JOURNAL OF ZOOLOGICAL SYSTEMATICS AND EVOLUTIONARY RESEARCH

Accepted on 15 December 2012
J Zoolog Syst Evol Res doi: 10.1111/jzs.12014

[1]*Laboratório de Mastozoologia e Biogeografia, Departamento de Ciências Biológicas, Universidade Federal do Espírito Santo, Vitória ES, Brazil;* [2]*Department of Biology, City College of New York, City University of New York, New York NY, USA;* [3]*Department of Biology, University of Richmond, Richmond VA, USA*

Phylogeographic structure is strong in the Atlantic Forest; predictive power of correlative paleodistribution models, not always

João Filipe Riva Tonini[1,2,3], Leonora Pires Costa[1] and Ana Carolina Carnaval[2]

Abstract

We assess whether correlative paleoclimatic models of species ranges accurately predict genetic diversity patterns in species of distinct life histories traits in the Atlantic Forest (AF) of Brazil. To this end, we use sequences of the mitochondrial gene *ND2* from *Dendropsophus elegans* and *Chiasmocleis carvalhoi* – summarized in the shape of phylogenies and population genetic statistics – and maximum entropy models of species distributions under current, 21 kya BP and 120 kya BP climatic reconstructions. The two target species have distinct ranges, habitat tolerances, rates of reproduction and dispersal abilities, yet are endemic to the AF. Although the more restricted and semi-fossorial *C. carvalhoi* is associated with forested habitats and thought to be a poor disperser, the widely ranged arboreal *D. elegans* inhabits open areas such as pastures and human-impacted regions of the AF, and is easily found perched on herbaceous vegetation in inundated areas. We had anticipated that correlative distribution models of the broadly distributed *D. elegans* would perform better then models of the narrowly ranged *C. carvalhoi*, thus better predicting current patterns of genetic diversity. The results demonstrate poor predictive ability of climate-based models of *C. carvalhoi* under current climatic conditions, suggesting that factors such as biotic interactions or dispersal ability may be playing a central role in defining this species distribution – both now and in the recent past. Models under current climate are nonetheless accurate in the broadly ranged *D. elegans*. As a corollary, paleoclimatic models accurately predicted patterns of diversity of the *ND2* mitochondrial gene in *D. elegans*, but not in *C. carvalhoi*. We attribute these distinct responses to the poor explanatory power of paleodistributions models when applied to species that violate the basic assumption of the environment as main driver of distribution patterns. This calls for a careful use of distribution models for the purpose of evolutionary biogeographical inference. Like *C. carvalhoi*, other species whose ranges are not yet at equilibrium, or which are impacted by competitor, parasite or pathogen presence, may not be suitable to the combined use of paleoclimatic-model based phylogeographic inference, as here implemented – despite relatively high area under the curve values.

Key words: Anurans – brazil – phylogeography – correlative models – hylidae – microhylidae – mitochondrial

Introduction

Due to their low mobility and high philopatry to birthplace – both of which facilitate population isolation and subsequent genetic divergence – anurans have been widely used to study the effects of environmental change on past population dynamics (Beebee 1996; van der Meijden et al. 2007; Zeisset and Beebee 2008). To promote further insight on this subject, herpetologists worldwide are combining the tools of correlative distribution modelling under paleoclimatic scenarios with analyses of molecular data (e.g. Graham et al. 2004; Carnaval et al. 2009; Thomé et al. 2010; Hoskin et al. 2011; Amaro et al. 2012). In these studies, explicit scenarios of biological responses to past climate change are said to be supported under a hypothesis-testing framework when there is mutual agreement between reconstructed paleoclimatic ranges and current levels and patterns of diversity (Carnaval et al. 2009; Martins 2011).

In the Atlantic Forest (AF) of Brazil, for instance, a molecular genetic analysis of three broadly ranged species of tree-frogs showed high levels of diversity in the central and northern forests (Carnaval et al. 2009). These regions were also predicted as suitable for the target taxa under current, Holocene (6 kya), and Last Glacial Maximum (LGM) (21 kya) climates – hence being inferred as climatically stable, and then identified as putative climatic refugia. A large central Late Quaternary refuge (from the state of Rio de Janeiro to Bahia) and a small northern refuge (in Pernambuco and Alagoas) were then recognized (Carnaval and Moritz 2008). In contrast, paleoclimatic range reconstructions of the same taxa indicated climatic instability south of Rio de Janeiro – a result consistent with the low levels of genetic diversity observed in southern populations.

We here ask whether this combination of tools is useful and results in mutual agreement in studies of species with distinct life-history traits, specifically dispersal ability and tolerance to habitat shifts. Were Carnaval et al. (2009) simply lucky to find consistency between spatial models and molecules because they focussed on broadly ranged species? Distribution model approaches usually assume that range changes are mostly determined by the availability of climatically suitable habitat, without additional limitations imposed by dispersal ability, life history or biotic interactions (Hampe 2004; Angert et al. 2011; Peterson et al. 2011). As dispersal ability and degree of ecological generalization facilitate range shift and the tracking of suitable climatic conditions in times of environmental change (Pöyry et al. 2009), allowing species ranges to reach equilibrium more readily than poor dispersers or specialist species, we predict that a combination of the tools of past range reconstructions and genetic analyses will be more useful in the study of broadly ranged, generalist species – for which climate is the main determinant of species distributions.

In this study, we use molecular data to test species-specific hypotheses about the location of Late Quaternary refugia in two endemic frogs of the AF that differ widely in their dispersal abilities and life histories. They are the widespread *Dendropsophus elegans* (Wied-Neuwied, 1824), and the narrowly distributed *Chiasmocleis carvalhoi* Cruz, Caramaschi & Izecksohn, 1997. While the more restricted, semi-fossorial *C. carvalhoi* is associated with forested habitats, the arboreal *D. elegans* inhabits open areas such as pastures and human-impacted regions of the AF, and is easily found perched on herbaceous vegetation in inundated areas (Izeckson and Carvalho-e-Silva 2001). Our prediction is that agreement will be observed between distribution models and genetic data in the case of *D. elegans*, which we expect to resemble the results of Carnaval et al. (2009), but not necessarily

Corresponding author: João Filipe Riva Tonini (jfrtonini@gmail.com)
Contributing authors: Leonora Pires Costa (leonoracosta@yahoo.com), Ana Carolina Carnaval (acarnaval@ccny.cuny.edu)

in *C. carvalhoi*. For the latter, we expect, the use of the combined tools of phylogeography and climate-based models of species distribution under past climates may generate more discord than agreement.

