ORIGINAL ARTICLE

Ancient divergence and recent population expansion in a leaf frog endemic to the southern Brazilian Atlantic forest

Tuliana O. Brunes[1,2] · Maria Tereza C. Thomé[3] · João Alexandrino[4] ·
Célio F. B. Haddad[3] · Fernando Sequeira[1]

Abstract The evolutionary history of Neotropical organisms has been often interpreted through broad-scale generalizations. The most accepted model of diversification for the Brazilian Atlantic forest (BAF) rely on putative historical stability of northern areas and massive past habitat replacement of its southern range. Here, we use the leaf frog *Phyllomedusa distincta*, endemic to the southern BAF, to better understand diversification patterns within this underexplored rainforest region. We used an integrative approach coupling fine-scale sampling and multilocus sequence data, with traditional and statistical phylogeographic (multilocus approximate Bayesian computation) methods to explore alternative hypotheses of diversification. We also employed species paleodistribution modeling to independently verify habitat stability upon a spatially explicit model. Our data support two divergent lineages with coherent geographic distribution that span throughout northern and southern ranges. Demographic estimates suggested the Southern lineage has experienced a recent population expansion, whereas the Northern lineage remained more stable. Hypothesis testing supports a scenario of ancient vicariance with recent population expansion. The paleodistribution model revealed habitat discontinuity during the Last Glacial Maximum (LGM) with one area of putative stability within the range of the Northern lineage. Evidence on genetic structure, demography, and paleodistribution of *P. distincta* support a historically heterogeneous landscape for the southern BAF, with both areas of forest stability and regions where forest occupation is probably recent. We also associate the southern end of the Cubatão shear zone with a phylogeographic break in the BAF. Taken together, our results argue for the idea of multiple mechanisms generating diversity in this biome and underscore the need of fine-scale data in revealing more detailed pictures.

Keywords Amphibians · Approximate Bayesian computation · Diversification · Phylogeography · Multilocus · Paleodistribution modeling

✉ Tuliana O. Brunes
brunestuliana@gmail.com

[1] CIBIO-InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Laboratório Associado, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal

[2] Departamento de Biologia, Faculdade de Ciências, Universidade do Porto (UP), 4169-007 Porto, Portugal

[3] Departamento de Zoologia, Instituto de Biociências, Universidade Estadual Paulista (UNESP), 13506-900 Rio Claro, São Paulo, Brazil

[4] Departamento de Ciências Biológicas, Universidade Federal de São Paulo (UNIFESP), 09972-270 Diadema, Brazil

Introduction

To describe and understand patterns of biological diversity in the Neotropics has been a challenge for many evolutionary biologists, with the focus being on unveiling the tempo and mode of diversification processes (Rull 2008). Encompassing enormous habitat heterogeneity, high numbers of endemic species, and a frightening level of deforestation (Myers et al. 2000; Ribeiro et al. 2009), the Brazilian Atlantic forest (BAF) has been the focus of most recent studies in diversification of South American taxa (Turchetto-Zolet et al. 2013). Advances in genetic and spatial analyses fomented a substantial increase in our knowledge of species diversification, especially by combining statistical phylogeography and species distribution modeling

(SDM; reviewed in Alvarado-Serrano and Knowles 2014). In the BAF, such tools allowed for the construction of a spatially explicit hypothesis of habitat persistence (Carnaval and Moritz 2008) and its rigorous testing based on genetic data (Carnaval et al. 2009). This resulted in a model where organisms found refuge in the north and central areas of the BAF during the Last Glacial Maximum (LGM) and subsequently colonized the southern region after Holocene rainforest recovery (Carnaval and Moritz 2008; Carnaval et al. 2009). Nonetheless, a set of studied taxa show patterns that support southern habitat persistence, highlighting instead the role of rivers and neotectonic activity during the Late-Tertiary and the Quaternary as main promoters of diversification (Amaral et al. 2013; Grazziotin et al. 2006; Pellegrino et al. 2005; Thomé et al. 2010).

Organisms widespread in the BAF have been frequently considered as best models to uncover the evolutionary history of the region. However, a recent ordination-based analysis of climatic data identified the northern and southern BAF as two distinct climatic regions (Carnaval et al. 2014), raising the question whether studies of taxa of more restricted distributions may be more appropriate to unveil the main diversification patterns of each region. Furthermore, whereas studies on widespread taxa frequently suffer from coarse sampling, extensive revisions suggest a need for fine-scale data to avoid broad generalizations over more complex histories (Martins 2011; Silva et al. 2012). Here, we use the leaf frog *Phyllomedusa distincta* Lutz, 1950 to better understand diversification patterns of species endemic to southern BAF. *Phyllomedusa distincta* is the only one of five species in the *Phyllomedusa burmeisteri* group with a range restricted to ombrophilous dense forest. A previous study found two highly supported mitochondrial clades showing contrasting levels of genetic variability. Insufficient sampling prevented further investigation of the demographics causing this pattern, and authors tentatively associated this divergence with geomorphologic events in southern BAF (Brunes et al. 2010). Here, we substantially extended the geographic and genetic sampling of *P. distincta* to combine traditional and statistical methods of phylogeographic analysis with species paleodistribution modeling. Our specific goals include (1) to verify if previously reported mtDNA structure is also presented in nuclear DNA; (2) estimate historical demographic parameters to construct and test hypotheses of contrasting scenarios of diversification; and, finally, (3) compare phylogeographic inference with independent spatially explicit hypothesis of habitat persistence obtained through species paleodistribution modeling.

Materials and methods

Sampling and sequencing procedures

We obtained a total of 126 tissue samples of *P. distincta* from field collecting trips and tissue donations from Brazilian herpetological collections (see Acknowledgments and Online Resource 1, Table S1). Samples consisted of liver, toe clips, or muscle preserved in 100 % ethanol. To avoid misclassification due to the presence of hybrid individuals between *P. distincta* and *Phyllomedusa tetraploidea* that occur in Ribeirão Branco, south of São Paulo state (Haddad et al. 1994), we previously confirmed the ploidy number of all individuals from this locality by cytogenetic analysis (Gruber et al. 2013; Sanae Kasahara, unpublished). We worked with a total of 31 localities (sampling localities, SL) with elevation ranging from sea level to ~915 m altitude (Table 1 and Fig. 1). We georeferenced all samples with GPS or used the coordinates of the closest town.

Genetic markers included one fragment of the mitochondrial NADH dehydrogenase subunit 2 gene (ND2), and three nuclear fragments: (1) a segment of exon 2 and intron 2 of the cellular myelocytomatosis (C-myc2), (2) a segment of exon 2 of chemokine receptor 4 (CXCR4), and (3) a segment of β-fibrinogen intron 7 (β-fibint7). We obtained whole genomic DNA through tissue samples digested in lysis buffer and Proteinase K and using QIA Quick DNEasy columns (Qiagen, Inc.) according to the manufacturer's protocol. We performed PCR amplification and sequencing of ND2 and β-fibint7, and of C-myc2 and CXCR4 following Brunes et al. (2010) and Brunes et al. (2014), respectively. We downloaded additional sequences from GenBank (Faivovich et al. 2010; Brunes et al. 2010, 2014; see Online Resource 1, Table S1). We edited alignments and corrected by eye in BIOEDIT 7.0.5.2 (Hall 1999).

Polymorphisms, pairwise distance, and recombination

We calculated polymorphism values in DNAsp 5.10 (Librado and Rozas 2009) for both the entire dataset and within main mitochondrial and nuclear groups including the number of segregating sites (S), number of haplotypes (h), haplotype diversity (Hd), number of singleton nucleotide variants (Sn), and the population mutation parameter theta, computed from the number of S (θw; Watterson 1975). We generated 95 % confidence intervals with 10,000 coalescent simulations under the standard neutral model.

To resolve the allele phase of nuclear markers we used the software PHASE 2.1 (Stephens et al. 2001) with the assistance of SeqPHASE (Flot 2010). We performed multiple independent runs for each gene with different seeds for the random-number generator and 1.0×10^6 iterations with the default values. We checked haplotype estimation through analysis of consistency across runs. We selected the haplotype reconstructions with higher probabilities. We further evaluated the presence of recombination events through the difference of sums of squares (DSS) test implemented in TOPALi 2.5 (Milne et al. 2004) using all phased haplotypes, a window size of 70 bp, steps 10 bp long, and 100

Table 1 Sampling locality information, sample size, and mitochondrial haplotype code of *Phyllomedusa distincta*

