. Journal of Biogeography 52:e15087.
Reilly, S. B., A. L. Stubbs, B. R. Karin, E. Arida, U. Arifin, A. Hamidy, H. Kaiser, K. Bi, A. Riyanto, D. T. Iskandar, and J. A. McGuire. 2023: Bewildering biogeography: Waves of dispersal and

diversification across southern Wallacea by bent-toed geckos (genus: *Cyrtodactylus*). Molecular Phylogenetics and Evolution 186:107853.
Reusser, L., P. Bierman, and D. Rood. 2015: Quantifying human impacts on rates of erosion and sediment transport at a landscape scale. Geology 43:171–174.
Revell, L. J. 2024: phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). PeerJ 12:e16505.
Revell, L. J., K. P. Schliep, D. L. Mahler, and T. Ingram. 2024: Testing for heterogeneous rates of discrete character evolution on phylogenies. Journal of Evolutionary Biology 37:1591–1602.
Rocha, S., D. Posada, M. A. Carretero, and D. J. Harris. 2007: Phylogenetic affinities of Comoroan and East African day geckos (genus Phelsuma): Multiple natural colonisations, introductions and island radiations. Molecular Phylogenetics and Evolution 43:685–692.
Roelants, K., D. J. Gower, M. Wilkinson, S. P. Loader, S. D. Biju, K. Guillaume, L. Moriau, and F. Bossuyt. 2007: Global patterns of diversification in the history of modern amphibians. Proceedings of the National Academy of Sciences 104:887–892.
Ruta, M., and M. I. Coates. 2007: Dates, nodes and character conflict: Addressing the Lissamphibian origin problem. Journal of Systematic Palaeontology 5:69–122.
Safonova, Y., A. Bankevich, and P. A. Pevzner. 2015: dipSPAdes: Assembler for Highly Polymorphic Diploid Genomes. Journal of Computational Biology 22:528–545.
Saitou, N., and M. Nei. 1987: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406–425.
Sakata, S., and T. Hikida. 2003: A new fossorial scincine lizard of the genus Pseudoacontias (Reptilia: Squamata: Scincidae) from Nosy Be, Madagascar. Amphibia-Reptilia 24:57–64.
Salcedo, M., J. Hoffmann, S. Donoughe, and L. Mahadevan. 2019: Computational analysis of size, shape and structure of insect wings. Biology Open 8:bio040774.
Scarpetta, S. G., B. R. Karin, A. Corl, and J. A. McGuire. 2025: The Evolutionary History of Dragon Lizards (Squamata: Agamidae) Revealed by Phylogenomics. Systematic Biology:syaf088.
Scherz, M. D., J. D. Daza, J. Köhler, M. Vences, and F. Glaw. 2017: Off the scale: a new species of fish-scale gecko (Squamata: Gekkonidae: Geckolepis) with exceptionally large scales. PeerJ 5:e2955.
Schlegel, H. 1858: 2. Natuurkundige Leercursus, Koninlijke Militarie Akademie. Pp.*in* Handleiding Tot de Beoefening Der Dierkunde. .
Schliep, K. P. 2011: phangorn: phylogenetic analysis in R. Bioinformatics 27:592–593.
Schultze, H.-P. 1995: Terrestrial biota in coastal marine deposits: fossil-Lagerstätten in the Pennsylvanian of Kansas, USA. Palaeogeography, Palaeoclimatology, Palaeoecology 19:255–273.
Sepkoski, J. J. 1984: A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. Paleobiology 10:246–267.
Simões, T. R., and R. A. Pyron. 2021: The Squamate Tree of Life. Bulletin of the Museum of Comparative Zoology 163:47–95.
Singhal, S., M. Grundler, G. Colli, and D. L. Rabosky. 2017: Squamate Conserved Loci (SqCL): A unified set of conserved loci for phylogenomics and population genetics of squamate reptiles. Molecular Ecology Resources 17:e12–e24.
Singhal, S., T. J. Colston, M. R. Grundler, S. A. Smith, G. C. Costa, G. R. Colli, C. Moritz, R. A. Pyron, and D. L. Rabosky. 2021: Congruence and Conflict in the Higher-Level Phylogenetics of Squamate Reptiles: An Expanded Phylogenomic Perspective. Systematic Biology 70:542–557.

Skeels, A., D. Esquerré, D. Lipsky, L. Pellissier, and L. M. Boschman. 2023: Elevational Goldilocks zone underlies the exceptional diversity of a large lizard radiation (Liolaemus; Liolaemidae). Evolution 77:2672–2686.
Smith, M. A. 1921: New or little-known Reptiles and Batrachians from Southern Annam (Indochina). Proceedings of the Zoological Society of London:431–432.
Song, H., D. B. Kemp, L. Tian, D. Chu, H. Song, and X. Dai. 2021: Thresholds of temperature change for mass extinctions. Nature Communications 12:4694.
Spalding, C., and P. M. Hull. 2021: Towards quantifying the mass extinction debt of the Anthropocene. Proceedings of the Royal Society B: Biological Sciences 288:20202332.
Streicher, J. W., and J. J. Wiens. 2017: Phylogenomic analyses of more than 4000 nuclear loci resolve the origin of snakes among lizard families. Biology Letters 13:20170393.
Stuart-Smith, R. D., C. Mellin, A. E. Bates, and G. J. Edgar. 2021: Habitat loss and range shifts contribute to ecological generalization among reef fishes. Nature Ecology & Evolution 5:656–662.
Tänzler, R., E. Toussaint, Y. Suhardjono, M. Balke, and A. Riedel. 2014: Multiple transgressions of Wallace's Line explain diversity of flightless Trigonopterus weevils on Bali. Proceedings of the Royal Society B: Biological Sciences 281:20132528.
Taylor, E. H. 1915: New Species of Philippine Lizards. Philippine Journal of Science 10:89–109.
Taylor, E. H. 1962: New Oriental Reptiles. The University of kansas Science Bulletin 43:209–263.
Tietje, M., and M.-O. Rödel. 2017: Contradicting habitat type-extinction risk relationships between living and fossil amphibians. Royal Society Open Science 4:170051.
Title, P. O., S. Singhal, M. C. Grundler, G. C. Costa, R. A. Pyron, T. J. Colston, M. R. Grundler, I. Prates, N. Stepanova, M. E. H. Jones, L. B. Q. Cavalcanti, G. R. Colli, N. Di-Poï, S. C. Donnellan, C. Moritz, D. O. Mesquita, E. R. Pianka, S. A. Smith, L. J. Vitt, and D. L. Rabosky. 2024: The macroevolutionary singularity of snakes. Science 383:918–923.
Toivonen, J., M. Fortelius, and I. Žliobaitė. 2022: Do species factories exist? Detecting exceptional patterns of evolution in the mammalian fossil record. Proceedings of the Royal Society B: Biological Sciences 289:20212294.
Tolley, K. A., T. M. Townsend, and M. Vences. 2013: Large-scale phylogeny of chameleons suggests African origins and Eocene diversification. Proceedings of the Royal Society B: Biological Sciences 280:20130184.
Tomović, L., M. A. Carretero, R. Ajtić, and J. Crnobrnja-Isailović. 2008: Evidence for post-natal instability of head scalation in the meadow viper (Vipera ursinii) – patterns and taxonomic implications. Amphibia-Reptilia 29:61–70.
Tonini, J. F. R., K. H. Beard, R. B. Ferreira, W. Jetz, and R. A. Pyron. 2016: Fully-sampled phylogenies of squamates reveal evolutionary patterns in threat status. Biological Conservation 204:23–31.
Townsend, T. M., D. H. Leavitt, and T. W. Reeder. 2011: Intercontinental dispersal by a microendemic burrowing reptile (Dibamidae). Proceedings of the Royal Society B: Biological Sciences 278:2568–2574.
Tzika, A. C., A. Ullate-Agote, S. Zakany, M. Kummrow, and M. C. Milinkovitch. 2023: Somitic positional information guides self-organized patterning of snake scales. Science Advances 9:eadf8834.
Uhen, M. D., B. Allen, N. Behboudi, M. E. Clapham, E. Dunne, A. Hendy, P. A. Holroyd, M. Hopkins, P. Mannion, P. Novack-Gottshall, C. Pimiento, and P. Wagner. 2023: Paleobiology Database User Guide Version 1.0. PaleoBios 40.

