


Environmental correlates of phylogenetic endemism in amphibians and the conservation of refugia in the Coastal Forests of Eastern Africa

Christopher D. Barratt^{1,2}  | Beryl A. Bwong^{1,3} | Renske E. Onstein² | Dan F. Rosauer⁴ | Michele Menegon⁵ | Nike Doggart⁶ | Peter Nagel¹ | W. Daniel Kissling² | Simon P. Loader^{1,7,8}

¹Biogeography Research Group, Department of Environmental Sciences, University of Basel, Basel, Switzerland

²Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Amsterdam, The Netherlands

³Herpetology Section, National Museums of Kenya, Nairobi, Kenya

⁴Research School of Biology and Centre for Biodiversity Analysis, Australian National University, Acton, ACT, Australia

⁵Tropical Biodiversity Section, Museo delle Scienze, Trento, Italy

⁶Tanzania Forest Conservation Group, Dar es Salaam, Tanzania

⁷Department of Life Sciences, University of Roehampton, London, UK

⁸Department of Life Sciences, Natural History Museum, London, UK

Correspondence

Christopher D. Barratt, Biogeography Research Group, Department of Environmental Sciences, University of Basel, Basel, Switzerland.
Email: c.d.barratt@gmail.com

Funding information

Humer Foundation via the Centre for African Studies Basel; Stipendienkommission für Nachwuchskräfte; COSTECH, Grant/Award Number: 2013-341-NA-2013-121; Freiwillige Akademische Gesellschaft Basel (FAG); Kenya Wildlife Service, Grant/Award Number: KWS/BRM/5001. University of Basel Travel Fund. Kenya Forest service permit, MUS/1/KFS/VOL.II/4

Editor: Jeremy Austin

Abstract

Aims: To quantify the spatial distribution of amphibian phylogenetic endemism (PE), an indicator of potential refugia, to test PE for correlations with current and historical environmental predictors, and to evaluate the effectiveness of current protected areas at conserving evolutionary history.

Location: Coastal Forests of Eastern Africa (CFEA) and the adjacent low-elevation Eastern Afromontane (EA).

Methods: We integrated new and existing spatial and phylogenetic data to map PE for almost the full amphibian assemblage (41 of 55 species), including 35 intraspecific lineages from several species and complexes showing high phylogeographic structure. Using spatial and non-spatial regressive models, we tested whether PE can be predicted by measures of Quaternary climate change, forest stability, topographic heterogeneity and current climate. PE results were intersected with the protected area network to evaluate current conservation effectiveness.

Results: We detect refugia in Tanzania and coastal Kenya previously identified as CFEA centres of endemism but also new areas (lowland Tanga region and Pangani river, Zaraninge forest, Mafia island, Matumbi hills). Results show that refugia for amphibians (high PE) are located in areas with long-term Quaternary climate stability and benign current climate (high precipitation of driest quarter, high annual precipitation), with climatically unstable areas demonstrating low PE. Conservation analyses revealed that ten PE hotspots account for over 25% of the total PE, but only small parts of these areas are under conservation protection.

Main Conclusions: Utilizing cryptic diversity from novel phylogeographic data and distribution modelling improves our understanding of endemism patterns, with climate stability being strongly correlated with the distribution of PE. Our analyses point towards high PE areas being refugia, which require an urgent need to consolidate protected areas within centres of endemism in this highly threatened biodiversity hotspot.

KEYWORDS

amphibians, cryptic diversity, niche models, refugia, spatial phylogenetics

1 | INTRODUCTION

Biological diversity is concentrated in the tropics but faces unprecedented anthropogenic impact, which has led to the earth being described as in the midst of a major sixth extinction event (Kolbert, 2014; Ceballos et al., 2015). Mitigating the loss of biodiversity is therefore an urgent priority, and explaining why some areas are biologically richer than others is a vital step towards prioritizing what should be conserved and why. At global and regional scales, high biodiversity and species endemism have been shown to be correlated with contemporary and historical climate regimes and topography (Kissling et al., 2012; Kissling, Blach-Overgaard, Zwaan, & Wagner, 2016; Ricklefs, Bermingham, Dick, & Moritz, 2005; Sandel et al., 2011), but comprehensive tests within biodiversity hotspots are mostly lacking. Africa holds rich biodiversity with over a quarter of the world's biodiversity

hotspots (Myers, Mittermeier, Mittermeier, Fonseca, & Kent, 2000), though often in small and fragmented areas. Major advances in understanding Africa's biodiversity have been made in recent years (e.g., Burgess, Mwasumbi, Hawthorne, Dickinson, & Doggett, 1992; Burgess, Clarke, & Rodgers, 1998; Jetz, Rahbek, & Colwell, 2004; Linder et al., 2012; Lorenzen, Heller, & Siegmund, 2012; Levinsky et al., 2013; Rovero et al., 2014), although many taxonomic groups remain under-sampled, with cryptic diversity often under-represented. Incorporating this intraspecific diversity within species is particularly important for understanding the correlates of biodiversity patterns and prioritizing conservation efforts at local scales.

The Coastal Forests of Eastern Africa (CFEA, Figure 1) are a global biodiversity hotspot, containing over 1,500 endemic plant species, and are regarded as a conservation priority. However, they remain severely threatened (Azeria, Sanmartin, As, Carlson, & Burgess, 2007; Burgess

