Phylogeographic model selection leads to insight into the evolutionary history of four-eyed frogs

Maria-Tereza C. Thomé[a] and Bryan C. Carstens[b,1]

[a]Departamento de Zoologia, Instituto de Biociências, Universidade Estadual Paulista, Campus Rio Claro, 13506900 Rio Claro, SP, Brazil; and [b]Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Columbus, OH 43210

Edited by John C. Avise, University of California, Irvine, CA, and approved April 12, 2016 (received for review February 11, 2016)

Phylogeographic research investigates biodiversity at the interface between populations and species, in a temporal and geographic context. Phylogeography has benefited from analytical approaches that allow empiricists to estimate parameters of interest from the genetic data (e.g., $\theta = 4N_e\mu$, population divergence, gene flow), and the widespread availability of genomic data allow such parameters to be estimated with greater precision. However, the actual inferences made by phylogeographers remain dependent on qualitative interpretations derived from these parameters' values and as such may be subject to overinterpretation and confirmation bias. Here we argue in favor of using an objective approach to phylogeographic inference that proceeds by calculating the probability of multiple demographic models given the data and the subsequent ranking of these models using information theory. We illustrate this approach by investigating the diversification of two sister species of four-eyed frogs of northeastern Brazil using single nucleotide polymorphisms obtained via restriction-associated digest sequencing. We estimate the composite likelihood of the observed data given nine demographic models and then rank these models using Akaike information criterion. We demonstrate that estimating parameters under a model that is a poor fit to the data is likely to produce values that lead to spurious phylogeographic inferences. Our results strongly imply that identifying which parameters to estimate from a given system is a key step in the process of phylogeographic inference and is at least as important as being able to generate precise estimates of these parameters. They also illustrate that the incorporation of model uncertainty should be a component of phylogeographic hypothesis tests.

information theory | model selection | *Pleurodema* | site frequency spectrum | Caatinga

In biological populations with interbreeding individuals, allele frequencies will inevitably change with time, both in stochastic and systematic manners, through neutral and adaptive processes. These processes—genetic drift, gene flow, mutation, recombination, and natural selection — constitute observable phenomena that lead directly to population structure, population divergence, and eventually speciation. Phylogeography is ideally situated to investigate systems where the microevolutionary processes that act within gene pools begin to form macroevolutionary patterns and has been described as the bridge between population genetics and phylogenetics (1). The power of the discipline comes from the consideration of geographic origin of individuals and populations along the continuum between populations and species (2, 3).

Phylogeographic research has progressed through several stages since Avise et al. (1) introduced the term. Initial studies were based on information that can be gathered from the genetic data under few assumptions, for example by calculating summary statistics or estimating gene trees. Inferences were then derived from qualitative interpretations about what that information implied about the evolutionary history of the system (e.g., refs. 4 and 5). This approach has been criticized as being prone to overinterpretation, because researchers are inclined to propose more detailed and complex historical scenarios than are actually supported by the data (6). The general response to such criticisms has been the widespread adoption of model-based methods to analyze phylogeographic data, particularly models that incorporate coalescent theory (7) to estimate parameters of interest under a formal framework. Model-based methods of phylogeographic inference clearly represent an advance to the field, but making inferences from these parameter estimates still forces researchers to make subjective decisions. Despite the potential complexity of the demographic models, the actual process of phylogeographic inference remains largely analogous to that of earlier investigations: The relative influence of evolutionary processes is derived from the magnitude of numeric values estimated for parameters that measure what the researchers believe to be important evolutionary processes. For example, subjective decisions regarding estimated rates of gene flow are commonly used to determine whether populations are reproductively isolated from their sister taxa (e.g., ref. 8) or conspecifics (e.g., ref. 9).

Once efficient algorithms and computational power became available, researchers applied model-based methods to phylogeographic research with little hesitation (but see ref. 10), with models implemented in software packages being particularly popular. For example, the paper describing a popular method that estimates temporal divergence with gene flow has been cited in more than 500 studies to date (11). Simulation-based techniques are also commonly applied to empirical systems, either to test competing hypotheses such as introgression and lineage sorting (e.g., refs. 12–14) or to test phylogeographic hypotheses against a null model (e.g., refs. 15–17). Such methods have been widely adopted by the phylogeographic community because model-based methods offer a path toward estimating putatively relevant parameters, and because the models themselves can be tailored to the particulars of a given system (e.g., refs. 18 and 19). Phylogeographic inferences are more transparent when based on parameters estimated under these models, and arguably less subjective. However, simply using a complex demographic model to analyze genetic data is not a guarantee that phylogeographic inferences will be correct.

In the cognitive sciences, researchers have long been mindful of confirmation bias, the tendency to interpret novel information in a manner consistent with preconceived ideas (20). People tend to seek out information that supports their preexisting beliefs and are unlikely to consider contradictory information. Particularly problematic is the primacy effect, in which the information that is learned first effectively has more emphasis than information that is obtained

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "In the Light of Evolution X: Comparative Phylogeography," held January 8–9, 2016, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/ILE_X_Comparative_Phylogeography.

Author contributions: M.-T.C.T. and B.C.C. designed research; M.-T.C.T. performed research; M.-T.C.T. collected field samples; M.-T.C.T. and B.C.C. analyzed data; and M.-T.C.T. and B.C.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: Files used in the analysis have been deposited at DRYAD (accession no. Q:7 ■■■).

[1]To whom correspondence should be addressed. Email: carstens.12@osu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1601064113/-/DCSupplemental.

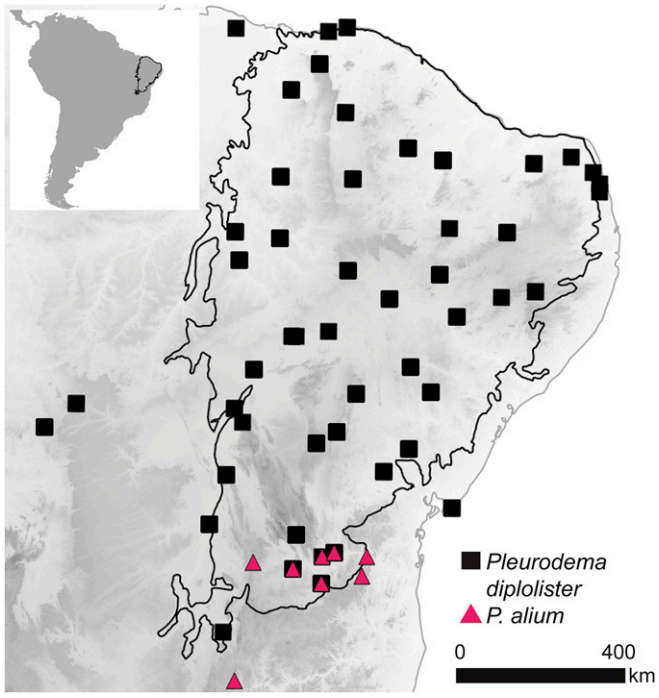

Fig. 1. Map of the sampling localities. The outline of the Caatinga is shown on an elevation map of northeastern Brazil, where darker shading corresponds to higher elevation. *P. diplolister* localities are marked with a dark square, *P. alium* localities with a pink triangle.

at a later date (20). Confirmation bias is likely prevalent in phylogeographic research (21), influencing phylogeographic inference by shaping the very questions that are asked by researchers. For example, if initial investigations into a given system used gene trees and phylogenetic thinking, researchers may not consider population processes such as gene flow as being potentially important, and choose to estimate divergence times under a species tree model, which may not actually fit the data (e.g., ref. 22). Researchers working in temperate systems in the Northern Hemisphere may assume that postglacial expansion is an important process and choose to estimate effective population size under growth models (e.g., ref. 23), whereas those working on focal taxa that inhabit island systems are likely to consider dispersal to be a key process shaping allele frequencies, and estimate effective population sizes under migration models (e.g., ref. 24). Such assumptions will guide choices about which models and software should be used to analyze the data and might also bias their interpretation of the values of parameters estimated under these models. Objective assessment of model fit should be an important component of phylogeographic research, particularly in systems where there is little preexisting information about the demographic history.

