GENOMIC STRUCTURE AND DIVERSITY IN NEOTROPICAL HERPETOFAUNA: THE ROLE OF DEMOGRAPHIC HISTORY AND GENE FLOW

by

DANIELLE GEORGE RIVERA

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

AUGUST 2021

Acknowledgements

I would very much like to acknowledge my dissertation supervisor and mentor, Dr. Matthew Fujita, for always being a constant source of support and guidance throughout my time at UTA. I am one among the many voices of my peers in signifying that Dr. Fujita's dedication to other's well-being, creativity, and science is unparalleled and we are all better scientists, and people, for having someone like him on our side. From my first day in his lab, Dr. Fujita has tailored his mentoring to my specific needs and has gone above and beyond to aid in my success. I have learned so much about patience, leadership, and caring from my experiences with him. He has been an immensely important catalyst and source for positive motivation and leadership. I can only hope to one day emulate his success – after all… I'm no Matt Fujita.

I express my gratitude to my Dissertation committee, Dr. Jeff Demuth, Dr. Esther Betran, Dr. Todd Castoe, and Dr. Eric Smith, for their valuable input, guidance, and constructive comments throughout my PhD. Thanks to the greatest collaborator who ever lived, Dr. Miguel Rodrigues, for your trust, knowledge, and many many tissues!!! Without you, I would certainly not have a dissertation! And to Dr. Ivan Prates, your guidance, both scientific and emotional, constant support, and amazing manuscript edits, have elevated my PhD to a level I could only dream of reaching. Special and unending thanks to the mentor who started my entire scientific career, Dr. Ana Carnaval. Dr. Carnaval spoiled me with good mentorship, humor, and Brazilian culture – my abilities in the lab, as a writer, and in the field, were all built by her very capable, boss-lady hands.

To my lab-mate, roommate, and bestie, Kathleen Currie, I want to thank you so much for being the same kind of weird and crazy as I am – without your support and friendship, this would have been a very different journey for me, and I am eternally grateful that I met you. I know our friendship doesn't end here; like a serious disease, we can never get rid of each other. To my other lab-mates: TJ Firneno, Jose Maldonado, and Dan Dudeck, I thank you for the support you've always offered, academically and personally, that has helped me progress successfully while at UTA. Thanks to Dr. Corey Roelke, honorary Fujita Lab member, for the knowledge, laughs, and insane conversations – I know those things have helped prepare me for whatever comes next.

To my fellow grad students and postdocs, both past and present at UTA, CCNY, and the AMNH, I give thanks for the unending support, and the many hours studying, writing, commiserating, drinking, and eating together. Without these shared experiences, I would never have grown into the person I have become, and I'm very glad that I did. Special thanks to Drs. "Guru" Dan Portik, Shannon Beston, Jared Goos, and Nicky Hales as well as my lovely Anna Penna for their friendships and particularly weird and savvy life-guidance and support. To the friends I've made both here and abroad: thank you so much!

Finally, to my family, including Josh Rivera, Yvette Rivera-Dale, Brittany Decker, Megan Decker, Caitlin Burke, and my little nephews Owen and Lucas, thank you for always being there for me. You all shaped who I was as a person, so that I could be a better scientist.

Dedication

I dedicate my PhD to my father, George Rivera, who left us far too soon. While he never fully understood the scientific jargon, he never stopped supporting me in whatever way I needed. Even in his passing, he has enabled me to reach further than I could have hoped. Bendiciónes, y muchas gracias por tu amor.

TABLE OF CONTENTS

Introduction

The tropics contain some of the world's most diverse ecological communities (Myers, et al., 2000). Identifying the evolutionary and ecological mechanisms responsible for the origination and persistence of this rich biodiversity has played a central role in our understanding of both local and global diversification patterns. Understanding what factors promote lineage persistence over evolutionary time, as well as the accumulation of evolutionary potential in geographic space, is key in the conservation of nature (Carnaval, et al., 2009; Oaks, 2019). The Neotropical region houses some of the world's most diverse and threatened ecosystems, but the historical and contemporary processes that have led to its high species richness and endemism remain relatively poorly known (Carnaval et al., 2009, 2014).

A number of challenges surrounding the characterization of Neotropical biodiversity involve our limitations in the use and interpretation of the biological data collected. For instance, while reproductive isolation has long been viewed as the primary factor behind lineage divergence and stable boundaries between closely related species, how introgression affects reproductive isolation and speciation has remained an enduring question in evolutionary biology (Avise et al., 1998; Mayr, 1963; Rabosky, 2016). When closely related populations come into contact, gene flow via hybridization can lead to the introgression of alleles (Mallet, 2005; O'Connell et al., 2021). Introgression levels can vary starkly across genome regions, leading to phenomena such as observed mito-nuclear discordance and confounding estimates of demographic histories. In Chapters I and II of this dissertation, we focus on two clades of neotropical toads -- the *Rhinella marina* and *Rhinella granulosa* species complexes, respectively -- to assess the true extent of genetic introgression and resolve mito-nuclear discordance across species thought to hybridize to extreme degrees based on natural history observations and multi-locus analyses. Use of multi-locus genetic datasets of select loci have uncovered patterns of

complex demographic histories and phylogeographic study of a number of Neotropical species (Firneno et al., 2020; Rivera, et al., 2020). Gene tree discordance has made phylogeographic reconstruction challenging, but this issue has been improved by the use of high throughput, reduced representation, or whole genome sequencing (Firneno et al., 2020; Graham et al., 2018). With the addition of genome-scale data, the identification of the patterns and mechanisms that both drive and maintain Neotropical biodiversity have become more tractable. Most hypotheses proposed to explain spatial biodiversity patterns in the Neotropics have invoked landscape configuration and change as key drivers of dispersal, range limitation, lineage divergence, and speciation (Carnaval et al., 2014; Dal Vechio, et al., 2019; Prates, et al., 2016; Rivera et al., 2020). These hypotheses have often been applied to explain current species distribution patterns and assemblage composition in other regions, becoming central to biogeographic investigations worldwide (Leaché et al., 2019; Potter et al., 2019). In Chapter III of this dissertation, we use high throughput sequencing data from co-distributed Neotropical species to infer species-specific patterns of widespread mechanisms, such as historical introgressive hybridization, and to investigate the contribution that landscape features may have on population co-divergence and demographic change on the basis of a comparative phylogeographic approach.

Our results point to highly heterogeneous levels of cross-species genetic introgression and hybridization in the evolutionary history of even closely related clades, as well as highly discordant patterns of demographic change among co-distributed taxa in response to physiographic barriers and former climatic change. Our combined approach illustrates the value of molecular evolution and comparative phylogeography in understanding how population processes and landscape clines have contributed to present-day patterns of biodiversity. Our results also challenge simplistic views about the role of hybridization and assemblage-level

demographic change in species formation and persistence, suggesting that models developed in regions with less complex biotas are often insufficient to explain patterns of biological diversification in the world's tropical regions.

References

Carnaval, A. C., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T., & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, 323(5915), 785–789.

Carnaval, A. C., Waltari, E., Rodrigues, M. T., Rosauer, D., VanDerWal, J., Damasceno, R., … Moritz, C. (2014). Prediction of phylogeographic endemism in an environmentally complex biome. *Proceedings. Biological Sciences / The Royal Society*, 281(1792). doi: 10.1098/rspb.2014.1461

Dal Vechio, F., Prates, I., Grazziotin, F. G., Zaher, H., & Mt, G. R. R. (2019). Rain forest shifts through time and riverine barriers shaped the diversification of South American terrestrial pit vipers (*Bothrops jararacussu* species group). *Journal of Biogeography*, 47, 516–526.

Firneno, T. J., Jr, O'Neill, J. R., Portik, D. M., Emery, A. H., Townsend, J. H., & Fujita, M. K. (2020). Finding complexity in complexes: Assessing the causes of mitonuclear discordance in a problematic species complex of Mesoamerican toads. *Molecular Ecology*, 29(18), 3543–3559.

Graham, A. M., Lavretsky, P., Muñoz-Fuentes, V., Green, A. J., Wilson, R. E., & McCracken, K. G. (2018). Migration-Selection Balance Drives Genetic Differentiation in Genes Associated with High-Altitude Function in the Speckled Teal (*Anas flavirostris*) in the Andes. *Genome Biology and Evolution*, 10(1), 14–32.

Leaché, A. D., Portik, D. M., Rivera, D., Rödel, M., Penner, J., Gvoždík, V., … Fujita, M. K. (2019). Exploring rain forest diversification using demographic model testing in the African foam-nest treefrog *Chiromantis rufescens*. *Journal of Biogeography*, 46(12), 2706–2721.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. Nature, 403(6772), 853–858.

Oaks, J. R. (2019). Full Bayesian Comparative Phylogeography from Genomic Data. *Systematic Biology*, 68(3), 371–395.

Potter, S., Afonso Silva, A. C., Bragg, J. G., Catalano, S. R., Donnellan, S., Doughty, P., … Moritz, C. (2019). Contrasting scales of local persistence between monsoonal and arid biomes in closely related, low-dispersal vertebrates. *Journal of Biogeography*, 46(11), 2506–2519.

Prates, I., Rivera, D., Rodrigues, M. T., & Carnaval, A. C. (2016). A mid-Pleistocene rainforest corridor enabled synchronous invasions of the Atlantic Forest by Amazonian anole lizards. *Molecular Ecology*, 25(20), 5174–5186.

Rivera, D., Prates, I., Rodrigues, M. T., & Carnaval, A. C. (2020). Effects of climate and geography on spatial patterns of genetic structure in tropical skinks. *Molecular Phylogenetics and Evolution*, 143, 106661.

Chapter I: Phylogenomics, introgression, and demographic history of South American true toads (*Rhinella*)

Abstract

The effects of genetic introgression on species boundaries and how they affect species' integrity and persistence over evolutionary time have received increased attention. The increasing availability of genomic data has revealed contrasting patterns of gene flow across genomic regions, which impose challenges to inferences of evolutionary relationships and of patterns of genetic admixture across lineages. By characterizing patterns of variation across thousands of genomic loci in a widespread complex of true toads (*Rhinella*), we assess the true extent of genetic introgression across species thought to hybridize to extreme degrees based on natural history observations and multi-locus analyses. Comprehensive geographic sampling of five large-ranged Neotropical taxa revealed multiple distinct evolutionary lineages that span large geographic areas and, at times, distinct biomes. The inferred major clades and genetic clusters largely correspond to currently recognized taxa within *Rhinella*; however, we also found evidence of cryptic diversity within taxa. Phylogenetic analyses revealed extensive mito-nuclear discordance, while genetic clustering analyses uncovered several admixed individuals within major genetic groups. Accordingly, historical demographic analyses supported that the evolutionary history of these toads involved cross-taxon gene flow both at ancient and recent times. Lastly, ABBA-BABA tests revealed widespread allele sharing across species boundaries, a pattern that can be confidently attributed to genetic introgression as opposed to incomplete lineage sorting. These results confirm previous assertions that the evolutionary history of *Rhinella* was characterized by various levels of hybridization even across environmentally

heterogeneous regions, posing exciting questions about what factors prevent complete fusion of diverging yet highly interdependent evolutionary lineages.

Introduction

How introgression affects reproductive isolation and speciation is an enduring question in evolutionary biology. Reproductive isolation has long been viewed as the primary factor behind lineage divergence and stable boundaries between closely related species (Avise et al., 1998; Mayr, 1963; Rabosky, 2016). When closely related populations come into contact, however, gene flow via hybridization can lead to the introgression of alleles (Mallet, 2005; O'Connell et al., 2021). Introgression levels can vary starkly across genome regions. In particular, in the presence of strong divergent selection, those loci underlying adaptive phenotypes can maintain marked differentiation even with extensive gene flow among closely related populations (Feder et al., 2012). Thus, these varying degrees of isolation across the genome may contribute to the maintenance of species boundaries despite the homogenizing effects of gene flow (Yeaman & Whitlock, 2011).

Differential introgression across genomic regions can lead to dramatic topological discordance between genealogies inferred from distinct genes, as illustrated by instances of mito-nuclear discordance (Bernardo et al., 2019; Bessa-Silva et al., 2020; Firneno et al., 2020). This gene-tree heterogeneity must be accounted for as it can make reconstructing evolutionary relationships and historical demography challenging (Carstens & Knowles, 2007; Firneno et al., 2020; Liu et al., 2010). The increasing availability of high-throughput sequencing datasets for non-model organisms has improved our ability to discern patterns of introgression in closely related species or populations (Firneno et al., 2020; Graham et al., 2018; Lavretsky et al., 2016) and thus clarify phylogenetic relationships and species limits. This is especially so in large,

widely distributed species complexes with limited variation in external morphological traits and hybridization blurring species limits (Guo et al., 2016; Phuong et al., 2017; Potter et al., 2016).

The increasing availability of genome-scale datasets has also fostered the development of model-based approaches to infer historical demographic events such as population size shifts and pulses of gene flow (Portik, et al., 2017; Prates, Xue et al., 2016). These approaches have transformed our understanding of how landscape and climate changes have contributed to the assembly of regional species pools, for instance by limiting dispersal, promoting speciation, or leading to lineage fusion (Graham et al., 2018; Lavretsky et al., 2016; Leaché et al., 2019; Portik, et al., 2017) One flexible approach involves simulating population histories to compare the fit of empirical genome-scale data to data simulated under alternative biogeographical scenarios (Dal Vechio et al., 2019; Portik, et al., 2017; Prates et al., 2016). This modeling framework can facilitate hypothesis testing, such as how climate-driven habitat shifts may have led to migration, introgression, or isolation across geographic regions. These approaches have been instrumental to shed light on the historical factors behind present-day spatial biodiversity patterns in regions that concentrate large proportions of biodiversity. This is the case of the Neotropics, where demographic inference has supported that Late-Quaternary climate fluctuations and Neogene geomorphological change have played a major role in shaping species range limits, genetic diversity levels, and lineage divergence (Gehara et al., 2017; Pirani et al., 2020; Prates, Xue et al., 2016). Nevertheless, biogeographic investigations in the Neotropics have often shown geographic and taxonomic bias, which questions the generality of the mechanisms invoked to explain species richness and distributions. For instance, taxa with wide ranges across South America's open vegetation biomes – the dry and highly seasonal Cerrado, Caatinga, and Chaco –

have received relatively less attention than rainforest biotas (Fonseca et al., 2018; Gehara et al., 2017; Werneck, 2011).

One example of a Neotropical clade whose biogeography history remains poorly known is the true South American toads, genus *Rhinella* (Bufonidae). Despite being the focus of a handful of phylogeographic studies, the evolutionary relationships and species limits between these toads remain elusive, perhaps due to wildly varying patterns of introgression and hybridization across species (Maciel et al., 2010; Pereyra et al., 2016; Pereyra et al., 2021; Sequeira et al., 2011; Vallinoto et al., 2009). As such, not only the evolutionary history of this group is unclear, but so are the environmental and geographic factors that may have favored introgression and its variation, or how hybridization may have contributed to lineage divergence or fusion (Azevedo et al., 2003; Correa et al., 2012; Malone & Fontenot, 2008; Pereyra et al., 2016; Sequeira et al., 2011). *Rhinella* is composed of multiple species complexes that are each distributed across much of the Neotropics. These groups are known to harbor high levels of cryptic lineage diversity, as revealed by single and multi-locus genetic analyses (Maciel et al., 2010; Pereyra et al. 2016; Pereyra et al. 2021; Vallinoto et al., 2009). Among them is the *Rhinella marina* group, best known for the globally invasive species *R. marina*. Previous studies of this group have identified both mitochondrial and nuclear introgression across species (Azevedo et al., 2003; Maciel et al., 2010; Vallinoto et al., 2009). However, lack of data about persisting genetically admixed populations in the wild makes it difficult to assess the magnitude of presumed hybridization and how it affects species boundaries (Azevedo et al., 2003; Malone & Fontenot, 2008; Pereyra et al. 2021). Despite the ecological diversity seen in *Rhinella*, with taxa that span savannas, rainforests, and xeric shrublands, biogeographic analyses have largely focused on taxa occurring within a single biome (Sequeira et al., 2011; Thomé et al., 2010),

which is also the case of other South American anuran clades (Fonseca et al., 2018; Gehara et al., 2017; Oliveira et al., 2018). As a result, how habitat transitions may contribute to patterns of gene flow and species range limits remains unclear.

In this investigation, we focus on the *R. marina* group to investigate evolutionary relationships, quantify the extent of hybridization, and examine whether landscape transitions among South America's biomes impose limits to gene flow and species ranges. For this purpose, we focus on *R. marina*, *R. poeppigii*, *R. horribilis*, *R. jimi*, and *R. schneideri*, which have established contact zones throughout the continent. We infer population structure, gene flow, and relationships based on geographically comprehensive sampling of genomic variation within each taxon. We then proceed to test alternative historical hypotheses to quantify plausible demographic events such as population size shifts and historical gene flow. With this approach, we seek to answer the following questions: what are the levels of genetic structure across and within each species? Do genomic data corroborate a pattern of widespread admixture or introgression across these species, as previously suggested based on only a few loci? Lastly, what historical demographic processes may explain species distributions and genetic diversity patterns within this clade?

Material and Methods

Sample collection

Our sampling included 185 individuals belonging to the *Rhinella marina* species group, as follows: 67 *R. marina*, 39 *R. schneideri*, 22 *R. horribilis*, 11 *R. jimi*, and nine *R. cf. poeppigii*,

four *R. veredas*, eight *R. rubescens*, and 25 *R. icterica*. We also included samples from the *Rhinella granulosa* and *R. margaritifera* major clades within *Rhinella* as outgroups in the divergence time estimation analyses (see below). Within each species, we sample multiple individuals from each locality across their known ranges, with the exception of *R. cf. poeppigii*, which was identified as distinct from *R. marina a posteriori* based on the genetic data (see Results). Tissue samples were obtained from the MTR herpetological tissue collection hosted at Instituto de Biociências, University of São Paulo (IBUSP) with vouchers at Museum of Zoology, University of São Paulo, as well as from the Amphibian and Reptile Diversity Research Center (ARDRC), and the Louisiana State University Museum of Natural Science (LSUMNS).

DNA extraction, amplification, & sequencing

We extracted genomic DNA using a standard phenol-chloroform extraction protocol (Sambrook & Russell, 2006). Fragments of the mitochondrial 16S were amplified using 16Sar and 16Sbr primers and sequenced on an ABI 3730xL (Primer information and PCR conditions in the Supplementary Text S1). Sequences were edited and aligned in Geneious Prime 2020.0.4 (Identification and Accession numbers in Supplementary Table S1). We generated double-digest restriction-site associated DNA sequencing (ddRADseq) data following (Peterson, et al., 2012), with modifications as described in Streicher et al. (2014). Briefly, 200-500 ng of DNA were digested using the *Sbf*I (restriction site 5′-CCTGCAGG-3′) and *Msp*I (restriction site 5′-CCGG-3′) restriction enzymes in a single reaction using the manufacturer's recommended buffer (New England Biolabs) for 5 hr at 37°C. Digested DNA was bead-purified before ligating barcodes and index adaptors, then samples with the same index were pooled and size-selected (415-515

bp) on a Blue Pippin Prep size selector (Sage Science). Final library preparation was analyzed and quantified on a BioAnalyzer (Agilent) and Qubit Fluorometer 4 (Thermo Fisher Scientific). The resulting 100 bp single-end libraries were sequenced at MedGenome on an Illumina HiSeq2500.

We used the command line version of ipyrad v. 0.9.45 (Eaton & Overcast, 2020) (available at https://ipyrad.readthedocs.io) to de-multiplex and assign reads to individuals based on sequence barcodes (allowing no mismatches from individual barcodes), perform *reference* read assembly (minimum clustering similarity threshold = 0.90), align reads into loci, and call single nucleotide polymorphisms (SNPs). As a reference, we used the *Rhinella marina* genome (Edwards et al., 2018). A minimum Phred quality score (= 33), sequence coverage (= 6x), read length (= 35 bp), and maximum proportion of heterozygous sites per locus (= 0.5) were enforced, while ensuring that variable sites had no more than two alleles (i.e., a diploid genome). Following the initial assembly, we used Matrix Condenser (de Medeiros & Farrell, 2018) to assess levels of missing data across samples and then re-assembled our dataset to ensure a minimum sample coverage of less than 35% missing loci within each sample and at least 75% of samples at each locus. This strategy resulted in a final dataset composed of 49,376 SNPs at 3,318 RAD loci with less than 12% missing data. Additionally, Weir and Cockerham mean F_{ST} estimates for the ddRADseq dataset using VCFTools (Danecek et al., 2011) and Nei's G_{ST} for the mitochondrial dataset were calculated using the R package `mmod` (Winter, 2012).

Inferring population structure and genetic admixture

Based on the ddRAD data, we used a genetic clustering approach to estimate the number of demes and if admixture was present among them. We assembled a SNP dataset as described

above but excluding outgroups and using only one SNP per RAD locus to maximize sampling of independent SNPs. This approach resulted in a dataset composed of 3,314 SNPs. Genetic clustering was performed using the maximum likelihood method `ADMIXTURE`, testing up to 15 populations with 20 replicates per K and a 10-fold cross-validation (Alexander, et al., 2009; Portik, 2016). The best K was determined by assessing the replicate with the lowest cross-validation error. To further characterize population structure, we used the non-parametric method of discriminant analysis of principal components (DAPC), implemented in the R package `adegenet` (Jombart & Ahmed, 2011; Jombart, et al., 2010). The *find.clusters* function was used to test the fit of 1-15 clusters (K). The K with the lowest Bayesian information criterion (BIC) score was considered the best-fit number of demes. The resulting ancestry coefficient matrices (Q-matrices) were then imported into QGIS (QGIS Development Team 2020. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org) to make average-per-locality pie-charts indicating admixture levels at each sampled locality for each species.

Phylogenetic Analyses

We reconstructed maximum likelihood phylogenies for both the mitochondrial and the unlinked SNP ddRADseq datasets using `IQTREE` v2.1.2, utilizing the built-in model selection tool ModelFinder Plus, implementing 1000 ultrafast bootstraps (Hoang et al., 2018; Kalyaanamoorthy et al., 2017; Nguyen et al., 2015). We specified that all partitions share the same branch lengths and selected the best-fit partitioning scheme by merging partitions (which implements the greedy algorithm of PartitionFinder), testing the “MrBayes" substitution model set and considering the top 10% partition schemes using the fast relaxed clustering algorithm

from PartitionFinder2 to save computational time (Chernomor et al., 2016; Lanfear et al., 2012; Lanfear et al., 2014; Lanfear et al., 2017). In addition, we performed phylogenetic inference under a Bayesian framework for both datasets using MrBayes 3.2.6 (Ronquist et al. 2012), implementing three independent runs of four Markov chains of 10 million generations each and sampling every 1,000 generations with the first 25% generations discarded as burn-in. We used Tracer 1.7 (Rambaut et al. 2018) to assess whether Markov chain mixing was adequate (effective sample sizes > 200) and to visually assess model parameter stationarity and convergence between runs. We then summarized a 50% majority-rule consensus tree.

To estimate divergence dates and inform the delimitation of species boundaries, we conducted Bayesian divergence dating analyses based on the mtDNA dataset in BEAST2 using an HKY model of nucleotide substitution, a log-normal relaxed molecular clock, and a Yule process speciation model. We follow Pramuk et al. (2008) by enforcing a minimum age for the root node between the *Rhinella marina* and *R. granulosa* species complexes based on a *Rhinella marina* fossil from the Clarendonian North American Stage of the middle Miocene (ca. 11 mya), described by Sanchiz (1998), and employed a normally distributed prior with a standard deviation of 0.5. We ran this analysis for 20 million generations sampling every 1000 generations. Runs were assessed using TRACER v1.6 (Rambaut & Drummond, 2009) to examine convergence. We then summarized a maximum clade credibility tree using TreeAnnotator discarding the first 25% of trees as burn-in (Bouckaert et al., 2019; Stamatakis, 2014). All phylogenetic tree-based methods were analyzed on Cipres (Miller et al., 2010).

Demographic modeling with ∂a∂i

We used the diffusion-approximation method ∂a∂i (Gutenkunst, et al., 2009) to test alternative hypotheses of population history within the *Rhinella marina* clade. Using both two- and three-dimensional joint site frequency spectra (2D- and 3D-JSFS), we divided the dataset into two population subsets: one comprised of *R. marina, R. horribilis*, and *R. jimi*; and another comprised of *R. schneideri* and *R. cf. poeppigii*. Folded-JSFS datasets were used in all ∂a∂i analyses.

We filtered the ddRAD data to allow no more than 35% missing data from any sample, removed singletons, and selected one SNP per locus using VCFtools (Danecek et al., 2011; Gutenkunst et al., 2009; Portik et al., 2017). We then used the *stacks_pipeline* Python script from Portik et al. (2017) to create the SNP input file for ∂a∂i. We used the python script easySFS (https://github.com/isaacovercast/easySFS) to determine the projection size of each population, which was determined by balancing a downscaled sample size that maximized the number of segregating sites (Gutenkunst et al., 2009; Marth et al., 2004). We then tested a range of extrapolation grid sizes (40-100 in 10-unit increments, e.g., 50, 60, 70 to 100, 110, 120) in the divergence-with-no-migration model to determine the appropriate grid size by selecting the model with the highest log-likelihood, implementing 4 rounds of optimization totaling 100 replicates. Once an optimal grid size was determined, each tested model was run 3 times independently.

For the subset composed of *R. marina, R. horribilis*, and *R. jimi*, we used a 3D-JSFS to test models incorporating gene flow at different times, including those accounting for ancient migration, recent secondary contact, and past simultaneous divergence of all lineages (Fig. S5). In addition to a model of 1) divergence with no migration, we tested the following models: 2) divergence with continuous symmetric gene flow between all populations; 3) divergence with

continuous symmetric gene flow between geographically adjacent populations; 4) isolation followed by secondary contact; 5) simultaneous divergence in isolation followed by more recent secondary contact between adjacent populations; 6) simultaneous divergence with continuous symmetric migration between adjacent populations; 7) ancient migration with very recent isolation; 8) ancient migration with a longer period of recent isolation; 9) a short ancient period of migration followed by a long period of isolation; and 10) ancient migration followed by lineage isolation and population size change across two epochs (Barratt et al., 2018; Portik et al., 2017).

For the subset composed of *R. schneideri* and *R. cf. poeppigii*, we tested 2D-JSFS models incorporating differing migration levels at different time periods (Fig. S6). In addition to a model of 1) divergence with no migration, we tested the following models: 2) divergence with continuous symmetric migration; 3) divergence with continuous asymmetric migration; 4) divergence with continuous symmetric migration and a varying rate of migration across two epochs; 5) divergence with continuous asymmetric migration and a varying rate of migration across two epochs; 6) divergence in isolation, followed by symmetric secondary contact; 7) divergence in isolation, followed by asymmetric secondary contact; 8) ancient symmetric migration then subsequent isolation; 9) ancient asymmetric migration then subsequent isolation; 10) divergence in isolation followed by symmetric secondary contact with subsequent isolation; and 11) divergence in isolation followed by asymmetric secondary contact with subsequent isolation (Charles et al., 2018; Portik et al., 2017).

Best-fit models were chosen based on log-likelihood values, which we assumed to be the true likelihood (and not composite likelihood) given that we have kept only one SNP per RAD

locus. Replicates with the consistently highest likelihood scores were used to calculate and compare models using the Akaike information criterion (AIC).

Inferring gene flow

To further explore potential hybridization between taxa, we inferred Patterson's D statistic, or the ABBA-BABA statistic, and the related admixture fraction estimates, or f_4-ratio statistics, based on the ddRAD data using Dsuite (Malinsky, et al., 2020; Patterson et al., 2012). Tests were designed with a 4-taxon fixed phylogeny (((P1,P2)P3)O), wherein a typical ancestral ("A") and derived ("B") allele pattern should follow BBAA. Under incomplete lineage sorting, conflicting ABBA and BABA patterns should occur in equal frequencies, resulting in a D statistic = 0. If, however, introgression between P3 and P1 or P2 has occurred, there should be an excess of these patterns and a D statistic significantly different from 0, with significance detected using a block-jackknifing approach (Durand, et al., 2011; Green et al., 2010; Malinsky et al., 2020; Patterson et al., 2012). We used the *f*-branch or f_b(C) metric to tease apart potentially correlated f_4-ratio statistics and estimate gene flow events between internal branches on the phylogeny (Malinsky et al., 2018; Martin et al., 2013). `Dsuite` uses a VCF file and a jackknifing approach to assess correlations in allele frequencies between closely-related species (Malinsky et al., 2020). Within `Dsuite`, we used the *Dtrios* and *Fbranch* programs to identify introgression between all combinations of species, as well as potential direction of gene flow, specifying *Rhinella veredas* as an outgroup and applying the Benjamini-Hochberg (BH) correction to control for the false discovery rate.

Results

Phylogenetic relationships

The 16S phylogeny suggested little phylogenetic structure within the *Rhinella marina* complex. One clade included most of the *R. horribilis* samples, while individuals from the remaining taxa formed a polytomy (Fig. S1). Maximum likelihood and Bayesian phylogenies based on the ddRADseq dataset resulted in fully concordant phylogenies (Fig. 1). These analyses inferred six highly supported clades, two corresponding to *R. marina* and the other four corresponding to *R. schneideri*, *R. horribilis*, *R. jimi*, and a clade tentatively assigned to *R. cf. poeppigii* (BS = 100; PP = 1.0; Fig. 1-3). These putative *R. poeppigii* samples were originally identified as *R. marina*, which would render *R. marina* to be paraphyletic; however, after re-examining these specimens morphologically, we were able to positively identify samples from western Amazonia in Brazil's state of Acre as *R. poeppigii*, while closely related samples from eastern localities in the state of Pará were morphologically more similar to *R. marina* (Fig. S7). Pairwise Nei's G_{ST} estimates for the 16S data were much lower than the Weir and Cockerham weighted F_{ST} estimates for the ddRADseq data. Across all taxa, the average pairwise G_{ST} for the mitochondrial data was 0.117 (0.025-0.228) while the average pairwise F_{ST} for the nuclear data was 0.506 (0.379-0.843) (Table S2).

The time-calibrated phylogeny based on the 16S mitochondrial data dated the root of *Rhinella marina* at 8.96 mybp (95% HPD: 6.342-11.477 mybp; Fig 4). Though many relationships have poor support due to lack of variability within the locus, some clades showed

high support, including a clade with most of the *R. horribilis* samples, which is dated at 4.28 mybp (95% HPD: 1.821-7.158 mybp). Two samples not included in this clade are samples distributed in the northern Andes, which cluster with other *R. marina* samples (Fig 4). *Rhinella cf. poeppigii* samples from eastern Amazonia form a highly supported clade with a divergence date of 1.59 mybp (95% HPD: .338-3.409 mybp), while the western Amazonia *R. poeppigii* sample clusters with other *R. marina* in southern Amazonia (Fig. 4). Additionally, *R. granulosa* is estimated to be sister to the *R. marina* complex, with *R. margaritifera* more distantly related. Due to the lack of variation within the *R. marina* group, we interpret dates within this complex with caution.

Population structure

Despite the high posterior probabilities of each clade in our ddRAD tree, the ADMIXTURE results supported genetic admixture both within and across multiple taxa within the *Rhinella marina* complex (Fig. 1), with a best-fit K of 7. Each clade corresponded to a cluster, except for the *R. schneideri* clade which consisted of two clusters. *Rhinella horribilis* (blue, Fig. 1-2) showed admixture from the northern cluster of *R. marina* into one northern Andes locality. One cluster of *R. marina* was relegated to northern Amazonia (light green, Fig. 1-2), while the other cluster showed a cline of admixture across its western and southern Amazonia clades (light green to purple, Fig. 1-2) and admixture from *R. jimi* and *R. cf. poeppigii* (dark green and orange, Fig. 1). The two genetic clusters within *R. schneideri* (pink and yellow, Fig. 1,3) followed an east-west admixture gradient across the Cerrado to the northern Atlantic Forest, as well as intermediate ecotones. *Rhinella jimi* occurs mostly in the semi-arid Caatinga shrublands of

northeastern Brazil, but also in the adjacent coastal rainforest (dark green, Fig. 1-2). The DAPC analysis supported this clustering scheme as well; however, BIC scores suggested similar support for six to eight clusters (Fig. S4). The seven clusters recovered were concordant with phylogenetic structure (Fig. 1).

Demographic inference

For the subset composed of *R. marina, R. horribilis*, and *R. jimi*, the best 3D-JFSF model was one that incorporated ancient migration with a short period of recent isolation since divergence, with a log-likelihood of -1572.69 and AIC of 3165.38 (Fig. 5, Table S3). This model included an ancient period of migration between all lineages (mA, Fig. 5), then another period of migration between geographically adjacent species after the divergence between *R. marina* and *R. jimi*, and then subsequent lineage isolation. Parameter estimates indicated a much longer ancient period of migration between all lineages with smaller migration rates (T1 = 10.82; mA = 0.05) compared to the shorter time of adjacent-species migration with higher rates of migration (T2 = 0.12; m1 = 1.36; m2 = 0.85) and the shortest period of isolation (T3 = 0.10) (Table S3).

For the subset composed of *R. schneideri* and *R. cf. poeppigii*, the best 2D-JFSF model incorporated divergence in isolation followed by secondary contact with asymmetric gene flow, with a log-likelihood of -539.27 and AIC of 1090.54 (Fig. 5, Table S3). Parameter estimates inferred a period of divergence in isolation (T1 = 0.07) with a shorter period of secondary contact (T2 = 0.01) and a much higher rate of migration from *R. cf. poeppigii* into *R. schneideri* (m12 = 15.5) than from *R. schneideri* into *R. cf. poeppigii* (m21 = 1.82) (Table S3).

D-statistics

Nearly all topological trios tested (((P1,P2)P3)O) had significant *D*-statistics (Table S4). The *R. jimi-marina-horribilis* trio was not significant ($p > 0.05$), indicating that we cannot reject the null hypothesis of no gene flow for that trio, which assumes that any ABBA-BABA patterns arose solely due to incomplete lineage sorting (Malinsky et al., 2020). *D*-statistics for all significant trios ranged from 0.12 to 0.49 (Table S4). The highest *D*-statistics were for *R. horribilis-jimi-schneideri* (0.49), *R. marina-jimi-schneideri* (0.37), and *R. schneideri-poeppigii-marina* (0.30). The f_b(C) statistic is a summary of f_4 admixture ratios and shows excess allele sharing between the branch on the y-axis and the sample on the x-axis (Malinsky et al. 2018). The f_b(C) statistics indicated the highest percentages of gene flow between *R. cf. poeppigii* and *R. marina* (11%), between *R. cf. poeppigii* and *R. horribilis* (8%), and between *R. jimi* and *R. schneideri* (7%) (Fig. 6, Table S5).

Discussion

Based on comprehensive geographic and genomic sampling within a clade of South American toads, this investigation found evidence of multiple distinct evolutionary lineages that span large geographic areas and, at times, distinct biomes. The inferred major clades and genetic clusters largely correspond to currently recognized taxa within *Rhinella*; however, we also found evidence of potentially cryptic diversity within *R. marina*, *R. schneideri*, and potentially *R. poeppigii*. Genetic clustering analyses suggested that many of the inferred groups include admixed individuals. Accordingly, demographic analyses supported that the evolutionary history

of these toads involved cross-taxon gene flow both at ancient (in the case of *R. marina, R. horribilis*, and *R. jimi*) and recent (in the case of *R. schneideri* and *R. cf. poeppigii*) times. Both demographic inference and ABBA-BABA tests inferred patterns of genetic introgression across species, supporting previous assertions that the evolutionary history of *Rhinella* was characterized by various levels of hybridization (Pereyra et al. 2016; Sequeira et al., 2011).