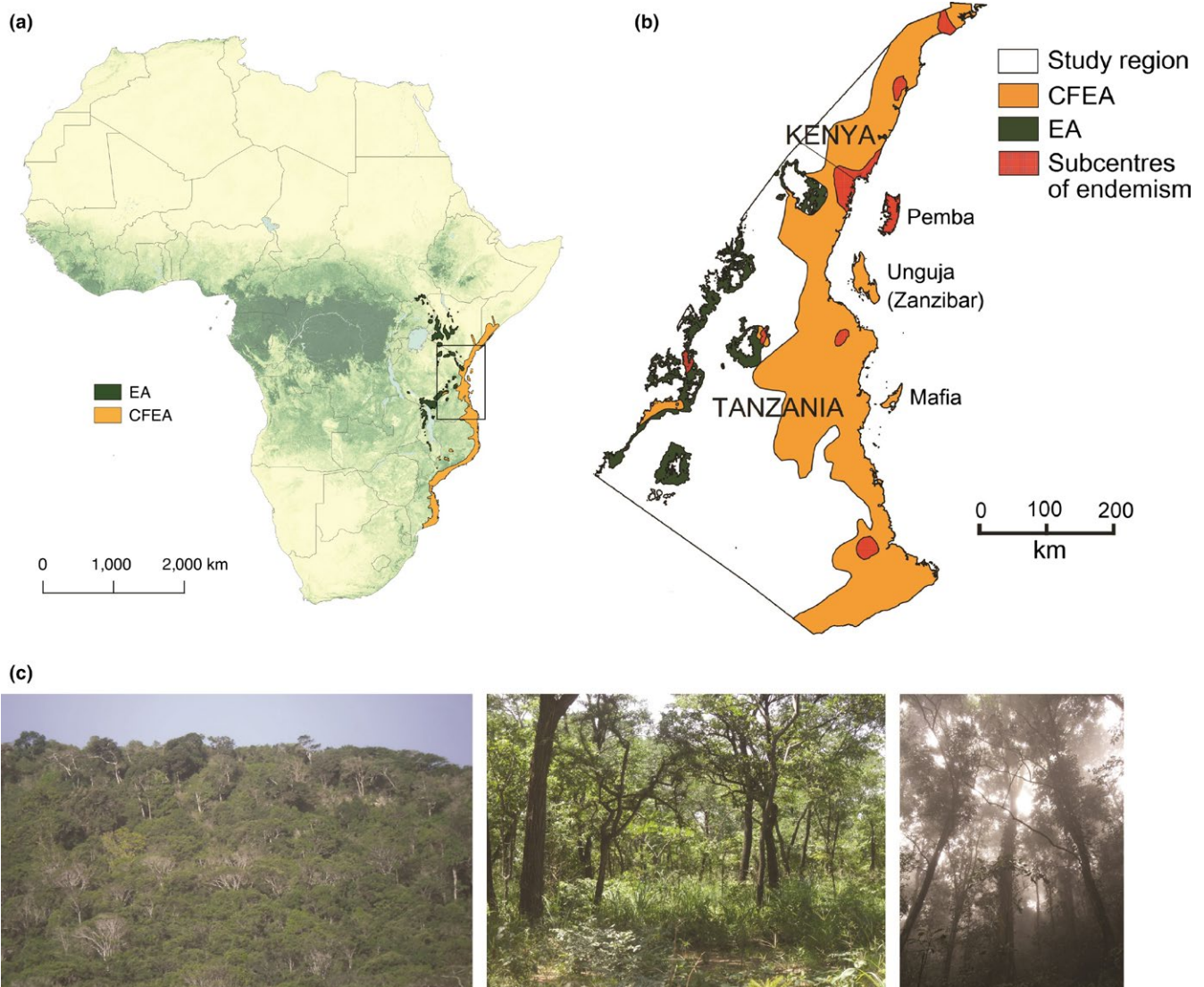


FIGURE 1 Distribution of forests in the study region within East Africa. (a) Map showing the location of the Coastal Forests of Eastern Africa (CFEA) and their proximity to the Eastern Afromontane (EA) region. (b) Study region encompassing major elements of CFEA and lowland EA forests as well as intervening habitat in Kenya and Tanzania, including sub-centres of endemism identified by Burgess et al. (1998). (c) Examples of CFEA, with coastal dry forest in Shimba hills, Kenya (left), miombo woodland in Kiwengoma FR, Tanzania (middle), and EA habitat represented by montane forest in Kitolomero FR, Udzungwa (right)

et al., 2007) and have been described as a “vanishing refuge” (Burgess et al., 1998). The majority of the CFEA hotspot is lowland (<300 m a.s.l.) with several raised areas along the coastline and inland where the CFEA overlaps with parts of the Eastern Afromontane (EA) region up to around 1,000 m (Udzungwa, Uluguru, Usambara in the Eastern Arc mountains, Tanzania, see Figure 1a,b). It is thought the fragmented lowland forests of the CFEA and EA region are the remnants of a once continuous forest that covered tropical Africa during the Early Tertiary (Couvreur, Chatrou, Sosef, & Richardson, 2008; Kissling et al., 2012). Historical climate change since the Miocene led to significant contraction and expansion of forests during glacial and interglacial periods (Axelrod & Raven, 1978; Trauth, Maslin, Deino, & Strecker, 2005; Maslin et al., 2014), subjecting many low-elevation areas to repeated inundation from sea level fluctuations (Burgess & Clarke, 2000; Kent, Hunt, & Johnstone, 1971). Climate stability and topography are therefore likely to be important factors in the persistence of this region's biodiversity and the location of refugia (e.g., Haffer, 1969; Moreau, 1933; Mayr & O'Hara, 1986).

Previously, sub-centres of endemism within the Swahili regional centre of endemism across the CFEA were identified using species distributional data from a number of plant, vertebrate and invertebrate groups (Burgess et al., 1998; Figure 1c). However, phylogenetic information and environmental data have not yet been integrated to validate these findings. The inclusion of phylogenetic information to supplement traditional biodiversity measures such as species richness was popularized by Faith (1992), using phylogenetic diversity (PD) to measure an area's evolutionary history based on the sum of the branch lengths connecting each tip on a phylogenetic tree to its root. Rosauer, Laffan, Crisp, Donnellan, and Cook (2009) extended PD's utility by combining it with weighted endemism (WE; Crisp, Laffan, Linder, & Monro, 2001), to derive phylogenetic endemism (PE). By subdividing the length of each phylogenetic branch across the areas it occurs, PE indicates where substantial components of PD are spatially restricted. A particularly desirable property of PE is that it circumvents the need for formal taxonomic classification by utilizing branch lengths rather than species units. As PE can incorporate cryptic diversity, it has gained popularity in an increasing number of macroecological and conservation studies (Carnaval et al., 2014; Gudde, Joy, & Mooers, 2013; Laity et al., 2015; Rosauer, Catullo, Vanderwal, & Moussalli, 2015; Rosauer et al., 2016). Areas of high PE are typically found where phylogenetic branches with few close relatives are geographically restricted, and are important in a conservation context because they have been used to indicate refugia where a large proportion of evolutionary history has accumulated over time (Carnaval et al., 2014; Rosauer et al., 2009; Rosauer & Jetz, 2015).

Amphibians offer an excellent model to investigate biodiversity patterns due to their relatively poor dispersal abilities and sensitivity to habitat changes compared with other taxa (but see Van Bocxlaer et al., 2010). The known CFEA species assemblage of Tanzania consists of 51 species (Poynton et al., 2007), with several additional species from coastal Kenya (Harper, Measey, Patrick, Menegon, & Vonesh, 2010) or recently described (Barratt et al., 2017) inflating that number to at least 55. Although several of these species are narrow-ranged

endemics, many are widespread and occur across adjacent areas of suitable habitat (e.g., savanna and bushland). The systematics of many CFEA species are very poorly understood, but some show strong phylogeographic structure with poorly defined species boundaries (e.g., Channing et al., 2013). Phylogeographic data from amphibians have never been utilized for understanding biodiversity patterns in the CFEA, but given their suitability compared to traditional measures, they are likely to provide vital insights into this threatened biodiversity hotspot.

In this study, we integrate new and existing phylogenetic and spatial data for amphibians to investigate PE across a major part of the CFEA spanning Tanzania and Kenya. Utilizing cryptic diversity for the first time, we focus on combining species data with intraspecific diversity in widespread species (hereafter referred to as “lineages”) which likely indicate refugia where biodiversity has persisted during times of climatic instability. Given the extreme climate oscillations and forest cover change across the CFEA since at least the Miocene, we hypothesize that measures of climate stability have a major influence on phylogeographical patterns and the distribution of refugial areas for amphibians, and are good predictors of PE. Using almost the full species assemblage with spatial and phylogenetic data and including intraspecific diversity within multiple species and species complexes, we map the geographic distribution of evolutionary history (PE) and test our two main hypotheses: (1) that PE predicts refugia, and is positively correlated with long-term climatic stability, and (2) the current protected area network in the region does not effectively conserve refugia for amphibians.

