Phylogeography of the Bothrops jararaca complex (Serpentes: Viperidae): past fragmentation and island colonization in the Brazilian Atlantic Forest

FELIPE G. GRAZZIOTIN,* MARKUS MONZEL,† SERGIO ECHEVERRIGARAY‡ and SANDRO L. BONATTO*

*Centro de Biologia Genômica e Molecular, PUCRS, 90619-900, Porto Alegre, RS, Brazil, †University of Trier, Department of Biogeography, D-54296 Trier, Germany, ‡Laboratório de Biotecnologia Vegetal e Microbiologia Aplicada, UCS, 95070-560, Caxias do Sul, RS, Brazil

Abstract

The Brazilian Atlantic Forest is one of the world’s major biodiversity hotspots and is threatened by a severe habitat loss. Yet little is known about the processes that originated its remarkable richness of endemic species. Here we present results of a large-scale survey of the genetic variation at the mitochondrial cytochrome b gene of the pitviper, jararaca lancehead (Bothrops jararaca), and two closely related insular species (Bothrops insularis and Bothrops alcatraz), endemic of this region. Phylogenetic and network analyses revealed the existence of two well-supported clades, exhibiting a southern and a northern distribution. The divergence time of these two phylogroups was estimated at 3.8 million years ago, in the Pliocene, a period of intense climatic changes and frequent fragmentation of the tropical rainforest. Our data also suggest that the two groups underwent a large size expansion between 50 000 and 100 000 years ago. However, the southern group showed a more marked signal of population size fluctuation than the northern group, corroborating evidences that southern forests may have suffered a more pronounced reduction in area in the late Pleistocene. The insular species B. alcatraz and B. insularis presented very low diversity, each one sharing haplotypes with mainland individuals placed in different subclades. Despite their marked morphological and behavioural uniqueness, these two insular species seem to have originated very recently and most likely from distinct coastal B. jararaca populations, possibly associated with late Pleistocene or Holocene sea level fluctuations.

Keywords: Bothrops alcatraz, Bothrops insularis, Bothrops jararaca, Brazilian Atlantic Forest, island speciation, phylogeography

Received 4 April 2006; revision accepted 19 June 2006

Introduction

The Brazilian Atlantic Forest has been considered one of the eight major biodiversity hotspots for conservation (Myers et al. 2000), but the processes through which its striking diversity was formed are controversial, and the knowledge about patterns of endemism and species phylogeography is poor in contrast to some other areas (see Moritz et al. 2000). The most well-known hypothesis about the origin of Brazilian Atlantic Forest diversity (as well as most of tropical forests) is based on the classical Pleistocene Refugia Model (Haffer 1969; Vanzolini & Williams 1981), which rests on the assumption that rainforest oscillations triggered by the climatic changes of the Pleistocene (glacial periods) promoted speciation. In this model, alternate cycles of fragmentation and reconnection of forests during the Pleistocene would result in repeated vicariant events with the consequent differentiation of species or populations. An alternative model, although also based on the principle of allopatric diversification, invokes older phenomena like the orogeny in the Tertiary to produce this diversity (Simpson 1979). Under this latter model, the uplift of the Brazilian east coast in the Tertiary would have produced geographic, hydrographic, and climatic modifications in the Brazilian Atlantic Forest.
region, which might have caused forest fragmentation with divergence in the associated fauna. Some recent genetic studies in Brazilian rainforest species (mostly Amazon’s species, e.g. Ribas & Miyaki 2004; Wüster et al. 2005b) have found results in agreement with Pleistocene refugia while others suggested a major influence of Pliocene (or pre-Pleistocene) events in the diversification of species or phylogroups (e.g. Lara & Patton 2000; Glor et al. 2001). On the other hand, a comprehensive study on the biogeography of Neotropical mammals (Costa 2003) suggested that their origin would not be explained by a single model of old vicariance or recent climatic changes. A more complex model of diversification implies a larger number of idiosyncratic patterns, so that a comprehensive understanding of general patterns would require numerous studies with different species. Unfortunately, the number of phylogeographic studies with adequate sampling in the Brazilian Atlantic Forest is still very limited, especially in its southern and central regions.

We focused our study on the Bothrops jararaca species complex, which comprises three closely related Atlantic Forest species of lancehead pitvipers (Marques et al. 2002): the mainland B. jararaca and the two insular species Bothrops insularis and Bothrops alcatraz. Bothrops jararaca (jararaca lancehead) is a semiarboreal pitviper distributed primarily in southern and southeastern Brazil (Fig. 1, Campbell & Lamar 2004), a region that comprises the most populated areas of Brazil, in which less than 7.5% of the original vegetation remains (Myers et al. 2000). Despite this fact, this species is abundant in this area, being the principal cause of snakebite envenomation in Brazil (Ribeiro & Jorge 1990). Bothrops jararaca is a forest dweller and may be found in evergreen or semideciduous broadleaf forest, closed scrub and in highly degraded formations and even cultivated fields (Sazima 1992). The ventral and dorsal colours, and patterns and number of scales are highly variable within and among populations, and also among some geographical regions (Hoge et al. 1977). The existence of some geographical structure in the morphology of B. jararaca prompted Salomão et al. (1997) to suggest that this taxon may represent a complex of several species. In the only study presenting information on genetic diversity within B. jararaca (Wüster et al. 2005a), there was already some indication of the existence of genetic structure in this species, although the sample size was limited to only nine individuals. In spite of these interesting features, current knowledge on the genetic diversity and the evolutionary history of B. jararaca and related species is still extremely limited.

