Received: 25 January 2017 | Accepted: 17 July 2017

DOI: 10.1111/mec.14266

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

# Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection

Daniel M. Portik[1] | Adam D. Leaché[2,3] | Danielle Rivera[1] | Michael F. Barej[4] | Marius Burger[5,6] | Mareike Hirschfeld[4] | Mark-Oliver Rödel[4] | David C. Blackburn[7] | Matthew K. Fujita[8]

[1]Department of Biology, The University of Texas at Arlington, Arlington, TX, USA

[2]Department of Biology, University of Washington, Seattle, WA, USA

[3]Burke Museum of Natural History and Culture, University of Washington, Seattle, WA, USA

[4]Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany

[5]African Amphibian Conservation Research Group, Unit for Environmental Sciences and Management, North-West University, Potchefstroom, South Africa

[6]Flora Fauna & Man, Ecological Services Ltd., Tortola, British Virgin Island

[7]Florida Museum of Natural History, University of Florida, Gainesville, FL, USA

[8]Amphibian and Reptile Diversity Research Center, The University of Texas at Arlington, Arlington, TX, USA

**Correspondence**
Daniel M. Portik and Matthew K. Fujita, Department of Biology, The University of Texas at Arlington, Arlington, TX, USA.
Emails: daniel.portik@uta.edu; mkfujita@uta.edu

**Funding information**
National Science Foundation, Grant/Award Number: DEB # 1457232, DEB #1202609; National Science Foundation GRFP; Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH, Grant/Award Number: S10 OD018174; Museum of Comparative Zoology (Harvard University)

## Abstract

The accumulation of biodiversity in tropical forests can occur through multiple allopatric and parapatric models of diversification, including forest refugia, riverine barriers and ecological gradients. Considerable debate surrounds the major diversification process, particularly in the West African Lower Guinea forests, which contain a complex geographic arrangement of topographic features and historical refugia. We used genomic data to investigate alternative mechanisms of diversification in the Gaboon forest frog, *Scotobleps gabonicus*, by first identifying population structure and then performing demographic model selection and spatially explicit analyses. We found that a majority of population divergences are best explained by allopatric models consistent with the forest refugia hypothesis and involve divergence in isolation with subsequent expansion and gene flow. These population divergences occurred simultaneously and conform to predictions based on climatically stable regions inferred through ecological niche modelling. Although forest refugia played a prominent role in the intraspecific diversification of *S. gabonicus*, we also find evidence for potential interactions between landscape features and historical refugia, including major rivers and elevational barriers such as the Cameroonian Volcanic Line. We outline the advantages of using genomewide variation in a model-testing framework to distinguish between alternative allopatric hypotheses, and the pitfalls of limited geographic and molecular sampling. Although phylogeographic patterns are often species-specific and related to life-history traits, additional comparative studies incorporating genomic data are necessary for separating shared historical processes from idiosyncratic responses to environmental, climatic and geological influences on diversification.

**KEYWORDS**
Africa, amphibians, ecological niche model, forest refugia, historical demography, Lower Guinea, phylogeography, riverine barrier

  

## 1 | INTRODUCTION

The Guineo-Congolian tropical rainforests of Africa contain exceptional species richness and endemism characteristic of biodiversity hot spots (Myers, Mittermeier, Mittermeier, da Fonseca, & Kent, 2000). The Lower Guinea (LG) region is a block of the Guineo-Congolian forest extending from southern Nigeria to southern Gabon, which can be divided into two biogeographic regions based on species turnover coinciding with the Sanaga River in Cameroon: the Gulf of Guinea forests to the north and the Atlantic Equatorial and Congolian forests to the south (Droissart et al., 2011; Linder et al., 2012; Oates, Bergl, & Linder, 2004). The LG region is notable for its high biodiversity, complex terrain and dynamic forest history, yet despite these qualities, the evolutionary processes responsible for generating and maintaining this diversity are not well understood. Several mechanisms of diversification in tropical forests have been proposed and adapted for Africa, including Pleistocene forest refugia, riverine barriers and ecological gradients (Crowe & Crowe, 1982; Endler, 1982; Haffer, 1997; Mayr & O'Hara, 1986; Moritz, Patton, Schneider, & Smith, 2000; Plana, 2004; Smith, Wayne, Girman, & Bruford, 1997). While controversy still exists over which diversification process is most prevalent for African forests, each of these models has been demonstrated to underlie speciation and population differentiation across a variety of taxa. Important to several of these hypotheses are the climatic oscillations associated with Pleistocene glacial cycles that caused cyclical changes in forest and savannah habitats (deMenocal, 2004), resulting in the formation of discrete forest refugia and shifting ecotones (Anhuf et al., 2006; Cowling et al., 2008; Maley, 1996). Many empirical studies of LG have focused on examining patterns of allopatric divergence associated with Pleistocene forest refugia or riverine barriers in vertebrates (Anthony et al., 2007; Clifford et al., 2004; Jacquet et al., 2014; Nicolas et al., 2008, 2011; Telfer et al., 2003) and in plants (Bohoussou et al., 2015; Born et al., 2011; Budde, González-Martínez, Hardy, & Heuertz, 2013; Daïnou et al., 2010; Dauby, Duminil, Heuertz, & Hardy, 2010; Dauby et al., 2014; Debout, Doucet, & Hardy, 2011; Duminil et al., 2010, 2015; Faye et al., 2016; Gomez et al., 2009; Hardy et al., 2013; Koffi, Hardy, Doumenge, Cruaud, & Heuertz, 2011; Ley, Heuertz, & Hardy, 2016; Ley et al., 2014; Lowe, Harris, Dormontt, & Dawson, 2010; Muloko-Ntoutoume, Petit, White, & Abernethy, 2000). However, the role of divergent selection along ecological gradients for parapatric speciation is becoming increasingly recognized (Freedman, Thomassen, Buermann, & Smith, 2010; Heuertz, Duminil, Dauby, Savolainen, & Hardy, 2014; Jacquet et al., 2015; Mitchell, Locatelli, Clee, Thomassen, & Gonder, 2015; Mitchell, Locatelli, Ghobrial et al., 2015; Smith et al., 2005, 2011). Although multiple mechanisms of diversification have been demonstrated in tropical Africa, few studies have taken an explicit demographic model-testing approach to distinguish between alternative scenarios. Here, we briefly review the major allopatric and parapatric diversification processes, including empirical support and associated limitations, before introducing our model-testing framework and aims for this study.

### 1.1 | Alternative models of diversification for tropical Africa

The Pleistocene forest refugia hypothesis asserts that fragmented forest patches formed during glacial maxima produced isolated subpopulations of forest-restricted taxa (Endler, 1982; Haffer, 1969, 1997; Mayr & O'Hara, 1986; Moritz et al., 2000; Plana, 2004). A corollary to this hypothesis is that during warmer and moister interglacial periods, forest habitats expanded and enabled secondary contact between formerly isolated populations. Several forest refugia have been proposed in LG based on limited paleopalynological data or correlations between gradients of species richness and endemism, which often coincide with regions of higher elevation (Küper et al., 2004; Maley, 1996; Mayr & O'Hara, 1986; Sosef, 1996). Although the existence of forest refugia is widely accepted, the locations of these LG refugia remain contentious because high species richness may also be derived from ecological diversity associated with topographic complexity (Hofer, Bersier, & Borcard, 1999, 2000), rather than accumulation through historical isolation. Additionally, a lack of palynological fossil data prevents accurate spatial resolution of forest change, and although phylogeographic studies have been used to support proposed refugia, these interpretations can be subjective (see Discussion). An improved method for testing the location and effects of refugia involves inferring areas of habitat stability through environmental niche modelling (ENM) and coupling these results with phylogeographic and demographic inferences (Carnaval, Hickerson, Haddad, Rodrigues, & Moritz, 2009; Knowles, Carstens, & Keat, 2007; Moussalli, Moritz, Williams, & Carnaval, 2009). The identification of species-specific regions of stability can provide strong evidence for the impacts of forest refugia on diversification patterns, but this approach has so far been rare for African taxa (but see Faye et al., 2016).

An alternative allopatric model of divergence is the riverine barrier hypothesis, which predicts that species and populations are delimited by tropical rivers, with sister lineages occurring on opposite banks (Gascon et al., 2000; Haffer, 1997; Moritz et al., 2000; Plana, 2004; Wallace, 1852). Although clear differentiation should occur in lower parts of the river, genetic similarity is expected to increase towards the headwaters where opportunities for dispersal and gene flow are higher. Several major rivers occur in LG, including the Cross, Sanaga and Ogooué Rivers (Figure 1). Although the geochronological histories of specific rivers are not strongly resolved, major drainage basins in LG were formed during the Early Cenozoic and they have remained largely intact (Goudie 2005). The rivers of LG appear to be important barriers partitioning several populations or species of vertebrates (Anthony et al., 2007; Bohoussou et al., 2015; Gonder et al., 2011; Jacquet et al., 2015; Leaché & Fujita, 2010; Mitchell, Locatelli, Clee et al., 2015; Telfer et al., 2003), but have been dismissed as largely unimportant for plants (Dauby et al., 2014; Debout et al., 2011; Hardy et al., 2013; Ley et al., 2014; Lowe et al., 2010). Previous work has also demonstrated that genetic patterns of variation may also differ between organelle and nuclear genomes with regard to rivers, highlighting the limitations of

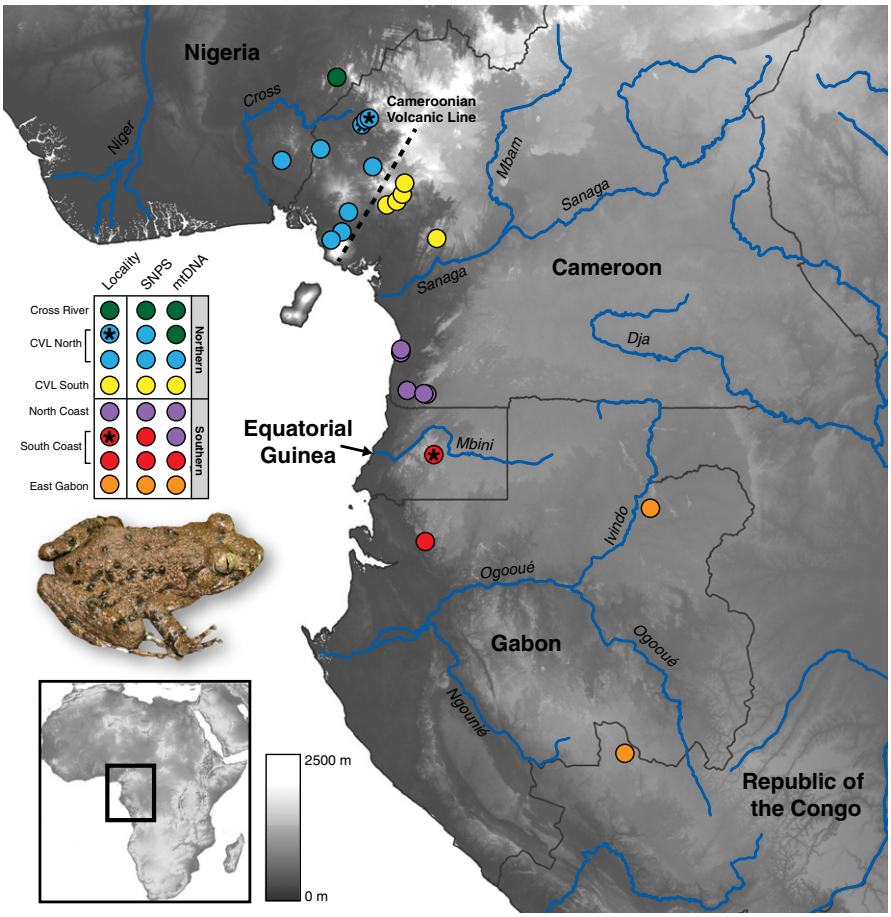


**FIGURE 1** Sampling localities and the geographic distribution of six genetically distinct populations of *Scotobleps gabonicus*. Locality colours correspond to assignment based on genomic data, which matched assignment based on mtDNA unless indicated by asterisks. Major rivers are depicted, along with relevant topographic features such as the Cameroonian Volcanic Line

using organelle loci to differentiate between scenarios of gene flow and incomplete lineage sorting (Anthony et al., 2007; Clifford et al., 2004; Gonder et al., 2011; Mitchell, Locatelli, Clee et al., 2015). The locations of major rivers between proposed refugia also necessitate consideration of more complex scenarios, such as secondary contact of expanding refugial populations being limited or prevented by river barriers. Unlike the forest refugia hypothesis, there is no expectation for population contraction or expansion under the riverine barrier hypothesis.

