Environmental variation predicts patterns of phenotypic and genomic variation in an African tropical forest frog

Courtney Miller[1], Geraud Tasse Taboue[2], Eric Fokam[2], Katy Morgan[1], Ying Zhen[3], Ryan Harrigan[3], Vinh Le Underwood[3], Kristen Ruegg[3], Paul Sesink Clee[4], Stephan Ntie[5], Patrick Mickala[5], Jean Mboumba[5], Trevon Fuller[6], Breda Zimkus[7], Thomas Smith[3], and Nicola Anthony[1]

[1]University of New Orleans
[2]University of Buea
[3]University of California Los Angeles
[4]Drexel University
[5]Universite des Sciences et Techniques de Masuku
[6]University of California, Los Angeles
[7]Harvard University

August 3, 2020

Abstract

Central African rainforests are predicted to be disproportionately affected by future climate change. How species will cope with these changes is unclear, but rapid environmental changes will likely impose strong selection pressures. Here we examined environmental drivers of phenotypic and genomic variation in the central African puddle frog (Phrynobatrachus auritus) to identify areas of elevated environmentally-associated turnover where populations may have the greatest capacity to adapt. We also compared current and future climate models to pinpoint areas of high genomic vulnerability where allele frequencies will have to shift the most in order to keep pace with future climate change. Analyses of body size, relative leg length, and head shape suggest that seasonal aspects of temperature and precipitation significantly influence phenotypic variation, whereas geographic distance and precipitation seasonality are the most important drivers of SNP allele frequency variation. However, neither landscape barriers nor the effects of past Pleistocene refugia had any influence on genomic differentiation. Most phenotypic and genomic differentiation coincided with key ecological gradients across the forest-savanna ecotone, montane areas and a coastal to interior rainfall gradient. Areas of greatest vulnerability were found in the lower Sanaga basin and southeastern region of Cameroon. In contrast with past conservation efforts that have focused on hotspots of species richness or endemism, our findings highlight the importance of preserving environmentally heterogeneous landscapes to preserve putatively adaptive variation and ongoing evolutionary processes in the face of climate change.

Running title: Mapping environmentally-associated variation in a tropical forest frog

Courtney A. Miller[1][*], Geraud Canis Tasse Taboue[2,3], Eric B. Fokam[2], Katy Morgan[1], Ying Zhen[4], Ryan J. Harrigan[4], Vinh Le Underwood[4], Kristen Ruegg[4], Paul R. Sesink Clee[5], Stephan Ntie[6], Patrick Mickala[6], Jean Francois Mboumba[6], Trevon Fuller[4], Breda M. Zimkus[7], Thomas B. Smith[4,8], Nicola M. Anthony[1]

1. Department of Biological Sciences, University of New Orleans, New Orleans, LA, USA 2. Department of Zoology and Animal Physiology, University of Buea, Buea, Cameroon 3. Institute of Geological and Mining

1

Research, Yaoundé, Cameroon

4. Institute of Environment and Sustainability, University of California, Los Angeles, CA, USA

5. Department of Biology, Drexel University, Philadelphia, Pennsylvania, USA

6. Department of Biology, University of Science and Technology of Masuku, Franceville, Gabon

7. Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA

8. Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA

ABSTRACT Central African rainforests are predicted to be disproportionately affected by future climate change. How species will cope with these changes is unclear, but rapid environmental changes will likely impose strong selection pressures. Here we examined environmental drivers of phenotypic and genomic variation in the central African puddle frog (*Phrynobatrachus auritus*) to identify areas of elevated environmentally-associated turnover where populations may have the greatest capacity to adapt. We also compared current and future climate models to pinpoint areas of high genomic vulnerability where allele frequencies will have to shift the most in order to keep pace with future climate change. Analyses of body size, relative leg length, and head shape suggest that seasonal aspects of temperature and precipitation significantly influence phenotypic variation, whereas geographic distance and precipitation seasonality are the most important drivers of SNP allele frequency variation. However, neither landscape barriers nor the effects of past Pleistocene refugia had any influence on genomic differentiation. Most phenotypic and genomic differentiation coincided with key ecological gradients across the forest-savanna ecotone, montane areas and a coastal to interior rainfall gradient. Areas of greatest vulnerability were found in the lower Sanaga basin and southeastern region of Cameroon. In contrast with past conservation efforts that have focused on hotspots of species richness or endemism, our findings highlight the importance of preserving environmentally heterogeneous landscapes to preserve putatively adaptive variation and ongoing evolutionary processes in the face of climate change.

1. INTRODUCTION

The tropical forests of the Congo Basin and Gulf of Guinea represent one of the most biologically diverse regions in the world. This region ranks third in plant, mammal, bird, and amphibian species richness after the Amazon and New Guinea (Mittermeier et al., 2003). With respect to amphibians, the Cameroonian highlands are recognized as one of the world's most important biodiversity hotspots (Amiet, 2008; Gvoždík et al., 2020; Herrmann et al., 2005). Several hypotheses have been advanced to explain the high biodiversity in this region. Previous phylogeographic studies have shown that past Pleistocene refugia shaped population structure in several central African rainforest species, providing support for their role as potential engines of diversification (Anthony et al., 2007; Born et al., 2011; Murienne et al., 2013; Nicolas et al., 2011; Plana, 2004; Quérouil et al., 2003). Alternatively, the riverine barrier hypothesis has argued that rivers could have led to the isolation and diversification of tropical forest species (Colyn et al., 1991; Wallace, 1854). Support for this hypothesis has been found in primates (Anthony et al., 2007; Mitchell et al., 2015; Telfer et al., 2003), birds (Aleixo, 2004) and rodents (Nicolas et al., 2011). In contrast, far less attention has been paid to the impact of historical barriers to gene flow or the effects of environmental variation (i.e. precipitation, temperature, vegetation density, etc.) on patterns of amphibian diversification in this region.

Several general hypotheses have been advanced to explain body size variation in amphibians. First, intraspecific amphibian body size has been shown to increase with decreasing temperatures, consistent with the heat balance hypothesis which predicts a negative relationship between body size and measures of energy availability (Boaratti & Da Silva, 2015; Ficetola et al., 2010; Rivas et al., 2018), although this has been found to not always be the case (Goldberg et al., 2018; Guo et al., 2019). Second, several studies have shown that larger bodied animals are found in drier areas (Goldberg et al., 2018; Guo et al., 2019; Olalla-Tárraga et al.,

2009) providing support for the water availability hypothesis which predicts that larger bodied animals are less prone to desiccation due to reduced surface-to-volume ratio. In contrast, the converse water availability hypothesis argues that body size is positively related to water availability such that larger animals will be found in areas with greater precipitation (Zug et al., 2001). Finally, the primary productivity hypothesis posits that larger body sizes will be found in areas with greater resource availability (Olalla-Tárraga & Rodríguez, 2007). In support of this hypothesis, Ficetola et al. (2010) found that body size increased in areas of higher primary productivity suggesting that less productive areas may have insufficient food supplies to support larger body sizes. However, aside from a study by (Charles et al. (2018) on the African leaf folding frog *Afrixalus paradorsalis* , few studies have addressed morphological variation in central African tropical frogs.

Environmental variation can act as a strong agent of diversifying selection, particularly in areas of high environmental heterogeneity (Endler, 1973), such as that observed across ecotones (Freedman et al., 2010; Smith et al., 1997; Termignoni-García et al., 2017) or across different levels of elevation (Thomassen et al., 2011). In central Africa, environmental variation has been shown to explain patterns of genetic differentiation in olive sunbirds (Smith et al., 2011), little greenbuls (Smith et al., 1997; Zhen et al., 2017), skinks (Freedman et al., 2010), chimpanzees (Mitchell et al., 2015), forest antelope (Ntie et al., 2017), soft-furred mice (Morgan et al., 2020), and reed frogs (Bell et al., 2017). These heterogeneous environments may capture ecological and evolutionary processes that are fundamental to maintaining and generating biological diversity (Moritz et al., 2000).

One major challenge is being able to effectively partition the effects of isolation by environment (IBE) from other potential drivers of population differentiation, namely: isolation by distance (IBD) and isolation by resistance due to landscape barriers (IBB). Advances in landscape genomics can be used to simultaneously assess the relative importance of competing ecological and historical drivers on genomic differentiation (Manthey & Moyle, 2015; Termignoni-García et al., 2017). Specifically, Fitzpatrick and Keller (2015) have shown that Generalized Dissimilarity Modelling (GDM; Ferrier et al., 2007) and Gradient Forests (GF; Ellis et al., 2012) can be powerful tools for analyzing gene–environment associations at the landscape level. Under a model of IBE, genetic differentiation increases with environmental differences, independent of geographic distance (Shafer & Wolf, 2013; Wang & Bradburd, 2014). In contrast, under a model of IBD, genetic differentiation is predicted to increase as a function of geographic distance whereas IBB is driven by landscape barriers to animal or plant dispersal (Balkenhol et al., 2017; Cushman & Schwartz, 2006). Resistance distances due to barriers between populations can be based on landscape features that may influence gene flow, such as habitat type (Tucker et al., 2017), physical barriers such as rivers (Mitchell et al., 2015) or the effects of past refugia (Ntie et al., 2017), which hereafter will be referred to as isolation by Pleistocene refugia (IBP). In many of these cases, circuit theory is used to incorporate IBE and IBB into models of population connectivity and identify which variables are the most important predictors of gene flow (McRae & Beier, 2007).

