Geographic patterns of speciation in West and Central Africa: The role of rivers, mountains, and refugia

By

Kaitlin Elizabeth Allen

Submitted to the graduate degree program in Ecology and Evolutionary Biology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Co-Chair: Rafe M. Brown

Co-Chair: A. Townsend Peterson

William Leo Smith

Deborah R. Smith

Peter C. Ojiambo

David C. Blackburn

Date Defended: 8 October 2021

The dissertation committee for Kaitlin Elizabeth Allen certifies that this is the approved version of the following dissertation:

Geographic patterns of speciation in West and Central Africa: The role of rivers, mountains and refugia

_____

Co-Chair: Rafe M. Brown

_____

Co-Chair: A. Townsend Peterson

Date Approved: 8 October 2021

Abstract

The tropical rainforests of Africa make up one of the largest rainforest blocks in the world, second only to the Amazon Rainforest of South America. This area hosts an incredible diversity of plants and animals and contains three of the world's biodiversity hotspots, indicative of both its high levels of endemism and its high conservation priority. Despite this, the study of biodiversity in the African rainforest lags far behind that of almost every other major rainforest in the world. This is widely due to logistical issues driven by infrastructure, political instability, and fragmented political boundaries.

The high levels of diversity in the African rainforest have been shaped by a variety of geological factors including the formation of two mountain ranges, four islands, and six major rivers, as well as continent wide climate events such as the Miocene aridification and the Pleistocene glaciation cycles. However, the roles of these factors in driving diversification in Central Africa are generally poorly understood. A major sampling gap exists for many taxa in the Congo Basin, and pollen core and fossil data tend to be poorly preserved in the tropics. In addition, many climatic and geologic events overlap spatially making it difficult to tease apart their combined/confounding effects. This dissertation addresses these difficulties by filling major sampling gaps in Central Africa and by using a combination of phylogenomic, population genomic, demographic and niche modeling analyses to gain a better understanding of the responses of West and Central African taxa to the major climatic and geographic events shaping African rainforest biota.

Chapter 1 examines the role that the Pleistocene glaciation cycles played in the distribution of mountain endemic taxa across the Afromontane archipelago. Since the 1800s biogeographers have noted the high floral and faunal similarity between Afromontane regions

when compared to the intervening lowlands. The prevailing hypothesis is that this similarity results from the spread of montane habitat into lowland areas during the glacial cycles of the Pleistocene. However, a lack of pollen core data from the Central African lowlands has made it difficult to determine if habitat shifts were extensive enough to allow biotic exchange between montane regions. In this study I use ecological niche modeling for sixteen Afromontane endemic bird, tree, and shrub species to infer the extent and most likely geographic positions of suitable habitat for Afromontane taxa during the present, the mid-Holocene and the Last Glacial Maximum (LGM). The results suggest widespread climatic suitability across lowland Central Africa for all taxa in our study during the LGM and, to a lesser extent, during the Mid-Holocene. Suitable areas connected all Afromontane regions in sub-Saharan African except for the Ethiopian Highlands, supporting the hypothesis that faunal and floral similarity between Afromontane regions is a result of species distributional changes during recent periods of global glaciation.

Chapter 2 takes a modern, integrative approach to testing classical hypotheses of rainforest speciation in lowland West and Central Africa. This chapter tests predictions derived from the river, refuge, and river-refuge hypotheses of tropical species diversification proposed for the Amazon rainforest in the 1960s. Evolutionary and biogeographic patterns inferred from these hypotheses were originally based solely on species distribution data and suggest that most tropical speciation events are driven by allopatry associated with the formation of rivers, Pleistocene refugia, or a combination of the two. However, application of these hypotheses to the African rainforest was limited. In this study, I use phylogenomic and population genomic methods, combined with ecological niche model projections to the LGM and mid-Holocene, to assess the predictions of these classical hypotheses in the arboreal sub-Saharan snake genus

Toxicodryas. Rivers were found to represent strong barriers to gene flow among populations of this genus, and no support was found for a major contraction of suitable habitat during the LGM. These results allow the rejection of both the refuge and river-refuge hypotheses in favor of the river hypothesis for diversification in this genus and have led to the description of two new species of *Toxicodryas* from the Congo Basin.

Chapter 3 takes a deeper look at population genetic patterns associated with African Pleistocene refugia. Based on limited pollen core records, Pleistocene refugia have been proposed for several localities across West and Central Africa. Many of these refugia are associated with areas of high surface relief, such as montane regions, but also include a fluvial refuge in gallery forests around the Congo River. However, the degree to which these refugia have driven species diversification patterns in the African rainforest has been widely debated, often because of conflicting phylogeographic or species distribution data. In this study I test the predictions of the Pleistocene refugial diversification hypothesis in a widespread tropical African skink species, *Trachylepis maculilabris*, using a combination of dated phylogeographic analysis, population demography, and ecological niche modeling with projections to the LGM and mid-Holocene. I found five distinct genetic clusters with divergence times in the Pliocene and Pleistocene. Demographic analyses suggest population expansion as the earth warmed after the LGM and allowed for the determination of the most likely locations of Pleistocene refugia. These results suggest that the Congo River fluvial refuge played a role in species diversification patterns in *T. maculilabris*, and that mountain ranges played complex roles, with the Albertine Rift acting as a refugium and the Cameroon Volcanic Line acting as both a population barrier and a refugium.

Chapter 4 examines the colonization history of two species in the *Trachylepis maculilabris* complex in the Gulf of Guinea islands. These islands are an offshore extension of the Cameroon Volcanic Line and consist of one continental and three oceanic islands of varying sizes and distances from the mainland. The oceanic islands have unusually high levels of biodiversity, resulting from dispersal from the mainland combined with in-situ diversification. Dispersal to the islands likely occurs through rafting via three major river systems and two oceanic currents that result in complex colonization patterns from both West and Central Africa. In this study I use dated phylogenomic, historical demographic, and ancestral geographic range reconstruction analyses to determine the most likely colonization history for this complex in the Gulf of Guinea islands. I found that the island species were monophyletic, indicating a single colonization event on São Tomé Island from the eastern Congo Basin around 14 million years ago. Colonization of the islands then likely proceeded sequentially from São Tomé to Príncipe to Tinhosa Grande, with the colonization of Tinhosa Grande occurring via a land bridge from Príncipe during the Pleistocene glaciation cycles.

Chapter 5 focuses on the role that polyploidization plays in climatic niche evolution in the sub-Saharan African frog genus *Xenopus*. There are 29 species of *Xenopus* in two sub-genera. One of these species is diploid, 16 are tetraploid, seven are octoploid, four are dodecaploid and one is unknown. There are clear geographic patterns to the distributions of different ploidy levels within the genus, such that octoploids and dodecaploids are found mainly in and around the Cameroon Volcanic Line and the Albertine Rift. Polyploid plant literature has suggested a relationship between polyploidization and adaptation to new or extreme environments, potentially explaining these geographic patterns. In order to examine the relationship between polyploidization and niche evolution in *Xenopus*, I conducted ancestral

state reconstructions of each species' realized niche on a reticulate phylogeny for the genus. I found that niche expansion and contraction around the ancestral niche was common in this genus, while niche shift and niche novelty were relatively rare. While niche expansion and contraction were equally distributed throughout the phylogeny, niche novelty was more likely to be associated with higher polyploidy species.

Acknowledgments

I would like to thank the KU Biodiversity Institute and the department of Ecology and Evolutionary Biology for providing me with the opportunity and support necessary to pursue my PhD at the University of Kansas. I would especially like to thank the previous director of the Biodiversity Institute Dr. Leonard Krishtalka and the interim director Dr. Jorge Soberón, the previous Chairs of EEB Dr. Christopher Haufler and Dr. Maria Orive, the current Chair of EEB Dr. Lena Hileman, the Associate Chair of Graduate Studies Dr. Andrew Short, the previous EEB Graduate Coordinator Aagje Ashe and the current coordinator Ryan Zeigler, as well as Lori Schlenker, Jamie Keeler, and Rick Evanhoe for all of their help and kindness throughout my time at KU. I would also like to thank the University of Kansas for funding through the Ida Hyde Scholarship, Biodiversity Institute Panorama Grant, Open Access Author Fund, International Programs Pre-dissertation Travel Grant, Graduate Studies Doctoral Student Research Fund, Summer Research Support funding, and Summer Research Scholarship; the American Society of Ichthyologists and Herpetologists for funding through the Gaige Fund; National Geographic through the Young Explorers Grant; The Explorers Club through the Mamont Grant; and APS through the Lewis and Clark Fund for Exploration and Field Research.

I would especially like to thank my PhD advisors Dr. Rafe Brown and Dr. Town Peterson for their invaluable guidance, support and countless manuscript edits throughout the last six years. I would also like to thank my committee members Dr. David Blackburn, Dr. Peter Ojiambo, Dr. Deb Smith, Dr. Leo Smith, and Dr. Joy Ward for their advice and support. In addition, I would like to thank my undergraduate and masters advisors Dr. Chad Montgomery, Dr. Bob Powell, and Dr. Aaron Bauer. Without their mentorship and guidance, I would never be where I am today. Also Dr. Eli Greenbaum and Dr. Ben Evans who, although they had no formal

role, were always willing to advise and collaborate, and without them much of this dissertation would be sorely lacking. I would like to thank all the lab mates, collaborators and friends that have supported me throughout my PhD including: Robin Abraham, Alana Alexander, Abdu Alkishe, Stephen Baca, Lindsay Campbell, Jackie Childers, Kevin Chovanec, Kerry Cobb, Marlon Cobos, Jacob Cooper, Luke DeCicco, Pietro de Mello, Bill Duellman, Doug and Maria Eifler, Divine Fotibu, Ben Freeman, Rich Glor, Nono Gonwouo, Johana Goyes Vallejos, Jesse Grismer, William Hatungimana, Mark Herr, Paul Hime, Carl Hutter, Kate Ingenloff, Grant Johnson, Marina Kameni, Alyssa Leinweber, Fernando Machado-Stredel, Melissa Mayhew, Pietro de Mello, Camila Meneses, Ana Motta, Haruka Nagao, Karen Olson, Chan Kin Onn, Katie Peters, Nikki Realubit, Mark Robbins, Abdallah Samy, Hannah Som, Viktoria Sterkhova, Christian Supsup, Javier Torres, Scott Travers, Linda Trueb, Jeff Weinell, Luke Welton, and Perry Wood Jr.

I would like to thank my family in the US for all their love and support over the years, for fostering my love of nature and for listening to countless hours of lizard talk over homemade wine. I would like to thank my family in Cameroon for accepting me with love and support and for always being good sports about the language barriers. I'd like to thank my cats for keeping me company during countless hours of writing. But most of all I'd like to thank my husband, Walter, who has always had my back through the ups and downs of this journey, and who continuously brings adventure and joy to my life.

Table of Contents

Introduction

The latitudinal diversity gradient, one of the most pervasive patterns in biogeography, is a strong trend of increasing species diversity from the poles toward the Equator (Darlington, 1957; Pianka, 1966; Wallace, 1878). To explain this phenomenon, many researchers have focused on speciation processes in major tropical rainforests because of their high biodiversity (e.g., Jablonski et al., 2006; Rohde, 1997; Stevens, 1989). However, due to a variety of logistical constraints pertaining to infrastructure, political instability, and fragmented political boundaries, our understanding of the patterns and processes of speciation in the African rainforest lags well behind that of tropical rainforests in South America and Asia (Greenbaum, 2017; Malhi et al., 2015; Siddig, 2019; Tolley et al., 2016).

The African rainforest consists of several major forest blocks, including the Upper Guinean forests of West Africa and the Lower Guinean and Congolian forests of Central Africa. Together they make up the second largest rainforest in the world and contain three recognized biodiversity hotspots (Mittermeier et al., 2011; Myers et al., 2000). Areas of high biodiversity and endemicity tend to be centered around the coastal regions of Cameroon, Gabon, and the Republic of Congo (Küper et al., 2004; Sosef et al., 2017), as well as the mountains of the Cameroon Volcanic Line and the Albertine Rift (Antonelli et al., 2018; Dagallier et al., 2020; Fjeldså & Lovett, 1997). The Congo Basin tends to exhibit lower biodiversity, but also represents a major sampling gap in many studies (De Klerk et al., 2002; Greenbaum, 2017, Lewin et al., 2016; Van de Perre et al., 2019).

A number of geologic factors have played roles in generating patterns of biodiversity in tropical African rainforest (reviewed in Couvreur et al., 2021). Two major mountain chains can be found in Central Africa, the Cameroon Volcanic Line in the west and the Albertine Rift in the

east, both of which experienced much of their geologic development during the Miocene (Burke, 2001; Macgregor, 2015; Marzoli et al., 2000; Paul et al., 2014; Reusch et al., 2010). Six large river systems run through these rainforests including the Volta, Sanaga, Ogooué, Congo, Niger, and Cross Rivers. Many of these rivers date back to the Late Mesozoic to Early Cenozoic (80–35 mya; Couvreur et al., 2021; Goudie, 2005; Stankiewicz & de Wit, 2006), with the exception that the present course of the Congo River dates to the mid to late Miocene, corresponding with the uplift of the East African Rift (Flügel et al., 2015; Stankiewicz & de Wit, 2006; Takemoto et al., 2015). Off of the mainland, four islands extending off of the Cameroon Volcanic Line into the Gulf of Guinea arose at various times in the last 5 to 30 million years (Aka et al., 2001; Burke, 2001; Caldeira & Munhá, 2002; Déruelle et al., 1991).

Past and present climate events have also greatly influenced biodiversity patterns in Central Africa (reviewed in Couvreur et al., 2021). the Dahomey Gap, a savannah mosaic that separates the Upper and Lower Guinean rainforests, has existed in some form for at least the last 150,000 years (Dupont & Weinelt, 1996; Allen et al., 2020) and is the result of weather patterns over North and West Africa (Salzmann & Hoelzmann, 2005). The Central African climatic hinge, located between southern Cameroon and Gabon, corresponds to an inversion of the dry and rainy seasons between West-Central and Central Africa (Gonmadje et al., 2012; Hardy et al., 2013). The Miocene aridification is associated with the expansion of savanna habitat in sub-Saharan Africa around six to eight million years ago (Couvreur et al., 2021; Feakins et al., 2013; Herbert et al., 2016; Pokorny et al., 2015; Ségalen et al., 2007; Senut et al., 2009; Sepulchre et al., 2006). The Pleistocene glaciation cycles took place from 0.012 to 2.5 mya and are hypothesized to have caused tropical rainforests to shrink into multiple climatically stable, isolated "refugia" during glacial periods (Maley, 1996). The full effects of these cyclic

expansions and contractions of rainforest habitat on species diversification have been widely debated (Amorim, 1991; Colinvaux et al., 2001; DeMenocal, 2004; Haffer, 1969, 1997; Mayr & O'Hara, 1986; Vitorino et al., 2016)

The role of many of these climatic and geologic events in driving diversification in Central Africa are poorly understood. Often multiple events will have impacted an area at different times, causing combined or confounding effects on speciation or population genetic patterns (Budde et al., 2013; Edwards et al., 2006; McKinnon et al., 2004; Slager et al., 2020; Tzedakis et al., 2013). In addition, the responses of individual species to climate or habitat changes may be highly dependent on their fundamental ecological niches, thus appearing idiosyncratic (Bell et al., 2017; Budde et al., 2013; Hardy et al., 2013; Jackson et al., 2009; Leaché et al., 2020; Nogués-Bravo, 2009). Major sampling gaps for many species in the Congo Basin and Nigeria have also made some broad scale, pan-African inferences about speciation patterns difficult (e.g., Allen et al., 2019; Greenbaum, 2017; Huntley & Voelkner; Jesus et al., 2005a, 2007).

In this dissertation, I aim to conduct an extensive analysis of the roles of these geologic and climatic events in driving species diversification patterns in widespread African rainforest taxa. This study provides one of the most comprehensive geographic analyses to date by filling the Congo Basin sampling gap for several widespread taxa. It also takes advantage of recent advances in paleo-climate modeling and genome-scale DNA sequence analysis to gain an understanding of the complex processes that have driven species diversification in the world's second largest but most poorly understood rainforest. Each chapter will contribute to this understanding as follows: **Chapter 1** – Examines the effects of glaciation cycles during the Pleistocene on potential habitat suitability shifts in the Afromontane archipelago and the Congo

Basin for a variety of mountain endemic birds, trees, and shrubs. **Chapter 2** – Determines whether rivers formed during the Miocene or rainforest distribution changes during the Pleistocene had a greater effect than on the phylogenetic diversification of an arboreal West and Central African snake genus. **Chapter 3** – Shows the combined effects of rivers, mountains, and Pleistocene refugia on the population genetics and historical demography of a widespread African skink species. **Chapter 4** – Highlights the role that rivers, ocean currents and Pleistocene sea level shifts have played in the colonization of skinks on the Gulf of Guinea islands. **Chapter 5** – Examines the effects of polyploidization on niche evolution in the African clawed frog and discusses the role that climate changes during the Miocene and Pleistocene may have played on the formation of polyploid species.

CHAPTER 1

Modeling potential Pleistocene habitat corridors between Afromontane forest regions

Allen, K. E., Tapondjou N., W.P., Freeman, B., Cooper, J.C., Brown, R.M., Peterson, A.T. 2021. Modeling potential Pleistocene habitat corridors between Afromontane forest regions. Biodiversity and Conservation 30, 2361–2375.

1.1 Abstract

The unusually high floral and faunal similarity between the different regions of the Afromontane archipelago has been noted by biogeographers since the late 1800s. A possible explanation for this similarity is the spread of montane habitat into the intervening lowlands during the glacial periods of the Pleistocene, allowing biotic exchange between mountain ranges. In this study, we sought to infer the existence and most likely positions of these potential habitat corridors. We focused on sixteen Afromontane endemic tree, shrub, and bird species in the Cameroon Volcanic Line, East African Rift and Great Escarpment. Species were chosen based on distribution above 1200–1500 m in at least two of the major Afromontane regions. Ecological niche models were developed for each species in the present and projected to the mid-Holocene and the Last Glacial Maximum (LGM). Models were thresholded to create binary maps of presence/absence and then summed across taxa to estimate potential LGM and mid-Holocene distributions. We found widespread climatic suitability for our montane taxa throughout the lowlands of Central Africa during the LGM, connecting all regions of the Afromontane archipelago except the Ethiopian Highlands and the Dahomey Gap. During the mid-Holocene, we noted more limited climatic

suitability for fewer species in lowland areas. Although we set out to test predictions derived from alternatively hypothesized corridors, we instead found widespread climatic suitability connecting Afromontane regions across the entire Congo Basin for all species.

1.2 Introduction

Afromontane forest is characterized by a cooler, more humid climate than the surrounding lowlands. Typically found above 1200–1500 m in elevation, this ecoregion is highly segmented and is often referred to as the "Afromontane archipelago," representing montane areas as islands separated by hotter, more arid lowlands (Kingdon, 1989; Moreau, 1963, 1966; White, 1978, 1981). The major geographic areas of the archipelago are the East African Rift (including the Ethiopian Highlands, Albertine Rift and Eastern Arc Range) in East Africa, the Cameroon Volcanic Line in West/Central Africa, and the Great Escarpment in South Africa, with small, isolated patches also occurring in the highlands of Guinea, Sierra Leone and Liberia. These areas tend to have high species diversity (Bussmann, 2004; Mutke & Barthlott, 2005; White, 1981) and are all considered to be biodiversity hotspots of high conservation priority (Mittermeier et al., 2011; Myers et al., 2000; Schmitt et al., 2010; White, 1981).

Since the late 1800s, biogeographers have noted striking floral and faunal similarity among the three disparate Afromontane forest regions, suggesting that they were connected in the recent geological past (Hall, 1973; Moreau, 1966; Prigogine, 1987; Reichenow, 1900; Sharpe, 1893; White, 1981). Pollen core data suggest that montane vegetation extended much further into the lowlands during the glacial periods of the Pleistocene than at present (Caratini & Giresse, 1979; Maley, 1987, 1989; Van Zinderen Bakker & Clark, 1962), and it is possible that montane forests expanded to form habitat corridors, or biogeographical conduits for biotic

exchange, that allowed species to expand their ranges among montane regions during the colder and drier periods of the Pleistocene. These ranges would have contracted subsequently, becoming interrupted and eventually separated by lowland evergreen forest during interglacial periods and the Holocene (Carlquist, 1966; Darlington, 1957; Hall, 1973; Kadu et al., 2011; Livingstone, 1975; Maley, 1991; Moreau, 1966; Tosi, 2008). It has been estimated that a decrease in temperature of 5°C during a glacial maximum would lower montane forest limits from ~1500 m to 500–700 m in elevation, allowing extensive spread of montane forests in lowland areas (Moreau, 1963). This effect is illustrated in the present day by near-sea-level distributions of montane plants and animals on the coastal slopes of Mount Cameroon, where climate patterns enable these species to persist at low elevations (Serle, 1964).

However, the location and extent of these corridors has been debated widely for lack of pollen core data from Central Africa (Coetzee & Van Zinderen Bakker, 1970; Hedberg, 1969; Wickens, 1976). Three possible corridors have been proposed: one across the northern rim of the Congo Basin (Moreau, 1963, 1966), and one through patchy montane areas in the southern Democratic Republic of the Congo, Zambia, northern Angola, and Gabon (White, 1981, 1983, 1993; Figure 1.1), both of which connect the Cameroon Volcanic Line to the East African Rift; and an additional corridor connecting the East African Rift and the Great Escarpment in southern Africa (Linder, 1983; Figure 1.1). The goal of this study was to distinguish among these classic hypotheses of Pleistocene range expansion by examining the extent and location of climatic suitability for montane species through time using paleo-projections of ecological niche models for montane species during the Last Glacial Maximum (LGM), ~22,000 years ago, and the mid-Holocene, ~6000 years ago.

We focused on montane tree, shrub, and bird species, as they have relatively high dispersal capabilities. We analyzed single species or species complexes with ranges that span multiple Afromontane regions (Moreau, 1966; White, 1981) and use these taxa as proxies for understanding the potential distribution of Afromontane forests as a whole (Kadu et al., 2011; Tosi, 2008). These proxies reflect the important roles that tree, shrub, and bird species play in their ecosystems, including providing habitat (Cooper et al., 2017; Tesfaye et al., 2013), food (Kaplin et al., 1998; Menke et al., 2012), and pollination services (Farwig et al., 2006; Janeček et al., 2007; Johnson & Brown, 2004; Newmark et al., 2020). We found widespread climatic suitability connecting the different Afromontane regions across Central and East Africa for all taxa in our data set during the LGM, but more limited suitability for many taxa during the mid-Holocene. This result suggests a climatic basis for the hypothesis that species were able to expand their geographic ranges among Afromontane regions during the glacial cycles of the Pleistocene.

1.3 Materials and Methods

1.3.1 Species selection

Bird, tree, and shrub species were used only if they were highland endemics (i.e., found exclusively above elevations of 1200–1500 m) distributed in two or more of the major Afromontane regions (Figure 1.2). Special focus was given to species found on both the eastern and western sides of the Congo Basin, but species that were found in highland regions south of the Congo Basin (i.e., *Cinnyris ludovicensis*), and in South and East Africa (i.e., *Rapanea melanophloeos*) were also included. Bird species were identified using distributional information

in published summaries (del Hoyo et al., 2004, 2005, 2006, 2007, 2008, 2009, 2010); tree and shrub species were identified from White (1981) and through online searches.

Distributional data were downloaded from GBIF; species with the highest representative sampling were used in this study, giving a total of 10 bird species (*Apalis cinerea*, *Ceblepyris caesius*, *Cinnyris ludovicensis*, *Cisticola chubbi*, *Kakamega poliothorax*, *Onychognathus walleri*, *Oreolais pulcher*, *Ploceus baglafecht*, *Ploceus insignis*, *Sylvia abyssinica*) and 6 tree and shrub species (*Agarista salicifolia*, *Hypericum revolutum*, *Ilex mitis*, *Podocarpus latifolius*, *Rapanea melanophloeos, Ternstroemia polypetala*). Data were then mapped in QGIS v3.4 to check for accuracy and cleaned, removing records with no locality info or locality info far outside of the species known range, records that included only a country code or province name, or records with uncertain IDs; records with accurate, but text-only locality information were georeferenced using Google Earth (earth.google.com) and Falling Rain Global Gazetteer Version 2.3 (www.fallingrain.com)

1.3.2 Ecological niche modeling

Environmental data were obtained from the WorldClim database v. 1.4 (Hijmans et al., 2005). Only mainland Africa was included in the analyses. Fifteen of the 19 bioclim variables were used in analysis at a 2.5' (~4.5 km) spatial resolution. The remaining four variables (bio8, bio9, bio18, and bio19) were excluded because they are known to hold artifacts in the form of discontinuities in the midst of continuous environmental gradients. The 15 variables were used for the LGM and the mid-Holocene under three general circulation models (GCMs): CCSM4, MIROC-ESM, and MPI-ESM-P. To reduce collinearity effects on the resulting models, a

principal component analysis (PCA) was performed on present-day bioclimatic variables and transferred to the past versions of the same set of variables.

Model calibration areas were defined as a 1000 km buffer around the known occurrence points for each species. Model calibration, creation, projection, and evaluation were carried out using the R package 'kuenm' (Cobos et al., 2019). To calibrate models, we created 765 candidate models for each species. Candidate models included all combinations of three sets of environmental predictors (PCAs 1–6, 1–5, 1–4), 17 regularization multiplier values (0.1–1.0 at intervals of 0.1, 2–6 at intervals of 1, and 8 and 10), and all combinations of four feature classes (linear = l, quadratic = q, product = p, and hinge = h; Cobos et al., 2019). Models were run using maximum entropy machine-learning algorithms in MAXENT v. 3.4 (Phillips et al., 2006); best models were chosen based on significant partial ROC scores (Peterson et al., 2008; 500 bootstrap replicates, $E = 5\%$), omission rates < 5% based on a modified least training presence thresholding (Anderson et al., 2003; $E = 5\%$), and AICc scores of ≤ 2 to minimize model complexity (Warren & Seifert, 2011). This model selection procedure identified the optimal parameter set to be used for final model creation.

Final models were created for each species using the full set of occurrence records and the parameters chosen during model calibration (see above), also using the 'kuenm' package. Models were run in MAXENT with 10 bootstrap replicates and logistic output formats. After models were run in the present, they were checked for consistency with the species known distribution and transferred to the LGM and mid-Holocene for each of the three GCMs. The mobility-oriented parity (MOP) index was used to test for model extrapolation (Owens et al., 2013). Models were visualized in QGIS 3.4 and thresholded using a modified least training presence thresholding ($E = 5\%$) to create binary maps of presence/absence. Models from each

time period were summed to estimate potential LGM and mid-Holocene distributions (Devitt et al., 2013; Yannic et al., 2014).

1.4 Results

For all species analyzed, our models suggested suitable climate, defined as climate for which a species was projected to be able to be present after binary thresholding, connecting Afromontane regions on either side of the Congo Basin during the LGM. Indeed, 20–40% of the species showed connections across the Congo Basin during the mid-Holocene (Figures 1.3 and 1.4). Broadly, trees/shrubs and birds showed different distributional histories but overall similar patterns. The southern route between the Cameroon Volcanic Line and the East African Rift appears to have been climatically suitable for a higher proportion of bird species than plant species during the mid-Holocene, but both routes were climatically suitable for nearly all species during the LGM. In tree and shrub species, both the northern and the southern routes were climatically suitable for all species during the mid-Holocene. We noted widespread climatic suitability across lowland regions of West and Central Africa during the LGM. Large areas of unsuitable climate separated the East African Rift and the Great Escarpment for all species during the mid-Holocene; however, during the LGM, continuous suitable habitat was inferred between these two regions for 30-40% of the bird species in this study and all tree/shrub species.

While we did not focus on species occurring in the Angolan and Ethiopian Highlands, they were included in the actual and potential distributions of many species. Our models suggested that climatically suitable habitat connected the Angolan Highlands to the rest of the Afromontane system during LGM, but such was not the case for the Ethiopian Highlands, which were separated by climatically unsuitable habitat for all species during all three time periods

included in this study. Similarly, the Dahomey Gap in Benin and Togo in West Africa remained unsuitable from LGM to present for all species.

The averaged MOP index identified several areas of heightened or strict extrapolation across taxa, none of which significantly impacted our study system (Figures 1.3 and 1.4). Areas of extrapolation included the Sahara, Kalahari and Namib deserts and the westernmost tip of west Africa. These areas should be interpreted with caution due to increased model uncertainty (Owens et al. 2013), but do not affect the results presented above.

1.5 Discussion

The extensive literature regarding terrestrial montane biogeography of the African continent historically has identified three potential routes for biotic exchange among Afromontane forest regions during the Pleistocene glaciations and the Holocene (Hall, 1973; Moreau, 1966; White, 1981). All three have been invoked to explain the unusual floral and faunal similarity among the Cameroon Volcanic Line, East African Rift and, Great Escarpment (Linder, 1983; Moreau, 1963, 1966; White, 1981, 1983, 1993; Figure 1.1). However, the location, extent, and even existence of these corridors have been debated (Coetzee &Van Zinderen Bakker, 1970; Hedberg, 1969; Wickens, 1976), particularly given the difficulty of obtaining pollen core data from Central Africa. Here, we investigated the climatic feasibility of these habitat corridors using species ecological niche modeling for a variety of Afromontane tree, shrub, and bird species during the Last Glacial Maximum (~22,000 yrs. ago), the mid-Holocene (~6000 years ago), and the present.

It is likely that several of the species in this study represent cryptic species complexes. For example, the west African population of *Cisticola chubbi* is considered by some authors to

be a separate species, *C. discolor* (e.g., Languy & Motombe, 2003; Maisels & Forboseh, 1999) but is considered conspecific by others based on vocal data (Urban et al., 1997). Few dated phylogeographic studies have been done on many of these taxa, but all have exhibited patterns of diversification during the Pleistocene (Cooper et al., 2021; Kadu et al., 2011; Meseguer et al., 2013; Migliore et al., 2020). Earlier divergence dates would not be unexpected (e.g., Mikula et al., 2021; Pokorny et al., 2015; Taylor et al., 2014; Ting, 2008) but also would not preclude the possibility of gene flow between these montane areas during the glaciation cycles (Feder et al., 2012; Martin et al., 2013; Nosil, 2008). If these taxa represent widespread species, as opposed to species complexes, they may be more generalist than the small range endemics found in the same mountain ranges (Büchi & Vuilleumier, 2014; Hanski & Gyllenberg, 1997; Jetz et al., 2007; Wilson et al., 2004), or at least have much greater dispersal abilities (Büchi & Vuilleumier, 2014; Lester et al., 2007). These abilities may have allowed them to take advantage of a wider range of potentially suitable habitats connecting different Afromontane regions, especially if these species are able to inhabit both high elevation rainforest and grassland (Jetz et al., 2007; Wilson et al., 2004).

Although we set out to discriminate among alternate positions of hypothesized corridors of suitable climate north and south of the Congo Basin during the LGM, we found widespread climatic suitability connecting Afromontane regions across the entire Congo Basin for all taxa included in this study (Figures 1.3 and 1.4). Evidence pointing to the existence of climatically suitable corridors, particularly the southern corridor, was available for some species during the mid-Holocene (Figures 1.3 and 1.4), when the global climate was approximately midway between the present and the LGM, suggesting that climate was suitable for some species to expand their geographic ranges through biotic exchange among Afromontane regions as recently

as 6000 years ago. A similar pattern was also seen in the corridor proposed to link the Eastern Arc Range and the Great Escarpment (Figure 1.3), where climatic connectivity between these ranges existed for some species during the mid-Holocene, and widespread connectivity was available for all species during the LGM (Figures 1.3 and 1.4).

The possibility of widespread habitat suitability for high elevation forest types in lowland Central Africa during the Pleistocene aligns with what little is known about this region from pollen records. Pollen from the genus *Podocarpus*, normally associated with montane cloud forest, has been detected in palynological records from the Congo River fan as early as 1.05 million years ago (Dupont et al., 2001). Records of *Podocarpus* and another montane species used in this study, *Ilex mitis*, have been found in pollen cores at 700m in the Batéké plateau in the Republic of the Congo that date back to the LGM (Elenga et al., 1994). *Podocarpus* pollen continued to be found in lowland Congo Basin palynological records until the beginning of the late-Holocene and occurred simultaneously with pollen from wet evergreen taxa (Maley, 1997). These palynological findings, as well as modelled broad-scale vegetation changes across Africa, appear to be closely linked to glacial cycles (Anhuf et al., 2000, 2006; deMenocal, 2004; Dupont et al., 2001; Maley, 1997) and, as a result, the extension of *Podocarpus* into lower elevations likely occurred synchronously between Central and East Africa (Maley, 1997). Although only one high elevation genus is strongly represented in pollen core records, this genus is indicative of shifts in entire Afromontane forest communities (Dupont et al., 2001; Maley, 1997). As a dioecious species, *Podocarpus* is prone to overrepresentation in the palynological records compared to many other high elevation tree and shrub species which, depending on their pollination syndromes, may be palynologically silent (Bush & Rivera, 2001; Dupont et al., 2001;

Farwig et al., 2006; Janeček et al., 2007; Johnson & Brown, 2004; Newmark et al., 2020; Nsor et al., 2019).

More broadly across Africa, some areas that present climatic barriers to biotic exchange today continued to present barriers throughout the mid-Holocene and the LGM, including the Dahomey Gap, a relatively dry, savannah region separating the upper and lower Guinean rainforests of West Africa (Figures 1.3 and 1.4). Pollen cores indicate that this savannah has remained stable since at least the mid-Holocene (Marchant & Hooghiemstra, 2004; Salzmann & Hoelzmann, 2005), but phylogeographic analyses with species divergence dates estimate that this savannah likely has been present since the Miocene (e.g., Demenou et al., 2018; Fuchs & Bowie, 2015). The Ethiopian Highlands also remained separated, in this case by a climatic barrier, from the rest of the Afromontane archipelago (Figures 1.3 and 1.4), possibly explaining the biological distinctness of this region (Poynton, 1999; Prigogine, 1987; Williams et al., 2004), and the relatively deep genetic divergences inferred among its endemic species and their close relatives in other Afromontane areas (Reyes-Velasco et al., 2018; Šmíd et al., 2019; Zinner et al., 2018).

Although correlative ecological niche models effectively characterize the geographic positions and extent of areas of climatic suitability, they do not allow us to determine direction of dispersal, historical timeline of colonization, nor polarity of range expansion. However, several dated phylogeographic studies of Afromontane species have not only suggested similar areas of habitat suitability, but also included directionality. For instance, a phylogeographic studies of the Afromontane tree species, *Prunus africana* and *Podocarpus latifolius*, suggested colonization of western Central Africa from East Africa through a northern dispersal route for *P. africana*, and through northern and southern dispersal routes for *P. latifolius* with divergence dates around the LGM for both species (Kadu et al., 2011; Migliore et al., 2020). A study of distributional

patterns of the Afromontane avifauna also suggested a general east-to-west pattern of range expansion (Prigogine, 1987). Phylogenetic data appear to support this pattern, with Cameroonian Afromontane bird populations from a variety of species embedded within Angolan-East African clades (Vaz da Silva, 2015), and a Pleistocene divergence between East African and Angolan Turaco species (Njabo & Sorenson, 2009). A phylogeographic study of *Chlorocebus* monkeys suggested faunal exchange via the northern route during the LGM but, in this case, dispersal from the west, resulting in colonization of the east (Tosi, 2008). Dispersal from the East African Rift southward to the Great Escarpment was supported by a phylogeographic study on *Lygodactylus* geckos (Travers et al., 2014), from the Eastern Arc Mountains southwards to the Rift Mountains in *Arthroleptis* frogs (Blackburn & Measey, 2009) and from the Great Escarpment northward to the East African Rift for *Euryops* plants (Devos et al., 2010), likely in all cases during globally cool-climate periods. These previous studies thus support the existence of Pleistocene habitat corridors among Afromontane regions and suggest multi-directional exchanges among regions.

Studies of the effects of Pleistocene glaciations on tropical montane taxa are only now accumulating, resulting in remarkable insights into some of Africa's most understudied, highly endemic, terrestrial flora and fauna—much of which has been formally classified as high-value conservation urgency (Blackburn & Measey, 2009; IUCN, 2020; Kadu et al., 2011; Mittermeier et al., 2011; Myers et al., 2000; Schmitt et al., 2010; Vaz da Silva, 2015). Although the overwhelming majority of research in the tropics has focused on diversification in hypothesized lowland rainforest refugia (Born et al., 2011; Budde et al., 2013; Daïnou et al., 2010; Dauby et al., 2010; Duminil et al., 2015; Faye et al., 2016; Fjeldså & Bowie, 2008; Gomez et al., 2009; Hardy et al., 2013; Ley et al., 2014, 2016; Lowe et al., 2010; Silva et al., 2019), this study

suggests that the Pleistocene climatic oscillations also served as an historical opportunity for faunal admixture, and possibly even a homogenizing period among disparate montane areas, at least for species included here. These results accord well with predictions developed by Peterson and Ammann (2013), in which certain world regions exhibited phases of isolation and speciation during the glacial periods (e.g., the Amazon Basin) and other regions were characterized by isolation/speciation phases during the *inter*-glacial periods (e.g., tropical montane regions).

Although the extent of potential habitat change in the Central African lowlands suggested by our models is high, our general results are in line with previous studies based on pollen core data which have concluded a spread of montane forest into the lowlands during the LGM, with the caveat that sampling is limited (Caratini & Giresse, 1979; Dupont et al., 2001; Maley, 1987, 1989, 1997; Van Zinderen Bakker & Clark, 1962; reviewed in Couvreur et al., 2021), and with recent global paleoclimate models that suggest a transition from tropical rainforest to warm temperate woodland and temperate broadleaf evergreen forest in Central Africa during the LGM (Allen et al., 2020). These recent models are a distinct shift away from previous hypotheses of a transition from lowland rainforest to strict savannah in Central Africa during the Pleistocene (Anhuf, 2000, 2006; Coetzee, 1964; van Zinderen Bakker & Coetzee, 1972), with some authors now suggesting that a widespread aridification and expansion of grasses during the glaciation cycles may have been generally restricted to eastern Africa (Ravelo et al., 2004; reviewed in Couvreur et al., 2021). Dupont et al (2001) notes that it is difficult to distinguish if grass (Poaceae) and sedge (Cyperaceae) pollen in his Central African palynological records originated in swampy vegetation along rivers, or in open savannah/woodland, and suggested that abundant *Podocarpus* pollen in lowland Central Africa during glacial cycles supported a cool, yet humid

climate. Our results, based on three widely used general circulation models, also suggest that these climate projections support a cool, humid climate for Central Africa during the LGM.

In addition to testing classical hypotheses of Afromontane biogeography, in this study we also demonstrate ways in which paleo-distributional projections of ecological niche models can be employed for practical purposes, or strategically employed in particular geographic regions where *in situ* data collection has been historically challenging or is logistically unfeasible today. Finally, the results presented here suggest fruitful opportunities for field studies and surveys targeting the unstudied, or intervening highlands among Central Africa's reasonably well-known montane biotic regions. This study lends urgency to the task of collecting corroborating pollen-core data, as well as safe-guarding the few formally designated protected areas of the central Afromontane highlands.

1.6 Figures

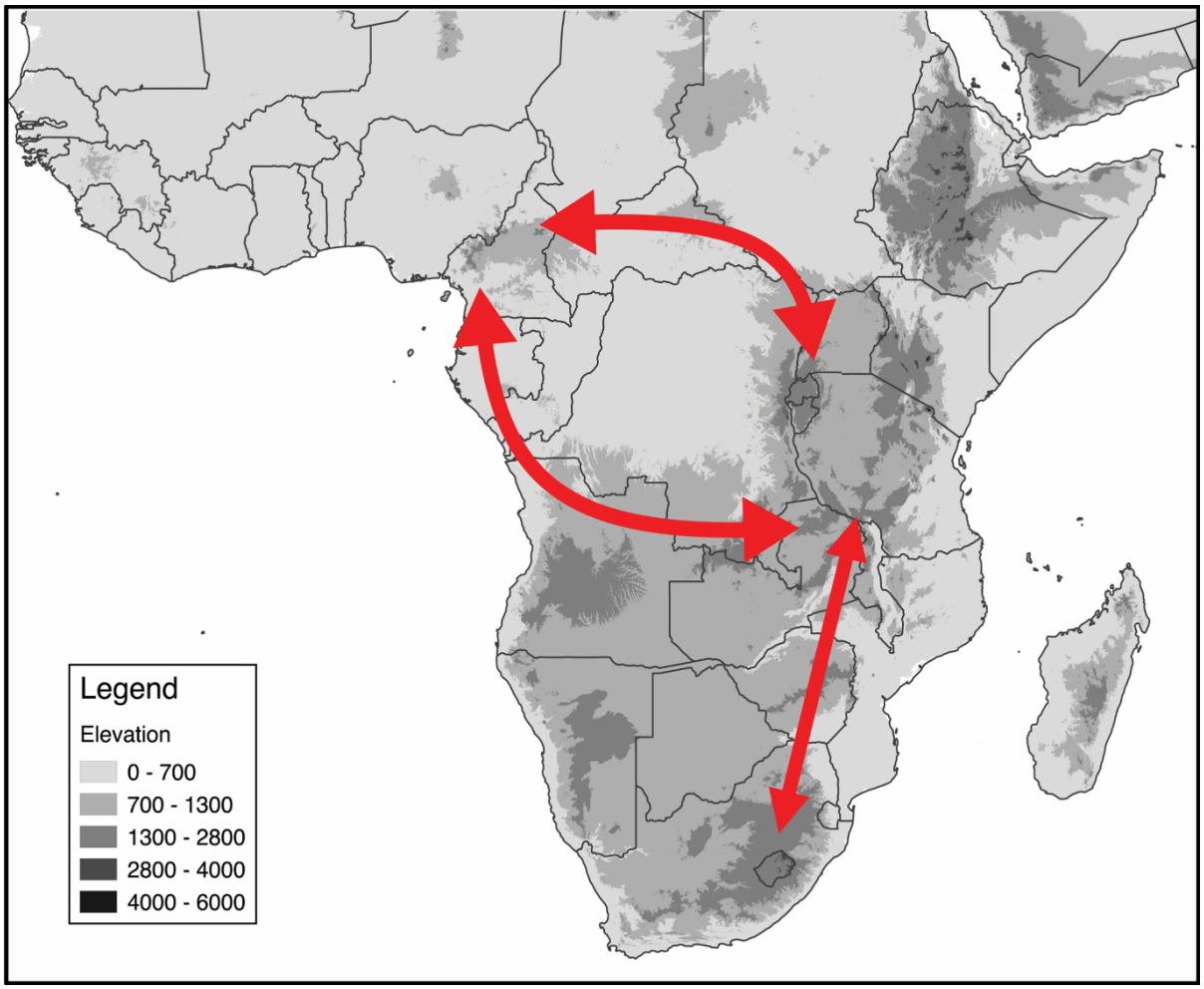


Figure 1.1 Three hypothesized habitat corridors between western, eastern, and southern Afromontane regions.

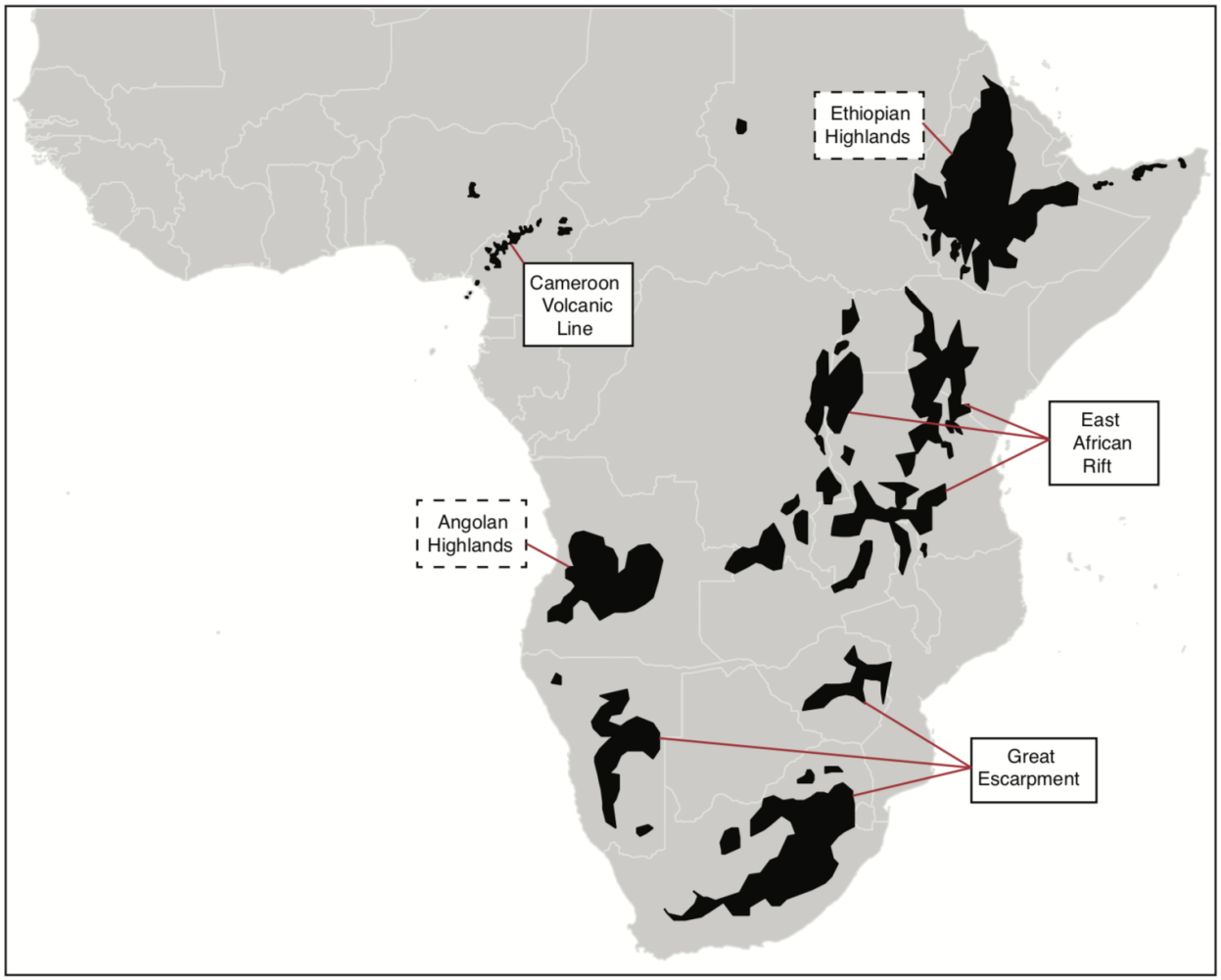


Figure 1.2 The Afromontane archipelago. The Cameroon Volcanic Line, East African Rift, and Great Escarpment were the focus of this study, but the Angolan Highlands and the Ethiopian Highlands were included in the distributional areas of several of our focal species.

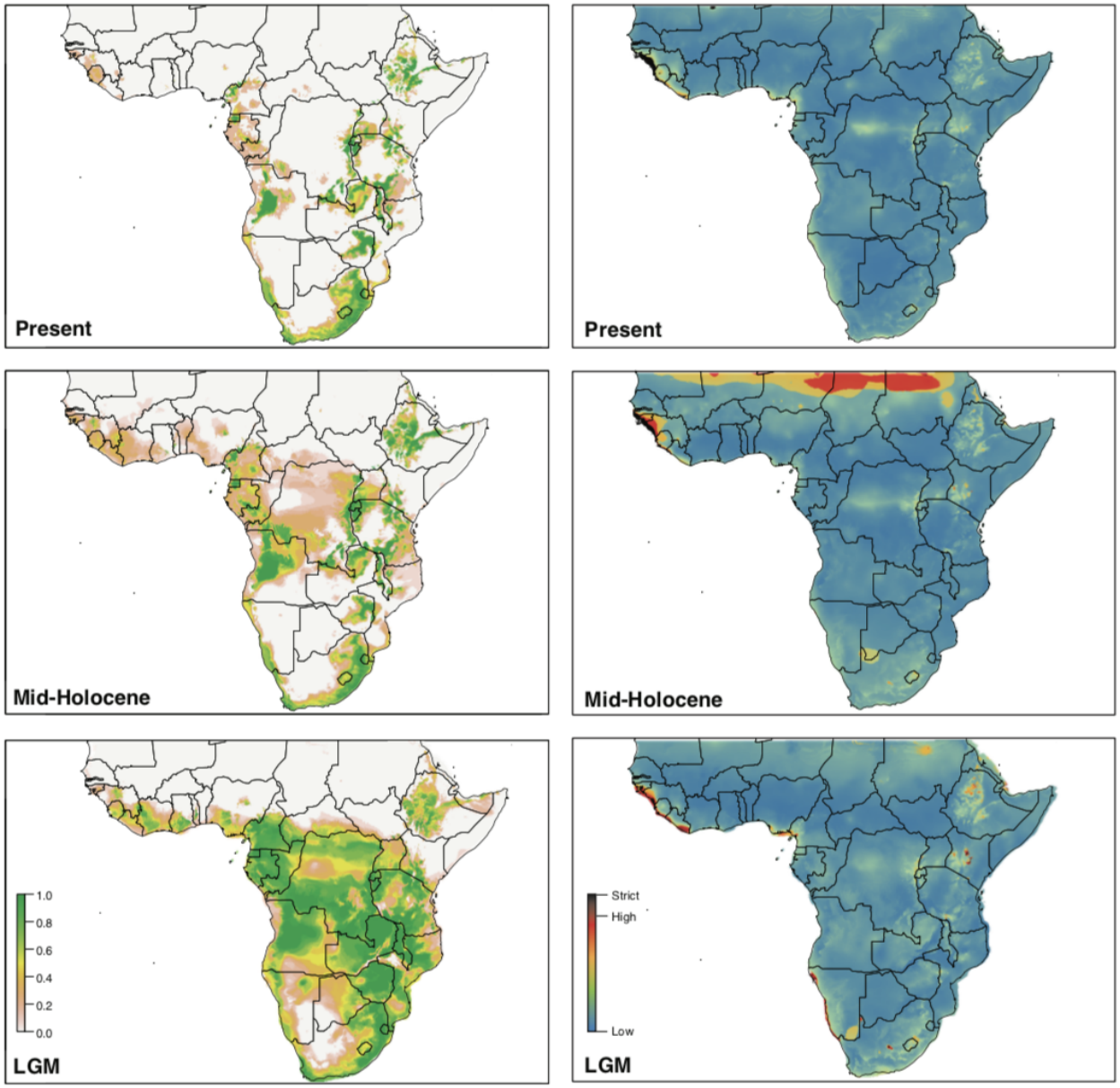


Figure 1.3 Left: Present-day and paleo-projections of ecological niche models for a variety of Afromontane tree and shrub species during the present, the mid-Holocene, and the Last Glacial Maximum. The scale bar represents the proportion of study species found in that area for each time period. Right: An averaged MOP analysis for extrapolation for each of the three time periods. The scale bar represents the extrapolation risk from strict to low extrapolation.

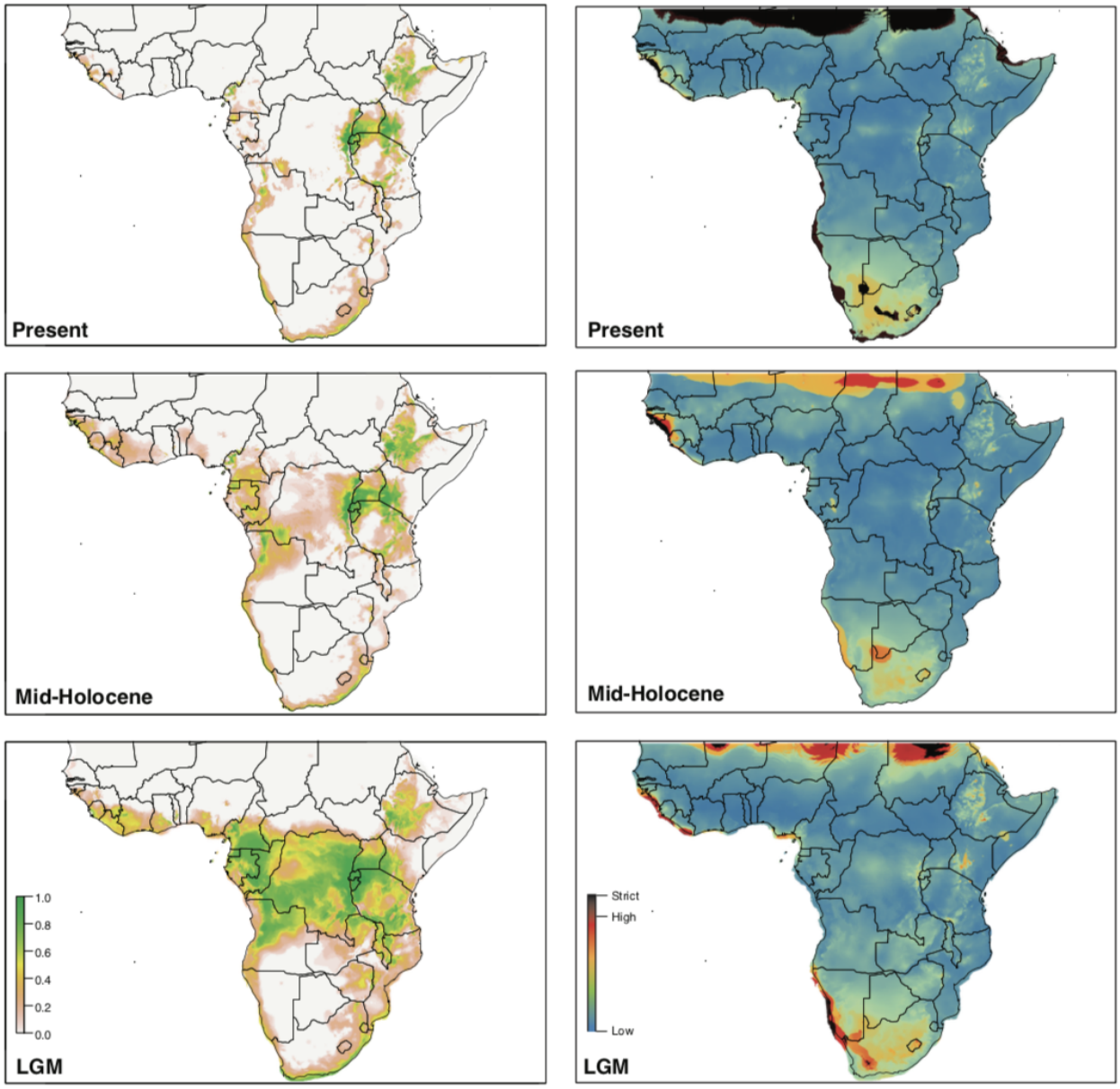


Figure 1.4 Left: Present-day and paleo-projections of ecological niche models for a variety of Afromontane avifauna during the present, the mid-Holocene, and the Last Glacial Maximum. The scale bar represents the proportion of study species found in that area for each time period. Right: An averaged MOP analysis for extrapolation for each of the three time periods. The scale bar represents the extrapolation risk from strict to low extrapolation.

CHAPTER 2

Rivers, not refugia, drove diversification in arboreal, sub-Saharan African snakes

Allen, K.E., Greenbaum, E., Hime, P.M., Tapondjou N., W.P., Sterkhova, V.V., Kusamba, C., Rödel, M.-O., Penner, J., Peterson, A.T., Brown, R.M. 2021. Rivers, not refugia, drove diversification in arboreal, sub-Saharan African snakes. Ecology and Evolution. 11, 6133–6152.

2.1 Abstract

The relative roles of rivers versus refugia in shaping the high levels of species diversity in tropical rainforests has been widely debated for decades. Only recently has it become possible to take an integrative approach to test predictions derived from these hypotheses using genomic sequencing and ecological niche models with projections to historic climates. Herein, we tested the predictions of the classic river, refuge, and river-refuge hypotheses on diversification in the arboreal sub-Saharan African snake genus *Toxicodryas*. We used dated phylogeographic inferences, population clustering analyses, demographic model selection, and ecological niche models with projections to the LGM and mid-Holocene to conduct a phylogenomic and historical demographic analysis of this genus. Our results revealed significant population genetic structure within both *Toxicodryas* species, corresponding geographically to river barriers and divergence times from the mid-Miocene to Pliocene. Our demographic analyses supported the interpretation that rivers are indications of strong barriers to gene flow among populations since their divergence. Additionally, we found no support for a major contraction of suitable habitat during the Last Glacial Maximum, allowing us to reject both the refuge and river-refuge

hypotheses in favor of the river barrier hypothesis. Based on conservative interpretations of our species delimitation analyses with the Sanger and ddRAD data sets, two newly described cryptic species are identified from east-central Africa. This study highlights the complexity of diversification dynamics in the African tropics and the advantages of integrative approaches to studying speciation in tropical regions.

2.2 Introduction

Three major allopatric diversification mechanisms have been proposed in the classical literature to explain species diversity in the tropics. The "river hypothesis" in which species and populations diverged across riverine barriers (Ayres & Clutton-Brock, 1992; Bates, 1863; Hershkovitz, 1977; Mayr, 1942; Sick, 1967; Wallace, 1853); the "refuge hypothesis" in which forests fragmented during the cold, dry Pleistocene glaciation cycles, causing isolation and divergence in small forest patches (Haffer, 1969, 1974, 1982; Prance, 1982; Vanzolini, 1973; Vanzolini & Williams, 1970); and an amalgamate "river-refuge hypothesis," in which speciation was promoted by a combination of river barriers and climate-driven vegetation changes (Ayres & Clutton-Brock, 1992; Haffer, 1992, 1993). These hypotheses have been widely employed as the context for studies of Neotropical biodiversity and the mechanisms of its production (e.g., Gascon et al., 2000; Haffer, 2008; Patton & Silva, 2005; Richardson et al., 2001; Weir, 2006). However, because the early scientific focus was primarily on the Amazon (Amorim, 1991; Cracraft, 1985; DeMenocal, 2004; Haffer, 1969, 1997; Plana, 2004; but see Fjeldså, 1994; Mayr & O'Hara, 1986) and given political instability in tropical Africa (Greenbaum, 2017; Siddig, 2019; Tolley et al., 2016), rigorous testing of the predictions stemming from these hypotheses has been neglected for the West and Central African rainforests until only recently.

Based on pollen core records (Brenac, 1988; Bonnefille & Riollet, 1988; Girese et al., 1994; Maley, 1987, 1989, 1991; Maley & Brenac, 1987; Maley & Livingstone, 1983; Sowunmi, 1991) and species distribution data (Colyn, 1987, 1991; Rietkerk et al., 1995; Richards, 1963; Sosef, 1991), Maley (1996) proposed several Pleistocene rainforest refugia for sub-Saharan Africa that are still considered today (e.g., Bell et al., 2017; Hughes et al., 2017; Huntley et al., 2019; Jongsma et al., 2018; Larson et al., 2016; Penner et al., 2011, Portik et al., 2017; Figure 2.1). Many of these hypothesized refugia are located in highland areas (e.g., the Cameroon Volcanic Line and the Albertine Rift; Figure 2.1, refugia 4 and 10 respectively). However, a major fluvial refuge, i.e., a refugium associated with a river, located in the gallery forests around the Congo River (Figure 2.1, refugium 9), has been supported by pollen core data (Maley, 1996), and distributional patterns of multiple bird (Huntley et al., 2018; Levinsky et al., 2013), mammal (Colyn et al., 1991; Levinsky et al., 2013), and plant taxa (Robbrecht, 1996).

