

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

The Roles of Ecological and Evolutionary Processes in Shaping Diversity Patterns Across Different Spatial Scales in a Speciose Group of Snakes

Permalink

<https://escholarship.org/uc/item/8kv5v7m2>

Author

Ramirez Ramirez, Juan P

Publication Date

2023

Supplemental Material

<https://escholarship.org/uc/item/8kv5v7m2#supplemental>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE
AND
SAN DIEGO STATE UNIVERSITY

The Roles of Ecological and Evolutionary Processes in Shaping Diversity Patterns across
Different Spatial Scales in a Speciose Group of Snakes

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Evolutionary Biology

by

Juan Ramirez Ramirez

June 2023

Dissertation Committee:

Dr. Tod Reeder, Co-Chairperson
Dr. Marko Spasojevic, Co-Chairperson
Dr. David Reznick
Dr. Jeet Sukumaran
Dr. Kevin Burns

The Dissertation of Juan Ramirez Ramirez is approved:

Committee Co-Chairperson

Committee Co-Chairperson

University of California, Riverside
San Diego State University

ACKNOWLEDGEMENTS

First of all, I want to thank my advisor, Tod Reeder, for his continued support and encouragement during my stay in the Program. I am grateful for all his generosity and understanding, and it has been a pleasure to be part of his lab and to have him as my advisor. I also want to thank my co-advisor, Marko Spasojevic, for receiving me at his lab at UCR and teaching me the advanced ecological and statistical concepts and methodologies that served as the basis for a considerable portion of this dissertation. My stay in his lab was a great experience, and he was always very attentive and helpful to me. I am grateful to the other members of my dissertation committee too (Kevin Burns, David Reznick, and Jeet Sukumaran) for their insightful feedback on past and present ideas developed during my work on this dissertation. I also acknowledge the Program coordinator, Liz Waters, for helping me to keep track of the goals set by the Program and for her continued confidence in my ability to reach them.

On the other hand, I want to thank the current members of the Reeder Lab (Dillon Jones and Julia Soares-Parreiras) for their friendship and the feedback they provided to previous versions of the several presentations I did of the manuscripts that this dissertation comprises. I am especially thankful to Julia and her husband, G. Azevedo, for inviting me to the multiple Thanksgiving and Christmas gatherings they organized here in San Diego or other beautiful locations where it would be very difficult for me to go by myself. I also thank the members of the Reeder Lab at the time I started my stay in the Program (Mike Burger, Sam Fellows, Alex Sumarli, and Sean Harrington) for their help with some research questions and feedback on presentations, as well as for their support

and friendship. It was a lot of fun to join them on their birding or herping trips, and I am grateful to them for inviting me. In that regard, I am also thankful to other fellow graduate students who kindly invited me to their field trips (Craig Fischer, Casey Richart, and Nick Vinciguerra). Finally, I wanted to thank the many past and present students in the Program who helped me navigate the Program and were always very kind to me, especially to Tierney Bougie, Ryan Buck, Agnese Lanzetti, and B. Myers.

I am also grateful to the members of the Spajosevic lab at UCR I had the chance to meet (Jared Anderson-Huxley, Tesa Madsen-Hepp, and Erin McCann) for their insightful feedback on previous versions of my work on the first chapter of my dissertation and for being very kind to me. I am indebted as well to the staff members at both SDSU and UCR who helped me have a smooth stay in both universities. I also thank other colleagues and friends in the herpetological and evolutionary biology communities for their help and feedback provided to my research.

Finally, I wanted to thank both my close and extended family for their continued personal and economic support during these years; without them, I would not be able to be a part of the Program in the first place. During these six years, I am especially thankful to my mother and brother for taking care of me when I had to undergo surgery. My aunt Pilar and her husband Ricardo also kindly invited and received me in their place and on trips to stunning sites. My father was also always willing to talk with me and give me insightful advice about life in the US and other aspects.

ABSTRACT OF THE DISSERTATION

The Roles of Ecological and Evolutionary Processes in Shaping Diversity Patterns across Different Spatial Scales in a Speciose Group of Snakes

by

Juan Ramirez Ramirez

Doctor of Philosophy, Graduate Program in Evolutionary Biology
University of California, Riverside and San Diego State University, June 2023
Dr. Tod Reeder and Dr. Marko Spasojevic, Co-Chairpersons

In general terms, the overall goal of my dissertation is to acquire a better understanding of the phylogenetic relationships within the most diverse group of snakes in the Neotropics (i.e., the family Dipsadidae) and then use this information to identify the drivers of diversity patterns of the group at multiple spatial scales. To reach this objective, I integrate methodologies, concepts, and hypotheses from phylogenetics, macroecology, and community ecology. For Chapter 1 of my dissertation, I explored the potential causes of phylogenetic uncertainty in one of the two main groups of the Dipsadidae (i.e., the Xenodontinae) using phylogenomic data and a careful examination of the observed gene tree discordance and its potential impact on species tree inference. Although I did not recover a strong resolution of several recalcitrant nodes, I identified the potential factors driving gene tree discordance in the group, apart from finding strong and consistent support for other dipsadid lineages. For Chapter 2, I evaluated the potential drivers of species richness of the Dipsadidae across the entire American continent. In this regard, I recovered that regional diversification had the most important

role in determining species richness variation across the distribution of the group. Finally, for Chapter 3 of my dissertation, I investigated the influence of a large-scale biogeographic event (the Great American Biotic Interchange, GABI) on the current species and functional diversity of a series of communities of the Dipsadidae across the Neotropics. Despite the increased input of taxa following the GABI, I did not recover that the communities involved in this event ended up being more diverse at this current time. Such a result indicates that in these communities, an equilibrium point has been reached between the processes of colonization and biotic interactions. In general, my studies continue with and expand upon the increasing recognition of the importance of phylogenetics in different areas of ecology and vice versa. Further, it is based on methodologies that can easily be implemented in other lineages and thus contribute to identifying more generalizable patterns about the drivers of species richness at the local and regional scales.

TABLE OF CONTENTS

Introduction of the Dissertation.....	1
 Chapter 1:	
Exploring the Drivers of Phylogenomic Uncertainty in the Hyper-diverse Neotropical Snake Subfamily Xenodontinae (Dipsadidae)	
Abstract.....	7
Introduction.....	9
Materials and Methods.....	15
Results.....	24
Discussion.....	30
Conclusions.....	39
References.....	41
Tables.....	51
Figures.....	57
 Chapter 2:	
Regional speciation drives spatial variation in species richness in the most diverse family of Neotropical snakes (Colubroidea: Dipsadidae)	
Abstract.....	64
Introduction.....	66
Materials and Methods.....	72
Results.....	81
Discussion.....	83
References.....	96
Tables.....	107

Figures.....	108
Chapter 3:	
Extinction debt and functional traits mediate community saturation over large spatiotemporal scales	
Abstract.....	112
Introduction.....	114
Materials and Methods.....	118
Results.....	123
Discussion.....	124
References.....	134
Tables.....	142
Figures.....	144
Conclusions of the Dissertation.....	146
Appendices.....	150
Chapter 1	150
Chapter 2	178
Chapter 3	212
Captions of Supplemental Files.....	275

LIST OF TABLES

Chapter 1

Table 1.1. Definition of the trait metrics indicative of Gene Tree Estimation Error calculated for each locus.....	51
--	----

Table 1.2. P-values of the multinomial regressions performed between the trait values of each locus against whether they support or not the inferred relationships of interest.....	54
---	----

Table 1.3. Values of the Δ metric of deep introgression.....	56
---	----

Appendix Table A1.1. Calibration points used for calculating phylogenetic informativeness values and the time when they reach their maximum.....	150
--	-----

Appendix Table A1.2. Voucher information of the sequence data available for the species included in this study.....	151
---	-----

Appendix Table A1.3. Summary of the Principal Component (PC) Axes of the PC analysis performed with the traits of the loci included herein.....	154
---	-----

Appendix Table A1.4. First five Loadings of Principal Component (PC) Axes of the PC analysis performed with the traits of the loci included herein.....	154
---	-----

Chapter 2

Table 2.1. Main explanations of associations between regional climate and species richness patterns.....	107
--	-----

Appendix Table A2.1. Best-fit partitioning scheme and substitution models for the protein-coding genes used in this study.....	184
--	-----

Appendix Table A2.2. Calibration points used for obtaining the timetree used for this work.....	185
---	-----

Appendix Table A2.3. Dutilleul's modified t-test between the explanatory variables of the theoretical structural equation model tested for this study.....	186
--	-----

Appendix Table A2.4. Model selection of the spatial weight matrices associated with the equations of the structural equation models tested.....	188
---	-----

Appendix Table A2.5. Results of the d-separation tests we performed with the theoretical structural equation model evaluated herein.....	190
--	-----

Appendix Table A2.6. Standardized estimates of the best-fitted structural equation model of the factors driving species richness patterns in the Dipsadidae.....	191
--	-----

Chapter 3

Table 3.1. Results of the generalized linear model performed between the species richness of 67 communities of the Dipsadidae and how many species colonized them across the Panamanian Isthmus.....	142
--	-----

Table 3.2. Results of the Beta regression performed between the functional dispersion of 67 communities of the Dipsadidae and how many species colonized them across the Panamanian Isthmus.....	143
--	-----

Appendix Table A3.1. Information about the 67 communities of the Dipsadidae included in this study.....	230
---	-----

Appendix Table A3.2. Loadings of the first four Principal Component (PC) Axes of the PC analyses performed with the climatic variables of the sites included herein...	241
--	-----

Appendix Table A3.3. Summary of the Principal Component (PC) Axes of the PC analyses performed with the climatic variables of the sites included herein.....	242
--	-----

Appendix Table A3.4. Species richness and functional dispersion of the sympatric snakes present in the 67 communities of the Dipsadidae included in this study.....	243
---	-----

Appendix Table A3.5. Comparison of biogeographic models performed with ‘BioGeoBEARS’	249
--	-----

Appendix Table A3.6. List of dispersal events of the Dipsadidae across the Panamanian Isthmus.....	250
--	-----

Appendix Table A3.7. Best-fitting partitioning scheme and substitution models for the protein-coding genes used in this study.....	253
--	-----

Appendix Table A3.8. Calibration points used for obtaining the timetree used for this work.....	254
---	-----

Appendix Table A3.9. Values of per-species diversification, speciation, and extinction rates averaged across taxa from the communities of the Dipsadidae considered herein.....	255
---	-----

Appendix Table A3.10. Values of the functional dispersion of the communities of dipsadid snakes included herein recalculated after removing the species colonizing them from across the Panamanian Isthmus.....	258
---	-----

Appendix Table A3.11. Results of t-tests testing for differences between communities of dipsadid snakes involved or not in the Great American Biotic Interchange.....	261
---	-----

LIST OF FIGURES

Chapter 1

Figure 1.1. Summary of the phylogenetic relationships within the Dipsadidae recovered in past studies.....	57
Figure 1.2. Coalescent-based species tree of the Dipsadidae.....	58
Figure 1.3. Per-node levels of incongruence between the coalescent-based species tree of the Dipsadidae and the inferred gene trees.....	59
Figure 1.4. Results of the Likelihood Mapping Analyses we performed for each of the nodes of interest using the concatenation of the genes employed in this study.....	60
Figure 1.5. Results of the Principal Component (PC) analysis performed with the traits of the loci included herein.....	61
Figure 1.6. Distributions of the Robinson-Foulds distances between the coalescent-based species tree we inferred and the observed or simulated gene trees.....	62
Figure 1.7. Comparison between the observed and expected values of the Δ metric of deep introgression.....	62
Appendix Figure A1.1. Coalescent-based species trees of the Dipsadidae obtained using each of the locus types employed in this study by separate.....	156
Appendix Figure A1.2. Concatenation-based species trees of the Dipsadidae obtained using each of the locus types employed in this study by separate.....	158
Appendix Figure A1.3. Coalescent-based species trees of the Dipsadidae obtained using different thresholds of the minimum taxonomic completeness considered.....	161
Appendix Figure A1.4. Coalescent-based species trees of the Dipsadidae obtained using different thresholds of the minimum alignment length considered.....	162
Appendix Figure A1.5. Coalescent-based species trees of the Dipsadidae obtained using different thresholds of the minimum percentage of missing data considered.....	164
Appendix Figure A1.6. Coalescent-based species trees of the Dipsadidae obtained using different thresholds of the minimum Ultra-Fast bootstrap support value where gene tree branches were collapsed.....	168
Appendix Figure A1.7. Coalescent-based species trees of the Dipsadidae where the values of gene and site concordance factors were mapped.....	170

Appendix Figure A1.8. Per-node levels of incongruence between the coalescent-based species tree of the Dipsadidae and the inferred gene trees obtained with different thresholds of the minimum Ultra-Fast bootstrap support value where the branches of the gene trees were collapsed.....172

Appendix Figure A1.9. Comparison between the values of the traits of each locus against whether they support or not the sister relationships between the tribe Pseudoboini and the rest of the Xenodontinae.....174

Appendix Figure A1.10. Comparison between the values of the traits of each locus against whether they support or not the sister relationships between the Xenodontini and the Psomophini tribes.....175

Appendix Figure A1.11. Comparison between the values of the traits of each locus against whether they support or not the sister relationships between the Philodryadini and the Echinantherini+Elapomorphini tribes.....176

Appendix Figure A1.12. Comparison between the values of the traits of each locus against whether they support or not the sister relationships between the Echinantherini and the Elapomorphini tribes.....177

Chapter 2

Figure 2.1. Patterns of species richness of the Dipsadidae across the New World, as estimated with Roll et al. (2017) distribution maps.....108

Figure 2.2. Patterns of geographic variation of the per-cell "Diversification" Rate (DR) and Mean Phylogenetic Diversity (MPD) metrics across the distribution of the Dipsadidae in the New World using the timetree of the group we had inferred after incorporating taxa without genetic data.....109

Figure 2.3. Theoretical structural equation model tested in this work to understand the drivers of species richness in the Dipsadidae.....110

Figure 2.4. Resulting piecewise structural equation model obtained in this work regarding the factors influencing regional species richness variation in the Dipsadidae.....111

Appendix Figure A2.1. Timetree of the family Dipsadidae derived from a relaxed-clock MCMC Bayesian analysis using sequences available up to 2021.....193

Appendix Figure A2.2. Patterns of species richness of the Dipsadidae across the New World, as estimated with Acevedo et al. (2017) distribution maps.....198

Appendix Figure A2.3. Patterns of geographic variation of the per-cell "Diversification" Rate (DR) and Mean Phylogenetic Diversity (MPD) metrics across the distribution of the Dipsadidae in the New World using the timetree of the group we had inferred.....	199
Appendix Figure A2.4. Geographic variation of the per-cell net diversification, speciation, and extinction rates across the distribution of the Dipsadidae in the New World, as obtained with the software BAMM and the timetree of the group we had inferred after incorporating taxa without genetic data.....	200
Appendix Figure A2.5. Moran's I correlograms indicating spatial autocorrelation in the residuals of the equations of the piecewise structural equation model we evaluated herein.....	202
Appendix Figure A2.6. Relationships between values of the DR metric and the per-cell net diversification and speciation rates obtained with the software BAMM and the timetree of the group we had inferred after incorporating taxa without genetic data.....	204
Appendix Figure A2.7. Relationships between values of the DR metric and the per-cell net diversification and speciation rates obtained with the software BAMM and the timetree of the group we had inferred.....	206
Appendix Figure A2.8. Timetree of the Dipsadidae incorporating taxa without genetic data where the net diversification rates estimated with the software BAMM were mapped.....	208
Appendix Figure A2.9. Timetree of the Dipsadidae where the net diversification rates estimated with the software BAMM were mapped.....	209
Appendix Figure A2.10. Patterns of species richness of the subfamilies of the Dipsadidae across the New World, as estimated with Roll et al. (2017) distribution maps.....	210
Chapter 3	
Figure 3.1. Location of the 67 communities of the Dipsadidae analyzed in this study, along with the values of species richness and functional dispersion they exhibit.....	144
Figure 3.2. Principal Component Analysis (PCA) of the climatic variables of the communities of the Dipsadidae studied herein.....	145

Appendix Figure A3.1. Relationship between the species richness or functional diversity of communities and whether they were colonized or not across the Panamanian Isthmus.....	269
Appendix Figure A3.2. Timetree of the family Dipsadidae derived from a relaxed-clock MCMC Bayesian analysis using sequences available up to 2019.....	270
Appendix Figure A3.3. Map of the biogeographic regions considered herein....	274

INTRODUCTION OF THE DISSERTATION

Despite developing separately during most of the twentieth century, the last two decades have seen an increased interest in the integration of evolutionary processes in ecology and vice versa (Ricklefs, 1987; Cadotte & Davies, 2016; Swenson, 2019). Such integration has happened on two main fronts. In the first place, the field of *eco-evolutionary dynamics* recognizes that evolutionary changes can occur even in short timescales where only ecological processes were expected to act (Hendry, 2017). On the other hand, *eco-phylogenetics* has emerged from the increased recognition that phylogenetics can provide important insights about processes occurring at larger temporal spans and across multiple spatial scales (Cadotte & Davies, 2016). Indeed, phylogenetic data is fundamental for answering classical ecological questions such as which taxa and regions are of conservation interest, the evolution of ecologically relevant traits, and the disciplines of macroecology and community ecology with their goals to understand the drivers of biodiversity patterns at large or small spatial scales (Tucker et al., 2016; Swenson, 2019). That is the case, as phylogenies reflect the accumulation of phenotypic, genetic, or behavioral differences between lineages that can describe or explain ecological patterns and the processes that generate them (Tucker et al., 2016). Thus, it is fundamental for eco-phylogenetic studies to have accurate inferences of the evolutionary relationships between the taxa of interest. In this regard, improvements in computational capabilities, the continuous development of new software, and the increased accessibility of genomic-scale data facilitate obtaining robust inferences of phylogenetic relationships between lineages (Goldman & Yang, 2008; Lemmon and Lemmon, 2012). That is

especially the case of modern high-throughput DNA sequencing techniques, which allow the collection of sequence data for hundreds or thousands of loci for non-model or model organisms alike in a more cost-effective way than the traditionally used Sanger sequencing technique (Ekblom & Galindo, 2011, Zhang et al., 2011). These methodological and technological advances have occurred along with the increase in the availability of fine-scaled datasets at the global scale on spatial, environmental, and vegetation coverage variables, as well as detailed data on the geographic distribution and functional traits of taxa (Hawkins & Diniz-Filho, 2004; Hortal et al., 2015). In combination, all of these developments have resulted in unprecedented growth in the area of eco-phylogenetics and the questions it can address (Cadotte & Davies, 2016). Despite this, many studies in eco-phylogenetics only use phylogenetic information as a proxy of ecological similarity, which does not recognize that only a few traits can determine whether taxa coexist instead of their overall similarity (Swenson, 2019). Thus, such studies fail to really incorporate evolutionary history to answer ecological questions (Swenson, 2019). Instead, eco-phylogenetic studies need to infer, instead of assume, the biogeographic events, diversification rates, or trait or niche shifts that could have influenced current-day regions or communities (Gerhold et al., 2015; Swenson, 2019).

Snakes of the primarily Neotropical family Dipsadidae represent a system of considerable interest for performing studies in eco-phylogenetics and, more specifically, for using phylogenetic data to test hypotheses in areas such as macroecology and community ecology. That is the case because of the high levels of species richness and extensive range of distribution exhibited by this clade. Indeed, the Dipsadidae is a

hyper-diverse group consisting of > 800 species and 112 genera (Uetz et al., 2023) inhabiting many different environments from Canada to Argentina, as well as four species found in Eastern Asia. Further, the high diversity of dipsadids can be observed not only in terms of the total number of species it exhibits but also regarding the considerable variety of life history strategies exhibited within this diverse clade of snakes. Unfortunately, despite the phylogenetic relationships between many of the dipsadid lineages being reconstructed consistently between studies, other relationships remain recalcitrant to being resolved even after using modern genomic data. Further, studies inspecting the diversity patterns of the Dipsadidae and their potential ecological and evolutionary drivers at either large or small spatial scales are very limited in number. The few studies in the community ecology or macroecology of the group are either based on all neotropical groups of snakes as a whole (e.g., Guedes et al., 2018; Azevedo et al., 2019; Vasquez-Restrepo et al., 2022), are restricted to specific geographical regions (e.g., Bellini et al., 2015; Cavalheri et al., 2015), or do not even incorporate phylogenetic information (e.g., Franca et al., 2007; Hartmann et al., 2009; but see Cadle & Greene, 1993 who use an incomplete and outdated phylogeny of the group). Thus, further studies addressing these questions with updated methodologies and datasets are badly needed. Doing so is not only necessary to increase our knowledge about the ecology and macroevolution of this group, but it will most likely have broader implications for other groups of organisms and parts of the world as well.

Specifically, this dissertation is set to evaluate the influence that different ecological (such as biotic interactions) and evolutionary processes (such as long-term

dispersal, speciation, and extinction) have had on determining the species richness patterns in the Dipsadidae across the entire neotropics. However, as undertaking this task requires having adequate knowledge of the phylogenetic relationships within the group, I also carefully inferred and inspected these relationships using the most up-to-date information available of the group until late 2022. Indeed, the first goal of my dissertation is to investigate the causes of our current inability to resolve some of the most important evolutionary relationships within the Dipsadidae, even when using modern genomic data. I expected that doing so would allow me to find a satisfactory and well-supported resolution to the recalcitrant relationships within one of its more speciose subclades. On the other hand, for the second goal of my thesis, I used phylogenetic data of the entire family, apart from updated distribution maps of the group, to understand the causes of species variation at the regional (i.e., on units of 10^3 square kilometers) spatial scale. Finally, the last objective of my dissertation was to identify signatures of the influence of a large-scale biogeographical event in the current species richness and trait diversity of a series of communities (i.e., at a spatial scale of less than 10^2 square kilometers) of dipsadid snakes across the Neotropics.

References

- Azevedo, J. A., Guedes, T. B., Nogueira, C. D. C., Passos, P., Sawaya, R. J., Prudente, A. L., ... & Antonelli, A. (2020). Museums and cradles of diversity are geographically coincident for narrowly distributed Neotropical snakes. *Ecography*, 43(2), 328-339.
- Bellini, G. P., Giraudo, A. R., Arzamendia, V., & Etchepare, E. G. (2015). Temperate snake community in South America: is diet determined by phylogeny or ecology?. *PLoS One*, 10(5), e0123237.
- Cadle, J.E. & Greene, H.W. (1993). Phylogenetic patterns, biogeography, and the ecological structure of Neotropical snake assemblages, pp. 281–293. In: R.E. Ricklefs and D. Schluter (eds.), *Species Diversity in Ecological Communities*. The University of Chicago Press, Chicago, Illinois, USA.
- Cadotte, M. W. & Davies, T. J. (2016). *Phylogenies in ecology: a guide to concepts and methods*. Princeton University Press, Princeton, New Jersey, USA.
- Cavalleri, H., Both, C., & Martins, M. (2015). The interplay between environmental filtering and spatial processes in structuring communities: The case of neotropical snake communities. *PloS One*, 10(6), e0127959.
- Ekblom, R., & Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, 107(1), 1–15.
- França, F. G. R., & Araújo, A. F. B. D. (2007). Are there co-occurrence patterns that structure snake communities in Central Brazil?. *Brazilian Journal of Biology*, 67, 33-40.
- Gerhold, P., Cahill Jr, J. F., Winter, M., Bartish, I. V., & Prinzing, A. (2015). Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Functional Ecology*, 29(5), 600-614.
- Goldman, N., & Yang, Z. (2008). Introduction. Statistical and computational challenges in molecular phylogenetics and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1512), 3889-3892.
- Guedes, T. B., Sawaya, R. J., Zizka, A., Laffan, S., Faurby, S., Pyron, R. A., ... & Antonelli, A. (2018). Patterns, biases and prospects in the distribution and diversity of Neotropical snakes. *Global Ecology and Biogeography*, 27(1), 14-21.
- Hawkins, B. A., & Diniz-Filho, J. A. F. (2004). 'Latitude' and geographic patterns in species richness. *Ecography*, 268-272.

- Hendry, A.P., 2017. *Eco-evolutionary Dynamics*. Princeton University Press, Princeton, New Jersey, USA.
- Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., & Ladle, R. J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 46, 523-549.
- Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61(5), 727-744.
- Ricklefs, R. E. (1987). Community diversity: relative roles of local and regional processes. *Science*, 235(4785), 167-171.
- Swenson, N.G. (2019). *Phylogenetic Ecology: A History, Critique, and Remodeling*. University of Chicago Press, Chicago, Illinois, USA.
- Tucker, C. M., Cadotte, M. W., Carvalho, S. B., Davies, T. J., Ferrier, S., Fritz, S. A., ... & Mazel, F. (2017). A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Reviews*, 92(2), 698-715.
- Uetz, P., Freed, P. & Hošek, J. (Eds.) (2023). The Reptile Database, <http://www.reptile-database.org>, accessed 23 January 2023.
- Vásquez-Restrepo, J. D., Ochoa-Ochoa, L. M., Flores-Villela, O., & Velasco, J. A. (2023). Deconstructing the dimensions of alpha diversity in squamate reptiles (Reptilia: Squamata) across the Americas. *Global Ecology and Biogeography*, 32(2), 250-266.
- Zhang, J., Chiodini, R., Badr, A., & Zhang, G. (2011). The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics*, 38(3), 95–109.

CHAPTER 1

Exploring the Drivers of Phylogenomic Uncertainty in the Hyper-diverse Neotropical Snake Subfamily Xenodontinae (Dipsadidae)

Abstract

Even with the application of genomic-scale data and up-to-date methods, certain phylogenetic relationships are still unsolved in many groups of organisms. Such uncertainty results from both the influence of methodological limitations (e.g., gene tree estimation error [GTEE] or unrecognized paralogy) and/or actual biological processes (e.g., Incomplete Lineage Sorting [ILS] or deep introgression). Respectively, these factors generate either artifactual or real gene tree discordance (between themselves and/or with regard to the inferred species tree), which are difficult to distinguish. Indeed, simultaneously accounting for multiple factors driving gene tree discordance is very computationally intensive, and it remains more feasible to inspect their influence separately. In this study, we investigated the origins of the uncertainty in the high-level phylogenetic relationships of a hyperdiverse snake subfamily (Xenodontinae of the Dipsadidae). To do so, we reanalyzed a recently published genomic dataset from the group containing a total of 4896 loci and 18 species across the subfamily's diversity. We evaluated the potential drivers of gene tree discordance by comparing the values of traits associated with GTEE between concordant and discordant gene trees, apart from performing coalescent simulations and tests for introgression. We recovered that the phylogenetic relationships within the Xenodontinae are highly supported and concordant

with previous studies, except for two consecutive deep branches, which were poorly supported and generated extensive gene tree discordance. We also found that this discordance is most likely caused by high levels of ILS, with no evidence of either systematic GTEE or hybridization. Our results demonstrate that increasing the amount of genomic data does not necessarily result in the resolution of recalcitrant branches. However, carefully inspecting the influence of the drivers of gene tree discordance allows for proposing more informed hypotheses about how the taxa of interest are related, which can later be tested with increased taxonomic sampling and more informative genetic markers.

Introduction

It is increasingly recognized that the phylogenetic relationships of some parts of the tree of life are recalcitrant to being solved, even after using hundreds or thousands of loci that modern genomic techniques allow us to obtain and analyze (Karin et al., 2019; Kulkarni et al., 2021; Singhal et al., 2021). In general, phylogenetic uncertainty in species tree reconstruction appears to be most frequently observed when there are high levels of discordance amongst topologies inferred with individual loci (=gene trees) (Lopes et al., 2021). Although many factors could be responsible for such increased levels of gene tree discordance, most correspond to methodological artifacts or limitations that result in the inference of incorrect gene trees in what is known as Gene Tree Estimation Error (GTEE). Examples of factors causing GTEE (in a broad sense) include errors in the alignment of DNA sequences, failure to recognize the paralogy of loci, having genetic data with reduced phylogenetic information, or using inadequate substitution models (Lopes et al., 2021; Singhal et al., 2021, Yan et al., 2022). However, discordance between gene trees (and when compared with the inferred species tree) can also result from biological processes, such as Incomplete Lineage Sorting (ILS), deep introgression, gene duplication, and horizontal gene transfer (Maddison, 1997; Yan et al., 2022).

Multiple approaches have been used to identify and account for the influence of GTEE, including gene filtering and Gene Genealogy Inspection (GGI). In the first former, filtering loci before species tree inference is commonly done to reduce GTEE (Alda et al., 2019). Such a process consists of removing outlier loci or those whose

corresponding alignments or gene trees have characteristics frequently considered indicative of GTEE (Betancur et al., 2019). Examples of these traits include alignment size, the amount of missing data, or the number of parsimony informative sites (Alda et al., 2019). However, a limitation of this approach is that considering that different genes potentially inform the reconstruction of different parts of a species tree; thus, the removal of some loci could negatively impact solving other relationships apart from those of the main interest (Singhal et al., 2021). In this regard, a "targeted" filtering approach was recently proposed, which uses metrics (such as phylogenetic informativeness) with the ability to determine the depth of the tree in which each locus is more informative (Alda et al., 2019). On the other hand, other studies (e.g., Ballesteros & Sharma, 2019; Singhal et al., 2021) compare the traits of loci (e.g., missing data percentage, composition heterogeneity, average number of species per alignment, or GC content) whose corresponding gene trees possibly conflict with the relationships of interest inferred in the species tree. Finding that low-quality loci produce gene trees that tend to be concordant with the reconstruction from the species tree would support the idea that this topology is potentially an artifact resulting from GTEE. Finally, the GGI approach of Arcila et al. (2017) evaluates whether the alignments used to infer gene trees are informative enough to distinguish between a series of alternate topologies (in most cases, all possible reconstructions of the node of interest, Arcila et al., 2017; Betancur et al., 2019). Such a possibility is tested by inferring gene trees constrained to these topologies and then comparing their likelihood values to each other with the Approximately-Unbiased (AU) test (Shimodaira, 2002). Loci in which only a few topologies have a significantly worse

likelihood that the rest have a reduced ability to distinguish between alternate phylogenetic hypotheses. Thus, GGI permits identifying and quantifying the combined influence of factors, such as limited informativeness or model misidentification on gene tree discordance (Alda et al., 2019). However, when combined with a filtering approach, this approach can also be used to reduce the influence of GTEE. Indeed, the loci for which all topologies are significantly worse than the one exhibiting the highest likelihood value can be filtered and used to infer a new species tree which is then compared with the original species tree (e.g., Arcila et al., 2017; Alda et al., 2019; Betancur et al., 2019). The constrained topologies for all loci of the highest likelihood values are also frequently employed to infer a species tree (Arcila et al., 2017; Mirarab, 2017).

The influence of ILS on gene tree discordance has been the focus of much attention, with several phylogenetic inference methods being proposed to account for it (Liu et al., 2019; Morales-Briones et al., 2021). Some of these approaches simultaneously infer both the species tree and the gene trees (known as "one-step" coalescent methods), such as BEST or StarBeast (Liu et al., 2019; Douglas et al., 2022). However, despite recent advances (e.g., StarBeast3 can now be used for datasets of hundreds of genes and dozens of species; Douglas et al., 2022), these methods remain computationally demanding. However, other approaches accounting for ILS exist ("summary" coalescent methods such as ASTRAL, NJ-EST, or STEAC; Zhang et al., 2018a; Liu et al., 2019), which estimate species trees using gene trees previously inferred using other software (Zhang et al., 2018a). For these "summary" coalescent methods, statistical consistency is only guaranteed for error-free gene trees, so their accuracy can be negatively affected by

GTEE (Zhang et al., 2018a; Cai et al., 2021). Further, both the one-step and summary coalescent methods assume that no gene flow between tree branches exists (Leaché et al., 2014; Solís-Lemus et al., 2016; Long & Kubatko, 2018; Liu et al., 2019).

Because of the potential impact of past gene flow between deeper branches, several methods that identify signals of introgression have been proposed (see Hibbins & Hahn, 2022, for a thorough review). These methods test for deviations from the expectations of the MSC on aspects such as allelic site patterns (such as the D-statistic or the HyDe metric of Blischak et al., 2018; Hibbins & Hahn, 2022), the number of gene trees supporting or conflicting with the inferred species tree (e.g., the Δ metric of Huson et al., 2005), or the coalescent times of non-sister pairs of taxa (as in the approach implemented in the software QuIBL of Edelman et al. 2019). Methods that can account for both ILS and introgression have also been designed (e.g., those implemented in the software SNaQ and PhyloNet; Solís-Lemus & Ane, 2016, Wen et al., 2018).

Unfortunately, these approaches are also computationally intensive, limiting their use to datasets of a reduced number of species and/or loci (Hibbins & Hahn, 2022). Thus, in many cases, studies on gene tree discordance need to separately test the influence of GTEE, ILS, and introgression when attempting to resolve conflicting phylogenetic relationships (Morales-Briones et al., 2021). In this regard, using subsets of the total number of loci and/or species of phylogenetic datasets allows for the analysis of such data with the more computationally demanding software. However, as mentioned above, removing suboptimal gene trees has the potential of eliminating loci that could provide important information about the resolution of certain parts of the tree.

Rapid evolutionary radiations favor the existence of gene tree discordance and likely, species tree uncertainty as a result. Such is the case, as the fast origination of closely related species makes it more likely for incomplete lineage sorting and introgression to occur (Cai et al., 2021). Further, rapid splitting events limit the accumulation of DNA substitutions along a branch (i.e., within an evolutionary species), resulting in the low phylogenetic signal of many loci and thus making gene tree estimation error more likely to occur (Giarlia & Esselstyn, 2015; Cai et al., 2021). Taking this into account, with its rapid diversification and considerable uncertainty between species trees, the snake family Dipsadidae represents an ideal system for exploring the roles of the different processes causing gene tree discordance. Indeed, with >800 species and 97 genera and inhabiting many different environments from Canada to Argentina, the Dipsadidae is the most diverse group of snakes in the New World (Uetz et al., 2023). Further, most of the dipsadid suprageneric clades in the Neotropics originated ~20 million years ago (Ramirez et al., submitted). Apart from a small group of North American taxa and two small Tibetan genera, the Dipsadidae comprises two main subfamilies of comparable species richness: the Dipsadinae and the Xenodontinae (Uetz et al., 2023). Both subfamilies have a complicated taxonomic history, and a consensus on their intrasubfamilial phylogenetic relationships is far from being reached. Indeed, previous investigations based on loci obtained with traditional Sanger sequencing technique have shown widespread paraphyly at the generic level within the Dipsadinae (e.g., Sheehy, 2012; Pyron et al., 2015, Ramirez et al., submitted; Ramirez & Reeder, in preparation). On the other hand, although most genera of the Xenodontinae appear to be

monophyletic, past studies have recovered well-supported but conflicting and, in some cases, poorly supported reconstructions of the relationships between most xenodontine subclades (e.g., Zaher et al., 2019; Pyron et al., 2015; Ramirez et al., submitted; Ramirez & Reeder, in preparation; see Figure 1.1 for a summary of the relationships within the Xenodontinae obtained among previous studies). Importantly, such a result is not only obtained through the analysis of traditional sequence data but is also recovered in phylogenies recently obtained using different types of phylogenomic data (Singhal et al., 2017).

In this study, we explored how the previous phylogenetic uncertainty in the high-level relationships within the Xenodontinae could result from the influence of processes causing discordance between gene trees and an inferred species tree. To do so, we first reconstructed a coalescent species tree of the Xenodontinae using the phylogenomic datasets compiled in two recent studies (Singhal et al., 2017, 2021). We then tested the informativeness of the alignments used to reconstruct the gene trees needed for such analyses by implementing a GGI approach. Afterward, we explored if gene tree discordance could result from GTEE by filtering loci by the size of their alignments, the number of species they contain, and the overall percentage of missing data. We also tested if gene trees supporting and conflicting with the inferred relationships of the species tree in the resolution of recalcitrant nodes significantly differ in the values of a total of 13 alignment and gene tree traits (listed in Table 1.1). We also performed coalescent simulations to evaluate if the observed levels of discordance can be expected to result from ILS. Finally, we investigated the potential role of introgression by

determining if the number of gene trees congruent or conflicting with the species tree differs from that expected if ILS were the sole cause of gene tree discordance through the use of the Δ metric of Huson et al., (2005).

Materials and Methods

Phylogenomic data compilation and assembly

To explore the high-level phylogenetic relationships between the members of the Dipsadidae (with a focus on the Xenodontinae), we compiled genome-wide sequence data from two recent publications (Singhal et al., 2017, 2021). These sequences were obtained with a target capture approach that uses a set of probes known as the Squamate Conserved Loci Set (SqCL; Singhal et al., 2017). The SqCL allows the recovery of sequence data from the loci used by the main approaches in vertebrate phylogenomics: Anchored-hybrid enrichment (AHE, Lemmon et al., 2012) and Ultraconserved elements (UCEs, Smith et al., 2014). Also, this set of probes includes 47 nuclear genes traditionally obtained with the Sanger sequencing technique and used in previous phylogenetic studies of the Squamata (e.g., Wiens et al., 2012; Pyron et al., 2013). We did not generate any new sequence data for this study. In total, the compiled sequence data correspond to 27 snake species: seven being outgroups while the remaining 20 are members of the Dipsadidae, with three samples from *Pseudoboa nigra*. Among the latter, four species are part of the Dipsadinae, and 18 represent the Xenodontinae. SqCL sequence data were available for six of the 16 tribes (i.e., 37.5%) currently recognized for the Xenodontinae (13 tribes according to Zaher et al., 2009, plus the addition of the

Amnesteophiini, the Eutrachelophiini and the Incaspidini by Myers, 2011; Myers & McDowell, 2014; and Arredondo et al., 2020, respectively). However, it needs to be recognized that most of the xenodontine tribes for which no SqCL data are available contain less than four species (except for the Alsophini, Hydropsini, Tachymenini, and the non-monophyletic Saphenophiini, which have between 12 and 47 species, Uetz et al., 2023). In contrast, the xenodontine tribes included in this study contain between 15 and 79 species, with only the Psomophini having three species (Uetz et al., 2023). Also, despite Singhal et al. (2017, 2021) report having SqCL data from one species of the Tachymenini (i.e., *Thamnodynastes hypoconia*), we found that this species sample clustered within those from the Philodryadini in all analyses performed herein (and those from Singhal et al., 2017), rendering this tribe to be non-monophyletic. Further, the identification of this sample was corrected as *Philodryas nattereri* in Singhal et al. (2022), and we follow this change in all analyses below. Similarly, the individual identified as *Pseudoboa newwiedii* by Singhal et al. (2017) appears to correspond instead to a *P. nigra* (Singhal et al., 2022). We then assembled, trimmed, and aligned the compiled phylogenomic data using the pipeline PHYLUCE-1.7.1 (Faircloth, 2016). We selected the option that performs sequence alignment with the software MAFFT 7.475 (Katoh and Standley, 2013), while internal trimming was performed with the software Gblocks 0.91b (Castresana, 2000).

Gene and species tree inference

We inferred individual Maximum Likelihood (ML) trees for every locus (=gene trees) with the software IQ-TREE 2.2.0 (Minh et al., 2020) and identified the best-fitting substitution model using Model Finder Plus (MFP, Kalyaanamoorthy et al., 2017), also implemented in IQ-TREE. We also performed 1000 ultrafast bootstrap replicates (UFBoot2, Minh et al., 2013; Hoang et al., 2018) to determine branch support. The rest of parameters were left in their default values. It is important to recognize that the ultrafast bootstrap values are interpreted differently than the traditional bootstrap (Felsenstein, 1985), as only values >95 are indicative of a clade being well-supported (Hoang et al., 2018). The UFBoot method is much faster and less computationally intensive than the traditional bootstrap and does not underestimate the probability that a clade is correct (i.e., it provides relatively unbiased support values, Minh et al., 2013; Hoang et al., 2018). We selected all gene trees with a 75% value of taxonomic completeness (i.e., minimum percentages of missing taxa per locus) and used them to infer a species tree of the Dipsadidae using Multi-threaded ASTRAL 5.15.4 (ASTRAL-MP; Zhang et al., 2018a; Yin et al., 2019), with branch support estimated as local posterior probabilities (Sayyari & Mirarab, 2016). Local posterior probabilities of 0.95 or greater were considered strongly supported (Sayyari & Mirarab, 2016). Using the dataset of 75% taxonomic completeness, we also estimated a concatenated species tree of the Dipsadidae using IQ-TREE 2.2.0. For this analysis, the same steps and parameters described above were used for inferring individual gene trees. Finally, both concatenated and coalescent species trees were inferred separately for each gene type (i.e., AHEs, UCEs, and Sanger

loci). Unless noted, for all downstream analyses, we used the species tree we obtained using ASTRAL and the dataset of 75% taxonomic completeness.

We then collapsed the branches of the gene trees with ultrafast bootstrap support values lower than 50, 75, and 95 using the software TreeCollapserCL4 (available from: <http://emmahodcroft.com/TreeCollapseCL.html>). These gene trees were then used to generate coalescent species trees of the Dipsadidae using ASTRAL). The topologies and support values of the resulting species trees were then compared. Further, we then filtered all the gene trees without collapsed branches according to three strategies. In the first place, we selected loci with taxonomic completeness of more than 25%, 50%, 75%, or 93%. We then filtered alignments larger than 474, 572, and 665base pairs, with the latter three corresponding to the 25%, 50%, and 75% percentiles of the distribution of alignment lengths, respectively. Finally, we filtered loci by missing data percentages, again using the 25%, 50%, and 75% percentiles of this variable. As described above, we then inferred the corresponding species trees for every one of these datasets and compared them to each other.

Finally, with the software phyparts 0.0.1 (Smith et al., 2015), we then determined the number and identity of the gene trees that either support or conflict with how each node was reconstructed in the ASTRAL species tree. Using that information, we identified those nodes for which more gene trees resolve the relationships of the subtending branches differently than in the species tree. Those branches are likely to be in the anomaly zone, a region of the tree space in which successive internal edges are so

short that it is more likely for gene trees to be discordant rather than concordant with the species tree (Degnan & Rosenberg 2006; Mendes & Hahn, 2018). To confirm if that is the case, we used scripts from Linkem et al. (2016) to evaluate whether any of the internal branches from the species tree is shorter than the minimum branch length (in coalescent units) expected if they were under the anomaly zone. Using the ASTRAL tree, we had inferred, we also obtained the corresponding values of gene and site concordant factors (gCF and sCF factors, respectively), which represent the percentage of decisive loci or alignment sites supporting the resolution of each of the branches of a given phylogeny (Minh et al., 2020). Phyparts analyses were performed after collapsing branches with ultrafast bootstrap supports of 50%, 75%, or 95%, and the results were then compared.

Exploring nonbiological causes of gene tree discordance

We then explored the potential causes of gene tree discordance in the intertribal relationships within the Xenodontinae. First, we explored the potential role that systematic gene tree estimation error could have on gene tree-species tree discordance. To do so, we first evaluated if the compiled genomic data is informative enough for distinguishing between alternate resolutions of the branches underlying the intertribal relationships within the Xenodontinae. This was performed using with Quartet Likelihood Mapping analyses (Strimmer & von Haeseler, 1997) for each branch of interest using IQ-TREE 2.2.0; after calculating gene and site concordance factors (gCFs and sCFs) also with IQ-TREE. We set the number of quartets randomly selected for

mapping or estimating site concordance factors as 2000 and 200, respectively. For Likelihood Mapping analyses, each of the three possible topologies between four taxa is represented as the corners of a triangle. The probability that each quartet of a set of gene trees supports these alternate topologies is mapped in the interior surface of the triangle (Strimmer & von Haeseler, 1997). Finding that most quartets are located towards the center of the triangle indicates that the data available are not very informative, as they have a reduced ability to differentiate between the alternate reconstructions for a branch of interest (Strimmer & von Haeseler, 1997).

We then used the results of the phyparts analysis considering only gene tree branches with support larger than 50 to identify which gene trees support, conflict, or are uninformative with how the intertribal relationships within the Xenodontinae are reconstructed in the coalescent species tree. For every gene tree the metric values of 13 traits frequently used as indicative of alignment or tree quality were then calculated (listed and defined in Table 1.1). Specifically, with regards to the traits of Maximum Phylogenetic Informativeness and the time in which this variable reaches a maximum were obtained with a modified version of the ‘informativeness.profile’ function of the R package PhyInformR (Dornburg et al., 2016). The inputs required by this function correspond to a calibrated phylogeny and maximum likelihood estimated site rates. The latter were obtained with the software HYPHY 2.5.32 (Hypothesis Testing using Phylogenies; Kosakovsky Pond et al., 2020). The calibrated tree was obtained by dating the ASTRAL species tree with the ‘chronos’ function of the R package ape (Paradis & Schliep, 2019). To do so, we implemented a correlated model of substitution rate

variation among branches and a smoothing parameter of 1. See Appendix Table A1.1 for details about the calibration points used to date this phylogeny. Details about how other traits were calculated are summarized in Table 1.1.

We then summarized all 13 calculated trait values by performing a Principal Components Analysis (PCA) with the function ‘prcomp’ of the software R 4.2.1 (R Core Team, 2021). The first five PC axes were selected since they explain ~60% of the compiled trait data variation. We then used the values of these PC axes corresponding to each locus, the type of gene they are part of (i.e., either AHE, Sanger loci, or UCE), and the interaction between these two factors as the independent variables of a series of multinomial logistic regressions we performed with the software R 4.2.1 (R Core Team, 2021). These logistic regressions were performed to test whether gene tree-species tree conflict could be predicted by the traits associated with each locus. For these regressions, the dependent variable was whether each gene tree supports, conflicts with, or is uninformative with the resolution of the species tree of any of the nodes underlying the high-level relationships within the Xenodontinae. In general, finding that, for a given PC axis and relationship of interest, the gene trees having values associated with high-quality alignments or trees tend to conflict with its resolution in the species tree would put the accuracy of this reconstruction into doubt. Instead, if loci with characteristics associated with low-quality alignments or gene trees tend to support the reconstruction of a given branch in the species tree, that would support the idea that alternate reconstructions of this branch are most likely artifacts. Finally, if the support or conflict of gene trees with the species tree is unrelated to the traits they (or their alignments) exhibit, that would

suggest that gene tree discordance instead results from random gene tree estimation error or the influence of processes such as Incomplete Lineage Sorting or deep introgression.

Exploring biological causes of gene tree discordance

We then evaluated if Incomplete Lineage Sorting (ILS) can explain by itself the observed levels of gene tree-species tree discordance. Assuming this to be the case, we simulated a total of 4896 gene trees (=the same number of gene trees as in the dataset of 75% taxonomic completeness) under the multispecies coalescent model, as implemented in the software Dendropy 4.5.2 (Sukumaran & Holder, 2010). We performed this simulation a total of 1000 times, considering for each replicate that the correct species tree was the ASTRAL phylogeny we had inferred. The ultrametricity of the ASTRAL species tree (as required for coalescent simulations in Dendropy) was guaranteed by extending its terminal branches with the function ‘force.ultrametric’ (method= “extend”) of the R package phytools (Revell, 2012). Given that the internal branch lengths of the species tree (in coalescent units) were not affected by this procedure, they were used as the population parameter of the simulations. For each replicate of the simulations we performed, we calculated the Information-Content corrected Robinson-Foulds (R-F) distances (Smith, 2020) between the ASTRAL species tree and with both 1) the gene trees we had inferred and 2) all simulated gene trees. However, given that for calculating R-F distance values, all trees were required to contain the same taxa, we had to prune beforehand both the species tree and each gene tree (either if obtained through simulation or inferred). While pruning these phylogenies, we made sure to retain at least one

member of the clades of interest (i.e., an outgroup, a member of the Dipsadinae, and one member of each of the sampled tribes of the Xenodontinae [Xenodontini, Psomophini, Elapomorphini, Echinantherini, Philodryadini, and Pseudoboini]). We filtered out trees lacking at least one member of these groups. Finally, we converted the average value of the R-F distances between the inferred gene trees to a standardized Z-score. We then obtained the p-value of that Z-score under a two-tailed test; for which a Z-score larger than 1.96 would correspond to statistical significance under a p-value of 0.05.

Finally, we evaluated whether the discordance we observed between gene trees could have resulted from introgression events at high-level phylogenetic scales. To do so, we calculated the value of Huson et al. (2005)'s Δ metric (as modified by Vanderpool et al., 2020) for each of the internal branches of the ASTRAL species tree. The Δ metric has been previously used in other studies testing for deep introgression events by using gene trees (e.g., Vanderpool et al., 2020; Cunha et al., 2022; Yu et al., 2022). The Δ metric is calculated by subtracting the number of gene trees supporting the two most common resolutions of a given branch and then dividing the absolute value of that number over the total of gene trees supporting either of these resolutions. If Incomplete Lineage Sorting were the only source of discordance, these two alternative resolutions would be in equal frequencies, and Δ would be equal to 0. However, if introgression is present, one of the two most common resolutions of a given branch should be more common than the other, and a Δ value larger than 0 would be expected. We calculated the values of Δ by implementing a modified script from Cunha et al. (2022) and using the gene concordance factors (gCFs) we had previously obtained. To determine if the observed Δ values were

significantly different from what would be expected by chance, we first obtained a null distribution by performing a total of 1000 resamplings (with repetition) of the 4896 gene trees we had inferred. We then calculated the gCFs of each resulting tree and obtained the values of Δ across every one of their branches. Finally, for each branch for which more than 5% of the inferred gene trees conflict with the species tree, we compared the observed values of Δ to those obtained through permutation. We did so by converting each inferred Δ into a standardized Z-score and calculating their corresponding p-values as a one-tailed test. Under this test, Z-scores equal to or larger than 1.65 would correspond to p-values of 0.05 or lower, thus indicating the existence of deep introgression in the branch of interest.

Results

Gene and species tree inference

The dataset of 75% taxonomic completeness contains a total of 2,650,049 base pairs corresponding to 4333 loci: 342 being AHEs, 3962 UCEs, and 28 corresponding to the traditionally used Sanger loci. On average, the compiled loci are 611.60 base pairs (1287.02 for the AHEs, 804.11 for the Sanger loci, and 551.95 for the UCEs) and have 11.77% missing data, including gaps (7.93% for the AHEs, 9.21% for the Sanger loci, 12.56% for the UCEs). Voucher numbers, number of loci, and average alignment length for all taxa included in this dataset are compiled in Appendix Table A1.2.

Phylogenomic data compilation and assembly

In general, most of the nodes of our inferred ASTRAL species tree (Figure 1.2) are strongly supported and concordant with how they are reconstructed in previous studies based on either traditional Sanger sequence data (e.g., Zaher et al., 2009; Pyron et al., 2015, Zaher et al., 2018; Ramirez et al., submitted; Ramirez & Reeder, in preparation) or with the SqCL loci set using coalescent-based methods (i.e., Singhal et al., 2017). Specifically, there is strong support for the monophyly of the Dipsadidae, its two main subfamilies, and all recognized tribes within the Xenodontinae (Figure 1.2). We also recovered strong support for sister taxa relationships between 1) Psomophini and Xenodontini, and 2) Echinantherini and Elapomorphini (Figure 1.2); both frequently found in previous studies (except for Pyron et al., 2013 and Ramirez & Reeder, in preparation, who recovered Psomophini as the earliest-diverging clade of the Xenodontinae followed by the Xenodontinae, Figure 1.1). However, the remaining intertribal relationships within the Xenodontinae are inconsistently resolved between studies, and most are weakly supported, as summarized in Figure 1.1. Specifically, for the inferred ASTRAL species tree, we found the Pseudoboini as weakly supported as the sister taxon to the remaining Xenodontinae. Within this more exclusive clade of Xenodontinae, there is a strongly supported clade containing Xenodontini+Psomophini that is sister to a weakly supported group containing Philodryadini, Elapomorphini, and Echinantherini, with the latter two tribes being each other closest relatives (Figure 1.2). The weakly supported relatively basal branches within Xenodontinae are extremely

short (≤ 0.01 coalescent units). Thus, at the present being they are best considered as polytomies.

The ASTRAL species tree obtained using only the UCE loci was almost identical to this topology based on all the data (Figure 1.2), only differing in the local posterior probability support values of the deep nodes within the Xenodontinae (Appendix Figure A1.1). The same topology was also obtained in the concatenated species tree obtained with all loci of the dataset of 75% taxonomic completeness and with only UCE loci (Appendix Figure A1.2). However, for both of these concatenated phylogenies, node support was high (=almost all ultrafast bootstrap support values were larger than 94). In contrast, for both coalescent and concatenated analyses obtained using AHEs, we recovered poorly supported sister group relationships for 1) the Philodryadini and the Psomophini+Xenodontini and 2) the Pseudoboini and the Echinantherini+Elapomorophini (Appendix Figures A1.1 and A1.2). Topologies obtained only using the Sanger loci with the concatenation and coalescent species approaches were very poorly supported and differed between themselves and the previously described topologies (Appendix Figures A1.1 and A1.2).

On the other hand, filtering loci by different values of alignment length or taxonomic completeness did not result in any noteworthy change in the inferred relationships or their support values in the corresponding species trees and when compared with the ASTRAL species tree of 75% taxonomic completeness (Appendix Figures A1.3 and A1.4). However, that was not the case when loci were filtered by the

percentage of missing data, where in all cases, Psomophini+Xenodontini was recovered as the sister group of all other xenodontines, and Pseudoboini (or Philodryadini when loci with missing data proportions of 25% or more were considered) as the next group to diverge from the rest (Appendix Figure A1.5). Finally, changing the minimum ultrafast bootstrap value where gene tree branches were collapsed resulted only in a change from the topology obtained in the ASTRAL tree when this metric was above 75. Interestingly, when the cutoff value of this metric was 95, the resulting topology was the same as that of the AHE species tree we obtained either through concatenation or with ASTRAL.

Exploring nonbiological causes of gene tree discordance

Apart from the MRCAs of the Xenodontini+Psomophini clade, the Echinantherini+Elapomorphini clade, and of *Pseudoboa nigra* and *Oxyrhopus petolarius*, all other nodes within the Xenodontinae have low values of site and gene concordance factors (<30 or 50, respectively, Appendix Figure A1.7). Thus, indicating that these nodes are associated with high levels of discordance between gene trees (Minh et al., 2020). Gene tree-species tree discordance is also elevated for these nodes according to phyparts, with more gene trees conflicting rather than supporting those nodes resolved in the ASTRAL species tree (Figure 1.3), independently of the ultrafast bootstrap cut-off value used to collapse gene tree branches (Appendix Figure A1.8). Such a result indicates that the branches underlying these nodes are in the anomaly zone, as confirmed by using a formula that tests for the presence of an anomaly zone (i.e., equation 1 from Linkem et al., 2016, see also Appendix Figure A1.7).

Results of the Likelihood Mapping analyses show that all quartets from the gene trees we inferred were located in the corners of each of the triangles in which they were mapped (Figure 1.4). These results indicate that, as a whole, the genomic dataset we compiled has the ability to distinguish between alternate phylogenetic reconstructions for each of the nodes of interest (Strimmer & von Haeseler, 1997; Minh et al., 2020).

In general, the first two axes of the PCA we performed with the 13 traits we compiled were primarily influenced by traits of phylogenetic informativeness (positive loadings of PI and of time of maximum PI, and negative loading of alignment length) and of tree quality (e.g., positive loadings of tree length, root tip-variation, and PI and time of maximum PI), respectively (Figure 1.5). On the other hand, the variables of alignment length, PI, and time of maximum PI are negatively loaded in PC Axis 3. PC axes have a limited ability to explain our trait data for any nodes considered, with only ~32% and ~61% cumulative proportion of data explained after including PC axes 1 through 2 or 5, respectively. Loadings and summary values for all PCAs performed are summarized in Appendix Tables A1.3 and A1.4, respectively. We found that only two of the five trait PC axes associated with alignment or tree quality had some ability to predict whether gene trees conflict, support, or neither with the ASTRAL species tree for each of the phylogenetic relationships of interest, with gene type (i.e., AHE, Sanger loci, or UCEs) only having a limited influence (Table 1.2). In summary, we demonstrated that values of PC1 and PC3 significantly differ between gene trees supporting and conflicting with the reconstruction in the species tree in the relationships between 1) Pseudoboini and the remainder of the Xenodontinae and 2) Psomophini and the Xenodontinae (Table 1.2).

The interaction between PC3 and whether loci are UCEs or AHEs also predicts whether loci support or are uninformative with the resolution of the relationships between Echinantherini and Elapomorphini and these two groups with the Philodryadini (Table 1.2). Despite these results, we found that in all cases, values of the first two PC axes overlap widely between gene trees supporting, conflicting, or neither with the relationships supported in the species trees (Figure 1.5). Similar results were obtained even after considering PCA axes 3-6 on each of the four relationships of interest (results not shown). Looking at each trait separately only reveals a few obvious differences between the values exhibited by loci conflicting, supporting, or neither with the relationship of interest in the species tree (Appendix Figures A1.9-A1.12). The most notable difference is that for all four relationships of interest, loci whose gene trees have greater values of UF bootstrap support averaged across all branches tend to support the reconstruction observed in the species tree (Appendix Figures A1.9-A1.12). In general, all the results described above tend to be maintained after collapsing gene tree branches with ultrafast bootstrap supports of 75% or 95% (results not shown).

Exploring biological causes of gene tree discordance

The average Information-Content corrected R-F distances between the inferred ASTRAL species tree and each rooted gene tree were completely overlapping and thus not significantly different from corresponding R-F distances obtained with the simulated gene trees (Figure 1.6). Such a result indicates that ILS is unable to explain the observed levels of gene tree-species tree discordance. Finally, none of the internal branches of the

ASTRAL species tree exhibited Δ metric values that were significantly different from zero (maximum obtained Z-value = 0.03; Z-values larger than 1.65 would indicate significance) (Table 1.3; Figure 1.7). Thus, the two most common resolutions of each branch of the species tree found across the inferred gene trees are found in similar frequencies, as expected under a null model with no introgression. This confirms that deep introgression does not appear to have had any role in explaining the observed levels of discordance between gene trees.

Discussion

In this study, we used the most up-to-date phylogenomic dataset available for the snake family Dipsadidae to explore the possible causes of continued uncertainty in the phylogenetic relationships of the group. In summary, our results indicate that with the currently available genomic data, even though the sampled dipsadid genera and tribes are recovered as monophyletic and strongly supported, the interrelationships between the tribes within the Xenodontinae remain essentially unresolved. Whereas the inferred relationships between the xenodontine tribes remain almost the same after using coalescent or concatenation-based species approaches and implementing several filtering schemes it needs to be noted that in all of these phylogenies, the branches underlying these relationships are extremely short, and weakly supported. Further, we found that independent of the species tree approach used, the topology obtained with the UCE loci is almost the same as the one used with all locus types from the SQCL loci set. Thus, indicating that the signal from the AHE and Sanger loci is being obscured by the more

numerous, albeit shorter UCEs genes. We also demonstrated that this lack of resolution in the relationships within the Xenodontinae does not appear to be caused by the lack of informativeness in the loci comprising our genomic dataset but instead because of extensive gene tree discordance. Our results also indicate that this discordance most likely results from the effects of both gene tree estimation error and extensive Incomplete Lineage Sorting (ILS) and not from deep introgression between the sampled lineages within the Xenodontinae.

Phylogenomics of the Xenodontinae

In general, apart from our inability to resolve the deep relationships within the Xenodontinae, our results consistently and strongly supported the monophyly of the Dipsadidae and its two main subfamilies and their tribes. As previously mentioned, these results have also been supported in all the previous phylogenies of the group obtained using both traditional Sanger sequence and genomic data, as well as concatenation and coalescent-based tree inference methods. Unfortunately, the very limited taxon sampling available within the subfamily Dipsadinae (two tribes sampled with two species each) prevented us from effectively testing the monophyly and interrelationships of the dipsadine tribes. However, conflict in the resolution of dipsadine tribes in previous studies has been limited and mostly corresponds to the non-monophyly or variable phylogenetic positions of several diverse and problematic genera (e.g., *Geophis*, *Tropidodipsas*, and *Nothopsis*). Whether this relative higher-level phylogenetic

consistency will continue to be observed in phylogenomic studies needs to be explored in the future.

On the other hand, our results conform to the trend among previous studies in finding that the inferred high-level relationships within the Xenodontinae are poorly supported and unstable. Indeed, and surprisingly, we even obtained different results from Singhal et al. (2017), who recovered a different set of relationships despite using practically the same dataset (although including taxa from all across the Squamata) and phylogenetic inference methodology as ours. However, considering the short branch lengths and poor support for the relationships inferred by Singhal et al. (2017), they are also best interpreted as essentially unresolved in the basal regions of the Xenodontinae phylogeny. More generally, the difficulty in resolving the relationships between the xenodontine tribes has not only involved the tribes included in this study. Almost all the tribes of the Xenodontinae from which phylogenomic data remain unavailable are not recovered in consistent phylogenetic position across the previous studies based on Sanger sequence data. The only relationships between the xenodontine tribes consistently resolved across all of these previous studies correspond to the sister group relationships between Echinantherini and Elapomorphini and between Tropidodryadini and Philodryadini. Further, Xenodontini is also consistently recovered to be closely related to the tribe Alsophini and several small tribes (i.e., the Eutrachelophiini, the Amnesteophiini, and the genus *Pseudalsophis* (e.g., Pyron et al., 2015; Zaher et al., 2018; Moraes et al., 2021; Ramirez et al., submitted). The clade represented by the Xenodontini and these other subtribes is also recovered as the sister group of the Psomophini in all

studies except for Pyron et al., 2013 and Ramirez & Reeder, in preparation). Thus, it is possible that even after increasing the representation of the different xenodontine tribes, their phylogenetic relationships would continue to be best described as unresolved. Such a possibility needs to be evaluated in future studies.

Exploring nonbiological causes of gene tree discordance

We also investigated the potential influence of Gene Tree Estimation Error (GTEE) on determining the discordance we observed between the topologies of the species trees and the gene trees. Our results find some limited evidence suggesting that the conflict or discord between gene trees and the species tree is more likely to be misconstrued or associated with low-quality alignments when compared to gene trees congruent with the species tree. That evidence indicates that loci whose gene trees conflict with the resolution of the species tree for the relationships of interest differ from those congruent with it (and/or those that are uninformative) in variables such as alignment length and the maximum value of phylogenetic informativeness. In this regard, filtering the gene tree dataset by the values of traits associated with GTEE would be expected to highlight any potential differences between gene trees conflicting or in concordance with the species tree. However, we did find this to be the case after filtering by alignment length. Further, filtering by other traits or changing the minimum support value of gene tree branches result in poorly supported and conflicted relationships within the Xenodontinae. In conjunction, these results suggest that GTEE is most likely closer to

being randomly distributed across gene trees independently of the reconstruction of the branch of interest they support.

Further, our findings contrast with that of several other studies, which have found that filtering loci (by traits such as the amount of missing data or parsimony informative sites on their alignments or their length) diminishes gene tree conflict and increases node support in the resulting species trees (e.g., Hosner et al., 2016; Blom et al., 2017; Alda et al., 2019, McLean et al., 2019; Mongiardino Koch, 2021, and Molloy and Warnow, 2018, for when levels of ILS and GTEE are low or moderate). Despite this, our study is not the first to illustrate that removing loci with characteristics associated with GTEE had a limited influence on species tree inference (e.g., Chan et al., 2020a; Chan et al., 2020b, and Molloy and Warnow, 2018, for when both levels of ILS and GTEE are elevated). Further, even among studies where filtering has been found to successfully diminish gene tree heterogeneity, there does not appear to be any trait that consistently outperforms the rest across different studies (Alda et al., 2019). Indeed, contradicting results have been obtained even after using metrics such as phylogenetic informativeness, which are expected to outperform the rest as they allow inspecting nodes at specific tree depths (Alda et al., 2019). Although this idea was corroborated by Alda et al. (2019), neither our study nor Singhal et al. (2021) found that gene trees supporting or conflicting with the species tree differ in their phylogenetic informativeness or in the timespan in which this metric reached its maximum value. It is apparent that the performance of the gene-filtering approach varies between datasets and lineages with different evolutionary histories, even when they are filtered by the same trait (Alda et al., 2019). In general, we

consider that the filtering approach is likely to be most useful when there is a tendency for gene tree-species tree conflict to occur more frequently amongst the loci which consistently exhibit traits associated with GTEE. Also, removing loci is probably more effective when gene trees tend to conflict with the resolution of clades consistently found to be monophyletic across many previous studies. Further studies are needed to evaluate whether these hypotheses are supported or not.

Exploring biological causes of gene tree discordance

Several lines of evidence obtained herein support the idea that the levels of discordance between gene trees (and with the species tree) are likely to result from extensive Incomplete Lineage Sorting. Indeed, in the first place, we found that more gene trees conflict rather than support the relationships of interest as recovered in the ASTRAL species tree. Such a result indicates that the branches underlying the relationships of interest are short enough to fall in the anomaly zone, as confirmed by the analyses we performed using the script from Linkem et al. (2016). In general, only a few other examples exist of studies recovering evidence that the relationships of interest are in the anomaly zone (such as Linkem et al., 2016; Cloutier et al., 2019; Chan et al., 2020; Singhal et al., 2021; and Morales-Briones et al., 2021). Finding this to be the case likely results from the few studies testing this phenomenon and the rarity in which it is expected to occur in real-life datasets, as demonstrated in Huang & Knowles's (2009) simulation study. Indeed, according to these authors, patterns consistent with the existence of an anomaly zone could also occur when gene trees are so uninformative regarding the

relationships of interest that they are best considered to correspond to polytomies (Huang & Knowles, 2009; Morales-Briones et al., 2021). Further supporting this result, we did not recover any noteworthy change in the resulting species tree after forcing poorly supported branches from the inferred gene trees to be polytomies. Despite this, and as in Cloutier et al. (2019), we recognize that it is difficult to determine whether short internodes in gene trees correspond to GTEE or a true anomaly zone, for which increased taxonomic sampling and the use of methods that simultaneously estimate gene trees and species trees will provide important insights.