Phylogenetic patterns and species boundaries

The phylogenetic findings of this study improve our knowledge about species diversity and distributions in South America. Our sampling validates previous reports of *Rhinella poeppigii* present in western Amazonia (Venâncio et al., 2017). *Rhinella poeppigii* has a history of both taxonomic uncertainty and misidentification, due to its similarity to *R. marina* (De la Riva, 2002; Venâncio et al., 2017; Venegas & Ron, 2014). After the first individuals were identified and collected in Ecuador, subsequent specimens collected in the region that were previously misidentified were discovered at Museo de Zoología, Pontificia Universidad Católica del Ecuador (QCAZ) (Venegas & Ron, 2014). In this study we included another individual from Porto Walter, Acre, Brazil, which further corroborates *R. poeppigii* extending into Brazil. Furthermore, we uncovered a group of *R. cf. poeppigii* specimens in eastern Amazonia near the Belo Monte Hydroelectric dam on the Xingu River (Fig. 3). These samples, however, do not display distinct *R. poeppigii* morphology, and in fact are more similar morphologically to *R. marina*, to which they were originally assigned (Fig. S7). Unfortunately, as sampling of this clade was initially unintentional, we did not sample specimens from across the range of *R. poeppigii*, which may be misrepresenting the genetic admixture visualized within this clade (Fig. 1,3). Given this restricted sampling and the more than 2,000 km distance in sampled individuals,

it may be that eastern *R. cf. poeppigii* is actually a yet undescribed cryptic species within the *Rhinella marina* complex.

The mitochondrial 16S rRNA marker has been used extensively for identification and barcoding of amphibians (Maya-Soriano et al., 2012; Rockney et al., 2015; Vences et al., 2005). Despite this marker being extremely useful in taxonomic identification for a number of closely related species (Firneno & Townsend, 2019), even within the *Rhinella* genus (Pereyra et al. 2016), there is an inherent lack of diversity recovered across all focal species within the *R. marina* complex (Fig. S1-S2). It is possible that purifying selection has acted on this region of the mitochondrial genome, thereby greatly reducing genetic diversity across the complex (Charlesworth et al., 1995; Cvijović et al., 2018). Considering that processes like purifying selection can also reduce genetic diversity at linked neutral sites, previous estimates of potential introgression within *Rhinella* species using mitochondrial data may be similarly affected (Cvijović et al., 2018). This phenomenon could have resulted in an overestimation of shared loci by any other means, such as hybridization, as opposed to a constraint on particular loci. With the 16S fragment sequenced being relatively short (~480 bp), an analysis of the entire 16S rRNA gene or even the whole-mitochondrial genome in this group could prove useful in disentangling the reasons for such low genetic diversity seen here.

By contrast, despite evidence of admixture both within and between species, nuclear data estimated a phylogeny with substantial structure and support (Fig. 1). When compared to other phylogenies generated with single or multi-locus datasets, high-throughput sequencing of the *Rhinella marina* complex has revealed a surprising amount of genetic complexity, introgression, and interspecific resolution (Bessa-Silva et al., 2020; Maciel et al., 2010; Vallinoto et al., 2009). These patterns suggest that in groups with such complex demographic histories, and especially

those with a likelihood of hybridization between divergent populations or species, large-scale genetic data can be very useful in disentangling relationships and histories.

Biogeographic drivers of species range limits

Inferred species range limits can be attributed to both present-day spatial environmental gradients and the history of topographic change in South America, as suggested for a number of other South American taxa (Carnaval et al., 2009; Fonseca et al., 2018; Gehara et al., 2014; Prates, Rivera et al., 2016). Mitochondrial divergence time analyses are consistent with the idea that the Andean uplift contributed to divergence between *R. marina* and *R. horribilis* (Fig. 4); pronounced genetic divergence between populations on each side of the Andean chain supports the recent recognition of *R. horribilis* as a taxon distinct from *R. marina* (Vallinoto et al., 2009). While the Andes likely limits contemporary gene flow between these two taxa, our finding of admixture between them suggests that the northern Andes may be a semi-permeable barrier (Fig. 2), in agreement with patterns seen in other organisms (Acevedo et al., 2016; Bessa-Silva et al., 2020; Maciel et al., 2010). Additionally, like other amphibians (Noonan & Wray, 2006) and reptiles (Gamble et al., 2008), the extensive fluvial network formed in western Amazonia by periodic Miocene flooding, known as the Pebas formation, may have contributed to divergence not only between *R. horribilis* and *R. marina*, but also between the northeast and south-southwestern Amazonian clades within *R. marina* (Vallinoto et al., 2009; Wesselingh & Salo, 2006). *Rhinella marina*, which is comprised of two well-supported clades, is distributed across Amazonian climates, which are known to have asynchronous historical eastern-western climatic cycles and have had an effect on species composition and genetic diversity within the biome (Cheng et al., 2013; Prates, Rivera et al., 2016). Considering that the distinct clades have a

northern-southern distribution, as opposed to an eastern-western distribution, however, it may be more plausible that geographic barriers, such as fluctuating fluvial networks from the Miocene through the Pleistocene (Cooke et al., 2012; Lundberg et al., 1998), have had a higher impact in promoting divergence between these clades within *R. marina*.

Similar to what is observed within *R. marina*, we see patterns of species distributed across environmental gradients repeated across the phylogeny; *R. schneideri* is distributed across the Cerrado, through Cerrado-Caatinga-Atlantic Forest ecotones, and into the northern Atlantic Forest, with an east-west gradient of admixture (Fig. 1,3). The Seasonally Dry Tropical Forests and savannas of South America have been known to harbor complex and cryptic genetic diversity and have been especially affected by Quaternary climate fluctuations (Bandeira et al., 2021; Fonseca et al., 2018; Gehara et al., 2017; Werneck et al., 2015). Considering the phylogenetic pattern that we see within *R. schneideri*, we can posit that this species expanded eastward during Plio-Pleistocene climate change (Bandeira et al., 2021; Lisiecki & Raymo, 2007). Paleoclimatic modeling of the biogeographic history and niche of *R. schneideri* on a finer scale is recommended to validate this hypothesis.

A puzzling biogeographic pattern that emerged from our results is the extremely disjunct distribution between *R. poeppigii* in western Amazonia and its sister clade, *R. cf. poeppigii*, from eastern Brazilian Amazonia, more than 2,000 km apart. This mysterious pattern has also been reported for other herpetofaunal species, including the lizards *Anolis* trachyderma (Ribeiro-Júnior, 2015) and *Potamites ecpleopus* (Ribeiro-Júnior & Amaral, 2017) and the horned treefrog *Hemiphractus scutatus* (de Lima Moraes & Pavan, 2018). Despite this large geographic distance, as well as the effects of contrasting climatic seasonality between the eastern and western localities in this region on other herpetofauna (Cheng et al., 2013; Prates, Rivera et al., 2016;

Wang et al., 2017), this and other studies indicate limited genetic divergence across disjunct regions (de Lima Moraes & Pavan, 2018). A comprehensive analysis of museum specimens and available tissues from these areas, in conjunction with a more thorough sampling of *R. poeppigii* across its range, will be required to confirm this unexpected pattern of genetic divergence within this group.

Hybridization and introgression

The interspecific relationships inferred with historical demographic modeling suggest extremely varied patterns of migration and hybridization through time within the *Rhinella marina* complex. Our study indicates that species within this group have diverged across multiple biomes and amassed significant genetic differentiation despite continuous gene flow among species (Fig. 5). Many of the species within the *R. marina* complex also have a shared introgressive history (Fig. 1,6; Table S4-S5). Hypothesis testing of demographic models suggests that the *R. marina-horribilis-jimi* clade continued to exchange genes throughout its dispersion across the continent, and species within this clade exchanged genes with other species within the complex (Fig. 1,5-6, Table S5). Despite evidence of gene flow between species, there was no evidence of population-wide hybridization or the presence of hybrid species within our sampling. Potential proposed hybridization events have been reported within or between *Rhinella* species groups, such as within the *R. granulosa* complex (Guerra et al., 2011; Pereyra et al. 2016) and the *R. crucifer* complex (Júnior et al., 2004; Thomé et al., 2012), where either instances of morphologically intermediate individuals or hybrid populations have been reported. Much of the speculation surrounding hybridization in Neotropical toads has been accompanied by a lack of data from natural populations to assess the biological reality of presumed hybrid species

(Fontenot et al., 2011; Malone & Fontenot, 2008; Thomé et al., 2012). Within the *R. marina* group, however, we found that recurrent gene flow between species at low levels is much more prevalent than the persistence of supposed hybrid "swarms".

Acknowledgements

We thank all the Brazilian collectors who made this study possible and particularly students in MTR's laboratory for field collection and support. Brazil's Instituto Chico Mendes de Conservação da Biodiversidade issued collecting permits (SISBIO 36753-1, 36753-4, and 27290-3). This work was co-funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; BIOTA 2013/50297-0), the National Science Foundation (DEB 1343578), and the National Aeronautics and Space Administration through the Dimensions of Biodiversity Program. MTR acknowledges additional funding from FAPESP grants 2003/10335-8, 2011/50146-6, 2011/50206-9, 2012/15754-8, and 2017/08357-6. DR was funded by NSF GRFP. IP was funded by NSF grant DEB 1754398.

References

Acevedo, A. A., Lampo, M., & Cipriani, R. (2016). The cane or marine toad, *Rhinella marina* (Anura, Bufonidae): two genetically and morphologically distinct species. *Zootaxa*, *4103*(6), 574–586.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664.

Avise, J. C., Walker, D., & Johns, G. C. (1998). Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings. Biological Sciences / The Royal Society*, *265*(1407), 1707–1712.

Azevedo, M. F. C., Foresti, F., Ramos, P. R. R., & Jim, J. (2003). Comparative cytogenetic studies of *Bufo ictericus*, *B. paracnemis* (Amphibia, Anura) and an intermediate form in sympatry. *Genetics and Molecular Biology*, *26*(3), 289–294.

Bandeira, L. N., Villalobos, F., Werneck, F. P., Peterson, A. T., & Anciães, M. (2021). Different elevational environments dictate contrasting patterns of niche evolution in Neotropical *Pithecopus* treefrog species. *Biotropica*, (btp.12929). doi: 10.1111/btp.12929

Barratt, C. D., Bwong, B. A., Jehle, R., Liedtke, H. C., Nagel, P., Onstein, R. E., … Loader, S. P. (2018). Vanishing refuge? Testing the forest refuge hypothesis in coastal East Africa using genome-wide sequence data for seven amphibians. *Molecular Ecology*, *27*(21), 4289–4308.

Bernardo, P. H., Sánchez-Ramírez, S., Sánchez-Pacheco, S. J., Álvarez-Castañeda, S. T., Aguilera-Miller, E. F., Mendez-de la Cruz, F. R., & Murphy, R. W. (2019). Extreme mito-nuclear discordance in a peninsular lizard: the role of drift, selection, and climate. *Heredity*, *123*(3), 359–370.

Bessa-Silva, A., Vallinoto, M., Sampaio, I., Flores-Villela, O. A., Smith, E. N., & Sequeira, F. (2020). The roles of vicariance and dispersal in the differentiation of two species of the *Rhinella marina* species complex. *Molecular Phylogenetics and Evolution*, *145*, 106723.

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., … Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *15*(4), e1006650.

Carnaval, A. C., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T., & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, *323*(5915), 785–789.

Carstens, B. C., & Knowles, L. L. (2007). Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. *Systematic Biology*, *56*(3), 400–411.

Charles, K. L., Bell, R. C., Blackburn, D. C., Burger, M., Fujita, M. K., Gvoždík, V., … Portik, D. M. (2018). Sky, sea, and forest islands: Diversification in the African leaf-folding frog *Afrixalus paradorsalis* (Anura: Hyperoliidae) of the Lower Guineo-Congolian rain forest. *Journal of Biogeography*, *45*(8), 1781–1794.

Charlesworth, D., Charlesworth, B., & Morgan, M. T. (1995). The pattern of neutral molecular variation under the background selection model. *Genetics*, *141*(4), 1619–1632.

Cheng, H., Sinha, A., Cruz, F. W., Wang, X., Edwards, R. L., d'Horta, F. M., … Auler, A. S. (2013). Climate change patterns in Amazonia and biodiversity. *Nature Communications*, *4*, 1411.

Chernomor, O., von Haeseler, A., & Minh, B. Q. (2016). Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices. *Systematic Biology*, *65*(6), 997–1008.

Cooke, G. M., Chao, N. L., & Beheregaray, L. B. (2012). Marine incursions, cryptic species and ecological diversification in Amazonia: the biogeographic history of the croaker genus *Plagioscion* (Sciaenidae). *Journal of Biogeography*, *39*(4), 724–738.

Correa, C. L., Méndez, M. A., Veloso, A., & Sallaberry, M. (2012). Genetic and Reproductive Evidence of Natural Hybridization between the Sister Species *Rhinella atacamensis* and

Rhinella arunco (Anura, Bufonidae). *Journal of Herpetology*, Vol. 46, pp. 568–577. doi: 10.1670/10-266

Cvijović, I., Good, B. H., & Desai, M. M. (2018). The Effect of Strong Purifying Selection on Genetic Diversity. *Genetics*, *209*(4), 1235–1278.

Dal Vechio, F., Prates, I., Grazziotin, F. G., Zaher, H., Grabosky, R., & Rodrigues, M. T. (2019). Rain forest shifts through time and riverine barriers shaped the diversification of South American terrestrial pit vipers (*Bothrops jararacussu* species group). *Journal of Biogeography*, *47*(2), 516–526.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics, 27*(15), 2156–2158.

De la Riva, I. (2002). Taxonomy and distribution of the South American toad *Bufo poeppigii* Tschudi, 1845 (Amphibia, Anura, Bufonidae). *Graellsia / Editada Por El Instituto Espanol de Entomologia Del Consejo Superior de Investigaciones Cientificas*, *58*(1), 49–57.

de Lima Moraes, L. J. C., & Pavan, D. (2018). On the occurrence of *Hemiphractus scutatus* (Spix, 1824) (Anura: Hemiphractidae) in eastern Amazonia. *Amphibian & Reptile Conservation*, *12*(1), 5–14 (e151).

de Medeiros, B. A. S., & Farrell, B. D. (2018). Whole-genome amplification in double-digest RADseq results in adequate libraries but fewer sequenced loci. *PeerJ*, *6*, e5089.

Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution*, *28*(8), 2239–2252.

Eaton, D. A. R., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics, 36*(8), 2592–2594.

Edwards, R. J., Tuipulotu, D. E., Amos, T. G., O'Meally, D., Richardson, M. F., Russell, T. L., … White, P. A. (2018). Draft genome assembly of the invasive cane toad, *Rhinella marina*. *GigaScience*, Vol. 7. doi: 10.1093/gigascience/giy095

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics: TIG*, *28*(7), 342–350.

Firneno, T. J., Jr, O'Neill, J. R., Portik, D. M., Emery, A. H., Townsend, J. H., & Fujita, M. K. (2020). Finding complexity in complexes: Assessing the causes of mitonuclear discordance in a problematic species complex of Mesoamerican toads. *Molecular Ecology*, *29*(18), 3543–3559.

Firneno, T. J., Jr., & Townsend, J. H. (2019). Evaluation of species boundaries in sympatric and parapatric populations of Mesoamerican toads. *Zoologica Scripta*, *48*(4), 454–465.

Fonseca, E. M., Gehara, M., Werneck, F. P., Lanna, F. M., Colli, G. R., Sites, J. W., Jr, … Garda, A. A. (2018). Diversification with gene flow and niche divergence in a lizard species along the South American "diagonal of open formations." *Journal of Biogeography*, *45*(7), 1688–1700.

Fontenot, B. E., Makowsky, R., & Chippindale, P. T. (2011). Nuclear–mitochondrial discordance and gene flow in a recent radiation of toads. *Molecular Phylogenetics and Evolution*, *59*(1), 66–80.

Gamble, T., Simons, A. M., Colli, G. R., & Vitt, L. J. (2008). Tertiary climate change and the diversification of the Amazonian gecko genus *Gonatodes* (Sphaerodactylidae, Squamata). *Molecular Phylogenetics and Evolution*, *46*(1), 269–277.

Gehara, M., Crawford, A. J., Orrico, V. G. D., Rodríguez, A., Lötters, S., Fouquet, A., … Köhler, J. (2014). High levels of diversity uncovered in a widespread nominal taxon:

continental phylogeography of the neotropical tree frog *Dendropsophus minutus*. *PloS One*, *9*(9), e103958.

Gehara, M., Garda, A. A., Werneck, F. P., Oliveira, E. F., da Fonseca, E. M., Camurugi, F., … Burbrink, F. T. (2017). Estimating synchronous demographic changes across populations using hABC and its application for a herpetological community from northeastern Brazil. *Molecular Ecology*, *26*(18), 4756–4771.

Graham, A. M., Lavretsky, P., Muñoz-Fuentes, V., Green, A. J., Wilson, R. E., & McCracken, K. G. (2018). Migration-Selection Balance Drives Genetic Differentiation in Genes Associated with High-Altitude Function in the Speckled Teal (*Anas flavirostris*) in the Andes. *Genome Biology and Evolution*, *10*(1), 14–32.

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., … Pääbo, S. (2010). A draft sequence of the Neandertal genome. *Science*, *328*(5979), 710–722.

Guerra, C., Baldo, D., Rosset, S., Borteiro, C., & Kolenc, F. (2011). Advertisement and release calls in Neotropical toads of the *Rhinella granulosa* group and evidence of natural hybridization between *R. bergi* and *R. major* (Anura: Bufonidae). *Zootaxa*, *3092*, 26–42.

Guo, B., Lu, D., Liao, W. B., & Merilä, J. (2016). Genomewide scan for adaptive differentiation along altitudinal gradient in the Andrew's toad *Bufo andrewsi*. *Molecular Ecology*, *25*(16), 3884–3900.

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, *5*(10), e1000695.

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.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics, 27*(21), 3070–3071.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 94.

Júnior, F. A. B., Caramaschi, U., & Haddad, C. F. B. (2004). Review of the *Bufo crucifer* species group, with descriptions of two new related species (Amphibia, Anura, Bufonidae). *Arquivos Do Museu Nacional. Museu Nacional*, *62*(3), 255–282.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589.

Lanfear, R., Calcott, B., Ho, S. Y. W., & 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., Calcott, B., Kainer, D., Mayer, C., & Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, *14*, 82.

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017). PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. *Molecular Biology and Evolution*, *34*(3), 772–773.

Lavretsky, P., Peters, J. L., Winker, K., Bahn, V., Kulikova, I., Zhuravlev, Y. N., … McCracken, K. G. (2016). Becoming pure: identifying generational classes of admixed individuals within lesser and greater scaup populations. *Molecular Ecology*, *25*(3), 661–674.

Leaché, A. D., Portik, D. M., Rivera, D., Rödel, M., Penner, J., Gvoždík, V., … Fujita, M. K. (2019). Exploring rain forest diversification using demographic model testing in the African foam-nest treefrog *Chiromantis rufescens*. *Journal of Biogeography*, *46*(12), 2706–2721.

Lisiecki, L. E., & Raymo, M. E. (2007). Plio–Pleistocene climate evolution: trends and transitions in glacial cycle dynamics. *Quaternary Science Reviews*, *26*(1), 56–69.

Liu, K., Wang, F., Chen, W., Tu, L., Min, M.-S., Bi, K., & Fu, J. (2010). Rampant historical mitochondrial genome introgression between two species of green pond frogs, *Pelophylax nigromaculatus* and *P. plancyi*. *BMC Evolutionary Biology*, *10*, 201.

Lundberg, J. G., Marshall, L. G., Guerrero, J., Horton, B., Malabarba, M., & Wesselingh, F. (1998). The stage for Neotropical fish diversification: a history of tropical South American rivers. *Phylogeny and Classification of Neotropical Fishes*, *27*, 13–48.

Maciel, N. M., Collevatti, R. G., Colli, G. R., & Schwartz, E. F. (2010). Late Miocene diversification and phylogenetic relationships of the huge toads in the *Rhinella marina* (Linnaeus, 1758) species group (Anura: Bufonidae). *Molecular Phylogenetics and Evolution*, *57*(2), 787–797.

Malinsky, M., Matschiner, M., & Svardal, H. (2020). Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*. doi: 10.1111/1755-0998.13265

Malinsky, M., Svardal, H., Tyers, A. M., Miska, E. A., Genner, M. J., Turner, G. F., & Durbin, R. (2018). Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nature Ecology & Evolution*, *2*(12), 1940–1955.

Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, *20*(5), 229–237.

Malone, J. H., & Fontenot, B. E. (2008). Patterns of reproductive isolation in toads. *PloS One*, *3*(12), e3900.

Marth, G. T., Czabarka, E., Murvai, J., & Sherry, S. T. (2004). The allele frequency spectrum in genome-wide human variation data reveals signals of differential demographic history in three large world populations. *Genetics*, *166*(1), 351–372.

Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., … Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research*, *23*(11), 1817–1828.

Maya-Soriano, M. J., Holt, W. V., & Lloyd, R. E. (2012). Biobanked Amphibian Samples Confirmed To Species Level Using 16S rRNA DNA Barcodes. *Biopreservation and Biobanking*, *10*(1), 22–28.

Mayr, E. (1963). *Animal Species and Evolution*. Harvard University Press.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gateway Computing Environments Workshop (GCE)*. doi: 10.1109/gce.2010.5676129

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274.

Noonan, B. P., & Wray, K. P. (2006). Neotropical diversification: the effects of a complex history on diversity within the poison frog genus *Dendrobates*. *Journal of Biogeography*, *33*(6), 1007–1020.

O'Connell, K. A., Prates, I., Scheinberg, L. A., Mulder, K. P., & Bell, R. C. (2021). Speciation and secondary contact in a fossorial island endemic, the São Tomé caecilian. *Molecular Ecology*, *30*(12), 2859–2871.

Oliveira, E. F., Gehara, M., São-Pedro, V. A., Costa, G. C., Burbrink, F. T., Colli, G. R., … Garda, A. A. (2018). Phylogeography of Muller's termite frog suggests the vicariant role of the Central Brazilian Plateau. *Journal of Biogeography*, *45*(11), 2508–2519.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., … Reich, D. (2012). Ancient admixture in human history. *Genetics*, *192*(3), 1065–1093.

Pereyra, M. O., Baldo, D., Blotto, B. L., Iglesias, P. P., Thome, M. T. C., Haddad, C. F. B., … Faivovich, J. (2016). Phylogenetic relationships of toads of the *Rhinella granulosa* group (Anura: Bufonidae): a molecular perspective with comments on hybridization and introgression. *Cladistics*, *32*, 36–53.

Pereyra, M. O., Blotto, B. L., Baldo, D., Chaparro, J. C., Ron, S. R., Elias-Costa, A. J., … Faivovich, J. (2021). Evolution in the Genus *Rhinella*: A Total Evidence Phylogenetic Analysis of Neotropical True Toads (Anura: Bufonidae). *Bulletin of the American Museum of Natural History*, Vol. 447. doi: 10.1206/0003-0090.447.1.1

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, *7*(5).

Phuong, M. A., Bi, K., & Moritz, C. (2017). Range instability leads to cytonuclear discordance in a morphologically cryptic ground squirrel species complex. *Molecular Ecology*, *26*(18), 4743–4755.

Pirani, R. M., Peloso, P. L. V., Prado, J. R., Polo, É. M., Knowles, L. L., Ron, S. R., … Werneck, F. P. (2020). Diversification history of clown tree frogs in Neotropical rainforests (Anura, Hylidae, *Dendropsophus leucophyllatus* group). *Molecular Phylogenetics and Evolution*, *150*, 106877.

Portik, D. (2016). *Phylo_Wrapper_Scripts*. Retrieved from https://github.com/dportik/Phylo_Wrapper_Scripts

Portik, D. M., Leaché, A. D., Rivera, D., Barej, M. F., Burger, M., Hirschfeld, M., … Fujita, M. K. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Molecular Ecology*, *26*(19), 5245–5263.

Potter, S., Bragg, J. G., Peter, B. M., Bi, K., & Moritz, C. (2016). Phylogenomics at the tips: inferring lineages and their demographic history in a tropical lizard, *Carlia amax*. *Molecular Ecology*, *25*(6), 1367–1380.

Prates, I., Rivera, D., Rodrigues, M. T., & Carnaval, A. C. (2016). A mid-Pleistocene rainforest corridor enabled synchronous invasions of the Atlantic Forest by Amazonian anole lizards. *Molecular Ecology*, *25*(20), 5174–5186.

Prates, I., Xue, A. T., Brown, J. L., Alvarado-Serrano, D. F., Rodrigues, M. T., Hickerson, M. J., & Carnaval, A. C. (2016). Inferring responses to climate dynamics from historical demography in neotropical forest lizards. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(29), 7978–7985.

Rabosky, D. L. (2016). Reproductive isolation and the causes of speciation rate variation in nature. *Biological Journal of the Linnean Society. Linnean Society of London*, *118*(1), 13–25.

Ribeiro-Júnior, M. A. (2015). Catalogue of distribution of lizards (Reptilia: Squamata) from the Brazilian Amazonia. I. Dactyloidae, Hoplocercidae, Iguanidae, Leiosauridae, Polychrotidae, Tropiduridae. *Zootaxa*, *3983*(1), 1–110.

Ribeiro-Júnior, M. A., & Amaral, S. (2017). Catalogue of distribution of lizards (Reptilia: Squamata) from the Brazilian Amazonia. IV. Alopoglossidae, Gymnophthalmidae. *Zootaxa*, *4269*(2), 151–196.

Rockney, H. J., Ofori-Boateng, C., Porcino, N., & Leaché, A. D. (2015). A comparison of DNA barcoding markers in West African frogs. *African Journal of Herpetology: Journal of the Herpetological Association of Africa*, *64*(2), 135–147.

Sambrook, J., & Russell, D. W. (2006). Purification of nucleic acids by extraction with phenol:chloroform. *CSH Protocols*, *2006*(1). doi: 10.1101/pdb.prot4455

Sequeira, F., Sodré, D., Ferrand, N., Bernardi, J. A. R., Sampaio, I., Schneider, H., & Vallinoto, M. (2011). Hybridization and massive mtDNA unidirectional introgression between the closely related Neotropical toads *Rhinella marina* and *R. schneideri* inferred from mtDNA and nuclear markers. *Biomedical Chromatography: BMC*, *11*(264).

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics, 30*(9), 1312–1313.

Thomé, M. T. C., Zamudio, K. R., Giovanelli, J. G. R., Haddad, C. F. B., Baldissera, F. A., Jr, & Alexandrino, J. (2010). Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. *Molecular Phylogenetics and Evolution*, *55*(3), 1018–1031.

Thomé, M. T. C., Zamudio, K. R., Haddad, C. F. B., & Alexandrino, J. (2012). Delimiting genetic units in Neotropical toads under incomplete lineage sorting and hybridization. *BMC Evolutionary Biology*, *12*, 242.

Vallinoto, M., Sequeira, F., Sodre, D., Bernardi, J. A. R., Sampaio, I., & Schneider, H. (2009). Phylogeny and biogeography of the *Rhinella marina* species complex (Amphibia, Bufonidae) revisited: implications for Neotropical diversification hypotheses. *Zoologica Scripta*, *39*(2), 128–140.

Venâncio, N. M., Antonio de Freitas, M., Abegg, A. D., & Kokubum, M. N. de C. (2017). First record of *Rhinella poeppigii* (Tschudi, 1845) in Brazil (Anura, Bufonidae). *Check List*, *13*(6), 747–750.

Vences, M., Thomas, M., van der Meijden, A., Chiari, Y., & Vieites, D. R. (2005). Comparative performance of the 16S rRNA gene in DNA barcoding of amphibians. *Frontiers in Zoology*, *2*(1), 5.

Venegas, P. J., & Ron, S. R. (2014). First records of *Rhinella poeppigii* (Tschudi 1845) from Ecuador, with a distribution map (Anura: Bufonidae). *Herpetology Notes*, *7*, 713–716.

Wang, X., Lawrence Edwards, R., Auler, A. S., Cheng, H., Kong, X., Wang, Y., … Chiang, H.-W. (2017). Hydroclimate changes across the Amazon lowlands over the past 45,000 years. *Nature*, *541*(7636), 204–207.

Werneck, F. P. (2011). The diversification of eastern South American open vegetation biomes: Historical biogeography and perspectives. *Quaternary Science Reviews*, *30*(13), 1630–1648.

Werneck, F. P., Leite, R. N., Geurgas, S. R., & Rodrigues, M. T. (2015). Biogeographic history and cryptic diversity of saxicolous Tropiduridae lizards endemic to the semiarid Caatinga. *BMC Evolutionary Biology*, *15*, 94.

Wesselingh, F. P., & Salo, J. A. (2006). A Miocene perspective on the evolution of the Amazonian biota. *Scripta Geologica*, *133*, 439–458.

Winter, D. J. (2012). MMOD: an R library for the calculation of population differentiation statistics. *Molecular Ecology Resources*, *12*(6), 1158–1160.

Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under migration-selection balance: The genetic architecture of local adaptation. *Evolution; International Journal of Organic Evolution*, *65*(7), 1897–1911.

Figures & Tables

TABLE 1 Optimal demographic models and estimated parameters. Abbr: LL = log-likelihood; θ ($4N_{ref}\mu L$) = the effective mutation rate of the reference population (ancestral population); nu1, nu2 = effective population sizes under the constant population size model; nuA = effective population sizes of the ancestral population; mA, m1, m2 = migration rates between the ancestral (A), first (1) or second (2) population; m12 = migration rate from population two to population one; m21 = migration rate from population one to population two; T1, T2, T3 = unscaled time between demographic events.

Model													
2D	LL	AIC	θ	nu1	nu2		m12	m21			T1	T2	
divergence in isolation with continuous asymmetric secondary contact	-539.27	1090.54	1424.5	0.01	0.05		15.5	1.82			0.07	0.01	
3D	LL	AIC	θ	nu1	nu2	nu3	nuA	mA	m1	m2	T1	T2	T3
ancient migration with shortest isolation	-1572.69	3165.38	95.33	0.57	0.15	1.31	6.15	0.05	1.36	0.85	10.82	0.12	0.10

FIGURE 1 (A) Maximum likelihood phylogeny of *Rhinella marina* complex focal species using ddRADseq data and corresponding ADMIXTURE plot (K=7). Black circles on the phylogeny denote ML bootstrap support (BS) >95 and bayesian posterior probability (PP) >95. (B) DAPC plot (K=7).

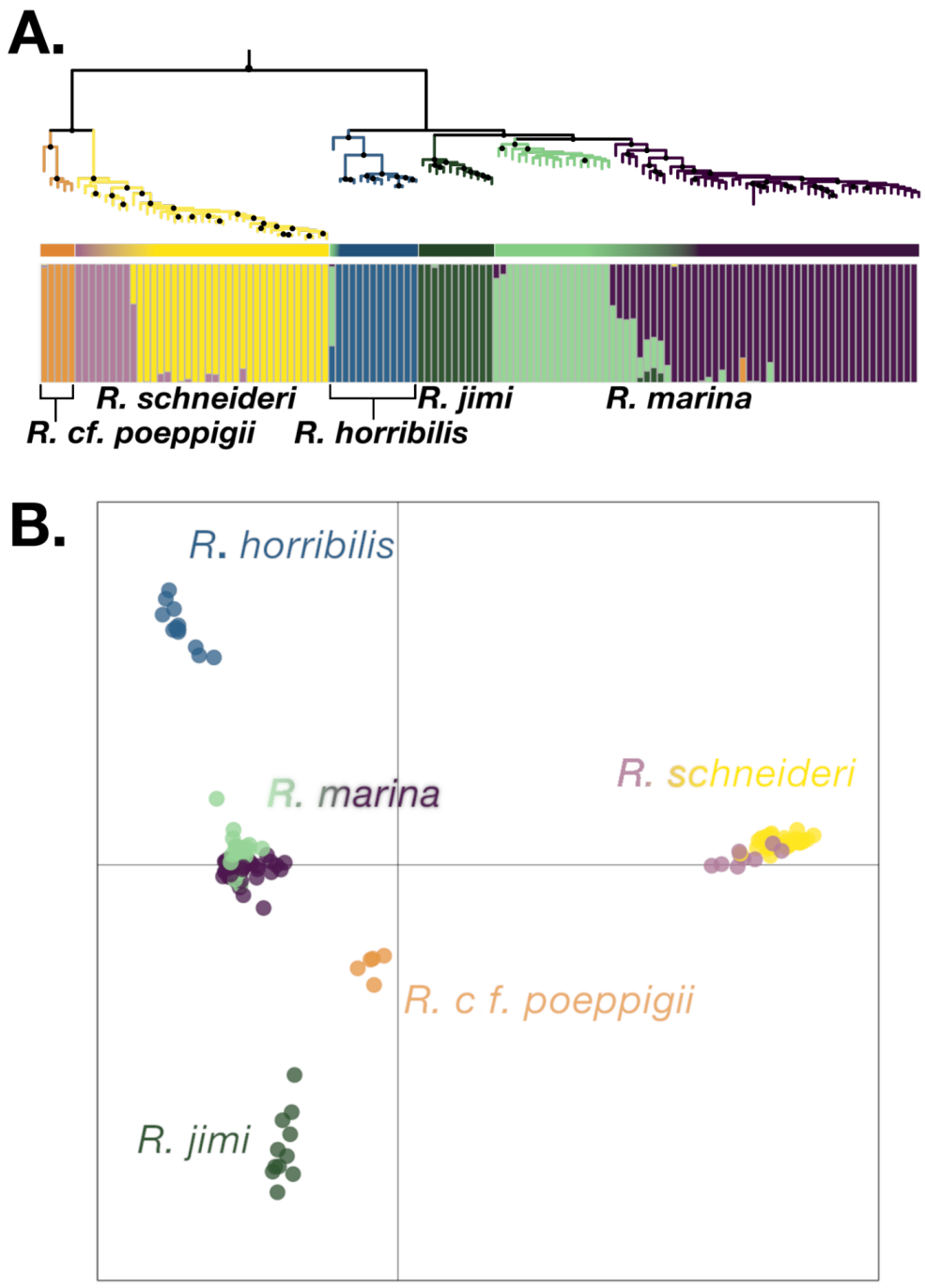


FIGURE 2 Locality map for the subset depicting average ADMIXTURE cluster assignments per locality (K=7) for *Rhinella horribilis*, *R. marina*, and *R. jimi*. Colors correspond to Figure 1. Map partitioned into biomes (Central America, Northern Andes, Northern Amazonia, Western Amazonia, Eastern Amazonia, Southern Amazonia, Pantanal, Chaco, Cerrado, Caatinga, Northern Atlantic Forest, Southern Atlantic Forest).