2 | METHODS

2.1 | Lineage discovery and phylogeny

We aimed to integrate as much of the CFEA amphibian assemblage as possible for PE analyses. We therefore included molecular data from species but also from intraspecific lineages within species where genetic and spatial sampling permitted. To build a phylogeny representing the amphibian assemblage of the study region (Figure 2), we first compiled genetic data from recent fieldwork and museum sampling. Genetic samples were collected across the CFEA and lowland parts of the EA over a total of ten non-consecutive field seasons spanning 2001–2015. DNA was extracted from fresh tissue samples (leg muscle, liver or toe clip) collected in the field (2013–2015), or from samples held in collections (collected between 2001 and 2012) stored in 100% ethanol. We built a DNA barcoding database using the 16S rRNA mitochondrial gene fragment, screening over 1,500 individual samples. Species identifications were verified against published GenBank sequences, and our own and collaborators unpublished sequences (B. Zimkus, D. Portik, personal communication, September 1st 2015). We examined thoroughly sampled species and complexes with adequate spatial and genetic data to define the intraspecific lineages present within each group based on 16S data. Based on structure in neighbor-joining trees from 16S data, we then amplified an additional mitochondrial gene for a single representative per lineage

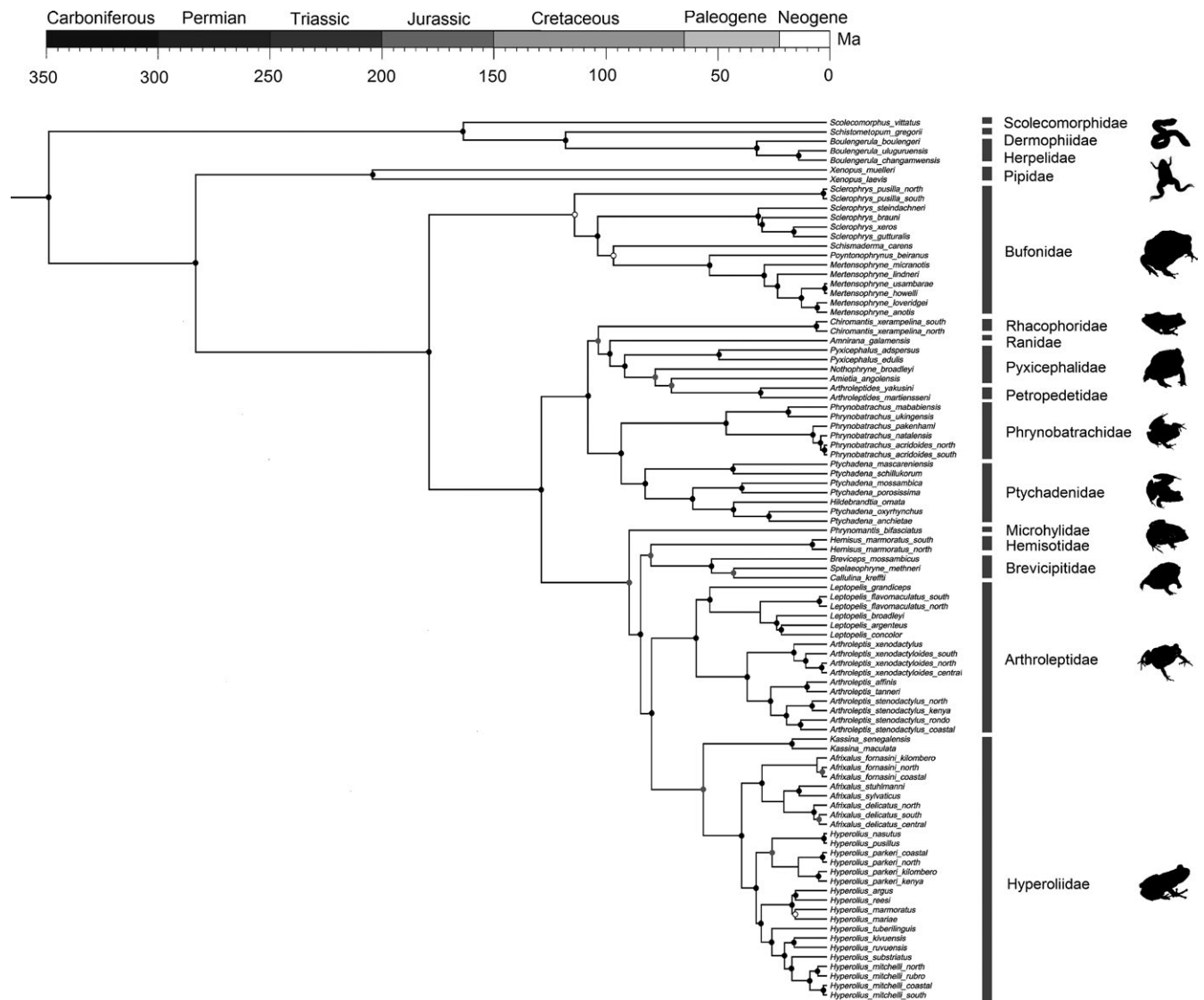


FIGURE 2 Phylogeny representing the amphibian species assemblage for the Coastal Forests of Eastern Africa and surrounding lowland Eastern Afromontane region based on genetic sampling. Intraspecific lineages ($n = 35$) are also represented in the phylogeny and have been included in calculations of phylogenetic endemism. Posterior probability for each node is represented by circles (black, $pp > 0.95$, grey, $pp 0.5-0.94$, or white, $pp < 0.5$)

(cytochrome oxidase subunit I; COI). Intraspecific lineages were defined using a minimum of 2% sequence divergence across a concatenated alignment of both genes, a reasonable threshold for quantifying divergence at the intraspecific level (see Chambers & Hebert, 2016; Fouquet et al. 2014; Vences et al., 2005). To ensure our analyses were not affected by oversplitting the lineages within species, we ran sensitivity analyses whereby lineages were defined by a more conservative threshold of 5% sequence divergence across the two genes. All subsequent analyses were repeated on this dataset and are included in the supplemental information. Details of the molecular procedures follow Barratt et al. (2017) for 16S and Poynton et al. (2016) for COI. All DNA sequences are deposited in GenBank, and accession codes for newly generated sequences from this study can be found in Table S1. Sequences were edited in GENEIOUS 6 and aligned with MUSCLE before Bayesian analyses were performed in BEAST 2.1.3 (Bouckaert

et al., 2014) on the concatenated alignment, using the optimal models of evolution per partition (Table S1) according to the Bayesian information criterion determined by PARTITIONFINDER 1.1.1 (Lanfear, Calcott, Ho, & Guindon, 2012). A full description of the parameters used for the BEAST analysis can be found under the subheading *Phylogenetic tree* in Appendix S1.

2.2 | Distribution data

Spatial sampling covered a significant portion of the CFEA across Tanzania and Kenya, including overlapping areas of the adjacent EA region up to 1,000 m elevation. With recent fieldwork, this area is more comprehensively sampled compared to the adjacent CFEA in Mozambique and Somalia. We collated data for (1) known locations of sequenced specimens (identified to species, or to intraspecific lineage

where possible) based on molecular data, and (2) known locations for specimens from museums, GBIF records and CFEA literature without molecular data. These were included in species distribution modelling (SDM), to estimate which lineage they belong to provided we could definitively identify to species or species complex. We conducted a rigorous filtering and correction procedure to our spatial dataset prior to analyses (see *Spatial data filtering*, Appendix S1). The final dataset comprised of 9,184 occurrence records (2,161 unique points; Tables S2 and S3).