What If the Phylogeographic Model Is Wrong?

There is a great asymmetry in terms of the amount of available background information between model and nonmodel systems. In the extreme case of *Homo sapiens*, the analytical models used for data analysis are informed by the academic output of entire disciplines (e.g., anthropology) as well as thousands of previous genetic investigations. In contrast, the average phylogeographer likely knows very little about the focal organism before an investigation, save what can be inferred from its taxonomic placement and general habitat. This asymmetry is exacerbated for researchers interested in tropical diversity, which account for the vast majority of organisms: Chances are that even the most basic natural history traits (area of occurrence, density, feeding habitats, maturation age, and reproductive mode) are unknown to science. Given this paucity of information, how should researchers determine which models to use in data analysis?

In their review of statistical methods in phylogeography, Nielsen and Beaumont (25) argue strongly that population parameters should be estimated under appropriate models to avoid bias in the parameter estimates: "A clear limitation of any model-based method is that the model might be wrong. In fact, the real complexity of the demography of natural populations is unlikely to be captured by any simple model we could propose. In some cases, this may not affect inferences much, but in other cases it will." If phylogeographic inferences are largely derived from parameter estimates made under complex models, then such inferences are implicitly conditioned on the statistical fit of the model used to estimate these parameters to the empirical data collected from the focal system. To date, there has been too little attention devoted to methods for assessing the statistical fit of phylogeographic models to the data.

Statistical Frameworks for Phylogeography

Phylogeographic research is a historical discipline rather than an experimental one, and evolutionary history cannot be replicated. Because the experimental controls used in classical hypothesis testing are not available (e.g., ref. 26), testing hypotheses, even with parametric simulation (e.g., refs. 15 and 27), forces the phylogeography to conform to a statistical framework that may not be suited to historical research (28). A more promising strategy for phylogeographic data analysis is to proceed by identifying which of many possible models of historical demography offer the best statistical fit to the observed data, rather than testing null hypotheses, where rejection only tells us that the model representing the hypothesis is a poor fit to the data. If the goal of phylogeography is to infer the evolutionary history of the focal taxon, then ranking a set of models that represent alternative evolutionary scenarios provides a rigorous tool for inference because it will help researchers to avoid confirmation bias. Because the parameters in each model correspond to

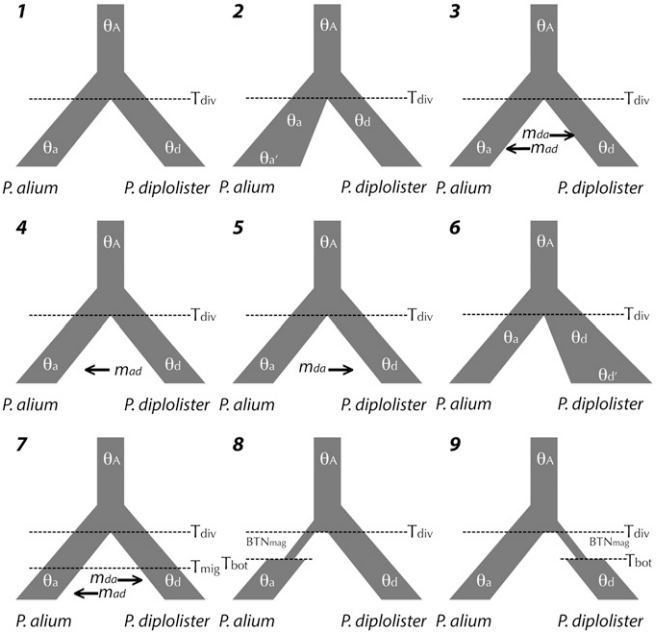

Fig. 2. Nine demographic models used in model selection are shown. Parameters abbreviations include genetic diversity of *P. alium* and *P. diplolister* (θ_a, θ_d), ancestral genetic diversity (θ_A), the timing of population divergence (Tdiv), migration between diverging lineages (m_{ad}, m_{da}), the rate of population expansion (exp), the timing of migration (T_{mig}), and bottlenecks (T_{bot}).

Table 1. Comparison of parameter estimated using FSC2 under four models

Model (w_i)	$N_{ancestral}$	N_{alium}	$N_{diplolister}$	T_{div}	$2Nm_{12}$	$2Nm_{21}$
3 (0.21)	1.48×10^4	6.86×10^4	134×10^6	5.86×10^4	0.069	0.904
4 (0.56)	1.43×10^4	6.98×10^4	1.33×10^6	5.88×10^4	0.072	n/a
7 (0.23)	5.59×10^3	6.92×10^4	1.36×10^6	5.93×10^4	0.078	0.738
		T_{MIG}				
		2.97×10^4				
6 (0.00)	2.65×10^2	8.20×10^4	2.3×10^6	3.28×10^6	n/a	n/a
	N_{found}	T_{exp}	G_{exp}			
	73	1.09×10^4	-4.6×10^{-5}			
Model average	1.25×10^4	6.94×10^4	1.34×10^6	5.89×10^4	0.073	0.783
Lower confidence interval	1.12×10^4	6.61×10^4	1.31×10^6	5.75×10^4	0.063	0.643
Upper confidence interval	1.37×10^4	7.26×10^4	1.37×10^6	6.02×10^4	0.083	0.887

Shown are estimates of population sizes ($N_{ancestral}$, N_{alium}, $N_{diplolister}$, and N_{found}), estimates of population divergence (T_{div}), the time that gene flow begins (T_{MIG}), the time that expansion begins (T_{exp}), gene flow ($2N_m$), and the magnitude of population size change (G_{exp}). The model probability of each model is shown in parentheses after the model number. All parameters were converted to real units assuming a mutation rate of 2.1×10^{-9}. See Table S1 for additional information regarding the results from all models. n/a, not assessed.

various evolutionary processes, the relative influence of particular evolutionary processes to the empirical system can be assessed by considering the set of parameters included in the model that offers the best fit to the data. Model selection is a useful framework for phylogeographic inference because it offers an approach that accounts for the uncertainty in the models used to analyze the data.

