Check for updates

Received: 8 June 2020 | Revised: 7 December 2020 | Accepted: 15 December 2020

DOI: 10.1111/jzs.12454

JOURNAL OF ZOOLOGICAL SYSTEMATICS AND EVOLUTIONARY RESEARCH

WILEY

SHORT COMMUNICATION

Late Pliocene population divergence and persistence despite Pleistocene climatic fluctuations in the Rio Doce snouted Treefrog (*Ololygon carnevallii*)

Edward A. Myers[1] | **Henrique Folly**[2,3] | **Eric Ragalzi**[3] | **Renato Neves Feio**[2] | **Diego José Santana**[3]

[1]Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

[2]Departamento de Biologia Animal, Museu de Zoologia João Moojen, Univerisdade Federal de Viçosa, Viçosa, MG, Brazil

[3]Instituto de Biociências, Laboratório de Zoologia, Universidade Federal de Mato Grosso do Sul, Campo Grande, MS, Brazil

Correspondence
Edward A. Myers, Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA.
Email: eddie.a.myers@gmail.com

Abstract

The causes of population genetic divergence within the Atlantic Forest of South America are diverse. For example, studies have pointed to the importance of regions of stable suitable habitat throughout the Pleistocene, large rivers acting as biogeographic barriers, as well as changes in elevation. Here, we generate a phylogeographic dataset for the Rio Doce Snouted Treefrog (*Ololygon carnevallii*), a species that is endemic to a narrow portion of the Atlantic Forest. Gene-tree analyses demonstrate that this species is composed of three distinct lineages that diverged from one another during the Pliocene. Ecological niche models projected to climate during the Pleistocene and mid-Holocene suggest that regions of suitable habitat would have shifted through time since the last glacial maximum, and regions of stable habitat were identified. Using generalized dissimilarity modeling, we find no association between genetic divergence and ecological niche models, riverine barriers, elevation, slope, or current climate suggesting that none of these variables have been responsible for lineage formation in the Rio Doce Snouted Treefrog. We suggest that additional phylogeographic studies of narrowly endemic species within the Atlantic Forest are needed to better understand the drivers of diversification and accumulation of biodiversity.

KEYWORDS
Atlantic Forest, DNA barcode, phylogeography, Pliocene divergence

Resumo

A divergência genética populacional de vários organismos na Mata Atlântica é resultado de diversas causas. Por exemplo, estudos têm apontado para a importância de regiões com habitat adequado estável ao longo do Pleistoceno, grandes rios atuando como barreiras biogeográficas, bem como mudanças na topografia. Neste trabalho nós geramos um conjunto de dados filogeográficos para a Perereca-de-Inverno-do-Rio-Doce (Ololygon carnevallii), uma espécie que é endêmica de uma porção estreita da Mata Atlântica. Análises de árvores de gene demonstram que esta espécie é composta por três linhagens distintas que divergiram durante o Plioceno. Os modelos de nicho ecológico projetados no Pleistoceno e Holoceno médio sugerem que as regiões de habitat adequado teriam mudado ao longo do tempo desde o último máximo glacial,

 © 2021 Wiley-VCH GmbH

e então as regiões de habitat estável foram identificadas. Usando uma modelagem de dissimilaridade generalizada não encontramos associação entre divergência genética e modelos de nicho ecológico, rios como barreiras, elevação, declive ou clima atual, sugerindo que nenhuma dessas variáveis foi responsável pela formação das linhagens dessa espécie. Sugerimos que estudos filogeográficos adicionais de espécies estreitamente endêmicas dentro da Mata Atlântica são necessários para melhor compreender as causas da diversificação e acúmulo de biodiversidade nessa ecorregião.

1 | INTRODUCTION

Determining the factors that have promoted population genetic divergence is important for understanding the process of speciation and the accumulation of biodiversity. It is well documented that genetic divergence accumulates between allopatric populations across a biogeographic barrier (Wallace, 1854). This can be for example rivers that bisects the distribution of a species (e.g., the Amazon river; Hayes & Sewlal, 2004) or can be stretches of unsuitable habitat (e.g., rainforest bisected by savannas; Lorenzen et al., 2012). Changes in climate and associated changes in habitat throughout the Pleistocene have also been responsible for promoting allopatric divergence on a global scale (Hewitt, 2000). Alternatively, genetic divergence can occur because of environmental differences with local adaptation and reduced gene flow between regions (Wang & Bradburd, 2014). This can occur across ecotones or because of heterogeneity in climate across the distribution of a species (Cooke et al., 2012; Ribeiro et al., 2016). Genetic differentiation across the distribution of a species can also result from limited dispersal resulting in a pattern of isolation-by-distance, where genetic distance is correlated with geographic distances between populations (Wright, 1943). These factors can work together to influence genetic divergence across the distribution of a species (Mitchell et al., 2015) and different factors may be responsible for influencing divergence in codistributed species (Myers et al., 2019).

The Atlantic Forest (AF) is a biodiversity hotspot of great conservation concern (Myers et al., 2000), and many studies have sought to understand the high levels of endemic diversity. An overview of these studies suggests that diversification within the AF has been complex and that it is unlikely that any two codistributed species will share the same evolutionary history (Carnaval et al., 2014; Turchetto-Zolet et al., 2012). Refugia of stable forest through time have been hypothesized to be important in driving allopatric divergence in many taxa (Batalha-Filho & Miyaki, 2016; Carnaval & Moritz, 2008). However, the timing of divergence in some species predates the Pleistocene glacial periods suggesting that climate cycles may not have been the sole factor driving diversification (Paz et al., 2019; Thomé et al., 2010). It has also been suggested that population expansion on to the Brazilian continental shelf during periods of lower sea levels has been important in structuring genetic diversity as opposed to the hypothesis that mesic forests in the southern AF were replaced by drier vegetation (Leite et al., 2016). Numerous studies have also suggested that biogeographic barriers, for example, large rivers within the AF, have been responsible for driving population divergence and speciation (Amaro et al., 2012; Thomé et al., 2010), yet other studies show little support for vicariance across riverine barriers (Colombi et al., 2010). It is also possible that multiple factors are interacting to structure genetic diversity in this biome (Sotelo-Muñoz et al., 2020). For example, elevation and historical climate change have influenced population genetic structure in codistributed birds (Thom et al., 2020) and forest refugia and rivers interact in promoting divergence in numerous taxa (Mascarenhas et al., 2019; Menezes et al., 2016). Therefore, to understand diversification within the AF multiple mechanisms need to be considered (Brunes et al., 2015).

The AF can be subdivided into two distinct bioclimatic domains with unique assemblages of plants and animals that are separated by the Rio Doce (Carnaval et al., 2014). Regions north of the river are warmer and the biotic communities are characterized by widespread lowland species with affinities to eastern Amazonia (Batalha-Filho et al., 2013). The southern portion of the AF is more seasonal in climate with more montane and subtropical taxa that are more shared with Andean lineages (Batalha-Filho et al., 2013; Carnaval et al., 2014). Phylogeographic patterns between these two forest blocks have also been shown to differ where lineages have differential responses to climatic cycling between these domains (Paz et al., 2019).