Major river barriers in West and Central Africa include the Volta, Sanaga, Ogooué, Congo, Niger, and Cross Rivers (Figure 2.1). The exact ages of many of these rivers are unknown but are generally estimated to date back to the Late Mesozoic to Early Cenozoic (80–35 mya; Couvreur et al., 2021; Goudie, 2005; Stankiewicz & de Wit, 2006). However, while the Congo basin is quite old (Flügel et al., 2015; Stankiewicz & de Wit, 2006), the present course of the Congo River appears to be much younger, dating to the mid to late Miocene and corresponding to the uplift of the East African Rift (Flügel et al., 2015; Stankiewicz & de Wit, 2006; Takemoto et al., 2015).

Numerous phylogeographic studies have supported the importance of rivers, refugia, or both as drivers of diversification across disparate plant and animal species. Rivers alone have been shown to be important barriers for some species of primates (Mitchell et al., 2015; Telfer et

al., 2003), shrews (Jacquet et al., 2015), and frogs (Charles et al., 2018; Penner et al., 2011, 2019; Wieczorek et al., 2000; Zimkus et al., 2010), but do not appear to represent an important barrier for many plant species (Dauby et al., 2014; Debout et al., 2011; Hardy et al., 2013; Ley et al., 2014; Lowe et al., 2010). Refugia are suggested to have played an important role in the diversification of rodents (Bohoussou et al., 2015; Nicolas et al., 2011; Nicolas et al., 2012), primates (Clifford et al., 2004; Haus et al., 2013; Tosi, 2008), frogs (Bell et al., 2017; Jongsma et al., 2018), lizards (Allen et al., 2019; Leaché et al., 2017), birds (Fjeldså & Bowie, 2008), pangolins (Gaubert et al., 2016), and rainforest plants (Born et al., 2011; Budde et al., 2013; Daïnou et al., 2010; Dauby et al., 2010; Duminil et al., 2015; Faye et al., 2016; Gomez et al., 2009; Hardy et al., 2013; Ley et al., 2014, 2016; Lowe et al., 2010). In some cases, divergence patterns match both refugial and riverine predictions (Anthony et al., 2007; Barej et al., 2011; Bohoussou et al., 2015; Gonder et al., 2011; Jacquet et al., 2014; Jongsma et al., 2018; Leaché et al., 2019; Leaché & Fujita, 2010; Marks, 2010; Portik et al., 2017), suggesting that both may have played roles simultaneously—or in combination—in evolutionary diversification. However, because of the spatial overlap of refugia with montane and riverine systems (Hofer et al., 1999, 2000), and the sparse pollen core and fossil records for the tropics (Colinvaux et al., 1996; Maley & Brenac, 1998), distinguishing between these three hypotheses has been difficult, especially when relying on phylogeographic data alone.

The three major allopatric diversification hypotheses make the following predictions regarding species diversification patterns in tropical African forests (1) river hypothesis: boundaries between population distributions should correspond to riverine barriers and the ages of populations should be relatively old, corresponding to the ages of the rivers; (2) refuge hypothesis: population distributions should be concordant with locations of hypothesized

rainforest refugia during cold, dry periods and populations are predicted to be relatively young, corresponding to the Pleistocene glaciation cycles; (3) river-refugia hypothesis: population distributions should be correlated with the locations of rainforest refugia and bounded by rivers barriers, or will have been confined to refugial locations and additionally subdivided by rivers. Under this scenario the timing of population splits should correspond to ages of rivers but would be expected to show patterns of range expansion and contraction for niche model projections to the Pleistocene.

In this study, we use the snake genus *Toxicodryas* as a model system to test multiple predictions derived from these hypotheses. The genus *Toxicodryas* consists of two species of large-bodied, rear-fanged, venomous sub-Saharan African snakes, *T. blandingii* and *T. pulverulenta* (Figure 2.2). For most of the 20th century, these taxa were placed in the Asian genus *Boiga* (Schmidt, 1923), and some authors still classify them as such, but recent phylogenetic analyses recover them as the sister genus to the African egg-eating snakes, *Dasypeltis* (Pyron et al., 2013; Weinell et al., 2020). No in-depth phylogenetic or phylogeographic analysis has been done within *Toxicodryas*. Both known species are primarily arboreal, feeding mainly on birds, bats, frogs, and arboreal lizards (Akani et al., 1998; Chippaux & Jackson, 2019; Nagy et al., 2011; Spawls et al., 2018). Because of their general arboreality, these species are predicted to have distributions strongly correlated with forest and woodland habitats. In addition, *Toxicodryas* is widely distributed within the Congo Basin and broadly across West and Central Africa, making this genus a suitable system for testing the competing predictions of the river, refugia, and river-refugia hypotheses.

Recent advances in paleo-climate modeling and genome-scale DNA sequencing have opened new avenues to testing classic hypotheses of tropical rainforest speciation (Bell et al.,

2017; Leaché et al., 2019; Portik et al., 2017). In this study, we integrate dated phylogeographic inference, population structure analyses, and machine learning-based demographic modeling to identify the timing of divergence as well as the location and permeability of past and present dispersal barriers. These genetic data are combined with ecological niche projections and climate stability modeling to determine the congruence of historical distributions with the refugial and river-refugial hypotheses. Our results demonstrate that, although population distributions alone could be congruent with any of the three hypotheses, diversification times predate the Pleistocene, a finding that aligns with predictions of the river-barrier hypothesis. Moreover, historical demographic analyses support models of no migration among populations since the time of divergence, and no signatures of population bottleneck and subsequent expansion were identified, as predicted under the refuge hypothesis. Additionally, ecological niche projections and climate stability modeling show no suggestion of suitable habitat contraction during or since the Pleistocene. Together, our analyses allow us to reject predictions of refugia hypotheses in favor of the prevailing role of riverine barriers in shaping, structuring, and maintaining diversity in this generally arboreal, forest-associated group of endemic African snakes.

2.3 Materials and Methods

2.3.1 Sampling

We obtained 20 specimens of *Toxicodryas* (seven *T. blandingii* and 13 *T. pulverulenta*) through fieldwork and from various museums (see Appendix I Table S1). Sampling was representative of the known range of each species throughout the upper and lower Guinean forest blocks of West and Central Africa including the countries of Guinea, Liberia, Ghana, Cameroon,

Gabon, and Democratic Republic of the Congo (DRC). Museum catalog numbers, GenBank accession numbers, and locality data for each specimen are presented in Appendix I Table S1.

2.3.2 Genetic data collection and processing

Tissue samples were preserved in 95% ethanol or RNAlater™ (Sigma-Aldrich) and genomic DNA (gDNA) was extracted using the Maxwell RSC system (Promega). The nuclear gene c-mos and the mitochondrial gene cytochrome *b* (cyt *b*) were PCR-amplified for each individual using standard primers (c-mos: S67, S68; Lawson et al., 2005; cyt *b*: L4910B, H15720; Burbrink et al., 2000) and sequenced on an ABI 3730 capillary electrophoresis system (Applied BioSciences®). Electropherograms were edited manually in Geneious v5.6.7 (http://www.geneious.com, Kearse et al., 2012) and resulting sequences were aligned in MAFFT v.5 with default parameters (Katoh & Kuma, 2002).

We also sequenced genome-wide anonymous nuclear markers for each individual following a modified version of the ddRADseq protocol of Peterson et al. (2012). For each individual, a total of 300–500 ng of gDNA was double digested using the restriction enzymes *Sbf*I (restriction site 5'- CCTGCAGG-3') and *Msp*I (restriction site 5'-CCGG-3'). The resulting double digestion products were then bead-cleaned with AmpureXP beads (Agencourt) and individually barcoded using custom oligonucleotide adapters. Pooled samples were size-selected to a mean insert length of 541 base pairs (bp) (487–595 bp range) with internal standards with a Pippin Prep™ (Sage Science, Beverly, MA). Resulting post-ligation products were amplified for eight cycles with a high-fidelity polymerase (Phusion™, New England Biolabs). An Agilent TapeStation was used to determine the final fragment size distribution and concentration of each

pool. Library pools were combined in equimolar amounts for sequencing on one Illumina HiSeqX lane (with a 10% *Phi*X spike-in and 150 bp paired-end reads).

Illumina reads from the ddRAD libraries were processed using STACKS v. 2.4 (Catchen et al., 2013). Because the ddRAD protocol generates strand-specific libraries, prior to read filtering and assembly, we used a read-stitching approach (Hime et al., 2019) to join the first read from an Illumina read pair with the reverse complement of the second, recapitulating the original orientation of fragments in the genome. Stitched reads were quality-filtered and demultiplexed by individual with the process_radtags function in STACKS with the following parameters: demultiplex each library by in-line barcode, check for both restriction enzyme cut sites, remove any read with an uncalled base, rescue barcodes and RAD-Tags, and discard any read with average Phred quality score < 20 over sliding windows of 15% of the total read length. Next, we used STACKS to *de novo* assemble filtered and stitched Illumina read pairs.

We aimed to produce three separate ddRAD data sets, including one for *T. blandingii*, one for *T. pulverulenta*, and a combined data set comprising both species. Because the optimal *de novo* assembly of ddRADseq data can vary widely across taxa (Paris et al., 2017; Shafer et al., 2017), we tested a range of assembly parameters to optimize the recovery of putatively single-copy orthologous loci. Final assembly parameters were selected based on the methods of Paris et al. (2017). According to their recommendations, in USTACKS, we kept m (the minimum number of reads needed to form a stack) at 3 while in CSTACKS, we varied M (the number of mismatches allowed during loci formation) and n (the number of mismatches allowed during catalog formation) until we identified the parameters at which the maximum number of polymorphic loci were available across 80% (r = 0.8) of the population. For our data, this was found to be M = 5 and n = 15. Further parameters were tested in POPULATIONS separately for

each species and for the genus as a whole in order to balance missing data and number of polymorphic loci. Within *T. blandingii* and *T. pulverulenta*, the percent missing data was low (5% and 7.3% missing data respectively) and no further processing was needed, and r = 0.8 was used. Because of dissimilarity between the two species causing high levels of missing data in the combined data set, further restrictions were implemented. For the genus-wide data set, we set r = 0.5 and p = 4 [p is the minimum number of populations in which a locus must be present (here 4/5)]. This approach increased the number of informative loci, but also the amount of missing data. For each of our three separate data sets, we generated a data set comprising only a single random SNP per locus (for population clustering analyses and demographic modeling), and another data set comprising full-length sequences for all loci (for use in phylogenetic reconstruction).

2.3.3 Phylogenetic analyses

We concatenated our Sanger data set (c-mos and cyt *b*) and implemented bModelTest in Beast v. 2.6.2 to assess all possible substitution models for each gene using a Bayesian approach (Bouckaert & Drummond, 2017). We conducted a time-calibrated analysis on our partitioned data set in Beast v. 2.6.2 (Bouckaert et al., 2019) using a relaxed log-normal clock and a Yule tree prior assuming a constant lineage birth rate. Dating analyses were based on three fossils for calibration in the outgroup, one at the Elapoidea + Colubridae node (minimum age: 30.9 Mya), one at the *Heterodon* + *Farancia* node (minimum age: 12.08 Mya), and one at the *Naja* + *Hemachatus* node (minimum age: 17.0 Mya). Fossil ages and placement were based on Head et al. (2016). Two runs of 100,000,000 generations were conducted and logged every 10,000 generations. Convergence was assessed using Tracer v. 1.7 (Rambaut et al., 2018). A burn-in of

10% was used to create a maximum clade credibility tree. Node ages are based on median tree heights.

We analyzed our SNP data set, including all samples of both species of *Toxicodryas*, using both species-tree summary quartet and maximum likelihood phylogenetic methods. The quartet method was implemented through SVDquartets (Chifman & Kubatko, 2014) in PAUP* v. 4.1a166 (Swofford, 2003). We sampled all possible quartets and assessed support using 100 nonparametric bootstraps, and species tree topology was summarized with DendroPy v. 4.4.0 (Sukumaran & Holder, 2010). We ran a maximum likelihood analysis of our genus-wide SNP data set in IQtree v. 1.6.12 (Nguyen et al., 2014) using 10,000 ultrafast bootstraps (Hoang et al., 2018) and the ModelFinder function to choose the best substitution model (Kalyaanamoorthy et al., 2017). As no outgroups were included in our SNP data set, for both SVDquartets and IQtree the placement of the root of each phylogeny was chosen to match that of the Sanger phylogeny.

2.3.4 Assessing genetic structure

We used multivariate, Bayesian, and admixture-based analyses to assess population structure. In all analyses, clustering algorithms were run on three data sets separately for comparison (*T. blandingii*, *T. pulverulenta*, and both species combined [genus *Toxicodryas*]). A discriminant analysis of principal components (DAPC) was run using Adegenet v. 2.1.1 (Jombart & Ahmed, 2011). This approach uses discriminant functions to maximize variation among clusters and minimize variation within clusters. The best clustering scheme was chosen based on Bayesian information criterion (BIC) scores. Numbers of clusters (K) ranging from 1–10 for the genus level analysis and 1–5 for each species were evaluated and a discriminant function analysis of principal components (DAPC) was performed based on the number of suggested

clusters. Ancestry proportions of all individuals were inferred using LEA v. 1.6.0 (Frichot & François, 2015) through the Bioconductor v. 3.4 package. The sNMF function was used to assess K values from 1–10 for the genus level analysis and 1–5 for each species, with 20 replicates, estimated individual admixture coefficients, and selected the value of K that minimized cross entropy (François, 2016; Frichot et al., 2014). Population structure and admixture were also tested using the Bayesian method STRUCTURE v. 2.3.4 (Falush et al., 2003; Pritchard et al., 2000). Each data set was evaluated for K=1–10 for the genus level analysis and 1–5 for each species, with 10 runs per K and a MCMC burn-in of 10,000 steps followed by 100,000 steps (Porras-Hurtado et al., 2013). Results were evaluated using the Evanno method (Evanno, 2005) and plotted through the R package pophelper v. 2.3.0 (Francis, 2017).

2.3.5 Demographic modeling

To test for historic and recent gene flow between our populations, we used the R package delimitR (Smith & Carstens, 2020; https://github.com/meganlsmith/delimitR). This program uses a binned multidimensional folded site frequency spectrum (bSFS; Smith et al., 2017) and a random forest machine learning algorithm to compare speciation models such as no divergence, divergence with and without gene flow, and divergence with secondary contact (Smith & Carstens, 2020). A bSFS was used because it stores the observed frequencies of the minor alleles for multiple populations and bins them to avoid inference problems associated with sampling too few segregating sites (Smith et al., 2017; Terhost & Song, 2015). Demographic histories are simulated using the multi-species coalescent model implemented through fastsimcoal2 (Excoffier et al., 2013) under a user-specified guide tree and set of priors on divergence times, population sizes, and migration rates. The random forest classifier then creates a user-defined

number of decision trees from a subset of the prior. Each decision tree compares the empirical bSFS to the SFS of each simulated speciation model and votes for the most similar model. The demographic model with the largest number of votes is chosen as the best model. Out-of-bag error rates are used to assess the power of the random forest classifier. The posterior probability of the selected model is then calculated by regressing against the out-of-bag error rates following Pudlo et al. (2015).

We created folded multi-dimensional site frequency spectrums for the two *T. blandingii* clades and the two Central African *T. pulverulenta* clades using easySFS (https://github.com/isaacovercast/easySFS), a wrapper for ∂a∂i (Gutenkunst et al., 2009). The West African *T. pulverulenta* clade was not included because of the low sample size available for this lineage. We simulated 100,000 data sets under four models: no divergence (Model 1), divergence without gene flow (Model 2), divergence with secondary contact (Model 3), and divergence with gene flow (Model 4). Priors for both models were drawn from uniform distributions for population size: 10,000 – 1,000,000 haploid individuals (twice the number of estimated diploid individuals), divergence time: 20,000 –2,000,000 generations, migration rate: 0.000005–0.005 corresponding to 0.05–5 migrants per generation. We then coarsened our empirical site frequency spectra to 10 bins each. Our out-of-bag error rates were calculated, and 500 random-forest classifiers were simulated using 100,000 pseudo-observed data sets for each model. A confusion matrix was calculated to determine how often the correct model was selected and posterior probability for the "best" model was estimated for each species.

The R package rangeExpansion (Peter & Slatkin, 2013, 2015) was used to assess signatures of population size change in the two *T. blandingii* clades and the two Central African *T. pulverulenta* clades. The West African *T. pulverulenta* clade was excluded because of its

small sample size. This program implements a founder effect algorithm using a stepping stone model, assuming that each colonizing event is associated with a founder event, to determine whether a population shows signatures of expansion or equilibrium isolation-by-distance (Peter & Slatkin, 2013, 2015). If population expansion is identified, the program will infer the strength of the founder effect and the most likely expansion origin (Peter & Slatkin, 2013, 2015).

2.3.6 Species delimitation

We conducted a species delimitation analysis on our Sanger data set using a Bayesian approach through BPP v.4.2.9 (Flouri et al., 2018; Yang, 2015) and a user specified guide tree (Rannala & Yang, 2013; Yang & Rannala, 2010). Following Yang & Flouri (2020), we used the default prior for theta ($\theta = 0.002$) and calculated the tau prior based on the estimated divergence times ($\tau = 0.036$). We used a burn-in of 20000, a sampling size of 200000, and a sampling frequency of two. The analysis was run twice with a random seed to ensure consistency.

Additionally, the package DelimitR (Smith & Carstens, 2020; https://github.com/meganlsmith/delimitR) was used to conduct a species delimitation analysis on our ddRAD data set. As described above for the demographic analysis, this program uses a binned multidimensional folded site frequency spectrum (bSFS; Smith et al., 2017) and a random forest machine learning algorithm to compare the speciation models: no divergence, divergence with gene flow, divergence with secondary contact, and divergence without gene flow (Smith & Carstens, 2020). Each scenario is simulated using a multi-species coalescent model implemented through fastsimcoal2 (Excoffier et al., 2013) and user specified priors on divergence times, population sizes, and migration rates. The empirical bSFS to the SFS of each simulated

speciation model is then compared by the random forest classifier and posterior probabilities and out-of-bag error rates are calculated.

2.3.7 Ecological niche modeling

Occurrence data for each species were obtained from the specimens used in this study, "expert" identified individual occurrences from the Global Biodiversity Information Facility (GBIF), and research-grade locality records from iNaturalist (www.inaturalist.org) that could be visually identified by the authors. Duplicate records were removed, and points were thinned within a distance of 10 kilometers using the spThin package (Aiello-Lammens et al., 2015) in R v. 3.4.4 (R Core Team, 2018). This resulted in a total of 43 *T. blandingii* localities and 30 *T. pulverulenta* localities (Appendix I Figure S1). A subset of points from each data set was set aside for model calibration (75%) and internal testing (25%) following Cobos et al. (2019).

Environmental data were obtained from the WorldClim database v. 1.4 (Hijmans et al., 2005). Fifteen of the 19 bioclim variables were downloaded at a 2.5-minute resolution. We excluded bio8 (Mean Temperature of Wettest Quarter), bio9 (Mean Temperature of Driest Quarter), bio8 (Precipitation of Warmest Quarter), and bio19 (Precipitation of Coldest Quarter) which, as combinations of other variables, are known to create artifacts in ecological niche models (Escobar et al. 2014). The same 15 variables were used for the Last Glacial Maximum (LGM) under three general circulation models (GCMs): CCSM4, MIROC-ESM, and MPI-ESM-P. In order to reduce spatial autocorrelation, principal component analyses (PCAs) were performed on present bioclim variables and projected to the LGM for the extent of sub-Saharan Africa.

Model calibration areas were defined as a 1,000-kilometer buffer around occurrence points for each species. Model calibration, creation, projection, and evaluation were done using

the R package kuenm (Cobos et al., 2019). In order to calibrate our models, we created 1,479 candidate models for each species by combining three sets of environmental predictors (PCAs 1–6, 1–5, 1–4), 17 possible regularization multipliers (0.1–1.0 at intervals of 0.1, 2–6 at intervals of 1, and 8 and 10), and all combinations of five feature classes (linear = l, quadratic = q, product = p, threshold = t, and hinge = h; Cobos et al. 2019).

Candidate models were run in Maxent (Phillips et al., 2006) and chosen based on significant partial Receiver Operating Characteristic (ROC) scores (Peterson et al., 2008), omission rates of $E \leq 5\%$ (Anderson et al., 2003), and corrected Akaike Information Criterion $AICc \leq 2$ to minimize model complexity (Warren & Seifert, 2011). These models determined the parameter set used for final model creation.

Final models were created for each species using the full set of occurrence records and the parameters chosen during model calibration. Models were run in Maxent with ten bootstrap replicates and logistic outputs. After models were run in the present, they were projected to the LGM and mid-Holocene for the three GCMs. The mobility-oriented parity (MOP) index was used to test for model extrapolation (Soberón & Peterson, 2005). Models were visualized in QGIS 3.4 and thresholded to 5% to create presence-absence maps. Models from each time period were summed to estimate potential LGM and mid-Holocene distributions as well as continuous stability maps (Devitt et al., 2013; Yannic et al., 2014).

2.4 Results

2.4.1 Genetic data collection and processing

Our concatenated c-mos and cyt *b* data set (Sanger data set hereafter) consisted of 1,237 bp, including indels. Both genes were represented in all samples with the exception of c-mos for

the outgroup *Contia longicaudae*. Information on samples used in the Sanger analysis, including museum catalog number and GenBank accession number, can be found in Appendix I Table S1. After filtering (see Methods, above), our genus-level ddRAD data set consisted of 2,848 loci with 20.7% missing data (here defined as proportion of missing loci across all individuals), and an effective mean per-sample depth of coverage of 78.7x ± 13.6x. Our *T. blandingii* data set consisted of 7,231 loci with 5.0% missing data, and an effective mean per-sample depth of coverage of 83.6x ± 12.0x. Our *T. pulverulenta* data set consisted of 4,471 loci with 7.3% missing data, and an effective per-sample mean depth of coverage of 77.9x ± 14.6x. The concatenated ddRAD data set used for phylogenetic analyses had a length of 450,512 bp and 3,024 SNPs.

2.4.2 Phylogenetic structure and divergence dating

Broad-scale phylogenetic relationships estimated in analyses of our Sanger and SNP data sets were identical in topology, with strongly supported internal nodes throughout (Figure 2.3; Appendix I Figure S2). Our two-locus Sanger tree and our 2,848-locus ddRAD SNP trees both supported two divergent lineages of *T. blandingii*, in West and Central Africa, respectively, and three divergent lineages of *T. pulverulenta*, one from West Africa and two in Central Africa, north and south of the Congo River (Figure 2.3). Fossil-calibrated divergence dating suggests that *T. blandingii* and *T. pulverulenta* diverged in the mid-Miocene (median age 12.2 Mya). Diversification within each species is estimated to have taken place primarily in the Pliocene, with the two clades in *T. blandingii* diverging around 4.3 Mya, the West African clade of *T. pulverulenta* diverging around 3.3 Mya, and the two Central African clades diverging around 1.9 Mya (Figure 2.3).

2.4.3 Population structure

A comparison of BIC values from the genus-level DAPC analyses suggested a total of five genetic clusters, with two populations in *T. blandingii* and three in *T. pulverulenta*, matching the clades identified in the phylogenetic analyses (Appendix I Figure S3). Our admixture-based method, LEA, identified two distinct genetic clusters at the genus level, corresponding to the two *Toxicodryas* species, and the same two populations for *T. blandingii* and three populations for *T. pulverulenta* as suggested by DAPC (Figure 2.4). A low amount of admixture was identified in the Cameroonian sample of *T. blandingii*, and varying levels of admixture were suggested for the Gabonese samples of *T. pulverulenta* (Figure 2.4). The population assignment of individuals between the two clustering methods was identical; however, admixture between populations was not detected by DAPC. Similarly, STRUCTURE suggested two populations at the genus level, and two in *T. blandingii*, but combined the Central African clades and suggested two populations, instead of three, for *T. pulverulenta*. Three populations were supported as the second highest ΔK and showed identical admixture proportions to those from LEA. We used five populations for our remaining analyses because multivariate-based analyses such as LEA and DAPC do not make assumptions about Hardy-Weinberg equilibrium and may be preferable over Bayesian methods such as STRUCTURE when sample sizes are small or uneven (Puechmaille, 2016).

2.4.4 Demographic modeling

Using machine learning-based demographic model selection, we identified divergence without gene flow as the best model for *T. blandingii* with a posterior probability of 0.68, and

divergence with gene flow for *T. pulverulenta* with a posterior probability of 0.63 (Figure 2.5). For both species, models representing no divergence and divergence with secondary contact received very low support (Appendix I Tables S2 and S3). The out-of-bag error rate for *T. blandingii* was 17.3% and 22.8% for *T. pulverulenta*, with all of the misclassifications being between highly similar models (i.e. between divergence with or without gene flow, as opposed to between divergence and no divergence). While it is possible that small sample sizes for several of our populations may have negatively impacted the power of our demographic analyses, our values for posterior probability and out-of-bag error rate are similar to those obtained by Smith and Carstens (2020), suggesting that the impact was minimal. The confusion matrix and number of votes per model can be found in Appendix I Tables S2 and S3.

Population size analyses for demographic signatures of range expansion versus equilibrium isolation-by-distance strongly rejected the expansion model for each of the four populations tested (West African *T. blandingii*: p=6.38; Central African *T. blandingii*: p=28.01; Gabon *T. pulverulenta*: p=833.43*;* Congo Basin *T. pulverulenta*: p=0.10). Accordingly, the strength of the founder effect (q) for each population was generally small, and the founder distance (d) was large (West African *T. blandingii*: q=0.000037, d=137.46; Central African *T. blandingii*: q=0.000028, d=180.61; Gabon *T. pulverulenta:* q=0.000042, d=120.65*;* Congo Basin *T. pulverulenta*: q=0.00017, d=29.35).

2.4.5 Species delimitation

We tested five distinct clades of *Toxicodryas*: West African *T. blandingii*, Central African *T. blandingii*, West African *T. pulverulenta* and two Central African clades of *T. pulverulenta* using the Bayesian species delimitation method BPP v. 4.2 (Flouri et al., 2018;

Yang, 2015). Our analysis supported all five clades of *Toxicodryas* as distinct species with high posterior probability (pp=0.98). Our analysis was run twice with random seeds to check for consistency, and both runs gave highly similar results. Similarly, DelimitR identified divergence without gene flow as the best model for *T. blandingii* and divergence with gene flow for *T. pulverulenta*. For both species, models suggesting lack of divergence or present-day gene flow had very low probabilities (Appendix I Tables S2 and S3).

2.4.6 Ecological niche modeling

Ecological niche models suggested widely overlapping ranges for *T. blandingii* and *T. pulverulenta*, with both species documented from both rainforest and woodland habitats (Figure 2.6). Projections to the LGM suggested a slight northern and southern contraction of suitable habitat for the genus in West and Central Africa. *Toxicodryas pulverulenta* showed evidence of a slight southward range expansion into Angola, while the range of *T. blandingii* remained stable (Figure 2.6a). The mid-Holocene distribution was highly similar to the present-day distribution for all data sets (Figure 2.6b).

Continuous climate stability maps estimating the areas of persistent suitable habitat from the LGM to the present suggest that the core distribution of each species has remained stable through time (Figure 2.6c). Instability in suitable habitat is only found on the edges of the species range, with the greatest potential for distribution change in southern Central Africa and northern West Africa.

2.5 Discussion

The relative roles of rivers and refugia in shaping the high levels of species diversity in tropical rainforests has been widely debated for decades (e.g., Amorim, 1991; Colinvaux et al., 2001; DeMenocal, 2004; Haffer, 1969, 1997; Mayr & O'Hara, 1986; Vitorino et al., 2016; reviewed in Couvreur et al., 2021). Only recently has it become possible to take an integrative approach to answering these questions with genomic sequencing and ecological niche projections (Portik et al., 2017; Leaché et al., 2019). Herein we tested alternate predictions of the classic river, refuge, and river-refuge hypotheses for terrestrial faunal diversification using a novel study system: the arboreal African snake genus *Toxicodryas*. We found strong support for predictions derived from the river hypothesis over the refuge and river-refuge hypotheses, based on the ages, locations, and timing of gene flow between each of our populations, as well as a lack of support for suitable habitat and population size contraction during the Last Glacial Maximum.

2.5.1 Species diversification

This study represents the first phylogenetic analysis of the genus *Toxicodryas*. Phylogenetic analyses of our two-locus Sanger data set and 2,848-locus RADseq SNP data set reveal two deeply divergent, strongly supported lineages in *T. blandingii* and three in *T. pulverulenta* (Figure 2.3; Appendix I Figure S2). Although the two recognized species are broadly sympatric, clades within each species are generally situated allopatrically across river barriers. Of the three *T. pulverulenta* clades, one is distributed in West Africa (albeit with limited sampling) and two are distributed in Central Africa, separated by the western Congo River. The boundary between the two clades within *T. blandingii* is delimited somewhere between the Sanaga River in Cameroon and the Congo River in the DRC, and, while the species distribution reaches this area, no genetic sampling is available. Both rivers have frequently been interpreted

as population barriers in other terrestrial vertebrates (Blackburn, 2008; Jongsma et al., 2018; Leaché et al., 2019; Leaché & Fujita, 2010; Portik et al., 2017), and a recent morphological analysis by Greenbaum et al. (2021) of all available *Toxicodryas* material, including vouchers specimens of the genetic samples used herein, has identified the most likely biogeographic barriers within the genus to be the confluence of the Congo and Ubangi Rivers for the Central African populations of *T. blandingii* and *T. pulverulenta*, and the Niger Delta as the barrier for the West and Central African populations of *T. pulverulenta*.

Our population structure analyses are concordant with phylogenetic analyses supporting five distinct genetic clusters (Figure 2.4), and our species delimitation analyses suggest that all of these clusters represent distinct, independently evolving lineages. Minor levels of admixture seem to have occurred between the *T. pulverulenta* clades separated by the western Congo River, and between the two clades of *T. blandingii* in the sample collected at the Sanaga River (Figure 2.4), however, our demographic analyses do not support contemporary gene flow, suggesting that this admixture is a result of historic introgression or incomplete lineage sorting. In both species, the Congo River barrier seems to be stronger in the west where the river is wider, and the current is stronger. In eastern DRC, samples of clades from both species can be found on either side of this river (Figure 2.4)

2.5.2 Divergence time estimates

Divergence time estimates from a time-calibrated phylogeny also support predictions derived from the river-barrier hypothesis. *Toxicodryas blandingii* and *T. pulverulenta* diverged in the mid-Miocene, and subsequent intraspecific diversification took place in the Pliocene (Figure 2.3). If the Pleistocene rainforest refugia hypothesis were supported we would expect

diversification times dating to the Pleistocene glaciation cycles, < 2.5 million years ago. Similar mid-Miocene and Pliocene divergence times have been noted for other widespread West and Central African taxa including frogs (Bell et al., 2017; Jongsma et al., 2018; Zimkus et al., 2017), and terrestrial snakes (Portillo et al., 2019), and similar West to Central African distribution splits have been seen in forest cobras (Wüster et al., 2018), frogs (Leaché et al., 2019), lizards (Allen et al., 2019), and shrews (Jacquet et al., 2015).

Greenbaum et al. (2021) identified two major river systems, the Congo/Ubangi and the Niger Delta, as likely barriers between *Toxicodryas* clades based on a combined molecular and morphological analysis of the genus. The Congo River and the Ubangi River, one of the Congo's tributaries, date back to the mid–late Miocene (Flügel et al., 2015; Stankiewicz & de Wit, 2006). The Congo is one of the largest rivers in the world, second only to the Amazon in discharge volume, and first in the world for depth. Much of its length, especially in the lower Congo, is characterized by swift-flowing currents and waterfalls. The depth and intensity of this river has rendered it a well-known barrier to many species including primates (Harcourt & Wood, 2012; Mitchell et al., 2015; Takemoto et al., 2015; Telfer et al., 2003), shrews (Jacquet et al., 2015), and other snakes (Portillo et al., 2019). The Niger is the third largest river in Africa and it likely originated approximately 29 – 34 million years ago (Reijers, 2011; Chardon et al., 2016), reaching its full extent by the mid-Miocene as continental uplift progressed (Reijers, 2011). It is a biogeographic barrier for several species of frogs (Onadeko & Rödel, 2009; Rödel et al., 2014), primates (Eriksson et al., 2004), and shrews (Igbokwe et al., 2019).

2.5.3 Alternate biogeographic barriers

While the timing and locations of population divergences in this study correspond with river barriers, the Miocene was also a time of global climatic change characterized by dramatic cooling and vegetation shifts throughout sub-Saharan Africa (Couvreur et al., 2021; Herbert et al., 2016; Jacobs, 2004; Menegon et al., 2014). Although most research surrounding the role of refugia in driving diversification has focused on the dramatic climate oscillations of the Pleistocene, it is likely that refugia are able to form during any period of climatic change (Haffer, 1997; Hampe & Jump, 2011; Jansson & Dynesius, 2002), but the role of possibly older refugia has received little attention in the literature (Couvreur et al., 2021; Hampe & Jump, 2011).

It is interesting to note that a number of biogeographic barriers that often play a role in species diversification in Central Africa do not appear to influence the genetic patterns seen in *Toxicodryas*. One of these biogeographic barriers is the Dahomey Gap, a dry forest-savanna mosaic that separates the Upper and Lower Guinean forest blocks (Figure 2.1). Based on pollen cores and climatic modeling, the Dahomey Gap has existed in some form for at least the last 150,000 years (Dupont & Weinelt, 1996; Allen et al., 2020) and has played a variable role in species diversification in West Africa (White, 1979; Nicolas et al., 2010; Penner et al., 2011; Linder et al., 2012; Droissart et al., 2018). The West African clade of *T. blandingii* crosses the Dahomey Gap and our species distribution analyses suggest that conditions are suitable for both species in that area, likely as a result of the forest-mosaic providing suitable habitat (Chippaux & Jackson, 2019).

In addition, the two mountain ranges that bisect both species of *Toxicodryas'* ranges, the Cameroon Volcanic Line and the Albertine Rift, do not appear to impact the genetic structure of these species. Both mountain ranges originated approximately 30 million years ago with major geological developments occurring within the Miocene (Marzoli et al., 2000; Burke, 2001;

Reusch et al., 2010; Paul et al., 2014; Macgregor, 2015) and both are continuing to undergo uplift and volcanism. Both are also biogeographically important in Central Africa as proposed Pleistocene refugia (Maley, 1996; Figure 2.1, refugia 4 and 10) and as barriers to dispersal for a variety of herpetofaunal species (e.g., Evans et al., 2011; Zimkus & Gvoždík, 2013; Menegon et al., 2014; Greenbaum et al., 2015; Wüster et al., 2018). However, they seem to have played neither role in the evolution of the genus *Toxicodryas*.

2.5.4 Population demography

We used machine learning-based demographic model selection to test different gene flow scenarios between the populations in each of our two species. If river formation was the major driver of diversification in these species, we might expect to see divergence with low or no gene flow at the time of divergence as the river was forming (Figure 2.5: models 2 and 4), but we would be less likely to see recent gene flow. Alternatively, we might be more likely to observe recent gene flow if the Pleistocene refugia hypothesis was supported, as populations diverge and reunite during glaciation cycles (Smith & Carstens, 2020; Figure 2.5: model 3). Our demographic analyses indicate divergence without gene flow between the two *T. blandingii* clades and divergence with minor gene flow across the Congo River in the two Central African *T. pulverulenta* clades (Figure 2.5). Recent gene flow was ruled out with high confidence in both species (Appendix I Table S3), further supporting the river hypothesis over the refuge hypothesis. In addition, population expansion analyses found none of the signatures of population bottleneck with subsequent expansion that would be expected under the refuge hypothesis, and strongly rejected this hypothesis in favor of equilibrium isolation-by-distance in all of our populations.

In light of the Miocene/Pliocene divergence times and lack of gene flow between these five clades, it is likely that they represent distinct evolutionary lineages. Based on these molecular results, significant differences in scale counts, and even venom toxicity, the Central African populations of *T. blandingii* (east of the Congo/Ubangi Rivers) and *T. pulverulenta* (east of the Niger Delta) represent cryptic new species that are now formally described (Greenbaum et al., 2021). The two Central African clades of *T. pulverulenta*, separated by the Congo River were conservatively described as one species based on the results of our STRUCTURE v. 2.3.4 analysis (Falush et al., 2003; Pritchard et al., 2000).

2.5.5 Ecological niche projections and habitat stability

The nature of the intervening habitat surrounding rainforest refugia during the Pleistocene has been widely debated (reviewed in Couvreur et al., 2021). Some authors have argued that much of the Central African rainforest was replaced by savannas (DeMenocal, 2004; Maley, 1996; Maley & Brenac, 1998), whereas others have emphasized the possibility of more subtle shifts in forest composition (i.e., from wet to dry forest; Colinvaux et al., 1996, 2001; White, 1981). A shift from rainforest to warm temperate woodland and temperate broadleaf evergreen forest, as opposed to savannas, during the Last Glacial Maximum was strongly supported by recent, comprehensive global vegetation models (Allen et al., 2020).

Toxicodryas species are generally characterized as arboreal across rainforest and woodland habitats and the two species exhibit widely overlapping distributions in West and Central Africa (Chippaux & Jackson, 2019). Our projected ecological niche models suggested that no substantial contraction of suitable climate occurred for either species during the LGM (Figure 2.6a), and our habitat stability mapping suggested that core ranges of both species have

remained stable for the past 22,000 years (Figure 2.6c). The greatest potential for habitat expansion in this species appears to be to the south into northern Angola and southern DRC (Figure 2.6).

Similar studies on frogs have suggested substantial habitat contraction in Central Africa during the Pleistocene (Leaché et al., 2019; Portik et al., 2017). In contrast, our inferred widespread habitat stability in *Toxicodryas* may be a result of the relatively reduced dependence of arboreal snakes on moist habitats, as reflected by their distribution in both woodland and rainforest. The stability of *Toxicodryas* habitat through the Pleistocene supports the hypothesis that rainforest composition shifted to dryer woodlands surrounding rainforest refugia, instead of a more dramatic shift to strict savanna habitat. Southward shifts in species suitability may correspond with predicted forest distribution shifts of White (1981) and Allen et al. (2020), suggesting a replacement of lowland rainforest with cooler, more temperate forest.

2.6 Conclusion

The complexity of geographic barriers in West and Central Africa, and the association of refugia with areas of high surface relief or riparian zones (Hofer et al., 1999, 2000; Figure 2.1), makes it extremely difficult to untangle the relative importance of different diversification mechanisms with distribution data alone (Leaché et al., 2019; Portik et al., 2017). This difficulty is particularly salient in our study system, where distribution data could suggest the association of populations with hypothesized refugia around the Congo River, Gabon, and in West Africa (refugia 9, 5–8, and 1–3 respectively, Figures 2.2 and 2.3). Yet, our dated phylogenies and ecological niche projections reject the Pleistocene population age and habitat contraction predictions of the refugial hypotheses in favor of the river barrier hypothesis.

It is important to note, however, that as a result of the difficulty in finding and obtaining tissues from arboreal, venomous, Central African snakes, the small sample sizes used in this analysis may have influenced the inference of our population demographic parameters (Hale et al., 2012; Nelson et al., 2012; Subramanian, 2016; Zwickl & Hillis, 2002). These impacts would have mostly likely led to an underestimation of the effects of more recent events in the population's history by missing alleles present at low frequencies (Hale et al., 2012; Nelson et al., 2012; Subramanian, 2016). While similar sample sizes have been used in phylogenomic studies of other difficult-to-find or difficult-to-sample organisms (e.g., Muniz et al., 2018; Nash et al., 2018; Frugone et al., 2019), and simulation studies on empirical data have suggested that small sample sizes (as small as three to four samples per population) can give accurate results if the number of SNPs in the data set is large (Landguth et al., 2012; Jeffries et al., 2016; Nazareno et al., 2017; Qu et al., 2020), more dense sampling of the genus *Toxicodryas* from throughout their range would be ideal to fully understanding the demographic history of these species.

These results highlight the importance of using an integrative, multidisciplinary approach to statistically distinguish among competing hypotheses to explain high levels of geographically concentrated species biodiversity. Moving beyond pure pattern-based inference, a deeper and more nuanced understanding of the production, partitioning, and maintenance of diversity in complex landscapes may lead to inference of environmental and evolutionary processes that accumulate terrestrial biodiversity in tropical areas, which coincide in many cases with Global Biodiversity Conservation Hotspots (Hrdina & Romportl, 2017; Mittermeier et al., 2000, 2011; Myers, 1988) and other imperiled ecosystems on Earth.

2.7 Figures

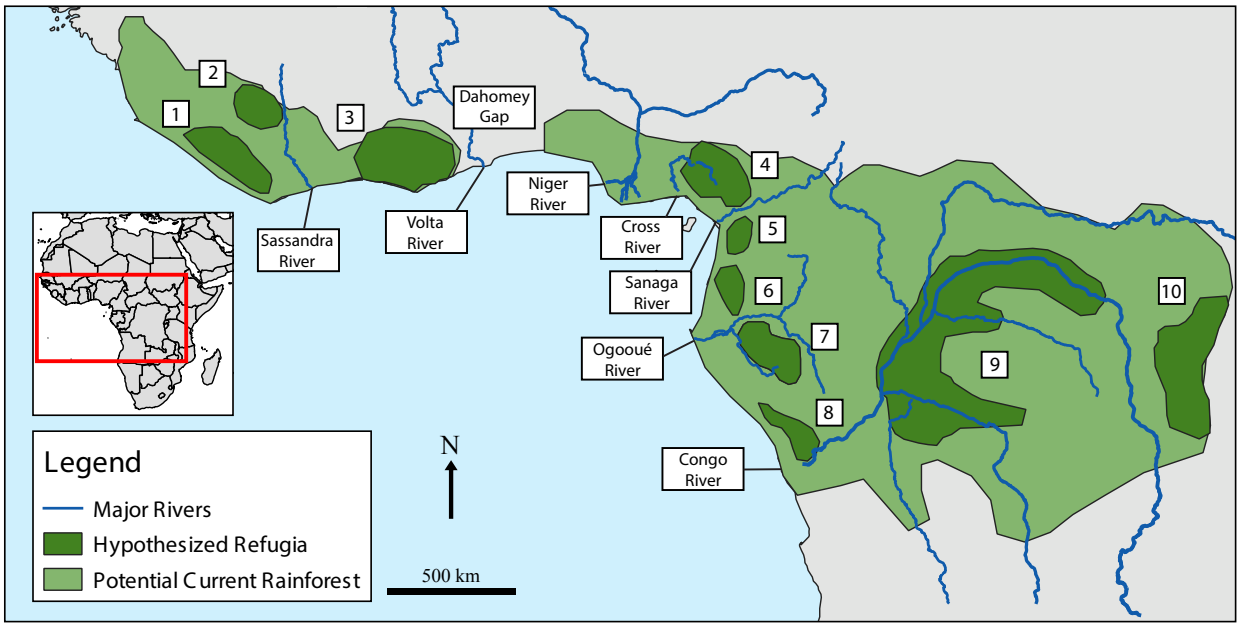


Figure 2.1 Locations of major rivers and hypothesized refugia (labeled 1–10) in West and Central Africa, adapted from Maley (1996).

Figure 2.2 Left: *Toxicodryas pulverulenta*. Right: *Toxicodryas blandingii* (male). Both photographs were taken in Banalia, Tshopo Province, DR Congo. Photo credit Konrad Mebert.

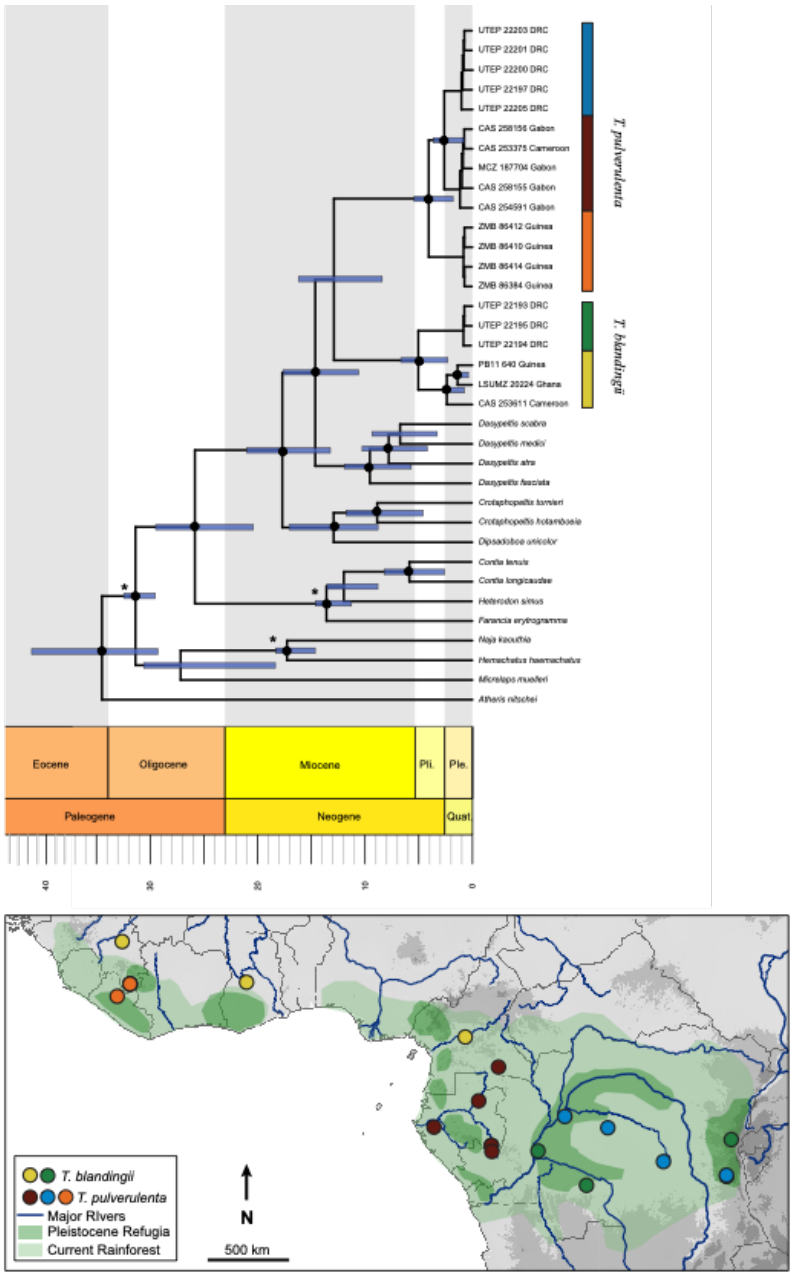


Figure 2.3 Top: A time-calibrated Bayesian phylogeny for *Toxicodryas* with c-mos and cyt b genes. Highly supported nodes (PP ≥0.9) are denoted with a black circle. Fossil-calibrated nodes are denoted with an asterisk. Node bars represent 95% confidence intervals. RADseq phylogenies showed identical topologies. Bottom: *Toxicodryas* clade distributions overlaid onto a map of elevation, major rivers and hypothesized rainforest refugia.

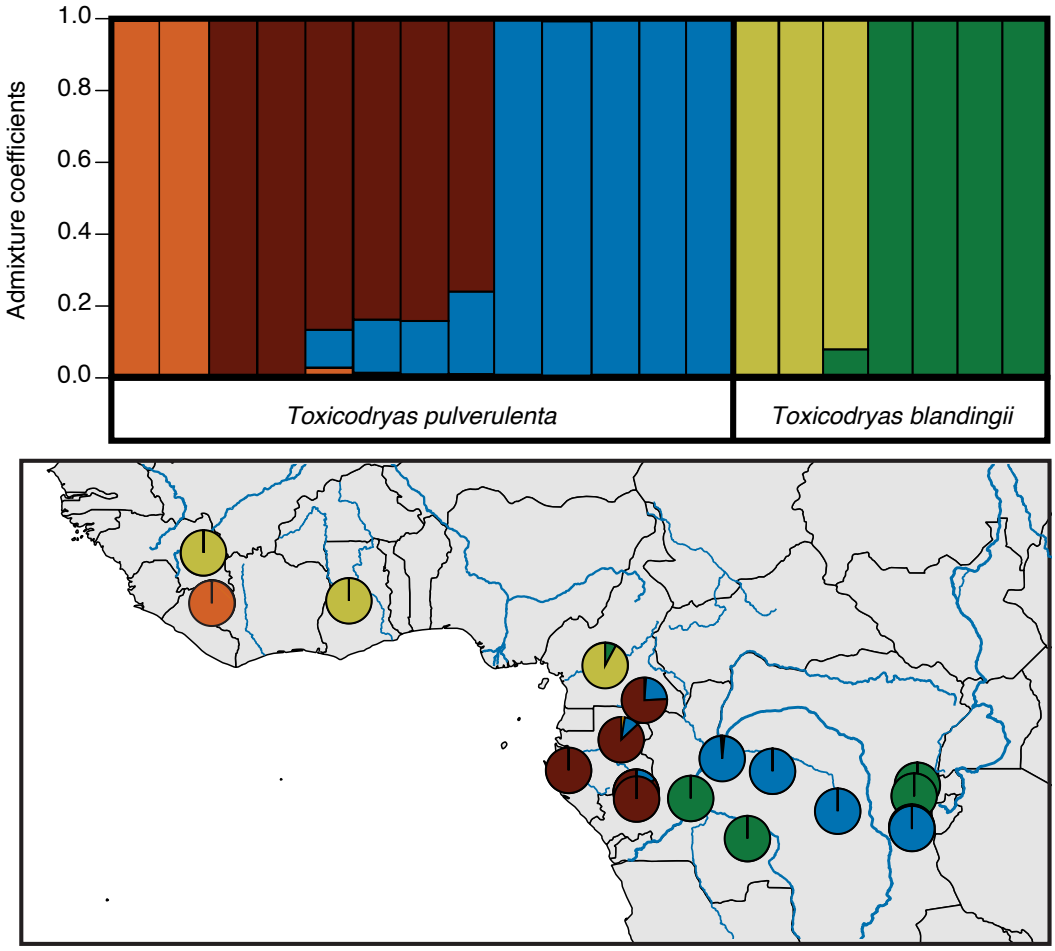


Figure 2.4 Population structure of the genus *Toxicodryas*. Top: Bar plot of population structure and membership probabilities for K=5 analyzed in LEA. Bottom: Geographic representation of population structure for K=5 overlaid onto a map of major rivers.

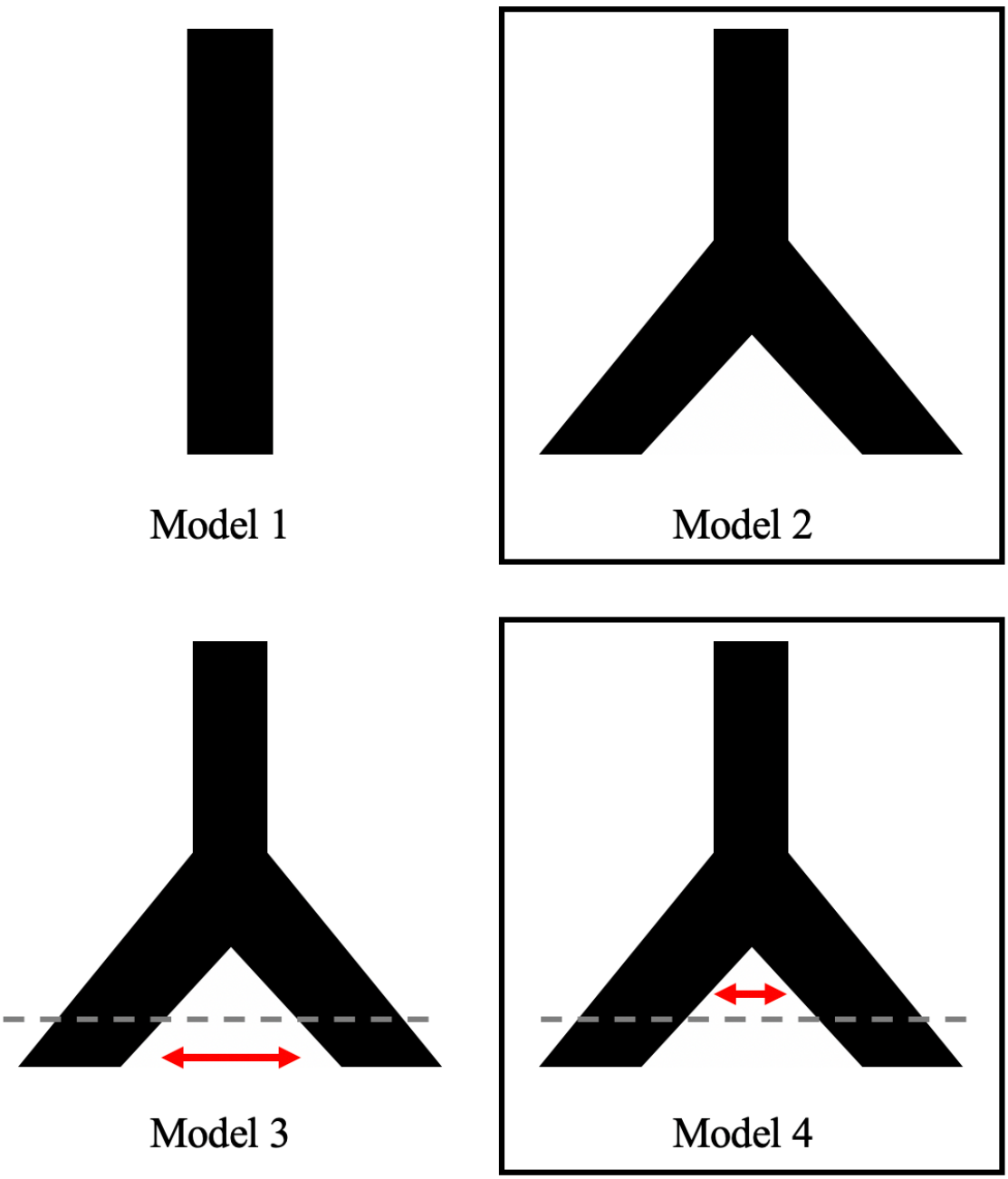


Figure 2.5 Four demographic models tested using DelimitR. Model 1: no divergence, Model 2: divergence without gene flow, Model 3: divergence with secondary contact, and Model 4: divergence with gene flow. Model 2 was chosen for *Toxicodryas blandingii*, and Model 4 was chosen for the two Central African clades of *T. pulverulenta*.

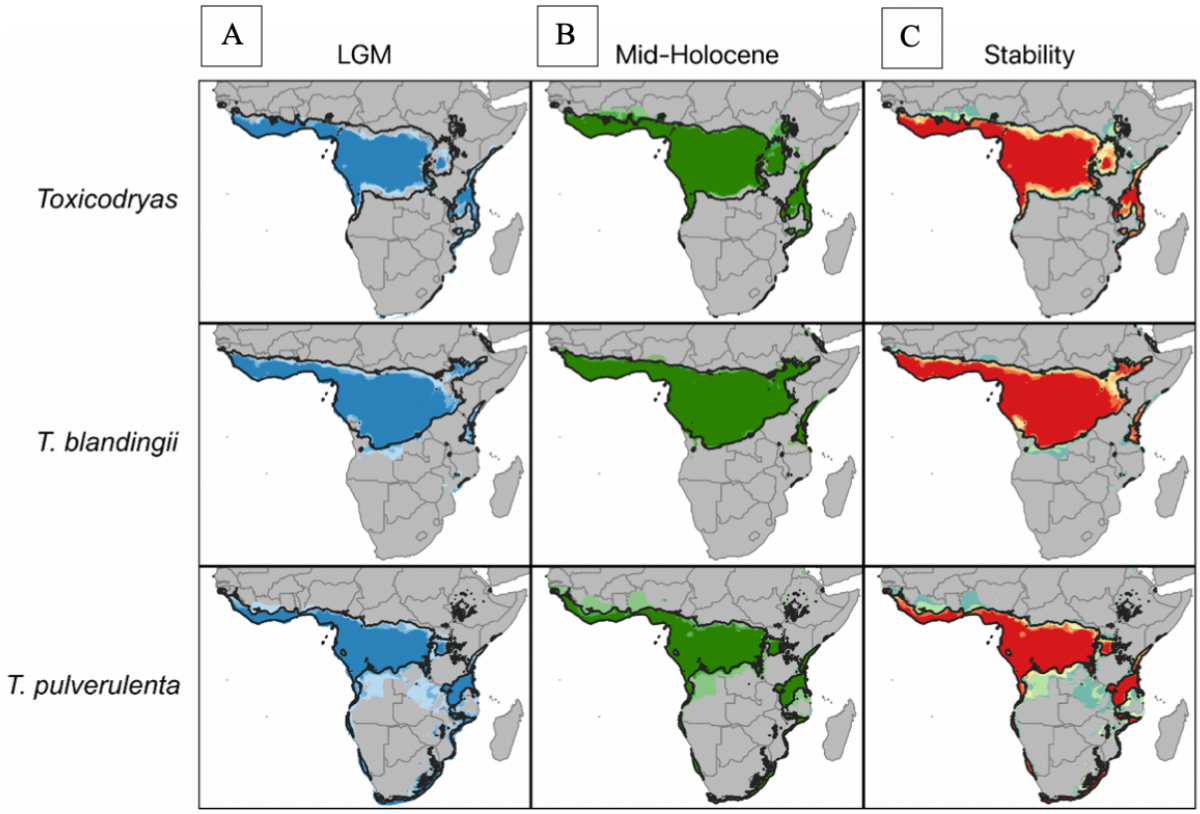


Figure 2.6 Ecological niche models showing A) the suitable habitat for *Toxicodryas* during the Last Glacial Maximum (LGM). The shade of blue represents agreement between global climate models (GCMs) with the darkest blue indicating agreement between all three GCMs and the lightest blue indicating support from only one GCM. B) The suitable habitat for *Toxicodryas* during the mid-Holocene. The shade of green represents agreement between GCMs with the darkest green indicating agreement between all three GCMs and the lightest green indicating support from only one GCM. C) The stability of suitable habitat across the LGM, mid-Holocene, and present, with red indicating high stability and blue indicating low stability.

CHAPTER 3

Pleistocene refugial diversification across highland and river systems in a widespread Central African skink

Allen, K.E., Tapondjou N., W.P., Greenbaum, E. Bauer, A.M., Ceríaco, L.M.P., Kusamba, C., Kameni N., M.M., Peterson, A.T., Brown, R.M. Pleistocene refugial diversification across highland and river systems in a widespread Central African skink. Molecular Ecology. *In review*.