In general, independently of whether the branches of interest are located or not in the anomaly zone, the results of our coalescent simulations indicate that ILS has a considerable role in determining the observed levels of gene tree discordance. Further, this is also supported by the analysis we performed to test for deep introgression discussed below, as we did not recover any significant deviation from its null expectation that ILS drives gene tree discordance. It remains unclear what caused the levels of ILS observed in the high-level relationships within the Xenodontinae. We are unaware of any key trait innovation that could have fostered the rapid radiation of the group, from which no morphological or life history synapomorphies are known to exist (Zaher et al., 2009). Further, there does not appear to be an obvious geological or biogeographical event that would be expected to be responsible for the fast origination of its high-level subclades. Indeed, the origin and early diversification of the Xenodontidae occurred in the Chacoan subregion of Morrone (2014) about 26-21 million years ago (Ramirez et al., submitted). Such a region mostly overlaps with the Brazilian shelf, known for being ancient (older

than 600 million years ago, Lovejoy et al., 2010) and very geologically stable during the entire Cenozoic (Wesselingh et al., 2010). Further, this geological structure is mostly below 1000 meters of elevation, and thus, it probably does not correspond to a strong dispersal barrier either (Wesselingh et al., 2010; Kay, 2015). The Brazilian shelf did suffer from several changes in sea levels and oceanic transgressions during the late Oligocene and Miocene (see Ortiz-Jaureguizar & Cladera, 2006; Rosetti et al., 2013) that surely had an important role in the diversification of the Xenodontinae, including one that occurred around the time in which the family originated. However, in this transgression (i.e., the Paranaense transgression), only the southernmost part of the Chacoan subregion was submerged (i.e., portions of the Chacoan and Pampa provinces of Morrone, 2015) (Ortiz-Jaureguizar & Cladera, 2006; Rosetti et al., 2013; Kay, 2015). The Chacoan subregion did appear to get isolated from its neighboring regions (i.e., the Guianan shelf, the basin of the proto-Amazon river, and the Andes) in the middle Miocene (Ortiz-Jaureguizar & Cladera, 2006; Morrone, 2014). However, as mentioned before, the main clades of the Xenodontinae are thought to have already emerged at that time. Of course, it is still possible that this geological event did promote the early diversification of the Xenodontinae but that the time estimates of when this occurred were overestimated by Ramirez et al. (submitted). However, no evidence at present suggests this to be the case. In general, further and more detailed biogeographic studies are needed to get a better understanding of the possible causes of the rapid origination of the high-level relationships within the Xenodontinae.

Finally, another potential biological cause of the extensive gene tree discordance recovered in this study is hybridization. The role of this process is increasingly recognized in many groups of organisms and has been reported in lineages diverging as far as ~20 million years ago (e.g., Vanderpool et al., 2020; Cai et al., 2021; MacGuigan & Near, 2019; DeBaun et al., 2023; but see Cunha et al., 2022 or Hughes et al., 2022 for counterexamples). For example, many introgression events were recently identified in the evolution of the Pseudoxyrhopiidae, an unrelated group of Malagasy snakes that share some similarities with the evolutionary history of the Dipsadidae (DeBaun et al., 2023). Indeed, this group originated around the same time as the Dipsadidae (~22 million years ago) after colonizing from mainland Africa, and both groups have evolved into many ecological roles (DeBaun et al., 2023, Ramirez et al., submitted). Despite this, when testing for the existence of introgression on the phylogeny of the Xenodontinae using the Δ metric of Huson et al.(2005), we did not find this to be the case at any of its branches. Thus, no current evidence supports the idea that the gene tree discordance observed in the Xenodontinae results from the influence of deep introgression. However, it needs to be considered that it is possible that events of introgression of this snake group did occur during the early evolution of the group but were not recovered in our analysis because they involved members of the tribes of this subfamily that we did not include herein or that have gotten extinct. Also, the Δ metric, as most other approaches testing for introgression at the phylogenetic scale, cannot identify these types of events when they occur between sister taxa (Hibbins & Hahn, 2022). Thus, it is possible that introgression between the members of the Xenodontinae included herein did occur, but that was not

detected because it involved sister lineages. In this regard, methods of network inference that account for this limitation (such as PhyloNet's InferNetwork_ML of Yu et al., 2014; and BEAST2's SpeciesNetwork from Zhang et al., 2018b) could be used to test this possibility. However, the implementation of these methods is limited by their computational intensiveness (Hibbins & Hahn, 2022), and it is unclear how they are affected by missing taxa. Thus, their application for inferring the potential existence of deep introgression in the Xenodontinae is deferred for future studies.

Conclusions

In conclusion, the results recovered in this study are an example that for some parts of the Tree of Life, simply increasing the amount of sequence data without a representative taxonomic sampling will not be sufficient to resolve the phylogenetic relationships amongst their members. Our results also support the idea that even upon careful inspection of the signatures from the main drivers of phylogenetic uncertainty caused by gene tree discordance, some recalcitrant nodes could still remain unsolved. Despite this, our study demonstrated that even under these circumstances, genomic-scale data could have a considerable role in corroborating and establishing support for other parts of the phylogeny. Further, the methodology we followed herein also allows gaining important insights about which specific factors likely determine gene tree discordance in the clade of interest and how they can influence our ability to infer an accurate species tree. Indeed, we recovered that Incomplete Lineage Sorting has a considerable role in determining the observed discordance between gene trees and with regards to the species

tree, with only limited evidence of a systematic bias in gene tree inference and none suggesting the influence of deep introgression. If this is indeed the case, one would expect ASTRAL to infer relatively accurate species trees, as this approach is robust to high levels of ILS (even when gene trees lie within the limits of the anomaly zone, Liu et al., 2019; Islam et al., 2020) when GTEE is reduced, and no hybridization occurs (Solís-Lemus et al., 2016; Long & Kubatko, 2018; Liu et al., 2019). However, testing whether this is the case can only be effectively tested by increasing the taxonomic sampling within the Xenodontinae and through the use of potentially more informative genomic information (such as rapidly evolving exons, Karin et al., 2019; or rare genomic changes such as intron indels, insertions of retroposons, Rokas & Holland, 2000; Cloutier et al., 2019), apart from implementing approaches that can reduce GTEE and more accurately identify introgression events (as described above). Doing so would not only be of interest to phylogeneticists, but it will also have implications for studies that use phylogenies to test hypotheses in areas such as trait evolution, spatial phylogenetics, and biogeography.

References

- Alda, F., Tagliacollo, V. A., Bernt, M. J., Waltz, B. T., Ludt, W. B., Faircloth, B. C., ... & Chakrabarty, P. (2019). Resolving deep nodes in an ancient radiation of neotropical fishes in the presence of conflicting signals from incomplete lineage sorting. *Systematic Biology*, 68(4), 573-593.
- Arcila, D., Ortí, G., Vari, R., Armbruster, J. W., Stiassny, M. L., Ko, K. D., ... & Betancur-R, R. (2017). Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. *Nature Ecology & Evolution*, 1(2), 0020.
- Arredondo, J. C., Grazziotin, F. G., Scrocchi, G. J., Rodrigues, M. T., Bonatto, S. L., & Zaher, H. (2020). Molecular phylogeny of the tribe Philodryadini Cope, 1886 (Dipsadidae: Xenodontinae): Rediscovering the diversity of the South American Racers. *Papéis Avulsos de Zoologia*, 60, e20206053-e20206053.
- Ballesteros, J. A., & Sharma, P. P. (2019). A critical appraisal of the placement of Xiphosura (Chelicerata) with account of known sources of phylogenetic error. *Systematic Biology*, 68(6), 896-917.
- Betancur-R, R., Arcila, D., Vari, R. P., Hughes, L. C., Oliveira, C., Sabaj, M. H., & Ortí, G. (2019). Phylogenomic incongruence, hypothesis testing, and taxonomic sampling: The monophyly of characiform fishes. *Evolution*, 73(2), 329-345.
- Blischak, P. D., Chifman, J., Wolfe, A. D., & Kubatko, L. S. (2018). HyDe: a Python package for genome-scale hybridization detection. *Systematic Biology*, 67(5), 821-829.
- Blom, M. P., Bragg, J. G., Potter, S., & Moritz, C. (2017). Accounting for uncertainty in gene tree estimation: summary-coalescent species tree inference in a challenging radiation of Australian lizards. *Systematic Biology*, 66(3), 352-366.
- Bossert, S., Murray, E. A., Pauly, A., Chernyshov, K., Brady, S. G., & Danforth, B. N. (2021). Gene tree estimation error with ultraconserved elements: An empirical study on *Pseudapis* bees. *Systematic Biology*, 70(4), 803-821.
- Brown, J. W., Walker, J. F., & Smith, S. A. (2017). Phyx: phylogenetic tools for unix. *Bioinformatics*, 33(12), 1886-1888.
- Cai, L., Xi, Z., Lemmon, E. M., Lemmon, A. R., Mast, A., Buddenhagen, C. E., ... & Davis, C. C. (2021). The perfect storm: gene tree estimation error, incomplete lineage sorting, and ancient gene flow explain the most recalcitrant ancient angiosperm clade, Malpighiales. *Systematic Biology*, 70(3), 491-507.

- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17(4), 540-552.
- Chan, K. O., Hutter, C. R., Wood Jr, P. L., Grismer, L. L., & Brown, R. M. (2020). Larger, unfiltered datasets are more effective at resolving phylogenetic conflict: Introns, exons, and UCEs resolve ambiguities in Golden-backed frogs (Anura: Ranidae; genus *Hylarana*). *Molecular Phylogenetics and Evolution*, 151, 106899.
- Chan, K. O., Hutter, C. R., Wood Jr, P. L., Grismer, L. L., & Brown, R. M. (2020). Target-capture phylogenomics provide insights on gene and species tree discordances in Old World treefrogs (Anura: Rhacophoridae). *Proceedings of the Royal Society B*, 287(1940), 1-9.
- Cloutier, A., Sackton, T. B., Grayson, P., Clamp, M., Baker, A. J., & Edwards, S. V. (2019). Whole-genome analyses resolve the phylogeny of flightless birds (Palaeognathae) in the presence of an empirical anomaly zone. *Systematic Biology*, 68(6), 937-955.
- Cunha, T. J., Reimer, J. D., & Giribet, G. (2022). Investigating sources of conflict in deep phylogenomics of vetigastropod snails. *Systematic Biology*, 71(4), 1009-1022.
- DeBaun, D., Rabibisoa, N., Raselimanana, A. P., Raxworthy, C. J., & Burbrink, F. T. (2023). Widespread reticulate evolution in an adaptive radiation. *Evolution*, 77(4), 931-945.
- Degnan, J. H., & Rosenberg, N. A. (2006). Discordance of species trees with their most likely gene trees. *PLoS Genetics*, 2(5), e68.
- Dornburg, A., Fisk, J. N., Tamagnan, J., & Townsend, J. P. (2016). PhyInformR: phylogenetic experimental design and phylogenomic data exploration in R. *BMC Evolutionary Biology*, 16(1), 1-7.
- Douglas, J., Jiménez-Silva, C. L., & Bouckaert, R. (2022). StarBeast3: Adaptive Parallelized Bayesian Inference under the Multispecies Coalescent. *Systematic Biology*, 71 (4): 901–916.
- Edelman, N. B., Frandsen, P. B., Miyagi, M., Clavijo, B., Davey, J., Dikow, R. B., ... & Mallet, J. (2019). Genomic architecture and introgression shape a butterfly radiation. *Science*, 366(6465), 594-599.
- Faircloth, B. C. (2016). PHYLUCS is a software package for the analysis of conserved genomic loci. *Bioinformatics*, 32(5), 786-788.

- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39(4), 783-791.
- Giarla, T. C., & Esselstyn, J. A. (2015). The challenges of resolving a rapid, recent radiation: empirical and simulated phylogenomics of Philippine shrews. *Systematic Biology*, 64(5), 727-740.
- Hibbins, M. S., & Hahn, M. W. (2022). Phylogenomic approaches to detecting and characterizing introgression. *Genetics*, 220(2), iyab173.
- Hime, P. M., Lemmon, A. R., Lemmon, E. C. M., Prendini, E., Brown, J. M., Thomson, R. C., ... & Weisrock, D. W. (2021). Phylogenomics reveals ancient gene tree discordance in the amphibian tree of life. *Systematic Biology*, 70(1), 49-66.
- Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35(2), 518-522.
- Hosner, P. A., Faircloth, B. C., Glenn, T. C., Braun, E. L., & Kimball, R. T. (2016). Avoiding missing data biases in phylogenomic inference: an empirical study in the landfowl (Aves: Galliformes). *Molecular Biology and Evolution*, 33(4), 1110-1125.
- Huang, H., & Knowles, L. L. (2009). What is the danger of the anomaly zone for empirical phylogenetics?. *Systematic Biology*, 58(5), 527-536.
- Hughes, L. C., Nash, C. M., White, W. T., & Westneat, M. W. (2022). Concordance and Discordance in the Phylogenomics of the Wrasses and Parrotfishes (Teleostei: Labridae). *Systematic Biology*. syac072. Advance online publication. <https://doi.org/10.1093/sysbio/syac072>
- Huson, D. H., Klöpper, T., Lockhart, P. J., & Steel, M. A. (2005, May). Reconstruction of reticulate networks from gene trees. In *Annual International Conference on Research in Computational Molecular Biology* (pp. 233-249). Springer, Berlin, Heidelberg.
- Islam, M., Sarker, K., Das, T., Reaz, R., & Bayzid, M. (2020). STELAR: a statistically consistent coalescent-based species tree estimation method by maximizing triplet consistency. *BMC Genomics*, 21(1), 1-13.
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A., & Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587-589.

- Karin, B. R., Gamble, T., & Jackman, T. R. (2020). Optimizing phylogenomics with rapidly evolving long exons: comparison with anchored hybrid enrichment and ultraconserved elements. *Molecular Biology and Evolution*, 37(3), 904-922.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772-780.
- Kay, R. F. (2015). Biogeography in deep time—What do phylogenetics, geology, and paleoclimate tell us about early platyrrhine evolution?. *Molecular Phylogenetics and Evolution*, 82, 358-374.
- Kosakovsky Pond, S. L., Poon, A. F., Velazquez, R., Weaver, S., Hepler, N. L., Murrell, B., ... & Muse, S. V. (2020). HyPhy 2.5—a customizable platform for evolutionary hypothesis testing using phylogenies. *Molecular Biology and Evolution*, 37(1), 295-299.
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21), 4453-4455.
- Kück, P., & Struck, T. H. (2014). BaCoCa—A heuristic software tool for the parallel assessment of sequence biases in hundreds of gene and taxon partitions. *Molecular Phylogenetics and Evolution*, 70, 94-98.
- Kulkarni, S., Kallal, R. J., Wood, H., Dimitrov, D., Giribet, G., & Hormiga, G. (2021). Interrogating genomic-scale data to resolve recalcitrant nodes in the spider tree of life. *Molecular Biology and Evolution*, 38(3), 891-903.
- Leaché, A. D., Harris, R. B., Rannala, B., & Yang, Z. (2014). The influence of gene flow on species tree estimation: a simulation study. *Systematic Biology*, 63(1), 17-30.
- Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61(5), 727-744.
- Linkem, C. W., Minin, V. N., & Leaché, A. D. (2016). Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (Squamata: Scincidae). *Systematic Biology*, 65(3), 465-477.
- Liu, L., Anderson, C., Pearl, D., Edwards, S.V. 2019. Modern phylogenomics: building phylogenetic trees using the multispecies coalescent model. In: Anisimova, M, editor. *Evolutionary genomics: statistical and computational methods*. New York, NY: Springer New York. p. 211–239.

- Long, C., & Kubatko, L. (2018). The effect of gene flow on coalescent-based species-tree inference. *Systematic Biology*, 67(5), 770-785.
- Lopes, F., Oliveira, L. R., Kessler, A., Beux, Y., Crespo, E., Cárdenas-Alayza, S., ... & Bonatto, S. L. (2021). Phylogenomic discordance in the eared seals is best explained by incomplete lineage sorting following explosive radiation in the southern hemisphere. *Systematic Biology*, 70(4), 786-802.
- Lovejoy, N. R., Willis, S. C., & Albert, J. S. (2010). Molecular signatures of Neogene biogeographical events in the Amazon fish fauna. In C. Hoorn & F. P. Wesselingh (Eds.), *Amazonia: Landscape and species evolution: A look into the past* (pp. 405– 417). Wiley-Blackwell.
- MacGuigan, D. J., & Near, T. J. (2019). Phylogenomic signatures of ancient introgression in a rogue lineage of darters (Teleostei: Percidae). *Systematic Biology*, 68(2), 329-346.
- Mclean, B. S., Bell, K. C., Allen, J. M., Helgen, K. M., & Cook, J. A. (2019). Impacts of inference method and data set filtering on phylogenomic resolution in a rapid radiation of ground squirrels (Xerinae: Marmotini). *Systematic Biology*, 68(2), 298-316.
- Mendes, F. K., & Hahn, M. W. (2018). Why concatenation fails near the anomaly zone. *Systematic Biology*, 67(1), 158-169.
- Minh, B. Q., Nguyen, M. A. T., & Von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*, 30(5), 1188-1195.
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37(5), 1530-1534.
- Mirarab, S. (2017). Phylogenomics: constrained gene tree inference. *Nature Ecology & Evolution*, 1(2), 0056.
- Mongiardino Koch, N. (2021). Phylogenomic subsampling and the search for phylogenetically reliable loci. *Molecular Biology and Evolution*, 38(9), 4025-4038.
- Molloy, E. K., & Warnow, T. (2018). To include or not to include: the impact of gene filtering on species tree estimation methods. *Systematic Biology*, 67(2), 285-303.

- Moraes, L. J., Entiauspe-Neto, O. M., de Fraga, R., Fernandes, I. Y., & Werneck, F. P. (2021). Systematics of the rare Amazonian genus *Eutrachelophis* (Serpentes: Dipsadidae), with an emended diagnosis for *Eutrachelophis papilio*. *Zoologischer Anzeiger*, 295, 191-204.
- Morales-Briones, D. F., Kadereit, G., Tefarikis, D. T., Moore, M. J., Smith, S. A., Brockington, S. F., ... & Yang, Y. (2021). Disentangling sources of gene tree discordance in phylogenomic data sets: testing ancient hybridizations in *Amaranthaceae* sl. *Systematic Biology*, 70(2), 219-235.
- Morrone, J. J. (2014). Biogeographical regionalisation of the Neotropical region. *Zootaxa*, 3782(1), 1-110.
- Myers, C. W. (2011). A new genus and new tribe for *Enicognathus melanauchen* Jan, 1863, a neglected South American snake (Colubridae: Xenodontinae), with taxonomic notes on some Dipsadinae. *American Museum Novitates*, 2011(3715), 1-33.
- Myers, C. W., & McDowell, S. B. (2014). New taxa and cryptic species of Neotropical snakes (Xenodontinae), with commentary on hemipenes as generic and specific characters. *Bulletin of the American Museum of Natural History*, 2014(385), 1-112.
- Ortiz-Jaureguizar, E., & Cladera, G. A. (2006). Paleoenvironmental evolution of southern South America during the Cenozoic. *Journal of Arid Environments*, 66(3), 498-532.
- Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526-528.
- Pyron, R. A., Burbrink, F. T., & Wiens, J. J. (2013). A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evolutionary Biology*, 13(1), 1-54.
- Pyron, R. A., Guayasamin, J. M., Peñafiel, N., Bustamante, L., & Arteaga, A. (2015). Systematics of Nothopsini (Serpentes, Dipsadidae), with a new species of *Synophis* from the Pacific Andean slopes of southwestern Ecuador. *ZooKeys*, (541), 109.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

- Rokas, A., & Holland, P. W. (2000). Rare genomic changes as a tool for phylogenetics. *Trends in Ecology & Evolution*, 15(11), 454-459.
- Singhal, S., Grundler, M., Colli, G., & Rabosky, D. L. (2017). Squamate Conserved Loci (Sq CL): A unified set of conserved loci for phylogenomics and population genetics of squamate reptiles. *Molecular Ecology Resources*, 17(6), e12-e24.
- Singhal, S., Colston, T. J., Grundler, M. R., Smith, S. A., Costa, G. C., Colli, G. R., ... & Rabosky, D. L. (2021). Congruence and conflict in the higher-level phylogenetics of squamate reptiles: an expanded phylogenomic perspective. *Systematic Biology*, 70(3), 542-557.
- Smith, M. R. (2020). Information theoretic generalized Robinson–Foulds metrics for comparing phylogenetic trees. *Bioinformatics*, 36(20), 5007-5013.
- Solís-Lemus, C., Yang, M., & Ané, C. (2016). Inconsistency of species tree methods under gene flow. *Systematic Biology*, 65(5), 843-851.
- Ramirez, J.P. & T.W. Reeder (in preparation). Regional speciation drives spatial variation in species richness in the most diverse family of Neotropical snakes (Colubroidea: Dipsadidae).
- Ramirez, J.P., Reeder, T.W., & M.J. Spasojevic (submitted). Extinction debt and functional traits mediate community saturation over large spatiotemporal scales. Target: *Journal of Animal Ecology*
- Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, (2), 217-223.
- Rossetti, D. F., Bezerra, F. H., & Dominguez, J. M. (2013). Late Oligocene–Miocene transgressions along the equatorial and eastern margins of Brazil. *Earth-Science Reviews*, 123, 87-112.
- Shimodaira, H. (2002). An approximately unbiased test of phylogenetic tree selection. *Systematic Biology*, 51(3), 492-508.
- Smith, M. R. (2020). Information theoretic generalized Robinson–Foulds metrics for comparing phylogenetic trees. *Bioinformatics*, 36(20), 5007-5013.
- Sayyari, E., & Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. *Molecular Biology and Evolution*, 33(7), 1654-1668.

- Sheehy, C.M. III. (2012) Phylogenetic relationships and feeding behavior of Neotropical snail-eating snakes (Dipsadinae, Dipsadini). PhD Thesis, UT Arlington, Texas, USA, 135 pp.
- Singhal, S., Grundler, M., Colli, G., & Rabosky, D. L. (2017). Squamate conserved loci (Sq CL): a unified set of conserved loci for phylogenomics and population genetics of squamate reptiles. *Molecular Ecology Resources*, 17(6), e12-e24.
- Singhal, S., Colston, T. J., Grundler, M. R., Smith, S. A., Costa, G. C., Colli, G. R., ... & Rabosky, D. L. (2021). Congruence and conflict in the higher-level phylogenetics of squamate reptiles: an expanded phylogenomic perspective. *Systematic Biology*, 70(3), 542-557.
- Singhal, S., Colli, G. R., Grundler, M. R., Costa, G. C., Prates, I., & Rabosky, D. L. (2022). No link between population isolation and speciation rate in squamate reptiles. *Proceedings of the National Academy of Sciences*, 119(4), e2113388119.
- Smith, B. T., Harvey, M. G., Faircloth, B. C., Glenn, T. C., & Brumfield, R. T. (2014). Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. *Systematic Biology*, 63(1), 83-95.
- Smith, S. A., Moore, M. J., Brown, J. W., & Yang, Y. (2015). Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. *BMC Evolutionary Biology*, 15(1), 1-15.
- Solis-Lemus, C., & Ané, C. (2016). Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLoS Genetics*, 12(3), e1005896.
- Solis-Lemus, C., Yang, M., & Ané, C. (2016). Inconsistency of species tree methods under gene flow. *Systematic Biology*, 65(5), 843-851.
- Strimmer, K., & Von Haeseler, A. (1997). Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proceedings of the National Academy of Sciences*, 94(13), 6815-6819.
- Sukumaran, J., & Holder, M. T. (2010). DendroPy: a Python library for phylogenetic computing. *Bioinformatics*, 26(12), 1569-1571.
- Uetz, P., Freed, P. & Hošek, J. (Eds.) (2023). The Reptile Database, <http://www.reptile-database.org>, accessed 23 January 2023.

- Vanderpool, D., Minh, B. Q., Lanfear, R., Hughes, D., Murali, S., Harris, R. A., ... & Hahn, M. W. (2020). Primate phylogenomics uncovers multiple rapid radiations and ancient interspecific introgression. *PLoS Biology*, 18(12), e3000954.
- Wen, D., Yu, Y., Zhu, J., & Nakhleh, L. (2018). Inferring phylogenetic networks using PhyloNet. *Systematic Biology*, 67(4), 735-740.
- Wesselingh, F. P., Hoorn, C., Kroonenberg, S. B., Antonelli, A., Lundberg, J.G., Vonhof, H. B & Hooghiemstra, H. (2010). On the origin of Amazonian landscapes and biodiversity: a synthesis. In: C. Hoorn, F. P. Wesselingh. *Amazonia: Landscape and Species Evolution: A Look into the Past*. Blackwell Publishing Ltd.
- Wiens, J. J., Hutter, C. R., Mulcahy, D. G., Noonan, B. P., Townsend, T. M., Sites Jr, J. W., & Reeder, T. W. (2012). Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. *Biology Letters*, 8(6), 1043-1046.
- Yan, Z., Smith, M. L., Du, P., Hahn, M. W., & Nakhleh, L. (2022). Species tree inference methods intended to deal with incomplete lineage sorting are robust to the presence of paralogs. *Systematic Biology*, 71(2), 367-381.
- Yin, J., Zhang, C., & Mirarab, S. (2019). ASTRAL-MP: scaling ASTRAL to very large datasets using randomization and parallelization. *Bioinformatics*, 35(20), 3961-3969
- Yu, Y., Dong, J., Liu, K. J., & Nakhleh, L. (2014). Maximum likelihood inference of reticulate evolutionary histories. *Proceedings of the National Academy of Sciences*, 111(46), 16448-16453.
- Yu, D., Ding, Y., Tihelka, E., Cai, C., Hu, F., Liu, M., & Zhang, F. (2022). Phylogenomics of elongate-bodied springtails reveals independent transitions from aboveground to belowground habitats in deep time. *Systematic Biology*, 71(5), 1023-1031.
- Zaher, H., Grazziotin, F. G., Cadle, J. E., Murphy, R. W., Moura-Leite, J. C. D., & Bonatto, S. L. (2009). Molecular phylogeny of advanced snakes (Serpentes, Caenophidia) with an emphasis on South American Xenodontines: a revised classification and descriptions of new taxa. *Papéis Avulsos de Zoologia*, 49, 115-153.
- Zaher, H., Yáñez-Muñoz, M. H., Rodrigues, M. T., Graboski, R., Machado, F. A., Altamirano-Benavides, M., ... & Grazziotin, F. G. (2018). Origin and hidden diversity within the poorly known Galápagos snake radiation (Serpentes: Dipsadidae). *Systematics and Biodiversity*, 16(7), 614-642.

- Zaher, H., Murphy, R. W., Arredondo, J. C., Graboski, R., Machado-Filho, P. R., Mahlow, K., ... & Grazziotin, F. G. (2019). Large-scale molecular phylogeny, morphology, divergence-time estimation, and the fossil record of advanced caenophidian snakes (Squamata: Serpentes). *PloS One*, *14*(5), e0216148.
- Zhang, C., Ogilvie, H. A., Drummond, A. J., & Stadler, T. (2018). Bayesian inference of species networks from multilocus sequence data. *Molecular Biology and Evolution*, *35*(2), 504-517.
- Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC bioinformatics*, *19*(6), 15-30.

Tables

Table 1.1. List and definition of the thirteen trait metrics considered indicative of Gene Tree Estimation Error (GTEE) calculated for each locus. Reported values are those expected to correspond to increased GTEE levels. Modified from Singhal et al. (2020).

Trait metric	Gene tree or alignment trait	Definition	Preferred values and justification	Methodological details
Average UF bootstrap support	Tree	Average of the Ultrafast bootstrap (UF) support values of all the branches of a given gene tree.	Higher values are preferred. Lower values indicate less certainty in the relationships inferred on each gene tree and can affect species tree inference.	None
Branch outliers	Tree	Number of branches more than 5X larger than mean branch lengths.	Smaller values are preferred. Presence of many outliers can indicate that alignments are of low quality.	None
GC content	Alignment	Mean GC proportion of sequences.	Reduced values are preferred. High GC content may impact accuracy gene tree inferences.	Obtained with the function GC.content of R package ape (Paradis & Schliep, 2019).
Alignment length	Alignment	Mean alignment length.	Ambiguous. Short alignments are less likely to have undergone recombination but are potentially less informative than longer alignments.	None
Maximum phylogenetic informativeness (PI)	Tree	Maximum value of Phylogenetic Informativeness (PI).	Ambiguous. Depends on the depth of the nodes of interest.	See Appendix Table A1.1.

Maximum PI time period	Tree	Time (in millions of years) in which PI reached its maximum value.	Ambiguous. Depends on the depth of the nodes of interest.	See Appendix Table A1.1
Mean residuals	Tree	Mean residual of the regression of the root-tip length of gene trees against their root-tip node depth.	Lower residuals are preferred. Higher values can indicate higher evolutionary rate heterogeneity.	None
Missing data percentage	Alignment	Average percentage of missing data in each alignment.	Less missing data is preferred. Higher percentage of missing data can bias gene tree inference.	Obtained with the function base.freq of R package ape (Paradis & Schliep, 2019).
Occupancy	Alignment	Number of species with sequence data available for each alignment.	Higher values are preferred. Lower occupancy can bias gene tree inference.	None
Relative composition frequency variability (RFCV)	Alignment	Measures nucleotide compositional bias of sequences.	Less heterogeneity is preferred. Higher values may impact accuracy of gene tree inference because of violation of the assumption of nucleotide homogeneity.	Obtained with software BaCoCa (Base Composition Calculator; Kück & Struck, 2014).
Root-tip variance	Tree	Variance in the distance between the tips of gene trees and their root.	Lower values are preferred. Higher root-tip variances can indicate high evolutionary rate heterogeneity.	None
Saturation C value	Alignment	Measures sequence nucleotide composition saturation.	Low values are preferred. High values of saturation difficult the accuracy of gene tree inference.	Obtained with the software BaCoCa (Base Composition Calculator; Kück & Struck, 2014).

Tree length	Tree	Sum of all branch lengths.	Ambiguous. Shorter trees are less informative (as short branches generally indicate lesser numbers of nucleotide substitutions), but longer tree lengths could also result from having incorrect alignments.	None
--------------------	------	----------------------------	--	------

Table 1.2. P-values of the multinomial regressions from the PC analysis performed using the dataset of 4333 genes from the Squamate Conserved Loci set with 75% or more taxonomic completeness. Total of 4 multinomial regressions performed, one for each of the relationships of interest: a) relationship between Pseudoboini and remaining members of the Xenodontinae, b) Psomophini+Xenodontini, c) Echinantherini+Elapomorphini, and d) Philodryadini sister to Echinantherini+Elapomorphini. For these regressions, we evaluated whether loci supporting, conflicting, or uninformative with regards to the resolution of a given relationship in the species tree could be predicted by the first five axes of the trait Principal Component analyses performed, gene type (either AHE, Sanger, or UCE loci), or the interaction between these two factors). Significant p-values are indicated in bold.

	a)		b)		c)		d)	
	Conflicti ng	Unformati ve	Conflicti ng	Unformati ve	Conflicti ng	Unformati ve	Conflicti ng	Uniforma tive
Intercept	1.62E-06	2.16E-05	0.001	0.192	0.001	0.017	1.91E-05	0.129
PC Axis 1 value	0.026	0.030	0.028	0.511	0.496	0.514	0.076	0.205
Gene type: gene	0.572	0.666	0.851	0.169	0.544	0.592	0.393	0.651
Gene type: UCE	0.054	0.044	0.663	0.738	0.993	0.819	0.914	0.154
PC Axis 2 value	0.729	0.303	0.559	0.141	0.794	0.709	0.733	0.187
PC Axis 3 value	0.049	0.104	0.010	0.016	0.211	0.055	0.826	0.209
PC Axis 4 value	0.202	0.114	0.077	0.216	0.737	0.747	0.439	0.259
PC Axis 5 value	0.684	0.367	0.546	0.090	0.632	0.490	0.289	0.111
Interaction PC Axis 1 value*Gene type: gene	0.793	0.775	0.508	0.319	0.796	0.779	0.685	0.896
Interaction PC Axis 1 value*Gene type: UCE	0.279	0.239	0.531	0.093	0.617	0.763	0.675	0.063
Interaction PC Axis 2 value*Gene type: gene	0.892	0.915	0.391	0.406	0.927	0.897	0.673	0.566
Interaction PC Axis 2 value*Gene type: UCE	0.672	0.787	0.017	0.615	0.216	0.717	0.101	0.697
Interaction PC Axis 3 value*Gene type: gene	0.953	0.928	0.554	0.348	0.992	0.988	0.328	0.307

Interaction PC Axis 3 value*Gene type: UCE	0.878	0.998	0.236	0.482	0.151	0.039	0.039	2.22E-06
Interaction PC Axis 4 value*Gene type: gene	0.886	0.856	0.840	0.379	0.900	0.897	0.074	0.885
Interaction PC Axis 4 value*Gene type: UCE	0.446	0.343	0.312	0.281	0.479	0.935	0.145	0.305
Interaction PC Axis 5 value*Gene type: gene	0.991	0.948	0.598	0.670	0.979	0.939	0.222	0.552
Interaction PC Axis 5 value*Gene type: UCE	0.659	0.974	0.380	0.736	0.866	0.895	0.845	0.141

Table 1.3. Values of the Huson et al. (2005) Δ metric of deep introgression for the internal branches of the coalescent-based species tree. Branches for which less than 5% of gene trees were reconstructed differently from the species tree were not included.

Node ID.	Delta value	Z-score	P-value
MRCA of Colubriiformes (=Caenophidia except for the Xenodermidae and <i>Acrochordus</i>)	0.3	0.018672	0.492552
MRCA of Endoglyptodonta (=Colubriiformes except for the Pareidae)	0.048048	-0.10216	0.540686
MRCA of Homalopsidae+Elapoidea	0.12987	-0.05498	0.521923
MRCA of Elapoidea	0.027763	-0.14263	0.556709
MRCA of all Dipsadinae	0.121019	0.007438	0.497033
MRCA of Imantodini	0.099338	-0.16334	0.564874
MRCA of Dipsadini	0.041775	-0.13099	0.552109
MRCA of Pseudoboini (except for <i>Oxyrhopus trigeminus</i>)	0.018792	-0.36316	0.641759
MRCA of Pseudoboa+ <i>Oxyrhopus</i> (except for <i>O. trigeminus</i>)	0.048843	-0.23599	0.593281
MRCA of <i>Oxyrhopus petolarius</i> , <i>Pseudoboa neuwiedii</i> , and <i>P. nigr</i> 1	0.071085	-0.01223	0.504878
MRCA of all Xenodontinae excluding the Pseudoboini	0.024523	-0.62054	0.73255
MRCA of <i>Erythrolamprus</i> + <i>Xenodon</i>	0.136812	-0.00886	0.503534
MRCA of <i>Erythrolamprus</i>	0.203883	-0.00085	0.50034
MRCA of <i>Erythrolamprus poecilogyrus</i> and <i>E. almadensis</i>	0.133949	-0.01738	0.506935
MRCA of Philodryadini+Echinantherini+Elapomorphini	0.205128	-0.0082	0.503272
MRCA of Echinantherini+Elapomorphini	0.090312	-0.0121	0.504827

Figures

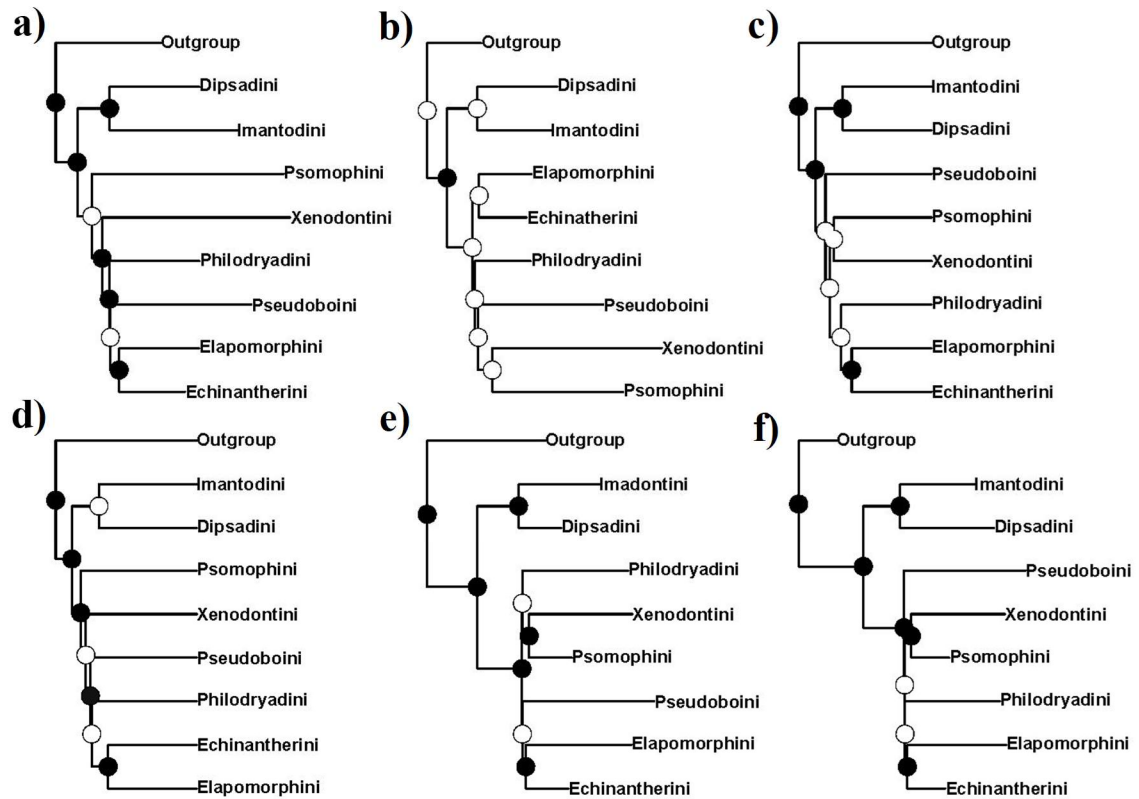


Figure 1.1. Summary of the phylogenetic relationships between the members of the Dipsadidae recovered in previous studies. Lineages of the Dipsadidae not sampled in this work were pruned from the phylogenies to facilitate the comparison between studies. High and low-supported nodes are indicated with black and white circles, respectively. (a) Pyron et al. (2013) concatenated maximum likelihood (ML) analysis of five mitochondrial and seven nuclear loci. (b) Zaher et al. (2018) concatenated ML analysis of three mitochondrial and three nuclear loci. (c) Ramirez et al. (submitted) and (d) Ramirez & Reeder (in preparation) time-calibrated concatenated Bayesian analysis (BA) of five mitochondrial and four nuclear loci. The node indicated with a gray circle has a posterior probability of 0.93. (e) Singhal et al. (2017) coalescent-based species tree analysis using ASTRAL and 5462 nuclear loci. (f). Coalescent-based species tree obtained for this study. Support values correspond to non-parametric bootstrap values (for a and b), Bayesian posterior probabilities (for c and d), or local posterior probabilities (for e and f).

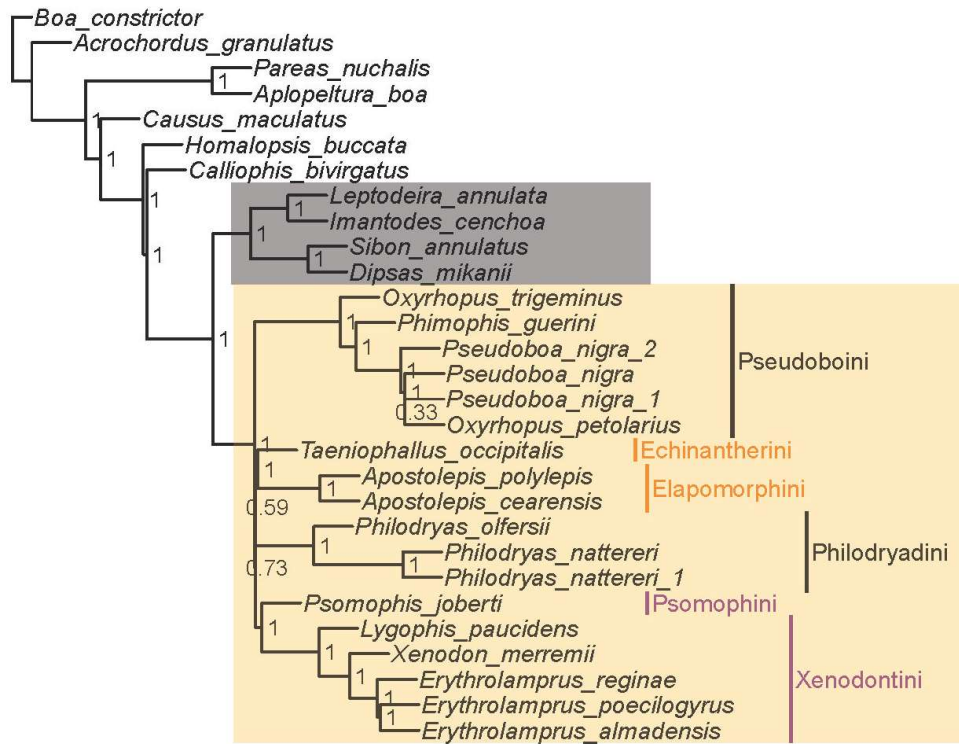


Figure 1.2. Phylogenetic relationships of the snake family Dipsadidae, as inferred with coalescent-based species tree analyses performed with the software ASTRAL-III. Numbers on each node correspond to local posterior probabilities values. The two main subfamilies of the Dipsadidae are highlighted with dark grey for the case of the Dipsadinae or pale gold for the Xenodontinae. The tribes of the Xenodontinae included in this study are also labeled. The labels corresponding to the two pairs of tribes consistently found to be related in all analyses are shown in orange (for Echinantherini+Elapomorpha) or dark magenta (for Psomophini+Xenodontini).

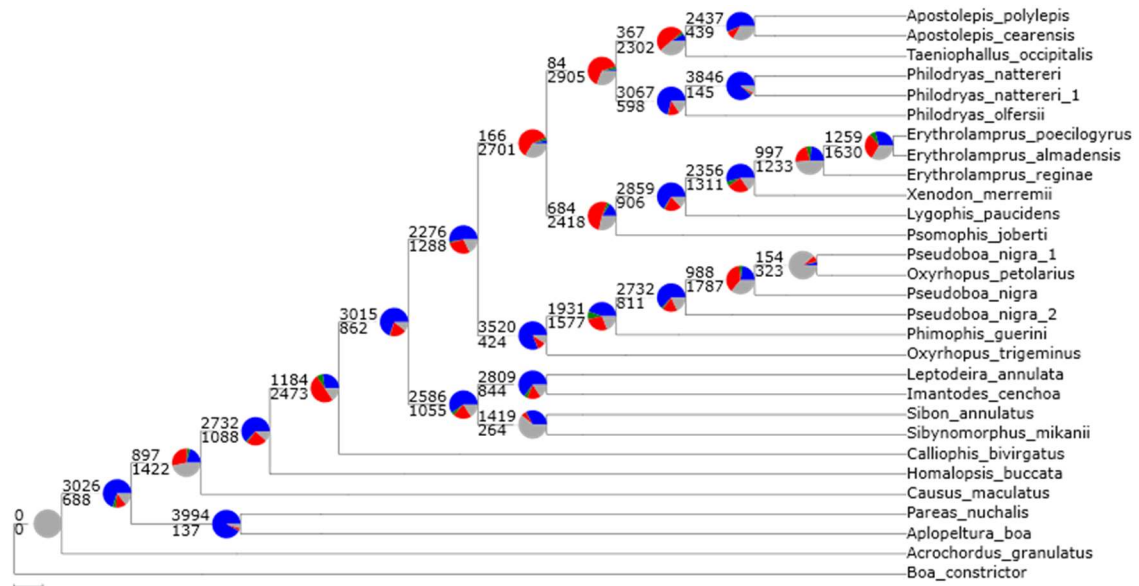


Figure 1.3. Species tree topology of the Dipsadidae obtained with ASTRAL-III indicating the levels of incongruence with the inferred gene trees for all nodes. Pie chart colors represent the proportion of gene trees supporting how a given node is resolved in the species tree (shown in blue), the most frequently found conflicting resolution (shown in green), all other conflicting possible reconstructions (shown in red), and non-informative gene trees (shown in grey). For this analysis, branches were collapsed on all gene trees with ultrafast bootstrap (UFboot) values of less than 50. Numbers above or below branches indicate the number of gene trees concordant or conflicting with the species tree.

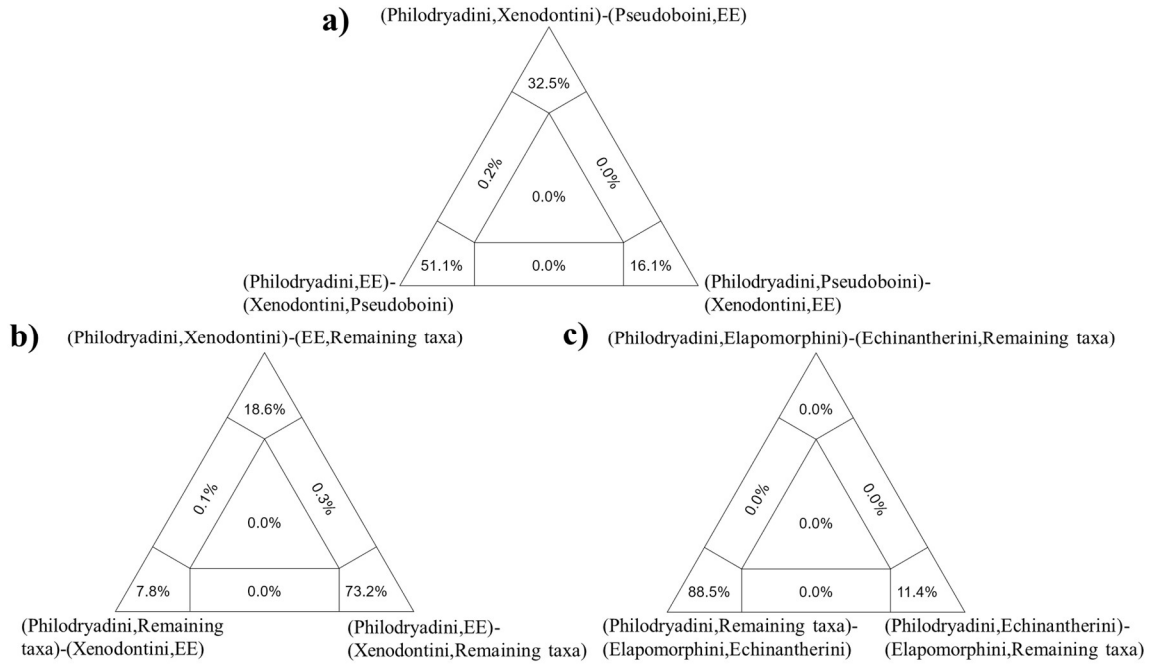


Figure 1.4. Results of the Likelihood Mapping Analyses we performed on the concatenated dataset used for this study (with 5462 loci) for each of the nodes of interest recovered in the ASTRAL species trees. These nodes correspond to a) the Most Recent Common Ancestor (MRCA) of Pseudoboini, Xenodontini, Philodryadini, and the Echinantherini+Elapomorphini (EE); b) the MRCA of the Xenodontini, Philodryadini, and the Echinantherini+Elapomorphini; and c) the MRCA of Philodryadini and the Echinantherini+Elapomorphini. The numbers in the corners of each triangle represent the percentages of quartets supporting every possible reconstruction of a given node, with the number in the central area corresponding to the percentage of uninformative quartets.

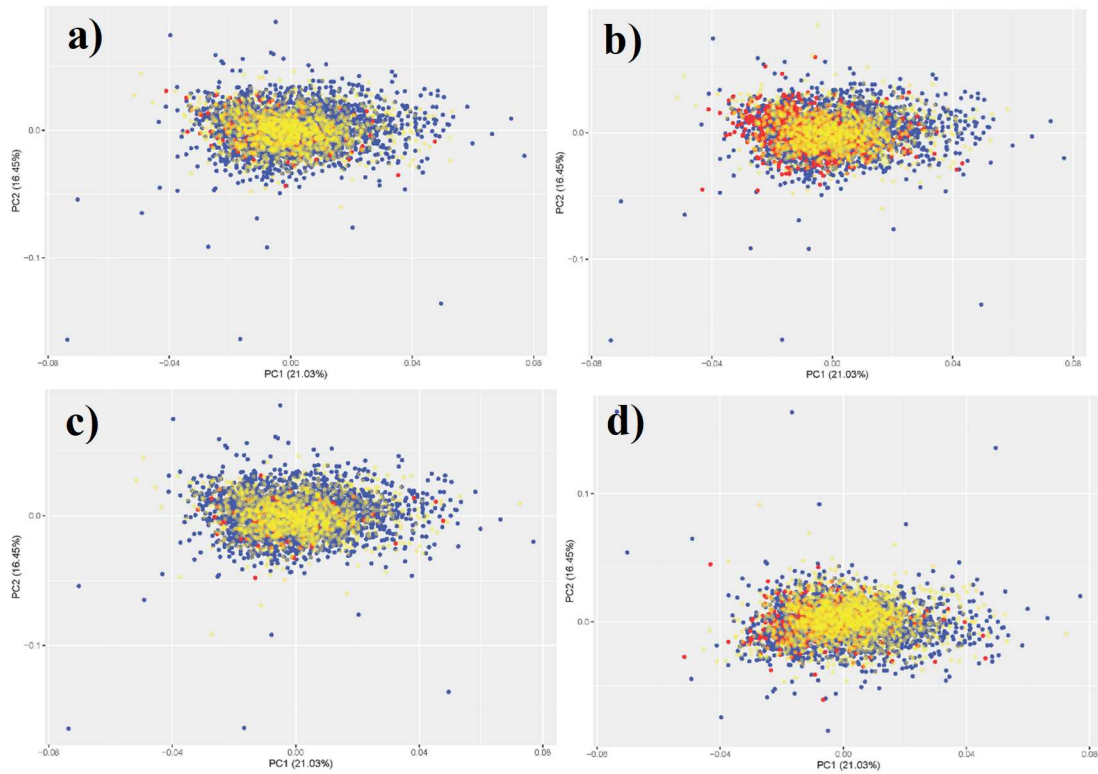


Figure 1.5. Results of the Principal Component Analyses (PCA) summarizing the values of the 13 traits for the alignments or the gene trees corresponding to the 4333 loci analyzed for this study. Data points correspond to each locus and are colored (red, blue, or yellow, respectively) according to whether their corresponding gene trees support, conflict, or are uninformative regarding each of the relationships of interest inferred in the species tree. These relationships are as follows: a) All Xenodontinae with the exclusion of the Pseudoboini, b) Xenodontini+Psomophini, c) Philodryadini+(Echinantherini+Elapomorphini), and d) Echinantherini+Elapomorphini. For simplicity, the arrows corresponding to the loadings of each characteristic are not shown. For this analysis, branches were collapsed on all gene trees with ultrafast bootstrap (UFboot) values of less than 50. All values were log-transformed to facilitate visualization.

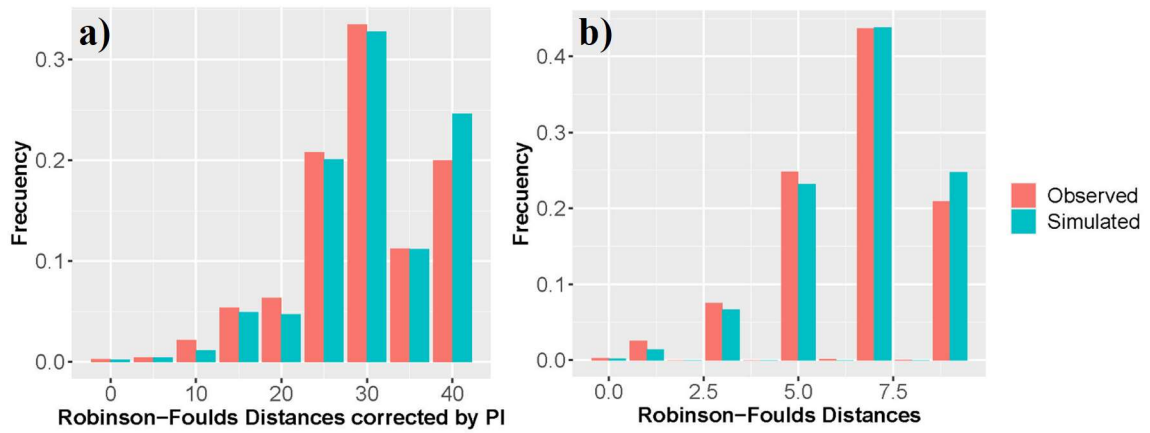


Figure 1.6. Distribution of a) Robinson-Foulds (RF) distance values corrected by loci phylogenetic information content, or b) raw RF distances between the observed or simulated gene trees and the ASTRAL species tree we obtained with the 4333 loci from the Squamate Conserved Loci set with 75% or more taxonomic completeness. Simulated values of these two metrics were obtained after performing 1000 replications of a scenario in which gene tree discordance is solely the result of Incomplete Lineage Sorting (ISL) and where the inferred species tree correctly represents the relationships between the clades of interest.

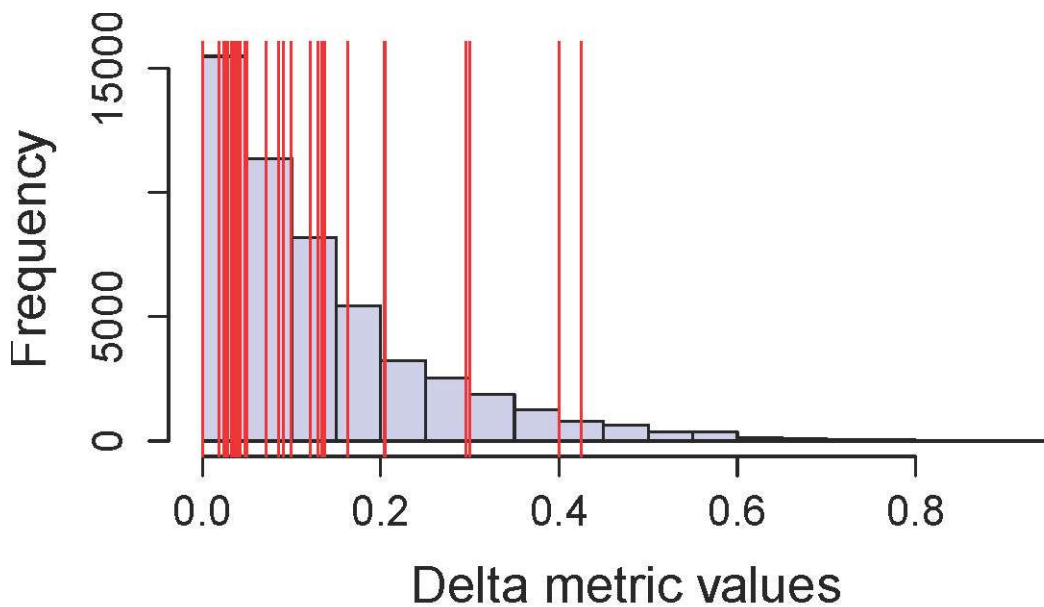


Figure 1.7. Comparison between the observed and expected values of the Huson et al. (2005) Δ metric of deep introgression for the internal branches of the coalescent-based species tree obtained with ASTRAL-III. Simulated values of the Δ metric are represented as pale purple bars and were obtained after performing 2000 re-samplings of the original collection of gene trees. Observed Δ metric values are shown with red vertical lines. Values from the branches for which less than 5% of gene trees reconstructed differently than in the species tree were not graphed.

CHAPTER 2

Regional speciation drives spatial variation in species richness in the most diverse family of Neotropical snakes (Colubroidea: Dipsadidae)

Abstract

Species richness gradients are frequently associated with spatial variation of climatic conditions. However, understanding how regional climate influences richness gradients is an ongoing discussion topic, with three non-exclusive explanations being proposed. Under these hypotheses, climate either has a 1) direct influence on the richness of regions by determining their carrying capacity; or an indirect role through its influence on either, 2) lineage diversification (i.e., rates of speciation and/or extinction), or 3) the time in which regions have been accumulating species. Unfortunately, few studies have simultaneously evaluated and compared the role of these mechanisms in determining richness gradients. Thus, studies accounting for this limitation are needed. Here, we explored if spatial variation in the species richness of the hyper-diverse snake family Dipsadidae is determined by the direct effect of current climate seasonality and past climate stability or through their influence on regional speciation or species accumulation times. To do so, we implemented piecewise structural equation models using recently published distribution maps and a newly inferred time-calibrated Bayesian phylogeny. Our results indicate that variation in species richness in the Dipsadidae through space is

strongly and directly predicted by regional speciation rates, with climate and species accumulation times having a limited influence. Such a result contrasts with the tendency of past studies to find that regional climate strongly predicts species gradients, albeit indirectly through its impact on species accumulation times. In contrast, previous support for climate-driven diversification or speciation in determining richness patterns is mixed. Further research is needed to determine if this lack of consistency is because of intrinsic differences between clades or from past studies only testing the influence of regional climate through proxies, among other methodological differences. Doing so would allow us to identify potential generalizations about the role of regional climate in determining richness gradients.

Introduction

Understanding how observed variation of species richness through space originated and has been maintained through time has been the focus of much interest for at least the last two centuries (von Humboldt, 1808; Hawkins, 2001). As a result of this interest, plenty of different hypotheses (at least 120; Palmer, 1994) having the goal of explaining the origin of species richness gradients have been proposed, each one invoking the role of a different ecological and/or evolutionary process (Palmer, 1994). We now recognize that many of these hypotheses are not independent of each other, or they cannot explain species richness by themselves (Rohde, 1992; Palmer, 1994). Despite this, there is still an ongoing debate about which processes have a more important role in determining the existence of species richness gradients (Marin et al., 2018). Traditionally, spatial variation in species richness has been considered as resulting from the influence of climatic variables (e.g., temperature, precipitation, productivity, and their short- and long-term variation) on the geographic range limits of individual species, on the niche space or the carrying capacity of regions, or from both (Roy & Goldberg, 2007; Stevens, 2011; Storch & Okie, 2019). In support of this idea, many studies have found strong, significant relationships between values of different environmental variables and species richness (Ricklefs, 2006; Wiens, 2007; Tello & Stevens, 2010). However, both the direction and strength of the associations between these variables exhibit considerable variation across regions and between taxa (Roy & Goldberg, 2007). Thus, this suggests that even patterns shared among many lineages (such as the latitudinal and altitudinal species richness

gradients) do not result from a single mechanism (Tello & Stevens, 2010; Cerezer et al., 2022).

Further, finding evidence of an association between species richness gradients and the spatial variation in one or more climatic variables does not necessarily indicate that such a relationship is causal (Kozak & Wiens, 2012). Instead, it is now recognized that associations between species richness and climatic variables need to be mediated by the only processes that can directly change the number of species existing in a given region: speciation, extinction, and dispersal (Wiens & Donoghue, 2004; Wiens, 2011). As a result, understanding the mechanisms responsible for large-scale species richness patterns requires integrating these three evolutionary processes with the environmental variables correlating with species richness (Wiens, 2007). Despite this, it is not uncommon, even among recent studies, to consider that species gradients can result from the ability of climate to directly determine the carrying capacity of regions, with some finding evidence in support of this possibility (Harvey et al., 2020; Machac, 2020; Garcia-Andrade et al., 2021). However, other studies have not found this to be the case (e.g., Qian et al., 2015; Gamisch et al., 2019, see Cerezer et al., 2022). In general, apart from the possibility of the direct influence of regional climate, there are two other non-exclusive explanations for the existence of patterns of spatial variation in species richness: the *diversification rates* and *time* hypotheses (Wiens, 2011; Kozak & Wiens, 2012; Marin & Hedges, 2016; Kozak, 2017). Under these hypotheses, regional climatic conditions either do not influence species richness gradients or do so indirectly. Indeed, under the diversification rate hypothesis, regions differ in factors (such as their area or

environmental conditions) that influence the net diversification (=speciation-extinction) rates they exhibit, which themselves impact the numbers of species they harbor (Kozak & Wiens, 2012; Qian et al., 2015). Thus, regions having conditions that promote lineage diversification are expected to contain increased levels of species richness.

On the other hand, under the time hypothesis, the increased species richness observed in certain regions results from them having characteristics (again, as the area or the climatic conditions they exhibit) that promote species accumulation for longer timespans (Kozak & Wiens, 2012; Qian et al., 2015; Cai et al., 2020). According to this hypothesis, regions colonized earlier or containing older taxa than elsewhere are expected to have accumulated more species, as they had more time for speciation events to occur (Kozak, 2017; Cai et al., 2020). As a result, this explanation of species richness patterns is also frequently known as the *time-for-speciation* hypothesis. However, limited dispersal between different regions can also cause variation in the species accumulation times they exhibit; and, thus, in their species richness (Mittelbach et al., 2007; Kozak, 2017). That could be the case if taxa have a strong tendency to conserve their ancestral climatic niches that limits their ability to disperse to regions outside their center of origin (i.e., if they exhibit niche conservatism; Wiens & Donoghue, 2004; Wiens et al., 2010). In the context of the latitudinal gradient of species richness, this possibility is known as the *tropical conservatism hypothesis* (Wiens & Donoghue, 2004). According to this hypothesis, the increased richness exhibited by many tropical lineages is expected to result from them originating in that part of the world and having a limited ability to disperse to temperate regions due to niche conservatism (Wiens & Donoghue, 2004).

Many studies have evaluated if more species-rich regions also have increased diversification rates or species accumulation times (i.e., tested the diversification rate and time hypotheses). However, it is not uncommon for these studies to not explore the influence of climate on regional diversification rates or species accumulation times altogether or to assume that such an effect occurs without explicitly testing if that is the case (Kozak & Wiens, 2012). Further, even for studies investigating how climate could determine richness gradients through its influence on regional diversification or species accumulation times, many use proxies of this variable instead, such as latitude or elevation. Although this is probably not an unrealistic assumption to make, it does not consider that other variables different from regional climate could also be correlated with latitude and be the ones driving species richness gradients (Gillman & Wright, 2013). For example, greater values of diversification rates in taxa inhabiting certain regions could occur due to the evolution of key innovations that allow them to exploit parts of the niche space better than other lineages (Heard & Hauser, 1995; Laudanno et al., 2021). If the clades with such key innovations tend to be distributed in specific regions, these regions would be expected to have higher diversification rates amongst their members than elsewhere. In addition, both the emergence and disappearance of barriers to dispersal could also promote the concentration of increased diversification or dispersal rates and species richness in the same regions, not necessarily through the influence of climate. In the first place, the disappearance of barriers to dispersal favors the immigration of taxa to new environments, resulting in range expansions that could also promote regional diversification (Heard & Hauser, 1995; Laudanno et al., 2021). The emergence of

dispersal barriers, as in the rise of mountain ranges, could also promote vicariant speciation or limit the dispersal of lineages (Sanmartin, 2012). As a result, one could expect that species richness can accumulate in regions containing such barriers, especially among taxa with limited dispersal abilities, without necessarily being the result of the influence of regional climate.

More generally, most studies exploring the causes of richness gradients do not simultaneously evaluate the roles of climatic conditions, regional diversification, or species accumulation, even when it is increasingly recognized that these factors act in combination (Qian et al., 2014; Garcia-Andrade et al., 2021). Instead, many of these works focus on exploring whether regional diversification can explain richness gradients (independently if it is driven or not by climate) (Li & Wiens, 2019). In contrast, few investigate the influence of species accumulation times or the potential direct role of regional environmental conditions (Qian et al., 2014; Garcia-Andrade et al., 2021; Li & Wiens, 2019). The reduced number of studies taking these considerations into account has limited our ability to reach more general conclusions about how frequently climate really determines regional richness through its influence on diversification rates or species accumulation times. As a result, our knowledge also remains limited about the factors potentially determining when this occurs, such as the different events shaping the evolutionary history of lineages and their dispersal abilities.

Following this line of thought, we simultaneously investigated the potential mechanisms explaining the influence of both past and present climate conditions on the observed patterns of species richness of the snake family Dipsadidae. With their >800

species and 97 genera inhabiting a wide variety of environments from Canada to Argentina, and with two small genera in Eastern Asia (Uetz et al., 2023), snakes of this family are a good model for answering questions in spatial phylogenetics. For this work, we evaluated if spatial variation in species richness across the distribution of the Dipsadidae results from the direct impact of regional climate or through its influence on either species accumulation or origination, in agreement with the time and diversification hypotheses, respectively. To do so, we implemented a piecewise structural equation model (psem) approach using the distribution maps of members of the Dipsadidae from Roll et al. (2017) and the most complete to date Bayesian time-calibrating phylogeny of the group, which we inferred herein. We incorporated into this phylogeny taxa without genetic data using a stochastic polytomy resolving software. Using a psem framework, we evaluated whether current seasonality and past stability of annual mean temperature and precipitation directly determined the regional (=grid cells of 50 km x 50 km) species richness of the Dipsadidae. Simultaneously, we investigated the influence of these climatic variables on the time in which cells have been accumulating species; or the averaged speciation rates amongst the species inhabiting them (as measured with appropriate phylogenetic metrics: the MPD and DR metrics, respectively). Further, we also explored the direct impact that these rates and accumulation times could have on cell richness.