One of these related species is B. insularis, the golden lancehead, which is restricted to Queimada Grande Island, about 30 km off the coast of São Paulo State (SP), southeastern Brazil. This species is more arboreal than mainland

---

Fig. 1 Geographical distribution of the Bothrops jararaca complex with sampling localities. Light grey area represents distribution of B. jararaca. Squares indicate individuals from the northern clade and crosses indicate individuals from the southern clade. Solid line and dashed line indicate the main genetic barrier as defined by samova and ais, respectively. Brazilian states: RS, Rio Grande do Sul; SC, Santa Catarina; PR, Paraná; SP, São Paulo; RJ, Rio de Janeiro; ES, Espírito Santo; MG, Minas Gerais; MS, Mato Grosso do Sul; and GO, Goiás.
**B. jararaca**; its diet is predominantly based on migrant passerine birds (Martins et al. 2001), its venom being very effective in killing this prey (Cogo et al. 1993). This species was considered by Salomão et al. (1997) the sister taxon to **B. jararaca**, but they noted that **B. insularis** may be placed within **B. jararaca** in the molecular phylogeny, rendering **B. jararaca** a paraphyletic taxon. Wüster et al. (2005a) found that the **B. insularis** haplotype is actually rooted within **B. jararaca**, but the very limited sampling in their study precluded an assessment of the phylogenetic relationships between **B. insularis** and **B. jararaca** haplotypes. The second related insular species is the recently described **B. alcatraz** (Alcatraz lancehead, Marques et al. 2002), endemic to the Alcatrazes Archipelago, also about 30 km off the coast of SP and more than 100 km north of Queimada Grande Island. **Bothrops alcatraz** is paedomorphic, has a smaller adult size and larger eyes than the continental **B. jararaca** and feeds primarily on ectothermic prey (centipedes and lizards), a characteristic of juvenile mainland vipers, and its venom is similar to that of juveniles of **B. jararaca** (Marques et al. 2002). Although Furtado (2005) has reported quantitative and qualitative differences between the venom of the Alcatraz lancehead and that of mainland **B. jararaca**, nothing is known about the origin or genetic variability of this species. Both insular species are categorized as critically endangered in the Red Lists of IUCN and IBAMA (Brazil), due to its high endemism and the continuous decline in habitat quality by human impact on both islands.

The aim of the present study was to assess the genetic variability of the mitochondrial cytochrome b (cyt b) gene in a large sample throughout most of the range of **B. jararaca** and the two insular species, and to infer their genetic structure and evolutionary history. The results were interpreted considering the Pleistocene vs. Tertiary competing scenarios for the origin of Brazilian Atlantic Forest biodiversity. Several other questions related to the evolutionary history of **B. jararaca** complex arise: (i) What is the degree of genetic diversity of each species within the complex? (ii) Are there signals of fluctuation in population size and what are the approximated dates for these fluctuations? (iii) What is the extent of the geographical structure in the mitochondrial DNA (mtDNA) variability? (iv) What are the approximate dates of divergence among the phylogroups? (v) What are the phylogenetic relationships among the three species of the complex, and what are their taxonomic implications? (vi) How and when did **B. insularis** and **B. alcatraz** colonize the islands? (vii) What are the implications for conservation?

**Materials and methods**

**Population sampling and molecular methods**

A total of 159 specimens of **Bothrops jararaca** from 94 localities were sampled covering most of the species’ range (Fig. 1, Table S1, Supplementary material) plus seven **Bothrops insularis** and five **Bothrops alcatraz** individuals. DNA was extracted from scales, blood, liver or shed skin following published protocols specific for each tissue (Bricker et al. 1996; Hillis et al. 1996). DNA from a few formalin-fixed museum specimens was also extracted using a modified protocol from Chatignon (2000).

A 726-bp region of the cytochrome b gene was amplified via polymerase chain reaction (PCR) using the primers and protocols described by Pook et al. (2000). Amplicons were purified with shrimp alkaline phosphatase and exonuclease I (Amersham Biosciences) and sequenced using the DYEnamic ET Dye Terminator Cycle Sequencing Kit (Amersham Biosciences) in a MegaBACE 1000 automated sequencer (Amersham Biosciences) following the manufacturer’s protocols. Chromatograms were checked with the CHROMAS software (Technelysium) and sequences were manually edited using BIOEDIT 6.0.7 (Hall 1999).