Upham, N. S., J. A. Esselstyn, and W. Jetz. 2019: Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. PLOS Biology 17:e3000494.
Valdez-Villaviceno, J. H., E. García-Padilla, and V. Mata-Silva. 2016: Anelytropsis papillosus Cope, 1885 (Squamata: Dibamidae), an overlooked species in the state of Oaxaca, Mexico. Mesoamerican Herpetology 3.
Wagner, P. J., and J. D. Marcot. 2013: Modelling distributions of fossil sampling rates over time, space and taxa: assessment and implications for macroevolutionary studies. Methods in Ecology and Evolution 4:703–713.
Wake, D. B. and V. T. V. 2008: Are we in the midst of the sixth mass extinction? A view from the world of amphibians. Proceedings of the National Academy of Sciences, U.S.A. (PNAS) 105: 11466-11473.
Wang, Z., W. Wu, J. Ren, C. Peng, D. Jiang, and J. Li. 2021: Evolution of Phenotype and Mitochondrial Genome Reveals Limbless and Body-elongated Squamates may Change Their Energy Basis for Locomotion. .
Warren, D., and R. Dinnage. 2025: ENMTools: Analysis of Niche Evolution using Niche and Distribution Models. .
Warren, D. L., R. E. Glor, and M. Turelli. 2008: Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. Evolution 62:2868–2883.
Weinell, J. L., W. R. Branch, T. J. Colston, T. R. Jackman, A. Kuhn, W. Conradie, and A. M. Bauer. 2019: A species-level phylogeny of *Trachylepis* (Scincidae: Mabuyinae) provides insight into their reproductive mode evolution. Molecular Phylogenetics and Evolution 136:183–195.
Wickert, A. D. 2016: Reconstruction of North American drainage basins and river discharge since the Last Glacial Maximum. Earth Surface Dynamics 4:831–869.
Wiens, J. J., M. C. Brandley, and T. W. Reeder. 2006: Why does a trait evolve multiple times within a clade? Repeated evolution of a snakelike body form in squamate reptiles. Evolution 60:123–141.
Wiens, J. J., Gabriela Parra-Olea, Mario Garcia-Paris, and David B. Wake. 2007: Phylogenetic history underlies elevational biodiversity patterns in tropical salamanders. Proceedings of the Royal Society B 274: 919-928.
Yao, L., and Y. Yuan. 2025: A Unified Method for Detecting Phylogenetic Signals in Continuous, Discrete, and Multiple Trait Combinations. Ecology and Evolution 15:e71106.
Yoder, A. D., and M. D. Nowak. 2006: Has Vicariance or Dispersal Been the Predominant Biogeographic Force in Madagascar? Only Time Will Tell. Annual Review of Ecology, Evolution, and Systematics 37:405–431.
Zheng, Y., and J. J. Wiens. 2016: Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4162 species. Molecular Phylogenetics and Evolution 94:537–547.

Appendix 1: Supplementary Materials for Chapter 2

Table S1. Metadata on all specimens included in this study

Specimen number	Orginal species assignment	Provisional species assignment	Tree name	Locality	Clade
ABV-01519	D. annae	D. annae	D_annae_ABV_01519	Vietnam, Gia Lai, Kon Ka Kinh	montanus
AMNH 1264	D. cf. celebensis	unassigned	Not included in tree	Central Core	cf. South & Central
AMNH 142976	D. cf. celebensis	unassigned	Not included in tree	Central Core	cf. South & Central
AMNH 142977	D. cf. celebensis	unassigned	Not included in tree	Central Core	cf. South & Central
AMNH 142978	D. cf. celebensis	unassigned	Not included in tree	Central Core	cf. South & Central
AMNH 142979	D. cf. celebensis	unassigned	Not included in tree	Central Core	cf. South & Central
FMNH 258662	D. montanus	D. greeri	D_greri_FMNH_258662	Champsak Province, Laos	bourreti
JAM 15211	D. celebensis	D. "katopasa"	D_"katopasa"_JAM_15211	Gunung Katopasa Camp 1	South & Central
JAM 15630	D. celebensis	D. "torumpupu"	D_"torumpupu"_JAM_15630	Gunung Torumpupu Camp 1, Kecamantan Kulawi	South & Central
JAM 17371	D. celebensis	D. “big klabat”	D_"klabat"_JAM_17371	Gunung Klabat, Sulawesi Utara	seramensis
JAM 17429	D. celebensis	D. "soputan"	D_"klabat"_JAM_17429	Gunung Klabat, Sulawesi Utara	Soputan
JAM 17430	D. celebensis	D. "soputan"	D_"klabat"_JAM_17430	Gunung Klabat, Sulawesi Utara	Soputan
JAM 2195/ TNHC 059521	D. seramensis	D. seramensis	D_seramensis_JAM_2195	Kecamatan Kairatu, Desa Lohiatala	seramensis
LSU 4676	D. tiomanensis	D. tiomanensis	D_tiomanensis_LSUHC_4676	Malaysia, Pulau Tioman	alfredi
MCZ 7623	D. seramensis	D. seramensis	Not included in tree	Seram Island	seramensis