Materials and Methods

Estimation of species richness values per cell

To explore how the species richness of the snakes of the family Dipsadidae varies across the Americas, we calculated the number of dipsadid species inhabiting each cell of a 50 km x 50 km grid. To do so, we overlaid this grid into the distribution maps of all species of Dipsadidae included in the Roll et al. (2017) dataset after reprojecting them to the Equal Earth map projection of Šavrič et al. (2019). We then removed cells having less than 25% land coverage and identified the number of species whose distribution overlaps with each cell. Calculating the per-cell species richness was done with the package `letsR` 4.0 (Vilela & Villalobos, 2015) of the statistical software R 4.2.2 (R Development Core Team, 2020). The range maps for the dipsadid species analyzed herein correspond to 61.99% of the total species richness of Dipsadidae (up to January 2023, Uetz et al., 2023). The resulting presence-absence matrix contains a total of 519 species across 10658 cells.

Estimation of ecological and evolutionary drivers of regional richness

To explore how past and current weather influenced cell species richness, we first downloaded the variables from the WorldClim 2.0 repository (Fick & Hijksmans, 2017) corresponding to each of the hypotheses we tested. These hypotheses consider the influence of the current climate (=temperature and precipitation) on regional species richness and how it has changed since either the Last-Interglacial (IGM) or the Last Maximum (LGM) periods. Further information about the hypotheses we tested and the

climatic variables they include is included in Table 2.1. We then aggregated (and reprojected) the values of the variables corresponding to each hypothesis to the same resolution as that of the grid cells used to calculate regional species richness.

Phylogeny of the Dipsadidae

To calculate the phylogenetic metrics employed in this study, we first inferred a Bayesian time-calibrated phylogenetic tree of the Dipsadidae. To do so, we compiled all sequences corresponding to five mitochondrial (12S, 16S, *cyt b*, ND2, and ND4) and four nuclear (BDNF, *c-mos*, NT3, and RAG2) genes publicly deposited in GenBank (up until March 7, 2021). We also incorporated sequences available in Zaher et al. (2018), not uploaded into GenBank. We also included sequences from 25 outgroup species from 14 different families of Alethinophidia. Sequence curation involved the identification and removal of sequences not correctly assigned to species (i.e., likely cases of contamination, mislabeling of sequences, or misidentifications of specimens), as well as the deletion of ambiguously aligned positions of the 12S and 16S ribosomal genes. Furthermore, sequences of terminals recovered in phylogenetically unexpected places (or associated with very long branch lengths) were compared with GenBank sequences by using the Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990), and we filtered those erroneously assigned to species. After deleting sequences recovered as being incorrectly identified, we concatenated all remaining sequences with the Sequence Matrix software v1.7.6 (Vaidya et al., 2011). Detailed information on all the samples used for this work is available in Suppl. Table S2.1.

We then aligned sequences with the online version of MAFFT v7 (Katoh & Standley, 2013) under default parameters, except for the case of the 12S and 16S ribosomal genes, for which we used the Q-INS-i algorithm (which considers the secondary structure of RNA sequences, Katoh & Toh, 2008). We posteriorly inspected alignments with the software Mesquite v3.2 (Maddison & Maddison, 2017). The final aligned matrix consists of 6038 nucleotide positions and a total of 437 species, of which 414 are dipsadids. These species represent 49.70% of the total richness of the group (Uetz et al., 2023). The software PartitionFinder 2 (Lanfear et al., 2016) was used to select the best-fitting partition scheme for the alignment according to the corrected Akaike Information Criterion (AICc) and the greedy algorithm of Lanfear et al. (2012). Partitions were defined a priori by gene identity, as well as by codon position for protein-coding genes (all included genes except 12S and 16S). The best partitioning scheme for each of the subsets is shown in Appendix Table A2.1.

We performed Bayesian phylogenetic analysis (BA) with BEAST 2 v4.8 (Bouckaert et al., 2014). The bModelTest package of BEAST 2 was applied to simultaneously explore the substitution model space and estimate both the corresponding model parameters and the phylogeny (Bouckaert & Drummond, 2017). We obtained the posterior distribution of trees, including dates and rates, by assuming a pure-birth or Yule model tree prior (a diversification model that considers extinction rates as negligible in comparison to speciation rates; Yule, 1924; Ritchie et al., 2017). A relaxed normal clock model of evolution was used and calibrated with five fossil calibration points (detailed in Appendix Table A2.2). As a starting tree for this analysis, we used the phylogeny with

the highest Maximum Likelihood (ML) value we obtained after running on the concatenated dataset the software RAXML-HPC V.8.2.12 (Stamatakis, 2014), as implemented in the CIPRES Science Gateway v. 3.3 (Miller et al., 2010). To obtain this phylogeny, we performed 20 alternative ML runs, each on a different starting tree and using the GTRCAT model of sequence evolution for all partitions. Node support was established with 1000 bootstrap replicates. The resulting phylogeny was then converted to an ultrametric tree using the previously mentioned calibration points and the nonparametric rate smoothing method implemented in the function 'chronos' of the R package ape 5.5 (Paradis & Schliep, 2019). The sampling of the Beast 2 analysis used two independent MCMC chains of 50 million generations each, saving one sampled tree per 1000 generations and with the first 50% of trees discarded as burn-in. Stationarity and mixing of the log-likelihood values and parameter estimates were evaluated with Tracer v1.6 (Rambaut et al., 2018) and corroborated by verifying that all parameters of the post-burn-in set of trees had effective sample sizes > 200. The posterior distribution of trees was summarized as a Maximum Clade Credibility (MCC) timetree using TreeAnnotator v2.4.8, part of the BEAST 2 package. The resulting calibrated phylogeny is shown in Appendix Figure A2.1.

We incorporated the species not included in this phylogeny with the software TACT (Taxonomic Addition for Complete Trees, Chang et al., 2020). TACT is a stochastic polytomy resolution algorithm that first evaluates if user-defined taxonomic groups are recovered as monophyletic in a given phylogenetic tree. Then, it sorts taxa without phylogenetic data to the least-inclusive monophyletic group they are thought to

belong to and assigns them to a random position within that clade (Chang et al., 2020). TACT also estimates the diversification rates of these monophyletic groups, which are used to generate waiting times for the unsampled species (Chang et al., 2020). To account for the stochastic variation in the position of unsampled taxa, we generated 100 completely sampled phylogenies of the Dipsadidae with TACT. To do so, we followed the taxonomy of Uetz et al. (2023) and assigned taxa to the taxonomic categories of subfamily, tribe, and genera. However, we changed the taxonomic assignment of a few taxa recovered in unexpected positions (i.e., conflicting with all previous phylogenies) in the original calibrated Bayesian phylogeny we had inferred (see Suppl. Table S2.2 for further details). In the end, distribution and phylogenetic data were available for a total of 517 species (61.99%).

Calculation of per-cell estimates of times of species accumulation and speciation rates

To estimate the per-species speciation rate of the taxa included in this study, we used the DR metric of Jetz et al. (2012), also known as the DivRate (e.g., Oliveira et al., 2016) or the λ_{DR} metric (e.g., Title et al., 2019). To do so, we used the R function DivRate from Oliveira et al. (2018) and then averaged the resulting per-species DR metric values across all the species inhabiting each cell of the study area. We performed this procedure using each of the 100 phylogenies we had obtained with TACT and then averaged the resulting per-cell values. Although originally considered to measure per-species net diversification rates, it has been shown that the DR metric performs better as a

speciation rate estimator (e.g., Quintero & Jetz, 2018; Title et al., 2019). In any case, studies have shown that DR predicts simulated values of either per-species diversification or speciation rates more accurately than other phylogenetic metrics (e.g., Oliveira et al., 2016; Title et al., 2019). However, the tip-diversification or speciation rates that can be obtained using the model-based approach implemented in software such as BAMM (Rabosky, 2014) appear to be even more accurate than the values of the DR metric (e.g., Title et al., 2019). Despite this, we decided to use this metric herein as its performance is similar to that of BAMM, especially when extinction rates are low (Title et al., 2019). Further, in general, values of extinction rates obtained from time-calibrated phylogenies are less reliable than speciation rates, thus, limiting our ability to estimate accurate diversification rates (Rabosky, 2016; Louca & Pennell, 2020). In any case, we calculated per-species speciation, extinction, and net diversification rates using the software BAMM (v. 2.5.0) and either the original phylogeny we inferred or one of the trees we had obtained using TACT selected randomly. We then used these rates to repeat all analyses we performed with the DR metric, as described below. Further details about the methods we used for the BAMM runs and the results we obtained are presented in Appendix 2.1.

On the other hand, to obtain an approximation of the time in which members of the Dipsadidae started to inhabit each of the geographic cells covering the study area, we calculated the mean pairwise phylogenetic distance (MPD; Webb, 2000) across all the species present on each cell. We calculated the MPD values of all cells across the entire study area using each of the 100 completely sampled phylogenies we generated and then averaged the resulting values. We chose the MPD metric to represent species

accumulation time following the results of Oliveira et al. (2016), which found that it was the most adequate for measuring clade age. MPD measures how different the members of a local or regional assembly are, with high values of MPD corresponding to longer time divergences between taxa (Oliveira et al., 2016). Per cell-MPD values were calculated with the function `mpd` of the R package `picante` 1.8.2 (Kembel et al., 2010).

Exploring the potential mechanisms of regional richness variation

To investigate the potential direct and indirect influences that past and present climatic conditions, regional speciation, and times of species accumulation have on cell species richness, we used piecewise structural equation modeling (pSEM). For the pSEM model we constructed, we treated species accumulation times, speciation rates, and current climate (=temperature and precipitation) seasonality as endogenous variables (i.e., those whose values could be determined by the other variables of the model). We also considered how regional climate (both temperature and precipitation) changed since the LGM and the IGM as exogenous variables (i.e., those being known not to result from the other variables in the model). Further, as in Garcia-Andrade et al. (2021), we considered that current climatic conditions could not have predicted regional species accumulation times due to them being at different temporal scales, so we did not consider this possibility in the pSEM model. To prevent collinearity, for each equation of the model, we also disregarded variables found to be strongly correlated (=absolute correlation coefficient greater than 0.8) to other variables from the same pathway. To perform these correlations, we used the modified version of the t-test proposed by Dutilleul (1993), as

implemented in the command `modified.ttest` of the R package `SpatialPack` 0.4 (Vallejos et al., 2020). The modified test accounts for the spatial autocorrelation between two spatial processes by adjusting the number of effective degrees of freedom (Dutilleul, 1993). Finally, we normalized the values of all variables (i.e., changed them to a mean of zero and a standard deviation of one) to make the coefficients of the equations of the final pSEM model comparable. The correlations between all the variables considered for inclusion in the theoretical pSEM model we tested are shown in the Appendix Table A2.3. The equations of this model after removing variables found to be autocorrelated are presented in Appendix 2.

We then used simultaneous autoregressive models (SARs) to account for spatial autocorrelation between the variables on each of the equations of the pSEM model. SARs expand upon Ordinary Least Regressions (OLS) by incorporating an error term that accounts for spatial dependency with a spatial weight matrix. Such a matrix indicates the influence on the residuals corresponding to each site by those of its neighbors (Dormann et al., 2007; Skeels et al., 2020). For each equation of the pSEM we tested, we inferred their corresponding spatial weights matrix by employing three weighted schemes: row standardized (W), globally standardized (C), and variance-stabilizing (S), and by using a distance of 50 km between the centroids of grid cells. We also repeated the estimation of the spatial weight matrices using these three weighted schemes but using a distance of 269.26 km instead (=the maximum distance between centroids). Finally, we determined which of the six resulting spatial weights matrices (plus an OLS regression model) best accounts for the spatial autocorrelation of the variables on each equation of interest by

comparing the corresponding values of the Akaike Information Criteria (AIC) and Nagelkerke pseudo-R². We then used the best-fitting spatial weights matrices corresponding to each equation (as shown in Appendix Table A2.4) when implementing the pSEM approach. We used the function `nb2listw` of the R package `spdep` (v. 1.2.7; Bivand et al., 2013; Bivand & Wong, 2018) for calculating the spatial weight matrices, which we then incorporated into SARs with the function `errorsarm` of the `spatialreg` package (v. 1.2.6, Bivand et al., 2013).

Finally, to evaluate the goodness of fit of the piecewise structural equation model we had defined, we used the command `summary.psem` from the R package `piecewiseSEM` (v. 2.1.2., Lefcheck, 2016). Specifically, we verified that the p-value of the estimated Fisher's C global goodness of fit metric was greater than 0.05. Using that function also allowed us to obtain the values of d-separation tests, which indicate if there are additional variables remaining to be included in a given pathway of the original pSEM model (Shipley, 2000; Lefcheck, 2016). We modified the equations of the pSEM model according to the results of the d-separation tests until none of their results (nor the global Fisher's C value) were significant. All analyses described above were repeated after using the original phylogeny of the Dipsadidae we inferred without accounting for taxa without phylogenetic data.

Results

Estimation of values of species richness per cell

Hotspots of species richness in the Dipsadidae concentrate in central and southeastern South America. More specifically, the highest richness values (over 60 species per cell and up to a maximum of 81) were observed in the Parana dominion of Morrone (2014), followed by most of the Amazonian region (especially in its southern part), where cells have between 50 and 70 species. Many of the cells between the Cerrado and Rondônia provinces of Morrone (2014) also have richness values of 50-60 species. The remaining parts of the dipsadid distribution have less than 30 species, except for a few cells in northern Costa Rica and the southwestern Andes of Colombia, which have up to 40 species (Figure 2.1). Similar patterns of species richness were observed when considering the dipsadid distribution maps compiled by Azevedo et al. (2019) (Appendix Figure A2.2).

We also mapped the mean per-species speciation rates for every cell across the entire study area. Geographic patterns of per-cell speciation rates were very similar to those of species richness. The higher values of these metrics concentrate in the Parana dominion, and to a lesser degree, across the entire Amazonian region (Figure 2.2). Speciation rates decrease when getting afar from those regions, with the lower values concentrating on North America and the Pacific coast of South America, south of Ecuador (Figure 2.2). Similar results were obtained using the DR metric calculated with the original phylogeny we had inferred (Appendix Figure A2.3). Also, the results described above do not change when using the per-cell diversification and speciation

rates obtained with the software BAMM and a randomly selected phylogeny obtained with TACT (Appendix Figure A2.4). However, spatial patterns in per-cell extinction rates were slightly different, with cells with relatively low extinction rates scattered across the Parana and Amazonian regions (Appendix Figure A2.4).

MPD values for the cells across the entire study area are shown in Figure 2.2. Cells exhibited MPD values of 45-50 over the entire Neotropical region, with only the West Indies (except for most parts of Cuba) having MPD values lower than 40 (Figure 2.2). Most of the eastern part of the United States (and the northern part of the west coast) also has MPD values of 45-50, but there is one patch each in the central and southwestern part of the country with MPDs of 35-45 (Figure 2.2). In contrast, the majority of the cells in the southeastern and central parts of the country and of northern and Central Mexico have high MPD values (i.e., above 50, up to a maximum of 59.75). The remaining parts of Mexico exhibit MPDs of 35-45, with the Baja California peninsula having values of less than 30 (Figure 2.2). Finally, MPD values could not be calculated for extensive areas in the central United States and Mexico, as they only contain one species of the Dipsadidae or none at all.

Of all the potential predictors of regional species richness of the Dipsadidae we considered, the only strong and significant relationship we recovered was between temperature anomaly since the LGM and both temperature seasonality and the temperature anomaly since the IGM (Appendix Table A2.3). As a result, we did not consider this variable in any of the equations of the original theoretical pSEM model we tested, which is shown in Figure 2.3. After evaluating this model and according to the

results of the d-separation tests we performed, we recovered that some of its pathways needed to be modified (Appendix Table A2.5). The resulting pSEM model obtained after performing these modifications is shown in Figure 2.4 and Appendix 2.2. According to this corrected and final model (Fisher's C = 9.798, degrees of freedom = 12, p-value = 0.634), the following variables had the highest absolute values of the direct standardized effect sizes: regional speciation rates (0.9), current temperature seasonality (0.0384), and temperature anomalies since the IGM (0.0267). The absolute value of the direct effect of the other explanatory variables is smaller than 0.02, even being non-significant for the case of precipitation seasonality and the precipitation anomalies since the IGM (Figure 2.4, Appendix Table A2.6). In general, the final model we evaluated was able to account for spatial autocorrelation between each of the variables in the final pSEM model we tested (Appendix Figure A2.5).

The confirmatory analyses we performed revealed that our results did not differ whether we used the DR metric or the diversification or speciation rates obtained with BAMM version 2.5.0 (results not shown). We also found that all these metrics were highly correlated, independently of whether they were obtained using the original phylogeny we inferred or the 100 trees generated with TACT (Appendix Figure A2.6 and A2.7, respectively).

Discussion

In this study, we used a piecewise structural equation model (psem) approach to evaluate the role of evolutionary and ecological factors in determining the observed

spatial variation of species richness in the Dipsadidae. In summary, we found that current climate seasonality and past climate stability practically do not have an impact on the regional richness of the Dipsadidae, neither directly nor through the influence of these variables on regional speciation rates or species accumulation times. Instead, the richness gradient in this group is the direct result of spatial variation in speciation rates, in agreement with the diversification rate hypothesis. Although strictly speaking, such a result does not necessarily indicate that regional diversification drives species richness gradients in the Dipsadidae, our confirmatory analyses performed using diversification rates obtained using BAMM support this to be the case. Unfortunately, our ability to compare our results to previous studies is limited, as only a few of them explicitly test whether climatic conditions can directly (i.e., by themselves determining the carrying capacities of regions) drive how species richness varies through space (=Qian et al., 2014; Gamisch et al., 2019; Machac, 2020; Harvey et al., 2020; Garcia-Andrade et al., 2021; Cerezer et al., 2022). Instead, as described below, most studies either evaluate if diversification rates and species accumulation times impact regional species richness or if these variables are themselves influenced by climate. Further, there does not appear to be a general trend among the studies that do test the potential role of regional environmental conditions in directly determining richness gradients. Indeed, while some studies recover evidence indicating this to be the case (e.g., Machac, 2020; Garcia-Andrade et al., 2021), others do not (e.g., Qian et al., 2014, and many of the taxa analyzed in Cerezer et al., 2022). Despite this lack of concordance between studies, it has been suggested that support for climate being a direct driver of regional richness may be more prevalent

among taxa with more extensive distribution ranges or with increased dispersion abilities, as in marine mammals (Cerezer et al., 2022). As shown by our results, the Dipsadidae is not an exception to this apparent tendency, considering that many species of this group have reduced ranges (i.e., a median range of 33486.99 Km² according to Roll et al., 2017's distribution maps, or about 0.13% of the total area covered herein); and probably, limited vagilities too. However, the evidence in other groups is not conclusive, as a lack of support for the direct role of climate on regional richness is found both among lineages with high (as angiosperms or aquatic birds; Qian et al., 2014; Cerezer et al., 2022) and low (as salamanders and lagomorphs; Cerezer et al., 2022) dispersal abilities. In contrast, all taxa in agreement with this expectation do have increased dispersal abilities (such as neotropical birds, freshwater fish, or marine mammals; Machac, 2020; Garcia-Andrade et al., 2021; Cerezer et al., 2022).

However, in general, whether regional climate could really directly generate species richness gradients is controversial. Indeed, some authors argue that this possibility cannot be an alternative to the diversification rate and time hypotheses (e.g., Wiens, 2011; Qian et al., 2014; Pontarp & Wiens, 2017; Wu & Wiens, 2022). Instead, simulations and theoretical considerations support the idea that even the carrying capacities of regions can only result from the processes of speciation, dispersal, and extinction (e.g., Wiens, 2011; Kozak & Wiens, 2012; Pontarp & Wiens, 2017). In contrast, other authors consider that the lack of support for both the diversification and time hypotheses found in multiple studies (e.g., Belmaker & Jetz, 2015; Oliveira et al., 2016; and the works cited above) supports the idea that richness gradients could result

from differences in carrying capacities between regions (e.g., Storch & Okie, 2019; Machac, 2020; Cerezer et al., 2022). However, it should be considered that there is another possible explanation for this lack of support for the diversification rates and the time hypotheses. Indeed, richness gradients could also originate from elevated dispersal causing the rapid assembly of species into regions with different climatic conditions (Qian et al., 2014). In other words, if the species within a given clade disperse widely and quickly enough between different regions by tracking their climatic preferences, it can result in the homogenization of their diversification rates and species accumulation times (Velasco et al., 2018). Thus, one would expect that a lack of support for both the diversification rates and time hypotheses would be expected to be more common among lineages with high dispersal abilities. As mentioned above, this prediction appears to be supported by the empirical evidence available. However, whether such a result is caused by climatic conditions determining the carrying capacities of regions or by the rapid and extensive dispersion of the lineages involved remains unknown. Further studies are needed to differentiate between these possibilities, although our ability to perform them is limited by the increasing difficulty of reconstructing dispersal events when more regions are considered (Sanmartín, 2012; Matzke, 2013).

Finding that speciation rates drive the regional species richness in the Dipsadidae is not surprising if it is recognized that, after accounting for the effects of spatial autocorrelation, only this variable had a strong and significant correlation with cell species richness. Further, this can also be observed in that hotspots of species richness tend to be the same as those having increased speciation rates. These areas correspond to

the Parana and Atlantic forests in Eastern Brazil, most of the Amazon region, and to a lesser degree, northern Costa Rica and the southwestern Andes of Colombia.

Interestingly, most of these areas also contain multiple species of the group we recovered as being subjected to a diversification rate shift (Appendix Figure A2.8 and Appendix 2.1 but see Appendix Figure A2.9 for a contradictory result). Such a group corresponds to the genus *Atractus* (except for a few species recovered within *Geophis*, most likely as a mistake; see Suppl. Table S2.2). However, there are two exceptions to this tendency. Firstly, despite being inhabited by numerous species of *Atractus* (e.g., Passos, 2009; Arteaga et al., 2017), we did not recover the Pacific region of Morrone (2014) as either a hotspot of species richness or diversification. That is probably the case because this region contains very few species of the Xenodontinae and many from the Dipsadinae (see Appendix Figure A2.10). In contrast, the species composition of the areas we recovered as hotspots of species richness in the entire family is not particularly biased toward any of its subfamilies. The only exception to this pattern is northern Costa Rica, which is particularly unusual because despite being recovered as a minor hotspot of species richness, it only has a few species of the Xenodontinae and lacks any species of *Atractus* altogether (Myers, 2003; Köhler, 2008). It is unclear why this is the case. However, it is interesting to note that the remaining parts of southern Central America have slightly lower species richness values than northern Costa Rica, in both cases higher than those from northern Central America. Thus, it is possible that with increased taxonomic, phylogenetic, or geographic sampling, the entirety of southern Central

America will become a richness hotspot due to it having more species of the Xenodontinae (a largely South American clade) than the rest of the subcontinent.

In any case, it is apparent from our results that the increase in diversification rates occurring within *Atractus* (see Appendix 2.1) had a considerable impact on determining the spatial variation in the species richness of the entire Dipsadidae. Such a result is not unexpected considering that this genus is the most species-rich of the family, being distributed over most of South America and containing many micro-endemic species (Pomar-Gómez et al., 2021; Uetz et al., 2023). Determining what caused the shift in diversification rates observed in *Atractus* is out of the scope of this paper. However, we consider that such an increase in diversification rates likely resulted from the rapid formation of barriers of dispersal in northern and central South America resulting from the continued rising of the Andes during the late Miocene and Pliocene or the change in water courses occurring as a result. Such is the case, as these events occurred in a similar timeframe as when *Atractus* originated and diversified (i.e., from ~14 million years to the present time, Ramirez et al. submitted). Indeed, although the uplift of the northern Andes started during the late Cretaceous, it had already reached half of its current elevation at ~19-13 million years ago, with the other half being attained ~4.5 million years ago (Rodriguez-Muñoz et al., 2022; Cassemiro et al., 2023). These changes resulted in the termination of the Pebas mega wetland system and the formation of most river systems in Northern South America, including the Amazon River (Hoorn et al., 2010; Hoorn et al., 2017; Cassemiro et al., 2023). In combination with the fossorial habits of members of *Atractus* (e.g., Myers, 2003), these geological events probably caused a burst of

speciation events (either through vicariance or dispersal) not observed in other dipsadid snakes with different habitat use and increased dispersal abilities. However, further studies are needed to put these hypotheses to the test.

In general, in contrast with our results, regional diversification or speciation rates do not tend to be found to determine species richness patterns in other clades (e.g., Jansson et al., 2013; Marin & Hedges, 2016; Economo et al., 2018; Economo et al., 2019; Li & Wiens, 2019; Igea & Tanentzap, 2020; Tietje et al., 2022; but see Pyron, 2014; Rolland et al., 2014; or Shrestha et al., 2018 for contrary results). Further, most recent investigations evaluating the influence of regional climate on diversification rates do not find this to be the case either (e.g., Qian et al., 2015; Rabosky et al., 2018; Harvey et al., 2020, Morales-Barbero et al., 2020; Sun et al., 2020; Wu & Wiens, 2022). However, both of these results need to be interpreted with some caution, as most older studies using latitude as a proxy for regional climate find that this variable determines richness gradients by influencing regional diversification (see Currie et al., 2004 and Gillman & Wright, 2013). Of course, using latitude instead of regional climatic conditions has its own limitations, such as failing to recognize that the variables and processes shaping species richness patterns exhibit complex spatial patterns and instead assuming they occur on a single dimension (Hawkins & Diniz-Filho, 2004). Further, it ignores the potential influence of other variables that, similar to climate, also vary along with latitude (Gillman & Wright, 2013). However, despite these limitations, it is noteworthy how consistent appears to be the support for the diversification rate hypothesis among studies evaluating it using latitude as a proxy of regional climate. In conclusion, whether driven

by climate or not, the role of regional diversification in driving species richness gradients is better considered as having mixed support, and evidence favoring this possibility could be more widespread than we thought. Further studies and the reanalysis of past data are needed to determine if this is the case.

In addition, finding that species accumulation times have a limited influence on the regional richness of the Dipsadidae is most likely a consequence of the frequent dispersal events between Central and South America (and between smaller biogeographic areas) that have occurred through the group history (Ramirez et al., submitted). Indeed, several dipsadid species have extremely extensive distributions (such as *Clelia clelia*, *Oxyrhopus petola*, *Imantodes cenchoa*, *Sibon nebulatus*, or *Xenodon rabdocephalus*; Köhler, 2008; Guedes et al., 2018), which they most likely colonized over very short timespans (Ramirez et al., submitted). In contrast, there are also several dipsadid groups whose dispersal seems to have been limited, as their current distribution is relatively restricted. For example, the tribes Psomophini, Echinantherini, and Elapomorphini of the Xenodontinae, which are restricted to central and eastern South America, and the tribe Diaporolepini of the Dipsadinae, which is exclusively found in the Pacific lowlands of northern South America. Speciose genera such as *Rhadinaea*, *Rhadinella*, or the polyphyletic *Geophis* are also found in Central America almost exclusively (Köhler, 2008; Guedes et al., 2018). As a result, many of the cells used for this study contain a mix of species that either colonized them recently or have inhabited them for relatively long timespans. Such tendency can be observed in that most of the cells across the distribution of the Dipsadidae exhibit similar values of species accumulation times, with only the

cells at its periphery exhibiting either increased or decreased values. Interestingly, similar results were obtained in a recent study on the lizard genus *Anolis* (Velasco et al., 2018), which has a similar spatial pattern of species richness to the Dipsadidae. Such a study also used a similar methodology to the one implemented herein (i.e., testing the diversification rates and time hypothesis, also using phylogenetic metrics and grid cells). As in the Dipsadidae, species accumulation times do not vary much across the distribution of *Anolis*, with extreme values concentrating around its edges. Again, this finding seems to result from the frequent dispersal events occurring throughout the group's history (Poe et al., 2017; Velasco et al., 2018). Such a result further supports the idea that frequent dispersal events between regions explain why the time hypothesis is not supported in either the Dipsadidae or *Anolis*, despite both groups having suffered from similar biogeographical events (such as the elevation of the Andes and the closure of the Panamanian Isthmus). However, it is possible that when considering more extensive regions than the ones used herein and in Velasco et al. (2018) [e.g., at the scale of the Morrone (2014) bioregions of the Neotropics], one could start observing some influence of species accumulation times in determining regional richness. That possibility is more likely for the case of *Anolis*, in which species with ranges as extensive as those observed in some dipsadid species are not observed (compare the range maps of the species from both groups, as shown in Roll et al., 2017).

More generally, however, our results contrast with most other studies testing if species richness patterns result from differences in species accumulation times between regions (i.e., the *time* hypothesis). That is the case, as they tend to recover positive

associations between species accumulation times and richness patterns (e.g., Pyron & Burbrink, 2009; Marin & Hedges, 2016; Economo et al., 2018; García-Andrade et al., 2021; and the studies listed in Li & Wiens, 2019). Despite this, there does not appear to be a clear trend in whether regional climatic conditions are the factors driving these associations between species accumulation times and species richness. Indeed, although some investigations find this to be supported (e.g., Kozak and Wiens, 2012; Qian et al., 2015; Marin et al., 2018; Wu & Wiens, 2022), others do not (e.g., Velasco et al., 2018; Machac, 2020, and Cerezer et al., 2022). Thus, further studies are needed to confirm the potential existence of general patterns about the role of species accumulation times in driving richness gradients and the influence of regional climate.

It has also been reported that of the few studies that do not find relationships between regional species richness and species accumulation times (i.e., Pyron & Wiens, 2013; Hutter et al., 2017), they involve lineages extending over longer timespans and containing more species (Li & Wiens, 2019). For such clades, associations between species richness and diversification or speciation rates (not necessarily mediated by climate) are observed instead (Pontarp & Wiens, 2017; Schluter & Pennell, 2017; Li & Wiens, 2019). A likely explanation for these patterns is that differences in diversification rates between regions require several successive speciation events to occur, a process that can require millions of years to unfold (Li & Wiens, 2019). In contrast, variation in species accumulation times can produce richness gradients in much shorter timespans through the process of dispersal (Li & Wiens, 2019). Despite this, if dispersal between regions occurs frequently enough, it can erode the role that species accumulation times

could have in causing species richness patterns (e.g., Velasco et al., 2018). On the other hand, we consider that another potential explanation of why the diversification rate hypothesis tends to be supported more among more species-rich lineages is that by containing more species, diversification rate shifts are more likely to have occurred through the story of the group just by chance. Our results support these expectations, considering that the Dipsadidae corresponds to a relatively large and old clade. Indeed, with over 800 species and having inhabited the New World for at least ~30 million years ago (Zaher et al., 2019; Ramirez et al., submitted), we would expect the diversification rates hypothesis to be supported but not the time hypothesis; thus, in agreement with our results. However, in contrast to this apparent tendency, several investigations have found support for the time hypothesis among old and species-rich clades (e.g., Qian et al., 2015; Marin & Hedges, 2016; Economo et al., 2018; Rabosky et al., 2018; Harvey et al., 2020) or to the diversification rate hypothesis in recently evolved lineages (e.g., Shrestha et al., 2018; and a few of the studies cited in Li & Wiens, 2019). Also, it remains unknown if the generalizations described above about the influence of clade size and age hold when considering only the studies in which the effects of regional climatic conditions on diversification or speciation rates is considered.

In summary, our finding that spatial variation of species richness in the Dipsadidae results from regional diversification (but not from species accumulation times nor the direct or indirect influence of regional climate) contrasts with the results of studies on many other taxa. That is the case, as these works tend to find that richness gradients are determined by species accumulation times while support for the influence of

regional diversification rates is mixed. However, our findings agree with the apparent tendency among previous studies to find that richness patterns are more commonly determined by regional diversification in large and old lineages. Instead, support for the role and species accumulation times is more prevalent among smaller and younger clades. Despite this, further studies and reanalysis of past works (as Liu & Wiens, 2019 or Velasco & Pinto-Ledesma, 2022) are needed to confirm the validity of these apparent generalizations. That is the case, as the studies serving as the basis for proposing the existence of these patterns differ in multiple aspects that may impact or bias their results and our interpretation of them, such as whether they use climatic variables or proxies as latitude or elevation. Further, they also differ in the methods used to estimate diversification rates or species accumulation times (i.e., using phylogenetic metrics or model-based approaches such as BAMM or the software GeoSSE, Goldberg et al., 2011). Also, there is variation in the size and shape of the regions considered (i.e., whether studies use latitudinal bands or midpoints, irregularly shaped areas, or fixed-sized cells; Cardillo, 2002; Willig & Presley, 2018). The last factor is especially relevant, considering that the relative roles of different processes in determining richness patterns could depend on the spatial scale being considered (Cerezer et al., 2022).

Apart from accounting for the influence of these factors, future studies would also greatly benefit from the use of methods that can simultaneously test the roles of different ecological and evolutionary drivers in generating richness gradients (such as structural equation models and variation partitioning). Doing so would permit evaluating the controversial possibility that climate can generate richness gradients solely through its

influence on the carrying capacity of regions. However, in general, it can be asked why we should expect that generalizable patterns regarding the processes that shape regional richness exist and what implications would be if that is the case. In this regard, the existence of regional richness patterns shared across many taxonomic groups (such as the latitudinal and altitudinal gradients of species richness) suggests that at least some commonalities in the relative importance of the factors driving them should exist. In any case, confirming the validity of potential generalizations about the conditions under which the diversification rates and time hypothesis are supported would also allow us to evaluate more specific explanations of the origin of richness gradients. One example of this is the *evolutionary speed* hypothesis (under which higher temperatures are thought to influence regional diversification by increasing mutation rates or rates of selection through the reduction of generation times), which has the diversification rates hypothesis as an underlying prediction (Rohde, 1992; Gilman & Wright, 2013). In contrast, the time hypothesis is a logical deduction for the *tropical conservatism hypothesis* (according to which niche conservatism limits taxa from regions with less extreme environments to disperse elsewhere, thus resulting in the accumulation of species richness) (Wiens & Donoghue, 2004; Wiens et al., 2010). Further, confirming whether there are common patterns in the factors determining richness gradients across different taxa would also allow us to get better insights into how the outcomes of such processes could change as a result of the impacts of human activities on future climate.

References

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403-410.
- Arteaga, A., Mebert, K., Valencia, J. H., Cisneros-Heredia, D. F., Peñafiel, N., Reyes-Puig, C., ... & Guayasamin, J. M. (2017). Molecular phylogeny of *Atractus* (Serpentes, Dipsadidae), with emphasis on Ecuadorian species and the description of three new taxa. *ZooKeys*, (661), 91.
- Azevedo, J. A., Guedes, T. B., Nogueira, C. D. C., Passos, P., Sawaya, R. J., Prudente, A. L., ... & Giraudo, A. R. (2019). Museums and cradles of diversity are geographically coincident for narrowly distributed Neotropical snakes. *Ecography*, 43(2), 328-339.
- Belmaker, J., & Jetz, W. (2015). Relative roles of ecological and energetic constraints, diversification rates and region history on global species richness gradients. *Ecology Letters*, 18(6), 563-571.
- Bivand, R., Pebesma, E. J., & Gomez-Rubio, V. (2013). *Applied spatial data analysis with R* (2nd ed.). Springer, New York, New York, United States of America
- Bivand, R. S., & Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *Test*, 27, 716-748
- Bouckaert, R. R., & Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, 17, 1-11.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., ... & Drummond, A. J. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10(4), e1003537.
- Cai, T., Shao, S., Kennedy, J. D., Alström, P., Moyle, R. G., Qu, Y., ... & Fjeldså, J. (2020). The role of evolutionary time, diversification rates and dispersal in determining the global diversity of a large radiation of passerine birds. *Journal of Biogeography*, 47(7), 1612-1625.

- Cardillo, M. (2002). The life-history basis of latitudinal diversity gradients: how do species traits vary from the poles to the equator?. *Journal of Animal Ecology*, 71(1), 79-87.
- Cassemiro, F. A., Albert, J. S., Antonelli, A., Menegotto, A., Wüest, R. O., Cerezer, F., ... & Graham, C. H. (2023). Landscape dynamics and diversification of the megadiverse South American freshwater fish fauna. *Proceedings of the National Academy of Sciences*, 120(2), e2211974120.
- Cerezer, F. O., Machac, A., Rangel, T. F., & Dambros, C. S. (2022). Exceptions to the rule: Relative roles of time, diversification rates and regional energy in shaping the inverse latitudinal diversity gradient. *Global Ecology and Biogeography*, 31(9), 1794-1809.
- Chang, J., Rabosky, D.L., & Alfaro, M. E. (2020). Estimating diversification rates on incompletely sampled phylogenies: theoretical concerns and practical solutions. *Systematic Biology*, 1669 (3), 602–611
- Currie, D. J., Mittelbach, G. G., Cornell, H. V., Field, R., Guégan, J. F., Hawkins, B. A., ... & Turner, J. R. G. (2004). Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, 7(12), 1121-1134.
- Dormann, C.F., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., ... & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30(5), 609-628.
- Dutilleul, P., (1993). Modifying the t test for assessing the correlation between two spatial processes. *Biometrics* 49(1), 305-314.
- Economo, E. P., Narula, N., Friedman, N. R., Weiser, M. D., & Guénard, B. (2018). Macroecology and macroevolution of the latitudinal diversity gradient in ants. *Nature Communications*, 9(1), 1-8.
- Economo, E. P., Huang, J. P., Fischer, G., Sarnat, E. M., Narula, N., Janda, M., ... & Knowles, L. L. (2019). Evolution of the latitudinal diversity gradient in the hyperdiverse ant genus *Pheidole*. *Global Ecology and Biogeography*, 28(4), 456-470.

- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302-4315.
- Gamisch, A., & Comes, H. P. (2019). Clade-age-dependent diversification under high species turnover shapes species richness disparities among tropical rainforest lineages of *Bulbophyllum* (Orchidaceae). *BMC Evolutionary Biology*, 19(1), 1-16.
- García-Andrade, A. B., Carvajal-Quintero, J. D., Tedesco, P. A., & Villalobos, F. (2021). Evolutionary and environmental drivers of species richness in poeciliid fishes across the Americas. *Global Ecology and Biogeography*, 30(6), 1245-1257.
- Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., ... & Zhang, M. (2011). The community climate system model version 4. *Journal of Climate*, 24(19), 4973-4991.
- Gillman, L. N., & Wright, S. D. (2014). Species richness and evolutionary speed: the influence of temperature, water and area. *Journal of Biogeography*, 41(1), 39-51.
- Goldberg, E. E., Lancaster, L. T., & Ree, R. H. (2011). Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology*, 60(4), 451-465.
- Gotelli, N. J., Anderson, M. J., Arita, H. T., Chao, A., Colwell, R. K., Connolly, S. R., ... & Grytnes, J. A. (2009). Patterns and causes of species richness: a general simulation model for macroecology. *Ecology Letters*, 12(9), 873-886.
- Guedes, T. B., Sawaya, R. J., Zizka, A., Laffan, S., Faurby, S., Pyron, R. A., ... & Antonelli, A. (2018). Patterns, biases and prospects in the distribution and diversity of Neotropical snakes. *Global Ecology and Biogeography*, 27(1), 14-21.
- Harvey, M. G., Bravo, G. A., Claramunt, S., Cuervo, A. M., Derryberry, G. E., Battilana, J., ... & Derryberry, E. P. (2020). The evolution of a tropical biodiversity hotspot. *Science*, 370(6522), 1343-1348.
- Hawkins, B. A. (2001). Ecology's oldest pattern?. *Trends in Ecology & Evolution*, 16(8), 470.

- Hawkins, B. A., & Diniz-Filho, J. A. F. (2004). 'Latitude' and geographic patterns in species richness. *Ecography*, 268-272.
- Heard, S. B., & Hauser, D. L. (1995). Key evolutionary innovations and their ecological mechanisms. *Historical Biology*, 10(2), 151-173.
- Hoorn, C., Wesselingh, F. P., Ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., ... & Antonelli, A. (2010). Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, 330(6006), 927-931.
- Hoorn, C., Bogotá-A, G. R., Romero-Baez, M., Lammertsma, E. I., Flantua, S. G., Dantas, E. L., ... & Chemale Jr, F. (2017). The Amazon at sea: Onset and stages of the Amazon River from a marine record, with special reference to Neogene plant turnover in the drainage basin. *Global and Planetary Change*, 153, 51-65.
- Hutter, C. R., Lambert, S. M., & Wiens, J. J. (2017). Rapid diversification and time explain amphibian richness at different scales in the Tropical Andes, Earth's most biodiverse hotspot. *The American Naturalist*, 190(6), 828-843.
- Igea, J. & Tanentzap, A.J. (2019). Multiple macroevolutionary routes to becoming a biodiversity hotspot. *Science Advances*, 5(2), eaau806
- Igea, J., & Tanentzap, A. J. (2020). Angiosperm speciation cools down in the tropics. *Ecology Letters*, 23(4), 692-700.
- Jansson, R., Rodríguez-Castañeda, G., & Harding, L. E. (2013). What can multiple phylogenies say about the latitudinal diversity gradient? A new look at the tropical conservatism, out of the tropics, and diversification rate hypotheses. *Evolution*, 67(6), 1741-1755.
- Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*, 491(7424), 444-448.
- Katoh, K., & Toh, H. (2008). Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC Bioinformatics*, 9, 1-13.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772-780.

- Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., ... & Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26(11), 1463-1464.
- Köhler, G. (2008). *Reptiles of Central America*. 2nd edition. Herpeton, Offenbach, Germany. 400 p.
- Kozak, K. H. (2017). What drives variation in plethodontid salamander species richness over space and time?. *Herpetologica*, 73(3), 220-228.
- Kozak, K. H., & Wiens, J. J. (2012). Phylogeny, ecology, and the origins of climate–richness relationships. *Ecology*, 93(8), S167-S181.
- Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29(6), 1695-1701.
- Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2016). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34(3), 772-773.
- Laudanno, G., Haegeman, B., Rabosky, D. L., & Etienne, R. S. (2021). Detecting lineage-specific shifts in diversification: a proper likelihood approach. *Systematic Biology*, 70(2), 389-407.
- Lefcheck, J. S. (2016). piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, 7(5), 573-579.
- Li, H., & Wiens, J. J. (2019). Time explains regional richness patterns within clades more often than diversification rates or area. *The American Naturalist*, 193(4), 514-529.
- Louca, S., & Pennell, M. W. (2020). Extant timetrees are consistent with a myriad of diversification histories. *Nature*, 580(7804), 502-505.
- Machac, A. (2020). The dynamics of bird diversity in the New World. *Systematic Biology*, 69(6), 1180-1199.
- Maddison, W.P., & Maddison, D.R. (2017). Mesquite: a modular system for evolutionary analysis. Version 3.2. Available at: <http://mesquiteproject.org>
- Marin, J., & Hedges, S. B. (2016). Time best explains global variation in species richness of amphibians, birds and mammals. *Journal of Biogeography*, 43(6), 1069-1079.

- Marin, J., Rapacciuolo, G., Costa, G. C., Graham, C. H., Brooks, T. M., Young, B. E., ... & Hedges, S. B. (2018). Evolutionary time drives global tetrapod diversity. *Proceedings of the Royal Society B: Biological Sciences*, 285(1872), 20172378.
- Matzke, N. J. (2013). Probabilistic historical biogeography: New models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers in Biogeography*, 5(4), 242–248
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2012). The CIPRES science gateway: enabling high-impact science for phylogenetics researchers with limited resources. In *Proceedings of the 1st Conference of the Extreme Science and Engineering Discovery Environment: Bridging from the extreme to the campus and beyond* (pp. 1-8).
- Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., ... & Turelli, M. (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*, 10(4), 315-331.
- Morales-Barbero, J., Gouveia, S. F., & Martinez, P. A. (2021). Historical climatic instability predicts the inverse latitudinal pattern in speciation rate of modern mammalian biota. *Journal of Evolutionary Biology*, 34(2), 339-351.
- Morrone, J. J. (2014). Biogeographical regionalisation of the Neotropical region. *Zootaxa*, 3782(1), 1-110.
- Myers, C. W. (2003). Rare snakes—five new species from eastern Panama: reviews of northern *Atractus* and southern *Geophis* (Colubridae: Dipsadinae). *American Museum Novitates*, 2003(3391), 1-47.
- Oliveira, B. F., Machac, A., Costa, G. C., Brooks, T. M., Davidson, A. D., Rondinini, C., & Graham, C. H. (2016). Species and functional diversity accumulate differently in mammals. *Global Ecology and Biogeography*, 25(9), 1119-1130.
- Otto-Bliesner, B. L., Marshall, S. J., Overpeck, J. T., Miller, G. H., & Hu, A. (2006). Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science*, 311(5768), 1751-1753.
- Palmer, M. W. (1994). Variation in species richness: towards a unification of hypotheses. *Folia Geobotanica et Phytotaxonomica*, 29(4), 511.
- Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526-528.

- Passos, P., Mueses-Cisneros, J. J., Lynch, J. D., & Fernandes, R. (2009). Pacific lowland snakes of the genus *Atractus* (Serpentes: Dipsadidae), with description of three new species. *Zootaxa*, 2293(1), 1-34.
- Poe, S., Nieto-Montes de Oca, A., Torres-Carvajal, O., De Queiroz, K., Velasco, J. A., Truett, B., ... & Latella, I. (2017). A phylogenetic, biogeographic, and taxonomic study of all extant species of *Anolis* (Squamata; Iguanidae). *Systematic Biology*, 66(5), 663-697.
- Pomar-Gómez, A. D., Cordeiro, P., Guedes, T. B., & Passos, P. (2021). The striking endemism pattern of the species-richest snake genus *Atractus* (Dipsadidae: Serpentes) highlights the hidden diversity in the Andes. *Amphibia-Reptilia*, 42(2), 189-200.
- Pontarp, M., & Wiens, J. J. (2017). The origin of species richness patterns along environmental gradients: uniting explanations based on time, diversification rate and carrying capacity. *Journal of Biogeography*, 44(4), 722-735.
- Pyron, R. A. (2014). Temperate extinction in squamate reptiles and the roots of latitudinal diversity gradients. *Global Ecology and Biogeography*, 23(10), 1126-1134.
- Pyron, R. A., & Burbrink, F. T. (2009). Can the tropical conservatism hypothesis explain temperate species richness patterns? An inverse latitudinal biodiversity gradient in the New World snake tribe Lampropeltini. *Global Ecology and Biogeography*, 18(4), 406-415.
- Pyron, R. A., & Wiens, J. J. (2013). Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proceedings of the Royal Society B: Biological Sciences*, 280(1770), 20131622.
- Qian, H., Wiens, J. J., Zhang, J., & Zhang, Y. (2015). Evolutionary and ecological causes of species richness patterns in North American angiosperm trees. *Ecography*, 38(3), 241- 250.
- Quintero, I., & Jetz, W. (2018). Global elevational diversity and diversification of birds. *Nature*, 555(7695), 246-250.
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5), 901-904.

- R Development Core Team (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rabosky, D. L. (2014). Automatic detection of key innovations, rate shifts, and diversity dependence on phylogenetic trees. *PloS One*, 9(2). e89543.
- Rabosky, D. L. (2016). Challenges in the estimation of extinction from molecular phylogenies: a response to Beaulieu and O'Meara. *Evolution*, 70(1), 218-228.
- Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., ... & Alfaro, M. E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, 559(7714), 392-395.
- Ramirez, J.P., Reeder, T.W., & M.J. Spasojevic (submitted). Extinction debt and functional traits mediate community saturation over large spatiotemporal scales. Target: *Journal of Animal Ecology*
- Ricklefs, R. E. (2006). Evolutionary diversification and the origin of the diversity–environment relationship. *Ecology*, 87(7), S3-S13.
- Ritchie, A. M., Lo, N., & Ho, S. Y. (2017). The impact of the tree prior on molecular dating of data sets containing a mixture of inter-and intraspecies sampling. *Systematic Biology*, 66(3), 413-425.
- Rodriguez-Muñoz, E., Montes, C., Rojas-Runjaic, F. J., & Crawford, A. J. (2022). Synthesis of geological data and comparative phylogeography of lowland tetrapods suggests recent dispersal through lowland portals crossing the Eastern Andean Cordillera. *PeerJ*, 10, e13186.
- Rohde, K. (1992). Latitudinal gradients in species diversity: the search for the primary cause. *Oikos*, 514-527.
- Roll, U., Feldman, A., Novosolov, M., Allison, A., Bauer, A. M., Bernard, R., ... & Colli, G. R. (2017). The global distribution of tetrapods reveals a need for targeted reptile conservation. *Nature Ecology & Evolution*, 1(11), 1677-1682.
- Rolland, J., Condamine, F.L., Jiguet, F. & Morlon, H. (2014) Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biology*, 12, e1001775.
- Roy, K., & Goldberg, E. E. (2007). Origination, extinction, and dispersal: integrative models for understanding present-day diversity gradients. *The American Naturalist*, 170, S71-85.

- Sanmartín, I. (2012). Historical biogeography: evolution in time and space. *Evolution: Education and Outreach*, 5(4), 555-568.
- Šavrič, B., Patterson, T., & Jenny, B. (2019). The equal earth map projection. *International Journal of Geographical Information Science*, 33(3), 454-465.
- Schluter, D., & Pennell, M. W. (2017). Speciation gradients and the distribution of biodiversity. *Nature*, 546(7656), 48-55.
- Shrestha, N., Wang, Z., Su, X., Xu, X., Lyu, L., Liu, Y., ... & Feng, X. (2018). Global patterns of *Rhododendron* diversity: The role of evolutionary time and diversification rates. *Global Ecology and Biogeography*, 27(8), 913-924.
- Shipley, B. (2000). A new inferential test for path models based on directed acyclic graphs. *Structural Equation Modeling*, 7(2), 206-218.
- Skeels, A., Esquerré, D., & Cardillo, M. (2020). Alternative pathways to diversity across ecologically distinct lizard radiations. *Global Ecology and Biogeography*, 29(3), 454-469.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312-1313.
- Stevens, R. D. (2011). Relative effects of time for speciation and tropical niche conservatism on the latitudinal diversity gradient of phyllostomid bats. *Proceedings of the Royal Society B: Biological Sciences*, 278(1717), 2528-2536.
- Storch, D., & Okie, J. G. (2019). The carrying capacity for species richness. *Global Ecology and Biogeography*, 28(10), 1519-1532.
- Sun, M., Folk, R. A., Gitzendanner, M. A., Soltis, P. S., Chen, Z., Soltis, D. E., & Guralnick, R. P. (2020). Recent accelerated diversification in rosids occurred outside the tropics. *Nature Communications*, 11(1), 1-12.
- Tello, J. S., & Stevens, R. D. (2010). Multiple environmental determinants of regional species richness and effects of geographic range size. *Ecography*, 33(4), 796-808.
- Title, P. O., & Rabosky, D. L. (2019). Tip rates, phylogenies and diversification: what are we estimating, and how good are the estimates? *Methods in Ecology and Evolution*, 10(6), 821-834.

- Tietje, M., Antonelli, A., Baker, W. J., Govaerts, R., Smith, S. A., & Eiserhardt, W. L. (2022). Global variation in diversification rate and species richness are unlinked in plants. *Proceedings of the National Academy of Sciences*, 119(27), e2120662119.
- Uetz, P., Freed, P. & Hošek, J. (Eds.) (2023). The Reptile Database, <http://www.reptile-database.org>, accessed 23 January 2023.
- Vaidya, G., Lohman, D. J., & Meier, R. (2011). SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, 27(2), 171-180.
- Vallejos, R., Osorio, F., & Bevilacqua, M. (2018). *Spatial Relationships Between Two Georeferenced Variables: with Applications in R*. Springer, New York, USA.
- Velasco, J. A., Villalobos, F., Diniz-Filho, J. A., Algar, A. C., Flores-Villela, O., Köhler, G., ... & Martinez-Meyer, E. (2018). Climatic and evolutionary factors shaping geographical gradients of species richness in *Anolis* lizards. *Biological Journal of the Linnean Society*, 123(3), 615-627.
- Velasco, J. A., & Pinto-Ledezma, J. N. (2022). Mapping species diversification metrics in macroecology: Prospects and challenges. *Frontiers in Ecology and Evolution*, 10, 951271.
- Vilela, B., & Villalobos, F. (2015). letsR: a new R package for data handling and analysis in macroecology. *Methods in Ecology and Evolution*, 6(10), 1229-1234.
- von Humboldt, A. (1808). *Ansichten der Natur mit wissen-schaftlichen Erläuterungen*. J. G. Cotta, Tübingen, Germany.
- Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist*, 156(2), 145-155.
- Wiens, J. J. (2007). Global patterns of diversification and species richness in amphibians. *The American Naturalist*, 170(S2), S86-S106.
- Wiens, J. J. (2011). The causes of species richness patterns across space, time, and clades and the role of “ecological limits”. *The Quarterly Review of Biology*, 86(2), 75-96.
- Wiens, J. J., & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, 19(12), 639-644.

- Wiens, J. J., Ackerly, D. D., Allen, A. P., Anacker, B. L., Buckley, L. B., Cornell, H. V., ... & Hawkins, B. A. (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, 13(10), 1310-1324.
- Willig, M.R. & Presley, S.J. (2018) Latitudinal gradients of biodiversity: theory and empirical patterns, pp 13–19 In: DellaSala, D. & Goldstein, M. (eds) *Encyclopedia of the Anthropocene*. Elsevier, Amsterdam, the Netherlands.
- Wu, G., & Wiens, J. J. (2022). The origins of climate-diversity relationships and richness patterns in Chinese plants. *Ecology and Evolution*, 12(12), e9607.
- Yule, G. U. (1925). A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis, FR S. *Philosophical Transactions of the Royal Society of London. Series B* 213(402-410), 21-87.
- Zaher, H., Murphy, R. W., Arredondo, J. C., Graboski, R., Machado-Filho, P. R., Mahlow, K., ... & Grazziotin, F. G. (2019). Large-scale molecular phylogeny, morphology, divergence-time estimation, and the fossil record of advanced caenophidian snakes (Squamata: Serpentes). *PloS One*, 14(5), e0216148.

Tables

Table 2.1. List and description of the main ecological and evolutionary explanations of the associations between regional climatic conditions and patterns of species richness. Modified after Velasco et al. (2018) and Garcia-Andrade et al. (2020). Please note that for this study, we treated separately the influence of the two factors included in hypotheses 1, 2, and 4 (i.e., precipitation and temperature) on regional species richness.

Hypothesis	WorldClim Variables	Explanation	Sources from climatic data
1 Water-energy or climate seasonality	Temperature seasonality (BIO4) or Precipitation seasonality (BIO15)	The species richness of a region is expected to be inversely proportional to the temperature or precipitation seasonality it exhibits.	WorldClim
2 Environmental heterogeneity	Elevation was used as a proxy of environmental heterogeneity	The species richness of a region is expected to be directly proportional to the climatic or spatial heterogeneity it exhibits.	WorldClim
3 Climate stability (or climatic anomalies) since the Last Inter-Glacial period (IGM)	Annual mean temperature (BIO1) or annual precipitation (BIO12) values, both from the present time and the Last Inter-Glacial period.	Regions where their temperature or precipitation has changed the less (=where the climatic anomalies are smaller) since the IGM are expected to contain more species	WorldClim. Data from the Last Inter-Glacial period (IGM; ~130 000 years before present) was obtained from Otto-Bliesner et al. (2006)
4 Climate stability (or climatic anomalies) since the Last Glacial Maximum (LGM)	Annual mean temperature (BIO1) or annual precipitation (BIO12) values, both from the present time and the Last Glacial Maximum period	Regions where their temperature or precipitation has changed the less (=where the climatic anomalies are smaller) since the LGM are expected to contain more species	Worldclim. Data from the Last Glacial Maximum (LGM; ~21000 years before present) was obtained from the Community Climate System Model, version 4 (CCSM4; Gent et al., 2011)

Figures

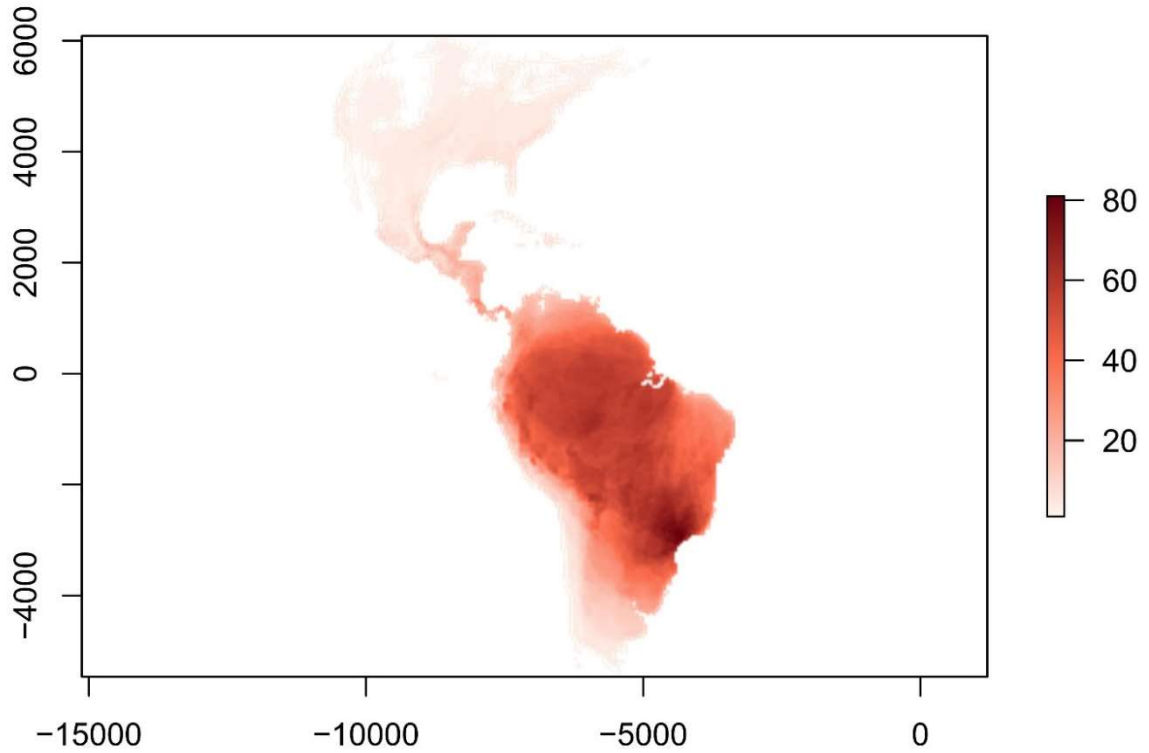


Figure 2.1. Patterns of species richness of the snake family Dipsadidae across the New World, as estimated from the range maps compiled by Roll et al. (2017). Darker and paler tones of red correspond to areas with higher and lower species richness of Dipsadidae, respectively. The map is overlaid on a 50 km x 50 km grid using an equal-earth projection. Cells with less than 25% of land surface are not mapped.

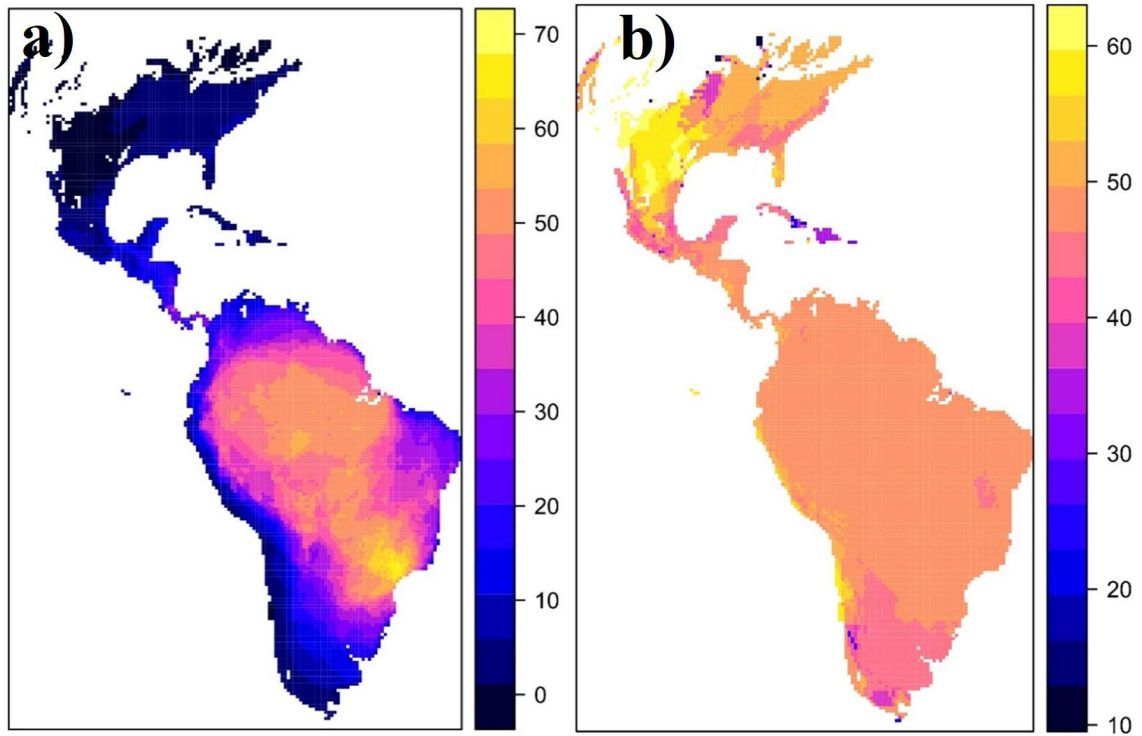


Figure 2.2. Patterns of geographic variation of the values of per-cell a) DR metric, and b) Mean Phylogenetic Diversity (MPD) of Jetz et al. (2012) across the distribution of the family Dipsadidae in the New World, as obtained with the 100 completely sampled phylogenies we generated with the stochastic polytomy resolving software TACT. To calculate these values, we also used the range maps compiled by Roll et al. (2017). Warmer and color colors correspond to higher and lower values of each metric. The map is overlaid on a 50 km x 50 km grid using an equal-earth projection. Cells with less than 25% of land surface are not mapped.

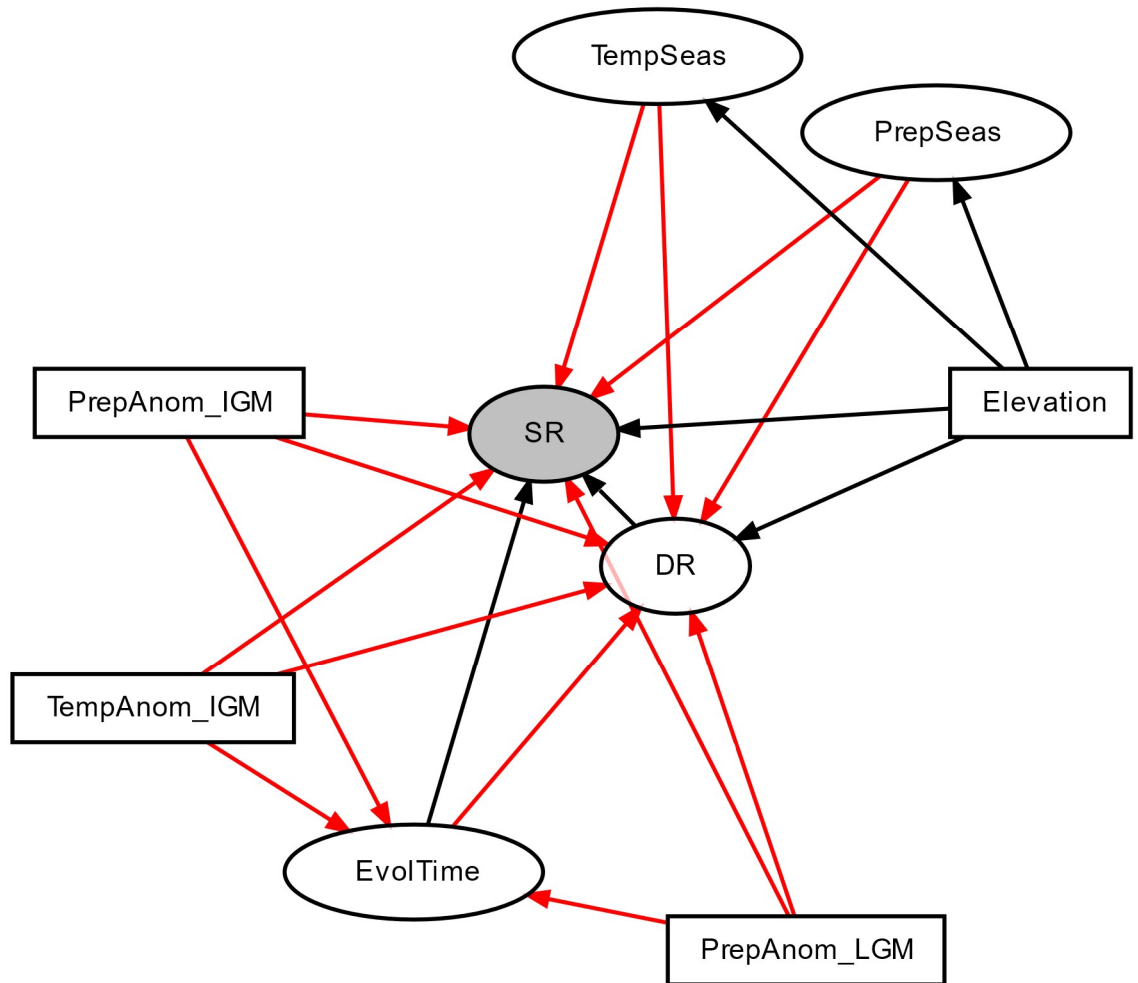


Figure 2.3. Theoretical structural equation model tested in this work for understanding the direct and indirect influences of past and current climate, species accumulation times, and speciation rates on the regional species richness of the snake family Dipsadidae. Exogenous and endogenous variables of the model are indicated with rectangles and ovals, respectively. In that order, black and red arrows indicate whether the linked variables are expected to be positively or negatively correlated. Abbreviations as follows: *SR* = Species richness, *EvolTime* = Evolutionary time (as measured with the MPD metric), *DR* = speciation rate (as measured with the DR metric), *PrepSeas* = current precipitation seasonality, *TempSeas* = current temperature seasonality, *PrepAnom* = precipitation anomaly, *TempAnom* = temperature anomaly, *LGM* = Last Glacial Maximum, *IGM* = Last Interglacial Maximum.