We used SDM to estimate the geographic distribution of each species and for lineages by following the framework outlined by Rosauer et al. (2015). The method uses a modelling approach informed by known species and lineage point data, partitioning a SDM into its constituent lineage distribution models (LDMs) which are informed by the habitat suitability of grid cells between points where lineages are known from (see example in Figure 3). Lineage distribution probability for a grid cell is conditional on (1) the habitat suitability of the cell for the species as a whole and (2) the habitat connectivity of that cell with known locations of that lineage. We constructed SDMs in MaxEnt 3.3.3k, which uses a machine-learning algorithm based on the principles of maximum entropy (Phillips et al., 2006), shown to be highly effective at predicting distributions using presence only data (Elith et al., 2011). We followed recommendations by Merow, Smith, and Silander (2013) to avoid bias in our SDMs. We used six bioclimatic layers and a digital elevation model

(DEM) that were not closely correlated (Pearson's $r < .6$); bio2 (mean diurnal temperature range), bio4 (temperature seasonality), bio5 (max temperature of warmest month), bio12 (annual precipitation), bio14 (precipitation of driest month) and bio18 (precipitation of warmest quarter). These variables were obtained from the WorldClim database based on the Community Climate System Model (CCSM) global circulation model (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005), and the USGS, available from: (<http://csgtm.iscgm.org/dataset/gtopo30>) at 30 arc-second resolution (approximately 1 km² grid cells). Due to the moderate numbers of presence points per species, SDMs used the sub-sample algorithm with a logistic output. This modelling algorithm allows the data to be partitioned for model testing and training across multiple replicates, and accounts for the spatial sorting bias that may be present in cross-validated models, especially in narrow-ranged species (Hijmans, 2012). We tested a range of regularization multiplier values between 0 and 3. We chose to use a regularization multiplier of 1 for our SDMs as it produced the most accurate distribution maps based on our own knowledge and IUCN range maps, whilst minimizing overfitting of the models. We ensured SDMs used an adequate number of unique presence locations (minimum 10 points except for endemics known to have narrow ranges, *Mertensophryne loveridgeri*, *Mertensophryne lindneri*, *Spelaophryne methneri*, *Hyperolius reesi* or species with highly clustered sampling points—*Phrynobatrachus natalensis*, *Ptychadena oxyrhynchus*, *M. lindneri*, see Table S2). Background data used 10,000 points within

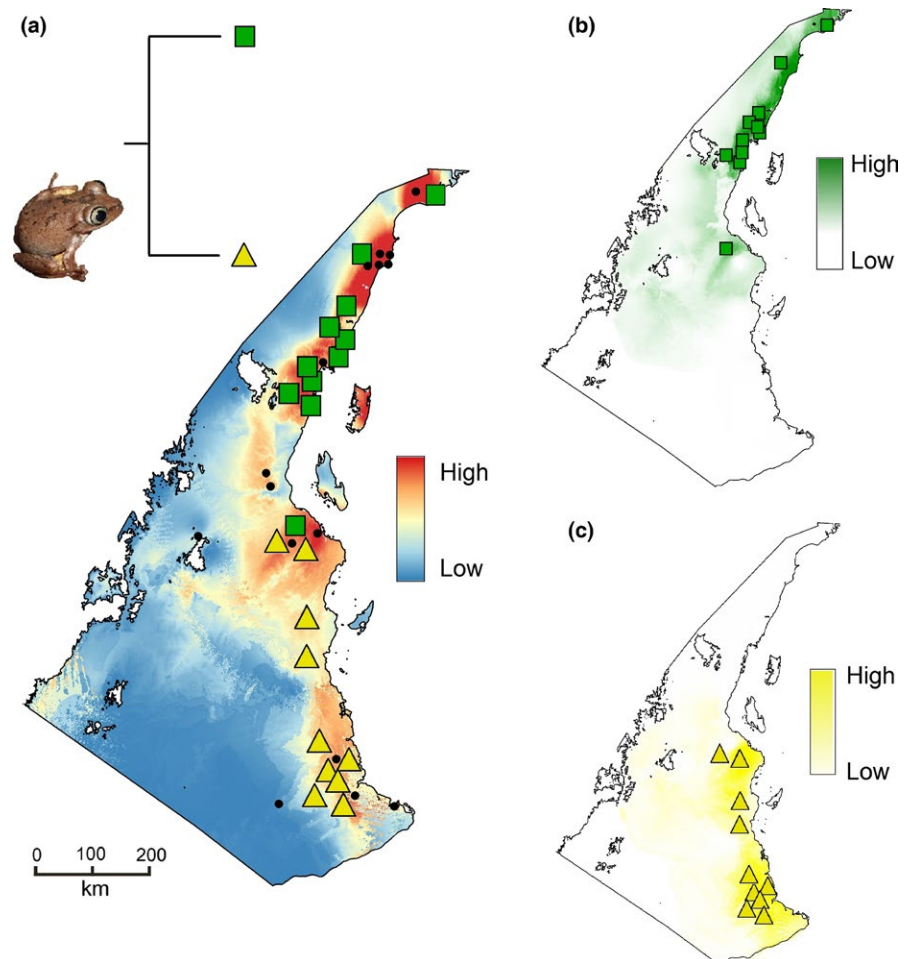


FIGURE 3 Example of a poorly understood species group (*Leptopelis argenteus*) with two intraspecific lineages. (a) Schematic representation of the two clades present in this species and its corresponding species distribution model (green squares = northern lineage, yellow triangles = southern lineage, black dots = species records without DNA to verify to which lineage they belong). (b) and (c) Lineage distribution models for each of the two intraspecific lineages after partitioning the species distribution modelling following Rosauer et al. (2015). For each lineage, the relative likelihood of occurrence is shown, demonstrating the overlap of likelihood occurrences for both lineages towards the centre of the map where both have been recorded

a buffered radius of land 100 km around all species' presence points to emphasize factors locally relevant in distinguishing suitable sites. To account for sampling bias (see Kramer-Schadt et al., 2013; Merow et al., 2013, 2014), presence data were rarefied prior to creating SDMs, with a minimum of 10 km distance allowed between retained points. This distance was chosen as it reduced the spatial autocorrelation of environmental data used to build the models whilst retaining a sufficient number of points per species to model an adequate number of species. We acknowledge that SDMs using low sample sizes (i.e., those listed above, with <10 spatial records) are likely to be less robust than those created for other species with higher numbers of spatial records. We took the average of 25 model replicates for each SDM, evaluating model performance using the area under the curve (AUC) statistic. Models were corrected for overprediction using the minimum convex polygon method in SDMTOOLBOX 1.1c (Brown, Cameron, Yoder, & Vences, 2014) supplemented by our own knowledge of coastal forest amphibians and species range maps (IUCN, 2016). Each SDM was subsequently partitioned into its constituent LDMs using cost distances, jointly informed by the species ecological niche and known locations of each lineage based on geo-referenced genetic data (Rosauer et al., 2015). LDMs were estimated using customized python scripts (<https://github.com/DanRosauer/phylospatial>). We retained suitability surfaces (likelihood of occurrence per grid cell) per species and lineage as a continuous variable between 0 and 1, which were then used to estimate PE.

2.3 | Phylogenetic endemism

We customized R scripts (<https://github.com/DanRosauer/phylospatial>) to estimate PE, using the phylogeny shown in Figure 2 and the spatial occurrence of each tree tip (represented by SDM and LDM scores per grid cell at 1 km² resolution as described above). Each of the branch lengths of the phylogeny included in this study was partitioned into the grid cells where it occurs based on the SDM and LDM suitability scores ranging between 0 and 1. For LDMs, the method uses the formula below where b is one of n branches linking lineages in a grid cell to the root, with local suitability as the LDM score for that grid cell, and the total suitability as the sum of all grid cells within the SDM:

$$\text{Model_weighted_PE} = \sum_{1,n}^b (\text{localsuitability}[b] / \text{totalsuitability}[b]) \times \text{length}[b].$$

The output from these calculations is a map of the PE per grid cell, where the values of all cells sum to the total length of all corresponding branches on the phylogeny. We follow the basis of Laity et al. (2015) and Rosauer et al. (2015, 2016), utilizing summed PE per grid cell across the region for further analyses.