The Rio Doce Snouted Treefrog (*Ololygon carnevallii* Caramaschi and Kisteumacher, 1989) is a small (<3.2 cm) frog that is narrowly endemic within the Bahia refuge of central AF (Carnaval & Moritz, 2008) at the turnover of the two distinct bioclimatic domains and its distribution is bisected by the Doce River. Given that this species is distributed across environmental turnover and across a well-documented riverine barrier, it may be expected to show population structure with these variables. Here, we generate barcode sequence data sampled from across the known distribution of *O. carnevallii* to assess population genetic structure, date the timing of divergence, and assess what factors have been responsible for population genetic divergence.

2 | MATERIALS AND METHODS

2.1 | Sample collection

We obtained a total of 30 tissue samples of *Ololygon carnevallii* from 12 localities to generate molecular data, by collecting individuals in the field from August 2015 to December 2016 (Figure 1, Table 1). Specimens were euthanized using 5% lidocaine chlorhydrate, fixed with 10% formalin, and transferred for permanent storage to 70% ethanol. Tissue samples (muscle or liver) were stored in 100% ethanol before specimen fixation in formalin (permits issued SISBIO 52251). We also verified 34 collection localities of this taxon by examining 429 specimens housed at several natural history collections for ecological niche modeling (see Table S1).

2.2 | Laboratory procedures

Total DNA from the samples was extracted using the Blood & Tissue DNA Mini Kit (Ludwig Biotec, Alvorada, Brazil) from a small piece of muscle or liver tissue preserved in ethanol according to standard DNA barcoding methods for anurans (Koroiva et al., 2020). A total of 658 bp were amplified from the 5′ region of the mitochondrial cytochrome c oxidase subunit 1 (*COI*) gene using the T3-AnF1 (5-ATT AAC CCT CAC TAA AGA CHA AYC AYA AAG AYA TYG G-3′) and T7-AnR1 (‘5- AAT ACG ACT CAC TAT AGC CRA ARA ATC ARA ADA RRT GTT G-3′) primers (Lyra et al., 2017). PCR conditions for amplification consisted of 1× buffer (Colorless GoTaq® Flexi Buffer; Promega Corp., Madison, WI), 0.2 mM dNTP mix, 0.2 µM of each primer, 2 mM MgCl2, 1U Taq polymerase (GoTaq® G2 hot start polymerase, Promega Corp., Madison, WI), and 2 µl of template DNA, in a total reaction volume of 25 µl. The PCR cycling program was run as follows: initial denaturation step with 3 min at 95 °C, 35 cycles of denaturation for 20 s at 95 °C, annealing for 20 s at 50 °C and extension for 1 min at 60 °C, and final extension for 5 min at 60 °C (see Lyra et al., 2017). PCR products were purified with Ethanol/Sodium Acetate and sequenced in both directions on an ABI 3130 Genetic Analyzer (Applied Biosystems). These sequences were aligned in Geneious v 9.0.5 (Biomatters Ltd.) using Muscle v3.11 (Edgar, 2004). The number of variable sites was calculated using *strataG* (Archer et al., 2017) in R, and the number of haplotypes was calculated with *haplotypes* in R (Aktas, 2015).

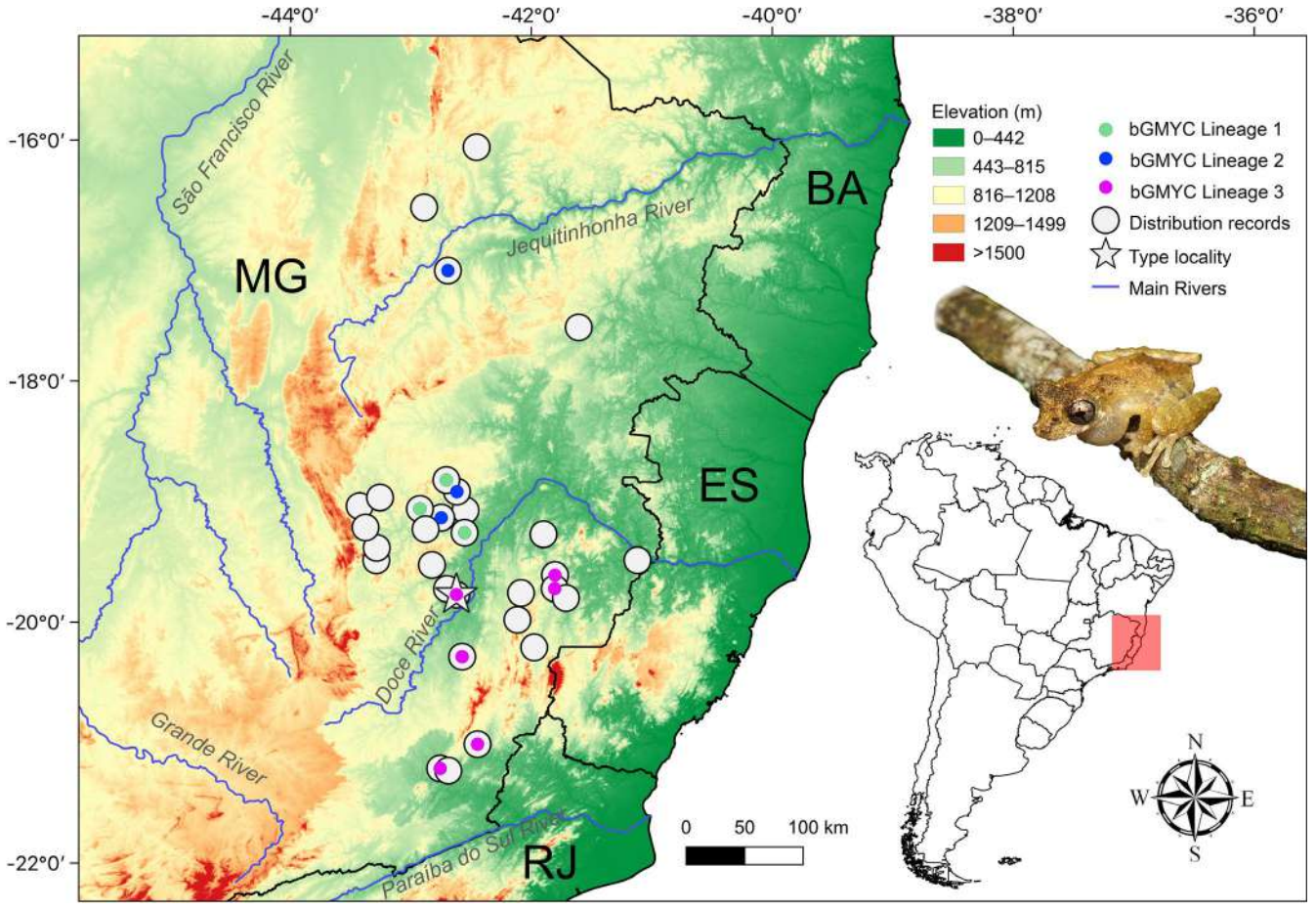

FIGURE 1 The known geographic distribution of *Ololygon carnevallii* is shown in gray circles, all genetic samples used in this study are illustrated by colored circles. All major rivers within southeastern Brazil are illustrated. Abbreviations are as follows, MG—Minas Gerais, BA—Bahia, ES—Espírito Santo, RJ—Rio de Janeiro. Inset is a picture of a live *O. carnevallii* (photo taken by P.S. Hote)