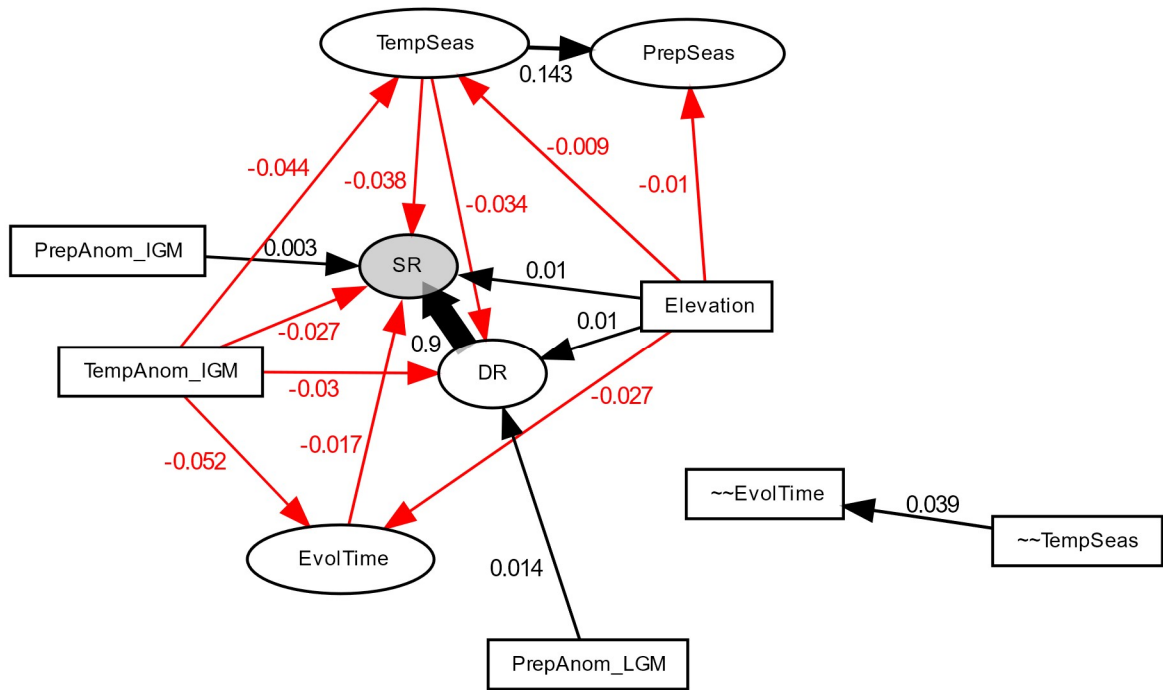


Figure 2.4. Resulting piecewise structural equation model obtained in this work regarding the factors influencing regional species richness variation in the snake family Dipsadidae, using Roll et al. (2017)’s dataset of species distributions. For this model, we used the values of the DR and MPD metrics obtained using the 100 completely sampled phylogenies we generated with the stochastic polytomy resolving software TACT. Arrows between variables indicate the pathways of the final model, with the numbers in-between (and the thickness of the arrows) representing the corresponding values of the *direct* standardized effect sizes. Non-significant paths were removed to facilitate the graphical representation of the model. Abbreviations as in Figure 2.3.

CHAPTER 3

Extinction debt and functional traits mediate community saturation over large spatiotemporal scales

Abstract

Determining if ecological communities are saturated (have a limit to the number of species they can support) has important implications for understanding community assembly, species invasions, and climate change. However, previous studies have generally been limited to short time frames that overlook extinction debt and have not explicitly considered how functional trait diversity may mediate patterns of community saturation. Here we combine data from biodiversity surveys with functional and phylogenetic data to explore if the colonization events after the Great American Biotic Interchange (closure of the Panamanian Isthmus) resulted in increases in species richness of communities of the snake family Dipsadidae. We determined the number and the direction of dispersal events between Central and South America by estimating ancestral areas based on a Bayesian time-calibrated phylogenetic analysis. We then evaluated whether variation in community saturation was mediated by the functional similarity of six traits for the resident and colonizing snakes and/or local environmental conditions. We found that colonized communities did not support more species than those that were not colonized. Moreover, we did not find an association between the functional

dispersion across sites and whether they were colonized by species dispersing across the Isthmus or not. Instead, variation in species richness was predicted best by covariates such as time since colonization and local environment. Taken together, our results suggest that snake communities of the Dipsadidae across the neotropics are saturated. Moreover, our research highlights two important factors to consider in studies of community saturation: extinction debt and the functional differences and similarities in species ecological roles.

Introduction

Understanding the processes responsible for determining how many species can coexist at a given site is a key goal for community ecology (Kraft et al., 2007; Vellend, 2010), with important implications for invasion biology (MacDougall et al., 2009) and predicting how communities will respond to global change (Chen et al. 2011, HilleRisLambers et al. 2013). Classically, much of the focus of such research has been on community saturation - if novel species can freely establish in a new community or if the structure of the community prevents their establishment unless they replace native species (Elton, 1958; MacArthur, 1965; MacArthur & Levins, 1967, Ricklefs, 1987). To date, many studies have used biological invasions to test for saturation and have only focused on species richness (e.g., Sax et al., 2002; Myers & Harms, 2009; and studies cited in Sax & Gaines, 2003; see also Olivares et al., 2018). The assumption of many of these studies is that if the species richness of a community does not increase after the introduction of species from other regions (either naturally or through human action), the community is saturated, while if richness does increase, the community is unsaturated (Olivares et al., 2018). For example, the lack of extinction of native species after the introduction of non-native species, in many cases, has led many researchers to suggest that many non-island communities are unsaturated (Gurevitch & Padilla, 2004; Sax et al., 2002; Smith & Shurin, 2006; Sax et al., 2007; Stohlgren et al., 2008).

Despite decades of research on saturation (e.g., Elton, 1950; MacArthur & Wilson, 1963; Terborgh & Faarborg, 1980), evidence for saturation is mixed (Olivares et al., 2018), with many studies confounding saturation with stability or equilibrium

(Olivares et al., 2018). More critically, much of the previous research on saturation suffers from two key issues that limit generalization. First, much of the research on saturation is based on exploring the consequences of biotic introductions without considering *extinction debt* – that the exclusion of natives by introduced species (or vice-versa) is not observed because of the long-time frame required for competitive exclusion to occur (Gilbert & Levine, 2013). Indeed, most studies testing for saturation based on evaluating the impacts of colonizing taxa have done so in the context of modern-day biological invasions (see Stachowicz & Tilman, 2005; Olivares et al., 2018). While monitoring the effects of species introductions through long periods of time (i.e., over hundreds to millions of years) is obviously not feasible, integrating historical biogeography into studies of community saturation can provide important insights into the long-term implications of species invasions (Pinto-Sanchez et al., 2014) and resolve the issue of extinction debt. Specifically, historical biogeography allows us to determine the origin and timing of colonization events that occurred in biological communities over thousands or millions of years ago. Using ancestral area estimation to infer the geographical origin of taxa makes it possible to infer if communities that had a greater number of long-term colonization events are also more speciose- a result that would support the hypothesis that communities are not saturated (Pinto-Sánchez et al., 2014).

A second limitation for assessing saturation is that determining if the species richness of communities differs if they were colonized or not does not account for the differences and similarities in species' ecological roles. The number of ecological roles in a given community can be determined by the size of the trait space filled by species in a

community (i.e., functional richness; Cornwell et al. 2006) or the dissimilarity in functional strategies among species (i.e., functional dispersion; Laliberté & Legendre, 2010). Thus, examining how the functional diversity of communities relates to whether they were colonized or not can help infer the mechanisms by which these species were able to invade in the first place. For example, finding that novel species increase community functional richness (the total multivariate volume of the trait space for all the species in a community; Mason et al., 2005; Villéger et al., 2008) or functional dispersion (mean distance of all species to the weighted centroid of the community in trait space, Laliberté & Legendre, 2010; Valdivia et al., 2017) would suggest that the new species bring new trait states to the community, expanding its original trait space (Gerhold et al., 2011). On the other hand, finding that the trait-space of a community either diminishes or does not change upon invasion would suggest that the novel colonizing species either use portions of the niche space that were either empty or previously occupied by native species (with the later suggesting niche limitation).

Here, we combine community composition data with a calibrated phylogeny and functional trait data to infer the processes underlying variation in community saturation after one of the largest biotic interchanges in recent Earth's history – the Great American Biotic Interchange (GABI) – the exchanges between the long-isolated North and South American biotas after the closure of the Panamanian Isthmus (Simpson, 1980; Stehli & Webb, 1985). The GABI, which occurred either 15 or 3 million years ago (see Montes et al., 2012; O'Dea et al., 2016 for debate about the exact date), provides unique insights for understanding the processes determining community saturation (Vermeij, 1991) and a

complementary point of view from the one given by the study of human-mediated species introductions (Stachowicz & Tilman, 2005). Although the impacts on the native biotas of the taxa involved in this large-scale event have been discussed at the continental or regional scales (e.g., Marshall et al., 1982), they remain mostly unexplored at the community scale (Pinto-Sánchez et al., 2014). Here, we focus on the hyper-diverse snake family Dipsadidae, with its >800 species and 112 genera inhabiting many different environments from Canada to Argentina (Uetz et al., 2023). Members of the Dipsadidae are a major component of Neotropical snake communities, and a variety of ecological morphotypes can live in sympatry. First, we used a calibrated phylogeny of the group to determine the timing and direction of the dispersion events of the dipsadid subclades across the Panamanian Isthmus. We then compiled species composition data for 67 communities of dipsadid snakes ranging from Mexico to Argentina, as well as data for six life-history traits thought to mediate their ecological interactions. We then analyzed if the total number of species within a community differs between those receiving or not colonizers from across the Isthmus to infer if dipsadid snake communities were saturated. Lastly, we tested if community functional trait diversity varies between those that received or not species from across the Isthmus to assess if colonizing species either were associated with an increase or not in the overall functional diversity of the communities they invaded.

Materials and Methods

Data set.

We established a database on local species composition and richness of snakes of the family Dipsadidae by compiling data from the literature for a total of 67 sites located between Mexico and Argentina, with 21 of them in North and Central America and 46 in South America (Figure 3.1a, see Suppl. Table S3.1 for species composition of each site). Sites ranged between 2-30 species per site, with an average of 9.69 ± 6.67 species (Suppl. Table S3.1, Figure 3.1b), adding together to a total of 215 species across all sites. For each of the 67 sites included, we also compiled data on variables that may influence the species richness or the functional trait space of each site (summarized in Appendix Table A3.1). We first included variables related to sampling effort that may bias our analyses, such as area sampled, sampling effort (in units of person-hours), time of first colonization (see below), or the type of sampling methodology used (i.e., if sampling methods different from visual searches, such as pitfall traps or drift fence arrays, were used or not).

Next, we included environmental variables for each site that may influence variation in richness or functional diversity. We first obtained site elevation from the original sources (or from Google Maps if not reported). We then extracted the 19 climatic variables available from the WorldClim 2.0 repository using a resolution of 2.5 minutes of a degree (Fick & Hijksmans, 2017). As many of these variables are correlated, we used a Principal Component Analysis (PCA) to describe the environment in four key axes of variation (Figure 3.2). PC1 explained 43.22% of the variation in our data where high

values of PC1 correspond with warm winter conditions (BIO6, 11, and 9) and low values indicate cold winter conditions. PC2 explained 22.12% of the variation in our data where high values of PC2 correspond with drier and warmer conditions during the summer (BIO 7, 14, and 5), and low values indicate rainier and colder summer conditions. PC3 explained 11.39% of the variation in our data where high values of PC3 correspond with low-elevation sites with high summer temperatures but reduced precipitation seasonality (BIO15, elevation, and BIO10). Finally, PC4 explained 8.20% of the variation in our data, where high values of PC4 correspond with low precipitation during the winter (BIO13, 3, and 16) and low values correspond with increased winter rainfall. PCA loadings, summary values, and scores are summarized in Appendix Tables A3.2, A3.3, and Suppl. Table S3.2, respectively. Lastly, both the species richness and functional diversity of dipsadid snakes present at a given site are also influenced by their interactions with species from other groups of snakes living in sympatry. Thus, we generated a list of the snake species from families other than the Dipsadidae found at each site (Suppl. Table S3.3) and calculated their species richness (Appendix Table A3.4).

Phylogenetic reconstruction and ancestral area estimation

First, to build a phylogenetic tree of snakes of the family Dipsadidae, we compiled all sequences corresponding to five mitochondrial (12S, 16S, *cyt b*, ND2, and ND4) and four nuclear (BDNF, *c-mos*, NT3, and RAG2) genes publicly deposited in GenBank (up until October 2018). We also incorporated sequences from 24 species from

14 different families of Caenophidia that were used as outgroups. In total, we compiled sequence data from a total of 341 species. We then aligned sequences with the online version of MAFFT v7 (Katoh & Standley, 2013) under default parameters, except for the case of the 12S and 16S ribosomal genes, for which we used the Q-INS-i algorithm (which considers the secondary structure of RNA sequences, Katoh & Toh, 2008). We then performed a Bayesian phylogenetic analysis (BA) with BEAST 2 v4.8 (Bouckaert et al., 2014). We applied the bModelTest package of BEAST 2 to simultaneously explore the substitution model space and estimate both the corresponding model parameters and the phylogeny (Bouckaert & Drummond, 2017). The sampling of the Beast 2 analysis used two independent MCMC chains of 50 million generations each, with the first 50% of trees discarded as burn-in. The posterior distribution of trees was summarized as a Maximum Clade Credibility (MCC) timetree using TreeAnnotator v2.4.8, part of the BEAST 2 package. See Appendix 3.1 for further details.

Next, to determine the number and the direction of dispersal events between Central America and South America, we estimated ancestral areas by implementing a model-based approach (the R package BioGeoBEARS, Matzke, 2013a, 2014) using the resulting Bayesian timetree (see details in Appendix 3.1). Critically, these results permitted identifying the geographic origin of each of the New World lineages comprising the snake family Dipsadidae. We considered taxa as being involved in the GABI if they are part of a lineage whose ancestral area is in a different subcontinent than what was inferred for its immediate ancestor and if that dispersal event occurred more recently than 15 million years ago. Using a cut-off of 3 million years did not change the

results of the analyses performed for this study (results not shown). The results of the ancestral area estimation also allowed us to obtain a rough estimate of the time when each site was first colonized by members of the Dipsadidae. To do so, we assumed that the time when dipsadids first arrived at a given site was the same as when the region where it is located was first colonized. We used the time of first colonization of each region and not the ages of all posterior colonization events, following Pinto-Sanchez et al. (2014) and the studies cited therein, which show that both approaches give similar results.

Functional traits.

To examine variation in functional diversity among sites, we extracted trait data from the literature for six life-history traits of the species they contain: microhabitat use, time of activity, reproductive mode, diet, total length, and total mass (See details in Appendix 3.1 and data in Suppl. Table S3.4). We selected these traits due to them being commonly considered as important components of the life history of snakes (e.g., Henderson et al. 1979; Shine, 1994; Pianka et al., 2017). We then calculated functional dispersion (fDis) based on the presence of species at each site using the function *dbFD* in the 'FD' package (Laliberté & Shipley, 2011) in R (R Development Core Team, 2020; 4.0.3 summarized in Appendix Table A3.1). Although there are many metrics of functional diversity (reviewed in Mouchet et al., 2010, Schleuter et al., 2010), we focused here on fDis because it is independent of species richness. Further, it can be used on both continuous and discrete characters, as well as for single or multiple traits (Laliberte &

Legendre, 2010). Finally, we also compiled data of the same six functional trait data mentioned above from species of snake families different from the Dipsadidae (Suppl. Table S3.5). We then used those values to estimate the fDis of non-Dipsadidae snakes found at each site (Figure 3.1c; Appendix. Table 3.4).

Statistical analyses

To determine if sites containing species colonizing them across the Panamanian Isthmus exhibit greater values of species richness than those that were not colonized, we used a generalized linear model (GLM). The model included species richness as a response variable and whether each site received or not species from across the Isthmus as a fixed independent variable. We also included covariates of species richness such as trap use, site area, sampling effort, local environmental conditions (as summarized into the four PC axes of climatic variation), time since first colonization by members of Dipsadidae, as well as the species richness and functional dispersion of other groups of snakes as additional fixed independent predictors. We used a quasipoisson error distribution and log link function, as the response variable (=the species richness of each site) only exhibits discrete values, and the data was overdispersed.

To determine if sites receiving species dispersing from across the Panamanian Isthmus also exhibit greater values of functional dispersion, we regressed the values of these two variables by using a Beta regression with a logit link function. Again, we considered the same six predictor variables described above, along with if sites were colonized or not by species from the opposite side of the Isthmus, as fixed independent

variables. We used a Beta regression (as implemented in the R package ‘betareg’; Cribari-Neto & Zeileis, 2010), as the response variable (fDis of each site) is continuous and ranges between zero and one without including these values.

For both the GLM and Beta regression analyses we performed, we first determined if they conform to their assumption of no-multicollinearity by using the command *cor.test* of R. We also calculated the values of Variance Inflation Factors (VIF) of each of their independent variables by using the command *vif* of R package ‘car’. After excluding the variables of site area and the first climate axis, all independent variables had VIFs lesser than the values commonly considered as indicating multicollinearity (VIF values of either 4 or 10, depending on the author; O'Brien, 2014).

Results

Phylogenetic reconstruction and ancestral area estimation.

The model with the best fit for our phylogenetic and distributional data corresponded to the BAYAREALIKE biogeographical model (a likelihood version of the full BayArea model of Landis et al., 2013; see Matzke, 2013b) with the jump parameter (+j, Matzke, 2014). (Appendix Table A3.5). The results of the Ancestral Area Estimation indicate the existence of at least 24 dispersal events across the Panamanian Isthmus by members of the Dipsadidae (Appendix Table A3.6), with 19 of them estimated to have occurred after 15 million years ago. Of the latter, 8 of them occurred from North to South America (all of them corresponding to members of Dipsadinae), and 11 of them in the opposite direction (with three of them corresponding to members of the Dipsadinae, and

eight from members of the Xenodontinae). A more detailed description of the results of the ancestral area estimation analysis is presented in Appendix 3.2.

Community saturation.

We found that there was no relationship between local species richness and whether communities received or not species from the other side of the Isthmus (z -value=1.086, $p=0.286$, $df=41$) as expected if communities were saturated. Instead, variation in species richness was best predicted by the species richness of non-dipsadid snakes (z -value=3.765; $p=6.98E-04$; Table 3.1; Appendix Figure A3.1). Similarly, we did not find an association between the functional dispersion of communities and if they were colonized or not by taxa dispersing across the Isthmus (z -value=-0.914; $p=0.361$; $df=12$), further in accordance with what would be expected if communities were saturated.

Discussion

While many short-term studies focusing on recent biological invasions have found that communities are unsaturated (e.g., Sax et al., 2002; Stohlgren et al., 2008; Myers & Harms, 2009; and the studies cited in Sax & Gaines, 2003), ours is one of the few using patterns of species introductions to find evidence for saturation over long temporal scales. Critically, most studies based on the impacts of modern-day biotic introduction events have generally not considered extinction debt or functional traits. As a result, the unsaturation of modern-day biological communities could correspond to a temporary artifact in which extinction debts will eventually be paid. In line with these modern-day

studies, most studies investigating the impacts of biotic introductions that occurred millions of years ago also find patterns of unsaturation (Vermeij, 1991; Lindberg, 1991; Flynn et al., 1991, Patzkowski & Holland 2007; other references cited in Smith & Shurin, 2006 and Pinto-Sanchez et al., 2014). However, these studies are generally based on evidence from the fossil record, so determining the coexistence of taxa at the community scale (and their identification at the species level) is difficult at best (Pinto-Sanchez et al., 2014). Interestingly, the only other study using historical biogeography for directly testing for community saturation at large temporal scales (i.e., Pinto-Sanchez et al., 2014) did not find evidence for community saturation. This discrepancy is surprising as both studies are based on analyzing the impacts of the GABI on the species richness of Neotropical communities. However, the difference between our results and those of Pinto-Sanchez et al. (2014) may be due to the high levels of similarity in the ecological roles (e.g., functional traits) exhibited by the group of frogs (Terrarana) used by Pinto-Sanchez et al. (2014) as compared to the snakes in the Dipsadidae. In parts of their distribution, many species of these frogs can live on sympatry without any apparent partitioning in habitat use or diet (e.g., Lieberman, 1986; Arroyo et al., 2008). The results of Pinto-Sanchez et al. (2014) likely indicate that, albeit the communities of frogs they studied are unsaturated in terms of species counts, these communities have already reached the limit of ecological entities they can harbor.

By merging historical biogeographic inferences (e.g., ancestral area reconstructions) with modern ecological approaches (e.g., functional traits), we provide a robust approach for examining community saturation that suggests that community

saturation may occur more frequently than suggested by the current evidence. Critically, our framework satisfies the conditions needed to assess saturation recently proposed by Olivares et al. (2018): the existence of both *niche limitation* and either *equilibrium* or *stability* in numbers of species or ecological entities after accounting for potential covariates (e.g., local climate, time of first colonization, site area, and sampling effort). Importantly, we explicitly quantified the functional space occupied by species. Metrics of the functional trait space (such as functional dispersion) consider the ecological redundancy and the roles that organisms have on ecosystems; aspects not considered by metrics of taxonomic diversity (Canning-Clode et al., 2010; Schleuter et al., 2010; Cadotte et al., 2011). However, functional trait space has only been quantified in a few studies testing for community saturation (e.g., Canning-Clode et al., 2010), none of which are based on studying the impacts of biotic introductions. As a result, it is possible that the common finding of community species richness being unsaturated may be obscured by not considering functional diversity.

We found that using metrics of the functional trait diversity while testing for community saturation is indicative of *niche limitation* in communities of Dipsadidae across the neotropics, as we focused on functional traits associated with different niche axes in snake communities. First, body length is among the most important traits not only in snakes but in most other groups of organisms (see Shine, 1994; Boback & Guyer, 2003; Feldman et al., 2016; and references in Burbrink et al., 2012). Body length in snakes is correlated to their mass (Boback & Guyer, 2003; Feldman & Meiri, 2012), but more importantly, it has been shown that the diet of snakes is highly dependent on their

body size. Indeed, the diet of larger snake species contains increased percentages of endotherms, while smaller species feed more on invertebrates (e.g., Shine, 1994; Pyron & Burbrink, 2009). Thus, it is likely that competition for food among snakes of different body lengths is limited. This is especially true considering the very specialized diets exhibited by some snakes, including several members of the Dipsadidae (such as the mollusk and annelid specialists *Dipsas* and *Sibon* [Ray et al., 2012], the small reptile eggs specialists *Enuliulus* and *Enuliophis* [Savage, 2002; Brown et al., 2018], or the crayfish specialist *Farancia* [Godley et al., 1984]). The role of predatory interactions between the snakes considered in this study (facilitated by differences in body size between species) is also likely to be reduced, as only a few of them seem to be ophiophagous (snake eating). Similarly, two other traits considered herein (time of activity and microhabitat use) are also commonly considered to influence the biotic interactions between coexisting snake taxa (e.g., Henderson et al., 1979; Vitt, 1987; Burbrink & Myers, 2015). Thus, snake species similar in most other respects (including body length) could coexist if they differ in the times of activity and/or the resources they use due to inhabiting different microhabitats. Finally, both oviparity and viviparity can provide different competitive advantages, but the relative advantage of each reproductive strategy often depends on the local climatic conditions (Blackburn, 1999). For example, viviparity is considered more advantageous in colder climates, as it allows development to occur faster, increasing the survival of young individuals (Bleu et al., 2012). By focusing on multiple niche axes, the traits we analyzed are likely to link to the many biotic interactions that may influence the saturation of neotropical snake communities (Mlambo, 2014).

In addition to niche limitation, we also addressed the second condition needed to assess saturation recognized by Olivares et al. (2018): *equilibrium* or *stability* in numbers of species or ecological entities. For the case of contemporary biotic invasions, the term stability refers to whether the species richness of communities does not significantly change even as the colonization events increase the size of the regional species pool (Olivares et al., 2018). In other words, the richness of sites would be considered stable if it remains unchanged despite having the opportunity to increase along with the number of colonizing lineages. Our results are consistent with this expectation, as we found that the species richness of the communities of interest did not depend on whether they were colonized or not by taxa dispersing from across the Panamanian Isthmus. Although it is impossible to know how the species richness of individual communities has changed through time, our results indicate that the general tendency was for the richness of colonized sites not to increase to a level that surpasses that of uncolonized sites. However, it is important to note that the lack of a significant difference in the species richness between colonized and uncolonized communities does not necessarily indicate that they are saturated. Indeed, it is still possible that the increased dispersal (and posterior diversification) following the GABI could have increased the richness of colonized sites to a level comparable to that of uncolonized communities. Despite this, we did not find evidence supporting the idea that colonized communities harbored more taxa than uncolonized sites before the GABI occurred. Instead, we found that across the several variables we compared between these two groups (see Appendix 3.2), the few that differ between them suggest that colonized sites used to harbor more species than

uncolonized communities. Indeed, we recovered that colonized communities tend to occur at lower latitudes and have higher winter temperatures than uncolonized sites (see Appendix 3.2). Considering that dipsadid snakes are ectotherms and currently exhibit a latitudinal species gradient (LDG) (Ramirez & Reeder, in prep.), it is likely that such a pattern already existed before the GABI occurred. Such is especially the case when considering that in other taxa, LDGs have existed for the last ~30-40 million years (Manion, 2020). Finally, we also found that colonized sites tend to have been inhabited by members of the Dipsadidae for longer (Appendix 3.2), making it more likely that before the GABI occurred those sites had accumulated more taxa than uncolonized communities.

Importantly, saturation can only occur when stability or equilibrium in species numbers or ecological entities results from niche limitation and not from stochastic mechanisms (Olivares et al., 2018). Supporting this idea, we also recovered that there were no significant differences in the functional dispersion of the sites involved in the GABI and those that were not. Such a result indicates that colonizing lineages did not tend to increase the trait space of most of the communities to where they dispersed. As a result, in those cases, taxa dispersing across the Panamanian Isthmus had to either replace native taxa or use unoccupied portions of the trait space so that they could colonize while keeping it relatively unchanged (Gerhold et al., 2011), with both scenarios being indicative of niche limitation. Despite this, it is also possible that uncolonized communities used to have a reduced functional dispersion, which increased to levels comparable to that of the trait space of colonized sites after the GABI. However, as

discussed above regarding site species richness, we did not find support for the idea that uncolonized communities would have reduced trait spaces than colonized communities or that they have reached comparable levels afterward. Further, if that were the case, one would expect that colonizing lineages represent a considerable portion of the trait space of at least some of the communities of interest. Despite this, we found that the functional dispersion of sites involved in the GABI does not significantly differ whether calculated with or without considering the taxa that colonized them (see Appendix 3.2). In summary, alternative explanations to our findings of a lack of association between site species richness or functional dispersion and whether communities were colonized are unlikely, thus, supporting the idea that these patterns are indeed indicative of community saturation.

Finally, it is also possible that the discrepancy between our results and those of Pinto-Sanchez et al. (2014), who did not find evidence for community saturation in Neotropical frogs after the GABI, may be due to the potential increased dispersal ability of the members of the Dipsadidae as compared to Terrarana frogs. Indeed, frogs are generally considered to have a more limited dispersal ability than other groups of organisms due to their dependency on water sources and sensitivity to saltwater (Warburg, 1997; Wells, 2007). Thus, the Dipsadidae could have dispersed across the Panamanian Isthmus at earlier dates than groups with more limited dispersal ability, such as the Terrarana. By colonizing earlier, interactions between native species and colonizers potentially had more time to unfold, and communities colonized by taxa of increased dispersal abilities (such as the Dipsadidae) are more likely to be saturated as a

result. Despite these expectations, there is no obvious difference in the time in which dipsadid snakes and Terrarana frogs crossed the Panamanian Isthmus (compare the ages reported herein to those reported by Pinto-Sanchez et al., 2014). Instead, the difference in dispersal abilities between these two groups seems to be better reflected in that the lineages of the Dipsadidae involved in the GABI exhibit a much more extensive distribution in the New World than the corresponding groups of the Terrarana. For example, species as *Clelia clelia*, *Oxyrhopus petola*, *Imantodes cenchoa*, *Sibon nebulatus*, as well as the genera *Dipsas* and *Sibon*, are now distributed from Mexico to Argentina; while the hyperdiverse *Atractus* is distributed from Panama to Argentina (Savage, 2002; Kohler, 2009). Contrastingly, the few Terrarana lineages involved in the GABI have only reached Eastern Honduras to the north (*Pristimantis*; Hedges et al., 2008) and Northwestern Ecuador to the south (*Craugastor*; Frost, 2021). In general, the dispersal ability of taxa is commonly considered as being positively related to the extent of their distribution. The reduced dispersal abilities of taxa limit them from expanding their ranges, even to the degree that they may be unable to reach areas having suitable conditions for them to survive (Dullinger et al., 2012; Slatyer et al., 2013). As a result, if no other species perform the ecological roles of dispersal-limited taxa in the communities they fail to colonize, it is unlikely that those communities contain the total number of taxa they could harbor (i.e., they are unsaturated). By contrast, the high dispersal rates of taxa of increased dispersal ability facilitate their interaction with taxa performing similar ecological roles, thus, potentially limiting the local accumulation of diversity (MacArthur & Levins, 1967; Ashby et al., 2019). However, although these predictions agree with the

patterns of community saturation observed in the Terrarana and the Dipsadidae, they are based on assuming that range extent is a proxy for dispersal ability. Although this is an intuitive assumption to make, the results of a recent meta-analysis indicated that range size and the dispersal ability of taxa (measured as the average distance of dispersion of the specimens of a given species) are rarely associated, at least for marine organisms (Lester et al., 2007).

In summary, our finding that communities of dipsadid snakes across the Neotropics are saturated (both in terms of species counts and ecological entities) contrasts with previous studies using biotic introductions to test for community saturation. This discrepancy likely results from other studies not considering the existence of extinction debts or that saturation could be mediated by functional traits. Extinction debts can require hundreds or even thousand years to unfold (Kitzes & Hartze, 2015; Figueiredo et al., 2019), with their extent determined by factors that can be taxon-specific or even fortuitous. Some of these factors include how frequently disturbances occur, the abundance distribution of the species involved, and if they exhibit traits associated with longer timespans (such as having increased longevity or dispersal ability, Kitzes & Hartze, 2015; Figueiredo et al., 2019; Lalechere et al. 2019). As a result, assuming that unsaturation or saturation is the rule at any spatial or temporal scale may be unjustified and premature. Indeed, unsaturation could simply correspond to a temporary unstable stage in which communities or regions stay after suffering from a considerable disturbance before reaching a potential stable stage (i.e., transient stable states, Fukami & Nakajima, 2011). Moving forward, it is important to note that studies testing for

community saturation across large temporal scales can also provide insights into the long-term impacts of modern-day species introductions and other disturbances not possible to be obtained from other methods. Such studies could also allow us to predict how the richness and functional diversity of communities and regions might be influenced by the predicted migrations of taxa caused by climate change and habitat homogenization around the world.

References

- Ashby, B., Shaw, A. K., & Kokko, H. (2020). An inordinate fondness for species with intermediate dispersal abilities. *Oikos*, 129(3), 311-319.
- Arroyo, S.B., Serrano-Cardozo, V.H., Ramírez-Pinilla, M.P. (2008). Diet, microhabitat and time of activity in a *Pristimantis* (Anura, Strabomantidae) assemblage. *Phyllomedusa*, 7: 109-119
- Blackburn, D.G. (1999). Viviparity and oviparity: evolution and reproductive strategies. Pp. 994-1003, in T.E. Knobil and J.D. Neill, eds. *Encyclopedia of Reproduction*. Academic Press, New York
- Bleu, J., Heulin, B., Haussy, C., Meylan, S., & Massot, M. (2012). Experimental evidence of early costs of reproduction in conspecific viviparous and oviparous lizards. *Journal of Evolutionary Biology*, 25(7), 1264-1274.
- Boback, S. M., & Guyer, C. (2003). Empirical evidence for an optimal body size in snakes. *Evolution*, 57(2), 345-451.
- Bouckaert, R. R., & Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, 17(1), 42.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., ... & Drummond, A. J. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *pLoS Computational Biology*, 10(4), e1003537.
- Brown, T. W., van den Burg, M. P., Maryon, D. F., & Arrivillaga, C. (2018). Arboreality and diet in Pacific Long-tailed Snakes, *Enulius flavitorques* (Squamata: Dipsadidae), and a potential adaptive hypothesis for egg attendance in Honduran Leaf-toed Geckos, *Phyllodactylus palmeus* (Squamata: Phyllodactylidae). *Reptiles & Amphibians*, 25(1), 31-34.
- Burbrink, F. T., & Myers, E. A. (2015). Both traits and phylogenetic history influence community structure in snakes over steep environmental gradients. *Ecography*, 38(10), 1036-1048.
- Burbrink, F. T., Chen, X., Myers, E. A., Brandley, M. C., & Pyron, R. A. (2012). Evidence for determinism in species diversification and contingency in phenotypic evolution during adaptive radiation. *Proceedings of the Royal Society B: Biological Sciences*, 279(1748), 4817-4826.

- Cadotte, M. W., Carscadden, K., & Mirotchnick, N. (2011). Beyond species: functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology*, 48(5), 1079-1087.
- Canning-Clode, J., Maloney, K. O., McMahon, S. M., & Wahl, M. (2010). Expanded view of the local–regional richness relationship by incorporating functional richness and time: a large-scale perspective. *Global Ecology and Biogeography*, 19(6), 875-885.
- Chen, I.C., Hill, J.K., Ohlemuller, R., Roy, D.B. & Thomas, C.D. (2011). Rapid range shifts of species associated with high levels of climate warming. *Science*, 333, 1024–1026.
- Cribari-Neto, F. & Zeileis, A (2010). Beta Regression in R. *Journal of Statistical Software*, 34(2), 1–24
- Elton, C. (1950). *The ecology of animals*. Methuen, London, United Kingdom.
- Elton, C. S. (1958). *The ecology of invasions by animals and plants*. London, United Kingdom: Methuen.
- Feldman, A., & Meiri, S. (2012). Length–mass allometry in snakes. *Biological Journal of the Linnean Society*, 108(1), 161-172.
- Feldman, A., Sabath, N., Pyron, R. A., Mayrose, I., & Meiri, S. (2016). Body sizes and diversification rates of lizards, snakes, amphisbaenians and the tuatara. *Global Ecology and Biogeography*, 25(2), 187-197.
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302-4315.
- Figueiredo, L., Krauss, J., Steffan-Dewenter, I., & Sarmiento Cabral, J. (2019). Understanding extinction debts: spatio–temporal scales, mechanisms and a roadmap for future research. *Ecography*, 42(12), 1973-1990.
- Flynn, L. J., Tedford, R. H., & Zhanxiang, Q. (1991). Enrichment and stability in the Pliocene mammalian fauna of North China. *Paleobiology*, 246-265.
- Frost, D. R. (2021). Amphibian Species of the World: An Online Reference. Version 6.1 (accessed 07/17/2021). Electronic Database accessible at <https://amphibiansoftheworld.amnh.org/index.php>. American Museum of Natural History, New York, USA.

- Fukami, T., & Nakajima, M. (2011). Community assembly: alternative stable states or alternative transient states?. *Ecology Letters*, 14(10), 973-984.
- Gerhold, P., Pärtel, M., Tackenberg, O., Hennekens, S. M., Bartish, I., Schaminée, J. H., ... & Prinzing, A. (2011). Phylogenetically poor plant communities receive more alien species, which more easily coexist with natives. *The American Naturalist*, 177(5), 668-680.
- Gilbert, B., & Levine, J. M. (2013). Plant invasions and extinction debts. *Proceedings of the National Academy of Sciences*, 110(5), 1744-1749.
- Godley, J. S., McDiarmid, R. W., & Rojas, N. N. (1984). Estimating prey size and number in crayfish-eating snakes, genus *Regina*. *Herpetologica* 40(1), 82-88.
- Gurevitch, J., & Padilla, D. K. (2004). Are invasive species a major cause of extinctions?. *Trends in Ecology & Evolution*, 19(9), 470-474.
- Hedges, S. B., Duellman, W. E., & Heinicke, M. P. (2008). New World direct-developing frogs (Anura: Terrarana): molecular phylogeny, classification, biogeography, and conservation. *Zootaxa*, 1737(1), 1-182.
- Henderson, R.W., J.R. Dixon, and P. Soini. 1979. Resource partitioning in Amazonian snake communities. *Milwaukee Public Museum, Contributions in Biology and Geology* 22:1-11
- HilleRisLambers, J., Harsch, M.A., Ettinger, A.K., Ford, K.R. & Theobald, E.J. (2013). How will biotic interactions influence climate change-induced range shifts? *Annals of the New York Academy of Sciences*, 1297, 112-125.
- Head, J. J., Mahlow, K., & Müller, J. (2016). Fossil calibration dates for molecular phylogenetic analysis of snakes 2: Caenophidia, Colubroidea, Elapoidea, Colubridae. *Palaeontologia Electronica*, 19 (2FC), 1-21.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high-resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965-1978.
- Katoh, K., & Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, 9(4), 286-298.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772-780.

- Kitzes, J., & Harte, J. (2015). Predicting extinction debt from community patterns. *Ecology*, 96(8), 2127-2136.
- Köhler G. (2008). *Reptiles of Central America*. 2nd edition. Herpeton, Offenback, Germany. 400 p.
- Kraft, N. J., Cornwell, W. K., Webb, C. O., & Ackerly, D. D. (2007). Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *The American Naturalist*, 170(2), 271-283.
- Lalechère, E., Archaux, F., & Jabot, F. (2019). Relative importance of landscape and species characteristics on extinction debt, immigration credit and relaxation time after habitat turnover. *Population Ecology*, 61(4), 383-395.
- Laliberté, E., & Legendre, P. (2010). A distance-based framework for measuring functional diversity from multiple traits. *Ecology*, 91(1), 299-305.
- Laliberté, E., & Shipley, B. (2011). FD: measuring functional diversity from multiple traits, and other tools for functional ecology. R package version 1.0– 12. R Project for Statistical Computing, Vienna, Austria.
- Landis, M. J., Matzke, N. J., Moore, B. R., & Huelsenbeck, J. P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology*, 62(6), 789-804.
- Lester, S. E., Ruttenberg, B. I., Gaines, S. D., & Kinlan, B. P. (2007). The relationship between dispersal ability and geographic range size. *Ecology Letters*, 10(8), 745-758.
- Lieberman, S.S. (1986). Ecology of the leaf litter herpetofauna of a Neotropical rain forest. La Selva, Costa Rica. *Acta Zoologica Mexicana*, 15, 1–72.
- Lindberg, D. R. (1991). Marine biotic interchange between the northern and southern hemispheres. *Paleobiology*, 308-324.
- MacArthur, R. H. (1965). Patterns of species diversity. *Biological Reviews*, 40(4), 510-533.
- MacArthur, R. H., & Wilson, E. O. (1963). An equilibrium theory of insular zoogeography. *Evolution*, 373-387.
- MacArthur, R., & Levins, R. (1967). The limiting similarity, convergence, and divergence of coexisting species. *The American Naturalist*, 101(921), 377-385.

- MacDougall, A.S., Gilbert, B. & Levine, J.M. (2009). Plant invasion and the niche. *Journal of Ecology*, 97, 609–615.
- Mannion, P. D. (2020). A deep-time perspective on the latitudinal diversity gradient. *Proceedings of the National Academy of Sciences*, 117(30), 17479–17481.
- Marshall, L., Webb, S., Sepkoski, J. Jr & Raup, D. (1982). Mammalian evolution and the great American interchange. *Science*, 215, 1351–1357.
- Mason, N. W., Mouillot, D., Lee, W. G., & Wilson, J. B. (2005). Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos*, 111(1), 112–118.
- Matzke, N.J. (2013a) ‘BioGeoBEARS’: biogeography with Bayesian (and likelihood) evolutionary analysis in R Scripts. University of California, Berkeley, CA. R package, version 0.2.1. Available at: www.R-project.org/package=BioGeoBEARS (accessed 21 February 2018).
- Matzke, N. J. (2013b). Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, 5(4).
- Matzke, N.J. (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in Island Clades. *Systematic Biology* 63, 951–970.
- MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton University. Princeton, New Jersey, USA.
- Mlambo, M. C. (2014). Not all traits are ‘functional’: insights from taxonomy and biodiversity-ecosystem functioning research. *Biodiversity and Conservation*, 23(3), 781–790.
- Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J.C., Valencia, V., Ayala, C., Pérez-Angel, L.C., Ramírez, V. & Niño, H. (2015), Middle Miocene closure of the Central American Seaway. *Science* 348, 226–229.
- Mouchet, M. A., Villéger, S., Mason, N. W., & Mouillot, D. (2010). Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. *Functional Ecology*, 24(4), 867–876.
- Myers, J. A., & Harms, K. E. (2009). Seed arrival, ecological filters, and plant species richness: a meta-analysis. *Ecology letters*, 12(11), 1250–1260.

- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673-690.
- O'Dea, A., Lessios, H. A., Coates, A. G., Eytan, R. I., Restrepo-Moreno, S. A., Cione, A. L., ... & Stallard, R. F. (2016). Formation of the Isthmus of Panama. *Science Advances*, 2(8), e1600883.
- Olivares, I., Karger, D. N., & Kessler, M. (2018). Assessing species saturation: conceptual and methodological challenges. *Biological Reviews*, 93(4), 1874-1890.
- Patzkowsky, M. E., & Holland, S. M. (2007). Diversity partitioning of a Late Ordovician marine biotic invasion: controls on diversity in regional ecosystems. *Paleobiology*, 33(2), 295-309.
- Pianka, E. R., Vitt, L. J., Pelegrin, N., Fitzgerald, D. B., & Winemiller, K. O. (2017). Toward a periodic table of niches, or exploring the lizard niche hypervolume. *The American Naturalist*, 190(5), 601-616.
- Pinto-Sánchez, N. R., Crawford, A. J., & Wiens, J. J. (2014). Using historical biogeography to test for community saturation. *Ecology Letters*, 17(9), 1077-1085.
- R Development Core Team (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Ray, J. M., Montgomery, C. E., Mahon, H. K., Savitzky, A. H., & Lips, K. R. (2012). Goo-eaters: diets of the neotropical snakes *Dipsas* and *Sibon* in Central Panama. *Copeia*, 2012(2), 197-202.
- Ricklefs, R. E. (1987). Community diversity: relative roles of local and regional processes. *Science*, 235(4785), 167-171.
- Savage, J.M. (2002) *The Amphibians and Reptiles of Costa Rica*. University of Chicago Press, Chicago, 934 pp.
- Sax, D. F., Gaines, S. D., & Brown, J. H. (2002). Species invasions exceed extinctions on islands worldwide: a comparative study of plants and birds. *The American Naturalist*, 160(6), 766-783.
- Sax, D. F., & Gaines, S. D. (2003). Species diversity: from global decreases to local increases. *Trends in Ecology & Evolution*, 18(11), 561-566.

- Sax, D. F., Stachowicz, J. J., Brown, J. H., Bruno, J. F., Dawson, M. N., Gaines, S. D. & O'Connor, M. I. (2007). Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution*, 22(9), 465-471.
- Schleuter, D., Daufresne, M., Massol, F., & Argillier, C. (2010). A user's guide to functional diversity indices. *Ecological Monographs*, 80(3), 469-484.
- Shine, R. (1994). Allometric patterns in the ecology of Australian snakes. *Copeia*, 851-867.
- Simpson, G. G. (1980) *Splendid isolation: the curious history of South American mammals*. Yale Univ. Press, New Haven, CT, USA.
- Slatyer, R. A., Hirst, M., & Sexton, J. P. (2013). Niche breadth predicts geographical range size: a general ecological pattern. *Ecology Letters*, 16(8), 1104-1114.
- Smith, S.A. & Shurin, J.B. (2006). Room for one more? Evidence of invasibility and saturation in ecological communities. Pages 423-447, in Cadotte, M.W., MacMahon, S.M. and Fukami, T. (eds). *Conceptual ecology and invasion biology: reciprocal approaches to nature*. Springer, Dordrecht, Germany.
- Stachowicz, J. J., & Tilman, D. (2005). Species invasions and the relationships between species diversity, community saturation, and ecosystem functioning. pp. 41-64 In: Sax, D. F., Stachowicz, J. J. and S. D. Gaines (eds). *Species invasions: insights into ecology, evolution, and biogeography*. Sinauer Associates Inc., MA, USA.
- Stehli, F.G. & Webb, S.D. (eds) (1985) *The great American biotic interchange*. Topics in Geobiology. Plenum Press, New York, NY, USA.
- Stohlgren, T. J., Barnett, D. T., Jarnevich, C. S., Flather, C., & Kartesz, J. (2008). The myth of plant species saturation. *Ecology Letters*, 11(4), 313-322.
- Terborgh, J. W., & Faaborg, J. (1980). Saturation of bird communities in the West Indies. *The American Naturalist*, 116(2), 178-195.
- Uetz, P., Freed, P. & Hošek, J. (Eds.) (2023). The Reptile Database, <http://www.reptile-database.org>, accessed 23 January 2023.
- Valdivia, N., Segovia-Rivera, V., Fica, E., Bonta, C. C., Aguilera, M. A., & Broitman, B. R. (2017). Context-dependent functional dispersion across similar ranges of trait space covered by intertidal rocky shore communities. *Ecology and Evolution*, 7(6), 1882-1891.

- Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, 85(2), 183-206.
- Vitt, L. J. (1987). Communities. pp. 335-365 in R. A. Seigel, J. T. Collins, and S. S. Novak (eds.), *Snakes: Ecology and Evolutionary Biology*. McGraw- Hill, New York.
- Vermeij, G. J. (1991). When biotas meet: understanding biotic interchange. *Science*, 253(5024), 1099-1104.
- Villéger, S., Mason, N. W., & Mouillot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, 89(8), 2290-2301.
- Warburg, M. R. (1997). *Ecophysiology of amphibians inhabiting xeric environments*. Springer: Berlin, Germany
- Wells, K. D. (2007). *The ecology and behavior of amphibians*. The University of Chicago: Chicago, IL, USA.

Tables

Table 3.1. Results of the generalized linear model performed between the total species richness of each of the 67 communities of snakes of the family Dipsadidae included in this study, and the number of species dispersing across the Panamanian Isthmus they contain, incorporating the effects of several confounding variables. The variables indicated as bold are those recovered as being significant predictors of the dependent variable of the model. Parameters of the model are as follows: Null deviance: 151.998 on 41 degrees of freedom; Residual deviance: 70.954 on 31 degrees of freedom; Dispersion parameter: 2.297.

Variable	Estimate	Standard Error	z-values	P-value
Functional dispersion of non-dipsadid snakes	-0.170	2.230	-0.076	0.940
Presence of colonizing lineages	0.234	0.216	1.086	0.286
Environmental PC2	-0.056	0.041	-1.366	0.182
Environmental PC3	-0.021	0.112	-0.186	0.854
Environmental PC4	0.100	0.115	0.873	0.389
Total sampling effort	0.000	0.000	1.410	0.168
Species richness of non-dipsadid snakes	0.053	0.014	3.765	6.98E-04
Trap use	0.192	0.198	0.969	0.340
Time of first colonization	-0.035	0.063	-0.553	0.584
Elevation	-0.001	0.000	-1.913	0.065

Table 3.2. Results of the Beta regression performed between the total functional dispersion of each of the 67 communities of snakes of the family Dipsadidae included in this study and the number of species dispersing across the Panamanian Isthmus they contain, incorporating the effects of several confounding variables. The variables indicated as bold are those recovered as being significant predictors of the dependent variable of the Beta regression. Parameters of the model are as follows: Log-likelihood: 90.75 on 12 Degrees of Freedom; Pseudo R-squared: 0.256.

Variable	Estimate	Standard Error	z-values	P value
Presence of colonizing lineages	-0.068	0.075	-0.914	0.361
Functional dispersion of non-dipsadid snakes	-0.266	0.768	-0.346	0.730
Environmental PC2	0.015	0.015	0.978	0.328
Environmental PC3	0.015	0.042	0.370	0.712
Environmental PC4	-0.013	0.040	-0.337	0.736
Total sampling effort	0.000	0.000	0.462	0.644
Species richness of non-dipsadid snakes	0.008	0.006	1.378	0.168
Time of first colonization	0.041	0.023	1.797	0.072
Trap use	-0.052	0.068	-0.756	0.450
Elevation	0.000	0.000	-0.108	0.914

Figures

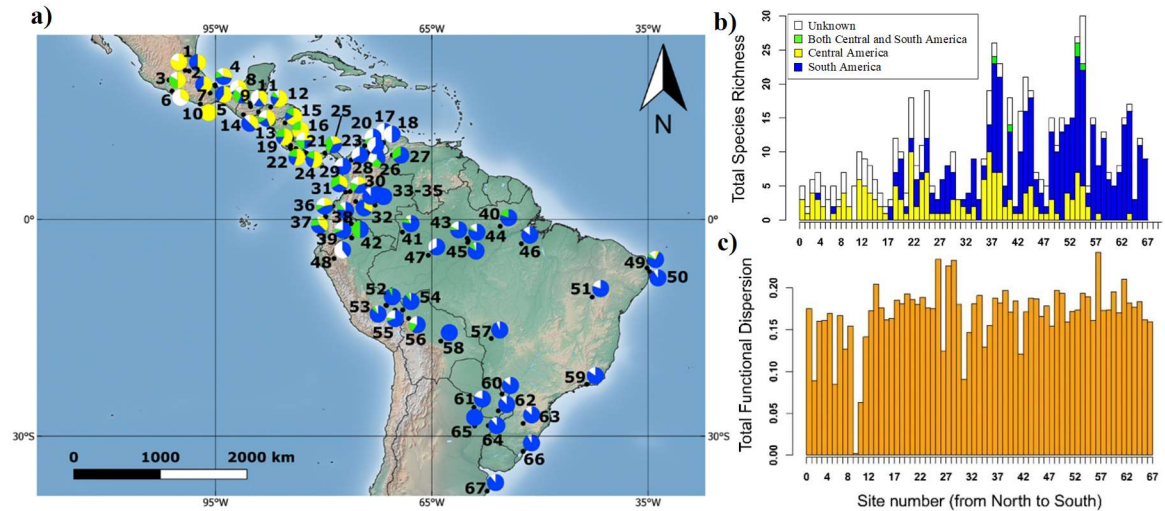


Figure 3.1. a) Map showing the location of the 67 communities of snakes of the family Dipsadidae analyzed in this study. Each site is represented with a circle, divided into yellow, blue, or green portions according to the proportion of taxa recovered by the analyses using BioGeoBEARS as having a Central American, South American, or both Central and South American ancestry, respectively. White portions of the circles indicate the proportion of species on each site from where there is not enough data to determine their continent of origin. b) Total species richness per site, separated into the geographical origin of each species, as recovered by BioGeoBEARS. The colors of the histogram bars are as in the pie charts. c). Total functional dispersion per site. Each number corresponds to one of the localities referred to in Appendix Table A3.1.

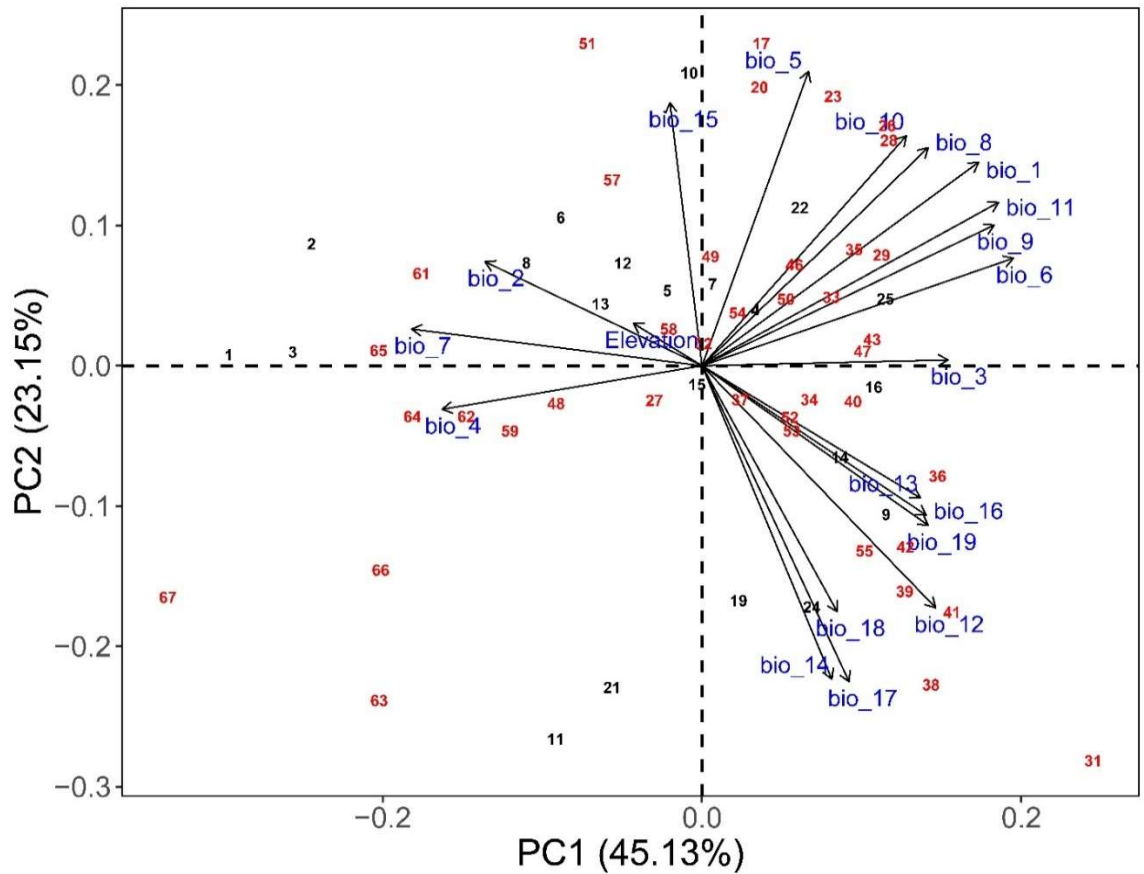


Figure 3.2. Principal Component Analysis (PCA) of the 19 WorldClim climatic variables (plus site elevation) of a series of communities of snakes of the family Dipsadidae across the Neotropics. Each community is indicated with a number, following the code used in Appendix Table A3.1. Numbers are colored (i.e., red or black) according to the continent where the corresponding communities are located (i.e., either Central or South America). Arrows correspond to the 19 WorldClim variables used or to site elevation.

CONCLUSIONS OF THE DISSERTATION

It is increasingly recognized that ecological and evolutionary factors do not necessarily act at different spatial and temporal scales, as has been commonly thought. Indeed, evolutionary processes can have a considerable role even at small spatial and temporal scales, and the same is true for ecological processes at larger scales. As a result of this recognition, ecological studies have incorporated phylogenetic information and vice versa for the last two decades. Indeed, incorporating phylogenetic information is leading to more insightful evaluations of the effects that historical events (such as biogeographic events, diversification shifts, and trait or niche evolution) have on current communities or regions. By extensively using multiple sources of phylogenetic and ecological data, my dissertation follows this line of inquiry to investigate fundamental questions at the interface between ecology and evolution that still remain relatively unexplored. Specifically, I tested for community saturation and a potential direct influence of regional climate or evolutionary time on regional species diversity in a group of neotropical snakes (family Dipsadidae) characterized by its elevated levels of diversity and extensive distribution. However, given that accuracy of such types of studies is dependent on having a robust and well-supported phylogeny, for Chapter 1 of my dissertation, I first explored the causes of phylogenetic uncertainty in the relationships within one of the two main groups of the Dipsadidae (=the Xenodontinae). Despite the application of genomic scale data, my results indicate that most deep relationships within this group are still best considered unresolved. A satisfactory and well-supported

resolution of the deeper phylogenetic relationships of the group could not be obtained even by including thousands of loci and a careful inspection of the factors potentially driving species tree-gene tree discordance. However, this study allowed me to identify that the observed levels of gene tree discordance in the group are most likely the result of a mixture of reduced phylogenetic information in the genomic data currently available and extensive Incomplete Lineage Sorting (ILS). My results demonstrate that increasing the number of loci analyzed does not guarantee the strong resolution of problematic phylogenetic relationships. Reaching such a goal may require a simultaneous increase of taxonomic and genomic sampling (e.g., whole genome sequencing).

For the remaining chapters of my dissertation, I generated and used the most taxonomically complete Bayesian calibrated timetree of the Dipsadidae based on all Sanger loci available to 2022 since it was more relevant to have an extensive taxonomic sampling. For Chapter 2, I investigated the regional species richness of the Dipsadidae and its origins, finding that regional species richness is not determined either by evolutionary time or a direct or indirect role of climate. Instead, regional richness gradients result from spatial variation in species diversification rates of unknown causes, which need to be explored in future studies. My results support the hypothesis proposed in the literature that diversification rates (but not evolutionary time) are more likely to explain richness patterns in more species-rich lineages. However, our ability to compare our findings to previous studies is limited by the considerable variation of methodologies used and the reduced importance given to evaluating the potential roles of evolutionary time and a direct influence of climate. In this regard, a detailed literature review or

metanalysis is badly needed to account for these limitations and identify potential generalities about the drivers of richness gradients across multiple regions or organisms. Finally, for Chapter 3, I explored the potential influence that a large-scale biogeographic event (the Great American Biotic Interchange, GABI) could have left on the current species richness and functional trait diversity of a series of communities of dipsadid snakes across the neotropics. I did not recover evidence that these species richness and functional diversity differed between communities involved or not in the GABI, indicating that colonization events following this geological event most likely involved the exclusion of native taxa. Thus, these communities are best considered saturated with species, a result that conflicts with the findings of previous studies testing for this possibility at both large and small spatial scales. Such a conflict is likely due to these previous studies not considering that community saturation can also occur at the functional diversity level or not recognizing that the biotic interactions determining if colonization events are successful may require a long time to unfold, respectively.

In summary, my dissertation provides new insights into the phylogenetics of the most diverse snake group in the Neotropics (the Dipsadidae) obtained by carefully inspecting and reanalyzing almost all previously published sequence data. I reconstructed phylogenies of the group that optimize either the taxonomic or genetic sampling, and although some relationships remain largely resolved, the resolution of others is becoming well established and corroborated. Future studies where a compromise can be reached between having more or better quality phylogenetic data and an increased number of lineages sampled are necessary. Doing so would corroborate my findings and those of

previous studies and will provide new insights into the phylogenetic relationships within the Dipsadidae that have remained recalcitrant to be resolved. The research I performed for my dissertation also integrates this updated phylogenetic information with multiple subdisciplines of ecology and evolution to better understand what drives the diversity patterns in the Dipsadidae across multiple spatial scales. The methodologies I followed and expanded upon to obtain these results have great potential to be implemented across many other lineages and regions to identify or corroborate the existence of general patterns about the main drivers of biodiversity gradients. Further, implementing and expanding upon this framework can provide further insights into how these biodiversity gradients could be affected by human activities and to better direct future conservation efforts to preserve them.

Appendices

Appendices Chapter 1

Tables

Appendix Table A1.1. Calibration points selected for obtaining the timetree used to calculate the values of phylogenetic informativeness and the time in which that metric reaches its maximum value.

	Calibration point	Minimum age (Ma)	Maximum age (Ma)	Fossil	References	Comments
1	stem Caenophidia (= crown Pan-Caenophidia)	66	73.75	<i>Krebsophis thobanus</i>	Rage & Werner (1999); Head et al. (2016)	Primary fossil calibration
2	stem Colubroidea (= stem Colubridiformes)	50.5	66	<i>Procerophis sahnii</i>	Rage et al. (2008) & Head et al. (2016)	Primary fossil calibration
3	MRCA of Elapidae+Colubridae+Natricidae+Dipsadidae	36.57	43.14	NA	NA	Secondary calibration, after Ramirez & Reeder (in prep.)
4	MRCA of Dipsadinae+Xenodontinae	26.52	29.87	NA	NA	Secondary calibration, after Ramirez & Reeder (in prep.)

Appendix Table A1.2. Voucher information and summary of the amount of DNA sequence data available for the snake species being included in this study. Museum numbers correspond to the catalog number of the museum where the samples were deposited, as follows: CHUNB (Coleção Herpetológica da Universidade de Brasília), GRCOLLI (Guarino Colli field number), MVZ-HERP (Herpetological Collection, Museum of Vertebrate Zoology, University of California Berkeley), UMFS (University of Michigan field series), UMMZ (University of Michigan Museum of Zoology). Taxa in bold were misidentified by Singhal et al. (2017); see Singhal et al. (2022). References are coded as follows: 1) Singhal et al. (2017); 2) Singhal et al. (2020); 3) Singhal et al. (2022).

Family: Subfamily	Tribe	Species	Museum number	Number of loci	Average loci length (bp)	SRA accession number	Reference
Boidae: NA	NA	<i>Boa constrictor</i>	UMFS_9997	4965	581.8	SAMN15718150	2
Acrochordidae : NA	NA	<i>Acrochordus granulatus</i>	UMFS_11534	4389	444.1	SAMN15718164	2
Pareidae: NA	NA	<i>Aplopeltura boa</i>	UMMZ_201904	4921	587.5	SAMN15718167	2
Pareidae: NA	NA	<i>Pareas nuchalis</i>	FK_2626	5008	595.3	SAMN15718165	2
Viperidae: NA	NA	<i>Causus maculatus</i>	MVZ_Herp_249821	2301	296.4	SAMN09221351	2
Homalopsidae : NA	NA	<i>Homalopsis buccata</i>	UMMZ_227683	4808	573.7	SAMN15718173	2
Elapidae: NA	NA	<i>Calliophis bivirgatus</i>	UMFS_10680	4996	595.8	SAMN15718157	2
Dipsadidae: Dipsadinae	Dipsadini	<i>Dipsas mikanii</i>	CHUNB_52075	4714	600.2	SAMN06705338	1, 3
Dipsadidae: Dipsadinae	Dipsadini	<i>Sibon annulatus</i>	MVZ_Herp_269290	2037	260.3	SAMN09221378	2
Dipsadidae: Dipsadinae	Imantodini	<i>Imantodes cenchoa</i>	CHUNB_52156	4546	589.19	SAMN06705356	1
Dipsadidae: Dipsadinae	Imantodini	<i>Leptodeira annulata</i>	GRCOLLI_26789	4536	580.1	SAMN06705371	1, 3

Dipsadidae: Xenodontinae	Pseudoboini	<i>Oxyrhopus trigeminus</i>	CHUNB_3213 3	4748	591.9	SAMN067053 20	1, 3
Dipsadidae: Xenodontinae	Pseudoboini	<i>Pseudoboa_ni gra</i> (not <i>P. neuwiedii</i>)	CHUNB_521 34	4749	577.3	SAMN067053 47	1, 3
Dipsadidae: Xenodontinae	Pseudoboini	<i>Pseudoboa nigra</i>	CHUNB_5215 0	4655	518.7	SAMN067053 53	1, 3
Dipsadidae: Xenodontinae	Pseudoboini	<i>Phimophis guerini</i>	CHUNB_5215 1	3967	600.8	SAMN067053 54	1, 3
Dipsadidae: Xenodontinae	Pseudoboini	<i>Oxyrhopus_pe tolarius</i>	GRCOLLI_26 733	4776	600.5	SAMN067053 69	1, 3
Dipsadidae: Xenodontinae	Pseudoboini	<i>Pseudoboa nigra</i>	CHUNB_5214 9	848	87.66	SAMN067053 52	1, 3
Dipsadidae: Xenodontinae	Psomophini	<i>Psomophis joberti</i>	CHUNB_5207 1	4780	596.8	SAMN067053 36	1, 3
Dipsadidae: Xenodontinae	Xenodontinae	<i>Erythrolampru s almadensis</i>	CHUNB_5205 6	4521	503.4	SAMN067053 30	1
Dipsadidae: Xenodontinae	Xenodontinae	<i>Xenodon_merr emii</i>	CHUNB_5214 4	4704	599.7	SAMN067053 49	1, 3
Dipsadidae: Xenodontinae	Xenodontinae	<i>Lygophis_pau cicens</i>	GRCOLLI_26 916	1843	586.7	SAMN067053 73	1, 3
Dipsadidae: Xenodontinae	Xenodontinae	<i>Erythrolampru s poecilogyrus</i>	CHUNB_5205 9	4065	391.1	SAMN067053 31	1, 3
Dipsadidae: Xenodontinae	Xenodontinae	<i>Erythrolampru s reginae</i>	CHUNB_5212 0	3384	199.1	SAMN067053 42	1, 3
Dipsadidae: Xenodontinae	Philodryadini	<i>Philodryas olfersii</i>	CHUNB_5211 6	4876	606.9	SAMN067053 40	1, 3
Dipsadidae: Xenodontinae	Philodryadini	<i>Philodryas nattereri</i>	CHUNB_5215 4	4816	606.4	SAMN067053 55	1, 3
Dipsadidae: Xenodontinae	Philodryadini	<i>Philodryas nattereri</i> (not <i>Thamnodynast es hypoconia</i>)	CHUNB_520 73	4843	606.9	SAMN067053 37	1, 3

Dipsadidae: Xenodontinae	Echinantherini	<i>Taeniophallus occipitalis</i>	GRCOLLI_26 711	4777	605.5	SAMN067053 68	1, 3
Dipsadidae: Xenodontinae	Elapomorphini	<i>Apostolepis cearensis</i>	CHUNB_5206 0	3760	464.6	SAMN067053 32	1
Dipsadidae: Xenodontinae	Elapomorphini	<i>Apostolepis polylepis</i>	CHUNB_3206 1	3478	418.3	SAMN092211 85	1

Appendix Table A1.3. Summary of the Principal Component (PC) Axes of the PC analysis performed with the 13 traits of the dataset of 4333 genes from the Squamate Conserved Loci set with 75% or more taxonomic completeness used in this study.