Similar to riverine barriers, regions of high elevation can also function as biogeographic barriers, resulting in similar predictions for population divergence as in the riverine barrier model. The lower elevation Gulf of Guinea forests is partially interrupted by a series of high-elevation volcanic ranges that comprise the Cameroon Volcanic Line (CVL), a region of high species richness and endemism (Amiet, 2012; Chirio & LeBreton, 2007; Oates et al., 2004; Figure 1). The CVL and surrounding lower elevation regions are proposed as a major forest refugium (Maley, 1996), a pattern supported in part by the elevated levels of genetic diversity within the region (Budde et al., 2013; Hardy et al., 2013); moreover, the volcanic peaks may also restrict gene flow between adjacent lowland forests. The CVL has been hypothesized to act as both refugium and elevational barrier (Budde et al., 2013; Hardy et al., 2013), yet previous studies lacked the dense population sampling necessary for studying fine-scale patterns of gene flow across the CVL.

The Lower Guinea region contains several ecotones, which can promote parapatric speciation through disruptive selection across ecological gradients (Moritz et al., 2000; Smith et al., 1997). Multiple vertebrate taxa exhibit a pattern of selection-driven phenotypic or genetic divergence across the conspicuous forest-savanna ecotone in central Cameroon, lending support to the ecological gradient hypothesis (Freedman et al., 2010; Mitchell, Locatelli, Clee et al., 2015; Mitchell, Locatelli, Ghobrial et al., 2015; Smith et al., 1997, 2005, 2011). Although this prominent ecotone is the best-characterized example, additional habitat gradients may influence phenotypic and lineage divergence within strictly forest-dwelling taxa of LG. The largely continuous forests to the south exhibit a longitudinal E-W gradient in precipitation, transitioning from high rainfall and low seasonality in the Atlantic Equatorial coastal forests to the more seasonal Congolian forests further inland (Heuertz et al., 2014; Olson et al., 2001). Another important physioclimatic feature of the Lower Guinea region is the climate hinge, a north–south seasonal inversion occurring at latitude 2°N, close to the southern border of Cameroon (Leroux, 1983; Vande Weghe, 2004). Both of these habitat features coincide with population structuring in plants (Debout et al., 2011; Duminil et al., 2010; Faye et al., 2016; Hardy et al., 2013; Heuertz et al., 2014), with potential phenological differences in plant reproduction occurring across the climatic hinge (Debout et al., 2011; Hardy et al., 2013). Under the ecological gradient hypothesis, selection can drive phenotypic divergence despite high gene flow, but

genetic divergence is expected to increase across the ecotone over time. Although population boundaries coinciding with ecological gradients offer indirect evidence of this process, more convincing evidence comes from phenotypic divergence and reproductive isolation associated with differences in habitats or through the detection of divergence with gene flow. Although these forest-specific physioclimatic features appear to play a strong role in driving divergences among plants, their effects on vertebrate taxa remain poorly studied.

### 1.2 | Study system and model-testing framework

The forest refugia, riverine barrier, elevational barrier and ecological gradient hypotheses each describe alternative scenarios for the formation of species and intraspecific population diversity. Considering the spatial extent and potential interactive effects of these evolutionary processes, it is best to study them in an integrative molecular framework using lineages broadly distributed throughout LG. Across vertebrate groups, amphibians display comparatively high diversity in LG (Amiet, 2012; Lawson, 1993; Oates et al., 2004), yet they are vastly understudied and the mechanisms driving their diversification are particularly poorly known. The LG endemic Gaboon forest frog, *Scotobleps gabonicus* Boulenger, 1900, represents a monotypic relict genus that likely diverged from other taxa in the family Arthroleptidae during the Oligocene (Portik & Blackburn, 2016). This stream-breeding species inhabits lowland rainforest and secondary forest, and although it is distributed throughout the CVL, it has not been recorded above 1,000 m elevation (Amiet, 1975; Portik et al., 2016). Given that *S. gabonicus* is a relatively small terrestrial species (<75 mm SVL) with low dispersal rates, we expect that phylogeographic patterns will reflect historical processes.

Using a combination of molecular data, including genomewide SNPs and mitochondrial DNA (mtDNA), we aim to identify discrete populations of *S. gabonicus*, reconstruct the phylogenetic relationships among these populations, and perform demographic modelling to compare alternative scenarios of population divergence. Although several phylogeographic studies in LG have relied solely on the arrangement of landscape features and population boundaries as evidence for particular scenarios of diversification, here we treat these patterns as indirect evidence requiring further evaluation. We assess whether population boundaries of *S. gabonicus* are located between ENM-inferred refugia, along major rivers, elevational barriers or across ecological gradients, and whether spatial patterns of effective migration and genetic diversity are consistent with predictions of these divergence scenarios. The latter approach is most salient for corroborating forest refugia, as we expect high migration and genetic diversity within refugial locations, and low genetic diversity in regions of expansion. We reconstruct the phylogenetic relationships among inferred populations and use them as a guide to examine the joint-demographic history of populations with demographic models derived from alternative divergence scenarios. These demographic models can be broadly classified either as allopatric, involving divergence in the absence of gene flow, or as parapatric, involving divergence with gene flow. Population divergences across riverine and elevational barriers fall under a general allopatric model of population splitting with no gene flow (with the exception of river headwaters, where migration is predicted) and without clear expectations for population size change or secondary contact. Forest refugia also involve a model of allopatric divergence, but here we expect the initial period of population isolation to be followed by size expansion and secondary contact. Finally, we represent divergence across ecological gradients with parapatric models involving either continuous gene flow, or historical gene flow accompanying divergence, followed by isolation. With this model-testing framework, we attempt to identify the mechanisms of population divergences within *S. gabonicus*, and determine the most prevalent mode for this species. We place our results in the context of other studies of LG taxa and offer directions for improving future studies of diversification in LG and other tropical regions.

## 2 | METHODS

### 2.1 | Sampling

We obtained 84 samples of *Scotobleps gabonicus* from 33 localities throughout its known range across southeastern Nigeria, Cameroon, Equatorial Guinea, Gabon and Republic of the Congo (Figure 1). Tissue samples (including liver, muscle or toe clips) were preserved in 95% ethanol or RNA later. Specimen vouchers are deposited in natural history museum collections listed in File S1.

### 2.2 | Molecular data

We extracted genomic DNA using QIAGEN DNeasy Blood and Tissue Kits and sequenced a continuous region of ~1,850 bp of mtDNA containing portions of the 12S and 16S ribosomal DNAs and intervening tRNA for Valine using standard primer pairs (12L1–16Sh, 12Sm–16Sa, 16Sc–16Sd; Darst & Cannatella, 2004). After editing the resulting sequences in GENEIOUS v.6.1 (Biomatters Ltd.), we generated a multiple alignment of 1,865 bp (including indels) for 84 samples in MAFFT v.5 using default parameters (Katoh & Kuma, 2002). The mtDNA sequences generated for this study are deposited in GenBank (accession numbers: MF11909–11993).

We collected ddRADseq data for 72 of the same samples following the protocol described by Peterson, Weber, Kay, Fisher, and Hoekstra (2012). For each sample, a total of 500 ng of genomic DNA was double-digested using 20 units of *Sbf*I (restriction site 5′-CCTGCAGG-3′) and *Msp*I (restriction site 5′-CCGG-3′) in a single reaction with the manufacturer's recommended buffer (New England Biolabs) for 5 hr at 37°C. The digested DNA was bead purified before ligating barcodes and Illumina adapters (Peterson et al., 2012). Samples receiving the same adapter sequences were pooled in equal concentrations and then size-selected (between 415 and 515 bp after accounting for adapter length) on a Blue Pippin Prep size selector (Sage Science). We used proofreading *Taq* and Illumina's indexed primers for the final library amplification step. An Agilent 2200 TapeStation was used to determine the final fragment size

distribution and concentration of each pool, and qPCR was performed to determine sequenceable library concentrations before multiplexing equimolar amounts of each pool for sequencing on one Illumina HiSeq 2500 lane (50-bp, single-end reads).

The raw Illumina reads were demultiplexed using STACKS v1.35 (Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011; Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013), and restriction site overhangs were subsequently removed using the fastx_trimmer module of the FASTX-TOOLKIT (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Sequencing statistics were inspected on a per sample basis using FASTQC v0.10.1 (https://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/). Loci were created from the demultiplexed trimmed reads using USTACKS with a minimum of 5X coverage and a maximum of 2-bp mismatches between the 39-bp fragments (representing a 5% nuclear divergence). A catalogue of loci was created from all samples using CSTACKS, and loci identifications were assigned to samples using SSTACKS. The POPULATIONS program was then used to generate alleles for loci present in 70% of all individuals, which resulted in the retention of 11,958 loci. We chose this threshold after examining alternative thresholds ranging from 50 to 100% (by increments of 10%). The haplotypes file was then filtered for invariant loci ($n$ = 2,974), nonbiallelic loci ($n$ = 11) and loci containing any number of samples with greater than two alleles ($n$ = 1,343). For loci containing multiple SNPs, we randomly chose a single SNP to be used for all subsequent analyses. Finally, any samples missing data for more than 50% of loci were removed, and the resulting haplotypes file was used to generate input files for subsequent analyses. After completing all filtering steps, our final SNP data set consisted of 66 samples and 7,633 loci. Using this final set of loci, we calculated the average sequencing depth per locus across all samples as well as the average sequencing depth of each sample (File S2). Python scripts to automate this workflow with STACKS, perform custom filtering steps, generate input files and calculate sequencing depths have been made available at: github.com/dportik/Stacks_pipeline.

## 2.3 | Population discovery

We determined the number of discrete populations present across the sampled range of *S. gabonicus* using a combination of Bayesian and likelihood clustering analyses, and multivariate methods. We used STRUCTURE v2.3.4 (Falush, Stephens, & Pritchard, 2003; Pritchard, Stephens, & Donnelly, 2000) to investigate the number of population clusters and potential admixture between populations in our data set using MCMC. Hierarchical analyses were performed for 10 runs per K, up to a maximum of 6 populations, and used the admixture model with a burn-in of 10,000 steps followed by 100,000 steps. We summarized our results using STRUCTUREHARVESTER (Earl & vonHoldt, 2012) and evaluated the likely number of populations based on inspection of likelihood plots and following Evanno, Regnaut, and Goudet (2005). We complemented these hierarchical analyses using the maximum-likelihood approach of ADMIXTURE (Alexander, Novembre, & Lange, 2009) and performed five replicate analyses to evaluate up to six populations. To assess the best value of K, we performed 10-fold cross-validation and determined the K value with the lowest cross-validation error. We also investigated the number of discrete populations using a discriminant analysis of principal components (DAPC) with ADEGENET v2.0.0 (Jombart, 2008; Jombart & Ahmed, 2011). A maximum of 20 clusters were investigated using the *k*-means algorithm. The preferred number of clusters was evaluated using BIC scores, and we ultimately explored a range of three to five clusters to describe using DAPC. To minimize overfitting, an initial DAPC was used to find the *a*-score for each set of clusters and this value was used to select the number of principal components to retain in a subsequent re-analysis (Jombart, 2008; Jombart & Ahmed, 2011). Group membership probabilities were then examined for each cluster.

## 2.4 | Spatial molecular variation and connectivity

We visualized spatial patterns in genetic diversity and regions of gene flow using the program EEMS (ESTIMATED EFFECTIVE MIGRATION SURFACES) (Petkova, Novembre, & Stephens, 2016). This method uses sampling localities and pairwise dissimilarity matrices calculated from SNP data to identify regions where genetic similarity decays more quickly than expected under isolation by distance. A user-selected deme size determines the geographic grid size and resulting set of migration routes, and the expected dissimilarity between two samples is approximated using resistance distance. These estimates are calculated without the need to include environmental variables or topographic information and are subsequently interpolated across the geographic space to provide a visual summary of observed genetic dissimilarities, including regions with higher and lower gene flow than expected. To use the EEMS pipeline, we converted our filtered STACKS haplotype file to plink format and translated it to a bed file format using PLINK (Chang et al., 2015). The bed file was used to calculate a dissimilarity matrix with the BED2DIFFS program included in the EEMS package (Petkova et al., 2016). An outer coordinate file was generated using the polyline method in the GOOGLE MAPS API V3 TOOL (http://www.birdtheme.org/useful/v3tool.html). We tested several deme sizes (150, 300, 500, 700) using the RUNEEMS_SNPS version of EEMS. For each deme size, we ran three independent analyses, with a burn-in of 1,000,000 and MCMC length of 20,000,000. The results were combined across the three independent analyses using the REEMSPLOTS R package (Petkova et al., 2016), and convergence of runs was assessed. Using this package, we plotted the geographic distance and genetic dissimilarity across demes and generated surfaces of effective diversity ($q$) and effective migration rates ($m$).