Model Selection in Bayesian and Information Theoretic Frameworks

Fagundes et al. (29) provided a compelling example of phylogeographic research using model selection in a Bayesian framework, using approximate Bayesian computation (ABC) to evaluate alternative models of human demographic history. Inspired by this work, many researchers have applied a similar approach to a wide range of nonmodel systems (e.g., refs. 30–34). However, as with any approach to data analysis, phylogeographic model choice using ABC has limitations, and decisions about which models to include in the comparison set can be challenging. Because ABC loses power to differentiate among models as the number of models in the comparison set increases (35), one cannot easily evaluate large numbers of models. Fagundes et al. (29) had the advantage of working in a model system where they could identify three types of models to test based on the results of hundreds of previous investigations, but the lack of similar information in nonmodel systems increases the odds of erroneous model choice and faulty phylogeographic inference.

A solution to evaluating a large number of models representing a great many possible demographic histories is to use information theory (36) to rank models. Information theory relies on the estimation of the Kullback–Leibler (37) information of a given model using the Akaike information criterion (AIC) (38), and the subsequent ranking of all models in the comparison set. The model ranking is achieved by calculating the difference between the AIC score of a particular model and the best model in the set (e.g., $\Delta_i = AIC_i - \min_{AIC}$), and subsequent transformation to model likelihoods (w_i) by normalizing AIC differences across the set of R models such that they sum to 1.0 [$w_i = exp\ (-1/2\Delta_i)/\sum_{r=1}^{R} exp\ (-1/2\Delta_r)$; see ref. 36]. A reasonable interpretation of these model probabilities is that they correspond to posterior probabilities under a uniform prior distribution (36). Information theory is commonly used to select models of DNA nucleotide substitution for analyses of sequence data (as in the software ModelTest; ref. 39), and has been effectively used to compare among large number of models in this context. To date, information theoretic approaches have been used in phylogeography to choose the best of several isolation-with-migration models (e.g., refs. 40 and 41), to evaluate models of postglacial expansion and colonization (21), and to evaluate models of source-sink migration (42, 43). In this paper, we briefly illustrate its application using data from the four-eyed frogs of northeastern Brazil.

Case Study: The *Pleurodema* System in the Brazilian Caatinga

Pleurodema alium and *Pleurodema diplolister* are sister species of four-eyed frogs that inhabit the Caatinga in northeastern Brazil (44). The Caatinga is a widespread xeric biome, surrounded by the extensive mesic environments of the Amazon, Cerrado, and Atlantic Rainforest. Its climate is highly seasonal and unpredictable, with severe droughts and rainless years. As is typical of amphibians from xeric habitats, *Pleurodema* persist throughout most of the year by burrowing underground, becoming active only after seasonal heavy rains create ephemeral pools for breeding. Even though the life cycle in *Pleurodema* depends on precipitation, these frogs cannot maintain populations in more mesic biomes and its distribution is restricted to the Caatinga xeric habitat.

Floristically, the Caatinga is one of the isolated nuclei in the Seasonally Dry Tropical Forests (SDTFs) of South America. The history of the SDTFs is debated, with some evidence suggesting that they were formerly continuous and recently fragmented [during the Last Glacial Maximum (LGM); ref. 45], and other evidence favoring an older (Tertiary) fragmentation (46). Environmental niche modeling results in contrasting maps ranging from a largely continuous to a fragmented Caatinga, depending on the approach used (47, 48). Regardless of the broader continental trends of the SDTFs, there is abundant geologic evidence that the Caatinga has been recurrently invaded (and at least partially replaced) by mesic forest throughout its history (49, 50).

P. alium and *P. diplolister* were recently the subject of phylogeographic investigation. Thomé et al. (51) collected >350 samples, sequenced the mitochondrial cytochrome oxidase I (COI) gene, and genotyped 12 microsatellite loci. Using these data, they were able to confirm that the species were distinct at the genetic level (both at COI and microsatellite markers), and that they have partly sympatric distributions: *P. alium* is restricted to the southern Caatinga, whereas *P. diplolister* is widespread in the biome, occurring also in pockets of Caatinga embedded within the Cerrado (Fig. 1). The population genetic structure within the broadly distributed *P. diplolister* reflected the distribution of its sister species, in that the *P. diplolister* samples that were sympatric with *P. alium* formed a separate genetic cluster.

Given the available information, a wide range of evolutionary processes (and therefore parameters) could be incorporated into a demographic model of *P. alium* and *P. diplolister*. Temporal divergence likely represents an important component, supported by

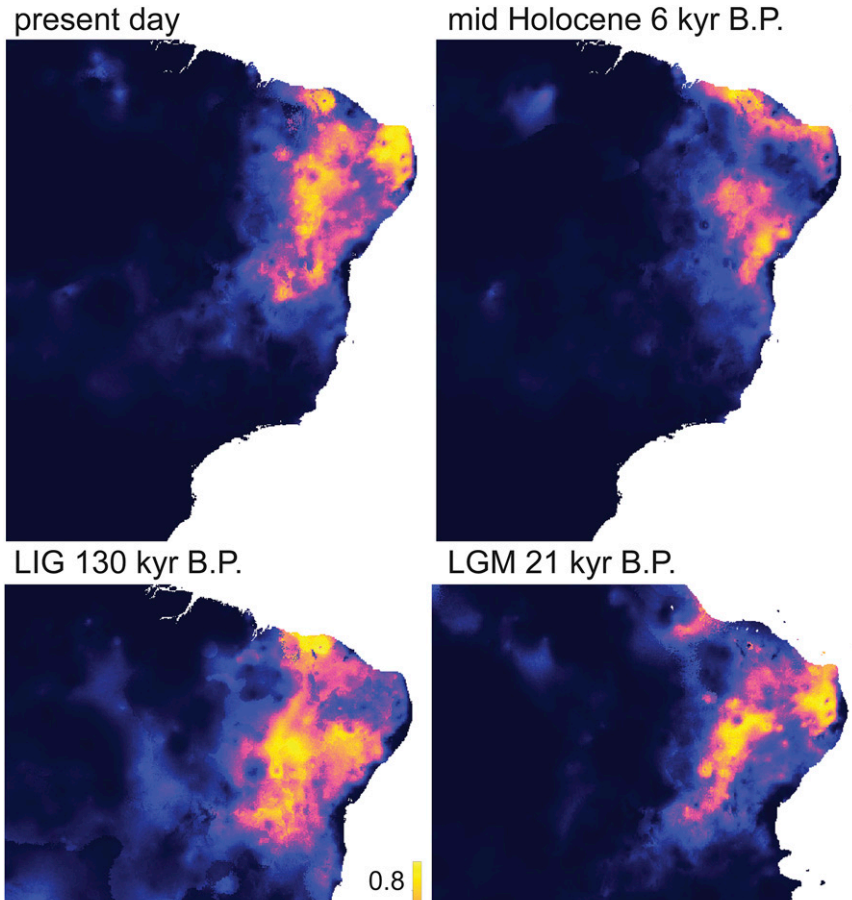

Fig. 3. Projections of suitable habitat for *P. alium* and *P. diplolister*. Shown clockwise from upper left are estimates of the current ecological niche, as well as projections of this niche onto past conditions of the mid Holocene, the LGM, and the LIG.

the deep divergence in the COI data (51). Effective population sizes are likely to differ between species, because *P. diplolister* has a much larger geographic range than *P. alium*, and probably a corresponding difference in census population size. Although range size and effective population size are not necessarily correlated, the difference in geographic range provides justification for allowing for the possibility of differences in effective population size among species, so long as we assume that the mutation rate does not vary between species. In addition to the processes of temporal divergence and different population sizes, other evolutionary processes could be important: population size change within species (such as population bottlenecks or exponential population growth), gene flow, and/or natural selection.