SL	Locality	ST	Longitude	Latitude	N	h (ND2)
1	Carapicuíba[a]	SP	−46.8356	−23.5225	2	1
2	São Paulo[a]	SP	−46.6279	−23.5752	1	1
3	Pilar do Sul	SP	−47.7164	−23.8131	1	2
4	Juquiá	SP	−47.6785	−24.3280	8	1, 2, 3
5	Eldorado	SP	−48.0804	−24.5337	5	1, 4
6	Pariquera-Açu[a]	SP	−47.8811	−24.7150	1	1
7	Iguape[a]	SP	−47.5414	−24.6981	4	1, 4
8	Iporanga[a]	SP	−48.7003	−24.5328	3	1, 4, 5
9	Ribeirão Branco[a]	SP	−48.7430	−24.3586	8	1, 2, 6
10	Adrianópolis	PR	−49.0315	−24.6404	6	1, 4, 5, 7
11	Piraquara	PR	−49.0444	−25.4422	2	8
12	Antonina[a]	PR	−48.7119	−25.4286	1	9
13	Guaratuba[a]	PR	−48.5747	−25.8828	4	2, 8
14	Garuva	SC	−48.8887	−25.9807	8	2, 10
15	São Bento do Sul	SC	−49.3786	−26.2503	2	10
16	Guaramirim[a]	SC	−49.0032	−26.5099	1	10
17	Timbó[a]	SC	−49.2594	−26.8075	5	10
18	Blumenau[a]	SC	−49.0911	−27.0289	3	10, 11
19	Apiúna[a]	SC	−49.4142	−27.0980	2	10, 12
20	Vidal Ramos[a]	SC	−49.3304	−27.3604	1	10
21	Barra Velha[a]	SC	−48.6918	−26.7242	4	10
22	Santa Lídia[a]	SC	−48.6537	−26.8330	8	10
23	Porto Belo[a]	SC	−48.5168	−27.1238	6	10
24	Florianópolis "North"	SC	−48.5069	−27.4585	11	10, 13
25	Florianópolis "Center"	SC	−48.5101	−27.5966	1	10
26	Florianópolis "South"	SC	−48.5438	−27.7307	1	14
27	Angelina	SC	−49.0049	−27.5885	3	10, 15
28	São Martinho	SC	−48.8998	−28.1880	1	10
29	Imbituba	SC	−48.8536	−28.3111	2	16, 17
30	Treviso[a]	SC	−49.4575	−28.5156	8	10
31	Dom Pedro de Alcântara[a]	RS	−49.8928	−29.3806	5	10

ST Brazilian states, *SP* São Paulo, *PR* Paraná, *SC* Santa Catarina, *RS* Rio Grande do Sul

[a] Indicates localities with individuals used in nuclear analyses (POFAD and STRUCTURE)

bootstrap repetitions. We used Alignment Transformation Environment, ALTER (Glez-Peňa et al. 2010), to transform datasets to the various input formats for subsequent analyses.

Genetic structure

Following Salzburger et al. (2011), we used the haplotype genealogy method to examine patterns of haplotype distribution of *P. distincta*. By turning phylogenetic trees into networks, this method avoids the presence of reticulation and, therefore, permits a clear visualization of relationships among highly similar haplotypes. Thus, we produced haplotype genealogies for both mitochondrial and nuclear fragments starting from a maximum parsimony (MP) tree though

DNAPARS available in PHYLIP package 3.69 (Felsenstein 2005). Then, we converted the most parsimonious trees into a haplotype genealogy in a beta version of Haploviewer (available at http://www.cibiv.at/~greg/haploviewer).

We evaluated the nuclear structure using a combined distance matrix (POFAD; Joly and Bruneau 2006) and a Bayesian clustering method (STRUCTURE; Pritchard et al. 2000). For both, we only included individuals sequenced for a least two of the three nuDNA fragments. This strategy resulted in a total of 22 individuals from 18 localities (Table 1). We calculated the *p*-distance (uncorrected) between individuals for each nuDNA fragment in MEGA 5.2.2 and combined into a multilocus distance matrix in POFAD. We evaluated both the original and a standardized matrix (all fragments are scaled

Fig. 1 Geographic distribution of *Phyllomedusa distincta* in the Southern Brazilian Atlantic forest (**a**) and mitochondrial haplotype genealogy derived from ND2 sequences (**b**). Sampling localities and haplotype numbers are indicated (see Table 1). The *circle* area of each haplotype is proportional to its frequency. Mutations are edges. Percentages indicate pairwise distance between haplogroups. Brazilian states: *SP* São Paulo, *PR* Paraná, *SC* Santa Catarina, *RS* Rio Grande do Sul
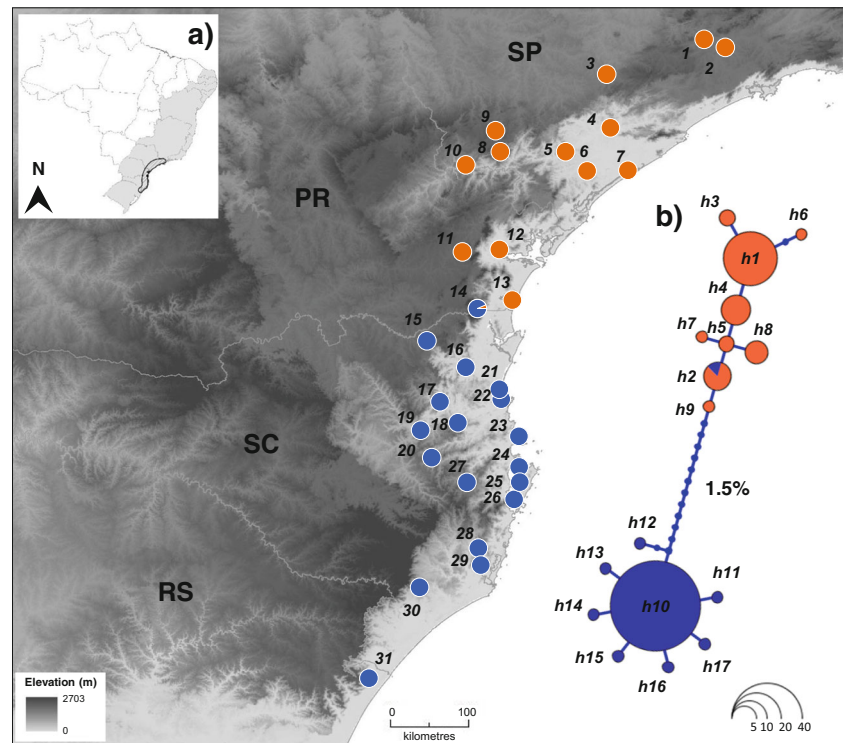
to the largest distance in the matrix). We visualized the multilocus distance network in Splits Tree 4.12.3 (Huson and Bryant 2006) via the NeighborNet method. We inferred population clusters with STRUCTURE (Pritchard et al. 2000) via allele frequency converted through the program xmfa2struct (available at http://www.xavierdidelot.xtreemhost.com/clonalframe.htm). We performed analysis under the admixture ancestry model, with five independent runs for each *K* ranging from 1 to 10. We discarded the first 1×10^5 MCMC iterations as burn-in and counted the next 25×10^4 generations. We found the best *K* value via the online program Structure Harvester v.0.6.93 (Earl and VonHoldt 2012) by monitoring the estimated log posterior probability of the data [Ln Pr $(X | K)$] (Pritchard et al. 2000) and estimating the second-order rate of change of the likelihood function (ΔK) (Evanno et al. 2005). Finally, we assembled the results of the five independent runs in the program CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007).

Divergence time, gene flow, and demographic parameters

We applied an isolation-with-migration model (IM; Wakeley and Hey 1997; Nielsen and Wakeley 2001; Hey and Nielsen 2007) to multilocus data as implemented in the program IMa2 (Hey 2010) to explore the speciation history of *P. distincta*. We used the HKY model of evolution (Hasegawa et al. 1985) for both mitochondrial and nuclear DNA. We then ran the program under Metropolis Coupled MCMC, using ten chains with linear heating mode. We conducted multiple preliminary

runs to assess mixing of the chains, as well as to determine appropriate priors for the parameters. After this, we ran the program three times with different random seeds. For each simulation, the length was $>10 \times 10^6$ steps, where the first 10×10^5 was discarded as burn-in. We checked convergence of the Markov chain simulations to the stationary distribution by monitoring multiple independent runs using different random number seeds (similar posterior distributions for each parameter across independent runs) and by assessing effective sample sizes values (ESS > 100), time series plots, and swapping rates between chains over the course of the run. We adopted a strict molecular clock and uniform priors on substitution rates of 0.0095 substitutions/site/year for ND2, estimated by Crawford (2003) for Craugastoridae species and previously used to estimate splits time in the *P. burmeisteri* species group (Brunes et al. 2010). For the nuclear genes, we used the mutation rate scalars estimated by the program. We converted divergence time, current and ancestral population sizes (θ), and population migration rates (2 Nm) to demographic units following author's instructions. We assumed a generation time of 1 year.

To evaluate if mitochondrial and nuclear groups present deviations from the neutral theory, we computed Tajima's *D* (Tajima 1989), Fu's *Fs* (Fu 1997), and R^2 test (Ramos-Onsins and Rozas 2002) with significance of values checked through 10,000 coalescent simulations in DnaSP 5.10 (Librado and Rozas 2009). We used the extended Bayesian skyline plot (EBSP) to reconstruct the historical demography over time of each lineage. This coalescent method is relatively sensitive

to estimate the population size history from multiple unlinked loci, even with a small number of individuals (Heled and Drummond 2008). The use of independent loci in the Bayesian skyline approach can reduce the coalescent error and increase the ability to detect population bottlenecks (Ho and Shapiro 2011; but see Heled and Drummond 2008), but the common violation of the panmixia assumption can produce misleading results (see Heller et al. 2013). To avoid this, we followed the advice of Heller et al. (2013) and carefully balanced our sampling by using individuals from several populations.