Central Africa faces a variety of threats from human activities and is especially vulnerable to climate change (Abernethy et al., 2013; James et al., 2013; Laporte et al., 2007; Oates et al., 2004). Temperatures are expected to rise along with potential shifts in rainfall patterns, including more intense dry seasons that could result in forest retreat (James et al., 2013). Species in this region, if they are to survive, would therefore be forced to respond to climate change either through dispersal, evolutionary adaptation or phenotypic plasticity (Davis et al., 2005; Holt, 1990). Given the threat that climate change poses to many species, there is now an increasing need to identify current and historical drivers of evolutionary diversification and recognize key areas for future conservation where species capacity to adapt is greatest (Anthony et al., 2015). Mapping landscape-level predictions of environmentally-associated genomic variation under both current and projected future environments can shed light on both the ability of populations to persist in their current state as well as their future capacity to respond to change through evolutionary adaptation (Gunderson, 2000; Sgro et al., 2010; Thrush et al., 2009). In this regard, the "genomic vulnerability" (Bay et al., 2018; Ruegg et al., 2018) has been used as a measure of the degree of "mismatch", or "offset", between current and future projections of environmentally-associated genomic variation and can be used as a proxy for population vulnerability to environmental change (Ruegg et al., 2018). Genomic vulnerability estimates how strongly allele frequencies

3

would have to change to keep track with the environmental changes predicted to occur at a certain location. Thus, the locations with the greatest vulnerability are those with the strongest predicted changes in allele frequencies, since these are the places where adaptation will have to keep up with environmental change the most.

In the present study, we used a combination of statistical methods to determine the potential drivers of phenotypic and genomic diversification in the widespread puddle frog, *Phrynobatrachus auritus* across its range in the west Central African countries of Cameroon, Equatorial Guinea, and Gabon. We used geospatial modeling to map patterns of genomic turnover (i.e. the change in allele frequencies with geographic distance) and predict areas of elevated genomic vulnerability. *P. auritus* serves as an ideal model for examining the effects of environmental heterogeneity because it occurs in a variety of forest types (Zimkus & Schick, 2010)and occupies a wide range of environmental conditions. Findings from these methods were then used to address the following hypotheses: 1) Temperature and precipitation are the most important drivers of morphological differentiation 2) IBE will influence genomic differentiation more than IBB or IBP 3) Areas of greatest environmentally-associated genomic turnover are associated with strong environmental gradients across the landscape 4) Patterns of environmentally-associated genomic variation reflect those observed for phenotypic variation 5) Genomic vulnerability will be highest in areas expected to undergo the greatest environmental change.

2. MATERIALS AND METHODS

2.1 Field sampling

Frogs were sampled between the months of March and July in 2013, 2014, and 2015. We collected a total of 191 *P. auritus* from four sites in Cameroon (Campo Ma'an (CM), N = 29; Ebo forest (EF), N = 24; Ndikiniméki (ND), N = 13; and Takamanda (TM), N = 15), and five sites in Gabon (Gamba Complex (GC), N = 24; Kessala (KS), N = 21; Lopé (LP), N = 18; Monts de Cristal (MC), N = 22; and Minkébé (MK), N = 25) (Fig. 1a).These sites encompassed a range of forest types, namely: lowland rainforest (CM, MK), sub-montane rainforest (EF, TM), forest-savanna ecotone (KS, MC), coastal rainforest (GC), and mixed lowland-agricultural forest (ND). After frogs were euthanized with MS222 solution, muscle and kidney tissue was placed in 95% ethanol for DNA extraction. Male frogs were differentiated from female frogs by the presence of a developed vocal sac, indicated by a dark throat and vocal folds. Frogs that could not be sexed by these secondary sexual characteristics were dissected and sexed by either the type of reproductive organ or by the presence of eggs. All animal handling procedures were carried out according to an approved Institutional Animal Care and Use Committee protocol 12-008. Specimens were deposited at the Museum of Comparative Zoology, Harvard University.

2.2 Environmental datasets

We extracted environmental data for each sampling location from publicly available geospatial raster layers (Table S1). These 26 environmental data layers included 19 bioclimatic variables summarizing the mean, maximum, minimum and range values of temperature and precipitation across the study region (Fick & Hijmans, 2017), four vegetation layers (Huete et al., 2002), elevation (Rodriguez et al., 2006), net primary productivity (NPP; Running et al., 2015), and potential evapotranspiration (PET; Mu et al., 2011). We then tested for collinearity in these raster layers (Table S2) using the 'removeCollinearity' function in the R package *virtualSpecies* to select a subset of variables where no two variables had a Pearson's correlation coefficient > 0.7, resulting in the following twelve uncorrelated variables: isothermality (i.e. temperature evenness; BIO3), minimum temperature of coldest month (BIO6), mean temperature of driest quarter (BIO9), precipitation seasonality (BIO15), precipitation of wettest quarter (BIO16), precipitation of warmest quarter (BIO18), precipitation of coldest quarter (BIO19), elevation, percent tree cover, leaf area index (LAI), NPP, and PET. Future projections of bioclimatic variables were taken from aggregated global climate models (Sesink Clee, 2017) for two representative concentration pathways (RCPs) 2.6 and 8.5, projected for year 2080 based on the Intergovernmental Panel on Climate Change (IPCC) 5th assessment report. RCP 2.6 represents a "best case" scenario in that global mean temperature is projected to rise by 0.3 to 1.7 °C by the late-21st

century, whereas RCP 8.5 represents a "worst case" scenario and projects global mean temperatures to rise by 2.6 to 4.8°C. All climate layers have a spatial resolution of 30 arc seconds (approximately $1km^2$).

2.3 Morphometric data and analysis

The following seven linear morphological characters were measured in 151 preserved adult frogs (100 males, 51 females): snout-vent length (SVL), tibia length (TL), femur length (FL), head length (HL), head width (HW), radio-ulnar length (RU), and hand length (HDL) according to Castellano and Giacoma (1998). Measurements were taken solely by the first author and repeated three times. We discarded the first set of measurements and tested the second and third measurements for repeatability (Pearson correlation coefficient ($r > 0.9$)). Once accepted, we used the third measurement for morphometric analyses.

We performed a PCA on the seven linear measurements in which PC1 was used as a proxy for body size. We also calculated relative leg length and generated a geometric morphometric variable for head shape using the program package SHAPE v.1.3 (Iwata & Ukai, 2002) based on photographs of 142 preserved frogs (97 males, 45 females) that were in sufficient condition. SHAPE traces the contour shape from an image, delineates the contour shape with elliptic Fourier descriptors (EFDs), and finally performs a principal component analysis of the EFDs to summarize the shape information (Iwata & Ukai, 2002). We retained the first principal component summarizing head shape for further analysis.

We used a random forest (RF) model within the *randomForest* R package (Liaw & Wiener, 2002) to determine the relative importance of each of the 12 uncorrelated environmental variables to body size, relative leg length, and head shape. RF is an ensemble learning method for nonlinear multiple regressions. When compared to similar approaches, RF consistently outperformed other methods (Cutler et al., 2007) and was among those least sensitive to spatial autocorrelation (Marmion et al., 2009). For each analysis, the data was first divided into training (70%) and testing (30%) sets to determine the optimal number of variables to split at each node in the tree before running a RF regression analysis based on 10,000 trees. Predictor variables were ranked in order of importance based on the number of times a given metric decreased the mean squared error (MSE) of the model. We then used the *rfPermute* R package to estimate the significance of importance metrics in all subsequent RF analyses. The response variable was permuted 1000 times on each of the 10,000 regression trees to create a null distribution against which the observed value was compared. Only significant ($p < 0.05$) environmental variables were retained in the final model that was used to extrapolate patterns of environmentally-associated morphological variation across the study area. To further explore the direction of associations between environmental predictor variables and body size, we performed a multiple linear regression model with significant predictor variables detected in RF modeling. Site origin was included as a factor in all analyses.