## 2.5 | Phylogenetic relationships and divergence dating

We estimated the phylogenetic relationships and associated divergence times of individuals and populations independently for our mtDNA and SNP data sets. We conducted Bayesian divergence-dating analyses with our concatenated mtDNA data set (12S, 16S) using

BEAST v1.8 (Drummond, Suchard, Xie, & Rambaut, 2012). We performed analyses using a coalescent tree prior, specifically the constant size growth prior, with a strict molecular clock calibrated with a 2% per Myr rate of divergence (Crawford, 2003). The analysis was run twice with 10,000,000 generations with sampling every 5,000 generations, producing a total of 2,000 trees. Runs were assessed using TRACER VERSION 1.6 (Rambaut & Drummond, 2009) to examine convergence. A burn-in of 10% was discarded, and maximum clade credibility (MCC) tree with median heights was created from the remaining 1,500 trees.

We investigated the evolutionary relationships among detected populations using the SNP data in a coalescent framework with SNAPP v1.3 (Bryant, Bouckaert, Felsenstein, Rosenberg, & RoyChoudhury, 2012) implemented in BEAST2 v2.4 (Bouckaert et al., 2014). The SNAPP model is based on the coalescent process and can accommodate incomplete lineage sorting; however, it assumes a lack of gene flow (Bryant et al., 2012). We applied additional missing data filters to our unlinked biallelic SNP data set, requiring loci to be present in 90% of the samples and requiring individuals to have at least 80% data, resulting in 2,466 SNPs (1,520 of which were retained in the analyses). To reduce run times, we subsampled each population to include 3–6 representatives, for a total of 25 individuals. We estimated the mutation rates ($u$ and $v$) from the data (0.997 and 1.002, respectively) and set the birth rate ($\lambda$) of the Yule prior to 25. We performed two independent runs with a chain length of one million generations, sampling every thousand generations. Runs were assessed using TRACER VERSION 1.6 (Rambaut & Drummond, 2009) to examine convergence, and tree topologies and node heights were visualized using DENSITREE (Bouckaert, 2010). We converted the resulting branch lengths of our SNAPP analyses to units of time using a mutation rate of $1 \times 10^{-8}$ (Lynch, 2010), but recognize this rate may be inaccurate because it was not obtained from amphibians.

## 2.6 | Joint-demographic history

To investigate hypotheses of alternative demographic histories of populations, we used the diffusion approximation method of $\delta a \delta i$ to analyse both two-dimensional and three-dimensional joint site frequency spectra (2D-JSFS, 3D-JSFS) (Gutenkunst, Hernandez, Williamson, & Bustamante, 2009). We lacked outgroup sequence information for our data and therefore created folded JSFS. We examined 14 alternative 2D models (represented visually in File S3) that differ in assumptions related to migration rates, periods of isolation, and population size changes. We performed 2D comparisons of the main northern–southern split and all geographically neighbouring populations. To account for missing data and to maximize the number of segregating sites used in analyses, these populations were projected down to smaller sample sizes for all 2D-JSFS (Northern: 40 alleles; Southern: 18 alleles; CVL North: 30 alleles; CVL South 18 alleles; Cross River: 14 alleles; North Coast: 10 alleles; South Coast/East Gabon: 12 alleles). Initial optimizations were performed by generating 50 sets of threefold randomly perturbed parameters and optimizing each parameter set using the Nelder–Mead method (*optimize_log_fmin*), running each optimization algorithm step for a maximum of 100 iterations. Each optimized parameter set was used to simulate the 2D-JSFS, and we used a multinomial approach to estimate the log-likelihood of the 2D-JSFS given the model. Parameters from the best scoring replicate were used as starting values for a second round of twofold perturbed parameters, and 50 replicates, and the parameter values from the best replicate were subsequently used to generate onefold perturbed starting parameters for a final set of 100 replicate optimizations to estimate the log-likelihood of the 2D-JSFS given the model.

We also explored the joint-demographic history of the entire Northern region by constructing a 3D-JSFS of the CVL South (14 alleles), CVL North (30 alleles) and Cross River (18 alleles) populations. We aimed to identify the main patterns of isolation and gene flow among these three populations. To narrow down the enormous number of possible models that capture variation in these factors, we created an informative model set based on the order of population divergence recovered in our SNAPP analyses and results of the 2D analyses. We focused on testing three specific variations of the forest refugia model, but also included more general 3D models to broaden our comparison, resulting in 11 models total (represented visually in File S4). We followed a similar approach as the 2D analyses, and for all models, we initially optimized threefold perturbed random starting parameters for up to 10 iterations per step across 20 replicates. The best scoring replicate parameters were used to create twofold perturbed starting parameters and optimized using a maximum of 10 iterations per algorithm step across 40 replicates. Finally, the best scoring replicate parameters of each model were used to create a onefold perturbed starting set of parameters and optimized using a maximum of 50 iterations per algorithm step across 50 replicates. Across all analyses, we used the optimized parameter sets of each replicate to simulate the 3D-JSFS, and the multinomial approach was used to estimate the log-likelihood of the 3D-JSFS given the model.

By keeping only a single SNP per RAD locus, we assumed our loci are unlinked, and the log-likelihood values returned are the true likelihood values rather than composite likelihoods resulting from linked SNPs. Models were compared using the Akaike information criterion (AIC), and the replicate with the highest likelihood for each model was used to calculate AIC scores, $\Delta$AIC scores and Akaike weights ($\omega_i$) (Burnham & Anderson, 2002). We did not transform parameters into biologically meaningful estimates because our primary aim was to perform model selection, and parameter values should ideally be estimated using a bootstrapping procedure to obtain confidence intervals (Gutenkunst et al., 2009). We provide an estimate of $\theta$, or the effective mutation rate of the reference population, which here corresponds to the ancestral population ($\theta = 4N_{ref}\mu l$, where l is the total length of sequenced region SNPs were ascertained from). Python scripts that define 2D and 3D models, perform model fitting and execute plotting functions are available at: github.com/dportik/dadi_pipeline.

## 2.7 | Ecological niche modelling

We compiled all available museum locality information for *Scotobleps gabonicus* using VERTNET (Constable et al., 2010), which resulted in an additional 161 records that were used for ecological niche models (ENM) (File S5). All ENMs were built with the 19 freely available bioclimatic layers from Worldclim.org (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) using MAXENT (Phillips, Anderson, & Schapire, 2006) implemented in SDMTOOLBOX v1.1c (Brown, 2014). To choose optimal model parameters, we tested five combinations of feature classes (Linear; Linear and Quadratic; Hinge; Linear, Quadratic and Hinge; Linear, Quadratic, Hinge, Product and Threshold) and regularization multipliers from 0.5 to 5, in increments of 0.5, and the model which had the lowest test omission (low omission rate), highest discrimination ability (high area under the curve, AUC) and lowest complexity, in that order, was used (Brown, 2014; Shcheglovitova & Anderson, 2013). A minimum convex polygon defined by a 100 km buffer around each occurrence record was used to sample background data, and occurrence records were spatially rarefied (10 km distance) and partitioned into three subsets of data for background testing and training to reduce the effects of spatial autocorrelation and overfitting (Boria, Olson, Goodman, & Anderson, 2014; Hijmans, 2012; Veloz, 2009). Models were projected to past climate data from the mid-Holocene (~6 kybp), the last glacial maximum (LGM; ~21 kybp) and the last interglacial (LIG; ~120 kya), and summed to estimate continuous stability maps (Devitt, Devitt, Hollingsworth, McGuire, & Moritz, 2013; Yannic et al., 2014). We constructed ENMs for the entire range and for two subsets of the data based on our molecular results: all populations north of the Sanaga River and all populations south of the Sanaga River. The decision to construct northern and southern models was based on the detection of divergent niches in chimpanzees similarly distributed across the Sanaga River, although their distribution extends much further east into the forest-savanna ecotone (Clee et al., 2015). We calculated niche overlap (Schoener's *D* statistic) between the northern and southern populations using continuous model outputs in the ENMTOOLS R package (Broennimann et al., 2012; Warren, 2016; Warren, Glor, & Turelli, 2008).

# 3 | RESULTS

## 3.1 | Population inference and spatial molecular patterns

The Bayesian population clustering analysis based on 7,633 unlinked SNPs sampled across the range of *Scotobleps gabonicus* resulted in the initial detection of two main populations (Figure 2a). The geographic distributions of these populations are allopatric and correspond to a Northern and Southern group with a latitudinal break across central Cameroon (Figure 1). We detected structuring within each of these main populations, and inspection of the likelihood plots and the highest $\Delta K$ in each analysis indicates support for three additional clusters in both the Southern and Northern groups. These results are also supported by the maximum-likelihood clustering analyses, for which the three population models exhibited the lowest cross-validation error for both groups and the population assignments of individuals are identical (Figure 2c). In the Northern group, the three populations are allopatric and include a population north of the Cross River in Nigeria and two populations occurring either

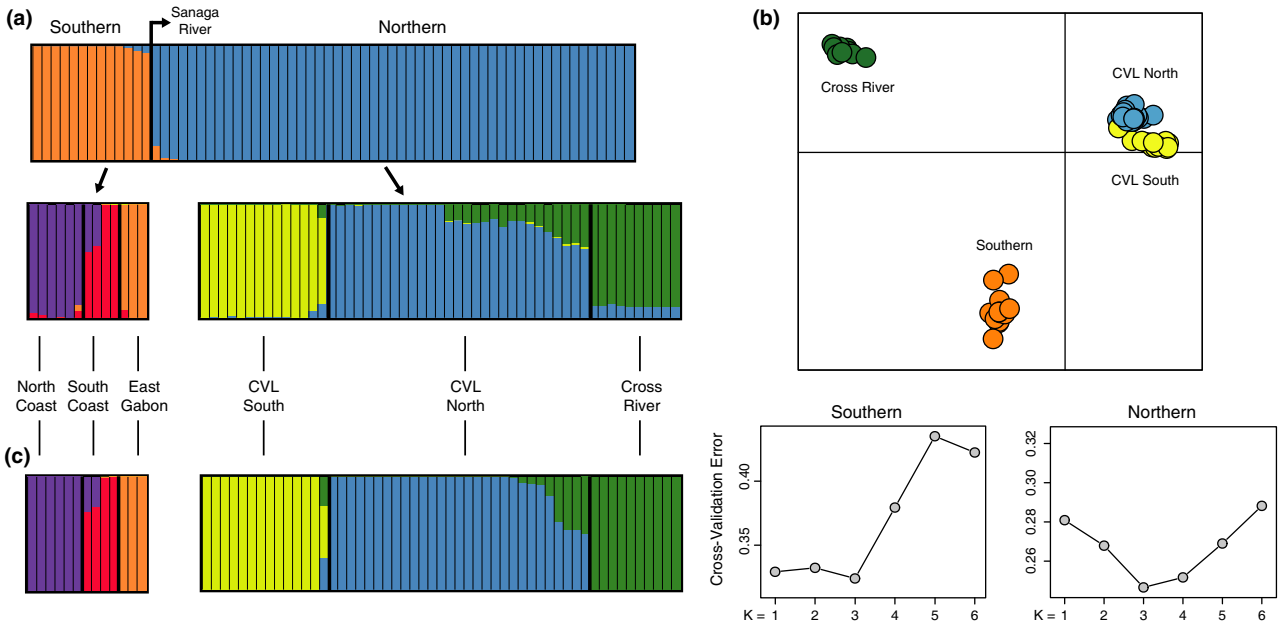


**FIGURE 2** Population assignment results for 66 individuals based on (a) hierarchical Bayesian population clustering using STRUCTURE, (b) discriminant analysis of principal components and (c) maximum-likelihood (ML) population clustering using ADMIXTURE. The cross-validation error plots associated with ML clustering analyses are shown. The spatial distribution of populations is presented with the same colour scheme in Figure 1

north or south of the CVL (Figure 1). There is a pattern consistent with admixture between the Cross River population and the northernmost localities of the CVL North population, particularly because Cross River mtDNA haplotypes are present in populations assigned to the CVL North using SNP data (see below). The Southern group also contains three allopatric populations, located in southwest coastal Cameroon (North Coast), coastal Equatorial Guinea and Gabon (South Coast) and eastern Gabon (Figure 1). The DAPC of the complete data set largely corroborated these results, and based on BIC, scores inferred four discrete populations (Figure 2b). The assignments of individuals matched results from the maximum-likelihood and Bayesian clustering analyses in characterizing the Cross River, CVL North and CVL South, but inferred a single Southern group rather than multiple populations. Under a five-population scenario with DAPC, the North Coast population is shown as distinct from the combined South Coast/East Gabon populations.

The EEMS analyses incorporating smaller deme sizes (<500) merged localities that we regarded as important to separate (particularly across the CVL), and therefore, we only present results from analyses using a deme size of 700. The spatial analyses highlight several barriers to migration resulting from either historical or contemporary patterns of gene flow (Figure 3a). There is evidence for restricted migration between the Cross River population and all neighbouring populations of the CVL North. A clear migration barrier is located at the centre of the CVL, starting inland and following the diagonal chain of high-elevation peaks to the coastline at Mt. Cameroon. A broad area with low migration rates is centrally located on the Sanaga River in Cameroon. To the south, there is evidence for a migration barrier between the North Coast and South Coast populations, and a large area of reduced migration in central Gabon that broadly overlaps the course of the Ogooué River (Figure 3a). Spatial analyses of genetic diversity highlight two main regions of exceptionally high diversity, one that covers a substantial portion of the CVL (but not extending to Mt. Cameroon) and a small region located in coastal southwest Cameroon largely corresponding to the location of the North Coast population (Figures 1 and 3b). Lower-than-expected genetic diversity occurs throughout southern Equatorial Guinea and all of Gabon, in coastal central Cameroon and in the Cross River population in Nigeria.