To this end, we sequenced a mitochondrial marker (NADH dehydrogenase subunit 2) from *D. elegans* and *C. carvalhoi*, and compared the genealogical signal and polymorphism data to climatic models of current and paleodistributions that we generated from available locality data. Going beyond a critical assessment of the usefulness of paleoclimatic models to guide biogeographical evolutionary research, this contribution adds new fine-scale molecular data for a biome that is becoming a model system for biodiversity studies in the Neotropics.

Materials and Methods

Study species and sampling

Dendropsophus elegans (Hylidae) is a small arboreal frog widely distributed in coastal Brazil, ranging from the northern state of Pernambuco to Rio Grande do Sul (Encarnação 2010). Individuals reproduce throughout the year, inhabit open areas along the forest edge and can easily colonize human-impacted environments (Camurugi et al. 2010; Forlani 2010). The species generally clings to herbaceous vegetation in the vicinity of lentic water, where females deposit their eggs (Prado and Pombal 2005). It is easily collected by hand along coastal lagoons and swamps, from sea level to montane altitudes of approximately 1300 m.a.s.l. We obtained 177 samples of *D. elegans* from 55 localities, sampling 1–16 individuals per site (Appendix S1).

Chiasmocleis carvalhoi (Microhylidae) is also small in size and lives mainly in forest environments. This semi-fossorial species inhabits leaf-litter or burrows within or along the edge of rainforest patches, occurring from sea level to ca. 550 m of altitude. It is however solely found in the central portion of the AF, from the southern portion of the state of Bahia to São Paulo (Forlani 2010). *Chiasmocleis carvalhoi* follows an explosive breeding pattern after the first heavy rains of the wet season, when hundreds of individuals reproduce at once in temporal ponds (Carvalho-e-Silva et al. 2008; Silva et al. 2008). We obtained 52 samples of *C. carvalhoi* from 11 localities, sampling 1–20 individuals per site (Appendix S1).

Paleoclimatic distribution models

To generate correlative models of species ranges under current and paleoclimates, we compiled point locality data from our own field-work and Brazilian scientific collections (Museu Nacional-UFRJ, Coleção Célio Fernando Batista Haddad-UNESP/RC, Instituto de Biologia-IB/USP, Coleção de Tecido e DNA-UFES, Museu de Zoologia-UFBA, Museu de Zoologia-UFV). We built correlative species distribution models with the maximum entropy algorithm implemented in Maxent 3.3.3h (Phillips et al. 2006; http://www.cs.princeton.edu/~schapire/maxent/), and trained with climatic data from the 19 bioclimatic variables available through the WorldClim database (Hijmans et al. 2005). Models were trained and projected into current climatic layers (1950–2000) at 1 km^2 spatial resolution, including the entire AF domain as delimited by decree 750/1993 (Ministério do Meio Ambiente, Brazil). Retrojections were also trained under current climatic conditions using the AF domain and projected to paleoclimatic scenarios of the Last Interglacial (LIG, 120 kya BP) and LGM (21 kya BP), provided under CCSM models at 5 km^2 spatial resolution (Hijmans et al. 2005; http://www.worldclim.org).

To increase model performance, we applied species-specific tuning under current climatic conditions (Anderson and Gonzalez 2011). We then used the best set of parameters as defined by the tuning in our retrojections, differing from Maxent's default settings. To avoid over-fitting, we tested different feature classes and regularization multipliers. The feature classes available in Maxent 3.3.3 h are linear (L), quadratic (Q), hinge (H), product (P) and threshold (T) (Phillips and Dudők 2008), and combinations between features are allowed. To minimize the production of overfit models, Maxent utilizes 'regularization', penalizing features that are weighted too strongly or complex models that include too many features. This forces Maxent to concentrate on the most important features – those with the highest explanatory ability (Phillips et al. 2006). According to Anderson and Gonzalez (2011), this process is similar to selecting optimal models using the Akaike Information Criterion (AIC; Warren and Seifert 2011).

We used 12 regularization multipliers from 0.5 to 6.0, in increments of 0.5 for each feature class tested. For the species with <10 occurrence records, we tested linear (L), linear and quadratic (LQ), hinge (H), linear, quadratic and hinge (LQH) and default; for the species with more than 11 occurrence records, we tested LQ, H, LQH, linear, quadratic and product (LQP), linear, quadratic, product and threshold (LQPT) and default. Point records used in the distribution models are from sample localities that we have tissues. For *D. elegans*, we used 52 point records partitioned in four replicate sets; for *C. carvalhoi*, we used 10 point records partitioned in ten replicates. Duplicate records per site, or records within 10 km apart, we reduced to one to minimize spatial autocorrelation (Anderson and Raza 2010).

We evaluated model performance using average values of 10% training presence test omission rates, and minimum training presence (MTP) test omission rates (both threshold dependent), as well as the area under the curve (AUC) of the receiver operating characteristic curve (threshold independent), identifying the lowest values of average differences between training and test AUCs (Anderson and Gonzalez 2011). Also, we used Maxent's multivariate environmental similarity surfaces to identify areas predicted as suitable that were affected by variables outside the training range (Elith et al. 2010).

For each model, we used Maxent's MTP threshold to obtain binary predictions; values bellow MTP threshold were converted to zero, and values above it converted to one. Binary predictions include all pixels deemed by each model to be at least as suitable as those where the species is known to occur (as per the training data set). For each data set, we used the intersection of the binary predictions under current and past climates to infer climatically stable areas (Waltari et al. 2007; Vences and Wake 2007; Carnaval and Moritz 2008).

Laboratory protocols

We extracted genomic DNA from ethanol-preserved muscle or liver tissues using a high salt extraction method (Bruford et al. 1992). For amplification of the mitochondrial gene *ND2*, we used primers L4437 and HND2PB (approximately 750 pb; Macey et al. 1997; Carnaval and Bates 2007) in *D. elegans*, and ND2ALAR and ND2ELEUF1 (approximately 900 pb; Carnaval and Bates 2007) in *C. carvalhoi*. Each polymerase chain reaction (PCR) had a final volume of 12.5 μl containing 1× buffer, 2.0 mM $MgCl_2$, 0.2 mM dNTP solution, 0.12 μM of each primer, 0.75 units of Platinum Taq (Invitrogen Co., Life Technologies, Grand Island, NY, USA) and 10–50 ng DNA. PCR profile included an initial denaturation step of 94°C for 5 min, 30–40 cycles of 94°C for 30 s, 48°C for 45 s and 72°C for 30 s. The final extension stage lasted 7 min at 72°C.