We specified nine demographic models for analysis, which were designed to represent a range of demographic histories. All models included lineage divergence between the sister taxa *P. alium* and *P. diploister* and some combination of the following demographic processes: population expansion or contraction, population bottlenecks, gene flow, and population-specific θ values (Fig. 2). There are hundreds of ways that the divergence of two species from a common ancestor could be parameterized (see ref. 35); here, we hope to specify models that span the range of possible models but include those that we believe to be plausible (e.g., we do not include *n*-island models that lack temporal divergence, because we consider divergence time to be an essential parameter to include in any model that contains sister species).

Sampling and Molecular Protocols. We sampled 183 individuals of *Pleurodema* from 55 locations in the core, isolates, or peripheral regions of the Caatinga, comprising most of its distribution in the Caatinga biome (see ref. 51). SNPs were collected via genome-wide sampling using restriction enzymes (double-digest RADseq; ref. 52). DNA digestion and barcode ligation were performed individually for each sample using 300 ng of freshly extracted DNA, the restriction enzymes Sbf1-HF and MspI, the ligation enzyme Ligase T4, and eight different barcoded Illumina adaptors. The digestion–ligation reactions were then pooled in groups of eight and purified with Agencourt AMPure beads, and PCR (12 cycles) was used to amplify the fragments containing barcodes using six different Illumina indexed primers and Phusion DNA polymerase. PCR products were quantified with Qubit Fluorometric Quantitation (Invitrogen), equimolar quantities of six groups containing eight samples each were pooled, and 250- to 500-bp fragments were selected using a Blue Pippin Prep. The fragment sizes were

confirmed with an Agilent 2100 Bioanalyzer (Agilent), and 100-bp, single-end, sequencing reactions were conducted using an Illumina HiSEq 2000 at Beckman Coulter Genomics.

Data Processing. Illumina outputs from *Pleurodema* samples were processed using the pyRAD pipeline (53). Except for the initial demultiplexing step, which was conducted separately on each library, we processed data for all samples together with the following parameter specifications: 10× minimal coverage, four or fewer unknown bases per sequence, minimum similarity of 0.90, a maximum ratio of shared polymorphisms of 20%, and a minimum coverage taxon of 70%. The number of reads that passed quality control was plotted against the number of loci obtained in each sample to establish a minimum number of reads for a sample to be considered. Because the number of loci stabilizes above 300,000 reads, we eliminated the 18 samples that were bellow this threshold before conducting a final SNP calling step in the remaining 165 samples. This scheme yielded 6,027 alignments containing SNPs.

Missing Data. After excluding SNPs that were possibly under selection (*Supporting Information*), our dataset consisted of 5,810 sequenced regions containing one or more SNPs. However, every region was not sequenced in each sample. Population-level data collected using RADseq and related protocols typically consist of data matrices with some degree of missing data (e.g., refs. 54 and 55), and these missing data can lead to biased estimates of effective population size and other parameters (56, 57). Missing data are likely to be particularly problematic for analytical methods that rely on estimates of allele frequencies because rare alleles may be undercounted. However, it is not clear how to best conduct analyses in a manner that accounts for the missing data. Missing data might be related to mutations in the recognition site of the enzymes, and removing all individuals that contain missing data about a certain threshold would be equal to removing the most divergent individuals, which could artificially homogenize the dataset and dramatically change the estimates of the number of rare alleles. Alternatively, removing all loci that contain missing data will dramatically reduce the size of any observed RADseq dataset and negate some of the advantages of collecting such data in the first place. Because we will analyze our data using a method that relies on estimates of the population site frequency spectra (discussed below), it is important to account for missing data in a manner that does not bias our estimate of these frequencies. To accomplish this, we choose SNPs (one per locus) and individuals at random from our full data and then replicated this downsampling 10 times using a Python script provided by Jordan D. Satler, The Ohio State University, Columbus, OH (*Supporting Information*). After the downsampling procedure, our replicate data matrixes consisted of approximately one-third of the total SNPs in one-half of the individuals and enabled us to calculate confidence intervals by comparing estimated parameters across replicates.

Model Selection. We estimated the composite likelihood of the probability of the observed data given the specified model using fastsimcoal2 (FSC2) (58). FSC2 estimates parameters specified by the user (including $\theta = 4N_e\mu$, population size change, gene flow, and population divergence) from the site frequency spectrum (SFS). Demographic processes will influence the site frequency distributions; for example, gene flow will produce an abundance of shared SNPs, population bottlenecks will result in a reduction of genetic diversity and thus fewer low-frequency SNPs, and so on. After the demographic model is specified, FSC2 selects initial parameter values at random from a range specified by the user and simulates data using the demographic model and parameter values. Composite likelihoods are calculated following Nielsen (59), who demonstrated that there is a relationship between the branch lengths of the genealogy and the probability of observing an SNP of a certain frequency distribution. Parameter optimization was conducted using the Brent algorithm implemented in FSC2, which identifies parameter values that maximize the likelihood estimate of the observed SFS given the demographic model. Finally, the maximized likelihood observed across all iterations is used in model comparison.

Using FSC2, the analysis of each of the 10 downsampled datasets was replicated 50 times (58). The individual run settings of each replicate included 100,000 simulations for the calculation of the composite likelihood and 50 cycles of the Brent algorithm (for parameter optimization). FSC2 analyses were conducted using massively parallel computing resources provided by the Ohio Supercomputer Center. After the maximum likelihood was estimated for each model in every replicate, we calculated the AIC scores and converted to model probabilities as above. This transformation allows us to measure the probability of each model given the observed data across replicates (e.g., Table S1), which we interpret as a measure of the degree of support for a particular model following ref. 60.

Results and Discussion

The results of the FSC2 analysis were consistent in the sense that only three models, all isolation with migration, have any appreciable model probability (i.e., >0.001; Table S1). The model with ongoing gene flow from *P. diplolister* to *P. alium* has the highest model probability. The secondary contact model and the model asymmetric gene flow between *P. diplolister* and *P. alium* have similar log-likelihoods given the data to the best model but lower AIC scores due to having additional parameters. Additionally, parameter estimates suggest that these models may be more similar than they seem (Table 1). For example, in the secondary contact model (i.e., model 7) parameter estimates of the time that gene flow begins are closer to the divergence of these species from their common ancestor than to the present, and in model 3 (i.e., the model with asymmetric gene flow) the rate of gene flow from *P. alium* to *P. diplolister* is estimated to be much lower than the rate of migration in the opposite direction (although these estimates are not perfectly comparable because the duration of gene flow is not the same under these models). Due to the similarity in parameters estimated by these models, our phylogeographic inferences are based on model-averaged parameter values (i.e., the value of a given parameter estimated under a particular model weighted by the model probability of that model, averaged across models that share the particular parameter; Table 1).

