Recent lineage diversification in a venomous snake through dispersal across the Amazon River

H. LISLE GIBBS[1,2*], MICHAEL SOVIC[1], DIANA AMAZONAS[3], HIPPOCRATES CHALKIDIS[4], DAVID SALAZAR-VALENZUELA[5] and ANA M. MOURA-DA-SILVA[3]

[1]*Department of Evolution, Ecology and Organismal Biology, Ohio State University, Columbus, OH 43210, USA*
[2]*Ohio Biodiversity Conservation Partnership, Ohio State University, Columbus, OH 43210, USA*
[3]*Laboratório de Imunopatologia, Instituto Butantan, São Paulo, Brazil*
[4]*Faculdades Integradas do Tapajós, Santarém, Pará, Brazil*
[5]*Centro de Investigación de la Biodiversidad y Cambio Climático, Universidad Tecnológica Indoamérica, Quito, Ecuador*

Identifying the evolutionary and ecological mechanisms that drive lineage diversification in the species-rich tropics is of broad interest to evolutionary biologists. Here, we use phylogeographical and demographic analyses of genome-scale RADseq data to assess the impact of a large geographical feature, the Amazon River, on lineage formation in a venomous pitviper, *Bothrops atrox*. We compared genetic differentiation in samples from four sites near Santarem, Brazil, that spanned the Amazon and represented major habitat types. A species delimitation analysis identified each population as a distinct evolutionary lineage while a species tree analysis with populations as taxa revealed a phylogenetic tree consistent with dispersal across the Amazon from north to south. Phylogenetic analyses of mitochondrial DNA variation confirmed this pattern and suggest that all lineages originated during the mid- to late Pleistocene. Historical demographic analyses support a population model of lineage formation through isolation between lineages with low ongoing migration between large populations and reject a model of differentiation through isolation by distance alone. The results provide a rare example of a phylogeographical pattern demonstrating dispersal over evolutionary timescales across a large tropical river and suggest a role for the Amazon River as a driver of *in situ* divergence both by impeding (but not preventing) gene flow and through parapatric differentiation along an ecological gradient.

ADDITIONAL KEYWORDS: Amazon River – ecological gradients – evolutionary lineages – historical demography – phylogeographical dispersal – RADseq – riverine barriers – venomous snakes.

INTRODUCTION

An outstanding challenge for evolutionary biology is to account for the origin of the exceptional species diversity in tropical regions such as the Amazon basin (Mittermeier, Gil & Mittermeier, 1997). The analysis of genome-scale molecular data, when combined with information on phenotypic variation and distributional and geographical information, can provide important information for evaluating hypotheses about speciation in the tropics (Mortiz *et al.*, 2000; Leite & Rogers, 2013).

Specifically, such analyses can generate: (1) information on relationships among historical (phylogeographical) lineages within species and in particular identification of sister groups in relation to geography and habitat/ecological attributes (Patton & Silva, 1998; Patton, Silva & Malcolm, 2000); (2) estimates of the timing of divergence events (Antonelli *et al.*, 2010); and (3) estimates of rates of gene flow among populations (Slatkin, 1994), against which patterns of phenotypic differentiation can be compared (Orr & Smith, 1998). These analyses can be used to assess which of two broad, non-exclusive classes of speciation mechanisms – divergence in allopatry or ecological speciation along gradients – better accounts for lineage diversification

*Corresponding author. E-mail: gibbs.128@osu.edu

in a given taxon with respect to specific large-scale geographical features in the tropical environment.

In Amazonia, major rivers and their associated habitats have long been suggested as large-scale geographical features that have had a significant impact on species diversification (Antonelli *et al.*, 2010). More specifically, in terms of allopatric mechanisms of divergence, Wallace (1852) originally proposed what is now known as the riverine barriers hypothesis (Patton *et al.*, 2000), which argues that rivers in this region may shape the biogeographical distribution of species by separating widespread organisms into isolated populations and impacting lineage formation (Mortiz *et al.*, 2000; Patton *et al.*, 2000). Previous studies have yielded mixed support for this hypothesis. For example, the effects of riverine barriers have been used to explain the distributional limits in several Amazonian vertebrates such as birds (e.g. Bates, Haffer & Grismer, 2004) and mammals (Ayres & Clutton-Brock, 1992; Rocha *et al.*, 2015) whereas other studies have shown no evidence major rivers have promoted diversification in other animal groups (Gascon *et al.*, 2000; see Antonelli *et al.*, 2010 for a recent review).

Large rivers in the Amazon basin can also contain ecologically diverse habitats and these have been associated with diversification in specific groups of vertebrates such as birds (Alexio & Rosetti, 2007; Harvey *et al.*, 2017). For example, in eastern Brazil where the Amazon River is widest, there are distinct habitats such as Terra Firme (upland forest) on either bank and floodplain habitats known as Várzea associated with the river (Pires & Prance, 1985). These represent distinct habitats that could generate diverse selection pressures over limited spatial scales on sedentary taxa, leading to lineage divergence under a broad form of an ecological gradient model of diversification. In support of this, Terra Firme and Várzea habitats have been associated with phyogeographical, demographic and morphological diversification in multiple species of Amazonian birds (Alexio & Rossetti, 2007; Harvey *et al.*, 2017).

For a species with a trans-river distribution and evolutionarily distinct lineages, the riverine barrier and ecological gradient hypotheses make different predictions about the phylogeographical patterns, levels of gene flow and patterns of phenotypic differentiation between populations (Moritz *et al.*, 2000; Patton *et al.*, 2000; Antonelli *et al.*, 2010), which are summarized in Table 1. In broad terms, these represent two general models of diversification: allopatric differentiation due to geographical isolation from a barrier to gene flow and parapatric differentiation due to habitat-specific selection pressures in the face of gene flow. We stress that these predictions apply to populations that span a river in a given location, which is the focus of this study. As discussed by Antonelli *et al.* (2010), other evolutionary processes such as the impact of Pleistocene refugia could afftect the evolutionary history of species with broad geographical distributions but we are unable to assess the impact of such mechanisms due to our limited sampling. Finally, these are not mutually exclusive mechanisms – both could operate simultaneously to varying extents to shape genetic structure in local populations that span a river.

These predictions can be examined by the application of model-based analyses of large-scale genetic data sets, an approach that has been rarely applied to studies of tropical vertebrates (but see Harvey & Brumfield, 2015; Weir *et al.*, 2015; Prates *et al.*, 2016; Harvey *et al.*, 2017). Such approaches allow a statistical evaluation of key features such as the number of phylogenetically distinct lineages present in a set of populations (Leaché *et al.*, 2014). Historical demographic analyses (Excoffier *et al.*, 2013) permit both the statistical assessment of the fit of the data to different models of population structure and also the estimation of key population parameters such as levels of gene flow and genetically effective population sizes from the best-fit model. These are useful for evaluating the evolutionary mechanisms that underlie patterns of genetic differentiation (e.g. Sovic, Fries & Gibbs, 2016) such as those outlined in Table 1. Leite &

Table 1. Phylogeographical, population genetic and phenotypic patterns predicted under the riverine barrier and ecological gradient mechanisms for the impact of a large tropical river on a species with a trans-river distribution and that shows phylogeographical structure; predictions are derived from Mortiz *et al.* (2000), Patton *et al.* (2000) and Antonelli *et al.* (2010)

Mechanism	Riverine barrier	Ecological gradient
Phylogeographical pattern	Reciprocal monophyly of lineages associated with opposite sides of river	Sister lineages occupy ecologically distinct habitats
Gene flow	Absent between lineages associated with opposite sides of river	Present between lineages
Phenotypic differentiation	If present then associated with historical relationships of lineages	Associated with ecologically distinct habitats

Rogers (2013) have emphasized how useful yet rarely applied such an approach would be for assessing the evolutionary mechanisms that generate diversity in the Amazon region.