We purified the PCR product with ExoSAP enzyme (GE Healthcare Life Sciences, Pittsburgh, PA, USA), and used BigDye v3.1 kit (Perkin Elmer, Applied Biosystems™, Foster City, CA, USA) for sequencing reaction. Template DNA was sequenced in both directions in automatic sequencer ABI 310 or ABI 3700 (Applied Biosystems, Life Technologies, Carlsbad, CA, USA). Laboratory work was performed at Núcleo de Genética Aplicada a Conservação da Biodiversidade (NGACB – UFES) and Empresa Brasileira de Pesquisa Agropecuária (Embrapa, Brasília, DF, Brazil). We aligned the sequences using ClustalW (Larkin et al. 2007) through the interface of mega 5.0 (Tamura et al. 2011), and edited them manually. All sequences are deposited in the GenBank (access numbers from JQ410480 to JQ410715).

Phylogeographic analysis

We rooted the mitochondrial tree of *C. carvalhoi* with sequences of *Stereocyclops incrassatus*, *Dermatonotus muelleri*, *Dasypops schirchi*, *Chiasmocleis leucosticta*, *Chiasmocleis mantiqueira* and *Chiasmocleis schubarti* (van der Meijden et al. 2007). Sequences of *Dendropsophus berthalutzae*, *Dendropsophus ebraccatus* (GenBank EU034096) and *Hypsiboas faber* (GenBank FJ502765) were used to root the ND2 genealogy of *D. elegans* (Wiens et al. 2010).

Maximum Likelihood (ML) and Bayesian Inference (BI) were used to reconstruct the *ND2* genealogy and to identify genetic structure within each species. Molecular evolution models were chosen by MRMODELTEST (Nylander 2004) using the AIC. BI and time of divergence between lineages were estimated in BEAST v1.6.2 (Drummond and Rambaut 2007). We performed two independent runs of 50 million generations each, with trees sampled every 1000 generations; Markov Chain Monte Carlo operators were optimized according to program output suggestions.

We checked the quality of the parameters of the Bayesian analysis with Tracer version 1.5, and generated consensus trees with TreeAnnotator (Drummond and Rambaut 2007). We discarded the first 5000 trees as burn-in. To allow for variation in the branch length, according to the log-normal distribution, we used an uncorrelated relaxed molecular clock (Drummond et al. 2006) calibrated with published rates of *ND2* evolution (0.957% per million years; Crawford 2003). This allowed us to roughly compare inferred times of diversification in *D. elegans* and *C. carvalhoi* with other anurans studies that used the same rate (Fitzpatrick et al. 2009; Brunes et al. 2010; Thomé et al. 2010; Amaro et al. 2012). A ML phylogeny was estimated in the PHYML portal (Guindon and Gascuel 2003) using the same evolution model selected by AIC. Clade support was estimated with 1000 bootstrap interactions.

Descriptive statistics are provided for the clades uncovered by the phylogenetic analysis. Haplotype and nucleotide diversity and demographic tests of neutrality or expansion, including Tajima' D (Tajima 1989) and Fu' Fs (Fu 1997), were calculated in DNASP v.5 (Librado and Rozas 2009) and ARLEQUIN 3.0 (Excoffier et al. 2005). Fixation indexes and Molecular Analyses of Variance (AMOVAS) were estimated in ARLEQUIN 3.0 (Excoffier et al. 2005). A value of $p < 0.05$ was considered significant in all aforementioned analyses, with the exception of Fs, where a value of $p < 0.02$ was considered significant given the power of the test (Ramos-Onsins and Rozas 2002). Percentage of p-uncorrected genetic distance was estimated in MEGA 5.0 (Tamura et al. 2011). We used a Mantel test to compare matrices genetic and geographic distances to test the hypothesis of isolation by distance within clades using the software PAST (Hammer et al. 2001). Geographic distances were estimated with Geographic Distance Matrix Generator v1.2.3 (http://biodiversityinformatics.amnh.org/open_source/gdmg/index.php).

Results

Models

The combination of feature classes and regularization multipliers, with either lowest values of 10% training presence test omission rates and MTP test omission rates, or least average differences between training and test AUCs, was LQP 2.0 for *D. elegans* and LQP 0.5 for *C. carvalhoi*. AUC values *D. elegans* and *C. carvalhoi* were 0.852 and 0.945 respectively (Fig. 1). Our predictions of suitable areas were not affected by variables outside the training range in both species and all time periods (results not shown).

Current and past climatic models for *D. elegans* predicted the existence of vast climatically stable areas throughout the Late Quaternary (Fig. 1), including large patches of stable areas in the central and northern region of the AF, and smaller stable areas along the southernmost coast. Yet, the distribution model developed under current climate, as well as model intersection, showed some over-prediction outside the known range of *D. elegans*, to the southwest.

Models of *C. carvalhoi* under current climatic conditions significantly over-predicted the range of this species, both to the north and south of expert-drawn ranges for this taxon (Fig. 1). The intersection of binary predictions under current and past climates predicted a continuous, narrow climatically stable area restricted to the coastal regions of the AF in southeastern Espírito Santo state, around the Doce River (Fig. 1).