There are several striking features of the divergence with gene flow models. Assuming a mutation rate of 2.1×10^{-9} substitutions per site per generation (61) to convert parameter estimates, the ancestral effective population size (averaged across replicates and models) was estimated to be small (~12,500 individuals). *P. alium* and *P. diplolister* began to diverge from their common ancestor during the last glacial cycle of the Pleistocene (~58,900 y B.P.) but continued to exchange alleles via migration. The rate of migration into each species from the other was not equal; roughly 10 times as many *P. diplolister* migrants entered the *P. alium* gene pool than the reverse ($2Nm_{da} = 0.78$; $2Nm_{ad} = 0.07$). Finally, whereas the current effective population size of each species is estimated to be larger than the ancestral population, current effective population sizes in *P. diplolister* are substantially larger than in *P. alium* ($N_d = 1.34 \times 10^6$; $N_a = 6.9 \times 10^4$), consistent with differences in their geographic ranges.

Perhaps the most surprising result from our analysis is how much parameter estimates depend on the model used to estimate the parameters. For example, divergence time is estimated to be two orders of magnitude more ancient when estimated under model 6 (~3,280,000 y B.P.) than under the best-ranked model (Table 1), whereas the ancestral effective population size was estimated to be much smaller (2.65×10^2). Given the lack of previous estimates for these parameters in this system, there would be little reason to be suspicious of these values absent an assessment of model fit. This example illustrates the importance of performing phylogeographic model selection before any attempt to make inferences about the

evolutionary history of a system, especially those based on parameter estimates.

There are several advantages to basing phylogeographic inferences on the results of model selection exercises. Such analyses allow researchers to identify which evolutionary processes have shaped genetic diversity. In *Pleurodema*, the divergence of the sister taxa *P. alium* and *P. diplolister* is occurring despite ongoing gene flow. This inference stems directly from results of the model selection exercise: All of the models that have good AIC scores and thus receive any appreciable support include some gene flow between these species. This inference is not based on the magnitude of the parameter estimates, but solely on the inclusion of the gene flow parameters in the highest-ranked models. In addition, the results of the model selection analysis prevent us from overinterpreting our data (*sensu* ref. 6). In *Pleurodema*, previously collected evidence suggested that population expansion could represent an important feature of this system (51), but none of the population size change or bottleneck models offered a good fit to the empirical data. As much as we expected expansion to be a dominant force shaping these data, there is no evidence for the influence of this process in the SNP dataset. We attribute this discrepancy to one of two causes. It could be that there is an actual difference in the signal between the SNP data analyzed here and the microsatellite and COI data analyzed by Thomé et al. (51). Each of these markers evolves at a different rate and thus will be informative at different timescales. Thus, it is possible that faster markers perform better in detecting demographic expansions as recent as 4,240 y B.P. (50). However, because these analyses differed in the number of individuals included (approximately three times as many in the microsatellite analysis as here), as well as in details of each analysis, this difference could result from some combination of these differences.

What factors may have caused the initial divergence of *P. alium* and *P. diplolister*? Results from analyses of environmental (climatic) niche modeling provide two important clues. First, the environmental niche of *P. alium* does not differ from that of *P. diplolister* (see Box 1). This makes it unlikely that these species are undergoing adaptive diversification, a result that is supported by an outlier loci analysis (for example, a Bayescan analysis detects only 14 out of 6,027 loci as being potentially under selection; *Supporting Information*). Second, species distribution modeling supports the hypothesis of a dynamic distribution for the Caatinga, as the predicted distribution of these species has changed over the last 130,000 y, including being notably smaller at the mid Holocene, and somewhat reduced at the LGM (Fig. 3). These historical distributions are at odds with previous paleomodelling of the SDTFs but consistent with the palynological record, which indicates that the present-day distribution of the Caatinga established very recently in the late Holocene (4,240 y B.P.; ref. 50). The dynamic range of these species supports the idea that these lineages have been periodically fragmented, possibly isolated, with secondary contact inhibiting the formation of reproductive isolation.

New Data, Better Methods, and Improved Inferences from Nonmodel Organisms. One of the pressing issues facing the discipline of phylogeography in the past was the limited amount of genetic data that could be collected from most systems, and the poor quality of parameter estimates that resulted from analysis of these data (62–64). In the last decade, advances in sequencing technology have led to dramatic improvements in the amount of data that can be collected from nonmodel systems (65, 66). Given modest levels of funding, researchers can now collect more data from any system than are likely required to accurately estimate parameters of interest (e.g., refs. 64 and 67). With next-generation datasets, phylogeography is well-positioned to address a more important question: Which parameters are important to estimate in a given system? Whereas many of the methods applied by phylogeographic investigations were developed initially for the analysis of data from model systems (e.g., ref. 58), scientists working in nonmodel systems have been forced to confront the question of model fit, and in response they are developing creative solutions to identifying models that fit a particular system.

Some approaches to model selection are built into the framework of existing analytical methods. For example IMa (68), which implements a divergence with gene flow model, can be used to conduct model selection using either likelihood ratio tests (e.g., ref. 68) or information theoretic approaches (69). Similarly Migrate-*n* (42), which implements an *n*-island model, can be used to select among many migration models (42, 43). In addition, there are a number of approaches to species delimitation that incorporate model selection. These include methods that identify the optimal species delimitation using likelihood ratio tests (70), reversible-jump Markov chain Monte Carlo (71, 72), information theory (73), ABC (74), and marginalized likelihoods (75). Methods for analyzing comparative phylogeographic data are also under active development, including the use of hierarchical Bayesian models to test simultaneous divergence (76, 77) or simultaneous population expansion (78, 79).

Although methods that implement model selection are extremely useful, they lack the flexibility of simulation-based approaches, which provide researchers with the capacity to customize their models to the particular details of nearly any empirical systems. ABC continues to be a useful approach to model selection, particularly when implemented in computational environments such as R (e.g., ref. 80) that can be easily used by researchers. Other methods are available that calculate the probability of SNP data. In addition to FSC2, used here, model selection can be conducted using diffusion approximation in the software dadi (81).