	Standard deviation	Proportion of Variance	Cumulative Proportion
PC1	1.588	0.194	0.194
PC2	1.290	0.128	0.322
PC3	1.186	0.108	0.430
PC4	1.093	0.092	0.522
PC5	1.079	0.090	0.612
PC6	1.016	0.079	0.691
PC7	0.970	0.072	0.763
PC8	0.949	0.069	0.833
PC9	0.893	0.061	0.894
PC10	0.820	0.052	0.946
PC11	0.704	0.038	0.984
PC12	0.456	0.016	1.000
PC13	3.57E-15	0.000	1.000

Appendix Table A1.4. First five Loadings of the Principal Component (PC) Axes of the PC analysis performed with the 13 traits of the dataset of 4333 genes from the Squamate Conserved Loci set with 75% or more taxonomic completeness used in this study.

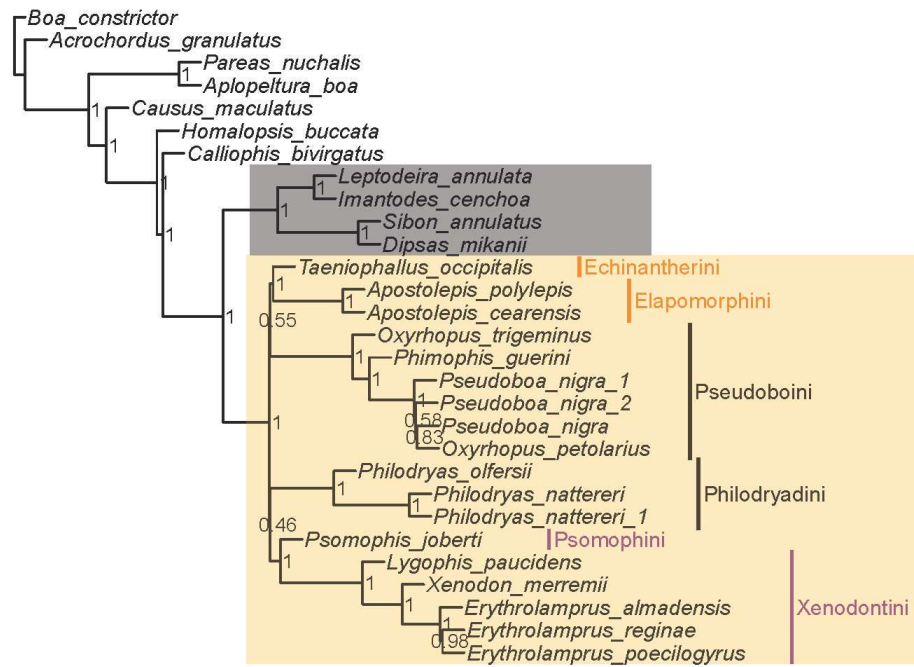
	Principal Component Axis				
	1	2	3	4	5
GC content	-0.23	0.15	-0.11	-0.11	0.22
Length alignment	-0.30	-0.11	-0.56	-0.13	0.02
Missing data proportion	0.15	0.08	0.15	-0.29	0.08
Occupancy	-0.25	-0.15	-0.38	-0.09	0.52
Relative composition frequency variability (RFCV)	0.02	0.14	0.26	-0.64	0.06
Saturation C value	0.07	-0.08	0.03	0.61	0.06
Maximum phylogenetic informativeness (PI)	0.51	0.24	-0.39	-0.04	-0.07
Maximum PI time period	0.51	0.24	-0.39	-0.04	-0.07
Branch outliers	0.19	-0.01	-0.03	0.06	0.66
Mean residuals	-0.02	0.15	0.05	-0.16	0.09
Root-tip variance	-0.11	0.58	0.13	0.22	0.23
Tree length	-0.25	0.64	0.04	0.12	0.01
Average UF bootstrap support	-0.36	0.18	-0.33	-0.03	-0.40

References used for constructing Appendix Table A1.2

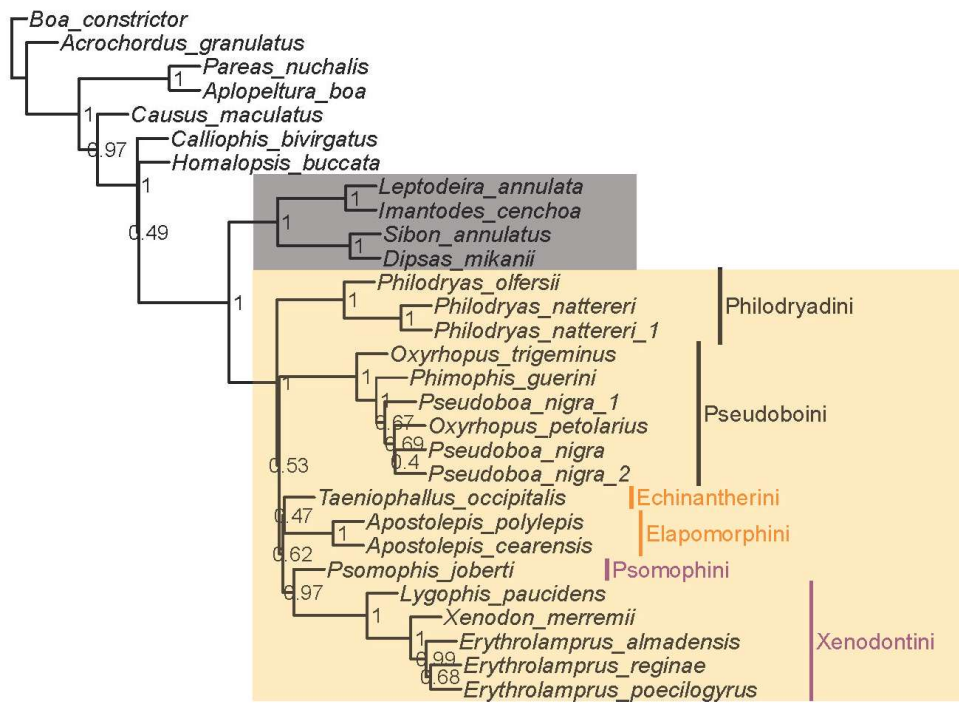
- Head, J. J., Mahlow, K., & Mueller, J. (2016). Fossil calibration dates for molecular phylogenetic analysis of snakes 2: Caenophidia, Colubroidea, Elapoidea, Colubridae. *Palaeontologia Electronica*, 19(2FC), 1-21.
- Rage, J.-C. and Werner, C. 1999. Mid-Cretaceous (Cenomanian) snakes from Wadi Abu Hashim, Sudan: The earliest snake assemblage. *Palaeontologia Africana*, 35:85-110
- Rage, J.-C., Folie, A., Rana, R S., Singh, H., Rose, K.D. and Smith, T. 2008. A diverse snake fauna from the early Eocene of Vastan Lignite Mine, Gujarat, India. *Acta Palaeontologica Polonica*, 53:391-403.
- Ramirez, J. & Reeder, T.W. (in preparation). Climate-driven regional diversification drives spatial variation in species richness in the most diverse family of Neotropical snakes (Colubroidea: Dipsadidae).

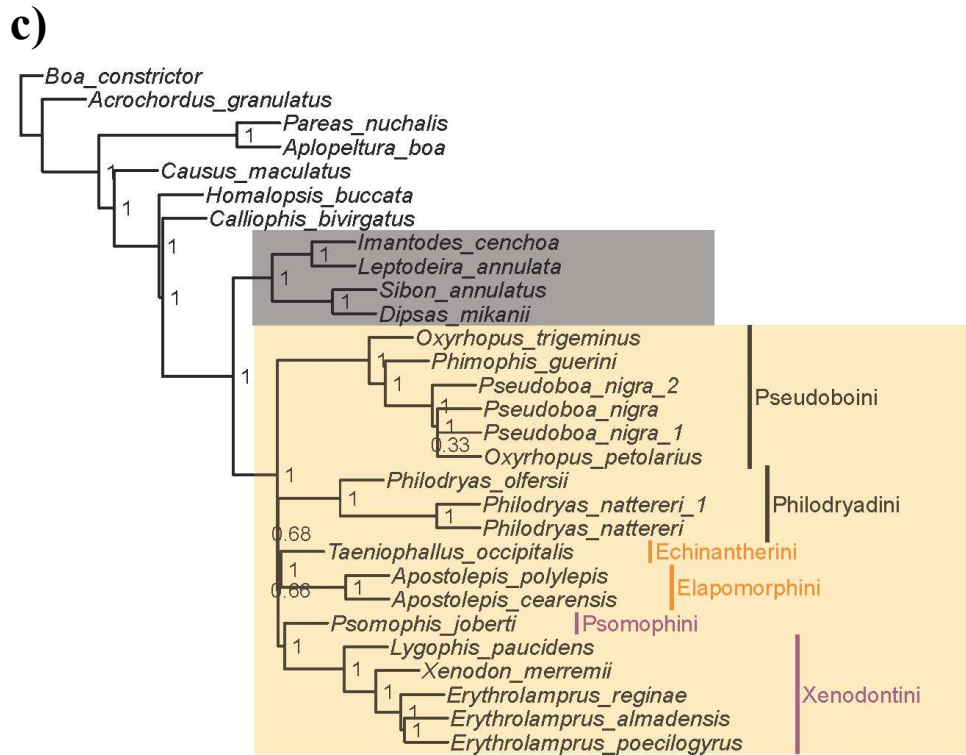
Figures

a)



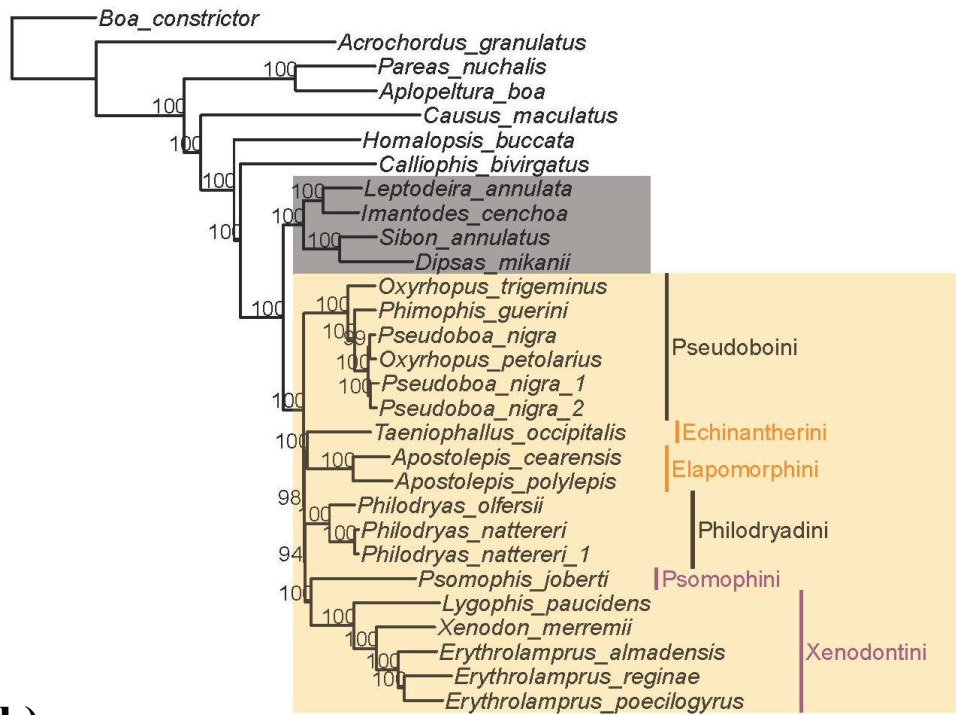
b)



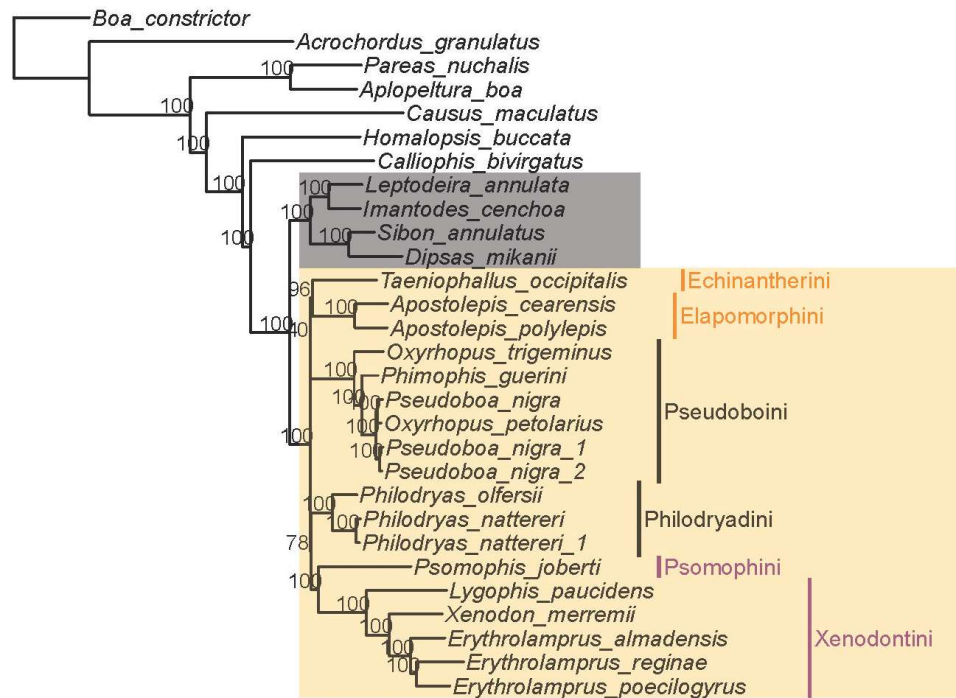


Appendix Figure A1.1. Phylogenetic relationships of the snake family Dipsadidae, as inferred with coalescent-based species tree analyses performed with the software ASTRAL-III. For these analyses, only the following subset of loci were included: a) Anchored Hybrid Enrichment (AHE), b) traditional Sanger sequencing, or c) Ultraconserved Elements (UCE) loci from the Squamate Conserved Loci dataset with 75% or more taxonomic completeness. Numbers on each node correspond to the local posterior probabilities of the clades they subtend. Color codes correspond to the same as in Figure 1.2.

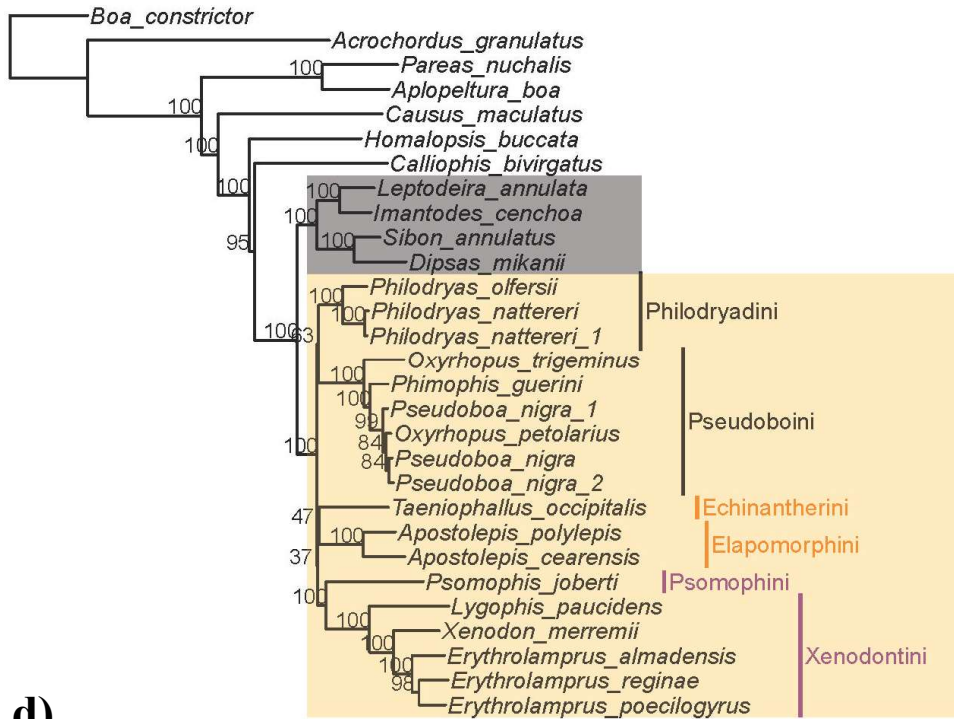
a)



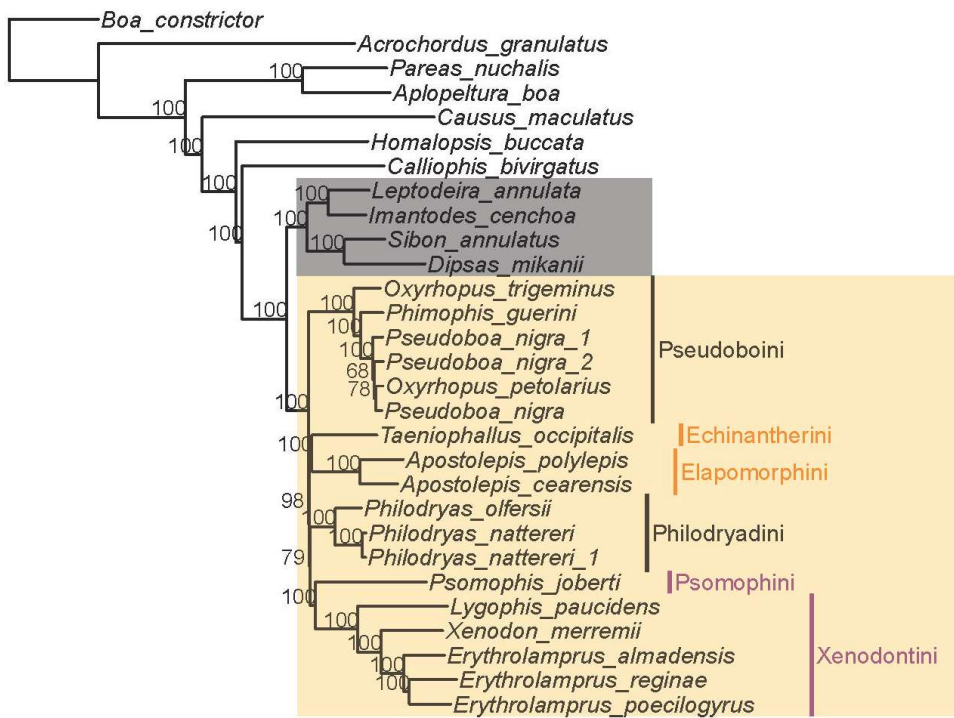
b)



c)

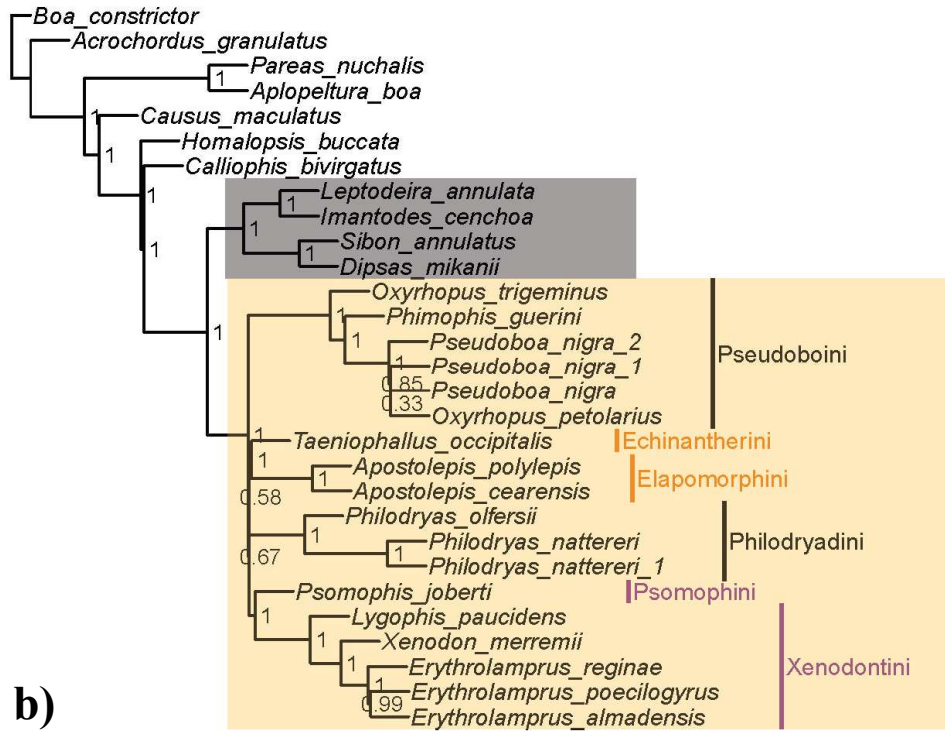


d)

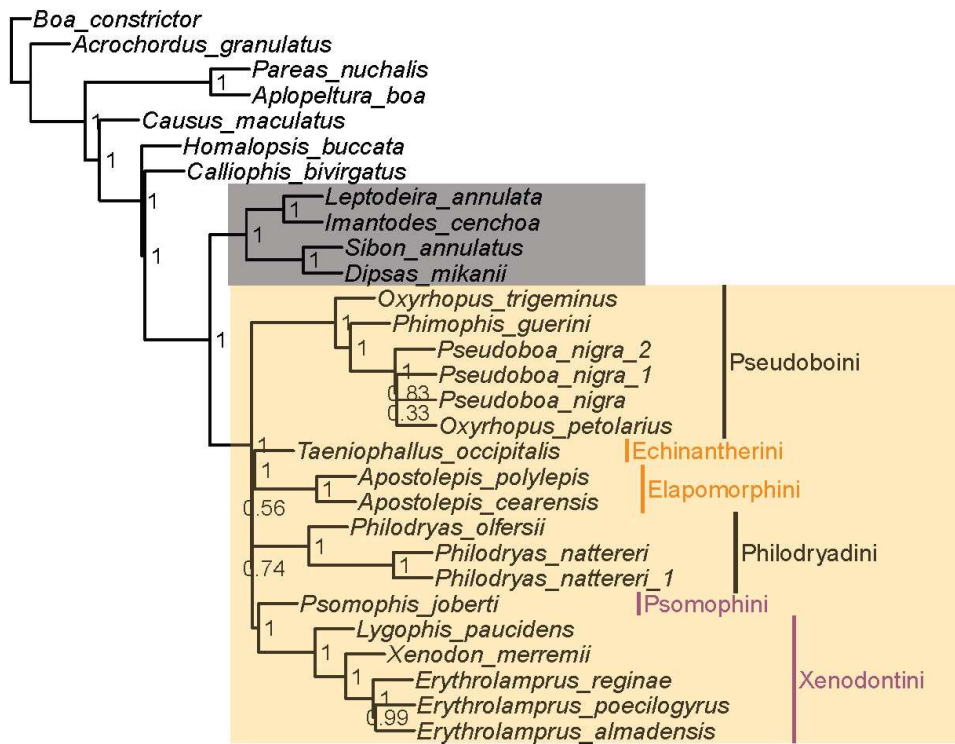


Appendix Figure A1.2. Phylogenetic relationships between the members of the Dipsadidae as inferred with concatenation-based species tree analyses performed with the software IQ-TREE 2. For these analyses, the following sets of loci were used: a) all the 4333 loci from the Squamate Conserved Loci set with 75% or more taxonomic completeness or only the b) Anchored Hybrid Enrichment (AHE), c) traditional Sanger sequencing or d) Ultraconserved Elements (UCE) loci from that dataset. Numbers on each node correspond to the local posterior probabilities of the clades they subtend. Color codes correspond to the same as in Figure 2.2.

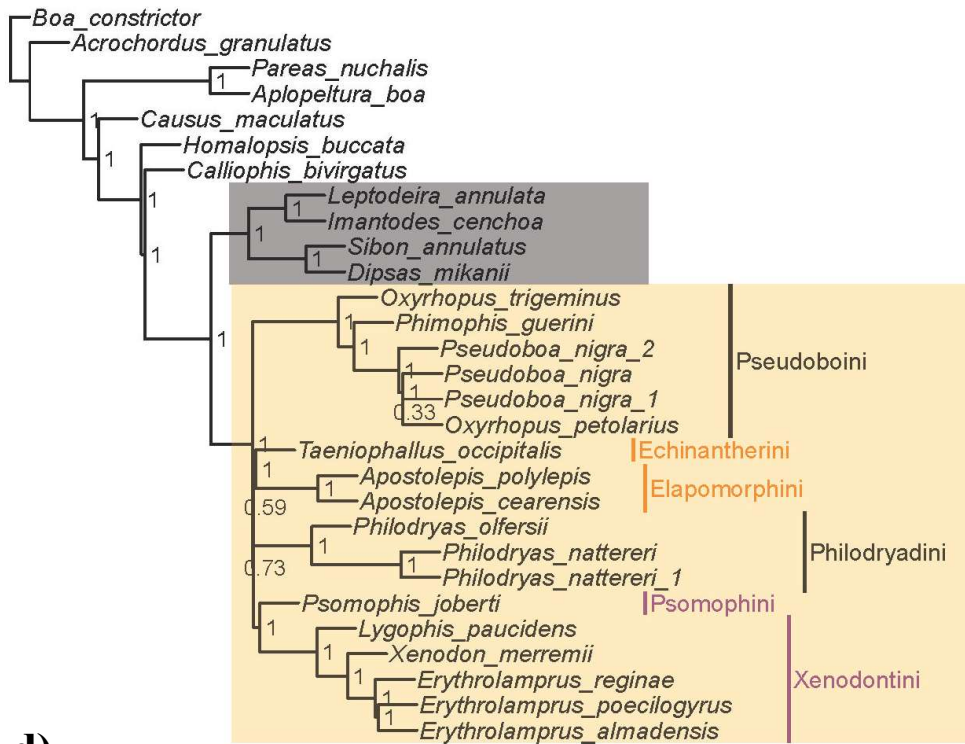
a)



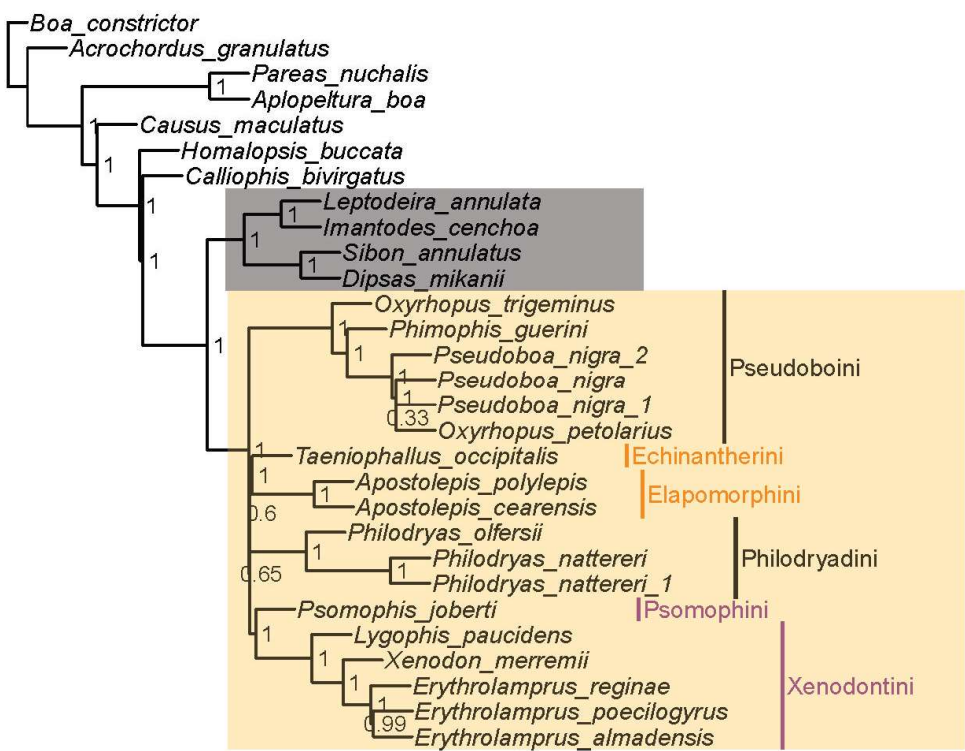
b)



c)

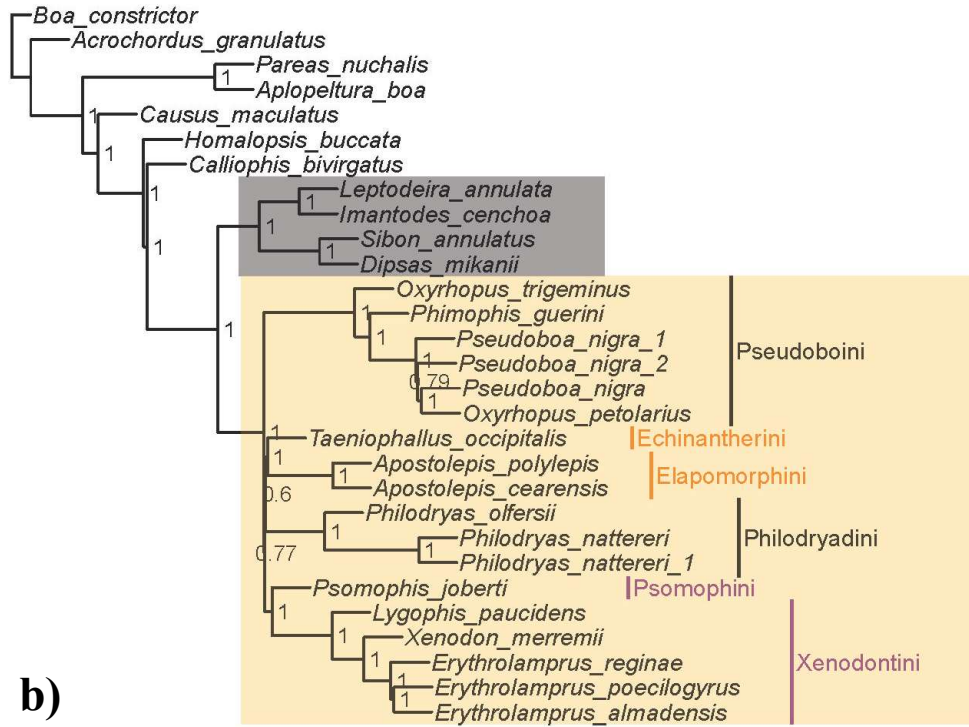


d)

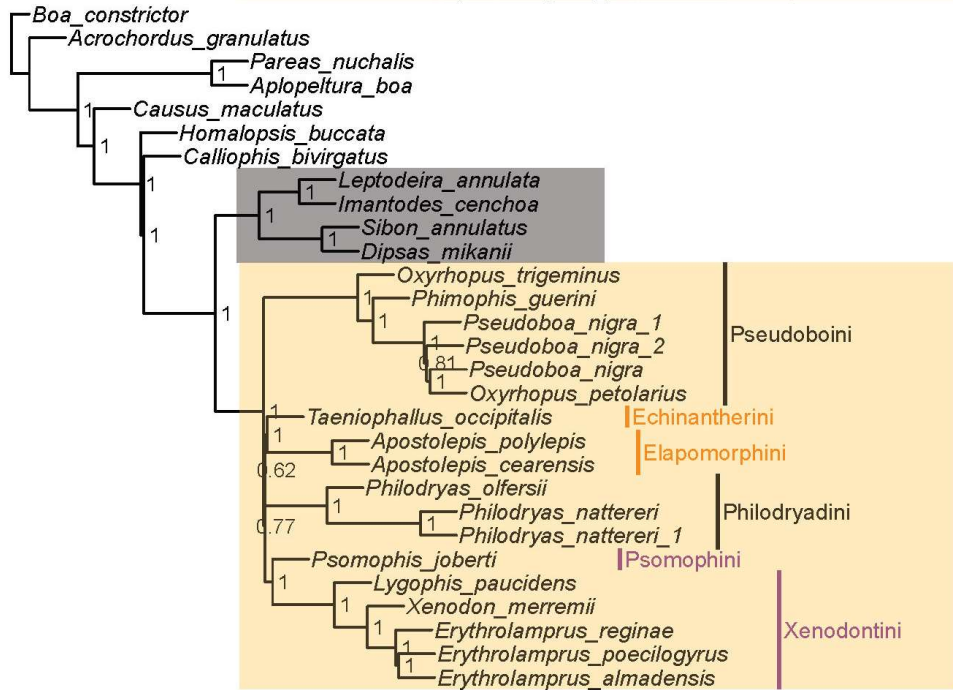


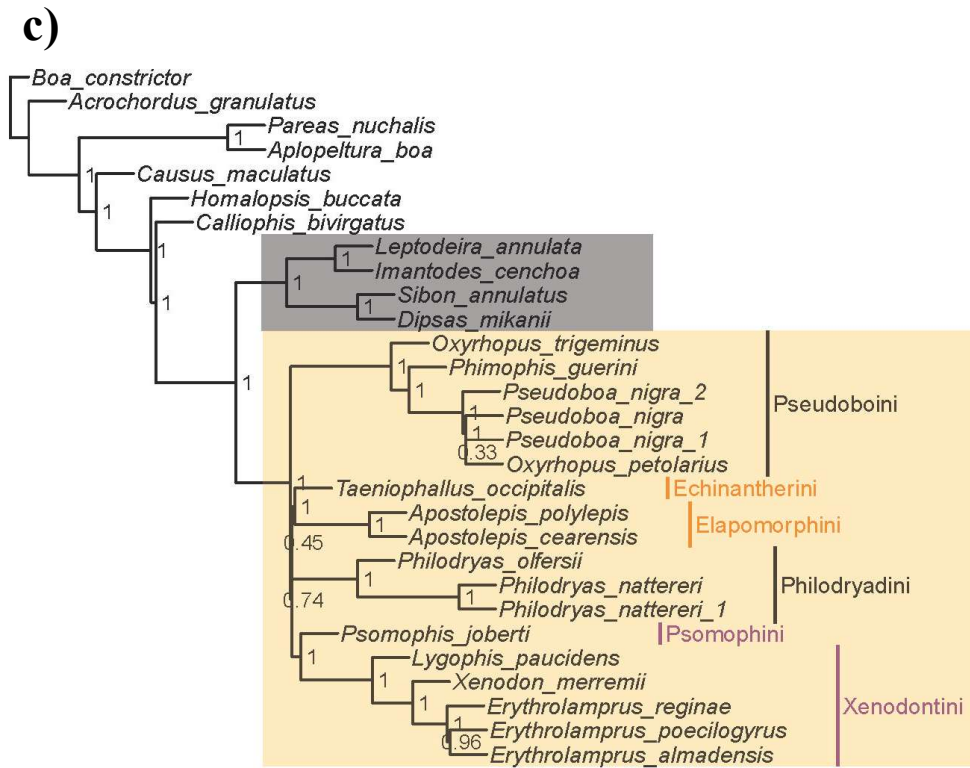
Appendix Figure A1.3. Phylogenetic relationships of the snake family Dipsadidae, as inferred with coalescent-based species tree analyses performed with the software ASTRAL-III. For these analyses, the loci from the Squamate Conserved Loci set were used for the following percentages of taxonomic completeness: a) 25, b) 50, c) 75, or d) 93. Color codes correspond to the same as in Figure 2.2.

a)



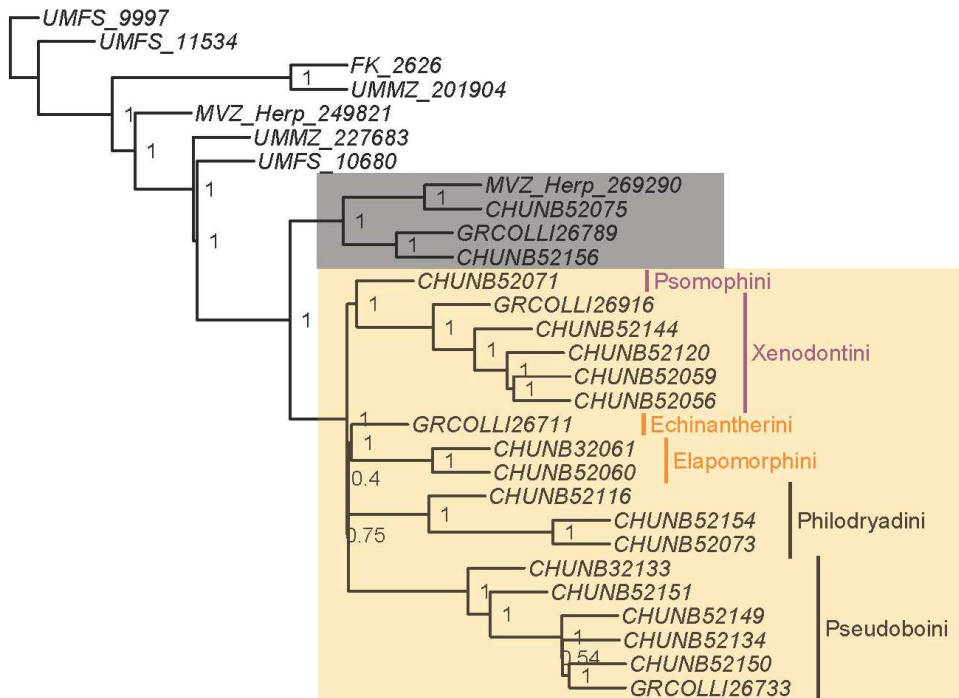
b)



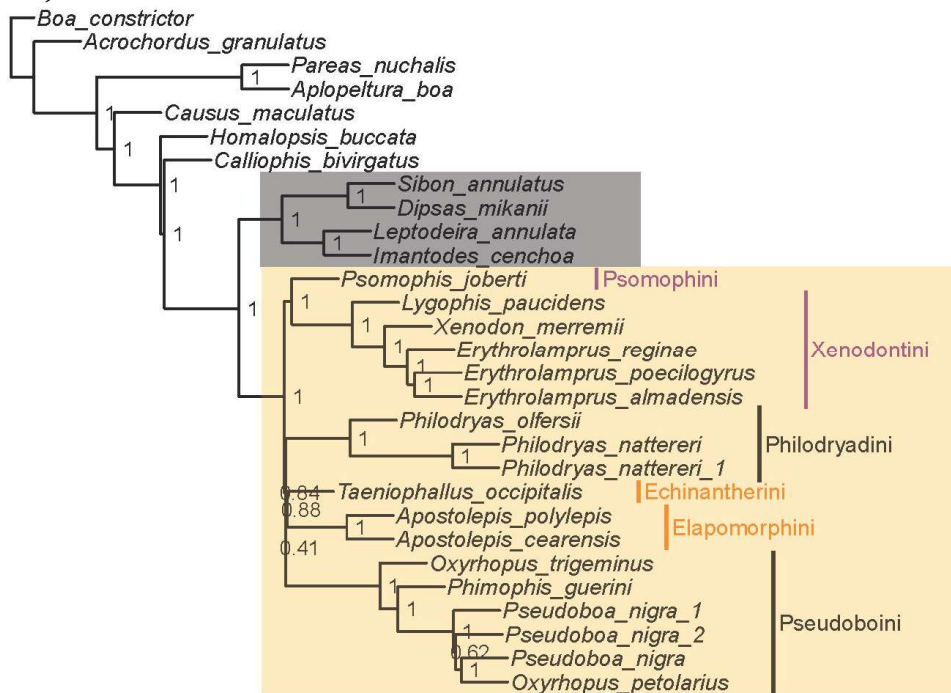


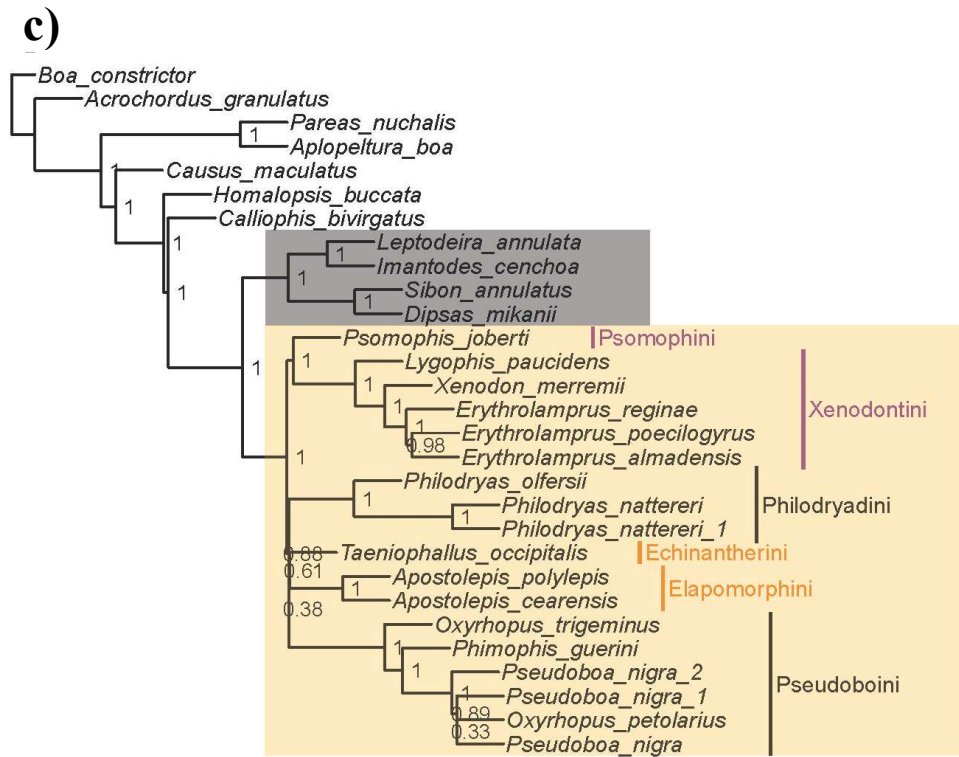
Appendix Figure A1.4. Phylogenetic relationships of the snake family Dipsadidae, as inferred with coalescent-based species tree analyses performed with the software ASTRAL-III. For these analyses, the dataset of 4333 loci from the Squamate Conserved Loci set were used with 75% or more taxonomic completeness after filtering those genes having alignment lengths larger than: a) 474, b) 572, and c) 665 base pairs. These values correspond to the four quantiles of alignment length. Color codes correspond to the same as in Figure 2.2.

a)



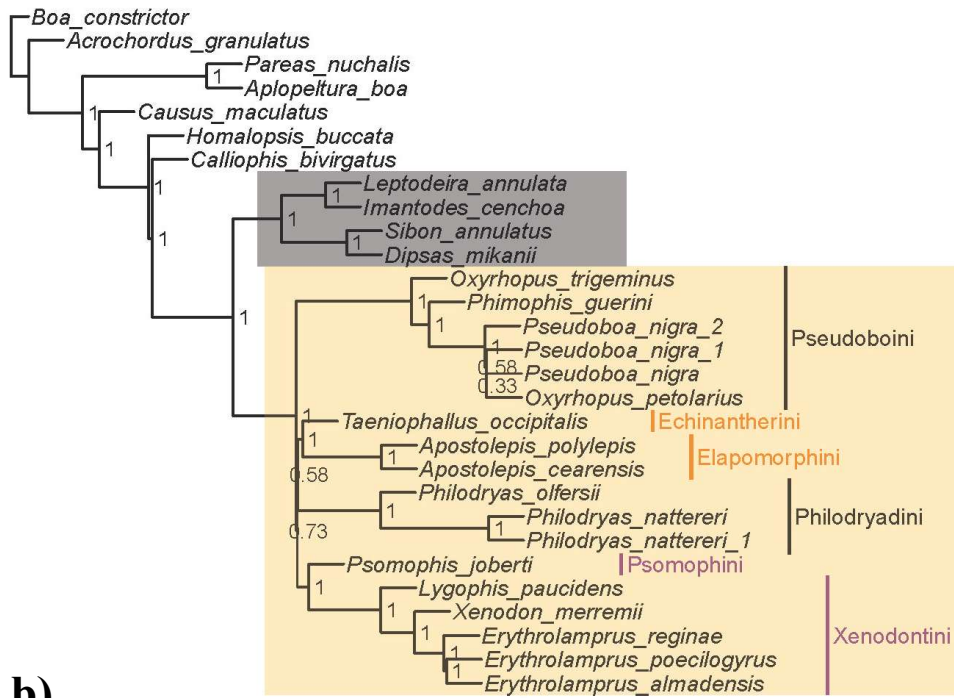
b)



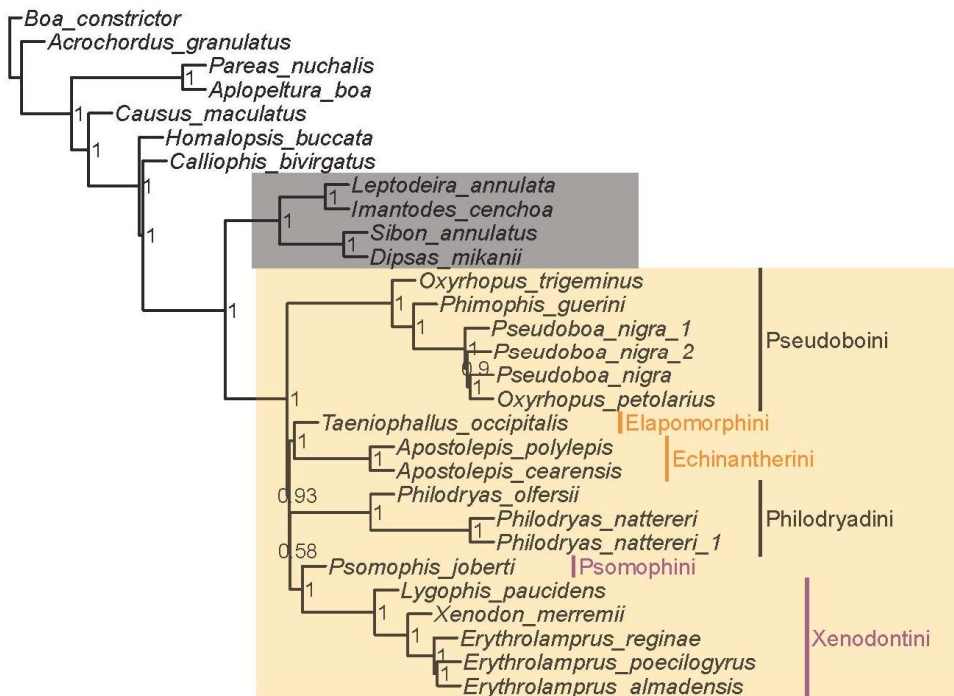


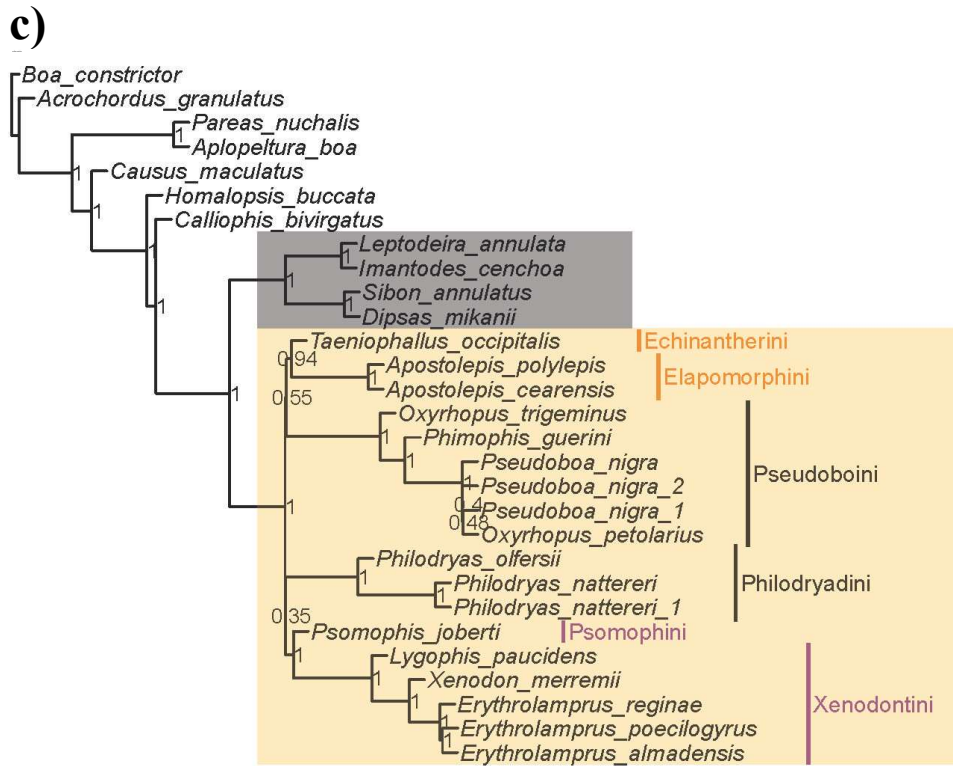
Appendix Figure A1.5. Phylogenetic relationships of the snake family Dipsadidae, as inferred with coalescent-based species tree analyses performed with the software ASTRAL-III. For these analyses, the dataset of 4333 loci from the Squamate Conserved Loci set with 75% or more taxonomic completeness was used after filtering those genes having less or equal than: a) 25, b) 50, c) 75 percentages of overall missing data. Color codes correspond to the same as in Figure 2.2.

a)

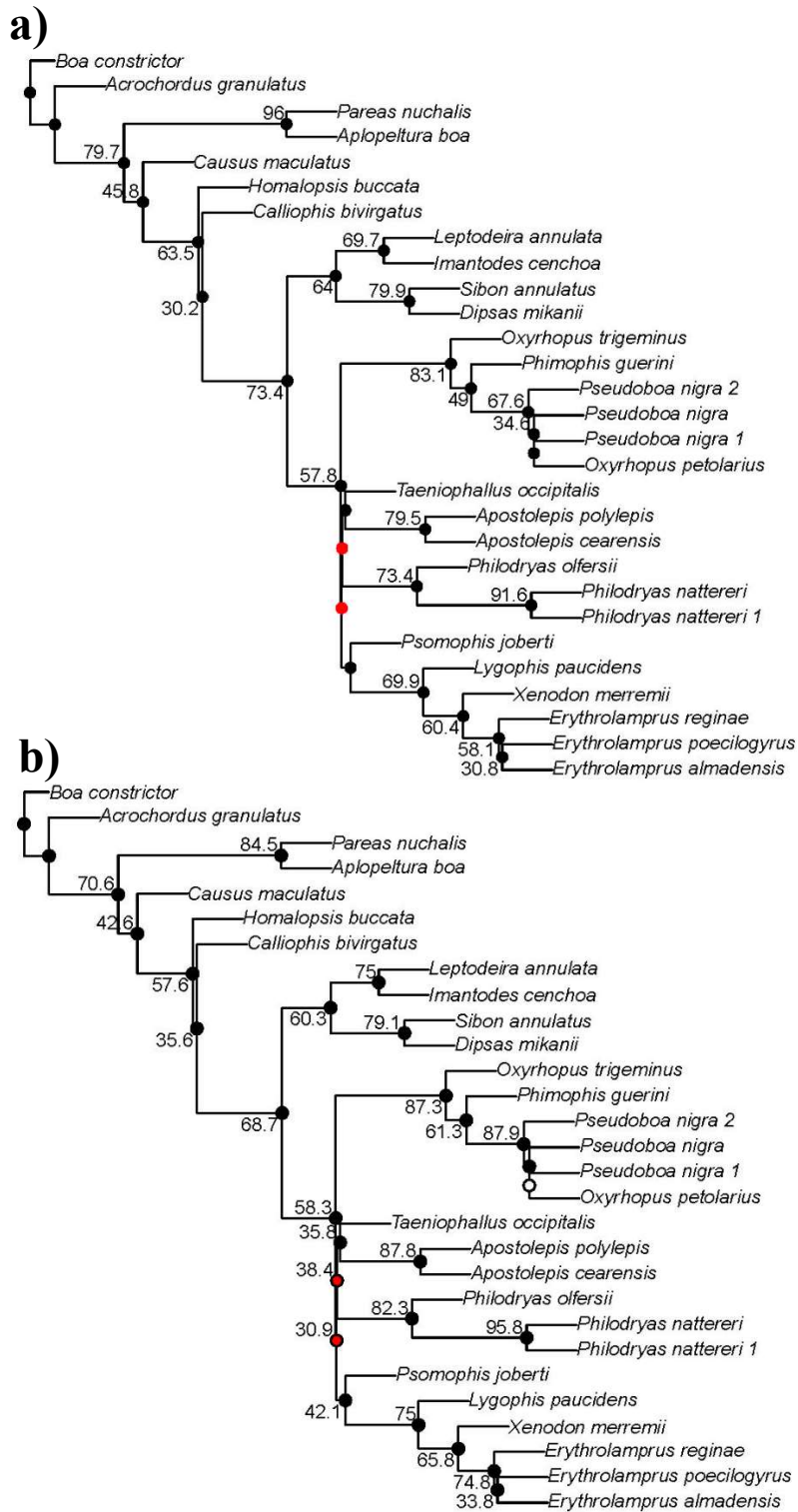


b)



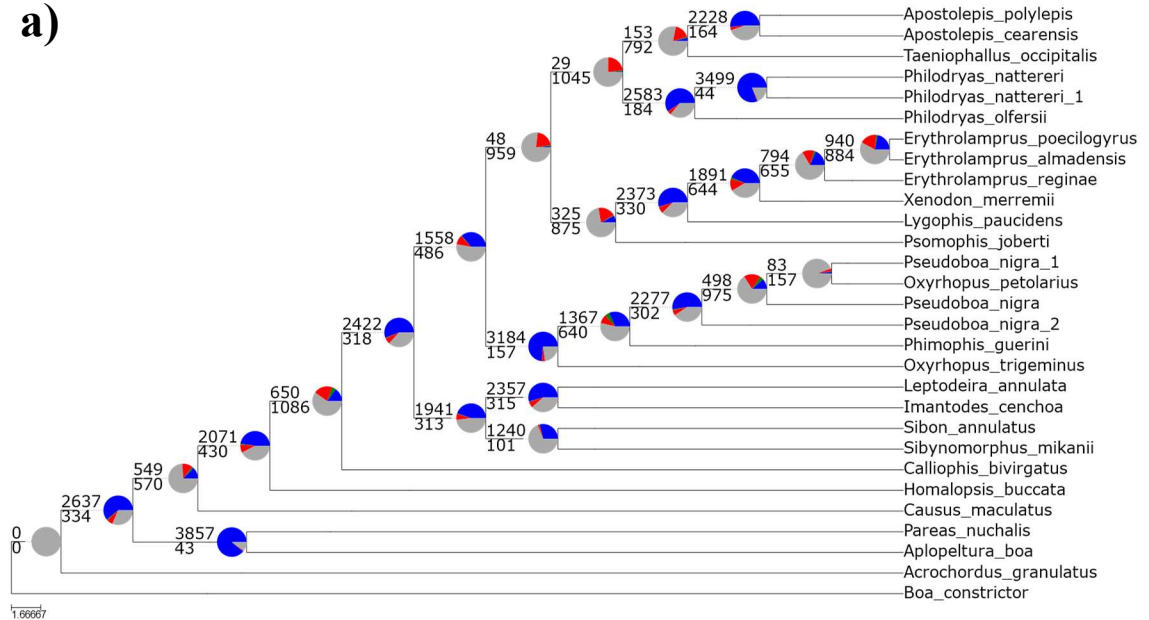


Appendix Figure A1.6. Phylogenetic relationships of the snake family Dipsadidae, as inferred with coalescent-based species tree analyses performed with the software ASTRAL-III. For these analyses, the dataset of 4333 loci from the Squamate Conserved Loci set with 75% or more taxonomic completeness was used after collapsing branches with values of ultrafast bootstrap (UFboot) of less than: a) 50, b) 75, or 95) from their corresponding gene trees. Color codes correspond to the same as in Figure 2.2.

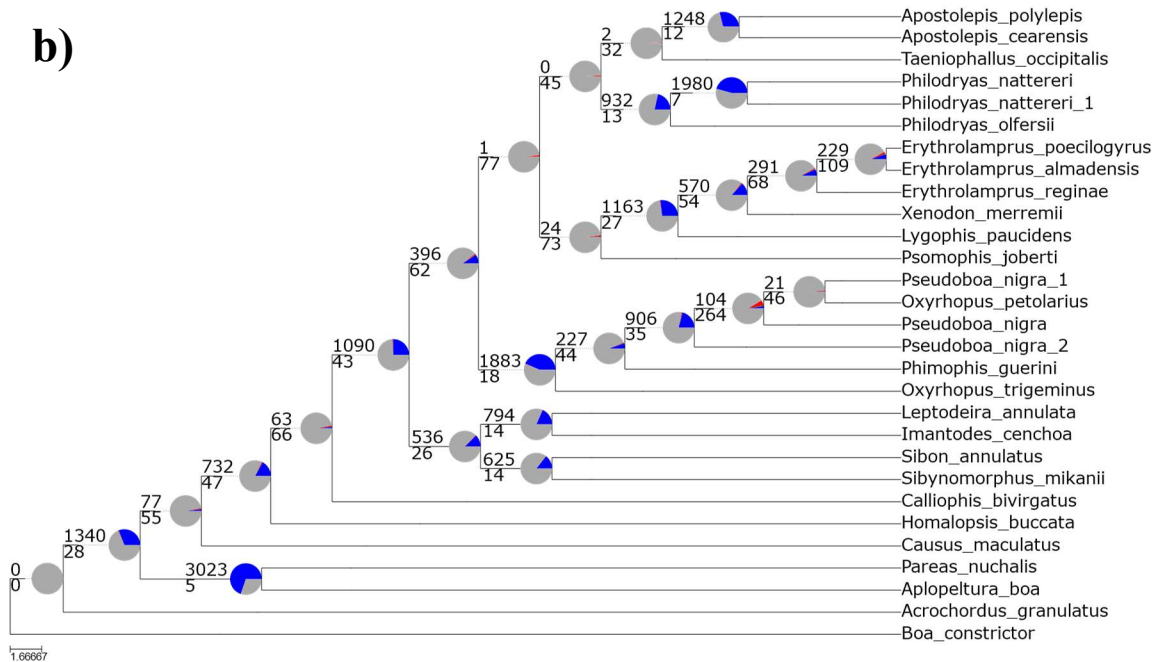


Appendix Figure A1.7. Phylogenetic relationships of the snake family Dipsadidae, as inferred with coalescent-based species tree analyses performed with the software ASTRAL-III. Node values correspond to the a) gene and b) site concordance factors (gCF and sCF, respectively). Nodes are colored black or white, depending on whether they have high or low local posterior probabilities or red if the subtending branches were recovered as being in the anomaly zone.

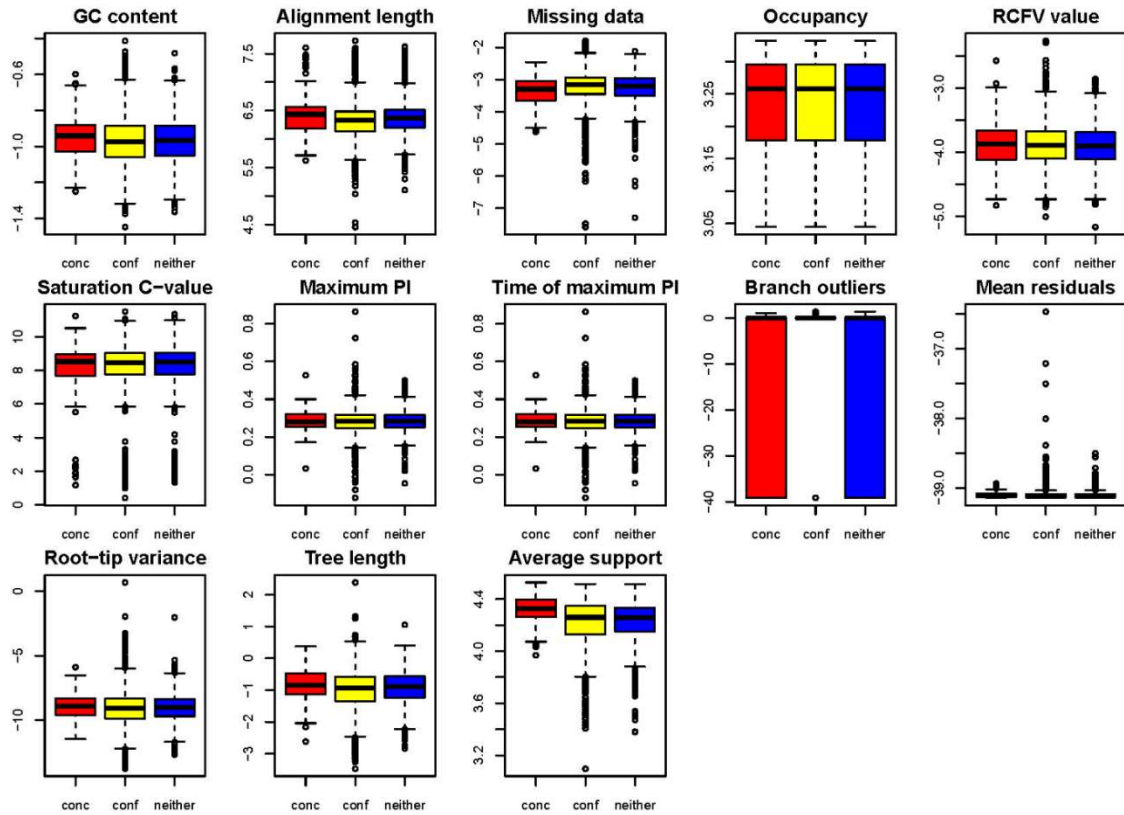
a)



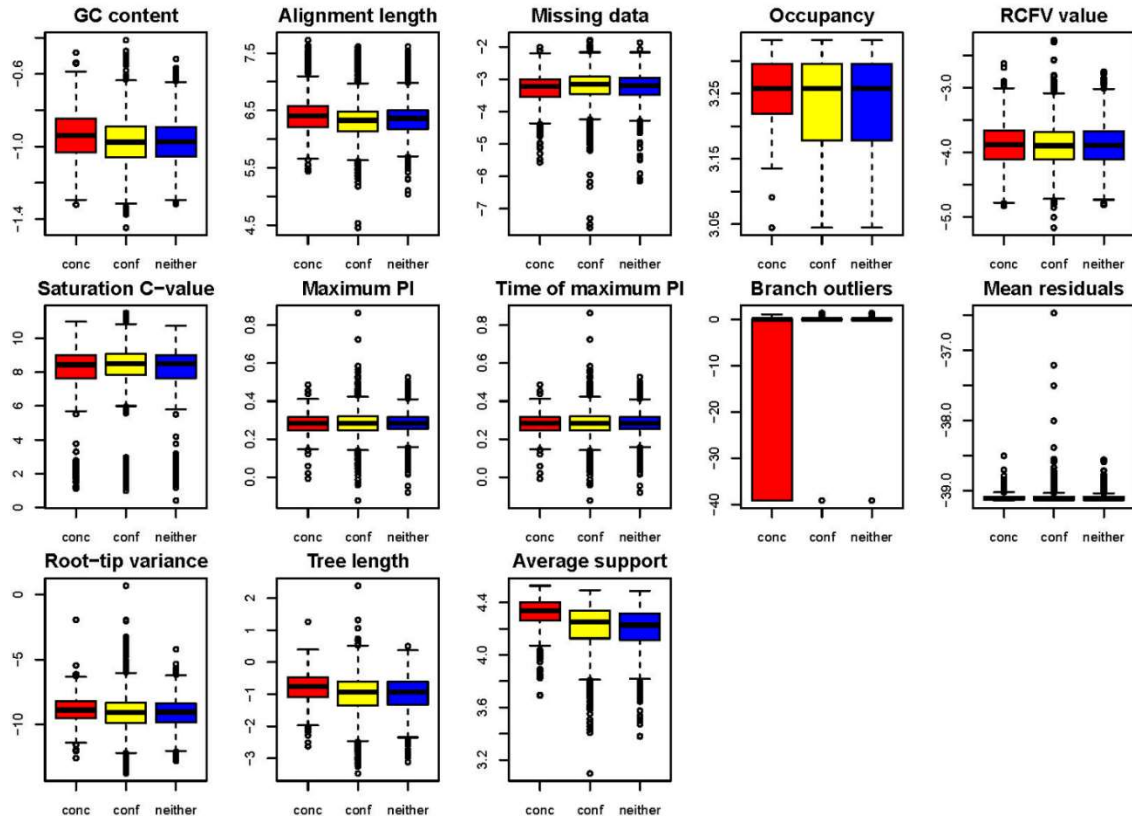
b)



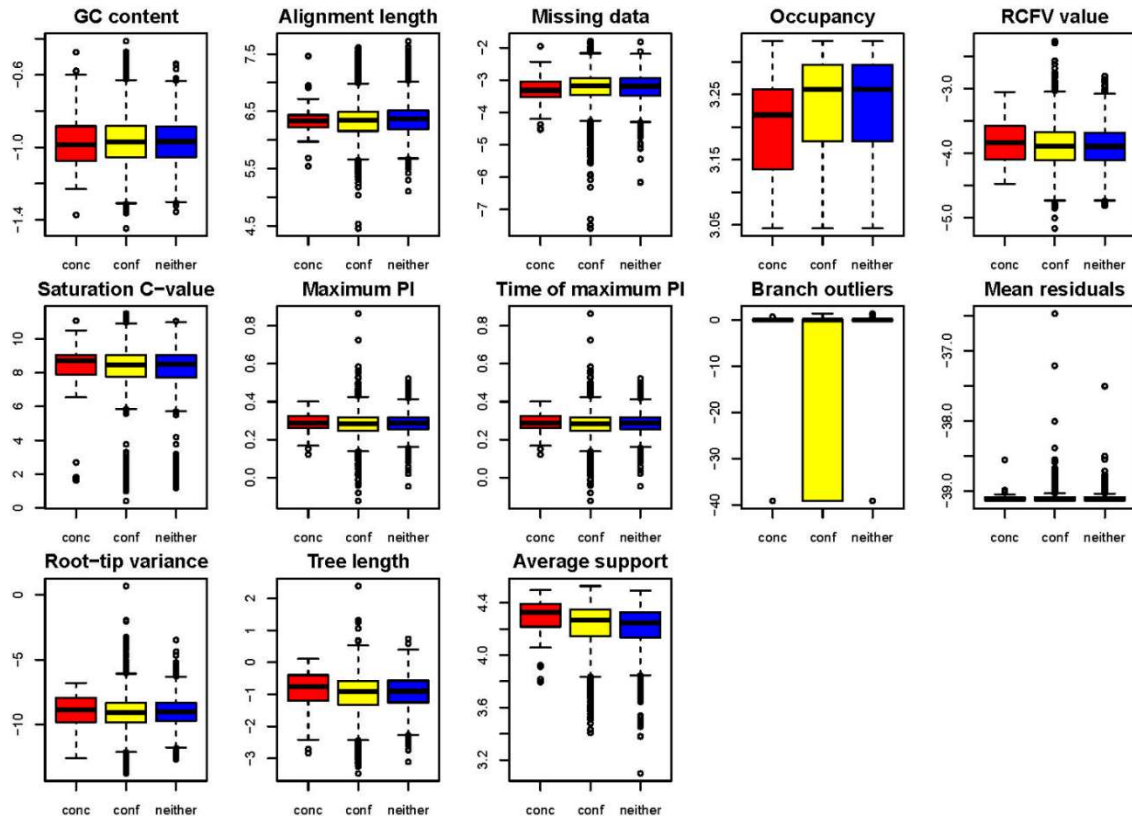
Appendix Figure A1.8. Species tree topology of the Dipsadidae obtained with ASTRAL-III indicating for every node the levels of incongruence with the inferred gene trees. Pie chart colors represent the proportion of gene trees supporting how a given node is resolved in the species tree (shown in blue), the most frequently found conflicting resolution (shown in green), all other conflicting possible reconstructions (shown in red), and non-informative gene trees (shown in grey). For this analysis, the branches on all gene trees were collapsed for having ultrafast bootstrap (UFboot) values of less than: a) 75, and b) 95. Numbers above or below branches indicate the number of gene trees concordant or conflicting with the species tree. It can be observed that many gene trees are uninformative regarding the resolution of a given node. However, the informative gene trees support the reconstruction inferred in the species tree for most nodes, except for the high-level relationships within the Xenodontinae.