Specimen number	Orginal species assignment	Provisional species assignment	Tree name	Locality	Clade
MVZ 253598	D. celebensis	D. "soputan"	D_"soputan"_MVZ_253598	Gunung Soputan, Kabupaten Minahasa Tengah	Soputan
MVZ 253599	D. celebensis	D. "tangkoko"	D_"tangkoko"_MVZ_253599	Tangkoko Nature Reserve, Kecamatan Bitung Tengah, Kabupaten Minahasa	seramensis
MVZ 253600	D. celebensis	D. "tangkoko 2"	D_"tangkoko"_MVZ_253600	Tangkoko Nature Reserve, Kecamatan Bitung Tengah, Kabupaten Minahasa	seramensis
MVZ 253601	D. celebensis	D. "tangkoko"	D_"gorontalo"_MVZ_253601	Desa Lombongo, Bogani Nani Wartabone Natl. Park, Kecamatan Suwawa, Kabupaten Bone Bolango	seramensis
MVZ 267791	D. noveaeguineae	D."tanahjampeah"	D_"tanahjampeah"_MVZ_267791	vicinity of Maminasa, Tanah Jampea Island	South & Central
MVZ 267792	D. noveaeguineae	D."tanahjampeah"	D_"tanahjampeah"_MVZ_267792	vicinity of Maminasa, Tanah Jampea Island	South & Central
MVZ 267793	D. noveaeguineae	D."tanahjampeah"	D_"tanahjampeah"_MVZ_267793	Telkom Tower Rd., Tanah Jampea Island	South & Central
MVZ 267794	D. noveaeguineae	D."tanahjampeah"	D_"tanahjampeah"_MVZ_267794	Telkom Tower Rd., Tanah Jampea Island	South & Central
MVZ 267795	D. noveaeguineae	D."tanahjampeah"	D_"tanahjampeah"_MVZ_267795	Telkom Tower Rd., Tanah Jampea Island	South & Central

Specimen number	Orginal species assignment	Provisional species assignment	Tree name	Locality	Clade
MVZ 267796	D. noveaeguineae	D."tanahjampeah"	D_"tanahjampeah"_MVZ_267796	Telkom Tower Rd., Tanah Jampea Island	South & Central
MVZ 267797	D. noveaeguineae	D."tanahjampeah"	D_"tanahjampeah"_MVZ_267797	Telkom Tower Rd., Tanah Jampea Island	South & Central
MVZ 267798	D. noveaeguineae	D."tanahjampeah"	D_"tanahjampeah"_MVZ_267798	Telkom Tower Rd., Tanah Jampea Island	South & Central
MVZ 284209	D. celebensis	D. "palopo"	D_"palopo"_MVZ_284209	Desa Puncak, Kota Palopo	South & Central
MVZ 284210	D. celebensis	D. "palopo"	D_"palopo"_MVZ_284210	Desa Puncak, Kota Palopo	South & Central
MVZ 284211	D. celebensis	D. "moramo cf oetamai"	D_"cf_oetamai"_MVZ_284211	Air Terjun Moramo, Desa Moramo, Kecamatan Moramo, Kabupaten Konawe Selatan	Southeast Sulawesi
MVZ 284212	D. celebensis	D. "moramo cf oetamai"	D_"cf_oetamai"_MVZ_284212	Air Terjun Moramo, Desa Moramo, Kecamatan Moramo, Kabupaten Konawe Selatan	Southeast Sulawesi
MVZ 284213	D. celebensis	D. "latawaro"	D_"latawaro"_MVZ_284213	Desa Latawaro, Kecamatan Lambai, Kabupaten Kolaka Utara	Southeast Sulawesi
MVZ 284214	D. celebensis	D. "latawaro"	D_"latawaro"_MVZ_284214_2	Desa Latawaro, Kecamatan Lambai, Kabupaten Kolaka Utara	Southeast Sulawesi

Specimen number	Orginal species assignment	Provisional species assignment	Tree name	Locality	Clade
MVZ 284215	D. celebensis	D. "latawaro"	D_"rante_limbong"_MVZ_284215	Desa Rante Limbong, Kecamatan Lasusua, Kabupaten Kolaka Utara	Southeast Sulawesi
MVZ 284216	D. celebensis	D. "tobela"	D_"tobela"_MVZ_284216	Desa Tobela, Kecamatan Porehu, Kabupaten Kolaka Utara	Southeast Sulawesi
MVZ 293570	D. taylori	D. taylori	D_taylori_Penida__MVZ_293570	Desa Tiagan at Puncak Mundi, Kecamatan Nusa Penida, Penida Island	taylori
MVZ 293571	D. taylori	D. “nusa penida”	D_taylori_Penida__MVZ_293571	Desa Tiagan at Puncak Mundi, Kecamatan Nusa Penida, Penida Island	taylori
MVZ 293572	D. taylori	D. “nusa penida”	D_taylori_Penida__MVZ_293572	Desa Tiagan at Puncak Mundi, Kecamatan Nusa Penida, Penida Island	taylori
MVZ 293573	D. taylori	D. “nusa penida”	D_taylori_Penida__MVZ_293573	Desa Tiagan at Puncak Mundi, Kecamatan Nusa Penida, Penida Island	taylori
MVZ 293574	D. taylori	D. “nusa penida”	D_taylori_Penida__MVZ_293574	Desa Tiagan at Puncak Mundi, Kecamatan Nusa Penida, Penida Island	taylori