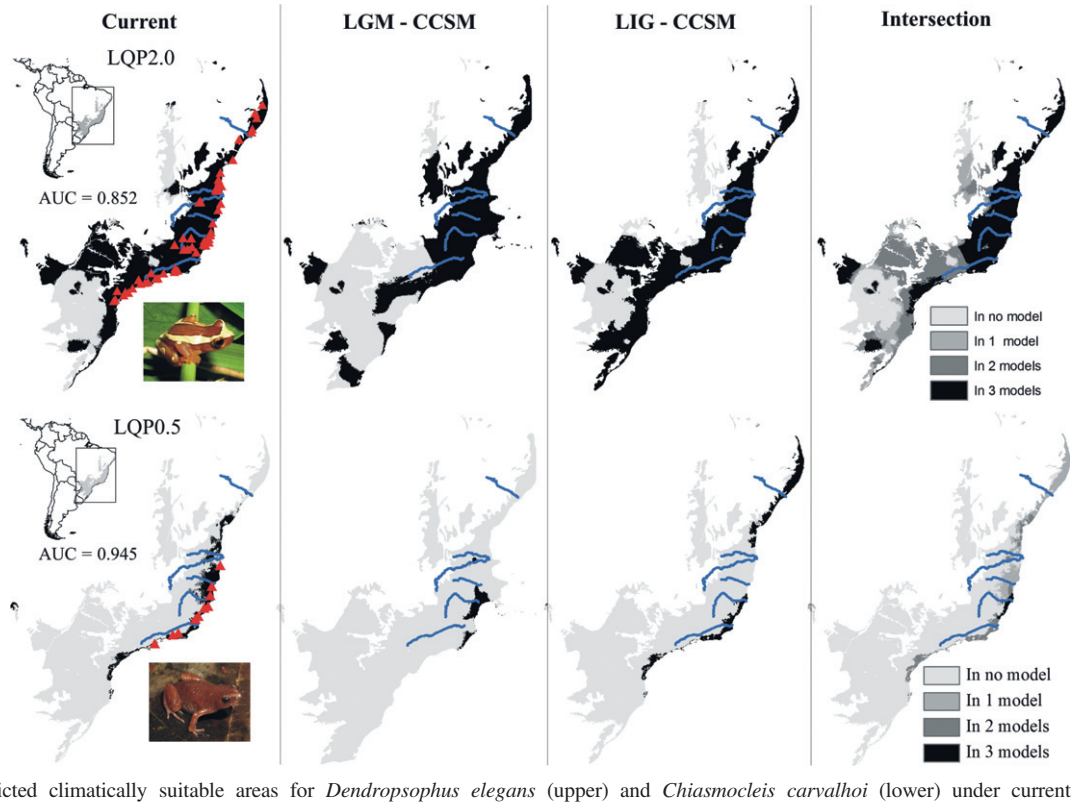

Fig. 1. Predicted climatically suitable areas for *Dendropsophus elegans* (upper) and *Chiasmocleis carvalhoi* (lower) under current and past climates, and the intersection between them. In the three columns to the left, suitable areas are shown in black, and unsuitable areas in grey. Grey scale in the Intersection show how many times a region was predicted suitable. Red triangles indicate available tissues. Blue lines represent major coastal rivers, from North to South: São Francisco, Mucuri, Doce and Paraíba do Sul. LGM-CCSM: Last Glacial Maximum; LIG-CCSM: Last Interglacial.

Phylogeographic analyses

The 177 samples sequenced for *D. elegans* resulted in 96 *ND2* haplotypes with 200 polymorphic sites (707 pb). The 52 individuals of *C. carvalhoi* resulted in 43 *ND2* haplotypes with 184 polymorphic sites (849 pb). Molecular evolution models chosen were GTR + G ($\gamma = 0.381$) and GTR + I + G (invariables sites = 0.351 and $\gamma = 1.119$) for *D. elegans* and *C. carvalhoi* respectively. BI and ML trees presented similar topology, with varying supports for the clades (Fig. 2; Appendix S2).

Phylogeographic histories are distinct between *D. elegans* and *C. carvalhoi*. Northern, central and southern clades were recovered for both species, yet clade support differed across taxa. Although *D. elegans* showed highly supported clades and clade relationships (Fig. 2a), the latter was not clear in *C. carvalhoi* – particularly regarding the central and southern clades (Fig. 2b). Phylogenetic structure and values of genetic distance were notably high both in *D. elegans* and *C. carvalhoi* (Fig. 2). Maximum genetic distance within the clades of *D. elegans* was 2.7% (clade N2) and 2.1% in *C. carvalhoi* (clade S1; Appendix S3).

Our tentative dating analyses suggest that initial splits happened during the Miocene and Pliocene both in *D. elegans* and *C. carvalhoi* (Table 1), yet contemporary lineages show more recent origin (Pliocene and Pleistocene; Table 1). In *D. elegans*, signature of population expansion was detected in the clades N4, C3, S3 and S4 (Table 2; Fig. 2a). Molecular data from *C. carvalhoi* clades N2 and C1 also showed signals of population expansion (Table 3; Fig. 2b). The latter overlap geographically with *D. elegans* clades N4 and C3 (Fig. 2), suggesting that historical events in this area may have affected population demography in both species.

Genetic diversity is high in *D. elegans*, varying from 0.95 to 0.89 (haplotypic), and from 0.01 to 0.04 (nucleotide; Table 2).

Table 1. Molecular clock estimated using the mitochondrial gene *ND2* in *Dendropsophus elegans* and *Chiasmocleis carvalhoi*

	Mean	±95% HPD	Geological time
D. elegans			
((N C)S)	6.56	4.92–8.37	Miocene
(N C)	4.80	4.68–6.04	Miocene
(N1 N2)	2.20	1.56–2.92	Plio-Pleistocene
(N3 N4)	2.06	1.26–2.87	Plio-Pleistocene
((N3 N4)(N1 N2))	3.45	2.58–4.39	Plio-Pleistocene
(C2 C3)	0.90	0.53–1.30	Pleistocene
((C2 C3)C4)	1.35	0.86–1.86	Pleistocene
(((C2 C3)C4)C1)	2.28	1.52–3.12	Plio-Pleistocene
(S2 S3)	0.93	0.55–1.33	Pleistocene
((S2 S3)S4)	0.98	0.61–1.38	Pleistocene
(((S2 S3)S4)S1)	1.44	0.92–2.04	Pleistocene
C. carvalhoi			
((C S)N)	5.72	3.20–8.40	Mio-Pliocene
(N1 N2)	3.27	1.41–5.33	Plio-Pleistocene
(C1 C2)	1.95	0.93–3.18	Plio-Pleistocene
(S1 S2)	4.06	1.81–6.67	Mio-Pleistocene

The AMOVA shows that 68% (d.f. = 2) of the genetic variation within this species can be attributed to differences between the northern, central and southern clades, whereas 27.8% (d.f. = 9) is attributed to diversity within these clades. Furthermore, clades northern, central and southern had genetic distance positively correlated with geographic distance and F_{ST} values were consistently >0.67 ($p < 0.01$) (Tables S3 and S5 in Appendix S3). *Chiasmocleis carvalhoi* shows extremely high haplotype diversity, ranging from 0.99 to 0.93 (haplotypic), and from 0.01 to 0.06 (nucleotide; Table 3). AMOVA results show that 87.17% (d.f. = 1) of the variation is attributed to differences between the northern,