Previous work on the impact of riverine barriers on diversification in Amazonian vertebrates has focused on birds, mammals and, to a lesser extent, amphibians (for a review see Antonelli *et al.*, 2010) whereas there are few studies on another species-rich vertebrate group, reptiles (but see Marques, Trefaut & Cohn-Haft, 2013). To address the question of how riverine barriers influence differentiation in tropical reptiles over evolutionary timescales, we examined genetic structure in a common venomous pitviper, the common lancehead (*Bothrops atrox*) near the city of Santarem which is situated at the habitat-diverse confluence of the Amazon and Tapajos Rivers in Para State, Brazil. *Bothrops atrox* is a medium-sized terrestrial pitviper found in tropical lowlands and rainforest of northern South America east of the Andes (Campbell & Lamar, 2004). Previously, Wüster *et al.* (1999) used mitochondrial DNA (mtDNA) and morphological data to examine the phylogenetic distinctiveness of geographically widespread populations of this complex including the form (*B. atrox*) studied here. There is evidence for at least four distinct mtDNA clades across the currently defined range of *B. atrox* (*s.s.*) including groups on either side of the Amazon (summarized by Werman 2005). Despite this variation and its wide range of ecological and geographical habitats, no subspecies are currently defined (Campbell & Lamar, 2004) although *B. atrox* (*s.l.*) is generally considered to consist of a complex of differentiated forms (Werman, 2005).

Here, we use *B. atrox* as a model to examine how large rivers influence diversification in Amazonian reptiles. Our approach differs from previous work in reptiles in that we conduct detailed sampling over a limited spatial scale that explicitly encompasses habitat variation across the Amazon River rather than limited sampling over broad geographical scales, which is more appropriate for assessing long-term biogeographical hypotheses (e.g. Werman, 2005). We analyse RADseq data (Andrew *et al.*, 2016) using a range of coalescent-based analyses (Excoffier *et al.*, 2013; Leaché *et al.*, 2014) to assess phylogeographical structure and the historical demography of these populations in light of the predictions in Table 1 with the goal of inferring the extent to which these mechanisms play a role in moulding lineage variation in this snake. We also use previously published data (Sousa *et al.*, 2017) on the functional characteristics of venom from snakes analysed to assess variation in an adaptive trait across these populations, which is relevant to assess potential mechanisms of diversification (Moritz *et al.*, 2000).

METHODS

SAMPLE LOCATIONS AND COLLECTION PROTOCOLS

Samples from individual *B. atrox* were collected from four locations that span the Amazon River near the cities of Santarem and Oriximiná, in the western region of the state of Pará, Brazil (Fig. 1A). These include the following sites located from north to south across the river. (1) *Oriximiná*: eight adult snakes were collected from a pasture area in the municipality of Oriximiná on the north shore of the Amazon River (GPS 01°46′03.39″S, 55°43′53.31″W). This site was historically Terra Firme forest habitat (Pires & Prance, 1985) that was cleared of forest ~ 20 years earlier. This habitat typically consists of dense 'ombrofila' upland forest, characterized by large trees with one or two species that create a uniform canopy layer, between 25 and 35 m tall. (2) *Várzea*: four adults were collected from a seasonally flooded island in the main course of the Amazon near the Oriximiná site (01°52′24.42″S, 55º 56′01.85″W) that is typical of floodplain habitat in this region (Pires & Prance, 1985). Várzea habitat is subject to periodic flooding by the Amazon River and is formed by the deposition of sediments that have led to the formation of many islands in this area. The typical vegetation consists of grasses which grow on highly fertile alluvial soils. (3) *Savanna*: four snakes were captured at a site along the Tapajos River that is representative of the distinct white sand Amazonian savannah habitat found in this region (Magnusson *et al.*, 2008) (S 02°26′49.97″S, 54°54′8.9″W). These regions have low-fertility quartz sand soils with limited vegetation consisting mostly of grasses with occasional bushes and small trees. (4) FLONA: 24 individuals were collected within or next to the Floresta Nacional do Tapajós, a protected area located in the municipality of Belterra next to the Tapajos River about 50 km south of the Amazon River (03°03′59.03″S, 54°58′8.94″W). This site also represents upland Terra Firme forest (Pires & Prance, 1985) but from the south shore of the Amazon (see Espirito-Santo *et al.*, 2005 for details).

At each of these sites individual snakes were collected using a combination of pitfall traps and opportunistic sampling. Blood was collected from each snake and stored in 95% EtOH for DNA analyses

GENETIC ANALYSES

Genomic DNA was extracted from samples and double-digest RADseq libraries (Peterson *et al.*, 2012) were generated following the protocol described by Sovic *et al.* (2016), which selects for fragments between 300 and 450 bp from libraries of DNA digested with *Sbf*I and *Eco*RI restriction enzymes. Individual libraries were pooled and sequenced as single-read 50-bp runs on an Illumina HiSeq 2000 system.
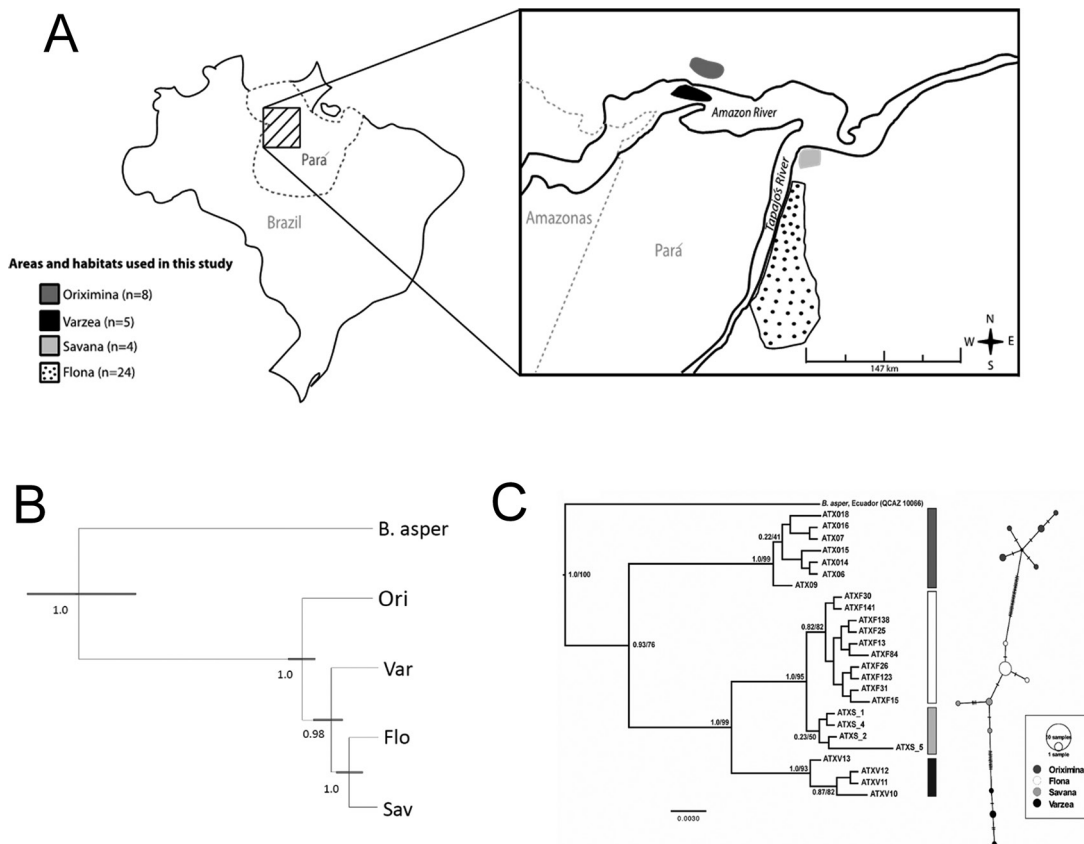
Figure 1. (A) Locations in eastern Amazonia from which *Bothrops atrox* was sampled. The number of individuals sampled is given next to habitat legends. (B) SNAPP tree summarizing phylogenetic relationships based on RAD data across four sampled populations of *B. atrox*. Values at nodes represent posterior support for clades after the first 20% of trees were removed as burnin. Values on the branches represent the minimum and maximum values observed in the set of 95% highest probability density values for branch length across the best supported BFD run. Population abbreviations are as follows (see Fig. 2): *B. asper* (outgroup); Ori (Oriximiná); Var (Várzea); Flo (FLONA); Sav (Savanna). (C) Bayesian phylogram and median joining haplotype network for unique mtDNA haplotypes from *B. atrox* samples. Bars of different shades indicate the geographical location of haplotypes from specific samples (see legend) and have the following order from top to bottom of the figure: Oriximiná, Várzea, Savanna and FLONA. For the phylogram, numbers above the branches are Bayesian posterior probabilities whereas numbers below are maximum likelihood bootstrap percentages. For the network, hatches represent the number of mutational steps between haplotypes.