3.1 Abstract

For decades, the roles of rivers and mountains in formation of Pleistocene refugia and generation of high levels of tropical species diversity have been debated. These debates have been largely driven by poor pollen core records and conflicting phylogeographic data from areas likely to be impacted by refugial dynamics. Only recently have genome-level population demography analyses and ecological niche projections become available as tools for studying the finer-scale effects of Pleistocene refugia on population dynamics. In this study, we tested the predictions of the refugial hypothesis on population diversification in a widespread skink species complex, *Trachylepis maculilabris*, by conducting a phylogenomic and historical demographic analysis across its range in West and Central Africa. Our analyses recovered five distinct genetic clusters, two in West Africa+Cameroon and three in Central Africa, with Pliocene and Pleistocene divergence date estimates. All populations show demographic signatures of population expansion after the Last Glacial Maximum, and a combination of niche and demographic modeling allowed us to narrow down the most likely refugial locations for this species. We

uncovered evidence for a Congo River fluvial refuge, and differing effects of mountain ranges, with the Cameroon Volcanic Line acting as both a barrier and a refugium, and the Albertine Rift acting as a refugium and a center of endemism. This study extends across a broad suite of Central African rivers and mountains, which leads to a refined hypothesis as to their likely role in the formation of Pleistocene refugia during the diversification of a widespread African forest vertebrate species.

3.2 Introduction

The refugial model of tropical speciation was proposed to explain Amazonian species diversification patterns in the 1960s (Haffer, 1969). However, until recently, rigorous tests of this hypothesis as it would apply to Central Africa have been less frequent (Bell et al., 2017; Hughes et al., 2017; Huntley et al., 2019; Jongsma et al., 2018; Larson et al., 2016; Penner et al., 2011; Portik et al., 2017). Based on pollen records and species distribution data, Pleistocene refugia have been proposed for several Central African localities (Bonnefille & Riollet, 1988; Brenac, 1988; Colyn, 1987, 1991; Girese et al., 1994; Maley, 1987, 1989, 1991; Maley & Brenac, 1987; Maley & Livingstone, 1983; Richards, 1963; Rietkerk et al., 1995; Sosef, 1991; Sowunmi, 1991), but the degree to which these refugia may have actually contributed to speciation has been widely debated (Amorim, 1991; Colinvaux et al., 2001; DeMenocal, 2004; Haffer, 1969, 1997; Mayr & O'Hara, 1986; Vitorino et al., 2016). Many hypothesized refugia are located in highland areas (i.e., the Cameroon Volcanic Line and the Itombwe Plateau); however, a major fluvial refuge located in the gallery forests around the Congo River has been supported by pollen core data (Maley, 1996), as well as the distributional patterns of primate subspecies (Colyn et al., 1991) and plant species (Robbrecht, 1996; Wieringa, 1999).

As a result of their small body size, low vagility, and ectothermic temperature regulation, reptiles and amphibians often exhibit strong population genetic responses to changes in climate or geography (Urban et al., 2014; Vences & Wake, 2007; Zeisset & Beebee, 2008). Numerous phylogeographic studies have supported the importance of rivers, mountains, and refugia (Figure 3.1) as drivers of diversification in West and Central African reptiles and amphibians. Biogeographically important rivers in the area include the Volta, Sanaga, Ogooué, Congo, Ubangi, Niger, and Cross Rivers—several of which have been shown to be barriers for frogs (Charles et al., 2018; Penner et al., 2011; Penner et al., 2019; Wieczorek et al., 2000; Zimkus et al., 2010) and snakes (Allen et al., 2021; Greenbaum et al., 2021). Refugia throughout West and Central Africa have played an important role in the diversification of both frogs (Bell et al., 2017; Jongsma et al., 2018) and lizards (Allen et al., 2019; Leaché et al., 2017). Similarly, two Central African mountain ranges, the Cameroon Volcanic Line and the Albertine Rift, have been interpreted as barriers for some species (Portik et al., 2017; Shirley et al., 2014; Smolensky et al., 2015; Wüster et al., 2018), and epicenters of diversification for others (Charles et al., 2018; Evans et al., 2011; Greenbaum et al., 2015; Greenbaum et al., 2012; Hughes et al., 2017; Portillo et al., 2015; Zimkus, 2009; Zimkus et al., 2010; Zimkus & Gvoždík, 2013). Both of these mountain ranges, as well as the Congo River, have been associated with the formation of refugia (Maley, 1996), and divergence patterns for some species are indicative of a combinatorial role of geographic and climatic events on diversification (Barej et al., 2011; Jongsma et al., 2018; Leaché & Fujita, 2010; Leaché et al., 2019; Portik et al., 2017). However, distinguishing between driving factors of diversification in the presence of multiple, overlapping potential causes can be difficult, especially if the cyclical nature of refugial diversification dynamics can

lead to conflicting genetic signals (Budde et al., 2013; Edwards et al., 2006; McKinnon et al., 2004; Slager et al., 2020; Tzedakis et al., 2013).

In this study, we use the *Trachylepis maculilabris* species complex of skinks to test the roles of Pleistocene refugial diversification across montane and river systems in tropical Africa. This complex is widely distributed throughout the rainforests, dry forests, and gallery forests of Central, West and East Africa at elevations from sea level to 2,550 meters (Chirio & Lebreton, 2007), and a recent, time-calibrated multi-locus genetic analysis of the species by Allen et al. (2019) recovered multiple Pleistocene age divergences within the Congo Basin, the Cameroon Volcanic Line, and the Albertine Rift, centering around previously hypothesized refugial locations (Maley, 1996). We use a combination of dated phylogeographic inference, population demography, and ecological niche models to test for association of these populations with glaciation cycles (particularly the Last Glacial Maximum and the present interglacial), by assessing if present-day population structure and distribution of genetic variation potentially were impacted by locations of refugia during cold, dry periods (Maley, 1996), and if demographic signals of bottleneck and expansion out of refugial locations can be detected.

With a novel data set consisting of 2,893 Radseq loci, we estimated divergence dates to the early Pleistocene, and found strong signatures of refugial population dynamics including signatures of population expansion after the end of the Last Glacial Maximum. Refugial locations were inferred in central Ghana, the southern Cameroon Volcanic Line, northern Democratic Republic of the Congo (DRC) near the Congo River, and both the northern and southern Albertine Rift. Taken together, our analyses suggest roles for both mountains and rivers in the formation of Pleistocene refugia in West and Central Africa. Phylogenetic and clustering analyses highlight a role for the Cameroon Volcanic Line as a barrier to dispersal between West

and Central Africa, as well as a potential refugium. In contrast, the Albertine Rift may have acted as a refugium for one lowland Congo Basin population, but possibly also as a center of endemism for another.

3.3. Materials and Methods

3.3.1 Sampling

We obtained 65 specimens of *Trachylepis maculilabris* from throughout the species range in West and Central Africa. Outgroups included 18 specimens of the closely related island species in the *T. maculilabris* complex (*T. adamastor* and *T. thomensis*), as well as three specimens of *T. affinis*. Specimens were obtained through fieldwork and from various museums. Museum catalog numbers, GenBank accession numbers, and locality data for each specimen are presented in Appendix II Table S1.

3.3.2 Genetic data collection and processing

Genomic DNA was extracted from tissue samples preserved in 95% ethanol or RNAlater (Sigma-Aldrich) using the Maxwell RSC system (Promega). Genome-wide anonymous nuclear markers were sequenced for each individual following a ddRADseq protocol modified from Peterson et al. (2012). For each individual, 300–500 ng of genomic DNA was double digested using the restriction enzymes *Sbf*I (restriction site 5'- CCTGCAGG-3') and *Msp*I (restriction site 5'-CCGG-3'). The resulting products were bead-cleaned with AmpureXP beads (Agencourt) and individually barcoded using custom oligonucleotide adapters. A Pippin Prep (Sage Science, Beverly, MA) was used to size select pooled samples to a mean insert length of 541 base pairs (bp) (487–595 bp range) using internal standards. Resulting post-ligation products were

amplified for eight cycles using a high-fidelity polymerase (Phusion, New England Biolabs). An Agilent TapeStation was used to determine the final fragment size distribution and concentration of each pool. Library pools were combined in equimolar amounts for sequencing on one Illumina HiSeqX lane (with a 10% *Phi*X spike-in and 150 bp paired-end reads).

STACKS v. 2.4 (Catchen et al., 2013) was used to process Illumina reads from the ddRAD libraries. Because the ddRAD protocol generates strand-specific libraries, we used a read-stitching approach following Hime et al. (2019) to join the first read from an Illumina read pair with the reverse complement of the second, recreating the original orientation of fragments in the genome. For each individual, stitched reads were quality-filtered and demultiplexed using the process_radtags function in STACKS with the following parameters: demultiplex each library by in-line barcode, check for both restriction enzyme cut sites, remove any read with an uncalled base, rescue barcodes and RAD-Tags, and discard any read with average Phred quality score < 20 over sliding windows of 15% of the total read length. We then used STACKS to assemble, *de novo,* filtered and stitched Illumina read pairs.

We produced two separate ddRAD data sets, one including outgroups for phylogenetic analyses, another including only *T. maculilabris* for population-level analyses. Because the optimal *de novo* assembly of ddRADseq data can vary widely across taxa (Paris et al., 2017; Shafer et al., 2017), we tested a range of assembly parameters to optimize the recovery of putatively single-copy orthologous loci. Final assembly parameters were selected based on the methods of Paris et al. (2017). According to their recommendations, in USTACKS, we kept m (the minimum number of reads needed to form a stack) at 3 while in CSTACKS, we varied M (the number of mismatches allowed during loci formation), and n (the number of mismatches allowed during catalog formation) until we identified the parameters at which the maximum

number of polymorphic loci were available across 80% (r = 0.8) of the population. For our data, we found this optimum at M = 8 for the phylogenetic data set and M = 6 for the population data set. The optimal n was 15 for all data sets. Further parameters were tested in POPULATIONS separately for each data set in order to balance missing data and number of polymorphic loci. For the phylogenetic data set we used r = 0.25 and p = 5 [p is the minimum number of populations in which a locus must be present (here 5/5)]. For the population data set, we used r = 0.65 (p was not necessary because only one population was considered). For each separate data set, we saved versions consisting of only a single random SNP per locus (for population clustering analyses and demographic modeling), and another copy comprising full-length sequences for all loci (for use in phylogenetic reconstruction).

3.3.3 Phylogenetic analysis

We analyzed our SNP data set using a maximum likelihood analysis in IQ-TREE v. 1.6.12 (Nguyen et al., 2014). We implemented 10,000 ultrafast bootstraps (Hoang et al., 2018) and the ModelFinder function to choose the best substitution model (Kalyaanamoorthy et al., 2017). The species *Trachylepis affinis, T. adamastor* and *T. thomensis* were used as outgroups (Allen et al., 2019; Ceríaco et al., 2016). We performed species tree analysis using SNAPP, as implemented in Beast v 2.6.3 (Bryant et al., 2012). We followed Stange et al., (2018) to date our phylogeny using a lognormal distribution with an offset of 0, a mean divergence time of 5.16 million years (standard deviation of 0.13) between the West Africa+Cameroon and Central African clades (Allen et al., 2019). Three non-admixed individuals per population were included, with two MCMC chains run for 1,000,000 generations each; the island species (*T. adamastor*

and *T. thomensis*) were constrained to be monophyletic. Convergence was visualized in Tracer v1.7.1 (Rambaut et al., 2018) after the first 10 percent of trees were discarded as burn-in.

3.3.4 Genetic structure

A discriminant analysis of principal components (DAPC) was conducted using Adegenet v. 2.1.1 (Jombart & Ahmed, 2011). The best-clustering scheme was chosen based on Bayesian information criterion (BIC) scores. Numbers of clusters (K) ranging from 1–10 were evaluated, and a discriminant function analysis of principal components (DAPC) served as the basis for selection of the number of clusters by maximizing variation among clusters and minimize variation within clusters. Ancestry proportions of all individuals were inferred using LEA v. 1.6.0 (Frichot & François, 2015) through the Bioconductor v. 3.4 package. The sNMF function was used to assess K values from 1–10, estimate individual admixture coefficients, and select the value of K that minimized cross entropy (Frichot et al., 2014; François, 2016).

3.3.5 Demographic modeling

We visualized spatial patterns of gene flow and genetic diversity using Estimated Effective Migration Surfaces (EEMS), an approach that identifies regions where genetic similarity decays more quickly than expected under isolation-by-distance (Petkova et al., 2016). We converted our filtered stacks output to the correct bed filetype using PLINK (Chang et al., 2015), and created a dissimilarity matrix using the BED2DIFFS program in the EEMS package (Petkova et al., 2016). The outer coordinate file was generated in QGIS v. 3.4. We used the RUNEEMS_SNPS script under several deme sizes (200, 300 and 500), and each deme was analyzed in three independent runs, with an MCMC length of 10,000,000 and a burn-in of

1,000,000. The results were combined, checked for convergence, and plotted using the REEMSPLOTS R package (Petkova et al., 2016), after which we selected the run with the lowest log posterior.

The R package rangeExpansion (Peter & Slatkin, 2013, 2015) was used to assess signatures of population size change in each of the major *T. maculilabris* clades identified in our phylogenetic and population structure analyses. One sample in Uganda (DFH 860) and the two southernmost samples in DRC (UTEP 21825 and UTEP 21826) were excluded from demographic analyses because they did not fall into primary clades identified in our phylogeny and were characterized as highly admixed in our cluster analysis. The rangeExpansion package implements a founder effect algorithm, using a stepping stone model and assuming that each colonizing event is associated with a founder event, to determine whether a population shows signatures of expansion or equilibrium isolation-by-distance (Peter & Slatkin, 2013, 2015). If population expansion is identified, the program will infer strength of the founder effect and a most-likely expansion origin (Peter & Slatkin, 2013, 2015).

We used the R package delimitR (Smith & Carstens, 2020; https://github.com/meganlsmith/delimitR) to test for historic and recent gene flow between our populations. This approach uses a binned multidimensional folded site frequency spectrum (bSFS; Smith et al., 2017) and a random forest machine learning algorithm to compare speciation models such as no divergence, divergence with gene flow, divergence without gene flow, and divergence with secondary contact (Smith & Carstens, 2020). A bSFS is used because it stores the observed frequencies of the minor alleles for multiple populations and bins them to avoid inference problems associated with sampling too few segregating sites (Smith et al., 2017; Terhost & Song, 2015). Demographic histories are simulated using the multi-species coalescent

model implemented through fastsimcoal2 (Excoffier et al., 2013) under a user-specified guide tree and set of priors on divergence times, population sizes, and migration rates. The random forest classifier then creates a user-defined number of decision trees from a subset of the prior. Each decision tree compares the empirical bSFS to the SFS of each simulated speciation model and votes for the most similar model. The demographic model with the largest number of votes is chosen as the best model. Out-of-bag error rates are used to assess power of the random forest classifier. The posterior probability of the selected model is then calculated by regressing against the out-of-bag error rates, following Pudlo et al. (2015).

We created folded multi-dimensional site frequency spectrums for the two West Africa+Cameroon *T. maculilabris* clades and the three Central African *T. maculilabris* clades using easySFS (https://github.com/isaacovercast/easySFS), a wrapper for ∂a∂i (Gutenkunst et al., 2009). Initially, we simulated 10,000 data sets under models of no divergence, divergence without gene flow, divergence with secondary contact, and divergence with gene flow between each population. Upon confirmation that each clade was recognized as distinct by DelimitR, we ran delimitR again on each clade, this time testing for population expansion. Based on a time-calibrated phylogeny from Allen et al. (2019) and an average generation time of 8 to 15 months for *Trachylepis* (Branch, 2014; Nassar et al., 2013), priors for the West Africa+Cameroon models were drawn from uniform distributions for population size: 300,000–3,000,000 haploid individuals (twice the number of estimated diploid individuals), divergence time: 3,000,000–7,000,000 generations, and migration rate: 0.00005–0.005. Priors for the Central African models were drawn from uniform distributions for population size: 300,000–3,000,000 haploid individuals (twice the number of estimated diploid individuals), divergence times: 1,000,000 – 3,000,000 and 2,000,000–5,000,000 generations, and migration rate: 0.0000005–0.005. Priors for

population size expansion were drawn from uniform distributions for exponential population growth rate: –0.001, –0.00035 between 15,000 and 4,900 generations ago, corresponding with the beginning of the present interglacial and the mid-Holocene.

For each set of analyses, we coarsened our empirical site frequency spectra to 10 bins. Our out-of-bag error rates were calculated, and 500 random-forest classifiers were simulated using 100,000 pseudo-observed data sets for each model. A confusion matrix was calculated to determine how often the correct model was selected, and posterior probability for the "best" model was estimated, as above, for each species.

3.3.6 Ecological niche models

We obtained occurrence data for each species from specimen-associated museum database records corresponding to genetic vouchers, individual occurrences from the Global Biodiversity Information Facility (GBIF), and research-grade locality records from iNaturalist (www.inaturalist.org), when we could visually corroborate the identification (Appendix II Figure S1). Duplicate records were removed, and the package spThin (Aiello-Lammens et al., 2015) was used to thin points within a distance of 10 kilometers in R v. 3.4.4 (R Core Team, 2018). Our record selection resulted in a total of 181 localities: 25 for the West African clade, 54 for the Cameroon clade, 85 for the Congo Basin clade (combining both Congo Basin populations), and 17 for the Albertine Rift clade. A subset of points from each data set was set aside for model calibration (75%) and internal testing (25%) following Cobos et al. (2019).

We obtained environmental data from the WorldClim database v. 1.4 (Hijmans et al., 2005). Fifteen of the 19 bioclim variables were downloaded at a 2.5’ (~4.5 km) spatial resolution. Mean temperature of wettest quarter, mean temperature of driest quarter, precipitation

of warmest quarter, and precipitation of coldest quarter were excluded because they are known to include odd spatial artifacts (Escobar et al., 2014). These 15 variables were also used for the Last Glacial Maximum (LGM) and the mid-Holocene under three general circulation models (GCMs): CCSM4, MIROC-ESM, and MPI-ESM-P. In order to reduce spatial autocorrelation, principal component analyses (PCAs) were performed on present bioclim variables and projected to the LGM for the extent of sub-Saharan Africa.

Model calibration areas were defined as a 300 to 400-km buffer around occurrence points for each species. Model calibration, creation, projection, and evaluation were done using the R package kuenm (Cobos et al., 2019). In order to calibrate our models, we created 714 candidate models for each clade by combining three sets of environmental predictors (PCAs 1–6, 1–5, 1–4), 17 possible regularization multipliers (0.1–1.0 at intervals of 0.1, 2–6 at intervals of 1, and 8 and 10), and all combinations of three feature classes (linear = l, quadratic = q, product = p; Cobos et al., 2019). Candidate models were run in Maxent (Phillips et al., 2006) and chosen based on significant partial receiver operating characteristic (ROC) scores (Peterson et al., 2008), omission rates of $E \leq 5\%$ (Anderson et al., 2003), and corrected Akaike information criterion $AICc \leq 2$ to minimize model complexity (Warren & Seifert, 2011). These models determined the parameter set used for final model creation.

Final models were created for each species using the full set of occurrence records and the parameters chosen during model calibration. Models were run in Maxent with ten bootstrap replicates and logistic outputs. After models were run in the present, they were projected to the LGM and mid-Holocene for the three GCMs. Models were visualized in QGIS 3.4 and thresholded to 5% to create presence-absence maps. Models from each time period were summed

to estimate potential LGM and mid-Holocene range areas as well as continuous stability maps (Devitt et al., 2013; Yannic et al., 2014).

3.4 Results

3.4.1 Genetic data collection and processing

After filtering, our phylogenetic ddRAD data set, including outgroups, consisted of 2,893 loci with 41.2% missing data (here defined as proportion of missing loci across all individuals), and an effective mean per-sample depth of coverage of 59.2 +/- 13.0. The concatenated ddRAD data set used for phylogenetic analyses had a length of 849,977 bp. Our population data set consisted of 2,763 loci with 19.5% missing data, and an effective mean per-sample depth of coverage of 62.3 +/- 13.6.

3.4.2 Phylogenetic analysis

Broad-scale phylogenetic relationships estimated from our SNP data sets were strongly supported at all internal nodes. *Trachylepis maculilabris* consists of geographically structured metapopulations, with a primary divergence between West Africa+Cameroon (Ghana through Cameroon) and Central Africa (Gabon, DRC, and Uganda). The West Africa+Cameroon clade consisted of two sub-clades, split at the border of Cameroon and Nigeria (the West African and Cameroon clades), and the Central African clade consisted broadly of two lowland Congo Basin clades, and one located in the Albertine Rift (Figure 3.2). Our SNAPP species tree analysis recovered identical relationships to those inferred by our IQ-TREE analysis, and all nodes were strongly supported. Our dating analysis estimated divergence between West Africa+Cameroon

and Central African clades, at 5.06 million years ago; diversification within each of these clades was estimated during the Pleistocene (1.6–2.6 mya; Figure 3.3).

3.4.3 Genetic structure

Our admixture-based genetic structure method, LEA, identified five distinct clusters. Populations identified by LEA generally coincided with clades found in our phylogenetic analyses, with the exception of two samples from south of the Congo Basin (UTEP 21825 and UTEP 21826) and one sample in northern Uganda (DFH 860); these three individuals did not cluster with the major clades discovered in our phylogenetic analysis but were not identified as distinct in the cluster analysis; instead, each was inferred as highly admixed. LEA uncovered high levels of genetic admixture in the Cameroon Volcanic Line, northern DRC, and in the vicinity of the Albertine Rift. Very limited admixture (or no admixture) was detected between the West Africa+Cameroon clade and the Central African clade (Figure 3.4). A comparison of BIC values from the genus-level DAPC analyses similarly suggested a total of five genetic clusters for these populations (Appendix II Figure S2).

3.4.4 Demographic modeling

Our EEMS analyses identified several barriers to migration, as well as areas of higher- or lower-than-expected genetic diversity. This analysis supported the interpretation of the Dahomey Gap and the Cameroon Volcanic Line as major barriers to dispersal. Areas of lower-than-expected genetic diversity were suggested for coastal Central Africa, and higher-than-expected diversity was detected in the upper Guinean rainforest of West Africa (Figure 3.5).

Analyses for distinguishing between demographic signatures of range expansion versus equilibrium isolation-by-distance, conducted in RangeExpansion, strongly supported the expansion model for eight of our nine populations. The only population for which it was not supported was the Cameroon clade. Accordingly, the strength of the founder effect (q) for each population was generally high, and the founder distance (d) was low (Appendix II Table S2). Population centers of origin and expansion were projected and visualized in Figure 3.3.

DelimitR supported the model for the West Africa+Cameroon clades as divergence with gene flow (330/500 votes) with a posterior probability of 0.85, divergence without gene flow received the second highest support (103/500 votes); models of no divergence (1/500 votes) or recent gene flow (45/500 votes) received very little support. The out-of-bag error rate was 26.7%, with all of the misclassifications being between highly similar models (i.e., between divergence with or without gene flow). When population size expansion was added, the most highly supported model was expansion during the present interglacial for both populations (257/500 votes), with a posterior probability of 0.60, with some support for expansion only in the West Africa clade (194/500 votes; Figure 3.6).

The most strongly supported models for the Central African clades (three populations) and divergence with gene flow between the lowland Congo Basin populations (155/500 votes) with a posterior probability of 0.80, followed closely by divergence without gene flow for all populations (131/500 votes). Models specifying no divergence, or recent gene flow, again received very little support. The out-of-bag error rate was 39.4%, with all of the misclassifications between highly similar models. When population size expansion was added, the most highly supported model was expansion during the present interglacial for all three populations (125/500 votes), with a posterior probability of 0.56 (Figure 3.6). The confusion

matrix and number of votes per model for each of these analyses can be found in Appendix II Tables S3–S7.

3.4.5 Ecological niche modeling

Ecological niche modeling suggested dynamic distribution changes from the LGM to the present in all clades except the West African clade (Figure 3.7). The West African clade had suitable range area throughout West and Central Africa and showed only minor contractions at the northern edge of its range during the LGM and mid-Holocene. Although this clade has widespread suitable range area, its range is restricted to the area west of the Cameroon Volcanic Line. Niche models for its sister clade, the Cameroon clade, were suggestive of contraction into the Cameroon Volcanic Line during the LGM and possible expansion outside of its current range during the mid-Holocene. Despite suitable range area availability in West Africa and Gabon, this clade currently is restricted to east of the Cameroon Volcanic Line and north of the Cameroon border with Gabon. Niche models for the lowland Congo-Basin clade were indicative of a north-south split in the species range during the LGM and the mid-Holocene. Suitable area for this clade was identified as well in Cameroon and in West Africa, as well as east of the Albertine Rift, however, its current range is restricted to the lowland Congo Basin. The Albertine Rift clade showed evidence of potential for expansion out of the Albertine Rift and into the lowland Congo Basin during the LGM and the Mid-Holocene. Although suitable areas were identified in western DRC and northern Uganda, this clade currently is restricted to the Albertine Rift (Figure 3.7).

3.5 Discussion

In this study, we tested the role of Pleistocene refugia on population diversification in the widespread African skink species *Trachylepis maculilabris*. We found strong signatures of refugial population dynamics, including divergence dates in the early Pleistocene and signatures of population expansion after the end of the Last Glacial Maximum. Our results corroborate the hypothesis that a fluvial refuge existed in the gallery forest surrounding the Congo River (Colyn et al., 1991; Maley, 1996; Robbrecht, 1996; Wieringa, 1999), and that refugia formed in areas of high topographic relief, such as the Cameroon Volcanic Line and the Albertine Rift (Bonnefille & Riollet, 1988; Brenac, 1988; Sosef, 1991; Colyn, 1987, 1991; Girese et al., 1994; Maley, 1987, 1989, 1991, 1996; Maley & Brenac, 1987; Maley & Livingstone, 1983; Richards, 1963; Rietkerk et al., 1995; Sowunmi, 1991). An additional potential refugium for this species was found in central Ghana, an area hypothesized to be a refugial location by Maley (1996) based on pollen core data.

The role of Pleistocene refugia in the diversification of tropical taxa has been debated in the past (Amorim, 1991; Colinvaux et al., 2001; DeMenocal, 2004; Haffer, 1969, 1997; Mayr & O'Hara, 1986; Vitorino et al., 2016). These debates are partly because of uncertainty surrounding the nature of habitat change in the lowland rainforest during glacial cycles (Allen et al., 2020; Anhuf, 2000, 2006; Coetzee, 1964; Dupont et al., 2001; Maley, 1991, 1996, 1997; Ravelo et al., 2004; van Zinderen Bakker & Coetzee, 1972), and partly because of idiosyncratic, species-specific responses to Pleistocene climate change (e.g., Bell et al., 2017; Budde et al., 2013). This study adds to a growing body of literature, which expands our understanding of the effects of glaciation cycles on diversification at the population level in Central Africa (e.g., Bell et al., 2017; Piñeiro et al., 2017; Portik et al., 2017; Charles et al., 2018; Leache et al., 2019; Helmstetter et al., 2020a,b).

3.5.1 Phylogenetic relationships

Our phylogenetic and population clustering analyses provide strong statistical support corroborating the view of *T. maculilabris* as a geographically highly-structured metapopulation, composed of five regionally defined clades in West and Central Africa (Figures 3.2–3.4). Two clades in West Africa and Cameroon form a lineage sister to another lineage in Central Africa composed of two clades in the lowlands, and one in the Albertine Rift (Figures 3.2 and 3.3). This West Africa+Cameroon clade was also found to include a single sample from Uganda, although this was not supported by the population clustering analyses. While more sampling is needed from the Ugandan population, it is interesting to note that similar distributions have been noted in a variety of species of frogs in the genera *Leptopelis*, *Ptychadena*, *Sclerophrys*, and *Xenopus* (Channing & Rödel, 2019; Evans et al., 2015).

Dating analyses estimate a mean divergence of 5.06 million years (3.8–6.5 mya) for the primary, west versus central African split. Moreover, diversification within each clade is placed statistically within the Pleistocene interval (1.6–2.6 mya; Figure 3.3). These divergence times and phylogenetic relationships approximate those of Allen et al. (2019) with the addition that the backbone of the phylogeny in Central Africa is now fully resolved. High levels of admixture were recovered within the West Africa+Cameroon and Central African clades, but very little admixture was found between them (Figure 3.4). Admixture levels were highest in the Cameroon Volcanic Line and in northeastern DRC (Figure 3.4). Our population demography analyses strongly suggested that this admixture was a result of divergence with gene flow, as opposed to recent gene flow (Figure 3.6). In light of the Miocene and early Pleistocene divergence times and the lack of gene flow between the West Africa+Cameroon, Central African and Albertine Rift

clades, it is likely that they represent distinct evolutionary lineages and, thus, surveys of morphological data and analyses of phenotypic variation are underway to determine if formal taxonomic revision is justified.

3.5.2 Geographic barriers

Because of the wide range occupied by *T. maculilabris*, we were able to observe phylogenetic and population genetic responses within the species to all major biogeographic barriers within West and Central Africa (Figure 3.1). These barriers include major river systems such as the Volta, Sanaga, Ogooué, Congo, Ubangi, Niger, and Cross Rivers (Charles et al., 2018; Jacquet et al., 2015; Mitchell et al., 2015; Penner et al., 2011, 2019; Telfer et al., 2003; Wieczorek et al., 2000; Zimkus et al., 2010); mountain systems including the Cameroon Volcanic Line and the Albertine Rift (Evans et al., 2011; Portik et al., 2017; Shirley et al., 2014; Smolensky et al., 2015; Wüster et al., 2018; Zimkus & Gvoždík, 2013); and ecological barriers including the Dahomey Gap (Droissart et al., 2018; Linder et al., 2012; Leaché et al., 2020; Nicolas et al., 2010; Penner et al., 2011; White, 1979) and the Central African climatic hinge between southern Cameroon and Gabon (Faye et al., 2016; Gonmadje et al., 2012; Hardy et al., 2013). Rivers were not identified as barriers for any of our *T. maculilabris* populations, which is unsurprising: the clade containing *T. maculilabris* is known for its overwater dispersal abilities, inferred previously for the species and its close relatives on many African islands (and even one South American island; Carranza & Arnold, 2003; Ceríaco, 2015; Ceríaco et al., 2016, 2020; Lima et al., 2013; Mausfeld & Vrcibradic, 2002; Rocha et al., 2010).

Mountain ranges may have played more extensive roles in the diversification of this species. The Cameroon Volcanic Line was identified by our phylogenetic, population clustering,

and migration surface analyses to pose a barrier between the West African and Cameroon clades, which appear to have diverged approximately 1.6 (0.96–2.29) million years ago (Figures 3.2–3.4; Appendix II Figure S3). If correct, such a timing for diversification began well after formation of this range itself, which likely initiated around 30 million years ago and continued into the Miocene (Burke, 2001; Marzoli et al., 2000; Paul et al., 2014; Reusch et al., 2010); alternatively, our findings may reflect a role for the Cameroon Volcanic Line in Pleistocene refugial dynamics. The Albertine Rift of eastern Central Africa apparently played a different role in population diversification in this species. Instead of acting as a barrier, it harbored its own deeply divergent lineage (Figures 3.2 and 3.3). This lineage was estimated to be ~2.6 million years old, corresponding to the beginning of the Pleistocene; again, the mountain range itself is much older (Macgregor, 2015; Paul et al., 2014), highlighting this area as a potential refuge during the glaciation cycles.

Ecological barriers played a more ambiguous role in population diversification in *T. maculilabris*. The Dahomey Gap was identified by our estimated migration surface analyses to constitute a gene flow barrier within the West African clade (Figure 3.5) but was not evident in our phylogenetic or population clustering analyses. This apparent disagreement among analyses may be due to the relatively young age of the Dahomey Gap (~150,000 years; Allen et al., 2020; Dupont & Weinelt, 1996), or it may represent only a very weak barrier to *T. maculilabris*. This geographic feature has already been identified to have played varying roles in the divergence of other West African populations (Demenou et al., 2018; Dongmo et al., 2019; Fuchs & Bowie, 2015; Leaché et al., 2020). The Central African climatic hinge, located between southern Cameroon and Gabon (Gonmadje et al., 2012; Hardy et al., 2013), corresponds to an inversion of the dry and rainy seasons between West-Central and Central Africa. This climatic and

geographic feature has been interpreted to have potentially played a role in diversification of a variety of species (Faye et al., 2016; Hardy et al., 2013; Helmstetter et al., 2020a). The location of this hinge corresponds to the divergence between our West Africa+Cameroon and Central African clades of *T. maculilabris*, and the absence of an alternative barrier makes it a candidate for a driver of diversification between these two clades, along with potential habitat changes, during the Miocene aridification (reviewed in Couvreur et al., 2021).

3.5.3 The role of refugia

The Pleistocene ages estimated for *T. maculilabris* populations, and their geographic overlap with previously hypothesized Pleistocene refugia (Maley, 1996), corroborate predictions that glacial cycles may have driven population diversification in this species. This interpretation is supported statistically by our population demography analyses, which suggests strong signals of population bottleneck and expansion out of these refugial locations (Figure 3.3 and 3.6; Appendix II Table S2), and our niche model projections that suggest population contraction during the Last Glacial Maximum and expansion through the mid-Holocene to present (Figure 3.7). Agreement among these disparate approaches highlight northern (and southern) DRC, the Cameroon Volcanic Line, and two locations in the Albertine Rift, as locations of potential refugia. An additional refugium was identified by our demography analyses in central Ghana but was not corroborated by our niche reconstructions; nevertheless, its location has been corroborated by pollen core data (Maley, 1996; Figures 3.3 and 3.7).

Hypothesized refugia have played roles in the historical understanding of patterns and processes of diversification of a large variety of other species in West and Central Africa. These include primates (Clifford et al., 2004; Haus et al., 2013; Tosi, 2008), birds (Fjeldså & Bowie,

2008; Huntley et al., 2018; Huntley & Voelker, 2017), rodents (Bohoussou et al., 2015; Mizerovská et al., 2019; Nicolas et al., 2012; Nicolas et al., 201), plants (Born et al., 2011; Budde et al., 2013; Daïnou et al., 2010; Dauby et al., 2010; Duminil et al., 2015; Faye et al., 2016; Gomez et al., 2009; Hardy et al., 2013; Helmstetter et al., 2020a,b; Ley et al., 2014; Ley, Heuertz, & Hardy, 2016; Lowe et al., 2010), frogs (Bell et al., 2017; Jongsma et al., 2018; Portik et al., 2017), and reptiles (Allen et al., 2019; Leaché et al., 2017, 2019). It is interesting to note that although both mountain ranges in our study were associated with refugial diversification, the Cameroon Volcanic Line acted as a population barrier to *T. maculilabris* in West-Central Africa, whereas the Albertine Rift harbored an endemic lineage dating to the beginning of the Pleistocene. Such a combined result potentially highlights these mountain ranges as having separate roles as either species "pumps" or species "cradles" during glaciation cycles (Dagallier et al., 2020; Fjeldså, Bowie, & Rahbek, 2012; Rahbek et al., 2019). Comparative studies of widespread Central African species offer opportunities for distinguishing between these (and other) predictions, underlying major hypotheses for generation of sub–Saharan African biodiversity.

3.6 Conclusions

Our study builds upon recent investigations of forest vertebrates, emphasizing the role of Pleistocene refugial dynamics in the formation of geographically-structured, widespread species metapopulations spanning West and Central Africa—and which have been profitable for advancing the evaluation of these concepts, through integration of phylogenomics, demography and ecological niche modeling (Bell et al., 2017; Leaché et al. 2019; Portik et al., 2017). However, many of these studies were only able to focus on a fraction of the range that we were

able to with the widespread *T. maculilabris*. Integration of methodological approaches holds promise for finer-scale resolution of the putative roles of climate cycles during population diversification in tropical Africa. Previous studies have shown the responses of individual species to be highly dependent on their habitat preferences, and fundamental ecological niches (Bell et al., 2017; Hardy et al., 2013; Jackson et al., 2009; Leaché et al., 2020; Nogués-Bravo, 2009). Perhaps these idiosyncratic responses, and a primary focus on species—as opposed to population—diversification may eventually be shown to account for conflict in the literature regarding the roles of glaciation cycles in tropical species diversification.

3.7 Figures

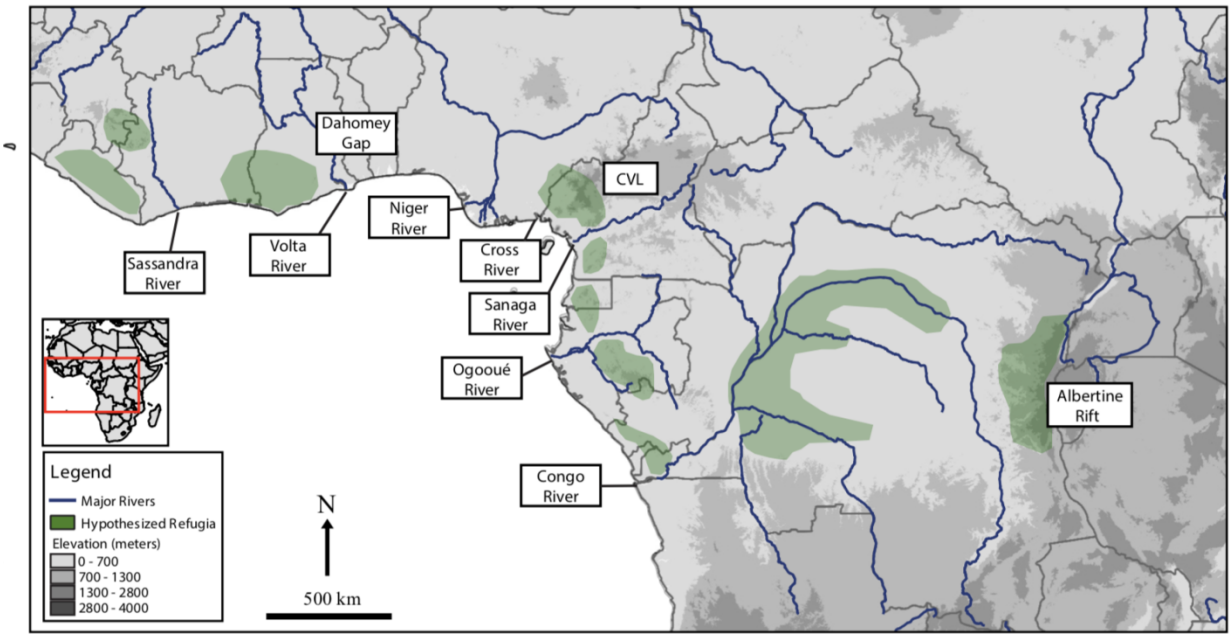


Figure 3.1 Major biogeographic barriers in West and Central Africa including rivers, refugia (adapted from Maley, 1996), and mountain ranges.

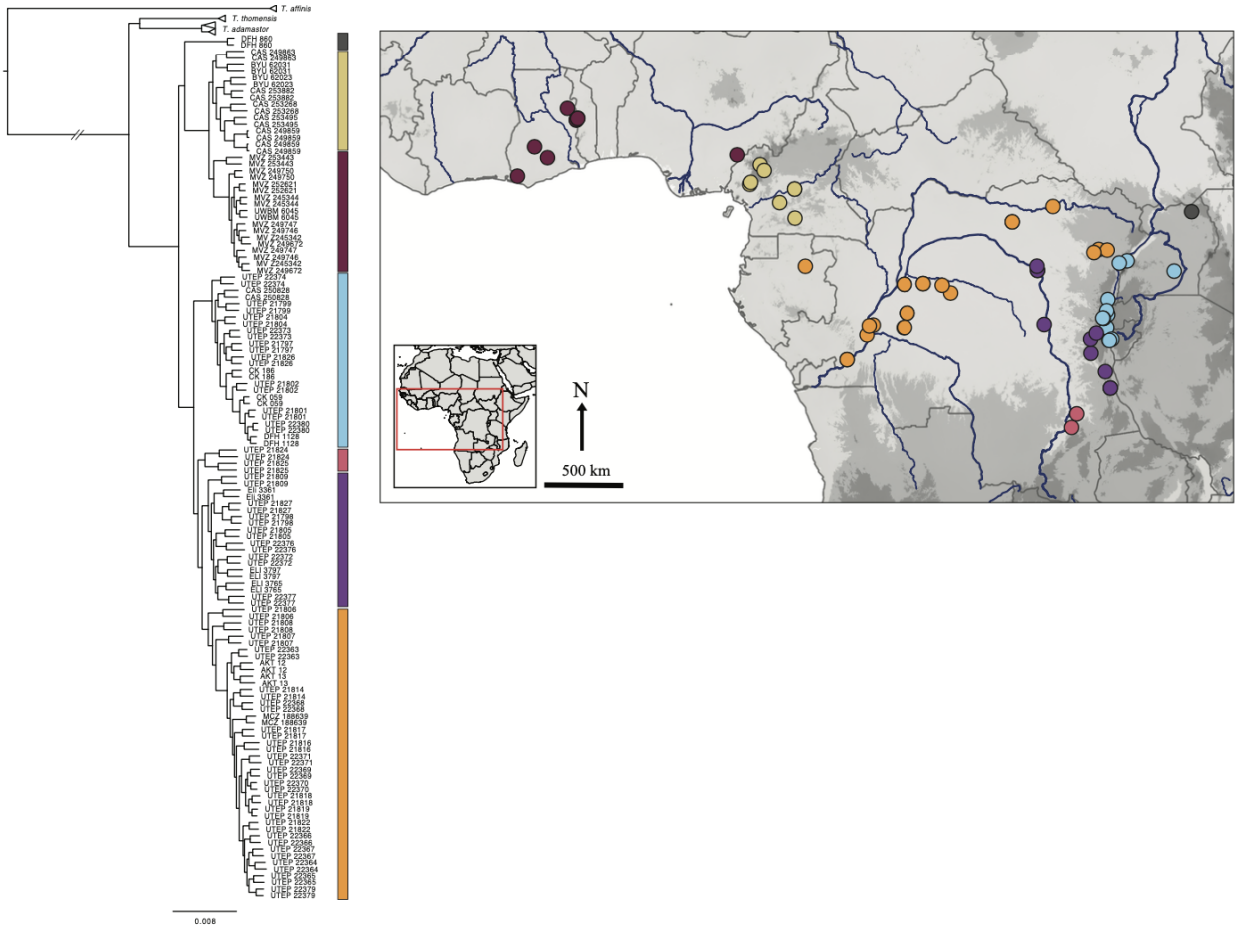


Figure 3.2 Intraspecific phylogenetic relationships of *Trachylepis maculilabris* clades across West and Central Africa based on maximum likelihood analysis (implemented in IQ-TREE) of SNP data. All nodes were highly supported.

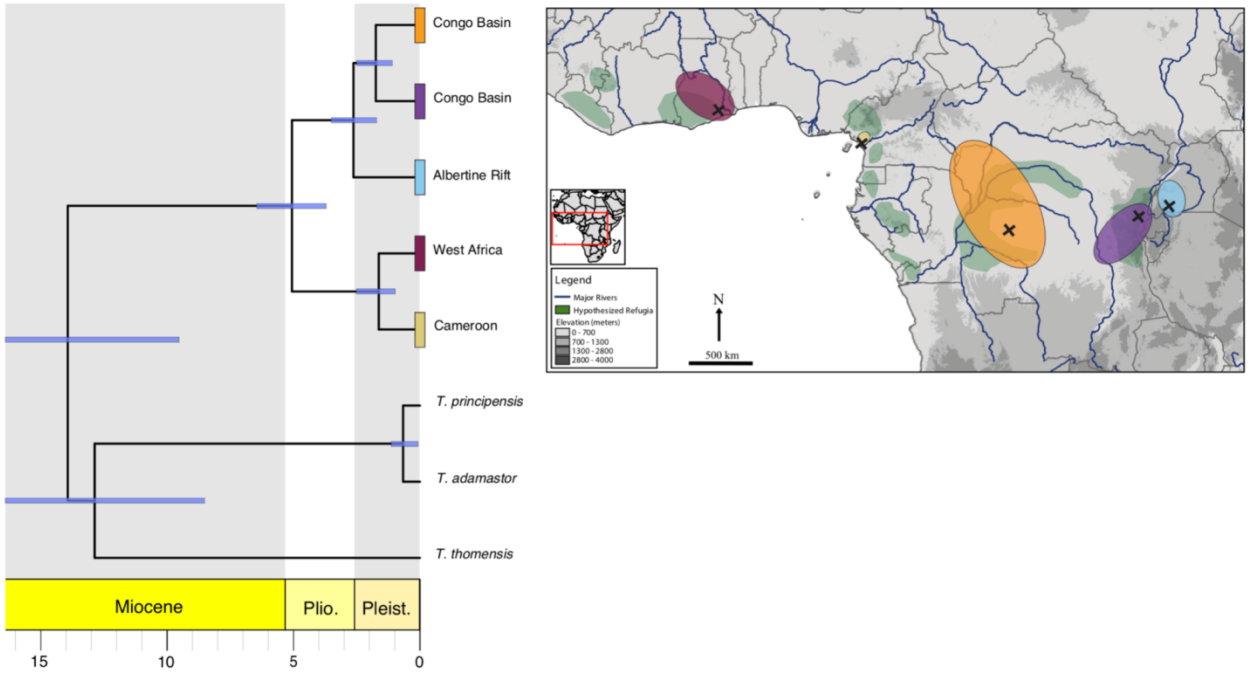


Figure 3.3 Left: Dated species tree analysis of the major *Trachylepis maculilabris* clades. All nodes were highly supported (pp>0.99). Right: Map of hypothesized Pleistocene refugia (Maley, 1996). The position of each geographic origin (indicated for each clade with a black "x") inferred from the expansion vs equilibrium isolation-by-distance analyses (implemented inRangeExpansion), accompanied by each clade's area of estimated population expansion (ellipses color-coded to match tree tip labels).

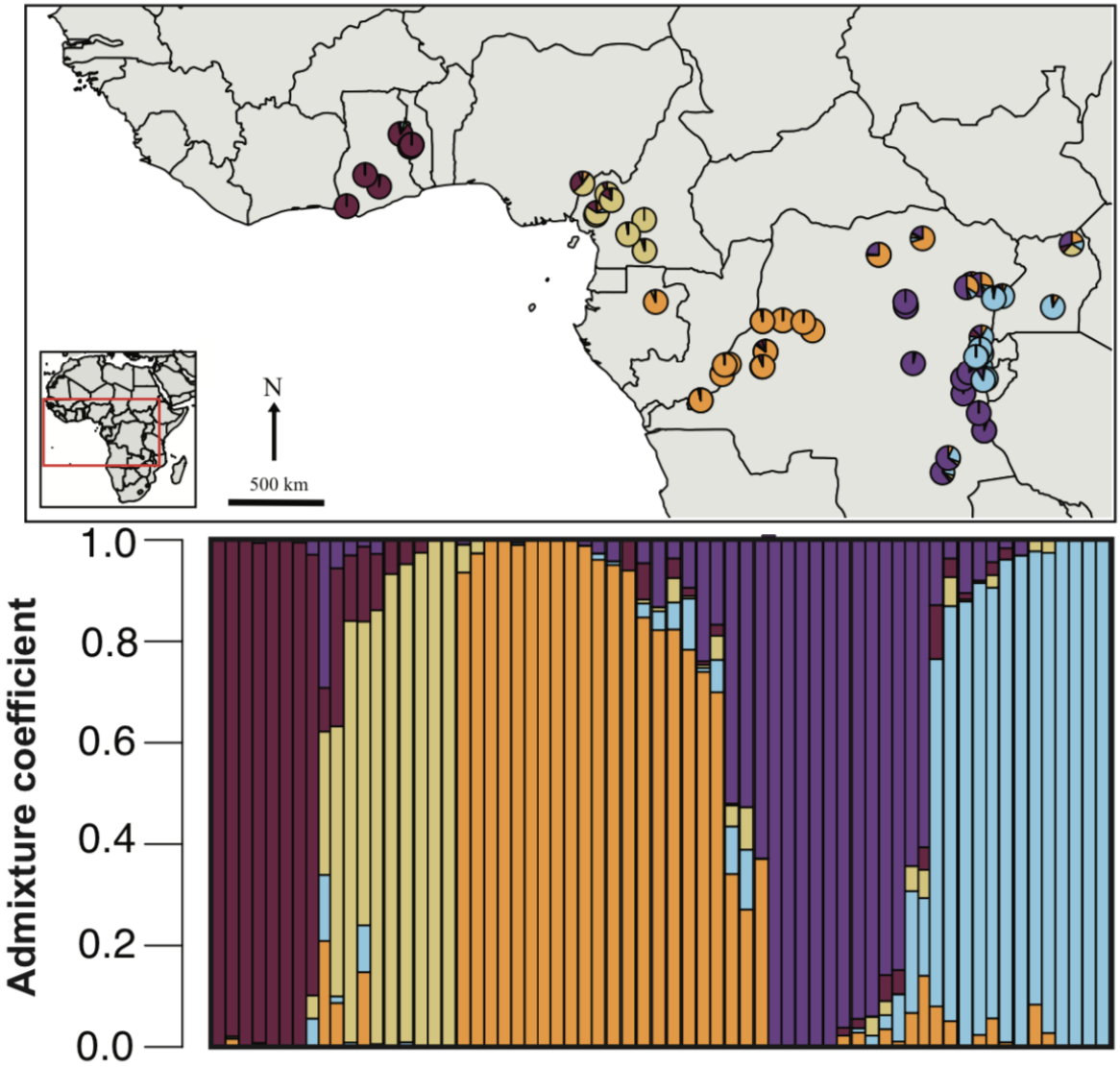


Figure 3.4 Population structure of *Trachylepis maculilabris*. Top: Geographic representation of population structure for K=5. Bottom: Bar plot of population structure and membership probabilities for K=5, analyzed in LEA.

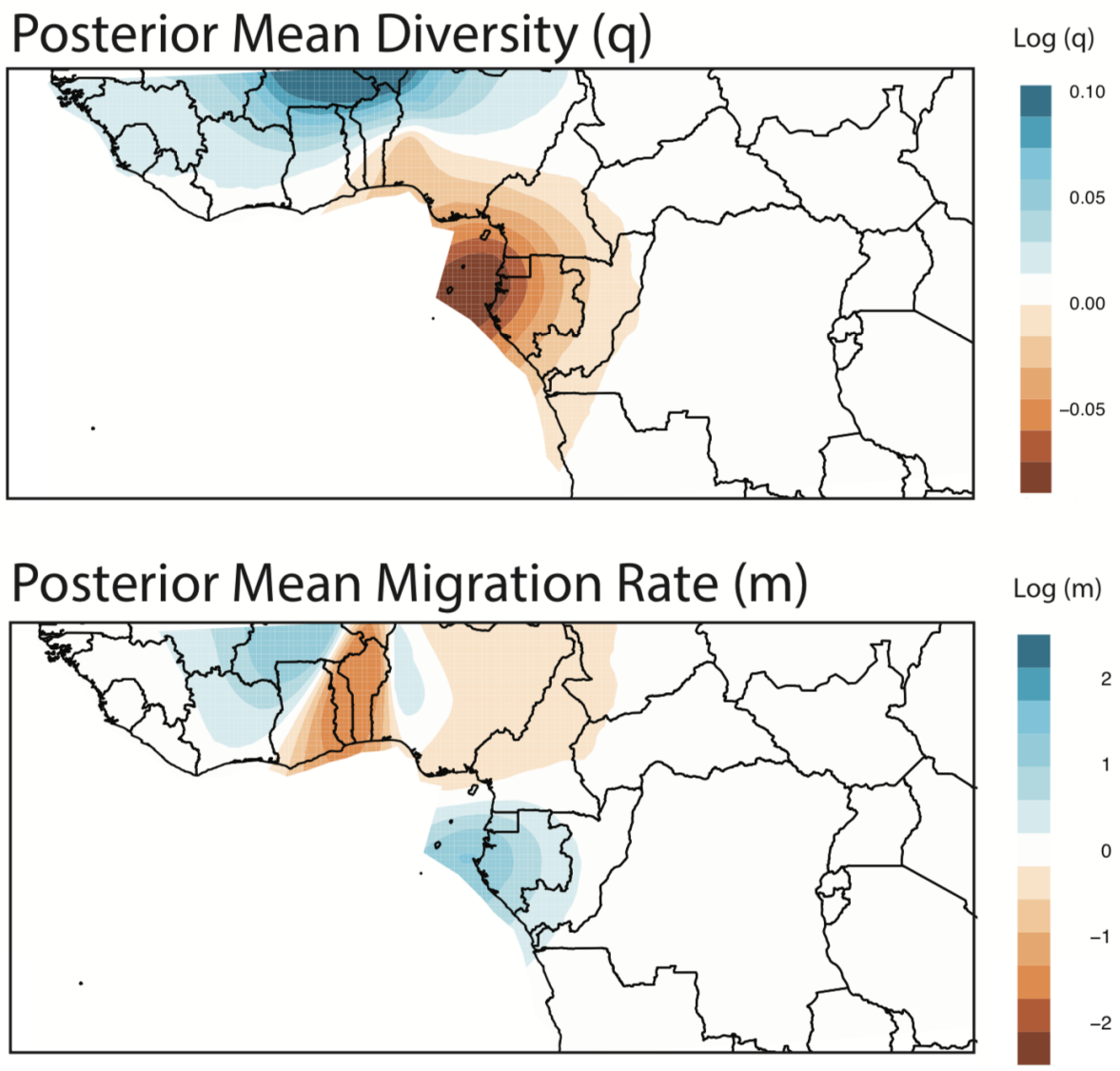


Figure 3.5 Estimated effective migration surfaces for *Trachylepis maculilabris*. Top: Posterior mean diversity surface identifying high diversity (blue) in West Africa and low diversity (orange) in coastal West-Central Africa. Bottom: Posterior mean migration rate showing high migration in West Africa and Gabon (blue) and low migration across the Dahomey Gap and in Central Nigeria through the Cameroon Volcanic Line (orange).

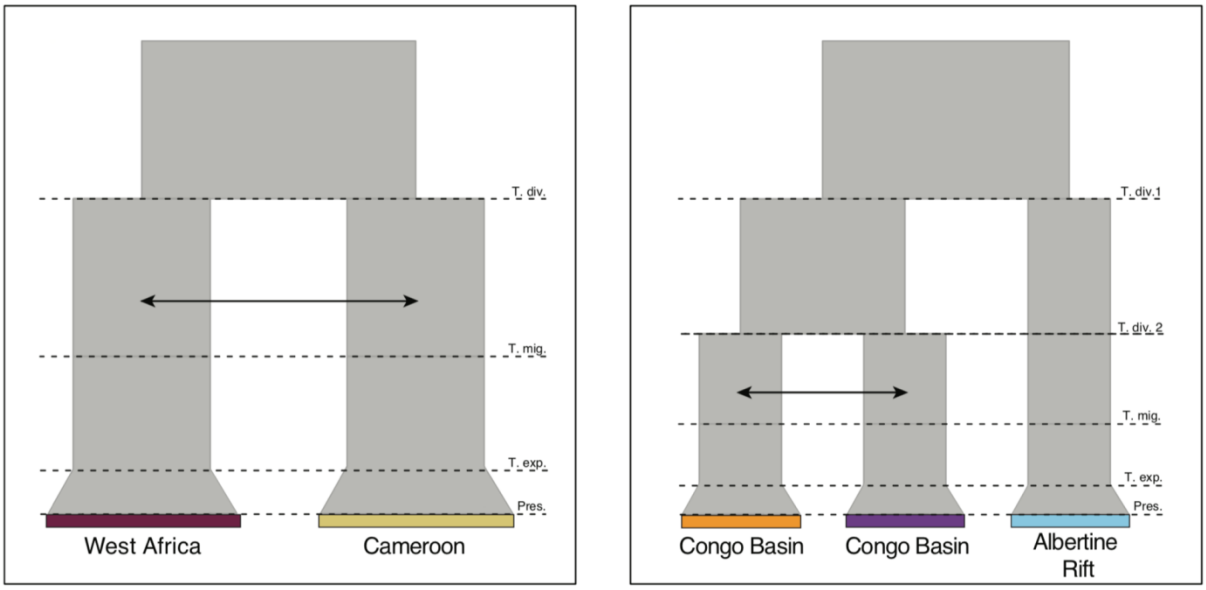


Figure 3.6 The most highly supported demographic models for each population chosen by demographic tests for migration and population expansion implemented in DelimitR. Left: For the West African/Cameroon clade divergence with gene flow and population expansion in both populations after the LGM were chosen. Right: For the Central African clade divergence with gene flow was chosen for the two lowland Congo Basin clades and population expansion after the LGM was chosen for all three clades.

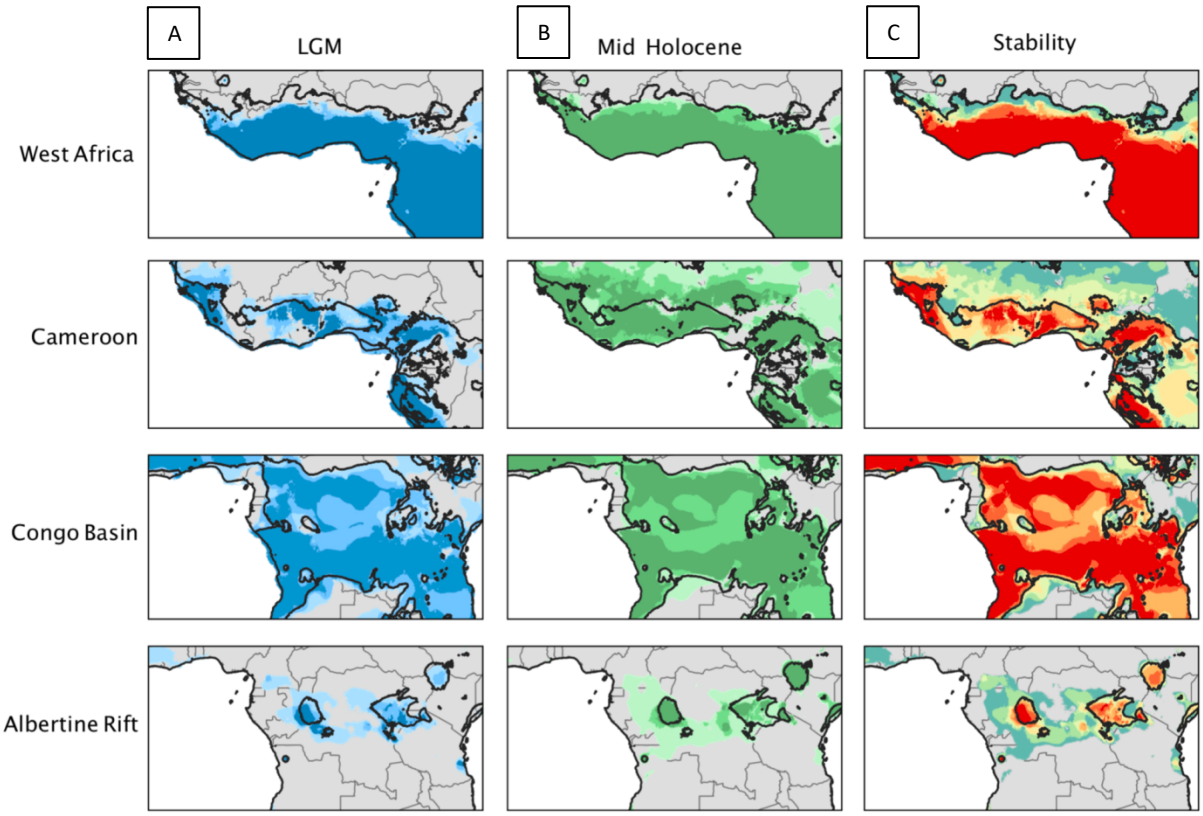


Figure 3.7 Niche models showing (A) suitable range area for *Trachylepis maculilabris* during the Last Glacial Maximum (LGM). The shade of blue represents agreement between global climate models (GCMs) with the darkest blue indicating agreement between all three GCMs and the lightest blue indicating support from only one GCM. (B) The suitable range area for *T, maculilabris* during the mid-Holocene. The shade of green represents agreement between GCMs with the darkest green indicating agreement between all three GCMs and the lightest green indicating support from only one GCM. (C) The stability of suitable range across the LGM, mid-Holocene, and present, with red indicating high stability and blue indicating low stability. In all figures black outlines delimit modeled present-day range.