2.4 RAD-seq data

We extracted genomic DNA from either kidney or muscle tissue using Qiagen DNEasy Blood and Tissue kit (Qiagen, CA), following the manufacturer's protocol. A total of 164 individuals had genomic DNA of sufficient quantity (> 50 ng) and quality needed for restriction site-associated sequencing (RAD-seq) as determined through gel electrophoresis (Davey et al., 2011). RAD-seq library preparation followed the BestRad protocol for Illumina sequencing as described in Ali et al. (2016). Briefly, genomic DNA (100 ng) was digested with 4.8 units of *Sbf* I-HF restriction enzyme (New England Biolabs NEB, R3632L) at 37°C for 1 h in a 12 µl reaction volume. Samples were heated to 65°C for 20 min and 4 µl of the indexed BestRad SbfI P1 RAD adapter (10 nM) was added to each sample. Ligation of inline barcoded P1 adaptors to digested genomic DNA was performed overnight at 20°C with 640 units of T4 DNA ligase (NEB, M0202M), followed by incubation at 20°C for 60 minutes, then 65°C for 20 min. Following ligation, 10 µl of each sample in each 48 well plate was pooled into a single tube and cleaned using 1x Agencourt AMPure XP beads (A63881; Beckman Coulter). Pooled DNA for each plate was then resuspended in 100 µl low TE and sheared to an average fragment size of 500 base pairs using a Bioruptor NGS sonicator (Diagenode). Sheared DNA was then concentrated to 55.5 µl using Ampure XP beads and used as the template in the NEBNext Ultra DNA Library Prep Kit for Illumina (NEB E7370L; v.1.2). The standard NEBNext protocol for library preparation was followed

5

except that we used custom P2 adaptors which were created by annealing a NEBNext Multiplex Oligo for Illumina (NEB, E7335L) to the oligonucleotide GATCGGAAGAGCACACGTCTGAACTCC AGTCACIIII-IIATCAGAACA*A (where * represents a phosphorothioate DNA base). In addition, instead of the USER® enzyme step, we used a universal P1 RAD primer (AATGATACGGCGACCACCGAGATCTAC ACTCTTT-CCCTACACGAC*G) and a universal P2 RAD primer (CAAGCAGAAGACGGCATACG*A) during final amplification. The final RAD library was cleaned using AMPure XP beads and sequenced at the UC Berkeley QB3 Vincent J Coates Genome Sequencing Laboratory (GSL) on an Illumina HiSeq2500: Rapid Run Mode (Illumina, San Diego, CA, USA) using paired-end 100-bp sequence reads.

2.5 Bioinformatics analysis of RAD-seq data

We used the bioinformatics software pipeline, STACKS v.1.44 (Catchen et al., 2013; Catchen et al., 2011) to process the restriction-site-associated DNA markers (RAD-tags) and generate single nucleotide polymorphism (SNP) datasets. First, we executed the "process_radtags" program in STACKS to demultiplex and trim sequence reads by the P1 barcodes and remove low quality reads (Phred quality score less than 20). After removing PCR duplicates with the "clone_filter" script, the processed reads were used to generate RAD loci without a reference genome using "denovo_map.pl" (parameter settings: $m = 3$ $M = 5$ $n = 4$). We empirically determined these parameters to limit the impact of over-splitting loci (see Harvey et al., 2015; Ilut et al., 2014). This involved running the *de novo* assembly over a wide range of values of M (1–8) with "ustacks". From these runs, we selected a value of $M = 5$ since we observed that the percentage of homozygous and heterozygous loci reached a plateau at this value and thus minimized over-splitting of alleles for the final SNP calling.

Stacks calls SNPs ("sstacks") within RAD loci using a multinomial-based likelihood model that estimates the likelihood of two most frequently observed genotypes at each site and performs a standard likelihood ratio test using a chi-square distribution (Catchen et al., 2011; Hohenlohe et al., 2010). For SNP inference, we used the default alpha significance level of 0.05. Paralogous loci that stacked together were identified and removed by subsequent quality control steps built into STACKS (max number of stacks per loci (m) = 3; Harvey et al., 2015; Ilut et al., 2014). After the preliminary assembly of catalog loci using "denovo_map.pl", we ran the STACKS correction mode (rxstacks-cstacks-sstacks) using the bounded SNP model with a 0.05 upper bound for the error rate. The "rxstacks" program made corrections to genotype and haplotype calls based on population information, rebuilt the catalog loci and filtered out loci with average log likelihood ratio of < 8.0.

We used three additional filtering steps to generate a set of high-quality RAD loci for down- stream population genetic analysis. First, we retained only RAD loci that were present in 80% of all samples. Second, we removed RAD loci that contained more than 40 SNPs, as these likely represented sequencing errors or over-clustering of paralogous loci. Lastly, we used the BLAT alignment algorithm (Kent, 2002) to *de novo* align the RAD loci and removed those that aligned to multiple positions. The final consensus set of RAD loci comprised SNP data from a total of 139 individuals. Genotypes were called, filtered, and bi-allelic SNPs were exported in VCF format using the STACKS "populations" program. SNPs from the last seven bp of the RAD loci were removed as this part of the locus is likely to contain sequence errors at the 3' end of the reads. The SNP dataset was further filtered with VCFtools v.0.1.14 (Danecek et al., 2011) to remove SNPs below a minor allele frequency (MAF) of 0.05 cutoff to reduce artifacts of sequence and assembly error. The dataset was also filtered to include only one random SNP per RAD locus for use in FastStructure (Raj et al., 2014) in order to avoid linkage disequilibrium between SNPs within RAD loci.

2.6 Analyses of population genomic structure

We performed a principal component analysis (PCA) using the Bioconductor package SNPRelate (Zheng et al., 2012); https://www.bioconductor.org/) to summarize population genomic structure. We used the program fastStructure to estimate the number of genetically distinct populations within the sampled *P. auritus* range. We tested a range of K values (where K denotes the number of inferred populations) from 1 to 10. The script "choose.py" included in the FastStructure package was used to determine the best

6

estimate of K that maximizes the marginal likelihood. We also calculated pairwise estimates of F_{ST} (Weir & Cockerham, 1984) among sites and among K populations inferred from FastStructure using VCFtools. To test for an IBD effect, a Mantel test was used to assess the correlation between pairwise F_{ST} values and geographic distance. Mantel tests were run with 999,999 permutations using VEGAN2.2-1 in R (Oksanen et al., 2018) and are reported using both raw F_{ST} and transformed $F_{ST}/(1-F_{ST})$ distances, as well as both raw Euclidian geographic distance and log-transformed Euclidian distances (Rousset, 1997; Slatkin, 1995).

2.7 Quantifying the relative impact of IBE on genomic differentiation

We used GDM to compare the importance of IBE to isolation by landscape barriers (IBB) or Pleistocene refugia (IBP) on patterns of genomic turnover, as detailed in Appendix 1. GDM is a matrix regression technique that evaluates the relationship between site-site dissimilarities in environmental or landscape 'predictor' variables and a biotic 'response' variables (e.g. pairwise genetic distances). A major advantage of GDM over other modelling methodologies is that it can fit non-linear relationships between environmental variables and the biological response variable through the use of I -spline basis functions (Ferrier et al., 2007). This approach can also incorporate a range of environmental data layers, resistance surfaces, and straight-line geographic distance as different predictors.

Pairwise dissimilarity in genomic composition between sites was modeled using two measures: 1) pairwise F_{ST} values and 2) a pairwise Bray-Curtis dissimilarity index based on the presence or absence of a SNP at each locus. IBE was represented by the set of 12 uncorrelated environmental variables described previously. In addition to these environmental variables, a set of predictor variables were generated to model the effect of landscape barriers (elevation and rivers) and hypothesized Pleistocene refugia under the Last Glacial Maximum (LGM) approximately 21,000 years ago. Pairwise resistance distances for IBB were generated by creating raster layers of resistance surfaces based on landscape features, elevation and rivers, using the raster calculator available in QGIS v.2.18. We then calculated pairwise resistance distances from these raster layers with CIRCUITSCAPE 4.0 (McRae et al., 2013). Two IBB matrices were generated, IBB1 and IBB2. For IBB1, resistance values increased with increasing elevation and rivers were treated as impenetrable. For IBB2, resistance increased with increasing elevation and also with Strahler order which reflects size and strength of perennial river systems. For IBP, we first projected habitat suitability for *P. auritus* under climate conditions during the LGM using two global climate models (MIROC and CCSM). We then created resistance surfaces where resistance was considered to be inversely proportional to habitat suitability, and finally, calculated pairwise resistance distances from this raster layer with CIRCUITSCAPE. Further details on how these predictor variables were generated can be found in Appendix 1. We ran four models for each genomic dataset with different configurations of these predictor variables: 1) IBE, IBB1, IBP-MIROC, 2) IBE, IBB1, IBP-CCSM, 3) IBE, IBB2, IBP-MIROC, 4) IBE, IBB2, IBP-CCSM.