BIOINFORMATIC METHODS

De novo locus assembly, identification of single nucleotide polymorphisms (SNPs) and genotyping were carried out on the raw fastq data using AftrRAD version 5.0 (Sovic, Fries & Gibbs, 2015). AftrRAD analyses were run across eight computation cores and carried out with default parameters. To reduce artefactual SNPs that accumulate toward the ends of reads due to assembly methods (see Sovic *et al.*, 2015), only SNPs that occurred in the first 31 positions along the read, after removal of barcodes and restriction sites, were retained for analyses (maxSNP-31). A minimum of five reads was required for a retained genotype call (MinReads-5). Loci with fewer than five reads in a sample were scored as missing data. Sovic *et al.* (2016)

provide a detailed rationale and justification for the methodological choices made in analysing the data set.

ASSESSING GENETIC STRUCTURE

We assessed genetic differentiation between samples collected from different sites in two ways. First, we estimated overall pairwise Fst values between each population using hierFST (Goudet, 2005). Second, to identify genetic clusters across all samples, we used the k-means clustering method implemented in adegenet 1.4-2 (Jombart & Ahmed, 2011). This program uses discriminant functions to maximize variation between clusters while minimizing variation within clusters and identifies the best supported clustering solutions based on Bayesian
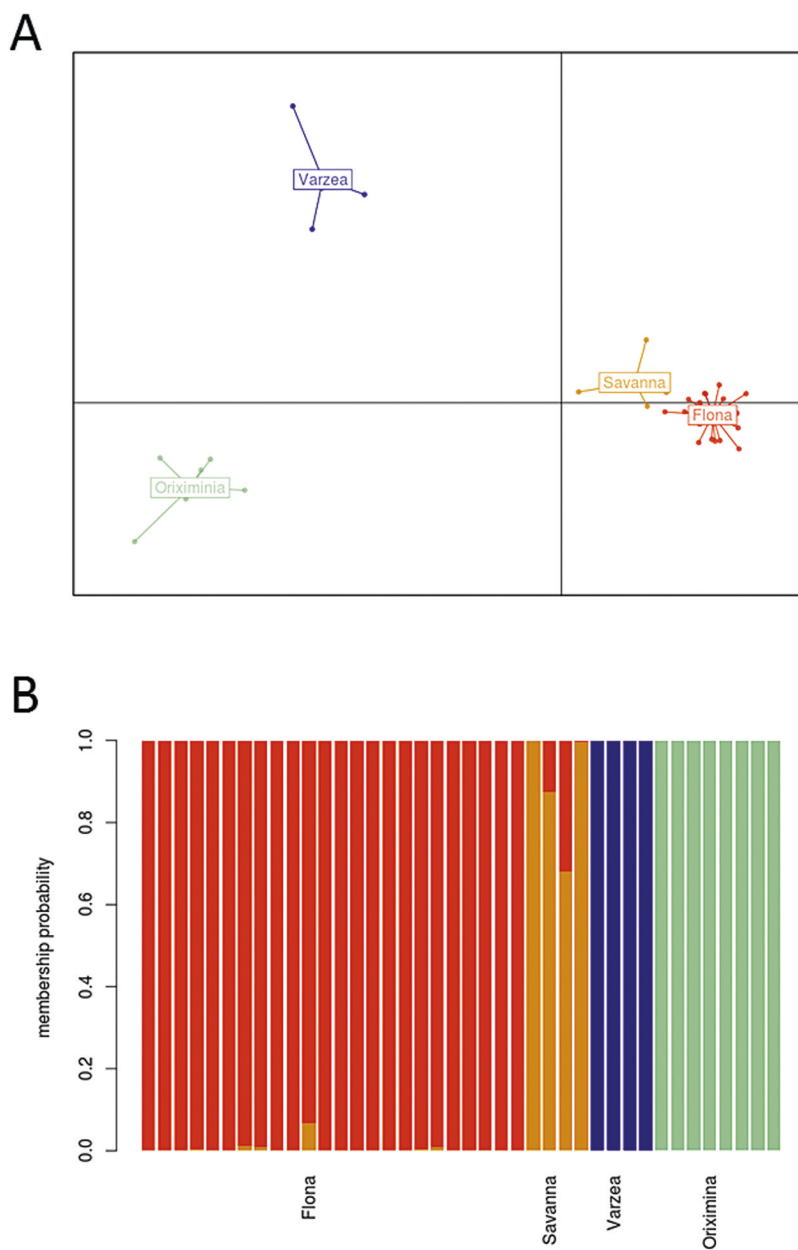
Figure 2. Results of DAPC analyses of *Bothrops atrox* genetic data under a *K* value of four. (A) Biplot of the results of a discriminant analysis of principal components summarizing genetic variation. Geographical locations of sample clusters are as in Figure 1. (B) Membership probabilities for individual samples from each of the four locations.

information criterion (BIC) scores from axes derived from a principal components analysis (PCA). We used this approach rather than the more commonly used program Structure (Pritchard, Stephens & Donnelly, 2000) because it does not rely on assumptions about Hardy Weinberg equilibrium that can lead to issues in defining clusters when sample sizes are small and uneven (Puechmaille, 2016). We evaluated *K* values ranging from 1 to 20 and then performed a discriminate function analysis of PCAs (DAPC) based on the clusters suggested in the previous step. Finally, as a complementary way of assessing how individuals from different locations were allocated to distinct genetic groups, we estimated individual sample assignment probabilities to defined clusters under a given *K* value using sample location as a prior in the analysis.

IDENTIFICATION OF DISTINCT LINEAGES

To test whether individuals from each population represent phylogenetically distinct lineages, we

conducted a species delimitation analysis as implemented in BFD* (Leaché *et al.*, 2014). To make the analysis computationally tractable, we randomly selected four individuals from each sample location for analysis and included a *B. asper* individual from the Amazon region of Ecuador as an outgroup. We then tested a set of nested models to evaluate whether these samples represented between one and four evolutionarily independent lineages (see Table 3). Models were ranked based on their marginal likelihood estimates, and Bayes factors for comparing relative likelihoods were calculated and interpreted following the guidelines of Kass & Raferty (1995). Finally, an SNAPP species tree (Bryant *et al.*, 2012) was generated using the lineages defined from the best-fit model to estimate phylogenetic relationships among lineages. BFD* and SNAPP runs were conducted using two prior settings for β: 10 000 (which corresponds to a θ value of 0.01) and 100 (θ = 0.0001).

HISTORICAL DEMOGRAPHIC MODELLING

Following Sovic *et al.* (2016), we used a model testing approach to evaluate the demographic processes that influenced the genetic distinctiveness of samples from each location. Specifically, we compared models that differed in the presence and direction of migration between populations and then estimated parameters from the best-fit model. For computational efficiency and model simplicity, we conducted three pairwise comparisons between samples in locations that were geographically proximate to each other: Oriximiná vs. Várzea; Várzea vs. Savanna; and Savanna vs. FLONA. For each comparison we evaluated three models: (1) an isolation-only model with no gene flow between populations in which the only parameters were divergence time and genetically effective population size (I model); (2) the same model but with the addition of migration (estimated separately in each direction) (IM model); and (3) a gene flow-only model that only included gene flow between proximate populations with no component of historical divergence (M model). The M model represents a stepping-stone model of population structure with no phylogenetic component of divergence and tests if isolation by distance (IBD) effects alone can account for the observed patterns of shared polymorphism between populations.