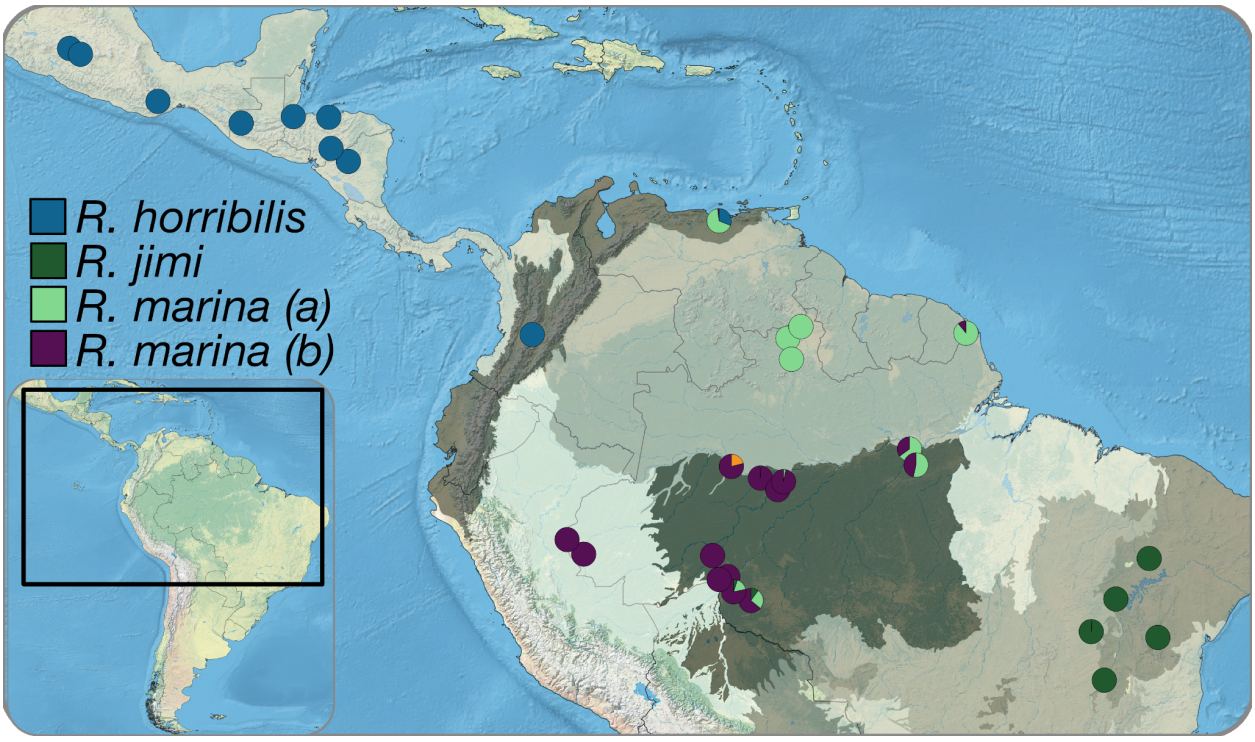


FIGURE 3 Locality map for the subset depicting average ADMIXTURE cluster assignments per locality (K=7) for *Rhinella poeppigii* and *R. schneideri*. Colors correspond to Figure 1. Map partitioned into biomes (Central America, Northern Andes, Northern Amazonia, Western Amazonia, Eastern Amazonia, Southern Amazonia, Pantanal, Chaco, Cerrado, Caatinga, Northern Atlantic Forest, Southern Atlantic Forest). *R. poeppigii* range adapted from (De la Riva, 2002; Venâncio et al., 2017; Venegas & Ron, 2014).

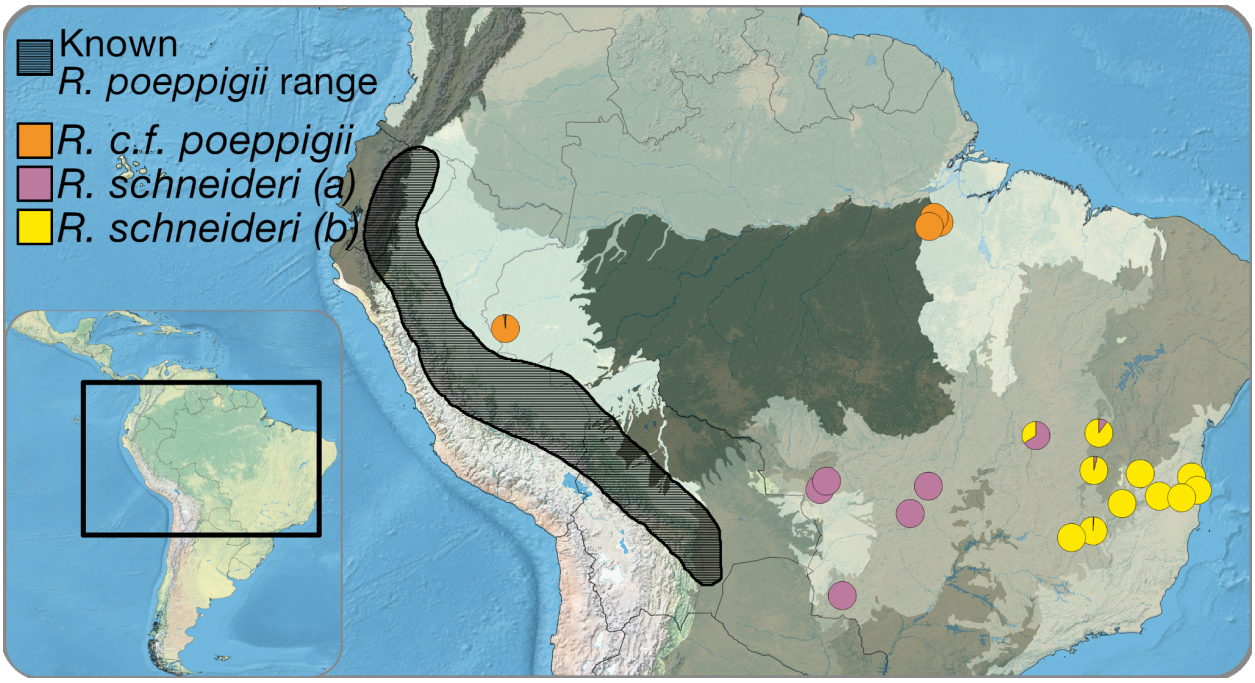


FIGURE 4 Time calibrated phylogeny based on mitochondrial 16S data. Black circles indicate PP > 0.90. Colors correspond to phylogeny in Figure 1.

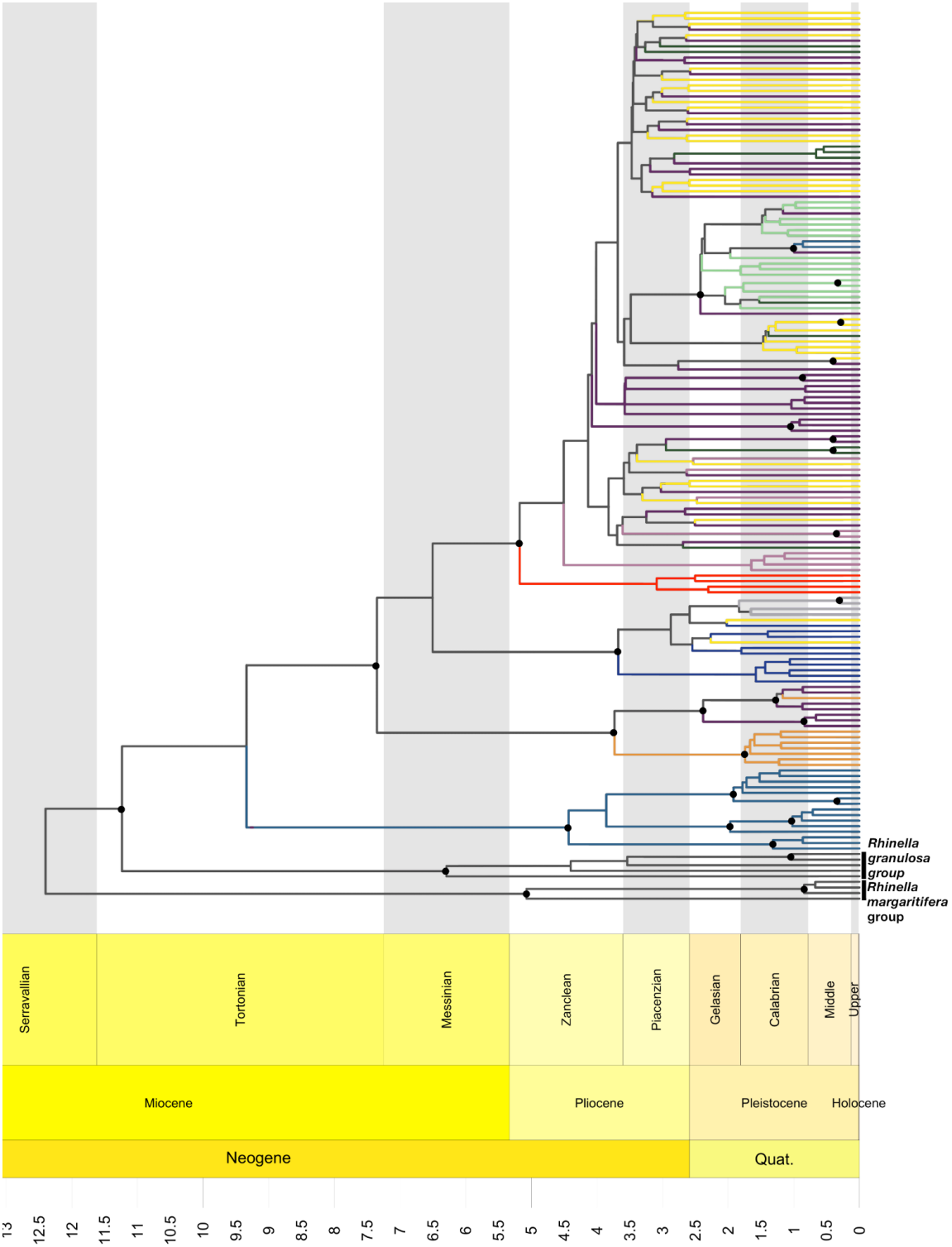


FIGURE 5 Optimal demographic models and residual plots for the (A) 3D-JSFS analysis of *Rhinella horribilis*, *R. marina*, and *R. jimi*, and (B) 2D-JSFS analysis of *R. poeppigii* and *R. schneideri*.

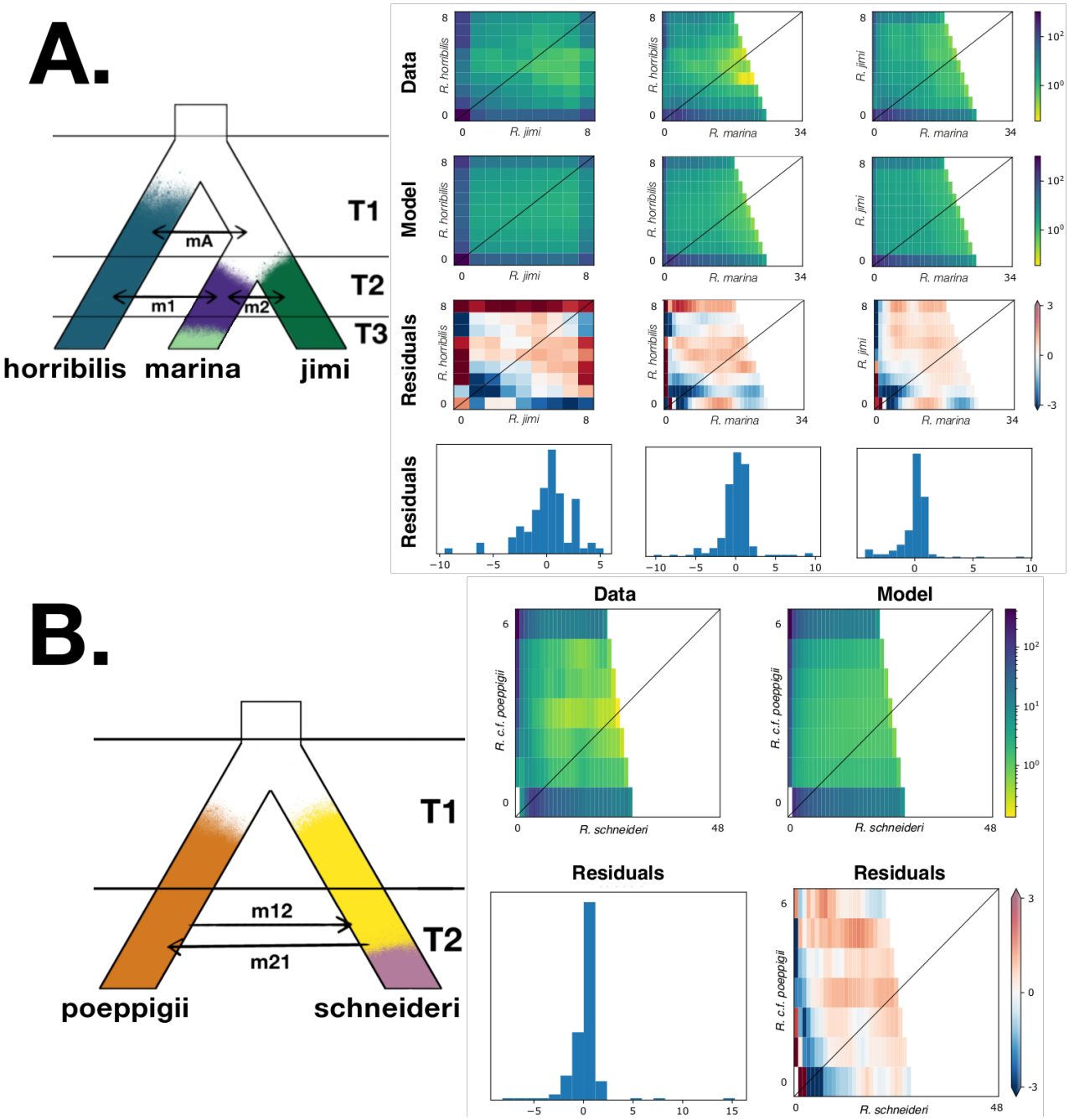
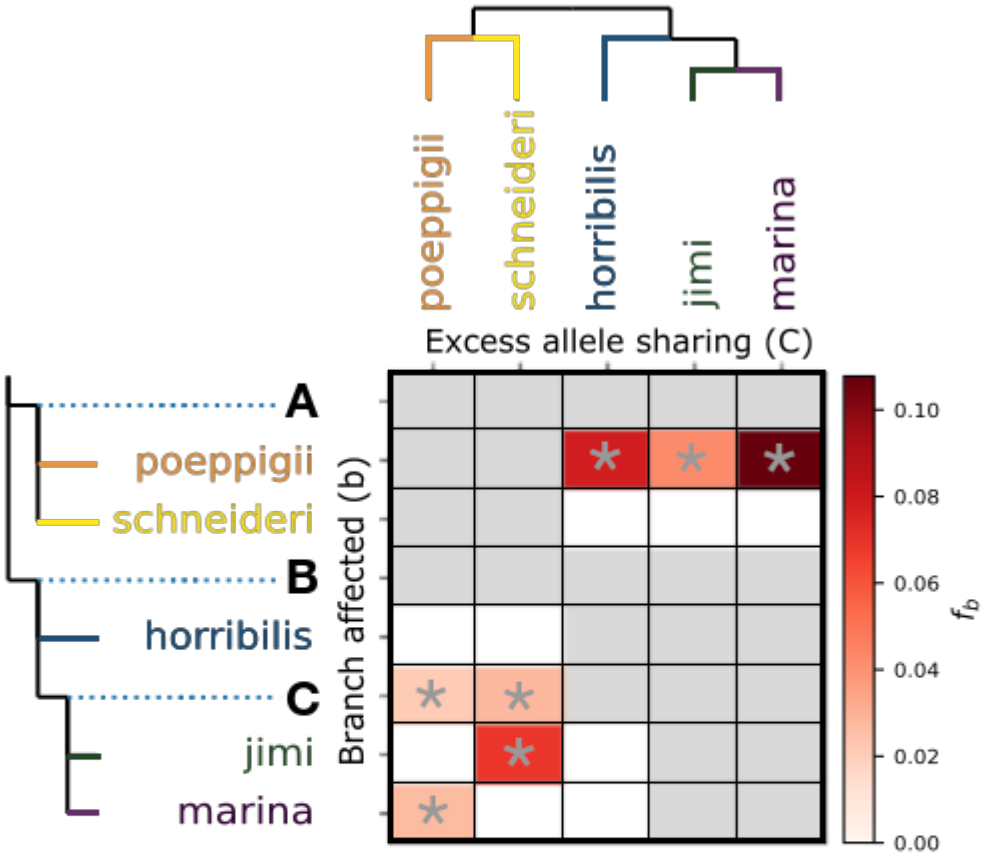


FIGURE 6 The f_b statistic (summary of f_4 admixture ratios). Grey color corresponds to tests that are not possible because of constraints on the phylogeny. * indicates a significant result.



Supplementary Figures

FIGURE S1 ML phylogeny based on 16S rRNA data. Black circles indicate nodal bootstrap support (BS) with BS > 80.

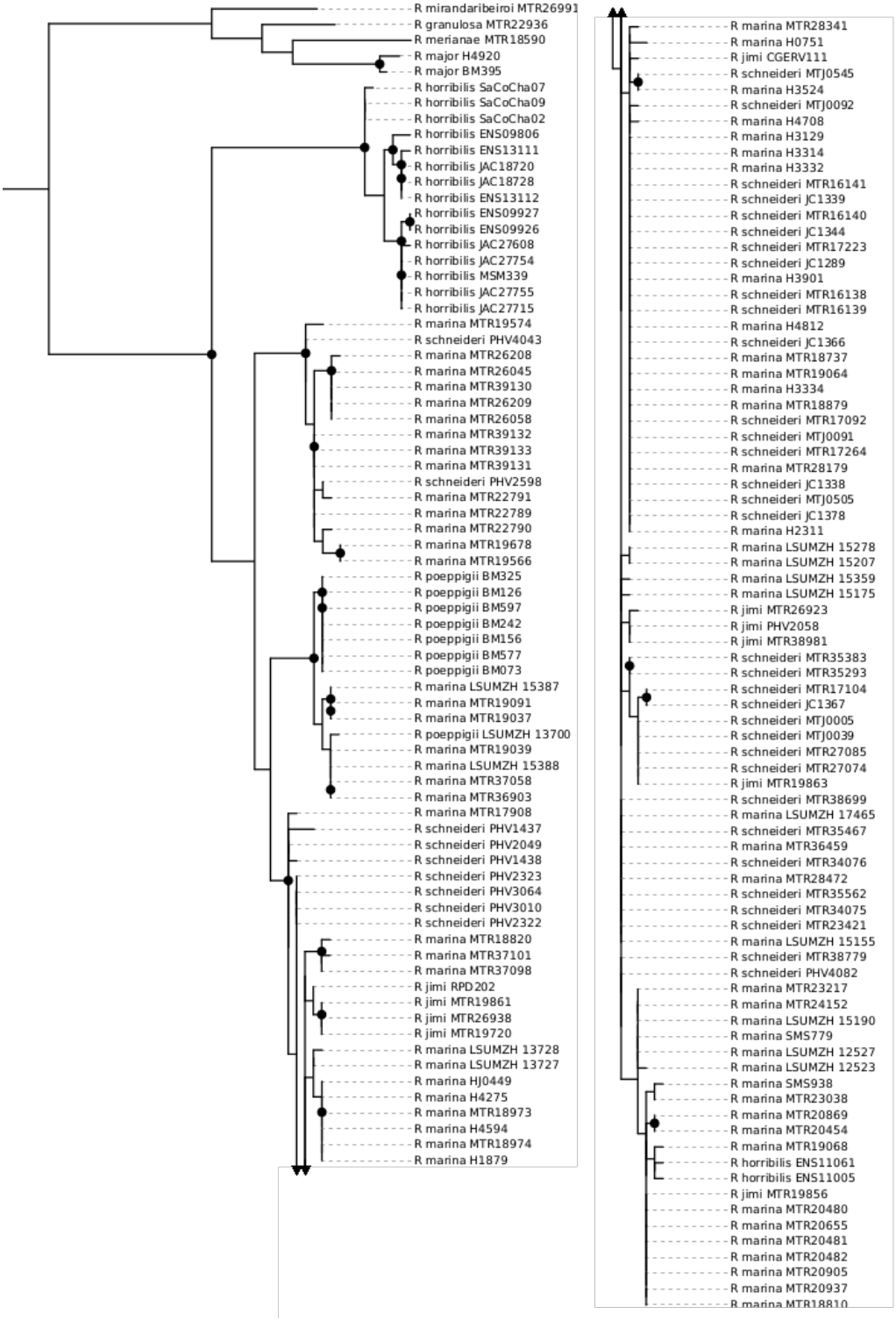


FIGURE S2 Time-calibrated phylogeny based on 16S rRNA data visualized as a cladogram.

Black circles indicate nodal bootstrap support (BS) with BS > 80. Node values = 95% HPD.

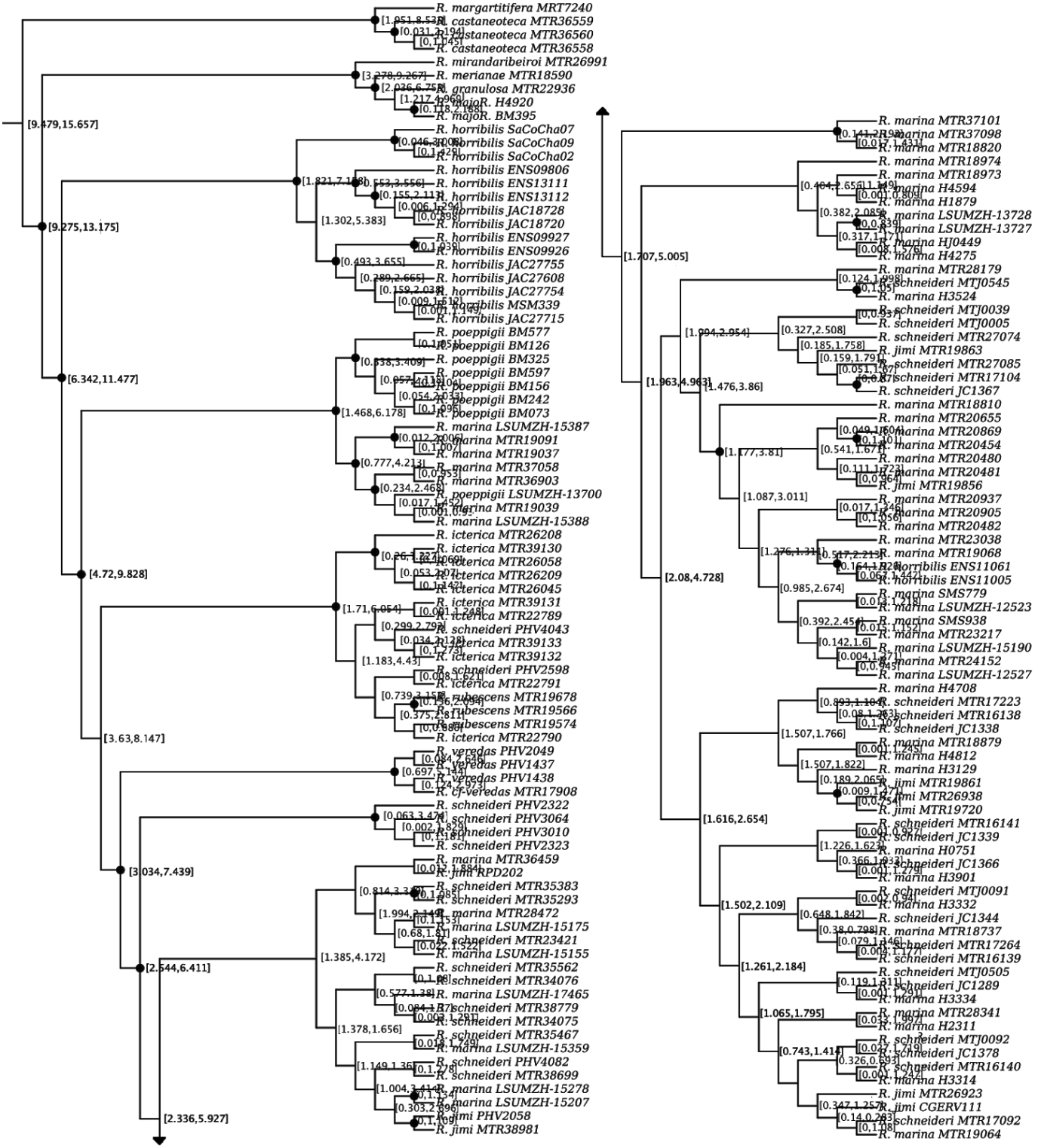


FIGURE S3 ADMIXTURE plots for K = 4-9.

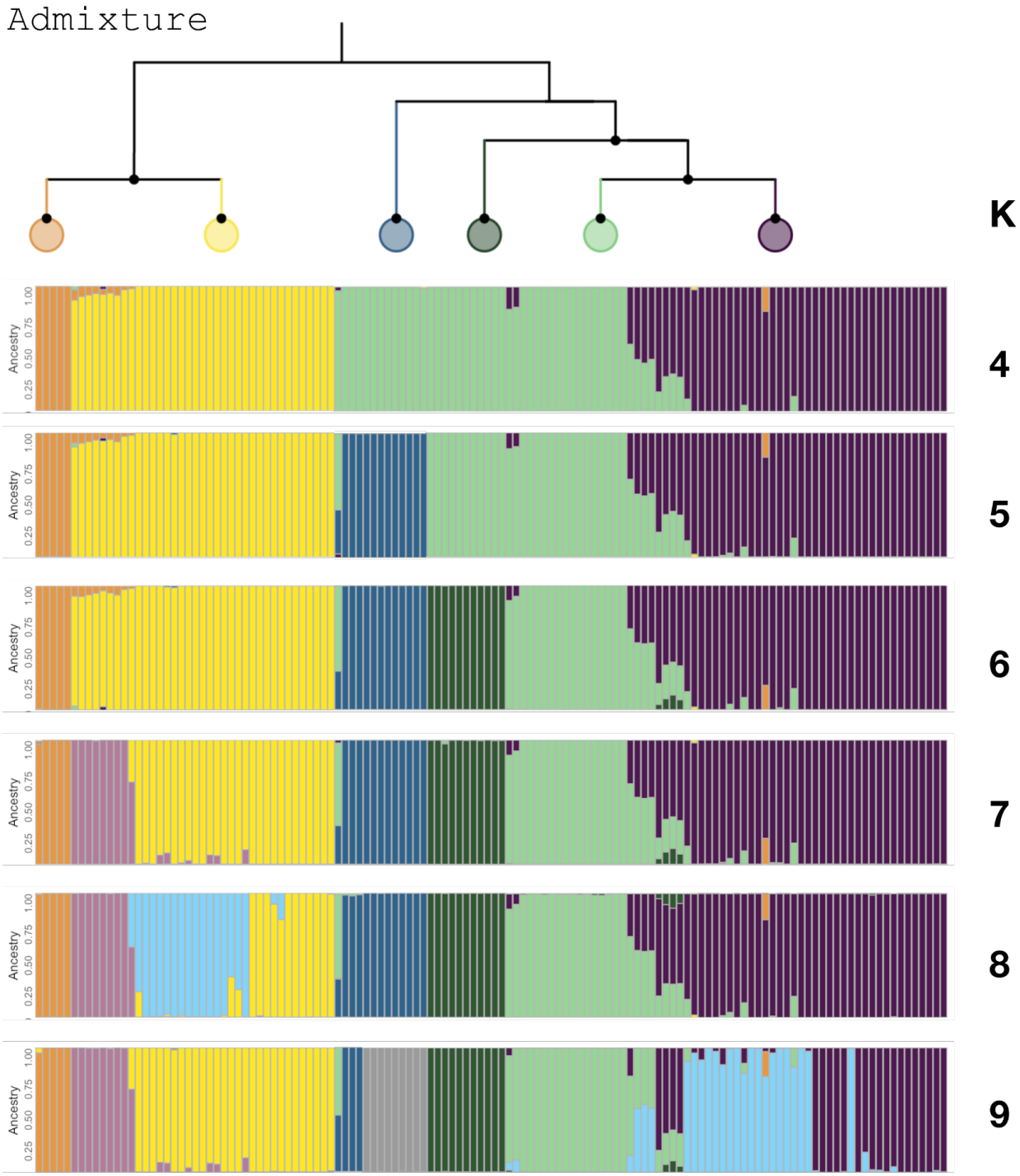


FIGURE S4 BIC values per cluster and associated plots for clusters 6-8. Blue arrow indicates cluster number (7) with lowest cross-validation error.

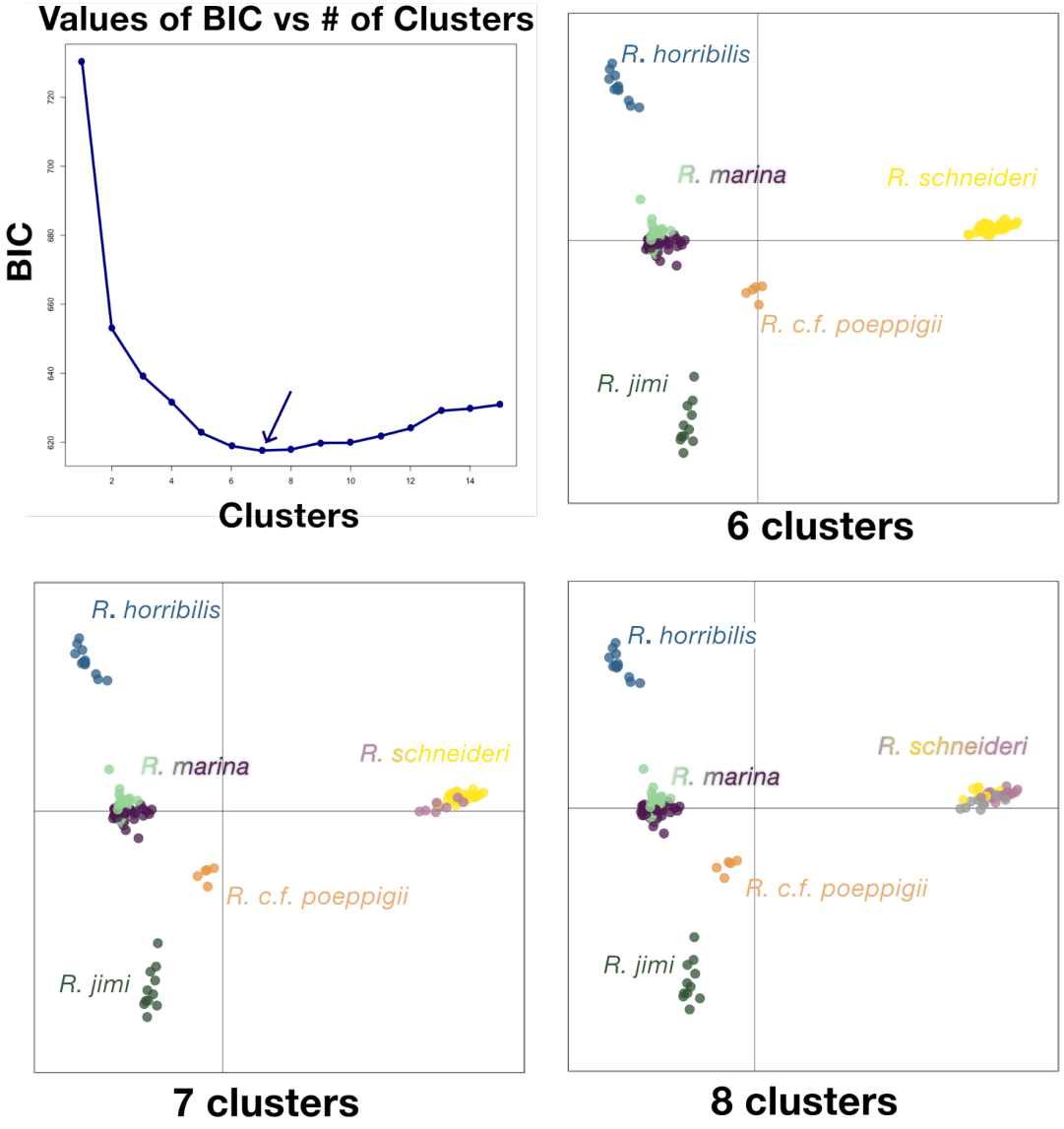


FIGURE S5 3D-JSFS Demographic Models tested in ∂a∂i. A. divergence with no migration, B. divergence with symmetric migration between all populations, C. divergence with symmetric migration between adjacent populations, D. isolation with secondary contact, E. simultaneous split with secondary contact between adjacent populations, F. simultaneous split with migration between adjacent populations, G. adjacent ancient migration with the shortest period of recent isolation, H. ancient migration with a short isolation and population size change across two epochs, I. brief ancient migration with longest period of isolation, J. ancient migration with a longer period of isolation. Arrows represent migration events.

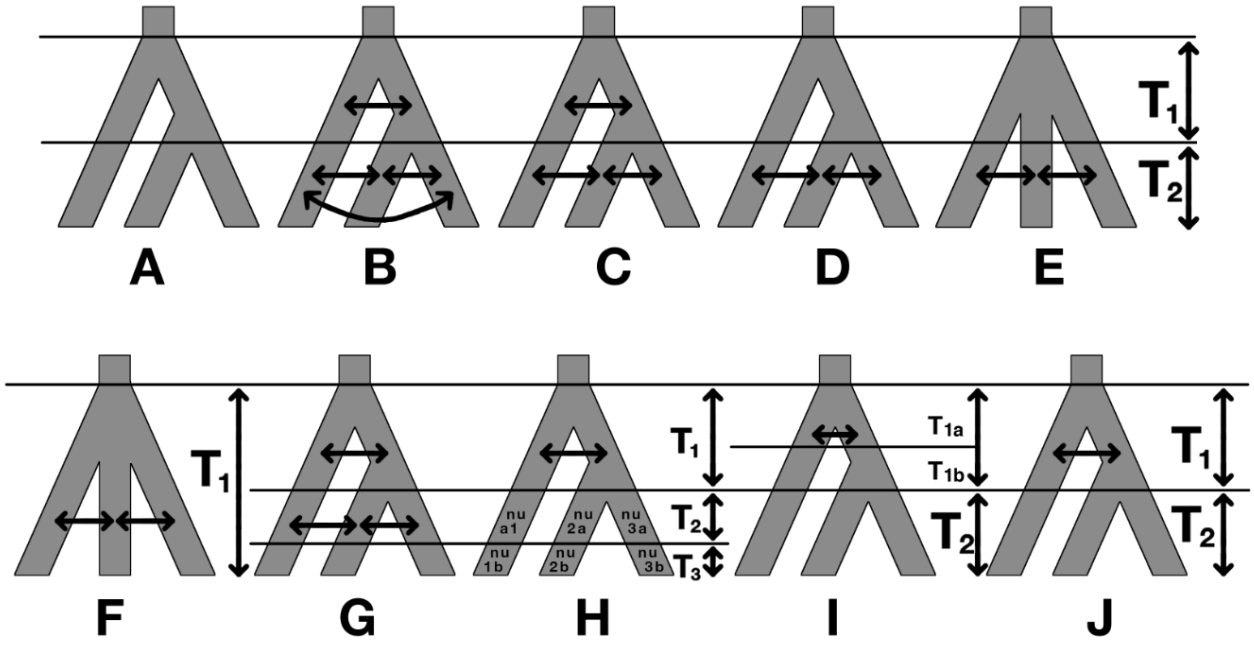


FIGURE S6 2D-JSFS Demographic Models tested in ∂a∂i. A. divergence with no migration, B. divergence with continuous symmetric migration, C. divergence with continuous asymmetric migration, D. divergence with continuous symmetric migration and a varying rate of migration across two epochs, E. divergence with continuous asymmetric migration and a varying rate of migration across two epochs, F. divergence in isolation, followed by symmetric secondary contact, G. divergence in isolation, followed by asymmetric secondary contact, H. ancient symmetric migration then subsequent isolation, I. ancient asymmetric migration then subsequent isolation, J. divergence in isolation followed by symmetric secondary contact with subsequent isolation, K. divergence in isolation followed by asymmetric secondary contact with subsequent isolation. Arrows correspond to migration events.