2.4 | Environmental correlates of phylogenetic endemism

To test for correlations with PE, we first prepared a selection of environmental predictors related to historical stability, topography,

and climate that could be expected to influence amphibian PE. For historical stability, we used two measures of Quaternary climate change, measured as the absolute difference in mean annual temperature and precipitation between the Last Glacial Maximum (LGM) and the present (see Kissling et al., 2012). These measures serve as proxies to represent the effect of climatic oscillations during the Quaternary (Eiserhardt, Borchsenius, Sandel, Kissling, & Svenning, 2015; Sandel et al., 2011). We also prepared a measure of historical forest stability over time following Graham, VanDerWal, Phillips, Moritz, and Williams (2010). In brief, forest points ($n = 500$) were randomly generated within an area classified as evergreen forest according to Tuanmu and Jetz (2014). We then constructed a distribution model in MAXENT using the same variables and model parameters used for SDMs. The model was then projected onto past climate data (mid-Holocene 6 ka BP, LGM 21 ka BP, Last Interglacial 120 ka BP). The forest stability measure was then obtained by averaging the log values of suitability for forest in each grid cell across each of the time periods (see *Forest modelling*, Appendix S1). Topographic heterogeneity was calculated as the difference between the minimum and maximum elevation appearing in each grid cell, using the GTOPO30 (USGS) 30 arc-second dataset (more detailed information can be found under the *Forest modelling* section of Appendix S1, with a map of modelled forest stability shown in Fig. S2). We included four bioclimatic variables using available climate data: bio1 (annual mean temperature), bio4 (temperature seasonality), bio12 (annual mean precipitation) and bio14 (precipitation of the driest month).

We used generalized linear models (GLM) and spatially autoregressive models (SAR) to test the relationship between PE (response variable) and the environmental variables mentioned above (predictor variables). We ensured predictor variables were not highly correlated with Pearson's $r < .7$ in all cases (Table S4). All rasters were resampled to 2.5 arc minutes (~5 km² resolution) for environmental predictor analyses to reduce computational requirements, and summarized in Fig. S3. We fitted GLMs for all grid cells in the "Glmulti" (Calcagno & de Mazancourt, 2010) R package, with log-transformed PE as the response and scaled predictors. We tested all combinations of predictors, with main effects only (level = 1), resulting in 128 possible models. We used the corrected Akaike information criterion (AICc) for model selection, assessing models using AICc weight. To account for spatial autocorrelation using an SAR, Moran's I was calculated for the log transformed PE results and for the residuals of the best GLM using correlograms (full details of the process are described in *Spatial autocorrelation*, Appendix S1). Distance classes were defined at 10-km intervals. We observed spatial autocorrelation at lower distance classes so prepared a spatial weights matrix and reran the best GLM with a SAR error model using the errorsarlm function in the R package "spdep" (Bivand & Piras, 2015). To define a spatial weights matrix, we used the kneareneigh function (with $k = 1$), deriving the minimum distance connecting each cell to at least one neighbour. Using this minimum distance, we then defined the neighbourhood structure using the dneareneigh function and created a spatial weights matrix using the nb2listw function (Kissling & Carl, 2008).

2.5 | Conservation

We follow Laity et al. (2015) and Rosauer et al. (2016) to demonstrate the effectiveness of the protected area network at conserving evolutionary history (PE). As PE is a measure of the spatial range of each branch on a phylogeny, summing the total PE across an area of interest estimates the total diversity found there (weighted for each branch by its distribution). Thus, by intersecting summed PE with shapefiles of the protected area network, it is possible to make conservation analyses of how much evolutionary history (i.e., refugia) is currently protected and where shortfalls may lie. We extracted the top 10% of all grid cell values for the PE results to reveal hotspots that contain a disproportionately high share of the total PE and are likely refugia using the *raster* package in R (Hijmans & van Etten, 2012). For each hotspot, we calculated the area size, summed PE and summed PE that falls within protected areas, representing this as a proportion of the whole study region. This was repeated for the top 2.5% and 5% of all grid cell values (Table S8). The shapefiles of the current protected area network across the study region to intersect PE results were obtained from the Critical Ecosystem Partnership Fund (<http://www.cepf.net>) and World Resources Institute (<http://www.wri.org>).

An extended version of these methods can be found in Appendix S1, with a schematic summary of the workflow in Fig. S4.

3 | RESULTS

3.1 | Lineage discovery and phylogeny

Within nine thoroughly sampled species (*Arthroleptis stenodactylus*, *Arthroleptis xenodactyloides*, *Leptopelis flavomaculatus*, *Sclerophrys pusilla*, *Hemisus marmoratus*, *Afixalus fornasini*, *Hyperolius parkeri*, *Phrynobatrachus acridoides* and *Chiromantis xerampelina*), we defined 24 intraspecific lineages. A further 11 lineages were identified within three poorly understood complexes which we refer to as *Leptopelis argenteus* (including *Leptopelis concolor*), *Afixalus stuhlmanni* (including *Afixalus delicatus* and *Afixalus sylvaticus*) and *Hyperolius mitchelli* (including *Hyperolius rubrovermiculatus*) (Fig. S1). The phylogeny (Figure 2) captures this cryptic diversity along with several closely related species from the surrounding area (several CFEA species known from Mozambique and the adjacent EA mountains).

3.2 | Distribution data

Species distribution models performed well, with mean test AUC values from 25 model runs >0.81 in all cases (range = 0.81–0.98, median = 0.89, mean = 0.89). Mean training AUC values were >0.75 in all cases (range = 0.75–0.98, median = 0.83, mean = 0.83; Table S5), indicating “good” model performance. Variable contributions towards each SDM were generally highest for precipitation of driest month (bio14) and elevation, with the mean contribution of these variables together affecting 59.66% of the predictions across all SDMs, but with high variation across species (Table S5). Forest models show similar variable contributions and are shown in Fig. S3. LDMs created from

each SDM clearly delineate the spatial distribution of each lineage (Figure 3), by accounting for probabilistic uncertainty with lower modelled suitability scores towards the periphery of each lineage range.

3.3 | Phylogenetic endemism

Phylogenetic endemism is unevenly distributed across our study region, with several aggregations of high PE in ten hotspots, covering a large part of coastal Kenya, and in Tanzania around the lowlands of the EA region surrounding Uluguru and East Usambara mountains (including the Pangani river), the Pugu hills, Matumbi hills and Mafia island. Several smaller hotspots of high PE are represented in Tanzania on Pemba Island and Zaraninge forest, and in the fragmented forests of Lindi region surrounding the Rondo Plateau (Figure 4b).

3.4 | Environmental correlates of phylogenetic endemism

Complex models that included many predictor variables performed best (Table S6) based on AICc scores. Of all possible GLM combinations we tested, the model with the highest Akaike weight (0.99) and lowest AICc score (delta AIC = 0) included all predictors. In this model, the most important predictors positively correlated with PE (in descending order) included precipitation of driest month (bio14), forest stability since the Last Interglacial, annual mean temperature, annual precipitation and topographic heterogeneity. Quaternary climate oscillations (anom_bio1, anom_bio12) and temperature seasonality (bio4) showed a negative correlation with PE (Figure 4c, Table S6). The spatial model of the best GLM fully accounted for spatial autocorrelation at all distance classes (Table S6) and confirmed the importance of most key predictors in the non-spatial model. Sensitivity analyses of the data using lineages defined by the more conservative 5% divergence cut-off showed the same results in terms of the importance and correlation of predictors (Table S7), indicating that results are consistent even when removing lineages with lower divergences from conspecifics (i.e., between 2% and 5%) from the analysis. This was also the case with the removal of bio14 (which contributes significantly to most SDMs and the forest model). Environmental correlates with amphibian PE were generally consistent between spatial and non-spatial regression models in all data sets including sensitivity analyses, although some predictor variables (forest stability and topographic heterogeneity) showed smaller effect sizes in the spatial models (Table S7).