We conducted analysis in BEAST 1.7.4 (Drummond et al. 2012) with the whole molecular dataset (one mtDNA and three nuDNA fragments) using the linear demographic function. We selected models of nucleotide evolution in jModelTest v.0.1.1 (Posada 2008) under the Akaike information criterion (AIC; Akaike 1974). We applied TN93 for ND2 and C-myc2, and HKY model for CXCR4 and β-fibint7. We then searched for the best model of molecular clock rate variation (Strict Clock or Relaxed Clock: Uncorrelated Lognormal) for each fragment performing preliminary analyses (data not shown). After checking the performance and accuracy of each preliminary analysis in Tracer 1.5 (available at http://tree.bio.ed.ac.uk/software/tracer/), we selected Strict Clock for ND2 and C-myc2, and Relaxed Clock: Uncorrelated Lognormal for β-fibint7 and CXCR4. We fitted the substitution rate with the ND2 calibration as in IMa2 analysis (see above). To estimate the prior for the mean distribution of population sizes, we performed a preliminary analysis as suggested by the authors (results not shown). We started by conducting an mtDNA run with a coalescent prior and constant population size, and then we used the obtained mean value (0.13) as prior in the final run with a log-normal distribution and standard deviation of 0.5. We performed three replicate runs with 100 million generations, with trees being sampled every 10,000 generations and the first 10 % of the samples being discarded as burn-in. We checked convergence through ESS values (>200) also in the software Tracer 1.5. We used Python scripts provided by the authors (available at http://beast.bio.ed.ac.uk/Tutorials) to plot the results and interpreted that assuming a generation time of 1 year.

Testing explicit scenarios of diversification

To apply a statistical phylogeographic approach, we performed a multilocus ABC analysis testing which of a set of alternative scenarios best fit the observed data. We designed four plausible scenarios for the diversification history of *P. distincta*. We gathered information on the habitat and studied system (e.g., Thomé et al. 2014), including suggestions of putative population structure based on mitochondrial data from previous work (Brunes et al. 2010), demographic changes expected under proposed LGM forest dynamics (Carnaval

and Moritz 2008), and the results of the initial demographic analyses performed here (population expansion and gene flow). The four scenarios (Fig. 2) were as follows: (A) a null hypothesis of absence of structure combined with little or no habitat fluctuation, where a single population shows long-term persistence possibly followed by moderate Holocene expansion ("long-term persistence"); (B) a combination of the effects of a barrier and little habitat fluctuation, in a scenario where two populations were formed by an ancient (pre-LGM) vicariant event possibly followed by moderate Pleistocene retraction and Holocene expansion ("ancient vicariance"); (C) a scenario where population structure is driven by recent habitat fragmentation, with an ancestral population splitting into two geographically restricted populations during the LGM, followed by strong Holocene expansion ("two-refuge"); and (D) a scenario where the formation of two populations occurs through the recent (post-LGM) colonization of a formerly inhabited region by migrants from a stable population (e.g., Carnaval et al. 2009; "single-refuge") (see below for scenarios details).

For each scenario, we simulated equal numbers of genealogies, selected the simulations that best fit the observed data (rejection step) with the aid of summary statistics (see below), and checked for the fit of accepted genealogies. We interpreted the proportion of selected genealogies from each alternative scenario as a direct measure of its posterior probability (see Pritchard et al. 2000) and calculated the Bayes factor to access inference power (Jeffreys 1961). The parameters included in each scenario were ancestral population size, divergence time, population size of each subpopulation, migration, and exponential growth. The priors for total population size and divergence time were provided per locus. For total population size, we first ran a short analysis with broad priors (0.1–50) sampling values from uniform distributions. After the rejection step, we extracted mean values and standard deviations from the posterior to be used as Gaussian priors for the definitive analysis (ND2, mean=13.01/SD= 11.69; β-fibint7, mean=12.22/SD=9.35; and CXCR4, mean=6.51/SD=5.84). For the divergence time in scenario B, we extracted mean values from IMa2 (t_A=600,000 years). Because the right tail of the approximate posterior density curve failed to reach zero (see "Results" section), we arbitrarily chose large intervals (SD=300,000 years). For LGM divergences (scenarios C and D) and time of expansion in all scenarios (growth), we used a mean of 21,000 years (t_R and t_E; SD=10,000 years). For all other parameters, we used uniform priors, with the following intervals: subpopulation sizes varied from 0.4 to 0.6 of total population size, asymmetrical migration (from the Northern to the Southern group; see IMa2 results) varied from zero to 4, and growth varied according to the scenarios. In long-term persistence and ancient vicariance (scenarios A and B), we set a growth ratio interval of 0.60–1 (60–100 % of current population sizes remained prior to
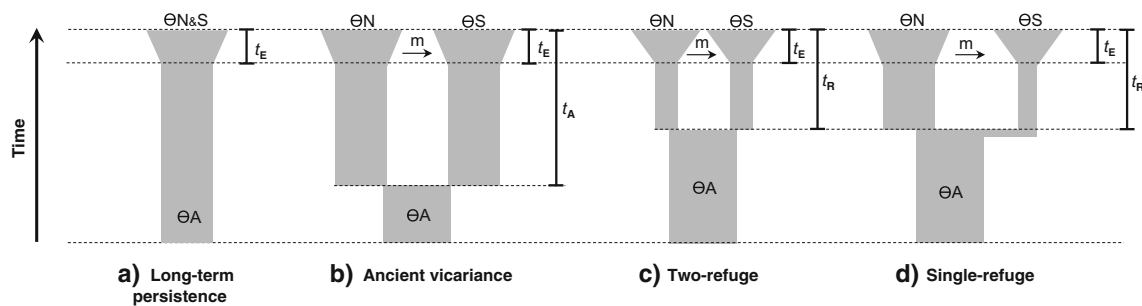
Fig. 2 Possible demographic models of the diversification of *Phyllomedusa distincta* in the southern Brazilian Atlantic forest used in the approximate Bayesian computation analysis. Model parameters: θ= 4 Neμ of ancestral (*A*), Northern (*N*), and Southern (*S*) groups (see text for details); m=2 Nm with *arrow* indicating direction of migration (IMa2 result: gene copies migrating from the Northern to the Southern group= 0.7; 95 % HPD=0–4); t_A=ancient divergence time (IMa2 result: 600,000 years; SD=300,000), t_R=refugia divergence time during the Last Glacial Maximum, and t_E=time of expansion starting in the Holocene up to the present (21,000 years; SD=10,000)

demographic expansion), which allows from moderate bottle-necks to total stability. For the two-refuge scenario, we set a growth ratio of 0.01–0.20 for both populations, simulating more drastic demographic reductions (only 1–20 % of their current effective size remained prior to demographic expansion), whereas for the single-refuge, we kept 0.60–1 for the first population (Northern) and 0.01–0.20 for the second population (Southern).

We generated 1,000,000 genealogies per locus for the definitive analysis (100,000 for prior selection) through the program ms (Hudson 2002), for which we obtained summary statistics with a modified version of SampleStats (Hudson 2002). We automated this process using Perl scripts (Online Resource 1, Perl script). After preliminary analyses, we removed the marker C-myc2 to avoid unrealistic results due the low number of gene copies and low variability. We performed the rejection step in msReject (Hickerson et al. 2007) with a tolerance of 0.0001 (0.01 for prior selection). Following Pelletier and Carstens (2014), we used three distinct measures of nucleotide diversity (within population one, within population two, and between populations) as observed summary statistics, which were calculated on a per locus basis in DnaSP 5.10 (Librado and Rozas 2009). Finally, we performed the rejection step both jointly and per locus, evaluating model fit using principal component analysis (PCA) of the summary statistics to check dispersion away from observed values (e.g., Camargo et al. 2012). With the PCA, we assume that goodness of fit of a particular scenario will be represented by a central position of the score of the observed data relative to the distribution of scores from the genealogies simulated for this scenario.

Species distribution models

To evaluate if climatically suitable areas for our focal taxa were stable during the Last Interglacial period (120 ka, LIG) and, subsequently, experienced drastic range reductions during the Last Glacial Maximum (21 ka, LGM) following theoretical expectations from the refuges theory (Haffer 1997), we generated correlative maps of the past and current potential distribution of *P. distincta* through the maximum entropy algorithm implemented by MaxEnt 3.3.3k (Phillips et al. 2006). MaxEnt's method has been widely used in SDMs combined with phylogeographic studies (Elith et al. 2006, 2011, and references therein). We used 31 occurrence points (Table 1 and Fig. 1) and the climatic layers available in the WorldClim database at ~5 km² spatial resolution (Hijmans et al. 2005). After preliminary evaluations (see below), we projected the current potential distribution for the LIG (120 ka) and for the LGM (21 ka), both models provided by the Community Climate System Model (CCSM) (Otto-Bliesner et al. 2006, 2007). Although the Model of Interdisciplinary Research on Climate (MIROC) is also extensively used to perform LGM projections (e.g., Camargo et al. 2013), we concentrated our LGM analysis in CCSM because it is derived from colder climatic conditions when compared to MIROC, which is initialized from climatic data similar to modern conditions (Weber et al. 2007). Following Anderson and Raza (2010), we calibrated the models in a defined study area. We used a buffer of 300 km from the current range distribution (endemic of the Southern Ombrophylous forest) of *P. distincta* (Fig. 6a). Under current climatic conditions variables, we selected five out of 19 bioclimatic with pairwise Pearson's correlation lower than 0.85 (data not shown). The selected variables included Mean Diurnal Range - BIO2, Temperature Seasonality - BIO4, Annual Precipitation - BIO12, Precipitation Seasonality - BIO15, and Precipitation of Warmest Quarter - BIO18. It is worth mentioning that some of these variables are well known to be highly associated with eco-physiological characteristics of frogs (e.g., temperature variation and distribution of precipitation). We estimated the contribution of each variable to explain the potential species distribution by performing a jackknife test. Following recent evidence that species-specific tuning can improve the performance of MaxEnt models (Anderson and Gonzalez 2011; Radosavljevic and Anderson 2013), we varied the regularization multiplier values (0.25,

0.50, 1.00, 1.50, 2.00, 4.00, 6.00, 8.00, and 10.00) to prevent overfitting. We performed statistical analysis with random test percentage of occurrence points (75 % for model prediction and 25 % for model validation) and ten replicates. We based quantitative model evaluation on the threshold-independent receiver operating characteristic (ROC) curve identifying the lowest values of average differences between training and test area under the curve (AUC). In addition, we also evaluated the model performance via (1) qualitative visual inspections of the response curves looking for smoother and regular patterns (e.g., Camargo et al. 2013) and (2) the predictive maps compared to expert knowledge of the distribution of BAF phytophysiognomies (IBGE 2012) and the known distribution of *P. distincta* (Haddad et al. 2008). We then used the best parameter set to generate the current and past climate projections of *P. distincta*. The final maps are presented with the original probability distribution generated by MaxEnt, considering the gradual changes from suitable to unsuitable habitats in nature. We avoided the use of an arbitrary threshold to produce binary maps of species presence/absence because thresholds may be subjective and constitute an additional source of error in ecological niche modeling (Liu et al. 2005; Phillips et al. 2006).