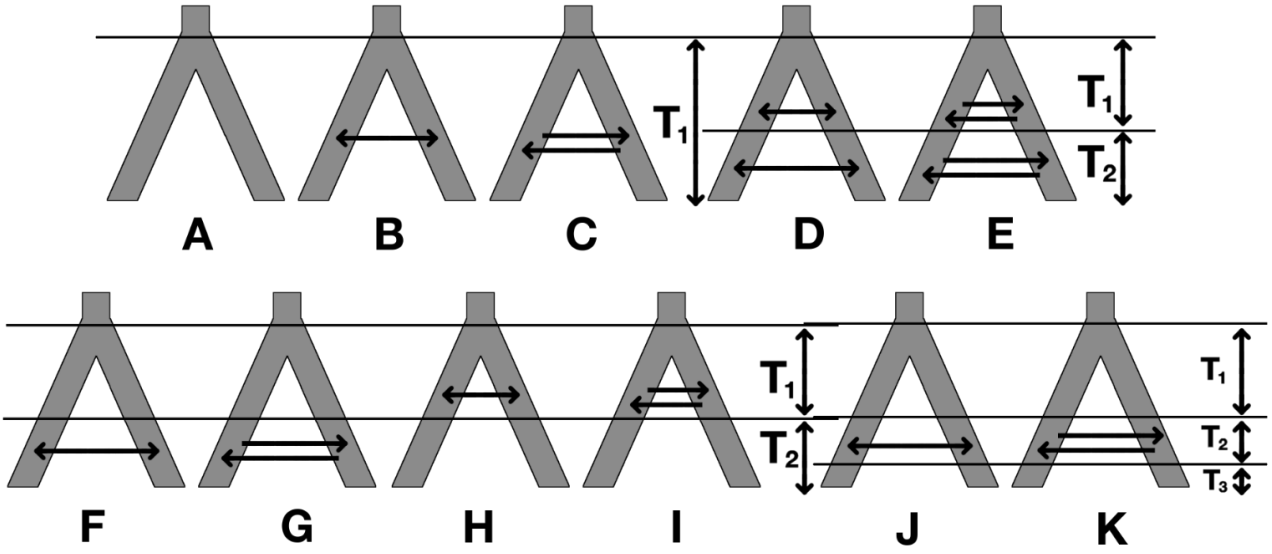


FIGURE S7 *Rhinella cf. poeppigii* dorsal and ventral images for specimens from the eastern locality of Pará, Brazil.

Supplementary Tables

TABLE S1 All sample identification information for individuals included in this study.

Species	Field Number	Locality	State/Country
R. horribilis	ENS08661	El Paraíso	EP - HN
R. horribilis	ENS08678	Atlántida	AT - HN
R. horribilis	ENS09806	Matagalpa	MT - NI
R. horribilis	ENS09926	Oaxaca	OA - MX
R. horribilis	ENS09927	Oaxaca	OA - MX
R. horribilis	ENS11005	El Limón	AR - VZ

R. horribilis	ENS11061	Clarinas	AN - VZ
R. horribilis	ENS13111	Carretera Remate-Gringo Perdido	PE - GT
R. horribilis	ENS13112	Carretera Remate-Gringo Perdido	PE - GT
R. horribilis	JAC18720	Zacapa	ZA - GT
R. horribilis	JAC18795	Izabal	IZ - GT
R. horribilis	JAC27218	Zacazonapan	EM - MX
R. horribilis	JAC27608	Huetamo - El Limon de Papatzingan	MI - MX
R. horribilis	JAC27715	Chilpancingo	GR - MX
R. horribilis	JAC27754	Ixtapa - Altamirano	GR - MX
R. horribilis	JAC27755	Ixtapa - Altamirano	GR - MX
R. horribilis	MSM339	Huehuetenango	HU - GT
R. horribilis	SaCoCha02	Tolima, Chaparral	CO - BR
R. horribilis	SaCoCha03	Tolima, Chaparral	CO - BR
R. horribilis	SaCoCha07	Tolima, Chaparral	CO - BR
R. horribilis	SaCoCha09	Tolima, Chaparral	CO - BR
R. horribilis	SaCoCha10	Tolima, Chaparral	CO - BR
R. jimi	CGERV111	Capitão Gervásio de Oliveira	PI - BR
R. jimi	MTR19720	Andaraí	BA - BR
R. jimi	MTR19856	Andaraí	BA - BR
R. jimi	MTR19861	Andaraí	BA - BR
R. jimi	MTR19862	Andaraí	BA - BR
R. jimi	MTR19863	Andaraí	BA - BR
R. jimi	MTR26923	São Desidério	BA - BR
R. jimi	MTR26938	São Desidério	BA - BR
R. jimi	MTR38981	Buritirama	BA - BR
R. jimi	PHV2058	São Desidério	BA - BR
R. jimi	RPD202	P.E. Sete Passagens, Miguel Calmon	BA - BR
R. marina	H0751	Caiçara	RO - BR
R. marina	H1879	Caiçara	RO - BR
R. marina	H2311	Abunã	RO - BR
R. marina	H3129	UHE Jirau, Mutum	RO - BR
R. marina	H3314	UHE Jirau	RO - BR
R. marina	H3332	UHE Jirau	RO - BR
R. marina	H3334	UHE Jirau	RO - BR
R. marina	H3524	UHE Jirau	RO - BR
R. marina	H3901	UHE Jirau, Mutum	RO - BR

R. marina	H4275	UHE Jirau, Mutum	RO - BR
R. marina	H4594	UHE Jirau, Mutum	RO - BR
R. marina	H4708	UHE Jirau, Caiçara	RO - BR
R. marina	H4812	UHE Jirau, Caiçara	RO - BR
R. marina	HJ0449	Abunã, Porto Velho	RO - BR
R. marina	LSUMZH-13727	Porto Walter	AC - BR
R. marina	LSUMZH-13728	Porto Walter	AC - BR
R. marina	LSUMZH-15155	Alter do Chao	PA - BR
R. marina	LSUMZH-15175	Santarem	PA - BR
R. marina	LSUMZH-15190	Santarem	PA - BR
R. marina	LSUMZH-15207	Santarem	PA - BR
R. marina	LSUMZH-15278	Santarem	PA - BR
R. marina	LSUMZH-15359	Rio Ituxi at the Madeireira Scheffer	AM - BR
R. marina	LSUMZH-15381	Rio Ituxi at the Madeireira Scheffer	AM - BR
R. marina	LSUMZH-15387	Rio Ituxi at the Madeireira Scheffer	AM - BR
R. marina	LSUMZH-15388	Rio Ituxi at the Madeireira Scheffer	AM - BR
R. marina	LSUMZH-15416	Rio Ituxi at the Madeireira Scheffer	AM - BR
R. marina	LSUMZH-17465	Parque Estadual Guajara-Mirim	RO - BR
R. marina	MTR18737	Lago Chaviana, Itapuru	AM - BR
R. marina	MTR18810	Lago Chaviana, Itapuru	AM - BR
R. marina	MTR18820	Lago Chaviana, Itapuru	AM - BR
R. marina	MTR18879	Lago Chaviana, Itapuru	AM - BR
R. marina	MTR18973	Lago Chaviana, Itapuru	AM - BR
R. marina	MTR18974	Lago Chaviana, Itapuru	AM - BR
R. marina	MTR19037	Moiobamba, Rio Purus	AM - BR
R. marina	MTR19039	Moiobamba, Rio Purus	AM - BR
R. marina	MTR19064	Moiobamba, Rio Purus	AM - BR
R. marina	MTR19068	Moiobamba, Rio Purus	AM - BR
R. marina	MTR19091	Moiobamba, Rio Purus	AM - BR
R. marina	MTR20454	E.E. Maracá	RR - BR
R. marina	MTR20480	E.E. Maracá	RR - BR
R. marina	MTR20481	E.E. Maracá	RR - BR
R. marina	MTR20482	E.E. Maracá	RR - BR
R. marina	MTR20655	Pacaraima	RR - BR
R. marina	MTR20869	Tepequém	RR - BR
R. marina	MTR20894	Tepequém	RR - BR

R. marina	MTR20905	Tepequém	RR - BR
R. marina	MTR20937	Tepequém	RR - BR
R. marina	MTR23038	Serra da Maroquinha	RR - BR
R. marina	MTR23039	Serra da Maroquinha	RR - BR
R. marina	MTR23058	Serra da Maroquinha	RR - BR
R. marina	MTR23217	Serra da Maroquinha	RR - BR
R. marina	MTR24152	Oiapoque	AP - BR
R. marina	MTR24183	Oiapoque	AP - BR
R. marina	MTR25670	Parque Nacional de Pacaás Novos	RO - BR
R. marina	MTR25697	Parque Nacional de Pacaás Novos	RO - BR
R. marina	MTR25710	Parque Nacional de Pacaás Novos	RO - BR
R. marina	MTR36459	São Pedro, Rio Içá	AM - BR
R. marina	MTR36903	Tefé	AM - BR
R. marina	MTR37058	Coari	AM - BR
R. marina	MTR37098	Coari	AM - BR
R. marina	MTR37101	Coari	AM - BR
R. marina	MTR37122	Pacairama	RR - BR
R. marina	MTR37150	ESEC Rio Acre, Assis Brasil	AC - BR
R. marina	MTR37196	Cantá	RR - BR
R. marina	MTR37197	Cantá	RR - BR
R. marina	SMS779	Serra do Apiaú	RR - BR
R. marina	SMS938	Serra do Apiaú	RR - BR
R. cf. poeppigii	BM073	UHE Belo Monte	PA - BR
R. cf. poeppigii	BM126	UHE Belo Monte	PA - BR
R. cf. poeppigii	BM156	UHE Belo Monte	PA - BR
R. cf. poeppigii	BM242	Vitória do Xingu	PA - BR
R. cf. poeppigii	BM325	Vitória do Xingu	PA - BR
R. cf. poeppigii	BM577	Altamira	PA - BR
R. cf. poeppigii	BM597	Altamira	PA - BR
R. poeppigii	LSUMZH-13700	Porto Walter	AC - BR
R. poeppigii	MTR37149	ESEC Rio Acre, Assis Brasil	AC - BR
R. schneideri	JC1289	Augusto de Lima, Hotel Amorim	MG - BR
R. schneideri	JC1338	Augusto de Lima, Hotel Amorim	MG - BR
R. schneideri	JC1339	Augusto de Lima, Hotel Amorim	MG - BR
R. schneideri	JC1344	Augusto de Lima, Serra do Cabral	MG - BR
R. schneideri	JC1366	Cristália	MG - BR

R. schneideri	JC1367	Cristália	MG - BR
R. schneideri	JC1378	Cristália	MG - BR
R. schneideri	MTJ0005	Parque Nacional Cavernas do Peruaçu	MG - BR
R. schneideri	MTJ0039	Parque Nacional Cavernas do Peruaçu	MG - BR
R. schneideri	MTJ0091	Parque Nacional Cavernas do Peruaçu	MG - BR
R. schneideri	MTJ0092	Parque Nacional Cavernas do Peruaçu	MG - BR
R. schneideri	MTJ0505	Parque Nacional Cavernas do Peruaçu	MG - BR
R. schneideri	MTJ0545	Parque Nacional Cavernas do Peruaçu	MG - BR
R. schneideri	MTR16138	Serra Bonita, Camacan	BA - BR
R. schneideri	MTR16139	Serra Bonita, Camacan	BA - BR
R. schneideri	MTR16140	Serra Bonita, Camacan	BA - BR
R. schneideri	MTR16141	Serra Bonita, Camacan	BA - BR
R. schneideri	MTR16500	São João do Paraíso, Povoado São Tiago	MG - BR
R. schneideri	MTR17092	Cidade de Jequitinhonha	MG - BR
R. schneideri	MTR17104	Jequitinhonha, Estrada	MG - BR
R. schneideri	MTR17223	Jequitinhonha, estrada	MG - BR
R. schneideri	MTR17264	Reserva Biológica da Mata Escura, Jequitinhonha	MG - BR
R. schneideri	MTR23421	ESEC Pirapitinga	MG - BR
R. schneideri	MTR23422	ESEC Pirapitinga	MG - BR
R. schneideri	MTR27074	Santa Maria da Vitória	BA - BR
R. schneideri	MTR27085	Santa Maria da Vitória	BA - BR
R. schneideri	MTR34075	E.E. Gregório Bondar, Barrolândia	BA - BR
R. schneideri	MTR34076	E.E. Gregório Bondar, Barrolândia	BA - BR
R. schneideri	MTR35293	E.E. Serra das Araras	MT - BR
R. schneideri	MTR35383	E.E. Serra das Araras	MT - BR
R. schneideri	MTR35467	Dolina águas Milagrosas	MT - BR
R. schneideri	MTR35562	Bonito	MS - BR
R. schneideri	MTR38699	Monte Alegre, Fazenda São Caetano	BA - BR
R. schneideri	MTR38779	Monte Alegre	BA - BR
R. schneideri	PHV2322	Barra do Garças	MT - BR
R. schneideri	PHV2323	Barra do Garças	MT - BR
R. schneideri	PHV3010	Alto Araguaia	MT - BR
R. schneideri	PHV3064	Alto Araguaia	MT - BR
R. schneideri	PHV4082	Monte Alegre de Goiás, Serra da Prata	GO - BR
R veredas	PHV1437	São Desidério	BA - BR
R veredas	PHV1438	São Desidério	BA - BR

R veredas	PHV2049	São Desidério	BA - BR
R cf. veredas	MTR17908	São Desidério	BA - BR
R rubescens	MTR19566	PARNA Serra do Cipó	MG - BR
R rubescens	MTR19574	PARNA Serra do Cipó	MG - BR
R rubescens	MTR19678	PARNA Serra do Cipó	MG - BR
R rubescens	MTR23695	Cabeça de Boi, Santana do Rio Preto	MG - BR
R rubescens	MTR23696	Cabeça de Boi, Santana do Rio Preto	MG - BR
R rubescens	MTR23701	Cabeça de Boi, Santana do Rio Preto	MG - BR
R rubescens	PHV2598	Alto Taquari	MT - BR
R rubescens	PHV4043	Niquelândia, Passa Sete	GO - BR
R icterica	MTR22789	Petrópolis	RJ - BR
R icterica	MTR22790	Petrópolis	RJ - BR
R icterica	MTR22791	Petrópolis	RJ - BR
R icterica	MTR26045	Parque Nacional do Itatiaia	RJ - BR
R icterica	MTR26057	Parque Nacional do Itatiaia	RJ - BR
R icterica	MTR26058	Parque Nacional do Itatiaia	RJ - BR
R icterica	MTR26208	Serra da Bocaina	RJ - BR
R icterica	MTR26209	Serra da Bocaina	SP - BR
R icterica	MTR26210	Serra da Bocaina	SP - BR
R icterica	MTR26211	Serra da Bocaina	SP - BR
R icterica	MTR26212	Serra da Bocaina	SP - BR
R icterica	MTR26213	Serra da Bocaina	SP - BR
R icterica	MTR26215	Serra da Bocaina	SP - BR
R icterica	MTR26484	RPPN Araucária, São Joaquim	SC - BR
R icterica	MTR26485	RPPN Araucária, São Joaquim	SC - BR
R icterica	MTR26486	RPPN Araucária, São Joaquim	SC - BR
R icterica	MTR26487	RPPN Araucária, São Joaquim	SC - BR
R icterica	MTR26489	RPPN Araucária, São Joaquim	SC - BR
R icterica	MTR26557	RPPN Araucária, São Joaquim	SC - BR
R icterica	MTR26684	Urubici	SC - BR
R icterica	MTR26752	Urubici	SC - BR
R icterica	MTR39130	Parque Estadual dos Tres Picos	RJ - BR
R icterica	MTR39131	Parque Estadual dos Tres Picos	RJ - BR
R icterica	MTR39132	Parque Estadual dos Tres Picos	RJ - BR
R icterica	MTR39133	Parque Estadual dos Tres Picos	RJ - BR
R castaneoteca	MTR36558	São Pedro, Rio Içá	AM - BR

R castaneoteca	MTR36559	São Pedro, Rio Içá	AM - BR
R castaneoteca	MTR36560	São Pedro, Rio Içá	AM - BR
R margartitifera	MRT7240	Guaraí	TO - BR
R granulosa	MTR22936	FLONA Contendas do Sincorá	BA - BR
R major	BM395	Altamira	PA - BR
R major	H4920	UHE Jirau, Mutum	RO - BR
R merianae	MTR18590	Lago Chaviana, Itapuru	AM - BR
R mirandaribeiroi	MTR26991	Barreiras	BA - BR
Gene/primers	16S / 16Sar-16Sbr		
Profile	[94°C (3:00); 94°C (0:45); 50°C (0:45); 72°C (0:45) for 37 cycles; 72°C (5:00)]		

TABLE S2 Pairwise Nei's G_{ST} and pairwise ddRADseq F_{ST} estimates. Values below diagonal are Nei's G_{ST} for the mitochondrial 16S data, and above diagonal in grey are Weir and Cockerham weighted F_{ST} estimates for the ddRADseq data.

	marina	***horribilis***	***jimi***	***poeppigii***	***schneideri***
marina	-	0.485	0.379	0.696	0.748
horribilis	0.044	-	0.68	0.843	0.783
jimi	0.053	0.063	-	0.84	0.762
poeppigii	0.183	0.198	0.212	-	0.517
schneideri	0.025	0.078	0.083	0.228	-

TABLE S3 Demographic model statistics. ⧫ Indicates the model with the highest log-likelihood.

3D-JSFS Models		**LL**	**AIC**
	no migration	-1699.9	3411.76
Continuous gene flow	*symmetric migration between all populations*	-1656.8	3333.5
	symmetric migration between adjacent populations	-1745.4	3508.88
	simultaneous split with continued migration between adjacent populations	-1632	3276.06
Secondary contact	*isolation with secondary contact*	-1840.2	3696.44
	simultaneous split with secondary contact between adjacent populations	-1920.5	3855.04
Ancient hybridization	*ancient migration with shortest isolation* ⧫	-1572.7	3165.38
	ancient migration with short isolation	-1706	3426.08
	ancient migration with longest isolation	-1668.7	3353.34
	ancient migration with a short isolation and population size change	-1815.2	3652.32
2D-JSFS Models			

	no migration	-758.16	1522.32
Continuous gene flow	*continuous symmetric migration*	-657.83	1323.66
	continuous asymmetric migration	-689.64	1389.28
	symmetric migration and a rate varying across two epochs	-696.19	1404.38
	asymmetric migration and a rate varying across two epochs	-684.68	1385.36
Secondary contact	*divergence in isolation with continuous symmetrical secondary contact*	-695.31	1400.62
	divergence in isolation with continuous asymmetrical secondary contact ♦	-566.14	1144.28
	divergence in isolation with continuous symmetrical secondary contact then subsequent isolation	-696.19	1404.38
	divergence in isolation with continuous asymmetrical secondary contact then subsequent isolation	-691.92	1397.84
Ancient hybridization	*divergence with ancient continuous symmetrical migration then subsequent isolation*	-811.51	1633.02
	divergence with ancient continuous asymmetrical migration then subsequent isolation	-932.44	1876.88

TABLE S4 D-statistics from ABBA-BABA tests.

P1	P2	P3	Dstatistic	Z-score	p-value
jimi	marina	horribilis	0.11	1.49	0.0683
horribilis	jimi	poeppigii	0.14	1.85	0.0321
horribilis	jimi	schneideri	0.49	11.35	0.00
horribilis	marina	poeppigii	0.24	3.47	0.0003
horribilis	marina	schneideri	0.17	2.37	0.0089
schneideri	poeppigii	horribilis	0.26	5.57	1.25E-08
jimi	marina	poeppigii	0.14	2.45	0.0070
marina	jimi	schneideri	0.37	5.93	1.51E-09
schneideri	poeppigii	jimi	0.12	2.26	0.0119
schneideri	poeppigii	marina	0.30	5.85	2.47E-09

* not significant

TABLE S5 F-branch statistics. Stats correspond to graphical representation in Figure 6.

Branch b	Branch C	fb	z-score
poeppigii	marina	0.11	5.85
poeppigii	horribilis	0.08	5.57
jimi	schneideri	0.07	5.93

poeppigii	jimi	0.04	2.26
C	schneideri	0.03	2.37
marina	poeppigii	0.03	2.45
C	poeppigii	0.02	1.85
schneideri	horribilis	0.00	0.00
schneideri	jimi	0.00	0.00
schneideri	marina	0.00	0.00
horribilis	poeppigii	0.00	0.00
horribilis	schneideri	0.00	0.00
jimi	poeppigii	0.00	0.00
jimi	horribilis	0.00	0.00
marina	schneideri	0.00	0.00
marina	horribilis	0.00	1.49

Chapter II: Phylogenomics and historical demography within the *Rhinella granulosa* toad species complex

Abstract

Mito-nuclear discordance, often identified through multi-locus sequencing of selected markers, presents particular difficulties in identifying historical processes which drive species diversity and boundaries. Mechanisms causing discordance, such as incomplete lineage sorting or introgression due to interspecific hybridization, are better identified based on population-level genomic datasets. In the toads of the *Rhinella granulosa* complex, patterns of mito-nuclear discordance and potential hybridization have been reported by several studies. However, these patterns were proposed based on few loci, such that alternative mechanisms behind gene-tree heterogeneity cannot be ruled out. Using genome-wide ddRADseq loci from a subset of species within this clade, we found only partial concordance between currently recognized species-level taxon boundaries and patterns of genetic structure. While most taxa within the *R. granulosa* species complex correspond to clades, genetic clustering analyses sometimes grouped distinct taxonomic units into a single cluster. Moreover, levels of admixture between inferred clusters were limited and restricted to a single taxon pair. In addition, D-statistics indicate that allele sharing across species is explained by incomplete lineage sorting as opposed to introgressive hybridization. These findings contradict previous assertions of widespread cryptic diversity in the *R. granulosa* clade and of disseminated gene flow. Lastly, our analyses suggest that diversification events within the *Rhinella granulosa* complex mostly dated back to the early Pliocene, being generally younger than species divergences in other closely related clades that present much higher levels of cross-species gene flow. This finding contradicts assertions that

the likelihood of hybridization scales negatively with levels of genetic divergence among species.

Introduction

Discordance between mitochondrial and nuclear genomes has been increasingly observed in numerous species over the past decades (Firneno et al., 2020; Ivanov, et al., 2018; Toews & Brelsford, 2012). This phenomenon has been explored more recently with high throughput sequencing, allowing the ability to analyze the nuclear genome at a larger scale compared to previous single or multi-locus studies (Firneno et al., 2020; Firneno & Townsend, 2019). Different patterns of differentiation between mitochondrial and nuclear genomes are commonly observed when comparing relationships among species or populations in phylogeographic studies (Firneno et al., 2020; Rivera, et al., 2020). Discordance can be caused by a variety of mechanisms, including differing selective pressures, sex-biased dispersal, neutral demographic processes, introgression or hybridization, incomplete lineage sorting, or a combination of mechanisms (Dufresnes et al., 2020; Firneno et al., 2020; Ivanov et al., 2018; Thielsch, et al., 2017; Toews & Brelsford, 2012). Gene tree discordance has made phylogenetic reconstruction challenging, but this issue has been improved by the use of high throughput, reduced representation, or whole genome sequencing (Firneno et al., 2020; Graham et al., 2018).

Mito-nuclear discordance, often identified through multi-locus sequencing of selected markers, presents particular difficulties in the identification of historical demographic processes that are driving species diversity and delimitation (Degnan & Rosenberg, 2009; Fujita, et al., 2012). The use of larger genomic datasets has helped determine the potential causes of mito-

nuclear discordance, even among species with relatively large genomes (Firneno et al., 2020; Hill et al., 2019). Moreover, the use of demographic model testing and tests for phylogenetic introgression between species or populations using this type of data provide the potential to uncover the root causes of genomic incongruence (Leaché et al., 2019; Portik et al., 2017). These methods can help to identify potential biogeographic, selective, or neutral processes that might be affecting phylogenetic structure in nuclear versus mitochondrial genomes.

Among the groups of organisms reported to have high levels of mito-nuclear discordance are the true toads (Bufonidae) (Azevedo, et al., 2003; Firneno et al., 2020; Pereyra et al., 2016; Sequeira et al., 2011). Widespread mito-nuclear discordance in toads was initially attributed to introgression of entire loci due to species hybridization, in agreement with the low selectivity of mating partners by these toads during their typically explosive breeding events (Abreu, Set al., 2021). However, a recent analysis using genome-wide loci found evidence of alternative mechanisms other than hybridization behind this discordance, in particular incomplete lineage sorting (Firneno et al., 2020). This analysis demonstrates the benefit of large genomic datasets when attempting to discern which mechanisms or processes may be driving patterns of among-locus heterogeneity in natural populations (Degnan & Rosenberg, 2009).

Within the true toad radiation, one clade in particular has been hypothesized to involve high levels of hybridization and poorly resolved intraspecific relationships due to mito-nuclear incongruencies: the *Rhinella granulosa* species complex (Narvaes & Rodrigues, 2009; Pereyra et al., 2016). Within this complex of South American toads, 13 nominal taxa are currently recognized: *R. granulosa*, *R. pygmaea*, *R. bergi*, *R. major*, *R. mirandaribeiroi*, *R. azarai*, *R. nattereri*, *R. fernandezae*, *R. dorbignyi*, *R. merianae*, *R. humboldti*, *R. bernardoi*, and *R. centralis* (Narvaes & Rodrigues, 2009). Although studies incorporating multi-locus datasets have found

marked genetic structure often corresponding to proposed taxon boundaries, several relationships remain poorly resolved, with conflicting patterns among loci. While these conflicts have at times been attributed to limited variability in the loci targeted or limited taxonomic sampling, some studies have proposed that poorly resolved relationships may stem from conflicting signals across markers due to hybridization between species (Pereyra et al., 2016; Simon, et al., 2016). Within the complex, potential hybridization has been hypothesized to occur between the taxon pairs *R. bergi* and *R. major*, *R. granulosa* and *R. mirandaribeiroi*, and *R. fernandezae* and *R. dorbignyi*. However, there has been no formal attempt to test the hybridization scenario versus alternative discordance-generating mechanisms based on an adequate number of independent genomic regions.

In this contribution, we implement a phylogenomic approach to formally test the hypothesis of introgressive hybridization as a source of mito-nuclear discordance across taxa within the *Rhinella granulosa* species complex. Based on comprehensive geographic sampling of genome-wide patterns of variation in multiple taxa across South America's major biomes, we infer the frequency and extent of introgression based on gene flow estimates and alternative models of historical demography. Based on this approach, we seek to address the following questions: Is there introgression between species within this complex, and how widespread is it? Are zones of introgression restricted to certain geographic regions, and where are they located? Lastly, how do patterns of introgression within the *R. granulosa* complex compare to other

similarly distributed toad clades where introgression has recently been demonstrated to be rampant?

Methods

Sampling of molecular data

Our genetic sampling included 45 individuals belonging to a subset of taxa in the *Rhinella granulosa* species complex, namely seven individuals of *R. granulosa*, eight *R. major*, 14 *R. mirandaribeiroi*, 13 *R. merianae*, one *R. centralis*, and two *R. humbolti*. As outgroups, we included four samples from the *Rhinella margaritifera* species complex. Tissue samples were obtained from the Museum of Zoology of the University of São Paulo (MZUSP), the Amphibian and Reptile Diversity Research Center (ARDRC) at the University of Texas in Arlington, and the Louisiana State University Museum of Natural Science (LSUMNS).

We extracted genomic DNA using a phenol-chloroform extraction protocol with SeraPure SpeedBead cleanup (Sambrook & Russell, 2006). Fragments of the mitochondrial 16S gene were PCR-amplified using the 16Sar and 16Sbr primers and sequenced on an ABI 3730xL sequencer following Rivera et al. (Chapter I of this Dissertation). We also supplemented the 16S dataset with additional sequences from *R. bernardoi*, *R. merianae*, *R. mirandaribeiroi*, *R. pygmaea*, *R. major*, *R. bergi*, *R. centralis*, *R. dorbignyi*, *R. fernandezae, R. granulosa*, and *R. humbolti* obtained from GenBank. The only species not sampled from this group for the 16S

phylogeny was *R. nattereri*. Sequences were edited and aligned in Geneious Prime 2020.0.4 (Identification and Accession numbers in Supplementary Table S1).

To characterize patterns of nuclear genetic structure, we generated a double-digest restriction-site associated DNA (ddRAD) dataset following Peterson et al. (2012). Briefly, 200-500 ng of genomic DNA were digested using the *Sbf*I (restriction site 5′-CCTGCAGG-3′) and *Msp*I (restriction site 5′-CCGG-3′) enzymes in a single reaction using the manufacturer's recommended buffer (New England Biolabs) for 5 hr at 37°C. Digested DNA was bead-purified before ligating barcodes and index adaptors, then samples with the same index were pooled and size-selected (415-515 bp) on a Blue Pippin Prep size selector (Sage Science). The quality and concentration of final libraries were analyzed and quantified on a BioAnalyzer (Agilent) and Qubit Fluorometer 4 (Thermo Fisher Scientific). The resulting 100 bp single-end libraries were sequenced at MedGenome on an Illumina HiSeq2500.

We used the command line version of ipyrad v. 0.9.45 (Eaton & Overcast, 2020) (available at https://ipyrad.readthedocs.io) to de-multiplex and assign reads to individuals based on sequence barcodes allowing no mismatches from individual barcodes, perform *reference* read assembly using a minimum clustering similarity threshold of 0.90, align the reads into loci, and call single nucleotide polymorphisms (SNPs). As a reference, we used the *Rhinella marina* genome (Edwards et al., 2018). A minimum Phred quality score (= 33), sequence coverage (= 6x), read length (= 35 bp), and maximum proportion of heterozygous sites per locus (= 0.5) were enforced, while ensuring that variable sites had no more than two alleles (i.e., a diploid genome). Following the initial assembly, we used Matrix Condenser (de Medeiros & Farrell, 2018) to assess levels of missing data across samples and then re-assembled our dataset to ensure a maximum of 25% missing data within each locus (i.e., each retained locus was sequenced in at

least 25% of the sampled individuals). This strategy resulted in 45 ingroup and four outgroup samples in a matrix with less than 15% total missing data. The final dataset was composed of 16,455 SNPs.

Phylogenetic relationships

We inferred maximum likelihood phylogenies for both the mitochondrial dataset and ddRADseq data using IQTREE v2.1.2, utilizing the built-in model selection tool ModelFinder Plus, implementing 1000 ultrafast bootstraps (Hoang, et al., 2018; Kalyaanamoorthy, et al., 2017; Minh et al., 2020). We utilized the greedy algorithm of PartitionFinder for model selection, only testing models of evolution available in MrBayes (Lanfear, et al., 2012; Lanfear, et al., 2014; Lanfear, et al., 2017).

In addition to the maximum likelihood analyses, we performed phylogenetic inference under a Bayesian framework using MrBayes 3.2.6 (Ronquist et al. 2012), implementing three independent runs of four Markov chains of 10 million generations each and sampling every 1,000 generations with the first 25% generations discarded as burn-in. We used the same best-fit model found by ModelFinder Plus in the Bayesian analyses. We used Tracer 1.7 (Rambaut et al. 2018) to assess whether Markov chain mixing was adequate (effective sample sizes > 200) and to visually assess model parameter stationarity and convergence between runs. We then summarized a 50% majority-rule consensus tree and used iTol to edit and visualize trees (Letunic & Bork, 2019).

Population genetic structure

To determine populations and admixture in the *R. granulosa* species complex, we filtered SNPS as described above but excluding outgroup samples using ipyrad and VCFtools (Danecek et al., 2011), resulting in a final dataset composed of 20,527 SNPs. We then determined the best-fit number of genetic clusters (K) using the maximum likelihood method ADMIXTURE with 20 replicates per K and a 10-fold cross-validation to determine the best-fit K (Alexander, et al., 2009). Specifically, the best-fit K was determined as the replicate with the lowest cross-validation error.

To further characterize population structure, we used the non-parametric method of discriminant analysis of principal components (DAPC), implemented in the R package adegenet (Jombart & Ahmed, 2011; Jombart, et al., 2010). The *find.clusters* function was used to test the fit of 1-20 clusters (K). The K with the lowest Bayesian information criterion (BIC) score was considered the best-fit number of clusters.

The resulting ancestry coefficient matrices (Q-matrices) from these clustering analyses were then imported into QGIS (QGIS Development Team, 2020) to allow visualizing patterns of genetic clustering and admixture in geographic space.

Gene flow

To infer potential hybridization or gene flow among taxa, we calculated Patterson's D statistic and the related admixture fraction estimates, or f_4-ratio statistics, using the program Dsuite (Malinsky, et al., 2020; Patterson et al., 2012). This approach, also known as an ABBA-BABA test, uses a 4-taxon fixed phylogeny -- in the form (((P1,P2)P3)O) -- to quantify the proportion of shared alleles that can be attributed to horizontal transfer among the populations

considered. This approach assumes that a typical ancestral ("A") and derived ("B") allele pattern among four terminal taxa should generate a BBAA structure. Under incomplete lineage sorting, conflicting ABBA and BABA patterns should occur in equal frequencies, resulting in a D statistic = 0. If, however, introgression between P3 and P1 or P2 has occurred, there should be an excess of one of these two patterns, and a thus D statistic significantly different from 0 (Durand, et al., 2011; Patterson et al., 2012). We used the *f*-branch or f_b(C) metric to tease apart potentially correlated f_4-ratio statistics and estimate gene flow events between internal branches on the phylogeny (Malinsky et al., 2018; Martin et al., 2013). Dsuite uses a VCF file and a jackknifing approach to assess correlations in allele frequencies between closely-related species (Malinsky et al., 2020). Within Dsuite, we used the *Dtrios* and *Fbranch* programs to identify introgression between all combinations of species, using *R. margaritifera* as the outgroup taxa, and applied the Benjamini-Hochberg (BH) correction to control for false discovery rate using a P value of 0.05.

Demographic modeling

We use the full-likelihood, multi-species coalescent method Generalized Phylogenetic Coalescent Sampler (G-PhoCS) v.1.3.2 on Cipres to estimate demographic parameters such as effective population sizes and divergence times along evolutionary history the *Rhinella granulosa* species complex (Gronau, et al., 2011; Miller, et al., 2010). For that, we used the phylogenetic topology inferred by our maximum likelihood and Bayesian analyses on the ddRAD data (see Results). To reduce computing time, we used a maximum of 10 randomly chosen individuals per delimited genetic cluster. G-PhoCS analyses used all 2,744 unlinked SNPs and an automatic fine-tuning for 500 steps, running the entire analysis twice independently for 500k generations, then again for one million generations, sampling every 1000 generations.