2.8 Mapping genomic turnover and predicting patterns of genomic vulnerability under future climate change

We evaluated the importance of environmental variables as predictors of environmentally-associated genomic turnover and spatialized these patterns across the study region using GF modeling within the package *gradientForest* in R (Ellis et al., 2012) (Appendix 1). Response variables were individual SNP minor allele frequencies within each population (only SNPs with MAF above 0.05 were used, N=3092). Predictor variables were represented by the same environmental variables that were included in the GDM along with latitude and longitude. GF uses a machine-learning algorithm to divide the biological data into different bins (i.e. different values of allele frequencies), with partitions occurring at several split values along each environmental variable. This binning is performed for every SNP, weighting each SNP individually according to its fit to the model (i.e. R^2) before aggregating across all SNPs. GF determines the "split importance' by measuring the amount of biological variation explained by a given split value (e.g. between 26 and 27°C), which is then cumulatively summed along each gradient to construct turnover functions (Fitzpatrick & Keller, 2015). The top three environmental variables in modeling genomic turnover from a total of 2000 regression trees were used to predict and map environmentally-associated turnover across the study region using a random grid of 100,000 sample points. To ensure that our GF model was performing better than random, we shuffled the environmental-predictor matrix to generate 200 randomized datasets and compared the number of SNP loci

7

with R^2 positive values to the mean R^2 value across SNP loci using GF models describing variation in the real versus randomized datasets.

Lastly, we predicted future environmentally-associated genomic variation based on GF models under projected climate change for the year 2080, RCPs 2.6 and 8.5, representing "best" and "worst" cases, respectively. To map predicted changes of genomic variation associated with environment, we subtracted future GF predictions from current predictions using QGIS v.2.18. Areas where environmentally-associated genomic variation changed the least are considered to have low genomic vulnerability while areas where they change the most are considered to have high genomic vulnerability.

3. RESULTS

3.1 Morphometric analyses

PC1 for linear body measurements described 86.4% of the variation and based on factor loadings, was used as a proxy for overall body size (Table S3). For head shape, PC1 explained 53% and was retained for further analyses. We found no significant correlation between pairwise differences in PC1 for linear body measurements and geographic distance between sites (Mantel r = 0.22; mantel simulated p-value > 0.05). Similarly, there was no relationship between geographic distance and relative leg length (Mantel r = 0.26; mantel simulated p-value > 0.05) or PC1 for head shape (Mantel r = 0.10; mantel simulated p-value > 0.05).

The RF model for body size explained 49% of the total variation. Significant environmental variables for this RF model comprised two measures of temperature variation (isothermality and minimum temperature of the coldest month), two measures of precipitation (precipitation seasonality and precipitation of the coldest quarter), elevation and latitude (Fig. S1). Site origin was also significant in the model, although it was the least important predictor. The results of the follow-up linear regression indicated that four predictors explained 50% of the variance in body size (R2 =0.50, F(7,143)=22.91, p<.001). These were: isothermality (β= 0.81, p<0.001), precipitation seasonality (β= -0.53, p<0.001), precipitation of the coldest quarter (β= -0.003, p<0.05), and site (β= -0.44, p<0.01). The RF model for relative leg length explained 44% of the variation with isothermality, precipitation seasonality, precipitation of the coldest quarter, precipitation of the warmest quarter, treecover and latitude as significant predictors. The RF model for head shape explained 59% of the variation and was explained by the following variables: isothermality, mean temperature of the driest quarter, minimum temperature of the coldest month, precipitation of the warmest quarter, wettest quarter, and coldest quarter, latitude and longitude (Fig. S1).

Projected variation in body size is greatest across the Cameroon highlands and forest-savanna ecotone as well as between the coast and interior of Gabon (Fig. 2a). Body size increases with distance from the equator with the exception of the southern coast of Gabon where body size decreased sharply towards the coast. Relative leg length also generally increased with distance from the equator, however it is relatively uniform in areas of continuous forest (Fig. 2b). This phenotypic variable also exhibits the greatest variation from the southern coast of Gabon moving inland. Head shape shares a similar pattern of variation across the forest-savanna ecotone of Cameroon as does body size but also exhibits a shift from wider to narrow heads moving away from the equator (Fig. 2c). In contrast to body size and relative leg length, there was little variation in head shape between the coast and interior of Gabon.

3.2 SNP variation

RAD-sequencing of 139 *P. auritus* samples generated a total of 838,425,400 reads after filtering out low-quality samples and reads. The number of raw sequence reads per sample ranged from 133,185 to 16 million. The mean coverage depth ranged from 5x to 26x across individual samples (mean = 8x, median = 7x, Appendix 2). From these reads, we assembled 2,979 high-quality RAD loci and a total of 32,966 SNPs that were present in 80% or more samples. Using a minor allele frequency cutoff of 5%, we retained 1631 RAD loci encompassing 3,092 SNPs. Across sampling sites, the number of segregating sites ranged from 1,834 to 13,335 (mean = 8,473) and nucleotide diversity (π) ranged from 0.0041 to 0.0062/bp (mean = 0.0055/bp).

3.3 Population genomic structure

8

The PCA identified significant population structure across the sampled range of *P. auritus* . PC1 explained 25.89% of the variation and separated the three northern sites (EF, ND, TM) from the remaining six sites (CM, MC, MK, LP, KS, GC). PC2 explained 9.26% of variation and separated GC, the southernmost coastal site, from all other sites (Fig. 1b). FastStructure analyses also revealed a pattern of population structure that is organized latitudinally into five distinct populations: 1) EF, ND, and TM, 2) CM, 3) MC and MK, 4) KS and LP, and 5) GC.

Pairwise F_{ST} values between sites ranged from 0.017 to 0.450 (mean = 0.234; Appendix 3), indicating low to moderate levels of genomic differentiation between sites. We found a significant correlation between pairwise F_{ST} and geographic distances between the sites (Mantel r = 0.6642; mantel simulated p-value = 0.001), suggesting a strong pattern of IBD.

3.4 Quantifying the relative impact of IBE, IBD, IBB, and IBP on genomic differentiation

The GDM based on pairwise F_{ST} values explained 83% of the variation in the model. Geographic distance, precipitation seasonality, precipitation of warmest quarter, precipitation of coldest quarter, and percent tree cover were all significant predictor variables in the full models. In contrast, when SNP presence-absence data was used as the biological response variable, the full models explained 97% of the variation and only geographic distance, precipitation of the warmest quarter, and precipitation of the coldest quarter were significant variables (Table 1). Neither IBB or IBP were significant in any of the models.

3.5 Patterns of genomic turnover and genomic vulnerability across environmental space

We used a GF approach to determine associations between SNPs and environmental variables and map environmentally-associated genomic turnover across the total study area. A total of 1396 SNPs (45% of all SNPs) had R^2 values > 0 (0.01-0.704, average 0.33). When testing model performance, the number of SNPs with R^2 values > 0 for all of the randomized datasets fell below the number observed for the real data (Fig. S2) and the mean R^2 value generated for the real dataset fell within the upper 95% quartile of values generated for the randomized datasets (Fig. S3), both indicating that the GF model shows a stronger association between environmental and genomic variation for our dataset relative to the set of randomized datasets. Precipitation of the coldest quarter, precipitation of the warmest quarter, and latitude were the most important environmental predictors of genomic turnover (Fig. 3b). Projected associations between allele frequencies and precipitation of the coldest quarter, precipitation of the warmest quarter, and latitude revealed areas of pronounced genomic turnover across the Cameroonian highlands (pink to orange), forest-savanna ecotone of south-central Cameroon (orange to green) and across the equator (green to blue) (Fig. 3c). There was also a moderate gradient from the coast to the interior of Gabon (dark blue to light blue).

Predictions of environmentally associated genomic turnover under future climate change projections showed the same general pattern of genotype-environment associations across the landscape relative to current predictions (Fig. S4). However, there were notable differences in patterns of genomic turnover between present and future climate change scenarios that were relatively consistent across both RCPs. Subtracting the current predictions from the future predictions under climate change projections (RCP 2.6 & RCP 8.5) for the year 2080 revealed two distinct hotspots of genomic vulnerability (Fig. 3d & Fig. S5): one centered around the lower Sanaga basin and a second in the far southeastern region of Cameroon.

4. DISCUSSION

We adopted a comprehensive statistical approach to disentangling the effects of environment, geographic distance, and landscape barriers on phenotypic and genomic variation in the African puddle frog *P. auritus* . Overall, we find that environmental variation plays an important role in shaping patterns of morphological and genomic differentiation. This is in addition to, but independent of, geographic distance. In particular, seasonal patterns of temperature and precipitation appear to be key in driving patterns of diversification in this tropical region, in keeping with a recent meta-analysis conducted of environmentally-mediated selection across the tropics (Siepielski et al., 2017). We also find that environmentally heterogeneous landscapes are important generators of patterns of high phenotypic and genomic variation suggesting that they may play

9

an important role in promoting and maintaining biodiversity.