TABLE 1 Collecting localities, GenBank accession numbers for *COI* barcode sequence data, the corresponding lineage for each individual sample sequenced, and the voucher accession information for all samples included in molecular analyses

Locality	Latitude	Longitude	GenBank accession number	bGMYC lineages	Voucher specimen
Braúnas	−19.134111	−42.747583	MT571617	2	MZUFV16555
Caratinga/FMA	−19.722674	−41.806127	MT571611	3	MZUFV17195
Caratinga/FMA	−19.722674	−41.806127	MT571612	3	MZUFV17198
Caratinga/FMA	−19.722674	−41.806127	MT571613	3	MZUFV17196
Caratinga/PCH A. Branca	−19.610900	−41.803900	MT571629	3	MAP-T425
Caratinga/UNEC	−19.722786	−41.806104	MT571630	3	UFMG-A17150
Caratinga/UNEC	−19.722786	−41.806104	MT571631	3	UFMG-A17151
Cataguases	−21.215544	−42.756769	MT571618	3	MZUFV16585
Cataguases	−21.215544	−42.756769	MT571619	3	MZUFV16586
Cataguases	−21.215544	−42.756769	MT571620	3	MZUFV16584
Cataguases	−21.215544	−42.756769	MT571638	3	CT2574
Dores de Guanhães	−19.060838	−42.924610	MT571634	1	UFMG-A17270
Dores de Guanhães	−19.060838	−42.924610	MT571635	1	UFMG-A17271
Leme do Prado	−17.083300	−42.692500	MT571632	2	UFMG-A14146
Leme do Prado	−17.083300	−42.692500	MT571633	2	UFMG-A14147
Marliéria	−19.772821	−42.622016	MT571621	3	MZUFV17163
Marliéria	−19.772821	−42.622016	MT571622	3	MZUFV17165
Marliéria	−19.772821	−42.622016	MT571623	3	MZUFV17151
Marliéria	−19.772821	−42.622016	MT571624	3	MZUFV17171
Mesquita	−19.259635	−42.554123	MT571614	1	MZUFV17160
Mesquita	−19.259635	−42.554123	MT571615	1	MZUFV17159
Mesquita	−19.259635	−42.554123	MT571616	1	MZUFV17169
Mesquita	−19.259635	−42.554123	MT576018	3	MZUFV17169
Muriaé	−21.013611	−42.446667	MT571637	3	CT2571
Santo A. do Grama	−20.287625	−42.574178	MT571625	3	MAP1312
Sapucaia de Guanhães	−18.918369	−42.617560	MT571626	2	MAP1297
Sapucaia de Guanhães	−18.918369	−42.617560	MT571627	2	MAP1298
Sapucaia de Guanhães	−18.918369	−42.617560	MT571628	2	MAP1299
Sapucaia de Guanhães	−18.918369	−42.617560	MT576017	2	MAP1299
Virginópolis	−18.822224	−42.705019	MT571636	1	UFMG-A17279

Abbreviations: CT, Cytogenetics Lab Tissue Collection in the Beagle Lab, Universidade Federal de Viçosa; MAP, Mapinguari Lab from Universidade Federal de Mato Grosso do Sul; MZUFV, Museu de Zoologia João Moojen, Universidade Federal de Viçosa; UFMG, Zoological Collecation of the Universidade Federal de Minas Gerais.

2.3 | Phylogenetic analyses

We generated a Bayesian phylogeny of all of the *O. carnevallii* sampled individuals to asses phylogeographic structure across its distribution using BEAST v 2.6 (Bouckaert et al., 2014). The most appropriate substitution model for *COI* was inferred using jModelTest2 v 2.1.10 (Darriba et al., 2012). A *COI* sequence for *Ololygon humilis* was downloaded from GenBank (accession number KU234705) and used as an outgroup to root the gene tree (Faivovich et al., 2005). We used a Yule speciation prior, implemented a strict clock rate of 2% per Myr (Crawford, 2003), and ran the analysis for 2 million generations sampling every 2,000 generations. Tracer v 1.7.1 (Rambaut et al., 2018) was used to assess effective sample sizes of estimated parameters and stationarity, ensuring that all effective samples sizes were >200.

We implemented the Bayesian version of the general mixed yule-coalescent model (*bGMYC*; Pons et al., 2006; Reid & Carstens, 2012), to objectively define populations given the posterior distribution of gene trees. We subsampled 100 trees from the posterior distribution of the BEAST analysis, after a burn-in of 10% using the R package *evobiR* (Blackmon & Adams, 2015). *bGMYC* was run for 50,000 generations, with a 40,000 generation burn-in, and a thinning interval of 100. The threshold parameters were set to $t1 = 2$ and $t2 = 100$. To generate a point estimate of the number of distinct lineages from the *bGMYC* analysis, we used a posterior probability cutoff of 0.05.

2.4 | Current and hindcast ecological niche models

To generate ecological niche models (ENM), we georeferenced all unique collection localities for voucher specimens of *O. carnevallii* (a total of 34 localities). These voucher specimens were verified by examining specimens at the following institutions: Museu de Zoologia João Moojen, Universidade Federal de Viçosa (MZUFV), amphibian collection of Universidade Federal de Minas Gerais (UFMGA), Museu de Ciências Naturais, Pontifícia Universidade Católica de Minas Gerais (MCNAM), and Coleção Zoológica da Universidade Federal de Mato Grosso do Sul (ZUFMS-AMP) (Table S1). Locality records were spatially thinned so that no localities were within 10 km of each other using the R package *spThin* (Aiello-Lammens et al., 2015). We downloaded the Worldclim v1.4 (Hijmans et al., 2005) 19 climate variables at 2.5 arcmin resolution. *ENMtools* (Warren et al., 2017) was used to test for correlations among these variables and all correlated variables but one was removed where Pearson's correlations >0.7. This resulted in a total of six bioclimatic variables: mean annual temperature, mean diurnal range, isothermality, temperature seasonality, annual precipitation, and precipitation of driest month. To identify optimal model parameters to use in our ENM, we tested all combinations of feature classes with regularization multipliers between 0.5 and 4 at 0.5 intervals in *ENMeval* (Muscarella et al., 2014). Ecological niche models were constructed using *Biomod2* (Thuiller et al., 2013), using the best fit model parameters. We sampled 10,000 pseudoabsence points in a region surrounding the geographic distribution of *O. carnevallii* (circumscribed within: -49, -36, -24, -12) and Maxent v3.4.1 (Phillips et al., 2006) was used to construct ENMs using the six uncorrelated bioclim variables. Maxent was run with 25 evaluation runs, replicated for 5,000 iterations, and 25% of samples were used as a training data set for model evaluation. We then produced an average of these ENMs runs and projected this model to both last glacial maximum (LGM; 21 kya) and Mid-Holocene (6 kya) climate conditions using the community climate system model general circulation model (CCSM4; Gent et al., 2011) with the same six non-correlated bioclim variables. Regions of habitat stability through time were identified by stacking and averaging the current and two projected-paleo climate ENMs. Regions highlighted in these stacked projects were inferred to be regions of climate refugia through time for *O. carnevallii*. All ENMs were normalized between 0 and 1 and exported as ASCII files.