We then merged runs using LogCombiner and discarded the first 10% of samples as burn-in (Alonso, et al., 2012; Drummond & Rambaut, 2007; Pramuk, et al., 2007). The Dsuite analysis determined no significant gene flow between populations (see Results), so we did not apply migration bands in G-PhoCS analyses. We followed Prates et al. (Prates, et al., 2018) to select distributions for the priors of the θ and τ parameters (scripts available at https://github.com/ivanprates/2018_Anolis_EcolEvol), applying a gamma distribution with parameters $\alpha = 2.0$ and $\beta = 30$. We used TRACER 1.7 to assess proper chain mixing and convergence based on the log output files (Rambaut, et al., 2018). We then converted θ estimates to Ne, in number of individuals, using the relationship $\theta = 4Ne\mu$, and τ estimates to T, in years, using the relationship $\tau = T\mu/g$. To apply these conversions, we used a nuclear mutation rate (μ) of 2.4 x 10-9 (Prates et al., 2018; Prates, Rivera, et al., 2016) and a generation time (g) of two years for bufonid toads (Lever, 2001).

Results

Phylogenetic analyses of the 16S mitochondrial gene based on both maximum likelihood and Bayesian approaches yielded identical relationships (Fig. 1). Overall, most species-level taxa currently recognized within the *R. granulosa* species complex were inferred as monophyletic. Mitochondrial analyses inferred 10 major clades, most of which corresponding to currently recognized taxa. Moreover, most of these clades were geographically coherent, with samples assigned to the same taxon generally clustering in geographic space, as follows: *R. dorbignyi* and *R. fernandezae* in the southern Pampas (Uruguay); *R. major* across southern and eastern Amazonia; *R. granulosa* in the Caatinga and northern Atlantic Forest; *R. centralis* in the northern

Andes, *R. humbolti* in northern Amazonia; *R. merianae* in central Amazonia; and *R. mirandaribeiroi* across the Cerrado.

The Bayesian mitochondrial analysis resulted in a phylogeny with higher overall relative support for most of the internal nodes (PP = 0.90 - 1.0) that correspond to relationships between species-level taxa. The only relationship that received lower support (PP = 0.64) was the position of *R. mirandaribeiroi* and the clade composed of *R. granulosa, R. centralis, R. humbolti,* and *R. merianae*. (Fig. 1). This phylogeny placed *R. major* as the sister to all other remaining taxa within this complex; within this clade, *R. pygmaea*, *R. azarai* and *R. bergi* formed their own respective clades, and *R. dorbingyi* and *R. fernandezae* formed a subclade. The other major clade within the complex included *R. mirandaribeiroi*, *R. granulosa*, *R. centralis*, *R. humbolti,* and *R. merianae*, nested in this order relative to the root.

Phylogenetic relationships based on the ddRADseq dataset -- which included only a subset of the taxa represented in the mitochondrial trees -- were overall highly supported (Fig. 1; Fig. 2A). Similar to the mitochondrial analyses, all of the currently recognized species were inferred to be monophyletic and geographically coherent, except for one *R. granulosa* individual, which was inferred as nested within the *R. mirandaribeiroi* clade (Fig. 1; Fig. 2A). The ddRAD phylogeny included two major clades: one composed of *R. granulosa* and *R. mirandaribeiroi*, and another composed of *R. major*, *R. merianae*, *R. humbolti*, and *R. centralis* (Fig. 1; Fig. 2A).

While the sample composition of clades was similar between mitochondrial and nuclear analyses, the relationships between taxon clades differed between marker types. Most notably, the affinity between *R. major* and the clade composed of *R. centralis*, *R. humbolti*, and *R. merianae* varied by marker type. *Rhinella major* was sister to all other species in the 16S

phylogeny while it was in a clade with *R. centralis*, *R. humbolti*, and *R. merianae* in the nuclear phylogeny (Fig. 1). Additionally, the placement of *R. mirandaribeiroi* in relation to *R. granulosa* differed between mitochondrial and nuclear markers as well. In the 16S phylogeny, *R. mirandaribeiroi* was sister to a clade with *R. granulosa*, *R. centralis*, *R. humbolti*, and *R. merianae*, while in the nuclear phylogeny, *R. mirandaribeiroi* shared a clade with *R. granulosa* only (Fig. 1).

Population Structure

ADMIXTURE results indicated the K value with the lowest cross-validation error was 4, which separated *R. major* (orange), *R. granulosa* (blue), *R. mirandaribeiroi* (purple), and *R. merianae* + *R. humbolti* + *R. centralis* (yellow; Fig. 2A). Two individuals - one *R. granulosa* and one *R. mirandaribeiroi* - were found to be admixed with the genetic clusters from the two species (Fig. 2A). The DAPC analysis yielded similar BIC scores for K = 4 - 6 (Fig. S2). K = 4 resulted in the same clusters identified by ADMIXTURE, while K = 5 split the ADMIXTURE cluster composed of *R. merianae* + *R. humbolti* + *R. centralis* into *R. merianae* and *R. humbolti* + *R. centralis* (Fig. 2), and K = 6 further subdivided *R. mirandaribeiroi* into two clusters, one of which included the *R. granulosa* individual inferred to be nested within *R. mirandaribeiroi* (Fig. 2).

Similar to the phylogenetic analyses, major genetic clusters inferred by ADMIXTURE were geographically coherent, each restricted to a certain portion of South America. Specifically, the cluster composed of samples assigned to *R. major* was restricted to the Amazon region; the cluster composed of *R. merianae*, *R. humbolti*, and *R. centralis* was distributed across the

Amazon and the Northern Andes; the cluster composed of samples assigned to *R. granulosa* was distributed across the Northern Atlantic Forest in eastern Brazil; and the cluster composed of samples assigned to *R. mirandaribeiroi* was distributed across the Cerrado in central Brazil (Fig. 3).

D-statistics

Analyses of introgression based on the ddRAD data using Dsuite resulted in only one trio ((*R. mirandaribeiroi*, *R. granulosa*) *R. major*) out of 20 trios tested that had a significant D-statistic (0.41; Table S2), pointing to a larger proportion of shared alleles than expected based solely on incomplete lineage sorting. However, after applying the Benjamini-Hochberg (BH) correction to control for the false discovery rate, none of the trios tested were significant (Table S2). These results were corroborated using the f_b(C) metric analysis, which showed an f_b statistic of 0.15 between *R. granulosa* and *R. major*, but it was not significant (Fig. 4, Table S3).

Demographic Inference

The G-PhoCS analysis estimated the *Rhinella granulosa* species complex to date back to about 5 million years before present (mybp) (mean value; 95% highest posterior density (HPD) = 4.817 - 5.092; Fig. 2; Table S4). Sister species divergences within the complex dated back to the mid-Pleistocene (Fig. 2; Table S4). The divergence between *R. granulosa* and *R. mirandaribeiroi* was estimated to date back to 1.026 mybp (HPD = 0.983 - 1.075), while the divergence between *R. humbolti* and *R. centralis* dated back to 1.496 mybp (HPD = 1.258 - 1.733). In turn, the most recent common ancestor of *R. merianae*, *R. humbolti*, and *R. centralis* dated back to 2.263 mybp

(2.158 - 2.358 95% HPD), while the ancestor of these three species and *R. major* dated back to 4.826 mybp (HPD = 4.625 - 4.983 95% HPD). The effective population size estimates resulted in a root population size of ~1.3 million, remaining stable along the *R. granulosa-mirandaribeiroi* branch while massively expanding to Amazonia along the *R. major-merianae-humbolti-centralis* branch, then experiencing a bottleneck after *R. major* diverged, then expanding again after *R. merianae* diverged (Fig. 2; Table S4). Estimates of current population sizes for each species ranged from ~145,833 - 1,093,750 (Table S4).

Discussion

Based on comprehensive geographic sampling of mitochondrial and nuclear genetic variation in a widespread species group of neotropical toads, we found only partial concordance between currently recognized species-level taxon boundaries and patterns of genetic structure. While most taxa within the *R. granulosa* species group corresponded to clades, genetic clustering analyses sometimes grouped distinct taxonomic units into a single cluster. Moreover, levels of admixture between inferred clusters were limited and restricted to a single sister taxon pair (*R. granulosa* and *R. mirandaribeiroi*), while D-statistics estimation provided no support for the hypothesis that allele sharing between non-sister taxa is a result of horizontal transfer. These findings contradict previous assertions of widespread cryptic diversity and of disseminated gene flow and introgression across taxa in the *R. granulosa* clade, a proposed scenario based on small numbers of loci (Guerra *et al.*, 2011; Pereyra *et al.*, 2016). Lastly, our analyses suggest that diversification events within the *Rhinella granulosa* clade mostly dated back to the early Pliocene, which is later than species divergences in other closely related clades, as observed in the *R. marina* species group (see Chapter I of this Dissertation).

Previous analyses of phylogenetic relationships in the *Rhinella granulosa* complex based on multi-locus datasets have resulted in unresolved interspecific relationships and evidence of mito-nuclear discordance (Azevedo et al., 2003; Firneno et al., 2020; Pereyra et al., 2016; Sequeira et al., 2011). By incorporating thousands of ddRADseq nuclear loci, however, we were able to improve phylogenetic resolution within this clade (Fig. 2). Previous studies have, in several cases, inferred relationships similar to ours. Specifically, both our study and that of Pereyra et al. (2016) found a sister relationship between *R. granulosa* and *R. mirandaribeiroi*, and well as phylogenetic clustering of *R. centralis*, *R. humbolti*, and *R. merianae*. However, while Pereyra et al (2016) found *R. major* to be the sister of all other species, our analysis inferred *R. major* as sister to the *R. centralis* + *R. humbolti* + *R. merianae* clade with high support (Fig. 2).

These discrepancies could have originated from the higher reliance of Pereyra et al.'s study in mitochondrial genes. Although we did not observe a clear signal of mitochondrial introgression across taxa based on the 16S locus, their analysis of other mitochondrial loci found several taxa to be non-monophyletic. This is the case, for instance, of *R. bergi*, *R. major*, and *R. fernandezae*, suggesting these taxa might have experienced past introgressive events (Pereyra et al., 2016). It is possible that some sympatric species in the *R. granulosa* complex have hybridized in the distant past, as recently reported in another clade of neotropical toads, the *R. marina* group (see Chapter I of this Dissertation). However, our analysis of genetic introgression based on extensive nuclear data found no support for hybridization as the main factor behind patterns of allele sharing across taxa. Considering that processes like purifying selection can also reduce genetic diversity at linked neutral sites (Cvijović, Good, & Desai, 2018), previous assertions of widespread introgression across species in the *R. granulosa* species complex based

on mitochondrial data may have resulted from an overestimation of shared loci through putative hybridization.

Instead, our analyses support incomplete lineage sorting as the main cause of allele sharing patterns. Incomplete lineage sorting may also explain our finding of apparent mito-nuclear discordance or incongruent species topologies across marker types (Fig. 2; Fig. 4; Table S2; Table S3). This is the case, for instance, of the *R. granulosa* and *R. mirandaribeiroi* pair. These taxa showed some admixture and displayed contrasting patterns of mitochondrial and nuclear phylogenetic structure. However, the D-statistics results supported that hypothesis that the proportion of shared alleles is no different from that expected under incomplete lineage sorting (Fig. 2; Fig. 4; Table S3). Taken together, our analyses based on genome-wide loci challenge previous assertions of rampant hybridization in the evolutionary history of the *R. granulosa* species complex. However, we note that our taxon sampling was more limited than that of previous studies due to challenges in accessing samples from throughout the vast distribution of this complex. It is possible that inclusion of the other species, as well as better sampling of each species across their respective ranges, especially near any potential contact zones, could reveal more complex patterns of admixture.

Our finding of negligible introgression across species in the *R. granulosa* complex contrasts with patterns observed in other closely related clades. For instance, the *Rhinella marina* species group experienced many instances of genetic admixture across species over its evolutionary history. Recent investigation based on thousands of genome-wide loci inferred multiple instances of admixture across and within species in the *R. marina* group, as well as past and ongoing pulses of gene flow (Chapter I, Fig. 1,2,5,6). Differential roles of hybridization inferred in the history of the *R. marina* and *R. granulosa* clades might have originated from

differences in taxonomic or geographic sampling of species, given that a relatively higher number of taxa and samples were included in the analysis of the *R. marina* group. Nevertheless, it is also possible that these patterns reflect biological differences across clades. Toads are generally reported to be prone to low selectivity of mating partners during their typically explosive breeding events, which also seems to apply to the *R. granulosa* complex; for instance, interspecific amplexus between co-distributed *Rhinella granulosa* and *R. crucifer* (a species from the *R. crucifer* species group) have been reported (Abreu et al., 2021). Nevertheless, our findings suggest that these apparently permeable prezygotic reproductive barriers do not necessarily lead to gene transfer across interbreeding species. If that is the case, it is possible that species in the *R. granulosa* complex have developed post-zygotic barriers to hybridization that species in other closely related clades have not, as it appears to be the case of the *R. marina* species group (Abreu et al., 2021; Pereyra et al., 2016). Interestingly, propensity for hybridization does not seem to be associated with the timing of species divergences (Singhal & Bi, 2017), as estimates for the *R. marina* clade suggest older divergences overall than those in the *R. granulosa* complex (see Chapter I of this Dissertation). Future comparative studies of genetic structure across clades will benefit from incorporating information on species' organismal attributes and natural history. Unfortunately, this much-needed information is largely lacking for species-rich clades of tropical organisms, as is the case of the intriguing true toads.

Acknowledgements

We thank all the international field biologists who made this study possible and particularly students in MTR's laboratory for field collection and support. Brazil's Instituto Chico Mendes de

Conservação da Biodiversidade issued collecting permits (SISBIO 36753-1, 36753-4, and 27290-3). This work was co-funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; BIOTA 2013/50297-0), the National Science Foundation (DEB 1343578), and the National Aeronautics and Space Administration through the Dimensions of Biodiversity Program. MTR acknowledges additional funding from FAPESP grants 2003/10335-8, 2011/50146-6, 2011/50206-9, 2012/15754-8, and 2017/08357-6. DR was funded by NSF GRFP. IP was funded by NSF grant DEB 1754398.

References

Abreu, L. M., Santos, J. E. M., Corrêa, A. T., & Ribeiro, M. V. (2021). Interspecific amplexus between two sympatric species, *Rhinella granulosa* (Spix, 1824) and *Rhinella crucifer* (Wied-Neuwied, 1821) (Anura: Bufonidae) in Ilhéus, Bahia, Brazil. *HNO*, *14*, 869–871.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664.

Alonso, R., Crawford, A. J., & Bermingham, E. (2012). Molecular phylogeny of an endemic radiation of Cuban toads (Bufonidae: *Peltophryne*) based on mitochondrial and nuclear genes: Origin and diversification of Cuban toads. *Journal of Biogeography*, *39*(3), 434–451.

Azevedo, M. F. C., Foresti, F., Ramos, P. R. R., & Jim, J. (2003). Comparative cytogenetic studies of *Bufo ictericus*, *B. paracnemis* (Amphibia, Anura) and an intermediate form in sympatry. *Genetics and Molecular Biology*, Vol. 26, pp. 289–294. doi: 10.1590/s1415-47572003000300012

Cvijović, I., Good, B. H., & Desai, M. M. (2018). The Effect of Strong Purifying Selection on Genetic Diversity. *Genetics*, *209*(4), 1235–1278.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158.

Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, *24*(6), 332–340.

de Medeiros, B. A. S., & Farrell, B. D. (2018). Whole-genome amplification in double-digest RADseq results in adequate libraries but fewer sequenced loci. *PeerJ*, *6*, e5089.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*, 214.

Dufresnes, C., Nicieza, A. G., Litvinchuk, S. N., Rodrigues, N., Jeffries, D. L., Vences, M., … Martínez-Solano, Í. (2020). Are glacial refugia hotspots of speciation and cytonuclear discordances? Answers from the genomic phylogeography of Spanish common frogs. *Molecular Ecology*, *29*(5), 986–1000.

Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution*, *28*(8), 2239–2252.

Eaton, D. A. R., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics, 36*(8), 2592–2594.

Edwards, R. J., Tuipulotu, D. E., Amos, T. G., O'Meally, D., Richardson, M. F., Russell, T. L., … White, P. A. (2018). Draft genome assembly of the invasive cane toad, *Rhinella marina*. *GigaScience*, Vol. 7. doi: 10.1093/gigascience/giy095

Firneno, T. J., Jr, O'Neill, J. R., Portik, D. M., Emery, A. H., Townsend, J. H., & Fujita, M. K. (2020). Finding complexity in complexes: Assessing the causes of mitonuclear

discordance in a problematic species complex of Mesoamerican toads. *Molecular Ecology*, *29*(18), 3543–3559.

Firneno, T. J., Jr., & Townsend, J. H. (2019). Evaluation of species boundaries in sympatric and parapatric populations of Mesoamerican toads. *Zoologica Scripta*, *48*(4), 454–465.

Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, *27*(9), 480–488.

Graham, A. M., Lavretsky, P., Muñoz-Fuentes, V., Green, A. J., Wilson, R. E., & McCracken, K. G. (2018). Migration-Selection Balance Drives Genetic Differentiation in Genes Associated with High-Altitude Function in the Speckled Teal (*Anas flavirostris*) in the Andes. *Genome Biology and Evolution*, *10*(1), 14–32.

Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G., & Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, *43*(10), 1031–1034.

Guerra, C., Baldo, D., Rosset, S., Borteiro, C., & Kolenc, F. (2011). Advertisement and release calls in Neotropical toads of the *Rhinella granulosa* group and evidence of natural hybridization between *R. bergi* and *R. major* (Anura: Bufonidae). *Zootaxa*, *3092*, 26–42.

Hill, G. E., Havird, J. C., Sloan, D. B., Burton, R. S., Greening, C., & Dowling, D. K. (2019). Assessing the fitness consequences of mitonuclear interactions in natural populations. *Biological Reviews of the Cambridge Philosophical Society*, *94*(3), 1089–1104.

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.

Ivanov, V., Lee, K. M., & Mutanen, M. (2018). Mitonuclear discordance in wolf spiders: Genomic evidence for species integrity and introgression. *Molecular Ecology*, *27*(7), 1681–1695.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 94.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589.

Lanfear, R., Calcott, B., Ho, S. Y. W., & 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., Calcott, B., Kainer, D., Mayer, C., & Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, *14*, 82.

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017). PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. *Molecular Biology and Evolution*, *34*(3), 772–773.

Leaché, A. D., Portik, D. M., Rivera, D., Rödel, M., Penner, J., Gvoždík, V., … Fujita, M. K. (2019). Exploring rain forest diversification using demographic model testing in the

African foam-nest treefrog *Chiromantis rufescens*. *Journal of Biogeography*, *46*(12), 2706–2721.

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, *47*(W1), W256–W259.

Lever, C. (2001). *The Cane Toad: The History and Ecology of a Successful Colonist*. Westbury Academic and Scientific Publishing.

Malinsky, M., Matschiner, M., & Svardal, H. (2020). Dsuite - Fast D -statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*. doi: 10.1111/1755-0998.13265

Malinsky, M., Svardal, H., Tyers, A. M., Miska, E. A., Genner, M. J., Turner, G. F., & Durbin, R. (2018). Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nature Ecology & Evolution*, *2*(12), 1940–1955.

Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., … Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research*, *23*(11), 1817–1828.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *2010 Gateway Computing Environments Workshop (GCE)*. doi: 10.1109/gce.2010.5676129

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.

Narvaes, P., & Rodrigues, M. T. (2009). Taxonomic revision of *Rhinella granulosa* species group (amphibia, anura, Bufonidae), with a description of a new species. *Volume*, *40*(1), 1–73.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., … Reich, D. (2012). Ancient admixture in human history. *Genetics*, *192*(3), 1065–1093.

Pereyra, M. O., Baldo, D., Blotto, B. L., Iglesias, P. P., Thomé, M. T. C., Haddad, C. F. B., … Faivovich, J. (2016). Phylogenetic relationships of toads of the *Rhinella granulosa* group (Anura: Bufonidae): a molecular perspective with comments on hybridization and introgression. *Cladistics: The International Journal of the Willi Hennig Society*, *32*(1), 36–53.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, *7*(5). Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/pmc3365034/

Portik, D. M., Leaché, A. D., Rivera, D., Barej, M. F., Burger, M., Hirschfeld, M., … Fujita, M. K. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Molecular Ecology*, *26*(19), 5245–5263.

Pramuk, J. B., Robertson, T., Sites, J. W., & Noonan, B. P. (2007). Around the world in 10 million years: biogeography of the nearly cosmopolitan true toads (Anura: Bufonidae). *Global Ecology and Biogeography 17*, 72-83.

Prates, I., Penna, A., Rodrigues, M. T., & Carnaval, A. C. (2018). Local adaptation in mainland anole lizards: Integrating population history and genome--environment associations. *Ecology and Evolution*, *8*(23), 11932–11944.

Prates, I., Rivera, D., Rodrigues, M. T., & Carnaval, A. C. (2016). A mid-Pleistocene rainforest corridor enabled synchronous invasions of the Atlantic Forest by Amazonian anole lizards. *Molecular Ecology*, *25*(20), 5174–5186.

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.

Rivera, D., Prates, I., Rodrigues, M. T., & Carnaval, A. C. (2020). Effects of climate and geography on spatial patterns of genetic structure in tropical skinks. *Molecular Phylogenetics and Evolution*, *143*, 106661.

Sambrook, J., & Russell, D. W. (2006). Purification of nucleic acids by extraction with phenol:chloroform. *CSH Protocols*, *2006*(1). doi: 10.1101/pdb.prot4455

Sequeira, F., Sodré, D., Ferrand, N., Bernardi, J. A. R., Sampaio, I., Schneider, H., & Vallinoto, M. (2011). Hybridization and massive mtDNA unidirectional introgression between the closely related Neotropical toads *Rhinella marina* and *R. schneideri* inferred from mtDNA and nuclear markers. *BMC Evolutionary Biology*, *11*, 264.

Simon, M. N., Machado, F. A., & Marroig, G. (2016). High evolutionary constraints limited adaptive responses to past climate changes in toad skulls. *Proceedings. Biological Sciences / The Royal Society*, *283*(1841). doi: 10.1098/rspb.2016.1783

Singhal, S., & Bi, K. (2017). History cleans up messes: The impact of time in driving divergence and introgression in a tropical suture zone. *Evolution; International Journal of Organic Evolution*, *71*(7), 1888–1899.

Thielsch, A., Knell, A., Mohammadyari, A., Petrusek, A., & Schwenk, K. (2017). Divergent clades or cryptic species? Mito-nuclear discordance in a *Daphnia* species complex. *BMC Evolutionary Biology*, *17*(1), 227.

Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, Vol. 21, pp. 3907–3930. doi: 10.1111/j.1365-294x.2012.05664.x

Figures

FIGURE 1 A comparison of the *Rhinella granulosa* complex 16S and ddRAD phylogenies. Support values are indicated above (PP) and below (BS) each node in grey. Red lines highlight mito-nuclear discordance.

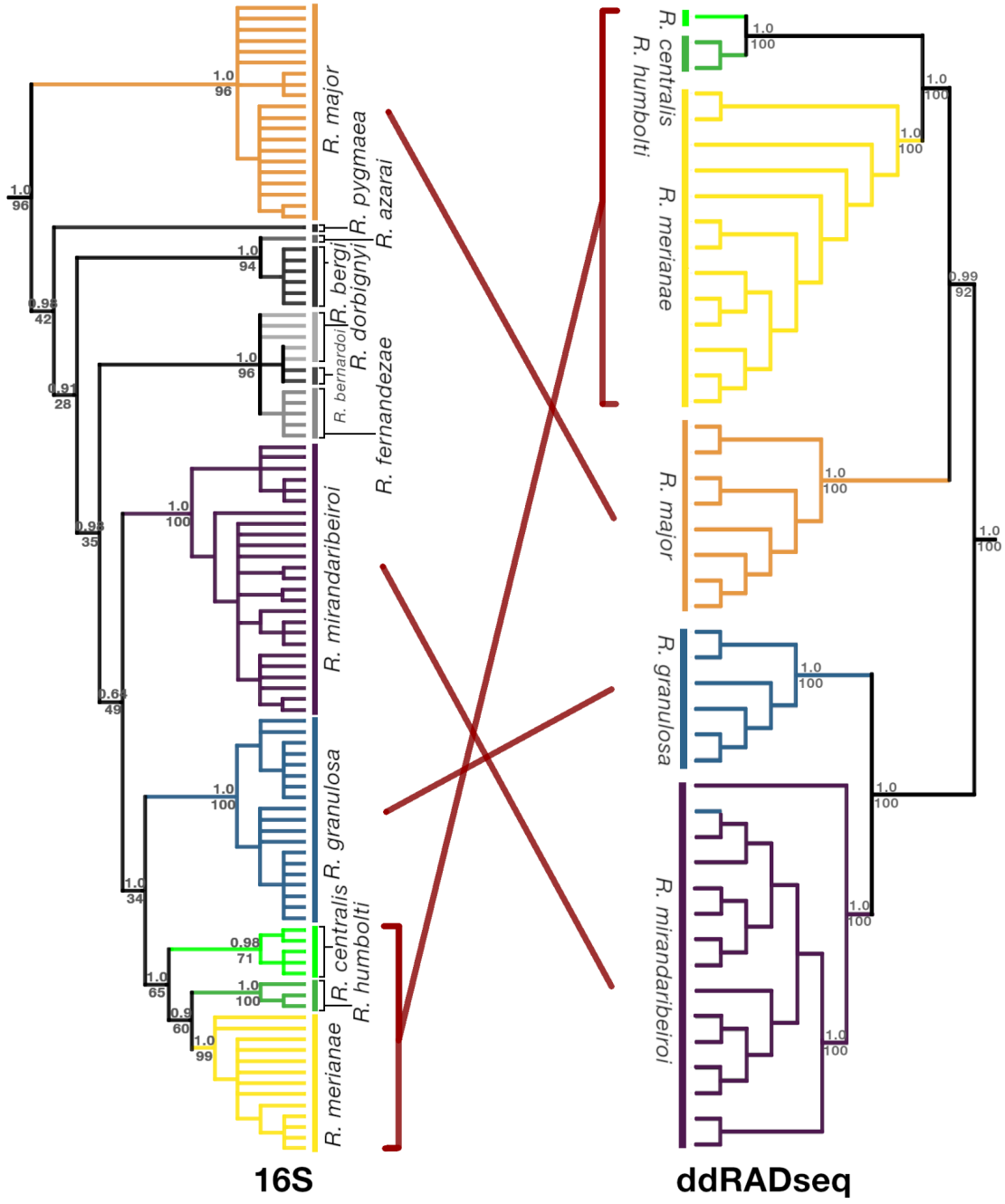


FIGURE 2 (A) Maximum likelihood/ Bayesian time-calibrated phylogeny (G-PhoCS) for *R. granulosa* complex focal species using ddRADseq data, and corresponding ADMIXTURE and DAPC plots. Branch widths are scaled to estimated population sizes (G-PhoCS) and * represent BS/PP > 95/0.99. (B) DAPC plots for K = 4 - 6.

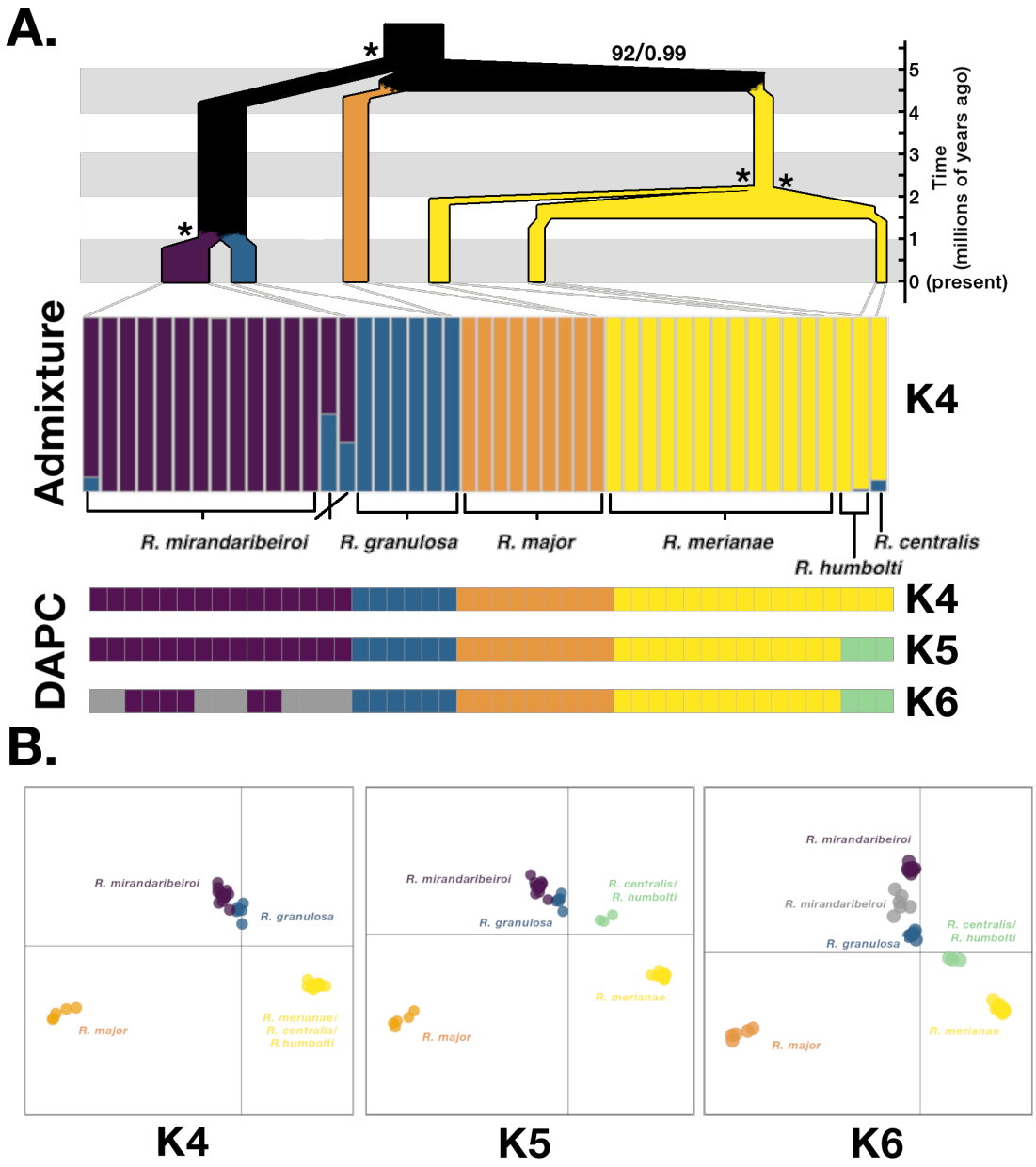


FIGURE 3 Locality map of focal species depicting average ADMIXTURE cluster assignments per locality (K=4). Map partitioned into biomes (Central America, Northern Andes, Northern Amazonia, Western Amazonia, Eastern Amazonia, Southern Amazonia, Pantanal, Chaco, Cerrado, Caatinga, Northern Atlantic Forest, Southern Atlantic Forest).

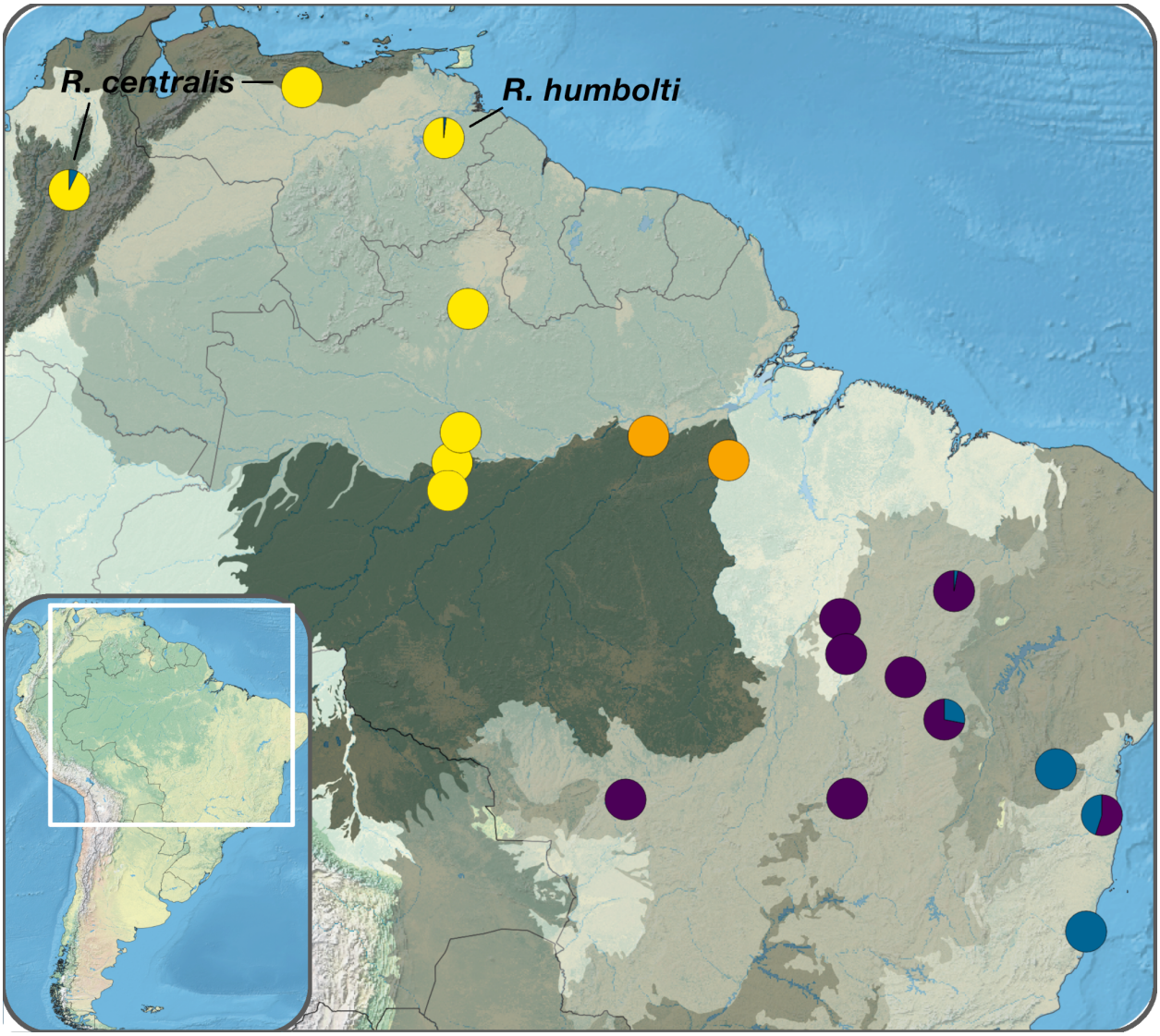
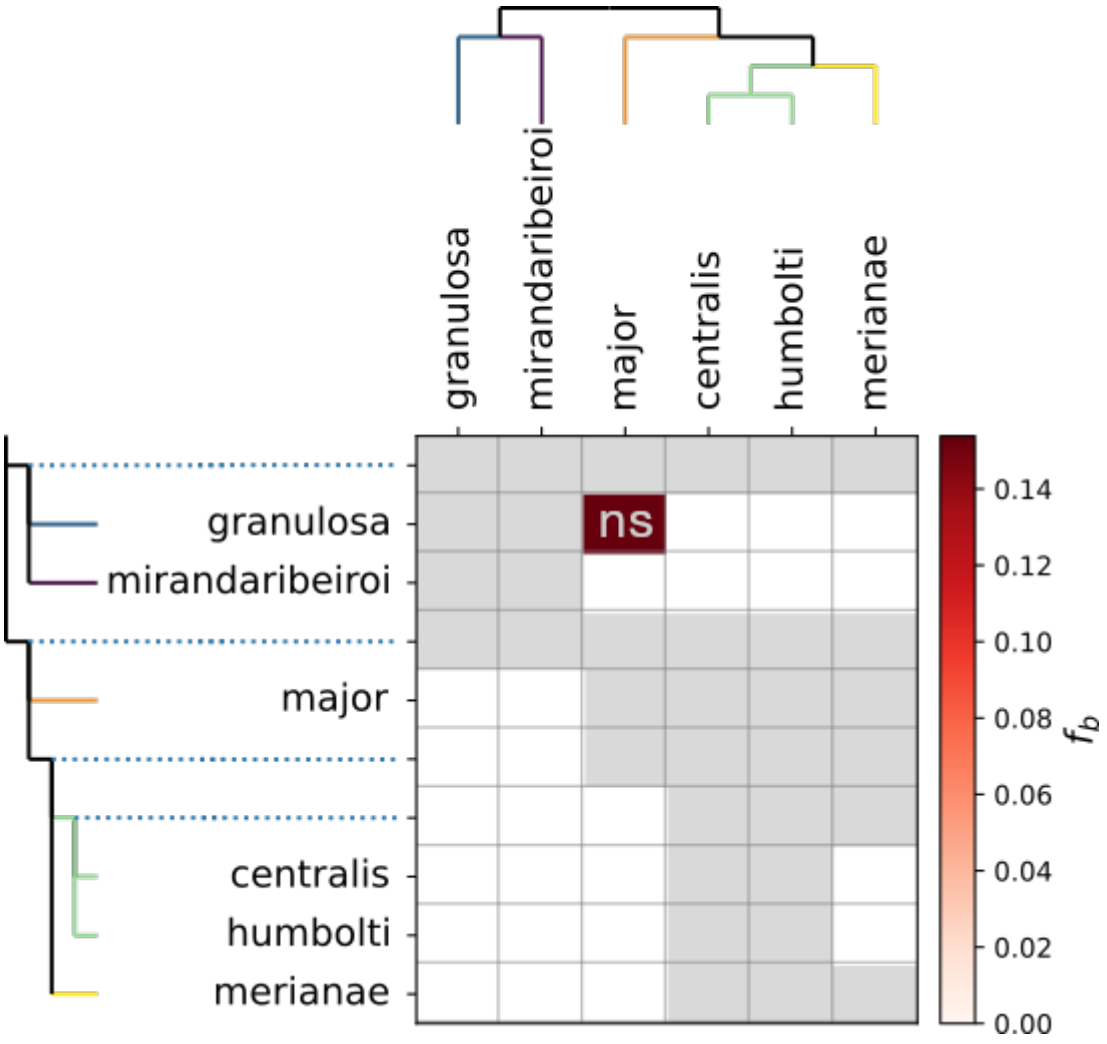


FIGURE 4 The f_b statistic (summary of f_4 admixture ratios). Grey color corresponds to tests that are not possible because of constraints on the phylogeny. ns = not significant.