**Sequence alignment and phylogenetic inferences**

Sequences were aligned using CLUSTAL_X 1.83 (Thompson et al. 1997) and corrected by eye. The substitution model used for the phylogenetic reconstructions was estimated with modeltest 3.7 (Posada & Crandall 1998) using the minimum theoretical information criterion (AIC) and the Bayesian information criterion (BIC) as suggested by Posada & Buckley (2004).

Phylogenetic trees were constructed by the maximum-likelihood (ML), maximum-parsimony (MP), and neighbour-joining (NJ) methods by using the program PAUP* 4.0b10 (Swofford 2002), following a ‘pluralist and critical approach’ as suggested by Thornton & Kolarczkowski (2005). Bayesian inference (BI) phylogenetic inference was also performed by using mrbayes 3.0b4 (Huelsenbeck & Ronquist 2001) with 1 000 000 cycles for the Markov chain Monte Carlo (MCMC) algorithm using flat priors. The posterior probabilities were calculated with only the last 9000 sampled trees after the log-likelihood values had stabilized. For the other three methods branch confidence values were estimated using 1000 bootstrap replicates. We inferred ML trees with a heuristic search [tree-bisection–reconnection (TBR)] option and an NJ starting tree; confidence was estimated by bootstrap using the nearest-neighbour interchange (NNI) heuristic search option. MP was performed by heuristic search (TBR) with starting trees produced by 1000 replications of random stepwise addition. To assess the MP statistical confidence, bootstrap replicates were conducted using the same heuristic search, but with starting trees obtained from simple stepwise addition. NJ analysis used the ML distance under the evolutionary model selected by modelltest. We used Bothrops erythromelas and Bothrops atrox as outgroup species, based on Wüster et al. (2002) results where **B. erythromelas** is one of the more
closely related species to the *B. jararaca* complex and *B. atrox* is a relatively more distant species.

We also constructed a haplotype network by the median-joining method (MJN) (Bandelt et al. 1999) using the program NETWORK 4.1.0.8 (www.fluxus-engineering.com) as it seems to give similar or better results than competing methods (Cassens et al. 2005).

Divergence times between clades were estimated using the approach employed by Wüster et al. (2005b) to calculate the divergence among clades of the Neotropical rattlesnake. We used the r8s 1.50 program (Sanderson 2003) with three calibration points: the divergence of South American population of the genus *Porthidium*, fixed in 3.5 million years ago (Ma); the divergence between Asian and African cobras, constrained to a minimum age of 16 Ma; and the divergence between Elapidae and Viperidae, constrained to a minimum age of 95 Ma. The confidence interval was estimated following the users manual and the rate of substitution per site per year was calculated using the penalized-likelihood method (Sanderson 2002) with the optimal smoothing value estimated by cross-validation.

**Population structure and phylogeographical history**

Individuals were grouped according the municipal limits. As the exact geographic location was not available for most of the samples, latitudinal and longitudinal positioning was inferred using the geographical centre of the town where the individuals were collected. To group samples in geographic populations and to estimate possible barriers to gene flow we used the spatial analysis of molecular variance (Dupanloup et al. 2002) implemented in samova 1.0 (http://web.unife.it/progetti/genetica/Isabelle/samova.html) and the Monmonier maximum difference algorithm (Monmonier 1973) implemented in the ari 1.0 program (Miller 2005). The genetic structures observed using the group definitions from these approaches were then contrasted using an analysis of molecular variance (amova, Excoffier et al. 1992) performed in arlequin 2.0 (Schneider et al. 2000).

Statistics such as nucleotide (π) and haplotype diversity (Hₒ), Tajima’s (Tajima 1983), Fu (Fu 1997) and Fu & Li’s (Fu & Li 1993) neutrality tests, F-statistics (Fₛₜ; Hudson et al. 1992), number of migrants per generation [Nm, using Fₛₜ = 1/(1 + 2Nm)], and mismatch distribution analyses (Rogers & Harpending 1992) were estimated using dnasp 4.0 (Rozas et al. 2003) and arlequin 2.0.

Female effective population size (*Nₑ*) was estimated using *Nₑ* = *θ* / 2 *µ* (*µ* = evolutionary rate per generation, *θ* = theta, see below). Generation time for *B. jararaca* was estimated as 5.5 years using the data from Sazima (1992), which suggested that *B. jararaca* attain adult size in its third or fourth year and it lives approximately 10–12 years (Sazima 1992). We estimated generation time as an average of the youngest reported age at maturity and the shortest reported life span minus 1 year as a compensation for probability of survival until old ages.

Several approaches were used to estimate and test different phylogeographical scenarios. First, we employed the program mdiv (Nielsen & Wakeley 2001) which uses an MCMC method within a likelihood framework to estimate the posterior distributions of: theta (*θ* = 2*Nₑµ*); the migration rate per generation (*Mₘdiv* = *Nₑµ*; *m*, migration rate); and the divergence time between populations (equations adjusted for mtDNA). The program also estimates the expected time to the most recent common ancestor (*T 이야기* for all sequences in the sample. Five runs were performed with 5 000 000 cycles each for the MCMC and the burn-in time of 10% as recommended by the program manual. The posterior probability distributions for all parameters are shown in the Supplementary material.