Specimen number	Orginal species assignment	Provisional species assignment	Tree name	Locality	Clade
MVZ 293578	D. taylori	D. “nusa penida”	D_taylori_Penida__MVZ_293578	Desa Tiagan at Puncak Mundi, Kecamatan Nusa Penida, Penida Island	taylori
MVZ 293579	D. taylori	D. taylori	D_taylori_Sumba__MVZ_293579	Taman Nasional Laiwangi Wanggameti, Desa Priangkareha, Kecamatan Tabundung	taylori
MVZ 293580	D. taylori	D. taylori	D_taylori_Sumba__MVZ_293580	Taman Nasional Laiwangi Wanggameti, Desa Priangkareha, Kecamatan Tabundung	taylori
MVZ 293583	D. taylori	D. taylori	Not included in tree	Sumba	taylori
MVZ 293584	D. taylori	D. taylori	D_taylori_Sumba__MVZ_293584	Taman Nasional Laiwangi Wanggameti, Desa Priangkareha, Kecamatan Tabundung	taylori
MVZ 293585	D. taylori	D. taylori	Not included in tree	Sumba	taylori
MVZ 293586	D. taylori	D. taylori	Not included in tree	Sumba	taylori
MVZ 293587	D. taylori	D. taylori	D_taylori_Sumba__MVZ_293587	Taman Nasional Laiwangi Wanggameti, Desa Priangkareha, Kecamatan Tabundung	taylori
MVZ 293588	D. taylori	D. taylori	Not included in tree	Sumba	taylori
MVZ 293589	D. taylori	D. taylori	Not included in tree	Sumba	taylori

Specimen number	Orginal species assignment	Provisional species assignment	Tree name	Locality	Clade
MVZ 293590	D. taylori	D. taylori	D_taylori_Sumba__MVZ_293590	Taman Nasional Laiwangi Wanggameti, Desa Priangkareha, Kecamatan Tabundung	taylori
MVZ 293591	D. taylori	D. "wetar"	D_"wetar"_MVZ_293591	Desa Ilwaki, Kecamatan Wetar, Wetar Island	alfredi
MVZ 293593	D. taylori	D. "wetar"	D_"wetar"_MVZ_293593	Desa Ilwaki, Kecamatan Wetar, Wetar Island	alfredi
MZB 14728	D. manadotuaensis	D. manadotuaensis	Not included in tree	Manado Tua	manadotuaensis
MZB 4273	D. oetamai	D. oetamai	Not included in tree	Buton Island	cf. Southeast Sulawesi
MZB 4274	D. oetamai	D. oetamai	Not included in tree	Buton Island	cf. Southeast Sulawesi
MZB 4276	D. oetamai	D. oetamai	Not included in tree	Buton Island	cf. Southeast Sulawesi
MZB 4278	D. oetamai	D. oetamai	Not included in tree	Buton Island	cf. Southeast Sulawesi
MZB 4279	D. oetamai	D. oetamai	Not included in tree	Buton Island	cf. Southeast Sulawesi
NAP 16013	D. greeri	D. greeri	D_greeri_NAP_16013	Vietnam, Binh Dinh	bourreti
NAP-01283	D. kondaoensis	D. kondaoensis	D_kondaoensis_NAP_01283	Vietnam, Ba Ria - Vung Tau, Con Dao, Bay Canh (Hon Ba) island	montanus
NAP-01301	D. kondaoensis	D. kondaoensis	D_kondaoensis_NAP_01301	Vietnam, Ba Ria - Vung Tau, Con Dao, Con Son Island	montanus

Specimen number	Orginal species assignment	Provisional species assignment	Tree name	Locality	Clade
NAP-02535	D. deharvengi	D. deharvengi	D_deharvengi_NAP_02535	Vietnam, Ba Ria - Vung Tau, Binh Chau - Phuoc Buu	montanus
NAP-02983	D. kondaoensis	D. kondaoensis	D_kondaoensis_NAP_02983	Vietnam, Ba Ria - Vung Tau, Con Dao, Con Son Island, Nui Chua Mt.	montanus
NAP-03563	D. bourreti	D. bourreti	D_bourreti_NAP_03563	Vietnam, Hai Phong, Cat Ba	bourreti
NAP-07697	D. greeri	D. greeri	D_greeri_NAP_07697	Vietnam, Gia Lai, Kon Chu Rang	bourreti
NAP-10061	D. deharvengi	D. deharvengi	D_deharvengi_NAP_10061	Vietnam, Ba Ria - Vung Tau, Binh Chau - Phuoc Buu	montanus
NAP-10923	D. bourreti	D. bourreti	Not included in tree	Vietnam, Vinh Phuc: Vinh Yen District, Tam Dao	bourreti
NAP-11665	D. alfredi	D. alfredi	D_alfredi_NAP_11665	Thailand, Yala, Than To Waterfall	alfredi
NAP-11666	D. alfredi	D. alfredi	D_alfredi_NAP_11666	Thailand, Yala, Than To Waterfall	alfredi
NAP-11667	D. somsaki	D. somsaki	D_somsaki_NAP_11667	Thailand, Chanthaburi, Khao Khitchakut N.P.	montanus
NAP-11875	D. elephantinus	D. elephantinus	D_elephantinus_NAP_11875	Cambodia, Kampot, Phnum Bokor	montanus
NAP-11942	D. elephantinus	D. elephantinus	D_elephantinus_NAP_11942	Cambodia, Kampot, Phnum Bokor	montanus
NAP-12898	D. montanus	D. montanus	D_montanus_NAP_12898	Cambodia, Kampot, Phnum Bokor	montanus
NAP-12989	D. montanus	D. montanus	D_montanus_NAP_12989	Vietnam, Lam Dong, Di Linh	montanus

Specimen number	Orginal species assignment	Provisional species assignment	Tree name	Locality	Clade
NAP-13142	D. tropcentr	D. tropcentr	D_tropcentr_NAP_13142	Vietnam, Ninh Thuan, Nui Chua, Da Vach	tropcentr
NAP-13250	D. greeri	D. greeri	D_greeri_NAP_13250	Vietnam, Da Nang, Son Tra	bourreti
NAP-13640	D. deimontis	D. deimontis	D_deimontis_NAP_13640	Vietnam, Ninh Thuan, Nui Chua Mt.	tropcentr
NAP-13641	D. deimontis	D. deimontis	D_deimontis_NAP_13641	Vietnam, Ninh Thuan, Nui Chua Mt.	tropcentr
NAP-13712	D. tropcentr	D. tropcentr	D_tropcentr_NAP_13712	Vietnam, Ninh Thuan, Nui Chua, Nuoc Ngot	tropcentr
NAP-15871	D. sp. 1	D. "irregularis"	D_sp_1_NAP_15871	Vietnam, Dak Lak, Chu Yang Sin N.P.	tropcentr
NAP-15949	D. sp. 1	D. "irregularis"	D_sp_1_NAP_15949	Vietnam, Dak Lak, Chu Yang Sin N.P.	tropcentr
NAP-16249	D. cf. Taylori	D. “nusa penida”	D_"cf_taylori"_NAP_16249	Indonesia, Nusa Penida, Tembeling	taylori
NAP-16470	D. greeri	D. greeri	D_greeri_NAP_16470	Vietnam, Binh Dinh	bourreti
NAP-16587	D. bourreti	D. bourreti	D_bourreti_NAP_16587	Vietnam, Lang Son	bourreti
NCSM 99089	D. elephantinus	D. elephantinus	D_elephantinus_NCSM_99089	Bokor National Park, Kampot, Cambodia	montanus
RMB 15295/ KUH 331727	D. novaeguineae	D. cf argenteus	D_"siquijor"_RMB_15295	Siquijor	Philippines
RMB 20796/ KUH 334071	D. leucurus	D. cf argenteus	D_"mindanao"_RMB_15296	Mindanao	Philippines
ROM 36056	D. bourreti	D. bourreti	Not included in tree	Cao Bang, Quang Thanh Village	Unassigned
TNHC 15479	D. noveaguineae	D. “negros”	D_"negros"_TNHC_15479	Negros Island	Philippines