Supplementary Figures

FIGURE S1 Locality map for sampled *R. granulosa* species. Colors correspond to mitochondrial 16S tree in Fig. 1.

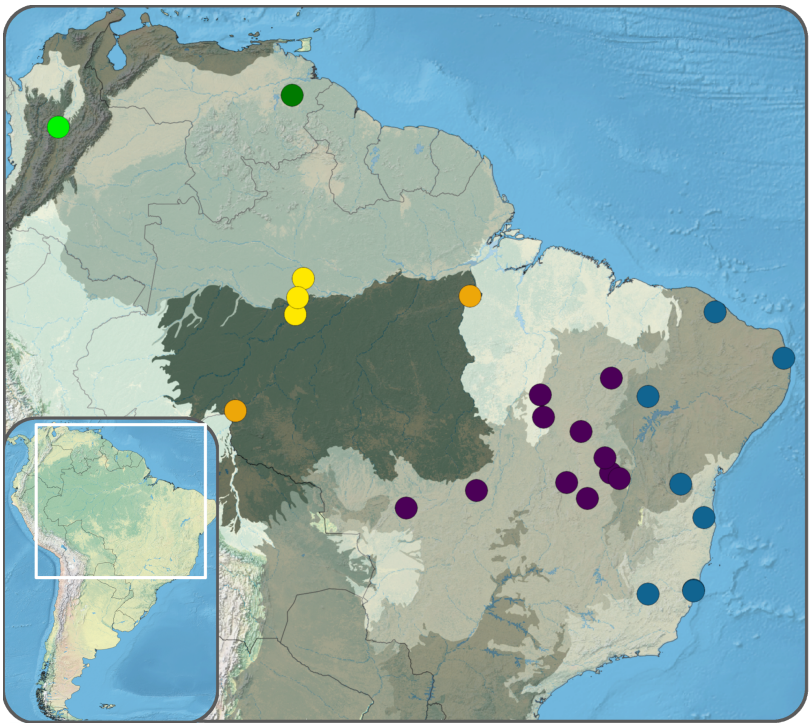
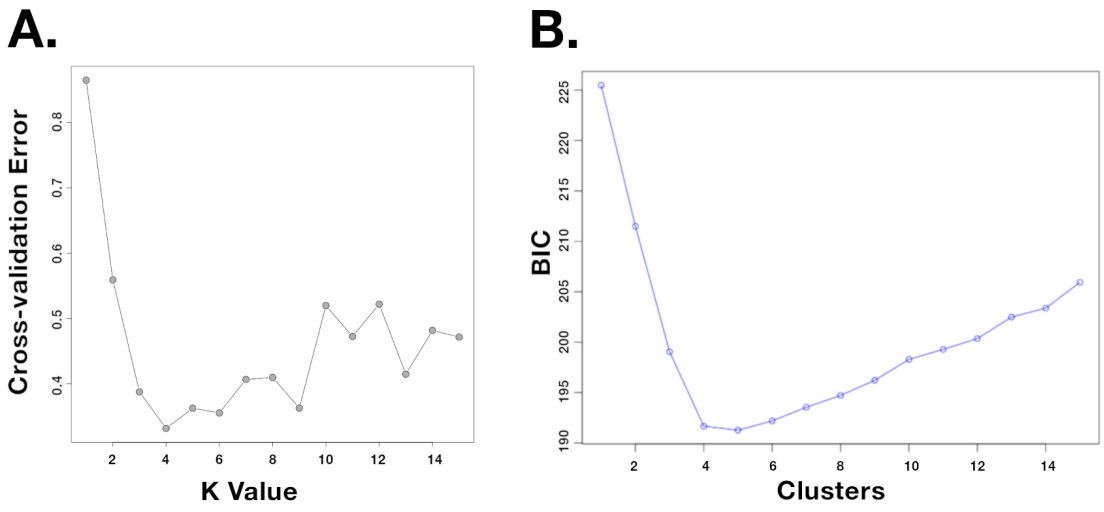


FIGURE S2 (A) Cross-validation error plot for `ADMIXTURE` analysis and (B) BIC plot for the DAPC analysis.



Supplementary Tables

TABLE S1. All sample identification information for individuals included in this study

Species	Field Number	Locality	State-Country
Rhinella centralis	ENS11296	Antioquia; Hotel El Lago	ANT - CO
Rhinella centralis	ENS11297	Hotel El Lago	ANT - CO
Rhinella centralis	ENS11298	Hotel El Lago	ANT - CO
Rhinella centralis	ENS11299	Hotel El Lago	ANT - CO
Rhinella dorbignyi	AF905	Uruguay	UY
Rhinella dorbingnyi	AF904	Uruguay	UY
Rhinella granulosa	AF909	Barra de Mamanguape	PB - BR
Rhinella granulosa	AF910	Barra de Mamanguape	PB - BR
Rhinella granulosa	MRT115	Pacoti	CE - BR
Rhinella granulosa	MTR12007	Linhares, Reserva da Companhia Vale do Rio Doce	ES - BR
Rhinella granulosa	MTR12026	Linhares, Reserva da Companhia Vale do Rio Doce	ES - BR
Rhinella granulosa	MTR12028	Linhares, Reserva da Companhia Vale do Rio Doce	ES - BR
Rhinella granulosa	MTR12059	Floresta Nacional de Goytacazes, Linhares	ES - BR
Rhinella granulosa	MTR12175	Floresta Nacional de Goytacazes, Linhares	ES - BR
Rhinella granulosa	MTR16052	Serra Bonita, Camacan	BA - BR
Rhinella granulosa	MTR16053	Serra Bonita, Camacan	BA - BR
Rhinella granulosa	MTR16054	Serra Bonita, Camacan	BA - BR
Rhinella granulosa	MTR16264	Serra Bonita, Camacan	BA - BR
Rhinella granulosa	MTR17662	Parque Estadual do Rio Doce, Marliéria	MG - BR
Rhinella granulosa	MTR17663	Parque Estadual do Rio Doce, Marliéria	MG - BR
Rhinella granulosa	MTR17665	Parque Estadual do Rio Doce, Marliéria	MG - BR
Rhinella granulosa	MTR22936	FLONA Contendas do Sincorá	BA - BR
Rhinella granulosa	MTR22988	FLONA Contendas do Sincorá	BA - BR
Rhinella granulosa	MTR22996	FLONA Contendas do Sincorá	BA - BR
Rhinella granulosa	MTR23002	FLONA Contendas do Sincorá	BA - BR
Rhinella granulosa	MTR23054	Serra da Maroquinha	RR - BR
Rhinella granulosa	MTR23531	Parque Nacional da Serra da Capivara	PI - BR
Rhinella granulosa	MTR23536	Parque Nacional da Serra da Capivara	PI - BR
Rhinella granulosa	MTR23547	Parque Nacional da Serra da Capivara	PI - BR
Rhinella granulosa	MTR23597	Parque Nacional da Serra da Capivara	PI - BR
Rhinella granulosa	MTR27075	Santa Maria da Vitória	BA - BR
Rhinella granulosa	PEU492	Ilhéus	BA - BR
Rhinella horribilis	ENS09806	Matagalpa	MT - NI
Rhinella horribilis	ENS11005	El Limón	AR - VE
Rhinella horribilis	ENS11061	Clarinas	AN - VE
Rhinella horribilis	ENS8661	El Paraíso	EP - HN
Rhinella horribilis	ENS8678	Atlántida	AT - HN
Rhinella horribilis	JAC18795	Izabal	IZ - GT
Rhinella humbolti	ENS11089	Guarico	AR - VE
Rhinella humbolti	ENS11242	Roscio	BO - VE
Rhinella jimi	MTR19863	Andaraí	BA - BR
Rhinella jimi	MTR26923	São Desidério	BA - BR
Rhinella major	BM377	Altamira	PA - BR
Rhinella major	BM395	Altamira	PA - BR
Rhinella major	BM410	Altamira	PA - BR
Rhinella major	H3058	UHE Jirau, Mutum	RO - BR
Rhinella major	H3078	UHE Jirau, Mutum	RO - BR

Rhinella major	H4920	UHE Jirau, Mutum	RO - BR
Rhinella major	H4930	UHE Jirau, Mutum	RO - BR
Rhinella major	LSUMZH-15118	Alter do Chao	PA - BR
Rhinella major	LSUMZH-15119	Alter do Chao	PA - BR
Rhinella major	LSUMZH-15120	Alter do Chao	PA - BR
Rhinella major	LSUMZH-15121	Alter do Chao	PA - BR
Rhinella major	LSUMZH-15122	Alter do Chao	PA - BR
Rhinella merianae	LSUMZH-12458	Fazenda Nova Esperanca	RR - BR
Rhinella merianae	LSUMZH-12483	Caracaraí	RR - BR
Rhinella merianae	LSUMZH-12485	Caracaraí	RR - BR
Rhinella merianae	LSUMZH-12488	Caracaraí	RR - BR
Rhinella merianae	LSUMZH12484	Caracarai	RR - BR
Rhinella merianae	LSUMZH12486	Caracarai	RR - BR
Rhinella merianae	LSUMZH12487	Caracarai	RR - BR
Rhinella merianae	LSUMZH12489	Caracarai	RR - BR
Rhinella merianae	LSUMZH12491	Caracarai	RR - BR
Rhinella merianae	LSUMZH12492	Caracarai	RR - BR
Rhinella merianae	MTR18590	Lago Chaviana, Itapuru, Rio Purus	AM - BR
Rhinella merianae	MTR18593	Lago Chaviana, Itapuru, Rio Purus	AM - BR
Rhinella merianae	MTR18594	Lago Chaviana, Itapuru, Rio Purus	AM - BR
Rhinella merianae	MTR18602	Lago Chaviana, Itapuru, Rio Purus	AM - BR
Rhinella merianae	MTR20516	Estação Ecológica Maracá	RR - BR
Rhinella merianae	MTR20519	Estação Ecológica Maracá	RR - BR
Rhinella merianae	MTR23032	Serra da Maroquinha	RR - BR
Rhinella merianae	MTR23148	Serra da Maroquinha	RR - BR
Rhinella merianae	MTR23210	Serra da Maroquinha	RR - BR
R. mirandaribeiroi	AF725	APM Manso	MT - BR
R. mirandaribeiroi	AF745	APM Manso	MT - BR
R. mirandaribeiroi	AF746	APM Manso	MT - BR
R. mirandaribeiroi	LAJ406	UHE Lajeado	TO - BR
R. mirandaribeiroi	MRT2120	Uruçuí	PI - BR
R. mirandaribeiroi	MRT2716	Uruçuí	PI - BR
R. mirandaribeiroi	MRT2917	Uruçuí	PI - BR
R. mirandaribeiroi	MRT7135	Guaraí	TO - BR
R. mirandaribeiroi	MRT7142	Guaraí	TO - BR
R. mirandaribeiroi	MRT8973	UHE Lajeado	TO - BR
R. mirandaribeiroi	MRT8997	UHE Lajeado	TO - BR
R. mirandaribeiroi	MRT9005	UHE Lajeado	TO - BR
R. mirandaribeiroi	MTR14850	Estação Ecológica Serra Geral do Tocantins	BA - BR
R. mirandaribeiroi	MTR14851	Estação Ecológica Serra Geral do Tocantins	BA - BR
R. mirandaribeiroi	MTR26991	Barreiras	BA - BR
R. mirandaribeiroi	MTR27049	Santa Maria da Vitória	BA - BR
R. mirandaribeiroi	PHV1431	São Desidério	BA - BR
R. mirandaribeiroi	PHV1432	São Desidério	BA - BR
R. mirandaribeiroi	PHV1433	São Desidério	BA - BR
R. mirandaribeiroi	PHV1434	São Desidério	BA - BR
R. mirandaribeiroi	PHV1461	São Desidério	BA - BR
R. mirandaribeiroi	PHV1525	Jaborandi, RVS Veredas	BA - BR
R. mirandaribeiroi	PHV1965	Jaborandi	BA - BR
R. mirandaribeiroi	PHV1966	Jaborandi	BA - BR
R. mirandaribeiroi	PHV2203	Estação Ecológica Serra Geral do Tocantins	BA - BR
R. mirandaribeiroi	PHV2376	Água Boa	MT - BR

R. mirandaribeiroi	PHV2424	Água Boa	MT - BR
R. mirandaribeiroi	PHV4041	Niquelândia, Acaba Vida	GO - BR
R. mirandaribeiroi	PHV4149	Monte Alegre de Goiás, Serra da Prata	GO - BR
Rhinella poeppigii	BM325	Vitória do Xingu	PA - BR
Rhinella poeppigii	BM577	Altamira	PA - BR
Rhinella poeppigii	LSUMZH-13700	Porto Walter	AC - BR
Rhinella schneideri	JC1366	Cristália	MG - BR
Rhinella schneideri	JC1367	Cristália	MG - BR
Rhinella schneideri	JC1378	Cristália	MG - BR
Rhinella schneideri	MTR34075	E.E. Gregório Bondar, Barrolândia	BA - BR
Rhinella schneideri	MTR34076	E.E. Gregório Bondar, Barrolândia	BA - BR
R. castaneoteca	MTR36558	São Pedro, Rio Içá	AM - BR
R. castaneoteca	MTR36559	São Pedro, Rio Içá	AM - BR
R. castaneoteca	MTR36560	São Pedro, Rio Içá	AM - BR
R. margaritifera	MRT7240	Guaraí	TO - BR
Rhinella marina	H3129	UHE Jirau, Mutum	RO - BR
Rhinella marina	H3314	UHE Jirau	RO - BR
Rhinella marina	H3332	UHE Jirau	RO - BR
Rhinella marina	H3334	UHE Jirau	RO - BR
Rhinella marina	H3524	UHE Jirau	RO - BR
Species	**Genbank Accession Number**		
Rhinella azarai	KP685186		
Rhinella bernardoi	KP685193		
Rhinella bernardoi	KP685194		
Rhinella centralis	KP685195		
Rhinella centralis	KP685196		
Rhinella dorbignyi	KP685197		
Rhinella dorbignyi	KP685198		
Rhinella dorbignyi	KP685199		
Rhinella fernandazae	KP685202		
Rhinella fernandezae	KP685200		
Rhinella fernandezae	KP685201		
Rhinella fernandezae	KP685203		
Rhinella fernandezae	KP685204		
Rhinella granulosa	KP685205		
Rhinella granulosa	KP685206		
Rhinella granulosa	KP685207		
Rhinella granulosa	KP685208		
Rhinella granulosa	KP685209		
Rhinella humbolti	KP685210		
Rhinella humbolti	KP685211		
Rhinella merianae	KP685220		
Rhinella merianae	KP685221		
Rhinella merianae	KP685222		
R. mirandaribeiroi	KP685223		
R. mirandaribeiroi	KP685224		
R. mirandaribeiroi	KP685225		
R. mirandaribeiroi	KP685226		
R. mirandaribeiroi	KP685227		
R. mirandaribeiroi	KP685228		
Rhinella pygmaea	KP685229		

TABLE S2. D-statistics. Grey values are not significant

P1	P2	P3	D-statistic	Z-score	p-value	f4-ratio	BBAA	ABBA	BABA
miranda-ribeiroi	granulosa	major	0.41	2.46	0.00688	0.15	14.99	7.77	3.24
humbolti	centralis	major	0.54	1.48	0.07	0.04	26.71	1.25	0.38
granulosa	miranda-ribeiroi	humbolti	0.22	1.21	0.11	0.08	12.65	5.56	3.52
centralis	humbolti	merianae	0.47	1.12	0.13	0.15	2.79	1.71	0.63
humbolti	merianae	major	0.27	1.08	0.14	0.03	27.19	2.01	1.16
granulosa	miranda-ribeiroi	merianae	0.19	1.05	0.15	0.06	13	5.94	4.05
centralis	merianae	miranda-ribeiroi	0.29	0.94	0.17	0.04	20.62	1.41	0.77
granulosa	miranda-ribeiroi	centralis	0.19	0.93	0.18	0.06	10.71	5.33	3.66
major	granulosa	humbolti	0.17	0.78	0.22	0.07	8.29	7.08	5.05
major	granulosa	merianae	0.14	0.72	0.23	0.05	8.49	6.64	4.97
major	granulosa	centralis	0.16	0.71	0.24	0.06	8.65	6.55	4.76
centralis	merianae	granulosa	0.22	0.67	0.25	0.03	22.41	1.42	0.9
humbolti	merianae	granulosa	0.18	0.64	0.26	0.03	25.27	1.82	1.26
humbolti	merianae	miranda-ribeiroi	0.16	0.62	0.27	0.02	22.77	1.85	1.35
miranda-ribeiroi	merianae	major	0.07	0.4	0.35	0.03	10.6	7.03	6.07
miranda-ribeiroi	centralis	major	0.04	0.22	0.41	0.02	10.18	6.73	6.18
miranda-ribeiroi	humbolti	major	0.03	0.19	0.42	0.02	10.47	6.4	5.97
centralis	merianae	major	0.06	0.17	0.43	0.01	24.66	1.53	1.36
humbolti	centralis	granulosa	0.02	0	nan	0	24.16	0.5	0.48
humbolti	centralis	miranda-ribeiroi	0.58	0	nan	0.02	21.84	0.46	0.12

*not significant after BH correction

TABLE S3. Fbranch statistics.

Branch b	Branch C	fb
granulosa	major	0.15
granulosa	merianae	0
granulosa	centralis	0
granulosa	humbolti	0
mirandaribeiroi	major	0
mirandaribeiroi	merianae	0
mirandaribeiroi	centralis	0
mirandaribeiroi	humbolti	0

*not significant after BH correction

major	granulosa	0
major	mirandaribeiroi	0
C	granulosa	0
C	mirandaribeiroi	0
D	granulosa	0
D	mirandaribeiroi	0
D	major	0
merianae	granulosa	0
merianae	mirandaribeiroi	0
merianae	major	0
centralis	granulosa	0
centralis	mirandaribeiroi	0
centralis	major	0
centralis	merianae	0
humbolti	granulosa	0
humbolti	mirandaribeiroi	0
humbolti	major	0
humbolti	merianae	0

Table S4. G-PhoCS results

	θ granulosa	**θ merianae**	**θ humbolti**	**θ centralis**
mean	4.61E-03	4.26E-03	3.00E-03	1.40E-03
stderr of mean	1.57E-06	1.15E-06	1.97E-06	8.38E-07
stdev	1.11E-04	9.71E-05	1.97E-04	9.27E-05
variance	1.22E-08	9.44E-09	3.87E-08	8.59E-09
median	4.61E-03	4.26E-03	3.00E-03	1.40E-03
geometric mean	4.61E-03	4.26E-03	3.00E-03	1.40E-03
95% HPD Interval	[4.38E-3, 4.81E-3]	[4.05E-3, 4.43E-3]	[2.62E-3, 3.39E-3]	[1.21E-3, 1.57E-3]
auto-correlation time (ACT)	324.5766	227.4016	161.8357	132.4802
effective sample size (ESS)	4991.1214	7123.966	10010.1589	12228.2504
Ne	**479,979**	**443,573**	**312,677**	**145,854**
	θ miranda-ribeiroi	**θ major**	**θ granulosa+ miranda-ribeiroi**	**θ humbolti+ centralis+ merianae**
mean	0.0105	4.65E-03	0.0127	3.07E-03
stderr of mean	2.74E-06	9.95E-07	3.11E-06	1.63E-06
stdev	2.17E-04	8.17E-05	2.76E-04	1.43E-04
variance	4.71E-08	6.67E-09	7.64E-08	2.03E-08

median	0.0105	4.65E-03	0.0127	3.07E-03	
geometric mean	0.0105	4.65E-03	0.0127	3.07E-03	
95% HPD Interval	[0.0101, 0.0109]	[4.5E-3, 4.81E-3]	[0.0122, 0.0132]	[2.79E-3, 3.34E-3]	
auto-correlation time (ACT)	257.0347	240.626	205.4787	211.3199	
effective sample size (ESS)	6302.654	6732.4441	7884.0354	7666.1064	
Ne	**1,093,750**	**484,448**	**1,322,917**	**319,760**	
	θ humbolti+ centralis+ merianae+ major		**θ root**		
mean	0.0896		0.0125		
stderr of mean	1.35E-03		3.14E-06		
stdev	0.0509		2.59E-04		
variance	2.59E-03		6.69E-08		
median	0.079		0.0125		
geometric mean	0.0761		0.0125		
95% HPD Interval	[0.0106, 0.1898]		[0.012, 0.013]		
auto-correlation time (ACT)	1142.4839		238.6483		
effective sample size (ESS)	1417.964		6788.2357		
Ne	**9,333,333**		**1,302,083**		

	τ granulosa-miranda-ribeiroi	**τ humbolti-centralis**	**τ humbolti-centralis-merianae**	**τ humbolti-centralis-merianae-major**	**τ root**
mean	1.23E-03	1.80E-03	2.72E-03	5.79E-03	5.95E-03
stderr of mean	5.26E-07	1.73E-06	5.35E-07	1.49E-06	1.80E-06
stdev	2.87E-05	1.48E-04	6.30E-05	1.11E-04	8.43E-05
variance	8.25E-10	2.18E-08	3.96E-09	1.23E-08	7.11E-09
median	1.23E-03	1.79E-03	2.72E-03	5.80E-03	5.95E-03
geometric mean	1.23E-03	1.79E-03	2.71E-03	5.79E-03	5.95E-03
95% HPD Interval	[1.18E-3, 1.29E-3]	[1.51E-3, 2.08E-3]	[2.59E-3, 2.83E-3]	[5.55E-3, 5.98E-3]	[5.78E-3, 6.11E-3]
auto-correlation time (ACT)	542.7848	221.4774	116.81	290.95	736.4447
effective sample size (ESS)	2984.6101	7314.5221	13868.6818	5567.97	2199.7593
T (years)					
mean	**1,026,083**	**1,496,167**	**2,263,167**	**4,826,083**	**4,957,833**
stderr of mean	880	2,739	882	2,344	3,019
stdev	24,062	123,367	52,121	92,475	71,602
variance	1	18	3	10	6
median	1,025,000	1,491,667	2,266,667	4,833,333	4,958,333
geometric mean	1,025,833	1,491,083	2,262,583	4,825,167	4,957,250
95% HPD Interval start	983,333	1,258,333	2,158,333	4,625,000	4,816,667

95% HPD Interval end	1,075,000	1,733,333	2,358,333	4,983,333	5,091,667

	Data-ld-ln	Full-ld-ln
mean	30.0265	-769522.8244
stderr of mean	6.47E-03	2.3346
stdev	0.558	186.5822
variance	0.3113	34812.9278
median	30.0244	-769522.4406
geometric mean	30.0213	n/a
95% HPD Interval	[28.9318, 31.1206]	[-769889.0853, -769158.2312]
auto-correlation time (ACT)	217.8016	253.6337
effective sample size (ESS)	7437.9668	6387.1678

μ	**θ**	**τ**
2.40E-09	4Neμ	Tμ/g

Chapter III: Comparative phylogeography and co-demographic change across the Neotropics

Abstract

Identifying the evolutionary and ecological mechanisms responsible for the origin and persistence of biodiversity has played a central role in our understanding of both local and global diversification patterns. Most hypotheses proposed to explain spatial biodiversity patterns in the Neotropics have called upon changes in climate and landscape change as key drivers of species dispersal, range limits, lineage divergence, and speciation. Here we utilize co-distributed Neotropical clades of reptiles and amphibians to investigate the contribution of landscape features to population divergence and demographic change utilizing a comparative phylogeographic approach. Using genome-wide loci from species that span large geographic areas and multiple biomes across the Neotropics, we test alternative scenarios of shared evolutionary history by estimating the timing and magnitude of demographic change across populations of multiple sympatric taxa. While our analyses did support several instances of temporally synchronous demographic events among co-distributed taxa which would point to shared responses, the majority of demographic events which clustered together in time suggested that they have not been driven by the same mechanisms, such as the action of new topographic barriers or climate-driven habitat shifts. Instead, they appear to have happened largely independently of one another, and perhaps have been driven by more species-specific mechanisms. Our results suggest that an initial focus on geographic or environmental mechanisms driving diversification on a continental scale, such as glaciation cycles, may have led to an overestimation of how congruent the evolutionary trajectories of sympatric species are.

Introduction

The world's most diverse ecological communities are concentrated in the tropics (Myers, Mittermeier, Mittermeier, da Fonseca, & Kent, 2000). Identifying the evolutionary and ecological mechanisms responsible for the origination and persistence of this rich biodiversity has played a central role in our understanding of both local and global diversification patterns. Understanding what factors promote lineage persistence over evolutionary time, as well as the accumulation of evolutionary potential in geographic space, is key for the conservation of nature (Carnaval, et al., 2009; Oaks, 2019). The Neotropical region houses some of the world's most diverse and threatened ecosystems, but the historical and contemporary processes that have led to its high species richness and endemism remain relatively poorly known (Carnaval et al., 2009, 2014). Most hypotheses proposed to explain spatial biodiversity patterns in the Neotropics have invoked landscape configuration and change as key drivers of dispersal (or limitation), lineage divergence, and speciation (Carnaval et al., 2014; Dal Vechio, et al., 2019; Prates, Rivera, et al., 2016; Rivera, et al., 2020). These hypotheses have often been applied to explain current species distributions patterns and assemblage composition in other regions, becoming central to biogeographic investigations worldwide (Leaché et al., 2019; Potter et al., 2019).

Multiple hypotheses have been proposed to explain how biodiversity has been generated and maintained across the Neotropics. Among them, one of the earliest and most influential is the theory of Pleistocene refugia. This hypothesis posits that the Amazonian rainforest as currently recognized was fragmented into multiple forest refugia during drier Pleistocene glacial periods, which promoted lineage divergence and speciation in isolated forested areas of higher stability, or refugia (Haffer, 1969; Haffer & Prance, 2001). Then, during interglacial periods, these forest fragments would expand, allowing species to increase their range and come into

contact with previously conspecific populations that had differentiated into new species (Haffer & Prance, 2001). This theory has since been expanded to propose similar climatically stable refugia promoting species accumulation outside of Amazonia, including other rainforest systems as well as open and drier habitat types. At times, studies invoking refugia have involved opposite demographic patterns, such as the expansion of montane species into the lowlands during glacial periods (Carnaval et al., 2009; Fenker et al., 2020; Fjeldsa, et al., 1999; Fontanella, et al., 2012; Werneck, et al., 2012).

Another factor frequently employed to explain the origin and maintenance of biodiversity in this region is geomorphological (and other landscape) change. For instance, the establishment and uplift of mountain ranges tied to tectonic plate activity have been hypothesized to lead to dispersal limitation, disruption of gene flow, and lineage divergence, leading to high species richness and endemism. For instance, this mechanism has been proposed to explain biodiversity patterns in the Andean Mountain chain and Brazil's southeastern Atlantic Forest (Brown & Twomey, 2009; Carnaval et al., 2014). Another hypothesis invoking landscape-driven changes proposes that the establishment of fluvial systems in South America -- in the form of both lacustrine environments and large rivers -- has historically led to genetic differentiation and speciation across banks while also preventing the homogenization of species pools in the present day (Hoorn et al., 2010; Ribas, et al., 2012; Thomé et al., 2010; Werneck, et al., 2015). While the actions of topographic and hydrographic barriers are not mutually exclusive, few studies have tried to address the contribution of both types of features in the same focal clade (Dal Vechio et al., 2019).

If geomorphological changes are important processes affecting species range limits and gene flow levels, we may expect entire species assemblages to be similarly affected by major

features such as mountain uplifts and river size fluctuations. In that case, co-distributed taxa may exhibit similar patterns of spatial genetic structure as determined by the action of a common environmental factor, such as a geographic barrier. This premise motivated the establishment of the field of phylogeography (Avise, 2009; Hickerson et al., 2010). While initially restricted to a description of concordant (or discordant) spatial phylogenetic patterns based on gene genealogies, comparative phylogeography has evolved to incorporate approaches aiming to infer shared patterns of demographic histories in response to habitat shifts, for instance by modeling synchronous pulses of population size change across taxa (Chan, et al., 2014; Xue & Hickerson, 2017). Because the climatic regimes that affect biome distributions and the demography of associated species frequently span large geographic areas in the Neotropics, taxa that occur in different regions -- even separated by thousands of kilometers -- may show congruent demographic change through time (Prates, Xue, et al., 2016; Xue & Hickerson, 2020). By revealing if sets of co-distributed species or populations co-diverged, co-expanded, or co-contracted, these emergent approaches have the potential to improve our knowledge of how organisms have responded to past climate and landscape change (Avise et al., 1987; Carnaval et al., 2009), as well as inform predictions of the future distribution and genetic diversity of species (Prates, Xue, et al., 2016).

In this investigation, we focus on co-distributed Neotropical clades to investigate the contribution of landscape features to population divergence and demographic change on the basis of a comparative phylogeographic approach. Using genome-wide loci from four clades of amphibians and reptiles that span large geographic areas and multiple biomes, we test alternative scenarios of shared evolutionary history by estimating the timing and magnitude of demographic change across populations of multiple sympatric taxa. Specifically, we explore whether and how

population trajectories have been affected by three major environmental factors: riverine barriers, mountain barriers, and climatic zones. To assess to what extent these factors may have led to assemblage-level versus species-specific idiosyncratic responses, we quantify the degree of synchronicity across taxa in population divergence and size changes, as well as levels of overlap in spatial patterns of genetic structure. Our analysis supports several instances of temporal concordance in the demographic history of co-distributed Neotropical taxa, illustrating the value of comparative phylogeography to our understanding of how landscape changes through time have contributed to present-day patterns of biodiversity.

Material and Methods

Sample collection

To infer patterns of co-demographic change across both closely and distantly related species, we focus on two clades of toads and two of lizards. We incorporated data from the *Rhinella marina* and *Rhinella granulosa* species complexes of toads generated by previous studies of these clades (see Chapters I and II of this Dissertation). Moreover, we sampled the lizard clades *Mabuya* (including *M. altamazonica* and the *M. nigropunctata* species complex) (Scincidae) and the *Gymnodactylus* genus (Phyllodactylidae). For the newly generated *Mabuya* dataset, we extracted whole genomic DNA from 131 samples belonging to *Mabuya altamazonica, M. nigropunctata*, *M. surinamensis*, and multiple *Mabuya* populations of uncertain taxonomic identity (see Results), as well as three *Mabuya frenata* specimens as outgroups. For the newly generated *Gymnodactylus* dataset, we sampled and extracted whole genomic DNA from 101 samples belonging to *G. amarali, G. geckoides, G. darwinii, G.*

guttulatus, G. vanzonini, and one *Gymnodactylus* population of unclear identity (see Results), as well as two specimens of *Thecadactylus rapicauda* as outgroups. All tissues were obtained from the Museum of Zoology of the University of São Paulo (MZUSP), the Amphibian and Reptile Diversity Research Center (ARDRC) at the University of Texas in Arlington, and the Louisiana State University Museum of Natural Science (LSUMNS).

DNA extraction, amplification, & sequencing

We extracted genomic DNA from the *Mabuya* and *Gymnodactylus* samples using standard protocols (Sambrook & Russell, 2006). DNA extractions were submitted to the Texas A&M AgriLife Genomic and Bioinformatics Service for library preparation and sequencing. Double-digest restriction-site associated DNA sequencing (ddRADseq) sample libraries were prepared using the *PstI* and *MspI* restriction enzymes and size-selected at 400-550 bp. The resulting 150 bp paired-end libraries were then sequenced using the Illumina NovaSeq S2. Sequence cluster identification, quality prefiltering, base calling, and uncertainty assessment were done in real time using Illumina's NCS 1.0.2 and RFV 1.0.2 software with default parameter settings. Sequencer *cbcl* basecall files were demultiplexed and formatted into *fastq* files using the bcl2fastq 2 2.19.0 script *configureBclToFastq.pl* (Identification and Accession numbers in Supplementary Table S1).