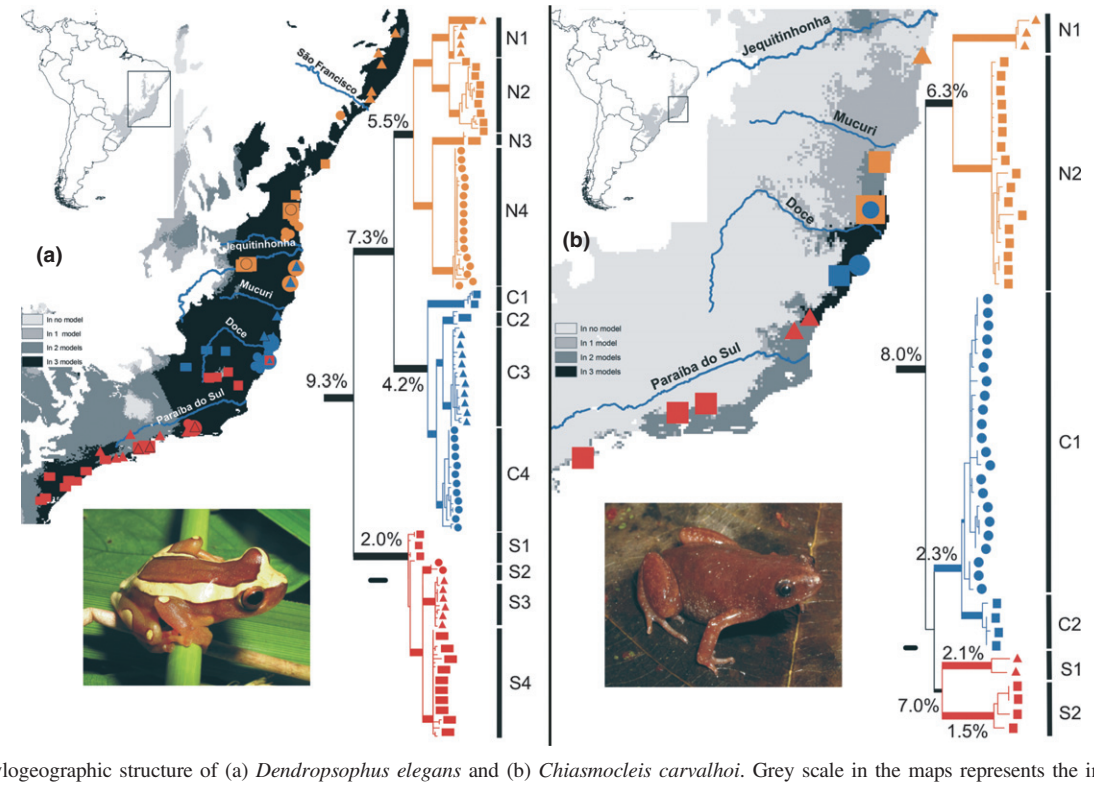

Fig. 2. Phylogeographic structure of (a) *Dendropsophus elegans* and (b) *Chiasmocleis carvalhoi*. Grey scale in the maps represents the intersection of climatic predictions. In the ML trees the northern clades are in orange, central clades are in blue, southern clades are in red. Presence of multiple symbols or colours in a given locality indicates that individuals of different clades are sympatric. Thick branches indicate support in both Bayesian Inference (BI) (>95 pp) and ML (>80% bootstrap support), and thin branches indicate support in BI only (>95 pp). Percentages in the branches indicate levels of genetic divergence (p-uncorrect) between groups within the respective clade. Scale bar = 0.01 substitutions/site in both phylogenies.

Table 2. DNA polymorphism results of the mitochondrial gene *ND2* for clades of *Dendropsophus elegans*

	N	*H*	Hd	Π	*D*	FS
N1	11	5	0.618 ± 0.164	0.007 ± 0.003	−0.591	2.243
N2	17	13	0.963 ± 0.033	0.010 ± 0.001	−0.691	−3.252
N3	4	2	0.500 ± 0.265	0.000 ± 0.000	−0.612	0.172
N4	35	11	0.758 ± 0.065	0.002 ± 0.000	−1.877*	−3.644*
C1	5	4	0.900 ± 0.161	0.002 ± 0.000	0.699	−1.405
C2	4	4	1.000 ± 0.177	0.003 ± 0.000	−0.796	−1.514
C3	25	19	0.967 ± 0.024	0.004 ± 0.000	−1.690	−16.263*
C4	25	11	0.767 ± 0.086	0.005 ± 0.000	−0.601	−1.969
S1	8	5	0.893 ± 0.086	0.003 ± 0.000	−0.581	−0.495
S2	3	1	NA	NA	NA	NA
S3	13	9	0.910 ± 0.068	0.002 ± 0.000	−1.907*	−2.530
S4	27	12	0.658 ± 0.106	0.002 ± 0.000	−2.283*	−4.619*

N, number of samples; H, number of haplotypes; Hd, haplotype diversity; π, nucleotide diversity per site; *D*, Tajima's D ($p < 0.05$); FS, Fu's FS ($p < 0.02$); NA, non applicable.
*Statistically significant.

Table 3. DNA polymorphism results of the mitochondrial gene *ND2* for clades of *Chiasmocleis carvalhoi*

	N	*H*	Hd	π	*D*	FS
N1	3	3	1.000 ± 0.272	0.005 ± 0.002	NA	0.308
N2	17	15	0.985 ± 0.025	0.006 ± 0.000	−1.849*	−8.445*
C1	22	17	0.974 ± 0.021	0.007 ± 0.000	−1.003	−6.573*
C2	4	3	0.833 ± 0.222	0.020 ± 0.015	−0.808	0.73
S1	2	2	1.000 ± 0.500	0.021 ± 0.010	NA	2.890
S2	4	3	0.833 ± 0.222	0.012 ± 0.003	0.134	2.706

N, number of samples; *H*, number of haplotypes; Hd, haplotype diversity; π, nucleotide diversity per site; *D*, Tajima's D ($p < 0.05$); FS, Fu's FS ($p < 0.02$); NA, not applicable.
*Statistically significant.

central and southern clades. Isolation by distance pattern was observed in clades northern and central; F_{ST} values among clades are higher than 0.68, and comparisons within clades also show significant results (Tables S4 and S6 in Appendix S3).

Spatial patterns of diversity in the AF are often discussed in the light of hydrology (river location), yet only a portion of the patterns observed in *D. elegans* and *C. carvalhoi* match the location of rivers in coastal Brazil. In *D. elegans*, the São Francisco River isolates clade N1 from N2, but clades around Jequitinhonha and Mucuri River have haplotypes on both margins. As well as clades C3 and C4 are separated by the Doce River, yet a mixture of haplotypes is observed across margins (Fig. 2a). To the south, the Paraíba do Sul River seems to split clades S1 from clades S2, S3 and S4. In *C. carvalhoi*, separation between clades N1 and N2 is coincident with the Mucuri River (Fig. 2b), and the boundary between clades C1 and C2 is located south of the Doce River. Yet, haplotypes belonging to clades N2 and C1 are found sympatrically. The two southern lineages in the latter species show a break around the Paraíba do Sul River.