Our first hypothesis posited that temperature and precipitation are the most important drivers of morphological differentiation. We found that temperature evenness and precipitation seasonality were significant predictors of all measures of phenotypic variation. Both temperature and precipitation are known to be important factors influencing amphibian development, growth, and population dynamics and are expected to be key determinants of survival under climate change (Ficetola & Maiorano, 2016; Pounds et al., 1999). Our linear regression results show that body size increases with temperature evenness and decreases with precipitation seasonality. Thus, larger bodies are expected to be found in more uniform habitats with less variation in temperature and precipitation. This is somewhat consistent with the converse water availability hypothesis such that in areas with less intense wet and dry periods there is likely more water available year-round, allowing for investment in growth.

Body size, relative leg length, and head shape are predicted to vary along the transition zone between forest and savanna in central Cameroon where patterns of elevated morphological divergence have also been reported in a sunbird (Smith et al., 2011). The Cameroonian highlands and coastal regions of the Gulf of Guinea also appear to be associated with variation in body size and head shape, possibly due to strong ecological gradients associated with elevation and precipitation in these regions. Differences in skull shape morphology have been linked to the type, size, and speed of prey consumed in frogs and other amphibians (Emerson, 1985; Kaczmarski et al., 2017; Van Buskirk & Schmidt, 2000; Vega-Trejo et al., 2014) indicating that head shape might be at least partly adaptive. Head morphology has also been shown to exhibit considerable developmental plasticity in response to changes in temperature and could have important consequences for post-larval survival (Tejedo et al., 2010). There are also strong gradients predicted in both body size and relative leg length differentiation from the coast to the interior of Gabon, likely influenced by the degree of variation found in the population from Gamba. These frogs may be an example of cryptic speciation considering they have smaller body sizes but larger relative leg lengths compared to the rest of the samples and also given their high levels of genomic differentiation from other sites. There is also evidence that *P. auritus* exhibits complex patterns of spatial niche partitioning (Zimkus et al., 2010) and extraordinary patterns of diversification (Gvoždík et al., 2020). Although we cannot disentangle the effects of genetic adaptation and phenotypic plasticity, it is important to note that tropical ectotherms are considered to be particularly sensitive to changes in temperature and/or precipitation so that even subtle shifts in these variables could have profound impacts on fitness (Deutsch et al., 2008; Ficetola & Maiorano, 2016).

Our second hypothesis stated that IBE will influence genomic differentiation more than IBB or IBP. Contrary to many phylogeographic studies that have been carried out previously in central Africa, we did not find evidence for an effect of landscape barriers or Pleistocene refugia on population genomic differentiation. These findings are in stark contrast to many previous studies that have placed emphasis on the role of Pleistocene refugia and/or rivers (Anthony et al., 2007; Bohoussou et al., 2015; Eriksson et al., 2004; Nicolas et al., 2011) with the exception of Bell et al. (2017) where rivers were not important in reed frog diversification. However, our findings provide strong support for the role of environment, specifically seasonal variation in patterns of precipitation, as the most important environmental factor. Geographic distance is also consistently identified as a strong predictor of genomic differentiation. The role of IBD was supported by findings from our Mantel tests, the significance of geographic distance in GDM, and the significance of latitude, but not longitude, in predicting genomic turnover in GF analyses.

Patterns of environmentally-associated genomic differentiation reported here are consistent with previous investigations of gene-environmental associations in this region. For example, precipitation has been shown to be an important predictor of patterns of genetic variation in central African lizards (Freedman et al., 2010), chimpanzees (Mitchell et al., 2015), birds (Smith et al., 2011), and forest antelope (Ntie et al., 2017). In the present study, precipitation of the coldest quarter is highest in the Cameroon highlands, and decreases progressively throughout central Cameroon and Gabon (Fig. S6a), mirroring shifts in genomic turnover observed in *P. auritus.* Conversely, precipitation of the warmest quarter is highest in most of Gabon, especially along the coast and decreases towards Cameroon (Fig. S6b). Both of these patterns demonstrate shifts in

genomic differentiation throughout the highlands, across the equator, and subtly from coastal to inland Gabon. Gradients in rainfall not only shape the distribution of forest cover but also present potentially strong selection pressures on the phenology of *P. auritus* since the timing and duration of amphibian reproductive events are very sensitive to rainfall levels (Corn, 2005; Ficetola & Maiorano, 2016).

GDM also identified precipitation seasonality as a significant predictor of genomic turnover. This environmental variable is linked to seasonal patterns in rainfall availability that are inverted across the Equator separating Cameroon and Gabon. Rainforests either side of the equator have their own distinct seasonal patterns of rainfall (Heuertz et al., 2014) such that the dry season in central Cameroon coincides with the rainy season in northern Gabon and vice versa. This seasonal inversion could be responsible for the shift in genomic variation observed in *P. auritus* across the equator. It has been hypothesized that these contrasting patterns of seasonal rainfall could lead to reproductive isolation and speciation across this region (Heuertz et al., 2014). Future work should look more closely at the seasonal inversion hypothesis and how heterogeneous annual patterns of rainfall influence genomic differentiation in other rainforest species.

Our third hypothesis was that areas of greatest environmentally-associated genomic turnover are associated with strong environmental gradients across the landscape. Areas of elevated genomic turnover in *P. auritus* appear to correspond to known ecological gradients. Genomic turnover is predicted to be high throughout the forest-savanna ecotone region south of the montane region in Cameroon where rainforest habitat in the south gradually transitions to savanna in the north. These findings are consistent with patterns of high intraspecific genomic diversity across this ecotonal region in the rainforest bird *Andropadus virens* (Zhen et al., 2017) and soft-furred mouse *Praomys misonnei* (Morgan et al., 2020). There is also high genomic turnover in *P. auritus* across the Cameroon highlands, reflecting both elevation and distance from the coast. The Cameroon highlands are a known biodiversity hotspot, especially for amphibian richness and endemism (Amiet, 2008; Herrmann et al., 2005; Pauwels & Rodel, 2007; Zimkus & Gvoždík, 2013) so that elevated genomic turnover in this region is to be expected. Mountain ranges and elevational gradients are often recognized as important drivers of genetic heterogeneity and, as is the case here, are important for the conservation of evolutionary potential. Overall, our results show a strong role for environment in shaping genomic differentiation such that areas of elevated genomic turnover span regions of strong ecological transition, providing further support for the role of environmental gradients and ecotones in shaping adaptive diversification.

Our fourth hypothesis stated that patterns of environmentally-associated genomic variation reflect those observed for phenotypic variation. Patterns of environmentally-associated morphological and genomic variation are relatively similar in Cameroon. RF and GF projections suggest that the forest-savanna ecotone in Cameroon is predicted to result in elevated morphological and genomic variation. These projections are based on the aggregated effects of the significant environmental variables which are primarily related to seasonality. Western Cameroon is characterized by more densely forested areas with pronounced precipitation seasonality, which then transitions south to habitats that include both forest and savanna and experience especially high seasonal variability in temperature and precipitation (Sesink Clee et al., 2015; Smith et al., 2011). Our findings are consistent with previous examples indicating that seasonality in moisture levels and precipitation are key explanatory variables for both morphological and genomic variation within this region (Smith et al., 2011), and thus provides further support for the role of environmental variation in driving diversification. In Gabon, we find relatively uniform patterns of genomic variation relative to patterns of morphological variation. While seasonal variation is less pronounced relative to Cameroon, Gabon harbors a variety of heterogeneous habitats, such as narrow, coastal alluvial plains, extensive wetlands, patches of savanna, and low elevation mountain zones (Lee et al., 2006), many of which may present unique selection pressures contributing to phenotypic variation.

Finally, we posited that genomic vulnerability is predicted to be highest in areas expected to undergo the greatest environmental change. We identified several areas of high genomic vulnerability where populations may be more susceptible to climate change under future projections. In the present study, the Sanaga River delta area in southwest Cameroon and an area in the southeast of the country, north of Lobéké and Nki National Parks, are predicted to be regions of greatest genomic vulnerability. While most of the

11

Sanaga River is unprotected, the Douala Edea Wildlife Reserve falls within this area and constitutes an important target for continued protection. Genomic vulnerability may be an important metric to incorporate into conservation prioritization as it may also indicate areas where populations are already susceptible to present-day environmental pressures. For example, Bay et al. (2018) have recently shown that yellow warbler (*Setophaga petechia*) populations with the highest genomic vulnerability were also experiencing the largest population declines. Therefore, areas of high genomic turnover and vulnerability may be important targets for future conservation efforts since the former serves as centers of high adaptive potential whereas the latter signal susceptibility to environmental change.