3.5 | Conservation

The ten identified hotspots of PE (Figure 5a) account for a relatively small area size but support a disproportionately high amount of the total PE found across the study region. Up to 25.5% of the total PE is accounted for by these hotspots based on using the upper 10% of all grid cells. (Table 1). The protected area network across this region is extensive. However, intersecting it with the PE results revealed that only a very small proportion of the amphibian evolutionary history in the region is formally protected. Overlaying the hotspots with

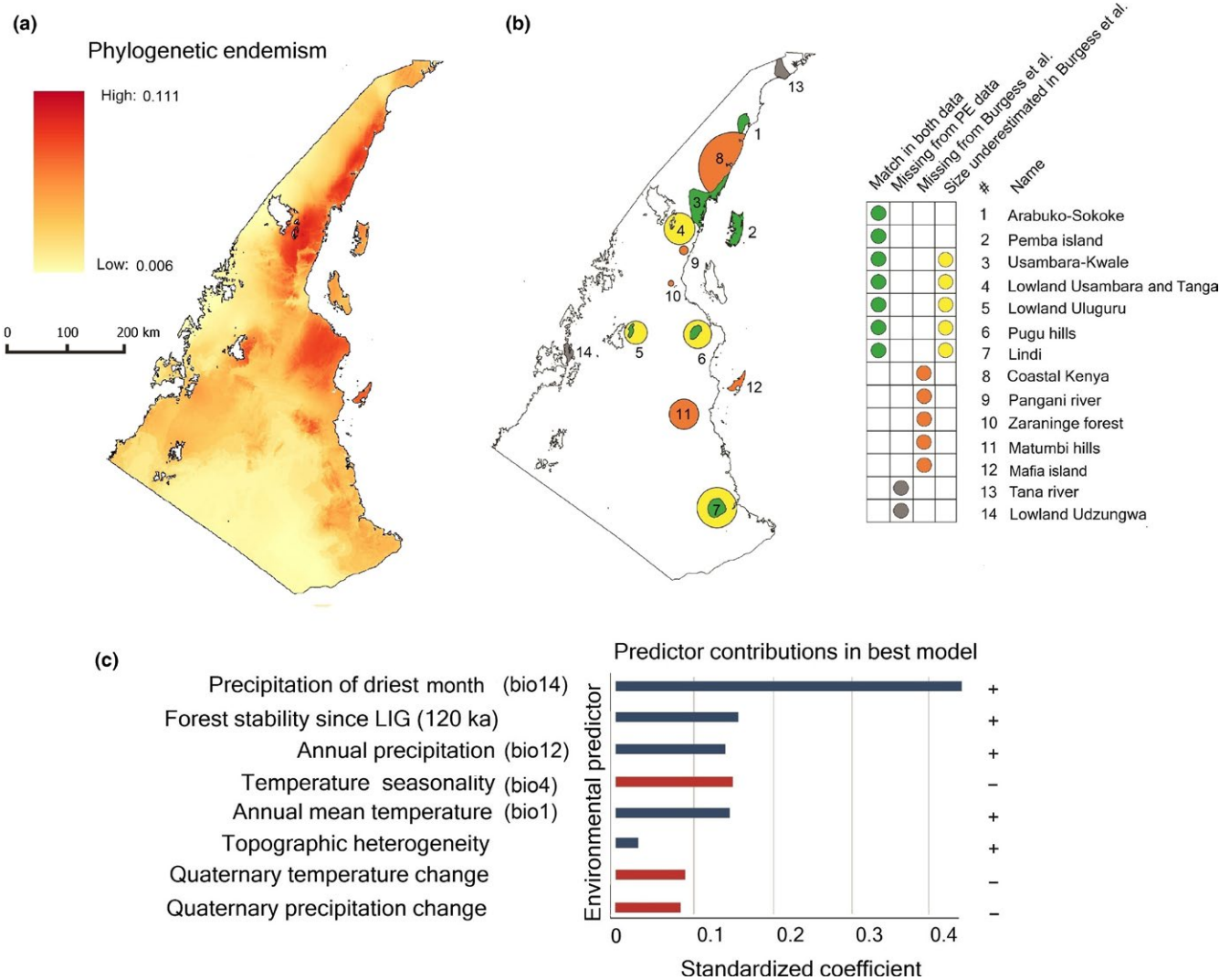


FIGURE 4 Distribution and drivers of phylogenetic endemism (PE) in the Coastal Forests of Eastern Africa (CFEA). (a) Summed PE of all species and intraspecific lineages used in this study. (b) Comparison of PE results with centres of endemism described in Burgess et al. (1998); green = match in both data, orange = missing from Burgess et al. (1998), yellow = size underestimated in Burgess et al. (1998), grey = area listed in Burgess et al. (1998) but not recovered in our PE analyses. (c) Standardized coefficients of predictor variables contributing to explain PE (the response variable) in the best generalized linear model. Blue bars denote positive effects on PE; red bars denote negative effects. Results of spatially autoregressive models are provided in Table S6

the protected area network demonstrated that only small parts of the areas containing high PE are covered (Figure 5), with <3% of the total PE in the study region protected. A list of major protected areas per hotspot is provided in Table S8. As with the environmental correlate analyses, sensitivity analyses using the 5% divergence cut-off dataset showed similar results (Table S9).

4 | DISCUSSION

This study integrates phylogenetic and spatial data to provide the first comprehensive analysis of CFEA biogeographical patterns. The CFEA is one of the highest priority ecosystems for conservation world-wide (Azeria et al., 2007) despite our limited knowledge of how inter- and intraspecific diversity is distributed. Phylogenetic results are largely

concordant with previous large-scale amphibian tree reconstructions in terms of topology and branch lengths, with high posterior probability; thus, we regard the tree as an accurate representation of the inter- and intrarelations of CFEA amphibians (Figure 2). Our results demonstrate the utility of measures such as PE that can provide more meaningful measures than species distribution data alone for local scale conservation efforts. Analyses of PE correlations with environmental predictors strongly support our hypothesis that high levels of Quaternary climate stability, in particular for precipitation, are positively correlated with PE, suggesting the presence of refugial areas. Conservation analyses using PE demonstrate the worrying reality that only a tiny proportion of the identified refugial areas (endemism hotspots) are formally protected for conservation. Based on these analyses, we suggest the protected areas covering high PE should be prioritized and consolidated to maintain the biodiversity that has accumulated and persisted there.