Results

Polymorphisms, pairwise distance, and recombination

Levels of polymorphism in mitochondrial and nuclear fragments showed substantial variation (Table 2). The 993-bp ND2 fragment revealed 17 haplotypes defined by 28 segregating sites. The total nuclear DNA variation based on each fragments revealed 26 (β-fibint7), 10 (CXCR4), and 4 (C-myc2) haplotypes defined by 20, 10, and 5 segregating sites, respectively. The highest nucleotide diversity was found in the ND2 fragment (0.0075) and the lowest in CXCR4 (0.0026). Both ND2 and β-fibint7 showed higher values of θw (0.0053 and 0.0056, respectively), followed by C-myc2 and CXCR4 (with 0.0045 and 0.0035). Comparing both groups and loci, nucleotide diversity was higher in the Northern group than in the Southern group, with the exception of CXCR4. For the ND2 fragment, the lowest nucleotide diversity (0.0003) was observed in the Southern group (ND2; 95 % CI 0–0.0011) based on 71 individuals. Comparatively, nucleotide diversity in the Northern was ~5-fold higher (0.0016; 95 % CI 0.0002–0.0043). The DSS test did not detect any recombination in any of the nuclear fragments after phasing. Due to the low number of individuals analyzed here, summary statistics of C-myc2 should be interpreted with caution.

Genetic structure

Single gene haplotype genealogies presented patterns that differ between mitochondrial and nuclear fragments (Fig. 1b and Online Resource 2, Fig. S1). The mtDNA ND2 haplotype genealogy showed two highly divergent clades (*p*-distance= 1.5 %), with one grouping most of the samples located in São Paulo and Paraná states (Northern lineage), and another grouping all individuals from Santa Catarina and Rio Grande do Sul states (Southern lineage). Shared haplotypes were found in only one locality on an intermediate geographic location (SL 14; one haplotype from the Northern lineage and seven from the Southern lineage). Haplotypes from the Northern lineage are slightly structured while the Southern lineage showed a star-like topology, with the exception of one unique internal haplotype from Apiúna (h12; Fig. 1b). The structure in nuclear fragments was not as geographically well defined, with all nuclear haplotype genealogies sharing variation between the two lineages defined by the mtDNA topology (see Online Resource 2, Fig. S1). Shared variation was less evident in C-myc2 and β-fibint7 than in CXCR4 as genealogies of these markers showed a tendency to separate haplotypes from both lineages. The β-fibint7 genealogy also presented some external haplotypes in the Southern region (SL 17, 18, 19, 23, 30) that grouped with the Northern ones. The group formed by the remaining haplotypes from the Southern region showed less genetic variation when compared with the Northern ones, resembling the pattern in the mtDNA structure.

When all the nuclear loci were jointly analyzed, POFAD showed a strong tendency to separate individuals from the Northern and Southern region with the presence of some intermediates (SL 18, 19–20; Fig. 3a). The STRUCTURE analysis recovered two clusters (*K*=2), according to the log posterior probability of the data [Ln Pr (*X* | *K*)] and ΔK (Online Resource 2, Fig. S2), with geographic distributions that also form a Northern and a Southern group (Fig. 3b). While the Northern cluster showed high exclusive values of assignments, the Southern cluster presented some intermediate individuals (SL 18–20, 23, and 30; ~*Q* values of 0.33–0.67) and one entirely allocated to the opposite cluster (SL 19).

Divergence time, gene flow, and demographic parameters

Analyses using the Isolation-with-Migration model showed posterior distribution of parameter estimates consistent across replicate runs with ESS values >120. In two cases (divergence time and theta), the right tail of the approximate posterior density curves (Online Resource 2, Fig. S3) failed to reach zero. Nevertheless, these analyses suggested the divergence between Northern and Southern groups occurred at ~0.61 Myr, during the Pleistocene. Estimates of effective population size for both groups were very similar (Ne ~2.6×10^5, 95 % HPD ~$1.2–4.1 \times 10^5$). The 95 % HPD of the ancestral

Table 2 Fragment information, summary statistics, and neutrality tests within the main mitochondrial clades and nuclear groups (see POFAD results) of *Phyllomedusa distincta*

Fragment	Groups	Length	Polymorphisms							Neutrality tests		
			N	SS	Sn	h	% Hd	θ [95 % CI]	π [95 % CI]	Tajima's D	Fu's Fs	R^2
ND2	All	993	118	28	10	17	66	0.005 [0.0041–0.0126]	0.0075 [0.0025–0.0172]	–	–	–
	North		47	9	4	9	72	0.002 [0.0004–0.0034]	0.0016 [0.0002–0.0043]	−0.6026 ns	−2.1381 ns	0.0885 ns
	South		71	9	9	8	19	0.0019 [0–0.0008]	0.0003 [0–0.0011]	−2.2743 ***	−9.4589***	0.0491 ns
β-fibint7	All	600–604	102	20	6	26	86	0.0056 [0.0025–0.01]	0.0056 [0.0014–0.0136]	–	–	–
	North		60	15	5	12	75	0.0053 [0.0018–0.0085]	0.0045 [0.001–0.0114]	−0.4381 ns	−1.6585 ns	0.0903 ns
	South		42	12	1	15	70	0.0046 [0.0011–0.0076]	0.0038 [0.0007–0.0096]	−0.5149 ns	−6.7597*	0.0955 ns
CXCR4	All	609	62	10	2	10	78	0.0035 [0.0007–0.0052]	0.0026 [0.0003–0.0069]	–	–	–
	North		26	3	2	4	44	0.0013 [0–0.0026]	0.0009 [0–0.0029]	−0.8212 ns	−1.2269 ns	0.1081 ns
	South		36	8	2	8	83	0.0032 [0.0008–0.0071]	0.0034 [0.0005–0.0088]	0.1187 ns	−1.6195 ns	0.1125 ns
C-myc2	All	318–323	18	5	1	4	65	0.0045 [0.0009–0.0108]	0.0044 [0.0003–0.0124]	–	–	–
	North		12	4	1	3	53	0.0041 [0–0.0113]	0.0045 [0–0.0128]	0.3853 ns	1.5263 ns	0.1770 ns
	South		6	1	0	2	53	0.0014 [0–0.0054]	0.0016 [0–0.0059]	0.8505 ns	0.6254 ns	0.2666 ns

N number of sequences or phases of nuclear genes, *SS* number of segregating sites, *Sn* number of singleton nucleotide variants, *h* number of haplotypes, *%Hd* haplotype diversity in percent, θ population mutation parameter and respective 95 % confidence interval, π nucleotide diversity and respective 95 % confidence interval

*$P<0.05$; ***$P<0.001$; *ns* no significance

effective population size starts in the upper limit of the current estimates (9×10^5, 95 % HPD=4.1×10^5–1.3×10^6), suggesting that both groups experienced a reduction of the effective population size at some period in the past. Estimates of gene flow showed an asymmetrical result, detecting gene copies migrating only from the Northern to the Southern group (2 Nm=0.7, 95 % HPD=0–4, LLR=5.623, $P<0.01$).

Deviations from the standard neutral model (Table 2) were significant only in the Southern group, as observed for ND2 (Tajima's D and Fu's Fs) and β-fibint7 (Fu's Fs).