Although we adopted a genome-wide approach in the present study, our SNP dataset is only likely to capture a fraction of the total number of loci in the genome that constitute targets for selection and/or regions of the genome that may be linked loci under selection. Further research should focus on linking genotypic variation to phenotypic traits under selection to more fully understand the evolutionary significance of divergence across ecological gradients as well as examine the relative importance of genetic versus environmental factors in contributing to the observed morphological variation.

Understanding the ecological and historical processes involved in diversification is important not only for increasing our knowledge of evolutionary mechanisms, but also for making evolutionarily informed conservation decisions to protect biodiversity and prioritize new area for preservation in the light of rapid climate change. By taking a robust statistical approach to disentangling competing drivers of differentiation, we show that environmental factors rather than historical barriers to gene flow are largely responsible for patterns of morphological differentiation and genomic turnover in our study species. These findings, therefore, highlight the importance of preserving heterogeneous environments, such as environmental gradients, in maintaining species adaptive evolutionary potential and underline the importance of considering evolutionary processes in the design of future protected areas.

REFERENCES

Abernethy, K. A., Coad, L., Taylor, G., Lee, M. E., & Maisels, F. (2013). Extent and ecological consequences of hunting in Central African rainforests in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* ,*368* (1625).

Aleixo, A. (2004). Historical diversification of a terra-firme forest bird superspecies: A phylogeographic perspective on the role of different hypotheses of amazonian diversification. *Evolution* ,*58* (6), 1303–1317.

Ali, O. A., O'Rourke, S. M., Amish, S. J., Meek, M. H., Luikart, G., Jeffres, C., & Miller, M. R. (2016). RAD capture (Rapture): flexible and efficient sequence-based genotyping. *Genetics* , *202* (2), 389–400.

Amiet, J. (2008). Frog Diversity in Cameroon. In S. Stuart, M. Hoffmann, J. Chanson, N. Cox, R. Berridge, P. Ramani, & Y. BE (Eds.),*Threatened amphibians of the world. Lynx Edicions with IUCN—The World Conservation Union, Conservation International and NatureServe*(p. 16).

Anthony, N. M., Atteke, C., Bruford, M. W., Dallmeier, F., Freedman, A., Hardy, O., Ibrahim, B., Jeffery, K. J., Johnson, M., Lahm, S. A., Lepengue, N., Lowenstein, J. H., Maisels, F., Mboumba, J. F., Mickala, P.,

Morgan, K., Ntie, S., Smith, T. B., Sullivan, J. P., . . . Gonder, M. K. (2015). Evolution and Conservation of Central African Biodiversity: Priorities for Future Research and Education in the Congo Basin and Gulf of Guinea. *Biotropica* , *47* (1), 6–17.

Anthony, N. M., Johnson-Bawe, M., Jeffery, K., Clifford, S. L., Abernethy, K. A., Tutin, C. E., Lahm, S. A., White, L. J. T., Utley, J. F., Wickings, E. J., & Bruford, M. W. (2007). The role of Pleistocene refugia and rivers in shaping gorilla genetic diversity in central Africa. *Proceedings of the National Academy of Sciences of the United States of America* , *104* (51), 20432–20436.

Balkenhol, N., Dudaniec, R. Y., Krutovsky, K. V, Johnson, J. S., Cairns, D. M., Segelbacher, G., Selkoe, K. A., von der Heyden, S., Wang, I. J., Selmoni, O., & Joost, S. (2017). Landscape Genomics: Understanding Relationships Between Environmental Heterogeneity and Genomic Characteristics of Populations. In O. P. Rajora (Ed.), *Population Genomics: Concepts, Approaches and Applications* (pp. 261–322). Springer International Publishing.

Bay, R. A., Harrigan, R. J., Underwood, V. Le, Gibbs, L., Smith, T. B., & Ruegg, K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* ,*361* (6401), 83–86.

Bell, R. C., Parra, J. L., Badjedjea, G., Barej, M. F., Blackburn, D. C., Burger, M., Channing, A., Dehling, J. M., Greenbaum, E., Gvoždík, V., Kielgast, J., Kusamba, C., Lötters, S., McLaughlin, P. J., Nagy, Z. T., Rödel, M. O., Portik, D. M., Stuart, B. L., VanDerWal, J., . . . 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* (19), 5223–5244.

Boaratti, A. Z., & Da Silva, F. R. (2015). Relationships between environmental gradients and geographic variation in the intraspecific body size of three species of frogs (Anura). *Austral Ecology* ,*40* (8), 869–876.

Bohoussou, K. H., Cornette, R., Akpatou, B., Colyn, M., Kerbis Peterhans, J., Kennis, J., Šumbera, R., Verheyen, E., N'Goran, E., & Katuala, P. (2015). The phylogeography of the rodent genus Malacomys suggests multiple Afrotropical Pleistocene lowland forest refugia.*Journal of Biogeography* , *42* (11), 2049–2061.

Born, C., Alvarez, N., Doyle, M., 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* (1), 131–142.

Castellano, S., & Giacoma, C. (1998). Stabilizing and directional female choice for male calls in the European green toad. *Animal Behaviour* , *56* (2), 275–287.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics.*Molecular Ecology* , *22* (11), 3124–3140.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J. H., & De Koning, D.-J. (2011). Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *Genes Genomes Genetics* ,*1* (3), 171–182.

Charles, K. L., Bell, R. C., Blackburn, D. C., Burger, M., Fujita, M. K., Gvoždík, V., Jongsma, G. F. M., Kouete, M. T., Leaché, A. D., & 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 rain forest. *Journal of Biogeography* , *45* (8), 1781–1794.

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* (4), 403–407.

Corn, P. S. (2005). Climate change and amphibians. *Animal Biodiversity and Conservation* , *28* (1), 59–67.

Cushman, S. A., & Schwartz, M. K. (2006). Gene Flow in Complex Landscapes: Testing Multiple Hypotheses with Causal Modeling.

ScholarWorks at University of Montana.

Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology* , *88* (11), 2783–2792.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics* , *27* (15), 2156–2158.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* , *12* (7), 499–510.

Davis, M. B., Shaw, R. G., & Etterson, J. R. (2005). Evolutionary responses to changing climate. *Ecology* , *86* (7), 1704–1714.

Deutsch, C. A., Tewksbury, J. J., Huey, R. B., Sheldon, K. S., Ghalambor, C. K., Haak, D. C., & Martin, P. R. (2008). Impacts of climate warming on terrestrial ectotherms across latitude.*Proceedings of the National Academy of Sciences of the United States of America* , *105* (18), 6668–6672.

Ellis, N., Smith, S. J., & Roland Pitcher, C. (2012). Gradient forests: Calculating importance gradients on physical predictors. *Ecology* ,*93* (1), 156–168.

Emerson, S. B. (1985). Skull Shape in Frogs: Correlations with Diet.*Herpetologica* , *41* (2), 177–188.

Endler, J. A. (1973). Gene Flow and Population Differentiation.*Science* , *179* (4070), 243–250.

Eriksson, J., Hohmann, G., Boesch, C., & Vigilant, L. (2004). Rivers influence the population genetic structure of bonobos (Pan paniscus).*Molecular Ecology* , *13* (11), 3425–3435.

Ferrier, S., Manion, G., Elith, J., & Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* , *13* (3), 252–264.

Ficetola, G. F., & Maiorano, L. (2016). Contrasting effects of temperature and precipitation change on amphibian phenology, abundance and performance. *Oecologia* , *181* (3), 683–693.

Ficetola, G. F., Scali, S., Denoël, M., Montinaro, G., Vukov, T. D., Zuffi, M. A. L., & Padoa-Schioppa, E. (2010). Ecogeographical variation of body size in the newt Triturus carnifex: Comparing the hypotheses using an information-theoretic approach. *Global Ecology and Biogeography* , *19* (4), 485–495.

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* (12), 4302–4315.

Fitzpatrick, M. C., & Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters* ,*18* , 1–16.

Freedman, A. H., Thomassen, H. a., Buermann, W., & Smith, T. B. (2010). Genomic signals of diversification along ecological gradients in a tropical lizard. *Molecular Ecology* , *19* (17), 3773–3788.

Goldberg, J., Cardozo, D., Brusquetti, F., Bueno Villafañe, D., Caballero Gini, A., & Bianchi, C. (2018). Body size variation and sexual size dimorphism across climatic gradients in the widespread treefrog Scinax fuscovarius (Anura, Hylidae). *Austral Ecology* ,*43* (1), 35–45.