Conclusions

Testing the statistical fit of our models given the data enabled us to address a major limitation of model-based phylogeography (19). By deriving our phylogeographic inferences from parameters estimated

Box 1 – Environmental Niche Models

We gathered 51 georeferenced occurrence points (2 for *P. alium* only, 44 for *P. diplolister* only, and 5 for both species) from sequenced samples collected in the core area of the Caatinga at a minimum distance of 8 km between points. We extracted climate information from 19 layers of bioclimatic variables available at the WorldClim website and used principal component analysis of occurrence data to compare their niches (82). Niche overlap was high ($D = 0.95$) and the hypothesis of niche equivalency could not be rejected ($P = 0.99$). The niches of the two species are more similar than would be by chance ($P = 0.0198$). To estimate past distributions we constructed correlative maps of potential distribution with the maximum entropy algorithm (83) and projected the model to past environmental conditions of the mid-Holocene (6,000 y B.P.) LGM at 21,000 y B.P. (MIROC4m general circulation model, Pliocene Model Intercomparison Project), and last interglacial (LIG) at 120,000 y B.P. (84). The study area encompasses current and putative past Caatinga distributions according to previous studies (47, 48). We selected eight uncorrelated variables (Pearson correlation <0.7) downloaded from Bioclim at 2.5 arc minutes resolution: mean diurnal range, isothermality, temperature seasonality, annual precipitation, precipitation of driest month, precipitation seasonality, precipitation of warmest quarter, and precipitation of coldest quarter. We used random training-test percentages (70% of observations for model training, and 30% for model testing), the auto features function, and the default regularization multiplier. The high mean value for the area under the receiver operating characteristics curve (AUC = 0.960, SD = 0.007, $n = 100$) indicates that the model performance was satisfactory. The most important variable was annual precipitation (evaluated with 100 iterations).

under suitable models, we avoided confirmation bias and overinterpretation. Parameter estimation was of central importance to our phylogeographic inference process, but only after we made an objective determination about which parameters to estimate. Perhaps the greatest advantage of this approach to phylogeography is that while the inferences themselves do not rely solely on parameter estimates, the parameters that are estimated via model averaging are likely to be more representative of the actual population values. It is incumbent on researchers who do not conduct model selection as part of their phylogeographic investigations to ask whether their phylogeographic inferences are based on a model of historical demography that is appropriate for their empirical system.

ACKNOWLEDGMENTS. We thank Célio F. B. Haddad, Miguel T. Rodrigues, José Pombal, Jr., and Marcelo Nápoli for donation of samples; ICMBio for the collecting permit (30512); and Francisco Brusquetti for help in the field. We also thank members of the B.C.C. laboratory and two reviewers for comments that improved this manuscript prior to publication. Financial support was provided by Fundação Grupo Boticário de Proteção à Natureza Grant 0909_20112 and São Paulo Research Foundation Grants 2012/50255-2, 2011/51392-0, and 2013/09088-8. Computational resources were provided by the Ohio Supercomputer Center.

1. Avise JC, et al. (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu Rev Ecol Evol Syst* 18:489–522.
2. Knowles LL (2004) The burgeoning field of statistical phylogeography. *J Evol Biol* 17(1):1–10.
3. Hickerson MJ, et al. (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Mol Phylogenet Evol* 54(1):291–301.
4. Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech *Fagus sylvatica* L. in Europe. *Evolution* 50:2515–2520.
5. Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Mol Ecol* 7:431–452.
6. Knowles LL, Maddison WP (2002) Statistical phylogeography. *Mol Ecol* 11(12): 2623–2635.
7. Kingman JFC (1982) The coalescent. *Stoch Proc Applic* 13:235–248.
8. Dolman G, Moritz C (2006) A multilocus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). *Evolution* 60(3):573–582.
9. Runemark A, Hey J, Hansson B, Svensson EI (2012) Vicariance divergence and gene flow among islet populations of an endemic lizard. *Mol Ecol* 21(1):117–129.
10. Templeton AR (2010) Coalescent-based, maximum likelihood inference in phylogeography. *Mol Ecol* 19(3):431–435, discussion 436–446.
11. Hey J (2010) Isolation with migration models for more than two populations. *Mol Biol Evol* 27(4):905–920.
12. Reid N, Demboski JR, Sullivan J (2012) Phylogeny estimation of the radiation of western North American chipmunks (*Tamias*) in the face of introgression using reproductive protein genes. *Syst Biol* 61(1):44–62.
13. Debiasse MB, Nelson BJ, Hellberg ME (2014) Evaluating summary statistics used to test for incomplete lineage sorting: Mito-nuclear discordance in the reef sponge *Callyspongia vaginalis*. *Mol Ecol* 23(1):225–238.
14. Grummer JA, et al. (2015) Estimating the temporal and spatial extent of gene flow among sympatric lizard populations (genus Sceloporus) in the southern Mexican highlands. *Mol Ecol* 24(7):1523–1542.
15. Knowles LL (2001) Did the pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshopprers. *Mol Ecol* 10(3):691–701.
16. DeChaine EG, Martin AP (2005) Historical biogeography of two alpine butterflies in the Rocky Mountains: Broad-scale concordance and local-scale discordance. *J Biogeogr* 32:1943–1956.
17. Smith CI, et al. (2011) Comparative phylogeography of a coevolved community: Concerted population expansions in Joshua trees and four yucca moths. *PLoS One* 6(10):e25628.
18. Knowles LL (2009) Statistical phylogeography. *Annu Rev Ecol Evol Syst* 40:593–612.
19. Beaumont MA, et al. (2010) In defense of model-based inference in phylogeography. *Mol Ecol* 19:436–446.
20. Nickerson RS (1998) Confirmation bias: A ubiquitous phenomenon in many guises. *Rev Gen Psychol* 2:175–220.
21. Carstens BC, et al. (2013) Model selection as a tool for phylogeographic inference: An example from the willow *Salix melanopsis*. *Mol Ecol* 22(15):4014–4028.
22. Reid NM, et al. (2014) Poor fit to the multispecies coalescent is widely detectable in empirical data. *Syst Biol* 63(3):322–333.
23. Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics* 22(6):768–770.
24. Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc Natl Acad Sci USA* 98(8):4563–4568.
25. Nielsen R, Beaumont MA (2009) Statistical inferences in phylogeography. *Mol Ecol* 18(6):1034–1047.
26. Neyman J, Pearson ES (1933) On the problem of the most efficient tests of statistical hypotheses. *Phil Trans R Soc A* 231:694–706.
27. Carstens BC, Stevenson AL, Degenhardt JD, Sullivan J (2004) Testing nested phylogenetic and phylogeographic hypotheses in the *Plethodon vandykei* species group. *Syst Biol* 53(5):781–792.
28. Cleland CA (2001) Historical science, experimental science, and the scientific method. *Geology* 29:987–990.
29. Fagundes NJ, et al. (2007) Statistical evaluation of alternative models of human evolution. *Proc Natl Acad Sci USA* 104(45):17614–17619.
30. Tsai Y-HE, Carstens BC (2013) Assessing model fit in phylogeographic investigations: An example from the North American willow *Salix melanopsis*. *J Biogeogr* 40: 131–141.
31. Espindola A, Carstens BC, Alvarez N (2014) Comparative phylogeography of mutualists and the effect of the host on the genetic structure of its partners. *Biol J Linn Soc Lond* 113:1021–1035.
32. Jamamillo-Correa JP, Gerardi S, Beaulieu J, Ledig FT, Bousquet J (2015) Inferring and outlining past population declines with linked microsatellites: A case study in two spruce species. *Tree Genet Genomes* 11:1–12.
33. Peres EA, et al. (2015) Pleistocene niche stability and lineage diversification in the subtropical spider Araneus omnicolor (Araneidae). *PLoS One* 10(4):e0121543.
34. Vera-Escalona I, Habit E, Ruzzante DE (2015) Echoes of a distant time: Effects of historical processes on contemporary genetic patterns in *Galaxias platei* in Patagonia. *Mol Ecol* 24(16):4112–4128.
35. Pelletier TA, Carstens BC (2014) Model choice for phylogeographic inference using a large set of models. *Mol Ecol* 23(12):3028–3043.
36. Burnham KP, Anderson DR (1998) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer, New York).
37. Kullback S, Leibler RA (1951) On information and sufficiency. *Ann Math Stat* 22:79–86.
38. Akaike H (1973) Information theory as an extension of the maximum likelihood principle. *Second International Symposium on Information Theory*, eds Petrov BN, Csaki F (Akademiai Kiado, Budapest), pp 267–281.
39. Posada D, Crandall KA (1998) MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14(9):817–818.
40. Koopman MM, Carstens BC (2010) Conservation genetic inferences in the carnivorous plant *Sarracenia alata* (Sarraceniaceae). *Conserv Genet* 11:2027–2038.
41. Rittmeyer EN, Austin CC (2015) Combined next-generation sequencing and morphology reveal fine-scale speciation in Crocodile Skinks (Squamata: Scincidae: Tribolonotus). *Mol Ecol* 24(2):466–483.
42. Beerli P, Palczewski M (2010) Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* 185(1):313–326.
43. Barrow LN, Bigelow AT, Phillips CA, Lemmon EM (2015) Phylogeographic inference using Bayesian model comparison across a fragmented chorus frog species complex. *Mol Ecol* 24(18):4739–4758.
44. Faivovich J, et al. (2012) A phylogenetic analysis of *Pleurodema* (Anura: Leptodactylidae: Leiuperinae) based on mitochondrial and nuclear gene sequences, with comments on the evolution of anuran foam nests. *Cladistics* 28:460–482.
45. Prado DE, Gibbs PE (1993) Patterns of species distributions in the dry seasonal forests of South America. *Ann Miss Bot Gard* 80(4):902–927.
46. Pennington RT, Prado DE, Pendry CA (2000) Neotropical seasonally dry forests and Quaternary vegetation changes. *J Biogeogr* 27:261–273.
47. Werneck FP, Costa GC, Colli GR, Prado DE, Sites JW (2011) Revisiting the historical distribution of Seasonally Dry Tropical Forests: New insights based on palaeodistribution modelling and palynological evidence. *Glob Ecol Biogeogr* 20:272–288.
48. Collevatti RG, et al. (2013) Drawbacks to palaeodistribution modelling: The case of South American seasonally dry forests. *J Biogeogr* 40:345–358.
49. Auler AS, et al. (2004) Quaternary ecological and geomorphic changes associated with rainfall events in presently semi-arid northeastern Brazil. *J Quaternary Sci* 19: 693–701.
50. de Oliveira PE, Barreto AMF, Suguio K (1999) Late Pleistocene/Holocene climatic and vegetational history of the Brazilian caatinga: The fossil dunes of the middle São Francisco River. *Palaeogeogr Palaeoclimatol Palaeoecol* 152:319–337.
51. Thomé MCT, et al. (2016) Recurrent connections between Amazon and Atlantic forests shaped diversity in Caatinga four-eyed frogs. *J Biogeography*, 10.1111/jbi.12685.
52. 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(5):e37135.
53. Eaton DAR (2014) PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30(13):1844–1849.
54. Rubin BER, Ree RH, Moreau CS (2012) Inferring phylogenies from RAD sequence data. *PLoS One* 7(4):e33394.
55. Wagner CE, et al. (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Mol Ecol* 22(3):787–798.
56. Arnold B, Corbett-Detig RB, Hartl D, Bomblies K (2013) RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Mol Ecol* 22(11):3179–3190.
57. Gautier M, et al. (2013) Estimation of population allele frequencies from next-generation sequencing data: Pool-versus individual-based genotyping. *Mol Ecol* 22(14): 3766–3779.
58. Excoffier L, Dupanloup I, Huerta-Sánchez E, Sousa VC, Foll M (2013) Robust demographic inference from genomic and SNP data. *PLoS Genet* 9(10):e1003905.
59. Nielsen R (2000) Estimation of population parameters and recombination rates from single nucleotide polymorphisms. *Genetics* 154(2):931–942.
60. Anderson DR (2008) *Model Based Inference in the Life Sciences* (Springer, New York).