Discussion

In general, the correlative distribution models of *D. elegans* under current climatic conditions predicted the known range of the species accurately. A slight over-prediction to the southwest of the AF was observed, similarly to what has been reported for other endemic and widespread anurans, such as the tree-frogs *Hypsiboas albomarginatus* and *H. faber* (Carnaval et al. 2009), and toads in the *Rhinella crucifer* group (Thomé et al. 2010). It is possible that the relatively poorer weather-station coverage in interior Brazil be responsible for such generalized over-predictions to the west.

The genetic diversity of *D. elegans* is strongly spatially structured. Combined with the paleoclimatic models, they are in general agreement with the scenario of Late Quaternary climatic stability proposed by Carnaval and Moritz (2008) and Carnaval et al. (2009). The northern and central clades, distributed within a large inferred refugium (or sanctuary, for those following the definition of Recuero and García-París 2011; Fig. 2a) showed deep divergences and higher levels of persistence and accumulation of genetic diversity relative to the southern clade. Yet, while our paleoclimatic models infer comparatively less stability in the south, they also suggest the existence of a small LGM refugium (or sanctuary as per Recuero and García-París 2011) along the coast, spanning part of the current range of the markedly divergent southern mtDNA clade.

In contrast, models of *C. carvalhoi* under current climatic conditions extensively over-predicted the range of this species – both to the north and to the south. Moreover, observed patterns of genetic diversity disagreed with the inferred location of historical refugia as per the paleoclimatic models. Although the latter predicted Late Quaternary stability by the Doce River in the state of Espórito Santo (Fig. 1), the genetic data suggested long-term presence of highly differentiated lineages in southern areas, which were inferred as unstable by the paleomodels (Fig. 2).

This mismatch between climate-based models and genetic patterns in *C. carvalhoi* agrees with our original expectations, and may have been driven by the comparatively limited predictive power of the climate-based correlative distribution models, as implemented here, when applied to extremely poor dispersers and narrowly distributed species (Hampe 2004; Nogués-Bravo 2009). Other narrowly ranged species that are phylogenetically related to *C. carvalhoi* are in fact distributed both to the north and to the south of the range of our target species (Forlani 2010). Because these taxa inhabit similar environments, it is possible that Maxent's over-predictions may be reflecting a (yet untested) retention of phylogenetic signal or niche conservatism across closely related species (e.g. Olalla-Tárraga et al. 2001). It remains to be seen whether the inability of *C. carvalhoi* to occupy climatically suitable sites to the south and north of its range are due to unmeasured abiotic factors or to biotic interactions such as competition with closely related species or failure to co-exist with more narrowly distributed lineages of parasites or pathogens (Ricklefs 2012). Violation of the basic assumption that environmental features, as measured here, are the main drivers of distribution patterns in narrow endemics requires careful use of distribution models for the purpose of evolutionary biogeographical inference. Like *C. carvalhoi*, other species whose ranges are not yet at equilibrium, or which are impacted by competitor, parasite or pathogen presence, may not be suitable to analyses such as the ones here implemented – despite relatively high AUC values tied to their inferred distribution models.

Rivers have been previously reported as phylogeographic markers for AF organisms (review by Costa and Leite 2012 and also Grazziotin et al. 2006; Cabanne et al. 2007, 2008; Fitzpatrick et al. 2009; Palma-Silva et al. 2009; Brunes et al. 2010; Resende et al. 2010; Thomé et al. 2010; Amaro et al. 2012). Pellegrino et al. (2005) were the first ones to specifically associate phylogeographic breaks with rivers in the AF, citing the Doce and the Paraíba do Sul while describing genetic diversity patterns in lizards. The same rivers correlate with two major breaks observed in *D. elegans*. It is still unclear whether these (and other) AF rivers were the primary drivers of divergence across mitochondrial haplogroups, whether they constitute barriers to gene flow by hindering further expansion from previous relict

areas (e.g. Nicolas et al. 2011; Fouquet et al. 2012; Gehring et al. 2012), or, alternatively, whether the perceived genetic structure is artificially caused by sampling issues or the stochasticity of the coalescent, given an underlying pattern of genetic isolation by distance (Irwin 2002).

Most phylogeographic data from AF species still come from mitochondrial sequencing and point to high levels of diversity genetic structure within taxa (Martins et al. 2007; Köhler et al. 2010; Chapple et al. 2011; Neaves et al. 2012). *Chiasmocleis carvalhoi* and *D. elegans* are no exception. Given what is known about the habitat specificity and breeding habitats of these two species, the *C. carvalhoi* results are not surprising. Yet, we expected to find less structure in the widely distributed and generalist *D. elegans*. Our data demonstrate that even broadly distributed AF species categorized as of 'least concern' by conservation biologists harbour cryptic diversity that merits further attention, especially in the face of the rampant landscape changes faced by this biome. Sequencing of multiple nuclear loci and phenotypic analyses will promote further insight on the processes that underscore biodiversity patterns in this and other species endemic to the threatened AF.

Acknowledgements

We thank M. Napoli (MZUFBA), M. Solé (UESC), C. F. B. Haddad (UNESP/RC), M. T. Rodrigues and R. C. Amaro (USP), V. Fagundes and J. L. Gasparini (UFES), R. Feio and M. Moura (MZUFV), J. Pombal Jr (MN) and P. S. Santos (UFMG) for providing tissue samples. M. Forlani and P. Peloso helped on the identification the specimens of *Chiasmocleis*. E. E. Gutiérrez, R. P. Anderson and E. Waltari provided comments and assistance on the climatic models. F. M. Martins and A. C. Pavan provided help on the genetic analysis. M. T. Rodrigues, Z. Spanos and B. Rizzo added useful comments on the manuscript. J. Justino, A. C. Loss and S. Ramira provided laboratory assistance. Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) provided collection permits. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) provided a masters scholarship to JFRT. National Science Foundation (DEB 1144692 to Rafael O. de Sá) currently support JFRT. This article was developed as part of the project 'Zoogeografia de vertebrados terrestres em uma zona de sutura no Corredor Central da Mata Atlântica' funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq process 481830/2008-1) to LPC. AC acknowledges funding from the National Science Foundation (DEB 0817035 and DEB 1120487).

References

Amaro RC, Rodrigues MT, Yonenaga-Yassuda Y, Carnaval AC (2012) Demographic processes in the montane Atlantic rainforest: molecular and cytogenetic evidence from the endemic frog *Proceratophrys boiei*. Mol Phylogenet Evol **62**:880–888.