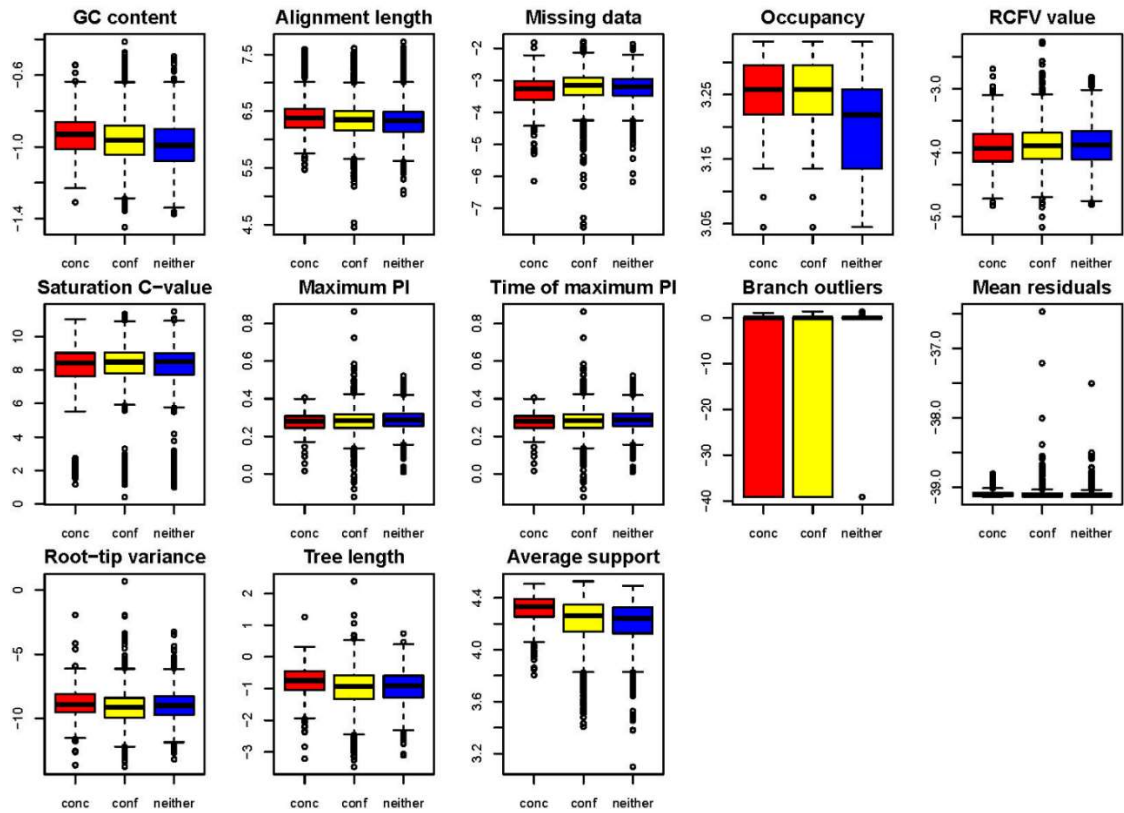
Appendix Figure A1.9. Comparison between the values of 13 traits indicative of gene tree estimation error between the loci whose corresponding gene trees support, conflict, or are uninformative regarding the reconstruction of the species tree of a sister relationship between the tribe Pseudoboini and the rest of the Xenodontinae. In that order, bars are colored red, blue, or yellow according to whether their corresponding gene trees support, conflict, or are uninformative regarding each of the relationships of interest inferred in the species tree. For this analysis, the branches on all gene trees having ultrafast bootstrap (UFboot) values of less than 50 were collapsed. All dependent variables were log-transformed to facilitate visualization.



Appendix Figure A1.10. Comparison between the values of 13 traits indicative of gene tree estimation error between the loci whose corresponding gene trees support, conflict, or are uninformative regarding the reconstruction of the species tree of a sister relationship between the Xenodontini and the Psomophini. In that order, bars are colored red, blue, or yellow according to whether their corresponding gene trees support, conflict, or are uninformative regarding each of the relationships of interest inferred in the species tree. For this analysis, the branches on all gene trees having ultrafast bootstrap (UFboot) values of less than 50 were collapsed. All dependent variables were log-transformed to facilitate visualization.



Appendix Figure A1.11. Comparison between the values of 13 traits indicative of gene tree estimation error between the loci whose corresponding gene trees support, conflict, or are uninformative regarding the reconstruction of the species tree of a sister relationship between the Philodryadini and the Echinantherini + Elapomorphini. In that order, bars are colored red, blue, or yellow according to whether their corresponding gene trees support, conflict, or are uninformative regarding each of the relationships of interest inferred in the species tree. For this analysis, the branches on all gene trees having ultrafast bootstrap (UFboot) values of less than 50 were collapsed. All dependent variables were log-transformed to facilitate visualization.



Appendix Figure A1.12. Comparison between the values of 13 traits indicative of gene tree estimation error between the loci whose corresponding gene trees support, conflict, or are uninformative regarding the reconstruction of the species tree of a sister relationship between the Echinantherini and the Elapomorphini. In that order, bars are colored red, blue, or yellow according to whether their corresponding gene trees support, conflict, or are uninformative regarding each of the relationships of interest inferred in the species tree. For this analysis, the branches on all gene trees having ultrafast bootstrap (UFboot) values of less than 50 were collapsed. All dependent variables were log-transformed to facilitate visualization.

Appendices Chapter 2

Appendix 2.1.

Methods and results of the BAMM analyses we performed.

We used the software BAMM 2.5.0 (Rabosky, 2014) to obtain values of the per-species speciation, extinction, and net diversification rates of the taxa included in this study. To do so, we performed BAMM analyses both with the original Bayesian calibrated timetree we estimated herein and a randomly selected tree of the 100 phylogenies obtained with TACT. The resulting values per-species rates were then averaged across the taxa inhabiting each cell and compared with the per-cell DR metric values we had previously obtained (Appendix Figures A2.6 and A2.7, respectively). In general, BAMM allows the simultaneous estimation of the total number of shifts in evolutionary rate regimes found across the dipsadid phylogenetic tree we inferred, as well as the extinction and diversification rates occurring between each shift. The Markov Chain Monte Carlo process implemented in BAMM allows estimating the posterior distribution of likely evolutionary rate regimes (=a mixture of density-dependent and constant diversification rates) and the number of shifts occurring among them (Rabosky, 2014; Seeholzer et al., 2017). The BAMM analysis we performed was run for ten million generations after identifying appropriate priors with the `setBAMMpriors` function of the R package `BAMMtools` 2.1.7 (Rabosky et al., 2014). We did not use the sampling fraction option of BAMM for either of the two separate analyses we performed. The convergence of the BAMM analyses was determined by verifying that the effective sample size (ESS) of the log-likelihood was greater than 200. We also followed the

guidelines described in <http://bamm-project.org> for ensuring adequate BAMM runs by exploring the credible set of shift configurations and computing marginal odds ratios. We used BAMMtools for visualizing the estimated evolutionary regime shifts and diversification rates across the input tree. Further, we obtained the speciation, extinction, and diversification rates of each species of the input trees by using the `getTipRates` function of BAMMtools.

The results of the BAMM analysis we performed with the randomly selected phylogeny of the Dipsadidae we had obtained with TACT indicate that the diversification of the group is characterized by two rate regime shifts. The first of these shifts corresponds to a decrease in the diversification rates occurring at the base of a clade formed by the Dipsadinae and Xenodontinae (Appendix Figure A2.8). This shift occurred about 29.87 million years ago (mya) (33.58-24.05, 95% confidence interval) during the Oligocene (Suppl. Figure A2.1). The mean speciation rate for the MRCA of the Dipsadinae and the Xenodontinae (0.187 species/millions of years, 0.167-0.213, 95% highest posterior density) is 0.865 times than the rate for the remaining members of Dipsadidae (0.217 species/millions of years, 0.107-0.406, 95% highest posterior density). The mean extinction rate for the Dipsadinae + Xenodontinae (0.044 species/millions of years, 0.013-0.083 of 95% highest posterior density) is 0.112 times than the corresponding rate for the remaining members of the family (0.135 species/millions of years (0.002-0.434 of 95% highest posterior density)). On the other hand, the other rate shift we recovered in this BAMM analysis corresponds to a clade containing the bulk of the genus *Atractus* (Appendix Figure A2.8). The placement of the few species of this

genus not recovered as part of this group most likely corresponds to artifacts due to limited data availability, as is further discussed in Suppl. Table 2A.2. This shift occurred about 13.67 million years ago (mya) (16.29-10.66, 95% confidence interval) during the Oligocene (Appendix Figure A2.8). In any case, the mean speciation rate of *Atractus* (0.442 species/millions of years, 0.302-0.625, 95% highest posterior density) is 2.72 times than the rate for the remaining members of Dipsadidae (0.162 species/millions of years, 0.145-0.185, 95% highest posterior density). The mean extinction rate of *Atractus* (0.259 species/millions of years, 0.036-0.522 of 95% highest posterior density) is 10.1 times than the corresponding rate for the remaining members of the family (0.026 species/millions of years (0.002-0.063 of 95% highest posterior density)).

On the other hand, the BAMM analyses we performed with the Bayesian timetree we inferred without including taxa without genetic data only recovered evidence for the same rate regime shift located at the base of a clade formed by the Dipsadinae and Xenodontinae (Appendix Figure A2.9). However, in this case, this shift corresponds to an increase in the diversification rates of the group. The mean speciation rate for the MRCA of the Dipsadinae and the Xenodontinae (0.115 species/millions of years, 0.101-0.132, 95% highest posterior density) is 0.72 times than the rate for the remaining members of Dipsadidae (0.160 species/millions of years, 0.067-0.294, 95% highest posterior density). On the other hand, the mean extinction rate for this clade (0.012 species/millions of years, 0.001-0.039 of 95% highest posterior density) is $\sim 1/8$ times than the corresponding rate for the remaining members of the family (0.098 species/millions of years (0.001-0.309 of 95% highest posterior density)). Finally, for both BAMM analyses, we extracted

the values of the per-species diversification, speciation, and extinction rates of every species included in this study (results not shown), which we then used to obtain Suppl. Figures S3.3, S3.5, and S3.6.

References not cited in the main text:

- Rabosky, D. L. (2014). Automatic detection of key innovations, rate shifts, and diversity dependence on phylogenetic trees. *PloS One*, 9(2). e89543.
- Rabosky, D. L., Grundler, M., Anderson, C., Title, P., Shi, J. J., Brown, J. W., ... & Larson, J. G. (2014). BAMM tools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*, 5(7), 701-707.
- Seeholzer, G. F., Claramunt, S., & Brumfield, R. T. (2017). Niche evolution and diversification in a Neotropical radiation of birds (Aves: Furnariidae). *Evolution*, 71(3), 702-715.

Appendix 2.2.

Equations tested in the a) theoretical and b) best-fitted structural equation models tested for this work. Differences in the variables considered by both models are indicated in bold. Abbreviations as in Figure 2.3:

a)

SR ~ TempSeas + PrepSeas + PrepAnom_IGM + EvolTime + DR +
TempAnom_IGM+ Elevation+ PrepAnom_LGM
DR ~ TempSeas + PrepSeas + PrepAnom_IGM + EvolTime+ TempAnom_IGM+
Elevation +PrepAnom_LGM
EvolTime ~ TempAnom_IGM + PrepAnom_IGM+PrepAnom_LGM
TempSeas ~ Elevation
PrepSeas ~ Elevation

b)

SR ~ TempSeas + PrepSeas + PrepAnom_IGM + EvolTime + DR +
TempAnom_IGM+ Elevation+ PrepAnom_LGM
DR ~ TempSeas + PrepSeas + PrepAnom_IGM + EvolTime+ TempAnom_IGM+
Elevation +PrepAnom_LGM
EvolTime ~ TempAnom_IGM + PrepAnom_IGM + PrepAnom_LGM + **Elevation**
TempSeas ~ Elevation + **TempAnom_IGM**
PrepSeas ~ Elevation + **TempSeas**
Correlated error: EvolTime ~ TempSeas

Tables

Appendix Table A2.1. Best-fit partitioning scheme for the protein-coding genes, and models of nucleotide substitution for each partition, based on the corrected Akaike information criterion (AICc) as implemented in PartitionFinder 2 (Lanfear et al., 2016). The abbreviation ‘pos.’ refers to codon position.

Subset	Best subset	Subset partitions	Subset sites
1	GTR+I+G	12S	1-358
2	GTR+I+G	16S	359-871
3	TRNEF+I+G	BDNF_3pos	872-1544\3
4	TVM+I	BDNF_1pos,	873-1544\3, 874-1544\3
5	GTR+I+G	ND2_1pos	1545-2288\3
6	GTR+I+G	ND4_2pos,	2289-2983\3, 1546-2288\3
7	GTR+I+G	ND2_3pos,	1547-2288\3, 3589-4703\3
8	GTR+G	ND4_3pos	2290-2983\3
9	GTR+I+G	ND4_1pos	2291-2983\3
10	TVM+G	cmos_3pos,	2984-3586\3, 5327-6038\3
11	TIM+I+G	rag2_1pos,	5325-6038\3, 2985-3586\3
12	TVM+I+G	cmos_2pos,	2986-3586\3, 5326-6038\3
13	GTR+I+G	cytb_1pos	3587-4703\3
14	GTR+I+G	cytb_2pos	3588-4703\3
15	TVM+G	nt3_1pos	4704-5324\3
16	SYM+I+G	nt3_2pos	4705-5324\3
17	TVMEF+G	nt3_3pos	4706-5324\3

Appendix Table A2.2. Calibration points used for obtaining the timetree used for this work and reported as Appendix Figure A2.1.

	Calibration point	Offset (Ma), Minimum age	95% quantile (Ma)	Fossil	References
1	Boidae:Boinae+Erycinae	58	67.5	<i>Titanoboa cerrejonensis</i>	Head et al. (2009); Head (2015)
2	stem Caenophidia (= crown Pan-Caenophidia)	66	88.5	<i>Krebsophis thobanus</i>	Rage & Werner (1999); Head et al. (2016)
3	stem Colubroidea (= stem Colubriformes)	50.5	85	<i>Procerophis sahnii</i>	Rage et al. (2008) & Head et al. (2016)
4	stem Elapidae	24.95	47.4	Elapidae indeterminate	McCartney et al. (2014); Zaher et al. (2019)
5	crown Viperidae	22.1	47	<i>Vipera cf. V. antiqua</i>	Zaher et al. (2019)*

*According to Head et al. (2017), this fossil cannot be used to calibrate the minimum age of the divergence between Viperidae and its sister group (i.e., Homalopsidae+Colubroidea). Such is the case, as this fossil is of a more recent date than the time of the first appearance of Homalopsidae or Colubroidea in the fossil record. However, the *Vipera cf. antiqua* fossil can still be used to calibrate the minimum age of the first divergence within Viperidae (=Azemiopinae vs. Viperinae+Crotalinae), as was done in Zaher et al. (2019) and followed herein.

Appendix Table A2.3. Dutilleul's modified t-test between all explanatory variables considered for the theoretical structural equation model tested for this study. For this model, we used the values of the DR and MPD metrics obtained using the 100 completely sampled phylogenies we generated with the stochastic polytomy resolving software TACT and the range maps compiled by Roll et al. (2017). Values on each cell correspond to the slope of the correlation between each combination of variables, with the corresponding p-values are shown being between parentheses. Strong correlations (=slopes>0.8) are indicated in bold. To prevent collinearity, we ensured that the equations from our pSEM model only incorporated one of the two variables found to be strongly correlated to each other (=temperature anomalies since the last glacial or interglacial maximum).

	Species Richness	Temperature Seasonality	Precipitation Seasonality	Precipitation Anomaly since LGM	Precipitation Anomaly since IGM	Evolutionary Time	DR metric	Elevation	Temperature Anomaly since LGM
Species Richness									
Temperature Seasonality	-0.787 (0.163)								
Precipitation Seasonality	0.070(0.72)	-0.318 (0.158)							
Precipitation Anomaly since LGM	-0.030 (0.686)	-0.051 (0.610)	0.004 (0.945)						
Precipitation Anomaly since IGM	0.280 (0.185)	-0.071 (0.739)	-0.060 (0.515)	0.199 (0)					
Evolutionary Time	-0.148 (NaN)	0.181 (NaN)	-0.021 (0.793)	-0.174 (0.007)	-0.086 (0.125)				
DR metric	0.953 (0.015)	-0.713 (0.203)	0.025 (0.905)	-0.101 (0.295)	0.303 (0.125)	-0.145 (0.005)			
Elevation	-0.216 (0.066)	-0.105 (0.416)	0.263 (0.026)	-0.009 (0.870)	-0.224 (2e- 04)	0.019(0.808)	-0.293 (0.020)		

Temperature								
Anomaly	-0.753		-0.216	-0.103	-0.074		-0.684	0.180
since LGM	(0.075)	0.851 (0.035)	(0.244)	(0.232)	(0.667)	0.191 (0)	(0.102)	(0.105)
Temperature								
Anomaly	-0.713		-0.092	-0.058	-0.104		-0.686	0.366
since IGM	(0.086)	0.729 (0.108)	(0.635)	(0.514)	(0.523)	0.126 (NaN)	(0.092)	(9e-04) 0.918 (7e-04)

Appendix Table A2.4. Model selection of the spatial weight matrices associated with each of the equations of the structural equation models tested herein using the range maps compiled by Roll et al. (2017). For each equation, we compared the model fit of spatial autoregressive analyses (SARs) performed using a total of six spatial weight matrices. These matrices correspond to all possible combinations between two neighborhood distances (d1=50 km, d2=269.2582 km) and three weight coding schemes (row standardized [“W”], globally standardized [“C”], and variance-stabilizing [“S”]). For this procedure, we used the values of the DR and MPD metrics obtained using the 100 completely sampled phylogenies we generated with the stochastic polytomy resolving software TACT.

Equation	Scheme	Model	R2	AIC value	Degrees of freedom
Equation 1	W	lm_mod	0.937874	2069.896	10
	W	error_d1	0.938161	2037.146	11
	W	error_d2	0.995113	-21395.8	11
	C	lm_mod	0.937874	2069.896	10
	C	error_d1	0.938161	2037.146	11
	C	error_d2	0.992442	-17370.1	11
	S	lm_mod	0.937874	2069.896	10
	S	error_d1	0.938161	2037.146	11
	S	error_d2	0.994186	-19791.5	11
Equation 2	W	lm_mod	0.711898	14096.53	9
	W	error_d1	0.713626	14050.01	10
	W	error_d2	0.992322	-19364	10
	C	lm_mod	0.711898	14096.53	9
	C	error_d1	0.713626	14050.01	10
	C	error_d2	0.982248	-11625	10
	S	lm_mod	0.711898	14096.53	9
	S	error_d1	0.713626	14050.01	10
	S	error_d2	0.989767	-16711.6	10
Equation 3	W	lm_mod	0.039329	25431.51	5
	W	error_d1	0.040127	25428.84	6
	W	error_d2	0.763563	12492.32	6
	C	lm_mod	0.039329	25431.51	5
	C	error_d1	0.040127	25428.84	6
	C	error_d2	0.603647	17262.3	6
	S	lm_mod	0.039329	25431.51	5
	S	error_d1	0.040127	25428.84	6
	S	error_d2	0.700906	14662.78	6
Equation 4	W	lm_mod	0.009451	22838.42	3
	W	error_d1	0.027877	22668.05	4

	W	error_d2	0.996971	-30617.3	4
	C	lm_mod	0.009451	22838.42	3
	C	error_d1	0.027877	22668.05	4
	C	error_d2	0.934899	-2293.79	4
	S	lm_mod	0.009451	22838.42	3
	S	error_d1	0.027877	22668.05	4
	S	error_d2	0.973208	-10491.1	4
Equation 5	W	lm_mod	0.070982	25609.86	3
	W	error_d1	0.078415	25538.69	4
	W	error_d2	0.954185	-2173.94	4
	C	lm_mod	0.070982	25609.86	3
	C	error_d1	0.078415	25538.69	4
	C	error_d2	0.906863	4376.447	4
	S	lm_mod	0.070982	25609.86	3
	S	error_d1	0.078415	25538.69	4
	S	error_d2	0.937635	673.4307	4

Appendix Table A2.5. Results of the d-separation tests we performed with the theoretical structural equation model being shown in Figure 2.3 and Appendix 2. Combinations of variables recovered as being significant are indicated in bold and represent correlations that need to be incorporated into the SEM model.

Independent claim	Critical value	P-value
TempSeas ~ PrepAnom_IGM	-0.492	0.6227
PrepSeas ~ PrepAnom_IGM	-1.2279	0.2195
TempSeas ~ TempAnom_IGM	-15.7429	0
PrepSeas ~ TempAnom_IGM	0.8646	0.3873
EvolTime ~ Elevation	-2.3224	0.0202
TempSeas ~ PrepAnom_LGM	1.7156	0.0862
PrepSeas ~ PrepAnom_LGM	-0.0428	0.9659
PrepSeas ~ TempSeas	3.5467	0.0004
EvolTime ~ TempSeas	3.3014	0.001
EvolTime ~ PrepSeas	0.6779	0.4978

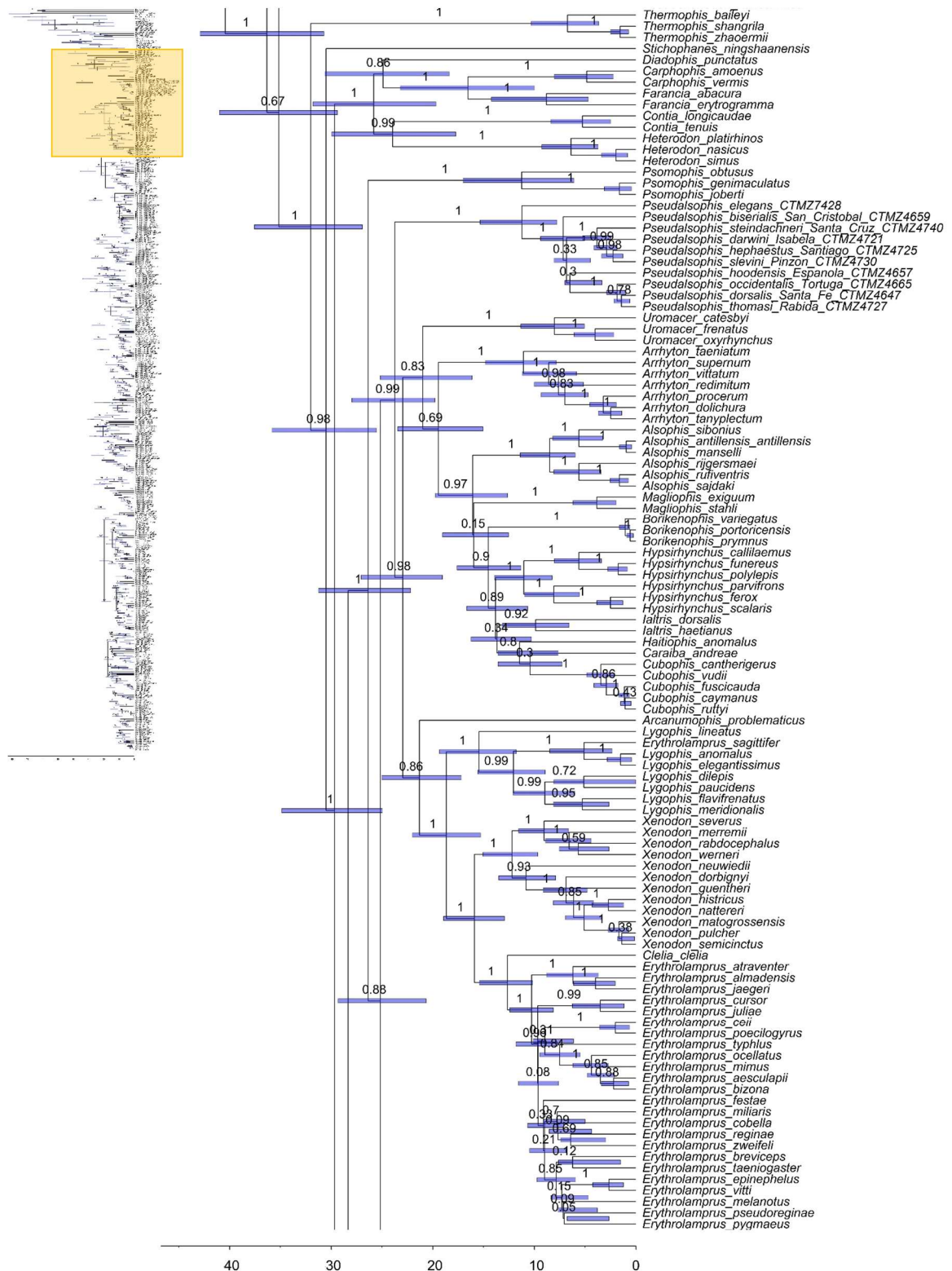
Appendix Table A2.6. Standardized estimates of the best-fitted structural equation model of the ecological and evolutionary factors driving species richness patterns in the Dipsadidae. For this model, we used the values of the DR and MPD metrics obtained using the calibrated Bayesian timetree we inferred herein after including taxa lacking genetic data with the software TACT and the range maps compiled by Roll et al. (2017). Abbreviations as in Figure 2.3.

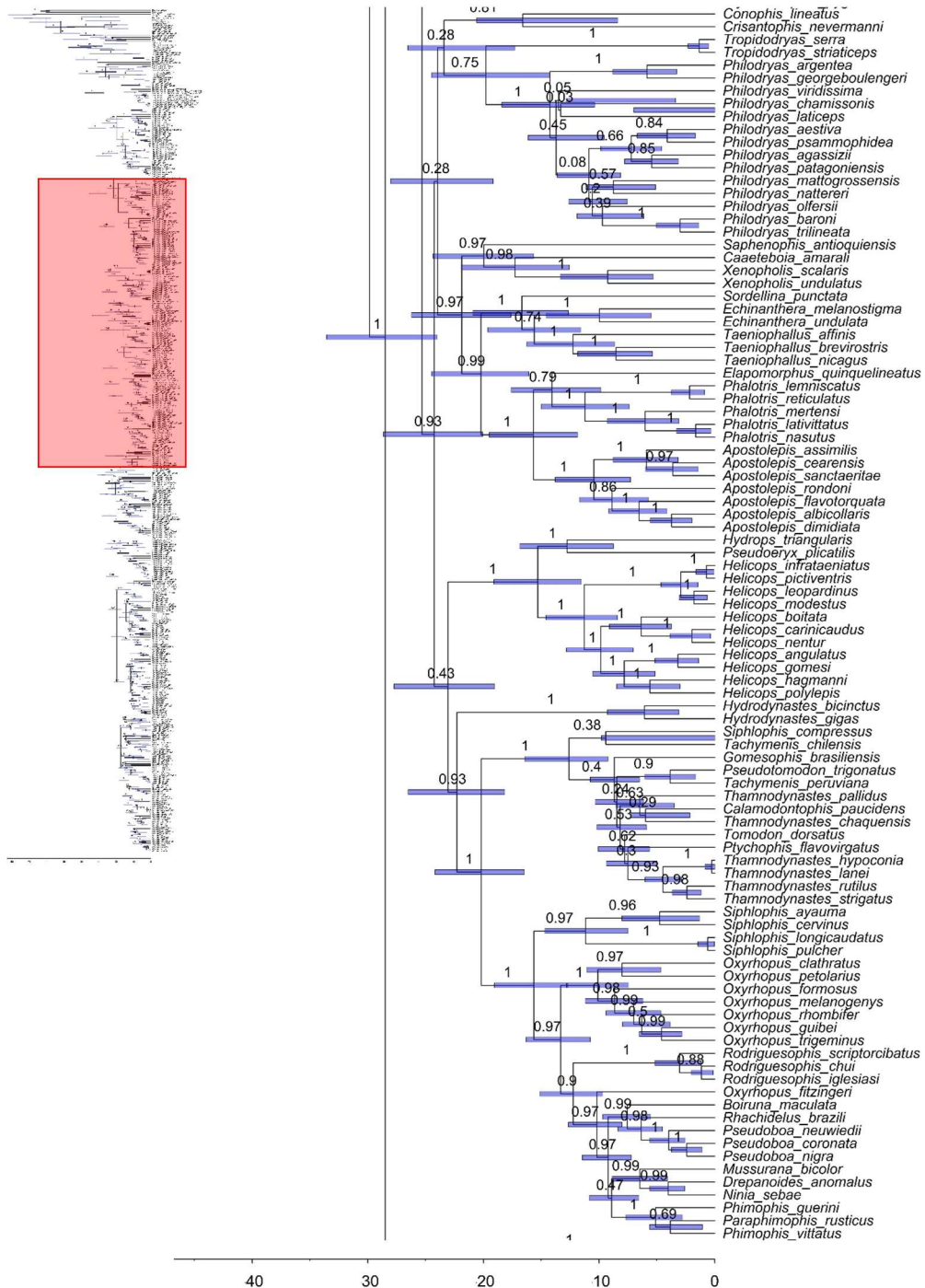
Response variable	Predictor variable	Standardized Estimate	P-value
SR	TempSeas	-0.0384	0.0035
SR	PrepSeas	0.0018	0.6045
SR	PrepAnom_IGM	0.0033	0.0446
SR	EvolTime	-0.0173	0
SR	DR	0.9005	0
SR	TempAnom_IGM	-0.0267	0
SR	Elevation	0.0095	0
SR	PrepAnom_LGM	-0.0041	0.0883
DR	TempSeas	-0.0335	0.0465
DR	PrepSeas	0.0012	0.7717
DR	PrepAnom_IGM	-0.0008	0.6881
DR	EvolTime	0.0009	0.6137
DR	TempAnom_IGM	-0.0298	0
DR	Elevation	0.0101	0
DR	PrepAnom_LGM	0.0142	0
EvolTime	TempAnom_IGM	-0.0524	0.0294
EvolTime	PrepAnom_IGM	-0.0082	0.4751
EvolTime	PrepAnom_LGM	0.0044	0.7862
EvolTime	Elevation	-0.0267	0.0202
TempSeas	Elevation	-0.0089	0
TempSeas	TempAnom_IGM	-0.0436	0
PrepSeas	Elevation	-0.01	0.036
PrepSeas	TempSeas	0.1429	0.0004
~EvolTime	~TempSeas	0.0389	0.0001

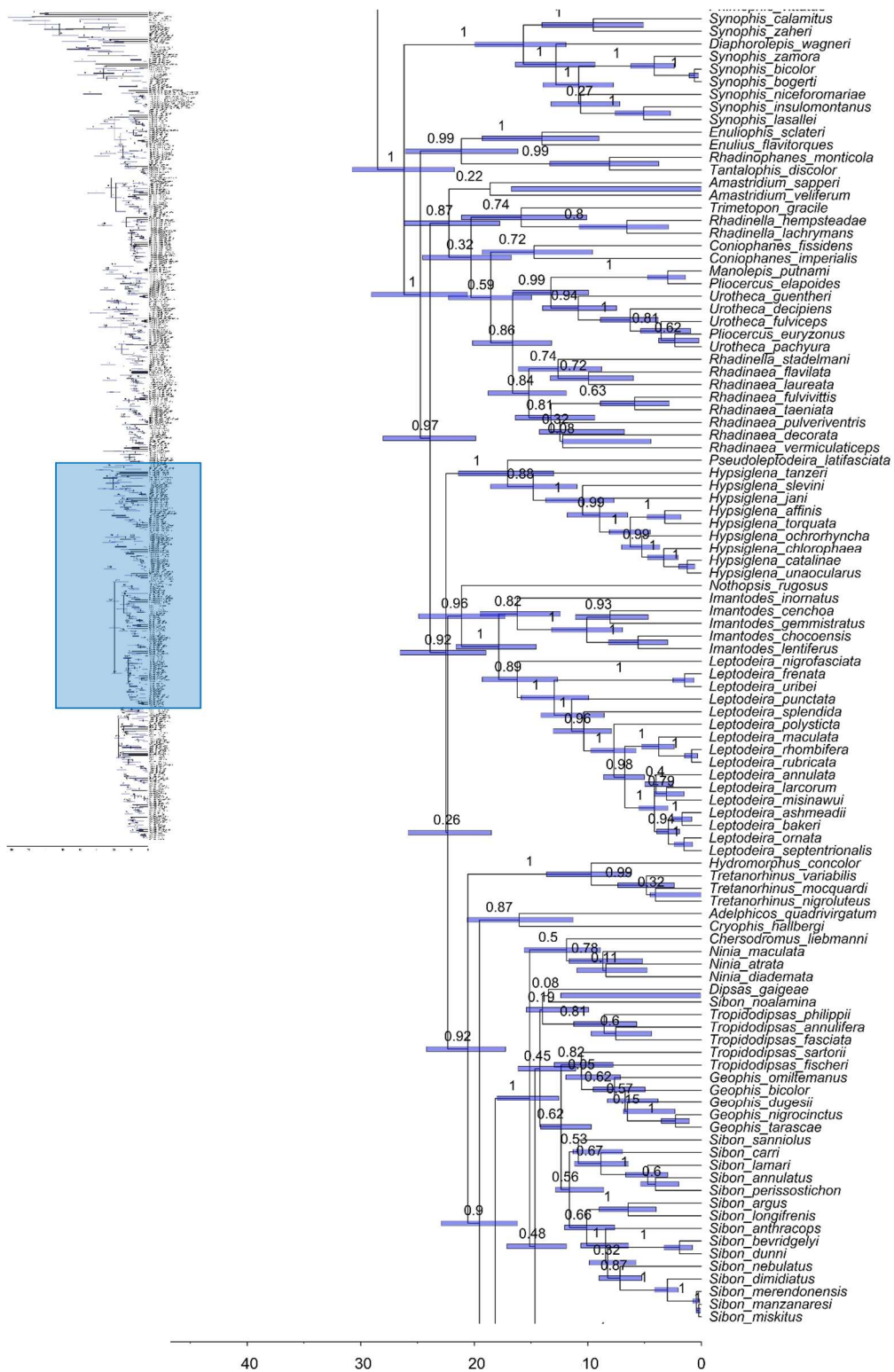
References used for constructing Appendix Table A2.2

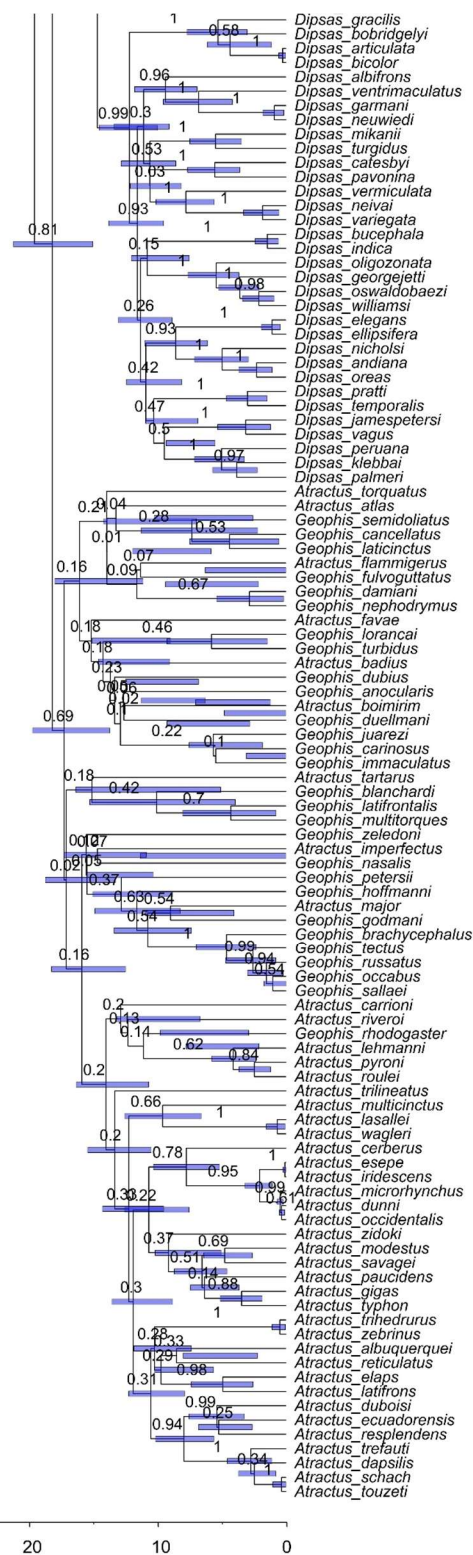
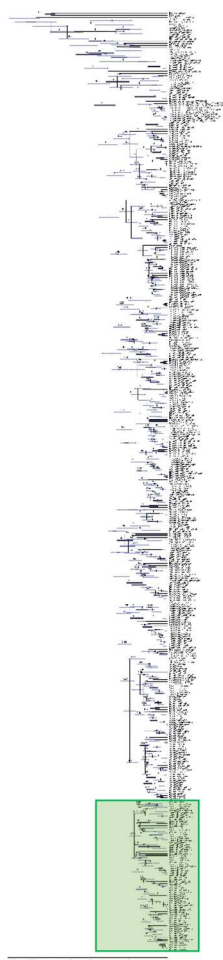
- Head, J.J., Bloch, J.I., Hastings, A.K., Bourque, J.R., Cadena, E.A., Herrera, F.A., Polly, P.D., and Jaramillo, C.A. 2009. Giant boid snake from the Paleocene neotropics reveals hotter past equatorial temperatures. *Nature*, 457:715-717
- Head, J. J. (2015). Fossil calibration dates for molecular phylogenetic analysis of snakes 1: Serpentes, Alethinophidia, Boidae, Pythonidae. *Palaeontologia Electronica*, 18(6FC), 1-17.
- Head, J. J., Mahlow, K., & Mueller, J. (2016). Fossil calibration dates for molecular phylogenetic analysis of snakes 2: Caenophidia, Colubroidea, Elapoidea, Colubridae. *Palaeontologia Electronica*, 19(2FC), 1-21.
- McCartney, J.A., Stevens, N.J., and O'Connor, P.M. 2014. The earliest colubroid-dominated snake fauna from Africa: perspectives from the Late Oligocene Nsungwe Formation of Southwestern Tanzania. *PLoS One* 9(3): e90415
- Rage, J.-C. and Werner, C. 1999. Mid-Cretaceous (Cenomanian) snakes from Wadi Abu Hashim, Sudan: The earliest snake assemblage. *Palaeontologia Africana*, 35:85-110
- Rage, J.-C., Folie, A., Rana, R S., Singh, H., Rose, K.D. and Smith, T. 2008. A diverse snake fauna from the early Eocene of Vastan Lignite Mine, Gujarat, India. *Acta Palaeontologica Polonica*, 53:391-403.
- Zaher, H., Murphy, R. W., Arredondo, J. C., Graboski, R., Machado-Filho, P. R., Mahlow, K., ... & Grazziotin, F. G. (2019). Large-scale molecular phylogeny, morphology, divergence-time estimation, and the fossil record of advanced caenophidian snakes (Squamata: Serpentes). *PloS One*, 14(5), e0216148.

Figures

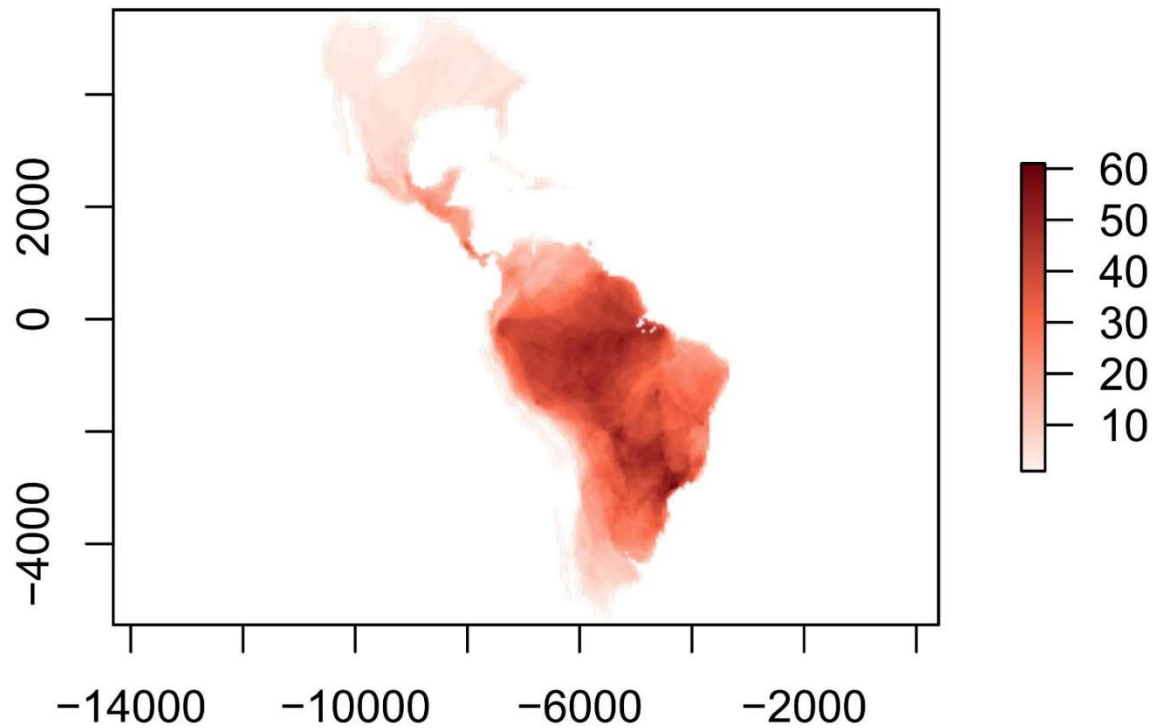




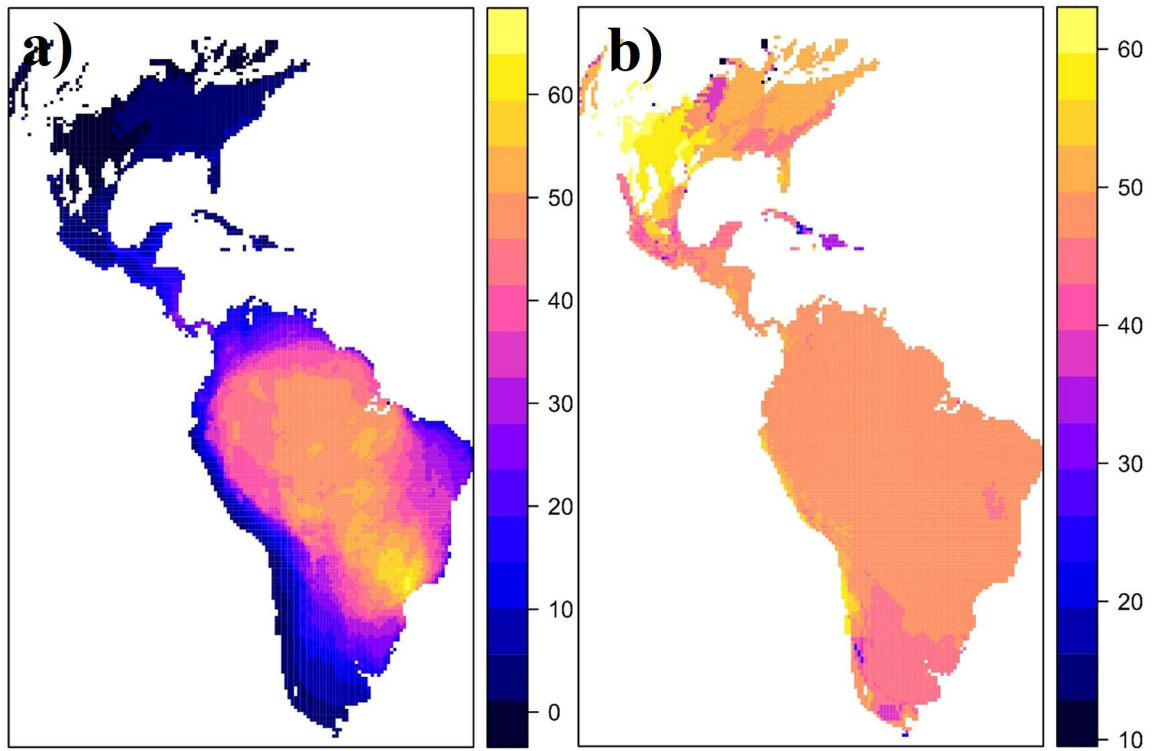




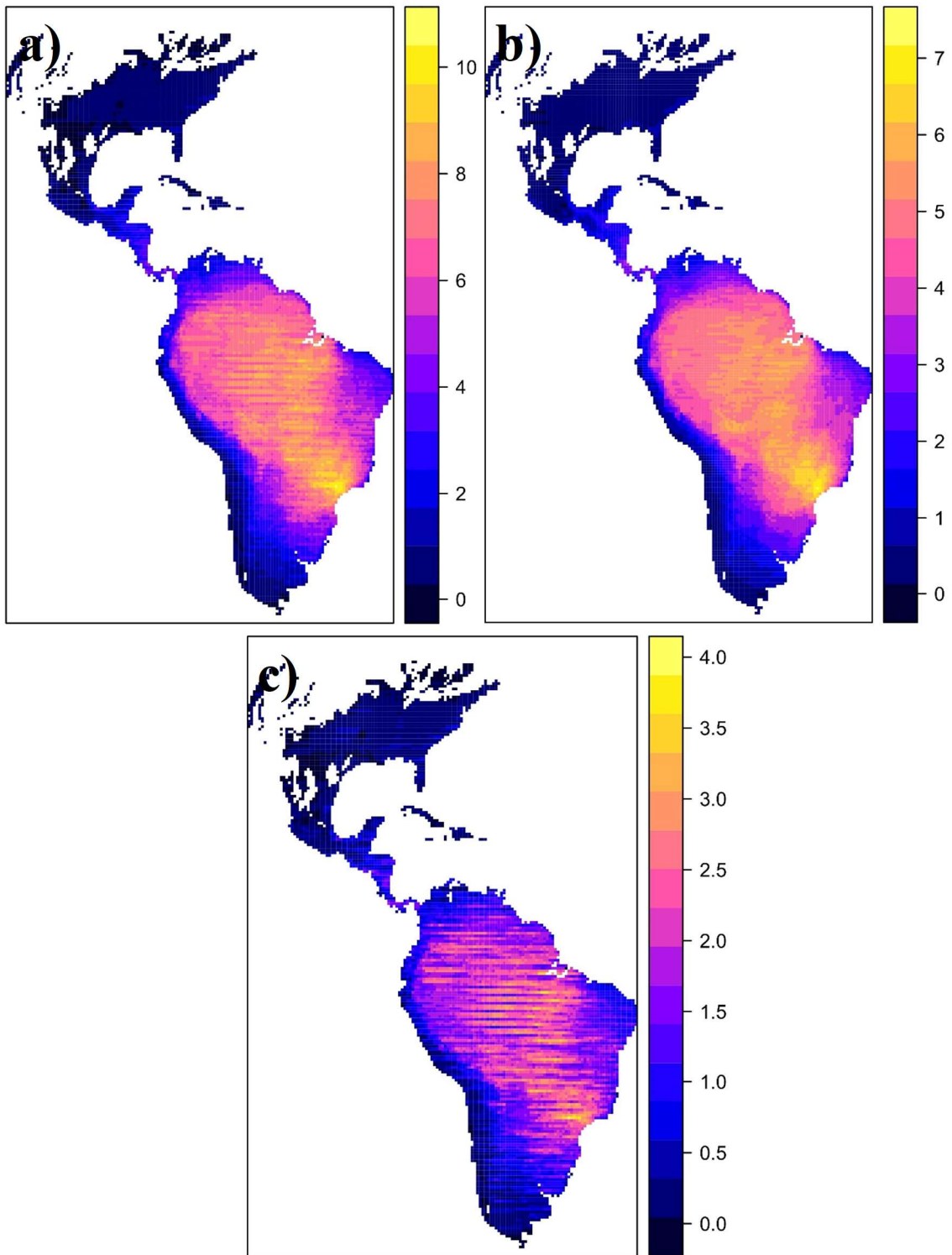
Appendix Figure A2.1. Timetree of the family Dipsadidae derived from a relaxed-clock MCMC Bayesian analysis using the software Beast 2, calibrated with the fossils presented in Appendix Table 2.2. Numbers above branches are the posterior probabilities of adjacent nodes. 95% credibility intervals of the age of each node are indicated with blue bars. The taxonomic arrangement of the members of Dipsadidae into tribes follows Zaher et al. (2009), and Sheehey (2013), incorporating recent taxonomic changes such as the dissolution of the Saphenophini tribe by Zaher et al. (2018). Red asterisks indicate taxa with a phylogenetic position that conflicts with their traditionally taxonomic placement, and thus, resulting in the non-monophyly of the genera in which they are currently placed. The units of the scale correspond to millions of years before the present time.



Appendix Figure A2.2. Patterns of species richness of the snake family Dipsadidae across the New World, as estimated from the range maps compiled by Acevedo et al. (2019). Darker and paler tones of red correspond to areas with higher and lower species richness of Dipsadidae, respectively. The map is overlaid on a 50 km x 50 km grid using an equal-earth projection. Cells with less than 25% of land surface are not mapped.

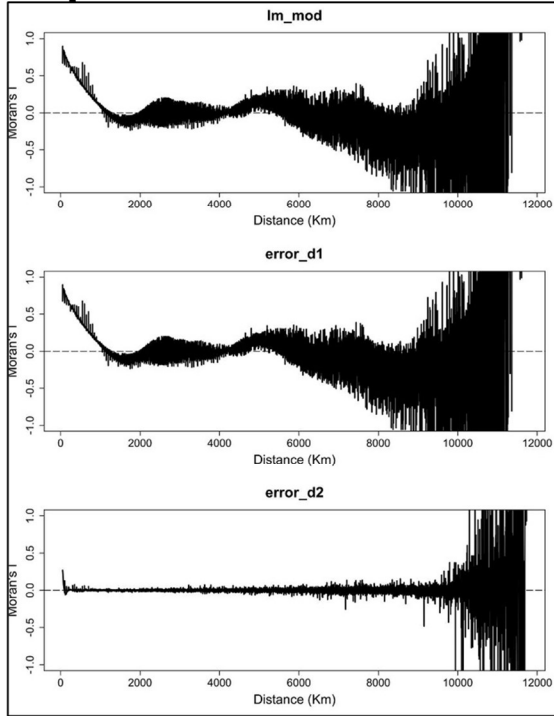


Appendix Figure A2.3. Patterns of geographic variation of the values of per-cell a) DR metric, and b) Mean Phylogenetic Diversity (MPD) of Jetz et al. (2012) across the distribution of the family Dipsadidae in the New World, obtained using the calibrated Bayesian calibrated phylogeny we inferred herein. To calculate these values, we used the range maps compiled by Roll et al. (2017). Warmer and color colors correspond to higher and lower values of each metric. The map is overlaid on a grid of 50 km x 50 km using an equal earth projection. Cells with less than 25% of land surface are not mapped.

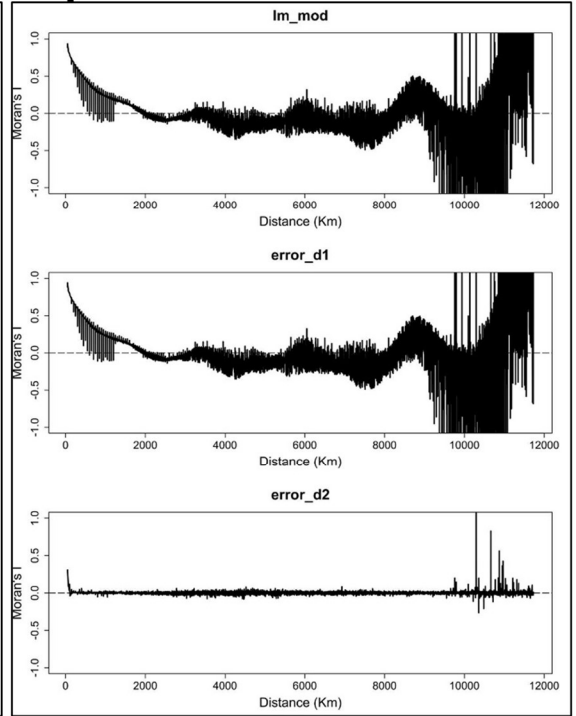


Appendix Figure A2.4. Patterns of geographic variation of the values of per-cell a-c) Net diversification, speciation, and extinction rates across the distribution of the family Dipsadidae in the New World, as obtained with the software BAMM. Values were obtained using one of the 100 completely sampled phylogenies we generated with the stochastic polytomy resolving software TACT and the range maps compiled by Roll et al. (2017). Warmer and color colors correspond to higher and lower values of each metric. The map is overlaid on a grid of 50 km x 50 km using an equal earth projection. Cells with less than 25% of land surface are not mapped.

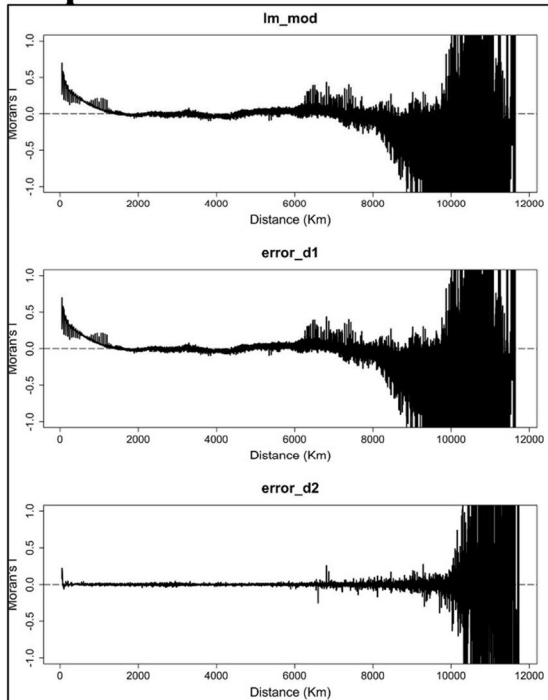
Equation 1



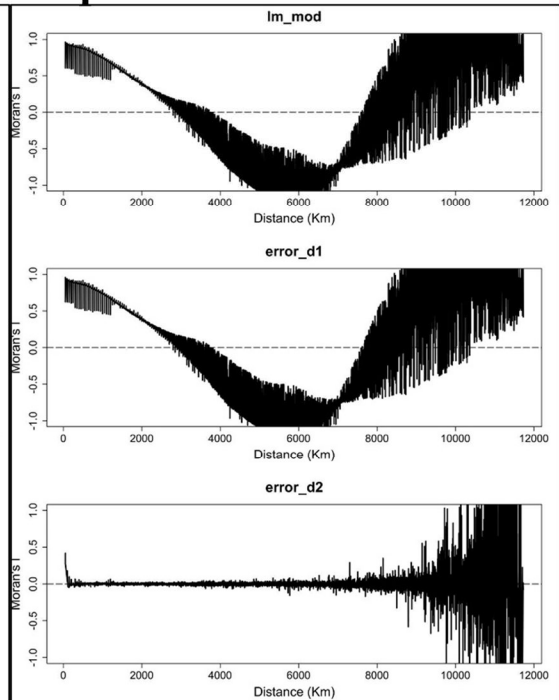
Equation 2



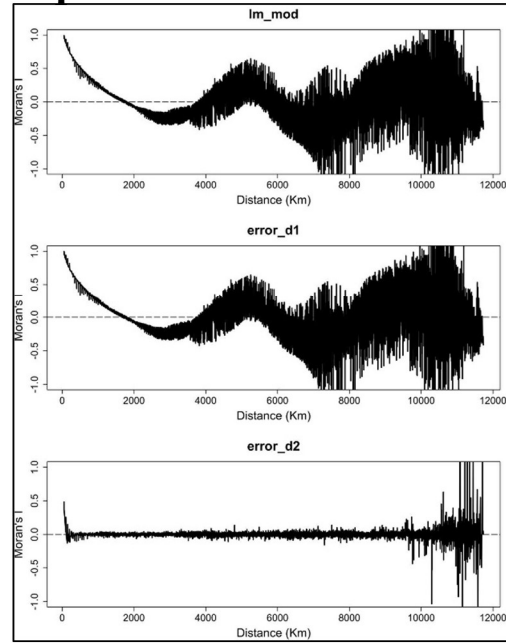
Equation 3



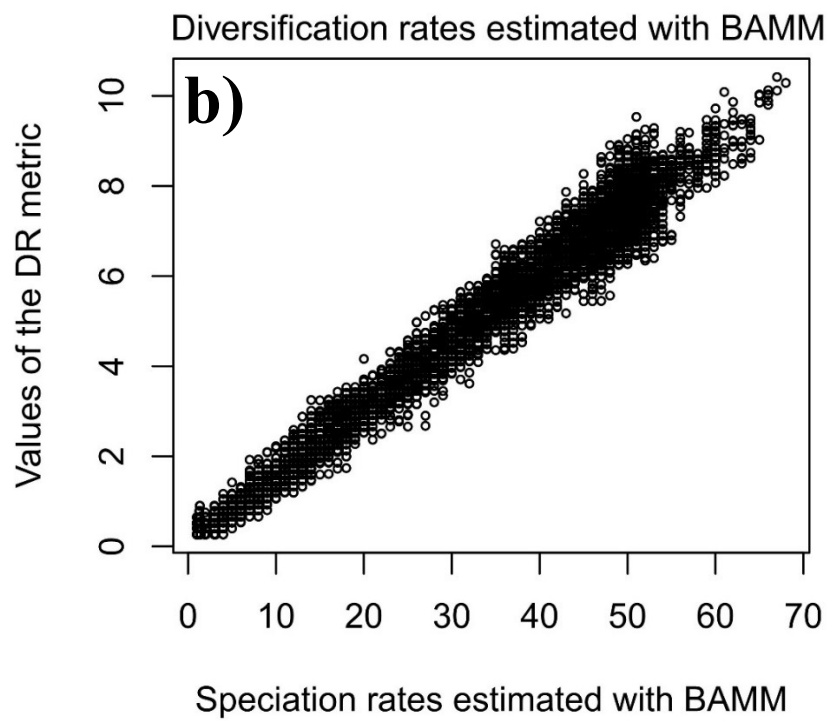
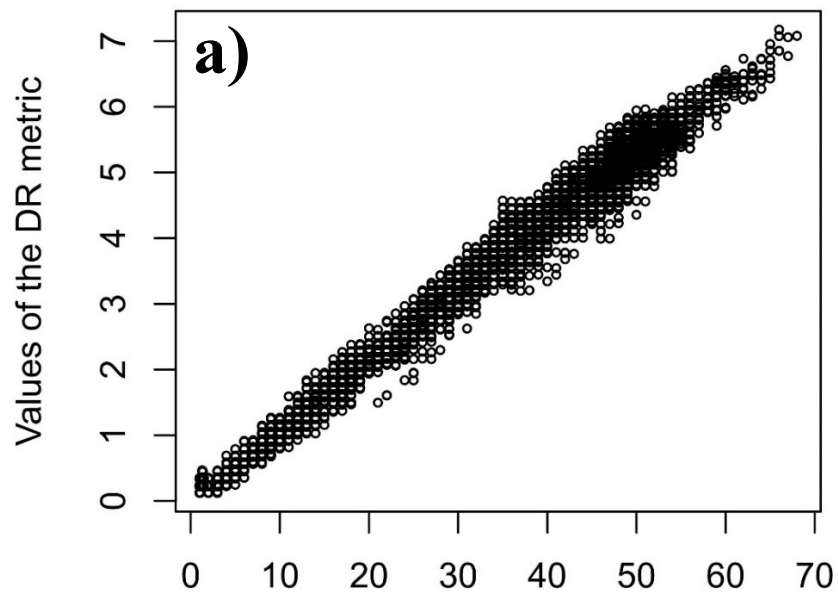
Equation 4



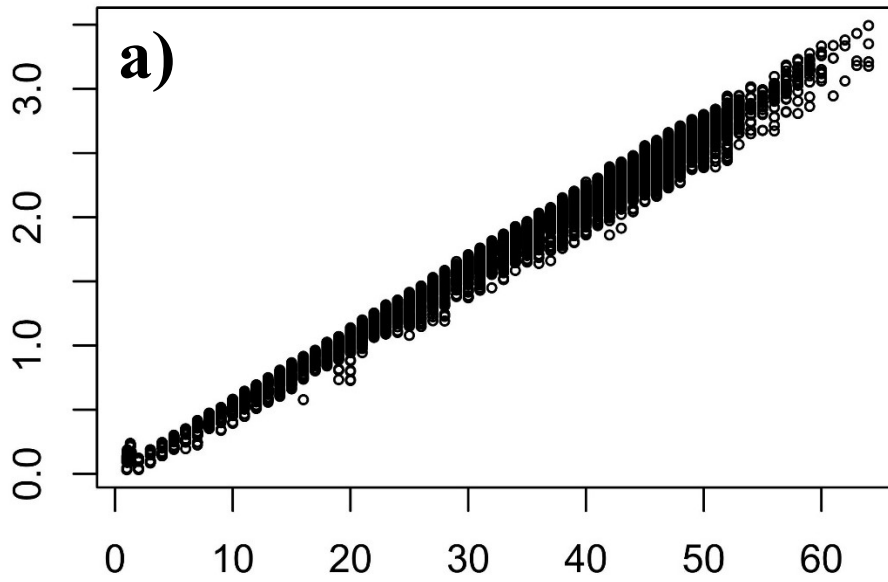
Equation 5



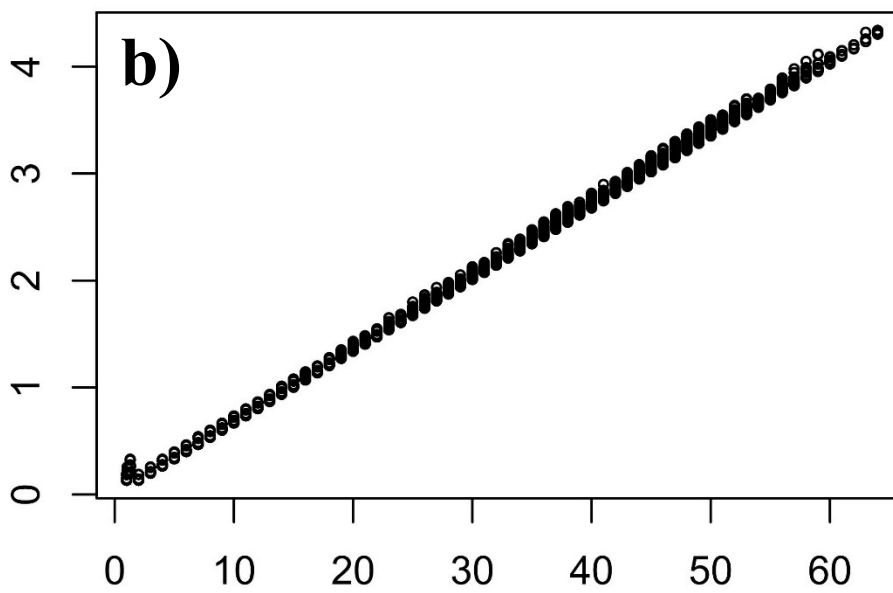
Appendix Figure A2.5. Moran's I correlograms indicating spatial autocorrelation in the residuals of the equations of the piecewise structural equation model we evaluated herein using the range maps compiled by Roll et al. (2017). For each of the equations of the model, we present the correlograms corresponding to the original Ordinary Least Squares (OLS) model and a Spatial Autoregressive model (SAR), performed using two weighted schemes (“error_d1” and “error_d2”) to account for the effect of spatial autocorrelation. To obtain these results, we used neighboring distances of 50 and 269.2582 km, respectively (d1 and d2).