61. Gottscho AD, Marks SB, Jennings WB (2014) Speciation, population structure, and demographic history of the Mojave Fringe-toed Lizard (*Uma scoparia*), a species of conservation concern. *Ecol Evol* 4(12):2546–2562.
62. Edwards SV, Beerli P (2000) Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54(6): 1839–1854.
63. Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends Ecol Evol* 18:249–256.
64. Felsenstein J (2006) Accuracy of coalescent likelihood estimates: Do we need more sites, more sequences, or more loci? *Mol Biol Evol* 23(3):691–700.
65. McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Mol Phylogenet Evol* 66(2):526–538.
66. Garrick RC, et al. (2015) The evolution of phylogeographic data sets. *Mol Ecol* 24(6): 1164–1171.
67. Carling MD, Brumfield RT (2007) Gene sampling strategies for multi-locus population estimates of genetic diversity (θ). *PLoS One* 2(1):e160.
68. Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc Natl Acad Sci USA* 104(8):2785–2790.
69. Carstens BC, Stoute HN, Reid NM (2009) An information-theoretical approach to phylogeography. *Mol Ecol* 18(20):4270–4282.
70. Knowles LL, Carstens BC (2007) Delimiting species without monophyletic gene trees. *Syst Biol* 56(6):887–895.
71. Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Proc Natl Acad Sci USA* 107(20):9264–9269.
72. Solís-Lemus C, Knowles LL, Ané C (2015) Bayesian species delimitation combining multiple genes and traits in a unified framework. *Evolution* 69(2):492–507.
73. Ence DD, Carstens BC (2011) SpedeSTEM: A rapid and accurate method for species delimitation. *Mol Ecol Resour* 11(3):473–480.
74. Camargo A, Morando M, Avila LJ, Sites JW, Jr (2012) Species delimitation with ABC and other coalescent-based methods: A test of accuracy with simulations and an empirical example with lizards of the *Liolaemus darwinii* complex (Squamata: Liolaemidae). *Evolution* 66(9):2834–2849.
75. Leaché AD, Fujita MK, Minin VN, Bouckaert RR (2014) Species delimitation using genome-wide SNP data. *Syst Biol* 63(4):534–542.
76. Hickerson MJ, Stahl E, Takebayashi N (2007) msBayes: Pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation. *BMC Bioinformatics* 8:268.
77. Oaks JR, et al. (2013) Evidence for climate-driven diversification? A caution for interpreting ABC inferences of simultaneous historical events. *Evolution* 67(4): 991–1010.
78. Chan YL, Schanzenbach D, Hickerson MJ (2014) Detecting concerted demographic response across community assemblages using hierarchical approximate Bayesian computation. *Mol Biol Evol* 31(9):2501–2515.
79. Xue AT, Hickerson MJ (2015) The aggregate site frequency spectrum for comparative population genomic inference. *Mol Ecol* 24(24):6223–6240.
80. Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends Ecol Evol* 25(7):410–418.
81. Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet* 5(10):e1000695.
82. Broennimann O, et al. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Glob Ecol Biogeogr* 21:481–497.
83. Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Modell* 190:231–259.
84. Otto-Bliesner BL, Marshall SJ, Overpeck JT, Miller GH, Hu A (2006) Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science* 311(5768): 1751–1753.
85. Hahn C, Bachmann L, Chevreux B (2013) Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads–A baiting and iterative mapping approach. *Nucleic Acids Res* 41(13):e129.
86. Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 180(2):977–993.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES 1

Q: 1_Please contact PNAS_Specialist.djs@sheridan.com if you have questions about the editorial changes, this list of queries, or the figures in your article. Please include your manuscript number in the subject line of all email correspondence; your manuscript number is 201601064.