Population size fluctuation and demographic parameters such as *Nₑ* and the migration rates for each population (*Mₖₚₛ* = *m*/*µ* ) were tested using the MCMC method implemented in the package lamarc 2.0.2 (Kuhner 2006). We used the substitution model selected in modeltest and set the initial parameters of gene flow among populations with the migration rates obtained with mdiv. Our search strategy was composed of three replicates of 10 initial chains and two long final chains. The initial chains were performed with 500 samples and a sampling interval of 20 (10 000 steps), using a burn-in of 1000 samples for each chain. The two final chains were carried out with the same burn-in and interval sampling, but with 10 000 samples (200 000 steps). The confidence interval for theta and growth rate was calculated using the percentile approach. To estimate the ancestral *Nₑ* at ‘t’ time ago we used the standard equation of population growth:

\[ \theta_t = \theta_{\text{now}} e^{-gt} \]

Where ‘*θ*’ is 2*Nₑµ*, ‘*t*’ is the time in mutational units, and ‘*g*’ is the exponential growth parameter.

We used the Bayesian Skyline Plot method implemented in the program BEAST 1.2 (Drummond & Rambaut 2003) to estimate the dynamics of the population size fluctuation along the time. This Bayesian approach incorporates the uncertainty in the genealogy by using MCMC integration under a coalescent model, where the timing of divergence dates provides information about effective population sizes through time. We used the evolutionary model suggested by modeltest and a length chain of 100 000 000. The *T 이야기* for all sequences and for all major clades were also estimated using this approach to compare with *T 이야기* from mdiv and the results of the methods implemented in r8s.
Two haplotypes were found in widespread common haplotype, identical to the most frequent and from mainland individuals and that is also identical to one individual (N26) found in the other six exclusive of this species and encountered in a single (Table 1). All five individuals of Bothrops jararaca variable sites (9.7%) and 45 haplotypes were found in the (GenBank Accession nos AY865653–AY865824). Sixty-one individuals of the Bothrops erythromelas two outgroups (Fig. 2). The median-joining network (Fig. 3) indicated 14 mutational steps between the southern and northern clades (Sc and Nc), in agreement with the phylogenies (Fig. 2). The subclades identified in the phylogenetic trees were also found in the network, sometimes as part of a reticulation. Again, individuals bj164 (N04) and bj144 (S12) were the only ones whose geographical locations do not agree with their haplotypes belonging to Nc and Sc clusters, respectively.

Within these two major clades there are some subclades that were well supported by most of the analyses (Fig. 2), such as the southern subclades Sc1 and Sc2 which did not present a clear geographic structure. In the northern clade, six geographically structured subclades were identified in all analyses (Nc1, Nc2, Nc3, Nc4a, Nc4b, and Nc5) with high support in at least one phylogeny. A seventh subclade (Ncmc) that comprises a large group of dispersed haplotypes was also found in some analyses, although with a low support. The relationship between the subclades within Nc varied considerably among the trees and received very low statistical support in most of them, in such a way that the branching order among these subclades could be regarded as unresolved.

The two haplotypes (N26 and N27) found in B. insularis from Queimada Grande Island grouped together with a haplotype (N31) found in one B. jararaca (bj329) from the São Bernardo do Campo region, in the strongly supported Nc4a subclade (Fig. 2). This corroborates the close relationship between the São Bernardo do Campo mainland population with the Queimada Grande Island population suggested by the presence of the shared haplotype (N26).

The median-joining network (Fig. 3) indicated 14 mutational steps between the southern and northern clades (Sc and Nc), in agreement with the phylogenies (Fig. 2). The subclades identified in the phylogenetic trees were also found in the network, sometimes as part of a reticulation. Again, individuals bj164 (N04) and bj144 (S12) were the only ones whose geographical locations do not agree with their haplotypes belonging to Nc and Sc clusters, respectively.

Population structure and genetic diversity
Both the SAMOVA and Monmonier approaches suggested that the main barrier to gene flow is localized approximately between the SP and PR states (Fig. 1), near the Paranaapanema River, a tributary of Paraná River, although it was not possible with the SAMOVA approach to determine the best number of geographic groups (see Discussion in Dupanloup et al. 2002). This putative barrier divides B. jararaca into a southern (SG) and a northern (NG) group, similarly as we found in the phylogenetic analyses. Based on the

Table 1 Summary statistics observed in southern group (SG), northern group (NG) and the whole distribution of Bothrops jararaca complex