As described by Sovic *et al.* (2016), we used the coalescent-based modelling package FastSimCoal252 (fsc252; Excoffier *et al.*, 2013) to evaluate these models and generate parameter estimates for divergence times, lineage-specific effective population sizes and migration rates. For each model, we performed 100 independent runs of fsc252 (30 ECM cycles and 1×10^5 simulations per run), and used the run with the highest likelihood for each model to perform model

selection with Akaike's information criterion (AIC). We used a prior range of 100 to 5×10^5 for both divergence times and effective population size estimates and a prior of 1×10^{-4} to 10 gene copies for migration rates. Confidence intervals for parameters were estimated based on an analysis of 50 bootstrapped resampled datasets for each population. Each resampled dataset was constructed by sampling with replacement from the original dataset and leaving out 10% of the original number of sites each time. We report the confidence intervals as the minimum and maximum values observed for each parameter from the set of runs for each data set. We are not aware of any direct estimates for genome-wide mutation rates in snakes or other squamates and so used a rate of 2.5×10^{-8}, as estimated for humans (Nachman & Crowell, 2000). This estimated mutation rate may be high given that endotherms are suggested to have higher mutation rates than ectotherms such as snakes (Gillooly *et al.*, 2005) and so parameter estimates that incorporate mutation rate need to be considered as tentative. Finally, we followed the approach of Lande, Engen & Saether (2003) to convert estimates of time from generations to years. Based on the most detailed information available for *Bothrops* species [age of maturity of 3 years (Sasa, Wasko & Lamar, 2009) and an annual adult survival rate of 0.67 (Guimarães *et al.*, 2014)] we estimated a generation time of 5.0 years using the equation $T = \alpha + [s/(1 - s)]$, where T is the generation time, α is the age at first maturity and s is the annual adult survival rate.

MTDNA ANALYSES

For comparison with the nuclear DNA results from RADseq data, we also sequenced a portion of the mtDNA cytochrome b (cyt-b) gene from the following numbers of individuals from each population: Oriximiná ($N = 8$), Várzea ($N = 4$), Savanna ($N = 4$) and FLONA ($N = 10$) and an outgroup (*B. asper*). Specifically, we amplified a 674-bp region of the cyt-b gene using the Gludg and AtrCB3 primer pairs (Parkinson, Campbell & Chippindale, 2002). Complementary sequences were assembled and edited with CodonCode Aligner 4 and we used MUSCLE (Edgar, 2004) in Geneious 7.0 to align the sequences.

We used both maximum-likelihood (ML) and Bayesian inference (BI) analyses to infer phylogenetic relationships among unique haplotypes. The ML analysis was conducted in RAxML 8.0 (Stamatakis, 2014) using a GTRGAMMA substitution model and performing 1000 bootstrap replicates in a rapid bootstrap analysis. For the BI analyses, we followed Castoe & Parkinson (2006) for the partition scheme of the cyt-b fragment and applied the best-fit substitution model as identified using the BIC

implemented in jModeltest 2.1.7 (Darriba *et al.*, 2012). Three independent runs, each with four Markov chains (one cold and three heated), were run for 20 million generations and sampled every 1000 generations in MrBayes v.3.2.3 (Ronquist *et al.*, 2012). Stationarity and effective sample sizes (ESS) for all parameters were assessed with Tracer v.1.5 (Drummond & Rambaut, 2007). Of the 20 000 trees resulting per run, the first 25% were discarded as burn-in. The remainder were used to calculate posterior probabilities (PP) for each bipartition in a 50% majority-rule consensus tree. We also used the program PopART (Leigh & Bryant, 2015) to construct a median joining network for the mtDNA dataset.

Finally, to generate additional estimates of population divergence times we calculated uncorrected pairwise distances between mtDNA haplotypes in each population (corrected for within-population variation) and applied a divergence rate of 2.4% per million years estimated for cyt-b variation in *B. jararraca* (Grazziotin *et al.*, 2006).

FUNCTIONAL DIFFERENCES IN VENOM

Venom variation in snakes is widely argued to represent an adaptation that has evolved to facilitate the capture and digestion of prey (Casewell *et al.*, 2013), and *in vitro* analyses of the enzymatic properties of whole venom can be used to directly assess the functional properties of venom as a molecular phenotype. Sousa *et al* (2017) report the results of functional analyses of pooled venom samples from individuals from three of the four sample locations for which we completed genetic analyses [FLONA (designated as the 'Forest' site in Sousa *et al.*, 2017), Várzea ('Floodplain') and Oriximiná ('Pasture')]. The FLONA and Oriximiná samples were from the same locations and included venom samples from some of the same individuals analysed here whereas the Várzea samples came from two locations within this habitat including individuals from the site were genetic samples were obtained [compare Fig. 1A with fig. 1 from Sousa *et al.* (2017]). For measures of phenotypic differentiation between populations that occupy ecologically distinct habitats [Floodplain (Várzea) vs. Upland Forest (FLONA and Oriximiná)], we use the results reported in figure 7 of Sousa *et al.* (2017) for six measures of major components of venom activity (haemorrhagic activity, metalloproteinase catalytic activity, coagulant activity, serine proteinase catalytic activity, myotoxic activity and PLA2 catalytic activity) and estimates of venom lethality (LD$_{50}$ estimates based on laboratory mice) and interpret these results in light of the genetic analyses presented here.

RESULTS

SEQUENCING AND SNP IDENTIFICATION

A mean of 1.47×10^6 sequence reads were assigned to each of the 41 samples (range 228 038–3 781 203) after quality filtering in AftrRAD. The mean read depth per locus was 140.3 reads while the median read depth was 62. A total of 3970 non-paralogous loci were identified. Of these, 1307 were scored in all samples with 681 loci being monomorphic and 626 being polymorphic. These polymorphic loci were used in subsequent analyses although, where appropriate, monomorphic loci were taken into account for parameter estimates based on overall levels of polymorphism.

GENETIC DIFFERENTIATION

Pairwise Fst values between all locations averaged 0.113 and were highest between samples from populations on either bank of the Amazon (FLONA and Oriximiná: 0.168) and lowest between the geographically proximate FLONA and Savanna sites (0.0396) (Table 1). DAPC analyses match the patterns of differentiation observed in Fst values. A comparison of BIC values for different *K* values in adegenet showed that the best supported solution was a *K* value of either 3 or 4 (Supporting Information, Fig. S1). Discriminant function analyses of principal components summarizing genetic variation in the data set show that for *K* = 3 there are three well-separated clusters consisting of individuals from Oriximiná, Várzea, and FLONA and Savanna combined (results not shown). The biplot for *K* = 4 shows a similar pattern with the difference being that the FLONA and Savanna samples now form distinct but related clusters (Fig. 2A). Finally, membership probabilities for individual samples from each of the four locations show a pattern that mirrors that of the discriminate function analysis biplot – samples from Oriximiná, Várzea and FLONA are clearly distinct with either no or little admixture present, whereas two of four Savanna samples show admixture with FLONA-specific genetic variation (Fig. 2B). Our broad interpretation is that there is evidence for four distinct genetic clusters but that two of these (FLONA and Savanna) are genetically similar to each other and could be interpreted as representing a single diverse cluster.

LINEAGE DELIMITATION ANALYSES

The most strongly supported model in the BFD* analyses using a β prior of 1000 is one in which each sampled population represents an evolutionarily distinct lineage based on Bayes factors calculated from model-specific ML estimates (Table 2). A four-lineage model is favoured over a three-lineage model

(which combines FLONA and Savanna populations into a single lineage) by a Bayes factor of 18.48, which represents 'decisive' support for the preferred model (Kass & Rafferty, 1995). When the analysis is repeated with a different β prior value of 100, the four-lineage model is still preferred but the Bayes factor is lower (4.83) relative to the second ranked model which remains the three-lineage model described above (results not shown).

PHYLOGENETIC ANALYSIS OF NUCLEAR AND MITOCHONDRIAL DNA

The SNAPP 'species' tree generated for RAD data from a subsample of individuals from each population shows a pattern consistent with a biogeographical dispersal event across the Amazon (Fig. 1B) from the north shore (represented by Oriximiná) to the south shore (represented by FLONA) with sequential differentiation in each of the intervening populations (Fig. 1B). All node support values are > 98%. An SNAPP tree generated from the BFD* results based on a β prior value of 100 shows the same topology but node support for the clade defining the relationship between Oriximiná and the other three populations was reduced (0.73).