2.5 | Determinates of genetic structure

We sought to test what factors have been responsible for promoting population genetic differentiation within *O. carnevallii* and focused on the following variables: (1) geographic distance, (2) climate variables, (3) the predicted ecological niche, (4) Pleistocene refugia during the LGM, (5) mid-Holocene distribution, (6) habitat stability through time, (7) rivers as biogeographic barriers, (8) rivers promoting gene flow, (9) elevation, and (10) slope isolating populations. Each of these variables have been suggested to be responsible for population genetic divergence and may be acting together to reduce gene flow between populations (e.g., Myers et al., 2019; García-Rodríguez et al., 2020).

We downloaded shapefiles for rivers from the HydroSHEDS project (Lehner et al., 2008) and elevation (Jarvis et al., 2008) to examine how these features of the landscape have shaped genetic diversity. Slope was derived from elevation using *raster* in R (Hijmans & van Etten, 2012). These shapefiles were imported into R, converted to ASCII format using *raster* (Hijmans & van Etten, 2012), and values were normalized between 0 and 1. Here, higher values represent higher costs to dispersal across the landscape. For example, when rivers were tested as biogeographic barriers the rivers were set to 1, whereas when rives were considered to act as conductance they had lower values compared with the rest of the landscape. Higher elevations and steeper slopes were set to higher values. To assess the effect of geographic distances, we created an ASCII file with equal values in each cell. We then used Circuitscape v0.1.0 (Anantharaman et al., 2019; McRae et al., 2008) implemented in Julia (Bezanson et al., 2017) to generate distance matrices for our sampled localities given all of these potential factors (e.g., ENMs, rivers, elevation, and slope). Current ENMs, LGM projected ENMs, stability ENMs, and rivers were used as conductance surfaces in Circuitscape, while rivers, elevation, and slope were used as resistance surfaces. We used rivers as both conductance and resistance because large rivers have been proposed to be barriers to gene flow within the AF; however, given that anurans have an aquatic larval stage, it is possible that rivers may facilitate gene flow. Resistance and conductance distances were calculated as pairwise in Circuitscape. The climate variables used were the same six non-correlated variables used in the ENMs and here were used as the values extracted from each specimen collecting locality.

Using generalized dissimilarity modeling (GDM; Ferrier et al., 2007), we tested how environmental variation between collection localities and models of resistance or conductance matrices have contributed to genetic differentiation. This approach is an extension of matrix regression and can accommodate non-linear relationships between variables (Ferrier et al., 2007). We used *hierfstat* (Goudet, 2005) in R to generate a matrix of Fst values between all sampled geographic localities. This Fst matrix was then used as the response variable, and the *gdm* R package (Manion et al., 2016) was used to fit generalized dissimilarity models. This allowed us to test for correlations between genetic distances, the seven distance matrices based on ENMs and geography, in addition to testing for correlations with genetic distance and geographic distances (these distances were estimated with the *gdm* function where 'geo' was set to true) and current climate variables. We then used the gdm.varImp function in the *gdm* R package to perform model and variable significance testing and to estimate variable importance in our GDM.

3 | RESULTS

3.1 | DNA sequence data and phylogenetic analyses

COI sequence data were generated for 30 *O. carnevallii* specimens (Table 1) for a total of 551 base pairs when trimmed to the shortest sequence, with 187 variable sites, and 23 unique haplotypes.