Specimen number	Orginal species assignment	Provisional species assignment	Tree name	Locality	Clade
TNHC 62475	D. novaeguineae	D. “negros”	D_"negros"_TNHC_62475	Negros Island	Philippines
UF 33073	D. taylori	D. taylori	Not included in tree	Komodo	Taylori
UF 33488	D. taylori	D. taylori	Not included in tree	Flores	Taylori
ZMFK 95559	D. manadotuaensis	D. manadotuaensis	Not included in tree	Manado Tua	manadotuaensis
ZMFK 95560	D. manadotuaensis	D. manadotuaensis	Not included in tree	Manado Tua	manadotuaensis
ZMMU NAP-14425	D. greeri	D. greeri	not included in tree	Vietnam, Thua Thien Hue, A Roang	Unassigned

Table S2. Morphological data for 42 dibamid specimens used in this study. All measurements are in millimeters

MVZ 253598	MVZ 224113	MVZ 224112	MCZ 7623	JAM 17430	JAM 17429	JAM 17371	Specimen
		F	F	F	M	M	Sex
180	123	168	203	140	121	183	SVL
21.34	53	46.2	22.33	19	18	10	TL
0.12	0.43	0.28	0.11	0.14	0.15	0.05	TL:SVL
5.66	3.59	4.57	7.27				BW
4.09	3.26	3.82					TW
5.46	3.03	na	5.12				HLL
4.25	2.25	3.27					Head depth
9.58	4.88	6.88	11.16				HL back of skull
5.75	3.98	3.07	6.2				HL from SL4
6.05	3.74	4.88	7.86				HW
2.89	1.49	2.13	3.78				EN
3.77	2.22	3.24	3.36				ES
1.29	1.16	1.47	1.59				IN
2.61	1.68	2.41	2.96				IO
1.07	0.8	0.84	0.87				FNSL
2.11	1.84	2.01	2.24				FNSW
1.17	1.13	1.73	0.99				FSL
2.69	2.33	2.74	1.55				FSW
0.8	1.66	2.05	1				IPW
0.86	0.73	0.92	0.81				NSW
1.27	1.27	1.36	0.69				Frontal:Frontonasal
0.93	2.27	2.23	1.23				Interparietal:Nuchal
34	28	28					ATSR
30	24	24	33				MBSR
26	22	26					PTSR
267	205	183	274	225	214	267	VEN
40	93	59+	40	37	38	`19+	SC
3	1	1	4	3	3	4	POS

MVZ 284212	MVZ 284209	MVZ 267797	MVZ 267795	MVZ 267793	MVZ 267791	MVZ 253600	Specimen
M	M						Sex
118	154	67	124	106	140	146	SVL
20.16	14.61	9.55	17.59	13.62	21.18	16.84	TL
0.17	0.09	0.14	0.14	0.13	0.15	0.12	TL:SVL
3.81	4.76	2.64	3.27	3.22	4.41	3.51	BW
3.39	3.56	1.61	2.8	2.37	3.7	3.11	TW
3.05	3.67	na	NA	na	3.69	3.52	HLL
2.57	2.99	1.79	2.55	2.14	2.86	2.89	Head depth
6.54	6.49	5.08	6.05	6.03	7.15	6.32	HL back of skull
3.92	4.79	2.5	3.38	3.36	4.61	4.74	HL from SL4
3.85	4.24	2.49	3.3	3.06	4.16	4.3	HW
2.04	2.23	1.33	1.88	1.65	2.39	2.41	EN
2.56	2.75	1.56	2.14	1.9	2.83	3.16	ES
1.14	1.31	0.78	0.83	0.68	0.88	1.11	IN
1.77	1.72	1.1	1.41	1.46	1.7	2.4	IO
0.77	1	0.37	0.71	0.63	0.7	0.88	FNSL
1.73	2.08	0.99	0.54	1.45	1.21	1.92	FNSW
0.96	1.32	0.59	1.05	0.62	0.76	0.76	FSL
1.86	2.34	1.27	1.68	1.59	0.73	1.29	FSW
0.51	1.06	0.37	0.52	0.61	0.73	0.67	IPW
0.59	0.53	0.37	0.46	0.6	0.61	0.61	NSW
1.08	1.13	1.28	3.11	1.10	0.60	0.67	Frontal:Frontonasal
0.86	2.00	1.00	1.13	1.02	1.20	1.10	Interparietal:Nuchal
24	30	25	29	25	26	32	ATSR
22	26	25	23	24	25	28	MBSR
21	26	19	21	21	21	24	PTSR
203	266	227	231	224	206	241	VEN
42	33	39	43	36	40	38	SC
2	2	3	3	2	3	3	POS