Gunderson, L. H. (2000). Ecological Resilience-In Theory and Application. *Annual Review of Ecology and Systematics* , *31* , 425–439.

Guo, C., Gao, S., Krzton, A., & Zhang, L. (2019). Geographic body size variation of a tropical anuran: Effects of water deficit and precipitation seasonality on Asian common toad from southern Asia. *BMC Evolutionary Biology* , *19* (1), 1–11.

Gvoždík, V., Nečas, T., Dolinay, M., Zimkus, B. M., Schmitz, A., & Fokam, E. B. (2020). Evolutionary history of the Cameroon radiation of puddle frogs (Phrynobatrachidae: Phrynobatrachus), with descriptions of two critically endangered new species from the northern Cameroon Volcanic Line. *PeerJ* , *2020* (3).

Harvey, M. G., Judy, C. D., Seeholzer, G. F., Maley, J. M., Graves, G. R., & Brumfield, R. T. (2015). Similarity thresholds used in DNA sequence assembly fromshort reads can reduce the comparability of population histories across species. *PeerJ* , *2015* (3), 1–16.

Herrmann, H.-W., Böhme, W., Euskirchen, O., Hermann, P. A., & Schmitz, A. (2005). African biodiversity hotspots: the amphibians of Mt Nlonako, Cameroon. *Salamandra* , *41* , 61–81.

Heuertz, M., Duminil, J., Dauby, G., Savolainen, V., & Hardy, O. J. (2014). Comparative phylogeography in rainforest trees from lower Guinea, Africa. *PLoS ONE* , *9* (1).

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* , *6* (2).

Holt, R. D. (1990). The microevolutionary consequences of climate change. *Trends in Ecology & Evolution* , *5* (9), 311–315.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* , *83* (1), 195–213.

Ilut, D. C., Nydam, M. L., & Hare, M. P. (2014). Defining loci in restriction-based reduced representation genomic data from nonmodel species: Sources of bias and diagnostics for optimal clustering. *BioMed Research International* , *2014* .

Iwata, H., & Ukai, Y. (2002). SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *The Journal of Heredity* , *93* (5), 384–385.

James, R., Washington, R., & Rowell, D. P. (2013). Implications of global warming for the climate of African rainforests. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* , *368* (1625), 20120298.

Jari Oksanen, F. G. B., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., & Stevens, M. H. H. (2018). Vegan: community ecology package. *R Package Version* , *2* (6).

Kaczmarski, M., Kubicka, A. M., Hromada, M., & Tryjanowski, P. (2017). Robustness of newt heads in condition of co-existence: a case of the Carpathian newt and the alpine newt. *Zoomorphology* , *136* (4), 511–521.

Kent, W. J. (2002). BLAT—The BLAST-Like Alignment Tool. *Genome Research* , *12* (4), 656–664.

Laporte, N. T., Stabach, J. A., Grosch, R., Lin, T. S., & Goetz, S. J. (2007). Expansion of Industrial Logging in Central Africa. *Science* , *316* , 1451.

Lee, M. E., Alonso, A., Dallmeier, F., Campbell, P., & Pauwels, O. S. (2006). The Gamba complex of protected areas: an illustration of Gabon's biodiversity. *Bulletin of the Biological Society of Washington* , *12* , 229–241.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News* , *2* (3), 18–22.

Manthey, J. D., & Moyle, R. G. (2015). Isolation by environment in white-breasted nuthatches (Sitta carolinensis) of the Madrean Archipelago sky islands: A landscape genomics approach. *Molecular Ecology* , *24* (14), 3628–3638.

Marmion, M., Luoto, M., Heikkinen, R. K., & Thuiller, W. (2009). The performance of state-of-the-art modelling techniques depends on geographical distribution of species. *Ecological Modelling* ,*220* (24), 3512–3520.

McRae, B. H., Shah, V. B., Mohapatra, T. K., & Anantharaman, R. (2013). Circuitscape 4. *Seattle WA: The Nature Conservancy* .

McRae, B. H., & Beier, P. (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America* , *104* (50), 19885–19890.

[dataset] Miller, Courtney (2020), Data from: Environmental variation predicts patterns of phenotypic and genomic variation in an African tropical forest frog, v2, Dryad, Dataset, https://doi.org/10.5061/dryad.m129c0s

Mitchell, M. W., Locatelli, S., Clee, P. R. S., Thomassen, H. A., & Gonder, M. K. (2015). Environmental variation and rivers govern the structure of chimpanzee genetic diversity in a biodiversity hotspot.*BMC Evolutionary Biology* , *15* (1), 1–13.

Mittermeier, R. A., Mittermeier, C. G., Brooks, T. M., Pilgrim, J. D., Konstant, W. R., Fonseca, G. A. B., & Kormos, C. (2003). Wilderness and biodiversity conservation. *Proceedings of the National Academy of Sciences of the United States of America* , *100* (18), 10309–10313.

Morgan, K., Mboumba, J. F., Ntie, S., Mickala, P., Miller, C. A., Zhen, Y., Harrigan, R. J., Le Underwood, V., Ruegg, K., Fokam, E. B., Tasse Taboue, G. C., Sesink Clee, P. R., Fuller, T., Smith, T. B., & Anthony, N. M. (2020). Precipitation and vegetation shape patterns of genomic and craniometric variation in the central African rodent Praomys misonnei.*Proceedings. Biological Sciences* , *287* , 20200449.

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

Mu, Q., Zhao, M., & Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment* , *115* (8), 1781–1800.

Murienne, J., Benavides, L. R., Prendini, L., Hormiga, G., & Giribet, G. (2013). Forest refugia in Western and Central Africa as "museums" of Mesozoic biodiversity. *Biology Letters* , *9* (1).

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* (1), 191–207.

Ntie, S., Davis, A. R., Hils, K., Mickala, P., Thomassen, H. A., Morgan, K., Vanthomme, H., Gonder, M. K., & Anthony, N. M. (2017). Evaluating the role of Pleistocene refugia, rivers and environmental variation in the diversification of central African duikers (genera Cephalophus and Philantomba). *BMC Evolutionary Biology* , *17* (212), 1–17.

Oates, J. F., Bergl, Ri. A., & Linder, J. M. (2004). Africa's Gulf of Guinea Forests: Biodiversity Patterns and Conservation Priorities.*Advances in Applied Biodiversity Science* , *6* , 1–95.

Olalla-Tárraga, M. Á., Diniz-Filho, J. A. F., Bastos, R. P., & Rodríguez, M. Á. (2009). Geographic body size gradients in tropical regions: Water deficit and anuran body size in the Brazilian Cerrado.*Ecography* , *32* (4), 581–590. https://doi.org/10.1111/j.1600-0587.2008.05632.x

Olalla-Tárraga, M. Á., & Rodríguez, M. Á. (2007). Energy and interspecific body size patterns of amphibian faunas in Europe and North America: Anurans follow Bergmann's rule, urodeles its converse.*Global Ecology and Biogeography* , *16* (5), 606–617.

Pauwels, O. S. G., & Rodel, M. (2007). Amphibians and National Parks in Gabon , western Central Africa Amphibien und Nationalparks in Gabun , westliches Zentralafrika. *Herpetozoa* , *19* , 135–148.

Plana, V. (2004). Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. *Philosophical Transactions of the Royal Society B: Biological Sciences* , *359* (1450), 1585–1594.

Pounds, J. A., Fogden, M. P. L., & Campbell, J. H. (1999). Biological response to climate change on a tropical mountain. *Nature* ,*398* (6728), 611–615.

Quérouil, S., Verheyen, E., Dillen, M., & Colyn, M. (2003). Patterns of diversification in two African forest shrews: Sylvisorex johnstoni and Sylvisorex ollula (Soricidae, Insectivora) in relation to paleo-environmental changes. *Molecular Phylogenetics and Evolution* , *28* (1), 24–37.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). FastSTRUCTURE: Variational inference of population structure in large SNP data sets.*Genetics* , *197* (2), 573–589.

Rivas, J., Quiero, A., Penna, M., & Velásquez, N. A. (2018). Body-Size Variation across Environmental Gradients in an Ectothermic Organism: An Intraspecific Approach to Ecogeographic Patterns. *Herpetologica* ,*74* (3), 191–198.

Rodriguez, E., Morris, C., & Belz, J. (2006). An assessment of the SRTM topographic products. *Photogrammetric Engineering and Remote Sensing* , *72* (3), 249–260.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* ,*145* (4), 1219–1228.

Ruegg, K., Bay, R. A., Anderson, E. C., Saracco, J. F., Harrigan, R. J., Whitfield, M., Paxton, E. H., & Smith, T. B. (2018). Ecological genomics predicts climate vulnerability in an endangered southwestern songbird. *Ecology Letters* , *21* (7), 1085–1096.