CHAPTER 4

Origin and colonization history of the *Trachylepis maculilabris* complex in the Gulf of Guinea

Allen, K.E., Ceríaco, L.M.P., Bauer, A.M., Greenbaum, Tapondjou N., W.P., E. Kusamba, C., Kameni N., M.M., Brown, R.M. Pleistocene refugial diversification across highland and river systems in a widespread Central African skink. *In prep*.

4.1 Abstract

The Gulf of Guinea islands formed through volcanic activity as an offshore extension of the Cameroon Volcanic Line in Central Africa. This small island chain has unusually high biodiversity, likely as a result of colonization from the mainland combined with in-situ diversification. Colonization histories for many species on the islands are complex and influenced by the proximity of several large river systems and oceanic currents. In this study, we use a combination of dated phylogenetic inference, population structure analyses, and machine-learning driven demographic analyses to determine the most likely colonization history for the widespread *Trachylepis maculilabris* complex in the Gulf of Guinea (comprised of *T. thomensis* on São Tomé and *T. adamastor* on Príncipe and Tinhosa Grande islands). We test hypotheses of single versus multiple colonization events, determine the geographic origin of colonization on the mainland, and estimate the patterns of colonization within the islands. We found high levels of genetic divergence between *T. thomensis*, *T. adamastor*, and *T. maculilabris* on the mainland,

with little genetic divergence between *T. adamastor* on Príncipe and Tinhosa Grande. In addition, we found that the island taxa were monophyletic, with a single colonization event most likely originating from the eastern Congo Basin via the Congo River. We found a divergence date from the mainland of around 14 million years, corresponding to the age of uplift of São Tomé Island. Our demographic analysis suggests island colonization in a successive manner from São Tomé to Príncipe to Tinhosa Grande, contrary to both the progression rule and the stepping stone model of island colonization.

4.2 Introduction

Oceanic island colonization patterns depend strongly on the physical characteristics of the island or archipelago being colonized. The size of the islands and their distance from the mainland (the theory of island biogeography; Hearney, 2000; McArthur & Wilson, 1967; Whittaker et al., 2008) and the ages of the islands (the progression rule; Juan et al., 2000; Roderick & Gillespie, 1998; Wagner & Funk, 1995) have been often cited as the largest determinates of the order in which islands are colonized, and thus which house the oldest phylogenetic lineages. In linear island systems, a stepping stone model can also be applied, in which colonization occurs on the closest island to the mainland first and then progresses up the island chain from nearest to farthest (Hisheh et al., 1998). This model assumes that all islands are already formed at the time of colonization and that island age is not a factor. While these models are often sufficient to explain island colonization history (e.g., Cowie & Holland, 2008; Juan et al., 2000; Parent et al., 2008; Shaw & Gillespie, 2016; Yang et al., 2018), they may also be confounded by outside factors such as wind or ocean currents (e.g., Arjona et al., 2018, 2020; Bell et al., 2015b; Reilly et al., 2019).

4.2.1 Gulf of Guinea Islands

The Gulf of Guinea islands of Central Africa are a linear island chain comprised of one continental and three main oceanic islands, formed as an offshore extension of the Cameroon Volcanic Line (Lee et al., 1994). The continental island, Bioko, is the youngest, having formed only 1 to 3 million years ago, and was periodically connected to the mainland by a land bridge during the glacial cycles of the Pleistocene (Figure 4.1; Deruelle et al., 1991; Marzoli et al., 2000; Meyers et al., 1998). The three oceanic islands, Príncipe, São Tomé, and Annobón, have never been connected to the mainland (Aka et al., 2001; Burke, 2001; Caldeira & Munhá, 2002; Déruelle et al., 1991). The closest island, Príncipe, is approximately 31 million years old and has an area of 128 km^2. The second closest, São Tomé, is younger at 14 million years old and much larger at 836 km^2. The final island, Annobón, is the smallest and the youngest at 5 million years old and 17 km^2 (Lee et al. 1994; Figure 4.1).

Despite their small size, these oceanic islands host a remarkable diversity of both endemic and widespread species (Jones, 1994), including at least 1,104 species of angiosperms (Figueiredo et al., 2011), 146 species of birds (de Lima & Melo, 2021), 24 species of reptiles (Ceríaco et al., 2018), seven species of frogs, and two species caecilians (AmphibiaWeb, 2021). This high biodiversity is a combined effect of dispersal from the mainland and in-situ diversification (Bell et al., 2015a, b; Jesus et al., 2009; Jones, 1994; Measey et al., 2007; Melo et al., 2011; Miller et al., 2012). Because many species are relatively intolerant of salt water (Balinsky, 1981), the most likely method of dispersal is via rafting on vegetation mats, swept out to sea during periods of heavy rain (De Balsac & Hutterer, 1982; Measey et al., 2007; De Queiroz 2014). Such vegetation rafts have been noted in major river systems of Africa and South

America (King, 1962; Renner, 2004), and have been found carrying a variety of terrestrial vertebrates (Archaval et al., 1979; Iherring, 1911; Schiesari et al., 2003). Such dispersal would be greatly assisted by massive decreases in ocean salinity, reported in the Gulf of Guinea, following heavy rainfall and high river discharge during the Central African rainy season (Measey et al., 2007).

Three major rivers flowing into the Gulf of Guinea provide potential sources for the vegetation rafts and freshwater plumes needed for overwater dispersal to the islands: the Niger in West Africa and the Congo and Ogooué in Central Africa (Bell et al., 2015a; Measey et al., 2007; Figure 4.1). The mouth of the Ogooué is the closest in proximity to the islands, but strong currents from both the south (the Benguela current) and the north (the Guinea current) have the potential to carry rafts from the Niger and the Congo rivers (Measey et al., 2007; Richardson & Walsh, 1986; Figure 4.1). In fact, phylogenetic evidence has suggested that discharge from the Congo River may be a major source of amphibian colonization of Gulf of Guinea islands; numerous island species exhibit close relationships with Central and East African species (Bell et al., 2015a; Drewes & Wilkinson, 2004; Loader et al., 2007; Measey et al., 2007; Uyeda et al., 2007; Wilkinson et al., 2003). This Central-East African relationship is also observed in the case of Príncipe endemic shrew, *Crocidura fingui* (Ceríaco et al., 2015; Nicolas et al., 2019). Reptiles on the islands show a more varied pattern, with some being more closely related to Central and East African species (Carranza & Arnold, 2006; Ceríaco et al., 2021; Jesus et al., 2005a, 2006, 2007; Kornilios et al., 2013; Soares et al., 2018), and some to West African species (Engelbrecht et al., 2019; Jesus et al., 2009; Wüster et al., 2018).

The phylogenetic patterns of colonization, speciation, and endemicity have also been found to vary widely across taxa within the islands, with some species showing strong evidence

of multiple colonization events, often from different geographic sources (Ceríaco et al. 2015; Jesus et al., 2005a; Soares et al., 2018). In contrast, other species show evidence of a single colonization from the mainland followed by stepping stone colonization of each of the other islands (Bell et al., 2015b; Jesus et al. 2007, 2009; Melo et al. 2011; Miller et al. 2012). However, many of these studies are limited by genetic, taxonomic, and/or geographic sampling of populations on the mainland, especially in the Congo River basin, and the exact origins and timing of dispersal for many of these taxa are not well understood.

4.2.2 Study system

In this study we focused on the widespread sub-Saharan African skink species complex, *Trachylepis maculilabris*. This complex is well known for its overwater dispersal abilities, having representative species on islands in the Gulf of Guinea, the Indian Ocean, and on Fernando de Noronha Island off the coast of Brazil (Bauer, 2003; Mausfeld et al., 2000; Weinell et al., 2019). Three species within this complex have been recently described from the islands in the Gulf of Guinea: *Trachylepis thomensis* from São Tomé (Ceríaco et al., 2016), *T. principensis* from Príncipe (Ceríaco et al., 2016), and *T. adamastor* from Tinhosa Grande Islet off the coast of Príncipe (Ceríaco, 2015). However, based on genetic work, *T. principensis* was synonymized with *T. adamastor* by Ceríaco et al. (2020), and the two populations will be referred to in this paper as the Príncipe and Tinhosa Grande populations.

A phylogenetic study of the Gulf of Guinea island species by Jesus et al. (2005a) found *T. thomensis* and *T. adamastor* to be highly genetically distinct from each other and from the mainland. However, the colonization history of these species is unclear. Whereas they are treated as monophyletic by Ceríaco et al. (2020) and were presumed to be monophyletic by Jesus et al.

(2005b), phylogenetic studies by Jesus et al. (2005a) and Weinell et al. (2019) suggest separate geographic origins in West and East Africa for each of the two island species. However, geographic and genetic sampling in each of these studies was sparse and, in each case, nodal support for the clade containing the Gulf of Guinea species was low.

4.2.3 Aims and hypotheses

In this study we used a combination of dated phylogenetic inference, population structure analyses and machine-learning driven demographic analyses on 2,893 ddRAD loci to test hypotheses of single versus multiple colonization events for the *Trachylepis maculilabris* complex in the Gulf of Guinea, to determine the mainland geographic origin of colonization, and to infer patterns of colonization within the islands. The level of genetic divergence between *T. thomensis*, *T. adamastor*, and the mainland was high. However, there was little genetic divergence between *T. adamastor* on Príncipe and on Tinhosa Grande, although each island population formed reciprocally monophyletic clades. Contrary to several previous studies based solely on traditional Sanger data and few loci, we found that the island taxa to be monophyletic, with a single colonization event, most likely from the eastern Congo Basin, and a divergence date of around 14 million years, corresponding to the age of uplift of São Tomé. Our demographic analysis suggests island colonization in a successive manner, from São Tomé to Príncipe to Tinhosa Grande, an order that contradicts both the progression rule and the stepping stone model of island colonization.

4.3 Materials and Methods

4.3.1 Genetic data collection and processing

We obtained 83 specimens of the *Trachylepis maculilabris* complex (*T. maculilabris*, *T. thomensis*, and *T. adamastor* from Príncipe and Tinhosa Grande), as well as three specimens of our outgroup *Trachylepis affinis*, through fieldwork and from various museums. Sampling was representative of the known range of this species complex throughout the upper and lower Guinean forest blocks of West and Central Africa and the islands in the Gulf of Guinea. Museum catalog numbers and locality data for each specimen are presented in Appendix III Table S1.

We extracted genomic DNA (gDNA) from tissue samples that were preserved in 95% ethanol or RNAlater (Sigma-Aldrich) using the Maxwell RSC system (Promega). A modified version of the ddRADseq protocol of Peterson et al. (2012) was used to sequence genome-wide anonymous nuclear markers for each individual. A total of 300–500 ng of gDNA was double digested from each sample using the restriction enzymes *Sbf*I (restriction site 5'- CCTGCAGG-3') and *Msp*I (restriction site 5'-CCGG-3'). The resulting double digestion products were then bead-cleaned with AmpureXP beads (Agencourt) and individually barcoded using custom oligonucleotide adapters. Size-selection to a mean insert length of 541 base pairs (bp) (487–595 bp range) with internal standards was performed on pooled samples using a Pippin Prep (Sage Science, Beverly, MA). Resulting post-ligation products were amplified for eight cycles with a high-fidelity polymerase (Phusion, New England Biolabs). An Agilent TapeStation was used to determine the final fragment size distribution and concentration of each pool. Library pools were combined in equimolar amounts for sequencing on one Illumina HiSeqX lane (with a 10% *Phi*X spike-in and 150 bp paired-end reads).

Illumina reads from the ddRAD libraries were processed using STACKS v. 2.4 (Catchen et al., 2013). To recapitulate the original orientation of fragments in the genome, we used a read-stitching approach following Hime et al. (2019) to join the first read from an Illumina read pair

with the reverse complement of the second. Stitched reads were quality-filtered and demultiplexed by individual with the process_radtags function in STACKS with the following parameters: demultiplex each library by in-line barcode, check for both restriction enzyme cut sites, remove any read with an uncalled base, rescue barcodes and RAD-Tags, and discard any read with average Phred quality score < 20 over sliding windows of 15% of the total read length. Next, we used STACKS to assemble, *de novo*, the filtered and stitched Illumina read pairs.

We produced two separate ddRAD data sets, one for the entire complex including mainland populations and one for the island populations. Because the optimal *de novo* assembly of ddRADseq data can vary widely across taxa (Paris et al., 2017; Shafer et al., 2017), we tested a range of assembly parameters to optimize the recovery of putatively single-copy orthologous loci. Final assembly parameters were selected based on the methods of Paris et al. (2017). According to their recommendations, in USTACKS, we kept m (the minimum number of reads needed to form a stack) at 3 while in CSTACKS, we varied M (the number of mismatches allowed during loci formation) and n (the number of mismatches allowed during catalog formation) until we identified the parameters at which the maximum number of polymorphic loci were available across 80% (r = 0.8) of the population. For both of our data sets, this was found to be M = 8 with an optimal n of 15. Further parameters were tested in POPULATIONS separately for each data set in order to balance missing data and number of polymorphic loci. For the entire complex we used r = 0.25 and p = 5 [p is the minimum number of populations in which a locus must be present (here 5/5)], and for the island data set we used r = 0.5 p = 3. For each population we generated a data set comprising only a single random SNP per locus for population clustering analyses and demographic modeling, and another data set comprising full-length sequences for all loci for use in phylogenetic reconstruction.

4.3.2 Phylogenetic relationships

We used a maximum likelihood analysis implemented in IQtree v. 1.6.12 to analyze our SNP data set (Nguyen et al., 2014). We implemented 10,000 ultrafast bootstraps (Hoang et al., 2018) and the ModelFinder function to choose the best substitution model (Kalyaanamoorthy et al., 2017). A species tree analysis was performed using SNAPP implemented through Beast v 2.6.3 (Bryant et al., 2012). We followed Stange et al. (2018) to date the phylogeny using a lognormal distribution with an offset of 0, a mean divergence time of 5.16 million years, and a standard deviation of 0.13 between the West and Central African populations as estimated in Allen et al. (2019). Three non-admixed individuals were used per population, with two MCMC chains run for 1,000,000 generations each, and the island species were constrained to be monophyletic based on the results of the maximum likelihood analysis. Convergence was visualized in Tracer v1.7.1 (Rambaut et al. 2018) with the first 10% discarded as burn-in.

4.3.3 Population structure

We used both multivariate and admixture-based analyses to assess population structure within the island populations. DAPC was run using Adegenet v. 2.1.1 (Jombart & Ahmed, 2011). This approach uses discriminant functions to maximize variation among clusters and minimize variation within clusters. Bayesian information criterion (BIC) scores were used to choose the best-clustering scheme with the number of clusters (K) ranging from 1–10. Ancestry proportions of all individuals were inferred using LEA v. 1.6.0 (Frichot & François, 2015) through the Bioconductor v. 3.4 package. The sNMF function was used to assess values of K from 1–5 with

200 replicates, estimate individual admixture coefficients, and select the value of K that minimized cross entropy (François, 2016; Frichot et al., 2014).

4.3.4 Geographic ancestral state estimation

The ancestral geographic range of the mainland and island populations was estimated using BioGeoBEARS v. 0.2.2–2 (Matzke, 2013) on our dated species tree. Geographic areas were divided into 12 ranges (Figure 4.4b) based on known dispersal barriers in Central and West Africa including the Dahomey Gap, Cameroon Volcanic Line, Ogooué River, Congo River, Albertine Rift, and the three islands. We ran six standard biogeographical models: DEC (Ree & Smith., 2008), DEC+J, DIVALIKE, (Ronquist, 1997), DIVALIKE+J, BAYAREALIKE (Landis et al., 2013), and BAYAREALIKE+J. The free parameter *j* models founder-event speciation at cladogenesis (Matzke, 2014). We set our maximum range size to three. Models were evaluated using maximum likelihood and AICc (Burnham & Anderson, 2002).

4.3.5 Inter-island colonization and gene flow

To model demographic processes among our island populations, we used the R package delimitR (Smith & Carstens, 2020; https://github.com/meganlsmith/delimitR). This program uses a binned multidimensional folded site frequency spectrum (bSFS; Smith et al., 2017) and a random forest machine learning algorithm to compare speciation models such as ‘no divergence,’ ‘divergence with and without gene flow,’ and ‘divergence with secondary contact’ (Smith & Carstens, 2020). A bSFS was used because it stores the observed frequencies of the minor alleles for multiple populations and bins them to avoid inference problems associated with sampling too few segregating sites (Smith et al., 2017; Terhost & Song, 2015). Demographic

histories are simulated using the multi-species coalescent model implemented through fastsimcoal2 (Excoffier et al., 2013) under a user-specified guide tree and set of priors on divergence times, population sizes, and migration rates. The random forest classifier then creates a user-defined number of decision trees from a subset of the prior. Each decision tree compares the empirical bSFS to the SFS of each simulated speciation model and votes for the most similar model. The demographic model with the largest number of votes is chosen as the best model. Out-of-bag error rates are used to assess the power of the random forest classifier. The posterior probability of the selected model is then calculated by regressing against the out-of-bag error rates following Pudlo et al. (2015).

We created folded multidimensional site frequency spectrums using easySFS (https://github.com/isaacovercast/easySFS), a wrapper for ∂a∂i (Gutenkunst et al., 2009). We simulated 10,000 data sets for between-island diversification under models for no divergence, divergence with gene flow, divergence with secondary contact, and divergence without gene flow. Priors were drawn from uniform distributions for population size: 2,000 – 12,000 for *T. thomensis*, 600 – 6,000 for *T. adamastor* (Príncipe), 300 – 3,000 for *T. adamastor* (Tinhosa Grande), and 300,000 – 3,000,000 for the mainland *T. maculilabris* populations. Population sizes represent haploid individuals (twice the number of estimated diploid individuals) and are based on island size and the known population density of island populations of *Trachylepis*, including the closely related *Trachylepis atlantica* on the island of Fernando de Noronha (de Oliveira Gasparotto et al., 2020) and *T. adamastor* on Tinhosa Grande (Sousa, 2017). Divergence times in number of generations, assuming one generation per year for *Trachylepis* (Goldberg, 2008): 20,000 – 2,000,000 generations for *T. adamastor* (Príncipe) – *T. adamastor* (Tinhosa Grande) and 3,000,000 – 7,000,000 for *T. adamastor – T. thomensis*; the migration rate ranged from

0.0000005 – 0.0005. For each set of analyses, we coarsened our empirical site frequency spectra to 10 bins each. Our out-of-bag error rates were calculated, and 500 random-forest classifiers were simulated using 100,000 pseudo-observed data sets for each model. A confusion matrix was calculated to determine how often the correct model was selected and posterior probability for the "best" model was estimated for each species.

4.4 Results

4.4.1 Genetic data collection and processing

After filtering, our complex level ddRAD data set, including the island species and *T. affinis* as an outgroup, consisted of 2,893 loci with 41.2% missing data (defined as proportion of missing loci across all individuals), and an effective mean per-sample depth of coverage of 59.2 +/- 13.0. Our island data set consisted of 2,817 loci with 15.3% missing data, and an effective mean per-sample depth of coverage of 62.3 +/- 13.6. The concatenated ddRAD data set used for phylogenetic analyses had a length of 849,977 bp.

4.4.2 Phylogenetic relationships

The broad-scale phylogenetic relationships estimated from our SNP data sets were strongly supported at all internal nodes. Our analysis supported two deeply divergent, monophyletic island clades representing *T. thomensis* and a combination of *T. adamastor* from Príncipe and Tinhosa Grande, which were found to be minimally divergent (Appendix III Figure S1). Our SNAPP species tree analysis recovered identical relationships to those of our IQtree analysis with all nodes being highly supported (Figure 4.2). Our dating analysis estimated divergence between the island and mainland species to be around 13.9 million years old, between

T. thomensis and *T. adamastor* to be around 12.9 million years old, and between *T. adamastor* from Príncipe and Tinhosa Grande to be around 0.67 million years old (Figure 4.2).

4.4.3 Population structure and demography

Our admixture-based genetic structure method, LEA (Frichot & François, 2015), identified two to three populations in our island data set with nearly equal probability, with the three-population model separating *T. adamastor* from Príncipe and Tinhosa Grande (Figure 4.3). A comparison of BIC values in our DAPC analysis (Jombart & Ahmed, 2011) strongly supported four populations, splitting *T. adamastor* (Príncipe) and *T. adamastor* (Tinhosa Grande) and implying two genetic clusters within *T. thomensis* (Appendix III Figure S2).

Analyses in DelimitR (Smith & Carstens, 2020) supported a successive pattern of island colonization with the most highly supported model being gene flow between *T. thomenis* and *T. adamastor* (Príncipe) and between *T. adamastor* (Príncipe) and *T. adamastor* (Tinhosa Grande; 119/500 votes) with a posterior probability of 0.69 (Figure 4.4c), genetic contact between all island populations received the second highest support (105/500 votes) with very little support for the remaining models (Appendix III Table S2). The out-of-bag error rate was 13.9%, with all misclassifications being between highly similar models.

4.4.4 Biogeographic origins of island species

The maximum likelihood score of our ancestral range reconstructions through BioGeoBEARS supported DEC+J and DIVALIKE+J equally, and in fact both model outputs were nearly identical, whereas the AICc score supported DIVALIKE as the best-fitting model (Appendix III Table S3). Both sets of supported models suggested different origins for the island

clade, with the 'jump' models supporting an origin in West Africa, and the DIVALIKE model supporting an origin in Central Africa (Figure 4.4). Within the mainland populations, the +J models supported population origins in West Africa with subsequent diversification in Cameroon, and a long-distance dispersal event, inferred from Cameroon to the Albertine Rift, followed by diversification throughout Central Africa. The DIVALIKE model supported a widespread ancestral population throughout Central and West Africa with subsequent vicariance into West and Central Africa, followed by diversification in each area. Given that the DIVALIKE model was the most highly supported by AICc over each of the other models (Appendix III Table S3), and that there are known statistical issues with the +J model (Ree & Sanmartín, 2018), the DIVALIKE model will be preferred in this study.

4.5 Discussion

In this study we set out to test hypotheses related to the origin and colonization history of the *Trachylepis maculilabris* complex in the Gulf of Guinea. We found, contrary to several previous studies (Jesus et al., 2005a; Weinell et al., 2019), that *T. thomensis* and the populations of *T. adamastor* on Príncipe and Tinhosa Grande are reciprocally monophyletic with respect to the *Trachylepis maculilabris* populations on the mainland (Figure 4.2). Our results suggest that these species originated from a single colonization event from East-Central Africa, which coincided with the timing of the uplift of São Tomé Island (Figure 4.4). Demographic analyses of each of these island populations suggest successive colonization from São Tomé to Príncipe, and then to Tinhosa Grande (Figure 4.4). Our results corroborated findings of Ceríaco et al. (2020) by identifying the *T. adamastor* population on Tinhosa Grande as being conspecific with *T. adamastor* on Príncipe, despite considerable phenotypic distinctiveness, such as greater body

size and extensive melanic pigmentation (Figure 4.2; Ceríaco, 2015). This study represents one of the most comprehensive analyses of a species complex in the Gulf of Guinea to date, both in terms of numbers of genetic loci and geographic sampling on the mainland, and adds to our growing understanding of how rivers, ocean currents, and in-situ diversification drive biodiversity patterns between the Gulf of Guinea island system and its mainland source populations (Bell et al., 2015a,b, 2017; Ceríaco, 2015; Ceríaco et al., 2015, 2016, 2020; Jesus et al., 2006, 2007, 2009; Measey et al., 2007; Soares et al., 2018).

4.5.1 Phylogenetic relationships

Phylogenetic analysis of this complex strongly supports the reciprocal monophyly of the Gulf of Guinea species with respect to the populations of *T. maculilabris* on the mainland in West and Central Africa (Figure 4.2). The divergence between the island species, and between the island species and the mainland, was deep, with a divergence time of approximately 13.9 million years between the islands and the mainland, and 12.9 million years between *T. thomensis* and *T. adamastor* (Figure 4.2). The two populations of *T. adamastor* from Príncipe and Tinhosa Grande were estimated to have diverged around 0.67 million years ago. This is in line with the findings of Ceríaco et al. (2020), who suggested that Príncipe and Tinhosa Grande were connected during periods of low sea levels throughout the Pleistocene, and that populations on each island were likely in contact during these periods. Despite their low genetic divergence, these two populations are reciprocally monophyletic in our phylogenetic analysis.

Our population structuring analysis, LEA (Frichot & François, 2015), gave nearly equal probability to the likelihood of two and three genetic clusters in the Gulf of Guinea (Figure 4.3), likely because of the distinct but shallow divergence between *T. adamastor* from Príncipe and

Tinhosa Grande. Interestingly, our DAPC (Jombart & Ahmed, 2011) analysis detected a fourth genetic cluster, identifying two clusters in *T. thomensis* (Appendix III Figure S2). This genetic signal was not identified by our phylogenetic analyses, but shallow genetic divergence has been identified along the island's elevational gradient in other species (Bell et al., 2015b), and further sampling would be needed to determine if this was the case in *T. thomensis*.

4.5.2 Colonization of the Gulf of Guinea

The monophyly of each island species strongly suggests a single colonization event for the *T. maculilabris* complex in the Gulf of Guinea (Figure 4.2), in contrast to several previous interpretations (Jesus et al., 2005a; Weinell et al., 2019). Both of the latter studies inferred multiple colonization events from the mainland. However, Jesus et al. (2005a) suggested an East African origin for *T. thomensis* and a West-Central African origin for *T. adamastor*, whereas Weinell et al. (2019) found the opposite geographic origin for each species. In our study, an ancestral biogeographic range reconstruction, including both island and mainland species and populations, is strongly indicative of an East-Central African origin for the island colonization event (Figure 4.4a,b). Both the ancestral biogeographic range reconstruction and the demographic analyses indicate São Tomé as the first island colonized (Figure 4.4c).

Similar colonization patterns have been found in many other taxa inhabiting the Gulf of Guinea islands. Sister species in East and Central Africa have been identified for a large majority of frog species (Bell et al., 2015a,b, 2017; Drewes & Wilkinson, 2004; Loader et al., 2007; Measey et al., 2007; Uyeda et al., 2007; Wilkinson et al., 2003; Zimkus et al., 2010), and a variety of lizard and snake species (Carranza & Arnold, 2006; Ceríaco et al., 2021; Jesus et al., 2005a, 2006, 2007; Kornilios et al., 2013; Soares et al., 2018). In all instances, the most likely

method of dispersal is overseas via rafts swept from the Congo River during periods of heavy rain and carried towards the island by the south to north Benguela current (Figure 4.1; Measey et al., 2007). At 836 km^2, São Tomé is the largest of the three islands, and is in the direct line of the Benguela current, making it a likely target for colonization (Figure 4.1; Bell et al., 2015a; Jesus et al., 2009; Measey et al., 2007).

4.5.3 Inter-island dispersal

Ancestral biogeographic range reconstructions and demographic analyses support a successive colonization of the island system from São Tomé to Príncipe to Tinhosa Grande (Figure 4.4). This appears to be a common pattern of dispersal in the Gulf of Guinea islands (Bell et al., 2015ab; Jesus et al., 2009; Measey et al., 2007), but it is unusual in that it goes against the widely accepted progression rule (Bell et al., 2015b; Wagner & Funk, 1995), in which islands are often colonized from oldest to youngest (Cowie & Holland, 2008; Juan et al., 2000; Parent et al., 2008), as well as the stepping stone model, in which islands are colonized from closest to furthest from the mainland (Hisheh et al., 1998). In the case of these islands, São Tomé is far younger than Príncipe (14 Mya vs. 31 Mya), but far larger (836 km^2 vs. 128 km^2), potentially increasing the likelihood of being intercepted by vegetative rafts (Figure 4.1; Lee et al. 1994; Measey et al., 2007).

Our dating analysis estimated divergence between the island and mainland species to be around 13.9 million years old, corresponding with the formation of São Tomé (Figure 4.2). The estimated divergence of 12.9 million years between *T. thomensis* and *T. adamastor* suggests a relatively swift colonization of Príncipe from São Tomé (Figure 4.2). Tinhosa Grande is a small islet adjacent to Príncipe and the two *T. adamastor* populations on each island were originally

described as separate species (*T. adamastor* on Tinhosa Grande and *T. principensis* on Príncipe; Ceríaco, 2015; Ceríaco et al., 2016) based on morphological differences since *T. adamastor* on Tinhosa Grande is significantly larger and melanistic (Ceríaco, 2015). Ceríaco et al. (2020) later determined that these species were not genetically distinct and speculated that they were likely connected by a land bridge during periods of low sea level during the Pleistocene (Figure 4.1). Our dating analysis supports this assertion, suggesting that divergence between these populations is approximately 0.67 million years old (Figure 4.2), and we assume that colonization of the islet was one of only a few population establishments in the island chain that did not occur via oceanic dispersal.

4.6 Conclusion

The islands of the Gulf of Guinea represent a remarkably complex system for studying mainland-to-island as well as inter-island colonization patterns. The presence of three major river drainages (the Niger, the Ogooué, and the Congo) as well as two converging ocean currents (the Benguela and the Guinea; Figure 4.1), combined with in situ speciation, have given the island chain high levels of biodiversity and a complex biogeographic history (Jones, 1994; Measey et al., 2007). Many studies have sought to tease apart the phylogenetic histories of endemic reptile and amphibian taxa found on these islands (Bell et al., 2015a; Carranza & Arnold, 2006; Drewes & Wilkinson, 2004; Fritz et al., 2011; Jesus et al., 2005a, 2007; Loader et al., 2007; Measey et al., 2007; Uyeda et al., 2007; Wilkinson et al., 2003; Zimkus et al., 2010), but have often been limited by sparse taxonomic and genetic sampling, and by poor representation from throughout the Congo Basin—which is critical to understanding the role that the Congo River may have played in island colonization. Future studies of this complex system will facilitate research not

only of the formation of biodiversity in the Gulf of Guinea, but in any island system impacted by large, discharging, adjacent continental river systems and strong surrounding ocean currents.

4.7 Figures

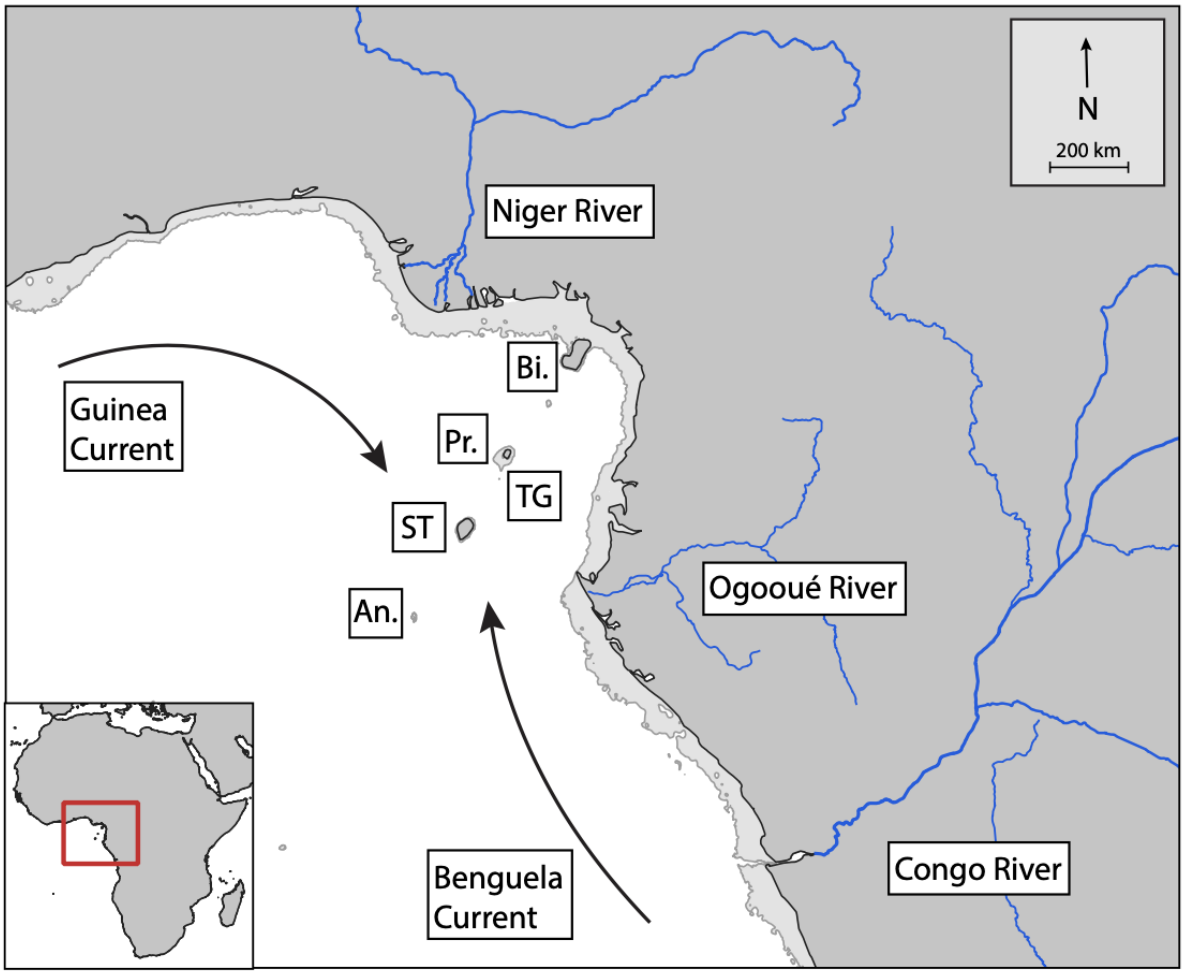


Figure 4.1 A map of the Gulf of Guinea showing major river systems and ocean currents. Estimated shoreline at the Last Glacial Maximum is shown in light grey. Islands are denoted as Bi.: Bioko, Pr.: Príncipe, TG: Tinhosa Grande, ST: São Tomé, An.: Annobón.

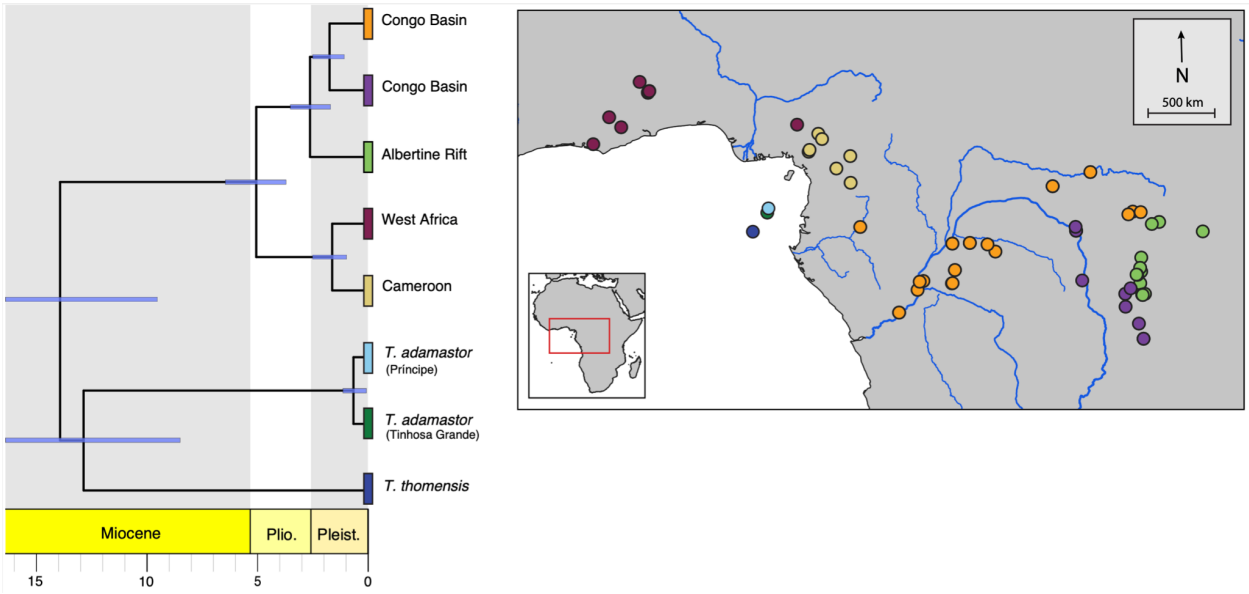


Figure 4.2 Dated species tree analysis and sampling localities of the major *Trachylepis maculilabris* clades found on the mainland and the species in the islands of the Gulf of Guinea. All nodes were highly supported (pp>0.99).

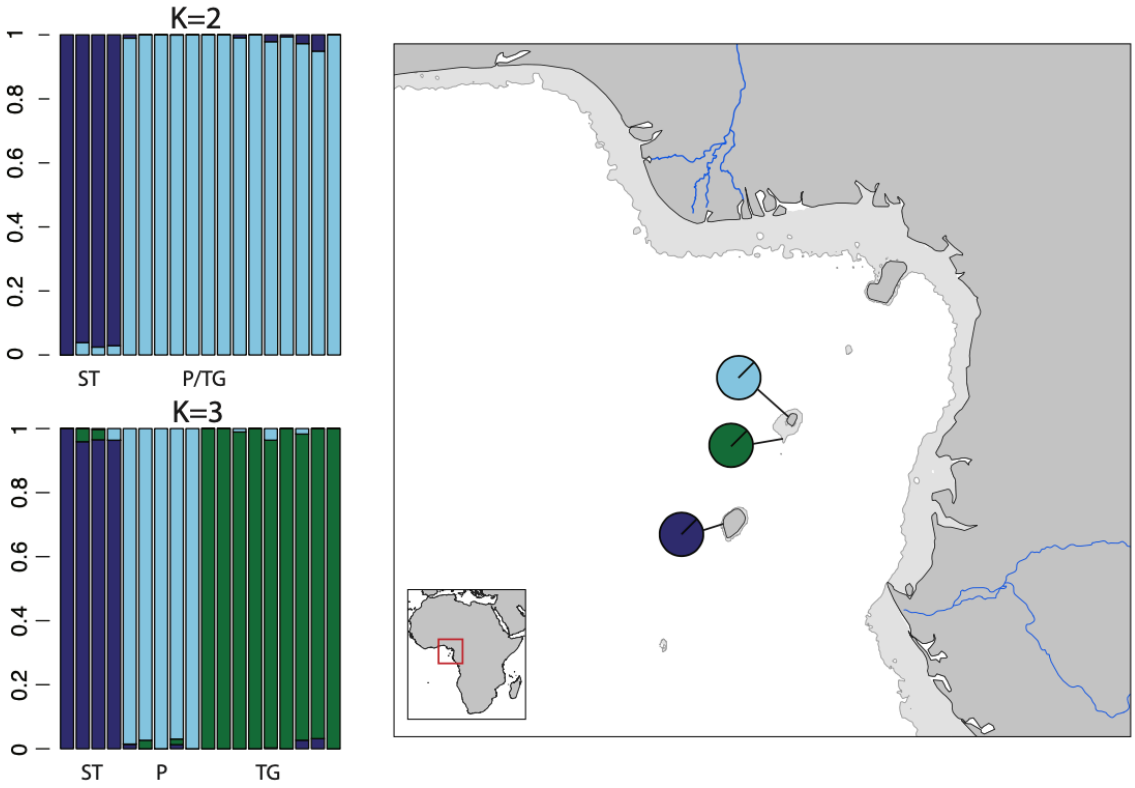


Figure 4.3 Population structure of the species in the *Trachylepis maculilabris* complex in the Gulf of Guinea for K=2 and K=3, analyzed in LEA.

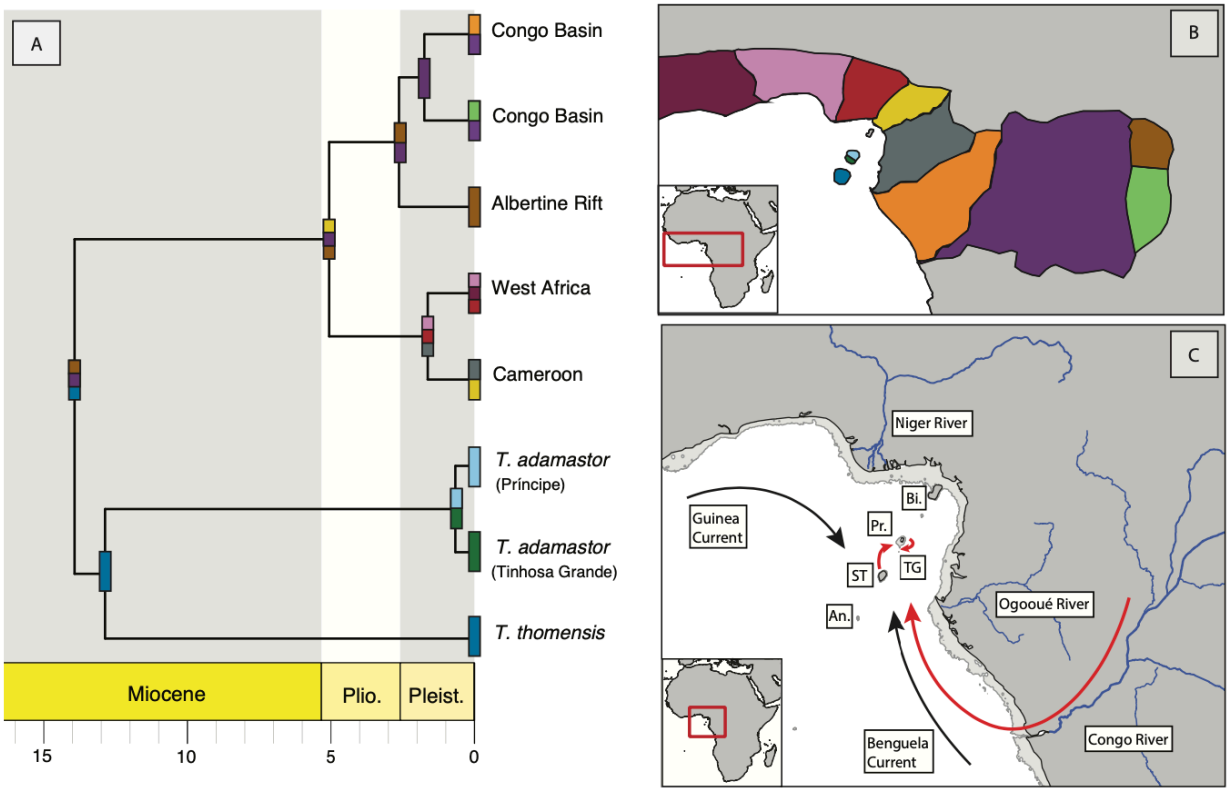


Figure 4.4 A) Ancestral distribution reconstruction created in BioGeoBEARS using the DIVALIKE model mapped onto the dated species tree of the major *Trachylepis maculilabris* clades and species in the islands of the Gulf of Guinea. B) Map of biogeographic areas used as input values for BioGeoBEARS the maximum number of areas per species was set to three. C) Most likely colonization scenario for *T. thomensis* and *T. adamastor* estimated from BioGeoBEARS and delimitR. Islands are denoted as Bi.: Bioko, Pr.: Príncipe, TG: Tinhosa Grande, ST: São Tomé, An.: Annobón.

CHAPTER 5

Allopolyploidization and niche evolution in African clawed frogs (*Xenopus*)

Allen, K.E., Evans, B.J., Tapondjou N., W.P., Ané, C., Brown, R.M., Peterson, A.T. Allopolyploidization and niche evolution in African clawed frogs (*Xenopus*). Evolution. *In review.*

5.1 Abstract

Divergence between fundamental ecological niches occupied by closely related species is relatively rare. One hypothesized exception to this is through formation of allopolyploid species, a process that is widespread in plants but is far less well known in animals. To test the prediction that allopolyploidization is coupled with fundamental niche shifts, we examined the African clawed frog genus *Xenopus*, whose evolution is characterized by allopolyploid speciation events interspersed with bifurcating speciation. Our ancestral niche reconstructions — estimated on a reticulate phylogeny and accounting for uncertainty in niche limits — correspond with geographic and habitat reconstructions for the genus, giving a geographic interpretation to observed niche evolution. We found that expansions and contractions around the ancestral niche were much more common than niche shifts, and that niche novelty was even rarer still. Niche evolution, primarily through expansion and contraction, occurred at equal frequencies in polyploid species formed by allopolyploidization compared to those evolving by bifurcating speciation. However, nearly all incidences of niche novelty were associated with allopolyploidization. These findings are largely in line with expectations developed from studies

on polyploid plants and underscore the contribution of polyploid speciation to rapid adaptation and novel ecological opportunity across plants and animals.

5.2 Introduction

5.2.1 Niche evolution

Many species do not occupy the full range of environmental conditions in which they could survive and reproduce (Hastings et al., 2005; Simberloff & Rejmánek, 2011). The fundamental ecological niche concept refers to the complete set of environmental conditions and resources that a species can utilize (Hutchinson, 1957), whereas the realized niche refers to the environmental conditions that a species actually occupies (Connell, 1961; Pulliam, 2000). Because species' geographic distributions are often limited by barriers to dispersal or competitive interactions with other organisms, realized niches are generally smaller than fundamental niches (Connell, 1961; Pulliam, 2000; Soberón & Peterson, 2005).

The fundamental niche of a species is genetically controlled and thus can be acted on by natural selection and studied in a phylogenetic context (e.g., Hansen, 1997; Harvey & Rambaut, 2000; Holt & Barfield, 2008; Holt & Gomulkiewicz, 1997; Labra et al., 2009; Owens et al., 2020; Peterson et al., 1999; Saupe et al., 2018; Wiens, 2004; Wiens et al., 2010; Wiens & Graham, 2005). Studies of fundamental niches evolving through time on dynamic landscapes have largely found them to be slowly evolving relative to ecological factors such as microhabitat use (Ackerly et al., 2006; Farallo et al., 2020; Ficetola et al., 2018) and life history (Duran et al., 2013; Rodrigues et al., 2019). This fundamental niche conservatism is likely a result of evolutionary constraints on development, behavior, physiology, or population genetic factors (Almendra et al., 2018; Araújo et al., 2013; Holt & Gomulkiewicz, 1997; Jakob et al., 2010;

Labra et al., 2009; Valladares et al., 2014; Wiens, 2004; Wiens et al., 2010; Wiens & Graham, 2005). Additionally, hard physiological limits may exist for some aspects of niche evolution, such as heat tolerance (reviewed in Araújo et al., 2013) or salinity (Feng & Papeş, 2017; Nougué et al., 2015). However, taxa that are able to diversify more rapidly into new niche spaces may benefit from being able to reduce interspecific competition with closely related species (Kozak & Wiens, 2010; Peers et al., 2013; Pitteloud et al., 2017; Wereszczuk & Zalewski, 2015).

5.2.2 Niche evolution in polyploids

An interesting, but little-studied, potential exception to the general rules of niche evolution is allopolyploid speciation, wherein a novel polyploid lineage is created through hybridization between two parental species (Abbott et al., 2013; Mallet, 2007; Otto, 2007; Stebbins, 1971). Allopolyploid speciation generally takes place in areas where the two parental species occur in sympatry (Levin, 2013; Marchant et al., 2016), and newly formed polyploid species may be immediately reproductively isolated from both parental species, due to meiotic incompatibilities between ploidy levels (Evans, 2008; Glover et al., 2016; Kobel, 1996; Kobel & Du Pasquier, 1986; Sigel, 2016). Over the long term, survival of these new polyploid species is presumably favored by geographic or ecological separation, which minimizes competition with the parental species (e.g., Fowler & Levin, 1984; Levin, 1975, 2004). Aiding in this separation, allopolyploid genomes inherit variation from both parental species and, thus, have more variation than either. Additionally, allopolyploid genomes may undergo rapid structural and transcriptomic changes, which may facilitate rapid adaptation to environments outside of the tolerance of either of their parental species, especially if beneficial variants are partially dominant (Adams &

Wendel, 2005; Arnold, 2004; Baniaga et al., 2020; Levy & Feldman, 2002; Otto, 2007; Otto et al., 2000; Van de Peer, 2021).

Habitat shifts associated with polyploid species formation have been found to alter, relative to the parental species, tolerance of extremes in precipitation (Donkpegan et al., 2017; Thompson et al., 2014; Visger et al., 2016) and temperature (David & Halanych, 2021; Nakagawa, 2006; Thompson et al., 2014), and are often also associated with changes in niche breadth (Brittingham et al., 2018; Marchant et al., 2016; Theodoridis et al., 2013). Three patterns are broadly associated with the relationship between the niches of allopolyploids compared to those of their parental species: (1) niche expansion or contraction, (2) niche intermediacy between parental species, or (3) niche novelty (Marchant et al., 2016). Although newly formed allopolyploid species may avoid competition by inhabiting ecological niches that are distinguished from their parental species (Fowler & Levin, 1984; Levin, 1975, 2004), they often exhibit partial niche overlap with at least one of those species (Marchant et al., 2016).

5.2.3 Polyploidy in Xenopus

Polyploidy is widespread in plants but is far less common in animals, where it is most common in parthenogenetic species (reviewed in Fujita & Moritz, 2009; Mable, 2004). The nature of polyploidy is unique in amphibians because they are the only known vertebrates that include multiple examples of closely related diploid and bisexual polyploid species, and nearly all instances represent independent evolutionary events (Mable et al., 2011). Among the ~50 documented species of polyploid amphibians (Evans et al., 2012; Schmid et al., 2015), the anuran genus *Xenopus* is one of the most intensively studied in terms of effects of

polyploidization on diversification and genomic evolution (Evans, 2008; Evans et al., 2004; Evans et al., 2015; Furman et al., 2018).

Xenopus frogs are widely distributed throughout much of sub-Saharan Africa. Several species are used as model organisms for human health-related research (Cannatella & de Sa, 1993; Shapiro & Zwarenstein, 1934); however, surprisingly little is known about their ecology, natural history, or behavior in nature. The genus is characterized by remarkably high levels of polyploidization: of its 29 species, only one is diploid, 16 are tetraploid, seven are octoploid, and four are dodecaploid; the ploidy of *X. fraseri* remains unknown but is probably tetraploid. These taxa were formed through multiple independent allopolyploidization events (Evans, 2008; Evans et al., 2015), most likely via sequential backcrossing of hybrid individuals with parental species and combined with aberrant oogenesis in hybrid individuals wherein eggs are produced that carry the complete genome of a hybrid female (Evans, 2008; Kobel, 1996; Kobel & Du Pasquier, 1986).

The genus *Xenopus* comprises two lineages recognized as subgenera, *Xenopus* (*Xenopus*) and *Xenopus* (*Silurana*) (Evans et al., 2015; Kobel et al., 1996). Both subgenera contain polyploids and have independent diploid ancestors inferred from their ancestral chromosome counts [$2n = 20$ in *Xenopus* (*Silurana*) and $2n = 18$ in *Xenopus* (*Xenopus*), where n is the haploid number of chromosomes]. Both diploid ancestors of subgenus *Xenopus* are extinct or undiscovered, and only one of the two diploid ancestors of the *Xenopus* (*Silurana*) allotetraploids (*X. tropicalis*) is known.

A clear geographic component exists in allopolyploidization in *Xenopus,* in which octoploids and dodecaploids are found only in and around Central African highlands, including the Cameroon Volcanic Line and the Albertine Rift (Channing & Rödel, 2019; Evans et al.,

2015; Figure 5.1). Several hypotheses have been formulated to explain this pattern, including (1) that species with higher ploidy levels and thus more genetic material, are able to adapt rapidly to montane environments (Ficetola & Stöck, 2016; Litvinchuk et al., 2011; Otto et al., 2007), or (2) that changes in species' distributions during the Pleistocene glaciation cycles allowed species' ranges to overlap and allopolyploidization to occur (Ehrendorfer, 1980; Ficetola & Stöck, 2016; Haufler et al., 1995; Stöck et al., 2010). Because many allopolyploidization events may have occurred prior to the Pleistocene, it is also possible that montane regions acted as "lifeboats," which allowed *Xenopus* species, including octoploid and dodecaploids, to persist, and that tetraploid species managed to disperse more broadly once suitable lowland habitat became available (Evans et al., 2004).

5.2.4 Niche evolution analyses

Geographic or biotic barriers may prevent a species from reaching all suitable areas such that accessible areas do not represent the full diversity of environmental conditions that it can potentially tolerate. Thus, characterization of a species' entire fundamental niche from occurrence records alone can be difficult, if not impossible (Owens et al., 2013; Saupe et al., 2012, 2018; Warren et al., 2014). This complication is especially true for species that are endemic to geographically discrete areas such as islands, lakes, mountain tops, or river systems (Barve et al., 2011), as are many of the mid- to high-elevation *Xenopus* species. This reduced representation of the fundamental niche owing to access limitations is referred to as the existing niche; species' environmental use may be further reduced by interaction with other species to the realized niche (Peterson, 2011).

Many previous niche evolution analyses have been conducted using characteristics of realized niches such as minimum, maximum, mean, and median values (e.g., Ackerly et al., 2006; Cooper et al., 2011; Graham et al., 2004; Kozak & Wiens, 2010; Yesson & Culham, 2006). However, to fully assess the degree of niche evolution associated with phylogenetic divergence, it is essential to conduct analyses on the fundamental niche because analyses of realized niches inevitably overestimate niche evolution (Broennimann et al., 2012; Godsoe, 2010; Saupe et al., 2018; Warren et al., 2008, 2014). To accomplish this, we use a recently developed method of fundamental niche characterization and reconstruction that involves the binning of climatic variables into niche spectra, with each bin marked as "present," "absent," or "uncertain" for each species (Owens et al., 2020). Accounting for uncertainty in our reconstructions allows us to account for the unknown nature of the complete fundamental niche for many species and should provide more realistic and conservative results (Owens et al., 2020).

5.2.5 Aims and hypotheses

Assessing niche evolution of allopolyploid species in a phylogenetic context has been hampered by the methodological challenge of appropriately accommodating non-independent evolutionary histories associated with reticulate evolution. In this paper, we combined the niche characterization and reconstruction techniques of Owens et al. (2020) with a maximum likelihood-based ancestral state reconstruction procedure created specifically for reticulate phylogenetic trees (Bastide et al., 2018; Solís-Lemus et al, 2017). This approach enabled us to explicitly test the hypothesis that allopolyploidization events are associated with increased instances of niche evolution, including niche shifts, changes in niche breadth, or movement into novel or extreme environments compared with bifurcating speciation events in the same genus

(Adams & Wendel, 2005; Ainouche & Wendel, 2014; Arnold, 2004; Levy & Feldman, 2002; Sigel, 2016; Soltis & Soltis, 2000).

We focused specifically on four metrics of temperature and precipitation that are frequently used in niche evolution research: mean annual temperature, temperature seasonality, annual precipitation, and precipitation seasonality (Donkpegan et al., 2017; Nakagawa, 2006; Thompson et al., 2014; Visger et al., 2016). Because *Xenopus* are aquatic, these metrics were also chosen because several of them have been shown to correlate well with water temperature and quality (Collins et al., 2019; Dokulil et al., 2021; Komatsu et al., 2007; Sharma et al., 2008) In addition, we conducted ancestral state reconstructions for geography and habitat affiliations on our phylogenetic network to determine whether niche reconstructions coincide with historic and present-day estimates of habitat and geography, thereby allowing us to assess whether phylogenetic niche reconstructions are congruent with the known natural history of *Xenopus*.

5.3 Materials and Methods

5.3.1 Phylogenetic network

Ancestral state reconstructions were performed on a phylogenetic network constructed from the multi-labeled (MUL) RAG1 and RAG2 gene tree from Evans et al. (2015), with additional relationships within the *laevis* and *muelleri* species groups based on Furman et al. (2015) and Evans et al. (2019). The MUL tree was converted to a network using the cluster-based algorithm in Dendroscope v. 3.7.2 (Huson & Scornavacca, 2012). To ensure correct parental relationships and inferred ploidy of species in our network, we added four nodes for extinct ancestral species, for which the sub-genomes are known only from their hybrid offspring, and which have already been inferred for the genus *Xenopus* (the “lost ancestors;” Evans et al.,

2015). These species are labeled Ancestor A, B, C and D in the phylogenetic networks, and were coded as missing data for all character states in our ancestral state reconstructions. All added taxa, including ancestral species and those of the *laevis* and *muelleri* species groups, were placed at midpoints of internal branches subtending sister taxa.

5.3.2 Geography and habitat reconstructions

We ran ancestral state reconstructions on geography and habitat use in the genus *Xenopus* in order to assess coincidence between these variables and niche evolution events. However, the reticulate nature of the evolutionary history of *Xenopus* violates assumptions of many phylogenetic comparative methods and biogeographic analyses (e.g., BioGeoBEARS; Matzke, 2013). For this reason, ancestral state reconstructions for geography and habitat were performed using maximum likelihood-based approaches under an equal-transition-rate evolutionary model in the package PhyloNetworks (Solís-Lemus et al., 2017; Bastide et al., 2018). Each species was coded for one of eight geographic regions (North Africa, South Africa, East Africa [including the Albertine Rift], Congo Basin, Cameroon Volcanic Line [CVL], Ethiopian highlands, and coastal West-Central Africa) and five habitat types (lowland rainforest, lowland savanna, montane forest, montane grassland, and mixed montane forest/grassland). Geographic and habitat classifications for each species were based on Channing & Rödel (2019), IUCN Red List species accounts (IUCN, 2020), and AmphibiaWeb species accounts (www.amphibiaweb.com).