Phylogenetic Reconstruction

For the new data generated in this study, we used the command line version of ipyrad v. 0.9.45 (Eaton & Overcast, 2020) to merge reads and perform de novo read assembly (using a minimum clustering similarity threshold of 0.90), align the reads into loci, and call single

nucleotide polymorphisms (SNPs). A minimum Phred quality score (= 33), sequence coverage (= 6x), read length (= 35 bp), and maximum proportion of heterozygous sites per locus (= 0.5) were enforced, while ensuring that variable sites had no more than two alleles (i.e., a diploid genome). Following the initial assembly, we used Matrix Condenser (de Medeiros & Farrell, 2018) to assess levels of missing data across samples and then re-assembled our dataset. For genus-level assemblies of *Mabuya* and *Gymnodactylus* datasets, we enforced no more than 30% missing data across the entire dataset, which included up to 70% missing data for some samples (including outgroups).

To characterize patterns of genetic structure that may be indicative of species boundaries, we inferred maximum likelihood phylogenies based on the ddRADseq data using IQTREE v2.1.2 using the built-in model selection tool ModelFinder Plus and implementing 1000 ultrafast bootstraps (Hoang, et al., 2018; Kalyaanamoorthy, et al., 2017; Minh et al., 2020). We employed the greedy algorithm of PartitionFinder for model selection, only testing models of evolution available in MrBayes (Lanfear, et al., 2017). To further delimit units, we implemented a genetic clustering approach based on the SNP data. To this purpose, we filtered SNPS following the steps described above (but excluding outgroups) for each genus using ipyrad and VCFtools (Danecek et al., 2011). We then used the maximum likelihood genetic clustering method ADMIXTURE by testing the best-fit number of genetic populations (K) from one to 20 populations with 20 replicates per K and a 10-fold cross-validation to assess model support (Alexander, et al., 2009). The best K was determined based on the replicate with the lowest cross-validation error.

Ecoevolity

To test for congruent patterns of population divergence across taxa, as well as estimate the timing of potentially synchronous population divergences, we used the full-likelihood method *ecoevolity* based on the SNP data (Bryant, et al., 2012; Oaks, 2019). After assessing phylogeographic breaks across species within the *Rhinella granulosa* complex, *Rhinella marina* complex, *Mabuya*, and *Gymnodactylus*, we chose population or species pairs that conformed to phylogeographic divergence across major environmental or geographic barriers, such as cross-biome or river divergence. Ultimately, we tested 16 different comparisons across taxa, which included: 1-3) *Mabuya surinamensis*, *M. altamazonica*, and *Rhinella marina*, which shared a pattern of intra-taxon population divergences across the Amazon River (north and south); 4-6) *Mabuya sp. I*, *M. surinamensis*, and *Rhinella poeppigii*, which shared a pattern of population divergence between eastern and western Amazonia, corresponding to the location of two major climatic systems that act in this region (Cheng et al., 2013; Prates, Xue, et al., 2016); 7) *Rhinella horribilis*, which showed a genetic break between the northern Andes and Central America; 8) *Rhinella merianae* and *R. centralis* + *R. humbolti*, which shared a genetic break between Amazonia and the northern Andes; 9-10) *Mabuya sp. I* and the *Mabuya sp. II* + *M. surinamensis*, which shared a genetic break between Amazonia and the Cerrado; 11-13) *R. schneideri*, *Rhinella mirandaribeiroi* + *R. granulosa*, and *Gymnodactylus darwinii* populations, which shared a break between the Cerrado and the Atlantic Forest; 14) *Mabuya sp. II*, which showed a genetic break between the Caatinga and the Pampas; 15) and *G. darwinii* populations, which showed genetic breaks between the Caatinga and northern Atlantic Forest, as well as 16) between the northern Atlantic Forest and southern Atlantic Forest. Maps and phylogenies for all sample comparisons were generated using QGIS (QGIS Development Team 2020. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org).

As *ecoevolity* does not require loci to be shared across taxa, we then generated assemblies for each population or species pair to maximize the number of loci used, which resulted in datasets with 0-15% missing data (Table S2). *Ecoevolity* also assumes no gene flow, therefore we excluded any admixed individuals (based on the clustering analyses) from the analyses. After testing a range of priors in preliminary analyses, we set the concentration prior for the number of divergent events to 5, using a gamma distribution with a shape = 10.0 and a prior mean number of events = 8.0. For the event time prior, we used an exponential distribution with rate of 1000 and set the population size prior to 0.002 with a gamma distribution with shape = 5.0 and scale = 0.0004. We ran the analysis twice independently, sampling every 100 steps for 100,000 iterations. We assessed convergence between runs using Tracer, combined the two runs, and discarded the first 10% of samples before using *pycoevolity* to process and plot the results. We used Bayes factors to determine support for the mean number of divergence events.

Results

Delimitation of coherent genetic lineages for downstream comparative analyses

Phylogenetic relationships for the *Mabuya* lizards were generally highly supported (Fig. 1). *Mabuya altamazonica*, which is composed of two highly supported clades, is sister to a clade composed of the other taxa in the *M. nigropunctata* species complex. Within this clade, three major clades were inferred. The first clade (the Occidental Clade) is composed of samples assigned to *M. nigropunctata* and sister to a clade composed of all the remaining species. The second major clade (the Meridional Clade) is composed of samples assigned to an unnamed species, which we refer to as *Mabuya sp. I*, as originally described by Miralles and Carranza

(Miralles & Carranza, 2010; Pinto-Sánchez, et al., 2015). The third major clade (the Oriental Clade) is composed of samples assigned to *M. surinamensis* and another unnamed putative species, which we refer to as *Mabuya sp. II* (Fig. 1).

Genetic clustering analysis of the *Mabuya* clade generally showed very little admixture between clusters (Fig. 1). *Mabuya altamazonica* was composed of two clusters, one of which is restricted to areas north of the Amazon River in northwestern Amazonia, and another occurs in southern Amazonia. *Mabuya nigropunctata* corresponded to a single genetic cluster distributed across southern Amazonia. *Mabuya sp. I* was composed of three clusters with very little admixture; one cluster was distributed in southwestern Amazonia, while the other two were distributed in southeastern Amazonia and the Cerrado, respectively (Fig. 1; Fig. S1). *Mabuya surinamensis* was composed of four genetic clusters: one in southern Amazonia, one in southern and southeastern Amazonia, one in northern Amazonia, and one from a single locality in northeastern Amazonia. *Mabuya sp. II* was composed of a single genetic cluster distributed across the Seasonally Dry Tropical Forests of the Caatinga and Cerrado (Fig. 1; Fig S1). Each of these clusters corresponded to a clade in the phylogenetic analyses of *Mabuya*, with the exception of a subclade within *M. surinamensis* that was composed of samples admixed between the three genetic clusters inferred within this taxon (Fig. 1).

In the case of *Gymnodactylus*, phylogenetic analyses resulted in high support across nearly all nodes (Fig. 2). The resulting phylogeny split the genus into two major clades: one containing *G. amarali*, *G. geckoides*, and two potentially unnamed *Gymnodactylus* species, here referred to as *Gymnodactylus sp. I* and *Gymnodactylus sp. II*; and another major clade containing *Gymnodactylus vanzolinii*, *G. guttulatus*, *G. darwinii*, and an unnamed population here referred

to at *Gymnodactylus sp. III* (Fig. 2). *Gymnodactylus darwinii*, which was split into three subclades, is sister to *Gymnodactylus sp. III*. In turn, the clade formed by these two species is sister to *G. guttulatus*, and these species are sister to *G. vanzolinii*. *Gymnodactylus geckoides* was inferred as sister to *Gymnodactylus sp. II*, the two forming a clade that is sister to *Gymnodactylus sp. I*. Lastly, the clade formed by those three taxa is sister to *G. amarali* (Fig 2).

Genetic clustering analysis of the *Gymnodactylus* genus was composed of 13 genetic clusters, with very little genetic admixture (Fig. 2). *Gymnodactylus amarali* was represented by one genetic cluster, with only one sample showing a minor amount of admixture; this cluster was distributed across the Cerrado and ecotonal region abutting eastern Amazonia (Fig. 2; Fig. S2). By contrast, this analysis supported extensive substructuring in *G. darwinii*, splitting it into eight genetic clusters, with only one sample being admixed between two clusters (Fig. 2). These eight clusters are restricted to different geographic regions and corresponded to three major clades inferred within this taxon, as follows: two clusters occurring in the Caatinga and across parts of the Caatinga-Cerrado ecotone (grouped in the North clade); two clusters occurring in the northern Atlantic Forest (grouped in the Central Clade); and four clusters occurring in the central and southern Atlantic Forest (grouped in the South Clade) (Fig. S2). In turn, samples assigned to *G. geckoides*, *Gymnodactylus sp. I*, and *Gymnodactylus sp. II* formed a single cluster, with the *Gymnodactylus sp. I* sample showing admixture with the *G. amarali* cluster. Lastly, *G. vanzolinii* and *G. guttulatus* were represented by their own individual genetic clusters, as was one unnamed population, here referred to as *Gymnodactylus sp. III*, distributed in the central Atlantic Forest (Fig. 2).

Synchronicity of divergences across co-distributed taxa

Analyses of synchronous divergence using *Ecoevolity* suggested that multiple population pairs co-diverged across different barriers through time (Fig. 3). Bayes factors indicated that, among the 16 pairs analyzed, divergences were clustered in four time periods (hereafter "co-divergence events") (Fig. 3). Interestingly, all four inferred co-divergence events included population pairs that occur in different (and often distant) geographic regions, suggesting no cross-taxon clustering of divergences within geographic regions. Instead, the timing of divergences for any given region were largely idiosyncratic across taxa, as follows.

The oldest co-divergence event involved three population pairs: the split between *Gymnodactylus darwinii* populations from northern versus southern Atlantic Forest; the split between *Mabuya sp. I* populations from eastern Amazonia versus the Cerrado; and the split between *Mabuya sp. I* populations from eastern versus western Amazonia (Fig. 3).

The second co-divergence event involved eight population pairs: the split between *Rhinella merianae* in Amazonia and the cluster formed by *R. centralis* and *R. humbolti*, in the northern Andes; the split between *R. granulosa* in the Atlantic Forest and *R. mirandaribeiroi* in the Cerrado; the split between *R. horribilis* in the northern Andes versus Central America; the split between *Gymnodactylus darwinii* in the Caatinga versus the northern Atlantic Forest; the split between *Mabuya altamazonica* north versus south of the Amazon River; the split between *M. surinamensis* north versus south of the Amazon River; the split between *M. surinamensis* in eastern versus western Amazonia; and the split between *Mabuya sp. II* in the Cerrado and *M. surinamensis* in Amazonia (Fig. 3).

The third co-divergence event involved two additional population pairs: the split between *R. marina* north versus south of the Amazon River; and the split between *Mabuya sp. II* in the Cerrado versus the Pampas region.

Lastly, the fourth and most-recent co-divergence event involved three additional population pairs: the split between *R. poeppigii* in eastern versus western Amazonia; the split between *R. schneideri* in the Cerrado versus the Atlantic Forest; and the split between *Mabuya sp. II* in the Cerrado versus the Caatinga (Fig. 3).

Population size shifts across co-distributed taxa

Among the population pairs diverging across eastern versus western Amazonia, which included *Rhinella poeppigii*, *Mabuya sp. I*, and *M. surinamensis*, most populations experienced a reduction in population size (Table 1). The western Amazonian M. *surinamensis* population experienced a population increase and eastern Amazonian *Mabuya sp. I* population remained relatively stable after divergence (Table 1).

Among the population pairs diverging across the Atlantic Forest and the Cerrado or Caatinga, which included *R. schneideri*, the clusters formed by *R. granulosa* and *R. mirandaribeiroi*, and *G. darwinii*, most populations experienced population reductions, except for *Rhinella mirandaribeiroi*, which experienced a population increase (Table 1).
Among the population pairs diverging across Amazonia and the Cerrado, which included *Mabuya sp. I*, *Mabuya sp. II*, and *M. surinamensis*, only the *Mabuya sp. II* Cerrado population experienced population decrease, while the other populations either increased or remained relatively stable (Table 1).

Among the population pairs diverging across (north and south of) the Amazon River, which included *R. marina*, *M. altamazonica*, and *M. surinamensis*, all experienced population declines post-divergence (Table 1).

Among the population pairs diverging across the North versus South Atlantic Forest (*G. darwinii*), the Caatinga versus Cerrado (*Mabuya sp. II*), the Caatinga versus Pampas (*Mabuya sp. II*), and Amazonia versus the northern Andes (*R. merianae* and *R. centralis* + *R. humbolti*), all involved population size increases or remained relatively stable post-divergence (Table 1). Lastly, the Central American population of *R. horribilis* experienced a size increase post-divergence, while the northern Andes population of this taxon remained relatively stable (Table 1).

Discussion

Notes on Mabuya and Gymnodactylus systematics

Our delimitation of major genetic groups based on phylogenetic and genetic clustering analyses revealed high levels of potentially cryptic divergence within both *Mabuya* and *Gymnodactylus* lizards. The *Mabuya nigropunctata* species complex, like many other clades within the speciose and taxonomically challenging radiation of Neotropical skinks, has unexpectedly high levels of genetic diversity, particularly considering how little morphological variation exists across the complex (Hedges, et al., 2012; Miralles & Carranza, 2010; Pinto-Sánchez et al., 2015). Some of the patterns of potentially cryptic diversity that we recovered here were already detected by previous investigations (Miralles & Carranza, 2010; Pinto-Sánchez et al., 2015). This is the case of the Occidental, Meridional, and Oriental major clades inferred within this complex, whose distributions agree with patterns from previous analyses based on multi-locus datasets (Miralles & Carranza, 2010), except for *Mabuya sp. II*, which we found to have a broad distribution ranging across dry forest habitats (Fig. S1). We were unable to

associate this unnamed candidate species with any other population from the literature. In addition, the unnamed *Mabuya sp. I* corresponded to a highly divergent lineage. While the taxonomy of *Mabuya* has received some attention over the past decade, these efforts have done little to address issues of unrecognized diversity in South American species, focusing instead on Central American and Caribbean species. (Hedges et al., 2012). The case of *Gymnodactylus* lizards is similar; we inferred a large number of geographically restricted clades that display enormous amounts of genetic structure with very little admixture, as seen in *G. darwinii*. Previous analyses of *Gymnodactylus* have also found evidence of large-scale cryptic diversity (Fig. 2; Fig S2)(Cassimiro & Rodrigues, 2009; Domingos et al., 2014; Pellegrino et al., 2005). Taken together, our results suggest that these South American lizards warrant dedicated investigations to properly characterize species limits based on comprehensive assessment of genetic and morphological variation.

Concordant and discordant species histories

Analyses of synchronous population divergence or size change have largely found discordant demographic patterns across co-distributed taxa in most of the geographic regions considered (Fig. 3; Table 1). We initially designed our co-divergence analyses without limiting comparisons to the same region, instead integrating different regions in the same analytical framework. With that, we expected that the timing of population divergences or size changes would cluster across taxa in correspondence with the number of geographic regions. We did find demographic events to cluster in time, inferring four different periods of co-divergence (Fig. 3). However, each of these periods involved populations often separated by large geographic distances and in distinct biomes (Table 1).

As an example of these discordant patterns, the second co-divergence event inferred by our analyses (B) involved eight population pairs. Half of them experienced size decreases in both populations compared. By contrast, the remaining population pairs experienced relatively stable population sizes or large increases (Table 1). This co-divergent event involved a wide range of habitats -- e.g., different portions of Amazonia, the Andean mountains, the Atlantic Forest, the Cerrado savannas --, as well as taxa -- e.g., toads in both the *R. granulosa* and *R. marina* clades, *Mabuya* skinks, and *Gymnodactylus* lizards. Similar patterns were inferred for the other co-divergence events as well. Individually, each of the inferred demographic events could be ascribed to a potential aspect of the evolutionary history of a species; for instance, population size increases may reflect range expansions tied to founder events after the colonization of new areas, while divergences involving no population size shifts may reflect vicariant separation of populations (Dal Vechio et al., 2019; Dal Vechio, et al., 2018; Prates, Rivera, et al., 2016). Nevertheless, the fact that such different demographic events clustered together in time suggests that they have not been driven by the same mechanisms, such as the action of new topographic barriers or climate-driven habitat shifts. Instead, they appear to have happened largely independently from one another.

Despite this general pattern of idiosyncratic species histories, we did find some instances of consistent demographic events. For instance, population pairs of *Mabuya sp. I* consistently had stable population sizes across regions as different as eastern Amazonia, western Amazonia, and the Cerrado, suggesting that the divergences between these population pairs may all reflect vicariant events, where no population bottlenecks were involved (e.g., in a scenario of dispersal). These time-concordant events sometimes involved closely related taxa, as in the toads *R. poeppigii* and *R. schneideri* which showed concordant population declines that were clustered in

time, albeit in different geographic regions (Table 1). In other cases, population pairs of closely related taxa showed similar responses in the same region but at different times, as is the case of *Mabuya sp. I*, *Mabuya sp. II*, and *M. surinamensis* populations in Amazonia and the Cerrado. If these concordant patterns within taxa or between closely related taxa were inferred correctly, they may reflect similar propensities of species to respond to environmental change, potentially as determined by organismal attributes. The role of traits mediating environmental tolerances or capacity for dispersal on patterns of genetic structure and demographic trends have received increased attention (Fenker, et al., 2021; Papadopoulou & Knowles, 2016; Zamudio, Bell, & Mason, 2016). For instance, such traits have been invoked to explain similarities and differences in the responses of different co-distributed Neotropical lizards in the face of shared patterns of climatic change over time (Prates, Xue, et al., 2016).

Congruent patterns of population divergences across co-distributed temperate zone taxa, traditionally inferred on the basis of single mitochondrial loci, have inspired the field of phylogeography (Avise et al., 1987; Hewitt, 2000). However, the increasing availability of genome-wide loci has revealed highly incongruent patterns of genetic structure and population divergence across taxa, not only in the Neotropics but also in several other tropical regions (Potter et al., 2018). Our results suggest that single-locus studies, as well as an initial focus on geographic regions prone to major cyclical climatic events (such as glaciations), may have led to an overestimation of how congruent the evolutionary trajectories of sympatric species are. Alternatively, currently available methods for co-demographic inference may still need to be substantially improved to allow proper integration of the high levels of genealogical heterogeneity revealed by genomic-scale datasets.

References

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 19(9), 1655–1664.

Avise, J. C. (2009). Phylogeography: retrospect and prospect. Journal of Biogeography, 36(1), 3–15.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., … Saunders, N. C. (1987). Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annual Review of Ecology and Systematics, 18(1), 489–522.

Brown, J. L., & Twomey, E. (2009). Complicated histories: three new species of poison frogs of the genus Ameerega (Anura: Dendrobatidae) from north-central Peru. Zootaxa, 2049(1), 1–38.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Molecular Biology and Evolution, 29(8), 1917–1932.

Carnaval, A. C., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T., & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. Science, 323(5915), 785–789.

Carnaval, A. C., Waltari, E., Rodrigues, M. T., Rosauer, D., VanDerWal, J., Damasceno, R., … Moritz, C. (2014). Prediction of phylogeographic endemism in an environmentally complex biome. Proceedings. Biological Sciences / The Royal Society, 281(1792). doi: 10.1098/rspb.2014.1461

Cassimiro, J., & Rodrigues, M. T. (2009). A new species of lizard genus Gymnodactylus Spix, 1825 (Squamata: Gekkota: Phyllodactylidae) from Serra do Sincorá, northeastern Brazil, and the status of G. carvalhoi Vanzolini, 2005. Zootaxa, 2008, 38–52.

Chan, Y. L., Schanzenbach, D., & Hickerson, M. J. (2014). Detecting concerted demographic response across community assemblages using hierarchical approximate Bayesian computation. Molecular Biology and Evolution, 31(9), 2501–2515.

Cheng, H., Sinha, A., Cruz, F. W., Wang, X., Edwards, R. L., d'Horta, F. M., … Auler, A. S. (2013). Climate change patterns in Amazonia and biodiversity. Nature Communications, 4, 1411.

Dal Vechio, F., Prates, I., & Grazziotin, F. G. (2019). Rain forest shifts through time and riverine barriers shaped the diversification of South American terrestrial pit vipers (Bothrops jararacussu species group). Journal of. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/jbi.13736?casa_token=b0jGobWgaScAAAAA:OlZrsqToibNROejKiJX_H6XRGvIJEQxbwHryE9YLUJGHDW_05ADo1e49kZJCvdDZBXKj52ZdJVX8xQ

Dal Vechio, F., Prates, I., Grazziotin, F. G., Zaher, H., & Rodrigues, M. T. (2018). Phylogeography and historical demography of the arboreal pit viper Bothrops bilineatus (Serpentes, Crotalinae) reveal multiple connections between Amazonian and Atlantic rain forests. Journal of Biogeography, 45(10), 2415–2426.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. Bioinformatics , 27(15), 2156–2158.

de Medeiros, B. A. S., & Farrell, B. D. (2018). Whole-genome amplification in double-digest RADseq results in adequate libraries but fewer sequenced loci. PeerJ, 6, e5089.

Domingos, F. M. C. B., Bosque, R. J., Cassimiro, J., Colli, G. R., Rodrigues, M. T., Santos, M. G., & Beheregaray, L. B. (2014). Out of the deep: cryptic speciation in a Neotropical gecko (Squamata, Phyllodactylidae) revealed by species delimitation methods. Molecular Phylogenetics and Evolution, 80, 113–124.

Eaton, D. A. R., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. Bioinformatics , 36(8), 2592–2594.

Fenker, J., Domingos, F. M. C. B., Tedeschi, L. G., Rosauer, D. F., Werneck, F. P., Colli, G. R., … Moritz, C. (2020). Evolutionary history of Neotropical savannas geographically concentrates species, phylogenetic and functional diversity of lizards. Journal of Biogeography, 47(5), 1130–1142.

Fenker, J., Tedeschi, L. G., Melville, J., & Moritz, C. (2021). Predictors of Phylogeographic Structure among co-distributed taxa across the complex Australian Monsoonal Tropics. Molecular Ecology. doi: 10.1111/mec.16057

Fjeldsa, J., Lambin, E., & Mertens, B. (1999). Correlation between endemism and local ecoclimatic stability documented by comparing Andean bird distributions and remotely sensed land surface data. Ecography, 22(1), 63–78.

Fontanella, F. M., Feltrin, N., Avila, L. J., Sites, J. W., & Morando, M. (2012). Early stages of divergence: phylogeography, climate modeling, and morphological differentiation in the South American lizard Liolaemus petrophilus (Squamata: Liolaemidae). Ecology and Evolution, 2(4), 792–808.

Haffer, J. (1969). Speciation in amazonian forest birds. Science, 165(3889), 131–137.

Haffer, J., & Prance, G. T. (2001). Climatic forcing of evolution in Amazonia during the Cenozoic: on the refuge theory of biotic differentiation. Amazoniana: Limnologia et Oecologia Regionalis Systematis Fluminis Amazonas, 16(3/4), 579–607.

Hedges, S. B., Blair Hedges, S., & Conn, C. E. (2012). A new skink fauna from Caribbean islands (Squamata, Mabuyidae, Mabuyinae). Zootaxa, Vol. 3288, p. 1. doi: 10.11646/zootaxa.3288.1.1

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. Nature, 405(6789), 907–913.

Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H., Johnson, J. B., … Yoder, A. D. (2010). Phylogeography's past, present, and future: 10 years after Avise, 2000. Molecular Phylogenetics and Evolution, Vol. 54, pp. 291–301. doi: 10.1016/j.ympev.2009.09.016

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.

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.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods, 14(6), 587–589.

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017). PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and

Morphological Phylogenetic Analyses. Molecular Biology and Evolution, 34(3), 772–773.

Leaché, A. D., Portik, D. M., Rivera, D., Rödel, M., Penner, J., Gvoždík, V., … Fujita, M. K. (2019). Exploring rain forest diversification using demographic model testing in the African foam-nest treefrog Chiromantis rufescens. Journal of Biogeography, 46(12), 2706–2721.

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.

Miralles, A., & Carranza, S. (2010). Systematics and biogeography of the Neotropical genus Mabuya, with special emphasis on the Amazonian skink Mabuya nigropunctata (Reptilia, Scincidae). Molecular Phylogenetics and Evolution, 54(3), 857–869.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. Nature, 403(6772), 853–858.

Oaks, J. R. (2019). Full Bayesian Comparative Phylogeography from Genomic Data. Systematic Biology, 68(3), 371–395.

Papadopoulou, A., & Knowles, L. L. (2016). Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. Proceedings of the National Academy of Sciences of the United States of America, 113(29), 8018–8024.

Pellegrino, K. C. M., Rodrigues, M. T., Waite, A. N., Morando, M., Yassuda, Y. Y., & Sites, J. W., JR. (2005). Phylogeography and species limits in the Gymnodactylus darwinii complex (Gekkonidae, Squamata): genetic structure coincides with river systems in the

Brazilian Atlantic Forest. Biological Journal of the Linnean Society. Linnean Society of London, 85, 13–26.

Pinto-Sánchez, N. R., Calderón-Espinosa, M. L., Miralles, A., Crawford, A. J., & Ramírez-Pinilla, M. P. (2015). Molecular phylogenetics and biogeography of the Neotropical skink genus Mabuya Fitzinger (Squamata: Scincidae) with emphasis on Colombian populations. Molecular Phylogenetics and Evolution, 93, 188–211.

Potter, S., Afonso Silva, A. C., Bragg, J. G., Catalano, S. R., Donnellan, S., Doughty, P., … Moritz, C. (2019). Contrasting scales of local persistence between monsoonal and arid biomes in closely related, low-dispersal vertebrates. Journal of Biogeography, 46(11), 2506–2519.

Potter, S., Xue, A. T., Bragg, J. G., Rosauer, D. F., Roycroft, E. J., & Moritz, C. (2018). Pleistocene climatic changes drive diversification across a tropical savanna. Molecular Ecology, 27(2), 520–532.

Prates, I., Rivera, D., Rodrigues, M. T., & Carnaval, A. C. (2016). A mid-Pleistocene rainforest corridor enabled synchronous invasions of the Atlantic Forest by Amazonian anole lizards. Molecular Ecology, 25(20), 5174–5186.

Prates, I., Xue, A. T., Brown, J. L., Alvarado-Serrano, D. F., Rodrigues, M. T., Hickerson, M. J., & Carnaval, A. C. (2016). Inferring responses to climate dynamics from historical demography in neotropical forest lizards. Proceedings of the National Academy of Sciences of the United States of America, 113(29), 7978–7985.

Ribas, C. C., Aleixo, A., Nogueira, A. C. R., Miyaki, C. Y., & Cracraft, J. (2012). A palaeobiogeographic model for biotic diversification within Amazonia over the past three

million years. Proceedings. Biological Sciences / The Royal Society, 279(1729), 681–689.

Rivera, D., Prates, I., Rodrigues, M. T., & Carnaval, A. C. (2020). Effects of climate and geography on spatial patterns of genetic structure in tropical skinks. Molecular Phylogenetics and Evolution, 143, 106661.

Sambrook, J., & Russell, D. W. (2006). Purification of nucleic acids by extraction with phenol:chloroform. CSH Protocols, 2006(1). doi: 10.1101/pdb.prot4455

Thomé, M. T. C., Zamudio, K. R., Giovanelli, J. G. R., Haddad, C. F. B., Baldissera, F. A., Jr, & Alexandrino, J. (2010). Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. Molecular Phylogenetics and Evolution, 55(3), 1018–1031.

Werneck, F. P., Gamble, T., Colli, G. R., Rodrigues, M. T., & Sites, J. W., Jr. (2012). Deep diversification and long-term persistence in the South American “dry diagonal”: integrating continent-wide phylogeography and distribution modeling of geckos. Evolution; International Journal of Organic Evolution, 66(10), 3014–3034.

Werneck, F. P., Leite, R. N., Geurgas, S. R., & Rodrigues, M. T. (2015). Biogeographic history and cryptic diversity of saxicolous Tropiduridae lizards endemic to the semiarid Caatinga. BMC Evolutionary Biology, 15, 94.

Xue, A. T., & Hickerson, M. J. (2017). Multi-dice : R package for comparative population genomic inference under hierarchical co-demographic models of independent single-population size changes. Molecular Ecology Resources, 17(6), e212–e224.

Xue, A. T., & Hickerson, M. J. (2020). Comparative phylogeographic inference with genome-wide data from aggregated population pairs. Evolution; International Journal of Organic Evolution, 74(5), 808–830.

Zamudio, K. R., Bell, R. C., & Mason, N. A. (2016). Phenotypes in phylogeography: Species' traits, environmental variation, and vertebrate diversification. Proceedings of the National Academy of Sciences of the United States of America, 113(29), 8041–8048.

Figures & Tables

FIGURE 1 Phylogenetic reconstruction for *Mabuya* species. Grey nodal circles represent BS > 95 and PP > 0.95. Bar plot and colors correspond to `ADMIXTURE` results.

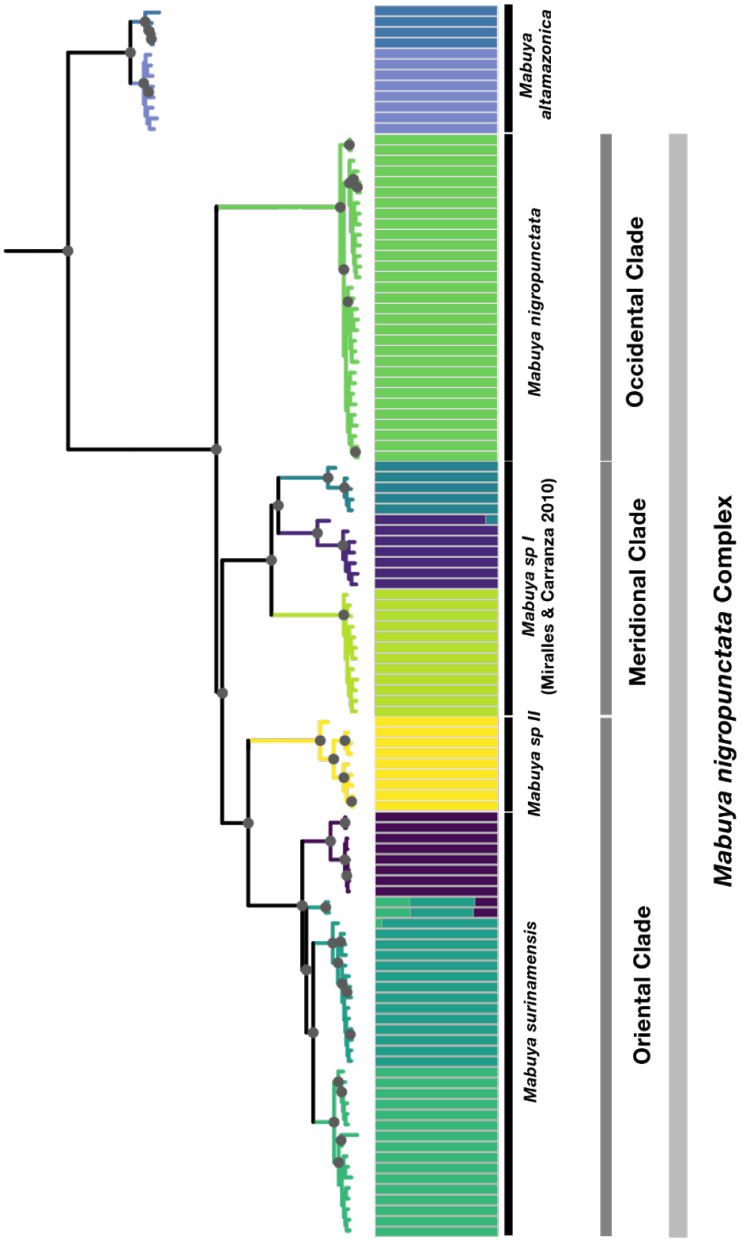


FIGURE 2 Phylogenetic reconstruction for *Gymnodactylus* species. Grey nodal circles represent BS > 95 and PP > 0.95. Bar plot and colors correspond to ADMIXTURE results.

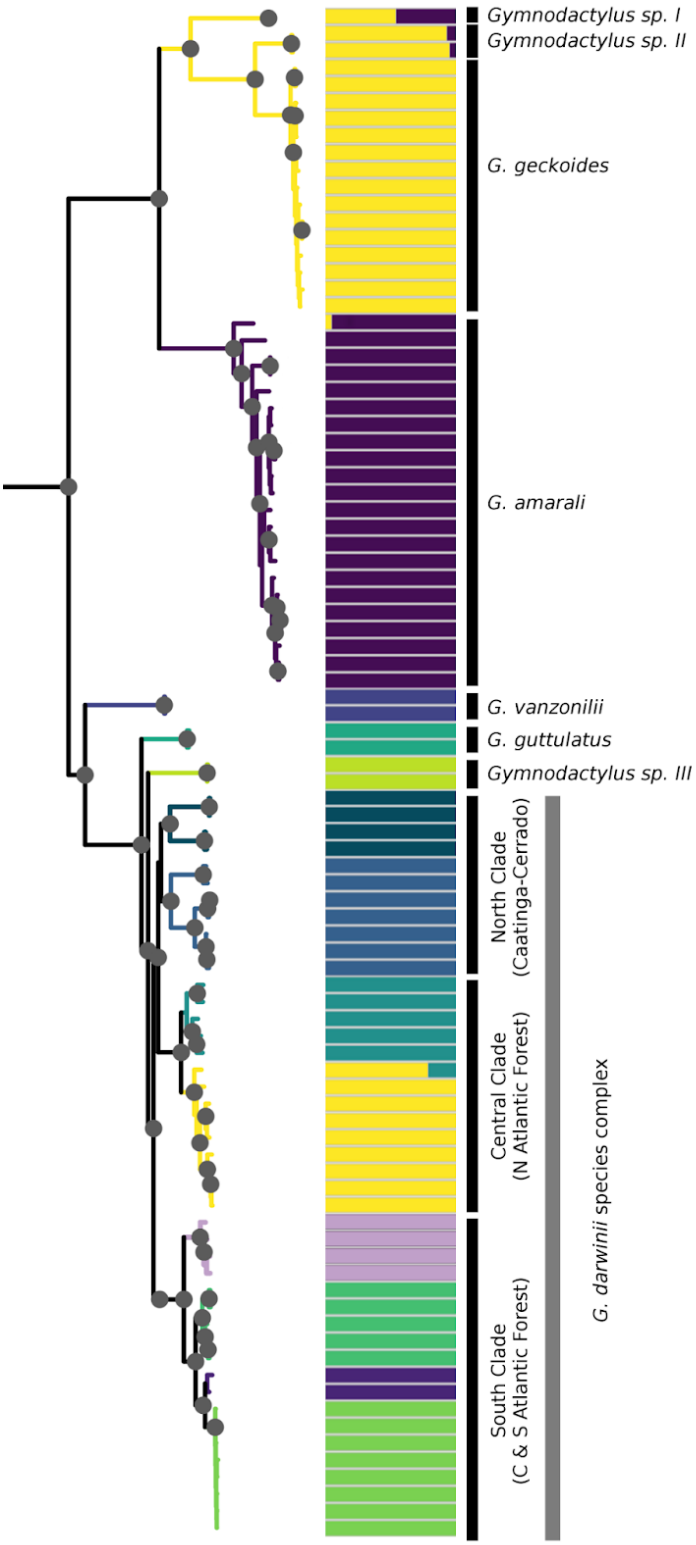


FIGURE 3 Approximate marginal posterior densities of divergence times for each tested species or population pair. Time is in units of expected substitutions per site

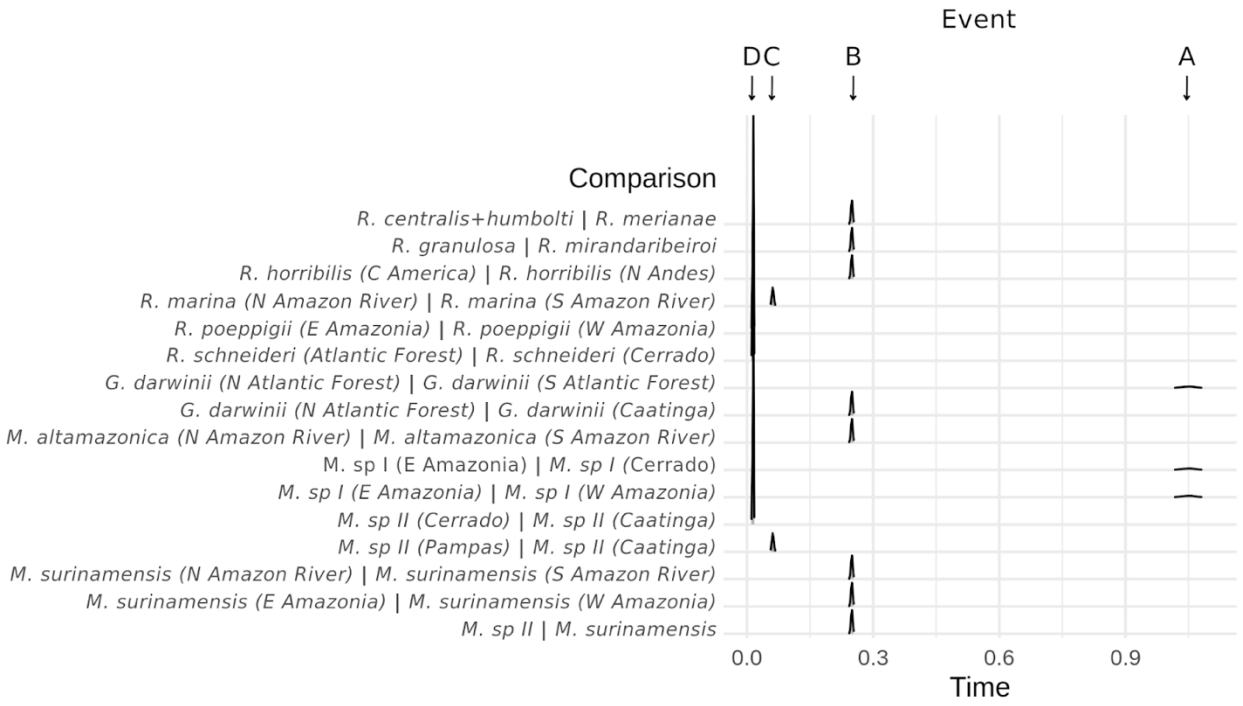


TABLE 1 Population size change (θ) and % change (Δ) for each comparison from the *Ecoevolity* analysis. Events refer to τ (expected substitutions per site): A = 1.0508; B = 0.2493; C = 0.0614; D = 0.0147.