Running, S., Mu, Q., & Zhao, M. (2015). MOD17A2H MODIS/terra gross primary productivity 8-day L4 global 500m SIN grid V006. *NASA EOSDIS Land Processes DAAC* .

Sesink Clee, P. R. (2017). Distributions, Drivers, and Risks of Wildlife Infectious Diseases across Africa: using geospatial analyses to elucidate disease occurrence in biodiversity hotspots. (Doctoral dissertation). Drexel University.

Sesink Clee, P. R., Abwe, E. E., Ambahe, R. D., Anthony, N. M., Fotso, R., Locatelli, S., Maisels, F., Mitchell, M. W., Morgan, B. J., Pokempner, A. A., & Gonder, M. K. (2015). Chimpanzee population structure in Cameroon and Nigeria is associated with habitat variation that may be lost under climate change. *BMC Evolutionary Biology* ,*15* (1), 1–13.

Sgro, C. M., Hoffmann, A. A., & Lowe, A. J. (2010). Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications* , *4* , 326–337.

Shafer, A. B. A., & Wolf, J. B. W. (2013). Widespread evidence for incipient ecological speciation: A meta-analysis of isolation-by-ecology. *Ecology Letters* , *16* (7), 940–950.

Siepielski, A. M., Morrissey, M. B., Buoro, M., Carlson, S. M., Caruso, C. M., Clegg, S. M., Coulson, T., DiBattista, J., Gotanda, K. M., Francis, C. D., Hereford, J., Kingsolver, J. G., Augustine, K. E., Kruuk, L. E. B., Martin, R. A., Sheldon, B. C., Sletvold, N., Svensson, E. I., Wade, M. J., & MacColl, A. D. C. (2017). Precipitation drives global variation in natural selection. *Science* , *355* (6328), 959–962.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* , *139* (1), 457–462.

Smith, T. B., Thomassen, H. A., Freedman, A. H., Sehgal, R. N. M., Buermann, W., Saatchi, S., Pollinger, J., Milá, B., Pires, D., Valkiunas, G., & Wayne, R. K. (2011). Patterns of divergence in the olive sunbird

17

Cyanomitra olivacea (Aves: Nectariniidae) across the African rainforest-savanna ecotone. *Biological Journal of the Linnean Society* , *103* (4), 821–835.

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

Tejedo, M., Marangoni, F., Pertoldi, C., Richter-Boix, A., Laurila, A., Orizaola, G., Nicieza, A. G., Álvarez, D., & Gomez-Mestre, I. (2010). Contrasting effects of environmental factors during larval stage on morphological plasticity in post-metamorphic frogs. *Climate Research* , *43* (1–2), 31–39.

Telfer, P. T., Souquière, S., Clifford, S. L., Abernethy, K. A., Bruford, M. W., Disotell, T. R., Sterner, K. N., Roques, P., Marx, P. A., & Wickings, E. J. (2003). Molecular evidence for deep phylogenetic divergence in Mandrillus sphinx. *Molecular Ecology* , *12* (7), 2019–2024.

Termignoni-García, F., Jaramillo-Correa, J. P., Chablé-Santos, J., Liu, M., Shultz, A. J., Edwards, S. V., & Escalante-Pliego, P. (2017). Genomic footprints of adaptation in a cooperatively breeding tropical bird across a vegetation gradient. *Molecular Ecology* ,*26* (17), 4483–4496.

Thomassen, H. a., Fuller, T., Buermann, W., Milá, B., Kieswetter, C. M., Jarrín-V., P., Cameron, S. E., Mason, E., Schweizer, R., Schlunegger, J., Chan, J., Wang, O., Peralvo, M., Schneider, C. J., Graham, C. H., Pollinger, J. P., Saatchi, S., Wayne, R. K., & Smith, T. B. (2011). Mapping evolutionary process: A multi-taxa approach to conservation prioritization. *Evolutionary Applications* , *4* (2), 397–413.

Thrush, S. F., Hewitt, J. E., Dayton, P. K., Coco, G., Lohrer, A. M., Norkko, A., Norkko, J., & Chiantore, M. (2009). Forecasting the limits of resilience: Integrating empirical research with theory.*Proceedings of the Royal Society B: Biological Sciences* ,*276* (1671), 3209–3217.

Tucker, J. M., Allendorf, F. W., Truex, R. L., & Schwartz, M. K. (2017). Sex-biased dispersal and spatial heterogeneity affect landscape resistance to gene flow in fisher. *Ecosphere* , *8* (6).

Van Buskirk, J., & Schmidt, B. R. (2000). Predator-induced phenotypic plasticity in larval newts: trade-offs, selection, and variation in nature. *Ecology* , *81* (11), 3009–3028.

Vega-Trejo, R., Zuniga-Vega, J. J., & Langerhans, R. B. (2014). Morphological differentiation among populations of Rhinella marina (Amphibia: Anura) in western Mexico. *Evolutionary Ecology* ,*28* (1), 69–88.

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment.*Molecular Ecology* , *23* (23), 5649–5662.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* , 1358–1370.

Zhen, Y., Ng, T. C., Harrigan, R. J., Ruegg, K. C., Anderson, E. C., Lohmueller, K. E., & Smith, T. B. (2017). Genomic divergence across ecological gradients in the Central African rainforest songbird (Andropadus virens). *Molecular Ecology* , *February* , 1–12.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* ,*28* (24), 3326–3328.

Zimkus, B. M., & Gvoždík, V. (2013). Sky Islands of the Cameroon Volcanic Line: a diversification hot spot for puddle frogs (Phrynobatrachidae: Phrynobatrachus). *Zoologica Scripta* , 1–21.

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* (3), 883–900.

Zimkus, B. M., & Schick, S. (2010). Light at the end of the tunnel: insights into the molecular systematics of East African puddle frogs (Anura: Phrynobatrachidae). *Systematics and Biodiversity* ,*8* (1), 39–47.

Zug, G. R., Vitt, L. J., & Caldwell, J. P. (2001). *Herpetology* . Academic Press.

DATA ASSESSIBILITY

Sampling locations, environmental data, morphometric data, and SNP data are available on Dryad at DOI:10.5061/dryad.m129c0s.

TABLES & FIGURES

Table 1. GDM results using genomic data and environmental variables across all models. P-values and relative importance of significant variables in modeling of both pairwise F_{ST} values and genetic dissimilarity based on presence or absence of a SNP at a particular geographic site.

	Significance (p-values)	Significance (p-values)	Importance	Importance
Variable	*FST*	*SNP Presence/absence*	*FST*	*SNP Presence/*
Geographic distance	< 0.001 - 0.001	< 0.001	1.56	0.59
Precipitation seasonality	0.022-0.029	NS	2.85	NA
Precipitation of warmest quarter	0.037 - 0.044	0.001 - 0.004	1.20	11.61
Precipitation of coldest quarter	0.007 - 0.010	0.001 - 0.004	4.70	2.67
Percent Treecover	0.0020 - 0.0026	NS	1.96	NA

Figure 1. Field sampling and inferred population structure of *P. auritus* . (a) Map of nine sampling locations: Takamanda (TM), Ndikiniméki (ND), Ebo Forest (EF), Campo Ma'an (CM), Monts de Cristal (MC), Minkébé (MK), Lopé (LP), Kessala (KS) and Gamba Complex (GC). Each point is a sample site and the colors correspond to the assigned population. (b) PCA using SNPs that have a minor allele frequency > 5%. Each point presents a sample, and samples are colored by site with similar color shades corresponding to their assigned populations. (c) FastStructure results using SNPs that have a minor allele frequency > 5% and filtered to have only one random SNP per locus. Site abbreviations are labeled within each population.

Figure 2. Morphological variation from RF projections using morphometric data and environmental variables (a) Variation in body size, represented by the first principle component of a PCA of all seven linear morphometric measurements. (b) Variation in relative leg length. (c) Variation in head shape, represented by the first principle component of a PCA of geomorphometric head shape calculated in a SHAPE analysis. For these maps, the greater the color difference, the greater difference in morphology.

Figure 3. GF results with maps of genomic variation and genomic vulnerability. (a) PC plot with labelled vectors indicating the direction and relative magnitude of environmental variables with the greatest contribution to the predicted patterns of SNP allele frequency differentiation. (b) Environmental and geographic variables ranked by their importance in explaining SNP allele frequency variation. (c) Map of the gradient forest model of environmentally-associated SNPs for *P. auritus* . Larger color differences between any two areas in the landscape correspond to larger genetic differences. Circles indicate sampling sites (d) Genomic vulnerability. Red indicates greater changes in allele frequencies and thus, higher genomic vulnerability, while blue indicates smaller changes in allele frequencies and less genomic vulnerability.

(a)

(b)

(c)