5.3.3 Ecological niche characterization

Because of the potential for species misidentification based solely on morphology, only genotyped and/or expert-identified specimens (Appendix IV Table S1) were used to characterize

niche space utilized by each species. The accessible area for each species, which is the area within which comparisons are made between presence sites and the background (Barve et al., 2011), was delimited as a buffer around presence points ranging from 50 km for range restricted species (e.g., *X. longipes, X. itombwensis*), to 500 km for widely distributed species (e.g., *X. laevis, X. muelleri*). For each species, several buffer sizes were tested, and we manually compared histograms of available versus used environmental parameters, to assure that an appropriate accessible area was implemented. In addition, each area was manually checked to ensure that no inaccessible areas were included within the distance-based buffer (e.g., islands, isolated mountaintops; see methods diagram in Appendix IV Figure S1).

Four climate variables, mean annual temperature, temperature seasonality, annual precipitation, and precipitation seasonality, were downloaded from the WorldClim v.1.4 database (Fick & Hijmans, 2017). Within the accessible area for each species, these climate variables were binned (e.g., by each degree Celsius) and niches were characterized using the recently developed R package Nichevol (Cobos et al., 2020; Owens et al., 2020). Species were scored as "present," "absent," or "uncertain" for each bin, with conditions for uncertain defined below. Binning climate variables gives a more complete understanding of species fundamental niche space than methods based solely on summary statistics (i.e., using the minimum, maximum, or median; e.g., Ackerly et al., 2006; Cooper et al., 2011; Graham et al., 2004; Yesson & Culham, 2006), and scoring environmental space as "uncertain" allows us to account for incomplete niche characterization.

5.3.4 Ecological niche evolution

Ancestral state reconstructions were performed on the *Xenopus* phylogenetic network using maximum likelihood-based approaches implemented by the package PhyloNetworks (Bastide et al., 2018; Solís-Lemus et al., 2017). Following Owens et al., (2020), bins were assembled into niche spectra showing "presence," "absence," or "uncertainty" along a climatic continuum (Appendix IV Figure S1). During ancestral state reconstruction, each bin was treated as a discrete character with two states (presence, absence) and assessed under an equal and reversible transition rate model of evolution. Uncertain bins were treated as missing. A Julia script to automate this process across bins in a niche spectrum is available in the Open Science Framework Digital Repository. Bins for which the probability of presence was greater than 0.7, were designated as "present," less than 0.3 were marked as "absent," and between 0.3 and 0.7 as "uncertain." These results were then smoothed (i.e., if an "unknown" bin was flanked by two "suitable" bins, it was coded as "suitable" following Owens et al. (2020); Appendix IV Figure S1). This increased continuity of areas of suitable and unsuitable habitat, based on the assumptions that species have unimodal responses to climatic conditions (Maguire, 1973) and that bins are not independent of one another within a single climate variable.

Each daughter species was then compared to its ancestral node with each of its respective parental species (one parental species if it arose by bifurcating speciation, or two parental species if it arose by allopolyploid speciation) to evaluate whether its niche is intermediate to that of its parentals (i.e., overlapping both by one or more bins), more similar to one parental (i.e., overlapping one by a greater number of bins than the other), or novel (not overlapping by any bins). Additionally, we evaluated niche expansion (i.e., possession of a niche that is wider than parental niche(s) by one or more bins on one or both sides of the niche spectrum), niche contraction (i.e., possession of a niche that is narrower than the parental niche by one or more

bins on one or both sides of the niche spectrum), and niche shift (i.e., expansion of a niche on one end of the spectrum and a contraction on the other by one or more bins). If the niche of parental species at a hybridization event was unknown, as for Ancestors A, B, C, and D, the projected niche for that node was used. Bins as defined above are 1°C for mean annual temperature, 100 mm for annual precipitation, 0.1 stdev for temperature seasonality and 10 cov for precipitation seasonality.

X^2 tests of goodness of fit were used to analyze the frequency of events in niche evolution categories (i.e., niche expansion, contraction, shift, or novelty), and a Z-test was used to compare proportions of niche evolution events associated with polyploidization versus those associated with bifurcating speciation. Both analyses were performed using R (RStudio Team, 2018).

5.4 Results

5.4.1 Geography and habitat reconstructions

We assessed geographic and habitat shifts through time in the genus *Xenopus* using ancestral state reconstructions on a reticulated phylogenetic network. The subgenus *Silurana* is currently distributed in lowland rainforest of the Congo Basin and coastal West-Central Africa, including on the continental island Bioko. Our ancestral state reconstructions of geography and habitat affiliations of this subgenus suggest that the ancestral habitat was lowland rainforest in coastal West-Central Africa (Figures 5.2 and 5.3). In the subgenus *Xenopus*, the inferred ancestral geographic range is either East Africa or the Ethiopian highlands, and the inferred ancestral habitat is montane grassland (Figures 5.2 and 5.3).

Within the subgenus *Xenopus*, the ancestral habitat of the *laevis* species group was reconstructed as lowland savannah, and the ancestral habitat for the *muelleri* species group as montane grassland (Figure 5.3). Although the habitat of these clades is generally similar among extant species, they have undergone extensive geographic distributional changes (Figure 5.2). The ancestor of the *muelleri* species group has a nearly equal likelihood of having originated in East Africa or the Ethiopian Highlands, with the remaining ancestral reconstructions in the clade predicting a shift from East Africa into North Africa. The ancestor of the *laevis* species group most likely originated in southern Africa, with subsequent diversification into East and Central Africa.

Our ancestral state reconstruction infers montane grassland habitat in the Albertine Rift as the ancestral state for the octoploid species *X. vestitus* and *X. lenduensis*, as well as for their two unknown tetraploid ancestral species, Ancestors A and B (Figures 5.2 and 5.3). Montane grassland was also inferred for the closely related sister species pair *X. itombwensis* and *X. wittei*, which share one tetraploid ancestor (Ancestor B) with *X. vestitus* and *X. lenduensis* (Figure 5.2), but with a shift or partial shift into cloud forest habitat (Figure 5.3).

Ancestors of a portion of the *amieti* species group (including *X. pygmaeus*, *X. allofraseri*, and *X. parafraseri;* Figure 5.2) are inferred to have returned to the lowland Congo Basin and coastal Central African rainforest. This lowland rainforest clade then hybridized multiple times with the Albertine Rift montane grassland clade. These hybridization events created a number of octoploid and dodecaploid species that are found in low elevation and high elevation forest habitats throughout the Congo Basin, coastal West-Central Africa, the Cameroon Volcanic Line, and the Albertine Rift (Figures 5.2 and 5.3). The only exceptions are *X. amieti*, which is found in

high elevation grasslands, and *X. eysoole* which can be found in either high elevation grassland or forest (Figure 5.3).

5.4.2 Ecological niche evolution

Compared to the available variation, African clawed frogs occupy a wide range of mean annual temperatures (14ºC to 28ºC used out of -3ºC to 31.6ºC that is available in Sub-Saharan Africa), annual precipitation (100 mm to 4100 mm used out of 8 mm to 4278 mm available), and precipitation seasonality (0 cv to 180 cv out of 0 cv to 180 cv available), but a low range of temperature seasonality (0.02 stdev to 5.95 stdev used out of 0 to 64.6 available; Fick & Hijmans, 2017). We will describe the niche evolution patterns for each climatic variable below.

Frequency of niche evolution. **–** We assessed the frequency of niche shift, expansion, contraction, and novelty in the genus *Xenopus* in the context of a reticulating phylogenetic network. In general, niche expansion and contraction were the most common patterns in our analyses, niche shifts were uncommon, and niche novelties were rare. For mean annual temperature, niche evolution categories were observed at significantly different frequencies, with niche contraction observed 12 times, niche expansion six times, niche shift two times, and niche novelty two times (Figure 5.4; $X^2 = 12.182$, $p = 0.0068$). Both instances of niche novelty have the caveat that they were observed in a comparison between either a dodecaploid or octoploid species and one of their parental species, but in both cases, there was still niche overlap with the other parental species. Niche shifts were inferred at the origin of subgenus *Xenopus* and for *X. laevis*. All niche evolution categories, except niche novelty, occurred at the same frequency between species formed by polyploidization and those formed by bifurcating evolution (niche shift: $z = -1.087$, $p = 0.28$; niche expansions: $z = 0.424$, $p = 0.67$; niche contractions: $z = -0.472$, p

= 0.64). Niche novelty occurred more often in species formed by polyploidization events (z = 1.94, p = 0.05). Overall, niche evolution for mean annual temperature occurred at equal rates between species formed by polyploidization and those formed by bifurcation (z = 0.264, p = 0.79).

For temperature seasonality, niche evolution categories were observed at significantly different frequencies, with niche contraction inferred nine times, niche expansion 16 times, niche shift two times, and niche novelty once (Appendix IV Figure S2; X^2 = 20.857, p = 0.00011). All shifts and novel niches were observed in the savanna-associated, broadly geographically distributed *laevis* and *muelleri* species groups, where the degrees of niche expansions and contractions were also the largest within African clawed frogs. All forms of niche evolution occurred statistically equally in species formed by polyploidization and those formed by bifurcating evolution (Niche shifts: z = -1.087, p = 0.28; niche expansions: z = -0.505, p is 0.62; niche contractions: z = -0.975, p = 0.33; niche novelty: z = -0.758, p = 0 .45). Overall, niche evolution for temperature seasonality occurred more often in species created through bifurcating evolution than those through polyploidization (z = -2.263, p = 0.02).

For annual precipitation, niche evolution categories were observed at significantly different frequencies, with niche contraction inferred 12 times, niche expansion 11 times, niche shift four times, and niche novelty once (Figure 5.5; X^2 = 12.286, p = 0.0065). One instance of niche shift and one instance or niche novelty were observed for the octoploid and dodecaploid hybrids, one shift was observed at origin of subgenus *Xenopus*, one at the origin of the *X. pygmaeus* clade, and one for *X. laevis*. All categories of niche evolution occurred at similar frequencies between species formed by polyploidization and those formed by bifurcating evolution (niche shifts: z = -0.480, p = 0.63; niche expansions: z = -0.704, p = 0.48; niche

contractions: z = -0.222, p = 0.83; niche novelty: z = 1.354, p = 0.18). Overall, niche evolution for annual precipitation occurred at roughly equivalent rates between species formed by polyploidization and those formed by bifurcation (z = -0.780, p = 0.44).

For precipitation seasonality, niche evolution categories were observed at significantly different frequencies, with niche contraction was seen ten times, niche expansion ten times, niche shift once, and no instances of niche novelty (Appendix IV Figure S3; X^2 = 17.286. p = 0.00062). The single instance of niche shift occurred in a dodecaploid species. All categories of niche evolution occurred at statistically equivalent rates between species formed by polyploidization and those formed by bifurcating evolution (niche shifts: z = 1.354, p = 0.18; niche expansion: z = -1.215, p = 0.22; niche contractions: z = -0.451, p = 0.65). Overall, niche evolution for this variable occurred at equal rates between species formed by polyploidization and those formed by bifurcation (z = -1.030, p = 0.30).

Across all variables, evolution occurred at similar rates between species formed by polyploidization and those formed by bifurcation (z = -1.836, p = 0.07), with approximately 55% of parental species comparisons associated with polyploidization (two per polyploid species) exhibiting some form of niche evolution (mean annual temperature, temperature seasonality, annual precipitation, and/or precipitation seasonality), while around 70% of comparisons did so for bifurcating evolution. However, when compared to both parental species simultaneously, 100% of polyploid species exhibited niche divergence for at least two climate variables from one or both parents.

Species group comparisons. –Within the subgenus *Silurana*, there was evidence for both niche stasis and niche shifts following allopolyploidization (Figure 5.4 and 5.5; Appendix IV Figures S2 and S3). The subgenus *Xenopus* exhibited a general trend towards narrowing of niche

breadth and colonization of novel niches as ploidy levels increased (Figure 5.4 and 5.5; Appendix IV Figures S2 and S3). A notable exception was the tetraploid *laevis* species group which exhibits remarkable variation in niche breadths among its very closely related species, ranging from one of the narrowest niches (*X. gilli*) to one of the most general (*X. laevis*; Figure 5.4 and 5.5; Appendix IV Figures S2 and S3). In the *amieti* species group, octoploids and dodecaploids showed very similar patterns by utilizing a variety of niches, but at generally narrow realized niche breadths (Figure 5.4 and 5.5; Appendix IV Figures S2 and S3). However, many of these higher ploidy species occupy restricted ranges, leading to high levels of uncertainty in their fundamental niche characterization. Additional discussion of niche evolution and biogeography of various species groups is provided in the Supplemental Results.

5.5 Discussion

In plants, polyploidy may be associated with rapid niche evolution and catalyzation of adaptation to novel climates (Brittingham et al., 2018; Donkpegan et al., 2017; Marchant et al., 2016; Nakagawa, 2006; Ramsey, 2011; Southgate et al., 2019; Theodoridis et al., 2013; Thompson et al., 2014; Visger et al., 2016) or unique/specialized conditions (Alix et al., 2017; Ekimova et al., 2012; Ramsey, 2011; Selmecki et al., 2015; Van de Peer et al., 2017). At the extreme, polyploidization may be a prelude to major adaptive radiations by opening new ecological opportunities and triggering shifts into different adaptive zones (Alix et al., 2017; de Bodt et al., 2005; Fawcett et al., 2009; Simpson, 1984; Van de Peer et al., 2017). Here, we used a combination of recently developed niche evolution analyses (Owens et al., 2020) and phylogenetic network ancestral state reconstructions (Bastide et al., 2018; Solís-Lemus et al., 2017) to explore this possibility in African clawed frogs (genus *Xenopus*), one of the few animal

groups with multiple polyploidization events that occurred independently and sequentially in closely related species.

By reconstructing niche evolution in the context of a highly reticulate phylogeny, the integration of these methods allowed us to account for estimates of the full spectrum of each species' fundamental niche, incorporate uncertainty where niche limits were unknown, and account for the unique nature of phylogenetic non-independence that is associated with the evolutionary history of allopolyploid species in this group. Our present-day niche characterizations coincide well with what is known about the physiology of the two *Xenopus* species that are used as model organisms (*X. laevis* and *X. tropicalis*; Khokha et al., 2002; Nagasawa et al., 2013). Our ancestral niche reconstructions also coincided with geographic and habitat reconstructions estimated from empirical tip values for the genus. Additionally, projected type and degree of niche evolution events were in line with expectations developed from studies on polyploid plants: expansions and contractions around the ancestral niche were the most common patterns, whereas niche shifts were rare, and niche novelty was rarer still.

Few other studies have explored the role of polyploidization in niche evolution in amphibian species. Otto et al. (2007) found significant shifts to colder, drier climates and higher elevations associated with allopolyploidization in the tetraploid North American gray tree frog *Hyla versicolor*. Tetraploid species in the genus *Bufotes* (Palearctic green toads) also show a significant shift to colder, drier climates and higher elevations (Ficetola & Stöck, 2016; Litvinchuk et al., 2011), and various polyploid frogs from South America have distributions influenced by temperature seasonality (David & Halanych, 2021). These studies strongly suggest that, like plants, allopolyploidization in amphibians can play a key role in adaptation to new environments.

5.5.1 Geography, habitat, and niche evolution

In this study, we aimed to phylogenetically test the hypothesis that higher levels of polyploidy are associated with shifts in habitat or movement into novel or more extreme environments (Adams & Wendel, 2005; Ainouche & Wendel, 2014; Arnold, 2004; Levy & Feldman, 2002; Sigel, 2016; Soltis & Soltis, 2000). The genus *Xenopus* has been well studied in terms of patterns of genome duplication (Chain et al., 2008, 2011; Evans, 2008; Evans & Kwon, 2015; Furman et al., 2015, 2018; Session et al., 2016), and its reticulate phylogeny has been carefully reconstructed (De Sa & Hillis, 1990; Evans, 2008; Evans et al., 2004, 2008, 2011a,b, 2015, 2019; Furman et al., 2015). The vast amount of work already done on this genus paves the way for the study of polyploid niche evolution in a new phylogenetic comparative context.

Geographic and habitat reconstructions for the genus *Xenopus* indicate multiple transitions between savanna and rainforest habitats and between lowland and high-elevation distributions (Figures 5.2 and 5.3). A major geographic and habitat divergence occurred with the split between the subgenera *Silurana*, a West-Central African rainforest group, and *Xenopus* whose ancestor was inferred to be an East African savanna-dwelling species (Figures 5.2 and 5.3). *Xenopus* (*Xenopus*) diversified and spread through the savannas of southern, eastern, and northern Africa before transitioning back into Central African rainforest with the evolution of the clade containing *X. pygmaeus*, *X. allofraseri*, and *X. parafraseri* (Figures 5.2 and 5.3). The ancestors of the high-elevation grassland Albertine Rift *X. lenduensis* - *X. vestitus* clade then hybridized with the ancestors of the lowland rainforest species *X. pygmaeus*, *X. allofraseri*, and *X. parafraseri* to produce multiple, independently evolved allopolyploid species (Evans et al.,

2015). These allopolyploid species spread throughout Central Africa to the Cameroon Volcanic Line, Albertine Rift, Congo Basin, and coastal Central Africa (Figures 5.2, 5.3, and 5.6).

These habitat and geographic transitions were reflected in our ancestral niche reconstructions. For instance, niche shifts are predicted to have occurred towards far lower mean annual temperatures and annual precipitation with the divergence between *Xenopus* (*Silurana*) and *Xenopus* (*Xenopus*) (Figures 5.4 and 5.5). Within *Xenopus* (*Xenopus*), the *muelleri* species group was inferred to have undergone almost stepwise shifts towards higher temperature tolerance, and to a lesser extent precipitation seasonality, as species shifted farther north into more arid habitats (Figure 5.4; Appendix IV Figure S3). The *laevis* species group exhibited a shift to higher precipitation seasonality and lower annual precipitation with its movement into southern Africa, with a subsequent return to low seasonality and higher precipitation as species evolved northward into the East African highlands and the Congo Basin (Figures 5.4 and 5.5). With the evolution of the *X. pygmaeus*, *X. allofraseri* and *X. parafraseri* clade in Central African lowland rainforests came a shift to higher precipitation and lower seasonality (Figure 5.5; Appendix IV Figure S3).

Variable niche evolution patterns were observed in the formation of the octoploid and dodecaploid species, with a general pattern of a highly pronounced niche and habitat shift away from one parental species while retaining similarity to the other parental species (Figure 5.6). However, no general trend was shown favoring one habitat type or parental classification (maternal vs. paternal) over the other. For example, the octoploid species clade including *X. boumbaensis*, *X. andrei*, and *X. amieti* shows a major shift in mean annual temperature, annual precipitation, and precipitation seasonality away from the high elevation Albertine Rift parental species, but only very minor niche contractions or no change in these variables when compared

to their lowland rainforest Central African parental species (Figures 5.4, 5.5, and 5.6; Appendix IV Figure S3). A second octoploid species group that includes *X. itombwensis* and *X. wittei* shows major shifts away from the lowland Central African parental niche, with little niche differentiation from their high-elevation Albertine Rift parental species.

As small, ectothermic, obligatorily aquatic animals, *Xenopus* would be expected to have relatively low vagility (Hillman et al., 2014; Measey, 2016). Animals with such low vagility often exhibit very few evolutionary transitions between habitats or biomes, especially at lower taxonomic scales (Anderson et al., 2002; Crisp et al., 2009). While this expectation is in direct contrast to our findings for *Xenopus*, these transitions have been recorded relatively frequently in smaller, less complex polyploid systems (Brittingham et al., 2018; Donkpegan et al., 2017; Marchant et al., 2016; Nakagawa, 2006; Southgate et al., 2019; Theodoridis et al., 2013; Visger et al., 2016), supporting the hypothesis that genome duplication coupled with competitive pressure from sympatric parental species could be driving these transitions (Fowler & Levin, 1984; Levin, 1975, 2004).

5.5.2 Timing of diversification

Evans et al. (2015) conducted a phylogenetic estimation of divergence dates on the multi-labeled gene tree that was used to infer our phylogenetic network. While it is difficult to precisely estimate the timing of allopolyploid species origins (Doyle & Egan, 2010; Kellogg, 2016), an upper limit can be determined using these dated multi-labeled trees (Marcussen et al., 2015). Evans et al. (2015) estimated the divergence between the subgenera *Silurana* and *Xenopus* to be between 35 and 40 mya, corresponding with a dramatic global cooling period triggering the breakup of a potential pan-African rainforest and spread of African savannas

(Couvreur et al., 2021). This period has been associated with the divergence of other West-Central and East African taxa (Couvreur et al., 2008; Tolley et al., 2013), and in our study corresponds with the evolution of the savanna adapted East African *Xenopus* (*Xenopus*) clade.

The tetraploid *muelleri* group diverged in the early Miocene and diversified throughout the mid-late Miocene. Our ancestral state reconstruction suggests that this primarily savanna clade arose somewhere in the East African or Ethiopian Highlands and subsequently diversified across western and eastern Africa. This period corresponds with the Mid-Miocene Climatic Optimum, and the uplift of the East African highlands and the Albertine Rift (Couvreur et al., 2021; Guillocheau et al., 2018; Simon et al., 2017). Within the *X. muelleri* group, *Xenopus fischbergi* and *X. fraseri* diverged approximately 5 million years ago (Evans et al., 2019), and are the only *Xenopus* species distributed in the Sahel (the savanna belt north of the Congo Basin). Their divergence may be associated with aridification and spread of grasslands across Africa during the Miocene aridification (Couvreur et al., 2021; Feakins et al., 2013; Herbert et al., 2016; Ségalen et al., 2007).

The allotetraploid *laevis* species group diverged in the late-Miocene. Diversification in this clade is likely associated with the expansion of savanna habitat during the Miocene aridification (Couvreur et al., 2021; Feakins et al., 2013; Herbert et al., 2016; Ségalen et al., 2007), and may also be associated with the presence of Pleistocene savanna refugia (Furman et al., 2015; Lorenzen et al., 2012), and with the uplift of geological barriers in South Africa such as the Great Escarpment (Furman et al., 2015).

This dating analysis places the timing of the formation of the Central African octoploids and dodecaploids within the last 2 to 10 million years, corresponding with the Miocene aridification (Pokorny et al., 2015; Senut et al., 2009; Sepulchre et al., 2006) and the Pleistocene

glaciation cycles (Maley, 1996). These climatic changes may have caused elevational shifts into the lowlands for the Albertine Rift highland ancestors of the *X. lenduensis* - *X. vestitus* clade and contact with the ancestors of the Central African lowland *X. pygmaeus*, *X. allofraseri,* and *X. parafraseri* clade, where hybridization and allopolyploidization occurred, creating many of the octoploids and dodecaploids in the genus. These polyploidization events were then followed by diversification into new geographic areas (i.e. the Cameroon Volcanic Line, coastal West-Central Africa, the Congo Basin, and the Albertine Rift), and the rapid evolution of shifted and novel niches.

5.5.3 Niche evolution patterns

Marchant et al. (2016) observed that, although newly formed allopolyploid plant species may avoid competition by inhabiting different ecological niches than either or both parental species (e.g., Fowler & Levin, 1984; Levin, 1975, 2004), these species often have partial niche overlap with at least one parental species. This pattern was also evident in *Xenopus*. We observed niche overlap among all octoploids and dodecaploids and only one of their parental species, but not overlapping or barely overlapping with the other parental species. All polyploid species showed niche divergence for at least two climate variables from one or both parents, and nearly all incidences of novel niche utilization were associated with allopolyploid species formation. In no instance was the climatic niche of an octoploid or dodecaploid completely novel from both parental species; however, this could change with the analysis of a greater number of climatic or abiotic environmental variables.

In addition, and also in congruence with Marchant et al. (2016), we observed that patterns of niche expansion, contraction, shift, and novelty occurred at different frequencies:

expansions and contractions were common (23.0% and 43.6% of niche evolution events respectively), niche shifts were rare (8.9% of niche evolution events), and niche novelty was rarer still (4.0% of niche evolution events, all but one of which was associated with allopolyploid species formation; Marchant et al., 2016). These patterns also occurred at different frequencies within each of our climate variables. For example, niche contraction was twice as prevalent as niche expansion in our mean annual temperature analysis (12 contractions vs. 6 expansions). The opposite was true for temperature seasonality (9 contractions vs. 16 expansions). However, for the precipitation variables, expansions and contractions occurred at nearly equal rates (annual precipitation 12 contractions vs 11 expansions; precipitation seasonality 10 contractions vs 10 expansions). These patterns of constrained evolution for mean annual temperature may be a result of physiological constraints on many animals' thermal tolerances (Araújo et al., 2013; Feder & Burggren, 1992; Wells, 2010).

Niche shift and niche novelty were observed at very low frequencies for all climate variables (mean annual temperature: 2 shifts and 2 novel; temperature seasonality: 2 shifts and 1 novel; annual precipitation: 4 shifts and 1 novel; precipitation seasonality: 1 shift and 0 novel). Niche shift was most likely to occur at the transition between habitat types. For instance, it occurs at the origin of subgenus *Xenopus* with the transition from rainforest to savanna, at the origin of the *laevis* species group with the movement to the colder more seasonal southern Africa, and at the origin of the *X. pygmaeus* clade with the transition back from savanna to rainforest. Niche shifts also occurred in association with octoploid or dodecaploid formation. Niche novelty was most likely to occur in association with octoploid or dodecaploid formation, with the exception of *X. muelleri* which shows a completely novel niche for temperature seasonality compared to its ancestral species. These results are suggestive that higher levels of

polyploidy are associated with increased levels of niche evolution and movement into novel environments (Adams & Wendel, 2005; Ainouche & Wendel, 2014; Arnold, 2004; Levy & Feldman, 2002; Sigel, 2016; Soltis & Soltis, 2000). However, these transitions were not common, and the octoploid and dodecaploid species exhibiting novel niches still maintained niche overlap with one parental species.

Although we found higher instances of niche novelty in species formed directly by allopolyploidization, high overall levels of niche evolution were also seen in tetraploid species that were formed by bifurcating speciation of tetraploid ancestors. The relatively older ages of these tetraploid speciation events may account for the high levels of niche evolution observed in these lineages (Evans et al., 2015). In addition, their duplicated (tetraploid) genomes may already allow for higher evolutionary and niche plasticity, even without the additional driver of competition with one or more sympatric parental species (Gaynor et al., 2018; Molina-Henao & Hopkins, 2019; Visger et al., 2016).

5.6 Conclusion

With its high levels of polyploidy, from 2n to 12n, and its wide range of habitat preferences, from rainforest to savanna, the genus *Xenopus* represents a unique system for the study of the effects of allopolyploidization on niche evolution. Our study suggests that allopolyploidy may be associated with high rates of niche evolution, raising the question of why this phenomenon is observed only rarely in vertebrate systems. Hypotheses for the relatively higher instance of polyploidization in plant than animal systems include disruptions in chromosomal sex determination (Gregory & Mable, 2005; Muller, 1925; Wertheim et al., 2013), imbalanced dosage compensation (Gregory & Mable, 2005; Orr, 1990; Wertheim et al., 2013),

and multivalent meiotic pairing (Gregory & Mable, 2005; Macgregor, 1993). However, species in the genus *Xenopus* have shown an unusual capacity to tolerate broad scale genomic changes and variation in gene dosage that arise after whole genome duplication (Evans, 2008; Evans & Kwon, 2015; Furman & Evans, 2016; Furman et al., 2018), which is seen to a lesser extent in other amphibians (Evans et al., 2012; Mable et al., 2011; Schmid et al., 2015), but never in mammals (Evans et al., 2017; Gregory & Mable, 2005; Mable et al., 2011). To the extent that polyploidization can catalyze adaptation, it may be that intolerance of variation in gene dosage prevents many animal groups from exploiting the adaptive potential of this genomic transition.

5.7 Figures

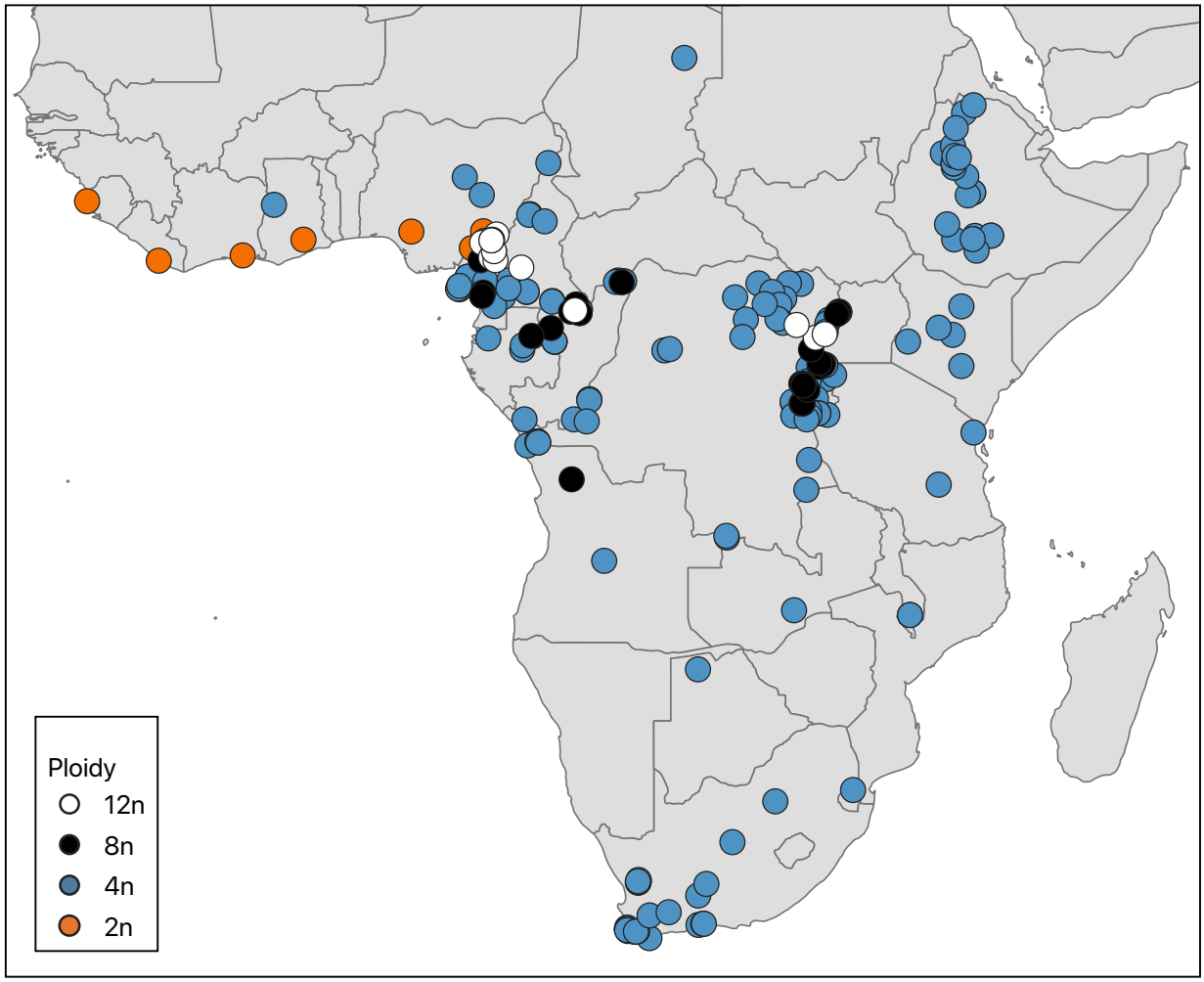


Figure 5.1 Geographic distribution of ploidy level in the genus *Xenopus* based on genotyped, expert-identified specimens used in this study.

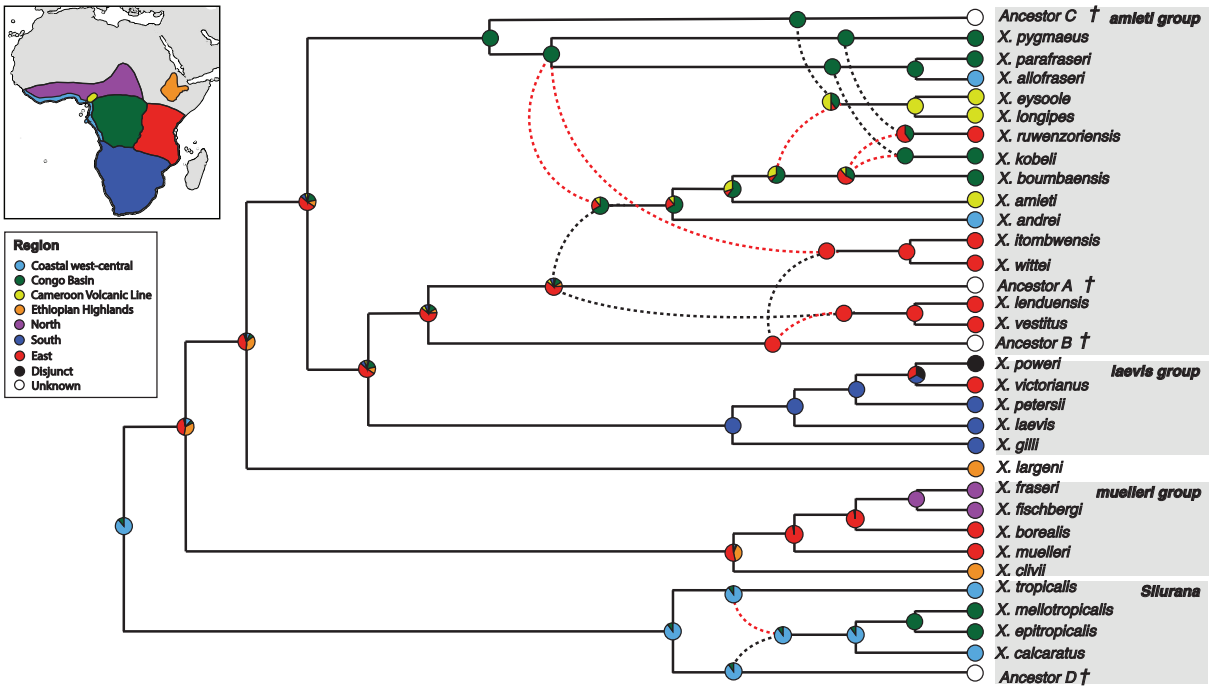


Figure 5.2 Geographic range reconstruction on a reticulate phylogenetic network for each species in the genus *Xenopus*. Species groups discussed in the text are highlighted in grey. Extinct or undiscovered ancestral species are denoted by †. Dotted lines indicate ancestry with red indicating maternal and black indicating paternal.

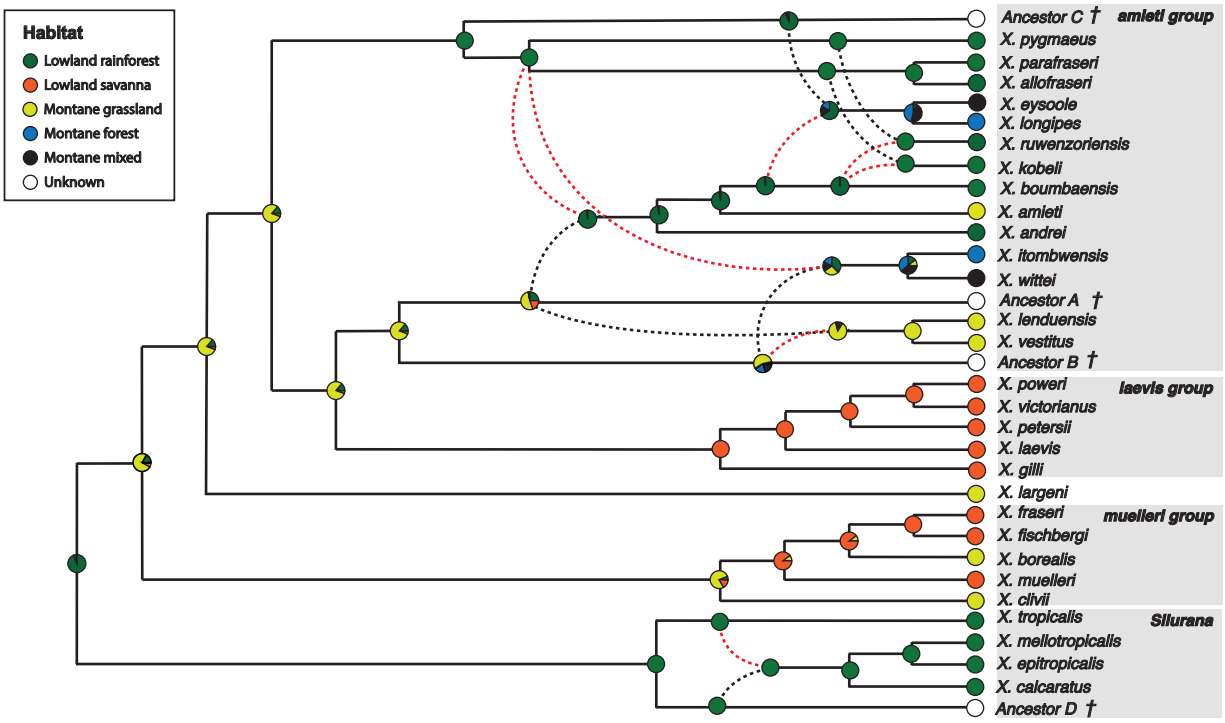


Figure 5.3 Habitat reconstruction on a reticulate phylogenetic network for each species in the genus *Xenopus*. Species groups discussed in the text are highlighted in grey. Extinct or undiscovered ancestral species are denoted by †. Dotted lines indicate ancestry with red indicating maternal and black indicating paternal.

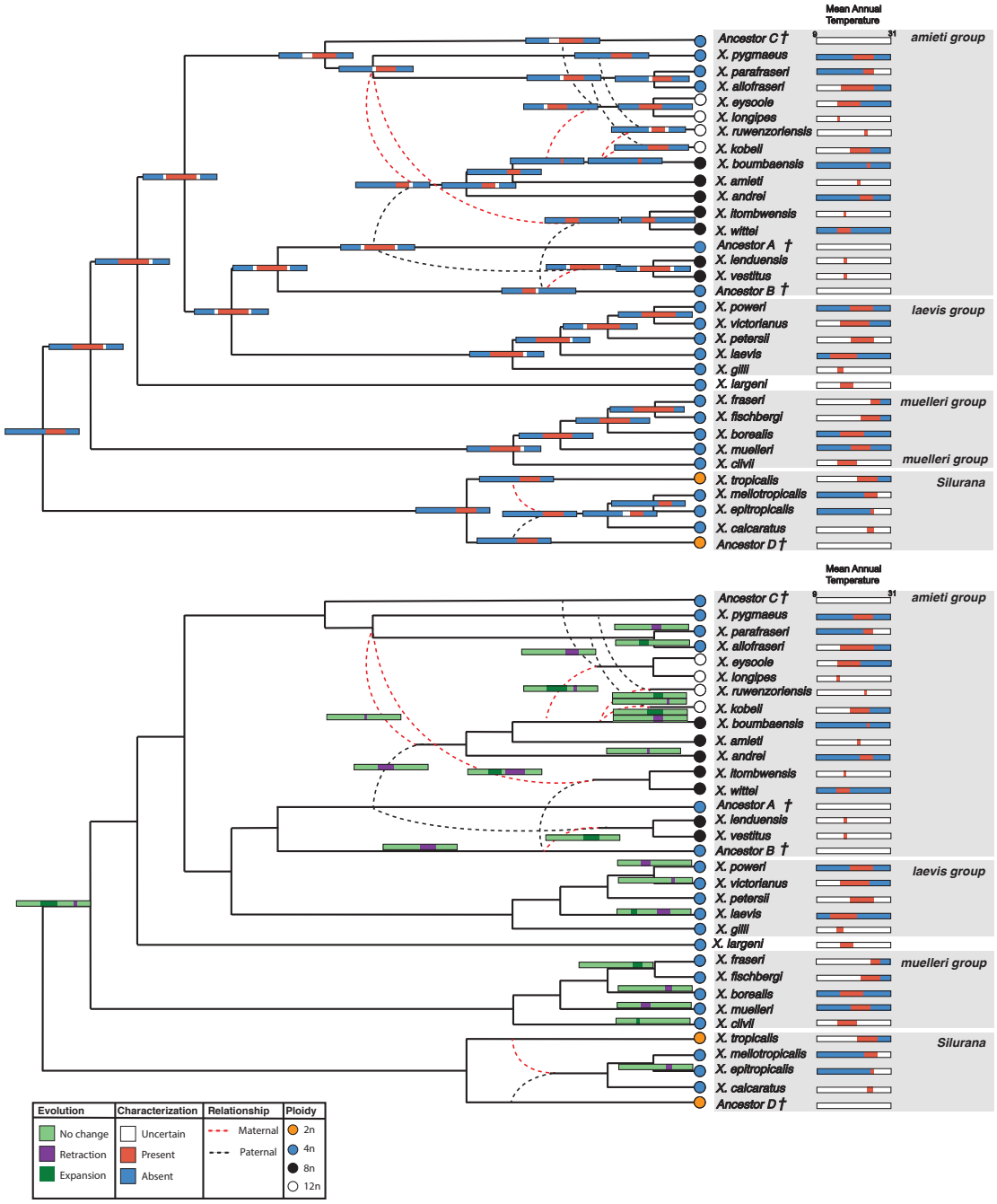


Figure 5.4 Top: Mean annual temperature reconstruction plotted on the nodes and Bottom: mean annual temperature evolution inferred from that reconstruction plotted on the branches of a reticulate phylogenetic network for each species in the genus *Xenopus*. Branches for which no niche evolution was inferred were left empty to increase readability. Species groups discussed in the text are highlighted in grey. Extinct or undiscovered ancestral species are denoted by †.

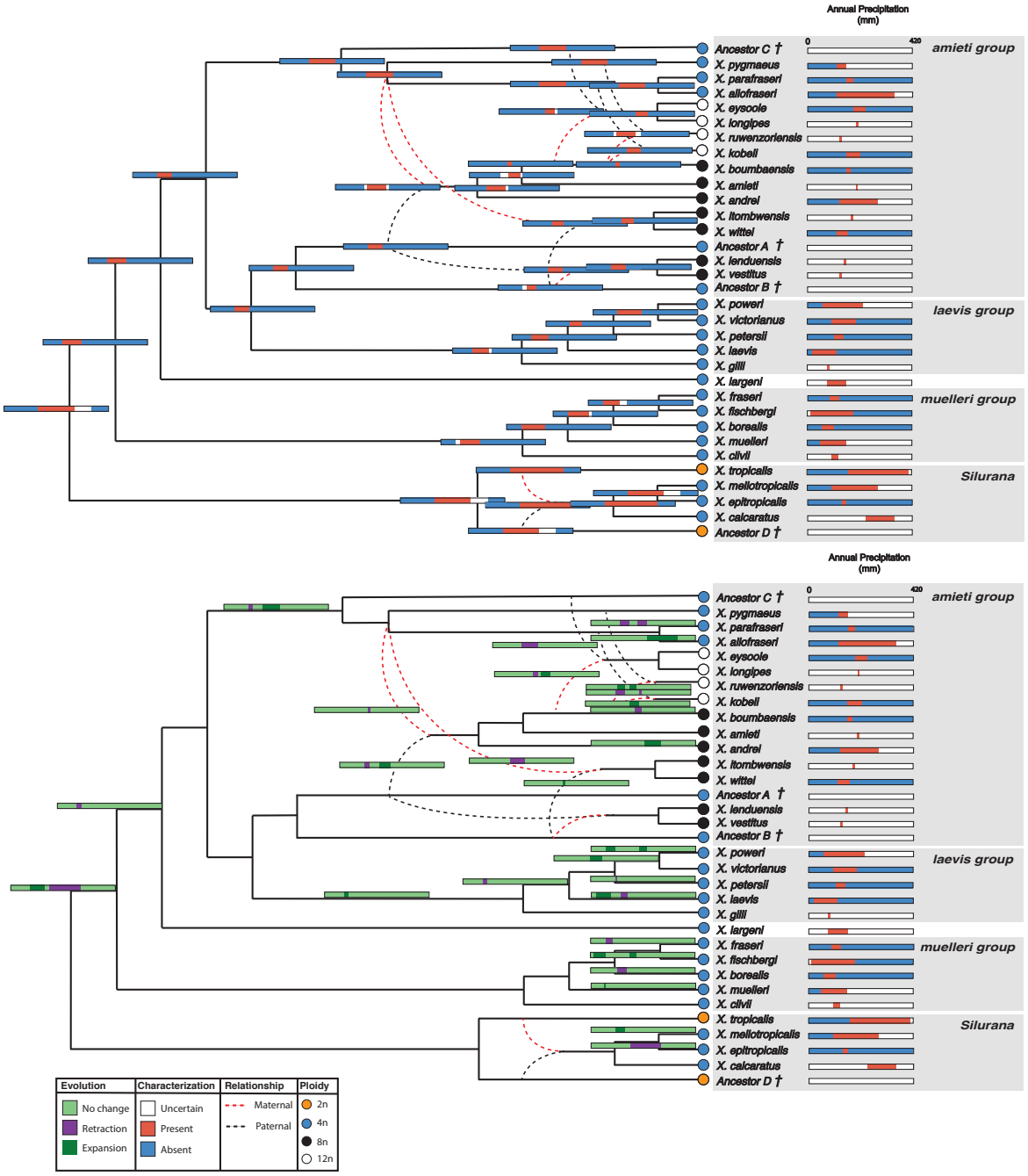


Figure 5.5 Top: Annual precipitation reconstruction plotted on the nodes and Bottom: annual precipitation reconstruction evolution inferred from that reconstruction plotted on the branches of a reticulate phylogenetic network for each species in the genus *Xenopus*. Branches for which no niche evolution was inferred were left empty to increase readability. Species groups discussed in the text are highlighted in grey. Extinct or undiscovered ancestral species are denoted by †.

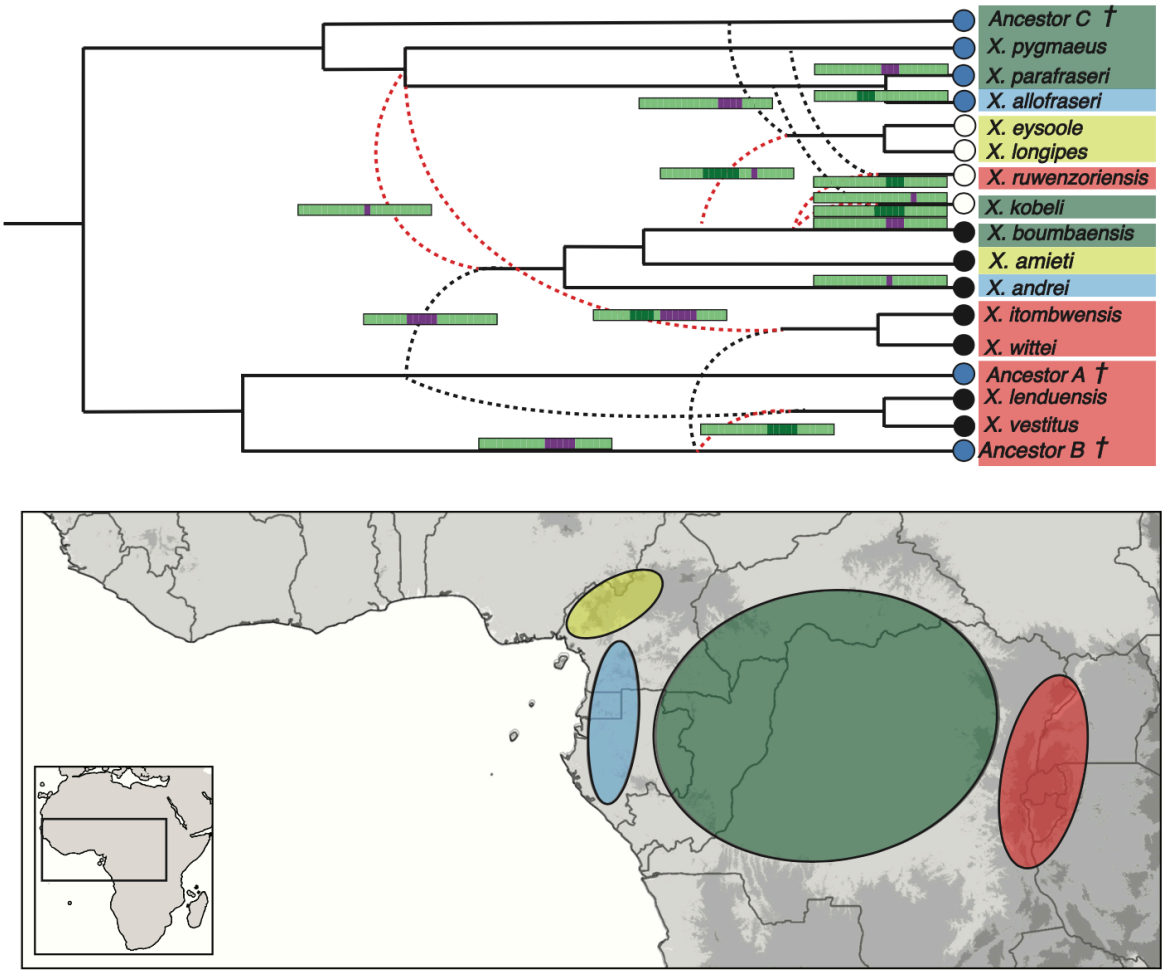


Figure 5.6 Geography and niche evolution for mean annual temperature of the *amieti* species group in *Xenopus*. Circles on taxa tips correspond to ploidy with blue = 4n, black = 8n, and white =12n. Niche evolution colors correspond to those in figures 5.4 and 5.5. Geographic distribution colors correspond roughly to those in Figure 5.2.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J., Bierne, N., ... & Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, 26, 229–246.

Achaval, F., González, J. G., Meneghel, M., & Megarejo, A. R. (1979). Lista comentada del material recogido en costas uruguayas, transportado por camalotes desde el Río Paraná. *Acta Zoológica Lilloana*, 195–200.

Ackerly, D. D., Schwilk, D. W., & Webb, C. O. (2006). Niche evolution and adaptive radiation: Testing the order of trait divergence. *Ecology*, *87*, S50–S61.

Adams, K. L., & Wendel, J. F. (2005). Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology*, *8*, 135–141.

Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography, 38*, 541–545.

Ainouche, M.L. & Wendel, J. F. (2014). Polyploid speciation and genome evolution: Lessons from recent allopolyploids. In: Pontarotti, P. (ed.) *Evolutionary Biology: Genome Evolution, Speciation, Coevolution, and Origin of Life*. New York, NY: Springer, pp. 87–113.

Aka, F. T., Kusakabe, M., Nagao, K., & Tanyileke, G. (2001). Noble gas isotopic compositions and water/gas chemistry of soda springs from the islands of Bioko, São Tomé and Annobón, along with Cameroon Volcanic Line, West Africa. *Applied Geochemistry*, *16*, 323–338.

Allen, K. E., Greenbaum, E., Hime, P.M., Tapondjou N., W.P., Sterkhova, V.V., Kusamba, C., … & Brown, R.M. (2021). Rivers, not refugia, drove diversification in arboreal, sub-Saharan African snakes. *Ecology and Evolution*, *11*, 6133–6152.

Allen, K. E., N. Tapondjou, W. P., Greenbaum, E., Welton, L. J., & Bauer, A. M. (2019). High levels of hidden phylogenetic structure within Central and West African *Trachylepis* skinks. *Salamandra*, *55*, 231–234.

Allen, J. R., Forrest, M., Hickler, T., Singarayer, J. S., Valdes, P. J., & Huntley, B. (2020). Global vegetation patterns of the past 140,000 years. *Journal of Biogeography*, *47*, 2073–2090.

Alix, K., Gérard, P. R., Schwarzacher, T., & Heslop-Harrison, J. S. (2017). Polyploidy and interspecific hybridization: Partners for adaptation, speciation and evolution in plants. *Annals of Botany*, *120*, 183–194.

Almendra, A. L., González-Cózatl, F. X., Engstrom, M. D., & Rogers, D. S. (2018). Evolutionary relationships and climatic niche evolution in the genus *Handleyomys* (Sigmodontinae: Oryzomyini). *Molecular Phylogenetics and Evolution*, *128*, 12–25.

Amorim, D. D. S. (1991). Refuge model simulations: Testing the theory. *Revista Brasileira de Entomologia*, *35*, 803–812.

AmphibiaWeb. 2021. <https://amphibiaweb.org> University of California, Berkeley, CA, USA. Accessed 17 Aug 2021.

Anderson, R. P., Lew, D., & Peterson, A. T. (2003). Evaluating predictive models of species' distributions: Criteria for selecting optimal models. *Ecological Modelling*, *162*, 211–232.

Anderson, R. P., Peterson, A. T., & Gómez-Laverde, M. (2002). Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos*, *98*, 3–16.

Anhuf, D., (2000). Vegetation history and climate changes in Africa north and south of the equator (10 S to 10 N) during the Last Glacial Maximum. In Smolka, P.P., Volkheimer, W. (Eds.) *Southern Hemisphere Paleo-and Neoclimates*. Berlin, Heidelberg.

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

Antonelli, A., Kissling, W. D., Flantua, S. G., Bermúdez, M. A., Mulch, A., Muellner-Riehl, A. N., ... & Hoorn, C. (2018). Geological and climatic influences on mountain biodiversity. *Nature Geoscience*, *11*, 718–725.

Araújo, M. B., Ferri-Yáñez, F., Bozinovic, F., Marquet, P. A., Valladares, F., & Chown, S. L. (2013). Heat freezes niche evolution. *Ecology Letters*, *16*, 1206–1219.

Arjona, Y., Fernández-López, J., Navascués, M., Alvarez, N., Nogales, M., & Vargas, P. (2020). Linking seascape with landscape genetics: Oceanic currents favour colonization across the Galápagos Islands by a coastal plant. *Journal of Biogeography*, *47*, 2622–2633.

Arjona, Y., Nogales, M., Heleno, R., & Vargas, P. (2018). Long-distance dispersal syndromes matter: diaspore–trait effect on shaping plant distribution across the Canary Islands. *Ecography*, 41, 805–814.

Arnold, M. L. (2004). Transfer and origin of adaptations through natural hybridization: Were Anderson and Stebbins right? *The Plant Cell*, *16*, 562–570.

Balinsky, J. B. (1981). Adaptation of nitrogen metabolism to hyperosmotic environment in Amphibia. *Journal of Experimental Zoology*, *215*, 335–350.

Baniaga, A. E., Marx, H. E., Arrigo, N., & Barker, M. S. (2020). Polyploid plants have faster rates of multivariate niche differentiation than their diploid relatives. *Ecology Letters*, *23*, 68–78.

Barej, M. F., Schmitz, A., Menegon, M., Hillers, A., Hinkel, H., Boehme, W., & Rödel, M.-O. (2011). Dusted off–the African *Amietophrynus superciliaris* species complex of giant toads. *Zootaxa*, *2772*, 1–32.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., Soberón, J & Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, *222*, 1810–1819.

Bates, H.W. (1863). *The Naturalist on the River Amazon.* London, UK: John Murray.

Bastide, P., Solís-Lemus, C., Kriebel, R., William Sparks, K., & Ané, C. (2018). Phylogenetic comparative methods on phylogenetic networks with reticulations. *Systematic Biology*, *67*, 800–820.

Bauer, A. M. (2003). On the identity of *Lacerta punctata* Linnaeus 1758, the type species of the genus *Euprepis* Wagler 1830, and the generic assignment of Afro-Malagasy skinks. *African Journal of Herpetology*, *52*, 1–7.

Bell, R. C., Drewes, R. C., Channing, A., Gvoždík, V., Kielgast, J., Lötters, S., ... & Zamudio, K. R. (2015a). Overseas dispersal of *Hyperolius* reed frogs from Central Africa to the oceanic islands of São Tomé and Príncipe. *Journal of Biogeography*, *42*, 65–75.

Bell, R. C., Drewes, R. C., & Zamudio, K. R. (2015b). Reed frog diversification in the Gulf of Guinea: Overseas dispersal, the progression rule, and in situ speciation. *Evolution*, *69*, 904–915.

Bell, R. C., Parra, J. L., Badjedjea, G., Barej, M. F., Blackburn, D. C., Burger, M., ... & Zamudio, K. R. (2017). Idiosyncratic responses to climate-driven forest fragmentation and marine incursions in reed frogs from Central Africa and the Gulf of Guinea Islands. *Molecular Ecology*, *26*, 5223–5244.

Blackburn, D. C. (2008). Biogeography and evolution of body size and life history of African frogs: Phylogeny of squeakers (*Arthroleptis*) and long-fingered frogs (*Cardioglossa*) estimated from mitochondrial data. *Molecular Phylogenetics and Evolution*, *49*, 806–826.

Blackburn, D. C., & Measey, G. J. (2009). Dispersal to or from an African biodiversity hotspot? *Molecular Ecology*, *18*, 1904–1915.

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.

Bonnefille, R. & Riollet, G. (1988). The Kashiru pollen sequence (Burundi): Paleoclimatic implications for the last 40,000 yr BP in tropical Africa. *Quaternary Research*, *30*, 19–35.

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., & Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, *17*, 1–11.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., … & Matschiner, M. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *15*, e1006650.

Branch, B. (2014). *Photographic guide to snakes, other reptiles and amphibians of East Africa*. Penguin Random House South Africa.

Brenac, P. (1988). Evolution de la végétation et du climat dans l'Ouest Cameroun entre 25.000 et 11.000 ans BP. Actes Xème Symposium Ass. Palynologues Langue Franwise. *Travaux de la Section des Sciences & Techniques de l'Institut Français de Pondichéry*, *25*, 91–103.

Brittingham, H. A., Koski, M. H., & Ashman, T. L. (2018). Higher ploidy is associated with reduced range breadth in the Potentilleae tribe. *American Journal of Botany*, *105*, 700–710.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N.E., Graham, C.H., Guisan, A. (2012). Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, *21*, 481–497.

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

Büchi, L., & Vuilleumier, S. (2014). Coexistence of specialist and generalist species is shaped by dispersal and environmental factors. *The American Naturalist*, *183*, 612–624.

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.

Burbrink, F. T., Lawson, R., & Slowinski, J. B. (2000). Mitochondrial DNA phylogeography of the polytypic North American Rat Snake (*Elaphe obsoleta*): A critique of the subspecies concept. *Evolution*, *54*, 2107–2118.

Burke, K. (2001). Origin of the Cameroon Line of volcano-capped swells. *The Journal of Geology*, *109*, 349–362.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and inference. A practical information-theoretic approach*. 2nd edition. Berlin, Heidelberg, New York: Springer.

Bush, M. B., & Rivera, R. (2001). Reproductive ecology and pollen representation among neotropical trees. *Global Ecology and Biogeography*, *10*, 359–367.

Bussmann, R. W. (2004). Regeneration and succession patterns in African, Andean and Pacific tropical mountain forests: The role of natural and anthropogenic disturbance. *Lyonia*, *6*, 93–111.

Caldeira, R. J. M. M., & Munhá, J. M. (2002). Petrology of ultramafic nodules from São Tomé island, Cameroon volcanic line (oceanic sector). *Journal of African Earth Sciences*, *34*, 231–246.

Cannatella, D. C., & de Sá, R. O. (1993). *Xenopus laevis* as a model organism. *Systematic Biology, 42,* 476–507.

Caratini, C., & Giresse, P. (1979). Contribution palynologique à la connaissance des environnements continentaux et marins du Congo à la fin du Quaternaire. *Comptes rendus de l'Académie des Sciences*, *288*, 379–382.