MVZ 293579	MVZ 293578	MVZ 293574	MVZ 293573	MVZ 293572	MVZ 284215	MVZ 284213	Specimen
					F		Sex
151	140	130	140	141	116	146	SVL
10.99	20.7	17.48	18.03	7.81+	13.9	11.86	TL
	0.15	0.13	0.13		0.12	0.08	TL:SVL
4.65	5.47	4.27	4.52	4.61	4.54	4.78	BW
3.21	3.42	3.5	3.92	3.6	3.09	3.91	TW
na	4.87	3.06	na	na	NA	4.47	HLL
3.22	3.17	3.25	3.16	3.36	2.85	3.43	Head depth
7.79	7.53	7.76	7.53	6.83	6.64	6.44	HL back of skull
4.81	4.41	4.2	4.37	4.9	3.92	4.9	HL from SL4
4.71	5.07	4.44	4.52	4.56	4.04	5	HW
2.5	2.15	2.1	2.61	2.44	1.86	2.62	EN
3.03	2.75	2.63	2.83	3.08	2.28	3.06	ES
0.85	0.91	0.91	1.11	1.03	1.11	1.21	IN
1.9	1.86	1.7	1.68	1.76	1.7	2.24	IO
0.83	1.05	0.76	0.93	1.03	0.63	1.21	FNSL
1.74	2.07	1.86	2.05	1.99	2.15	1.76	FNSW
0.5	0.69	0.69	0.74	1.05	0.64	0.51	FSL
1.09	0.91	1.04	0.9	0.8	1.25	1.23	FSW
0.85	0.73	0.71	0.9	0.6	0.67	0.83	IPW
0.73	0.88	0.67	0.63	0.68	0.64	0.61	NSW
0.63	0.44	0.56	0.44	0.40	0.58	0.70	Frontal:Frontonasal
1.16	0.83	1.06	1.43	0.88	1.05	1.36	Interparietal:Nuchal
29	26	30	26	29	30	28	ATSR
25	24	26	27	25	32	24	MBSR
23	22	23	22	21	26	20	PTSR
210	196	197	204	200	187	188	VEN
20+	37	37	38	15+	28	22+	SC
4	3	4	3	3L 4R	3	4	POS

MVZ 293588	MVZ 293587	MVZ 293586	MVZ 293585	MVZ 293584	MVZ 293583	MVZ 293580	Specimen
F	M	M	F	F	F	M	Sex
108	146	104	161	131	158	140	SVL
20.29	24.96	18.82	26.45	20.78	25.86	17.72+	TL
							TL:SVL
4.05	4.74	3.71	4.8	4.75	4.72	4.81	BW
2.7	3.01	2.29	3.24	2.67	3.34	3.7	TW
na	5.36	2.84	na	na	na	5.31	HLL
2.33	3.19	2.25	3.45	2.68	3.12	3.31	Head depth
5.76	8.13	5.72	7.11	6.54	7.54	7.24	HL back of skull
3.88	4.69	3.55	4.85	4.25	4.75	5.13	HL from SL4
3.76	4.24	3.37	4.86	3.58	4.55	5.09	HW
2.07	2.35	1.86	2.55	0.92	2.47	2.5	EN
2.45	2.83	2.23	3.09	2.06	2.98	3	ES
0.89	0.97	0.92	1.11	1.64	1.05	0.96	IN
1.38	1.9	1.36	2.06	1.47	1.85	2.02	IO
0.66	0.93	0.61	0.99	0.77	0.97	0.93	FNSL
0.6	1.99	1.56	2.06	1.6	1.81	2.06	FNSW
0.48	0.63	0.61	0.43	0.43	0.48	0.52	FSL
0.95	0.91	0.74	0.78	0.92	0.93	0.99	FSW
0.49	0.55	0.56	0.82	0.45	0.58	0.82	IPW
0.41	0.57	0.46	0.54	0.48	0.7	0.56	NSW
1.58	0.46	0.47	0.38	0.58	0.51	0.48	Frontal:Frontonasal
1.20	0.96	1.22	1.52	0.94	0.83	1.46	Interparietal:Nuchal
31	28	30	30	30	29	29	ATSR
24	24	26	26	26	24	26	MBSR
23	23	23	24	23	23	23	PTSR
190	202	198	208	203	196	199	VEN
47	43	47	43	43	44	33+	SC
4	4	3l 4r	4	4	4	4	POS

MZB 4274	MZB 4273	MZB 4272	MVZ 293593	MVZ 293591	MVZ 293590	MVZ 293589	Specimen
F	M	M(J)			F	M	Sex
138.59	103.45	62.37	137	120	162	146	SVL
6.8+	14.93	17.49	23.44	18.79	24.5	27.2	TL
	0.14	0.28	0.17	0.16			TL:SVL
3.76	4.15	2.55	3.93	4.01	4.81	5.28	BW
NA	NA	NA	3.06	2.53	3.39	3.12	TW
NA	2.28	0.98	3.76	na		5.33	HLL
			2.45	2.26	3.23	3.19	Head depth
			5.42	5.31	7.71	7.74	HL back of skull
4.2	2.99	2.83	3.67	3.14	4.92	4.63	HL from SL4
3.69	3.1	2.62	3.79	3.23	4.77	4.5	HW
2.04	1.61	1.2	1.82	1.74	2.43	2.29	EN
2.44	2.16	1.41	2.35	2.1	3.1	2.87	ES
1	0.85	0.85	1	0.79	0.98	0.95	IN
2.31	1.78	1.67	1.68	1.33	1.86	1.78	IO
1.76	1.69	1.06	0.8	0.64	0.76	0.86	FNSL
0.72	0.83	0.49	1.53	1.36	2.09	1.74	FNSW
1.94	1.95	1.59	0.6	0.56	0.49	0.61	FSL
0.89	1.05	0.63	0.79	0.66	1.01	0.95	FSW
0.79	0.45	0.65	0.61	0.66	0.66	0.52	IPW
			0.46	0.4	0.58	0.57	NSW
1.24	1.27	1.29	0.52	0.49	0.48	0.55	Frontal:Frontonasal
#DIV/0!	#DIV/0!	#DIV/0!	1.33	1.65	1.14	0.91	Interparietal:Nuchal
24	26	26	27	26	29	28	ATSR
23	23	24	23	22	25	26	MBSR
23	22	21	20	20	24	23	PTSR
214	197	?	201	209	204	201	VEN
17+	40	42	51	46	41	44	SC
			4l 3r	3	4	4	POS