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>S</th>
<th>h</th>
<th>H_d</th>
<th>π</th>
<th>D</th>
<th>F'</th>
<th>D'</th>
<th>F_s</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>76</td>
<td>32</td>
<td>15</td>
<td>0.862 ± 0.022</td>
<td>0.00761 ± 0.00074</td>
<td>−0.99754</td>
<td>−2.71499*</td>
<td>−3.05309*</td>
<td>−0.362</td>
</tr>
<tr>
<td>NG</td>
<td>95</td>
<td>37</td>
<td>31</td>
<td>0.904 ± 0.0061</td>
<td>0.01081 ± 0.00065</td>
<td>−0.27207</td>
<td>0.16639</td>
<td>0.40752</td>
<td>−6.355</td>
</tr>
<tr>
<td>whole</td>
<td>171</td>
<td>61</td>
<td>45</td>
<td>0.943 ± 0.0088</td>
<td>0.02068 ± 0.00047</td>
<td>0.47881</td>
<td>0.58120</td>
<td>0.48785</td>
<td>−3.763</td>
</tr>
</tbody>
</table>

N, number of sequences; S, number of variable sites; h, number of haplotypes; H_d, haplotype diversity; π, nucleotide diversity; D, Tajima’s D; F’, Fu and Li’s F; D’, Fu and Li’s D; F_s, Fu’s F_s. *P < 0.05.
concordance among all results, further population analyses were performed considering also these two groups of populations.

We found a high genetic diversity for the *B. jararaca* complex as a whole and also when the two geographic groups are considered separately (Table 1). Diversity values were not estimated for the insular species due to their small sample sizes. The $F_{ST}$ between the two groups of populations (SG and NG) was high (0.711), as expected due to their almost perfect allopatry, with a small estimated rate of migrants per generation ($N_m = 0.21$) between them. Migration rates estimated by LAMARC and MDIV presented similar values, although LAMARC results suggested an asymmetric gene flow: $N_m = 0.22$ with MDIV (see Supplementary material for posterior probability, Fig. S1), and $N_m = 0.54$ from NG to SG and 0.28 from SG to NG with LAMARC. Similarly, the AMOVA test detected significant structuring between SG and NG: 69.98% of the genetic variation is found between the two groups of populations while only 16.22% are found among populations within groups, and the remaining 13.8% was found within populations. All the above results suggest a strong genetic structure between the two groups and support a model of low gene flow.

**Coalescent analyses and divergence times**

The cyt $b$ evolutionary rate in *B. jararaca* complex was estimated here to be $3.69 \times 10^{-9}$ substitutions per site per year, which is similar to those found in other studies, which ranged from $2.35 \times 10^{-9}$ to $7.20 \times 10^{-9}$ (Zamudio & Greene 1997; Pook et al. 2000; Wüster et al. 2002). *Bothrops jararaca* complex initial diversification was dated about 5.56 Ma (confidence interval, CI 2.43–6.93) using the penalized likelihood (PL) method of $r8s$. The Bayesian
approach implemented in the program BEAST yielded an estimated coalescence time for the whole complex at about 6.0 Ma (CI, 3.7–8.9 Ma), while the likelihood approach implemented in MDIV resulted in 5.1 Ma. The divergence time between the two major groups of populations (NG and SG) was estimated with MDIV to have occurred at 1.9 coalescence units ago corresponding to 3.87 Ma. The diversification within clades began more recently, and interestingly, most of the main subclades of Nc and Sc seems to have originated almost simultaneously, between 0.14 and 0.76 Ma (Table 2).

Demographic history

For the overall B. jararaca sample and for the NG population group the neutrality tests were not significant, being significantly negative only for the SG (Fu and Li’s $D$ and $F$) (Table 1). The mismatch distribution taking into account the whole species or the two geographic groups separately showed multimodal distributions (Fig. S4, Supplementary material), in agreement with the mainly not significant results of the neutrality tests except for the SG group. The point estimate for the ancestral $\theta$ calculated by the MDIV program for the whole species (9.416; see Fig. S3, Supplementary material for posterior probability) was converted into a historical $N_e$ of about 370 000 individuals. The current $\theta$ estimated in LAMARC (37.50, SD 1.82) resulted in a present $N_e$ of approximately 1 480 000 individuals. Using the coalescence time calculated with MDIV and the $g$ value estimated from LAMARC analysis (122.4, SE 17.31), the ancestral $N_e$ resulted in approximately 148 000 individuals, 2.5 times smaller than MDIV estimation. For the
two groups, NG showed a similar pattern with a relative small growth rate. However, SG displayed a larger growth rate, in agreement with the significant negative values of the neutrality tests. Values for the demographic parameters were presented in Table 3.

The scenario for the whole species presented by the Bayesian skyline plot showed a dual pattern; a long history of constant population size followed by a very recent expansion (Fig. 4a) estimated to have occurred in the Pleistocene about 0.1 Ma. The population expanded in size approximately 10 times and there is only a weak signal of a population bottleneck preceding it. When the two geographic groups were analysed separately this recent expansion was found in both (Fig. 4b), although the SG expansion seems to have occurred more recently.