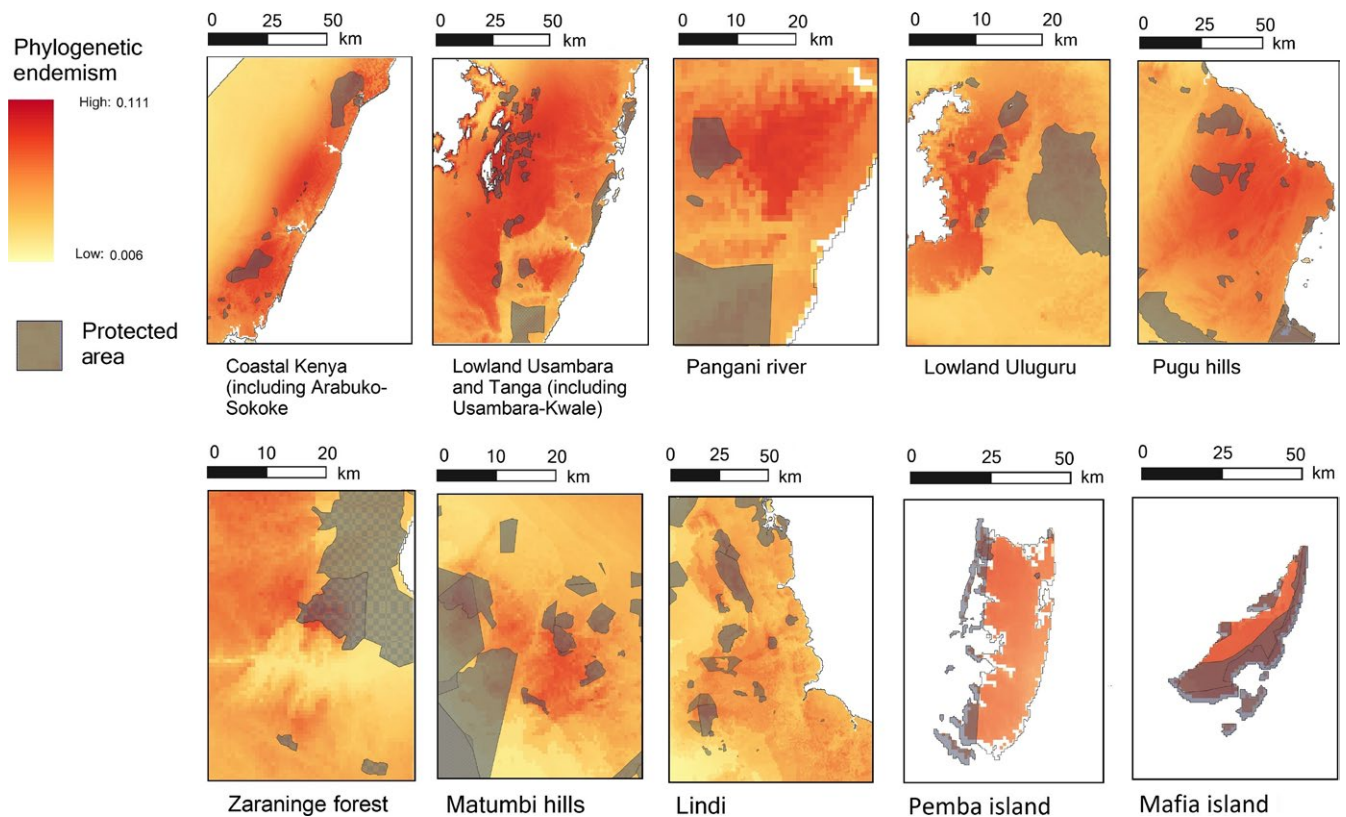


FIGURE 5 Phylogenetic endemism (PE) hotspots, with the protected area network displayed in blue. The location of hotspots within the region is illustrated in Figure 4b. Maps demonstrate that only a small proportion of high PE grid cells are protected, with results summarized for each hotspot in Table 1 and major protected areas per PE hotspot listed in Table S8

TABLE 1 Hotspots of phylogenetic endemism (PE) categorized using the top 10% of all PE scores across grid cells; numbers refer to Figure 4b. Size of each area is shown in km² and as a percentage of the total study area, with the protected area proportions (%) also indicated. Similarly, PE summed across each area is shown both as an absolute value and as a percentage, with the proportional representation (%) of the total indicated. Results are summarized in bold across all hotspots. Sensitivity analyses of these data can be found in Table S8

Hotspot	Area (km ²)	% of total area	% of total area currently protected	PE	% of total PE	% of total PE currently protected
1. and 8. Coastal Kenya (including Arabuko-Sokoke)	6,374	2.27	0.30	462.39	5.89	0.82
2. Pemba island	109	0.04	0.00	6.08	0.08	0.01
3. and 4. Lowland Usambara and Tanga (including Usambara-Kwale)	8,351	2.98	0.18	644.77	8.22	0.58
5. Lowland Uluguru	1,021	0.36	0.07	65.56	0.84	0.15
6. Pugu hills	10,028	3.58	0.35	682.76	8.70	0.87
7. Lindi	31	0.01	0.01	1.73	0.02	0.02
9. Pangani river	625	0.22	0.02	42.54	0.55	0.05
10. Zaraninge forest	398	0.14	0.02	23.59	0.30	0.04
11. Matumbi hills	661	0.24	0.04	40.41	0.51	0.10
12. Mafia island	481	0.17	0.11	30.95	0.39	0.26
Total	28,079	10.02	1.11	2,000.78	25.50	2.89

4.1 | Environmental correlates of PE

The lowland areas of the CFEA have been subjected to more severe and prolonged climatic changes compared to adjacent higher elevation

areas (Burgess & Clarke, 2000; Mumbi, Marchant, Hooghiemstra, & Wooller, 2008) which have likely influenced biodiversity in this region, probably since the Tertiary (Azeria et al., 2007). Environmental predictors (Fig. S4) illustrate the heterogeneous nature of the study region

with several areas that have remained more stable than others in terms of Quaternary precipitation and temperature change. These measures are a useful proxy for historical climatic stability through deeper time (see Sandel et al., 2011) and can help to explain biogeographic breaks and endemism at the species level not only in many amphibians (e.g., Barratt et al., 2017; Loader et al., 2014; Zimkus et al., 2017) but also across other taxonomic groups. Broad biogeographic breaks are common in many other vertebrate taxa across other parts of sub-Saharan Africa including ungulates (Arcander, Johansen, & Coutelec-Vreto, 1999; Lorenzen et al., 2012), small mammals (Mynhardt et al. 2015), birds (Habel et al., 2015) and reptiles (Tolley et al., 2011). This points towards the importance of refugia in the persistence of biodiversity during long-term climate change (Fjelds  & Lovett, 1997).

Our results on environmental correlates or PE are consistent with other literature linking climate and habitat stability to endemism patterns and glacial refugia globally (Sandel et al., 2011; Kissling et al., 2016), in other parts of the world (Bell et al., 2010; Carnaval et al., 2014; Rosauer et al., 2015; Weber, Vanderwal, Schmidt, McDonald, & Shoo, 2014), and in sub-Saharan Africa (e.g., Fjelds  & Lovett, 1997; Gehrke & Linder, 2011; Hamilton, 1976; Loader et al., 2014; Levinsky et al., 2013; Tolley et al., 2011). While our results can be indicative of endemism patterns in non-volant species with poor dispersal abilities (Zug et al., 2001), the integration of data from taxa across larger parts of the tree of life within the CFEA would be beneficial to fully describe biodiversity patterns and evolutionary processes (e.g., Gonz lez-Orozco et al., 2015).

4.2 | Conservation

Conservation prioritization across the CFEA, as in many other biodiversity hotspots, can be boosted by fine-scale data to determine how best to effectively apply the limited conservation resources that are allocated. To achieve this, policymakers require more detailed knowledge of biodiversity and endemism patterns than are currently available to supplement existing information. In this paper, we have shown that the integration of phylogenetic, spatial data and distribution modelling can include cryptic diversity in well sampled taxonomic groups and may be particularly useful for confirming known hotspots of endemism and highlighting new areas. The congruence of our PE results with those derived from existing species distribution data for this region (Burgess et al., 1998) is striking despite fundamental differences in the datasets used (both taxonomically and methodologically). Burgess et al.'s (1998) analysis investigated endemism using a large proportion of known endemic CFEA species (i.e., almost 800 endemic plants, invertebrates and vertebrates). Our results support the recognition of many of Burgess et al.'s (1998) centres of endemism, but show that several of these areas are likely underestimated in extent, and we highlight additional and previously unrecognized areas of high PE.