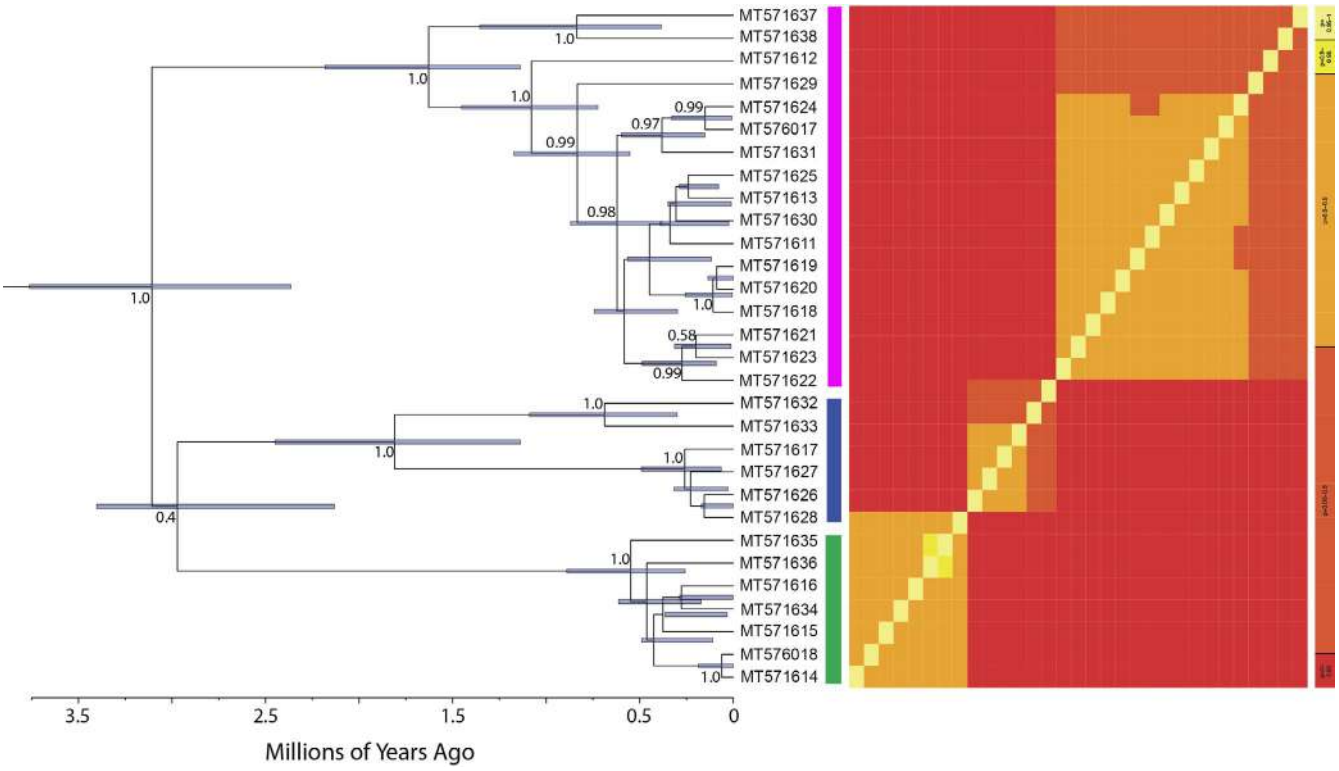

FIGURE 2 Phylogeny and divergence times estimated in Beast and results from bGMYC analysis. Tip labels correspond to samples listed in Table 1 and colored bars correspond to collecting localities in Figure 1. The colored bGMYC table corresponds to the posterior probability that the corresponding sequences are conspecific and illustrates any uncertainty in inferred population limits

These sequence data were uploaded to GenBank (Table 1), and the alignment is available in fasta format in the (Alignment S1). The best fit model of sequence evolution inferred from jModeltest2 was HKY+G. All effective sample size values from the BEAST gene-tree analysis were >200, suggesting stationarity. The timing of gene divergence within *O. carnevallii* began at ~3.5 mya (HPD: 2.6–4.2 mya; Figure 2). bGMYC analyses suggested that there are three distinct mtDNA lineages (Figure 2). These lineages are united with strong posterior probability; however, the relationships between them is unresolved with low posterior probability (Figure 2). Lineages 1 and 2 are distributed north of the Rio Doce and largely in sympatry, while lineage 3 is south of the Rio Doce with the exception of specimens from the type locality (Figure 1). Within lineage genetic variation is variable, with 10 haplotypes in lineage 3, 5 haplotypes in lineage 1, and 4 haplotypes in lineage 2.

3.2 | Ecological niche models

After thinning specimen locality data, we retained a total of 37 localities for ENM. The best fit feature class was hinge with a regularization multiplier of 1; however, the next best fit model differed by a deltaAIC of only 0.22 which was LQHP with a regularization multiplier of 1 (the AIC score of the next best model was >8; Table S2). We therefore ran two separate ENMs using both of these model settings. However, because the model projections were similar, we only discuss the Hinge 1 model (see Figure S1 for LQHP projections). The current ENMs performed well with AUC values of 0.91 and predicted the known geographic distribution of *O. carnevallii*. However, the models did predict suitable habitat to the southwest in the state of Minas Gerias and northern Sao Paulo where this taxon has not been collected (Figure 3). Both hindcast projections suggest shifts in suitable habitats for *O. carnevallii*. For example, at 6 kya the potential suitable habitat is much more expansive but with a shift in distribution to the interior of Brazil. While at 21 kya, the projected suitable habitat is shifted largely to the east of the current distribution (Figure 3). The stability map generated with all three projections suggests regions of suitable habitat in central and western Minas Gerais (Figure 3).

3.3 | Generalized dissimilarity modeling

The GDM with seven predictor variable matrices, current climate conditions, and geographic distance explained 25.8% of the genetic variation as measured by Fst. However, this model was not significant with a *p*-value of 0.12. The most important predictor variables were geographic distance and elevation; however, after

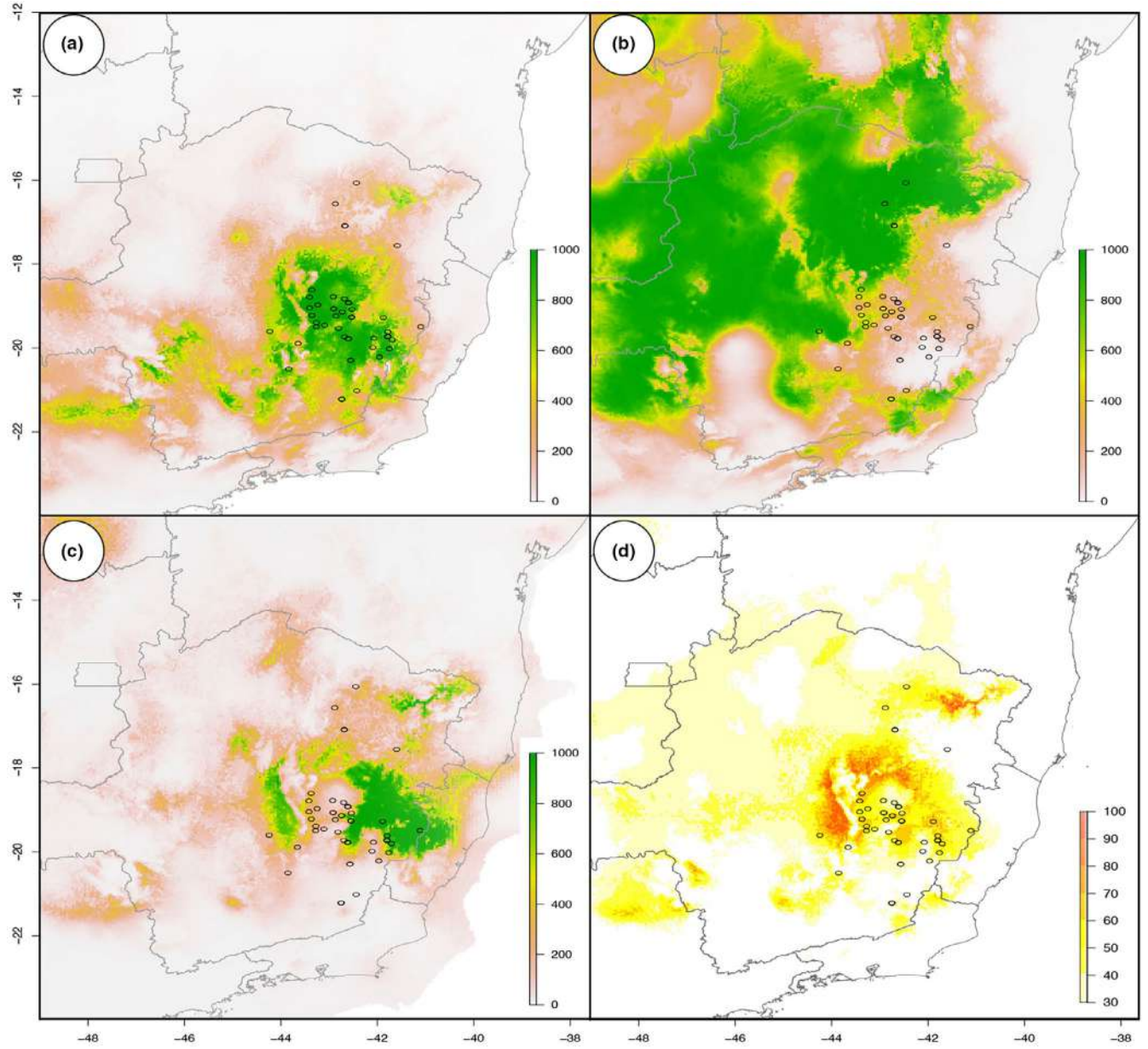

FIGURE 3 Ecological niche models for *Ololygon carnevallii*. (a) ENM under current climate conditions, (b) ENM projected on to mid-Holocene climate (6 kya), (c) ENM projected on to LGM climate conditions (21 kya), (d) Regions of climate suitability through time based on all three projections. In all projections, the distribution of *Ololygon carnevallii* is shown as hollow circles

permutation tests these were not significant (p-value > 0.14; for relationships between F_{ST} and predictor variables see Figure S2).