		Comparison		Event
eastern – western Amazonia Population size (θ)	(root)	R. poeppigii (eastern Amazonia)	R. poeppigii (western Amazonia)	D
	0.409	0.165	0.010	
	% Δ	-59.61	-97.55	
	(root)	*M. surinamensis (eastern Amazonia)*	*M. surinamensis (western Amazonia)*	B
	0.047	0.010	0.084	
	% Δ	-79.48	79.57	
	(root)	*Mabuya sp. I (eastern Amazonia)*	*Mabuya sp. I (western Amazonia)*	A
	0.058	0.061	0.039	
	% Δ	5.17	-32.93	
	(root)	*R. schneideri (Atlantic Forest)*	*R. schneideri (Cerrado)*	D

	0.050	0.026	0.025	
	% Δ	-47.39	-50.40	
Atlantic Forest – Cerrado/ Caatinga Population size (θ)	(root)	*R. granulosa (Atlantic Forest)*	*R. mirandaribeiroi (Cerrado)*	B
	0.074	0.066	0.092	
	% Δ	-10.33	24.32	
	(root)	*G. darwinii (Caatinga)*	*G. darwinii (northern Atlantic Forest)*	B
	0.140	0.014	0.042	
	% Δ	-90.30	-70.26	
North – South of Amazon River Population size (θ)	(root)	*R. marina (North of Amazon River)*	*R. marina (South of Amazon River)*	C
	0.136	0.040	0.051	
	% Δ	-70.80	-62.76	
	(root)	*M. altamazonica (North of Amazon River)*	*M. altamazonica (South Amazon River)*	B
	0.142	0.059	0.058	
	% Δ	-58.86	-59.14	
	(root)	*M. surinamensis (North of Amazon River)*	*M. surinamensis (South of Amazon River)*	B
	0.082	0.054	0.035	
	% Δ	-34.99	-57.47	
Amazonia - Cerrado Population size (θ)	(root)	*Mabuya sp. II (Cerrado)*	*M. surinamensis (Amazonia)*	B
	0.030	0.033	0.077	
	% Δ	10.81	158.78	
	(root)	*Mabuya sp. I (eastern Amazonia)*	*Mabuya sp. I (Cerrado)*	A
	0.066	0.082	0.051	
	% Δ	24.24	-23.02	
northern – southern Atlantic Forest Population size (θ)	(root)	*G. darwinii (southern Atlantic Forest)*	*G. darwinii (northern Atlantic Forest)*	A
	0.004	0.046	0.073	
	% Δ	974.41	1605.45	
	(root)	*Mabuya sp. II (Caatinga)*	*Mabuya sp. II (Cerrado)*	D

	0.006	0.383	0.202	
Caatinga – Cerrado – Pampas Population size (θ)	% Δ	6716.48	3501.60	
	(root)	*Mabuya sp. II (Pampas)*	*Mabuya sp. II (Caatinga)*	C
	0.017	0.771	0.281	
	% Δ	4491.07	1572.62	
northern Andes – Amazonia Population size (θ)	(root)	*R. centralis + R. humbolti (northern Andes)*	*R. merianae (Amazonia)*	B
	0.040	0.060	0.108	
	% Δ	48.14	167.99	
Central America – northern Andes Population size (θ)	(root)	*R. horribilis (Central America)*	*R. horribilis (northern Andes)*	B
	0.082	0.098	0.076	
	% Δ	19.56	-7.78	

Supplemental Figures

FIGURE S1 Locality maps for *Mabuya altamazonica* and *M. nigropuctata* complex species

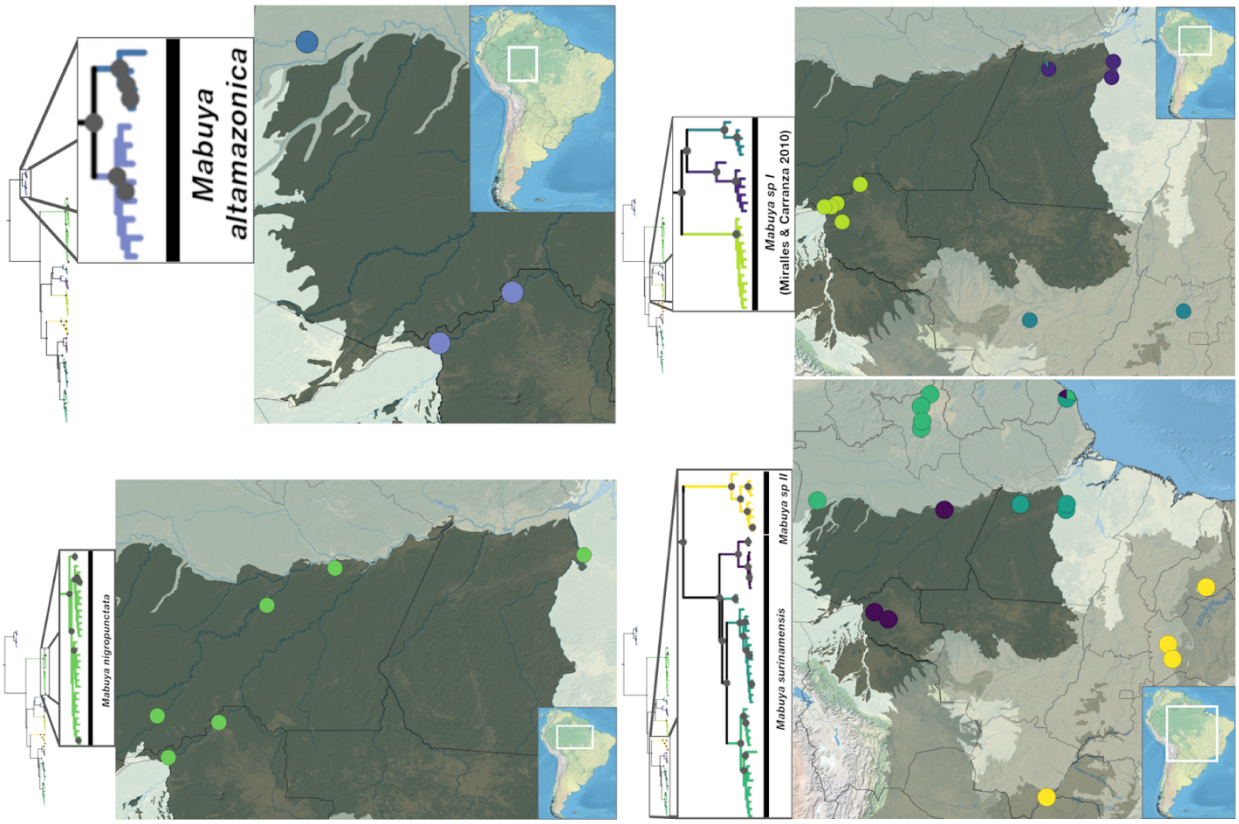
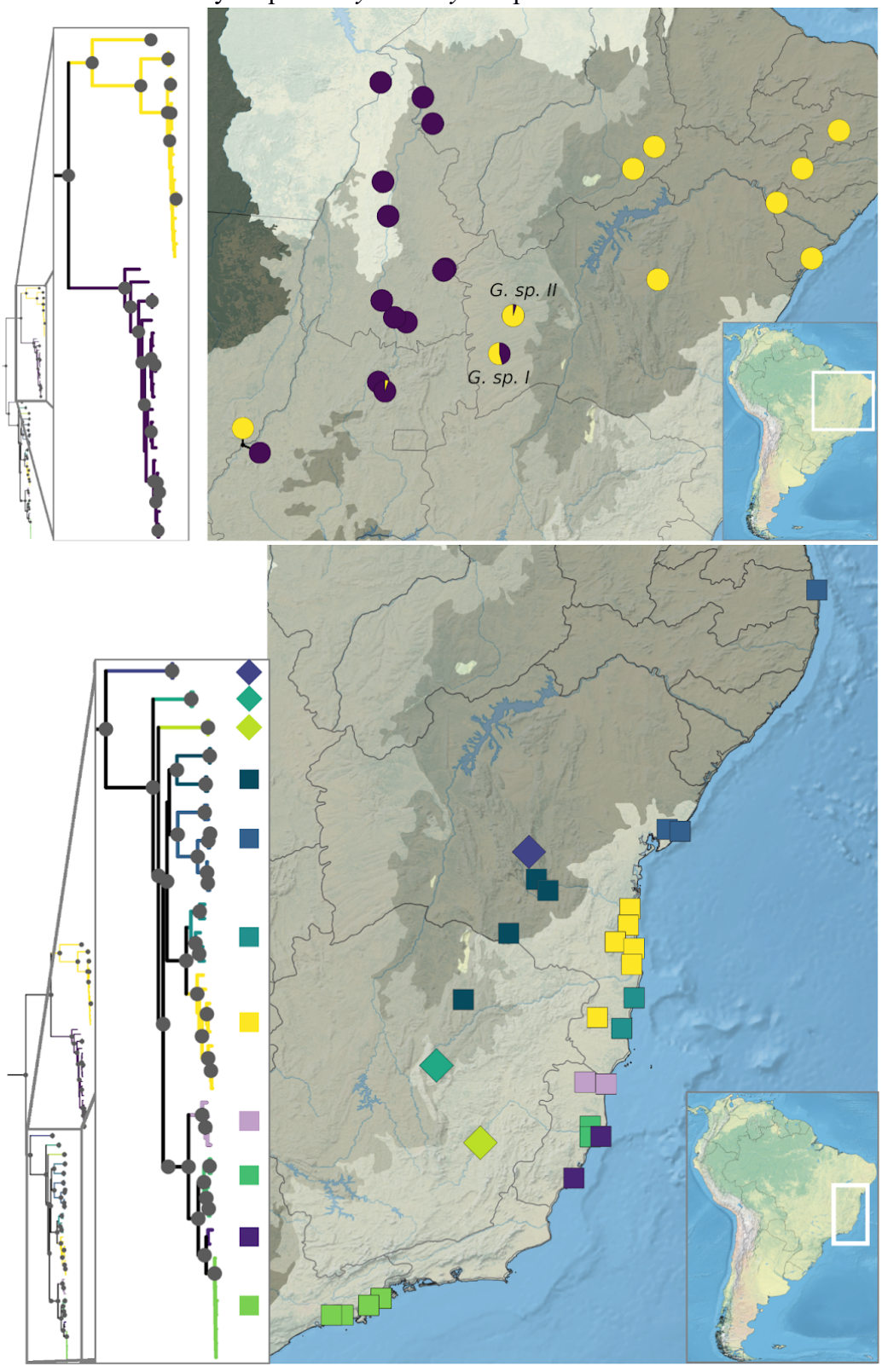


FIGURE S2 Locality maps for *Gymnodatylus* species.



Supplemental Tables

TABLE S1. All sample identification information for individuals included in this study

Species	Field Number	Locality
Mabuya altamazonica	H2044	Mutum, RO, BR
Mabuya altamazonica	MTR36127	Comunidade Cachoeirinha, Rio Içá, AM, BR
Mabuya altamazonica	MTR36242	Comunidade Cachoeirinha, Rio Içá, AM, BR
Mabuya altamazonica	H1231	Caiçara, RO, BR
Mabuya altamazonica	H2002	Caiçara, RO, BR
Mabuya altamazonica	H2275	Abunã, RO, BR
Mabuya altamazonica	H2290	Abunã, RO, BR
Mabuya altamazonica	H2701	UHE Jirau, Mutum, RO, BR
Mabuya altamazonica	H3043	UHE Jirau, Mutum, RO, BR
Mabuya altamazonica	HJ0664	Abunã, RO, BR
Mabuya altamazonica	MTR35769	Comunidade Cachoeirinha, Rio Içá, AM, BR
Mabuya altamazonica	MTR35892	Comunidade Cachoeirinha, Rio Içá, AM, BR
Mabuya altamazonica	MTR36075	Açaí, Rio Içá, AM, BR
Mabuya altamazonica	MTR36153	Açaí, Rio Içá, AM, BR
Mabuya nigropunctata	BM500	Anapu, PA, BR
Mabuya nigropunctata	H2065	Mutum, RO, BR
Mabuya nigropunctata	H953	Abunã, RO, BR
Mabuya nigropunctata	HJ0429	Abunã, RO, BR
Mabuya nigropunctata	HJ0722	Abunã, RO, BR
Mabuya nigropunctata	LSUMZH-14105	Rio Ituxi, AM, BR
Mabuya nigropunctata	LSUMZH-14106	Rio Ituxi, AM, BR
Mabuya nigropunctata	LSUMZH-14107	Rio Ituxi, AM, BR
Mabuya nigropunctata	LSUMZH-14108	Rio Ituxi, AM, BR
Mabuya nigropunctata	LSUMZH-14109	Rio Ituxi, AM, BR
Mabuya nigropunctata	LSUMZH-14112	Rio Ituxi, AM, BR
Mabuya nigropunctata	LSUMZH-14115	Across Rio Ituxi at the Madeireira Scheffer, AM, BR
Mabuya nigropunctata	LSUMZH-16426	Amazonas, AM, BR

Mabuya nigropunctata	LSUMZH-16427	Amazonas, AM, BR
Mabuya nigropunctata	LSUMZH-16441	Amazonas, AM, BR
Mabuya nigropunctata	LSUMZH-16446	Amazonas, AM, BR
Mabuya nigropunctata	LSUMZH-16452	Amazonas, AM, BR
Mabuya nigropunctata	LSUMZH-16468	Amazonas, AM, BR
Mabuya nigropunctata	LSUMZH-16489	Amazonas, AM, BR
Mabuya nigropunctata	LSUMZH-16490	Amazonas, AM, BR
Mabuya nigropunctata	MTR19033	Moiobamba, AM, BR
Mabuya nigropunctata	MTR19034	Moiobamba, AM, BR
Mabuya nigropunctata	MTR19035	Moiobamba, AM, BR
Mabuya nigropunctata	MTR19047	Moiobamba, AM, BR
Mabuya nigropunctata	MTR19093	Moiobamba, AM, BR
Mabuya nigropunctata	MTR19230	Moiobamba, AM, BR
Mabuya nigropunctata	MTR19253	Moiobamba, AM, BR
Mabuya nigropunctata	MTR19434	Moiobamba, AM, BR
Mabuya nigropunctata	MTR36509	São Pedro, Rio Içá, AM, BR
Mabuya nigropunctata	BM518	Anapu, PA, BR
Mabuya nigropunctata	BM537	Anapu, PA, BR
Mabuya nigropunctata	H2224	Abunã, RO, BR
Mabuya nigropunctata	MTR19232	Moiobamba, AM, BR
Mabuya sp. I	BM005	UHE Belo Monte, PA, BR
Mabuya sp. I	BM097	UHE Belo Monte, PA, BR
Mabuya sp. I	BM115	UHE Belo Monte, PA, BR
Mabuya sp. I	BM176	Vitória do Xingu, PA, BR
Mabuya sp. I	BM655	Vitória do Xingu, PA, BR
Mabuya sp. I	H1307	Caiçara, RO, BR
Mabuya sp. I	H1516	Mutum, RO, BR
Mabuya sp. I	H1526	Mutum, RO, BR
Mabuya sp. I	H1760	Mutum, RO, BR
Mabuya sp. I	H2667	UHE Jirau, Mutum, RO, BR
Mabuya sp. I	H3244	UHE Jirau, Abunã, RO, BR

Mabuya sp. I	H818	Caiçara, RO, BR
Mabuya sp. I	LG1085	Niquelândia, GO, BR
Mabuya sp. I	LG1558	APM Manso, MT, BR
Mabuya sp. I	LG1561	APM Manso, MT, BR
Mabuya sp. I	LSUMZH-14179	Agropecuaria Treviso LTDA, PA, BR
Mabuya sp. I	LSUMZH-17862	Rio Formoso, Parque Estadual Guajara-Mirim, RO, BR
Mabuya sp. I	LSUMZH-17863	Rio Formoso, Parque Estadual Guajara-Mirim, RO, BR
Mabuya sp. I	LSUMZH-17864	Rio Formoso, Parque Estadual Guajara-Mirim, RO, BR
Mabuya sp. I	LSUMZH-17865	Rio Formoso, Parque Estadual Guajara-Mirim, RO, BR
Mabuya sp. I	RGL1006	UHE Guaporé, MT, BR
Mabuya sp. I	RGL1024	UHE Guaporé, MT, BR
Mabuya sp. I	RGL1025	UHE Guaporé, MT, BR
Mabuya sp. I	RRT56	PCH Rondonopolis, Rondonopolis, MT, BR
Mabuya sp. I	BM082	UHE Belo Monte, PA, BR
Mabuya sp. I	BM658	Vitória do Xingu, PA, BR
Mabuya sp. I	H1884	Caiçara, RO, BR
Mabuya sp. I	LG1550	APM Manso, MT, BR
Mabuya sp. I	LG1568	APM Manso, MT, BR
Mabuya sp. II	2942	Rosana, SP, BR
Mabuya sp. II	MTR23467	Parque Nacional da Serra da Capivara, PI, BR
Mabuya sp. II	MTR23475	Parque Nacional da Serra da Capivara, PI, BR
Mabuya sp. II	MTR26337	Parque Nacional da Serra da Capivara, PI, BR
Mabuya sp. II	MTR26342	Parque Nacional da Serra da Capivara, PI, BR
Mabuya sp. II	MTR26979	São Desidério, BA, BR
Mabuya sp. II	MTR27143	Correntina, BA, BR
Mabuya sp. II	MTR26340	Parque Nacional da Serra da Capivara, PI, BR
Mabuya sp. II	MTR27140	Correntina, BA, BR
Mabuya sp. II	PHV3100	Alto Araguaia, MT, BR
Mabuya surinamensis	BM092	UHE Belo Monte, PA, BR
Mabuya surinamensis	BM285	Vitória do Xingu, PA, BR
Mabuya surinamensis	BM475	Vitória do Xingu, PA, BR

Mabuya surinamensis	BM692	Anapu, PA, BR
Mabuya surinamensis	LSUMZH-12297	Fazenda Nova Esperanca, RR, BR
Mabuya surinamensis	LSUMZH-12311	Fazenda Nova Esperanca, RR, BR
Mabuya surinamensis	LSUMZH-12332	Fazenda Nova Esperanca, RR, BR
Mabuya surinamensis	LSUMZH-12365	Fazenda Nova Esperanca, RR, BR
Mabuya surinamensis	LSUMZH-12369	Fazenda Nova Esperanca, RR, BR
Mabuya surinamensis	LSUMZH-14195	Agropecuaria Treviso LTDA, PA, BR
Mabuya surinamensis	LSUMZH-14206	Agropecuaria Treviso LTDA, PA, BR
Mabuya surinamensis	LSUMZH-14207	Agropecuaria Treviso LTDA, PA, BR
Mabuya surinamensis	LSUMZH-14223	Agropecuaria Treviso LTDA, PA, BR
Mabuya surinamensis	LSUMZH-14224	Agropecuaria Treviso LTDA, PA, BR
Mabuya surinamensis	LSUMZH-14238	Agropecuaria Treviso LTDA, PA, BR
Mabuya surinamensis	LSUMZH-14290	Agropecuaria Treviso LTDA, PA, BR
Mabuya surinamensis	LSUMZH-14337	Agropecuaria Treviso LTDA, PA, BR
Mabuya surinamensis	LSUMZH-14352	Agropecuaria Treviso LTDA, PA, BR
Mabuya surinamensis	LSUMZH-14358	Agropecuaria Treviso LTDA, PA, BR
Mabuya surinamensis	LSUMZH-16393	Amazonas, AM, BR
Mabuya surinamensis	LSUMZH-16399	Amazonas, AM, BR
Mabuya surinamensis	LSUMZH-17858	Rio Formoso, Parque Estadual Guajara-Mirim, RO, BR
Mabuya surinamensis	LSUMZH-17859	Rio Formoso, Parque Estadual Guajara-Mirim, RO, BR
Mabuya surinamensis	LSUMZH-17860	Rio Formoso, Parque Estadual Guajara-Mirim, RO, BR
Mabuya surinamensis	LSUMZH-17861	Rio Formoso, Parque Estadual Guajara-Mirim, RO, BR
Mabuya surinamensis	MTR20400	E.E. Maracá, RR, BR
Mabuya surinamensis	MTR20422	E.E. Maracá, RR, BR
Mabuya surinamensis	MTR20561	E.E. Maracá, RR, BR
Mabuya surinamensis	MTR20598	E.E. Maracá, RR, BR
Mabuya surinamensis	MTR20619	E.E. Maracá, RR, BR
Mabuya surinamensis	MTR20786	Pacaraima Gilberto Macuxi, RR, BR
Mabuya surinamensis	MTR23160	Serra da Maroquinha, RR, BR
Mabuya surinamensis	MTR23182	Serra da Maroquinha, RR, BR
Mabuya surinamensis	MTR23184	Serra da Maroquinha, RR, BR

Mabuya surinamensis	MTR24125	Oiapoque, AP, BR
Mabuya surinamensis	MTR24128	Oiapoque, AP, BR
Mabuya surinamensis	MTR25583	Parque Nacional de Pacaás Novos, RO, BR
Mabuya surinamensis	MTR36135	Comunidade Cachoeirinha, Rio Içá, AM, BR
Mabuya surinamensis	SMS931	Serra do Apiaú, RR, BR
Mabuya surinamensis	MTR23036	Serra da Maroquinha, RR, BR
Mabuya surinamensis	MTR25581	Parque Nacional de Pacaás Novos, RO, BR
Mabuya surinamensis	SMS031	Comunidade Projó, AM, BR
Mabuya frenata	MTR10568	UHE Ponte de Pedra, MS/MT, BR
Mabuya frenata	PHV2861	Alto Araguaia, MT, BR
Mabuya frenata	PHV2862	Alto Araguaia, MT, BR
Gymnodactylus amarali	A2261	Barra do Garças, MT, BR
Gymnodactylus amarali	ESTR00196	Carolina, MA, BR
Gymnodactylus amarali	ESTR00642	Estreito, MA, BR
Gymnodactylus amarali	ESTR01293	Estreito, MA, BR
Gymnodactylus amarali	LG0889	Barra do Garças, MT, BR
Gymnodactylus amarali	LG1075	Niquelândia, GO, BR
Gymnodactylus amarali	LG1313	Serra da Mesa, GO, BR
Gymnodactylus amarali	LG1314	Serra da Mesa, GO, BR
Gymnodactylus amarali	MSH10885	Serra Andorinhas, PA, BR
Gymnodactylus amarali	MTR03949	Peixe, TO, BR
Gymnodactylus amarali	MTR04052	Paranã, TO, BR
Gymnodactylus amarali	MTR04459	Peixe, TO, BR
Gymnodactylus amarali	MTR06428	São Salvador (Faz. Traçadal), TO, BR
Gymnodactylus amarali	MTR06433	São Salvador, TO, BR
Gymnodactylus amarali	MTR06630	UHE Lajeado, TO, BR
Gymnodactylus amarali	MTR06732	UHE Lajeado, TO, BR
Gymnodactylus amarali	MTR07479	Guaraí, TO, BR
Gymnodactylus amarali	MTR07542	Guaraí, TO, BR
Gymnodactylus amarali	MTR14255	Estação Ecológica Serra Geral do Tocantins, TO, BR
Gymnodactylus amarali	MTR14604	Estação Ecológica Serra Geral do Tocantins, TO, BR

Gymnodactylus amarali	MTR14605	Estação Ecológica Serra Geral do Tocantins, TO, BR
Gymnodactylus amarali	MTR14606	Estação Ecológica Serra Geral do Tocantins, TO, BR
Gymnodactylus darwinii	A1029	Ubatuba, SP, BR
Gymnodactylus darwinii	A2246	Praia do Forte, BA, BR
Gymnodactylus darwinii	A2247	Praia do Forte, BA, BR
Gymnodactylus darwinii	A2249	Mata de São João, BA, BR
Gymnodactylus darwinii	A8373	Vitória, ES, BR
Gymnodactylus darwinii	ABA16-1	C.E. Almada, Ilhéus, BA, BR
Gymnodactylus darwinii	ABA17	C.E. Almada, Ilhéus, BA, BR
Gymnodactylus darwinii	FSFL1445	Prado, BA, BR
Gymnodactylus darwinii	FSFL1491	Prado, BA, BR
Gymnodactylus darwinii	H557	Bertioga, SP, BR
Gymnodactylus darwinii	H570	Bertioga, SP, BR
Gymnodactylus darwinii	JC1512	Grão Mogol, MG, BR
Gymnodactylus darwinii	JC1515	Grão Mogol, MG, BR
Gymnodactylus darwinii	LG0802	Ubatuba, Ilha da Pesca, SP, BR
Gymnodactylus darwinii	LG0934	Ubatuba, Ilha da Pesca, SP, BR
Gymnodactylus darwinii	LG0935	Ubatuba, Ilha do Promirim, SP, BR
Gymnodactylus darwinii	LG0957	Porto Seguro, BA, BR
Gymnodactylus darwinii	LG0991	Porto Seguro, BA, BR
Gymnodactylus darwinii	LG1349	Una, BA, BR
Gymnodactylus darwinii	LG1372	Corcovado, Ubatuba, SP, BR
Gymnodactylus darwinii	LG1600	Barra do Una, SP, BR
Gymnodactylus darwinii	LG2064	Cabedelo, Mata do Amém, PB, BR
Gymnodactylus darwinii	LSH004	Guarapari, ES, BR
Gymnodactylus darwinii	MTR01266	UHE Rosal, ES, BR
Gymnodactylus darwinii	MTR06035	Serra do Teimoso, Jussari, BA, BR
Gymnodactylus darwinii	MTR06038	Serra do Teimoso, Jussari, BA, BR
Gymnodactylus darwinii	MTR10297	Parque Estadual Itaunas, ES, BR
Gymnodactylus darwinii	MTR10298	Parque Estadual Itaunas, ES, BR
Gymnodactylus darwinii	MTR11105	Ilhéus, BA, BR

Gymnodactylus darwinii	MTR11790	Itacaré, BA, BR
Gymnodactylus darwinii	MTR12058	Floresta Nacional de Goytacazes, Linhares, ES, BR
Gymnodactylus darwinii	MTR12182	Floresta Nacional de Goytacazes, Linhares, ES, BR
Gymnodactylus darwinii	MTR12235	Linhares, Reserva da Companhia Vale do Rio Doce, ES,
Gymnodactylus darwinii	MTR12431	Regência, ES, BR
Gymnodactylus darwinii	MTR12450	Linhares, Reserva da Companhia Vale do Rio Doce, ES,
Gymnodactylus darwinii	MTR13412	Trancoso (Fazenda Nova Alegria), BA, BR
Gymnodactylus darwinii	MTR16188	Santa Luzia, estrada Camacan - Canavieiras, BA, BR
Gymnodactylus darwinii	MTR16207	Canavieiras, BA, BR
Gymnodactylus darwinii	MTR16454	Condeuba, Fazenda Santo Antonio, BA, BR
Gymnodactylus darwinii	MTR21513	Pinheiros, Trilha da Anta, ES, BR
Gymnodactylus darwinii	MTR21514	Pinheiros, Água limpa, ES, BR
Gymnodactylus darwinii	MTR22917	FLONA Contendas do Sincorá, BA, BR
Gymnodactylus darwinii	MTR22945	Barra da Estiva, BA, BR
Gymnodactylus darwinii	MTR22950	Barra da Estiva, BA, BR
Gymnodactylus geckoides	CGERV075	Capitão Gervásio de Oliveira, PI, BR
Gymnodactylus geckoides	CGERV102	Capitão Gervásio de Oliveira, PI, BR
Gymnodactylus geckoides	LG0475	Jacobina, Serra do Ouro, PA, BR
Gymnodactylus geckoides	LG0495	Jacobina, PA, BR
Gymnodactylus geckoides	LG0804	Xingó, AL/SE, BR
Gymnodactylus geckoides	LG0912	Xingó, AL/SE, BR
Gymnodactylus geckoides	LG1050	Barra dos Coqueiros, SE, BR
Gymnodactylus geckoides	LG1051	Barra dos Coqueiros, SE, BR
Gymnodactylus geckoides	LG1130	Camaçari, BA, BR
Gymnodactylus geckoides	MTR15375	Parque Nacional do Catimbau (Fazenda Porto Seguro),
Gymnodactylus geckoides	MTR15395	Parque Nacional do Catimbau (Fazenda Porto Seguro),
Gymnodactylus geckoides	MTR887012	Cabaceiras, PB, BR
Gymnodactylus geckoides	MTR906096	Morro do Chapéu, BA, BR
Gymnodactylus geckoides	MTR906097	Morro do Chapéu, BA, BR
Gymnodactylus geckoides	MTR946147	Barra do Garças, MT, BR
Gymnodactylus guttulatus	JC1517	Sopa, prox. a Guinda, Diamantina, MG, BR

Gymnodactylus guttulatus	JC1518	Guinda, Diamantina, MG, BR
Gymnodactylus sp. I	MTR17942	Mata das Barrigudas, Correntina, BA, BR
Gymnodactylus sp. II	MTR17909	São Desidério, BA, BR
Gymnodactylus sp. II	MTR17910	São Desidério, BA, BR
Gymnodactylus sp. III	MTR17507	Parque Estadual do Rio Doce, Marliéria, MG, BR
Gymnodactylus sp. III	MTR17568	Parque Estadual do Rio Doce, Marliéria, MG, BR
Gymnodactylus vanzolinii	JC1207	Mucugê, BA, BR
Gymnodactylus vanzolinii	JC1249	Mucugê, BA, BR
Thecadactylus rapicauda	ENS7108	Izabal, GT
Thecadactylus rapicauda	ENS9222	Izabal, GT

TABLE S2 Metadata for *Ecoevolity* analysis.

Species/Population 1	**N**	**Species/Population 2**	**N**	**Phylogenetic Break**	**# Loci**
Gymnodactylus darwinii Northern Atlantic Forest	13	*Gymnodactylus darwinii* Southern Atlantic Forest	8	North Atlantic Forest - South Atlantic Forest	18,796
Gymnodactylus darwinii Northern Atlantic Forest	13	*Gymnodactylus darwinii* Caatinga	9	Caatinga - North Atlantic Forest	16,251
Mabuya altamazonica N of Amazon River	4	*Mabuya altamazonica* S of Amazon River	8	Northern Amazonia - Southern Amazonia	30,162
Mabuya sp. I Eastern Amazonia (Miralles & Carranza 2010)	7	*Mabuya sp. I* Cerrado (Miralles & Carranza 2010)	5	Eastern Amazonia - Cerrado	30,145
Mabuya sp. I Eastern Amazonia (Miralles & Carranza 2010)	7	*Mabuya sp. I* Western Amazonia (Miralles & Carranza 2010)	12	Eastern Amazonia - Western Amazonia	31,782
Mabuya sp II Caatinga	5	*Mabuya sp II* Cerrado	3	Caatinga – Cerrado	14,750
Mabuya sp II Caatinga	5	*Mabuya sp II* Pampas	1	Caatinga – Pampas	11,394
Mabuya surinamensis S of Amazon River	8	*Mabuya surinamensis* N of Amazon River	16	Northern Amazonia - Southern Amazonia	23,754
Mabuya surinamensis Eastern Amazonia	14	*Mabuya surinamensis* Western Amazonia	16	Eastern Amazonia - Western Amazonia	24,721
Mabuya sp. II	9	*Mabuya surinamensis*	40	Dry Diagonal (Caatinga+ Cerrado+ Chaco) - Amazonia	13,434
Rhinella centralis + humbolti	3	*Rhinella merianae*	13	Northern Andes - Northern Amazonia	2,116

Rhinella granulosa	7	*Rhinella mirandaribeiroi*	14	Cerrado - Atlantic Forest	2,183
Rhinella horribilis Northern Andes	4	*Rhinella horribilis* Central *America*	9	Northern Andes - Central America	3,169
Rhinella marina N of Amazon River	13	*Rhinella marina* S of Amazon River	49	Northern Amazonia - Southern Amazonia	3,049
Rhinella poeppigii Eastern Amazonia	6	*Rhinella poeppigii* Western Amazonia	1	Eastern Amazonia - Western Amazonia	3,311
Rhinella schneideri Atlantic Forest	30	*Rhinella schneideri* Cerrado	9	Cerrado - Atlantic Forest	13,910