Gene-tree-based analyses of mtDNA sequences from individuals in each population show a similar topology to the SNAPP tree (Fig. 1C). There was variation within mtDNA haplotypes from individuals collected at specific sites, suggesting sufficient time for intraspecific variation to accumulate over time. However, support values for clades made up by haplotypes from individuals from each location were substantial (> 0.85 branch support values and > 75% bootstrap support). As with the SNAPP analyses, the most weakly supported relationship was between Oriximiná and the other populations. Strikingly, the phylogenetic pattern between clades made up of haplotypes from each location matched that of the SNAPP tree, namely one of dispersal across the Amazon from the northern population to the southern population with sequential differentiation of populations found in distinct habitats between each bank.

Finally, point estimates of divergence times based on mtDNA sequence divergence between samples from each geographically proximate clade were ≤ 1.6 Mya (Oriximiná vs. Várzea = 1.61 Mya; Várzea vs. Savanna: 0.89 Mya ; Savanna vs. FLONA: 0.16 Mya), suggesting that these lineages originated in the mid- to late Pleistocene with a rough periodicity of ~700 000 years between the origin of each lineage.

HISTORICAL DEMOGRAPHY

Statistical comparisons based on AIC show that for all pairwise comparisons an Isolation with Migration (IM) model significantly better fits patterns of shared RADseq variation between populations than do Isolation only (I) or Migration only (M) models (Table 3). Parameter estimates from IM models indicate that all populations diverged recently, have relatively large effective population sizes, and experience low but roughly symmetrical levels of gene flow between geographically proximate populations and have diverged within the last 0.5 My (Fig. 3; Supporting Information, Table S1). Long-term effective population sizes (Ne) are large, ranging from a low of ~ 35 000 individuals in the Savanna population to a high of ~ 76 000 individuals in the Oriximiná population (Fig. 3). Two pairs of geographically close populations (Oriximiná and Várzea, and Várzea and Savanna) share migrants at roughly symmetric rates of ≤ 1 migrants per generation while the other population pair (FLONA and Savanna) show a higher rate of gene exchange (~ 2.5 migrants per generation) consistent with their more limited differentiation based on the genetic differentiation analyses (Fig. 3). Finally, point estimates of divergence times between pairs of populations range from 14 792 (FLONA–Savanna) to 95 518 (Várzea–Savanna) generations ago with the estimate for the divergence

Table 2. Pairwise Fst values between *Bothrops atrox* populations shown in Figure 2; all values are significantly different from zero ($P < 0.01$)

	Oriximiná	Várzea	Savanna	FLONA
Oriximiná		0.0791	0.124	0.168
Várzea			0.111	0.156
Savanna				0.0396
FLONA				

Table 3. Statistical ranking of models proposing different combinations of populations of *Bothrops atrox* as evolutionarily independent lineages based on an analysis using BFD* (Leaché *et al.*, 2014)

Model (proposed number of distinct lineages)	MLE	Bayes factor
1 – F, S, V, O (4)	−26569.76076	0.00
2 – (F+S), V, O (3)	−26579.00291	18.48
5 – (F+S), (V+O) (2)	−26628.39763	117.27
3 – (F+S+V), O (2)	−26698.75389	257.99
4 – (F+S+V+O) (1)	−26835.5118	531.50

Models evaluated between one and four potential independent lineages consisting of combinations of individuals from the FLONA (F), Savanna (S), Várzea (V) and Oriximiná (O) populations. Models are shown in order of their maximum likelihood estimate (MLE) and associated Bayes factor, which is based on a comparison of the MLE of the top ranked model with other models with lower MLE values.

Number of Migrants per Generation

Estimated Effective Population Sizes
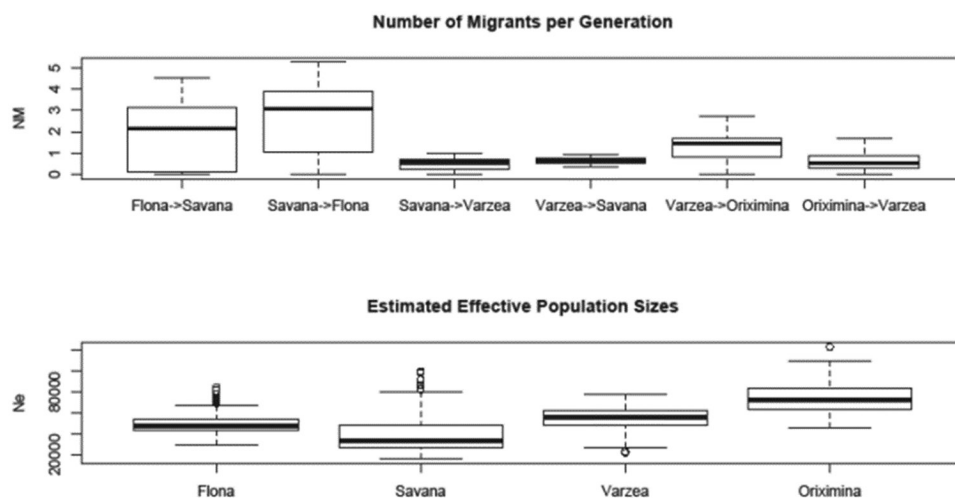
Figure 3. Bi-directional estimates of gene flow (NM, migrants per generation) and genetically effective population sizes (Ne, numbers of diploid individuals) between geographically proximate pairs of *Bothrops atrox* populations. Thick lines represent point estimates with bar plots indicating the confidence intervals for each estimate (see Table S1 for specific values). Populations are as indicated in Figure 1A.

between Oriximiná and Várzea being intermediate between these values (41 568 generations) (Table S1). This result is not consistent with an expectation from the branching pattern shown in both the SNAPP and the mtDNA tree, which predicts that divergence estimates should vary in magnitude in the following way: Ori–Var > Var–Sav > Sav–Flo. We have no explanation for this discrepancy. Applying our estimate of a generation time of 5 years gives divergence time estimates of < 500 000 years ago, consistent with the result from mtDNA of a recent Pleistocene origin for these lineages. Overall the picture is of large, well-established populations that have diverged from each over recent evolutionary timescales but which remain genetically connected by low levels of gene flow.
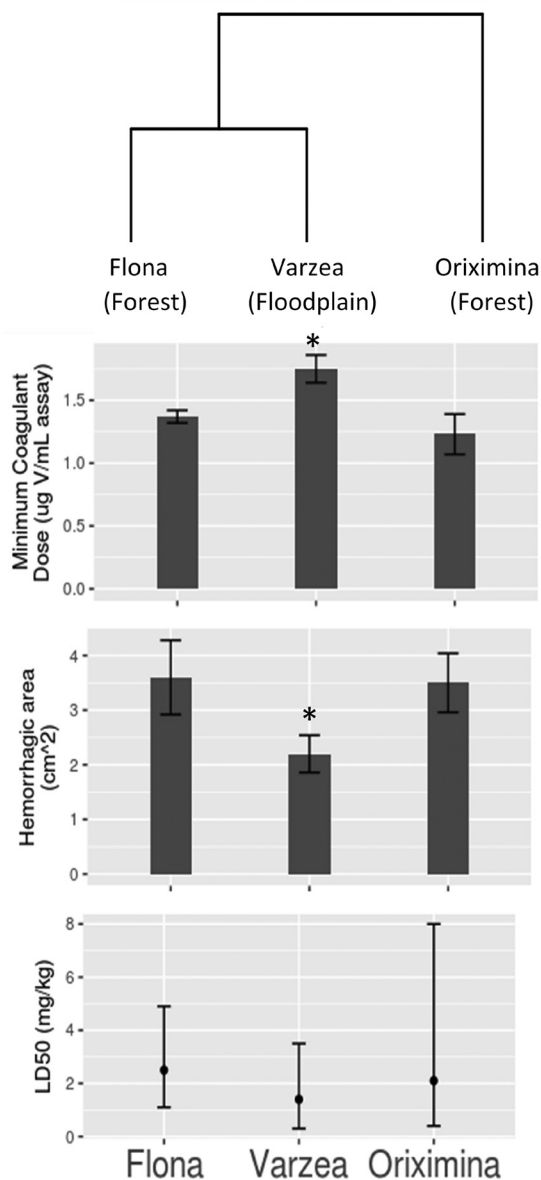
FUNCTIONAL DIFFERENCES IN VENOM COMPOSITION

Four of six measures of venom functional activity showed significant differences between floodplain (Várzea) venom and venom from the two upland forest (Terra Firme) sites but no difference between the two forest sites (haemorrhagic activity, metalloproteinase catalytic activity, coagulant activity, serine proteinase catalytic activity), while one was significantly different between Oriximiná and the other two populations (PLA2 catalytic activity) and one (myotoxic activity) showed no difference between populations (see Fig. 4 for examples of the first pattern). Mean LD_{50} values were similar between the two forest sites but lower for the Várzea population, but not significantly so probably due to the small sample sizes of mice used in the lethality tests (Fig. 4). Thus, the most common pattern of differences in activity are habitat-specific

differences between venom from floodplain and the two upland forest sites.