4 | DISCUSSION

The Rio Doce treefrog underwent diversification during the late Pliocene (~3.5 mya) forming three distinct mtDNA lineages. These lineages have remained distinct through multiple glacial cycles throughout the Pleistocene, irrespective of suitable habitats. Ecological niche models projected to the LGM and the mid-Holocene demonstrate that the potential suitable habitat for this taxon would have shifted through time, from east to west. However, given the timing of lineage formation, this was not a primary factor influencing divergence. While Pleistocene refugia have been important in promoting diversification in many AF taxa (Carnaval et al., 2009; De Mello Martins, 2011), many species have population divergence that predates this period. In more widespread species, population divergence into northern and southern lineages that diverged in the Pliocene has been demonstrated (Grazziotin et al., 2006; Thomé et al., 2010). Similarly, in these groups, hindcast ENMs have demonstrated reduced availability of suitable habitat during the Pleistocene but with limited congruence with population genetic structure (Thomé et al., 2010; Tonini et al., 2013). It is possible that

these earlier divergence times are due to Pliocene orogeny and habitat fragmentation with population divergence re-enforced by small scale habitat fragmentation during late Pleistocene climatic oscillations (Grazziotin et al., 2006; Thomé et al., 2010).

The timing of lineage formation with *O. carnevallii* further supports discordant diversification mechanisms in the AF. Because population and species divergence in the AF has been dated from the Pliocene to the Pleistocene, the mechanisms generating the species richness in this biome have been diverse. These seemingly constant rates of lineage divergence and accumulation in the AF biome are similar to patterns seen across the Neotropics, where the timing of divergence has been continuous throughout the Tertiary and Quaternary (Rull, 2008). These results from the AF also further support the notion that the processes that have given rise to the biodiversity of the Neotropics are complex (Costa, 2003; Turchetto-Zolet et al., 2013). These insights only point to the need for additional phylogeographic studies of AF endemics to better understand the complex diversification mechanisms of this biome.

Most phylogeographic studies within the AF have focused on the patterns and processes of diversification in widespread species (Carnaval et al., 2009; Paz et al., 2019; Thomé et al., 2010). The Rio Doce treefrog is narrowly distributed approximately within the Bahia refugium of the Atlantic Forest identified in Carnaval and Moritz (2008). Because the distribution of *O. carnevallii* is restricted to the Bahia refugium, it may be expected that there is no signature of Pleistocene climate fluctuations driving population divergence. Instead, we find deep divergence within this range restricted species that dates to the Pliocene. Because phylogeographic patterns are often species specific, associated with life-history traits and ecology (Papadopoulou and Knowles, 2016), additional phylogeographic studies of taxa with distributions restricted to these Pleistocene forest refugia are needed. This would ultimately clarify if shared historical processes versus species-specific responses have influenced diversification within a region of climatic stability. Furthermore, such studies on narrow endemics will shed light on what has been responsible for generating endemicity and influencing patterns of genetic variation at fine spatial scales. This may also highlight new regions of conservation concern that have been overlooked by broad-scale studies of widely distributed taxa.

The lineages we identify here began to diversify prior to the onset of global climatic fluctuations of the Pleistocene. However, irrespective of the timing of divergence among these clades, we explicitly tested for associations between patterns of genetic divergence and environmental suitability through time (Paz et al., 2019). ENMs do show reduced suitable habitat during the LGM and historically stable regions. However, there is no association between ENMs, whether current day, hindcast models or habitat stability through time. Furthermore, there was no association with other commonly cited variables that could be driving diversification (e.g., elevation, current climate variables, or rivers). While this suggests that these variables have not been important in driving genetic divergence within this taxon, it is also possible that there is not enough signal in the barcode sequence data that was generated here.

The Rio Doce is often cited as an important barrier to gene flow (Ribeiro et al., 2011; Thomé et al., 2010); however, this river has not been important in structuring population genetic divergence in *O. carnevallii*. This is demonstrated in both the geographic distribution of mtDNA lineages and results from generalized dissimilarity models. Therefore, it is possible that earlier tectonic uplift of montane regions within the Atlantic Forest were important for initial separation of lineages (Mello et al., 1999; Riccomini & Assumpção, 1999). Furthermore, because lineages 1 and 2 seem to co-occur north of the Rio Doce it is also possible that these lineages have diverged in sympatry or parapatry as the result of niche divergence. Future natural history studies focused on the differences in calls, timing of breeding, or habitat use could help to clarify divergence in niche between these two lineages. These lineages may have also diverged in the climatic niche that is modeled in ENMs. If this is the case, then treating all collecting localities as a single unit in the ENM could lead to over-projection of the model (e.g., Pearson et al., 2007). The future collection of specimens, and sequence data to assign these specimens to lineages, from new localities may allow for expanded tests of niche divergence. However, given the narrow distribution of this taxon new collecting localities may not lead to increased resolution in ENMs. A potential caveat here is that the phylogeographic patterns found in organelle versus nuclear genomes may not always be the same across a potential barrier (Toews & Brelsford, 2012). This highlights the need to generate additional genomic data to discriminate among opposing models of population divergence. Future studies incorporating genomic data (e.g., RADseq or sequence capture; Davey et al., 2011; Faircloth et al., 2012) with newer models of past climate (Brown et al., 2018) will be particularly important in understanding these processes.

ACKNOWLEDGEMENTS

EAM was funded by the Peter Buck and Rathbone Bacon Fellowship through the Smithsonian's National Museum of Natural History. HF and ER thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001. DJS thanks CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for his research fellowship (311492/2017-7). We thank Clodoaldo L Assis and Priscila Hote for their assistance with fieldwork and Paulo C. A. Garcia for providing *Ololygon carnevallii* tissues. We thank P. Hote for the photograph of these species in life.

DATA AVAILABILITY STATEMENT

The genetic data that support the findings of this study are openly available in GenBank, accession numbers MT571611–MT576018. An alignment of these sequences is also available in the supporting information.