Q: 2_Please (i) review the author affiliation and footnote symbols carefully, (ii) check the order of the author names, and (iii) check the spelling of all author names, initials, and affiliations. Please check with your coauthors about how they want their names and affiliations to appear. To confirm that the author and affiliation lines are correct, add the comment "OK" next to the author line. This is your final opportunity to correct any errors prior to publication. Misspelled names or missing initials will affect an author's searchability. Once a manuscript publishes online, any corrections (if approved) will require publishing an erratum; there is a processing fee for approved erratum. Author names are edited to exactly match the submission form. A hyphen was inserted in the name "Maria-Tereza." If it is deleted, please also delete it in the contributions footnote.

Q: 3_Please review and confirm your approval of the short title: Phylogeographic model selection. If you wish to make further changes, please adhere to the 50-character limit. (NOTE: The short title is used only for the mobile app and the RSS feed.)

Q: 4_Please review the information in the author contribution footnote carefully. Please make sure that the information is correct and that the correct author initials are listed. Note that the order of author initials matches the order of the author line per journal style. You may add contributions to the list in the footnote; however, funding should not be an author's only contribution to the work.

Q: 5_Please verify that all supporting information (SI) citations are correct. Note, however, that the hyperlinks for SI citations will not work until the article is published online. In addition, SI that is not composed in the main SI PDF (appendices, datasets, movies, and "Other Supporting Information Files") have not been changed from your originally submitted file and so are not included in this set of proofs. The proofs for any composed portion of your SI are included in this proof as subsequent pages following the last page of the main text. If you did not receive the proofs for your SI, please contact **PNAS_Specialist.djs@sheridan.com**.

Q: 6_Please check the order of your keywords and approve or reorder them as necessary. Note that PNAS allows up to five keywords; please do not add new keywords unless you wish to replace others.

Q: 7_Please indicate whether the data have been deposited at DRYAD or another publicly accessible database before your page proofs are returned. It is PNAS policy that the data be deposited BEFORE the paper can be published. Also, please supply the missing accession number, ID code, DOI, or other type of data identification and change the wording of the data deposition footnote as needed.

Q: 8_Please cite Dataset S1 and Dataset S2 in order in the main text.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES 2

Q: 9_Former reference 27 was a duplicate of 15 and has been renumbered as 15. All subsequent references beginning with 28 were renumbered.

Q: 10_Abbreviations for gene names must be defined if used more than once in the paper. In the sentence beginning “Thomé et al. (51) collected” the abbreviation “COI” was defined as “cytochrome oxidase I.” Please confirm this is correct or alter the definition.

Q: 11_Per PNAS policy institutions and locations are given for individual suppliers of gifts or information. In the sentence beginning " To accomplish this, we choose SNPs," the institution and location " The Ohio State University, Columbus, OH" were given for Jordan D. Satler. Please confirm this is correct or insert a different institution/location.

Q: 12_For all journal references in which the issue number is missing, please add the issue number if the journal assigns one.

Q: 13_Please check the legend for Fig. 2 and revise it if needed. The abbrevation “exp” defined in the legend does not appear in the figure, and the abbreviation “BTN_{mag}” is not defined in the legend.

Q: 14_In the legend for Table 1 the abbreviation “n/a” was defined as “not assessed.” Please confirm this is correct or alter the definition.

Q: 15_Single subheadings cannot be used, per PNAS style. Please either provide an additional subheading under “Results and Discussion” or delete the single subheading, “New Data, Better Methods, and Improved Inferences from Nonmodel Organisms.”

Q: 16_Duplicate headings in Table 1 were deleted per style. Please note that the information pertaining to T_{MIG}, N_{found}, T_{exp}, and G_{exp} must be removed from the table and placed into the table legend or new table footnotes to simplify the table; there cannot be two different types of information within a single column. Please make appropriate changes.

Supporting Information

Thomé and Carstens 10.1073/pnas.1601064113

Scans for Mitochondrial Fragments and Loci Under Selection

To verify the possible presence of the fragments of the mitochondrial genome, we performed the in silico digestion of the mitochondrial genome from a closely related *Pleurodema* species for which there is a complete mitochondria sequence available (*Pleurodema thaul*) in 2.0 Webcutter online program (RNA.lundberg.gu.se/cutter2/). We found cutting sites for the restriction enzymes used in the RADseq protocol to be extremely rare and producing fragments of sizes larger than the range we selected. We also aligned the raw sequences of some individuals over this genome using MITObim (85), with no significant matches.

We used the Bayesian approach in Bayescan 2.1 (86) to detect outlier loci under two different configurations. First, to avoid possible interference of loci under selection in the model selection approach, we defined populations according to the locations of origin of samples. This analysis yielded 217 loci possibly under selection with a 0.05 target false discovery rate. Second, we used Bayescan defining populations according to species assignments and co-occurrence (sympatry or allopatry), which reduced the number of outlier loci to 14.

Script Used to Downsample SNP Replicate Datasets

A Python script used to downsample SNP data and build AFS files for analysis in FSC2 was kindly provided by Jordan D. Satler and is available on GitHub at https://github.com/jordansatler/SNPtoAFS. Q:1

Table S1. Results of FSC2 analyses averaged across replicates

Model	lnL	k	AIC	Δ_i	w_i
1	−5,625.6470	4	11,259.294	83.77	0.00
2	−5,668.5443	7	11,351.087	175.57	0.00
3	−5,582.7282	6	11,177.456	1.94	0.21
4	−5,582.7601	5	11,175.520	0	0.56
5	−5,625.6169	5	11,261.234	85.71	0.00
6	−6,556.0953	7	13,126.191	1,950.67	0.00
7	−5,581.6987	7	11,177.397	1.88	0.23
8	−5,625.3260	6	11,262.652	87.13	0.00
9	−5,625.4058	6	11,262.812	87.29	0.00

Shown from left for each model (see Fig. 2) are the maximum likelihood estimate of the model (lnL), the number of parameters (k), the AIC score, the Aikaike differences (Δ_i), and model probabilities (w_i). Information theoretic calculations follow Anderson (60).

Other Supporting Information Files

Dataset S1 (TXT)
Dataset S2 (TXT)

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

Q: 1_The section “ReadMe: SNPtoAFS” was deleted because this information is given in the README file available on GitHub.