Appendix Figure A2.6. Relationships between the estimated values of the DR metric and the per-cell net diversification and speciation rates, as obtained with the software BAMM and the range maps compiled by Roll et al. (2017). Values of the DR metric were obtained with the 100 completely sampled phylogenies we generated with the stochastic polytomy resolving software TACT. The per-cell rates we had inferred with BAMM were obtained using one randomly selected tree from those obtained with TACT. After accounting for spatial autocorrelation, the DR metric is strongly correlated with both rates (for net diversification rates: correlation: 0.9952, F-statistic: 334.2807 on 1 and 3.2161 DF, p-value: 2e-04, for speciation rates: Dutilleul's modified t-test, correlation: 0.9909, F-statistic: 176.2375 on 1 and 3.2528 degrees of freedom, p-value: 6e-04).

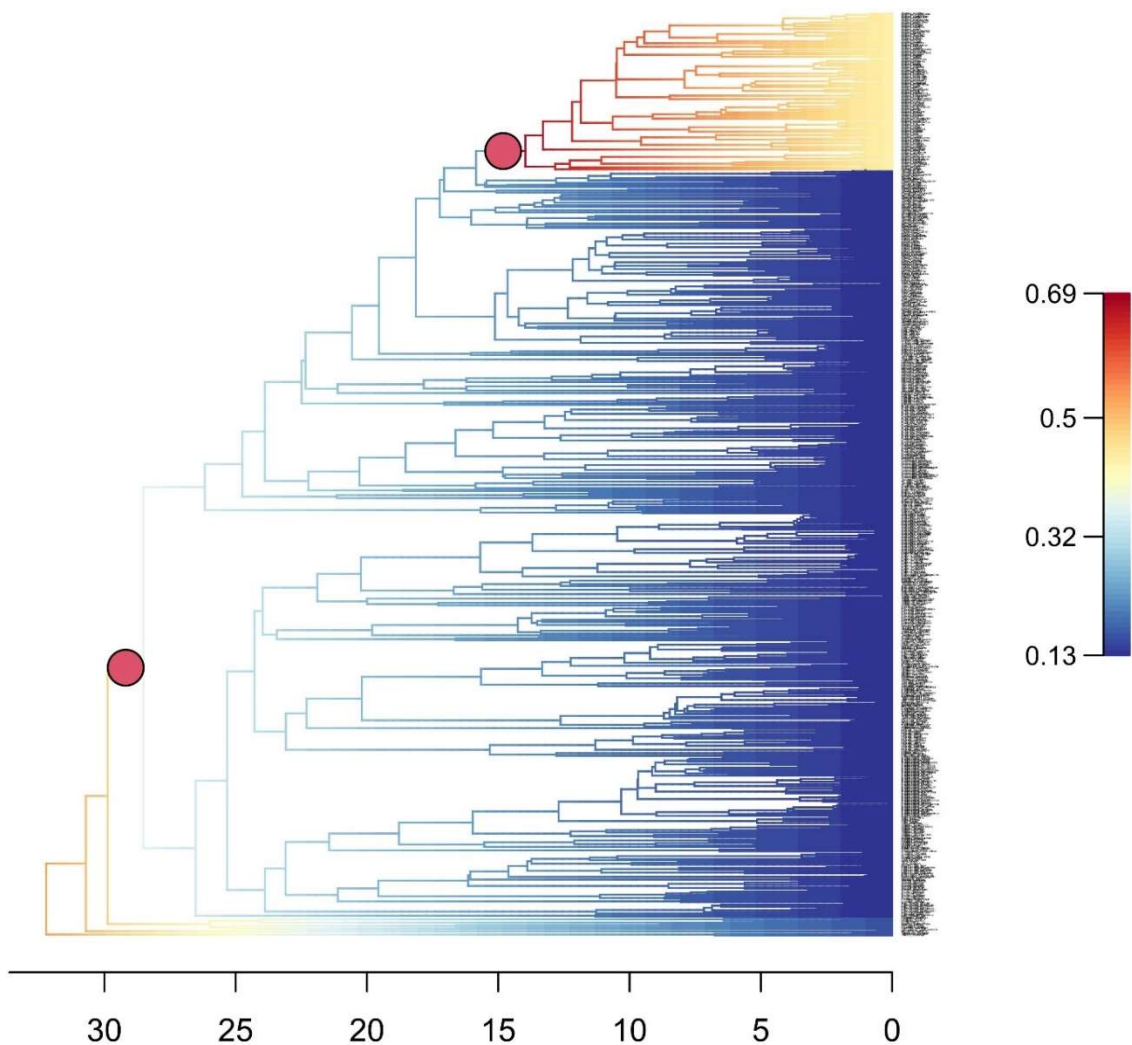


Diversification rates estimated with BAMM

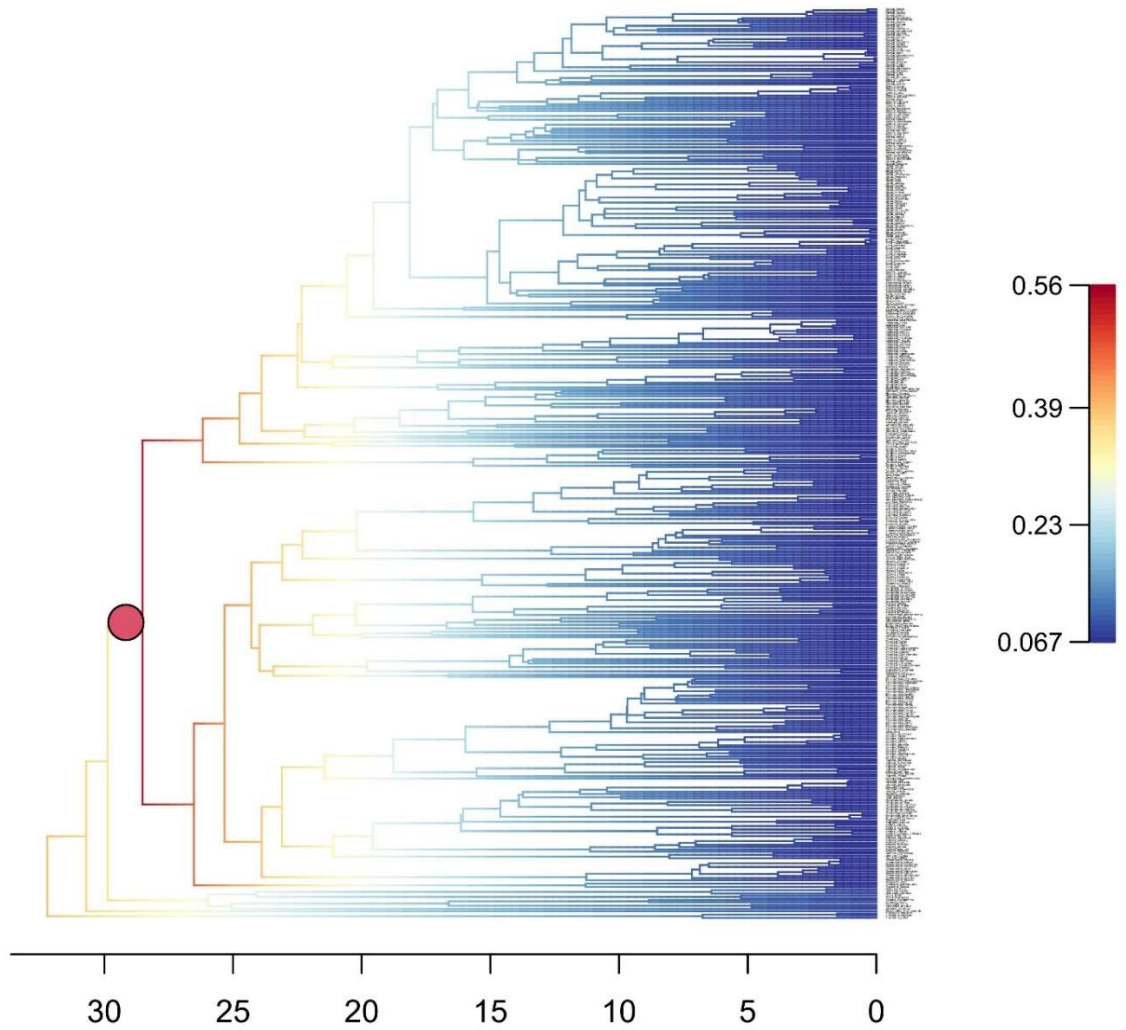


Speciation rates estimated with BAMM

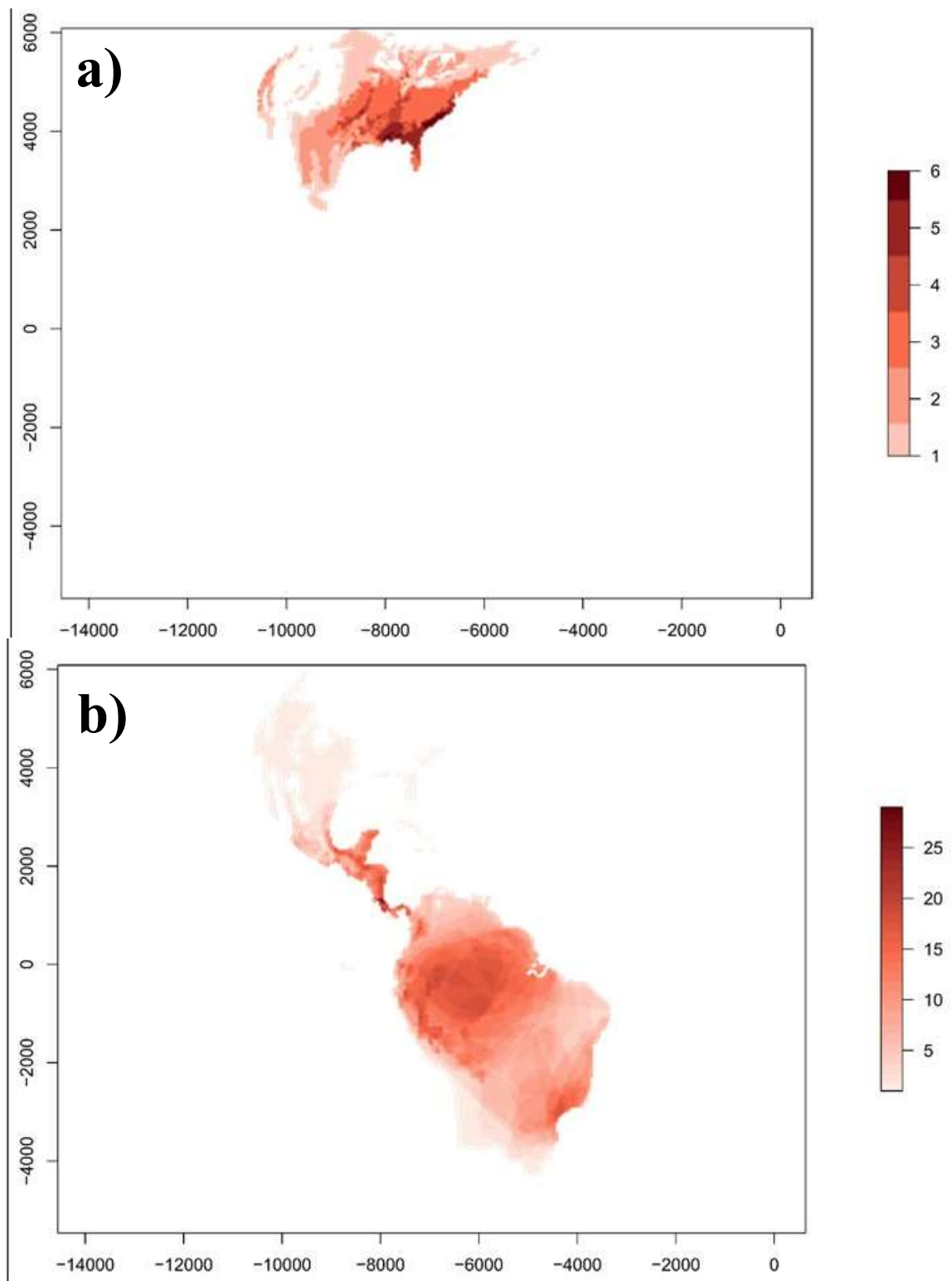
Appendix Figure A2.7. Relationships between the estimated values of the DR metric and those of the per-cell diversification and speciation rates obtained with BAMM. In both cases, we used the Bayesian calibrated phylogeny we estimated herein. To calculate these values, we used the range maps compiled by Roll et al. (2017). After accounting for spatial autocorrelation, the DR metric is strongly correlated with both rates (for net diversification rates: Dutilleul's modified t-test, correlation: 0.9974, F-statistic: 580.7129 on 1 and 3.086 degrees of freedom, p-value: 1e-04, for speciation rates: Dutilleul's modified t-test, correlation: 0.9996, F-statistic: 3741.193 on 1 and 3.0648 degrees of freedom, p-value: 0)

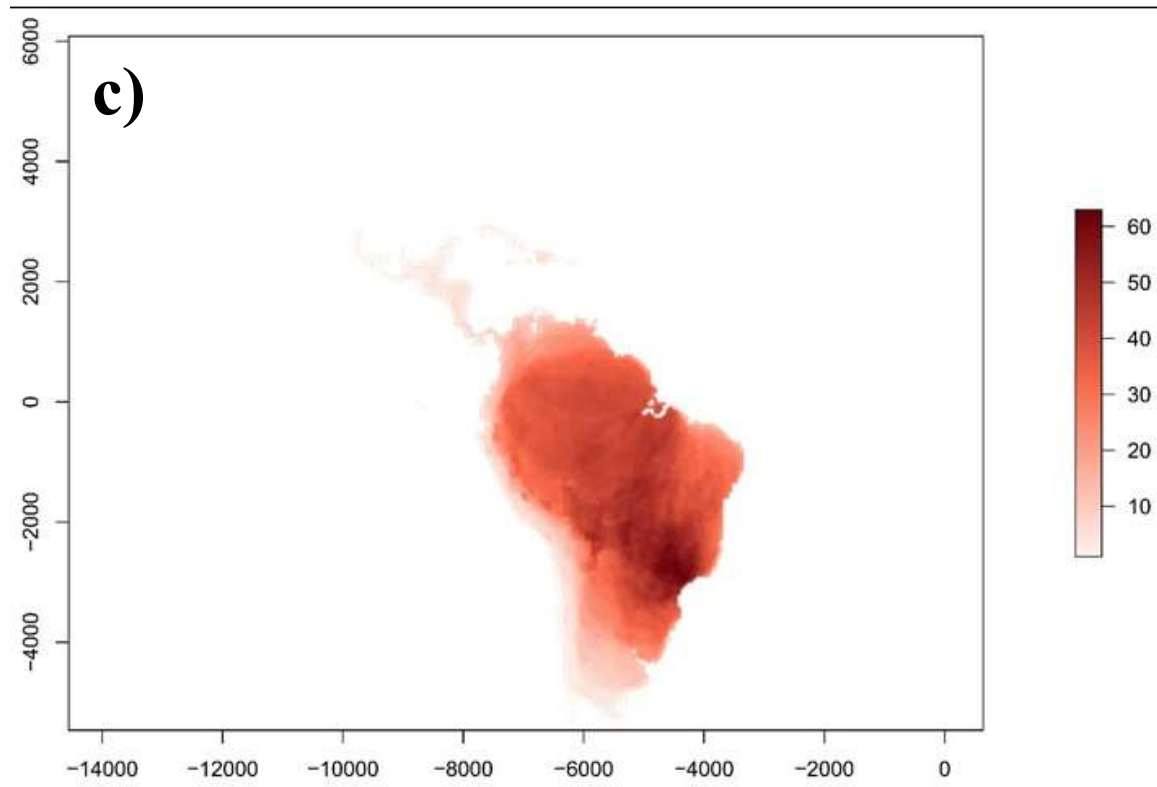


Appendix Figure A2.8. Timetree of the family Dipsadidae with branches colored by the values of net diversification rates (in units of new lineages per million years) estimated with the software BAMM. For this analysis, we accounted for the effect of incomplete taxa sampling by using a randomly selected phylogeny of the 100 completely sampled timetrees of the Dipsadidae we obtained with the software TACT. The red circle indicates the most supported shift in the diversification rates within the Dipsadidae. Warmer and color colors correspond to higher and lower diversification rates, respectively.



Appendix Figure A2.9. Timetree of the family Dipsadidae with branches colored by the values of net diversification rates (in units of new lineages per million years) estimated with the software BAMM. For this analysis, we did not account for the effect of incomplete taxa sampling. The red circle indicates the most supported shift in the diversification rates within the Dipsadidae. Warmer and color colors correspond to higher and lower diversification rates, respectively.





Appendix Figure A2.10. Patterns of species richness of the subfamilies of the snake family Dipsadidae across the New World, as estimated from the range maps compiled by Roll et al. (2017). Panel a) shows the distribution of the subfamily Carphophiinae and panels b) and c) to the distribution of the subfamilies Dipsadinae and Xenodontinae, respectively. Darker and paler tones of red correspond to areas with higher and lower species richness, respectively. The map is overlaid on a 50 km x 50 km grid using an equal-earth projection. Cells with less than 25% of land surface are not mapped.

Appendices Chapter 3

Appendix 3.1 – extended methods.

Phylogenetic reconstruction and ancestral area estimation.

First, to build a phylogenetic tree of snakes of the family Dipsadidae, we compiled all sequences corresponding to five mitochondrial (12S, 16S, *cyt b*, ND2, and ND4) and four nuclear (BDNF, *c-mos*, NT3, and RAG2) genes publicly deposited in GenBank (up until October 2018). We also incorporated sequences available in Zaher et al. (2018), not uploaded into GenBank. We also included sequences from 24 species from 14 different families of Caenophidia to be used as outgroups. Sequence curation involved the identification and removal of sequences not correctly assigned to species (i.e., likely cases of contamination, miss-labeling of sequences, or misidentifications of specimens), as well as the deletion of ambiguously aligned positions of the 12S and 16S ribosomal genes. Furthermore, sequences of terminals recovered in phylogenetically unexpected places (or associated with very long branch lengths) were compared with GenBank sequences by using the Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990), and we filtered those erroneously assigned to species. After deleting sequences recovered as being incorrectly identified, we concatenated all remaining sequences with the Sequence Matrix software v1.7.6 (Vaidya et al., 2011). Detailed information on all the samples used for this work is available in Suppl. Table S3.6.

We then aligned sequences with the online version of MAFFT v7 (Kato & Standley, 2013) under default parameters, except for the case of the 12S and 16S ribosomal genes, for which we used the Q-INS-i algorithm (which considers the

secondary structure of RNA sequences, Katoh & Toh, 2008). We posteriorly inspected alignments with the software Mesquite v3.2 (Maddison & Maddison, 2017). The final aligned matrix consists of 5978 nucleotide positions and a total of 341 terminals. The software PartitionFinder 2 (Lanfear et al., 2016) was used to select the best-fitting partition scheme for the alignment according to the corrected Akaike Information Criterion (AICc) and the greedy algorithm of Lanfear et al. (2012). Partitions were defined a priori by gene identity, as well as by codon position for protein-coding genes (all included genes except 12S and 16S). The best partitioning scheme for each of the subsets is shown in Appendix Table A3.7.

We performed Bayesian phylogenetic analysis (BA) with BEAST 2 v4.8 (Bouckaert et al., 2014). The bModelTest package of BEAST 2 was applied to simultaneously explore the substitution model space and estimate both the corresponding model parameters and the phylogeny (Bouckaert & Drummond, 2017). We estimated the posterior distribution of trees, including dates and rates, by assuming a pure-birth or Yule model tree prior (a diversification model that considers extinction rates as negligible in comparison to speciation rates; Yule, 1924; Ritchie et al., 2017). A relaxed normal clock model of evolution was used and calibrated with five fossil calibration points (detailed in Appendix Table A3.8). As a starting tree for this analysis, we used the phylogeny with the highest Maximum Likelihood value we obtained after running on the concatenated dataset the software RAXML-HPC V.8 (Stamatakis, 2014), as implemented in the CIPRES Science Gateway v. 3.3 (Miller et al., 2010). For obtaining this phylogeny, we used the rapid bootstrap analysis/search for best-scoring ML tree (-f a) algorithm with

1000 bootstrap replicates and the GTRCAT model of sequence evolution for all partitions. The resulting phylogeny was then converted to an ultrametric tree using the previously mentioned calibration points and the nonparametric rate smoothing method implemented in the function 'chronos' of the R package ape 5.5 (Paradis & Schliep, 2019). The sampling of the Beast 2 analysis used two independent MCMC chains of 50 million generations each, saving one sampled tree per 1000 generations and with the first 50% of trees discarded as burn-in. Stationarity and mixing of the log-likelihood values and parameter estimates were evaluated with Tracer v1.6 (Rambaut et al., 2018) and corroborated by verifying that all parameters of the post-burn-in set of trees had effective sample sizes > 200. The posterior distribution of trees was summarized as a Maximum Clade Credibility (MCC) timetree using TreeAnnotator v2.4.8, part of the BEAST 2 package. The resulting calibrated phylogeny is shown in Appendix Figure 3.2.

Next, to determine the number and the direction of dispersal events between Central America and South America, we estimated ancestral areas by implementing a model-based approach using the resulting timetree. Terminals of the tree were coded according to their range in eight different biogeographic areas (Nearctic region; Northern Central America; Antillean region; North Pacific region; South Pacific region; Amazonian region; Chaco-Parana region, and the southern South America region). The scheme of biogeographic areas used corresponds to a simplified version of regionalization of the Neotropics by Morrone (2014) (Appendix 2 and Appendix Figure 3.3 for details). For determining the areas inhabited by each species of our phylogeny, we used the distribution maps compiled by Roll et al. (2017) or the original descriptions of

recently described taxa. The areas inhabited by each species are summarized in Suppl. Table S3.7. We used the R package ‘BioGeoBEARS’ (Matzke, 2013a, 2014) to evaluate the fit of six biogeographical models (DEC, DEC-J, DIVA, DIVA-J, BAYAREA-LIKE, BAYAREALIKE-J) to our data based on the corrected Akaike information criterion (AICc). The model recovered as having the best fit was then used for conducting the ancestral area estimation is shown in Appendix Figure A3.2. We considered taxa as being involved in the GABI if they are part of a lineage whose ancestral area is located in a different subcontinent than what was inferred for its immediate ancestor and if that dispersal event occurred more recently than 15 million years ago. Using 3 million years ago as a cut-off did not change the results of the analyses performed for this study (results not shown). The results of the ancestral area estimation also allowed us to obtain a rough estimate of the time in which each of the included sites was colonized first by snakes of the family Dipsadidae. To do so, the time in which dipsadids first arrived at a given site was assumed to be the same as the time of the first colonization of the region in which it is located. We used the time of first colonization of each region and not the ages of all posterior colonization events, following Pinto-Sanchez et al. (2014) and the studies cited therein, which show that both approaches give similar results.

Functional traits.

Among the traits we considered, microhabitat use is a categorical trait that describes if a species is arboreal, semiarboreal, fossorial, semi fossorial, terrestrial, aquatic, or semiaquatic. Time of activity is also a discrete trait categorizing species as diurnal, nocturnal, or both. Reproductive mode considers taxa as oviparous, viviparous,

with a single species exhibiting both reproductive modes. Finally, the diet category contains the following non-exclusive subcategories: taxa feeding on amphibian eggs, reptile eggs, bird eggs, hard-bodied invertebrates (i.e., terrestrial arthropods), soft-bodied invertebrates (i.e., oligochaetes and terrestrial gastropods, independently if they have a shell or not), amphibians, non-avian reptiles, endotherms, and fishes. Values of total length and mass were obtained from Feldman et al. (2016). However, it is important to clarify that most body mass data correspond to estimates obtained with clade-specific equations that allow estimating the body mass of a given snake species using only its body length (see Feldman et al., 2016 for more details). These estimations were necessary given the paucity of data about the body mass of species of snakes in general (Feldman & Meiri, 2012).

We acknowledge that the different discrete categories in which we classified the behavior of taxa into the functional traits used herein are subjective to some degree. Specifically, incidental records (such as the recent observation of climbing, diurnal behavior in an aquatic, nocturnal snake; de Lima Moraes & de Fraga, 2015) could have biased how we coded the trait states of at least some taxa. Such limitation is especially relevant for poorly known species, for which there is less information to confirm if a given behavior is within the normal variation they exhibit. For accounting for this limitation, we only coded taxa into a particular dietary category if a given diet item corresponded to more than 10% in frequency to the total diet items reported for the species. Unfortunately, quantitative information about microhabitat usage or time of activity was more limited, preventing us from using similar criteria. On the other hand,

our classification of behaviors into trait states is also subjective in the sense that slightly different categories have been defined and used in previous studies (such as in Rodriguez-Robles & de Jesus-Escobar, 1999; Colston et al., 2010; Feldman & Meiri, 2012; Burbrink & Myers, 2015; Harrington et al., 2018). For example, in contrast to other trait compilations, we consider a diet based on endotherms as a single category instead of having two separate groupings, one for each of the two groups of endotherms that exist. However, after our literature review, we found that the majority of the snakes considered herein that eat birds also eat mammals. Similar reasoning was followed for defining other of the trait states used for this work. Finally, some of the traits considered herein can vary ontogenetically, seasonally, and between different populations of the same species (Henderson et al. 1979; Shine, 1994; Rodriguez-Robles & de Jesus-Escobar, 1999). Fortunately, for this study, we found limited evidence for such types of variability. However, we acknowledge that limitations in our functional data exist, and our patterns are based on the best information available to date.

Lastly, we recognize there are other traits that we did not consider that could also be mediating the interactions between dipsadid snakes and the environment they inhabit. Examples of these characters include other morphological traits (such as head width and length [e.g., Moon et al., 2019; Pianka et al., 2017], tail length [e.g., Feldman & Meiri, 2012], or thickness of the skin [e.g., Shine et al., 2019]), qualitative life-history traits (such as home-range size and movement, age of sexual maturity [e.g., Grimm et al., 2014], clutch size and its frequency [e.g., Shine, 1994; Pianka et al., 2017], either sit-and-wait or active foraging modes [e.g., Colston et al., 2010; Pianka et al., 2017]), and

physiological traits (such as resting metabolic rate, or active body temperature [e.g., Pianka et al., 2017]). Unfortunately, the limited availability of data about these traits for most members of the family prevented us from using them to calculate the functional dispersion of the communities used for this work. Moreover, extracting functional trait data from the literature results in additional potential sources of bias. Specifically, we would expect that intraspecific trait variation, misidentifications, and cryptic species could prevent us from correctly coding the traits of taxa inhabiting each community of interest. Although these limitations are more likely to occur in widespread taxa (e.g., Chenuil et al., 2019), they have the potential to bias any taxa. We set to minimize the impact of potential misidentifications by verifying that all the communities included herein are not outside the range of the species inhabiting them.

Finally, we coded the functional traits data of species not identified unambiguously (e.g., with abbreviations such as sp., cf., or aff.) based on their characteristics exhibited by their closest relatives. Specifically, for species identified only at the generic level, we assumed that their functional trait values correspond to the most common character states exhibited among their congeners. There were only 13 taxa for which we could not obtain or make assumptions about their traits and which were removed from all analyses involving functional dispersion as a result. Such is the case, as these species are of uncertain phylogenetic relationships or part of genera exhibiting a considerable variation in the functional traits amongst their members. These species correspond to four dipsadids (*Cercophis auratus*, *Thamnodynastes* sp. 2 and T. sp. 3., and *Pseudoleptodeira latifasciata*) and nine non-dipsadid snakes (Colubridae sp.nov.,

Colubrid sp., *Rena myopicus*, *Rena bressoni*, *Liotyphlops*
anops, *Epictia* cf. *signata*, *Epictia antoniogarciae*, *Epictia albifrons*, and *Micrurus*
tricolor).

References in Appendix 3.1 not included in the main text

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403-410.
- Chenuil, A., Cahill, A. E., Délémontey, N., Du Salliant, E. & Fanton, H. (2019). Problems and questions posed by cryptic species. A framework to guide future studies. Pages 77-106 in E. Casetta, J. Marques da Silva, & D. Vecchi (Eds.), *From assessing to conserving biodiversity*. History, philosophy and theory of the life sciences (vol 24). Springer Berlin, Germany.
- Colston, T. J., Costa, G. C., & Vitt, L. J. (2010). Snake diets and the deep history hypothesis. *Biological Journal of the Linnean Society*, 101(2), 476-486.
- de Lima Moraes, L. J. C., & de Fraga, R. (2015). Climbing behavior in the aquatic snake *Helicops hagmanni* Roux 1910 (Serpentes: Dipsadidae). *Herpetology Notes*, 8, 19-20.
- Grimm, A., Ramírez, A. M. P., Moulherat, S., Reynaud, J., & Henle, K. (2014). Life-history trait database of European reptile species. *Nature Conservation*, 9, 45.
- Harrington, S. M., De Haan, J. M., Shapiro, L., & Ruane, S. (2018). Habits and characteristics of arboreal snakes worldwide: arboreality constrains body size but does not affect lineage diversification. *Biological Journal of the Linnean Society*, 125(1), 61-71.
- Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29(6), 1695-1701.
- Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2016). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34(3), 772-773.
- Matzke, N.J. (2013a) 'BioGeoBEARS': biogeography with Bayesian (and likelihood) evolutionary analysis in R Scripts. University of California, Berkeley, CA. R package, version 0.2.1. Available at: <http://CRAN.R-project.org/package=BioGeoBEARS> (accessed 21 February 2018).
- Miller, M.A., Pfeiffer, W., & Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees, pages 1-8 in: *Proceedings of the Gateway Computing Environments Workshop (GCE)*, 14 Nov. 2010, New Orleans, LA.

- Moon, B. R., D. A. Penning, M. Segall & A. Herrel. 2019. Feeding in Snakes: form, function, and evolution of the feeding system. Pp. 527– 574 in V. Belsand I. Whishaw, eds. *Feeding in vertebrates*. Springer, Cham, Switzerland.
- Morrone, J. J. (2014). Biogeographical regionalisation of the Neotropical region. *Zootaxa*, 3782(1), 1-110.
- Maddison, W. P., & Maddison, D. R. (2017). Mesquite: a modular system for evolutionary analysis. Version 3.2. 2017.
- Paradis E & K. Schliep (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526-528.
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67(5), 901.
- Ritchie, A. M., Lo, N., & Ho, S. Y. (2017). The impact of the tree prior on molecular dating of data sets containing a mixture of inter-and intraspecies sampling. *Systematic Biology*, 66(3), 413-425.
- Rodriguez-Robles, J.A. & De Jesus-Escobar, J.M. (1999). Molecular systematics of New World lampropeltine snakes (Colubridae): Implications for biogeography and evolution of food habits. *Biological Journal of the Linnean Society* 68:355–385.
- Roll, U., Feldman, A., Novosolov, M., Allison, A., Bauer, A. M., Bernard, R., ... & Colli, G. R. (2017). The global distribution of tetrapods reveals a need for targeted reptile conservation. *Nature Ecology & Evolution*, 1(11), 1677.
- Shine, R., Goiran, C., Shilton, C., Meiri, S., & Brown, G. P. (2019). The life aquatic: an association between habitat type and skin thickness in snakes. *Biological Journal of the Linnean Society*, 128(4), 975-986.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312-1313.
- Vaidya, G., Lohman, D. J., & Meier, R. (2011). SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, 27(2), 171-180.
- Yule G.V. 1924. A mathematical theory of evolution based on the conclusions of Dr. J. C. Willis, P.R.S. *Philosophical Transactions of the Royal Society B* 213:21–87.

Zaher, H., Yáñez-Muñoz, M. H., Rodrigues, M. T., Graboski, R., Machado, F. A., Altamirano-Benavides, M., ... & Grazziotin, F. G. (2018). Origin and hidden diversity within the poorly known Galápagos snake radiation (Serpentes: Dipsadidae). *Systematics and Biodiversity*, 16(7), 614-642.

Appendix 3.2 – Further methodological details

Our definition of biogeographic areas was mostly defined based on Morrone's (2004) regionalization of the Neotropics. However, we separated his Pacific Dominion into two regions separated by the boundary between the Choco-Darien and the Western Ecuador provinces. We also did not consider the Paramo province as part of his South American transition zone. Instead, we classified the species found in that region as being found on the Pacific or the Boreal Brazilian Dominion if their range covers the Western or Eastern slopes of the Andes, respectively. We also merged several of the regions he defined to reduce the total number of areas considered. In the first place, we merged his Mexican transition zone with his Mesoamerican dominion to form a single area we name the Northern Central American region. We also combined his Chacoan and Parana dominions and his southern part of the South American Transition zone and the antarctic region, respectively, to reduce these four areas into only two. We named the resulting two regions as the Chaco-Parana region and the southern South American regions. Finally, we also merged both the Boreal Brazilian and Southern Brazilian Dominions with his South-Eastern Amazonian Dominion to generate an area we name the Amazonian region. In summary, the resulting eight areas are the following: 1) Nearctic region; 2) Northern Central America; 3) Antillean region; 4) North Pacific region; 4) South Pacific region; 5) Amazonian region; 6) Chaco-Parana region, and the 7) southern South America region.

Reducing the number of biogeographic areas was necessary, as the distribution range of Dipsadidae includes all the ten dominions and transition regions originally defined by Morrone (2014) for the Neotropics, as well as three extra-tropical areas (Holarctic, Antarctic, and Asian regions). Performing biogeographical analyses to a dataset containing such a number of areas and taxa will take unreasonable amounts of time and computational resources, which are further increased by the large number of species considered. Furthermore, as the range of most species covers at least two or three of these regions, ancestral range estimations would also be likely to be very broad and uninformative if many relatively small areas are defined.

Additional details of the results from the biogeographic analysis

We found that the two Neotropical subfamilies of Dipsadidae diverged at 26.98 million years ago (95% Confidence Interval, CI, 22.14-32.56 Mya), with the Most Recent Common Ancestor (MRCA) of both clades inhabiting South America (with 84.0% probability; Suppl. Figure S3.1). We recovered that both the MRCA of all Dipsadinae and the MRCA of the early-diverging tribe Diaphorolepini inhabited South America at 24.49 Mya (CI, 29.8-19.88 Mya) or 14.69 Mya (CI, 19.16-10.72 Mya), respectively (with probabilities of 78.6% and 93.3%, respectively; Suppl. Figure S3.1). The MRCA of the remaining members of the Dipsadinae was recovered as inhabiting Central America (with 83.9% probability; Suppl. Figure S3.1) at 22.88 Mya (CI, 27.97-18.51 Mya), with many descending nodes being recovered as exclusively inhabiting Central America. However, members of eight clades were recovered as later dispersing to South America after 14.23 Mya (summarized in Appendix Table A3.6), with one additional lineage (*Nothopsis*

rugosus) potentially doing so even as earlier as 19.33 Mya. Interestingly, within those secondarily South American groups, we found three instances of taxa dispersing back to Central America after 1.38 Mya, and potentially even after 12.68 Mya. In general, most dispersal events of members of Dipsadinae across the Panamanian Isthmus involved only a few species, except for the dispersal of the MRCAs of the members of *Dipsas* and *Atractus* (both genera have more than 30 species).

In contrast, we recovered the MRCA of all members of Xenodontinae as inhabiting South America (with 98.6% probability, Suppl. Figure S3.1) at 25.09 Mya (CI, 30.42-20.57 Mya). We found evidence of ten dispersal events from South to Central America involving members of Xenodontinae (Appendix Table A3.6), each made by clades containing only one or two species. Most of these dispersal events occurred after 8.02 Mya, but there are three taxa that could have dispersed as earlier as 20.48 Mya (Appendix Table A3.6). See Suppl. Figure S3.1 for the ancestral area estimations of other major nodes along the phylogeny of Dipsadidae, including those not involved in the GABI. In general, although we considered most of the dispersal events across the Isthmus made by members of Dipsadidae, there are still a few genera distributed in both North and South America that we could not include in our analyses (*Enulius*, *Enuliophis*, and *Urotheca* of the subfamily Dipsadinae). Despite this, preliminary results of an analysis performed including sequences of these genera, suggest that the inclusion of these taxa does not change the conclusions reported herein. Further, there are a few genera from which sequence data is only available from species present on one of the sides of the

Isthmus (e.g., *Rhadinaea* and *Geophis* of the subfamily Dipsadinae, and *Phimophis* of the subfamily Xenodontinae), and that would need to be included in future analyses.

In general, we could not infer the geographical origin of 98 species (46.7% of the total number of species across all communities), as there were no sequence data available that would allow us to include them in our phylogenetic analyses. However, for 16 of these species, we could infer their probable area of origin by assuming that they originated in the same continent (=either South America, Central America, or both) as their closest relatives. Indeed, for taxa identified with abbreviations such as cf. or aff. we assumed that their geographic origin was the same as that of their closest relatives. Further, for species not identified below the generic level (i.e., Genus sp.), we assumed that their geographic origin was the same as its congeners only if they are exclusively known from a given continent. Species for which this was not the case were categorized as of unknown origin. The results of our analyses did not vary depending on whether these 16 species were considered or not of unknown origin. In general, the number of taxa per site from which we could not infer or assume their biogeographic origin was variable ($\bar{x}$ =23.1%, range=0-100%, 83.58% of all communities have percentages of missing data lesser than 45%).

Exploring possible alternative patterns consistent with the idea of communities being saturated:

We also evaluated if on average, the taxa inhabiting sites involved in the GABI have increased diversification (or speciation) rates than uncolonized sites. To do so, we

estimated the per-species net diversification rates of the taxa inhabiting each of the communities of the Dipsadidae studied in this work using the software BAMM 2.5.0 (Rabosky, 2014). BAMM allows the simultaneous estimation of the total number of shifts in evolutionary rate regimes found across the dipsadid phylogenetic tree we inferred, as well as the diversification rates occurring between each shift. The Markov Chain Monte Carlo process implemented in BAMM allows determining the posterior distribution of likely evolutionary rate regimes (=a mixture of density-dependent and constant diversification rates) and the number of shifts occurring among them (Rabosky, 2014; Seeholzer et al., 2017). The BAMM analysis was run for ten million generations after a burn-in of 10%, with other parameters of the run were left in their default values. We determined the convergence of the BAMM analyses by verifying that the effective sample size (ESS) of the log-likelihood and of the estimated number of shifts were both greater than 200. We used the R package ‘BAMMtools’ (Rabosky et al., 2014) for visualizing the obtained evolutionary regime shifts and diversification rates across the input tree. We then obtained the net diversification rates (as well as the speciation and extinction rates by separate) of each species of the input tree by using the *getTipRates* function of BAMMtools (Rabosky et al., 2014). After obtaining the diversification rates of each species, we then averaged those values across the members of each of the communities of interest to obtain community-wide diversification rates (shown as Appendix Table A3.9).

On the other hand, we also evaluated if not finding differences in the functional dispersion between sites involved or not in the GABI could have a different explanation

from them being saturated. Such an explanation corresponds to the possibility that functional dispersion values increased more after colonization among communities that originally had more reduced trait spaces. Finally, an additional testable prediction of this explanation is that in at least some communities, the value of functional dispersion calculated while including the colonizing taxa would have been larger than when excluding such lineages. We tested this prediction by calculating the functional dispersion of each community after removing the taxa colonizing them after the GABI, following the methodology we described in the Materials and Methods of the main text. We then used a paired Welch's t-test (as implemented in the software R) to compare these values to those that we had obtained while not excluding the colonizers in each community. The results of this t-test ($t=0.846$, 76.71 degrees of freedom, $p\text{-value}=0.4$) indicate that excluding the species colonizing each community does not significantly change the values of functional dispersion obtained when considering such species in the calculation. The values of the functional dispersion of each community calculated after removing the species involved in the GABI are presented in Appendix. Table A3.10.

We looked for patterns in several variables that could suggest that the dispersal events following the GABI could have increased the species richness or functional dispersion of colonized sites to the level of uncolonized sites. In other words, we investigated if differences in several variables between communities involved or not in the GABI could explain why, at some point in the past, they might have differed in their richness or functional dispersion. The variables we compared between colonized and uncolonized sites are the same as the covariates of the linear models we used to test for

community saturation (i.e., trap use, sampling effort, local climatic conditions, time since first colonization by members of Dipsadidae, species richness and functional dispersion of other groups of snakes). However, we also compared three additional variables not included in the original models: site area, latitude, and distance to the Panamanian Isthmus. The last variable was calculated by measuring the distance between each site and a randomly selected point located within the limits of the “Choco-Darien” province of Morrone (2014). We then used paired Welch’s tests to compare the variables of interest between the sites that were involved or not in the GABI. The obtained results are summarized in Appendix Table A3.11.

References in Appendix 3.2 not included in the main text

- Rabosky, D. L. (2014). Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PloS One*, 9(2), e89543.
- Rabosky, D. L., Grundler, M., Anderson, C., Title, P., Shi, J. J., Brown, J. W., ... & Larson, J. G. (2014). BAMM tools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*, 5(7), 701-707.
- Seeholzer, G. F., Claramunt, S., & Brumfield, R. T. (2017). Niche evolution and diversification in a Neotropical radiation of birds (Aves: Furnariidae). *Evolution*, 71(3), 702-715.

Appendix Table A3.1. Information about the 67 communities of snakes of the family Dipsadidae included in this study, including their area, coordinates, elevation, species richness, functional dispersion, and the number of species from taxa colonizing from the other side of the Isthmus that they contain. The values of two climatic variables (Annual Precipitation, and Mean Temperature) of each site, as well as the sampling effort needed to estimate their species composition, are also presented. We also report the species richness and calculated functional dispersion of snakes from families different from Dipsadidae present on each site. The sources used for compiling the information reported in this table are also reported.

Num ber in map	Country	Locality	Latit ude	Longitu de	Elevati on (m)	Richne ss	FDi s	species from groups involv ed in the GABI	Mean Temperat ure (C)	Annual precipitati on (mm)	Area (ha)	Sampli ng effort (hours- person)	Source
1	Mexico	Hidalgo, Ixmiquilpan, valle del Mezquital	20.6	-99.2333	1500	5	0.1 6	0.00	17.31	467	96163	462	Fernández-Badillo et al. (2010)
2	Mexico	Hidalgo, Reserva de la Biosfera Barranca de Metztlán	20.49 736	-98.6678	250	2	0.1 9	0.00	19.95	600	NA	672	Vite-Silva et al. (2010)
3	Mexico	Michoacán, Tacámbaro	19.25 833	-101.475	1950	6	0.1 9	0.00	17.67	1065	1100	NA	Medina- Aguilar et al. (2011)
4	Mexico	Veracruz, Sierra de Los Tuxtlas	18.58 333	-95.075	205	7	0.1 6	1.00	25.47	2454	NA	NA	Urbina- Cardona & Reynoso (2005)

5	Mexico	Veracruz, Los Tuxtlas, Parque de Flora y Fauna Silvestre Tropical	18.45	-95.0333	2000	5	0.1 7	0.00	24.38	1976	221000	NA	Morales- Mávil et al. (2001)
6	Mexico	Guerrero, West Sierra Madre del Sur	17.74 741	-101.037	140	3	0.0 8	0.00	22.58	1264	NA	NA	Almazán- Núñez et al. (2018)
7	Mexico	Oaxaca, San Juan La Lana, 278 surroundings of Santiago Jalahui	17.45 278	-95.7692	1410	5	0.1 8	1.00	24.64	2478	NA	NA	Rendón et al. (1998)
8	Mexico	Chiapas, microcuenca Bombana	16.94 806	-93.02	777.5	3	0.1 5	0.00	22.51	1406	NA	2320	Percino- Daniel et al. (2013)
9	Guatemala	El Peten, Chinajá (town)	16.03 3	-90.217	1100	7	0.1 5	0.00	25.34	3981	59800	NA	Duellman (1963)
10	Mexico	Oaxaca, sierra de Cuatro Venados	15.93 523	-97.1325	500	2	0.1 5	0.00	26.21	892	2160	NA	Martín- Regalado et al. (2016)
11	Guatemala	Alta Verapaz, Finca Chichen	15.59 4	-90.149	443	5	0.0 6	0.00	17.57	2846	4	NA	Stuart (1948)

12	Honduras	Atlantida and Yoro, Cordillera Nombre de Dios, Refugio de Vida Silvestre Texíguat and Parque Nacional Pico Bonito	15.52 5	-87.37	60	10	0.1 4	0.00	23.87	1204	86	NA	Townsend et al. (2012)
13	Honduras	Copán, subcuenca del río Copán	14.84 167	-89.0583	1000	9	0.1 6	0.00	22.65	1537	7800	NA	Alemán et al. (2013)
14	Guatemala	Suchitepéquez, Patulul, Finca Universitaria San Julián	14.46 248	-91.1327	1725	8	0.1 6	0.00	24.06	3840	NA	7280	Guerra-Centeno et al. (2015)
15	Nicaragua	Región Autónoma del Atlántico Norte, Waslala	13.33 333	-85.3667	333	6	0.1 8	0.00	23.27	2073	5242	507.88	Orozco et al. (2015)
16	Nicaragua	Río San Juan, El Quebracho Private Wildlife Reserve	11.16 611	-84.4058	82	4	0.1 6	0.00	25.72	2973	1500	NA	Barquero et al. (2010)
17	Colombia	Guajira, arroyo Palomino	10.97 028	-72.7582	1200	2	0.0 8	2.00	27.90	1112	NA	313.6	Blanco-Torres et al. (2013)

18	Venezuela	Zulia, approximate y 20 Km S of Maracaibo	10.58 815	-71.71	1550	4	0.1 9	1.00	27.78	755	54700	60	Larreal et al. (2012)
19	Costa Rica	Alajuela, San Ramón, Reserva Biológica Alberto Brenes	10.21 9	-84.597	1300	12	0.1 8	1.00	21.43	3434	559	NA	Bolaños & Ehmcke (1996)
20	Colombia	Magdalena, Sabanas de San Ángel, “La Gloria project”	10.17 478	-74.3272	175	10	0.1 8	1.00	27.30	1386	NA	567	Angarita-M. et al. (2015)
21	Costa Rica	Cartago, Rio Macho Biological Station	9.766 67	-83.85	410	4	0.1 9	4.00	18.52	2850	2.00E+05	1440	Acosta-Chaves et al. (2015)
22	Costa Rica	San José & Puntarenas, Parque Nacional Carara	9.756 111	-84.6075	525	18	0.1 8	0.00	26.04	2193	3300	149	Laurencio & Malone (2009)
23	Colombia	Cesar, ciénaga de Zapatosa	9.291 667	-73.7917	340	8	0.2 0	0.00	28.29	1676	300	31200	Medina-Rangel et al. (2015)
24	Panama	Bocas del Toro, Bosque Protector Palo Seco	9.156 821	-82.5184	200	11	0.1 8	0.00	23.19	3023	NA	576	Hofer & Bersier (2001)

25	Panama	Panama, Barro Colorado Island	9.15	-79.85	27	19	0.1 6	0.00	26.42	2394	NA	NA	Rand & Myers (1990)
26	Colombia	Magdalena, El Banco, corregimiento Botillero	9.004 722	-73.9714	88	5	0.2 0	0.00	28.59	2024	200	110	Dueñez- Gómez et al. (2004)
27	Venezuel a	Barinas, Ramal de Calderas, Sector San Ramón	8.867 722	-70.4934	189	3	0.1 2	0.00	22.39	1459	3378	NA	Barrio- Amorós & Molina (2011)
28	Colombia	Cesar, Tamalemequ e	8.858 674	-73.8079	344	6	0.1 9	0.00	28.52	2104	NA	115	Ruiz (2011)
29	Colombia	Córdoba, Valencia, corregimiento de San Rafael del Pirú	8.202 767	-76.1776	325	8	0.2 0	0.00	27.02	2030	4000	NA	Ruiz (2014)
30	Colombia	Valle del Cauca, Reserva de Yotoco	3.833	-76.333	200	10	0.1 8	1.00	23.31	1389	196	25944	Castro- Herrera et al. (2007)
31	Colombia	Valle del Cauca, Buenaventura , Central Hidroeléctric a de bajo Anchicayá	3.731 667	-76.9583	116	3	0.0 9	1.00	25.97	4668	NA	NA	Vargas & Bolaños (1999)
32	Colombia	Huila,	2.454	-75.5594	217	3	0.1	1.00	24.02	1557	NA	NA	Moreno-

		Magdalena River dry valley, Gigante, Garzón, El Agrado and Paicol municipalities	667				5						Arias & Quintero- Corzo (2015)
33	Colombia	Caquetá, La Macarena, Caño Cristales, sendero Laguna El Turista	2.249	-73.758	300	4	0.1 7	0.00	26.22	2465	NA	600	Barrientos et al. (2017)
34	Colombia	Caquetá, San Vicente del Caguán	2.016 666	-74.6417	68	9	0.1 9	0.00	24.91	2429	0.3	NA	Cortes- Ávila & Toledo (2013)
35	Colombia	Guaviare, Calamar, cuenca rio Itilla	1.955 472	-72.8632	168	3	0.1 3	0.00	26.86	2467	7288	880	Osorno- Muñoz et al. (2019)
36	Colombia	Nariño, Tumaco, Dirección General Marítima station (DIMAR)	1.823 839	-78.7266	375	9	0.1 6	1.00	26.14	2658	NA	118	Pinto- Erazo et al. (2019)

37	Ecuador	Esmeraldas, Bilsa Biological Station	0.359 167	-79.7006	16	19	0.1 9	2.00	23.23	1830	14700	240	Ortega-Andrade et al. (2010)
38	Ecuador	Napo, Santa Cecilia region	0.05	-76.9833	139	26	0.1 8	0.00	24.65	3438	69000	32	Duellman (1978)
39	Ecuador	Orellana, Estación de Biodiversidad Tiputini	- 0.618 06	-76.1719	214	23	0.2 1	0.00	25.18	3232	64406	5040	Cisneros-Heredia (2006)
40	Brazil	Pará, Floresta Estadual do Trombetas	- 0.962 77	-55.5223	230	5	0.1 7	0.00	25.30	2382	650	640	Ávila-Pires et al. (2010)
41	Brazil	Amazonas, Japurá river	- 1.782 47	-69.0789	280	18	0.1 8	0.00	26.00	3314	NA	241	Frazão et al. (2020)
42	Peru	Loreto, Teniente Lopez (town)	- 2.594 56	-76.1158	34	2	0.1 2	0.00	25.77	2955	NA	360	Duellman & Mendelson (1995)
43	Brazil	Amazonas, Manaus, Experimental Farm area of the Federal University of Amazonas	- 2.641 46	-60.0954	2200	16	0.1 7	0.00	26.43	2356	NA	756	Masseli et al. (2019)
44	Brazil	Amazonas, Manaus, Reserva Ducke	- 2.966 67	-59.9333	48	21	0.1 9	0.00	26.47	2300	554	128	de Fraga et al. (2011)

45	Brazil	Amazonas, Manaus, INPA-WWF reserves	-3.217	-60.033	-1	19	0.19	0.00	26.80	2181	NA	7200	Zimmerman & Rodrigues (1990)
46	Brazil	Pará, Brasil Novo, Caverna Planaltina	-3.376 22	-52.5761	73	7	0.18	0.00	25.30	1898	450	576	Oliveira et al. (2012)
47	Brazil	Amazonas, upper Rio Urucu (3 Km south of headwaters Rio Urucu)	-5	-65.5	105	6	0.18	0.00	26.38	2504	3000	NA	Gascon & Pereira (1993)
48	Peru	Cajamarca, Santa Rosa de la Yunga	-5.434 47	-78.5715	1850	5	0.17	0.00	21.37	1131	NA	428.07	Koch et al. (2018)
49	Brazil	Paraíba, Rio Tinto and Mamanguape, Guaribas Biological Reserve	-6.740 54	-35.1423	125	15	0.18	1.00	24.96	1389	NA	640	Barbosa et al. (2015)
50	Brazil	Paraíba, João Pessoa municipality, Barra de Gramame	-7.233 47	-34.806	625	11	0.18	0.00	25.67	1854	NA	NA	Sampaio et al. (2018)
51	Brazil	Bahía, São Francisco river Quarternary dunes	-10.76 76	-42.7555	700.5	15	0.16	0.00	25.53	684	NA	528	Rodrigues (1996)

52	Peru	Madre de Dios, Manu National Park, Cocha Cashu	-11.9	-71.367	202	14	0.19	0.00	25.05	2594	124	267.3	Rodríguez et al. (1990)
53	Peru	Madre de Dios, Manu National Park, Pakitza	-11.9442	-71.2831	250	15	0.17	0.00	24.96	2665	NA	136	Morales & McDiarmid (1996)
54	Peru	Madre de Dios, Reserva Cusco Amazonico	-12.5562	-69.0718	805	27	0.19	0.00	25.40	2068	6657	320	Duellman (2005)
55	Peru	Manu, Los Amigos Biological Station	-12.5686	-70.0992	82	30	0.19	0.00	25.03	3455	200	NA	Rabosky et al. (2009)
56	Peru	Parque Nacional y Área Natural de Manejo Integrado Madidi, Serranía Tequeje	-13.7167	-68.2167	316	5	0.17	0.00	25.00	2010	8500	840	Cortez-Fernández (2005)
57	Brazil	Mato Grosso, near Poconé	-16.5	-56.75	250	15	0.19	0.00	25.43	1302	1000	NA	Strussmann & Sazima (1993)

58	Bolivia	Santa Cruz, Santa Rosa del Sara, propiedad Juan Deriba	- 16.86 67	-63.75	425	8	0.1 7	0.00	24.91	1797	NA	1190	Pinto- Viveros et al. (2015)
59	Brazil	Rio do Janeiro, Serra do Mendanha	- 22.82 5	-43.4917	125	13	0.1 9	0.00	20.78	1550	NA	40	Pontes et al. (2009)
60	Paraguay	Canindeyú, Reserva Natural del Bosque Mbaracayú, Aguará Ñu.	- 24.18 33	-55.2667	61	7	0.2 2	0.00	22.21	1626	NA	238	Cacciali et al. (2015)
61	Argentina	Formosa, Municipalida d de Colonia Villafañe, La Picadita, “Estancia La Ascensión”	- 26.03 78	-59.1572	37	5	0.1 8	0.00	22.03	1131	843	2520	Pesci et al. (2018)
62	Paraguay	Itapua, Parque Nacional San Rafael, Kangüery	- 26.51 17	-55.7889	685	8	0.1 9	0.00	21.63	1554	212	1128	Cacciali (2013)
63	Brazil	Rio Grande do Sur, Fazenda da Brigada Militar	- 28.24 42	-52.3283	6	15	0.2 0	0.00	17.23	1773	NA	1604	Zanella & Cechin (2006)

64	Argentina	Corrientes, Reserva Natural del Iberá, Departamento San Martín, colonia Pellegrini	-28.55	-57.1667	80	17	0.20	0.00	20.65	1455	NA	210	Etcheper et al. (2013)
65	Argentina	Corrientes, Bella Vista, Reserva Paleontológica del Arroyo Toropí	-28.5918	-59.0588	121	3	0.06	0.00	20.62	1063	10000	NA	Ingaramo et al. (2015)
66	Brazil	Rio Grande do Sur, Rio Grande, Balneário Cassino	-32.1318	-52.3482	188	12	0.18	0.00	18.60	1275	NA	192	Santos et al. (2012)
67	Argentina	Coastal dunes Buenos Aires province	-37.6167	-57.3333	1184	9	0.17	0.00	14.37	905	NA	212	Kacoliris et al. (2006)

Appendix Table A3.2. Loadings of the first four Principal Component (PC) Axes of the PC analyses performed with the elevation and the values of the 19 WorldClim variables corresponding to each of the sites included in this work.

Environmental Variable	Principal Component			
	1	2	3	4
Elevation (meters)	-0.070	0.049	-0.396	-0.107
Annual Mean Temperature (BIO1)	0.280	0.234	0.115	-0.052
Mean Diurnal Range (BIO2)	-0.220	0.120	-0.078	-0.127
Isothermality (BIO3)	0.249	0.007	-0.194	0.363
Temperature Seasonality (BIO4)	-0.263	-0.050	0.291	-0.289
Maximum Temperature of Warmest Month (BIO5)	0.108	0.339	0.278	-0.308
Minimum Temperature of Coldest Month (BIO6)	0.315	0.124	0.017	0.106
Temperature Annual Range (BIO7)	-0.294	0.042	0.128	-0.281
Mean Temperature of Wettest Quarter (BIO8)	0.229	0.251	0.149	-0.197
Mean Temperature of Driest Quarter (BIO9)	0.296	0.162	-0.010	0.079
Mean Temperature of Warmest Quarter (BIO10)	0.207	0.265	0.316	-0.212
Mean Temperature of Coldest Quarter (BIO11)	0.300	0.188	-0.012	0.072
Annual Precipitation (BIO12)	0.236	-0.279	-0.082	-0.223
Precipitation of Wettest Month (BIO13)	0.221	-0.152	-0.296	-0.367
Precipitation of Driest Month (BIO14)	0.131	-0.361	0.233	-0.011
Precipitation Seasonality (BIO15)	-0.032	0.303	-0.444	-0.182
Precipitation of Wettest Quarter (BIO16)	0.227	-0.172	-0.271	-0.356
Precipitation of Driest Quarter (BIO17)	0.149	-0.364	0.224	0.002
Precipitation of Warmest Quarter (BIO18)	0.137	-0.283	-0.024	-0.335
Precipitation of Coldest Quarter (BIO19)	0.229	-0.184	0.123	0.130

Appendix Table A3.3. Summary of the Principal Component (PC) Axes of the PC analyses performed with the elevation and the values of the 19 WorldClim variables corresponding to each of the sites included in this work.

Component	Total	Proportion of Variance	Cumulative Proportion
PC1	3.0044	0.4513	0.4513
PC2	2.152	0.2316	0.6829
PC3	1.37589	0.09465	0.77753
PC4	1.30871	0.08564	0.86316
PC5	0.93289	0.04351	0.90668
PC6	0.874	0.0382	0.9449
PC7	0.79507	0.03161	0.97648
PC8	0.43051	0.00927	0.98575
PC9	0.35563	0.00632	0.99207
PC10	0.27311	0.00373	0.9958
PC11	0.18619	0.00173	0.99754
PC12	0.13167	0.00087	0.9984
PC13	0.10371	0.00054	0.99894
PC14	0.08948	0.0004	0.99934
PC15	0.08336	0.00035	0.99969
PC16	0.05544	0.00015	0.99984
PC17	0.04842	0.00012	0.99996
PC18	0.02665	0.00004	1
PC19	0.009775	0	1
PC20	3.21E-08	0	1

Appendix Table A3.4. Species richness and functional dispersion of the sympatric snakes from other families present in the 67 communities of snakes of the family Dipsadidae included in this study.