UF 33488	UF 33073	TNHC 062475	TNHC 059689	TNHC 059521	MZB 4279	MZB 4278	MZB 4276	Specimen
F	M	M	F(J)	F	F	M	M	Sex
125.86	126.17	119.83	66.84	157.75	117.25	117.01	145.71	SVL
19.4	21.1	18.15	6.17	18.24	10.1	17.28	8.2+	TL
0.15	0.17	0.15	0.09	0.12	0.09	0.15		TL:SVL
4.2	4.02	4.45	2.28	5.1	3.93	4.21	3.86	BW
3.45	2.78	3.87	2.04	4.53	NA	NA	NA	TW
NA	3.1	3.83	NA	NA	NA	2.75	2.9	HLL
2.78	2.57		1.61	3.85				Head depth
6.4	5.67	6.49	3.87	8.64				HL back of skull
4.27	3.82	4.06	2.48	4.96	3.6	4	3.85	HL from SL4
4.07	3.48	4.05	2.43	5.06	3.83	3.83	3.81	HW
2.02	1.97	2.19	1.27	2.63	1.64	1.54	1.69	EN
2.52	2.44	2.3	1.7	3.28	2.23	1.71	1.7	ES
1.02	0.95	0.93	0.7	1.14	0.92	0.85	1.07	IN
1.77	1.71	2.12	1.47	2.3	2.86	2.3	2.27	IO
1.05	0.95	0.7	0.52	0.77	1.79	1.64	1.83	FNSL
1.88	1.72	1.62	1.23	1.99	0.43	0.75	0.77	FNSW
0.64	0.76	1.07	0.43	1.08	1.42	2.27	2.27	FSL
0.77	0.86	1.93	1.16	1.46	0.78	1	0.83	FSW
0.73	0.66	0.89	0.64	0.89	0.63	0.51	0.88	IPW
0.56	0.66	0.93	0.52	0.48				NSW
0.41	0.50	1.19	0.94	0.73	1.81	1.33	1.08	Frontal:Frontonasal
1.30	1.00	0.96	1.23	1.85	#DIV/0!	#DIV/0!	#DIV/0!	Interparietal:Nuchal
26	28	28	26	34	24	24	25	ATSR
22	24	23	23	30	23	22	24	MBSR
22	20	22	22	29	21	21	21	PTSR
208	206	202	228	244	221	194	192	VEN
43	44	40	35	40	28	41	18+	SC
3	3	2	1	4				POS

Appendix 2: supplementary materials for Chapter 3

Table S1: Most favored models for Principal Coordinate axes

model	k	AICc	AICc Weight	Cumulative weight	Log Likelihood
PCoA 1 ~ strength	3	111.73	0.89	0.89	-52.34
PCoA 1 ~ scale count	3	117.36	0.05	0.94	-55.16
PCoA 1 ~ strength + locality	5	117.48	0.05	0.99	-52.31
PCoA 1 ~ scale count + locality	5	121.22	0.01	1.00	-54.18
PCoA 2 ~ fusions + locality	5	117.17	1.00	1.00	-49.16

Table S2: Most favored models correlating maximum proportional tail length (TL) and snout vent length (SVL) to network summary statistics & principal coordinate values. Here, the locality of Dibamus species is used as a proxy for phylogenetic covariance.

model	k	AICc	AICc Weight	Cumulative weight	Log Likelihood
TL ~ PCoA 2 + locality	5	-106.74	0.28	0.28	59.80
TL ~ composite scale count + locality	5	-106.66	0.27	0.54	59.76
TL ~ fusions + locality	5	-105.84	0.18	0.72	59.35
TL ~ strength + locality	5	-104.82	0.11	0.83	58.84
TL ~ locality	4	-104.31	0.08	0.91	57.06
SVL ~ PCoA 2	3	259.88	0.59	0.59	-126.42
SVL ~ composite scale count	3	262.68	0.15	0.74	-127.82
SVL ~ PCoA 2 + locality	5	264.11	0.09	0.83	-125.62
SVL ~ composite scale count + locality	5	264.45	0.07	0.89	-125.80
SVL ~ fusions	3	264.69	0.05	0.95	-128.83

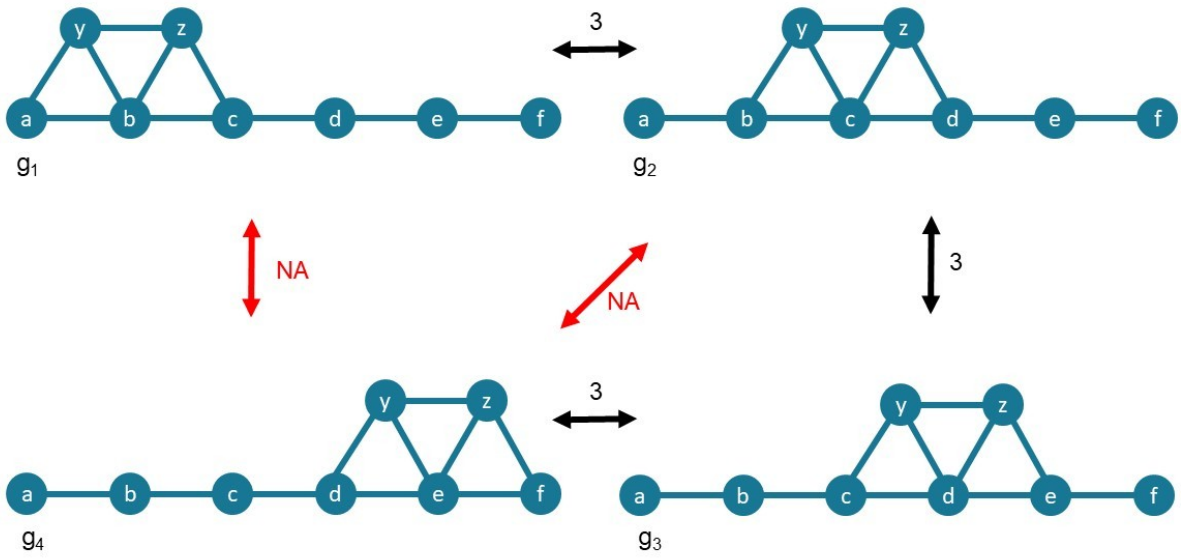

Figure S1: An example of topologies that cannot be compared using the algorithm described in this paper. While The graphs can be compared step-wise through this sequence of transformations, the algorithm described here cannot compare g1 to g3 or g4.

Appendix 3: Supplementary materials for Chapter 4

Diet hierarchy

- Omnivorous
  - carnivorous
    - invertebrates
      - insects
      - spiders
      - ants
      - termites
    - vertebrates
      - fish
      - frogs
      - lizards
      - snakes
      - mammals
      - birds
    - eggs
      - reptile eggs
      - bird eggs
      - amphibian eggs
  - herbivorous
      - fruit
      - nectar
      - honeydew
      - plant material

Table S1. References for phylogeny construction.