Although the R^2 test did not show a significant p value for the putative expansion of the Southern group as in other tests, values were lower in the Southern group than in the Northern group for the ND2 fragment. The multilocus EBSP analysis of historical demography over time showed a trend of smooth and gradual population increase for the Northern group starting around the early Holocene according to the median (Fig. 4a). Nearly at the same time, the Southern group started an abrupt population growth with a 10-fold increase in effective population size

Fig. 3 Multilocus nuclear structure of *Phyllomedusa distincta* based on the three nuclear fragments (CXCR4, β-fibint7, C-myc2): **a** genetic distance network; *circles* were colored following the two major mitochondrial groups (see Fig. 1); **b** STRUCTURE results based on nuclear allele frequencies for $K=$ 2. Individuals are represented as *bars*, with colors representing the proportion of assignment. Sampling localities are presented near the *circles* and *left* of each *bar* (see Table 1 and Fig. 1)
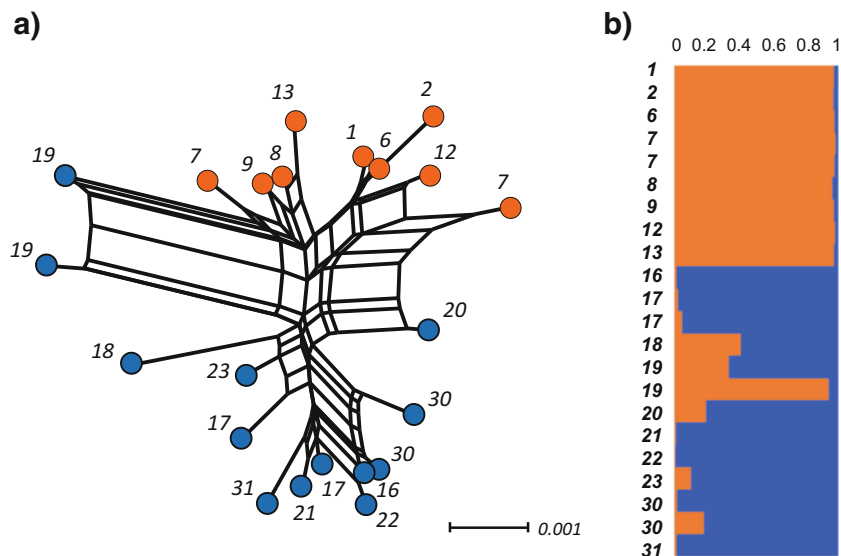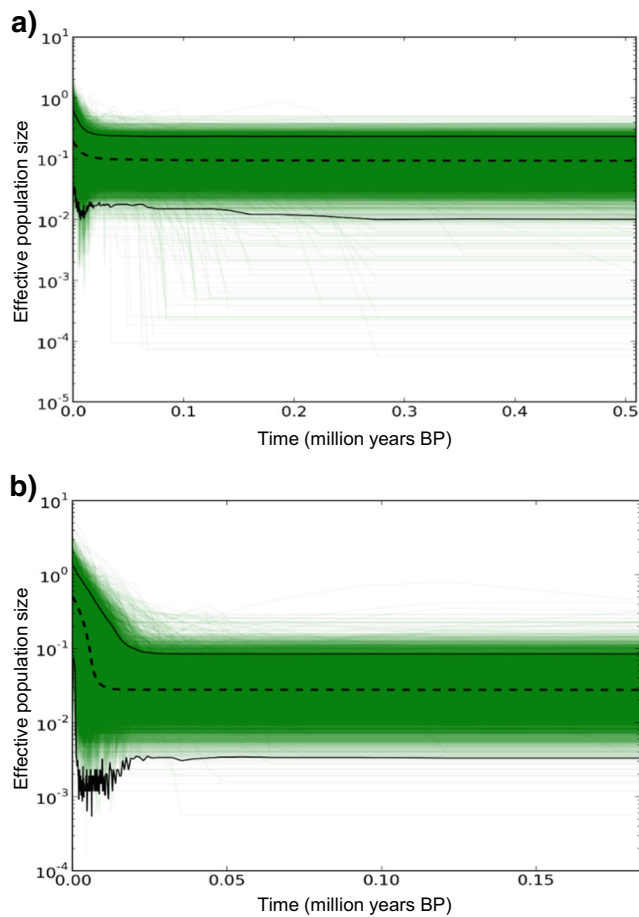
Fig. 4 Extended Bayesian Skyline plots of *Phyllomedusa distincta* groups: **a** Northern group and **b** Southern group. *Dashed line* represents the median population size (multiply by thousand and by generation time of 1 year) and the *lightly shaded gray* the 95 % HPD. The accumulation of *green lines* represents the full posterior distribution. The *y*-axis is in logarithmic scale

(Fig. 4b). Confidence intervals for both analyses are large and include the possibility of no growth.

Testing explicit scenarios of diversification

The scores representing the vectors of the summary statistics calculated from the observed data were central to the vectors of the simulated genealogies (Fig. 5) indicating a good fit of the models to the data for ABC analyses. After the rejection step, scenarios A (long-term persistence) and D ("single-refuge") received zero or very low posterior probability values in both per locus and multilocus analyses. The posterior probabilities of the models were congruent between β-fibint7 and CXCR4, supporting the ancient vicariance scenario (Fig. 2b), whereas ND2 rejection was inconclusive, presenting similar values for scenario B ("ancient vicariance") and C ("two-refuge"). Multilocus rejection was conclusive, supporting the scenario B with substantial support according to the Bayes factor (Table 3).

Species distribution models

The species-specific tuning results showed a slight quantitative variation, but a wide range of qualitative variation according to the regularization parameter used (Online Resource 2, Table S1). The average test AUC values were high (0.88–0.98), and the difference between AUC in training and test analyses varied from −0.019 to 0.04. Overall, the average omission rates based on the MTP test showed the lowest and more constant values with the regularization multiplier set to 1.0, similar to Radosavljevic and Anderson (2013). The species-specific tuning of *P. distincta* presented regular/smoother response curves and more realistic predictive suitability areas using the regularization parameter set to 1.0 (see Online Resource 2, Table S1 and Fig. S4 to response curves). The jackknife test identified the most important climatic variables to explain the distribution of *P. distincta* in the BAF (Online Resource 2, Fig. S5). Of the five climatic variables used, the mean diurnal range (BIO 2), the temperature seasonality (BIO 4), and the precipitation seasonality (BIO 15) presented smaller differences both in gain (test and training) and in AUC values.

The present time predictive map showed an area with high climatic suitability compatible with the current distribution of the *P. distincta*. Areas of overprediction appeared in central regions of Rio Grande do Sul and Rio de Janeiro states (Fig. 6a). Similar overpredictions were observed in both paleoclimatic projections (Fig. 6c). The present time and the LIG projections yielded similar distribution maps. The LGM projection predicted intense fragmentation with two small isolated areas within the current distribution of *P. distincta*, one in the current coastal region of Paraná state reaching to the north of Santa Catarina state, and another located in the limit of the LGM land masses in central Santa Catarina state (the sea level was lower than today).

Discussion

It is well demonstrated that phylogeographical investigations require extensive geographic sampling and multilocus analysis, and that they may be strengthened by using complementary paleomodeling of species distributions (Alvarado-Serrano and Knowles 2014). In this study, we were able to clarify the evolutionary history of *P. distincta*, an endemic leaf frog species from the Southern Brazilian Atlantic forest. Overall, our results indicate the presence of two divergent lineages with coherent geographic distribution within *P. distincta* range. Demographic estimates (from summary statistics, neutrality tests, and ESBP) suggested that the Southern lineage has experienced a recent population expansion
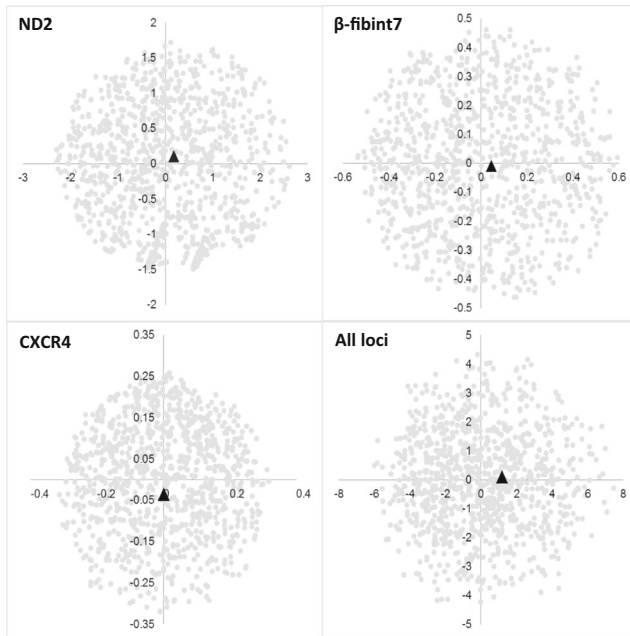
Fig. 5 Model checking in ABC analyses of *Phyllomedusa distincta*. Ordination plots show summary statistics vectors of simulated (*gray circles*) and observed (*black triangles*) data after the rejection step. We conducted the rejection step per locus and for all loci jointly (see text for details)

that started in the Holocene. The testing of alternative historical demographic models (ABC analysis) favored a scenario of ancient (mid-Pleistocene) vicariance with moderate Holocene population expansions, whereas paleodistribution modeling inferred past habitat fragmentation, corroborating the signals of recent demographic expansion for the Southern lineage. Comparing our results to findings from different co-distributed organisms, we provide a broader perspective of species diversification in southeastern BAF.

Genetic structure

The mtDNA ND2 haplotype genealogy confirmed the early findings of Brunes et al. (2010) in revealing the presence of

two highly divergent lineages (*p*-distance=1.5 %) with coherent geographic distribution within the *P. distincta* range (Fig. 1a, b). The substantial sampling increase relative to a previous study enabled us to detect one locality with the presence of haplotypes from both mtDNA lineages, which may indicate gene flow between lineages. In general, single locus nuclear genealogies did not recover straightforwardly the two lineages observed for mtDNA (Online Resource 2, Fig. S1), which could be explained by the retention of ancestral polymorphism and/or gene flow. Such incongruence between cytoplasmic and nuclear gene genealogies is expected due to the higher effective population sizes and lower mutation rates of nuclear loci as previously demonstrated by Edwards and Beerli (2000).