Anderson RP Gonzalez I Jr (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. Ecol Model **222**:2796–2811.

Anderson RP, 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. J Biogeogr **37**:1378–1393.

Angert AL, Crozier LG, Rissler LJ, Gilman SE, Tewksbury JJ, Chunco AJ (2011) Do species' traits predict recent shifts at expanding range edges? Ecol Lett **14**:677–689.

Beebee TJC (1996) Ecology and Conservation of Amphibians. Chapman & Hall, London.

Bruford MW, Hanotte O, Brookfield JFY, Burke T (1992) Single-locus and DNA fingerprinting. In: Hoelzel AR (ed.), Molecular Genetic Analyses of Populations: A Practical Approach. IRL Press, Oxford, pp 225–269.

Brunes TO, Sequeira F, Haddad CFB, Alexandrino J (2010) Gene and species trees of a Neotropical group of tree frogs: genetic diversification in the Brazilian Atlantic Forest and the origin of a polyploid species. Mol Phylogenet Evol **57**:1120–1133.

Cabanne GS, Santos FR, Miyaki CY (2007) Phylogeography of *Xiphorhynchus fuscus* (Passeriformes, Dendrocolaptidae): vicariance and recent demographic expansion in southern Atlantic forest. Biol J Linn Soc **91**:73–84.

Cabanne GS, D'Horta FM, Sari EHR, Santos FR, Miyaki CY (2008) Nuclear and mitochondrial phylogeography of the Atlantic Forest endemic *Xiphorhynchus fuscus* (Aves: Dendrocolaptidae): biogeography and systematics implications. Mol Phylogenet Evol **49**:760–773.

Camurugi F, Lima TM, Mercês EA, Juncá FA (2010) Anurans of the reserva ecológica da michelin, municipality of igrapiúna, state of Bahia, Brazil. Biota Neotrop **10**. http://www.biotaneotropica.org.br/v10n2/en/abstract?inventory+bn02810022010

Carnaval ACOQ, Bates J (2007) Amphibian DNA shows marked genetic structure and track Pleistocene climate change in northeastern Brazil. Evolution **61**:2942–2957.

Carnaval ACOQ, Moritz C (2008) Historical climate modeling predicts patterns of current biodiversity in the Brazilian Atlantic Forest. J Biogeogr **35**:1–15.

Carnaval ACOQ, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C (2009) Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. Science **323**:785–789.

Carvalho-e-Silva AMT, Silva GR, Carvalho-e-Silva SP (2008) Anurans at Rio das Pedras Reserve, Mangaratiba, RJ, Brazil. Biota Neotrop **8**. http://www.biotaneotropica.org.br/v8n1/en/abstract?inventory+bn02608012008

Chapple DG, Chapple SNJ, Thompson MB (2011) Biogeographic barriers in south-eastern Australia drive phylogeographic divergence in the garden skink, *Lampropholis guichenoti*. J Biogeogr **38**:1761–1775.

Costa LP, Leite YLR (2012) Historical fragmentation shaping vertebrate diversification in the Atlantic Forest biodiversity hotspot. Bones, clones, and biomes. In: Patterson BD, Costa LP (eds), The History and Geography of Recent Neotropical Mammals. Chicago Press, Chicago, IL, pp 283–306.

Crawford AJ (2003) Huge populations and old species of Costa Rican and Panamanian dirt frogs inferred from mitochondrial and nuclear gene sequences. Mol Ecol **12**:2525–2540.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol **7**:214.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLoS Biol **4**:699–710.

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. Meth Ecol Evol **1**:330–342.

Encarnação LC (2010) Reavaliação do status taxonômico de *Dendropsophus elegans* (Wied-Neuwied, 1824) (Amphibia: Anura: Hylidae). Master theses. Universidade Federal da Bahia, Salvador.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online **1**:47–50.

Fitzpatrick SW, Brasileiro CA, Haddad CFB, Zamudio KR (2009) Geographical variation in genetic structure of an Atlantic Coastal forest frog reveals regional differences in habitat stability. Mol Ecol **18**:2877–2896.

Forlani MC (2010) Morfologia do gênero Chiasmocleis Méhely, 1904 (Anura, Microhylidae, Gastrophryninae) e suas implicações filogenéticas. Master theses. Universidade de São Paulo, São Paulo.

Forlani MC, Bernardo PH, Haddad CFB, Zaher H (2010) Herpetofauna of the Carlos Botelho State Park, São Paulo State, Brazil. Biota Neotrop **10**. http://www.biotaneotropica.org.br/v10n3/en/abstract?inventory+bn00210032010

Fouquet A, Ledoux JB, Dubut V, Noonan BP, Scotti I (2012) The interplay of dispersal limitation, rivers, and historical events shapes the genetic structure of an Amazonian frog. Biol J Linn Soc **106**:356–373.

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

Gehring PS, Pabijan M, Randrianirina JE, Glaw F, Vences M (2012) The influence of riverine barriers on phylogeographic patterns of Malagasy reed frogs (Heterixalus). Mol Phylogenet Evol **64**:618–632.

Graham CH, Ron SR, Santos JC, Schneider CJ, Moriz C (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in Dendrobatidae frogs. Evolution **58**: 1781–1793.

Grazziotin FG, Monzel M, Echeverrigaray S, Bonatto SL (2006) Phylogeography of the *Bothrops jararaca* complex (Serpentes: Viperidae): past fragmentation and island colonization in the Brazilian Atlantic Forest. Mol Ecol **15**:3969–3982.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol **52**: 696–704.

Hammer O, Harper DAT, Ryan PD (2001) Past: Paleontological statistics software package for education and data analysis. Palaeontol Electron **4**:1–9.

Hampe A (2004) Bioclimatic models: what they detect and what they hide. Glob Ecol Biogeogr **11**:469–471.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol **25**:1965–1978.

Hoskin CJ, Tonione M, Higgie M, MacKenzie JB, Williams SE, VanDerWal J, Moritz C (2011) Persistence in peripheral refugia promotes phenotypic divergence and speciation in a rainforest frog. Am Nat **178**:561–578.

Irwin DE (2002) Phylogeography breaks without geograpic barriers to gene flow. Evolution **56**:2383–2394.

Izeckson E, Carvalho-e-Silva SP (2001) Anfíbios do Município do Rio de Janeiro. Editora UFRJ, Rio de Janeiro.