Carlquist, S. (1966). The biota of long-distance dispersal. I. Principles of dispersal and evolution. *The Quarterly Review of Biology*, *41*, 247–270.

Carranza, S., & Arnold, E. N. (2006). Systematics, biogeography, and evolution of *Hemidactylus* geckos (Reptilia: Gekkonidae) elucidated using mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, *38*, 531–545.

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

Ceríaco, L. M. (2015). Lost in the middle of the sea, found in the back of the shelf: A new giant species of *Trachylepis* (Squamata: Scincidae) from Tinhosa Grande islet, Gulf of Guinea. *Zootaxa*, *3973*, 511–527.

Ceríaco, L. M., Arellano, A. L., Jadin, R. C., Marques, M. P., Parrinha, D., & Hallermann, J. (2021). Taxonomic revision of the Jita snakes (Lamprophiidae: *Boaedon*) from São Tomé and Príncipe (Gulf of Guinea), with the description of a new species. *African Journal of Herpetology*, *70*, 1–31.

Ceríaco, L. M., Bernstein, J., Sousa, A. C., Marques, M. P., Bauer, A. M., & Norder, S. J. (2020). The reptiles of Tinhosa Grande islet (Gulf of Guinea): A taxonomic update and the role of Quaternary sea level fluctuations in their diversification. *African Journal of Herpetology*, *69*, 200–216.

Ceríaco, L. M., Marques, M. P., Jacquet, F., Nicolas, V., Colyn, M., Denys, C., ... & Bastos-Silveira, C. (2015). Description of a new endemic species of shrew (Mammalia, Soricomorpha) from Príncipe Island (Gulf of Guinea). *Mammalia*, *79*, 325–341.

Ceríaco, L. M., Marques, M. P., & Bauer, A. M. (2016). A review of the genus *Trachylepis* (Sauria: Scincidae) from the Gulf of Guinea, with descriptions of two new species in the *Trachylepis maculilabris* (Gray, 1845) species complex. *Zootaxa*, *4109*, 284–314.

Chain, F. J., Ilieva, D., & Evans, B. J. (2008). Duplicate gene evolution and expression in the wake of vertebrate allopolyploidization. *BMC Evolutionary Biology*, *8*, 1–16.

Chain, F. J., Dushoff, J., & Evans, B. J. (2011). The odds of duplicate gene persistence after polyploidization. *BMC Genomics*, *12*, 1–7.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer data sets. *Gigascience*, *4*, s13742-015.

Channing, A., & Rödel, M. O. (2019). *Field guide to the frogs & other amphibians of Africa*. Penguin Random House South Africa.

Chardon, D., Grimaud, J. L., Rouby, D., Beauvais, A., & Christophoul, F. (2016). Stabilization of large drainage basins over geological time scales: Cenozoic West Africa, hot spot swell growth, and the Niger River. *Geochemistry, Geophysics, Geosystems*, *17*, 1164–1181.

Charles, K. L., Bell, R. C., Blackburn, D. C., Burger, M., Fujita, M. K., Gvoždík, V., … Portik, D. M. (2018). Sky, sea, and forest islands: Diversification in the African leaf-folding frog *Afrixalus paradorsalis* (Anura: Hyperoliidae) of the Lower Guineo-Congolian rainforest. *Journal of Biogeography*, *45*, 1781–1794.

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics*, *30*, 3317–3324.

Chippaux, J. P., & Jackson, K. (2019). *Snakes of Central and Western Africa*. Baltimore, MD: John Hopkins University Press.

Chirio, L., & LeBreton, M. (2007). *Atlas des reptiles du Cameroun* (Vol. 67). IRD Editions.

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

Cobos, M. E., Peterson, A. T., Barve, N., & Osorio-Olvera, L. (2019). Kuenm: An R package for detailed development of ecological niche models using Maxent. *PeerJ*, *7*, e6281.

Coetzee, J. A. (1964). Evidence for a considerable depression of the vegetation belts during the Upper Pleistocene on the East African mountains. *Nature*, *204*, 564–566.

Coetzee, J. A., & van Zinderen Bakker, E. M. (1970). Palaeoecological problems of the Quaternary of Africa. *South African Journal of Science*, *66*, 78.

Colinvaux, P. A., De Oliveira, P. E., Moreno, J. E., Miller, M. C., & Bush, M. B. (1996). A long pollen record from lowland Amazonia: Forest and cooling in glacial times. *Science (Washington)*, *274*, 85–88.

Colinvaux, P. A., Irion, G., Räsänen, M. E., Bush, M. B., & De Mello, J. N. (2001). A paradigm to be discarded: Geological and paleoecological data falsify the Haffer & Prance refuge hypothesis of Amazonian speciation. *Amazoniana*, *16*, 609–646.

Collins, S. M., Yuan, S., Tan, P. N., Oliver, S. K., Lapierre, J. F., Cheruvelil, K. S., ... & Soranno, P. A. (2019). Winter precipitation and summer temperature predict lake water quality at macroscales. *Water Resources Research*, *55*, 2708–2721.

Colyn, M. M. (1987). Les Primates des forêts ombrophiles de la cuvette du Zaïre: Interprétations zoogéographiques des modèles de distribution. *Revue de Zoologie Africaine, Tervuren, 101*, 183–96.

Colyn, M. M. (1991). L'importance zoogéographique du bassin du fleuve Zaïre pour la spéciation: Le cas des Primates Siens. *Annales Musée Royal de I; Afrique Centrale, Section Zoologie, Tervuren, 264*, 1– 250.

Colyn, M., Gautier-Hion, A. & Verheyen, W. (1991). A re-appraisal of palaeoenvironmental history in Central Africa: Evidence for a major fluvial refuge in the Zaire Basin. *Journal of Biogeography*, *18*, 403–407.

Connell, J. H. (1961). The influence of interspecific competition and other factors on the distribution of the barnacle *Chthamalus stellatus*. *Ecology*, *42*, 710–723.

Cooper, N., Freckleton, R. P., & Jetz, W. (2011). Phylogenetic conservatism of environmental niches in mammals. *Proceedings of the Royal Society B*, *278*, 2384–2391.

Cooper, J. C., Maddox, J. D., McKague, K., & Bates, J. M. (2021). Multiple lines of evidence indicate ongoing allopatric and parapatric diversification in an Afromontane sunbird (*Cinnyris reichenowi*). *The Auk*, *138*, ukaa081.

Cooper, T. J., Wannenburgh, A. M., & Cherry, M. I. (2017). Atlas data indicate forest dependent bird species declines in South Africa. *Bird Conservation International*, *27*, 337–354.

Couvreur, T. L., Chatrou, L. W., Sosef, M. S., & Richardson, J. E. (2008). Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. *BMC Biology*, *6*, 54.

Couvreur, T. L. P., Dauby, G., Blach-Overgaard, A., Deblauwe, V., Dessein, S., Droissart, V., …& Sepulchre, P. (2021). Tectonics, climate and the diversification of the tropical African terrestrial flora and fauna. *Biological Reviews, 96,* 16–51.

Cowie, R. H., & Holland, B. S. (2006). Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. *Journal of Biogeography*, *33*, 193–198.

Cracraft, J. (1985). Historical biogeography and patterns of differentiation within the South American avifauna: Areas of endemism. *Ornithological Monographs*, *36*, 49–84.

Crisp, M. D., Arroyo, M. T. K., Cook, L. G., Gandolfo, M. A., Jordan, G. J., McGlone, M. S., … & Linder, H. P. (2009). Phylogenetic biome conservatism on a global scale. *Nature*, 458, 754–756.

Dagallier, L. P. M., Janssens, S. B., Dauby, G., Blach-Overgaard, A., Mackinder, B. A., Droissart, V., ... & Couvreur, T. L. (2020). Cradles and museums of generic plant diversity across tropical Africa. *New Phytologist*, *225*, 2196–2213.

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.

Darlington, P. J., Jr. (1957). *Zoogeography and the geographical distribution of animals*. New York, NY: John Wiley & Sons.

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.

David, K. T., & Halanych, K. M. (2021). Spatial proximity between polyploids across South American frog genera. *Journal of Biogeography, 48,* 991–1000.

de Balsac, H. H. & Hutterer, R. (1982) Les Soricidae (mammiferes insectivores) des îles du Golfe de Guinée; faites nouveaux et problémes biogéographiques. *Bonner Zoologische Beiträge*, *33*, 133–150.

de Bodt, S., Maere, S., & Van de Peer, Y. (2005). Genome duplication and the origin of angiosperms. *Trends in Ecology & Evolution*, *20*, 591–597.

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.

del Hoyo, J., Elliott, A., & Christie, D. (2004). *Handbook of the Birds of the World. Vol. 9: Cotingas to Pipits and Wagtails.* Barcelona, Spain: Lynx Edicions.

del Hoyo, J., Elliott, A., & Christie, D. (2005). *Handbook of the Birds of the World. Vol. 10: Cuckoo-shrikes to Thrushes.* Barcelona, Spain: Lynx Edicions.

del Hoyo, J., Elliott, A., & Christie, D. (2006). *Handbook of the Birds of the World. Vol. 11: Old World Flycatcher's to the Old World Warblers.* Barcelona, Spain: Lynx Edicions.

del Hoyo, J., Elliott, A., & Christie, D. (2007). *Handbook of the Birds of the World: Vol. 12: Picathartes to Tits and Chickadees.* Barcelona, Spain: Lynx Edicions.

del Hoyo, J., Elliott, A., & Christie, D. (2008). *Handbook of the Birds of the World: Vol. 13: Penduline-tits to Shrikes.* Barcelona, Spain: Lynx Edicions.

del Hoyo, J., Elliott, A., & Christie, D. (2009). *Handbook of the Birds of the World: Vol. 14: Bush-shrikes to Old World Sparrows*. Barcelona, Spain: Lynx Edicions.

del Hoyo, J., Elliott, A., & Christie, D. (2010). *Handbook of the Birds of the World: Vol. 15: Weavers to New World Warblers.* Barcelona, Spain: Lynx Edicions.

De Klerk, H. M., Crowe, T. M., Fjeldså, J., & Burgess, N. D. (2002). Patterns of species richness and narrow endemism of terrestrial bird species in the Afrotropical region. *Journal of Zoology*, *256*, 327–342.

e Lima, R. F., & Melo, M. (2021). A revised bird checklist for the oceanic islands of the Gulf of Guinea (Príncipe, São Tomé and Annobón). *Bulletin of the British Ornithologists' Club*, *141*, 179–198.

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

Demenou, B. B., Doucet, J. L., & Hardy, O. J. (2018). History of the fragmentation of the African rain forest in the Dahomey Gap: Insight from the demographic history of *Terminalia superba*. *Heredity*, *120*, 547–561.

de Oliveira Gasparotto, V. P., Micheletti, T., Mangini, P. R., & Dias, R. A. (2020). Impact of invasive species on the density and body size of an insular endemic lizard (*Trachylepis atlantica*). *ResearchSquare*. DOI: https://doi.org/10.21203/rs.3.rs-17579/v1

Deruelle, B., Moreau, C., Nkoumbou, C., Kambou, R., Lissom, J., Njofang, E., & Nono, A. (1991). The Cameroon Volcanic Line: A review. In A. B. Kampunzu, & R. T. Lubala (Eds.), *Magmatism in extensional structural settings*. Berlin: Springer

De Sa, R. O., & Hillis, D. M. (1990). Phylogenetic relationships of the pipid frogs *Xenopus* and *Silurana*: An integration of ribosomal DNA and morphology. *Molecular Biology and Evolution*, *7*, 365–376.

Devos, N., Barker, N. P., Nordenstam, B., & Mucina, L. (2010). A multi-locus phylogeny of *Euryops* (Asteraceae, Senecioneae) augments support for the" Cape to Cairo" hypothesis of floral migrations in Africa. *Taxon*, *59*, 57–67.

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.

Dokulil, M. T., de Eyto, E., Maberly, S. C., May, L., Weyhenmeyer, G. A., & Woolway, R. I. (2021). Increasing maximum lake surface temperature under climate change. *Climatic Change*, *165*, 1–17.

Dongmo, J. B., DaCosta, J. M., Djieto-Lordon, C., Ngassam, P., & Sorenson, M. D. (2019). Variable phylogeographic histories of five forest birds with populations in Upper and Lower Guinea: Implications for taxonomy and evolutionary conservation. *Ostrich*, *90*, 257–270.

Donkpegan, A. S., Doucet, J. L., Migliore, J., Duminil, J., Dainou, K., Piñeiro, R., … & Hardy, O. J. (2017). Evolution in African tropical trees displaying ploidy-habitat association: The genus *Afzelia* (Leguminosae). *Molecular Phylogenetics and Evolution*, *107*, 270–281.

Doyle, J.J., & Egan, A.N. 2010. Dating the origins of polyploidy events. *New Phytologist*, *186*, 73–85.

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.

Drewes, R.C. & Wilkinson, J.A. (2004) The California Academy of Sciences Gulf of Guinea expedition (2001) I. The taxonomic status of the genus *Nesionixalus* Perret, 1976 (Anura:

Hyperoliidae), treefrogs of São Tomé and Príncipe, with comments on the genus *Hyperolius*. *Proceedings of the California Academy of Sciences*, *55*, 395–407.

Droissart, V., Dauby, G., Hardy, O. J., Deblauwe, V., Harris, D. J., Janssens, S., ... & Couvreur, T. L. (2018). Beyond trees: Biogeographical regionalization of tropical Africa. *Journal of Biogeography*, *45*, 1153–1167.

Dupont, L. M., Donner, B., Schneider, R., & Wefer, G. (2001). Mid-Pleistocene environmental change in tropical Africa began as early as 1.05 Ma. *Geology*, *29*, 195–198.

Dupont, L. M., & Weinelt, M. (1996). Vegetation history of the savanna corridor between the Guinean and the Congolian rain forest during the last 150,000 years. *Vegetation History and Archaeobotany*, *5*, 273–292.

Duran, A., Meyer, A. L., & Pie, M. R. (2013). Climatic niche evolution in New World monkeys (Platyrrhini). *PLoS One*, *8*, e83684.

Edwards, C. E., Soltis, D. E., & Soltis, P. S. (2006). Molecular phylogeny of *Conradina* and other scrub mints (Lamiaceae) from the southeastern USA: Evidence for hybridization in Pleistocene refugia? *Systematic Botany*, *31*, 193–207.

Ehrendorfer, F. (1980). Polyploidy and distribution. In Lewis, W. (ed). *Polyploidy*. New York, NY; Plenum Press, pp. 45–60.

Ekimova, N. V., Muratova, E. N., & Silkin, P. P. (2012). The role of polyploidy in adaptation and settling of steppe shrubs in Central Asia. *Russian Journal of Genetics: Applied Research*, *2*, 1055–109.

Elenga, H., Schwartz, D., & Vincens, A. (1994). Pollen evidence of late Quaternary vegetation and inferred climate changes in Congo. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *109*, 345–356.

Engelbrecht, H. M., Branch, W. R., Greenbaum, E., Alexander, G. J., Jackson, K., Burger, M., ... & Tolley, K. A. (2019). Diversifying into the branches: Species boundaries in African green and bush snakes, *Philothamnus* (Serpentes: Colubridae). *Molecular Phylogenetics and Evolution*, *130*, 357–365.

Eriksson, J., Hohmann, G., Boesch, C. & Vigilant, L. (2004) Rivers influence the population genetic structure of bonobos (*Pan paniscus*). *Molecular Ecology*, *13*, 3425–3435.

Escobar, L. E., Lira-Noriega, A., Medina-Vogel, G., & Peterson, A.T. (2014). Potential for spread of the white-nose fungus (*Pseudogymnoascus destructans*) in the Americas: Use of Maxent and Niche A to assure strict model transference. *Geospatial Health*, *9*, 221–229.

Estrela, F. (1994). Diversity and endemism of angiosperms in the Gulf of Guinea islands. *Biodiversity & Conservation*, *3*, 785–793.

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.

Evans, B. J., Kelley, D. B., Tinsley, R. C., Melnick, D. J., & Cannatella, D. C. (2004). A mitochondrial DNA phylogeny of African clawed frogs: Phylogeography and implications for polyploid evolution. *Molecular Phylogenetics and Evolution*, *33*, 197–213.

Evans, B. J. (2008). Genome evolution and speciation genetics of clawed frogs (*Xenopus* and *Silurana*). *Frontiers in Bioscience*, *13*, 4687–4706.

Evans, B. J., Carter, T. F., Tobias, M. L., Kelley, D. B., Hanner, R., & Tinsley, R. C. (2008). A new species of clawed frog (genus *Xenopus*) from the Itombwe Massif, Democratic Republic of the Congo: Implications for DNA barcodes and biodiversity conservation. *Zootaxa*, *1780*, 55–68.

Evans, B. J., Greenbaum, E., Kusamba, C., Carter, T. F., Tobias, M. L., Mendel, S. A., & Kelley, D. B. (2011). Description of a new octoploid frog species (Anura: Pipidae: *Xenopus*) from the Democratic Republic of the Congo, with a discussion of the biogeography of African clawed frogs in the Albertine Rift. *Journal of Zoology*, *283*, 276–290.

Evans, B. J., Pyron, R. A., & Wiens, J. J. (2012). Polyploidization and sex chromosome evolution in amphibians. In Soltis P., & Soltis D. (eds.) *Polyploidy and Genome Evolution*. Berlin, Heidelberg, Germany; Springer, pp. 385–410.

Evans, B. J., Carter, T. F., Greenbaum, E., Gvoždík, V., Kelley, D. B., McLaughlin, P. J., Pauwels, O. G. S., Portik, D. M., Stanley, E. L., Tinsley, R. C., Tobias, M. L., & Blackburn, D. C. (2015). Genetics, morphology, advertisement calls, and historical records distinguish six new polyploid species of African clawed frog (*Xenopus,* Pipidae) from West and Central Africa. *PLoS One*, *10*, e0142823.

Evans, B. J., & Kwon, T. (2015). Molecular polymorphism and divergence of duplicated genes in tetraploid African clawed frogs (*Xenopus*). *Cytogenetic and Genome Research*, *145*, 243–252.

Evans, B. J., Upham, N. S., Golding, G. B., Ojeda, R. A., & Ojeda, A. A. (2017). Evolution of the largest mammalian genome. *Genome Biology and Evolution*, *9*, 1711–1724.

Evans, B. J., Gansauge, M. T., Stanley, E. L., Furman, B. L., Cauret, C. M., Ofori-Boateng, C., ... & Meyer, M. (2019). *Xenopus fraseri*: Mr. Fraser, where did your frog come from? *PLoS One*, *14*, e0220892.

Evans, M. E., Smith, S. A., Flynn, R. S., & Donoghue, M. J. (2009). Climate, niche evolution, and diversification of the “bird-cage” evening primroses (*Oenothera,* sections *Anogra* and *Kleinia*). *American Naturalist*, *173*, 225–240.

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V.C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, *9*, e1003905.

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.

Farallo, V. R., Muñoz, M. M., Uyeda, J. C., & Miles, D. B. (2020). Scaling between macro-to microscale climatic data reveals strong phylogenetic inertia in niche evolution in plethodontid salamanders. *Evolution*, *74*, 979–991.

Farwig, N., Böhning-Gaese, K., & Bleher, B. (2006). Enhanced seed dispersal of *Prunus africana* in fragmented and disturbed forests? *Oecologia*, *147*, 238–252.

Fawcett, J. A., Maere, S., & Van De Peer, Y. (2009). Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. *Proceedings of the National Academy of Sciences USA*, *106*, 57375–5742.

Faye, A., Deblauwe, V., Mariac, C., Richard, D., Sonké, B., Vigouroux, Y., & Couvreur, T. (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.

Feakins, S. J., Levin, N. E., Liddy, H. M., Sieracki, A., Eglinton, T. I., & Bonnefille, R. (2013). Northeast African vegetation change over 12 my. *Geology*, *41*, 295–298.

Feder, M. E., & Burggren, W. W. (Eds.). (1992). *Environmental physiology of the amphibians*. University of Chicago Press.

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*, 342–350.

Feng, X., & Papeş, M. (2017). Physiological limits in an ecological niche modeling framework: A case study of water temperature and salinity constraints of freshwater bivalves invasive in USA. *Ecological Modelling*, *346*, 48–57.

Ficetola, G. F., & Stöck, M. (2016). Do hybrid origin polyploid amphibians occupy transgressive or intermediate ecological niches compared to their diploid ancestors? *Journal of Biogeography*, *43*, 703–715.

Ficetola, G. F., Lunghi, E., Canedoli, C., Padoa-Schioppa, E., Pennati, R., & Manenti, R. (2018). Differences between microhabitat and broad-scale patterns of niche evolution in terrestrial salamanders. *Scientific Reports*, *8*, 1–12.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*, 4302–4315.

Figueiredo, E., Paiva, J., Stevart, T., Oliveira, F., & Smith, G. F. (2011). Annotated catalogue of the flowering plants of São Tomé and Príncipe. *Bothalia*, *41*, 41–82.

Fjeldså, J. (1994). Geographical patterns for relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodiversity & Conservation*, *3*, 207–226.

Fjeldså, J., & Bowie, R. C. (2008). New perspectives on the origin and diversification of Africa's forest avifauna. *African Journal of Ecology*, *46*, 235–247.

Fjeldså, J., Bowie, R. C., & Rahbek, C. (2012). The role of mountain ranges in the diversification of birds. *Annual Review of Ecology, Evolution, and Systematics*, *43*, 249–265.

Fjeldsaå, J., & Lovett, J. C. (1997). Geographical patterns of old and young species in African forest biota: the significance of specific montane areas as evolutionary centres. *Biodiversity & Conservation*, *6*, 325–346.

Flouri, T., Jiao, X., Rannala, B., & Yang, Z. (2018). Species tree inference with BPP using genomic sequences and the multispecies coalescent. *Molecular Biology and Evolution*, *35*, 2585–2593.

Flügel, T. J., Eckardt, F. D., & Cotterill, F. P. (2015). The present day drainage patterns of the Congo river system and their Neogene evolution. In De Wit, M. J., Guillocheau, F., & De Wit, M. C. (Eds.) *Geology and Resource Potential of the Congo Basin*. (pp. 315–337). Berlin, Germany: Springer.

Fowler, N. L., & Levin, D. A. (1984). Ecological constraints on the establishment of a novel polyploid in competition with its diploid progenitor. *American Naturalist*, *124*, 703–711.

Francis, R. M. (2017). Pophelper: An R package and web app to analyze and visualize population structure. *Molecular Ecology Resources*, *17*, 27–32.

François, O. (2016). Running structure-like population genetic analyses with R. *R tutorials in Population Genetics, U. Grenoble-Alpes*, 1–9.

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*, 925–929.

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G. & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*, 973–983.

Fritz, U., Branch, W. R., Hofmeyr, M. D., Maran, J., Prokop, H., Schleicher, A., ... & Hundsdoerfer, A. K. (2011). Molecular phylogeny of African hinged and helmeted terrapins (Testudines: Pelomedusidae: *Pelusios* and *Pelomedusa*). *Zoologica Scripta*, *40*, 115–125.

Frugone, M. J., López, M. E., Segovia, N. I., Cole, T. L., Lowther, A., Pistorius, P., ... & Vianna, J. A. (2019). More than the eye can see: Genomic insights into the drivers of genetic

differentiation in Royal/Macaroni penguins across the Southern Ocean. *Molecular Phylogenetics and Evolution*, *139*, 106563.

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

Fujita, M. K., & Moritz, C. (2009). Origin and evolution of parthenogenetic genomes in lizards: Current state and future directions. *Cytogenetic and Genome Research*, *127*, 261–272.

Furman, B. L., Bewick, A. J., Harrison, T. L., Greenbaum, E., Gvoždík, V., Kusamba, C., & Evans, B. J. (2015). Pan-African phylogeography of a model organism, the African clawed frog '*Xenopus laevis*'. *Molecular Ecology*, *24*, 909–925.

Furman, B. L., & Evans, B. J. (2016). Sequential turnovers of sex chromosomes in African clawed frogs (*Xenopus*) suggest some genomic regions are good at sex determination. *G3: Genes, Genomes, Genetics*, *6*, 3625–3633.

Furman, B. L., Dang, U. J., Evans, B. J., & Golding, G. B. (2018). Divergent subgenome evolution after allopolyploidization in African clawed frogs (*Xenopus*). *Journal of Evolutionary Biology*, *31*, 1945–1958.

Gascon, C., Malcolm, J. R., Patton, J. L., da Silva, M. N., 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.

Gaubert, P., Njiokou, F., Ngua, G., Afiademanyo, K., Dufour, S., Malekani, J., … & Antunes, A. (2016). Phylogeography of the heavily poached African common pangolin (Pholidota, *Manis tricuspis*) reveals six cryptic lineages as traceable signatures of Pleistocene diversification. *Molecular Ecology*, *25*, 5975–5993.

Gaynor, M. L., Marchant, D. B., Soltis, D. E., & Soltis, P. S. (2018). Climatic niche comparison among ploidal levels in the classic autopolyploid system, *Galax urceolata*. *American Journal of Botany*, *105*, 1631–1642.

Girese, P., Maley, J. & Brenac, P. (1994). Late Quaternary palaeoenvironments in the Lake Barombi Mbo (Cameroon) deduced from pollen and carbon isotopes of organic matter. *Palaeogeography, Palaeoclimatology, Palaeoecology, 107*, 65–78.

Glover, N. M., Redestig, H., & Dessimov, C. (2016). Homoeologs: What are they and how do we infer them? *Trends in Plant Science*, *21*, 609–621.

Godsoe W. (2010). Regional variation exaggerates ecological divergence in niche models. *Systematic Biology*, *59*, 298–306.

Goldberg, S. R. (2008). Reproductive cycle of the western three-striped skink, *Trachylepis occidentalis* (Squamata: Scincidae), from southern Africa. *Salamandra*, *44*, 1235–126.

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*, 4766–4771.

Gonmadje, C. F., Doumenge, C., Sunderland, T. C., Balinga, M. P., & Sonké, B. (2012). Phytogeographical analysis of Central African forests: The Ngovayang massif (Cameroon). *Plant Ecology and Evolution*, *145*, 152–164.

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

Graham, C. H., Ron, S. R., Santos, J. C., Schneider, C. J., & Moritz, C. (2004). Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution*, *58*, 1781–1793.

Greenbaum E. (2017). *Emerald Labyrinth: A Scientist's Adventures in the Jungles of the Congo*. Lebanon, NH: University Press of New England.

Greenbaum, E., Allen, K. E., Vaughan, E. R., Pauwels, O. S., Wallach, V. Kusamba, C., … & Brown, R.M. (2021). Night stalkers from above: A monograph of *Toxicodryas* tree snakes (Squamata: Colubridae) with descriptions of two new cryptic species from Central Africa. *Zootaxa*, *4965*, 1–44.

Greenbaum, E., Portillo, F., Jackson, K., & Kusamba, C. (2015). A phylogeny of Central African *Boaedon* (Serpentes: Lamprophiidae), with the description of a new cryptic species from the Albertine Rift. *African Journal of Herpetology*, *64*, 18-38.

Greenbaum, E., Tolley, K. A., Joma, A., & Kusamba, C. (2012). A new species of chameleon (Sauria: Chamaeleonidae: Kinyongia) from the northern Albertine Rift, Central Africa. *Herpetologica*, *68*, 60–75.

Gregory, T. R., & Mable, B. K. (2005). Polyploidy in animals. In Gregory, T. R. (ed.) *The Evolution of the Genome*. San Diego: Elsevier Academic Press, pp. 427–517.

Guillocheau, F., Simon, B., Baby, G., Bessin, P., Robin, C., & Dauteuil, O. (2018). Planation surfaces as a record of mantle dynamics: The case example of Africa. *Gondwana Research*, *53*, 82–98.

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

Haffer, J. R. (1974). Avian speciation in tropical South America. *Publications of the Nuttall Ornithological Club, 14*, 1–390.

Haffer, J. R. (1982). General aspects of the refuge theory. In Prance, G. T. (Ed.) *Biological Diversification in the Tropics*. (pp. 6–24). New York, NY: Columbia University Press.

Haffer, J. R. (1992). On the "river effect" in some forest birds of southern Amazonia. *Boletim do Museu Paraense Emílio Goeldi. Nova série. Zoologia*, *8*, 217–245.

Haffer, J. R. (1993). Time's cycle and time's arrow in the history of Amazonia. *Compte Rendu des Séances de la Société de Biogéographie*, *69*, 15–45.

Haffer, J. R. (1997). Alternative models of vertebrate speciation in Amazonia: An overview. *Biodiversity and Conservation*, *6*, 451–476.

Haffer J. R. (2008). Hypotheses to explain the origin of species in Amazonia. *Brazilian Journal of Biology*, *68*, 917– 947.

Hale, M. L., Burg, T. M., & Steeves, T. E. (2012). Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. *PLoS ONE*, *7*, e45170.

Hall, J. B. (1973). Vegetational zones on the southern slopes of Mount Cameroon. *Vegetatio*, *27*, 49–69.

Hampe, A., & Jump, A. S. (2011). Climate relicts: Past, present, future. *Annual Review of Ecology, Evolution, and Systematics*, *42*, 313–333.

Hansen, T. F. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution*, *51*, 1341–1351.

Hanski, I., & Gyllenberg, M. (1997). Uniting two general patterns in the distribution of species. *Science*, *275*, 397–400.

Harcourt, A. H., & Wood, M. A. (2012). Rivers as barriers to primate distributions in Africa. *International Journal of Primatology*, *33*, 168–183.

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.

Harvey, P. H., & Rambaut, A. (2000). Comparative analyses for adaptive radiations. *Philosophical Transactions of the Royal Society of London B*, *355*, 1599–1605.

Hastings, A., Cuddington, K., Davies, K. F., Dugaw, C. J., Elmendorf, S., Freestone, A., ... & Thomson, D. (2005). The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, *8*, 91–101.

Haufler, C. H., Windham, M. D., & Rabe, E. W. (1995). Reticulate evolution in the *Polypodium vulgare* complex. *Systematic Botany*, *20*, 89–109.

Haus, T., Akom, E., Agwanda, B., Hofreiter, M., Roos, C., & Zinner, D. (2013). Mitochondrial diversity and distribution of African green monkeys (*Chlorocebus* Gray, 1870). *American Journal of Primatology*, *75*, 350–360.

Head, J. J., Mahlow, K., & Müller, J. (2016). Fossil calibration dates for molecular phylogenetic analysis of snakes 2: Caenophidia, Colubroidea, Elapoidea, Colubridae. *Palaeontologia Electronica*, *19*, 1–21.

Heaney, L. R. (2000). Dynamic disequilibrium: A long-term, large-scale perspective on the equilibrium model of island biogeography. *Global Ecology and Biogeography*, *9*, 59–74.

Hedberg, O. (1969). Evolution and speciation in a tropical high mountain flora. *Biological Journal of the Linnean Society*, *1*, 135–148.

Helmstetter, A. J., Amoussou, B. E., Bethune, K., Kandem, N. G., Kakai, R. L. G., Sonké, B., & Couvreur, T. L. (2020a). Phylogenomic approaches reveal how a climatic inversion and glacial refugia shape patterns of diversity in an African rain forest tree species. *bioRxiv*, 807727.

Helmstetter, A. J., Béthune, K., Kamdem, N. G., Sonké, B., & Couvreur, T. L. (2020b). Individualistic evolutionary responses of Central African rain forest plants to Pleistocene climatic fluctuations. *Proceedings of the National Academy of Sciences*, *117*, 32509–32518.

Herbert, T. D., Lawrence, K. T., Tzanova, A., Peterson, L. C., Caballero-Gill, R., & Kelly, C. S. (2016). Late Miocene global cooling and the rise of modern ecosystems. *Nature Geoscience*, *9*, 843–847.

Hershkovitz, P. (1977). *Living New World Monkeys (Platyrrhini).* Chicago, IL: University of Chicago Press.

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.

Hillman, S. S., Drewes, R. C., Hedrick, M. S., & Hancock, T. V. (2014). Physiological vagility: Correlations with dispersal and population genetic structure of amphibians. *Physiological and Biochemical Zoology*, *87*, 105–112.

Hime, P. M., Briggler, J. T., Reece, J. S., & Weisrock, D. W. (2019). Genomic data reveal conserved female heterogamety in giant salamanders with gigantic nuclear genomes. *G3: Genes | Genomes | Genetics*, *9*, 3467–3476.

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*, 518–522.

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.

Holt, R. D., & Barfield, M. (2008). Habitat selection and niche conservatism. *Israel Journal of Ecology & Evolution*, *54*, 295–309.

Holt, R. D., & Gomulkiewicz, R. (1997). The evolution of species' niches: A population dynamic perspective. In Othmer, H. G., Adler, F. R., Lewis, M. A., & Dallon, J. C. (eds.) *Case Studies in Mathematical Modelling: Ecology, Physiology, and Cell Biology*. Saddle River, N.J.; Prentice-Hall, pp. 25–50.

Hrdina, A., & Romportl, D. (2017). Evaluating global biodiversity hotspots–Very rich and even more endangered. *Journal of Landscape Ecology*, *10*, 108–115.

Hughes, D. F., Kusamba, C., Behangana, M., & Greenbaum, E. (2017). Integrative taxonomy of the Central African forest chameleon, *Kinyongia adolfifriderici* (Sauria: Chamaeleonidae), reveals underestimated species diversity in the Albertine Rift. *Zoological Journal of the Linnean Society*, *181*, 400–438.

Huntley, J. W., Harvey, J. A., Pavia, M., Boano, G., & Voelker, G. (2018). The systematics and biogeography of the Bearded Greenbuls (Aves: Criniger) reveals the impact of Plio-Pleistocene forest fragmentation on Afro-tropical avian diversity. *Zoological Journal of the Linnean Society*, *183*, 672–686.

Huntley, J. W., Keith, K. D., Castellanos, A. A., Musher, L. J., & Voelker, G. (2019). Underestimated and cryptic diversification patterns across Afro-tropical lowland forests. *Journal of Biogeography*, *46*, 381–391.

Huntley, J. W., & Voelker, G. (2017). A tale of the nearly tail-less: the effects of Plio-Pleistocene climate change on the diversification of the African avian genus *Sylvietta*. *Zoologica Scripta*, *46*, 523–535.

Huson, D. H., & Scornavacca, C. (2012). Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Systematic Biology*, *61*, 1061–1067.

Hutchinson, G.E. (1957) Concluding remarks. Population Studies: Animal Ecology and Demography. *Cold Spring Harbor Symposium on Quantitative Biology*, 22, 415–457.

Igbokwe, J., Nicolas, V., Oyeyiola, A., Obadare, A., Adesina, A. S., Awodiran, M. O., ... & Olayemi, A. (2019). Molecular taxonomy of *Crocidura* species (Eulipotyphla: Soricidae) in a key biogeographical region for African shrews, Nigeria. *Comptes Rendus Biologies*, *342*, 108–117.

Iherring, R. (1911) Cobras e amphibios das ilhotas de Aguapé. *Revista do Museu Paulista*, *8*, 454–461.

IUCN (2020) The IUCN Red List of Threatened Species. Version 2020-2. https://www.iucnredlist.org.

Jablonski, D., Roy, K., & Valentine, J. W. (2006). Out of the tropics: Evolutionary dynamics of the latitudinal diversity gradient. *Science*, *314*, 102–106.

Jackson, S. T., Betancourt, J. L., Booth, R. K., & Gray, S. T. (2009). Ecology and the ratchet of events: climate variability, niche dimensions, and species distributions. *Proceedings of the National Academy of Sciences*, *106*, 19685–19692.

Jacobs, B.F. (2004). Paleobotanical studies from Tropical Africa: Relevance to the evolution of forest, woodland and savannah biomes. *Philosophical Transactions of the Royal Society B*., *359*, 1573–1583.

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.

Janecek, S., Hrazsky, Z., Bartos, M., Brom, J., Reif, J., Horák, D., ... & Pesata, M. (2007). Importance of big pollinators for the reproduction of two *Hypericum* species in Cameroon, West Africa. *African Journal of Ecology*, *45*, 607–613.

Jansson, R., & Dynesius, M. (2002). The fate of clades in a world of recurrent climatic change: Milankovitch oscillations and evolution. *Annual Review of Ecology and Systematics*, *33*, 741–777.

Jakob, S. S., Heibl, C., Rödder, D., & Blattner, F. R. (2010). Population demography influences climatic niche evolution: evidence from diploid American *Hordeum* species (Poaceae). *Molecular Ecology*, *19*, 1423–1438.

Jeffries, D. L., Copp, G. H., Lawson Handley, L., Olsén, K. H., Sayer, C. D., & Hänfling, B. (2016). Comparing RAD seq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Molecular Ecology*, *25*, 2997–3018.

Jesus, J., Brehm, A., & Harris, D. J. (2005a). Relationships of scincid lizards (*Mabuya* spp.) from the islands of the Gulf of Guinea based on mtDNA sequence data. *Amphibia-Reptilia*, *26*, 467–473.

Jesus, J., Brehm, A., & Harris, D. J. (2006). Phylogenetic relationships of *Lygodactylus* geckos from the Gulf of Guinea islands: rapid rates of mitochondrial DNA sequence evolution?. *The Herpetological Journal*, *16*, 291–295.

Jesus, J., Harris, D. J., & Brehm, A. (2005b). Phylogeography of *Mabuya maculilabris* (Reptilia) from São Tomé island (gulf of Guinea) inferred from mtDNA sequences. *Molecular Phylogenetics and Evolution*, *37*, 503–510.

Jesus, J., Harris, D. J., & Brehm, A. (2007). Relationships of *Afroablepharus* Greer, 1974 skinks from the Gulf of Guinea islands based on mitochondrial and nuclear DNA: Patterns of colonization and comments on taxonomy. *Molecular Phylogenetics and Evolution*, *45*, 904–914.

Jesus, J., Nagy, Z. T., Branch, W. R., Wink, M., Brehm, A., & Harris, D. J. (2009). Phylogenetic relationships of African green snakes (genera *Philothamnus* and *Hapsidophrys*) from Sao

Tome, Principe and Annobón islands based on mtDNA sequences, and comments on their colonization and taxonomy. *The Herpetological Journal*, *19*, 41–48.

Johnson, S. D., & Brown, M. (2004). Transfer of pollinaria on birds' feet: A new pollination system in orchids. *Plant Systematics and Evolution*, *244*, 181–188.

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

Jones, P. J. (1994). Biodiversity in the Gulf of Guinea: an overview. *Biodiversity & Conservation*, *3*, 772–784.

Jongsma, G. F. M., Barej, M. F., Barratt, C. D., Burger, M., Conradie, W., Ernst, R., … & Blackburn, D. C. (2018). Diversity and biogeography of frogs in the genus *Amnirana* (Anura: Ranidae) across sub-Saharan Africa. *Molecular Phylogenetics and Evolution*, *120*, 274–285.

Juan, C., Emerson, B. C., Oromí, P., & Hewitt, G. M. (2000). Colonization and diversification: Towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology & Evolution*, *15*, 104–109.

Kadu, C. A. C., Schueler, S., Konrad, H., Muluvi, G. M. M., Eyog-Matig, O., Muchugi, A., ... & Geburek, T. (2011). Phylogeography of the Afromontane *Prunus africana* reveals a former migration corridor between East and West African highlands. *Molecular Ecology*, *20*, 165–178.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*, 587.

Kaplin, B. A., Munyaligoga, V., & Moermond, T. C. (1998). The Influence of Temporal Changes in Fruit Availability on Diet Composition and Seed Handling in Blue Monkeys (*Cercopithecus mitis doggetti*). *Biotropica*, *30*, 56–71.

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.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... & Thierer, T. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649.

Kellogg, E.A. (2016). Has the connection between polyploidy and diversification actually been tested? *Current Opinion in Plant Biology*, *30*, 25–32.

Khokha, M. K., Chung, C., Bustamante, E. L., Gaw, L. W., Trott, K. A., Yeh, J., ... & Grammer, T. C. (2002). Techniques and probes for the study of *Xenopus tropicalis* development. *Developmental Dynamics*, *225*, 499–510.

King, W. (1962). The occurrence of rafts for dispersal of land animals into the West Indies. *Quarterly Journal of the Florida Academy of Sciences*, *25*, 45–52.

Kingdon, J. (1989) Island Africa. Princeton, NJ.

Kobel, H. R., & Du Pasquier, L. (1986). Genetics of polyploid *Xenopus*. *Trends in Genetics*, *2*, 310–315.

Kobel, H. R. (1996). Allopolyploid speciation. In: Tinsley, R. C. & Kobel, H. R. (eds.) *The Biology of Xenopus*. Oxford, UK: Clarendon Press, pp. 390–401.

Komatsu, E., Fukushima, T., & Harasawa, H. (2007). A modeling approach to forecast the effect of long-term climate change on lake water quality. *Ecological Modelling*, *209*, 351–366.

Kornilios, P., Giokas, S., Lymberakis, P., & Sindaco, R. (2013). Phylogenetic position, origin and biogeography of Palearctic and Socotran blind-snakes (Serpentes: Typhlopidae). *Molecular Phylogenetics and Evolution*, *68*, 35–41.

Kozak, K. H., & Wiens, J. J. (2010). Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecology Letters*, *13*, 1378–1389.

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.

Labra, A., Pienaar, J., & Hansen, T. F. (2009). Evolution of thermal physiology in *Liolaemus* lizards: Adaptation, phylogenetic inertia, and niche tracking. *The American Naturalist*, *174*, 204–220.

Landguth, E. L., Fedy, B. C., Oyler-McCance, S. J., Garey, A. L., Emel, S. L., Mumma, M., ... & Cushman, S. A. (2012). Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Molecular Ecology Resources*, *12*, 276–284.

Landis, M. J., Matzke, N. J., Moore, B. R., & Huelsenbeck, J. P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology*, *62*, 789–804.

Languy, M., & Motombe, F. N. (2003). Birds of Takamanda Forest Reserve, Cameroon. In Cominskey, J. A., Sunderland, T. C. H., Sunderland-Groves, J. L. (Eds.) *Takamanda: The biodiversity of an African Rainforest.* SI/MAB: Smithsonian Institution/Monitoring and Assessment of Biodiversity Series #8. Washington DC.

Larson, T. R., Castro, D., Behangana, M., & Greenbaum, E. (2016). Evolutionary history of the river frog genus *Amietia* (Anura: Pyxicephalidae) reveals extensive diversification in Central African highlands. *Molecular Phylogenetics and Evolution*, *99*, 168–181.

Lawson, R., Slowinski, J.B., Crother, B.I., & Burbrink, F.T. (2005). Phylogeny of the Colubroidea (Serpentes): New evidence from mitochondrial and nuclear genes. *Molecular Phylogenetics and Evolution*, *37*, 581–601.

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

Leaché, A. D., Grummer, J. A., Miller, M., Krishnan, S., Fujita, M. K., Böhme, W., ... Ofori-Boateng, C. (2017). Bayesian inference of species diffusion in the West African *Agama agama* species group (Reptilia, Agamidae). *Systematics and Biodiversity*, *15*, 192–203.

Leaché, A. D., Oaks, J. R., Ofori-Boateng, C., & Fujita, M. K. (2020). Comparative phylogeography of West African amphibians and reptiles. *Evolution*, *74*, 716–724.

Leaché, A. D., Portik, D. M., Rivera, D., Rödel, M.-O., Penner, J., Gvoždík, V., … & Fujita, M.K. (2019). Exploring rain forest diversification using demographic model testing in the African foam-nest treefrog *Chiromantis rufescens*. *Journal of Biogeography*, *46*, 2706–2721.

Lee, D. C., Halliday, A. N., Fitton, J. G., & Poli, G. (1994). Isotopic variations with distance and time in the volcanic islands of the Cameroon line: evidence for a mantle plume origin. *Earth and Planetary Science Letters*, *123*, 119–138.

Lester, S. E., Ruttenberg, B. I., Gaines, S. D., & Kinlan, B. P. (2007). The relationship between dispersal ability and geographic range size. *Ecology Letters*, *10*, 745–758.

Levin, D. A. (1975). Minority cytotype exclusion in local plant populations. *Taxon*, *24*, 35–43.

Levin, D. A. (2004). The ecological transition in speciation. *New Phytologist*, *161*, 91–96.

Levin, D. A. (2013). The timetable for allopolyploidy in flowering plants. *Annals of Botany*, *112*, 1201–1208.

Levinsky, I., Araújo, M. B., Nogués-Bravo, D., Haywood, A. M., Valdes, P. J., & Rahbek, C. (2013). Climate envelope models suggest spatio-temporal co-occurrence of refugia of African birds and mammals. *Global Ecology and Biogeography*, *22*, 351–363.

Levy, A. A., & Feldman, M. (2002). The impact of polyploidy on grass genome evolution. *Plant Physiology*, *130*, 1587–1593.

Lewin, A., Feldman, A., Bauer, A. M., Belmaker, J., Broadley, D. G., Chirio, L., ... & Meiri, S. (2016). Patterns of species richness, endemism and environmental gradients of African reptiles. *Journal of Biogeography*, *43*, 2380–2390.

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.

Lima, A., Harris, D. J., Rocha, S., Miralles, A., Glaw, F., & Vences, M. (2013). Phylogenetic relationships of *Trachylepis* skink species from Madagascar and the Seychelles (Squamata: Scincidae). *Molecular Phylogenetics and Evolution*, *67*, 615–620.

Linder, H. P. (1983). The historical phytogeography of the *Disinae* (Orchidaceae). *Bothalia*, *14*, 565–570.

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.

Litvinchuk, S. N., Mazepa, G. O., Pasynkova, R. A., Saidov, A., Satorov, T., Chikin, Y. A., Shabanov, D. A., Crottini, A., Borkin, L. J., Rosanov, J. M. & Stöck, M. (2011). Influence of environmental conditions on the distribution of Central Asian green toads with three ploidy levels. *Journal of Zoological Systematics and Evolutionary Research*, *49*, 233–239.

Livingstone, D. A. (1975). Late Quaternary climatic change in Africa. *Annual Review of Ecology and Systematics*, *6*, 249–280.

Loader, S. P., Pisani, D., Cotton, J. A., Gower, D. J., Day, J. J., & Wilkinson, M. (2007). Relative time scales reveal multiple origins of parallel disjunct distributions of African caecilian amphibians. *Biology Letters*, *3*, 505–508.

Lorenzen, E. D., Heller, R., & Siegismund, H. R. (2012). Comparative phylogeography of African et savanna ungulates. *Molecular Ecology*, *21*, 3656–3670.

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.

Mable, B. K. (2004). 'Why polyploidy is rarer in animals than in plants': myths and mechanisms. *Biological Journal of the Linnean Society*, *82*, 453–466.

Mable, B. K., Alexandrou, M. A., & Taylor, M. I. (2011). Genome duplication in amphibians and fish: An extended synthesis. *Journal of Zoology*, *284*, 151–182.

MacArthur, R.H. & Wilson, E.O. (1967). *The Theory of Island Biogeography*. Princeton, NJ: Princeton University Press.

Macgregor, D. (2015). History of the development of the East African Rift System: A series of interpreted maps through time. *Journal of African Earth Sciences*, *101*, 232–252.

Macgregor, H. C. (1993). *An Introduction to Animal Cytogenetics*. London: Chapman & Hall.

Maguire Jr, B. (1973). Niche response structure and the analytical potentials of its relationship to the habitat. *American Naturalist*, *107*, 213–246.

Maisels, F., & Forboseh, P. (1999). The Kilum/Ijim forest project: Biodiversity monitoring in the montane forests of Cameroon. *Bulletin of the African Bird Club*, *6*, 110-114.

Maley, J. (1987). Fragmentation de la forêt dense humide africaine et extension des biotopes montagnards au Quatemaire récent: Nouvelles données polliniques et chronologiques. Implications paléo-climatiques et biogéographiques. *Palaeoecology of Africa*, *18*, 307–334.

Maley, J. (1989). Late Quaternary climatic changes in the African rain forest: Forest refugia and the major role of sea surface temperature variations. In Leinen, M. & Sarnthein, M. (Eds.) *Paleoclimatology and Paleometeorology: Modern and Past Patterns of Global Atmospheric Transport* (pp. 585–616). Dordrecht, Netherlands: Springer.

Maley, J. (1991). The African rain forest vegetation and palaeoenvironments during late Quaternary. *Climatic Change*, *19*, 79–98.

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*, *104*, 31–73.

Maley, J. (1997). Middle to Late Holocene changes in tropical Africa and other continents: Paleomonsoon and sea surface temperature variations. In Weiss, H., Kukla, G., Dalfes, H. N. (Eds.) *Third millennium BC climate change and Old World collapse*. Berlin, Heidelberg: Springer.

Maley, J., & Brénac, P. (1987). Analyses polliniques préliminaires du Quaternaire récent de l'Ouest Cameroun: Mise en évidence de refuges forestiers et discussion des problèmes paléoclimatiques. *Mémoires et Travaux de l'Institut de Montpellier*, *17*, 129–142.

Maley, J., & Brenac, P. (1998). Vegetation dynamics, palaeoenvironments and climatic changes in the forests of western Cameroon during the last 28,000 years BP. *Review of Palaeobotany and Palynology*, *99*, 157–187.

Maley, J., & Livingstone, D. A. (1983). Extension d'un élément montagnard dans le Sud du Ghana (Afrique de l'Ouest) au Pléistocène supérieur et à l'Holocène inférieur: Premières données polliniques. *Comptes Rendus de l'Académie des Sciences*, *Paris*, *296*, 1287–1292.

Malhi, Y., Adu-Bredu, S., Asare, R. A., Lewis, S. L., & Mayaux, P. (2013). African rainforests: past, present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*, 20120312.

Mallet, J. (2007). Hybrid speciation. *Nature*, 446, 279–283.

Marchant, R., & Hooghiemstra, H. (2004). Rapid environmental change in African and South American tropics around 4000 years before present: A review. *Earth-Science Reviews*, *66*, 217–260.

Marchant, D. B, Soltis, D. E., & Soltis, P. S. (2016). Patterns of abiotic niche shifts in allopolyploids relative to their progenitors. *New Phytologist*, *212*, 708–718.

Marcussen, T., Heier, L., Brysting, A. K., Oxelman, B., & Jakobsen, K. S. (2015). From gene trees to a dated allopolyploid network: Insights from the angiosperm genus *Viola* (Violaceae). *Systematic Biology*, *64*, 84–101.

Marks, B. D. (2010). Are lowland rainforests really evolutionary museums? Phylogeography of the green hylia (*Hylia prasina*) in the Afrotropics. *Molecular Phylogenetics and Evolution*, *55*, 178–184.

Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., ... & Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research*, *23*, 1817–1828.

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

Matzke, N. J. (2013). Probabilistic historical biogeography: New models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers in Biogeography*, *5*, 242–248.

Matzke, N. J. (2014). Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology*, *63*, 951–970.

Mausfeld, P., Vences, M., Schmitz, A., & Veith, M. (2000). First data on the molecular phylogeography of scincid lizards of the genus *Mabuya*. *Molecular Phylogenetics and Evolution*, *17*, 11–14.

Mausfeld, P., & Vrcibradic, D. (2002). On the nomenclature of the skink (*Mabuya*) endemic to the Western Atlantic Archipelago of Fernando de Noronha, Brazil. *Journal of Herpetology*, *36*, 292–295.

Mayr, E. (1942). *Systematics and the Origin of Species from the Viewpoint of a Zoologist.* New York, NY: Columbia University Press.

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

McKinnon, G. E., Jordan, G. J., Vaillancourt, R. E., Steane, D. A., & Potts, B. M. (2004). Glacial refugia and reticulate evolution: The case of the Tasmanian eucalypts. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*, 275–284.

Measey, J. (2016). Overland movement in African clawed frogs (*Xenopus laevis*): A systematic review. *PeerJ*, *4*, e2474.

Measey, G. J., Vences, M., Drewes, R. C., Chiari, Y., Melo, M., & Bourles, B. (2007). Freshwater paths across the ocean: Molecular phylogeny of the frog *Ptychadena newtoni* gives insights into amphibian colonization of oceanic islands. *Journal of Biogeography*, *34*, 7–20.

Melo, M., Warren, B. H., & Jones, P. J. (2011). Rapid parallel evolution of aberrant traits in the diversification of the Gulf of Guinea white-eyes (Aves, Zosteropidae). *Molecular Ecology*, *20*, 4953–4967.

Menegon, M., Loader, S. P., Marsden, S. J., Branch, W. R., Davenport, T. R. B., & Ursenbacher, S. (2014). The genus *Atheris* (Serpentes: Viperidae) in East Africa: Phylogeny and the

role of rifting and climate in shaping the current pattern of species diversity. *Molecular Phylogenetics and Evolution*, *79*, 12–22.

Menke, S., Böhning-Gaese, K., & Schleuning, M. (2012). Plant–frugivore networks are less specialized and more robust at forest–farmland edges than in the interior of a tropical forest. *Oikos*, *121*, 1553–1566.

Meseguer, A. S., Aldasoro, J. J., & Sanmartín, I. (2013). Bayesian inference of phylogeny, morphology and range evolution reveals a complex evolutionary history in St. John's wort (*Hypericum*). *Molecular Phylogenetics and Evolution*, *67*, 379–403.

Meyers, J. B., Rosendahl, B. R., Harrison, C. G., & Ding, Z. D. (1998). Deep-imaging seismic and gravity results from the offshore Cameroon Volcanic Line, and speculation of African hotlines. *Tectonophysics*, *284*, 31–63.

Migliore, J., Lézine, A. M., & Hardy, O. J. (2020). The recent colonization history of the most widespread Podocarpus tree species in Afromontane forests. *Annals of Botany*, *126*, 73–83.

Mikula, O., Nicolas, V., Šumbera, R., Konečný, A., Denys, C., Verheyen, E., ... & Bryja, J. (2021). Nuclear phylogenomics, but not mitogenomics, resolves the most successful Late Miocene radiation of African mammals (Rodentia: Muridae: Arvicanthini). *Molecular Phylogenetics and Evolution*, *157*, 107069.

Miller, E. C., Sellas, A. B., & Drewes, R. C. (2012). A new species of *Hemidactylus* (Squamata: Gekkonidae) from Príncipe Island, Gulf of Guinea, West Africa with comments on the African-Atlantic clade of Hemidactylus geckos. *African Journal of Herpetology*, *61*, 40–57.

Mitchell, M. W., Locatelli, S., Ghobrial, L., Pokempner, A. A., Clee, P. R. S., Abwe, E. E., ... & Fotso, R. (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.

Mittermeier, R. A., Myers, N., & Mittermeier, C. G. (2000). *Hotspots: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions.* Chicago, IL: University of Chicago Press.

Mittermeier, R. A., Turner, W. R., Larsen, F. W., Brooks, T. M., & Gascon, C. (2011). Global biodiversity conservation: The critical role of hotspots. In Zachos, F. E. & Habel, J. C. (Eds.), *Biodiversity Hotspots – Distribution and Protection of Conservation Priority Areas*. (pp. 3–22). Berlin, Germany: Springer.

Mizerovská, D., Nicolas, V., Demos, T. C., Akaibe, D., Colyn, M., Denys, C., ... & Bryja, J. (2019). Genetic variation of the most abundant forest-dwelling rodents in Central Africa (P*raomys jacksoni* complex): Evidence for Pleistocene refugia in both montane and lowland forests. *Journal of Biogeography*, *46*, 1466–1478.

Molina-Henao, Y. F., & Hopkins, R. (2019). Autopolyploid lineage shows climatic niche expansion but not divergence in *Arabidopsis arenosa*. *American Journal of Botany*, *106*, 61–70.

Moreau, R. E. (1963). Vicissitudes of the African biomes in the late Pleistocene. *Proceedings of the Zoological Society of London*, *141*, 395–421.

Moreau, R. E. (1966). *The Birds Faunas of Africa and its Islands.* New York, NY: Academic Press.

Muller, H. J. (1925). Why polyploidy is rarer in animals than in plants. *American Naturalist*, *59*, 346–353.

Muniz, F. D. L., Campos, Z., Rangel, S. H., Martínez, J. G., Souza, B. C., De Thoisy, B., ... & Farias, I. P. (2018). Delimitation of evolutionary units in Cuvier's dwarf caiman, *Paleosuchus palpebrosus* (Cuvier, 1807): insights from conservation of a broadly distributed species. *Conservation Genetics*, *19*, 599–610.

Mutke, J., & Barthlott, W. (2005). Patterns of vascular plant diversity at continental to global scales. *Biologiske Skrifter*, *55*, 521–531.

Myers, N. (1988). Threatened biotas: "Hot spots" in tropical forests. *Environmentalist*, *8*, 187–208.

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

Nagasawa, K., Tanizaki, Y., Okui, T., Watarai, A., Ueda, S., & Kato, T. (2013). Significant modulation of the hepatic proteome induced by exposure to low temperature in *Xenopus laevis*. *Biology Open*, *2*, 1057–1069.

Nagy, Z. T., Kusamba, Z. C., Tungaluma, G. G., Lokasola, A. L., Kolby, J., & Kielgast, J. (2011). Foraging acrobatics of *Toxicodryas blandingii* in the Democratic Republic of the Congo. *Herpetology Notes*, *4*, 91–92.

Nakagawa, M. (2006). Ploidy, geographical distribution and morphological differentiation of *Parasenecio auriculata* (Senecioneae; Asteraceae) in Japan. *Journal of Plant Research*, *119*, 51–61.

Nash, H. C., Low, G. W., Choo, S. W., Chong, J. L., Semiadi, G., Hari, R., ... & Rheindt, F. E. (2018). Conservation genomics reveals possible illegal trade routes and admixture across pangolin lineages in Southeast Asia. *Conservation Genetics*, *19*, 1083–1095.

Nassar, F., Challita, M., Sadek, R., & Hraoui-Bloquet, S. (2013). Sexual dimorphism and female reproductive cycle in the scincid lizard *Trachylepis vittata* (Olivier, 1804) in Lebanon (Reptilia: Scincidae). *Zoology in the Middle East*, *59*, 297–301.

Nazareno, A. G., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: An empirical study from an Amazonian plant species. *Molecular Ecology Resources*, *17*, 1136–1147.

Nelson, M. R., Wegmann, D., Ehm, M. G., Kessner, D., Jean, P. S., Verzilli, C., ... & Mooser, V. (2012). An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science (Washington)*, *337*, 100–104.

Newmark, W. D., Mkongewa, V. J., Amundsen, D. L., & Welch, C. (2020). African sunbirds predominantly pollinate plants useful to humans. *The Condor*, *122*, duz070.

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2014). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*, 268–274.

Nicolas, V., Akpatou, B., Wendelen, W., Kerbis Peterhans, J.., 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., Jacquet, F., Hutterer, R., Konecný, A., Kouassi, S.K., Durnez, L. … & Denys, C. (2019). Multilocus phylogeny of the *Crocidura poensis* species complex (Mammalia,

Eulipotyphla): Influences of the palaeoclimate on its diversification and evolution. *Journal of Biogeography*, *46*, 871–883.