Multilocus approaches based on the coalescent theory and/or allele frequency have been successful in combining the information of different genes to reveal population groups even with reduced genome sampling and in the presence of incomplete lineage sorting and/or hybridization (Fusinatto et al. 2013; Thomé et al. 2012). Here, the two lineages of *P. distincta* revealed by mtDNA analysis were unambiguously corroborated by multilocus analyses, albeit with some evidence of intermediate genomes suggesting either incomplete lineage sorting or mixing (Fig. 3a, b). Whereas ancestral polymorphism is necessarily present at some time window along the divergence process itself, gene flow may or may not occur along the evolutionary scenario that generates deep divergence between populations. For the data presented here, the distinction between the processes was better investigated through multilocus models of isolation in the presence of gene flow and hypothesis testing of alternative historical demographic scenarios (IMa2 and ABC, see below).

Divergence time, gene flow, and demographic parameters

The isolation with migration model implemented in IMa2 inferred the mean divergence time between Northern and Southern group to be ~600,000 years, well into the Pleistocene. The boundary between the two diverged groups geographically coincides at the present day with the Cubatão shear zone, listed as a tectonic fault with Quaternary surface

Table 3 Posterior probabilities and model support of historical demographic models of *Phyllomedusa distincta* assessed by an approximate Bayesian computation analysis (see Fig. 2)

Scenario	ND2	β-fibint7	CXCR4	All loci
Long-term persistence	0	0	0.08	0
Ancient vicariance	0.40	0.74	0.55	0.86
Two-refuge	0.59	0.18	0.31	0.13
Single-refuge	0.01	0.08	0.06	0.01
Bayes factor *K*	1.5	4	1.8	6.3
Strength of evidence	Barely worth mentioning	Barely worth mentioning	Barely worth mentioning	Substantial

The rejection step was performed per locus and with the three loci jointly. Bayes factor $K=Pr(D \mid M_1)/Pr(D \mid M_2)$. Scale for interpretation of K follows Jeffreys (1961)
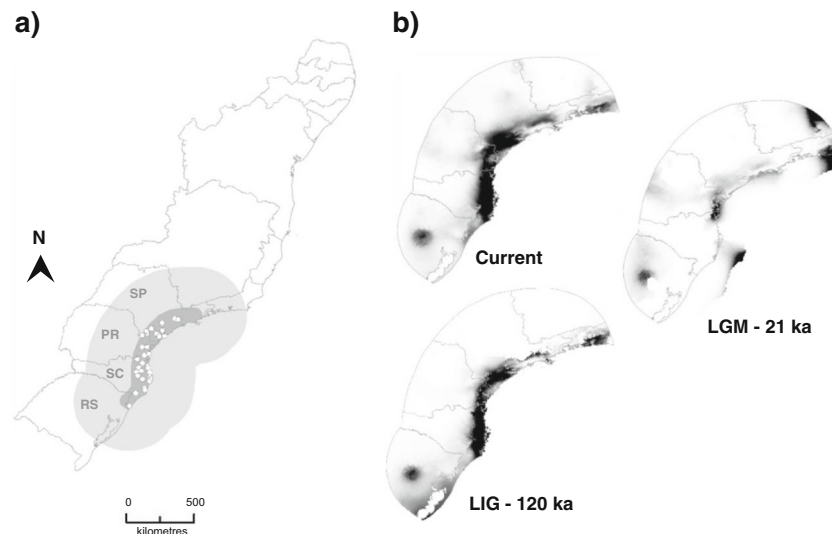
Fig. 6 MaxEnt models for *Phyllomedusa distincta* in southeastern Brazilian Atlantic forest: **a** study area in grayscale (see text for details) and **b** predicted climatically suitable areas for current time, Last Glacial Maximum period (LGM), and Last Interglacial period (LIG). *Dark colors* represent high climatic suitability and *soft colors* represent low climatic suitability. Both paleoclimatic models were performed under the Community Climate System Model (CCSM). *White dots* indicate occurrence points used in the analysis. Brazilian states: *SP* São Paulo, *PR* Paraná, *SC* Santa Catarina, *RS* Rio Grande do Sul

rupture (Saadi et al. 2002). No information is however available on the timing and extent of ruptures in this geological complex. Migration from Northern to Southern group was much higher than in the opposite direction. This result could perhaps be associated with higher demographic stability inferred for the Northern group, but to adequately address the hypothesis of asymmetric gene flow would require a detailed analysis of the contact zone, which is not within the scope of this study.

Multilocus estimates of demographical change across time suggest distinct histories for the two groups of populations of *P. distincta* in southern BAF. EBSP for both groups show large confidence intervals and also depict trends of recent growth that vary in intensity. Populations in the southern range of *P. distincta* (Southern group) may have experienced higher demographic instability than populations in the northern parts of the range (Northern group). Evidence from other analyses are in accordance with this result, suggesting at least some expansion for the Southern group as observed for ND2 (Tajima's *D* and Fu's *Fs*) and β-fibint7 (Fu's *Fs*). This result reflects the overall pattern of high values of segregating sites and haplotypes when compared with the nucleotide diversity. Comparing both groups and all loci, nucleotide diversity was higher in the Northern group than in the Southern group for most fragments, with the exception of CXCR4. The inconsistency of some diversity statistics across loci seems to reflect differences in mutation rates, as "faster" markers such as mtDNA and introns showed significant values more often, while differences among tests may be also explained by the better performance of Fu's *Fs* test with larger sample sizes (see Ramos-Onsins and Rozas 2002).

Our results are partially concordant with studies reporting on the demographics of other southern BAF taxa. In southern BAF, populations of some organisms were shown to have experienced more demographic instability than northern populations (treefrogs of the genus *Hypsiboas*, Carnaval et al. 2009; antbirds of the genus *Myrmeciza*, Amaral et al. 2013; mammals of the genus *Akodon*, Valdez and D'Elía 2013), while higher demographic stability was inferred for the southern populations of toads of the *Rhinella crucifer* group (Thomé et al. 2014). It is difficult, however, to exactly compare between these studies because the population boundaries are not the same. Most studies have revealed population boundaries to be in the central or southern state of São Paulo, while the genetic boundary here described is the first located in the more southern transition between the states of Paraná and Santa Catarina.

Testing explicit scenarios of diversification

Multilocus data and multidisciplinary approaches have been considered of major importance to design and test hypotheses exploring the evolutionary history of BAF organisms due to its geological/climatic complexity and scarcity of independent evidence (Amaral et al. 2013; Thomé et al. 2014). Here, we used multilocus data in an ABC framework to test for alternative diversification scenarios within the leaf frog species *P. distincta*. Because of its comparative nature (model probabilities are relative), the choice of scenarios is perhaps the most important decision in ABC. It is known that inference power decreases as more scenarios are included (Pelletier and Carstens 2014), whereas omitting scenarios may render high

probabilities to unrealistic scenarios, conditioned on the concurrent models. We balanced our decision with a complementary approach by including both scenarios based on the literature (Fig. 2a, c, and d) and models based on parameters estimated for the focus organism (Fig. 2b). The adequacy of our selected scenarios was verified after the model fit step. Our simulations showed a good fit to the observed data, demonstrating that acceptance of models was not a result of choosing the best among a set of poorly fitting scenarios.

Very low (or null) posterior probabilities both for the rejection performed per locus and jointly allowed us to clearly reject two scenarios for the evolution of *P. distincta* populations: (A) the long-term persistence of one single demographically stable population (Fig. 2a) and the recent colonization of southern areas of the range from northern populations (Fig. 2d). Two other scenarios could not be readily rejected as explanatory for the genetic diversity observed within *P. distincta*. The ancient divergence of two relatively stable populations (Fig. 2b) received substantial support from the joint loci rejection. On the other hand, the scenario of LGM divergence following severe bottleneck effects resulting from habitat fragmentation followed by Holocene demographic expansions (Fig. 2c) should not be completely ruled out because it had as much support as hypothesis A from the mtDNA data analysis. Despite this, the highest posterior probability obtained for joint loci rejection favors the ancient vicariance as the best fitting scenario for *P. distincta* evolutionary history.

Species distribution models

Paleoclimatic distribution modeling of *P. distincta* suggested that there was a dramatic reduction of suitable areas for the species occurrence during the LGM. Within the current distribution of *P. distincta*, one single area in coastal Paraná state would represent a refuge as the other putative area of LGM species persistence is today covered by the Atlantic Ocean. This result is not completely at odds with the findings from our genetic analyses. Rather, there is a geographic overlap of the single remaining area in coastal Paraná state and the distribution of the Northern lineage, which is also more according to EBSP and neutrality tests. However, our paleomodeling results predict hardly any habitat left for the Southern lineage during the LGM, which would imply a scenario of recent colonization that was formerly rejected in ABC. Thus, here we believe that the paleomodeling results are showing support for the impact of late Pleistocene climatic oscillations on demography of *P. distincta* (e.g., Álvarez-Presas et al. 2014), in particular, on populations from the Southern lineage where their location seems to have been much more dynamic than the Northern ones. Nevertheless, we cannot exclude the possibility that our analysis was affected by the coarse ability of species distribution paleomodeling to deliver detailed

predictions of historical distributions in the South America region (see Rojas et al. 2009 and Collevatti et al. 2013).