ORCID

Edward A. Myers https://orcid.org/0000-0001-5369-6329
Diego José Santana https://orcid.org/0000-0002-8789-3061

REFERENCES

Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). SpThin: An R Package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, *38*(5), 541–545.

Aktas, C. (2015). *Haplotypes: Haplotype Inference and Statistical Analysis of Genetic Variation*. R Package Version, 1.

Amaro, R. C., Rodrigues, M. T., Yonenaga-Yassuda, Y., & Carnaval, A. C. (2012). Demographic processes in the montane Atlantic rainforest: molecular and cytogenetic evidence from the endemic frog *Proceratophrys boiei*. *Molecular Phylogenetics and Evolution*, *62*(3), 880–888.

Anantharaman, R., Hall, K., Shah, V., & Edelman, A. (2019). Circuitscape in Julia: High Performance Connectivity Modelling to Support Conservation Decisions. *ArXiv Preprint ArXiv:1906.03542*.

Archer, F. I., Adams, P. E., & Schneiders, B. B. (2017). Stratag: An r package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources*, *17*(1), 5–11.

Batalha-Filho, H., Fjeldså, J., Fabre, P.-H., & Miyaki, C. Y. (2013). Connections between the Atlantic and the Amazonian Forest Avifaunas Represent Distinct Historical Events. *Journal of Ornithology*, *154*(1), 41–50.

Batalha-Filho, H., & Miyaki, C. Y. (2016). Late Pleistocene divergence and postglacial expansion in the Brazilian Atlantic Forest: Multilocus Phylogeography of Rhopias Gularis (Aves: Passeriformes). *Journal of Zoological Systematics and Evolutionary Research*, *54*(2), 137–147.

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, *59*(1), 65–98.

Blackmon, H., & Adams, R. H. (2015). EvobiR: Comparative and Population Genetic Analyses. R package version 1.1. https://CRAN.R-project.org/package=evobiR

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*(4), e1003537.

Brown, J. L., Hill, D. J., Dolan, A. M., Carnaval, A. C., & Haywood, A. M. (2018). PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Scientific Data*, *5*(1), 1–9.

Brunes, T. O., Maria, T. C., Thomé, J. A., Haddad, C. F. B., & Sequeira, F. (2015). Ancient divergence and recent population expansion in a leaf frog endemic to the southern Brazilian Atlantic forest. *Organisms Diversity & Evolution*, *15*(4), 695–710.

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., & Moritz, C. (2008). Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography*, *35*(7), 1187–1201.

Carnaval, A. C., Waltari, E., Rodrigues, M. T., Rosauer, D., VanDerWal, J., Damasceno, R., Prates, I., Strangas, M., Spanos, Z., Rivera, D., Pie, M. R., Firkowski, C. R., Bornschein, M. R., Ribeiro, L. F., & Moritz, C. (2014). Prediction of phylogeographic endemism in an environmentally complex biome. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1792), 20141461.

Colombi, V. H., Lopes, S. R., & Fagundes, V. (2010). Testing the Rio Doce as a riverine barrier in shaping the Atlantic rainforest population divergence in the Rodent Akodon Cursor. *Genetics and Molecular Biology*, *33*(4), 785–789.

Cooke, G. M., Chao, N. L., & Beheregaray, L. B. (2012). Divergent natural selection with gene flow along major environmental gradients in Amazonia: Insights from genome scans, population genetics and phylogeography of the characin fish Triportheus albus. *Molecular Ecology*, *21*(10), 2410–2427.

Costa, L. P. (2003). The Historical Bridge between the Amazon and the Atlantic Forest of Brazil: A study of molecular phylogeography with small mammals. *Journal of Biogeography*, *30*(1), 71–86.

Crawford, A. J. (2003). Relative rates of nucleotide substitution in frogs. *Journal of Molecular Evolution*, *57*(6), 636–641.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). JModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*(7), 499–510.

De mello martins, F. (2011). Historical biogeography of the Brazilian Atlantic Forest and the Carnaval-Moritz Model of Pleistocene Refugia: What Do Phylogeographical Studies Tell Us? *Biological Journal of the Linnean Society*, *104*(3), 499–509.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797.

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, *61*(5), 717–726.

Faivovich, J., Haddad, C. F. B., Garcia, P. C. A., Frost, D. R., Campbell, J. A., & Wheeler, W. C. (2005). Systematic review of the frog family hylidae, with special reference to hylinae: phylogenetic analysis and taxonomic revision. *Bulletin of the American Museum of Natural History*, *2005*(294), 1–240.

Ferrier, S., Manion, G., Elith, J., & Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*, *13*(3), 252–264.

García-Rodríguez, A., Guarnizo, C. E., Crawford, A. J., Garda, A. A., & Costa, G. C. (2020). Idiosyncratic responses to drivers of genetic differentiation in the complex landscapes of Isthmian Central America. *Heredity*, 1–15.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., & Vertenstein, M. (2011). The community climate system model version 4. *Journal of Climate*, *24*(19), 4973–4991.

Goudet, J. (2005). Hierfstat, a Package for R to Compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*(1), 184–186.

Grazziotin, F. G., Monzel, M., Echeverrigaray, S., & Bonatto, S. L. (2006). Phylogeography of the Bothrops Jararaca Complex (Serpentes: Viperidae): Past Fragmentation and Island Colonization in the Brazilian Atlantic Forest. *Molecular Ecology*, *15*(13), 3969–3982.

Hayes, F. E., & Sewlal, J.-A.- N. (2004). The Amazon River as a dispersal barrier to passerine birds: Effects of river width, habitat and taxonomy. *Journal of Biogeography*, *31*(11), 1809–1818.

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

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978.

Hijmans, R. J., & van Etten, J. (2012). *Raster: Geographic Analysis and Modeling with Raster Data*. R Package Version 1:9–92.

Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). *Hole-Filled SRTM for the Globe Version 4. CGIAR-CSI SRTM 90mDatabase*. Retrieved from http://Srtm.Csi.Cgiar.Org

Koroiva, R., Rodrigues, L. R. R. & Santana, D. J. (2020). DNA barcoding for identification of anuran species in the central region of South America. *PeerJ*, *8*, e10189.

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, *89*(10), 93–94.

Leite, Y. L. R., Costa, L. P., Loss, A. C., Rocha, R. G., Batalha-Filho, H., Bastos, A. C., Quaresma, V. S., Fagundes, V., Paresque, R., Passamani, M., & Pardini, R. (2016). Neotropical forest expansion during the last glacial period challenges refuge hypothesis. *Proceedings of the National Academy of Sciences*, *113*(4), 1008–1013

Lorenzen, E. D., Heller, R., & Siegismund, H. R. (2012). Comparative phylogeography of African Savannah Ungulates 1. *Molecular Ecology*, *21*(15), 3656–3670.