Clade	Reference Work
Squamata	(Tonini et al. 2016; Zheng and Wiens 2016)
Opluridae	(Tonini et al. 2016)
Scincidae	(Miralles et al. 2026)
Chamaeleonidae	(Scarpetta et al. 2025)
Pseudoxyrhophiiidae + *Mimophis*	(Burbrink et al. 2019)
Xenotyphlopidae, Typhlopidae	(Tonini et al. 2016; Miralles et al. 2018)
Gekkonidae	(Tonini et al. 2016)
Boidae	(Noonan and Chippindale 2006; Tonini et al. 2016)
Gerrhosauridae	(Blair et al. 2015)
Ebenavia	(Hawlitschek et al. 2018)
Lygodactylus	(Gippner et al. 2021) (divergence times within species groups spaced at even intervals)
Paragehyra	(Crottini et al. 2015)

Table S2. Environmental variables used in this study

#	Layer name	Description
1	2_M_Diu_Tmp_Rng	BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
2	3_Isothermality	BIO3 = Isothermality (BIO2/BIO7) (×100)
3	8_M_Tmp_Wet_Qt	BIO8 = Mean Temperature of Wettest Quarter
4	9_M_Tmp_Dry_Qt	BIO9 = Mean Temperature of Driest Quarter
5	13_Prc_Wet_Month	BIO13 = Precipitation of Wettest Month
6	14_Prc_Dry_Month	BIO14 = Precipitation of Driest Month
7	15_Prc_Seasonal	BIO15 = Precipitation Seasonality (Coefficient of Variation)
8	18_Prc_Warm_Qt	BIO18 = Precipitation of Warmest Quarter
9	19_Prc_Cold_Qt	BIO19 = Precipitation of Coldest Quarter
10	Elev	Elevation (m)
11	PC1_claysilt_v_sand	Physical soil PC 1 - sandy versus clay & silt dominated soils
12	PC2_coarse_v_dense	Physical soil PC 2 - coarse versus dense soils
13	PC3_coarseness	Physical soil PC 3 - soil coarseness
14	PC4_clay_v_deepdense	Physical soil PC 4 - clay-rich versus sandy & dense soils

Table S3. Phylogenetic signal of substrate use and limb development at the species level.

	Yao & Yuan's *M*	p-value
arboreality	0.83	0.001
terrestriality	0.75	0.001
fossoriality	0.93	0.001
saxicoly	0.84	0.001
four limbs	0.97	0.001

Table S4. Single rate-matrix models of character evolution for substrate use and limb development. Models include equal-rates (ER), all-rates-different (ARD), loss-only (loss), and gain-only (gain). Models favored by AIC are in bold.

trait	model	log Likelihood	d.f.	AIC	weight	gain	loss
Terrestrial	ER	-251.61	1	505.23	0	1	1
Terrestrial	ARD	-251.44	2	506.89	0	0.69	0.73
Terrestrial	loss	-237.17	1	476.34	1	0	0.02
Terrestrial	gain	-	-	-	-	-	-
Fossorial	ER	-64.12	1	130.25	0.02	0	0
Fossorial	ARD	-59.42	2	122.84	0.98	0	0.01
Fossorial	loss	-	-	-	-	-	-
Fossorial	gain	-78.29	1	158.57	0	-	-
Arboreal	ER	-144.18	1	290.35	0.7	0.01	0.01
Arboreal	ARD	-144.02	2	292.04	0.3	0.01	0.01
Arboreal	loss	-	-	-	-	-	-
Arboreal	gain	-196.87	1	395.75	0	-	-
Saxicolous	ER	-129.61	1	261.21	0	0.01	0.01
Saxicolous	ARD	-114.98	2	233.97	1	0	0.02
Saxicolous	loss	-	-	-	-	-	-
Saxicolous	gain	-136.75	1	275.5	0	-	-
four-legged	ER	-43.04	1	88.08	0.02	0	0
four-legged	ARD	-38.08	2	80.16	0.98	0	0
four-legged	loss	-46.09	1	94.18	0	0	0
four-legged	gain	-	-	-	-	-	-

Table S5. Phylogenetic signal of substrate use and limb development at the species level. Tests are marked with asterisks to denote levels of significance. * = p <0.05, ** = p <0.001.

trait	Yao & Yuan's *M*	p-value
Diurnality	0.65	0.0315*
Substrate Use	0.70	0.0012**
Diet PC axes 1 - 3	0.69	0.0057**
Climate MDS axes 1 - 3	0.61	0.3312
Body Size Distribution	0.66	0.016*
Range Size	0.61	0.0668
Net Diversification	0.59	0.1363

Table S6. Relationship between general predictors of diversity and biogeographic spread (clade age, body size) and measures of evolutionary success (diversity, net diversification rate, and log geographic range size).significant relationships (P<0.05) are in bold.

correlation	Value	pseudo-R^2	Std.Error	t.value	p.value	AIC
div. ~ clade age	-0.0259	0.0002	0.3066	-0.0844	0.9331	131.1997
div. ~ min. body size	-0.3226	0.1259	0.1267	-2.5454	0.0140	126.7597
div. ~ max. body size	0.3323	0.1373	0.1243	2.6734	0.0101	126.1986
net. div. rate ~ clade age	-0.6122	0.0402	0.4330	-1.4138	0.1636	165.7240
net. div. rate ~ min. body size	**-0.4630**	**0.1191**	**0.1826**	**-2.5358**	**0.0144**	**163.2800**
net. div. rate ~ max. body size	**0.6039**	**0.2080**	**0.1713**	**3.5251**	**0.0009**	**158.2575**
log range size ~ clade age	-0.4654	0.0341	0.3623	-1.2847	0.2048	147.8818
log range size ~ min. body size	0.1287	0.0135	0.1607	0.8009	0.4270	150.5134
log range size ~ max. body size	**0.6912**	**0.3999**	**0.1261**	**5.4798**	**0.0000**	**127.6600**

Table S7. Mantel tests of correlation of niche overlap with phylogenetic distance.

	Z-score	p-value
climate	228034.55	0.533
diet	291911.38	0.001**
substrate	276882.13	0.001**
activity	188434.94	0.177
body size	338017.91	0.113
combined niche	264656.18	0.001**

Table S8. Significant ($p<0.05$) correlation values from partial mantel tests of niche overlap using Schoener's D for 52 genera of malagasy squamates

	geography	climate	diet	substrate	activity	size
geography	1.000	0.672		0.194		0.048
climate	0.672	1.000		0.103		0.022
diet			1.000		0.126	0.165
substrate	0.194	0.103		1.000		
activity			0.126		1.000	
size	0.048	0.022	0.165			1.000