The evolutionary importance of the southern BAF region (including some southern areas) in shaping the present-day biodiversity has been highlighted since the end of the 1970s based on different types of data (Porto et al. 2013 and references therein). Interestingly, our result runs parallel with recent findings of Carnaval et al. (2014) based on modeling partitioned-BAF analysis. The potential LGM coastal refugia in Paraná state found here was recently considered an area with higher values of phylogeographic endemism (Carnaval et al. 2014). Data from pollen cores and soil chemical isotope analyses represent a more realistic, albeit spatially discontinuous, source of paleoecological evidence. Palynological evidence suggests that cooler and dryer climates had an effect of reducing semideciduous forests in the southern part of the BAF (e.g., Behling 2002), but *P. distincta* only occurs in the non-seasonal coastal and sub-montane rainforests. Data from chemical isotopes in speleothems located in non-seasonal rainforest areas suggest otherwise that the LGM climate may have been as wet as in the present (Cruz et al. 2005). Much more data will be needed to better understand what paleoecological conditions southern BAF species may have endured since the LGM.

Patterns of diversification in southern Brazilian Atlantic forest

There is an ongoing debate on the relative contribution of geographic barriers and Pleistocene refuges in determining diversification of forest-dependent taxa in the BAF (e.g., Carnaval et al. 2009; Thomé et al. 2010; Amaral et al. 2013). The challenge is now to undertake studies designed to distinguish between the alternative, but non-exclusive, forest refugia and barrier hypotheses to explain diversification patterns of BAF organisms. A previous account of the diversification patterns in the *P. burmeisteri* group throughout the BAF (Brunes et al. 2010) suggested that both the Tertiary and Quaternary were important for the diversification of species within the group, a result in agreement with other molecular studies of BAF organisms (e.g., Grazziotin et al. 2006; Thomé et al. 2010; Amaral et al. 2013). Phylogeographic breaks across the BAF have been shown to coincide mainly with major rivers (Doce river) and/or regions with recent neotectonic phenomena such as the Guapiara lineament near Ribeira de Iguape in the southern state of São Paulo (e.g., Saadi et al. 2002; Ribeiro 2006; Thomé et al. 2010). In this study, we report the divergence between two population groups of the species *P. distincta* in the southern BAF, which presently meet near the southern end of the Cubatão shear zone (see Fig. 1 and Saadi et al. 2002). We distinguished between refugial and barrier hypothesis to explain population divergence by contrasting

alternative demographic scenarios. The main difference between these two scenarios classes is that the first involves the vicariance between populations undergoing extensive demographic changes across a short time period (refugia), whereas the second vicariance between more or less stable populations occurs deep in the past (barrier). We concluded that ancient vicariance has undergone without major demographic changes, but minor expansions cannot be completely ruled out. It is therefore not yet clear whether the Cubatão shear zone has been a stable primary barrier or a zone of secondary contact resulting from Holocene range expansions from northern and southern refugial areas. However, it is worth mentioning that results from a recently multilocus ABC analyses of a BAF endemic toad have reinforced the role of geographic barriers in promoting the main divergences in *Rhinella crucifer* complex, with refugia presenting a secondary role by causing intraspecific structure (Thomé et al. 2014).

Conclusions

The southern end of the Cubatão shear zone, near the transitions between the states of Paraná and Santa Catarina, has been associated with a phylogeographic break in the BAF. This pattern adds to the mosaic that is arising from studies of diversification in several BAF organisms, replacing the elegant simplicity of the Carnaval et al. (2009) hypothesis with idiosyncratic complexity (see Thomé et al. 2010). The recent increase in the number of phylogeographic studies in the BAF associated with apparently idiosyncratic patterns argues for the need of a first meta-analysis following a rigorous hypothesis testing approach that may provisionally clarify the relative importance of refugia versus barriers in shaping lineage and species diversity in the megadiverse BAF.

References

Akaike, H. A. I. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716–723.

Alvarado-Serrano, D. F., & Knowles, L. L. (2014). Ecological niche models in phylogeographic studies: applications, advances and precautions. *Molecular Ecology Resources, 14*(2), 233–248. doi:10.1111/1755-0998.12184.

Álvarez-Presas, M., Sánchez-Gracia, A., Carbayo, F., Rozas, J., & Riutort, M. (2014). Insights into the origin and distribution of biodiversity in the Brazilian Atlantic forest hot spot: a statistical phylogeographic study using a low-dispersal organism. *Heredity, 112*(6), 656–665. doi:10.1038/hdy.2014.3.

Amaral, F. R., Albers, P. K., Edwards, S. V., & Miyaki, C. Y. (2013). Multilocus tests of Pleistocene refugia and ancient divergence in a pair of Atlantic Forest antbirds (*Myrmeciza*). *Molecular Ecology, 22*(15), 3996–4013. doi:10.1111/mec.12361.

Anderson, R. P., & Gonzalez, I. (2011). Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. *Ecological Modelling, 222*(15), 2796–2811. doi:10.1016/j.ecolmodel.2011.04.011.

Anderson, R. P., & Raza, A. (2010). The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography, 37*(7), 1378–1393. doi:10.1111/j.1365-2699.2010.02290.x.

Behling, H. (2002). South and southeast Brazilian grasslands during Late Quaternary times: a synthesis. *Palaeogeography, Palaeoclimatology, Palaeoecology, 177*, 19–27.

Brunes, T., Sequeira, F., Haddad, C., & Alexandrino, J. (2010). Gene and species trees of a Neotropical group of treefrogs: genetic diversification in the Brazilian Atlantic Forest and the origin of a polyploid species. *Molecular Phylogenetics and Evolution, 57*(3), 1120–1133. doi:10.1016/j.ympev.2010.08.026.

Brunes, T. O., Alexandrino, J., Baêta, D., Zina, J., Haddad, C. F. B., & Sequeira, F. (2014). Species limits, phylogeographic and hybridization patterns in Neotropical leaf frogs (Phyllomedusinae). *Zoologica Scripta, 43*(6), 586–604. doi:10.1111/zsc.12079.

Camargo, A., Morando, M., Avila, L., & Sites, J. (2012). Coalescent-based methods: a test of accuracy with simulations and an empirical example with lizards of the *Liolaemus darwinii* complex (Squamata: Liolaemidae). *Evolution, 66*(9), 2834–2849. doi:10.5061/dryad.4409k652.

Camargo, A., Werneck, F. P., Morando, M., Sites, J. W., Jr., & Avila, L. J. (2013). Quaternary range and demographic expansion of *Liolaemus darwinii* (Squamata: Liolaemidae) in the Monte Desert of Central Argentina using Bayesian phylogeography and ecological niche modelling. *Molecular Ecology, 22*(15), 4038–4054. doi:10.1111/mec.12369.

Carnaval, A. C., & Moritz, C. (2008). Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest.

Journal of Biogeography, 35, 1187–1201. doi:10.1111/j.1365-2699.2007.01870.x.

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. doi:10.1126/science.1166955.

Carnaval, A. C., Waltari, E., Rodrigues, M. T., Rosauer, D., Vanderwal, J., Damasceno, R., et al. (2014). Prediction of phylogeographic endemism in an environmentally complex biome. *Proceedings of the Royal Society B, 281*, 20141461.

Collevatti, R., Terribile, L., de Oliveira, G., Lima-Ribeiro, M., Nabout, J., Rangel, T., & Diniz-Filho, J. (2013). Drawbacks to palaeodistribution modelling: the case of South American seasonally dry forests. *Journal of Biogeography, 40*(2), 345–358. doi:10.1111/jbi.12005.

Crawford, A. J. (2003). Relative rates of nucleotide substitution in frogs. *Journal of Molecular Evolution, 57*(6), 636–641. doi:10.1007/s00239-003-2513-7.

Cruz, F. W., Burns, S. J., Karmann, I., Sharp, W. D., Vuille, M., Cardoso, A. O., et al. (2005). Insolation-driven changes in atmospheric circulation over the past 116,000 years in subtropical Brazil. *Nature, 434*, 63–66. doi:10.1029/2003JB002684.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution, 29*(8), 1969–1973. doi:10.1093/molbev/mss075.

Earl, D. A., & VonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources, 4*(2), 359–361. doi:10.1007/s12686-011-9548-7.

Edwards, S. V., & Beerli, P. (2000). Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution, 54*(6), 1839–1854. http://www.ncbi.nlm.nih.gov/pubmed/11209764.

Elith, J., Graham, C. H., Anderson, R. P., Dudı́, M., Ferrier, S., Guisan, A., et al. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography, 2*, 129–151.

Elith, J., Phillips, S. J., Hastie, T., Dudı́, M., Miroslav, D., Chee, Y. E., et al. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions, 17*(1), 43–57. doi:10.1111/j.1472-4642.2010.00725.x.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology, 14*(8), 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x.

Faivovich, J., Haddad, C. F. B., Baêta, D., Jungfer, K.-H., Álvares, G. F. R., Brandão, R. A., et al. (2010). The phylogenetic relationships of the charismatic poster frogs, Phyllomedusinae (Anura, Hylidae). *Cladistics, 26*, 227–261.

Felsenstein J (2005) PHYLIP (Phylogeny Inference Package) version 3.6. *Distributed by the author. Department of Genome Sciences, University of Washington, Seattle*

Flot, J.-F. (2010). Seqphase: a web tool for interconverting phase input/output files and fasta sequence alignments. *Molecular Ecology Resources, 10*(1), 162–166. doi:10.1111/j.1755-0998.2009.02732.x.

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics, 147*, 915–925.

Fusinatto, L. A., Alexandrino, J., Haddad, C. F. B., Brunes, T. O., Rocha, C. F. D., & Sequeira, F. (2013). Cryptic genetic diversity is paramount in small-bodied amphibians of the genus *Euparkerella* (Anura: Craugastoridae) endemic to the Brazilian Atlantic forest. *PloS One, 8*(11), e79504. doi:10.1371/journal.pone.0079504.