Discussion

Phylogeographic patterns

The phylogenetic and network analyses of cyt b sequences with a large sampling of Bothrops jararaca identified a highly diverse species composed of two geographically structured phylogroups, with a deep-rooted divergence. The genetic diversity found within B. jararaca complex is relatively high when compared with other snakes from the same genus. For example, Wüster et al. (1999) found an uncorrected distance (cyt b) within the Bothrops atrox complex ranging from 0.4% to 4.8%, while the B. jararaca complex showed a similar distance ranging from 0.2% to 5.0%. However, B. atrox complex has a much wider distribution than B. jararaca complex and includes a minimum of seven species (Werman 2005). The high variability in B. jararaca is the likely outcome of the presence of the two phylogroups that diverged about 3.8 Ma.

The allopatry of the two phylogroups is almost complete, with only two individuals found in geographic groups different to their clades. The phylogeographic pattern of these two haplotypes is consistent with recent gene flow and not with incomplete lineage sorting, e.g. they are very similar to haplotypes from the other group (one is actually identical and the other has only a single substitution from the most common haplotype from the NG, see Avise 2000 for similar examples). Additionally,
Fig. 4 Bayesian skyline plot showing the effective population size fluctuation throughout time. (a) whole *Bothrops jararaca* complex (solid line, median estimations; grey area, confidence interval; arrow shows the maximum likelihood divergence time), (b) sequence divided by clades [solid lines, median estimations; grey areas, confidence intervals (CI)].
the localities of the individuals that carried these two haplotypes were found not very distant (about 100 km) from the region where the putative genetic barrier was suggested to have existed. Although this putative genetic barrier approximately corresponds to the present course of the Paranapanema River, it is very difficult to assess its historical importance to the genetic differentiation of these two groups. The present drainage of the Paraná River system was probably established in the Mid-Tertiary (Potter 1998) and has undergone many fluctuations in water level and flow, which are compatible with our estimated divergence times (> 3 Ma). Interestingly, in a recent paper Pellegrino et al. (2005) show that the genetic structure of lizards of the Gymnodactylus darwinii complex coincides with the river system in the northern regions of the Brazilian Atlantic Forest and that major coastal rivers in this region may have played a key role in its diversification. Their results found support in studies on the herpetofauna of sand dune regions around the São Francisco River in the state of Bahia (see Rodrigues 1996) and other distributional data. Unfortunately, their estimated divergence times among G. darwinii clades presented very large confidence intervals and the species range have a small area of overlap with B. jararaca. Nevertheless, we found a clade in the region between the Rio Doce and Paraiba Rivers (Nc1) that corresponds with one of their clades, but this B. jararaca clade has an overlapped distribution with clade Nc5 and haplotype N10. Rivers do not seem to be effective dispersal barriers to B. jararaca, as other rivers of comparative size exist throughout the range of the species without inducing genetic differentiation. The differences between the two phylogeographic patterns could be explained by biological differences between the species. While geckos of the genus Gymnodactylus are small terrestrial lizards that occur mostly in lowland forests near sea level in the coastal area (Pellegrino et al. 2005), B. jararaca has a larger body and a much broader range that surpass many coastal river systems in the Brazilian Atlantic Florest.

Is there any correlation between these two groups with morphology or other described variation? The only association we observed was presented by Hoge et al. (1977), describing an interesting variation in the number of ventral scales in B. jararaca. They analysed specimens from most of the distribution of the species in Brazil, and distinguished two overlapping groups, one with fewer ventral scales in the southern parts of the distribution (SC and PR states; median = 188 scales), and the other with a larger number of ventral scales in the northern parts of the range (RJ, MG and ES states; median = 203 scales). Specimens from SP state were morphologically intermediate between the two patterns. Hoge et al. (1977) considered average annual temperature to be functionally linked to the observed variation and referred to a study that correlated temperature variation and somite formation in Thamnophis elegans populations (Fox & Fox 1961). On the other hand, considering the results of the present study, this morphological distinction between southern and northern populations may also be explained by a past fragmentation event of the species range (as also suggested by Wüster et al. 2005a), instead of a recent ecological effect. In this case, the intermediate morphologies found in SP state may represent a hybrid zone between the southern and northern populations. However, the inference of the causes underlying this geographic variation requires further studies based on a large sample size in this sympatric area and also using nuclear markers.

Single locus estimation of population divergence should always be treated with caution as it faces two main limitations (Jennings & Edwards 2005): (i) overestimation due to polymorphism in the ancestral population; and (ii) a large variance due to the stochastic nature of the lineage sorting process. We tried to minimize these problems by estimating the divergence time between the two main geographic groups using different approaches, including the MDIV method, which is a complex maximum-likelihood estimator that takes in account demographic parameters such as migration rates as well as the coalescence-based program BEAST. In addition, deep divergence times as estimated here are usually not as affected by ancestral polymorphism as more recent events (Arbogast et al. 2002). Therefore, we argue that 3.87 Ma is likely a good estimate of divergence time between the two populations groups of B. jararaca.