Köhler J, Koscinski D, Padial JM, Chaparro JC, Handford P, Lougheed SC, De la Riva I (2010) Systematics of Andean gladiator frogs of the *Hypsiboas pulchellus* species group (Anura, Hylidae). Zool Scr **39**: 572–590.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics **23**:2947–2948.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25**:1451–1452.

Macey JR, Larson A, Ananjeva NB, Fang Z, Papenfiss TJ (1997) Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Mol Biol Evol **14**:91–104.

Martins FM (2011) Historical biogeography of the Brazilian Atlantic Forest and the Carnaval–Moritz model of Pleistocene refugia: what do phylogeographical studies tell us? Biol J Linn Soc **104**:499–509.

Martins FM, Ditchfield AD, Meyer D, Morgante JS (2007) Mitochondrial DNA phylogeography reveals marked population structure in the common vampire bat, *Desmodus rotundus* (Phyllostomidae). J Zool Syst Evol Res **45**:372–378.

van der Meijden A, Vences M, Hoeeg S, Boistel R, Channing A, Meyer A (2007) Nuclear gene phylogeny of narrow-mouthed toads (Family: Microhylidae) and a discussion of competing hypotheses concerning their biogeographical origins. Mol Phylogenet Evol **44**:1017–1030.

Neaves LE, Zenger KR, Prince RIT, Eldridge MDB (2012) Impact of Pleistocene aridity oscillations on the population history of a widespread, vagile Australian mammal, *Macropus fuliginosus*. J Biogeogr **39**:1545–1563.

Nicolas V, Missoup AD, Denys C, Krebis Peterhans J, Katuala P, Couloux A, Colyn M (2011) The roles of rivers and Pleistocene refugia in shaping genetic diversity in *Praomys misonnei* in tropical Africa. J Biogeogr **38**:191–207.

Nogués-Bravo D (2009) Predicting the past distribution of species climatic niches. Glob Ecol Biogeogr **18**:521–531.

Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evoyylutionary Biology Center, Uppsala University, Uppsala.

Olalla-Tárraga MÁ, McInnes L, Bini LM, Diniz-Filho JAF, Fritz SA, Hawkins BA, Hortal J, Orme CDL, Rahbek C, Rodrőgues MÁ, Purvis A (2001) Climatic niche conservatism and the evolutionary dynamics in species range boundaries: global congruence across mammals and amphibians. J Biogeogr **38**:2237–2247.

Palma-Silva C, Lexer C, Paggi GM, Barbara T, Bered F, Bodanese-Zanettini MH (2009) Range-wide patterns of nuclear and chloroplast DNA diversity in *Vriesea gigantea* (Bromeliaceae), a neotropical forest species. Heredity **103**:503–512.

Pellegrino KCM, Rodrigues MT, Waite AN, Morando M, Yassuda YY, Sites JW Jr (2005) Phylogeography and species limits in the *Gymnodactylus darwinii* complex (Gekkonidae, Squamata): genetic structure coincides with river systems in the Brazilian Atlantic Forest. Biol J Linn Soc **85**:13–26.

Peterson AT, Soberón J, Pearson RG, Anderson RP, Martőnez-Meyer E, Nakamura M, Araújo MB (2011) Ecological Niches and Geographic Distributions. Princeton University Press, Princeton, NJ.

Phillips SJ, Dudők M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography **31**:161–175.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model **190**: 231–259.

Pöyry J, Luoto M, Heikkinen RK, Kuussaari M, Saarinen K (2009) Species traits explain recent range shifts of Finnish butterflies. Glob Change Biol **15**:732–743.

Prado GM, Pombal JP Jr (2005) Distribuição espacial e temporal dos anuros em um brejo da Reserva Biológica de Duas Bocas, sudeste do Brasil. Arquivos do Museu Nacional **63**:685–705.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. Mol Biol Evol **19**:2092–2100.

Recuero E, Garcőa-Parős M (2011) Evolutionary history of *Lissotriton helveticus*: multilocus assessment of ancestral vs. recent colonization of the Iberian Peninsula. Mol Phylogenet Evol **60**:170–182.

Resende HC, Yotoko KSC, Delabie JHC, Costa MA, Campiolo S, Tavares MG, Campos LAO, Fernandes-Salomão TM (2010) Pliocene and Pleistocene events shaping the genetic diversity within the central corridor of the Brazilian Atlantic Forest. Biol J Linn Soc **101**:949–960.

Ricklefs RE (2012) Host-pathogen coevolution, secondary sympatry and species diversification. Philos Trans R Soc Lond B Biol Sci **365**: 1139–1147.

Silva HR, Carvalho ALG, Bittencourt-Silva GB (2008) Frogs of Marambaia: a naturally isolated Restinga and Atlantic Forest remnant of southeastern Brazil. Biota Neotrop **8**. http://www.biotaneotropica.org.br/v8n4/en/abstract?inventory+bn01808042008.

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

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28**:2731–2739.

Thomé MTC, Zamudio KR, Giovanelli JGR, Haddad CFB, Baldissera FA Jr, Alexandrino J (2010) Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. Mol Phylogenet Evol **55**:1018–1031.

Vences M, Wake DB (2007) Speciation, species bound-aries and phylogeography of amphibians. In: Heatwole HH, Tyler M (eds), Amphibian Biology, Vol 6. Systematics Surrey Beatty & Sons, Chipping Norton, NSW, pp 2613–2669.

Waltari E, Hijmans RJ, Peterson A, Nyári AS, Perkins SL, Guralnick RP (2007) Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. PLoS ONE **2**:e563.

Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecol Appl **21**:335–342.

Wiens JJ, Kuczynski CA, Hua X, Moen DS (2010) An expanded phylogeny of treefrogs (Hylidae) based on nuclear and mitochondrial sequence data. Mol Phylogenet Evol **55**:871–882.

Zeisset I, Beebee TJC (2008) Amphibian phylogeography: a model for understanding historical aspects of species distributions. Heredity **101**:109–119.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Localities (specific locality, state) with sampled tissues and vouchers numbers.

Appendix S2. Phylogenetic trees.
Appendix S3. Genetic divergences and F_{ST} scores.
Table S1. Genetic divergences among clades of *D. elegans*.
Table S2. Genetic divergences among clades of *C. carvalhoi*.
Table S3. F_{ST} scores among clades of *D. elegans*.
Table S4. F_{ST} scores among clades of *C. carvalhoi*.
Table S5. DNA polymorphism and isolation by distance results of the mitochondrial gene *ND2* for clades of *D. elegans*.
Table S6. DNA polymorphism and isolation by distance results of the mitochondrial gene ND2 for clades of *C. carvalhoi*.