DISCUSSION

MECHANISMS OF DIFFERENTIATION

Phylogeographical analyses of variation in both nuclear DNA (RAD) and mtDNA demonstrate that these genetically differentiated populations that span the Amazon over a relatively limited geographical range (~200 km) also form evolutionarily distinct lineages. The pattern of phylogeographical relationships between populations clearly rejects the classical riverine barrier hypothesis, which proposes that tropical rivers such as the Amazon act as evolutionarily long-term barriers that isolate populations on either shore for long periods, leading to the development of reciprocally monophyletic clades on either shore. Instead, our results suggest that differentiation among *B. atrox* populations followed a dispersal model in which an ancestral population of *B. atrox* from terra firma forest on the north side of the Amazon sequentially colonized regions spanning the river before ending up on the south bank (see Patton *et al.*, 2000: fig. 164). This pattern clearly demonstrates that the river per se did not represent an insurmountable barrier to movement but rather over recent evolutionary timescales snakes were able to disperse across the river with lineage differentiation occurring as this process unfolded. Phylogenetic trees consistent with dispersal have rarely been observed in phylogeographical studies of Amazonian vertebrates in relation to riverine barriers as compared to more commonly observed patterns of reciprocal monophyly

Figure 4. Examples of assays of functional variation in venom that show significant differences in haemorrhagic and coagulant activity, and measures of lethality to laboratory mice (LD_{50}) between floodplain (Várzea) and upland forest (FLONA and Oriximiná) sites in relation to the evolutionary history of these populations. Panels from top to bottom are: (1) the inferred evolutionary relationships of the FLONA, Várzea and Oriximiná populations which occupy upland forest, floodplain and upland forest habitats, measures for (2) coagulation and (3) haemorrhagic activity, and (4) LD_{50} values. Asterisks denote significantly different values. The high-–low–high or low–high–ow pattern of activity is consistent with an evolutionary history of selection matching venom function to forest, then floodplain and then back to forest habitats as these locations are colonized by snakes dispersing across the Amazon. Data on activity and LD_{50} values are from figure 7 of Sousa *et al.* (2017).

with respect to a river (Rocha *et al.*, 2015) or the lack of any phylogeographical pattern (Patton *et al.*, 2000). This could be because previous studies on dispersal-limited taxa such as reptiles and amphibians have focused on regions of Amazonia where river barriers were relatively small and river-associated habitat diversity was low.

Our results also suggest that while the Amazon is not an absolute barrier to gene flow it nonetheless significantly impedes migration and this mechanism plays a role in allopatric lineage diversification in these snakes. First, the fact that we observed significant phylogeographical structure that has originated over thousands of years among populations spanning the river that are all within 200 km of each other argues for the river per se limiting gene flow between populations. Second, migration estimates between the two populations that share the south bank of the Amazon (FLONA and Savanna) also have levels of migration that are more than twice as high as the other pairwise comparisons, all of which include populations separated by the river albeit to varying distances. Although limited, this comparison suggests that a large river represents an environmental feature that restricts dispersal in these snakes, which are primarily terrestrial (Campbell & Lamar, 2004). Overall, our study suggests that in terms of the predictions in Table 1, although we reject a strict version of the barrier hypothesis our results support a modified version in terms of the Amazon impeding gene flow in these snakes in a way that contributes to lineage formation.

Multiple pieces of evidence also support a role for selection related to ecological differences between populations playing a role in lineage diversification (the ecological gradients mechanism in Table 1). Specifically, the phylogeographical pattern is consistent with a model of parapatric differentiation in which sister taxa occupy ecologically distinct habitats or the opposite ends of ecological gradients which impose different selection pressures on snakes in each population, in turn influencing lineage diversification (Patton & Silva, 1998). This interpretation is complicated by a weakness of our study design in that habitat differences and geographical distance are confounded due to the small number of populations sampled. As a result, IBD could potentially account for the divergence of these populations rather than habitat differences per se. However, our historical modelling rejects a pure IBD mechanism as solely responsible for the genetic divergence between geographically proximate populations, although it contributes to this process.

Historical demographic analysis shows that this divergence has occurred despite populations being connected via low levels of ongoing gene flow (~ 1 migrant per generation or higher in terms of the FLONA–Savanna populations). This means that any

Table 4. Results of FastSimCoal model comparisons for isolation only, migration only, and isolation and migration models for pairwise comparisons of geographically proximate populations shown in Figure 1A

Pair/model	Number of parameters	Ln Likelihood	AIC	Akaike weight
Flona vs. Savanna				
Migration only	4	−2680.86	12 353.80	0.01
Isolation only	4	−2681.51	12 356.79	0.00
Isolation and migration	6	−2680.86	12 344.25	0.99
Várzea vs. Savanna				
Migration only	4	−2992.76	13 790.19	0.01
Isolation only	4	−2992.45	13 788.76	0.01
Isolation and migration	6	−2989.70	13 780.05	0.98
Oriximiná vs. Várzea				
Migration only	4	−2914.79	13 431.11	0.01
Isolation only	4	−2915.28	13 433.37	0.01
Isolation and migration	6	−2912.03	13 422.37	0.98

adaptive or neutral forces influencing divergence have been sufficiently strong to overcome sustained genetic 'leakage' between geographically proximate lineages or that selection against maladapted migrants may limit genetically effective migration (Lenormand, 2002). The large, genetically effective sizes of individual lineages argue that the impact of genetic drift as a driver of diversification due to, for example, founder-based dispersal effects is minimal and that natural selection can potentially act efficiently in these populations to produce local adaptation (Petit & Barbadilla, 2009).

Models of differentiation via selection in relation to ecological gradients emphasize two phenomena: the maintenance of differentiation in the face of gene flow (see above), and evidence for phenotypic differentiation that matches ecological differentiation in habitats (Smith *et al.*, 2005). As shown in Figure 4, data from Sousa *et al.* (2017) demonstrate differences in venom function between snakes from the two upland forest sites (Oriximiná and FLONA) and the floodplain site (Várzea) and provide evidence for adaptive divergence between populations in multiple and distinct features of the venom phenotype that correspond to habitat differences. Particularly striking is that the phylogenetically distinct Oriximiná and FLONA snakes, which occupy similar habitats, are statistically indistinguishable in venom function whereas the phylogenetically intermediate yet ecologically distinct Várzea snakes show significant divergence in the functional properties of their venom. Mapping these differences in venom phenotype onto the inferred history of these lineages suggests an evolutionary scenario of rapid evolution of a distinct Várzea venom phenotype after colonization of the novel floodplain habit by ancestors with an 'upland' venom phenotype followed by re-evolution of the 'upland forest' venom phenotype when the habitat occupied by the FLONA snakes was subsequently colonized (Fig. 4).

This close matching of venom phenotype to habitat at least for the upland forest–floodplain snakes suggests a role for selection and adaptive diversification in the formation of this phylogenetically distinct lineage, consistent with an ecological gradients mechanism. As discussed above, venom is a molecular adaptation that allows venomous snakes to effectively kill and digest prey (Casewell *et al.*, 2013), and it shows geographical variation in many species (Chippaux, Williams & White, 1991), including *B. atrox*, albeit on a larger geographical scale than studied here (Calvete *et al.*, 2011). The habitats in which the *B. atrox* populations were located probably have different prey communities (H. Chalkidis, unpublished data) which could lead to divergent selection that could result in fine-scale adaptive differentiation in venom as observed in other venomous snakes (Gibbs & Chiucchi, 2011). We are not arguing that ecologically based selection pressures that impact venom are necessarily the only or even the major form of adaptive differentiation along this ecological gradient. Rather, it represents the only form of adaptive phenotypic variation that we were able to measure in these snakes in the present study. Of great interest would be to assess if morphological variation associated with habitat use in other populations of *B. atrox* (Wüster *et al.*, 1997) also differs among the populations studied here.