Demographic dynamics

Bothrops jararaca, considered as a single group, presented a long period with a large constant population size and an intensive growth during the last 100 000 years as estimated by the skyline plot (Fig. 4), although the absolute population size could be inflated by the presence of two disjunct phylogroups. Considering the SG phylogroup separately, most results support the scenario of a bottleneck accompanied by very recent size and range expansion with gene flow. The multimodal mismatch distribution (Supplementary material, Fig. S4) suggests the bottleneck was moderate as a ragged distribution is expected if genetically distinct founders remained in the population (Rogers & Harpending 1992). On the other hand, the NG phylogroup showed a more complex pattern, including apparently opposing signals for population size changes (e.g. a small g in LAMARC and nonsignificant values for the neutrality tests, but a large recent growth in BEAST). However, these conflicting results could be explained by the differences of the two approaches: while the growth rate estimated in LAMARC and the neutrality statistics are point estimates that average out the whole history of the population, the skyline plot estimates population changes through time. Therefore, as the demographic history of the NG seems
dominated by a stationary population size (Fig. 4b), only approaches such as the skyline plot could detect with statistical significance very recent size expansions.

**Diversification patterns in the Brazilian Atlantic Forest**

In the following section, we attempt to establish a coherent scenario for the diversification of this species group, although limited by the few similar studies in the Atlantic Forest. The Pliocene seems to have been a period of active differentiation for many groups in South America (e.g. Cortés-Ortiz et al. 2003; see Moritz et al. 2000 for additional molecular studies), possibly due to changes in humidity, which ultimately promoted changes in the phytogeographic domains (Simpson 1979). In the Brazilian Atlantic Forest these changes were likely caused by the uplift of the coastal Brazilian mountains and the consequent interruption of precipitation in southeastern Brazil by the early Pliocene at about 5.6 Ma, which coincides with the transition from tropical humid to semiarid or arid conditions (Simpson 1979; Vasconcelos et al. 1992). This orogenic process deeply changed the phytogeography and climatic conditions of south and southeast areas of Brazil, and consequently fragmented Brazilian Atlantic Forest with drier areas. This uniform stable dry climate persisted until the Middle to Upper Pliocene when a gradual increase in humidity would have occurred (Vasconcelos et al. 1992). This pre-Pleistocene hypothesis was used by Lara & Patton (2000) to explain the deep genetic divergence among three clades observed in Atlantic Forest spiny rats of the genus *Trisomy*. Interestingly, the spiny rats are a common prey item identified in stomach contents of adult *B. jararaca* (Sazima 1992). Considering that the divergence between the northern and southern phylogroups of *B. jararaca* likely occurred > 3 Ma, we suggest that it may have also been induced by these events in the Pliocene. As *B. jararaca* is a forest dweller (Sazima 1992), it is likely that the forest fragmentation separated a single widespread ancestral species into isolated populations. This scenario, including the assumption of restricted gene flow, could have persisted, with the ultimate survival of only two lineages, until the establishment of more humid climatic conditions in the Early Pleistocene.

On the other hand, climatic changes that occurred throughout the Pleistocene in the Neotropics, including the Atlantic Forest, that may have formed the putative forest refuges, could explain the origin and diversification of the subclades found in *B. jararaca*, especially those that are geographically structured and diverged more recently than 1 Ma (Table 2). This is in agreement with several molecular studies that have found divergence times between species/lineages that coincide with these events (e.g. Lara & Patton 2000; Ribas & Miyaki 2004; Wüster et al. 2005b). Similarly, the dramatic increase in the population size that occurred between 50 000 and 100 000 years ago may be explained by the climatic changes in the late Pleistocene. Polar advections causing strong impact on the climate and vegetation have been documented for the late Pleistocene in both southern and southeastern Brazil (Ledru et al. 1996; Behling & Lichte 1997). The contraction and expansion of the forest was a recurrent process, and this last population expansion of *B. jararaca* may likely be an outcome of one of these events. Although our estimate of the age of the population expansion has a large interval, not providing a point date to correlate with a more specific event, this suggests that the expansion had begun before the Last Glacial Maximum in South America (LGM, c. 27 500 to c. 14 500 14C years ago) and that therefore, the LGM had little impact in the overall historical demography of *B. jararaca*, in contrast to the large changes suggested to have occurred in the vegetation (Behling 2002; Behling & Negrelle 2001)

The contrast between the Southern and Northern populations in the effective population size fluctuation (Fig. 4b) could be explained by a differential effect of drier periods in the two regions. More severe climatic oscillations in southern regions of Brazil have been documented by various studies with palaeopalinology. For example, some studies suggest that in the late Pleistocene the Brazilian Atlantic Forest in that region was replaced by grassland and moved about 500 Km further north (Behling & Lichte 1997; Behling & Negrelle 2001). Consequently, it is likely that the southern group underwent a more pronounced bottleneck that resulted in a stronger and more recent signal of population expansion (Table 2 and Fig. 4b).