Glez-Peña, D., Gómez-Blanco, D., Reboiro-Jato, M., Fdez-Riverola, F., & Posada, D. (2010). ALTER: program-oriented conversion of DNA and protein alignments. *Nucleic Acids Research, 38*(Web Server issue), W14–W18. doi:10.1093/nar/gkq321.

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. doi:10.1111/j.1365-294X.2006.03057.x.

Gruber, S. L., Silva, A. P. Z., Haddad, C. F. B., & Kasahara, S. (2013). Cytogenetic analysis of *Phyllomedusa distincta* Lutz, 1950 (2n=2x=26), *P. tetraploidea* Pombal and Haddad, 1992 (2n=4x=52), and their natural triploid hybrids (2n=3x=39) (Anura, Hylidae, Phyllomedusinae). *BMC Genet, 14*, 75. doi:10.1186/1471-2156-14-75.

Haddad, C. F. B., Pombal, J. P. J., & Batistic, R. F. (1994). Natural hybridization between diploid and tetraploid species of leaf-frogs, genus *Phyllomedusa* (Amphibia). *Journal of Herpetology, 28*(4), 425–430.

Haddad, C. F. B., Toledo, L. F., & Prado, C. P. A. (2008). *Anfíbios da Mata Atlântica: guia dos anfíbios anuros da Mata Atlântica (1° ed.)*. São Paulo: Neotrópica.

Haffer, J. (1997). Alternative models of vertebrate speciation in Amazonia: an overview. *Biodiversity and Conservation, 6*, 451–476.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series, 41*, 95–98.

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution, 22*, 160–174.

Heled, J., & Drummond, A. J. (2008). Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology, 8*, 289. doi:10.1186/1471-2148-8-289.

Heller, R., Chikhi, L., & Siegismund, H. R. (2013). The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. *PloS One, 8*(5), e62992. doi:10.1371/journal.pone.0062992.

Hey, J. (2010). Isolation with migration models for more than two populations. *Molecular Biology and Evolution, 27*(4), 905–920. doi:10.1093/molbev/msp296.

Hey, J., & Nielsen, R. (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America, 104*(8), 2785–2790. doi:10.1073/pnas.0611164104.

Hickerson, M. J., Stahl, E., & Takebayashi, N. (2007). msBayes: pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation. *BMC Bioinformatics, 8*, 268. doi:10.1186/1471-2105-8-268

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. doi:10.1002/joc.1276.

Ho, S. Y. W., & Shapiro, B. (2011). Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources, 11*(3), 423–434. doi:10.1111/j.1755-0998.2011.02988.x.

Hudson, R. R. (2002). Generating samples under a Wright–Fisher neutral model of genetic variation. *Bioinformatics, 18*(2), 337–338.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution, 23*(2), 254–267. doi:10.1093/molbev/msj030.

IBGE. (2012). Mapa da área de aplicação da Lei n° 11.428 de 2006. *Instituto Brasileiro de Geografia e Estatística—IBGE*. http://www.ibge.gov.br/home/geociencias/recursosnaturais/mapas_doc6.shtm

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and

multimodality in analysis of population structure. *Bioinformatics, 23*(14), 1801–1806. doi:10.1093/bioinformatics/btm233.

Jeffreys, H. (1961). *The theory of probability.* Oxford: Oxford University Press.

Joly, S., & Bruneau, A. (2006). Incorporating allelic variation for reconstructing the evolutionary history of organisms from multiple genes: an example from *Rosa* in North America. *Systematic Biology, 55*(4), 623–636. http://www.ncbi.nlm.nih.gov/pubmed/16969938.

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics, 25*(11), 1451–1452. doi:10.1093/bioinformatics/btp187.

Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography, 3*, 385–393.

Lutz, B. (1950). Anfíbios anuros da coleção Adolpho Lutz. V. Locomoção e estruturadas extremidades. V.a *Phyllomedusa* (P.) *burmeisteri distincta* A. Lutz. V.b *Aplastodiscus perviridis* A. Lutz. *Memórias do Instituto Oswaldo Cruz, 48*, 599–637.

Martins, F. M. (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. doi:10.1111/j.1095-8312.2011.01745.x.

Milne, I., Wright, F., Rowe, G., Marshall, D. F., Husmeier, D., & McGuire, G. (2004). TOPALi: software for automatic identification of recombinant sequences within DNA multiple alignments. *Bioinformatics, 20*(11), 1806–1807. doi:10.1093/bioinformatics/bth155.

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. doi:10.1038/35002501.

Nielsen, R., & Wakeley, J. (2001). Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics, 158*(2), 885–896.

Otto-Bliesner, B. L., Marshall, S. J., Overpeck, J. T., Miller, G. H., & Hu, A. (2006). Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science, 311*(5768), 1751–1753. doi:10.1126/science.1120808.

Otto-Bliesner, B. L., Hewitt, C. D., Marchitto, T. M., Brady, E., Abe-Ouchi, A., Crucifix, M., et al. (2007). Last Glacial Maximum ocean thermohaline circulation: PMIP2 model intercomparisons and data constraints. *Geophysical Research Letters, 34*(12), L12706. doi:10.1029/2007GL029475.

Pellegrino, K. C. M., Rodrigues, M. T., Waite, A. N., Morando, M., Yassuda, Y. Y., & Sites, J. W. J. (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, 85*(1982), 13–26. doi:10.1016/j.ympev.2011.07.010.

Pelletier, T. A., & Carstens, B. C. (2014). Model choice for phylogeographic inference using a large set of models. *Molecular Ecology, 23*(12), 3028–3043. doi:10.1111/mec.12722.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling, 190*(3–4), 231–259. doi:10.1016/j.ecolmodel.2005.03.026.

Porto, T. J., Carnaval, A. C., & da Rocha, P. L. B. (2013). Evaluating forest refugial models using species distribution models, model filling and inclusion: a case study with 14 Brazilian species. *Diversity and Distributions, 19*(3), 330–340. doi:10.1111/j.1472-4642.2012.00944.x.

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution, 25*(7), 1253–1256. doi:10.1093/molbev/msn083.

Pritchard, J., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*, 945–959.

Radosavljevic, A., & Anderson, R. P. (2013). Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography, 41*, 629–643. doi:10.1111/jbi.12227.

Ramos-Onsins, S. E., & Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution, 19*(12), 2092–2100. http://www.ncbi.nlm.nih.gov/pubmed/12446801.

Ribeiro, A. C. (2006). Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: an example of faunal evolution associated with a divergent continental margin. *Neotropical Ichthyology, 4*(2), 225–246.

Ribeiro, M. C., Metzger, J. P., Martensen, A. C., Ponzoni, F. J., & Hirota, M. M. (2009). The Brazilian Atlantic Forest: how much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation, 142*(6), 1141–1153. doi:10.1016/j.biocon.2009.02.021.

Rojas, M., Moreno, Æ. P., Kageyama, Æ. M., Crucifix, M., Hewitt, Æ. C., Ohgaito, R., et al. (2009). The Southern Westerlies during the last glacial maximum in PMIP2 simulations. *Climate Dynamics, 32*, 525–548. doi:10.1007/s00382-008-0421-7.

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. doi:10.1111/j.1365-294X.2008.03789.x.

Saadi, B. A., Machette, M. N., Haller, K. M., Dart, R. L., Bradley, L., & Souza, A. M. P. D. De. (2002). Map and database of Quaternary faults and lineaments in Brazil. *U.S. Geological Survey, Open-File Report 02-230, version 1.0.*, 1–59.

Salzburger, W., Ewing, G. B., & von Haeseler, A. (2011). The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Molecular Ecology, 20*(9), 1952–1963. doi:10.1111/j.1365-294X.2011.05066.x.

Silva, S. M., Moraes-Barros, N., Ribas, C. C., Ferrand, N., & Morgante, J. S. (2012). Divide to conquer: a complex pattern of biodiversity depicted by vertebrate components in the Brazilian Atlantic Forest. *Biological Journal of the Linnean Society, 107*(1), 39–55. doi:10.1111/j.1095-8312.2012.01919.x.

Stephens, M., Smith, N. J., & Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics, 68*, 978–989.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics, 123*(3), 585–595.

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. doi:10.1016/j.ympev.2010.02.003.

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. doi:10.1186/1471-2148-12-242.

Thomé, M. T. C., Zamudio, K. R., Haddad, C. F. B., & Alexandrino, J. (2014). Barriers, rather than refugia, underlie the origin of diversity in toads endemic to the Brazilian Atlantic Forest. *Molecular Ecology, 23*(24), 6152–6164. doi:10.1111/mec.12986

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. doi:10.1111/mec.12164

Valdez, L., & D'Elía, G. (2013). Differentiation in the Atlantic Forest: phylogeography of *Akodon montensis* (Rodentia, Sigmodontinae) and the Carnaval–Moritz model of Pleistocene refugia. *Journal of Mammalogy, 94*(4), 911–922. doi:10.1644/12-MAMM-A-227.1.

Wakeley, J., & Hey, J. (1997). Estimating ancestral population parameters. *Genetics, 145*, 847–855.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology, 7*, 256–276.

Weber, S. L., Drijfhout, S. S., Abe-Ouchi, A., Crucifix, M., Eby, M., Ganopolski, A., et al. (2007). The modern and glacial overturning circulation in the Atlantic Ocean in PMIP coupled model simulations. *Climate of the Past, 3*, 51–64. doi:10.5194/cpd-2-923-2006.

GfBS