RIVERS AS DRIVERS OF TROPICAL DIVERSIFICATION

Taken as a whole, our results suggest new perspectives on how large tropical rivers can act as a driver of diversification. First, in addition to acting as barriers preventing or limiting gene flow, large rivers may also contribute to diversification by providing ecologically diverse habitats, which result in the origin of lineages through adaptation. It also seems likely that the origin of phylogenetically and phenotypically distinct

lineages can be due to the combined influence of different mechanisms. For example, in our study the fact that the Amazon may impede (but not prevent) gene flow between populations in different habitats may allow selection to generate habitat-specific adaptive differences by reducing migration load within each population (Lenormand, 2002). Fine-scale population analyses along ecological gradients in the Amazon in organisms with limited dispersal abilities may reveal other examples consistent with this mode of differentiation in this region of the tropics. Second, our results reinforce previous work on birds (Harvey *et al.*, 2017) showing that within the Amazon, the ecological differences between upland forest (Terra Firme) and floodplain (Várzea) habitats may represent a fundamental environmental gradient that has had repeated and significant impacts on the evolutionary histories of multiple groups of organisms that occupy the two habitats. Incorporating this perspective particularly in regions where floodplains are a significant component of habitats found within the river may aid in developing a more complete understanding of the mechanisms driving diversification in this region. Finally, our analyses also suggest that lineage diversification in this region may have occurred during the mid- to late Pleistocene or more recently than the pre-Pleistocene time frame suggested by earlier work for the diversification by most vertebrates in Amazonia (Moritz *et al.*, 2000).

IMPLICATIONS OF LINEAGE IDENTIFICATION

Other genetic studies of *B. atrox* in the Amazon are limited but we note that Wüster *et al.* (1997) reported that samples from *B. atrox* collected on the north (Bal sample) and south (Ita) of the Amazon River in the Western Amazon region had different mtDNA haplotypes. The link with this result and our study is unclear because of the limited spatial scale from which our samples were collected. Specifically, it is not certain whether our results from populations within 200 km of each other represent a range-wide picture of phylogeographical variation in *B. atrox* or reflect a local situation in a region of exceptionally high habitat diversity. Replicated sampling from *B. atrox* populations in different habitats at multiple transects spanning the Amazon and from populations to the north and south of the river would clarify how general our results are in terms of the directionality of the pattern of colonization suggested by the phylogeographical patterns in nuclear and mtDNA. Overall, our results support Wüster *et al.*'s (1997) conclusion that *B. atrox* represents a highly variable species with recent lineage differentiation associated with habitat diversity but we add to this result by

providing evidence for the presence of distinct lineages over a much finer geographical scale than previously recognized (see also Grazziotin *et al.*, 2006). Our findings complement recent work by Salazar-Valenzuela (2016) on a related widespread congeneric snake (*B. asper*) distributed from Central to northern South America, who used similar approaches to find evidence for multiple recently evolved and previously unrecognized lineages, including those associated with altitudinal habitat gradients.

BIOMEDICAL IMPLICATIONS

A practical implication of this study is that it stresses the importance of using individuals collected across diverse habitats when assembling venom pools for generating antivenoms for *B. atrox*. The incidence of snake bites is high in this region of Brazil (Pardal *et al.*, 2004). As such, developing an effective antivenom is a health priority in the region and the possible presence of lineage-specific differences in venom composition needs to be taken into account when developing appropriate treatments for this significant health issue.

ACKNOWLEDGEMENTS

REFERENCES

Alexio A, Rossetti DF. 2007. Avian gene trees, landscape evolution, and geology: towards a modern synthesis of Amazonian historical biogeography? *Journal of Orinthology* **148:** S443–S453.

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews. Genetics* **17:** 81–92.

Antonelli A, Quijada-Mascareñas A, Crawford AJ, Bates JM, Velazco PM, Wüster W. 2010. Molecular studies and paleogeography of Amazonian tetrapods and their relation to geological and climatic models. In: Hoorn C, Wesselingh FP, eds. *Amazonia: landscape and species evolution. A look into the past*. Chichester: Wiley-Blackwell, 386–404.

Ayres JM, Clutton-Brock TH. 1992. River boundaries and species range size in Amazonian primates. *The American Naturalist* **140:** 531–537.

Bates JM, Haffer J, Grismer E. 2004. Avian mitochondrial DNA sequence divergence across a headwater stream of the Rio Tapajos, a major Amazonian river. *Journal of Ornithology* **145:** 199–205

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

Campbell JA, Lamar WW. 2004. *The venomous reptiles of the Western Hemisphere*. Ithaca: Cornell University Press.

Calvete JJ, Sanz L, Pérez A, Borges A, Vargas AM, Lomonte B, Angulo Y, Gutiérrez JM, Chalkidis HM, Mourão RH, Furtado MF, Moura-Da-Silva AM. 2011. Snake population venomics and antivenomics of *Bothrops atrox*: paedomorphism along its transamazonian dispersal and implications of geographic venom variability on snakebite management. *Journal of Proteomics* **74:** 510–527.

Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG. 2013. Complex cocktails: the evolutionary novelty of venoms. *Trends in Ecology & Evolution* **28:** 219–229.

Castoe TA, Parkinson CL. 2006. Bayesian mixed models and the phylogeny of pitvipers (Viperidae: Serpentes). *Molecular Phylogenetics and Evolution* **39:** 91–110.

Chippaux JP, Williams V, White J. 1991. Snake venom variability: methods of study, results and interpretation. *Toxicon* **29:** 1279–1303.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9:** 772.

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

Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5:** 113.

Espiritio-Santo FDB, Shimabukuro YE, Aragao LEOC, Machado ELM. 2005. Análise da composição florística e fitossociológica da floresta nacional do Tapajós com o apoio geográfico de imagens de satélites. *Acta Amazonica* **3:** 155–173.

Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M. 2013. Robust demographic inference from genomic and SNP data. *PLoS Genetics* **9:** e1003905.

Gascon C, Malcolm JR, Patton JL, da Silva MN, Bogart JP, Lougheed SC, Peres CA, Neckel S, Boag PT. 2000. Riverine barriers and the geographic distributions of Amazonian species. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 13672–13677.

Gibbs HL, Chiucchi JE. 2011. Deconstructing a complex molecular phenotype: population-level variation in individual venom proteins in Eastern Massasauga Rattlesnakes (*Sistrurus c. catenatus*). *Journal of Molecular Evolution* **72:** 383–397.

Gibbs L, Sovic M, Amazonas D, Chalkidis H, Salazar-Valenzuela D, Moura-da-Silva A. 2017. Data from: Recent lineage diversification in a venomous snake through dispersal across the Amazon River. *Dryad Digital Repository* https://doi.org/10.5061/dryad.36sv8.

Gillooly JF, Allen AP, West GB, Brown JH. 2005. The rate of DNA evolution: Effects of body size and temperature on the molecular clock. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 140–145.

Goudet J. 2005. HIERFSTAT, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Resources* **5:** 184–186

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. *Molecular Ecology* **15:** 3969–3982.

Guimarães M, Munguía-Steyer R, Doherty PF Jr, Martins M, Sawaya RJ. 2014. Population dynamics of the critically endangered golden lancehead pitviper, *Bothrops insularis*: stability or decline? *PloS One* **9:** e95203.

Harvey MG, Brumfield RT. 2015. Genomic variation in a widespread Neotropical bird (*Xenops minutus*) reveals divergence, population expansion, and gene flow. *Molecular Phylogenetics and Evolution* **83:** 305–316.

Harvey MG, Aleixo A, Ribas CC, Brumfield RT. 2017. Habitat association predicts genetic diversity and population divergence in Amazonian birds. *The American Naturalist* **190:** 631–648.

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

Kass RE, Raferty AE. 1995. Bayes factors and model uncertainty. *Journal of the American Statistical Association* **90:** 773–795.

Lande R, Engen S, Saether BE. 2003. *Stochastic population dynamics in ecology and conservation*. Oxford: Oxford University Press.