Therefore, neither a scenario of Pleistocene forest refuges nor one of Pliocene orogeny in the Brazilian Atlantic Forest seem sufficient to explain alone our phylogeographic results with the *B. jararaca* complex. On the contrary, our results suggest that the divergence of the two main lineages of *B. jararaca* is associated with an old fragmentation event in the Brazilian Atlantic Forest during the Pliocene, but the estimated age of the diversification of the subclades and of the population expansion suggest that the genetic and geographic variability of subpopulations of *B. jararaca* may have been shaped by late Pleistocene climatic oscillations. These conclusions are in agreement with the remarks of Costa (2003) that the processes that originated the Neotropical biodiversity are complex, and that, for a complete picture to be obtained, many more phylogeographic studies are needed.

**The origin of the insular jararacas and the taxonomic and conservation status of the species of the complex**

According to Marques et al. (2002), the two insular species may have been originated from populations of a *B. jararaca*-like ancestor from the eastern coast of Brazil, which may have been isolated during one of the sea-level oscillations that occurred in the late Pleistocene. Our results (Figs 2
and 3) strongly corroborate the hypothesis of a very recent origin of *Bothrops insularis* and *Bothrops alcatraz*, but evidenced that these two species were derived from *B. jararaca* itself and not from any putative ancestral species. As the haplotypes found in the two insular species are different and belong to divergent and geographically separated subclades, it is likely that they have had independent origins. The very low genetic diversity found within each island suggests a possible founder effect for the origin of these insular species, although additional samples are needed to test this hypothesis.

Does the phylogenetic position of *B. insularis* and *B. alcatraz* mtDNA haplotypes inside *B. jararaca* clades represent a problem for the delimitation or validity of these species, as suggested by Salomão et al. (1997)? In fact, a gene genealogy pattern as found here is a very likely outcome when speciation involves small populations isolated from a larger ancestral one (peripatric or ‘budding’ speciation), as usually occurs in island colonization (Avise 2000). Actually, a large number of such ‘paraphyletic’ species phylogenies based on mtDNA are known; for instance, Funk & Omland (2003) found that 22.4% of the 147 reptile species surveyed are polyphylectic. More specifically, Johnson et al. (2000), discussing island biogeography models, showed that phylogenies of island populations showing paraphyly of mainland alleles occur when the mainland population is very large and exhibits geographic structure and/or when a short time has passed since colonization. Rapid speciation events were also recognized by Moritz (1994) as causing problems to the definition of Evolutionarily Significant Units based on the reciprocal monophyly concept, but argued that this does not affect conservation priorities since these taxa are frequently recognized as different species based on other biological criteria. Therefore, the nonmonophyly of the mtDNA haplotypes of the mainland *B. jararaca* and the very recent differentiation of the two island species as proposed here should not influence their taxonomic and conservation status. Many striking biological features observed separately in *B. insularis* and *B. alcatraz* (Marques et al. 2002) are not found in mainland *B. jararaca*, thus supporting their status as valid species and distinct evolutionary units for conservation.

Furthermore, the deep mtDNA lineage divergence found in *B. jararaca* suggests that, for conservation or medical purposes, this species might likely be divided into two separate geographical units, since 70% of its mitochondrial variability is found between the northern and southern groups. However, if dispersal in *B. jararaca* is mediated mainly by males, if females tend to be more sedentary as suggested by Szirmai (1992), this major phylogeographic pattern may not represent the main genetic constitution of the species. Further studies including biparentally inherited markers are necessary to better understand the current population structure as well as management and taxonomic implications for these two major evolutionary groups.

### Acknowledgements

We thank F.L. Franco, W. Fernandes and O.A.V. Marques (Instituto Butantan), A. Melgarejo (Instituto Vital Brazil), R. Maciel and G.A. Cota (FUNED), M. DiBernardo and G.F. Pontes (PUCRS), M. Leitão-de-Araújo (FZB-RS), P.J. Demeda (UCS) for collecting snakes and N.J.R. Fagundes, E. Eizirik, P. Müller, A. M. Solé-Cava, F. R. Santos, K. Zamudio, T. Lema, and the anonymous reviewers for suggestions on an earlier version of this manuscript. We thank to F.S.L. Vianna for laboratory assistance. K. Zamudio, M. Martins and O.A.V. Marques provided useful comments and collecting support (Fapesp, 1995/09642–5). F.G.G and M.M. were supported by a fellowship from CAPES, Brazil and DAAD, Germany, respectively. Funding was provided by CNPq and FAPERGS, Brazil, to S.L.B.

### Supplementary material

The supplementary material is available from http://www.blackwellpublishing.com/products/journals/suppmat/MEC/MEC3057/MEC3057sm.htm

**Table S1** Localities, haplotypes and vouchers for specimens used in this study

**Fig. S1** Posterior probability distribution for parameter M estimated with MDIV

**Fig. S2** Posterior probability distribution for parameter T estimated with MDIV

**Fig. S3** Posterior probability distribution for parameter Theta estimated with MDIV

**Fig. S4** Mismatch distribution for the Northern clade (A); Southern clade (B); whole distribution (C)

### References