## 3.2 | Phylogenetic relationships and divergence times

The divergence-dating analysis of mtDNA using a calibrated general mutation rate supports a Pliocene divergence (2.9 Mya) between the Northern and Southern regions (Figure 4). However, the associated confidence interval also includes divergence dates in the early Pleistocene (95% highest posterior density region [HPD] 2.5–3.3 Mya). Additional population structuring is apparent in both regions, and the TMRCA of the Southern clade is 2.6 Mya (95% HPD 2.2–3.0 Mya), and the TMRCA of the Northern clade is 2.7 Mya (95% HPD 2.3–3.1 Mya). These dates are largely overlapping with the initial split between the Northern and Southern regions. There are three well-supported subclades in the Southern region corresponding to the North Coast, South Coast and eastern Gabon populations (Figures 1 and 4). There are also three strongly supported subclades in the Northern region, corresponding to the Cross River (but with some nuclear CVL North individuals), CVL North and CVL South populations (Figures 1 and 4). Divergences between major subclades in the Northern and Southern regions occur in the mid-Pleistocene (1.5 and 1.4 Mya, respectively), and the TMRCA of individual widespread subclades ranges from 0.3 to 0.5 Mya.

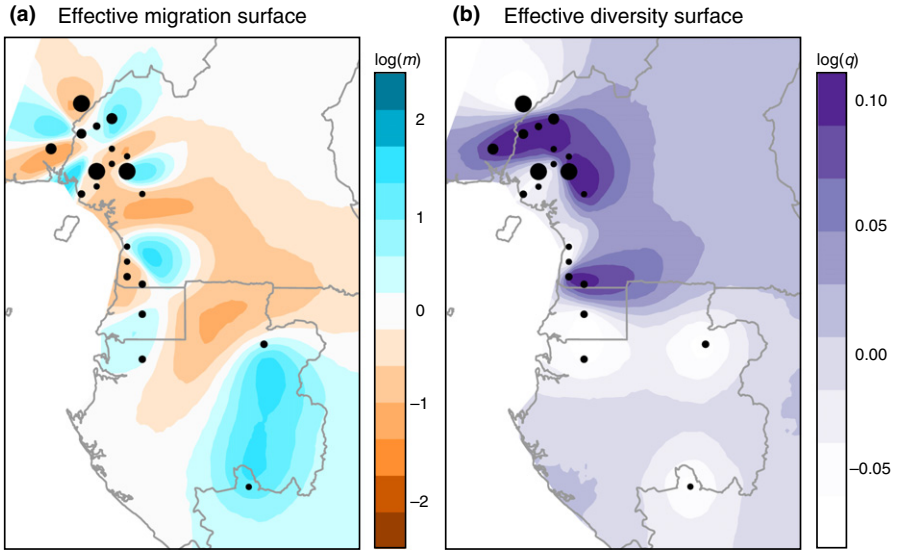


**FIGURE 3** Contour maps representing the posterior mean of (a) effective migration surface and (b) effective diversity surface, for all populations of *Scotobleps gabonicus*. In (a), blue colours represent areas of high migration, or dispersal corridors, whereas orange regions represent areas of low migration, or dispersal barriers. In (b), white colour indicates areas of lower-than-expected genetic diversity, and dark purple coloration represents higher levels of genetic diversity. Size of dots reflects number of samples in a merged locality

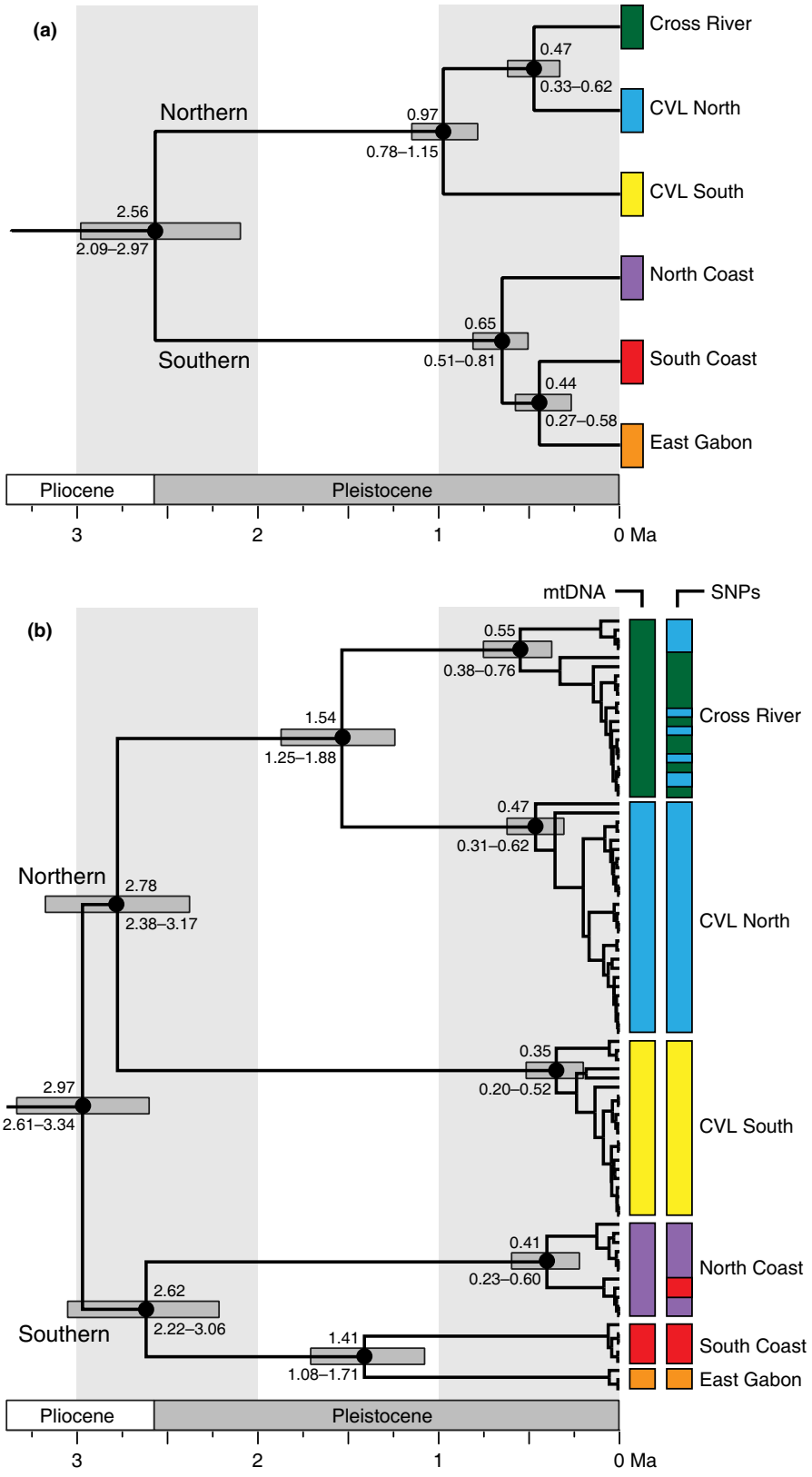


**FIGURE 4** Chronograms of *Scotobleps gabonicus* resulting from (a) Bayesian coalescent analysis of SNP data using SNAPP and (b) a BEAST analysis of mtDNA data calibrated with a mutation rate. Nodes with high support (posterior probability >0.9) are filled black. Median ages are provided above nodes, with 95% highest posterior densities (HPD) below. Error bars representing the 95% HPD are also shown on nodes. Coloured boxes show population assignments matching Figure 1 and are shown in (b) based on results for the mtDNA phylogeny and based on the population clustering analyses of genomewide SNP data

The phylogenetic relationships among populations inferred through Bayesian coalescent analyses of SNPs are identical to those inferred using mtDNA, including a primary division between all Northern and Southern populations (Figure 4). The population relationships within each respective region are also identical to those recovered with mtDNA, and all nodes received high support (>0.95 posterior probability). The divergence time inferred for the split between the Northern and Southern regions is 2.5 Mya (95% HPD

2.1–2.9 Mya), overlapping with the mtDNA estimate; however, there are substantial differences between all other time estimates. The TMRCA estimated for the Northern clade and for the Southern clade (0.9 and 0.6 Mya, respectively) are considerably younger than the dates obtained with mtDNA data (2.7 and 2.6 Mya, respectively) and do not overlap with the initial split between the Northern and Southern regions (Figure 4). This pattern of younger time estimates is also true for the additional divergences within each region (Figure 4). These discrepancies are almost certainly due to a violation of a primary assumption of the SNAPP model—that gene flow is not present among lineages. Violations of this assumption in an alternative coalescent Bayesian model (*BEAST; Heled & Drummond, 2010) consistently produced inflated estimates of population size and underestimated divergence times, the latter approaching zero under high migration rates (Leaché, Harris, Rannala, & Yang, 2014). Although a parallel study has not yet been conducted using the SNAPP model, the presence of gene flow is expected to overestimate population sizes and compress divergence times in a similar manner. We find evidence for gene flow across all of our population comparisons (see below) and therefore treat these nuclear divergence dates with caution.

### 3.3 | Joint-demographic history

We present detailed results of our demographic inferences in Table 1, Files S6 and S7, and visually in Figures 5 and 6. For both the 2D and 3D analyses, the variation in log-likelihood scores observed across the initial optimizations of highly perturbed random starting parameters decreased during subsequent analyses incorporating less perturbed parameters from previous replicates, producing more consistent log-likelihoods. Models with the lowest scoring log-likelihoods include no divergence or complete isolation. Refugial models provided the best fit to our observed 2D-JSFS for three of the population pairs, including the Northern and Southern, CVL North and CVL South, and North Coast and South Coast/East Gabon comparisons (Table 1). A refugial model involving expansion and secondary contact with symmetric migration is supported as the best fit for the Northern and Southern comparison ($\Delta$AIC = 13.6, $\omega_i$ = 0.99) (Figure 5a). A refugial model involving expansion and secondary contact with asymmetric migration is strongly supported as the best fit for the CVL North and CVL South comparison ($\Delta$AIC = 94.2, $\omega_i$ = 0.99) (Figure 5b). The historical and contemporary effective population size of the CVL North is larger than estimates for the CVL South, and during secondary contact, migration is higher from CVL North to the CVL South population (Table 1). Within the Southern region, a refugial model involving expansion and secondary contact with asymmetric migration is supported as the best fit for the North Coast and South Coast/East Gabon comparison ($\Delta$AIC = 6.5, $\omega_i$ = 0.94) (Figure 5d). For the Cross River and CVL North comparison, the data are explained equally well by (i) a refugial model with no size change and secondary contact ($\omega_i$ = 0.53) and (ii) a parapatric model of divergence with continuous gene flow and asymmetric migration ($\omega_i$ = 0.43) (Figure 5c). In both models, the Cross River population has a substantially smaller effective population size relative to the CVL North. Given the pattern of population expansion consistently detected in the CVL North in other

**TABLE 1** Demographic models and parameter values (unscaled) for pairwise population comparisons. A summary of all models evaluated is provided in File S6.

| Comparison | Model | Log-l | theta | nu1 | nu2 | nu1a | nu2a | nu1b | nu2b | m12 | m21 | T1 | T2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Northern, Southern | Split, Size Change, Secondary Contact, Symmetrical Gene Flow | −439.9 | 574.0 | – | – | 0.259 | 0.247 | 3.121 | 1.291 | 0.049 | m12 | 0.302 | 0.556 |
| CVL North, CVL South | Split, Size Change, Secondary Contact, Asymmetrical Gene Flow | −463.3 | 367.3 | – | – | 0.338 | 1.331 | 4.528 | 1.931 | 0.108 | 1.084 | 0.325 | 0.870 |
| Cross River, CVL North | Split, Secondary Contact, Asymmetrical Gene Flow | −378.0 | 288.0 | 0.350 | 6.746 | – | – | – | – | 1.272 | 0.221 | 0.495 | 0.402 |
| N. Coast, S. Coast + East Gabon | Split, Size Change, Secondary Contact, Asymmetrical Gene Flow | −144.8 | 305.6 | – | – | 1.732 | 3.029 | 3.480 | 2.357 | 0.202 | 0.645 | 2.229 | 0.307 |

Abbreviations are as follows: Theta ($4N_{ref}\mu l$), the effective mutation rate of the reference population (ancestral population); nu1, nu2, effective population sizes under the constant population size model; nu1a, nu2a, effective population sizes before instantaneous size change; nu1b, nu2b, effective population sizes after instantaneous size change; m12, migration rate from population two to population one; m21, migration rate from population one to population two; T1, unscaled time between population split and the present; T2, unscaled time of secondary contact or isolation interval.