Nicolas, V., Missoup, A. D., Colyn, M., Cruaud, C., & Denys, C. (2012). West-Central African Pleistocene lowland forest evolution revealed by the phylogeography of Misonne's soft-furred mouse. *African Zoology*, *47*, 100–112.

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

Njabo, K. Y., & Sorenson, M. D. (2009). Origin of Bannerman's turaco *Tauraco bannermani* in relation to historical climate change and the distribution of West African montane forests. *Ostrich-Journal of African Ornithology*, *80*, 1–7.

Nosil, P., (2008). Speciation with gene flow could be common. *Molecular Ecology*, *17*, 2103–2106.

Nsor, C. A., Godsoe, W., & Chapman, H. M. (2019). Promiscuous pollinators—Evidence from an Afromontane sunbird–plant pollen transport network. *Biotropica*, *51*, 538–548.

Nogués-Bravo, D. (2009). Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography*, *18*, 521–531.

Nougué, O., Gallet, R., Chevin, L. M., & Lenormand, T. (2015). Niche limits of symbiotic gut microbiota constrain the salinity tolerance of brine shrimp. *American Naturalist*, *186*, 390–403.

Onadeko, A.B. & Rödel, M.-O. (2009) Anuran surveys in south-western Nigeria. *Salamandra*, *45*, 1–14.

Orr, H. A. (1990). "Why polyploidy is rarer in animals than in plants" revisited. *American Naturalist*, *136*, 759–770.

Otto, S. P., & Whitton, J. (2000). Polyploid incidence and evolution. *Annual Review of Genetics*, *34*, 401–437.

Otto, S. P. (2007). The evolutionary consequences of polyploidy. *Cell*, *131*, 452–462.

Otto, C. R., Snodgrass, J. W., Forester, D. C., Mitchell, J. C., & Miller, R. W. (2007). Climatic variation and the distribution of an amphibian polyploid complex. *Journal of Animal Ecology*, *76*, 1053–1061.

Owens, H. L., Campbell, L. P., Dornak, L. L., Saupe, E. E., Barve, N., Soberón, J., Ingenloff, K., Lira-Noriega, A., Hensz, C. M., Myers, C. E., & Peterson, A. T. (2013). Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling*, *263*, 10–18.

Owens, H. L., Ribeiro, V., Saupe, E. E., Cobos, M. E., Hosner, P. A., Cooper, J. C., Samy, A. M., Barve, V., Barve, N., Muñoz-R., C. J., Peterson, A. T. (2020). Reconstructing ecological niche evolution via ancestral state reconstruction with uncertainty incorporated. *Ecology and Evolution*, *10*, 6967–6977.

Parent, C. E., Caccone, A., & Petren, K. (2008). Colonization and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*, 3347–3361.

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for stacks. *Methods in Ecology and Evolution*, *8*, 1360–1373.

Patton, J. L., & Silva, M. F. (2005). The history of Amazonian mammals: Mechanisms and timing of diversification. In Bermingham, E., Dick, C.W., & Moritz, C. (Eds.) *Tropical*

rainforests: Past, present and future. (pp. 107–126). Chicago, IL: The University of Chicago Press.

Paul, J. D., Roberts, G. G., & White, N. (2014). The African landscape through space and time. *Tectonics*, *33*, 898–935.

Peers, M. J., Thornton, D. H., & Murray, D. L. (2013). Evidence for large-scale effects of competition: Niche displacement in Canada lynx and bobcat. *Proceedings of the Royal Society B*, *280*, 20132495.

Penner, J., Augustin, M., & Rödel, M.-O. (2019) Modelling the spatial baseline for amphibian conservation in West Africa. *Acta Oecologica, 84*, 31–40.

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

Peter, B. M., & Slatkin, M. (2013). Detecting range expansions from genetic data. *Evolution*, *67*, 3274–3289.

Peter, B. M., & Slatkin, M. (2015). The effective founder effect in a spatially expanding population. *Evolution*, *69*, 721–734.

Peterson, A. T. (2011). Ecological niche conservatism: A time-structured review of evidence. *Journal of Biogeography*, *38*, 817–827.

Peterson, A. T., & Ammann, C. M. (2013). Global patterns of connectivity and isolation of populations of forest bird species in the late Pleistocene. *Global Ecology and Biogeography*, *22*, 596–606.

Peterson, A.T., Papes, ¸M., & Soberón, J. (2008). Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, *213*, 63–72.

Peterson, A. T., Soberón, J., & Sánchez-Cordero, V. (1999). Conservatism of ecological niches in evolutionary time. *Science*, *285*, 1265–1267.

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.

Pianka, E. R. (1966). Latitudinal gradients in species diversity: a review of concepts. *American Naturalist*, *100*, 33–46.

Piñeiro, R., Dauby, G., Kaymak, E., & Hardy, O. J. (2017). Pleistocene population expansions of shade-tolerant trees indicate fragmentation of the African rainforest during the Ice Ages. *Proceedings of the Royal Society B: Biological Sciences*, *284*, 20171800.

Pitteloud, C., Arrigo, N., Suchan, T., Mastretta-Yanes, A., Vila, R., Dincă, V., ... & Alvarez, N. (2017). Climatic niche evolution is faster in sympatric than allopatric lineages of the butterfly genus *Pyrgus*. *Proceedings of the Royal Society B*, *284*, 20170208.

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

Pokorny, L., Riina, R., Mairal, M., Meseguer, A. S., Culshaw, V., Cendoya, J., … & Sanmartín, I. (2015). Living on the edge: Timing of Rand Flora disjunctions congruent with ongoing aridification in Africa. *Frontiers in Genetics*, *6*, 154.

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, Á., & Lareu, M. (2013). An overview of STRUCTURE: applications, parameter settings, and supporting software. *Frontiers in Genetics*, *4*, 98.

Portik, D. M., Leaché, A. D., Rivera, D., Barej, M. F., Burger, M., Hirschfeld, M., …& Fujita, M. K. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Molecular Ecology*, *26*, 5245–5263.

Portillo, F., Greenbaum, E., Menegon, M., Kusamba, C., & Dehling, J. M. (2015). Phylogeography and species boundaries of *Leptopelis* (Anura: Arthroleptidae) from the Albertine Rift. *Molecular Phylogenetics and Evolution*, *82*, 75–86.

Portillo, F., Stanley, E. L., Branch, W. R., Conradie, W., Rödel, M.-O., Penner, J., ... Eli Greenbaum (2019). Evolutionary history of burrowing asps (Lamprophiidae: Atractaspidinae) with emphasis on fang evolution and prey selection. *PLoS ONE*, *14*, e0214889.

Poynton, J. C. (1999). Distribution of amphibians in sub-Saharan Africa, Madagascar, and Seychelles. In Duellman, W. D. (Ed.). *Patterns of Distribution of Amphibians, a Global Perspective.* Baltimore, MD: Johns Hopkins University Press.

Prance, G.T. (1982). *Biological Diversification in the Tropics*. New York, NY: Columbia University Press.

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

Prigogine, A., (1987). Disjunctions of montane forest birds in the Afrotropical Region. *Bonner zoologische Beiträge, 38*, 195–207.

Pudlo, P., Marin, J. M., Estoup, A., Cornuet, J. M., Gautier, M., & Robert, C. P. (2015). Reliable ABC model choice via random forests. *Bioinformatics*, *32*, 859–866.

Puechmaille, S. J. (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: Subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, *16*, 608–627.

Pulliam, H. R. (2000). On the relationship between niche and distribution. *Ecology Letters*, *3*, 349–361.

Pyron, R. A., Burbrink, F. T., & Wiens, J. J. (2013). A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evolutionary Biology*, *13*, 93.

QGIS Development Team (2020). QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org.

Qu, W. M., Liang, N., Wu, Z. K., Zhao, Y. G., & Chu, D. (2020). Minimum sample sizes for invasion genomics: Empirical investigation in an invasive whitefly. *Ecology and Evolution*, *10*, 38–49.

Rahbek, C., Borregaard, M. K., Antonelli, A., Colwell, R. K., Holt, B. G., Nogues-Bravo, D., ... & Fjeldså, J. (2019). Building mountain biodiversity: Geological and evolutionary processes. *Science*, *365*, 1114–1119.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*, 901–904.

Ramsey, J. (2011). Polyploidy and ecological adaptation in wild yarrow. *Proceedings of the National Academy of Sciences USA*, *108*, 7096–7101.

Rannala, B., & Yang, Z. (2013). Improved reversible jump algorithms for Bayesian species delimitation. *Genetics*, *194*, 245–253.

Ravelo, A. C., Andreasen, D. H., Lyle, M., Lyle, A. O., & Wara, M. W. (2004). Regional climate shifts caused by gradual global cooling in the Pliocene epoch. *Nature*, *429*, 263–267.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ree, R. H., & Sanmartín, I. (2018). Conceptual and statistical problems with the DEC+ J model of founder-event speciation and its comparison with DEC via model selection. *Journal of Biogeography*, *45*, 741–749.

Ree, R. H., & Smith, S. A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, *57*, 4–14.

Reilly, S. B., Stubbs, A. L., Karin, B. R., Bi, K., Arida, E., Iskandar, D. T., & McGuire, J. A. (2019). Leap-frog dispersal and mitochondrial introgression: Phylogenomics and biogeography of *Limnonectes* fanged frogs in the Lesser Sundas Archipelago of Wallacea. *Journal of Biogeography*, *46*, 757–769.

Reichenow, A., (1900). *Die Vögel Afrikas*. Neudamm, Germany: J. Neumann.

Renner, S. (2004). Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences*, *165*, S23–S33.

Reusch, A. M., Nyblade, A. A., Wiens, D. A., Shore, P. J., Ateba, B., Tabod, C. T., & Nnange, J. M. (2010). Upper mantle structure beneath Cameroon from body wave tomography and the origin of the Cameroon Volcanic Line. *Geochemistry, Geophysics, Geosystems*, *11*, Q10W07.

Reyes-Velasco, J., Manthey, J. D., Freilich, X., & Boissinot, S. (2018). Diversification of African tree frogs (genus *Leptopelis*) in the highlands of Ethiopia. *Molecular Ecology*, *27*, 2256–2270.

Richards, P. W. (1963). Ecological notes on West African vegetation. II. Lowland forest of the southern Bakundu Forest Reserve. *Journal of Ecology*, *51*, 123–149.

Richardson, J. E., Pennington, R. T., Pennington, T. D., & Hollingsworth, P. M. (2001). Recent and rapid diversification of a species-rich genus of Neotropical trees. *Science (Washington)*, *293*, 2242–2245.

Richardson, P. L., & Walsh, D. (1986). Mapping climatological seasonal variations of surface currents in the tropical Atlantic using ship drifts. *Journal of Geophysical Research: Oceans*, *91*, 10537–10550.

Rietkerk, M., Ketner, P., & De Wilde, J. (1995). Caesalpinioideae and the study of forest refuges in Gabon: Preliminary results. *Bulletin du Museum National d'Histoire Naturelle, Paris, Adansonia*, *17*, 95–105.

Robbrecht, E. (1996). Geography of African Rubiaceae with reference to glacial rain forest refuges. In van der Maesen, L. J. G., van der Burgt, X. M. & van Medenbach de Rooy, J.M. (Eds.), *The Biodiversity of African Plants*. (pp. 564–581). Dordrecht, Netherlands: Kluwer.

Rocha, S., Harris, D. J., & Carretero, M. (2010). Genetic diversity and phylogenetic relationships of Mabuya spp. (Squamata: Scincidae) from western Indian Ocean islands. *Amphibia-Reptilia*, *31*, 375–385.

Rödel, M.-O., Emmrich, M., Penner, J., Schmitz, A. & Barej, M. F. (2014). The taxonomic status of two West African *Leptopelis* species: *L. macrotis* Schiøtz, 1967 and *L. spiritusnoctis* Rödel, 2007 (Amphibia: Anura: Arthroleptidae). *Zoosystematics and Evolution*, *90,* 21–32.

Roderick, G. K., & Gillespie, R. G. (1998). Speciation and phylogeography of Hawaiian terrestrial arthropods. *Molecular Ecology*, *7*, 519–531.

Rodrigues, J. F. M., Villalobos, F., Iverson, J. B., & Diniz-Filho, J. A. F. (2019). Climatic niche evolution in turtles is characterized by phylogenetic conservatism for both aquatic and terrestrial species. *Journal of Evolutionary Biology*, *32*, 66–75.

Rohde, K. (1997). The larger area of the tropics does not explain latitudinal gradients in species diversity. *Oikos*, *79*, 169–172.

Ronquist, F. (1997). Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Systematic Biology*, *46*, 195–203.

RStudio Team (2018). RStudio: Integrated Development for R. RStudio, Inc., Boston, http://www.rstudio.com/.

Salzmann, U., & Hoelzmann, P. (2005). The Dahomey Gap: An abrupt climatically induced rain forest fragmentation in West Africa during the late Holocene. *The Holocene*, *15*, 190–199.

Saupe, E. E., Barve, V., Myers, C. E., Soberón, J., Barve, N., Hensz, C. M., Peterson, A.T., Owens, H. L. & Lira-Noriega, A. (2012). Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling*, *237*, 11–22.

Saupe, E. E., Barve, N., Owens, H. L., Cooper, J. C., Hosner, P. A., & Peterson, A. T. (2018). Reconstructing ecological niche evolution when niches are incompletely characterized. *Systematic Biology*, *67*, 428–438.

Schiesari, L., Zuanon, J., Azevedo-Ramos, C., Garcia, M., Gordo, M., Messias, M., & Vieira, E. M. (2003). Macrophyte rafts as dispersal vectors for fishes and amphibians in the Lower Solimões River, Central Amazon. *Journal of Tropical Ecology*, *19*, 333–336.

Schmid, M., Evans, B. J., & Bogart, J. P. (2015). Polyploidy in amphibia. *Cytogenetic and Genome Research*, *145*, 315–330.

Schmidt, K. P. (1923). Contributions to the herpetology of the Belgian Congo based on the collection of the American Museum Congo Expedition, 1909–1915. Part II. Snakes, with field notes by Herbert Lang and James P. Chapin. *Bulletin of the American Museum of Natural History, 49*, 1–146

Schmitt, C. B., Denich, M., Demissew, S., Friis, I., & Boehmer, H. J. (2010). Floristic diversity in fragmented Afromontane rainforests: Altitudinal variation and conservation importance. *Applied Vegetation Science*, *13*, 291–304.

Ségalen, L., Lee-Thorp, J. A., & Cerling, T. (2007). Timing of C4 grass expansion across sub-Saharan Africa. *Journal of Human Evolution*, *53*, 549–559.

Selmecki, A. M., Maruvka, Y. E., Richmond, P. A., Guillet, M., Shoresh, N., Sorenson, A. L., ... & Pellman, D. (2015). Polyploidy can drive rapid adaptation in yeast. *Nature*, *519*, 349–352.

Senut, B., Pickford, M., & Ségalen, L. (2009). Neogene desertification of Africa. *Comptes Rendus Geoscience*, *341*, 591–602.

Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tiercelin, J. J., & Brunet, M. (2006). Tectonic uplift and Eastern Africa aridification. *Science*, *313*, 1419–1423.

Serle, W. (1964). The lower altitudinal limit of the montane forest birds of the Cameroon Mountain, West Africa. *Bulletin of the British Ornithologists' Club*, *84*, 87–91.

Session, A. M., Uno, Y., Kwon, T., Chapman, J. A., Toyoda, A., Takahashi, S., ... & Rokhsar, D. S. (2016). Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature*, *538*, 336–343.

Shafer, A. B., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, J. B. (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, *8*, 907–917.

Shapiro, H. A., & Zwarenstein, H. (1934). A rapid test for pregnancy of *Xenopus laevis*. *Nature*, *133*, 762.

Sharma, S., Walker, S. C., & Jackson, D. A. (2008). Empirical modelling of lake water-temperature relationships: A comparison of approaches. *Freshwater Biology*, *53*, 897–911.

Sharpe, R. B. (1893). *On the Zoo-geographical Areas of the World, Illustrating the Distribution of Birds.* London, UK: Henderson & Co.

Shaw, K. L., & Gillespie, R. G. (2016). Comparative phylogeography of oceanic archipelagos: Hotspots for inferences of evolutionary process. *Proceedings of the National Academy of Sciences*, *113*, 7986–7993.

Shirley, M. H., Vliet, K. A., Carr, A. N., & Austin, J. D. (2014). Rigorous approaches to species delimitation have significant implications for African crocodilian systematics and conservation. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20132483.

Sick, H. (1967). Rios e enchentes na Amazônia como obstáculo para a avifauna**.** *Atas do Simpósio Sobre a Biota Amazônica*, *5*, 495–520.

Siddig, A. A. (2019). Why is biodiversity data-deficiency an ongoing conservation dilemma in Africa? *Journal for Nature Conservation*, *50*, 125719.

Sigel, E. M. (2016). Genetic and genomic aspects of hybridization in ferns. *Journal of Systematics and Evolution*, *54*, 638–655.

Silva, S. M., Peterson, A. T., Carneiro, L., Burlamaqui, T. C. T., Ribas, C. C., Sousa-Neves, T., ... & Aleixo, A. (2019). A dynamic continental moisture gradient drove Amazonian bird diversification. *Science Advances*, *5*, eaat5752.

Simberloff, D., & Rejmánek, M. (eds.). (2011). *Encyclopedia of Biological Invasions* (No. 3). Berkeley, CA; University of California Press.

Simon, B., Guillocheau, F., Robin, C., Dauteuil, O., Nalpas, T., Pickford, M., ... & Bez, M. (2017). Deformation and sedimentary evolution of the Lake Albert Rift (Uganda, East African rift system). *Marine and Petroleum Geology*, *86*, 17–37.

Simpson, G. G. (1984). *Tempo and Mode in Evolution*. New York, NY; Columbia University Press.

Slager, D. L., Epperly, K. L., Ha, R. R., Rohwer, S., Wood, C., Van Hemert, C., & Klicka, J. (2020). Cryptic and extensive hybridization between ancient lineages of American crows. *Molecular Ecology*, *29*, 956–969.

Šmíd, J., Mazuch, T., Nováková, L., Modrý, D., Malonza, P. K., Elmi, H. S. A., ... & Moravec, J. (2019). Phylogeny and systematic revision of the gecko genus *Hemidactylus* from the Horn of Africa (Squamata: Gekkonidae). *Herpetological Monographs*, *33*, 26–47.

Smith, M. L., & Carstens, B. C. (2020). Process-based species delimitation leads to identification of more biologically relevant species. *Evolution*, *74*, 216–229.

Smith, M. L., Ruffley, M., Tank, D. C., Sullivan, J., & Carstens, B. C. (2017). Demographic model selection using random forests and the site frequency spectrum. *Molecular Ecology*, *26*, 4562–4573.

Smolensky, N. L., Hurtado, L. A., & Fitzgerald, L. A. (2015). DNA barcoding of Cameroon samples enhances our knowledge on the distributional limits of putative species of *Osteolaemus* (African dwarf crocodiles). *Conservation Genetics*, *16*, 235–240.

Soares, L. B., Ceríaco, L. M. P., Marques, M. P., Bastos- Silveira, C., Scheinberg, L. A., Harris, D. J., Brehm, A. & Jesus, J. A. (2018). Review of the leaf-litter skinks (Scincidae: *Panaspis*) from the Gulf of Guinea oceanic islands, with the description of a new species. *African Journal of Herpetology*, *67*, 132–159.

Soberón, J., & Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, *2*, 1–10.

Solís-Lemus, C., Bastide, P., & Ané, C. (2017). PhyloNetworks: A package for phylogenetic networks. *Molecular Biology and Evolution*, *34*, 3292–3298.

Soltis, P. S. & Soltis, D. E. (2000). The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences USA, 97*, 7051–7057.

Sosef, M. S. M. (1991). New species of *Begonia* in Africa and their relevance to the study of glacial rain forest refuges. *Wageningen Agricultural University Papers*, *91*, 120–151.

Sosef, M. S., Dauby, G., Blach-Overgaard, A., van der Burgt, X., Catarino, L., Damen, T., ... & Couvreur, T. L. (2017). Exploring the floristic diversity of tropical Africa. *BMC Biology*, *15*, 1–23.

Sousa, A. C. A. D. (2017). *Phylogeography, ecology and conservation of skink Adamastor, Trachylepis Adamastor Ceríaco, 2015* (Master's Thesis, Universidade de Évora, Évora, Portugal). Retrieved from: https://dspace.uevora.pt

Southgate, M. W., Patel, N. R., & Barrington, D. S. (2019). Ecological outcome of allopolyploidy in *Adiantum* (Pteridaceae): Niche intermediacy and expansion into novel habitats. *Rhodora*, *121*, 108–135.

Sowunmi, M. A. (1991). Late Quaternary environments in equatorial Africa: Palynological evidence. *Palaeoecology of Africa*, *22*, 213–238.

Spawls, S., Howell, K., Hinkel, H., & Menegon, M. (2018). *Field Guide to East African Reptiles*. New York, NY: Bloomsbury Publishing.

Stange, M., Sánchez-Villagra, M. R., Salzburger, W., & Matschiner, M. (2018). Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfishes (Ariidae) supports Miocene closure of the Panamanian Isthmus. *Systematic Biology*, *67*, 681–699.

Stankiewicz, J., & de Wit, M. J. (2006). A proposed drainage evolution model for Central Africa—Did the Congo flow east? *Journal of African Earth Sciences*, *44*, 75–84.

Stebbins, G.L. (1971). *Processes of Organic Evolution (Second Edition)*, Englewood Cliffs, NJ: Prentice-Hall.

Stevens, G. C. (1989). The latitudinal gradient in geographical range: how so many species coexist in the tropics. *American naturalist*, *133*, 240–256.

Stöck, M., Ustinova, J., Lamatsch, D. K., Schartl, M., Perrin, N., & Moritz, C. (2010). A vertebrate reproductive system involving three ploidy levels: Hybrid origin of triploids in

a contact zone of diploid and tetraploid Palearctic green toads (*Bufo viridis* subgroup). *Evolution*, *64*, 944–959.

Subramanian, S. (2016). The effects of sample size on population genomic analyses–implications for the tests of neutrality. *BMC Genomics*, *17*, 1–13.

Sukumaran, J., & Holder, M. T. (2010). DendroPy: A Python library for phylogenetic computing. *Bioinformatics*, *26*, 1569–1571.

Swofford, D. L. (2003). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sunderland, MA: Sinauer Associates.

Takemoto, H., Kawamoto, Y., & Furuichi, T. (2015). How did bonobos come to range south of the Congo River? Reconsideration of the divergence of *Pan paniscus* from other *Pan* populations. *Evolutionary Anthropology: Issues, News, and Reviews*, *24*, 170–184.

Taylor, P. J., Maree, S., Cotterill, F. P., Missoup, A. D., Nicolas, V., & Denys, C. (2014). Molecular and morphological evidence for a Pleistocene radiation of laminate-toothed rats (*Otomys*: Rodentia) across a volcanic archipelago in equatorial Africa. *Biological Journal of the Linnean Society*, *113*, 320–344.

Telfer, P. T., Souquiere, 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.

Terhorst, J., & Song, Y. S. (2015). Fundamental limits on the accuracy of demographic inference based on the sample frequency spectrum. *Proceedings of the National Academy of Sciences, USA*, *112*, 7677– 7682.

Tesfaye, D., Fashing, P. J., Bekele, A., Mekonnen, A., & Atickem, A. (2013). Ecological flexibility in Boutourlini's blue monkeys (*Cercopithecus mitis boutourlinii*) in Jibat

Forest, Ethiopia: A comparison of habitat use, ranging behavior, and diet in intact and fragmented forest. *International Journal of Primatology*, *34*, 615–640.

Theodoridis, S., Randin, C., Broennimann, O., Patsiou, T., & Conti, E. (2013). Divergent and narrower climatic niches characterize polyploid species of European primroses in *Primula* sect. *Aleuritia*. *Journal of Biogeography*, *40*, 1278–1289.

Thompson, K. A., Husband, B. C., & Maherali, H. (2014). Climatic niche differences between diploid and tetraploid cytotypes of *Chamerion angustifolium* (Onagraceae). *American Journal of Botany*, *101*, 1868–1875.

Ting, N. (2008). Mitochondrial relationships and divergence dates of the African colobines: Evidence of Miocene origins for the living colobus monkeys. *Journal of Human Evolution*, *55*, 312–325.

Tolley, K. A., Alexander, G. J., Branch, W. R., Bowles, P., & Maritz, B. (2016). Conservation status and threats for African reptiles. *Biological Conservation*, *204*, 63–71.

Tolley, K. A., Townsend, T. M., & Vences, M. (2013). Large-scale phylogeny of chameleons suggests African origins and Eocene diversification. *Proceedings of the Royal Society B*, *280*, 20130184.

Tosi, A. J. (2008). Forest monkeys and Pleistocene refugia: A phylogeographic window onto the disjunct distribution of the *Chlorocebus lhoesti* species group. *Zoological Journal of the Linnean Society*, *154*, 408–418.

Travers, S. L., Jackman, T. R., & Bauer, A. M. (2014). A molecular phylogeny of Afromontane dwarf geckos (*Lygodactylus*) reveals a single radiation and increased species diversity in a South African montane center of endemism. *Molecular Phylogenetics and Evolution*, *80*, 31–42.

Tzedakis, P. C., Emerson, B. C., & Hewitt, G. M. (2013). Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution*, *28*, 696–704.

Urban, E. K., Fry, C. H., & Keith, S. (1997). *The Birds of Africa*. *Vol. V.* Princeton, NJ: Princeton University Press.

Urban, M. C., Richardson, J. L., & Freidenfelds, N. A. (2014). Plasticity and genetic adaptation mediate amphibian and reptile responses to climate change. *Evolutionary Applications*, *7*, 88–103.

Uyeda, J.C., Drewes, R.C. & Zimkus, B.M. (2007) The California Academy of Sciences Gulf of Guinea Expeditions (2001, 2006) VI. A new species of *Phrynobatrachus* from the Gulf of Guinea Islands and a reanalysis of *Phrynobatrachus dispar* and *P. feae* (Anura: Phrynobatrachidae). *Proceedings of the California Academy of Sciences*, *58*, 367–385.

Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M. B., Balaguer, L., Benito-Garzón, M., ... & Zavala, M. A. (2014). The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecology Letters*, *17*, 1351–1364.

Van de Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nature Reviews Genetics*, *18*, 411–424.

Van de Peer, Y., Ashman, T. L., Soltis, P. S., & Soltis, D. E. (2021). Polyploidy: an evolutionary and ecological force in stressful times. *The Plant Cell*, *33*, 11–26.

Van de Perre, F., Leirs, H., & Verheyen, E. (2019). Paleoclimate, ecoregion size, and degree of isolation explain regional biodiversity differences among terrestrial vertebrates within the Congo Basin. *Belgian Journal of Zoology*, *149*, 23–42.

Van Zinderen Bakker, E. M., & Clark, J. D. (1962). Pleistocene climates and cultures in north-eastern Angola. *Nature*, *196*, 639–642.

Van Zinderen Bakker, E. M., & Coetzee, J. A. (1972). A re-appraisal of late-Quaternary climatic evidence from tropical Africa. *Palaeoecology of Africa*, *7*, 151–181.

Vanzolini, P.E. & Williams, E.E. (1970). South American anoles: geographic differentiation and evolution of the *Anolis chrysolepis* species group (Sauria, Iguanidae). *Arquivos de Zoologia, São Paulo*, *19*, 291–298.

Vanzolini, P.E. (1973). Paleoclimates, relief, and species multiplication in equatorial forests. In Meggers, B. J., Ayensu, E. S., & Duckworth, W. D. (Eds.) *Tropical Forest Ecosystems in Africa and South America: A Comparative Review*. (pp. 255–258). Washington D. C.: Smithsonian Press.

Vaz da Silva, B. (2015). *Evolutionary History of the Birds of the Angolan Highlands–The Missing Piece to Understand the Biogeography of the Afromontane Forests* (Master's Thesis). University of Porto, Porto, Portugal.

Vences, M., & Wake, D. B. (2007). Speciation, species boundaries and phylogeography of amphibians. *Amphibian Biology*, *7*, 2613–2671.

Visger, C. J., Germain-Aubrey, C. C., Patel, M., Sessa, E. B., Soltis, P. S., & Soltis, D. E. (2016). Niche divergence between diploid and autotetraploid *Tolmiea*. *American Journal of Botany*, *103*, 1396–1406.

Vitorino, L. C., Lima-Ribeiro, M. S., Terribile, L. C., & Collevatti, R. G. (2016). Demographical history and palaeodistribution modelling show range shift towards Amazon Basin for a Neotropical tree species in the LGM. *BMC Evolutionary Biology*, *16*, 213–228.

Wagner, W. L., & Funk, V. A. (eds). (1995). *Hawaiian biogeography. Evolution on a hot spot archipelago*. Smithsonian Institution, Washington, DC.

Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. *Ecological Applications*, *21*, 335–342.

Wilkinson, M., Loader, S. P., Gower, D. J., Sheps, J. A., & Cohen, B. L. (2003). Phylogenetic relationships of African caecilians (Amphibia: Gymnophiona): insights from mitochondrial rRNA gene sequences. *African Journal of Herpetology*, *52*, 83–92.

Wallace, A. R. (1853). *A Narrative of Travels on the Amazon and Rio Negro*. London, UK: Reeve.

Wallace, A. R. (1878). *Tropical nature, and other essays*. London, UK: Macmillan and Co.

Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. *Ecological Applications*, *21*, 335–342.

Weir, J. T. (2006). Divergent timing and patterns of species accumulation in lowland and highland neotropical birds. *Evolution*, *60*, 842–855.

Whittaker, R. J., Triantis, K. A., & Ladle, R. J. (2008). A general dynamic theory of oceanic island biogeography. *Journal of Biogeography*, *35*, 977–994.

White, F. (1979). The Guineo-Congolian Region and its relationships to other phytochoria. *Bulletin du Jardin botanique national de Belgique/Bulletin van de Nationale Plantentuin van Belgie*, *49*, 11–55.

White, F. (1981). The history of the Afromontane archipelago and the scientific need for its conservation. *African Journal of Ecology*, *19*, 33–54.

Wieczorek, A. M., Drewes, R. C., & Channing, A. (2000). Biogeography and evolutionary history of *Hyperolius* species: application of molecular phylogeny. *Journal of Biogeography*, *27*, 1231–1243.

Wieringa, J. J. (1999). Monopetalanthus exit. A systematic study of Aphanocalyx, Bikinia, Icuria, Michelsonia and Tetraberlinia (Leguminosae, Caesalpinioideae). *Wageningen Agricultural University Papers*, 320p.

Weinell, J., Barley, A. J., Siler, C. D., Orlov, N. L., Ananjeva, N. B., Oaks, J. … Brown, R. M. (2020). Phylogenetic relationships and biogeographic range evolution in Cat-eyed Snakes *Boiga* (Serpentes: Colubridae). *Zoological Journal of the Linnaean Society*, *2020*, 1–16.

Wüster, W., Chirio, L., Trape, J. F., Ineich, I., Jackson, K., Greenbaum, E., ... Hall, C. (2018). Integration of nuclear and mitochondrial gene sequences and morphology reveals unexpected diversity in the forest cobra (*Naja melanoleuca*) species complex in Central and West Africa (Serpentes: Elapidae). *Zootaxa*, *4455*, 68–98.

Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, *61*, 854–865.

Yang, Z. & Flouri, T. (2020) BP&P version 4.2 documentation. https://github.com/bpp/bpp

Yang, S. F., Komaki, S., Brown, R. M., & Lin, S. M. (2018). Riding the Kuroshio Current: Stepping stone dispersal of the Okinawa tree lizard across the East Asian Island Arc. *Journal of Biogeography*, *45*, 37–50.

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences*, *USA*, *107*, 9264–9269.

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.

Zeisset, I., & Beebee, T. J. C. (2008). Amphibian phylogeography: A model for understanding historical aspects of species distributions. *Heredity*, *101*, 109–119.

Zimkus, B. M. (2009). Biogeographical analysis of Cameroonian puddle frogs and description of a new species of *Phrynobatrachus* (Anura: Phrynobatrachidae) endemic to Mount Oku, Cameroon. *Zoological Journal of the Linnean Society*, *157*, 795–813.

Zimkus, B. M., & Gvoždík, V. (2013). Sky Islands of the Cameroon Volcanic Line: a diversification hot spot for puddle frogs (Phrynobatrachidae: *Phrynobatrachus*). *Zoologica Scripta*, *42*, 591–611.

Zimkus, B. M., Lawson, L. P., Barej, M. F., Barratt, C. D., Channing, A., Dash, K. M., ... & Gvoždík, V. (2017). Leapfrogging into new territory: How Mascarene ridged frogs diversified across Africa and Madagascar to maintain their ecological niche. *Molecular Phylogenetics and Evolution*, *106*, 254–269.

Zimkus, B. M., Rödel, M. O., & Hillers, A. (2010). Complex patterns of continental speciation: molecular phylogenetics and biogeography of sub-Saharan puddle frogs (*Phrynobatrachus*). *Molecular Phylogenetics and Evolution*, *55*, 883–900.

Zwickl, D. J., & Hillis, D. M. (2002). Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology*, *51*, 588–598.

APPENDIX I

From:

CHAPTER 2

Rivers, not refugia, drove diversification in arboreal, sub-Saharan African snakes

Allen, K.E., Greenbaum, E., Hime, P.M., Tapondjou N., W.P., Sterkhova, V.V., Kusamba, C., Rödel, M.-O., Penner, J., Peterson, A.T., Brown, R.M. 2021. Rivers, not refugia, drove diversification in arboreal, sub-Saharan African snakes. Ecology and Evolution. 11, 6133–6152.

Table S1. Museum accession numbers, GenBank accession IDs, and sampling localities for all specimens used in this study.

Museum Number	**Species**	**Latitude**	**Longitude**	**c-mos**	**cyt *b***
LSUMZ 20224	*Toxicodryas blandingii*	7.861808	-1.593136	MW655852	MW655869
PB11 640	*Toxicodryas blandingii*	10.3801	-9.300244	MW655850	MW655870
CAS 253611	*Toxicodryas blandingii*	4.4732	11.9648889	MW655847	MW655868
UTEP 22193	*Toxicodryas blandingii*	-2.58617	16.47339	MW655849	MW655872
UTEP 22195	*Toxicodryas blandingii*	-1.878	28.4524	MW655848	MW655871
UTEP 22194	*Toxicodryas blandingii*	-4.71209	19.4802	MW655851	MW655873
CAS 258155	*Toxicodryas pulverulenta*	-2.2419	13.58746	MW655834	MW655856
MCZ 187704	*Toxicodryas pulverulenta*	0.51624	12.79457	MW655836	MW655858
USNM 584253	*Toxicodryas pulverulenta*	-2.612499	13.615022	MW655845	N/A
CAS 254591	*Toxicodryas pulverulenta*	-1.10775	10.02691	MW655853	MW655854
CAS 253375	*Toxicodryas pulverulenta*	2.6106	14.0234	MW655833	MW655855
CAS 258156	*Toxicodryas pulverulenta*	-2.2419	13.58746	MW655846	MW655857
ZMB 86412	*Toxicodryas pulverulenta*	7.755777	-8.814877	MW655843	MW655867
ZMB 86410	*Toxicodryas pulverulenta*	7.575188	-9.242767	MW655842	MW655864
ZMB 86414	*Toxicodryas pulverulenta*	7.585033	-9.224848	MW655844	MW655866
ZMB 86384	*Toxicodryas pulverulenta*	7.755777	-8.814877	MW655841	MW655865
UTEP 22205	*Toxicodryas pulverulenta*	-1.14679	20.80005	MW655835	MW655863
UTEP 22200	*Toxicodryas pulverulenta*	-4.16473	28.16529	MW655838	MW655860
UTEP 22197	*Toxicodryas pulverulenta*	-3.24329	24.24856	MW655837	MW655859
UTEP 22201	*Toxicodryas pulverulenta*	-4.08432	28.15373	MW655839	MW655862
UTEP 22203	*Toxicodryas pulverulenta*	-0.44791	18.1346	MW655840	MW655861
	Dasypeltis atra			AF471136	AF471065
	Dasypeltis scabra			AY611945	AY612036
	Contia tenuis			AF471134	AF471095
	Heterodon simus			AF471142	AF217840
	Hemachatus haemachatus			N/A	AF217821
	Atheris nietschei			AF471125	AF471070
	Naja Kaouthia			AY058938	AF217835
	Crotaphopeltis hoetamboeia			AY611973	AY612064
	Crotaphopeltis tornieri			AF471112	AF471093
	Dasypeltis medici			AY611990	AY612081
	Dasypeltis fasciata			KX660328	KX660463
	Dipsadoboa unicolor			AF471139	AF471062

Table S2. DelimitR confusion matrices. Red indicates the number out of 100,000 simulated data sets that were correctly classified by the random forest classifiers, orange and grey indicate the number of models that were incorrectly classified with grey indicating few to no misclassified models and orange indicating a higher number of misclassified models. Model 1: no divergence, model 2: divergence without gene flow, model 3: divergence with secondary contact, and model 4: divergence with gene flow.

T. blandingii	Model 1	Model 2	Model 3	Model 4
Model 1	91990	0	8010	0
Model 2	0	87719	13	12268
Model 3	36659	39	63201	101
Model 4	0	11847	160	87993

T. pulverulenta	Model 1	Model 2	Model 3	Model 4
Model 1	89710	0	10290	0
Model 2	0	81778	39	18183
Model 3	45225	48	54520	207
Model 4	2	17002	329	82667

Table S3. The number of 500 random forest classifiers that voted for each competing demographic model in delimitR. Model 1: no divergence, model 2: divergence without gene flow, model 3: divergence with secondary contact, and model 4: divergence with gene flow.

	Model 1	Model 2	Model 3	Model 4
T. blandingii	1	278	6	215
T. pulverulenta	0	150	81	269

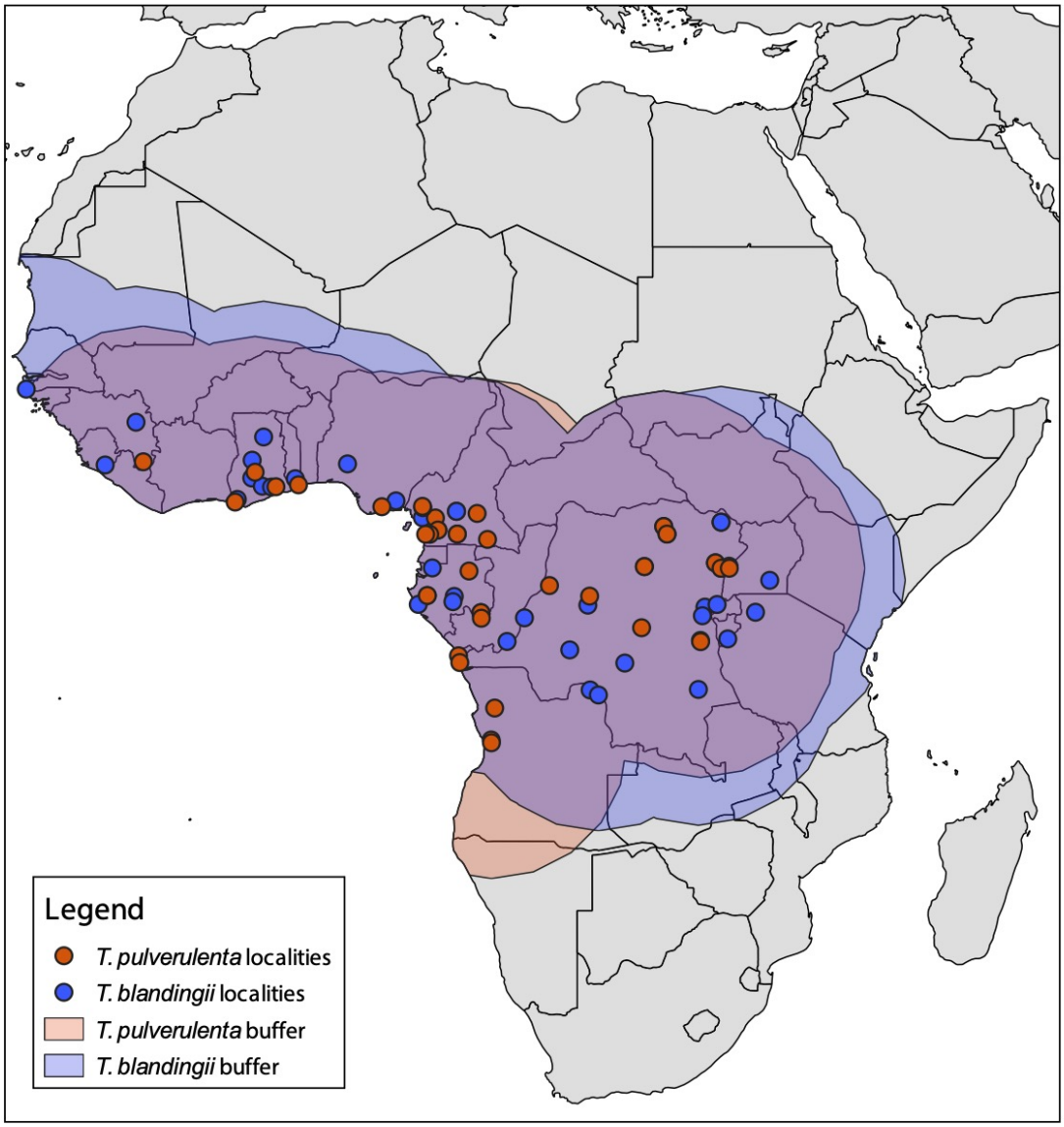


Figure S1. Points and 1000-kilometer buffers used for niche modeling parsed by species.

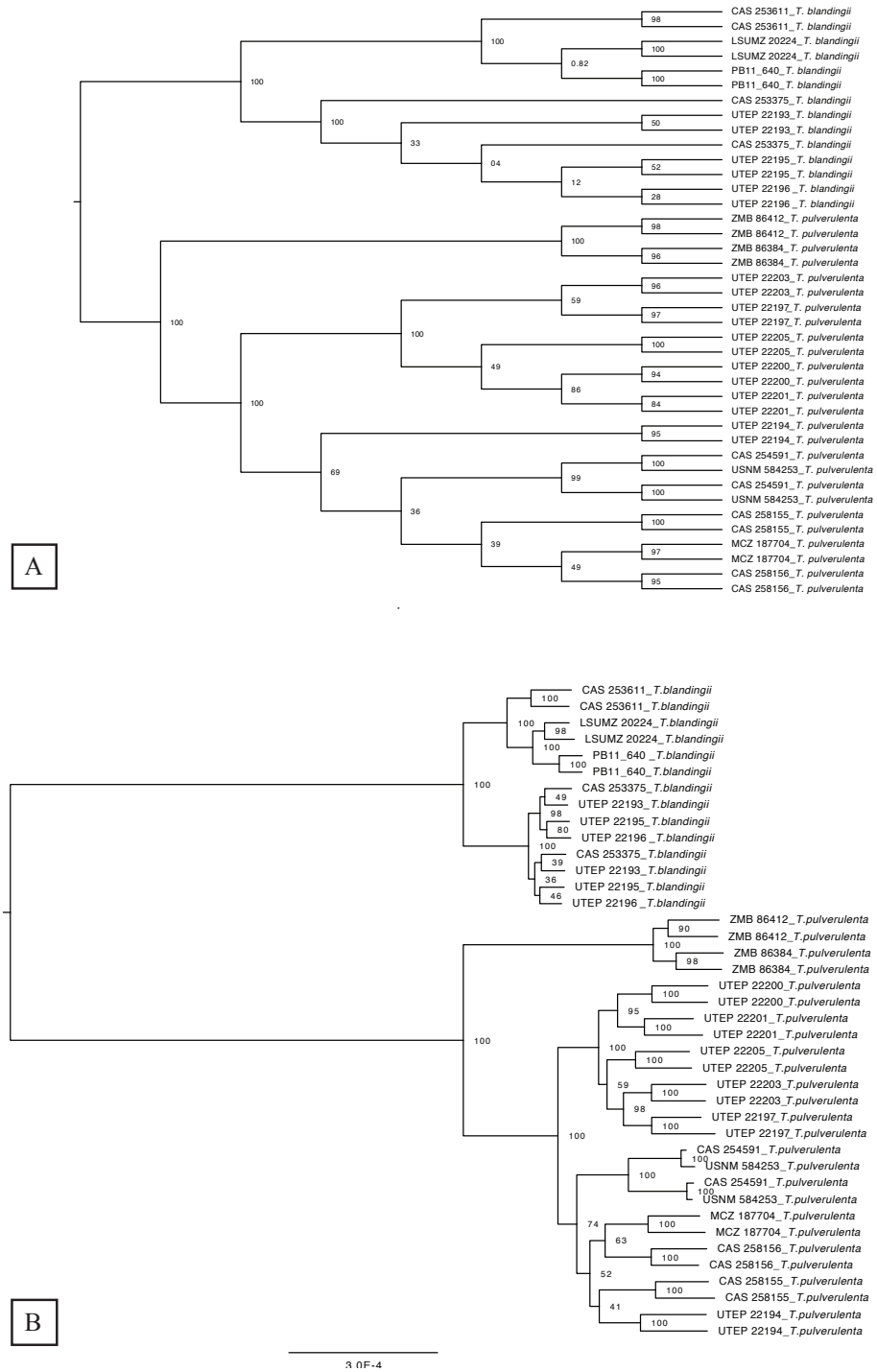


Figure S2. Phylogenetic trees estimated from full ddRADseq data sets in A) SVDquartets, and B) IQtree. Numerical nodal support values represent percentages of 100 nonparametric bootstrap replicates in (A) and percentages of 10,000 ultrafast bootstrap replicates in (B). Branch lengths in (B) are proportional to expected substitutions per site.

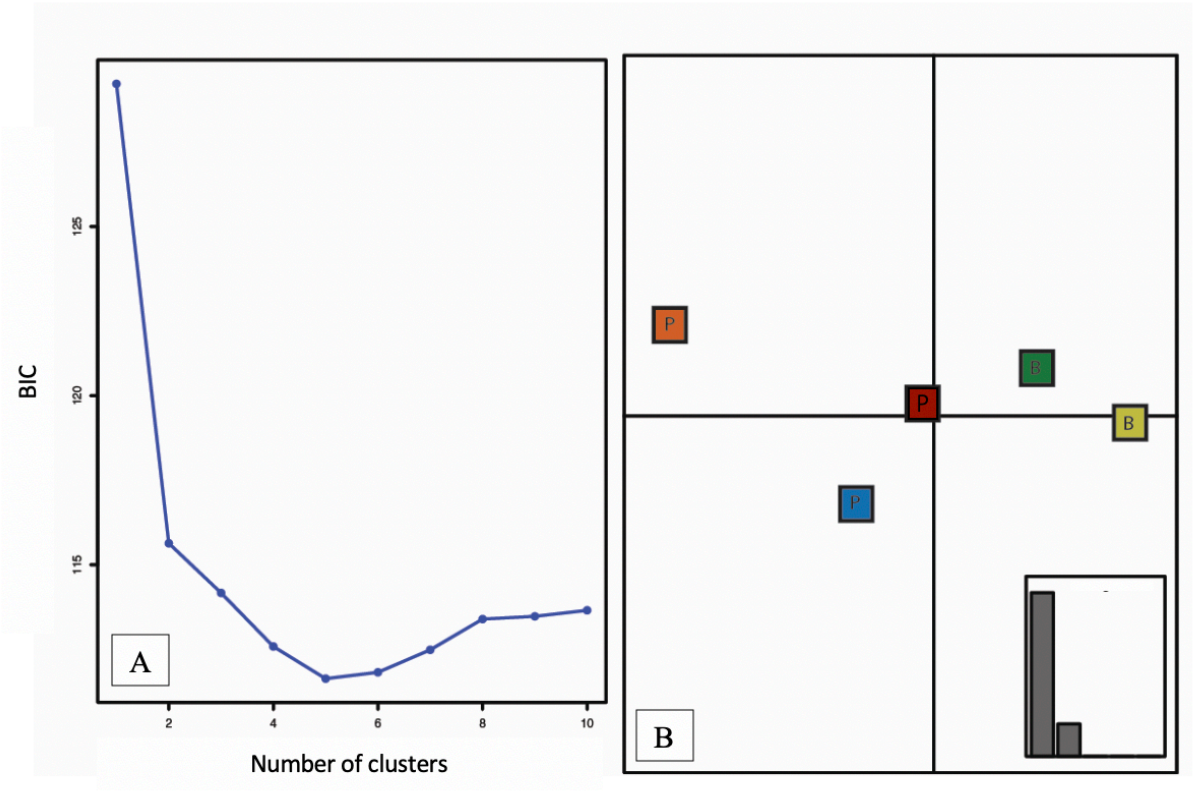


Figure S3. DAPC population structure analysis. A) inference of population number using BIC suggests 5 genetic clusters and B) scatterplot for those 5 clusters and the barplot of the DA eigenvalues. Colors correspond to clades in Figure 2.

APPENDIX II

From:

CHAPTER 3

Pleistocene refugial diversification across highland and river systems in a widespread Central African skink

Allen, K.E., Tapondjou N., W.P., Greenbaum, E. Bauer, A.M., Ceríaco, L.M.P., Kusamba, C., Kameni N., M.M., Peterson, A.T., Brown, R.M. Pleistocene refugial diversification across highland and river systems in a widespread Central African skink. Molecular Ecology. *In review*.

Table S1. Accession numbers and sampling localities for all specimens used in this study.

Species	Museum Accession #	Latitude	Longitude
T. maculilabris	UWBM 6045	6.8350	-1.7217
T. maculilabris	UTEP 22380	0.2556	32.6184
T. maculilabris	UTEP 22379	-0.9076	20.6252
T. maculilabris	UTEP 22378	0.2940	25.2951
T. maculilabris	UTEP 22377	0.5187	25.2602
T. maculilabris	UTEP 22376	-5.9302	29.1920
T. maculilabris	UTEP 22374	-3.3835	29.1124
T. maculilabris	UTEP 22373	-2.2402	28.8491
T. maculilabris	UTEP 22372	-4.0843	28.1537
T. maculilabris	UTEP 22371	-3.1153	16.1189
T. maculilabris	UTEP 22370	-1.9713	18.2814
T. maculilabris	UTEP 22369	-1.9713	18.2814
T. maculilabris	UTEP 22368	-4.4222	15.0490
T. maculilabris	UTEP 22367	-0.4479	18.1346
T. maculilabris	UTEP 22366	-0.4479	18.1346
T. maculilabris	UTEP 22365	-0.4075	19.1445
T. maculilabris	UTEP 22364	-0.4075	19.1445
T. maculilabris	UTEP 22363	3.6768	26.1174
T. maculilabris	UTEP 21827	-3.3412	28.1345
T. maculilabris	UTEP 21826	-2.2319	28.7864
T. maculilabris	UTEP 21825	-7.2936	27.3947
T. maculilabris	UTEP 21824	-8.0139	27.1129
T. maculilabris	UTEP 21822	-0.4945	20.1681
T. maculilabris	UTEP 21819	-2.7274	18.1443
T. maculilabris	UTEP 21818	-2.7410	18.1093
T. maculilabris	UTEP 21817	-2.6028	16.4547
T. maculilabris	UTEP 21816	-2.6430	16.2412
T. maculilabris	UTEP 21814	-4.4222	15.0490
T. maculilabris	UTEP 21809	-5.9302	29.1920
T. maculilabris	UTEP 21808	1.3984	28.5680
T. maculilabris	UTEP 21807	1.2385	28.3141
T. maculilabris	UTEP 21806	1.3669	29.0278
T. maculilabris	UTEP 21805	-5.0569	28.9236
T. maculilabris	UTEP 21804	-2.7325	29.0035
T. maculilabris	UTEP 21802	-1.8439	28.9957
T. maculilabris	UTEP 21801	0.6834	29.6714
T. maculilabris	UTEP 21800	-3.3835	29.1124
T. maculilabris	UTEP 21799	-3.4090	29.1427
T. maculilabris	UTEP 21798	-3.0261	28.4423
T. maculilabris	UTEP 21797	-2.2458	28.8218
T. maculilabris	MVZ 253443	6.4100	9.1500
T. maculilabris	MVZ 252621	6.2587	-1.0257
T. maculilabris	MVZ 249750	8.8729	0.0435

T. maculilabris	MVZ 249747	8.3338	0.5945
T. maculilabris	MVZ 249746	8.3484	0.6011
T. maculilabris	MVZ 249672	8.2665	0.5177
T. maculilabris	MVZ 245344	5.2818	-2.6414
T. maculilabris	MVZ 245342	8.2446	0.5243
T. maculilabris	MCZ 188639	0.5112	12.8028
T. maculilabris	ELI 3797	0.2942	25.2886
T. maculilabris	ELI 3361	-2.5723	25.6464
T. maculilabris	DFH 860	3.4025	33.5543
T. maculilabris	DFH 1128	0.8008	30.0944
T. maculilabris	CK 186	-2.0550	29.0548
T. maculilabris	CK 059	-1.2511	29.0599
T. maculilabris	CAS 253882	4.9554	9.8679
T. maculilabris	CAS 253495	3.0585	12.2468
T. maculilabris	CAS 253268	3.8821	11.4206
T. maculilabris	CAS 250828	-3.3430	29.2731
T. maculilabris	CAS 249863	5.5806	10.5908
T. maculilabris	CAS 249859	4.5973	12.2253
T. maculilabris	BYU 62031	5.8843	10.3766
T. maculilabris	BYU 62023	4.8478	9.8197
T. maculilabris	AKT 13	2.8611	23.9336
T. maculilabris	AKT 12	2.8611	23.9336
T. adamastor	MB03-000956	1.6365	7.4260
T. adamastor	MB03-000957	1.6365	7.4260
T. adamastor	MB03-001003	1.6365	7.4260
T. adamastor	MB03-001005	1.6365	7.4260
T. adamastor	MB03-001017	1.3441	7.2933
T. adamastor	MB03-001018	1.3441	7.2933
T. adamastor	MB03-001019	1.3441	7.2928
T. adamastor	MB03-001020	1.3433	7.2916
T. adamastor	MB03-001021	1.6259	7.4138
T. adamastor	MB03-001043	1.3414	7.2932
T. adamastor	MB03-001044	1.3438	7.2926
T. adamastor	MB03-001045	1.3437	7.2924
T. adamastor	MB03-001046	1.3436	7.2922
T. adamastor	MB03-001048	1.3427	7.2914
T. thomensis	MB03-000960	0.3382	6.7326
T. thomensis	CAS 219288	0.3297	6.7290
T. thomensis	CAS 219287	0.3297	6.7290
T. thomensis	CAS 218820	0.3257	6.5084
T. affinis	MB03-001022	1.6390	7.4236
T. affinis	MB03-001023	1.6390	7.4236
T. affinis	MB03-001024	1.6390	7.4236

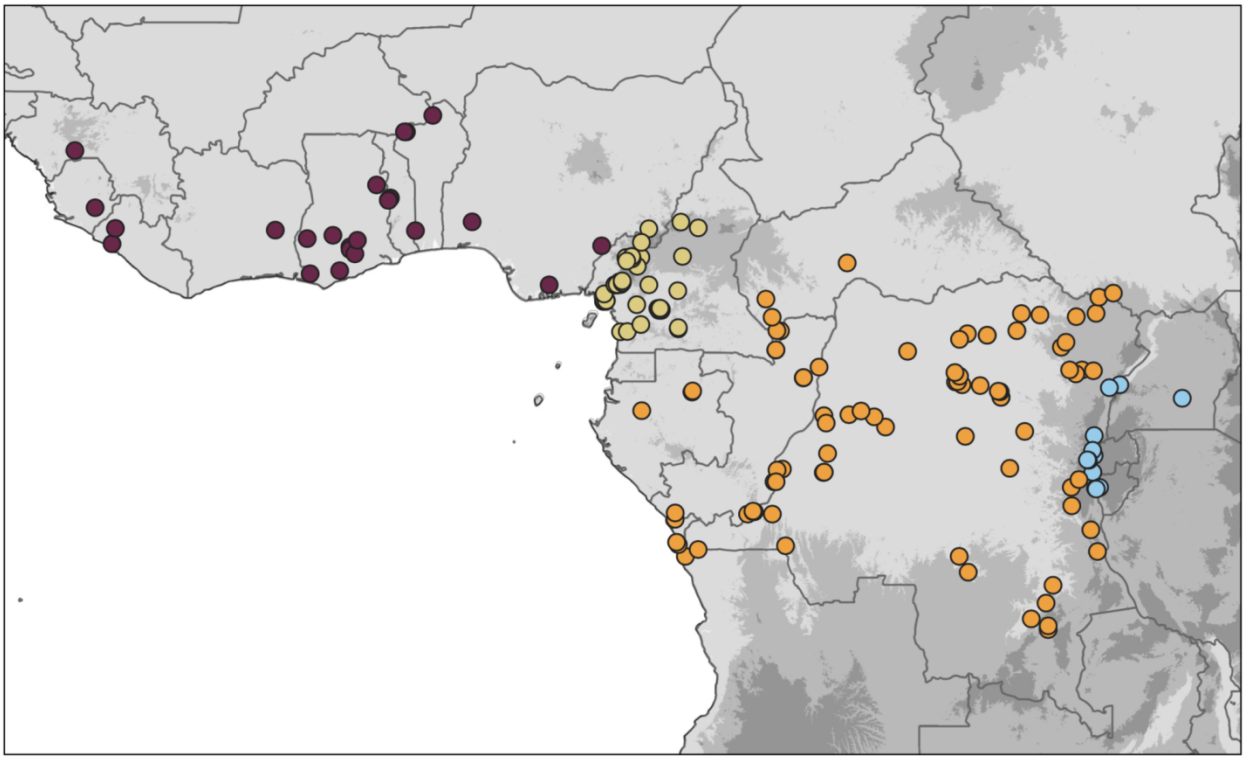


Figure S1. Niche model points, pre-thinning, parsed by clade, color-coded to match figures in the main text. The two lowland Central African clades were combined for niche modeling.