Number in map	Country	Locality	Latitude	Longitude	Elevation (m)	SR of snakes from other families	FDis of snakes from other families	Source
1	Mexico	Hidalgo, Ixmiquilpan, valle del Mezquital	20.6	-99.233	1500	13	0.17023	Fernández- Badillo et al. (2010)
2	Mexico	Hidalgo, Reserva de la Biosfera Barranca de Metztlán	20.4974	-98.668	250	9	0.16696	Vite-Silva et al. (2010)
3	Mexico	Michoacán, Tacámbaro	19.2583	-101.48	1950	27	0.21016	Medina- Aguilar et al. (2011)
4	Mexico	Veracruz, Sierra de Los Tuxtlas	18.5833	-95.075	205	6	0.20896	Urbina- Cardona & Reynoso (2005)
5	Mexico	Veracruz, Los Tuxtlas, Parque de Flora y Fauna Silvestre Tropical	18.45	-95.033	2000	7	0.11641	Morales-Mávil et al. (2001)
6	Mexico	Guerrero, West Sierra Madre del Sur	17.7474	-101.04	140	12	0.16901	Almazán- Núñez et al. (2018)
7	Mexico	Oaxaca, San Juan La Lana, surroundings of Santiago Jalahui	17.4528	-95.769	1410	7	0.18015	Rendón et al. (1998)
8	Mexico	Chiapas, microcuenca Bombana	16.9481	-93.02	777.5	7	0.12682	Percino-Daniel et al. (2013)

9	Guatemala	El Peten, Chinajá (town)	16.033	-90.217	1100	16	0.17534	Duellman (1963)
10	Mexico	Oaxaca, sierra de Cuatro Venados	15.9352	-97.133	500	8	0.15678	Martín-Regalado et al. (2016)
11	Guatemala	Alta Verapaz, Finca Chichen	15.594	-90.149	443	8	0.16396	Stuart (1948)
12	Honduras	Atlantida and Yoro, Cordillera Nombre de Dios, Refugio de Vida Silvestre Texíguat and Parque Nacional Pico Bonito	15.525	-87.37	60	10	0.16915	Townsend et al. (2012)
13	Honduras	Copán, subcuenca del río Copán	14.8417	-89.058	1000	23	0.16436	Alemán et al. (2013)
14	Guatemala	Suchitepéquez, Patulul, Finca Universitaria San Julián	14.4625	-91.133	1725	12	0.14249	Guerra-Centeno et al. (2015)
15	Nicaragua	Región Autónoma del Atlántico Norte, Waslala	13.3333	-85.367	333	5	0.13642	Orozco et al. (2015)
16	Nicaragua	Río San Juan, El Quebracho Private Wildlife Reserve	11.1661	-84.406	82	2	0.15	Barquero et al. (2010)
17	Colombia	Guajira, arroyo Palomino	10.9703	-72.758	1200	3	0.09234	Blanco-Torres et al. (2013)
18	Venezuela	Zulia, approximately 20 Km S of Maracaibo	10.5882	-71.71	1550	10	0.18681	Larreal et al. (2012)
19	Costa Rica	Alajuela, San Ramón, Reserva Biológica Alberto Brenes	10.219	-84.597	1300	9	0.16753	Bolaños & Ehmccke (1996)
20	Colombia	Magdalena, Sabanas de San Ángel, “La Gloria project”	10.1748	-74.327	175	10	0.18771	Angarita-M. et al. (2015)

21	Costa Rica	Cartago, Rio Macho Biological Station	9.76667	-83.85	410	4	0.14429	Acosta-Chaves et al. (2015)
22	Costa Rica	San José & Puntarenas, Parque Nacional Carara	9.75611	-84.608	525	16	0.16103	Laurencio & Malone (2009)
23	Colombia	Cesar, ciénaga de Zapatosa	9.29167	-73.792	340	11	0.16368	Medina-Rangel et al. (2015)
24	Panama	Bocas del Toro, Bosque Protector Palo Seco	9.15682	-82.518	200	7	0.1625	Hofer & Bersier (2001)
25	Panama	Panama, Barro Colorado Island	9.15	-79.85	27	21	0.17519	Rand & Myers (1990)
26	Colombia	Magdalena, El Banco, corregimiento Botillero	9.00472	-73.971	88	7	0.15301	Dueñez-Gómez et al. (2004)
27	Venezuela	Barinas, Ramal de Calderas, Sector San Ramón	8.86772	-70.493	189	5	0.14401	Barrio-Amorós & Molina (2011)
28	Colombia	Cesar, Tamalemeque	8.85867	-73.808	344	7	0.17777	Ruiz (2011)
29	Colombia	Córdoba, Valencia, corregimiento de San Rafael del Pirú	8.20277	-76.178	325	12	0.18501	Ruiz (2014)
30	Colombia	Valle del Cauca, Reserva de Yotoco	3.833	-76.333	200	11	0.14262	Castro-Herrera et al. (2007)
31	Colombia	Valle del Cauca, Buenaventura, Central Hidroeléctrica de bajo Anchicayá	3.73167	-76.958	116	10	0.1739	Vargas & Bolaños (1999)
32	Colombia	Huila, Magdalena River dry valley, Gigante, Garzón, El Agrado and Paicol municipalities	2.45467	-75.559	217	13	0.15106	Moreno-Arias & Quintero-Corzo (2015)
33	Colombia	Caquetá, La Macarena, Caño Cristales, sendero Laguna El Turista	2.249	-73.758	300	6	0.1434	Barrientos et al. (2017)

34	Colombia	Caquetá, San Vicente del Caguán	2.01667	-74.642	68	10	0.15478	Cortes-Ávila & Toledo (2013)
35	Colombia	Guaviare, Calamar, cuenca rio Itilla	1.95547	-72.863	168	3	0.17735	Osorno-Muñoz et al. (2019)
36	Colombia	Nariño, Tumaco, Dirección General Marítima station (DIMAR)	1.82384	-78.727	375	13	0.20534	Pinto-Erazo et al. (2019)
37	Ecuador	Esmeraldas, Bilsa Biological Station	0.35917	-79.701	16	23	0.16066	Ortega-Andrade et al. (2010)
38	Ecuador	Napo, Santa Cecilia region	0.05	-76.983	139	25	0.18817	Duellman (1978)
39	Ecuador	Orellana, Estación de Biodiversidad Tiputini	-0.6181	-76.172	214	27	0.19076	Cisneros-Heredia (2006)
40	Brazil	Pará, Floresta Estadual do Trombetas	-0.9628	-55.522	230	8	0.16609	Ávila-Pires et al. (2010)
41	Brazil	Amazonas, Japurá river	-1.7825	-69.079	280	7	0.18257	Frazão et al. (2020)
42	Peru	Loreto, Teniente Lopez (town)	-2.5946	-76.116	34	4	0.15171	Duellman & Mendelson (1995)
43	Brazil	Amazonas, Manaus, Experimental Farm area of the Federal University of Amazonas	-2.6415	-60.095	2200	13	0.18964	Masseli et al. (2019)
44	Brazil	Amazonas, Manaus, Reserva Ducke	-2.9667	-59.933	48	22	0.18721	de Fraga et al. (2011)
45	Brazil	Amazonas, Manaus, INPA-WWF reserves	-3.217	-60.033	-1	17	0.18546	Zimmerman & Rodrigues (1990)
46	Brazil	Pará, Brasil Novo, Caverna Planaltina	-3.3762	-52.576	73	6	0.18799	Oliveira et al. (2012)

47	Brazil	Amazonas, upper Rio Urucu (3 Km south of headwaters Rio Urucu)	-5	-65.5	105	4	0.0903	Gascon & Pereira (1993)
48	Peru	Cajamarca, Santa Rosa de la Yunga	-5.4345	-78.572	1850	6	0.09937	Koch et al. (2018)
49	Brazil	Paraíba, Rio Tinto and Mamanguape, Guaribas Biological Reserve	-6.7405	-35.142	125	14	0.17857	Barbosa et al. (2015)
50	Brazil	Paraíba, João Pessoa municipality, Barra de Gramame	-7.2335	-34.806	625	7	0.15071	Sampaio et al. (2018)
51	Brazil	Bahía, São Francisco river Quarternary dunes	-10.768	-42.756	700.5	10	0.17224	Rodrigues (1996)
52	Peru	Madre de Dios, Manu National Park, Cocha Cashu	-11.9	-71.367	202	9	0.19446	Rodríguez et al. (1990)
53	Peru	Madre de Dios, Manu National Park, Pakitza	-11.944	-71.283	250	16	0.19221	Morales & McDiarmid (1996)
54	Peru	Madre de Dios, Reserva Cusco Amazonico	-12.556	-69.072	805	24	0.18454	Duellman (2005)
55	Peru	Manu, Los Amigos Biological Station	-12.569	-70.099	82	30	0.19095	Rabosky et al. (2009)
56	Peru	Parque Nacional y Área Natural de Manejo Integrado Madidi, Serranía Tequeje	-13.717	-68.217	316	7	0.19442	Cortez-Fernandez (2005)
57	Brazil	Mato Grosso, near Poconé	-16.5	-56.75	250	12	0.17811	Strussmann & Sazima (1993)
58	Bolivia	Santa Cruz, Santa Rosa del Sara, propiedad Juan Deriba	-16.867	-63.75	425	4	0.15249	Pinto-Viveros et al. (2015)

59	Brazil	Rio do Janeiro, Serra do Mendanha	-22.825	-43.492	125	12	0.13623	Pontes et al. (2009)
60	Paraguay	Canindeyú, Reserva Natural del Bosque Mbaracayú, Aguará Ñu.	-24.183	-55.267	61	7	0.18697	Cacciali et al. (2015)
61	Argentina	Formosa, Municipalidad de Colonia Villafañe, La Picadita, “Estancia La Ascensión”	-26.038	-59.157	37	4	0.1572	Pesci et al. (2018)
62	Paraguay	Itapua, Parque Nacional San Rafael, Kangüery	-26.512	-55.789	685	6	0.15647	Cacciali (2013)
63	Brazil	Rio Grande do Sur, Fazenda da Brigada Militar	-28.244	-52.328	6	4	0.15427	Zanella & Cechin (2006)
64	Argentina	Corrientes, Reserva Natural del Iberá, Departamento San Martin, colonia Pellegrini	-28.55	-57.167	80	5	0.19235	Etchepare et al. (2013)
65	Argentina	Corrientes, Bella Vista, Reserva Paleontológica del Arroyo Toropí	-28.592	-59.059	121	1	0	Ingaramo et al. (2015)
66	Brazil	Rio Grande do Sur, Rio Grande, Balneário Cassino	-32.132	-52.348	188	1	0	Santos et al. (2012)
67	Argentina	Coastal dunes Buenos Aires province	-37.617	-57.333	1184	3	0.18044	Kacoliris et al. (2006)

Appendix Table A3.5. Comparison among biogeographic models, as evaluated with ‘BioGeoBEARS’ for the analysis performed as defined in the text. For each model, the following values are reported: log-likelihood (LnL), numbers of parameters, and their corresponding value of the corrected Akaike information criterion (AIC weights and AICc).

Biogeographical model	log-likelihood (LnL)	Number of parameters	d	e	j	AICc	AICc weights
DEC	-1013	2	0.012	0.0014	0	2031	5.40E-63
DEC+J	-1009	3	0.011	4.30E-09	0.0057	2024	1.80E-61
DIVALIKE	-1058	2	0.013	1.30E-03	0	2120	2.20E-82
DIVALIKE+J	-1052	3	0.013	1.00E-12	0.0047	2111	2.30E-80
BAYAREALIKE	-883	2	0.0032	0.049	0	1770	2.10E-06
BAYAREALIKE+J	-868.9	3	0.0038	0.032	0.0052	1744	1

Appendix Table A3.6. List of the dispersal events of Dipsadidae snakes across the Panamanian Isthmus, indicating the direction, and estimated age in which they occurred, according to the results of the BioGeoBEARS analyses.

Taxa dispersing across the Isthmus	Direction of dispersal event	Date of dispersal event	95% Confidence Interval (in millions of years)	Ancestral area estimation probabilities
MRCA of Carphophiinae	from South America to Central America	28.36937	23.29-34.12	EFG (61.46%); DEFG (5.98%); EF (4.00%)
MRCA of the Dipsadinae with the exclusion of the Diaphorolepini	from South America to Central America	24.49182	19.88-29.8	EFG (52.25%); EF (21.20%); BDEF (2.60%)
MRCA of <i>Coniophanes fissidens</i> and <i>Pliocercus euryzonus</i>	from Central America to both Central and South America	14.55257	9.99-18.77	BD (74.05%); BD (20.77%); B(0.98%)
MRCA of <i>Imantodes gemmistratus</i> , <i>Imantodes cenchoa</i> , <i>Imantodes lentiferus</i> , <i>Imantodes chocoensis</i> , and <i>Imantodes inornatus</i>	from Central America to both Central and South America	15.75877	11.44-19.11	BD (57.66%); BDE (28.14%); BDEF (3.34%)
MRCA of <i>Leptodeira septentrionalis</i> , <i>Leptodeira annulata</i> and <i>Leptodeira bakeri</i>	from Central America to both Central and South America	5.467573	3.68-7.14	AB (29.19%); BD (10.53%); B (2.53%)
MRCA of all <i>Atractus</i> (plus <i>Geophis godmani</i>)	from Central America to South America	15.03595	11.5-18.21	BD (65.13%); B (28.37%); E (3.10%)
MRCA of <i>Dipsas</i>	from Central America to South America	13.72452	10.6-16.37	BD (97.77%); E (0.76%); BDE (0.42%)
MRCA of <i>Dipsas neivai</i> , <i>Dipsas variegata</i>	from South America to both Central and South America	6.756322	4.44-9.34	EFH (83.50%); EF (8.24%); DEFH (5.43%)

MRCA of <i>Siphlophis_cervinus</i> and <i>Siphlophis_compressus</i>	from South America to both Central and South America	9.779022	6.65-13.48	EFG (51.88%); DEFG (23.52%); EFGH (14.34%)
<i>Diaphorolepis_wagneri</i>	from South America to both Central and South America	at some point after 12.6819	9.04-16.62	EF (82.35%); E (9.56%); DEF (4.56%)
<i>Nothopsis_rugosus</i>	from Central America to both Central and South America	at some point after 19.36305	15.31-23.85	BD (73.88%); BDE (22.37%); BDEF (0.59%)
<i>Ninia_atrata</i>	from Central America to both Central and South America	at some point after 14.13315	10.91-17.53	BD (97.07%); B (0.77%); BDE (0.77%)
<i>Sibon_bevridgelyi</i>	from Central America to South America	at some point after 1.699204	0.61-3.05	D (80.20%); BD (14.49%); E (2.69%)
<i>Sibon_nebulatus</i>	from Central America to both Central and South America	at some point after 8.162918	6.07-10.54	BD (96.94%); BDE (0.82%); B (0.01%)
<i>Dipsas_articulata</i>	from South America to Central America	at some point after 4.898102	2.49-7.6	E (77.51%); D (14.26%); BD (5.41%)
<i>Manolepis_putmani</i>	from South America to Central America	at some point after 22.08978	16.25-24.58	EFG (90.58%); EFGH (4.65%); FG (1.34%)
<i>Oxyrhopus_petolaris</i>	from South America to both Central and South America	at some point after 1.26657	0.35-2.42	EFGH (70.92%); DEFGH (13.16%); BDEFGH (2.49%)
<i>Clelia_clelia</i>	from South America to both Central and South America	at some point after 4.524012	2.29-6.62	EFGH (76.48%); DEFGH (5.66%); FGH (3.96%)

<i>Pseudoboa_neuwiedii</i>	from South America to both South and Central America	at some point after 3.566634	2.24-5.13	EFGH (47.15%); EFG (45.22%); DEFGH (1.68%)
<i>Crisantophis_nevermanni</i>	from South America to Central America	at some point after 16.95365	9.33-22.92	EFG (22.22%); EG (16.61%); EF (16.61%)
<i>Lygophis_lineatus</i>	from South America to both Central and South America	at some point after 14.66508	10.97-18.56	EFG (71.39%); FG (20.53%); EFGH (3.16%)
<i>Erythrolamprus_epinephelus</i>	from South America to both Central and South America	at some point after 5.492384	3.54-7.49	EF (78.15%); EFH (11.49%); EFG (6.94%)
<i>Erythrolamprus_aesculapii</i>	from South America to both Central and South America	at some point after 0.89041	0.14-1.8	EF (30.40%); DEF (22.46%); EFG (18.50%)
<i>Erythrolamprus_mimus</i>	from South America to both Central and South America	at some point after 1.494761	0.14-2.48	EF (50.98%); DEF (41.41%); EFG (1.81%)

Appendix Table A3.7. Best-fit partitioning scheme for the protein-coding genes, and models of nucleotide substitution for each partition, based on the corrected Akaike information criterion (AICc) as implemented in PartitionFinder 2 (Lanfear et al., 2016). The abbreviation ‘pos.’ refers to codon position.

Subset	Best Model	Subset Partitions	Subset Sites
1	GTR+I+G+X	12S	1-425
2	GTR+I+G+X	16S	426-980
3	TRN+I+G+X	bdnfcodon1	981-1653\3
4	GTR+I+X	bdnfcodon2, bdnfcodon3	983-1653\3, 982-1653\3
5	GTR+I+G+X	nd2codon1	1654-2694\3
6	GTR+I+G+X	nd2codon2	1655-2694\3
7	GTR+I+G+X	nd2codon3	1656-2694\3
8	GTR+I+G+X	nd4codon1, cytbcodon1	2695-3372\3, 4668-5438\3
9	GTR+I+G+X	nd4codon2, cytbcodon2	2696-3372\3, 4669-5438\3
10	GTR+G+X	nd4codon3	2697-3372\3
11	TRN+I+G+X	rag2codon1, cmoscodon2	3373-4086\3, 4088-4667\3
12	GTR+I+G+X	rag2codon2, cmoscodon3	3374-4086\3, 4089-4667\3
13	GTR+G+X	rag2codon3, cmoscodon1	3375-4086\3, 4087-4667\3
14	GTR+I+G+X	cytbcodon3	4670-5438\3
15	GTR+G+X	nt3codon1	5439-5978\3
16	SYM+I+G	nt3codon2	5440-5978\3
17	GTR+G+X	nt3codon3	5441-5978\3

Appendix Table A3.8. Calibration points used for obtaining the timetree used for this work and reported as Appendix Figure 3.2.

	Calibration point	Offset (Ma), Minimum age	95% quantile (Ma)	Fossil	References
1	Boidae:Boinae+Erycinae	58	67.5	<i>Titanoboa cerrejonensis</i>	Head et al. (2009); Head (2015)
2	stem Caenophidia (= crown Pan-Caenophidia)	66	88.5	<i>Krebsophis thobanus</i>	Rage & Werner (1999); Head et al. (2016)
3	stem Colubroidea (= stem Colubriiformes)	50.5	85	<i>Procerophis sahnii</i>	Rage et al. (2008) & Head et al. (2016)
4	stem Elapidae	24.95	47.4	Elapidae indeterminate	McCartney et al. (2014); Zaher et al. (2019)
5	crown Viperidae	22.1	47	<i>Vipera cf. V. antiqua</i>	Zaher et al. (2019)*

*According to Head et al. (2017), this fossil cannot be used to calibrate the minimum age of the divergence between Viperidae and its sister group (i.e., Homalopsidae+Colubroidea). Such is the case, as this fossil is of a more recent date than the time of the first appearance of Homalopsidae or Colubroidea in the fossil record. However, the *Vipera cf. antiqua* fossil can still be used to calibrate the minimum age of the first divergence within Viperidae (=Azemiopinae vs. Viperinae+Crotalinae), as was done in Zaher et al. (2019) and followed herein.

Appendix Table A3.9. Values of per-species diversification, speciation, and extinction rates averaged across all taxa inhabiting each of the communities of the Dipsadidae considered for this study.

Country	Locality	Average of per-species net diversification rates	Average of per-species speciation rates	Average of per-species extinction rates
Mexico	Hidalgo, Ixmiquilpan, valle del Mezquital	0.100	0.133	0.033
Mexico	Hidalgo, Reserva de la Biosfera Barranca de Metztitlán	0.114	0.135	0.021
Mexico	Michoacán, Tacámbaro	0.108	0.113	0.005
Mexico	Veracruz, Sierra de Los Tuxtlas	0.108	0.114	0.005
Mexico	Veracruz, Los Tuxtlas, Parque de Flora y Fauna Silvestre Tropical	0.108	0.114	0.006
Mexico	Guerrero, West Sierra Madre del Sur	NA	NA	NA
Mexico	Oaxaca, San Juan La Lana, surroundings of Santiago Jalahui	0.114	0.136	0.022
Mexico	Chiapas, microcuenca Bombana	0.112	0.136	0.024
Guatemala	El Peten, Chinajá (town)	0.108	0.113	0.005
Mexico	Oaxaca, sierra de Cuatro Venados	0.115	0.138	0.024
Guatemala	Alta Verapaz, Finca Chichen	NA	NA	NA
Honduras	Atlantida and Yoro, Cordillera Nombre de Dios, Refugio de Vida Silvestre Texíguat and Parque Nacional Pico Bonito	0.114	0.135	0.022
Honduras	Copán, subcuenca del río Copán	0.113	0.131	0.019
Guatemala	Suchitepéquez, Patulul, Finca Universitaria San Julián	0.111	0.125	0.014
Nicaragua	Región Autónoma del Atlántico Norte, Waslala	0.108	0.114	0.006
Nicaragua	Río San Juan, El Quebracho Private Wildlife Reserve	0.114	0.135	0.022
Colombia	Guajira, arroyo Palomino	0.108	0.113	0.005
Venezuela	Zulia, approximately 20 Km S of Maracaibo	0.107	0.114	0.008
Costa Rica	Alajuela, San Ramón, Reserva Biológica Alberto Brenes	0.112	0.128	0.017

Colombia	Magdalena, Sabanas de San Ángel, “La Gloria project”	0.108	0.114	0.006
Costa Rica	Cartago, Rio Macho Biological Station	0.109	0.117	0.008
Costa Rica	San José & Puntarenas, Parque Nacional Carara	0.111	0.125	0.014
Colombia	Cesar, ciénaga de Zapatosa	0.108	0.114	0.006
Panama	Bocas del Toro, Bosque Protector Palo Seco	0.111	0.127	0.016
Panama	Panama, Barro Colorado Island	0.109	0.115	0.007
Colombia	Magdalena, El Banco, corregimiento Botillero	0.108	0.114	0.005
Venezuela	Barinas, Ramal de Calderas, Sector San Ramón	0.108	0.114	0.005
Colombia	Cesar, Tamalemeque	0.108	0.115	0.007
Colombia	Córdoba, Valencia, corregimiento de San Rafael del Pirú	0.109	0.115	0.006
Colombia	Valle del Cauca, Reserva de Yotoco	0.111	0.123	0.012
Colombia	Valle del Cauca, Buenaventura, Central Hidroeléctrica de bajo Anchicayá	0.112	0.128	0.016
Colombia	Huila, Magdalena River dry valley, Gigante, Garzón, El Agrado and Paicol municipalities	0.112	0.128	0.016
Colombia	Caquetá, La Macarena, Caño Cristales, sendero Laguna El Turista	0.109	0.113	0.005
Colombia	Caquetá, San Vicente del Caguán	0.110	0.121	0.010
Colombia	Guaviare, Calamar, cuenca rio Itilla	0.109	0.116	0.007
Colombia	Nariño, Tumaco, Dirección General Marítima station (DIMAR)	0.127	0.164	0.037
Ecuador	Esmeraldas, Bilsa Biological Station	0.112	0.129	0.017
Ecuador	Napo, Santa Cecilia region	0.111	0.125	0.014
Ecuador	Orellana, Estación de Biodiversidad Tiputini	0.112	0.126	0.015
Brazil	Pará, Floresta Estadual do Trombetas	0.111	0.124	0.013
Brazil	Amazonas, Japurá river	0.110	0.121	0.011
Peru	Loreto, Teniente Lopez (town)	0.108	0.114	0.005
Brazil	Amazonas, Manaus, Experimental Farm area of the Federal University of Amazonas	0.109	0.115	0.007

Brazil	Amazonas, Manaus, Reserva Ducke	0.110	0.120	0.010
Brazil	Amazonas, Manaus, INPA-WWF reserves	0.111	0.124	0.013
Brazil	Pará, Brasil Novo, Caverna Planaltina	0.111	0.125	0.014
Brazil	Amazonas, upper Rio Urucu (3 Km south of headwaters Rio Urucu)	0.108	0.114	0.005
Peru	Cajamarca, Santa Rosa de la Yunga	0.114	0.135	0.021
Brazil	Paraíba, Rio Tinto and Mamanguape, Guaribas Biological Reserve	0.110	0.121	0.011
Brazil	Paraíba, João Pessoa municipality, Barra de Gramame	0.110	0.118	0.009
Brazil	Bahía, São Francisco river Quarternary dunes	0.109	0.115	0.006
Peru	Madre de Dios, Manu National Park, Cocha Cashu	0.109	0.117	0.008
Peru	Madre de Dios, Manu National Park, Pakitza	0.110	0.118	0.008
Peru	Madre de Dios, Reserva Cusco Amazonico	0.111	0.125	0.014
Peru	Manu, Los Amigos Biological Station	0.110	0.121	0.011
Peru	Parque Nacional y Área Natural de Manejo Integrado Madidi, Serranía Tequeje	0.108	0.113	0.005
Brazil	Mato Grosso, near Poconé	0.109	0.116	0.007
Bolivia	Santa Cruz, Santa Rosa del Sara, propiedad Juan Deriba	0.109	0.115	0.006
Brazil	Rio do Janeiro, Serra do Mendanha	0.109	0.115	0.007
Paraguay	Canindeyú, Reserva Natural del Bosque Mbaracayú, Aguará Ñu.	0.111	0.123	0.012
Argentina	Formosa, Municipalidad de Colonia Villafañe, La Picadita, “Estancia La Ascensión”	0.113	0.130	0.018
Paraguay	Itapua, Parque Nacional San Rafael, Kangüery	0.110	0.122	0.011
Brazil	Rio Grande do Sur, Fazenda da Brigada Militar	0.109	0.116	0.007
Argentina	Corrientes, Reserva Natural del Iberá, Departamento San Martin, colonia Pellegrini	0.111	0.123	0.012
Argentina	Corrientes, Bella Vista, Reserva Paleontológica del Arroyo Toropí	0.110	0.118	0.008
Brazil	Rio Grande do Sur, Rio Grande, Balneário Cassino	0.109	0.116	0.007
Argentina	Coastal dunes Buenos Aires province	0.109	0.116	0.007

Appendix Table A3.10. Values of functional dispersion of the 67 communities of dipsadid snakes included in this study calculated after removing the species colonizing them from across the Panamanian Isthmus.

Country	Locality	Functional dispersion
Mexico	Hidalgo, Ixmiquilpan, valle del Mezquital	0.1749
Mexico	Hidalgo, Reserva de la Biosfera Barranca de Metztitlán	0.0890
Mexico	Michoacán, Tacámbaro	0.1762
Mexico	Veracruz, Sierra de Los Tuxtlas	0.1469
Mexico	Veracruz, Los Tuxtlas, Parque de Flora y Fauna Silvestre Tropical	0.1697
Mexico	Guerrero, West Sierra Madre del Sur	0.0851
Mexico	Oaxaca, San Juan La Lana, surroundings of Santiago Jalahui	0.1471
Mexico	Chiapas, microcuenca Bombana	0.1293
Guatemala	El Peten, Chinajá (town)	0.1545
Mexico	Oaxaca, sierra de Cuatro Venados	0.0021
Guatemala	Alta Verapaz, Finca Chichen	0.0628
Honduras	Atlantida and Yoro, Cordillera Nombre de Dios, Refugio de Vida Silvestre Texíguat and Parque Nacional Pico Bonito	0.1413
Honduras	Copán, subcuenca del río Copán	0.1724
Guatemala	Suchitepéquez, Patulul, Finca Universitaria San Julián	0.2047
Nicaragua	Región Autónoma del Atlántico Norte, Waslala	0.1762
Nicaragua	Río San Juan, El Quebracho Private Wildlife Reserve	0.1617
Colombia	Guajira, arroyo Palomino	0.1637
Venezuela	Zulia, approximately 20 Km S of Maracaibo	0.1852
Costa Rica	Alajuela, San Ramón, Reserva Biológica Alberto Brenes	0.1742
Colombia	Magdalena, Sabanas de San Ángel, “La Gloria project”	0.1927
Costa Rica	Cartago, Rio Macho Biological Station	0.1672
Costa Rica	San José & Puntarenas, Parque Nacional Carara	0.1771
Colombia	Cesar, ciénaga de Zapatosa	0.1888
Panama	Bocas del Toro, Bosque Protector Palo Seco	0.1766
Panama	Panama, Barro Colorado Island	0.1686
Colombia	Magdalena, El Banco, corregimiento Botillero	0.2342
Venezuela	Barinas, Ramal de Calderas, Sector San Ramón	0.1245
Colombia	Cesar, Tamalemeque	0.2202
Colombia	Córdoba, Valencia, corregimiento de San Rafael del Pirú	0.2246
Colombia	Valle del Cauca, Reserva de Yotoco	0.1739

Colombia	Valle del Cauca, Buenaventura, Central Hidroeléctrica de bajo Anchicayá	0.0357
Colombia	Huila, Magdalena River dry valley, Gigante, Garzón, El Agrado and Paicol municipalities	0.1124
Colombia	Caquetá, La Macarena, Caño Cristales, sendero Laguna El Turista	0.1553
Colombia	Caquetá, San Vicente del Caguán	0.1908
Colombia	Guaviare, Calamar, cuenca rio Itilla	0.1290
Colombia	Nariño, Tumaco, Dirección General Marítima station (DIMAR)	0.1534
Ecuador	Esmeraldas, Bilsa Biological Station	0.1778
Ecuador	Napo, Santa Cecilia region	0.1821
Ecuador	Orellana, Estación de Biodiversidad Tiputini	0.1971
Brazil	Pará, Floresta Estadual do Trombetas	0.1712
Brazil	Amazonas, Japurá river	0.1848
Peru	Loreto, Teniente Lopez (town)	0.1208
Brazil	Amazonas, Manaus, Experimental Farm area of the Federal University of Amazonas	0.1714
Brazil	Amazonas, Manaus, Reserva Ducke	0.1880
Brazil	Amazonas, Manaus, INPA-WWF reserves	0.1867
Brazil	Pará, Brasil Novo, Caverna Planaltina	0.1648
Brazil	Amazonas, upper Rio Urucu (3 Km south of headwaters Rio Urucu)	0.1783
Peru	Cajamarca, Santa Rosa de la Yunga	0.1544
Brazil	Paraíba, Rio Tinto and Mamanguape, Guaribas Biological Reserve	0.1929
Brazil	Paraíba, João Pessoa municipality, Barra de Gramame	0.1934
Brazil	Bahía, São Francisco river Quarternary dunes	0.1767
Peru	Madre de Dios, Manu National Park, Cocha Cashu	0.1717
Peru	Madre de Dios, Manu National Park, Pakitza	0.1733
Peru	Madre de Dios, Reserva Cusco Amazonico	0.1938
Peru	Manu, Los Amigos Biological Station	0.1834
Peru	Parque Nacional y Área Natural de Manejo Integrado Madidi, Serranía Tequeje	0.1607
Brazil	Mato Grosso, near Poconé	0.2397
Bolivia	Santa Cruz, Santa Rosa del Sara, propiedad Juan Deriba	0.1731
Brazil	Rio do Janeiro, Serra do Mendanha	0.1573
Paraguay	Canindeyú, Reserva Natural del Bosque Mbaracayú, Aguará Ñu.	0.1954
Argentina	Formosa, Municipalidad de Colonia Villafañe, La Picadita, "Estancia La Ascensión"	0.1746
Paraguay	Itapua, Parque Nacional San Rafael, Kangüery	0.2103
Brazil	Rio Grande do Sur, Fazenda da Brigada Militar	0.1967
Argentina	Corrientes, Reserva Natural del Iberá, Departamento San Martin,	0.1766

colonia Pellegrini

Argentina	Corrientes, Bella Vista, Reserva Paleontológica del Arroyo Toropí	0.1838
Brazil	Rio Grande do Sur, Rio Grande, Balneário Cassino	0.1622
Argentina	Coastal dunes Buenos Aires province	0.1596

Appendix Table A3.11. Results and statistics of the t-tests performed to test for possible differences in several variables between communities of dipsadid snakes involved or not in the Great American Biotic Interchange

Variable	t-value	Degrees of freedom	p-value
Functional dispersion of non-dipsadid snakes	1.697	35.45	0.098
Environmental PC2	-0.546	51.13	0.587
Environmental PC3	0.502	35.74	0.619
Environmental PC4	1.892	60.47	0.063
Total sampling effort	0.859	35.33	0.396
Species richness of non-dipsadid snakes	1.485	58.52	0.143
Trap use	1.642	51.00	0.107
Time of first colonization	2.916	44.33	0.006
Elevation	-2.074	37.24	0.045
Environmental PC1	5.650	36.64	1.93E-06
Site Area	-0.756	24.81	0.457
Distance to the Isthmus	-3.863	47.09	3.40E-04
Latitude	-6.439	43.86	7.72E-08
Community-wide diversification rates	0.701	48.15	0.487

References used for constructing Appendix Tables A3.1 and A3.4 (and Supplemental Tables S3.1 and S3.3).

Acosta-Chaves, V. J., G. Chaves, J. G. Abarca, A. García Rodríguez, & F. Bolaños (2015). A checklist of the amphibians and reptiles of Río Macho Biological Station, Provincia de Cartago, Costa Rica. *Check List*, 11: 1,784.

Alemán-Mejía, J. B. (2011). Caracterización de reptiles y percepción local hacia las serpientes en fincas ganaderas de la subcuenca del Río Copán, Honduras. *Agroforestería en las Américas*, 48, 103-117.

Almazán-Núñez, R. C., Alvarez-Alvarez, E. A., Ruiz-Gutiérrez, F., Almazán-Juárez, Á., Sierra-Morales, P., & Toribio-Jiménez, S. (2018). Biological survey of a cloud forest in southwestern Mexico: plants, amphibians, reptiles, birds, and mammals. *Biota Neotropica*, 18(2). E20170444

Angarita-M, O., Montes-Correa, A. C., & Renjifo, J. M. (2015). Amphibians and reptiles of an agroforestry system in the Colombian Caribbean. *Amphibian & Reptile Conservation*, 8(1), 33-52.

Ávila-Pires TCS, Hoogmoed MS, da Rocha WA. (2010). Notes on the Vertebrates of northern Pará, Brazil: a forgotten part of the Guianan Region, I. Herpetofauna, *Boletim do Museu Paraense Emílio Goeldi*, 5(1): 13-112.

Barquero, M. D., Salazar-Saavedra, M., Sandoval, L., Brenes, D., Martínez, F., & Figueroa, A. (2010). Composition and species richness of herpetofauna in two isolated regions of southern Nicaragua. *Herpetology Notes*, 3, 341-352.

Barrientos, L. S., M. A. Morales-Betancourt, C. A. Lasso, G. Torres y P. Herrera. (2017). Anfíbios y reptiles. Pp. 97-123 in: Lasso, C. A. & M. A. Morales-Betancourt (Eds.). III. *Fauna de Caño Cristales, Sierra La Macarena, Meta, Colombia*. Serie Editorial Fauna Silvestre Neotropical. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt (IAvH). Bogotá, D. C., Colombia.

Barrio-Amorós, C.L. & Molina, C. (2010) Herpetofauna del Ramal de Calderas, Andes de Venezuela. In: Rial, B.A., Señaris, J.C., Lasso, C.A., & Flores, A. (Eds), *Evaluación Rápida de la Biodiversidad y Aspectos Socioecosistémicos del Ramal de Calderas. Andes de Venezuela*. RAP Bulletin of Biological Assessment 56. Conservation International, Arlington, VA. USA, pp. 74–80.

Blanco-Torres, A. B., Báez, L., Patiño-Flores, E., & Renjifo, J. M. (2013). Herpetofauna del valle medio del río Ranchería, La Guajira, Colombia. *Revista Biodiversidad Neotropical*, 3(2), 113-122.

Bolaños, F., & Ehmcke, J. (1996). Anfibios y reptiles de la reserva biológica San Ramón, Costa Rica. *Pensamiento Actual*, 2(2): 107-112

Cacciali, P. I. E. R. (2013). Diversidad y selección de hábitat de la fauna de serpientes en Kangüery (Área para Parque San Rafael). *Boletín del Museo Nacional de Historia Natural del Paraguay*, 17, 29-39.

Cacciali, P., Bauer, F., & Martínez, N. (2015). Herpetofauna de la Reserva Natural del Bosque Mbaracayú, Paraguay. *Kempffiana*, 11(1), 29-47.

Castro-Herrera F, Bolívar-García W, Herrera-Montes MI (2007) *Guía de anfibios y reptiles del bosque de Yotoco, Valle del Cauca, Colombia*. Grupo de Investigación Laboratorio de Herpetología, Universidad del Valle, Cali, Colombia, 70 pp.

Cisneros-Heredia, D.F. (2006) *La herpetofauna de la Estación de Biodiversidad Tiputini, Ecuador*. B.Sc. Thesis. Universidad San Francisco de Quito, Quito, 129 pp.

Cortes-Ávila, L., & Toledo, J. J. (2013). Estudio de la diversidad de serpientes en áreas de bosque perturbado y pastizal en San Vicente del Caguán (Caquetá), Colombia. *Actualidades Biológicas*, 35(99), 185-197.

Cortez-Fernandez, C. (2005). Herpetofauna de la zona norte del Parque Nacional y Area Natural de Manejo Integrado Madidi (PNANMI-Madidi). *Ecología en Bolivia*, 40(2), 10-26.

de Fraga, R., Lima, A. P., & Magnusson, W. E. (2011). Mesoscale spatial ecology of a tropical snake assemblage: the width of riparian corridors in central Amazonia. *The Herpetological Journal*, 21(1), 51-57.

Dueñez-Gómez, F., Muñoz-Guerrero, J., & Ramírez-Pinilla, M. P. (2004). Herpetofauna del corregimiento Botillero (El Banco, Magdalena) en la depresión Momposina de la región Caribe colombiana. *Actualidades Biológicas*, 26(81), 161-17.

Duellman, W. E. (1963). Amphibians and reptiles of the rainforests of southern El Petén, Guatemala. *University of Kansas Publications, Museum of Natural History* 15(5):205-249

Duellman, W.E. (1978). The biology of an equatorial herpetofauna in Amazonian Ecuador. *Miscellaneous Publications of the Museum of Natural History of the University of Kansas* 65:1-352

Duellman W.E., J. R. Mendelson III (1995) Amphibians and reptiles from the northern Departamento Loreto, Peru: taxonomy and biogeography. *University of Kansas Science Bulletin* 55:329-379

Duellman, W.E. (2005) *Cusco Amazónico: the lives of amphibians and reptiles in an Amazonian rainforest*. Cornell University Press, Ithaca and London, 472 p.

Etchepare, E. G., del Rosario Ingaramo, M., Porcel, E., & Álvarez, B. B. (2013). Diversidad de las comunidades de escamados en la Reserva Natural del Iberá, Corrientes, Argentina. *Revista Mexicana de Biodiversidad*, 84(4), 1273-1283.

Fernández-Badillo, L., & Goyenechea-Mayer Goyenechea, I. (2010). Anfibios y reptiles del valle del Mezquital, Hidalgo, México. *Revista Mexicana de Biodiversidad*, 81(3), 705-712.

Frazão, L., Oliveira, M. E., Menin, M., Campos, J., Almeida, A., Kaefer, I. L., & Hrbek, T. (2020). Species richness and composition of snake assemblages in poorly accessible areas in the Brazilian Amazonia. *Biota Neotropica*, 20(1): e20180661.

Gascon, C., & Pereira, O. D. S. (1993). Preliminary checklist of the herpetofauna of the upper Rio Urucu, Amazonas, Brazil. *Revista Brasileira de Zoologia*, 179-183.

Guerra-Centeno, D., Fuentes-Rousselin, H., Morán-Villatoro, D., & Valdéz-Sandoval, C. (2015). Riqueza de herpetofauna de la Finca Universitaria San Julián, Patulul, Suchitepéquez, Guatemala. *Ciencia, Tecnología y Salud*, 2(1), 13-24.

Hofer, U., & Bersier, L. F. (2001). Herpetofaunal diversity and abundance in tropical upland forests of Cameroon and Panama. *Biotropica*, 33(1), 142-152.

Ingaramo, M. D. R., Marangoni, F., & Cajade, R. (2015). Herpetofauna de la Reserva Paleontológica del Arroyo Toropí, Bella Vista, Corrientes, Argentina. *Cuadernos de Herpetología*, 29 (1): 69-75.

Kacoliris, F., Horlent, N., & Williams, J. (2006). Herpetofauna, Coastal Dunes, Buenos Aires Province, Argentina. *Check List*, 2(3): 15-21.

Koch, C., Venegas, P. J., Santa Cruz, R., & Boehme, W. (2018). Annotated checklist and key to the species of amphibians and reptiles inhabiting the northern Peruvian dry forest along the Andean valley of the Marañón River and its tributaries. *Zootaxa*, 4385(1), 1-101.

Larreal, J. T., Rivas, G. A., Portillo-Quintero, C., & Barros, T. R. (2012). Squamata reptiles of a fragment of tropical dry forest in northwestern Venezuela (Lake Maracaibo region). *Check List*, 8(6), 1220-1224.

Laurencio, D., & Malone, J. H. (2009). The amphibians and reptiles of Parque Nacional Carara, a transitional herpetofaunal assemblage in Costa Rica. *Herpetological Conservation and Biology*, 4(1), 120-131.

Martín-Regalado, N., Lavariega, M. C., Gómez-Ugalde, R. M., & Rodríguez-Pérez, C. (2016). Anfibios y reptiles de la sierra de Cuatro Venados, Oaxaca, México. *Arxius de Miscellània Zoològica*, 1, 217-232.

Masseli, G. S., Bruce, A. D., Santos, J. G., Vincent, T., & Kaefer, I. L. (2019). Composition and ecology of a snake assemblage in an upland forest from Central Amazonia. *Anais da Academia Brasileira de Ciências*, 91(Suppl. 1): e20190080

Medina-Aguilar, O., Alvarado-Díaz, J., & Suazo-Ortuño, I. (2011). Herpetofauna de Tacámbaro, Michoacán, México. *Revista Mexicana de Biodiversidad*, 82(4), 1194-1202.

Medina-Rangel, G. F. (2011). Diversidad alfa y beta de la comunidad de reptiles en el complejo cenagoso de Zapatos, Colombia. *Revista de Biología Tropical*, 59(2), 935-968.

Morales-Mávil, J. E., González, M. H., & Méndez, O. B. (2001). Biomasa de reptiles en un borde de selva de la región de Los Tuxtlas, Veracruz, México. *Foresta Veracruzana*, 3(2), 25-30.

Morales, V.R., y R.W. McDiarmid. 1996. Annotated checklist of the amphibians and reptiles of Pakitza, Manu National Park Reserve Zone, with comments on the herpetofauna of Madre de Dios, Peru. Pages 503- 522 In D.E. Wilson and A. Sandoval (editors) *Manu*. Smithsonian Institution Press, Washington, D.C.

Moreno-Arias, R. & Quintero-Corzo, S. (2015) Reptiles del valle seco del río Magdalena (Huila, Colombia). *Caldasia*, 37 (1): 183- 195. 21.

Oliveira, E. A., Ruz, E. H., & Barros, F. B. (2013). Herpetofauna de las proximidades de la Caverna Planaltina, Brasil Novo, Pará (Amazonia brasileña). *Herpetotropicos*, 9(1-2), 55-68.

Orozco, L., Molinares, B., Soto, G., Deheuvels, O., Rojas, J., López, A., & Vásquez, F. (2015). Diversidad de herpetofauna (anfibios y reptiles) en cacaotales y fragmentos de bosque en Waslala, Nicaragua. *Revista Agroforestería Neotropical*, 1(5):42-59.

Osorno-Muñoz, M., Caicedo-Portilla, R., & D. L. Gutierrez-Lamus (2019). Anfibios y reptiles de las cuencas altas de los ríos Itilla, Tunia y Yará. Pages 145-169 In: Rojas, A. et al. (eds). *Expediciones científicas en las nuevas áreas del PNN Serranía de Chiribiquete*. Fundación para la Conservación y el Desarrollo Sostenible. FCDS, Bogotá, D.C.

- Ortega-Andrade, H. M., Bermingham, J., Aulestia, C., & Paucar, C. (2010). Herpetofauna of the Bilsa Biological Station, province of Esmeraldas, Ecuador. *Check List*, 6(1), 119-154.
- Percino-Daniel, R., Cruz-Ocaña, E., Pozo-Ventura, W., & Velázquez-Velázquez, E. (2013). Diversidad de reptiles en dos microcuencas del río Grijalva, Chiapas, México. *Revista Mexicana de Biodiversidad*, 84(3), 938-948.
- Pesci, G. P., Sánchez, J. M., Muñoz Leao, S., & Pelegrin, N. 2018. Reptiles y anfibios de una localidad del Chaco Húmedo en Formosa, Argentina. *Cuadernos de Herpetología*, 32(1), 47–54.
- Pinto-Erazo, M. A., Espinosa, M. L. C., Rangel, G. F. M., & Galeano, M. Á. M. 2020. Herpetofauna de dos municipios del suroeste de Colombia. *Biota Colombiana*, 21(1), 41-57.
- Pinto Viveros, Marco & García-Durán, Dulce & Mano-Cuellar, Katherine & Pinto-Ledezma, Jesús. (2015). List of amphibians and reptiles of Juan Deriba property, Santa Cruz-Bolivia. *Kempffiana*. 11. 70-75.
- Pontes, J. A. L., Pontes, R. C., & Rocha, C. F. D. (2009). The snake community of Serra do Mendanha, in Rio de Janeiro State, southeastern Brazil: composition, abundance, richness and diversity in areas with different conservation degrees. *Brazilian Journal of Biology*, 69(3), 795-804.
- Rabosky, D. L., von May, R., Grundler, M. C., & Davis Rabosky, A. R. (2019). The western amazonian richness gradient for squamate reptiles: are there really fewer snakes and lizards in southwestern amazonian lowlands?. *Diversity*, 11(10), 199.
- Rand, A. S., and C. W. Myers. 1990. The herpetofauna of Barro Colorado Island, Panama: an ecological summary. Pp. 386– 409 in A. Gentry, ed. *Four neotropical rainforests*. Yale Univ. Press, New Haven , CT, USA.
- Rendón R., A., T. Álvarez, and O.A. Flores-Villela. (1998). Herpetofauna de Santiago Jalahui, Oaxaca, México. *Acta Zoologica Mexicana (n.s.)* 75: 17–45.
- Rodrigues, M. T. (1996). Lizards, snakes, and amphisbaenians from the Quaternary sand dunes of the middle Rio São Francisco, Bahia, Brazil. *Journal of Herpetology*, 513-523.
- Rodrigues, J. B., Gama, S. C. A., Pereira Filho, G. A., & França, F. G. R. (2015). Composition and ecological aspects of a snake assemblage on the savanna enclave of the Atlantic Forest of the Guaribas Biological Reserve in northeastern Brazil. *South American Journal of Herpetology*, 10(3), 157-164.

Rodriguez, L.B. & J.E. Cadle (1990). A preliminary overview of the herpetofauna of Cocha Cashu, Manu National Park, Peru. In: A.H. Gentry (ed.), *Four Neotropical Rainforests*, pp. 410-425. Yale University Press, New Haven, Connecticut.

Ruiz, O. (2012). Reporte de las serpientes del municipio de Tamalameque, Cesar-Colombia. *Revista MVZ Córdoba*, 2924-2927.

Ruiz, O. (2014). Ofidios del corregimiento de San Rafael de Pirú, Valencia, Córdoba-Colombia. *Revista Colombiana de Ciencia Animal-RECIA*, 3-13.

Sampaio, I. L., Santos, C. P., França, R. C., Pedrosa, I. M., Solé, M., & França, F. G. (2018). Ecological diversity of a snake assemblage from the Atlantic Forest at the south coast of Paraíba, northeast Brazil. *ZooKeys*, (787), 107-125.

Santos, M. B. D., Oliveira, M. C. L. M. D., & Tozetti, A. M. (2012). Diversity and habitat use by snakes and lizards in coastal environments of southernmost Brazil. *Biota Neotropica*, 12(3), 78-87.

Strussmann, C., & Sazima, I. (1993). The snake assemblage of the Pantanal at Poconé, western Brazil: faunal composition and ecological summary. *Studies on Neotropical Fauna and Environment*, 28(3), 157-168.

Stuart, L. C. (1948). The amphibians and reptiles of Alta Verapaz, Guatemala. *Miscellaneous Publications Museum of Zoology, University of Michigan*, 69:1- 109.

Townsend, J. H., Wilson, L. D., Medina-Flores, M., Aguilar-Urbina, E., Atkinson, B. K., Cerrato-Mendoza, C. A., ... & McKewy-Mejía, M. (2012). A premontane hotspot for herpetological endemism on the windward side of Refugio de Vida Silvestre Texíguat, Honduras. *Salamandra*, 48(2), 92-114.

Urbina-Cardona, J.N, Reynoso, V.H. 2005. Recambio de anfibios y reptiles en el gradiente potrero-borde-interior en la Reserva de Los Tuxtlas, Veracruz, México, pages 191–207, in: Halffter, G., Soberón, J., Koleff, P. & Melic, A. (eds.) *Sobre Diversidad Biológica: El significado de las Diversidades Alfa, Beta y Gamma*. CONABIO, SEA, DIVERSITAS & CONACyT. Volumen 4. Editorial Monografías Tercer Milenio, Zaragoza, España.

Vargas, F., & Bolaños, M. E. (1999). Anfibios y reptiles presentes en hábitats perturbados de selva lluviosa tropical en el bajo Anchicayá, Pacífico colombiano. *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, 23, 499-511.

Vite-Silva, V. D., Ramírez-Bautista, A., & Hernández-Salinas, U. (2010). Diversidad de anfibios y reptiles de la Reserva de la Biosfera Barranca de Metztitlán, Hidalgo, México. *Revista Mexicana de Biodiversidad*, 81(2), 473-485.

Zanella, N., & Cechin, S. Z. (2006). Taxocenose de serpentes no planalto médio do Rio Grande do Sul, Brasil. *Revista Brasileira de Zoologia*, 23(1), 211-217.

Zimmermann, B.L. & M.T. Rodrigues (1990). Frogs, snakes, and lizards of the INPA/WWF reserves near Manaus, Brazil. In: A.H. Gentry (ed.), *Four Neotropical Rainforests*, pp. 426-454. Yale University Press, New Haven, Connecticut.

References used for constructing Appendix Table A3.8

Head, J.J., Bloch, J.I., Hastings, A.K., Bourque, J.R., Cadena, E.A., Herrera, F.A., Polly, P.D., and Jaramillo, C.A. 2009. Giant boid snake from the Paleocene neotropics reveals hotter past equatorial temperatures. *Nature*, 457:715-717

Head, J. J. (2015). Fossil calibration dates for molecular phylogenetic analysis of snakes 1: Serpentes, Alethinophidia, Boidae, Pythonidae. *Palaeontologia Electronica*, 18(6FC), 1-17.

Head, J. J., Mahlow, K., & Mueller, J. (2016). Fossil calibration dates for molecular phylogenetic analysis of snakes 2: Caenophidia, Colubroidea, Elapoidea, Colubridae. *Palaeontologia Electronica*, 19(2FC), 1-21.

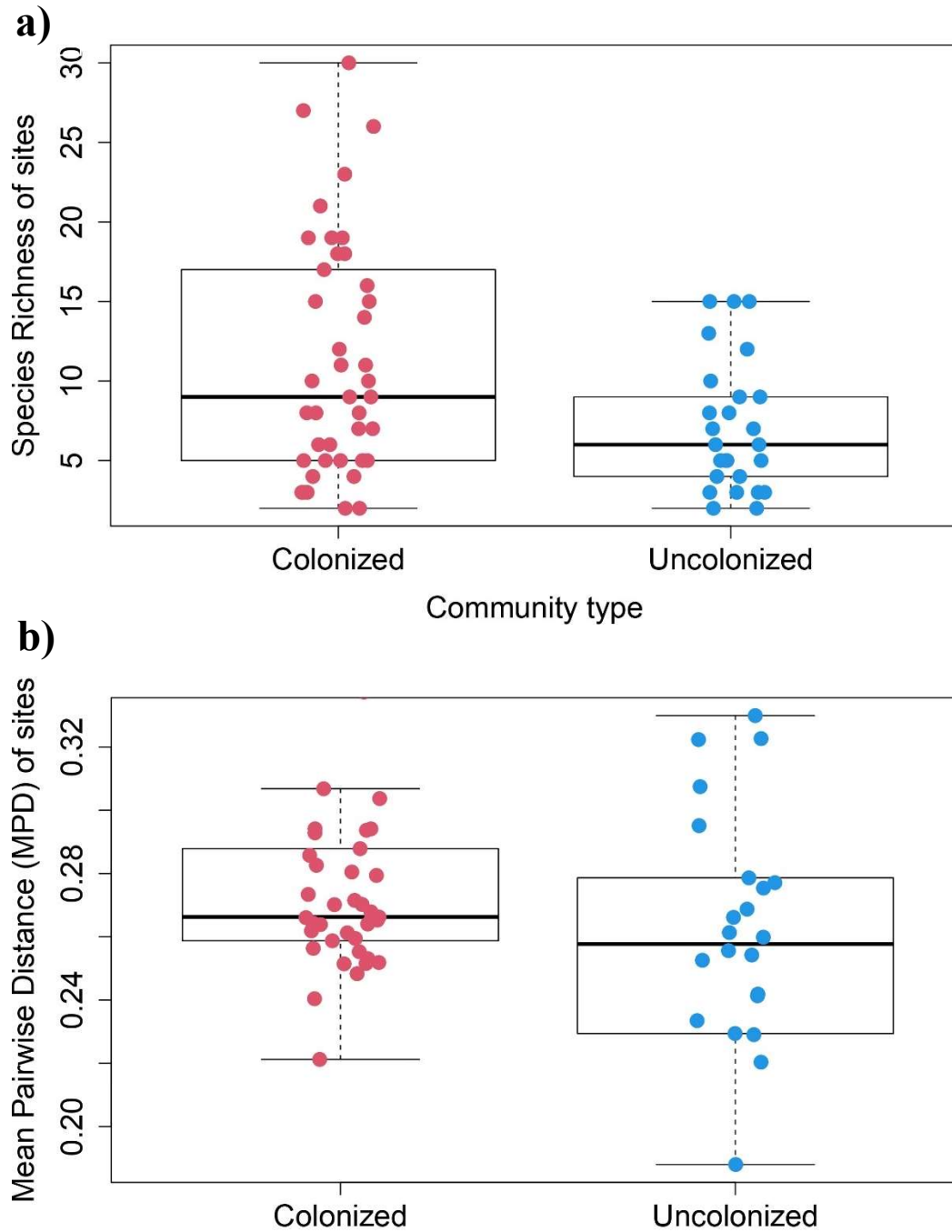
McCartney, J.A., Stevens, N.J., and O'Connor, P.M. 2014. The earliest colubroid-dominated snake fauna from Africa: perspectives from the Late Oligocene Nsungwe formation of southwestern Tanzania. *PLoS ONE* 9(3): e90415

Rage, J.-C. and Werner, C. 1999. Mid-Cretaceous (Cenomanian) snakes from Wadi Abu Hashim, Sudan: The earliest snake assemblage. *Palaeontologia Africana*, 35:85-110

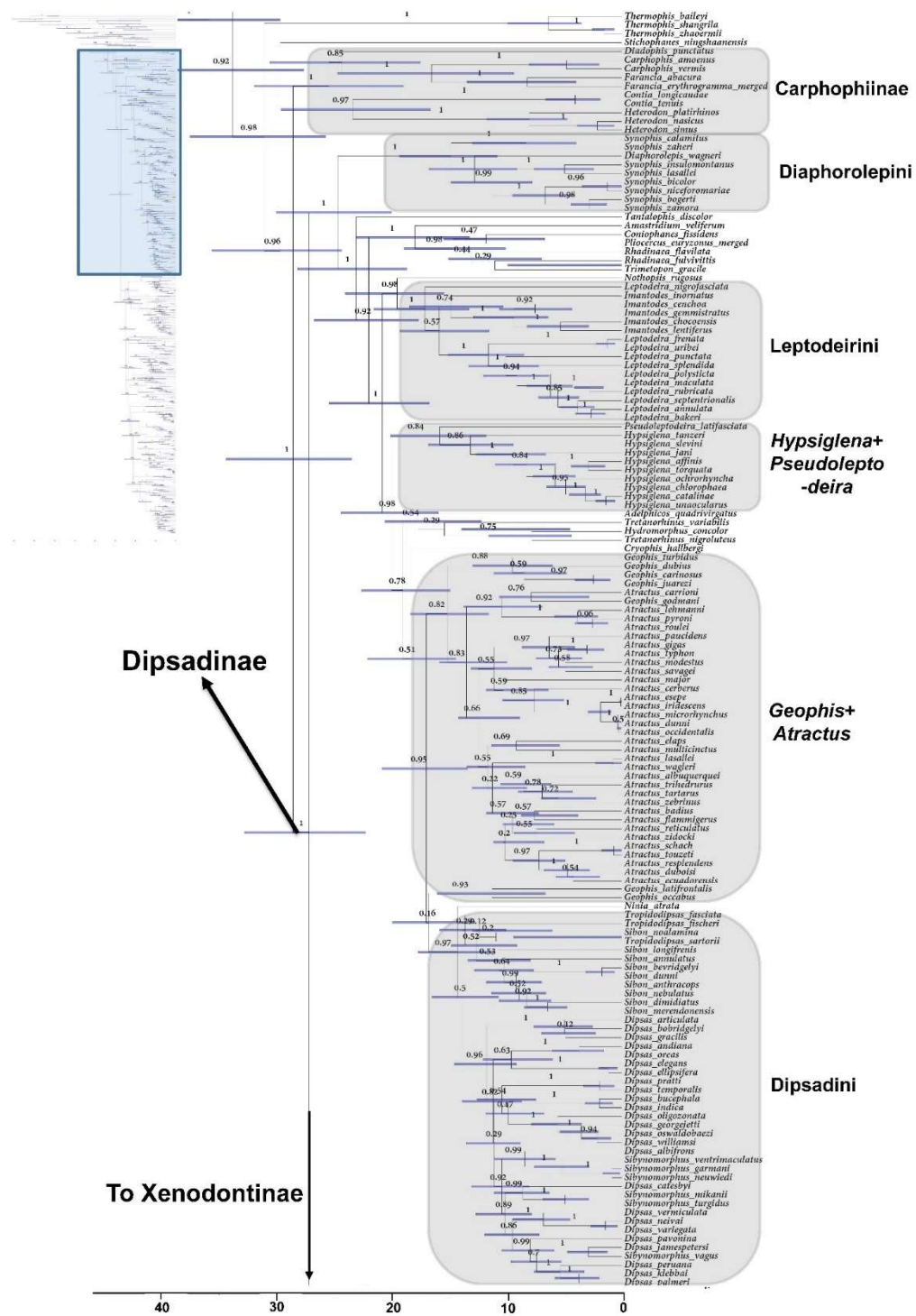
Rage, J.-C., Folie, A., Rana, R S., Singh, H., Rose, K.D. and Smith, T. 2008. A diverse snake fauna from the early Eocene of Vastan Lignite Mine, Gujarat, India. *Acta Palaeontologica Polonica*, 53:391-403.

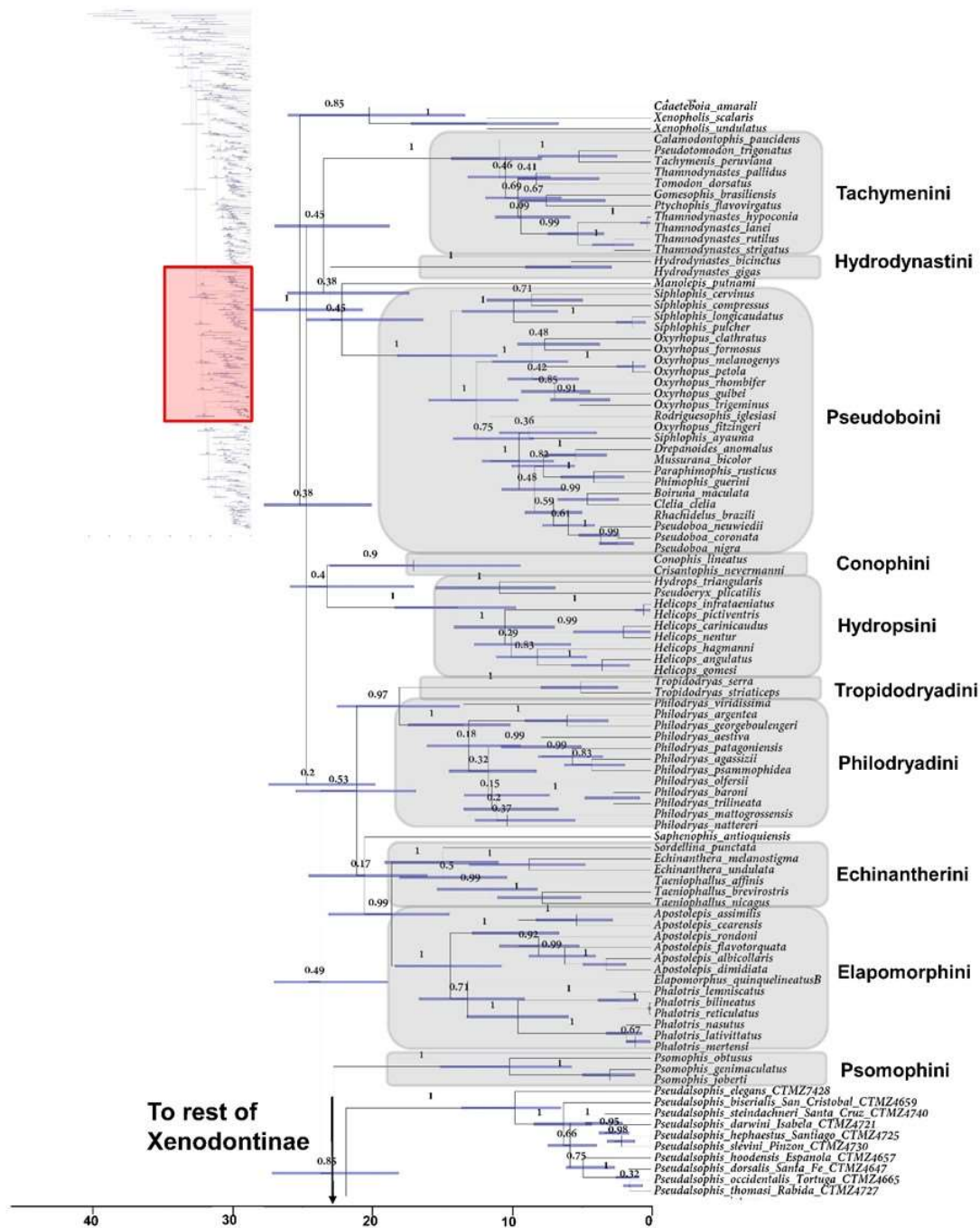
Zaher, H., Murphy, R. W., Arredondo, J. C., Graboski, R., Machado-Filho, P. R., Mahlow, K., ... & Grazziotin, F. G. (2019). Large-scale molecular phylogeny, morphology, divergence-time estimation, and the fossil record of advanced caenophidian snakes (Squamata: Serpentes). *PloS One*, 14(5), e0216148.

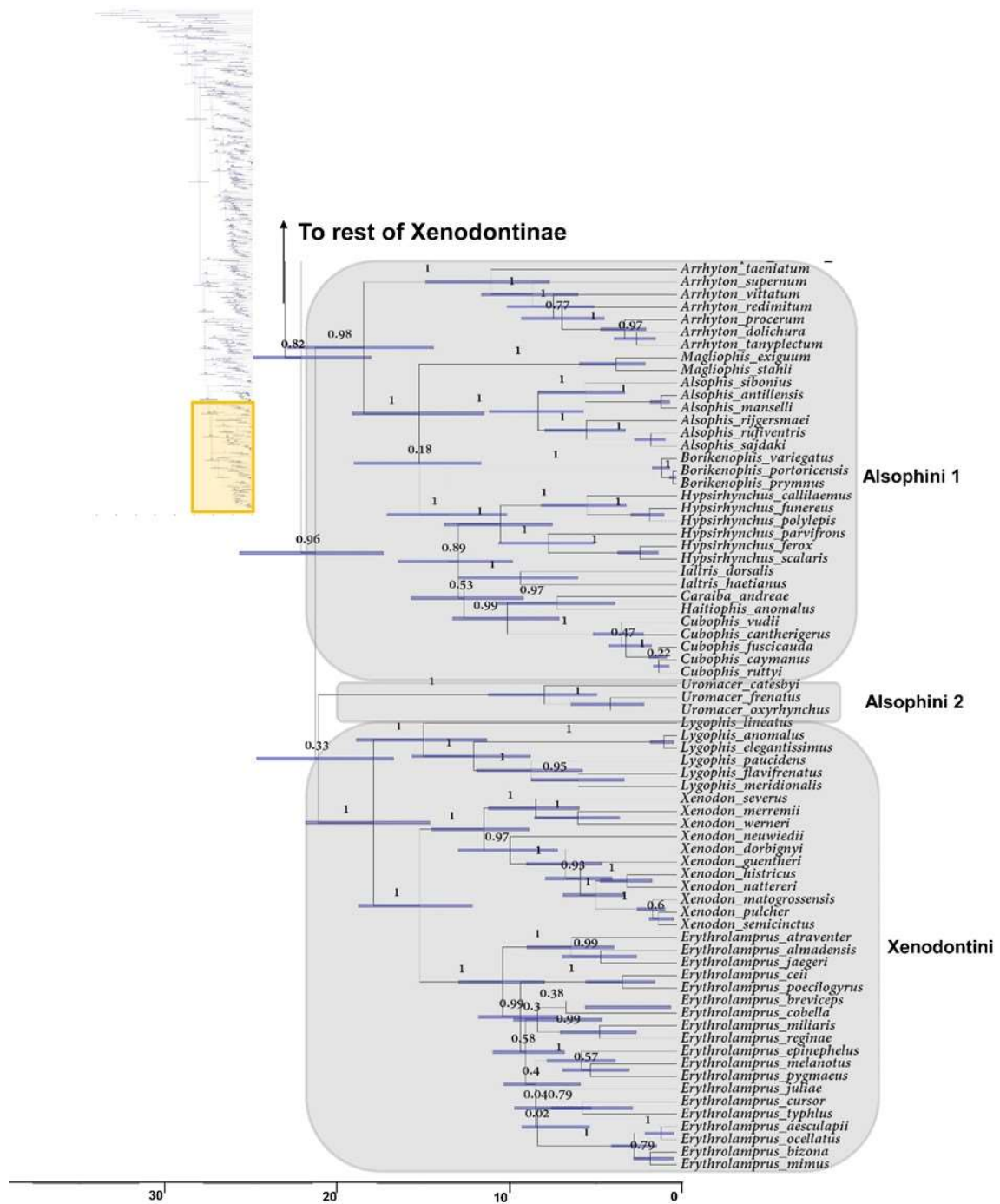
Appendix Figure A3.1.



Appendix Figure A3.1. Relationship between the a) species richness or b) functional diversity (as measured with the MPD metric) of the communities included in this study and whether they contain or not taxa colonizing them across the opposite side of the Panamanian Isthmus







Appendix Figure A3.2. Timetree of the family Dipsadidae derived from, calibrated with the fossils presented in Appendix Table A3.8 of the attached manuscript. Numbers above branches are the posterior probabilities of adjacent nodes. 95% credibility intervals of the age of each node are indicated with blue bars. The taxonomic arrangement of the members of Dipsadidae into tribes follows Zaher et al. (2009), incorporating recent taxonomic changes such as the dissolution of the Saphenophini tribe by Zaher et al. (2018). The units of the scale correspond to millions of years before the present time.



Appendix Figure A3.3. Map of the biogeographic regions considered in the biogeographic analyses performed herein (modified after Morrone, 2014). The location of the 67 sites or communities that served as the basis for our study is also indicated.

CAPTIONS OF SUPPLEMENTAL FILES ATTACHED TO THIS DISSERTATION

Chapter 2.

Supplemental Table S2.1. Accession numbers of the sequences used for constructing the timetree employed for this work

Supplemental Table S2.2. Taxonomic arrangement of the Dipsadidae used for the TACT analyses performed for this work, following Uetz et al. (2023).

References used for constructing Suppl. Table 2.2

Chapter 3.

Supplemental Tables:

Supplemental Table S3.1. Species composition data of the 67 sites included herein from members of the Dipsadidae.

Supplemental Table S3.2. Scores for the Principal Component (PC) Axes corresponding to the PC analyses performed with the climatic variables of the sites included in this work.

Supplemental Table S3.3. Species composition data of the 67 sites included herein from species from families different from Dipsadidae.

Supplemental Table S3.4. Functional data corresponding to each of the species found in the 67 sites included in this work.

Supplemental Table S3.5. Functional data corresponding to each of the species from families different from Dipsadidae found in the 67 sites included in this work.

Supplemental Table S3.6. GenBank accession numbers of each of the sequences used for constructing the timetree used for this work

Supplemental Table S3.7. Table indicating the areas inhabited by each of the species of the Dipsadidae included in the Bayesian timetree used for the ancestral area estimation performed herein

References used for constructing Supplemental Tables S3.8 and S3.9

Supplemental Figures:

Supplemental Figure S3.1. Results of the ancestral estimations as recovered by 'BioGeoBEARS'