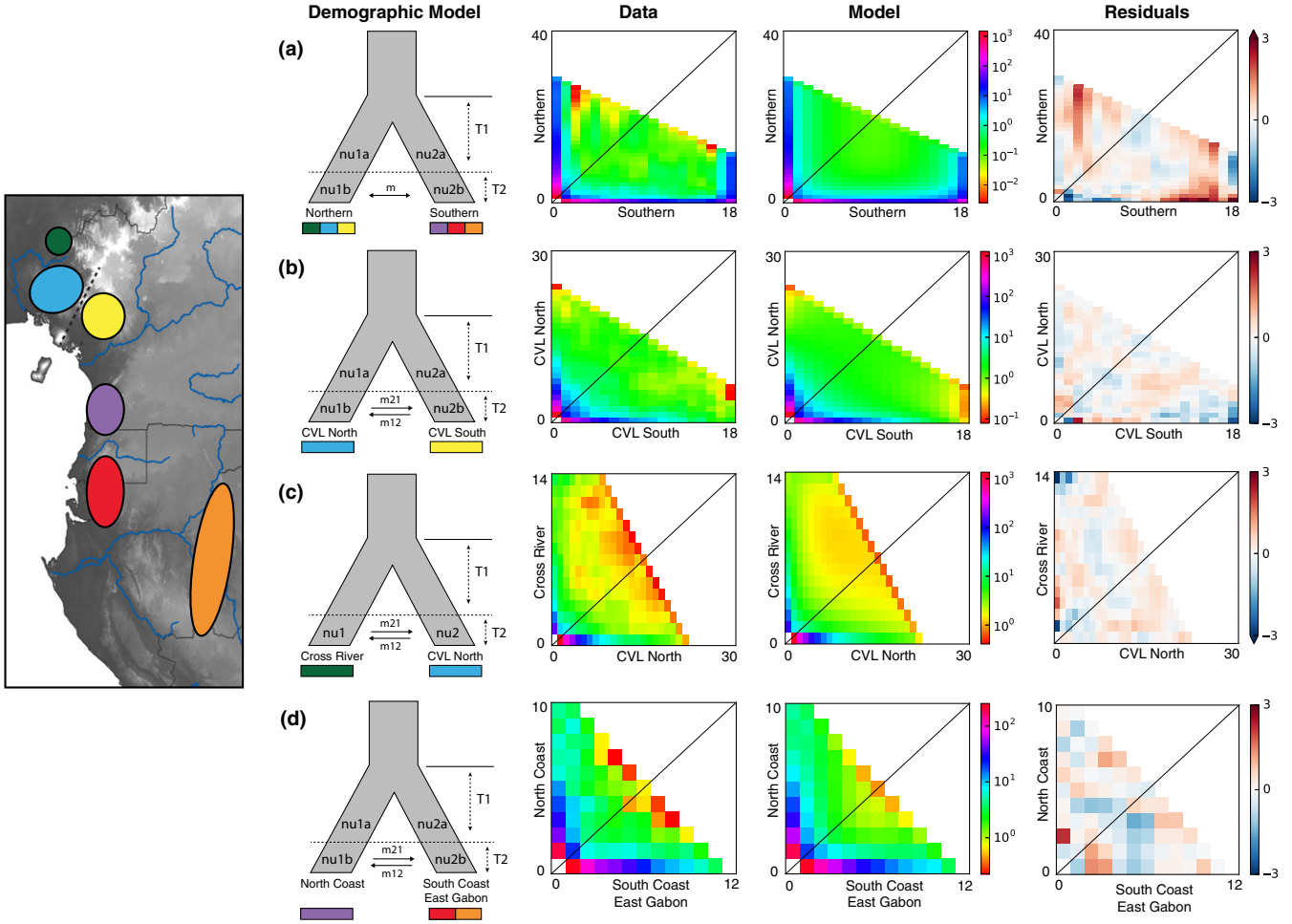


**FIGURE 5** Results of population genetic model comparisons using the two-dimensional site frequency spectrum (2D-SFS) between population sets including (a) Northern and Southern, (b) CVL North and CVL South, (c) the Cross River and CVL North and (d) the North Coast and South Coast/East Gabon. A simplified graphic of the best-fit model is depicted, along with comparisons of the 2D-SFS for the data, the model and resulting residuals. Additional models and parameter values are provided in Table 1 and File S6

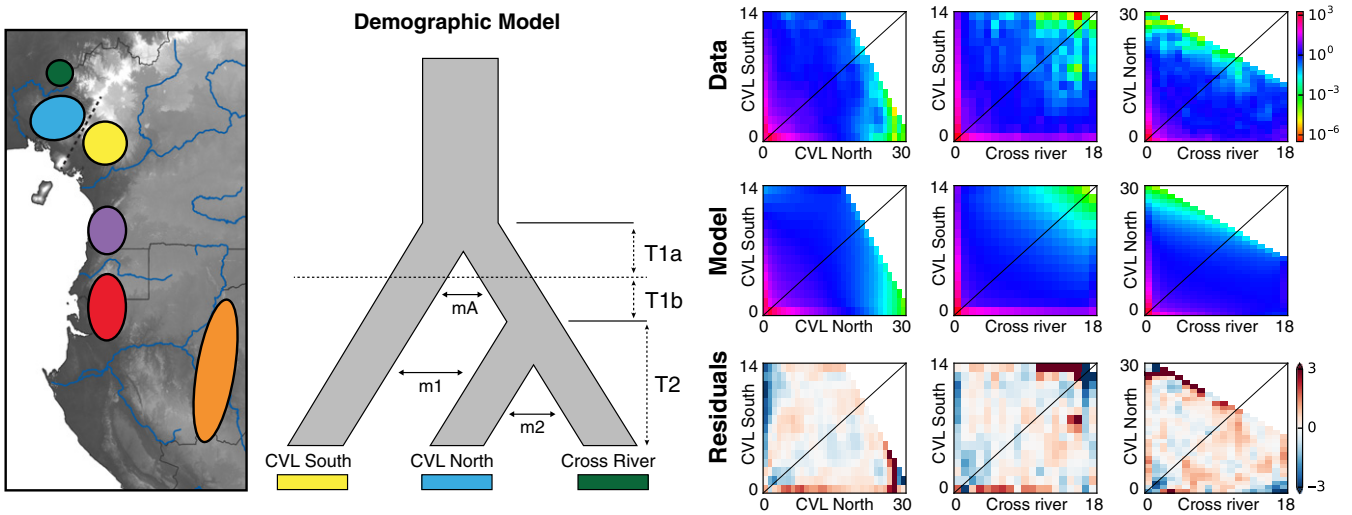


**FIGURE 6** Results of population genetic model comparisons using the three-dimensional site frequency spectrum (3D-SFS) between the Cross River, CVL North and CVL South populations. A simplified graphic of the best-fit model is depicted, along with comparisons of the pairwise 2D-SFS for the data, the model and resulting residuals for each population combination. Parameter values for this and additional models are provided in File S7

comparisons, we modified these two models to allow for instantaneous size change only in the CVL North, as an alternative to enforcing size change in both populations. These modified models did not explain additional features of the 2D-JSFS and were ranked below the models including size change for both populations (AIC = 775.7 and 778.6, respectively).

The 3D-JSFS analysis of the Northern region corroborated results of the pairwise 2D-JSFS analyses, but clarified gene flow occurred between the CVL South and the ancestral population of the CVL North and Cross River (Figure 6, File S7). The best model demonstrates a refugial scenario for the CVL South and this ancestral population, where an initial divergence in isolation was followed by secondary contact (model "Refugia 1", $\Delta$AIC = 24.8, $\omega_i$ = 0.99). The divergence between the CVL North and Cross River then followed and occurred with gene flow. In the model set examined, we also included a refugial model in which divergences among respective populations occurred in isolation, followed by a single period of simultaneous secondary contact (File S4). This particular model was intended to represent a single event (forest expansion) facilitating secondary contact among neighbouring populations simultaneously, yet it was not favoured above our top model. Taken together, the 2D-JSFS and 3D-JSFS demographic modelling of the Northern region suggests temporally distinct processes drove the divergence events between the three genetically distinct populations.

## 3.4 | Ecological niche models

The ENM we generated for the entire species had high discrimination ability, with an AUC statistic of 0.88 and annual precipitation (bio12) having the highest gain (most useful information for the model). The ENMs for the Northern and Southern regions had moderately high AUC statistics (0.74 and 0.83, respectively), and precipitation during the coldest quarter (bio19) contained the most useful information for both models. Niche overlap calculations between the Northern and Southern ENMs indicated a moderate amount of suitable habitat overlap (Schoener's $D$ = 0.76). Based on ENMs generated for the entire species, Northern region and Southern region, we identified concordant areas that are predicted to have remained climatically stable and suitable (Figure 7, File S8). The single species model predicted four main regions of habitat stability: a region in coastal central Nigeria that is outside the currently known distribution of *S. gabonicus*, a region surrounding the CVL in western Cameroon, a region between the Ngovayang Massif and coastline in southwest Cameroon that extends south in to Equatorial Guinea and a region located in southeast Gabon (Figure 7). Two of the stability zones in Cameroon are close to the classical refugia proposed by Maley (1996), but it is worth noting the stable habitat is not primarily located on the higher elevation regions or peaks of the CVL (Figure 7: refugium 1) or Ngovayang Massif (Figure 7: refugium 2), but rather mainly in the surrounding mid- to low-elevation regions. Our

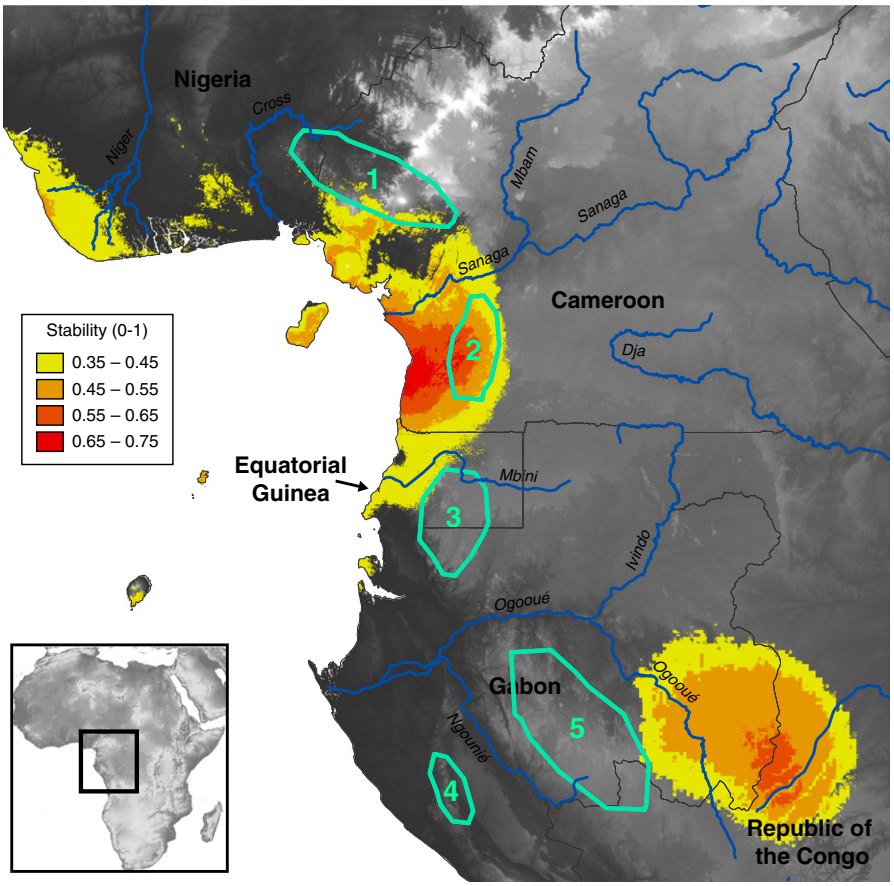


**FIGURE 7** Stability surfaces representing regions of persistent suitable habitat across LGM and current climate regimes. Red colour represents the highest habitat stability inferred. Putative refuge areas identified by Maley (1996) are outlined in green—1: the expanded Cameroonian Volcanic Line; 2: Ngovayang and surrounding massifs; 3: Monts de Cristal; 4: Monts Doudou; and 5: Massif du Chaillu

results indicate stable habitat associated with the CVL includes the area extending from the base of Mt. Kupe to Mt. Cameroon at the coastline (Figure 7).