Leaché AD, Fujita MK, Minin VN, Bouckaert RR. 2014. Species delimitation using genome-wide SNP data. *Systematic Biology* **63:** 534–542.

Leigh JW, Bryant DM. 2015. Popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6:** 1110–1116.

Leite RN, Rogers DS. 2013. Revisiting Amazonian phylogeography: insights into diversification hypotheses and novel perspectives. *Organisms, Diversity and Evolution* **13:** 639–664.

Lenormand T. 2002. Gene flow and the limit to natural selection. *Trends in Ecology & Evolution* **17:** 183–189.

Magnusson WE, Lima AP, Albernaz ALKM, Sanaiotti TM, Guillaumet JL. 2008. Composição florística e cobertura vegetal das savanas na região de Alter do Chão, Santarém – PA. *Revista Brasileira de Botânica*. **31:** 165–177.

Marques S, Trefaut M, Cohn-Haft M. 2013. Are Amazonian rivers biogeographic barriers for lizards? A study on the geographic variation of the Spectacled Lizard *Leposoma osvaldoi* Avila-Pires (Squamata, Gymnophthalmidae). *Journal of Herpetology* **47:** 511–519.

Mittermeier RA, Gil PR, Mittermeier CG. 1997. *Megadiversity - Earth's biologically wealthiest nations*. Washington: Conservation International.

Moritz C, Patton JL, Schneider CJ, Smith TB. 2000. Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology, Evolution, and Systematics* **31:** 533–563.

Nachman MW, Crowell SL. 2000. Estimate of the mutation rate per nucleotide in humans. *Genetics* **156:** 297–304.

Orr MR, Smith TB. 1998. Ecology and speciation. *Trends in Ecology & Evolution* **13:** 502–506.

Pardal PP, Souza SM, Monteiro MR, Fan HW, Cardoso JL, França FO, Tomy SC, Sano-Martins IS, de Sousa-e-Silva MC, Colombini M, Kodera NF, Moura-da-Silva AM, Cardoso DF, Velarde DT, Kamiguti AS, Theakston RD, Warrell DA. 2004. Clinical trial of two antivenoms for the treatment of *Bothrops* and *Lachesis* bites in the north eastern Amazon region of Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **98:** 28–42.

Parkinson CL, Campbell JA, Chippindale PT. 2002. Multigene phylogenetic analysis of pitvipers, with comments on their biogeography. In: Schuett GW, Höggren M, Douglas ME, Greene HW, eds. *Biology of the vipers*. Salt Lake City: Eagle Mountain Publishing, 93–110.

Patton JF, Silva MNF. 1998. Rivers, refuges, and ridges: the geography of speciation of Amazonian mammals. In: Howard D, Berlocher S, eds. *Endless forms: modes and mechanisms of speciation*. Oxford: Oxford University Press, 202–213.

Patton JL, Silva MNF, Malcolm JR. 2000. Mammals of the Rio Jurua and the evolutionary and ecological diversification of Amazonia. *Bulletin of the American Museum of Natural History* **244:** 1–306.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7:** e37135.

Petit N, Barbadilla A. 2009. Selection efficiency and effective population size in *Drosophila* species. *Journal of Evolutionary Biology* **22:** 515–526.

Pires JM, Prance GT. 1985. The vegetation types of the Brazilian Amazon. In: Prance GT, Lovejoy TE, eds. *Key environments: Amazonia*. New York: Pergamon Press, 109–145.

Prates I, Xue AT, Brown JL, Alvarado-Serrano DF, Rodrigues MT, Hickerson MJ, Carnaval AC. 2016. Inferring responses to climate dynamics from historical demography in neotropical forest lizards. *Proceedings of the National Academy of Sciences of the United States of America* **113:** 7978–7985.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Puechmaille SJ. 2016. The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular Ecology Resources* **16:** 608–627.

Rocha RG, Ferreira E, Loss AC, Heller R, Fonseca C, Costa LP. 2015. The Araguaia river as an important biogeographical divide for didelphid marsupials in central Brazil. *Journal of Heredity* **106:** 593–607.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61:** 539–542.

Salazar-Valenzuela D. 2016. *Diversification in the Neotropics: Insights from demographic and phylogenetic patterns of Lancehead Pitvipers (Bothrops spp.)*. PhD thesis, Ohio State University.

Sasa M, Wasko DK, Lamar WW. 2009. Natural history of the terciopelo *Bothrops asper* (Serpentes: Viperidae) in Costa Rica. *Toxicon* **54:** 904–922.

Slatkin M. 1994. Gene flow and population structure. In: Real L, ed. *Ecological genetics*. Princeton: Princeton University Press, 3–17.

Smith TB, Calsbeek R, Wayne RK, Holder KH, Pires D, Bardeleben C. 2005. Testing alternative mechanisms of evolutionary divergence in an African rain forest passerine bird. *Journal of Evolutionary Biology* **18:** 257–268.

Sousa LF, Portes-Junior JA, Nicolau CA, Bernardoni JL, Nishiyama MY Jr, Amazonas DR, Freitas-de-Sousa LA, Mourão RH, Chalkidis HM, Valente RH, Moura-da-Silva AM. 2017. Functional proteomic analyses of *Bothrops atrox* venom reveals phenotypes associated with habitat variation in the Amazon. *Journal of Proteomics* **159:** 32–46.

Sovic MG, Fries AC, Gibbs HL. 2015. AftrRAD: a pipeline for accurate and efficient de novo assembly of RADseq data. *Molecular Ecology Resources* **15:** 1163–1171.

Sovic MG, Fries AC, Gibbs HL. 2016. Origin of a cryptic lineage in a threatened reptile through isolation and historical hybridization. *Heredity* **117:** 358–366.

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

Wallace AR. 1852. On the monkeys of the Amazon. *Proceedings of the Zoological Society of London* **20:** 107–110.

Wasko DK, Sasa M. 2009. Activity patterns of a Neotropical ambush predator: Spatial ecology of the Fer-de-lance (*Bothrops asper*, Serpentes: Viperidae) in Costa Rica. *Biotropica* **41:** 241–249.

Weir JT, Faccio MS, Pulido-Santacruz P, Barrera-Guzmán AO, Aleixo A. 2015. Hybridization in headwater regions, and the role of rivers as drivers of speciation in Amazonian birds. *Evolution* **69:** 1823–1834.

Werman SD. 2005. Hypotheses on the historical biogeography of Bothropoid pitvipers and related genera of the Neotropics. In: Donnelly MA, Crother BI, Guyer C, Wake MH, White ME, eds. *Ecology and evolution in the tropics: a herpetological perspective*. Chicago: University of Chicago Press, 306–365.

Wüster W, Salomão MG, Duckett GJ; BBBSP. 1999. Mitochondrial DNA phylogeny of the Bothrops atrox species complex (Squamata: Serpentes: Viperidae). *Kaupia* **8:** 135–144.

Wüster W, Salomão MG, Thorpe RS, Puorto G, Furtado MFD, Hoge SA, Theakston RDG, Warrell DA. 1997. Systematics of the *Bothrops atrox* complex: new insights from multivariate analysis and mitochondrial DNA sequence information. In: Thorpe RS, Wüster W, Malhotra A, eds. *Venomous snakes. Ecology, evolution and snakebite*. New York: Oxford University Press, 99–113.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. BIC values for *K* values from 1 to 20 based on analyses of RADseq loci using adegenet.
Table S1. Point estimates and associated confidence intervals (based on results from 50 bootstrapped datasets for each pair of populations) from FastSimCoal analyses. Point estimates for genetically effective population size (Ne) represent the average value for the population across the runs in which it was included, and confidence intervals represent the range of values observed. Gene flow estimates are numbers of migrant individuals per generation with 'Nm 1→2' values representing migration from the first population to the second population, as listed in the first column of the table whereas 'Nm 2→1' represents migration in the opposite direction. Divergence estimates are given as generation times (G) and converted to years (Y) using a generation time estimate of 5 years (see Methods).

SHARED DATA

DNA sequences – mtDNA: GenBank accessions MG720268–MG720293.
DNA sequences – RADseq datasets: data deposited in the Dryad digital repository (Gibbs et al., 2017).
Historical demographic models used in the FastSimCoal: available at https://github.com/ mikesovic.