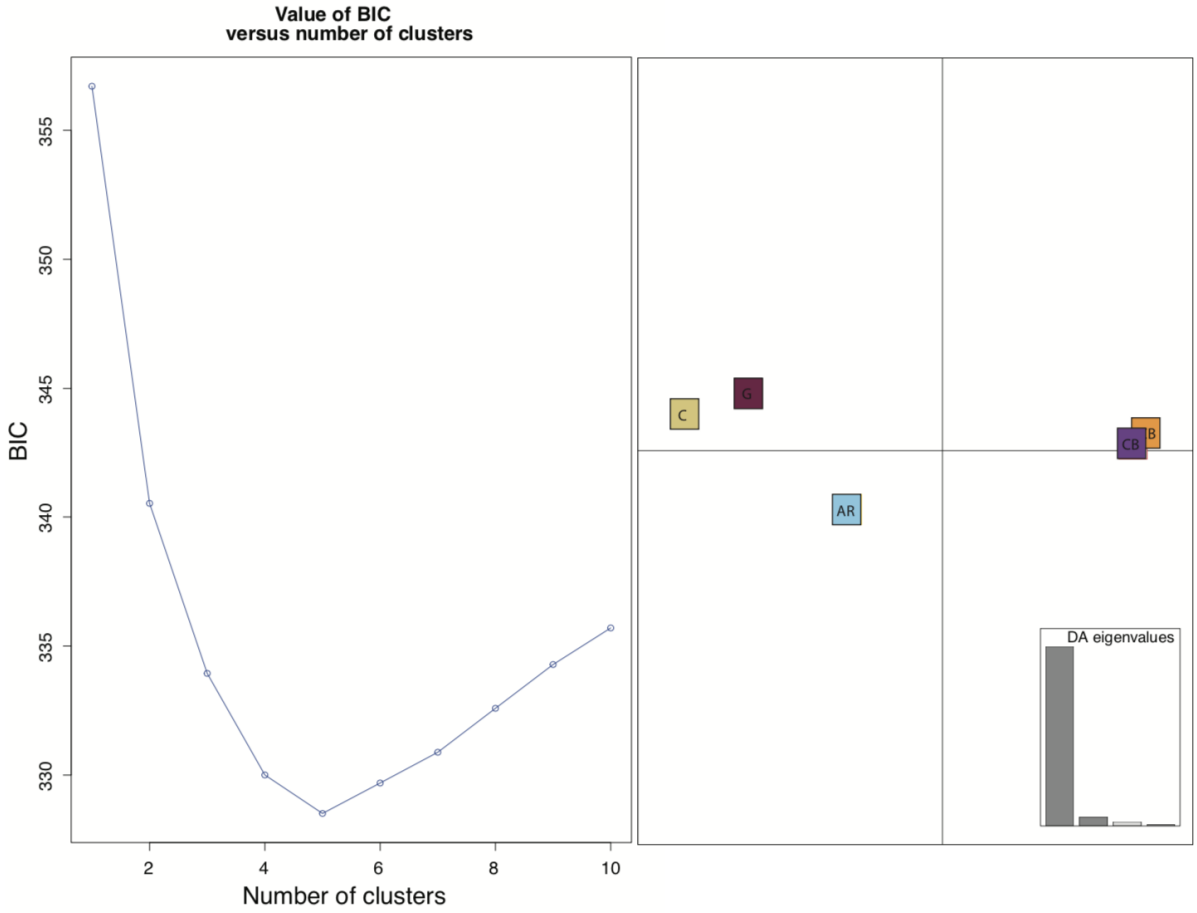


Figure S2. DAPC population structure analysis. (A) inference of population number using BIC suggests 5 genetic clusters and (B) scatterplot for those 5 clusters and the bar plot of the DA eigenvalues. Colors correspond to those used for each clade in the main text.

Table S2. Output of the RangeExpansion analysis showing the longitude and latitude of the projected point of origin, the strength of the founder effect (Q), the the decrease in diversity over 1/10/100 km (R1 - R100), the founder distance (D1), the correlation coefficient (Rsq) and correlation p-value (Pval) for the most likely origin.

	Cameroon	**Ghana**	**Albertine Rift**	**Congo Basin (P)**	**Congo Basin (O)**
Longitude	9.81967	-0.0210889	30.9153	28.7551378	20.0139
Latitude	3.80209979	6.41584526	-0.5286479	-1.1784011	-2.2908721
Q	0.00079551	0.00055242	0.00251631	0.00094779	0.00029919
R1	0.99841151	0.99889638	0.99499259	0.99810801	0.99940198
R10	0.98433897	0.98907236	0.95208522	0.98139688	0.99405178
R100	0.86273682	0.90050847	0.66522029	0.8406487	0.94354035
D1	6.34876232	9.1425321	2.00710993	5.32873034	16.8805696
Rsq	0.43314123	0.87414314	0.75046691	0.79693103	0.38045268
Pval	7.07136136	4.40E-13	7.85E-21	1.56E-09	7.81E-12

Table S3. DelimitR confusion matrices for the West Africa+Cameroon clade migration and expansion analyses. Red indicates the number out of 10,000 simulated data sets that were correctly classified by the random forest classifiers, orange and yellow indicate the number of models that were incorrectly classified with yellow indicating few to no misclassified models and orange indicating a higher number of misclassified models. Model codes are interpreted in Table S7 below.

West Africa+Cameroon–Migration

	M 1	**M 2**	**M 3**	**M 4**
M 1	7227	0	2773	0
M 2	0	9118	0	882
M 3	5100	0	4900	0
M 4	0	1932	0	8068

West Africa+Cameroon–Expansion

	M 1	**M 2**	**M 3**	**M 4**
M 1	8897	28	532	543
M 2	379	6593	1516	1512
M 3	1711	458	7728	103
M 4	1695	413	97	7795

Table S4. The number of 500 random forest classifiers that voted for each competing demographic model in delimitR. Model codes are interpreted in Table S7 below.

	M 1	**M 2**	**M 3**	**M 4**
Divergence	24	103	43	330
Expansion	29	257	194	20

Table S5. DelimitR confusion matrices for the Central African clade migration and expansion analyses. Red indicates the number out of 10,000 simulated data sets that were correctly classified by the random forest classifiers, orange and yellow indicate the number of models that were incorrectly classified with yellow indicating few to no misclassified models and orange indicating a higher number of misclassified models. Model codes are interpreted in Table S7 below.

Central Africa–Migration

	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13
M 1	8997	0	892	0	0	0	0	0	34	23	26	28	0
M 2	0	8462	0	1204	0	0	0	334	0	0	0	0	0
M 3	5088	0	1925	4	0	0	0	0	767	772	696	748	0
M 4	0	1719	0	7896	0	0	0	385	0	0	0	0	0
M 5	0	0	0	0	8216	0	0	0	0	0	0	0	1784
M 6	0	0	0	0	1	9999	0	0	0	0	0	0	0
M 7	0	0	0	0	0	0	10000	0	0	0	0	0	0
M 8	0	1992	0	479	0	0	0	7528	0	0	0	0	1
M 9	2463	0	1223	0	0	1	1	0	1573	1569	1552	1618	0
M 10	2364	0	1231	1	0	0	0	0	1480	1681	1642	1601	0
M 11	2419	0	1212	2	0	0	2	0	1467	1568	1679	1651	0

M 12	2427	0	1226	0	0	0	0	0	1502	1531	1563	1751	0
M 13	0	0	0	0	934	0	0	0	0	0	0	0	9066

Central Africa–Expansion

	M 1	**M 2**	**M 3**	**M 4**	**M 5**	**M 6**	**M 7**	**M 8**
M 1	4651	13	320	510	4062	39	258	147
M 2	53	6421	197	278	58	1188	203	1602
M 3	821	105	3901	66	874	405	3812	16
M 4	827	47	57	7186	821	488	61	513
M 5	4087	21	344	506	4612	45	234	151
M 6	181	394	719	1412	176	6303	749	66
M 7	818	128	3721	72	864	399	3978	20
M 8	226	401	9	1370	255	59	9	7671

Table S6. The number of 500 random forest classifiers that voted for each competing demographic model in delimitR. Model codes are interpreted in Table S7 below.

	M 1	**M 2**	**M 3**	**M 4**	**M 5**	**M 6**	**M 7**	**M 8**	**M 9**	**M 10**	**M 11**	**M 12**	**M 13**
Divergence	1	34	5	31	131	48	38	24	9	8	10	6	155
Expansion	34	125	66	36	38	99	80	22	--	--	--	--	--

Table S7. delimit R model code interpretation.

Model key

West Africa Migration

M1: 1 pop
M2: 2 pops, no gene flow
M3: 2 pops secondary contact
M4: 2 pops, divergence with gene flow

West Africa Expansion

M1: No expansion
M2: Expansion in both pops
M3: Expansion in Pop 0
M4: Expansion in pop 1

Central Africa Migration

M1: 1 pop
M2: 2 pops (0 and 1/2)
M3: 2 pops (0 and 1/2), secondary contact
M4: 2 pops (0 and 1/2) divergence with gene flow
M5: 3 pops
M6: 3 pops secondary contact 0 <->1
M7: 3 pops secondary contact 0 <-> 2
M8: 3 pops secondary contact 1 <-> 2
M9: 3 pops secondary contact 0 <->1, 0 <-> 2
M10: 3 pops secondary contact 0 <->1, 1<-> 2
M11: 3 pops secondary contact 0 <->2, 1<-> 2
M12: 3 pops secondary contact all pops
M13: 3 pops divergence with gene flow 1-2

Central Africa Expansion

M1: No expansion
M2: Expansion in all pops
M3: Expansion in pop 0
M4: Expansion in pop 1
M5: Expansion in pop 2
M6: Expansion in pop 0 & 1
M7: Expansion in pop 0 & 2
M8: Expansion in pop 1 & 2

APPENDIX III

From:

CHAPTER 4

Origin and colonization history of the *Trachylepis maculilabris* complex in the Gulf of Guinea

Allen, K.E., Ceríaco, L.M.P., Bauer, A.M., Greenbaum, Tapondjou N., W.P., E. Kusamba, C., Kameni N., M.M., Peterson, A.T., Brown, R.M. Pleistocene refugial diversification across highland and river systems in a widespread Central African skink. *In prep*.

Table S1. Accession numbers and sampling localities for all specimens used in this study.

Species	Museum Accession #	Latitude	Longitude
T. maculilabris	UWBM 6045	6.8350	-1.7217
T. maculilabris	UTEP 22380	0.2556	32.6184
T. maculilabris	UTEP 22379	-0.9076	20.6252
T. maculilabris	UTEP 22378	0.2940	25.2951
T. maculilabris	UTEP 22377	0.5187	25.2602
T. maculilabris	UTEP 22376	-5.9302	29.1920
T. maculilabris	UTEP 22374	-3.3835	29.1124
T. maculilabris	UTEP 22373	-2.2402	28.8491
T. maculilabris	UTEP 22372	-4.0843	28.1537
T. maculilabris	UTEP 22371	-3.1153	16.1189
T. maculilabris	UTEP 22370	-1.9713	18.2814
T. maculilabris	UTEP 22369	-1.9713	18.2814
T. maculilabris	UTEP 22368	-4.4222	15.0490
T. maculilabris	UTEP 22367	-0.4479	18.1346
T. maculilabris	UTEP 22366	-0.4479	18.1346
T. maculilabris	UTEP 22365	-0.4075	19.1445
T. maculilabris	UTEP 22364	-0.4075	19.1445
T. maculilabris	UTEP 22363	3.6768	26.1174
T. maculilabris	UTEP 21827	-3.3412	28.1345
T. maculilabris	UTEP 21826	-2.2319	28.7864
T. maculilabris	UTEP 21825	-7.2936	27.3947
T. maculilabris	UTEP 21824	-8.0139	27.1129
T. maculilabris	UTEP 21822	-0.4945	20.1681
T. maculilabris	UTEP 21819	-2.7274	18.1443
T. maculilabris	UTEP 21818	-2.7410	18.1093
T. maculilabris	UTEP 21817	-2.6028	16.4547
T. maculilabris	UTEP 21816	-2.6430	16.2412
T. maculilabris	UTEP 21814	-4.4222	15.0490
T. maculilabris	UTEP 21809	-5.9302	29.1920
T. maculilabris	UTEP 21808	1.3984	28.5680
T. maculilabris	UTEP 21807	1.2385	28.3141
T. maculilabris	UTEP 21806	1.3669	29.0278
T. maculilabris	UTEP 21805	-5.0569	28.9236
T. maculilabris	UTEP 21804	-2.7325	29.0035
T. maculilabris	UTEP 21802	-1.8439	28.9957
T. maculilabris	UTEP 21801	0.6834	29.6714
T. maculilabris	UTEP 21800	-3.3835	29.1124
T. maculilabris	UTEP 21799	-3.4090	29.1427
T. maculilabris	UTEP 21798	-3.0261	28.4423
T. maculilabris	UTEP 21797	-2.2458	28.8218
T. maculilabris	MVZ 253443	6.4100	9.1500
T. maculilabris	MVZ 252621	6.2587	-1.0257
T. maculilabris	MVZ 249750	8.8729	0.0435
T. maculilabris	MVZ 249747	8.3338	0.5945

T. maculilabris	MVZ 249746	8.3484	0.6011
T. maculilabris	MVZ 249672	8.2665	0.5177
T. maculilabris	MVZ 245344	5.2818	-2.6414
T. maculilabris	MVZ 245342	8.2446	0.5243
T. maculilabris	MCZ 188639	0.5112	12.8028
T. maculilabris	ELI 3797	0.2942	25.2886
T. maculilabris	ELI 3361	-2.5723	25.6464
T. maculilabris	DFH 860	3.4025	33.5543
T. maculilabris	DFH 1128	0.8008	30.0944
T. maculilabris	CK 186	-2.0550	29.0548
T. maculilabris	CK 059	-1.2511	29.0599
T. maculilabris	CAS 253882	4.9554	9.8679
T. maculilabris	CAS 253495	3.0585	12.2468
T. maculilabris	CAS 253268	3.8821	11.4206
T. maculilabris	CAS 250828	-3.3430	29.2731
T. maculilabris	CAS 249863	5.5806	10.5908
T. maculilabris	CAS 249859	4.5973	12.2253
T. maculilabris	BYU 62031	5.8843	10.3766
T. maculilabris	BYU 62023	4.8478	9.8197
T. maculilabris	AKT 13	2.8611	23.9336
T. maculilabris	AKT 12	2.8611	23.9336
T. adamastor	MB03-000956	1.6365	7.4260
T. adamastor	MB03-000957	1.6365	7.4260
T. adamastor	MB03-001003	1.6365	7.4260
T. adamastor	MB03-001005	1.6365	7.4260
T. adamastor	MB03-001017	1.3441	7.2933
T. adamastor	MB03-001018	1.3441	7.2933
T. adamastor	MB03-001019	1.3441	7.2928
T. adamastor	MB03-001020	1.3433	7.2916
T. adamastor	MB03-001021	1.6259	7.4138
T. adamastor	MB03-001043	1.3414	7.2932
T. adamastor	MB03-001044	1.3438	7.2926
T. adamastor	MB03-001045	1.3437	7.2924
T. adamastor	MB03-001046	1.3436	7.2922
T. adamastor	MB03-001048	1.3427	7.2914
T. thomensis	MB03-000960	0.3382	6.7326
T. thomensis	CAS 219288	0.3297	6.7290
T. thomensis	CAS 219287	0.3297	6.7290
T. thomensis	CAS 218820	0.3257	6.5084
T. affinis	MB03-001022	1.6390	7.4236
T. affinis	MB03-001023	1.6390	7.4236
T. affinis	MB03-001024	1.6390	7.4236

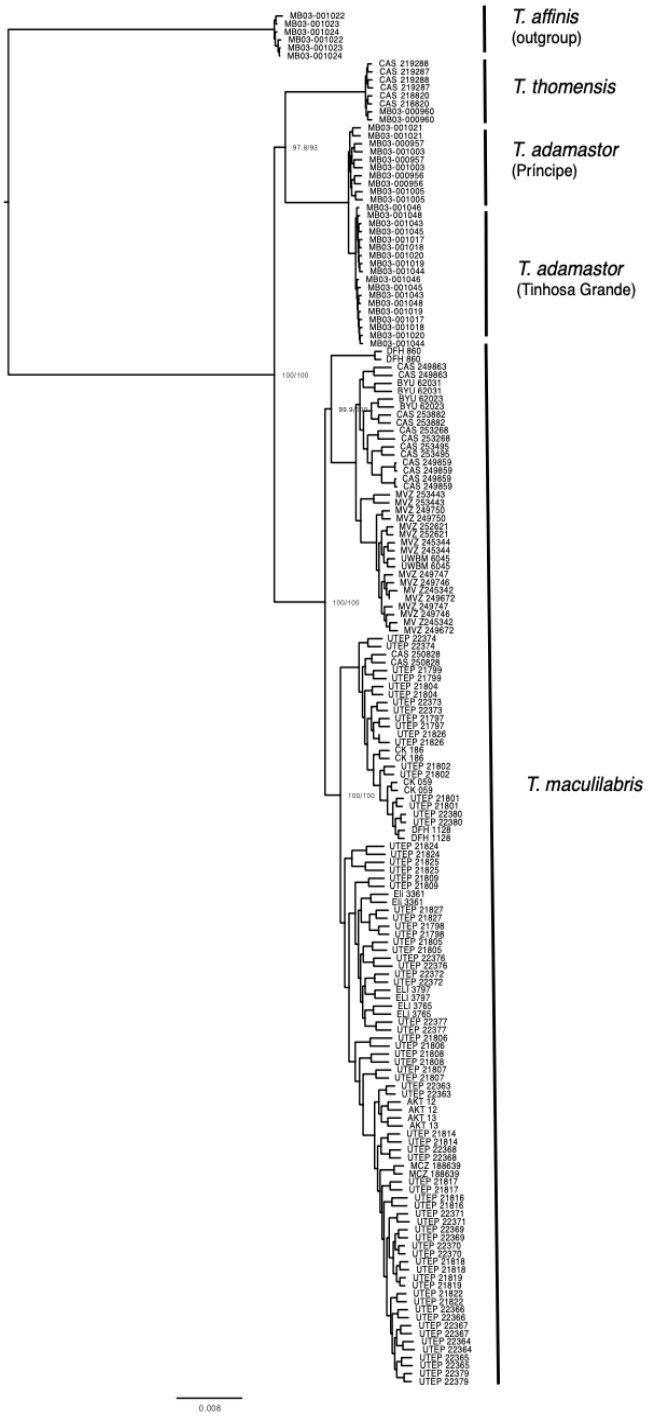


Figure S1. Intraspecific phylogenetic relationships of the *Trachylepis maculilabris* complex across West and Central Africa and the island species based on maximum likelihood analysis (implemented in IQ-TREE) of SNP data. All nodes were highly supported.

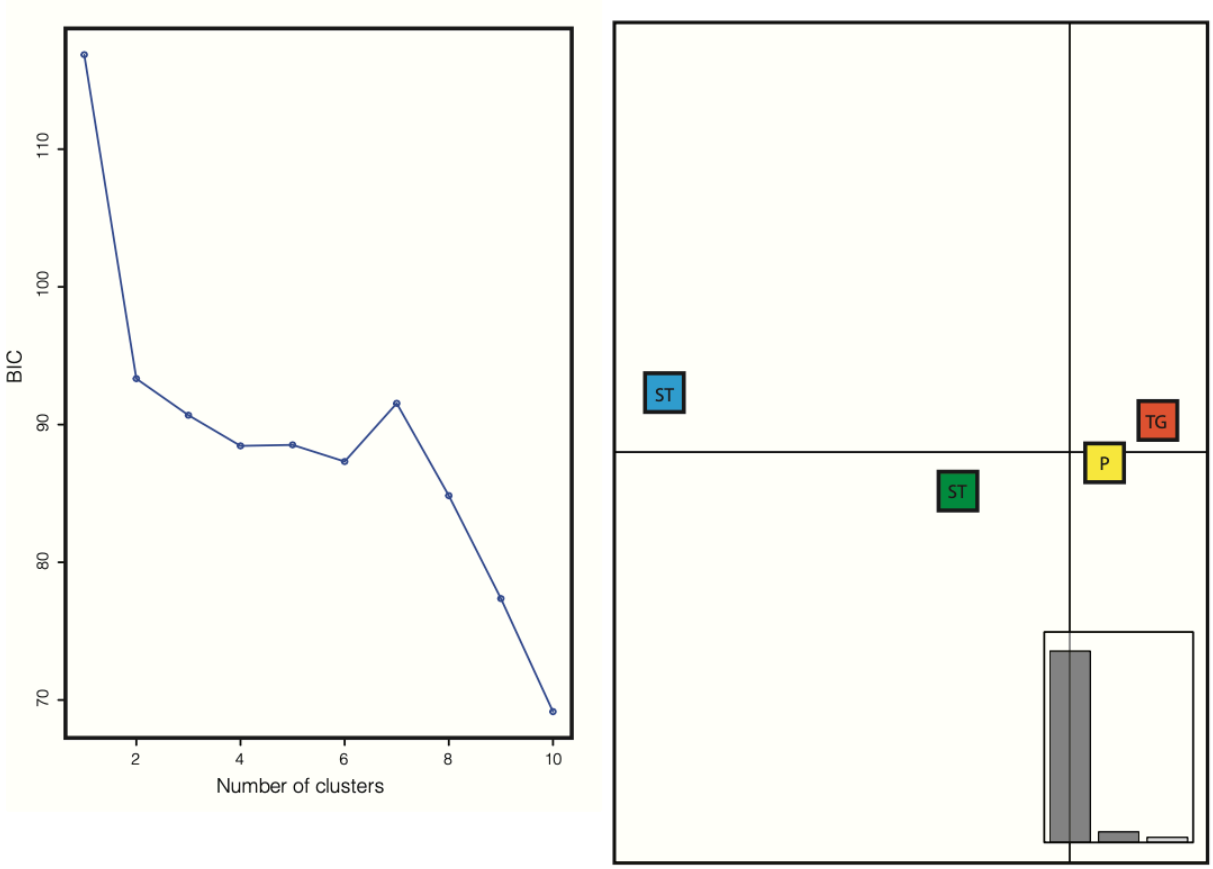


Figure S2. DAPC population structure analysis of *Trachylepis thomensis* and *T. adamastor*. (A) inference of population number using BIC suggests 4 genetic clusters and (B) scatterplot for those 4 clusters and the barplot of the DA eigenvalues. Islands are denoted as ST: São Tomé, P: Príncipe, TG: Tinhosa Grande.

Table S2. The number of 500 random forest classifiers that voted for each competing demographic model in delimitR for *Trachylepis thomensis* and *T. adamastor* on Príncipe and Tinhosa Grande. Model codes are interpreted below.

Model	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**	**M9**	**M10**	**M11**	**M12**	**M13**
Votes	3	2	23	7	2	20	42	9	95	119	68	105	5

Model interpretation:

M1: 1 pop

M2: 2 pops (0 and 1/2)

M3: 2 pops (0 and 1/2), secondary contact

M4: 2 pops (0 and 1/2) divergence with gene flow

M5: 3 pops

M6: 3 pops secondary contact 0 <->1

M7: 3 pops secondary contact 0 <-> 2

M8: 3 pops secondary contact 1 <-> 2

M9: 3 pops secondary contact 0 <->1, 0 <-> 2

M10: 3 pops secondary contact 0 <->1, 1<-> 2

M11: 3 pops secondary contact 0 <->2, 1<-> 2

M12: 3 pops secondary contact all pops

M13: 3 pops divergence with gene flow 1-2

Table S3. Ancestral geographic range reconstruction (BioGeoBEARS) parameters and statistical comparison across all populations of *T. maculilabris* on the mainland, *Trachylepis thomensis* and *T. adamastor* on Príncipe and Tinhosa Grande.

Model	**LnL**	**parameters**	**d**	**e**	**j**	**AICc**	**AICc**
DEC	-35.66	2	0.041	0.12	0	77.73	0.079
DEC+J	-32.04	3	0.0068	1.0e-12	3	76.07	0.18
DIVALIKE	-33.71	2	0.029	0.066	0	73.82	0.56
DIVALIKE+J	-32.04	3	0.0068	1.0e-12	2	76.07	0.18
BAYAREALIKE	-44.9	2	1.36	2.05	0	96.2	7.7e-06
BAYAREALIKE+J	-44.9	3	1.36	2.05	0.0001	101.8	4.7e-07

APPENDIX IV

From:

CHAPTER 5

Allopolyploidization and niche evolution in African clawed frogs (*Xenopus*)

Allen, K.E., Evans, B.J., Tapondjou N., W.P., Ané, C., Brown, R.M., Peterson, A.T. Allopolyploidization and niche evolution in African clawed frogs (*Xenopus*). Evolution. *In review.*

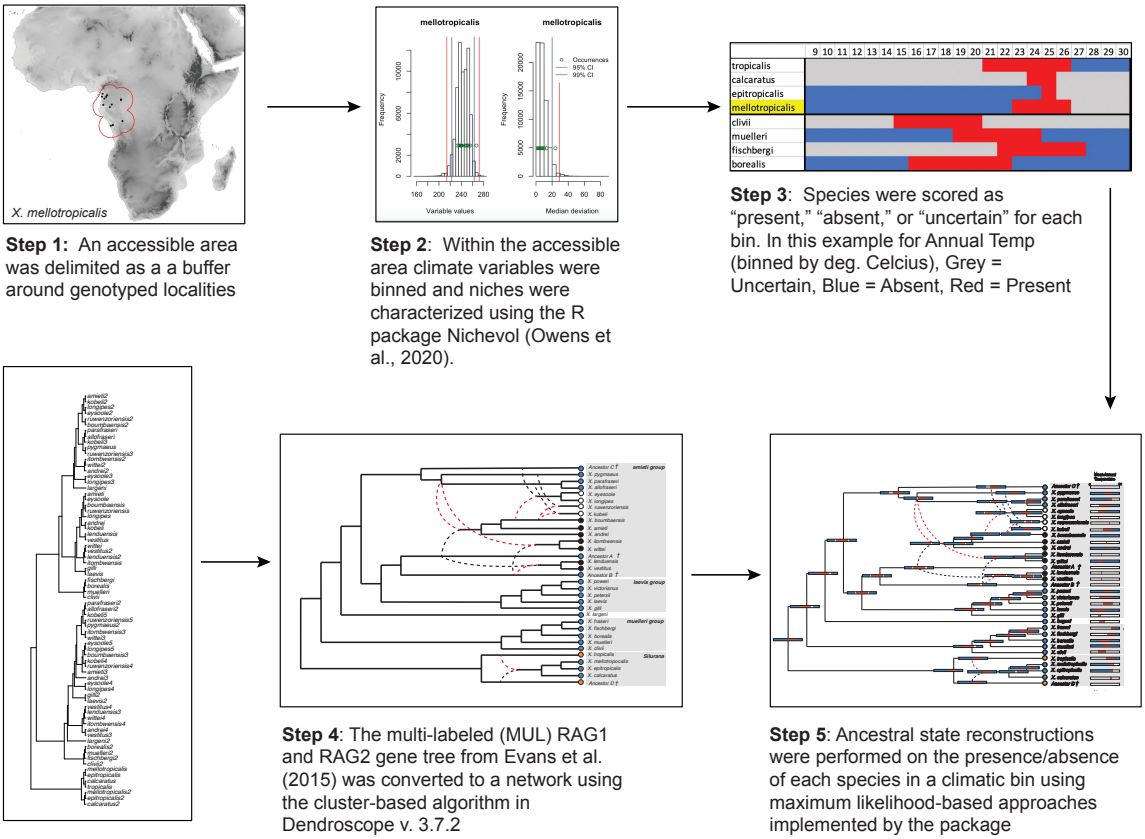
Table S1. Accession numbers, sample IDs and sampling localities for all specimens used in this study. Country codes: DRC – Democratic Republic of Congo, EG: Equatorial Guinea, SA: South Africa

Species	Museum ID	Sample ID	Country	Latitude	Longitude
X. allofraseri	–	BJE3461	EG	3.334018	8.586133
X. allofraseri	IRSNB	PM107	DRC	-5.59632	13.16028
X. allofraseri	CAS 207765	RCD13495	EG	3.52775	8.579389
X. allofraseri	–	VG09-357	Cameroon	4.06894	9.0684
X. allofraseri	–	X1-23	EG	3.366783	8.504117
X. amieti	MHNG 2644.54	AMNH17267	Cameroon	5.63	10.33
X. amieti	–	BJE3160	Cameroon	5.5703	10.60973
X. borealis	MHNG 2644.64	AMNH17312	Kenya	0.66	37.5
X. borealis	–	xen226	Kenya	-1.17	38.01
X. borealis	–	xen227	Kenya	2.33	38.01
X. borealis	–	xen230	Kenya	1.08	36.68
X. boumbaensis	MHNG 2644.57	AMNH17284	Cameroon	2.02	15.13
X. boumbaensis	NMP6V 74703/5	VG10-151	Cameroon	2.44217	15.43078
X. boumbaensis	NMP6V 74715/1	VG10-190	Cameroon	2.10276	15.35658
X. boumbaensis	NMP6V 74723/1	VG10-201	Cameroon	1.9419	15.62694
X. boumbaensis	NMP6V 74735/1	VG10-223	Cameroon	2.09175	15.41001
X. calcaratus	CAS 207759	RCD13494	EG	3.53	8.58
X. calcaratus	NMP6V 74746	VG05-S	Cameroon	4.06894	9.0684
X. clivii	MHNG 2644.50	AMNH17252	Ethiopia	9.02	38.7
X. clivii	MCZ cryogenic 350	BJE1522	Ethiopia	6.4818833	39.749817
X. epitropicalis	MHNG 2644.56	AMNH17275	DRC	-4.3	15.3
X. epitropicalis	UTEP 21188	ELI 1754	DRC	-3.1153	16.1989
X. epitropicalis	UTEP 21189	ELI 1799	DRC	-3.1852	16.1988
X. eysoole	MCZ A-148068	BJE3175	Cameroon	†6.2510	10.528017
X. eysoole	MCZ A-148094	BJE3217	Cameroon	6.5486167	10.759983
X. eysoole	MCZ A-138016	A-34586	Cameroon	6.2427778	10.500833
X. eysoole	MCZ A-148127	DMP580	Cameroon	6.01425	10.270267
X. eysoole	NMP6V 74555	VG10-07	Cameroon	6.05	10.43
X. eysoole	NMP6V 74745/1	VGCM12-01	Cameroon	6.061835	9.91576
X. fischbergi	UWBM 5964	ADL3542	Ghana	8.29078	-2.28506
X. fischbergi	MHNG 2644.60	AMNH17296	Nigeria	9.91667	8.9
X. fischbergi	AMNH 158377	MWK13208	Chad	16.91	21.77
X. fischbergi	NMP6V 74744/1	VG10-280	Cameroon	10.74	13.8
X. fischbergi	UTEP 21194	ELI 3206	DRC	3.67679	26.11742
X. gilli	–	rg(2-2)	South Africa	-34.25	18.41
X. gilli	–	xs(2-3)	SA	-34.37	18.92

X. itombwensis	MCZ A--138192	BJE00275	DRC	-3.35679	23.69011
X. kobeli	CAS 249964	BJE3071	Cameroon	4.60755	12.210283
X. kobeli	MCZ A-148061	BJE3153	Cameroon	5.204317	10.4289
X. kobeli	CAS 253768	DMP-1228	Cameroon	5.0415098	10.698480
X. kobeli	NMP6V 74714/1	VG10-188	Cameroon	2.10276	15.35658
X. kobeli	MCZ A--148065	BJE3166	Cameroon	5.5703	10.609733
X. laevis	–	EA12-1	SA	-34.35	18.8167
X. laevis	–	KML5	SA	-34.347	18.9833
X. laevis	–	LG12-3	SA	-34.268	18.4137
X. laevis	–	RGL1	SA	-34.337	18.9982
X. laevis	–	XSL15	SA	-34.3567	18.9333
X. laevis	–	BJE3505	SA	-33.410933	19.732583
X. laevis	MCZ A-149234	BJE3525	SA	-33.222967	20.851933
X. laevis	MCZ A-149256	BJE3547	SA	-33.966517	22.602067
X. laevis	–	BJE3550	SA	-33.892417	22.878033
X. laevis	MCZ A-149259	BJE3572	SA	-33.900867	22.912417
X. laevis	MCZ A-149243	BJE3534	SA	-32.24035	22.585783
X. laevis	MCZ A-149260	BJE3573	SA	-31.562133	23.061467
X. laevis	–	BJE3628	SA	-31.474717	19.060983
X. laevis	–	BJE3637	SA	-31.378817	19.103483
X. laevis	MCZ A-149265	BJE3639	SA	-31.471967	19.047483
X. laevis	MCZ A-149267	BJE3641	SA	-31.470583	19.074767
X. laevis	MCZ A-149268	BJE3642	SA	-31.469817	19.077933
X. laevis	MCZ A-149270	BJE3644	SA	-31.3281	19.083667
X. laevis	MCZ A-149273	BJE3647	SA	-31.402267	19.05005
X. laevis	MCZ A-149262	BJE3575	SA	-29.109267	24.589417
X. laevis	–	PF	SA	-26.715	27.1033
X. laevis	MHNG 2644.61	AMNH17301	Malawi	-15.78	35
X. laevis	–	lg(12-3)	SA	-34.2	18.4
X. largeni	MHNG 2644.59	AMNH17292	Ethiopia	5.6	38.9
X. largeni	MCZ cryogenic 409	BJE1581	Ethiopia	10.71493	37.55213
X. lenduensis	MCZ A-139853	BJE2957	DRC	2.01641	30.83111
X. longipes	AMNH A168447	BJE00238	Cameroon	6.202	10.45953
X. mellotropicalis	MHNG 2644.58	AMNH17288	Cameroon	2.88	11.15
X. mellotropicalis	–	BJE3013	Cameroon	3.808933	10.114833
X. mellotropicalis	CAS 255058	BJE3652	DRC	-5.83	12.57
X. mellotropicalis	NCSM 76797	BLS13506	Gabon	0.45358	10.27809
X. mellotropicalis	NCSM 78871	BLS14731	Gabon	-0.20175	12.26928
X. mellotropicalis	IRSNB	PM119	DRC	-5.6581	13.19947
X. mellotropicalis	NMP6V 74568	VG09-069	Cameroon	3.39134	11.46633

X. mellotropicalis	ZFMK 87790	VG10-04	Cameroon	2.35	10.616667
X. mellotropicalis	–	VG10-178	Cameroon	2.13881	15.65567
X. mellotropicalis	–	xen229	Cameroon	2.95	9.92
X. mellotropicalis	–	xen236	Gabon	0.57	12.87
X. mellotropicalis	UTEP 21191	ELI 1679	DRC	-4.41891	16.04641
X. muelleri	MHNG 2644.63	AMNH17308	Tanzania	-8.13	36.68
X. muelleri	–	RT3	Swaziland	-26.05	31.66
X. parafraseri	MCZ A--148029	BJE3063	Cameroon	3.87	11.52
X. parafraseri	CAS 253332	DCB-294	Cameroon	2.6445	14.03115
X. parafraseri	CAS 253366	DCB-279	Cameroon	2.6106	14.0234
X. parafraseri	CAS 253589	DCB-572	Cameroon	3.1737611	12.527139
X. parafraseri	CAS 253609	DCB-595	Cameroon	3.1983333	12.522778
X. parafraseri	NCSM 78872	BLS 14751	Gabon	0.04256	12.29834
X. parafraseri	NMP6V 75140/1	VGCG12-75	Congo Rep.	0.25175	14.16123
X. parafraseri	NMP6V 75140/6	VGCG12-32	Congo Rep.	0.26771	14.16181
X. parafraseri	NMP6V 75140/2	VGCG12-78	Congo Rep.	0.26861	14.15924
X. petersii	NMP6V 74142	VG08_81	Angola	-12.6	17.07
X. petersii	–	PM085	DRC	-5.61671	13.15926
X. petersii	MHNG 2644.67	AMNH17324	R. Congo	-4.3	12.4
X. poweri	–	RT4	Botswana	-18.97	22.57
X. poweri	MCZ A-148106	BJE3252	Cameroon	6.175283	10.07456
X. poweri	NMP6V 74582/2	VG09_100	Cameroon	7.74358	12.71618
X. poweri	NMP6V 74586/2	VG09_128	Cameroon	7.70149	12.67672
X. poweri	–	Xen058	Cameroon	7.32	13.58
X. poweri	MHNG 2644.52	AMNH17259	Nigeria	9.9	8.9
X. poweri	–	AMNH17260	Nigeria	9	8.9
X. poweri	MHNG 2644.53	AMNH17263	Zambia	-15.5	28.2
X. poweri	–	private 31	Zambia	-11.23793	24.268160
X. poweri	MCZ A-148108	BJE3254	Cameroon	6.175283	10.074567
X. pygmaeus	–	AMNH17323	DRC	-0.25	20.6
X. ruwenzoriensis	–	AMNH17316	Uganda	0.7	30
X. tropicalis	UWBM 5957	ADL3902	Ghana	6.24246	-0.5571
X. tropicalis	MHNG 2644.55	AMNH17271	Sierra Leone	8.5	-13.24
X. tropicalis	ROM 19161	tropicalis.L1720	Liberia	5.01	-9.04
X. tropicalis	ZFMK 87781	VG10-01	Cameroon	5.7512	9.3146
X. tropicalis	NMP6V 74659	VG10-40	Cameroon	6.73848	9.97818
X. tropicalis	–	xen228	Ivory Coast	5.32	-4.12
X. tropicalis	–	xen231	Nigeria	6.72	5.77
X. vestitus	–	RT2	Uganda	-1.22	29.71
X. victorianus	UTEP 21041	ELI1010	Burundi	-4.0277	30.15072

X. victorianus	UTEP 21036	ELI1064	Burundi	-3.35335	29.27119
X. victorianus	UTEP 21033	ELI1139	Burundi	-3.08188	29.40467
X. victorianus	UTEP 21034	ELI1141	Burundi	-3.06974	29.48445
X. victorianus	UTEP 21040	ELI938	Burundi	-3.94328	29.62828
X. victorianus	UTEP 21039	ELI965	Burundi	-3.94827	29.61908
X. victorianus	MCZ A-140118	BJE2897	DRC	1.9855667	30.861267
X. victorianus	UTEP 20170	EBG2329	DRC	1.5505	30.2529
X. victorianus	UTEP 20174	EBG2463	DRC	1.33869	30.15173
X. victorianus	UTEP 21046	CK003	DRC	-1.25114	29.0599
X. victorianus	MCZ A-138177	BJE260	DRC	-2.5028	28.87554
X. victorianus	MCZ A-138179	BJE262	DRC	-2.2455	28.81235
X. victorianus	UTEP 21050	CFS1090	DRC	-2.05496	29.05479
X. victorianus	UTEP 20165	EBG2147	DRC	-3.82118	29.09608
X. victorianus	UTEP 21047	ELI1298	DRC	-4.10784	29.09715
X. victorianus	UTEP 21048	ELI526	DRC	-3.34033	28.13085
X. victorianus	UTEP 21054	ELI502	DRC	-3.25732	28.11797
X. victorianus	UTEP 21052	ELI1369	DRC	-4.08432	28.15373
X. victorianus	UTEP 21056	ELI1461	DRC	-4.3	28.94133
X. victorianus	UTEP 21045	EBG2872	DRC	-6.682	29.0811
X. victorianus	–	ZFMK86159	Kenya	0.2666667	34.883333
X. victorianus	–	xen232	Rwanda	-1.7	30.55
X. victorianus	CAS168711	CAS168711	Tanzania	-5.07	38.72
X. victorianus	–	ZFMK63119	Uganda	0.8333	30.05
X. victorianus	–	xen234	Uganda	-1.116667	30.033333
X. victorianus	–	xen232	Rwanda	-1.95	30.06
X. victorianus	–	xen234	Uganda	-1.12	30.05
X. wittei	MHNG 2644.62	AMNH17304	Uganda	-1.06	29.9
X. wittei	CAS 201644	JVV4161	Uganda	-0.98	29.69
X. wittei	–	RT1	Rwanda	-2.57	28.98



Step 1: An accessible area was delimited as a a buffer around genotyped localities

Step 2: Within the accessible area climate variables were binned and niches were characterized using the R package Nichevol (Owens et al., 2020).

Step 3: Species were scored as "present," "absent," or "uncertain" for each bin. In this example for Annual Temp (binned by deg. Celcius), Grey = Uncertain, Blue = Absent, Red = Present

Step 4: The multi-labeled (MUL) RAG1 and RAG2 gene tree from Evans et al. (2015) was converted to a network using the cluster-based algorithm in Dendroscope v. 3.7.2

Step 5: Ancestral state reconstructions were performed on the presence/absence of each species in a climatic bin using maximum likelihood-based approaches implemented by the package PhyloNetworks

Figure S1. A diagrammatic representation of the niche evolution methods used in this study.

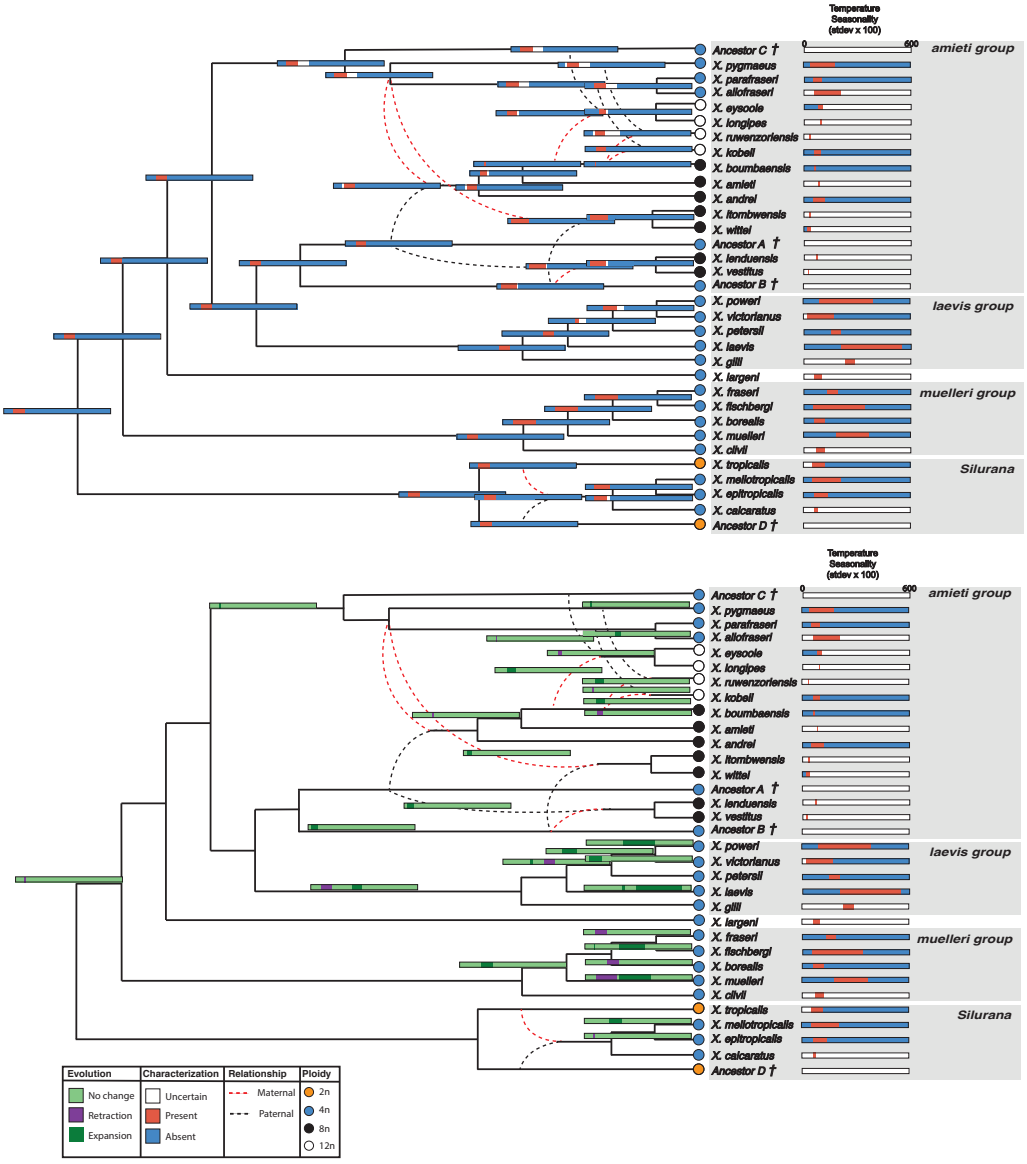


Figure S2. Top: Temperature seasonality reconstruction plotted on the nodes and Bottom: temperature seasonality evolution inferred from that reconstruction plotted on the branches of a reticulate phylogenetic network for each species in the genus *Xenopus*. Branches for which no niche evolution was inferred were left empty to increase readability. Species groups discussed in the text are highlighted in grey. Extinct or undiscovered ancestral species are denoted by †.

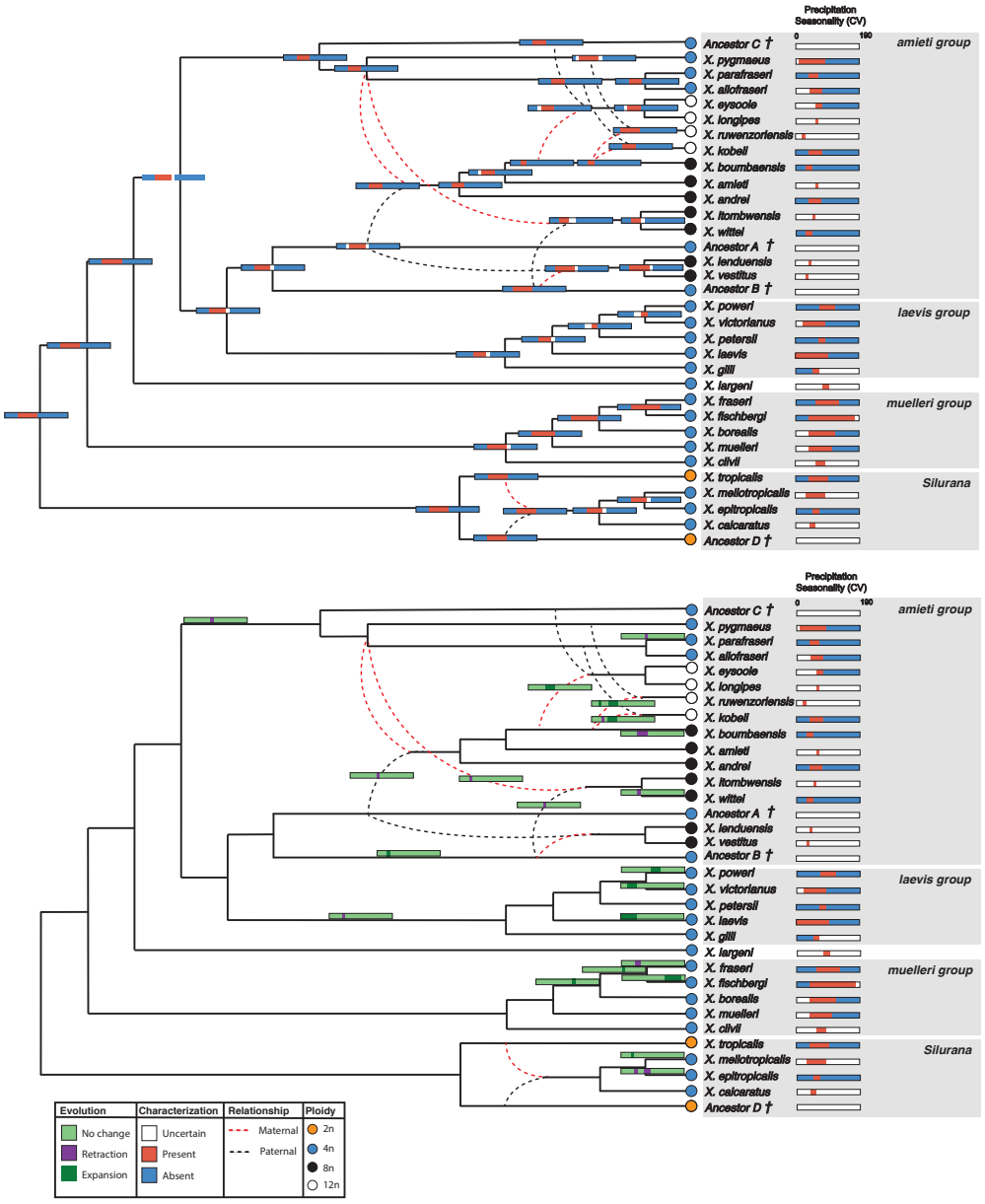


Figure S3. Top: Precipitation seasonality reconstruction plotted on the nodes and Bottom: precipitation seasonality evolution inferred from that reconstruction plotted on the branches of a reticulate phylogenetic network for each species in the genus *Xenopus*. Branches for which no niche evolution was inferred were left empty to increase readability. Species groups discussed in the text are highlighted in grey. Extinct or undiscovered ancestral species are denoted by †.

Supplemental Results

Clade-by-clade description of niche evolution

Subgenus *Silurana*

The subgenus *Silurana* occupies low-elevation rainforests in coastal and inland West and Central Africa. This is reflected in their niche characterization of relatively high mean annual temperatures with low temperature seasonality, and mid- to high precipitation with mid- to low seasonality (Figures 5.4 and 5.5; Appendix IV Figures 1 and 2). This species group is one of two in the genus with variation in ploidy level, which permits comparisons among closely related species of differing ploidy levels (here, a diploid and three tetraploid species). Compared to the coastal diploid species *X. tropicalis,* the coastal tetraploid species *X. calcaratus*, exhibited a very similar niche, with high levels of niche uncertainty, making further interpretation difficult. The two inland species of *Xenopus* (*Silurana*), *X. epitropicalis* and *X. mellotropicalis* – both tetraploids – show niche shifts towards lower annual precipitation and higher precipitation seasonality. *Xenopus epitropicalis* has a narrower niche breadth across all climatic variables compared to *X. mellotropicalis* (Figures 5.4 and 5.5; Appendix IV Figures 1 and 2). Thus, in this group there is evidence of niche stasis and niche shifts following allopolyploidization.

Subgenus *Xenopus*

Coinciding with the geographic and habitat shift between the subgenera *Xenopus* (*Silurana*) and *Xenopus* (*Xenopus*) from West and Central African rainforest to East African high elevation savanna (Figures 5.2 and 5.3), niche shifts are predicted by our ancestral state reconstructions to have occurred towards far lower mean annual temperatures and annual

precipitations (Figures 5.4 and 5.5), with slightly lower temperature seasonality and no projected change for precipitation seasonality (Appendix IV Figures 1 and 2). As ploidy levels increased in subgenus *Xenopus*, we observed a general trend towards narrowing of niche breadth and also colonization of novel niches.

***muelleri* species group**

The *muelleri* species group is probably composed only of allotetraploid species (ploidy of the species *X. fraseri* has not been confirmed but allotetraploidy is suspected (Evans et al., 2019)), and is broadly distributed in low elevation habitat but also includes two species with mid- to high elevation distributions (*X. borealis* and *X. clivii*). Species in this group are typically found in savanna and grassland habitats (Figures 5.2 and 5.3). Our geographic range reconstructions suggest an East African origin of this group with movement into the Ethiopian highlands and across the Sudanian savannas north of the Congo Basin (Figure 5.3). Consistent with these geographical inferences, our niche evolution analyses suggest corresponding shifts to higher temperature tolerance (Figure 5.4), and to a lesser extent precipitation seasonality (Appendix IV Figure 2), in the western and southern species of this clade. The exception to this is *X. borealis*, a high elevation species, whose mean annual temperature niche is shifted towards cooler temperatures (Figure 5.4). The base of the *muelleri* species group suggests a widening of the breadth of the temperature seasonality niche compared to its ancestor, and *X. muelleri* shows a completely non-overlapping niche shift compared to its ancestral species' projection for temperature seasonality (Appendix IV Figure 2). The projected temperature seasonality and annual precipitation niches for the remainder of the clade expand and contract around the

ancestral niche (Figures 5.5; Appendix IV Figure 2). Overall, the homoploid species of the *muelleri* species group exhibit high niches diversity along several axes.

***laevis* species group**

The *laevis* species group is composed only of allotetraploid species and is distributed in lowland savanna south of the Congo Basin (*X. laevis*), in the East African Rift (*X. victorianus*), the Congo Basin (*X. petersii*, *X. poweri*) and coastal fynbos habitat of southwest South Africa (*X. gilli*; Figure 5.2). All of the species in the clade overlap with inferred ancestral niche for mean annual temperature, with a niche shift towards cooler temperatures for *X. laevis*, and a niche contraction away from cooler temperatures for *X. poweri* (Figure 5.4). The base of the *laevis* species group is projected to shift towards higher temperature seasonality and expand into lower annual precipitation compared to its ancestor. *Xenopus laevis* utilizes the highest temperature seasonality and lowest annual precipitation of any *Xenopus* species (Figures 5.5, Appendix IV Figure 1). The inferred niche of the ancestor of *X. petersii*, *X. victorianus*, and *X. poweri* exhibits a contraction at the upper end of temperature seasonality and the lower end of annual precipitation, with a re-expansion into higher seasonality in *X. poweri* (Figures 5.5, Appendix IV Figure 1). *Xenopus gilli* utilizes narrower but overlapping niche space as *X. laevis* for all variables, however, uncertainty in its niche characterization makes it difficult to fully determine the breadth of its fundamental niche except in the case of precipitation seasonality, for which its narrower niche is confirmed. The homoploid *laevis* group therefore exhibits remarkable variation in niche breadths among its very closely related species, ranging from one of the narrowest niches (*X. gilli*) to one of the most general (*X. laevis*).

amieti species group

This species group contains tetraploid, octoploid, and dodecaploid species, the latter of which are entirely or almost entirely independently evolved (e.g., Evans et al., 2015). This provides an exceedingly unusual opportunity to explore vertebrate niche evolution, with replication, through multiple sequences of sequential allopolyploidization events. The general pattern observed in this group is a narrowing of niche breadth as ploidy level increases and several examples of niche shifts and niche novelty. The three extant allotetraploids in this group all inhabit the Central African Congo Basin and Coastal lowland rainforest of Central Africa: *X. pygmaeus*, *X. allofraseri* and *X. parafraseri*. Evolution of these species was accompanied by a shift from the ancestral niche to much higher annual precipitation (Figure 5.5), an expansion to slightly lower temperature seasonality (Appendix IV Figure 1), and a contraction to slightly lower precipitation seasonality (Appendix IV Figure 2). For all climatic variables tested, the ancestral niche of the clade was projected to be most similar to the inland species *X. pygmaeus*, with contractions to a smaller niche breadth for the other inland species *X. parafraseri* for all variables, and an expansion of niche breadth for the coastal species, *X. allofraseri*, to lower mean annual temperatures (Figures 5.4 and 5.6), higher temperature seasonality (Appendix IV Figure 1), and higher annual precipitation (Figure 5.5). The exception to this is in precipitation seasonality where *X. pygmaeus* expanded to a lower seasonality (Appendix IV Figure 2).

Octoploids in the *amieti* group

Two possibly extinct ancestors (Ancestors A and B), predicted by our analyses to inhabit high elevation grassland in the Albertine Rift, hybridized with ancestral members of the Central African low elevation rainforest *X. pygmaeus*, *X. allofraseri* and *X. parafraseri* clade to form

three sets of octoploid daughter species. The octoploid species *X. boumbaensis*, *X. andrei* and *X. amieti* are found in lowland rainforest of Central Africa, coastal West-Central Africa, and at high elevations in the Cameroon Volcanic Line (Figures 5.1 and 5.2). These species experienced major shifts in mean annual temperature (Figure 5.4), annual precipitation (Figure 5.5), and precipitation seasonality (Appendix IV Figure 2) compared to the high elevation Albertine Rift parental species, but only very minor niche contractions or no change in these variables when compared to their lowland rainforest Central African parental species. *Xenopus boumbaensis*, despite being a lowland species, has an extremely small known range size and thus a small niche space, and shows major contractions in niche in all of these variables. *Xenopus andrei* is a coastal species and experienced a major expansion into higher annual precipitations (Figure 5.5). *Xenopus amieti,* a species with a highly restricted range in the Cameroon Volcanic Line, exhibits high levels of uncertainty in its fundamental niche and little can be said about its niche evolution.

The fundamental niche of the high elevation Albertine Rift *X. lenduensis* - *X. vestitus* species pair was difficult to characterize because both species have extremely restricted (known) ranges with large amounts of niche uncertainty, and are parented by two unknown ancestral species, Ancestor A and Ancestor B. Based on the available information, we projected these species to occupy mid-level mean annual temperatures and precipitations (Figures 5.4 and 5.5), medium to low precipitation seasonalities, and low temperature seasonalities (Appendix IV Figure 1 and 2).

Another pair of octoploid species formed by the lowland Central African clade and the highland Albertine Rift clade is *X. itombwensis* and *X. wittei*. Both are high elevation forest and forest-grassland species in the Albertine Rift (Figures 5.1 and 52). In contrast to octoploids from west Central Africa (*X. boumbaensis*, *X. andrei* and *X. amieti*), *X. itombwensis* and *X. wittei*

show major niche shifts away from their lowland Central African parental species in terms of a shift to lower mean annual temperatures (Figure 5.4), expansion into lower temperature seasonality (Appendix IV Figure 1), and a contraction to lower annual precipitation (Figure 5.5), while showing little niche differentiation from their high elevation Albertine Rift parental species. Overall, octoploids inhabit a diversity of niches but show a general pattern of narrower realized niche breadth relative to tetraploids in subgenus *Xenopus,* however many occupy restricted ranges, leading to higher uncertainty in their fundamental niche characterization.

Dodecaploids in the *amieti* group

Three sets of dodecaploids were formed through allopolyploidization between various tetraploid ancestors of lowland Central African *X. pygmaeus*, *X. allofraseri* and/or *X. parafraseri* and various octoploid ancestors of *X. boumbaensis*, *X. amieti*, and/or *X. andrei* (Evans et al., 2015). The *X. eysoole-X. longipes* sister species pair was formed through hybridization, most likely between a low elevation rainforest octoploid ancestor and an unknown tetraploid Ancestor C (Evans et al., 2015), and both inhabits high elevation forest and forest-grassland in the Cameroon Volcanic Line. For mean annual temperature, the most recent common ancestor of this species pair shows a completely novel niche compared to *X. boumbaensis*, *X. amieti*, and *X. andrei*, that is greatly restricted to the lower end of the niche of the lowland Central African ancestor (Figure 5.4). A similar pattern is seen in annual precipitation with a completely novel niche compared to the octoploid ancestor and a contraction to higher precipitations compared to the lowland Central African clade (Figure 5.5). *Xenopus ruwenzoriensis*, an independently evolved dodecaploid, is an inhabitant of low elevation rainforest in the Albertine Rift. This species is a small range endemic with high levels of uncertainty in its niche characterization, it

was projected to have a niche expansion to lower precipitation seasonality compared to the ancestors of *X. boumbaensis*, *X. amieti*, and *X. andrei* (Figure 5.4; Appendix IV Figure 1), as well as an expansion on both ends of the niche spectrum for annual precipitation and precipitation seasonality (Figure 5.5, Appendix IV Figure 2). It was not projected to have major niche differences from *X. pygmaeus* for any of the niche variables. *Xenopus kobeli*, another independently evolved dodecaploid is an inhabitant of low elevation rainforest in the Congo Basin. This species shows expansions into lower mean annual temperature from the ancestor of *X. boumbaensis*, *X. amieti*, and *X. andrei* (Figure 5.4). A similar pattern is seen for annual precipitation but with a contraction into higher niche values compared to the octoploid ancestor, but a restricted niche compared to the ancestor of *X. parafraseri* and *X. allofraseri* (Figures 5.5, Appendix IV Figure 1). Overall, dodecaploids show a very similar pattern of niche utilization to octoploids in this species group by exhibiting utilization of a variety of niches, but at generally narrow realized niche breadths. Similar to octoploids, however, many occupy restricted ranges, leading to high levels of uncertainty in their fundamental niche characterization.