## 4 | DISCUSSION

Our study uses a combination of spatially explicit genetic analyses and demographic model selection to identify the mechanisms of diversification occurring within tropical forests. We applied our model-testing framework to a forest frog in the Lower Guinea region of tropical Africa to determine whether the formation of intraspecific populations resulted from allopatric scenarios such as forest refugia, rivers and elevational barriers or from divergence with gene flow across ecological gradients. We find allopatric models consistent with the forest refugia hypothesis explain a majority of population divergences in *Scotobleps gabonicus* and find no support for parapatric models of diversification associated with forest-specific ecological gradients. However, based on the observed genetic patterns, we find evidence for potential synergisms between landscape features and historical refugia, including the Sanaga River and elevational barriers such as the Cameroonian Volcanic Line. Below, we summarize these major findings on intraspecific diversification in Lower Guinea and offer a perspective for future work on diversification in the tropics.

### 4.1 | Forest refugia drove simultaneous divergences in Lower Guinea

Our investigation uncovered six genetically distinct and allopatric populations of the Gaboon forest frog, with an initial main divergence occurring between the Northern and Southern regions. Additional population splitting in each respective region occurred nearly simultaneously with this initial divergence in the Late Pliocene during a time of significant climate change (Figure 4b). These results indicate a single event may be underlying these population-splitting events, such as shifts in forest cover. Based on reconstructions of paleovegetation and current patterns of species richness and endemism, several Pleistocene forest refugia have been proposed for LG that vary in their location and extent (Anhuf et al., 2006; Cowling et al., 2008; Maley, 1996; deMenocal, 2004). An important consideration is that the locations of areas of habitat stability may vary greatly between species due to differences in their ecological traits. Because the geographic positions of forest refugia directly guide predictions about spatial genetic diversity and gene flow, the identification of species-specific climatically stable regions should ideally be used to inform phylogeographic hypotheses and evaluate the relevance of classically defined refugia (i.e., Maley, 1996). The three major areas of climatic stability that we inferred for *S. gabonicus* are restricted primarily to western Cameroon and southeast Gabon, which supports the presence of forest refugia for this species (Figure 7).

The putative population boundary between the Northern and Southern region is located near the Sanaga River (Figure 1), which also forms the central boundary between the two inferred Cameroonian forest refugia (Figure 7). Our spatial analyses identified a migration barrier with low genetic diversity overlapping the course of the Sanaga River, between the refugia (Figure 3). We emphasize these phylogeographic patterns, based on the coincidence of population boundaries and landscape features, are consistent with forest refugia or the Sanaga River underlying population divergence. To resolve this confounding issue, we used genomewide SNP data to assess demographic models derived from scenarios of forest refugia and riverine barriers. The set of forest refugia models include demographic components that are not expected for riverine barriers, such as size contraction and isolation, followed by population size increase with or without secondary contact (Haffer, 1997; Hardy et al., 2013; Moritz et al., 2000). Our demographic model selection strongly supports a refugial model for the Northern and Southern regions in which a period of divergence in isolation is followed by a size increase in both populations, with secondary contact and gene flow (Figure 5, Table 1). Importantly, we find evidence for contemporary gene flow across the Sanaga River, demonstrating the river is not a complete barrier to migration and cannot fully explain the historical isolation of these populations. Although a comprehensive geological history of this river is unavailable, preserved offshore fluvial paleodrainage systems indicate the location of the Sanaga River delta has been consistent since the Pleistocene (Ngueutchoua & Giresse, 2010). In addition, divergence estimates of freshwater fish endemic to the Sanaga River suggest these species were established well before the Pleistocene (Day et al., 2013; Goodier, Cotterill, O'Ryan, Skelton, & de Wit, 2011; Pinton, Agnèse, Paugy, & Otero, 2013). It follows that major changes in river course or flow are unlikely to have driven the initial isolation between these populations of *S. gabonicus*. In addition, population size changes are unexpected for divergences across a riverine barrier, but they are consistent with postrefugia expansion events. Although our results indicate forest refugia underlie the primary divergence between Northern and Southern populations of *S. gabonicus*, we propose the Sanaga River could have reinforced historical patterns of divergence by limiting contact between postexpansion populations by serving as a semi-permeable barrier to gene flow.

In the Northern region of LG, we detected two genetically distinct populations of *S. gabonicus* occurring primarily north and south of a region containing several high-elevation peaks in the Cameroonian Volcanic Line (Figures 1, 2 and 4). We inferred a spatial migration barrier along this line (Figure 3), suggesting low historical or contemporary migration between the CVL North and CVL South, yet both populations show high genetic diversity consistent with the refugium inferred for this location (Figure 7). Volcanic activity is not likely the primary driver of divergence because the peaks in the main range of the CVL were established 31–14 Ma (Marzoli et al., 2000), pre-dating the divergence time estimate for this population pair by an order of magnitude (2.4–3.2 Mya; Figure 4). Through our demographic modelling, we found strong support for a refugial model involving divergence in isolation, followed by population size increase and secondary contact (Figure 5, Table 1). Although we detected stable habitat for *S. gabonicus* throughout much of the CVL (Figure 6), the topography

of this region may have presented species-specific challenges for persistence across portions of this refugium. For example, although the reproductive biology of Gaboon forest frogs is poorly known, this species is commonly associated with slowly moving streams in which it is assumed to breed. This type of stream habitat is less prevalent or even absent from the steep slopes associated with the high-elevation sections of the CVL, potentially limiting persistence in otherwise suitable refugial habitat. The line of higher elevation reliefs present within the inferred forest refugium could have effectively divided it, particularly for species with low vagility or specialized habitat requirements. We therefore propose that the interactive effects of forest contractions and the topography of the CVL drove additional population structuring in *S. gabonicus* in this region. Although sampling is coarse and based on organellar loci, we note that some plant and rodent species (Ley et al., 2014; Nicolas et al., 2011) display a potential genetic break across this high-elevation section of the CVL. Similarly, Budde et al. (2013) detected the presence of distinct gene pools of rainforest trees occurring in sympatry at a single locality in the CVL (Mt. Kupe, in our CVL South) and suggested this pattern could have arisen through multiple areas of persistence or from secondary contact. Improved geographic sampling and the collection of genomic data in other codistributed lowland forest taxa can help clarify the role of the CVL as biogeographic barrier in LG, especially during refugial periods.

The remaining simultaneous population divergence occurred between the North Coast and South Coast/East Gabon lineages (Figures 1 and 4). We detected prominent regions of habitat stability in both southern Cameroon and in southeast Gabon, indicating forest refugia could have driven the divergence event between these populations. Our spatial analyses show high diversity present in the southern Cameroon refugium, but low diversity throughout Gabon. A migration barrier is present between the North Coast and South Coast populations, and a broad region of low migration is inferred across Gabon, coinciding with the Ogooué and Ivindo Rivers. In our comparison of the North Coast and combined South Coast and East Gabon populations, the demographic model with highest support is a refugial model involving divergence in isolation, followed by population size increase and secondary contact (Table 1). The gene flow detected between the North Coast and South Coast/East Gabon populations is also supported by a pattern of admixture indicated by the presence of North Coast mtDNA haplotypes in individuals with a South Coast nuclear background (Figures 1 and 4). This suggests a minor role of the Mbini River, if any, in limiting dispersal between these populations. Within the context of refugia, the low genetic diversity found in Gabon would suggest these populations are likely the result of expansion out of the isolated refugium (Figures 3 and 7), and future sampling near the refugial location may indeed reveal higher genetic diversity.

## 4.2 | A relationship between riverine barriers and recent divergences

Near the distribution limit of *S. gabonicus*, we detected a distinct population north of the Cross River in Nigeria that diverged much more recently than the splitting events associated with forest refugia (Figures 1 and 4b) and which has considerably lower genetic diversity than that of adjacent localities near the CVL (Figure 3). We did not detect any regions of habitat stability this far north and therefore did not strongly consider the refugia hypothesis to be a plausible scenario. The results from the 2D and 3D-JSFS modelling demonstrate that refugial models do not fully explain the genomic data; rather, we find stronger evidence for divergence with gene flow between the Cross River and CVL North. The riverine barrier hypothesis predicts allopatric divergence between populations surrounding the river, but opportunities for migration increase at the headwaters (Haffer, 1997; Moritz et al., 2000; Plana, 2004). Given the Cross River population is located at the headwaters, in this context a scenario of divergence with gene flow is actually consistent with the riverine barrier hypothesis. We observed Cross River mtDNA haplotypes present in populations of the CVL North located south of the headwaters, but not in populations along the course of the river in Western Cameroon or at the mouth of the river in southeastern Nigeria (Figure 1). These patterns are consistent with the spatial predictions of gene flow along riverine barriers. Taxa sympatric with Gaboon forest frogs, including chimpanzees, gorillas and several rainforest trees, show no signs of genetic breaks at the Cross River (Anthony et al., 2007; Clifford et al., 2004; Duminil et al., 2010, 2015; Gonder et al., 2011; Mitchell, Locatelli, Clee et al., 2015). Other studies with a larger geographic scope (that include Upper Guinea) have detected genetic breaks between populations in western Cameroon and those located in Nigeria, but west of the Niger River. The large sampling gap between the Niger and Cross Rivers in Nigeria leaves population boundaries unresolved, and whether these breaks coincide with the Cross River remains unknown (Dowell et al., 2016; Fuchs & Bowie, 2015; Fuchs, Fjeldså, & Bowie, 2017; Nicolas et al., 2008, 2010). These patterns could also be the result of postulated sea incursions, which may have been the proximate cause of divergence (Penner, Wegmann, Hillers, Schmidt, & Rödel, 2011). While the role of the Cross River in shaping intraspecific molecular patterns is not yet clear, our results demonstrate gene flow occurs near the headwaters as predicted by the riverine barrier hypothesis. Systematic sampling of additional taxa along the course of the Cross River, from the mouth to the headwaters, will be required to robustly test the riverine barrier hypothesis in future work.

## 4.3 | Moving forward with diversification in the Afrotropics

Our study has demonstrated the inferential power of using genomic data to test among alternative demographic scenarios, and we find strong evidence for forest refugia shaping the contemporary biodiversity patterns of an amphibian species in tropical Africa. These refugial patterns can be further evaluated through comparative work; however, this requires adequate field sampling, molecular data and an appropriate framework for testing alternative models. Several studies have demonstrated divergent haplotypes, rare alleles or

genetic differentiation between two classic Cameroonian refugia (CVL and Ngovayang Massif; Figure 6) (Bohoussou et al., 2015; Brouat, Gielly, & McKey, 2001; Budde et al., 2013; Dauby et al., 2014; Koffi et al., 2011; Ley et al., 2014; Lowe et al., 2010). Unfortunately, this pattern is generally attributed to processes related to forest refugia without testing the plausible alternative scenario of allopatric divergence across the Sanaga River, which bisects these refugial locations. In contrast, mtDNA population boundaries have sometimes been attributed to divergence driven by the Sanaga River without an explicit consideration of forest refugia (Bohoussou et al., 2015; Jacquet et al., 2015; Nesi et al., 2013). We emphasize that the geographic arrangement multiple landscape features in LG is inherently problematic for the interpretation of descriptive phylogeographic patterns (those based on genetic diversity or population boundaries), especially those inferred from organellar loci or limited geographic sampling. Although certain modes of divergence may ultimately prove to be more or less prevalent for specific taxonomic groups (such as a lack of riverine barrier divergences in plants), there is still a need to consider all hypotheses to prevent any biases in conclusions moving forward. The genomic data generated in our study provided the necessary signal to distinguish between demographic models derived from several divergence scenarios, including forest refugia and riverine barriers. Although the model-testing framework we have made available is one of many analytical options for investigating population divergence, future studies should ideally incorporate genomewide sampling of genetic variation to adequately test alternative mechanisms of divergence.

We found forest refugia drove multiple divergences in our focal taxon, but to determine the overall role of forest refugia for diversification in Lower Guinea future studies can assess whether refugia underlie simultaneous divergences within and across additional taxa. The aggregate site frequency spectrum (aSFS), which is built from the independent SNP data sets of multiple taxa, has recently been developed and used to examine the temporal synchronicity of expansion times for populations across species (Xue & Hickerson, 2015). The aSFS is also a promising tool for investigating synchronous divergences of populations within and among species, allowing a statistical assessment of whether species display shared responses to key historical events (Xue & Hickerson, 2015). As population divergences in tropical rainforests are likely a consequence of multifaceted environmental and geological factors, aggregate demographic history models could be used to investigate a wide range of diversification scenarios across sets of taxa for which genomic data are available. A combination of ecologically diverse flora and fauna can be studied in this comparative population genomic framework to better understand the major diversification processes in tropical Africa.