Lyra, M. L., Haddad, C. F. B., & de Azeredo-Espin, A. M. L. (2017). Meeting the challenge of DNA barcoding neotropical amphibians: polymerase chain reaction optimization and new COI primers. *Molecular Ecology Resources*, *17*(5), 966–980.

Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., & Fitzpatrick, M. C. (2016). *Gdm: Functions for Generalized Dissimilarity Modeling*. R Package.

Mascarenhas, R., Miyaki, C. Y., Dobrovolski, R., & Batalha-Filho, H. (2019). Late pleistocene climate change shapes population divergence of an Atlantic Forest Passerine: A model-based phylogeographic hypothesis test. *Journal of Ornithology*, *160*(3), 733–748.

McRae, B. H., Dickson, B. G., Keitt, T. H., & Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, *89*(10), 2712–2724.

Mello, C. L., Metelo, C. M. S., Suguio, K., & Kohler, H. C. (1999). Quaternary sedimentation, neotectonics and the evolution of the doce river middle valley lake system (Southeastern Brazil). *Revista do Instituto Geológico*, *20*(1-2), 29–36.

Menezes, L., Canedo, C., Batalha-Filho, H., Garda, A. A., Gehara, M., & Napoli, M. F. (2016). Multilocus Phylogeography of the Treefrog Scinax eurydice (Anura, Hylidae) Reveals a Plio-Pleistocene Diversification in the Atlantic Forest. *PLoS One*, *11*(6), e0154626.

Mitchell, M. W., Locatelli, S., Sesink Clee, P. R., Thomassen, H. A., & Gonder, M. (2015). Environmental variation and rivers govern the structure of chimpanzee genetic diversity in a biodiversity hotspot. *BMC Evolutionary Biology*, *15*(1), 1.

Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (2014). ENM Eval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for maxent ecological niche models. *Methods in Ecology and Evolution*, *5*(11), 1198–1205.

Myers, E. A., Xue, A. T., Gehara, M., Cox, C. L., Davis, A. R., Rabosky, J.-L.-E., Martínez-Gómez, J. E., & Burbrink, F. T. (2019). Environmental heterogeneity and not vicariant biogeographic barriers generate community-wide population structure in desert-adapted snakes. *Molecular Ecology*, *28*(20), 4535–4548.

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

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*, *113*(29), 8018–8024.

Paz, A., Spanos, Z., Brown, J. L., Lyra, M., Haddad, C., Rodrigues, M., & Carnaval, A. (2019). Phylogeography of Atlantic Forest Glassfrogs (Vitreorana): When geography, climate dynamics and rivers matter. *Heredity*, *122*(5), 545–557.

Pearson, R. G., Raxworthy, C. J., Nakamura, M., & Townsend Peterson, A. (2007). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of biogeography*, *34*(1), 102–117.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3–4), 231–259.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., & Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595–609.

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.

Reid, N. M., & Carstens, B. C. (2012). Phylogenetic estimation error can decrease the accuracy of species delimitation: A Bayesian implementation of the general mixed yule-coalescent model. *BMC Evolutionary Biology*, *12*(1), 196.

Ribeiro, C., Priciane, J. P., Lemos-Filho, R. S., de Oliveira Buzatti, M. B. L., & Myriam, H. (2016). Species-Specific Phylogeographical Patterns and Pleistocene East-West Divergence in Annona (Annonaceae) in the Brazilian Cerrado. *Botanical Journal of the Linnean Society*, *181*(1), 21–36.

Ribeiro, R. A., Lemos-Filho, J. P., Ramos, A. C. S., & Lovato, M. B. (2011). Phylogeography of the Endangered Rosewood Dalbergia Nigra (Fabaceae): Insights into the evolutionary history and conservation of the Brazilian Atlantic Forest. *Heredity*, *106*(1), 46–57.

Riccomini, C., & Assumpção, M. (1999). Quaternary tectonics in Brazil. *Episodes*, *22*(3), 221–225.

Rull, V. (2008). Speciation timing and neotropical biodiversity: the tertiary-quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology*, *17*(11), 2722–2729.

Sotelo-Muñoz, M., Maldonado-Coelho, M., Svensson-Coelho, M., dos Santos, S. S., & Miyaki, C. Y. (2020). Vicariance, dispersal, extinction and hybridization underlie the evolutionary history of Atlantic Forest Fire-Eye Antbirds (Aves: Thamnophilidae). *Molecular Phylogenetics and Evolution*, *148*, 106820.

Thom, G., Smith, B. T., Gehara, M., Montesanti, J., Lima-Ribeiro, M. S., Piacentini, V. Q., Miyaki, C. Y., & do Amaral, F. R. (2020). Climatic dynamics and topography control genetic variation in Atlantic Forest montane birds. *Molecular Phylogenetics and Evolution*, *148*, 106812

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

Thuiller, W., Georges, D., & Engler, R. (2013). *Biomod2: Ensemble Platform for Species Distribution Modeling*. R Package Version.

Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, *21*(16), 3907–3930.

Tonini, J. F. R., Costa, L. P., & Carnaval, A. C. (2013). Phylogeographic structure is strong in the Atlantic forest; predictive power of correlative paleodistribution models, not always. *Journal of Zoological Systematics and Evolutionary Research*, *51*(2), 114–121.

Turchetto-Zolet, A. C., Cruz, F., Vendramin, G. G., Simon, M. F., Salgueiro, F., Margis-Pinheiro, M., & Margis, R. (2012). Large-scale phylogeography of the disjunct neotropical tree species Schizolobium parahyba (Fabaceae-Caesalpinioideae). *Molecular Phylogenetics and Evolution*, *65*(1), 174–182.

Turchetto-Zolet, A. C., Pinheiro, F., Salgueiro, F., & Palma-Silva, C. (2013). Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, *22*(5), 1193–1213.

Wallace, A. R. (1854). On the monkeys of the Amazon. *Annals and Magazine of Natural History*, *14*(84), 451–454.

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, 5649–5662.

Warren, D., Dinnage, R., & Matzke, N. (2017). *ENMTools: Analysis of Niche Evolution Using Niche and Distribution Models*. R Package (Version 0.2).

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

Alignment S1. Sequence alignment of *COI* generated and used in this study.

Table S1. List of all museum numbers of *Ololygon carnevallii* specimens examined and georeferenced for ENMs.

Table S2. Model parameter fit for ENMs obtained from ENMeval.

Figure S1. Current ENM for *O. carnevallii* obtained using second best fit model parameters (see Table S2).

Figure S2. Correlations between F_{ST} and potential explanatory variables from GDM.

How to cite this article: Myers EA, Folly H, Ragalzi E, Feio RN, Santana DJ. Late Pliocene population divergence and persistence despite Pleistocene climatic fluctuations in the Rio Doce snouted Treefrog (*Ololygon carnevallii*). *J Zool Syst Evol Res*. 2021;00:1–11. https://doi.org/10.1111/jzs.12454