## ACKNOWLEDGEMENTS

We thank the staff and students at various scientific collections that helped to make both samples and associated data available to us, including F. Tillack (Museum für Naturkunde, Berlin), W. Böhme (Museum Koenig, Bonn), J. Rosado (Museum of Comparative Zoology, Cambridge, MA), B. Stuart (North Carolina Museum of Natural Sciences, Raleigh, NC), A. Schmitz (Museum d'Histoire naturelle, Genève), N. Smolensky (Texas A&M Biodiversity Research and Teaching Collection, College Station, TX) C. Spencer (Museum of Vertebrate Zoology, Berkeley, CA), and G. Watkins-Colwell (Yale Peabody Museum, New Haven, CT). We thank A. Sellas (Center for Comparative Genomics, California Academy of Sciences) for assistance in some laboratory components of this work. Genomic data collection was funded by National Science Foundation grant DEB # 1457232 to MKF and ADL. Niche modelling was funded by an National Science Foundation GRFP grant to DR. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. The Cameroon Ministry of Forests and Wildlife (MINFOF) and Ministry of Scientific Research and Innovation (MINRESI) provided necessary permits for conducting research and exportation to DCB, MFB and MH. Fieldwork was supported by National Science Foundation grant DEB #1202609 and two Putnam Expeditionary Grants from the Museum of Comparative Zoology (Harvard University) to DCB and under the approval of the Institutional Animal Care and Use Committee (2014-2) at the California Academy of Sciences. Field work of MH was supported by a Elsa Neumann Stipend (federal state of Berlin), the German Academic Exchange Service and the German Herpetological Society (DGHT Leipzig); MFB was supported by "Freunde und Förderer des Museums für Naturkunde e.V.". Fieldwork by MB in the Republic of Congo was part of a rapid biodiversity initiative, commissioned by Flora Fauna & Man, Ecological Services Ltd (FFMES). Jerome Gaugris of FFMES conducted the study organization and design. Permits were issued by the Groupe d'Etude et de Recherche sur la Diversité Biologique.

## DATA ACCESSIBILITY

The pipelines for processing ddRADseq data and performing demographic model selection are available at: https://github.com/dportik/. We include a large data package on DRYAD (https://doi.org/10.5061/dryad.7p7n2) that includes our final ddRADseq filtered haplotypes file and resulting input files for several programs. We also include data and results files for EEMS, STRUCTURE, ADMIXTURE, BEAST, SNAPP and *δaδi* in this DRYAD package.

## AUTHOR CONTRIBUTIONS

A.D.L., M.K.F., D.C.B., M.O.R. and D.M.P. designed the research project and collected molecular data. A.D.L., D.C.B., M.F.B., M.B., M.H. and M.O.R. performed fieldwork to obtain samples. D.M.P. performed all data processing and molecular analyses, and D.R. performed ecological niche modelling. D.M.P. wrote the manuscript with contributions from A.D.L., M.K.F., D.C.B., D.R. and M.O.R., and all authors approved the final manuscript.

## REFERENCES


Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664.

Amiet, J.-L. (1975). Ecologie et distribution des Amphibiens Anoures de la région de Nkongsamba (Cameroun). *Annales de la Faculté des Sciences de Yaoundé*, *20*, 33–107.

Amiet, J.-L. (2012). *Les Rainettes du Cameroun: (Amphibiens Anoures)* (p. 591). Saint-Nazaire: Jean-Louis Amiet & la Nef des livres.

Anhuf, D., Ledru, M.-P., Behling, H., Da Cruz, F. W. Jr, Cordeiro, R. C., Van der Hammen, T., … Da Silva Dias, P. L. (2006). Paleo-environmental change in Amazonian and African rainforest during the LGM. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *239*, 510–527.

Anthony, N. M., Johnson-Bawe, M., Jeffrey, K., Clifford, S. L., Abernethy, K. A., Tutin, C. E., … Bruford, M. W. (2007). The role of Pleistocene refugia and rivers in shaping gorilla genetic diversity in central Africa. *Proceedings of the National Academy of Sciences USA*, *104*, 20432–20436.

Bohoussou, K. H., Cornette, R., Akpatou, B., Colyn, M., Peterhans, J. K., Kennis, J., … Nicolas, V. (2015). The phylogeography of the rodent genus *Malacomys* suggests multiple Afrotropical Pleistocene lowland forest refugia. *Journal of Biogeography*, *42*, 2049–2061.

Boria, R. A., Olson, L. E., Goodman, S. M., & Anderson, R. P. (2014). Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, *275*, 73–77.

Born, C., Alvarez, N., McKey, D., Ossari, S., Wickings, E. J., Hossaert-McKey, M., & Chevallier, M.-H. (2011). Insights into the biogeographical history of the Lower Guinea Forest Domain: Evidence for the role of refugia in the intraspecific differentiation of *Aucoumea klaineana*. *Molecular Ecology*, *20*, 131–142.

Bouckaert, R. R. (2010). DensiTree: Making sense of sets of phylogenetic trees. *Bioinformatics*, *26*, 1372–1373.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., … Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*, e1003537.

Broennimann, O., Fitzpatrick, P., Pearman, P. B., Petitpierre, B., Pellissier, L., Yoccoz, N. G., … Guisan, A. (2012). Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, *21*, 481–497.

Brouat, C., Gielly, L., & McKey, D. (2001). Phylogenetic relationships in the genus *Leonardoxa* (Leguminosae: Caesalpinioideae) inferred from chloroplast *trnL* intron and *trnL-trnF* intergenic spacer sequences. *American Journal of Botany*, *88*, 143–149.

Brown, J. L. (2014). SDMtoolbox: A python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution*, *5*, 694–700.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & 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.

Budde, K. B., González-Martínez, S. C., Hardy, O. J., & Heuertz, M. (2013). The ancient tropical rainforest tree *Symphonia globulifera* L. f. (Clusiaceae) was not restricted to postulated Pleistocene refugia in Atlantic Equatorial Africa. *Heredity*, *111*, 66–76.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. New York, NY: Springer.

Carnaval, A. C., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T., & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. *Science*, *323*, 785–789.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*, 171–182.

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140.

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*, 7.

Chirio, L., & LeBreton, M. (2007). *Atlas des Reptiles du Cameroun*, Vol. 67 (p. 688). Paris: Collection Patrimoines Naturels, Muséum National d'Histoire Naturelle & IRD Editions.

Clee, P. R. S., Abwe, E. E., Ambahe, R. D., Anthony, N. M., Fotso, R., Locatelli, S., … Gonder, M. K. (2015). Chimpanzee population structure in Cameroon and Nigeria is associated with habitat variation that may be lost under climate change. *BMC Evolutionary Biology*, *15*, 2.

Clifford, S. L., Anthony, N. M., Bawe-Johnson, M., Abernethy, K. A., Tutin, C. E. G., White, L. J. T., … Wickings, E. J. (2004). Mitochondrial DNA phylogeography of western lowland gorillas (*Gorilla gorilla gorilla*). *Molecular Ecology*, *13*, 1551–1565.

Constable, H., Guralnick, R., Wieczorek, J., Spencer, C., Peterson, A. T., & Committee, T. V. N. S. (2010). VertNet: A new model for biodiversity data sharing. *PLoS Biology*, *8*, e1000309.

Cowling, S. A., Cox, P. M., Jones, C. D., Maslin, M. A., Peros, M., & Spall, S. A. (2008). Simulated glacial and interglacial vegetation across Africa: Implications for species phylogenies and trans-African migration of plants and animals. *Global Change Biology*, *14*, 827–840.

Crawford, A. J. (2003). Relative rates of nucleotide substitution in frogs. *Journal of Molecular Evolution*, *57*, 636–641.

Crowe, T. M., & Crowe, A. A. (1982). Patterns of distribution, diversity and endemism in Afrotropical birds. *Journal of Zoology*, *198*, 417–442.

Daïnou, K., Bizoux, J.-P., Doucet, J.-L., Mahy, G., Hardy, O. J., & Heuertz, M. (2010). Forest refugia revisited: nSSRs and cpDNA sequences support historical isolation in a wide-spread African tree with high colonization capacity, *Milicia excels* (Moracea). *Molecular Ecology*, *19*, 4462–4477.

Darst, C. R., & Cannatella, D. C. (2004). Novel relationships among hyloid frogs inferred from 12S and 16S mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, *31*, 462–475.

Dauby, G., Duminil, J., Heuertz, M., & Hardy, O. J. (2010). Chloroplast DNA polymorphism and phylogeography of a Central African tree species widespread in mature rainforest: *Greenwayodendron suaveolens* (Annonaceae). *Tropical Plant Biology*, *3*, 4–13.

Dauby, G., Duminil, J., Heuertz, M., Koffi, G. K., Stévart, T., & Hardy, O. J. (2014). Congruent phylogeographical patterns of eight tree species in Atlantic Central Africa provide insights into the past dynamics of forest cover. *Molecular Ecology*, *23*, 2299–2312.

Day, J. J., Peart, C. R., Brown, K. J., Friel, J. P., Bills, R., & Moritz, T. (2013). Continental diversification of an African catfish radiation (Mochokidae: *Synodontis*). *Systematic Biology*, *62*, 351–365.

Debout, G. D. C., Doucet, J.-L., & Hardy, O. J. (2011). Population history and gene dispersal inferred from spatial genetic structure of a Central African timber tree, *Distemonanthus benthamianus* (Caesalpinioideae). *Heredity*, *106*, 88–99.

Devitt, T. J., Devitt, S. E., Hollingsworth, B. D., McGuire, J. A., & Moritz, C. (2013). Montane refugia predict population genetic structure in the Large-blotched *Ensatina* salamander. *Molecular Ecology*, *22*, 1650–1665.

Dowell, S. A., Portik, D. M., de Buffrénil, V., Ineich, I., Greenbaum, E., Kolokotronis, S.-O., & Hekkala, E. R. (2016). Molecular data from contemporary and historical collections reveal a complex story of cryptic diversification in the *Varanus* (*Polydaedalus*) *niloticus* species group. *Molecular Phylogenetics and Evolution*, *94*, 591–604.

Droissart, V., Sonké, B., Hardy, O. J., Simo, M., Taedoumg, H., Nguembou, C. K., & Stévart, T. (2011). Do plant families with contrasting functional traits show similar patterns of endemism? A case study with Central African Orchidaceae and Rubiaceae. *Biodiversity and Conservation*, *20*, 1507–1531.

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

Duminil, J., Heuertz, M., Doucet, J.-L., Bourland, N., Cruaud, C., Gavory, F., … Hardy, O. J. (2010). CpDNA-based species identification and phylogeography: Application to African tropical tree species. *Molecular Ecology*, *19*, 5469–5483.

Duminil, J., Mona, S., Mardulyn, P., Doumenge, C., Walmacq, F., Doucet, J.-L., & Hardy, O. J. (2015). Late Pleistocene molecular dating of past population fragmentation and demographic changes in African rain forest tree species supports the forest refuge hypothesis. *Journal of Biogeography*, *42*, 1443–1454.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361.

Endler, J. A. (1982). Pleistocene forest refuges: Fact or fancy. In G. T. Prance (Ed.), *Biological diversification in the tropics* (pp. 641–657). New York, NY: Columbia University Press.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*, 2611–2620.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, *164*, 1567–1587.

Faye, A., Deblauwe, V., Mariac, C., Richard, D., Sonké, B., Vigouroux, Y., & Couvreur, T. L. P. (2016). Phylogeography of the genus *Podococcus* (Palmae/Arecaceae) in Central African rain forests: Climate stability predicts unique genetic diversity. *Molecular Phylogenetics and Evolution*, *105*, 126–138.

Freedman, A. H., Thomassen, H. A., Buermann, W., & Smith, T. B. (2010). Genomic signals of diversification along ecological gradients in a tropical lizard. *Molecular Ecology*, *19*, 3773–3788.

Fuchs, J., & Bowie, R. C. K. (2015). Concordant genetic structure in two species of woodpecker distributed across the primary West African biogeographic barriers. *Molecular Phylogenetics and Evolution*, *88*, 64–74.

Fuchs, J., Fjeldså, J., & Bowie, R. C. K. (2017). Diversification across major biogeographic breaks in the African Shining/Square-tailed Drongos complex (Passeriformes: Dicruridae). *Zoologica Scripta*, *46*, 27–41.

Gascon, C., Malcom, J. R., Patton, J. L., da Silva, M. N. F., Bogart, J. P., Lougheed, S. C., … Boag, P. T. (2000). Riverine barriers and the geographic distribution of Amazonian species. *Proceedings of the National Academy of Sciences USA*, *97*, 13672–13677.

Gomez, C., Dussert, S., Hamon, P., Hamon, S., de Kochko, A., & Poncet, V. (2009). Current genetic differentiation of *Coffea canephora* Pierre ex. A. Froehn in Guineo-Congolian African zone: Cumulative impact of ancient climatic changes and recent human activities. *BMC Evolutionary Biology*, *9*, 167.

Gonder, M. K., Locatelli, S., Ghobrial, L., Mitchell, M. W., Kujawski, J. T., Lankester, F. J., … Tishkoff, S. A. (2011). Evidence from Cameroon reveals differences in the genetic structure and histories of chimpanzee populations. *Proceedings of the National Academy of Sciences USA*, *108*, 4766–47771.

Goodier, S. A. M., Cotterill, F. P. D., O'Ryan, C., Skelton, P. H., & de Wit, M. J. (2011). Cryptic diversity of African Tigerfish (genus *Hydrocynus*) reveals palaeogeographic signatures of linked Neogene geotectonic events. *PLoS One*, *6*, e28775.

Goudie, A. S. (2005). The drainage of Africa since the Cretaceous. *Geomorphology*, *67*, 437–456.

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, *5*, e1000695.

Haffer, J. (1969). Speciation in Amazonian forest birds. *Science*, *165*, 131–137.

Haffer, J. (1997). Alternative models of vertebrate speciation in Amazonia: An overview. *Biodiversity and Conservation*, *6*, 451–477.

Hardy, O. J., Born, C., Budde, K., Daïnou, K., Dauby, G., Duminil, J., … Poncet, V. (2013). Comparative phylogeography of African rain forest trees: A review of genetic signatures of vegetation history in the Guineo-Congolian region. *Comptes Rendus Geoscience*, *345*, 284–296.

Heled, J., & Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, *27*, 570–580.

Heuertz, M., Duminil, J., Dauby, G., Savolainen, V., & Hardy, O. J. (2014). Comparative phylogeography in rainforest trees from Lower Guinea. *PLoS One*, *9*, e84307.

Hijmans, R. J. (2012). Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null model. *Ecology*, *93*, 679–688.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978.

Hofer, U., Bersier, L.-F., & Borcard, D. (1999). Spatial organization of a herpetofauna on an elevational gradient revealed by null model tests. *Ecology*, *80*, 976–988.

Hofer, U., Bersier, L.-F., & Borcard, D. (2000). Ecotones and gradient as determinants of herpetofaunal community structure in the primary forest of Mount Kupe, Cameroon. *Journal of Tropical Ecology*, *16*, 517–533.

Jacquet, F., Denys, C., Verheyen, E., Bryja, J., Hutterer, R., Kerbis Peterhans, J. C., … Nicolas, V. (2015). Phylogeography and evolutionary history of the *Crocidura olivieri* complex (Mammalia, Soricomorpha): From a forest origin to broad ecological expansion across Africa. *BMC Evolutionary Biology*, *15*, 71.

Jacquet, F., Nicolas, V., Colyn, M., Kadjo, B., Hutterer, R., Decher, J., … Denys, C. (2014). Forest refugia and riverine barriers promote diversification in the West African pygmy shrew (*Crocidura obscurior* complex, Soricomorpha). *Zoologica Scripta*, *43*, 131–148.

Jombart, T. (2008). *Adegenet*: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405.

Jombart, T., & Ahmed, I. (2011). *Adegenet 1.3-1*: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071.

Katoh, M., & Kuma, M. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*, 3059–3066.

Knowles, L. L., Carstens, B. C., & Keat, M. L. (2007). Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Current Biology*, *17*, 940–946.

Koffi, K. G., Hardy, O. J., Doumenge, C., Cruaud, C., & Heuertz, M. (2011). Diversity gradients and phylogeographic patterns in *Santiria trimera* (Burseraceae), a widespread African tree typical of mature rainforests. *American Journal of Botany*, *98*, 254–264.

Küper, W., Sommer, J. H., Lovett, J. C., Mutke, J., Linder, H. P., Beentje, H. J., … Barthlott, W. (2004). Africa's hotspots of biodiversity redefined. *Annals of the Missouri Botanical Garden*, *91*, 525–535.

Lawson, D. P. (1993). The reptiles and amphibians of Korup National Park Project, Cameroon. *Herpetological Natural History*, *1*, 27–90.

Leaché, A. D., & Fujita, M. K. (2010). Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proceedings of the Royal Society B: Biological Sciences*, *277*, 3071–3077.

Leaché, A. D., Harris, R. B., Rannala, B., & Yang, Z. (2014). The influence of gene flow on species tree estimation: A simulation study. *Systematic Biology*, *63*, 17–30.

Leroux, M. (1983). *Le Climat de l' Afrique Tropicale*, Vol. 1, 2. Paris: Champion Slatkine.

Ley, A. C., Dauby, G., Köhler, J., Wypior, C., Röser, M., & Hardy, O. J. (2014). Comparative phylogeography of eight herbs and lianas

(Marantaceae) in central African rainforests. *Frontiers in Genetics*, *5*, 403.

Ley, A. C., Heuertz, M., & Hardy, O. J. (2016). The evolutionary history of central African rain forest plants: Phylogeographical insights from sister species in the climber genus *Haumania* (Marantaceae). *Journal of Biogeography*, *44*, 308–321, Online Early Access.

Linder, H. P., de Klerk, H. M., Born, J., Burgess, N. D., Fjeldså, J., & Rahbek, C. (2012). The partitioning of Africa: Statistically defined biogeographical regions in Sub-Saharan Africa. *Journal of Biogeography*, *39*, 1189–1205.

Lowe, A. J., Harris, D., Dormontt, E., & Dawson, I. K. (2010). Testing putative African tropical forest refugia using chloroplast and nuclear DNA phylogeography. *Tropical Plant Biology*, *3*, 50–58.

Lynch, M. (2010). Rate, molecular spectrum, and consequences of human mutation. *Proceedings of the National Academy of Sciences USA*, *107*, 961–968.

Maley, J. (1996). The African rain forest – Main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh*, *104B*, 31–73.

Marzoli, A., Piccirillo, E. M., Renne, P. R., Bellieni, G., Iacumin, M., Nyobe, J. B., & Tongwa, A. T. (2000). The Cameroonian Volcanic Line revisited: Petrogenesis of continental basaltic magmas from lithospheric and asthenospheric mantle sources. *Journal of Petrology*, *41*, 87–109.

Mayr, E., & O'Hara, R. J. (1986). The biogeographical evidence supporting the Pleistocene forest refuge hypothesis. *Evolution*, *40*, 55–67.

deMenocal, P. B. (2004). African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters*, *220*, 3–24.

Mitchell, M. W., Locatelli, S., Clee, P. R. S., Thomassen, H. A., & Gonder, M. K. (2015). Environmental variation *and* rivers govern the structure of chimpanzee genetic diversity in a biodiversity hotspot. *BMC Evolutionary Biology*, *15*, 1.

Mitchell, M. W., Locatelli, S., Ghobrial, L., Pokempner, A. A., Clee, P. R. S., Abwe, E. E., … Gonder, M. K. (2015). The population genetics of wild chimpanzees in Cameroon and Nigeria suggests a positive role for selection in the evolution of chimpanzee subspecies. *BMC Evolutionary Biology*, *15*, 3.

Moritz, C., Patton, J. L., Schneider, C. J., & Smith, T. B. (2000). Diversification of rainforest faunas: And integrated molecular approach. *Annual Review of Ecology and Systematics*, *31*, 533–563.

Moussalli, A., Moritz, C., Williams, S. E., & Carnaval, A. C. (2009). Variable responses of skinks to a common history of rainforest fluctuation: Concordance between phylogeography and palaeo-distribution models. *Molecular Ecology*, *18*, 483–499.

Muloko-Ntoutoume, N., Petit, R. J., White, L., & Abernethy, K. (2000). Chloroplast DNA variation in a rainforest tree (*Aucoumea klaineana*, Burseraceae) in Gabon. *Molecular Ecology*, *9*, 359–363.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*, 853–858.

Nesi, N., Kadjo, B., Pourrut, X., Leroy, E., Shongo, C. P., Cruaud, C., & Hassanin, A. (2013). Molecular systematics and phylogeography of the tribe Myonycterini (Mammalia, Pteropodidae) inferred from mitochondrial and nuclear markers. *Molecular Phylogenetics and Evolution*, *66*, 126–137.

Ngueutchoua, G., Giresse, P. (2010). Sand bodies and incised valleys within the Late Quaternary Sanaga-Nyong delta complex on the middle continental shelf of Cameroon. *Marine and Petroleum Geology*, *27*, 2173–2188.

Nicolas, V., Akpatou, B., Wendelen, W., Peterhans, J. K., Olayemi, A., Decher, J., … Colyn, M. (2010). Molecular and morphometric variation in two sibling species of the genus *Praomys* (Rodentia: Muridae): Implications for biogeography. *Zoological Journal of the Linnean Society*, *160*, 397–419.

Nicolas, V., Mboumba, J.-F., Verheyen, E., Denys, C., Lecompte, E., Olayemi, A., … Colyn, M. (2008). Phylogeographic structure and regional history of *Lemniscomys striatus* (Rodentia: Muridae) in tropical Africa. *Journal of Biogeography*, *35*, 2074–2089.

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

Oates, J. F., Bergl, R. A., & Linder, J. M. (2004). Africa's Gulf of Guinea Forests: Biodiversity patterns and conservation priorities. *Advances in Applied Biodiversity Science*, *6*, 1–90.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., … Kassem, K. R. (2001). Terrestrial ecoregions of the world: A new map of life on Earth. *BioScience*, *51*, 933–938.

Penner, J., Wegmann, M., Hillers, A., Schmidt, M., & Rödel, M.-O. (2011). A hotspot revisited – A biogeographical analysis of West African amphibians. *Diversity and Distributions*, *17*, 1077–1088.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo discovery and genotyping in model and non-model species. *PLoS One*, *7*, e37135.

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*, 94–100.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*, 231–259.

Pinton, A., Agnèse, J.-F., Paugy, D., & Otero, O. (2013). A large-scale phylogeny of *Synodontis* (Mochokidae, Siluriformes) reveals the influence of geological events on continental diversity during the Cenozoic. *Molecular Phylogenetics and Evolution*, *66*, 1027–1040.

Plana, V. (2004). Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. *Philosophical Transactions of the Royal Society B*, *359*, 1585–1594.

Portik, D. M., & Blackburn, D. C. (2016). The evolution of reproductive diversity in Afrobatrachia: A phylogenetic comparative analysis of an extensive radiation of African frogs. *Evolution*, *70*, 2017–2032.

Portik, D. M., Jongsma, G. F. M., Kouete, M. T., Scheinberg, L. A., Freiermuth, B., Tapondjou, W. P., & Blackburn, D. C. (2016). A survey of amphibians and reptiles in the foothills of Mount Kupe, Cameroon. *Amphibian and Reptile Conservation*, *10*, 37–67 (e131).

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

Rambaut, A., & Drummond, A. J. (2009). Tracer v1.5.0. Available from: http://beast.bio.ed.ac.uk/

Shcheglovitova, M., & Anderson, R. P. (2013). Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. *Ecological Modelling*, *269*, 9–17.

Smith, T. B., Calsbeek, R., Wayne, R. K., Holder, K. H., 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.

Smith, T. B., Thomassen, H. A., Freedman, A. H., Sehgal, R. N. M., Buermann, W., Saatchi, S., … Wayne, R. K. (2011). Patterns of divergence in the olive sunbird *Cyanomitra olivacea* (Aves: Nectariniidae) across the African rainforest-savanna ecotone. *Biological Journal of the Linnean Society*, *103*, 821–835.

Smith, T. B., Wayne, R. K., Girman, D. J., & Bruford, M. W. (1997). A role for ecotones in generating rainforest biodiversity. *Science*, *276*, 1855–1857.

Sosef, M. S. M. (1996). Begonias and African rainforest refuges: General aspect and recent progress. In L. J. G. van der Maesen, X. M. van der

Burg & J. M. van medenbach de Rooy (Eds.), *The biodiversity of African plants* (pp. 602–611). Dordrecht: Kluwer.

Telfer, P. T., Souquière, S., Clifford, S. L., Abernethy, K. A., Bruford, M. W., Disotell, T. R., … Wickings, E. J. (2003). Molecular evidence for deep phylogenetic divergence in *Mandrillus sphinx*. *Molecular Ecology*, *12*, 2019–2024.

Vande Weghe, J. P. (2004). *Forests of Central Africa*. Tielt: Lannoo Publishers.

Veloz, S. D. (2009). Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, *36*, 2290–2299.

Wallace, A. R. (1852). On the monkeys of the Amazon. *Proceedings of the Zoological Society London*, *20*, 107–110.

Warren, D. L. (2016). ENMtools. Available from: https://github.com/danlwarren/ENMTools

Warren, D. L., Glor, R. E., & Turelli, M. (2008). Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution*, *62*, 2868–2883.

Xue, A. T., & Hickerson, M. J. (2015). The aggregate site frequency spectrum for comparative population genomic inference. *Molecular Ecology*, *24*, 6223–6240.

Yannic, G., Pellissier, L., Ortego, J., Lecomte, N., Couturier, S., Cuyler, C., … Cote, S. D. (2014). Genetic diversity in caribou linked to past and future climate change. *Nature Climate Change*, *4*, 132–137.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Portik DM, Leaché AD, Rivera D, et al. Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Mol Ecol*. 2017;00:1–19. https://doi.org/10.1111/mec.14266