Integrative measures such as PE that include cryptic diversity with distribution modelling can reveal finer scale endemism patterns than species occurrence data alone. Such data at finer resolution are particularly important for understanding local scale patterns and processes (e.g., Carnaval et al., 2014; Rosauer et al., 2015) and show the strength of using measures such as PE that are not reliant on formal taxonomic

classification. Although it appears that there is an extensive protected area network across the CFEA region, our data reveal large gaps in this network. Nevertheless, it should be noted that our data are modelled and as such do not account for the substantial anthropogenic landscape modification that has occurred (e.g., Godoy et al., 2012; Hall, Burgess, Lovett, Mbilinyi, & Gerau, 2009). Combined with relatively low levels of legal enforcement, anthropogenic modification poses a major conservation concern across the CFEA, which will be exacerbated by future population growth and climate change. The high levels of threatened biodiversity and the rapidly increasing human population make conservation efforts in this region extremely challenging, especially given the limited funding and resources.

Conservationists are rarely afforded the luxury of planning new protected areas, and in most cases, it is a race against time to protect whatever is possible before it disappears. Given the rapid loss of most forests outside governmental reserves, particularly in the East Usambaras and surrounding areas, our research suggests that strengthening protected areas that fall within identified PE hotspots would perhaps be the best strategy to conserve the biodiversity and evolutionary history of this region. We further suggest that existing frameworks for effective conservation prioritization based on phylogenetic and spatial data (e.g., Pollock et al., 2015) could be used to optimize conservation management efforts in this region. The conceptual framework in this manuscript is applicable to discover the distribution of biodiversity in any area, at any scale, and including cryptic diversity, allowing a flexible and objective means to identify important areas that should be considered for future conservation prioritization.

ACKNOWLEDGEMENTS

CDB is funded by a PhD scholarship from the Humer Foundation via the Centre for African Studies Basel. BAB is funded by Stipendienkommission f r Nachwuchskr fte. The University of Dar es Salaam, Frontier Tanzania and Tanzania Forest Conservation Group gave field advice. Natural History Museum, London, Field Museum of Natural History, Chicago, Museum of Vertebrate Zoology, California, Museum of Comparative Zoology, Harvard, and National Museums of Kenya, Nairobi, provided specimen loans. Tanzanian fieldwork led by CDB (COSTECH permit 2013-341-NA-2013-121), funded by Freiwilligen Akademischen Gesellschaft Basel. Kenyan fieldwork led by BAB assisted by Joash Nyamache (Kenya Wildlife Service permit KWS/BRM/5001) and Kenya Forest service permit MUS/1/KFS/VOL.II/4.

AUTHOR CONTRIBUTIONS

C.D.B., S.L., W.D.K. and R.E.O. conceived the ideas; C.D.B., B.A.B., S.P.L. and M.M. collected the data; C.D.B., B.A.B. and D.F.R. analysed the data. All authors contributed to the writing, led by C.D.B., S.P.L., W.D.K. and R.E.O.

CONFLICT OF INTEREST

There are no conflict of interests.

REFERENCES

- Arctander, P., Johansen, C., & Coutellec-Vreto, M.-A. (1999). Phylogeography of three closely related African bovids (tribe Alcelaphini). *Molecular Biology and Evolution*, 16, 1724–1739.
- Axelrod, D. I., & Raven, P. H. (1978). Late Cretaceous and Tertiary vegetation history of Africa. In M. J. A. Werger (Ed.), *Biogeography and ecology of Southern Africa* (pp. 77–130). The Hague, Junk.
- Azeria, E. T., Sanmartin, I., As, S., Carlson, A., & Burgess, N. (2007). Biogeographic patterns of the East African coastal forest vertebrate fauna. *Biodiversity and Conservation*, 16, 883–912.
- Barratt, C. D., Lawson, L. P., Bittencourt-Silva, G. B., Doggart, N., Morgan-Brown, T., Nagel, P., & Loader, S. P. (2017). A new, narrowly distributed, and critically endangered species of spiny-throated reed frog (Anura: Hyperoliidae) from a highly threatened coastal forest reserve in Tanzania. *The Herpetological Journal*, 27, 13–24.
- Bell, R. C., Parra, J. L., Tonione, M., Hoskin, C. J., MacKenzie, J. B., Williams, S. E., & Moritz, C. (2010). Patterns of persistence and isolation indicate resilience to climate change in montane rainforest lizards. *Molecular Ecology*, 19, 2531–2544.
- Bivand, R., & Piras, G. (2015). Comparing implementations of estimation methods for spatial econometrics. *Journal of Statistical Software*, 63, 1–36. Retrieved from <http://www.jstatsoft.org/v63/i18/>
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., ... Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10, e1003537. <https://doi.org/10.1371/journal.pcbi.1003537>
- Brown, J. L., Cameron, A., Yoder, A. D., & Vences, M. (2014). A necessarily complex model to explain the biogeography of the amphibians and reptiles of Madagascar. *Nature Communications*, 5, 5046.
- Burgess, N., & Clarke, G. P. (Eds.). (2000). *Coastal Forests of Eastern Africa*. Gland, Switzerland and Cambridge, UK: IUCN Forest Conservation Programme.
- Burgess, N. D., Clarke, G. P., & Rodgers, W. A. (1998). Coastal Forests of Eastern Africa: status, endemism patterns and their potential causes. *Biological Journal of the Linnean Society*, 64, 337–367.
- Burgess, N., Fjeldsø, J., Howell, K., Kilahama, F., Loader, S. P., Lovett, J. C., & Mbilinyi, B. (2007). The biological importance of the Eastern Arc Mountains of Tanzania and Kenya. *Biological Conservation*, 134, 209–231.
- Burgess, N. D., Mwasumbi, L. B., Hawthorne, W. J., Dickinson, A., & Doggett, R. A. (1992). Preliminary assessment of the distribution, status and biological importance of coastal forests in Tanzania. *Biological Conservation*, 62, 205–218.
- Calcagno, V., & de Mazancourt, C. (2010). glmulti: An R package for easy automated model selection with (generalized) linear models. *Journal of Statistical Software*, 34, 1–29.
- Carnaval, A. C., Waltari, E., Rodrigues, M. T., Rosauer, D. F., Vanderwal, J., Damasceno, R., ... Moritz, C. (2014). Prediction of phylogeographic endemism in an environmentally complex biome. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 281, 20141461. <https://doi.org/10.1098/rspb.2014.1461>
- Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, 1, 1–5.
- Chambers, E. A., & Hebert, P. D. N. (2016). Assessing DNA barcodes for species identification in North American reptiles and amphibians in natural history collections. *PLoS One*, 11, e0154363.
- Channing, A., Hillers, A., Lötters, S., Rödel, M.-O., Schick, S., Conradie, W., ... Burger, M. (2013). Taxonomy of the super-cryptic *Hyperolius nasutus* group of long reed frogs of Africa (Anura: Hyperoliidae), with descriptions of six new species. *Zootaxa*, 3620, 301–350.
- Couvreur, T. L. P., Chatrou, L. W., Sosef, M. S. M., & Richardson, J. E. (2008). Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. *BMC Biology*, 6, 54.
- Crisp, M. D., Laffan, S., Linder, H. P., & Monro, A. (2001). Endemism in the Australian flora. *Journal of Biogeography*, 28, 183–198.
- Eiserhardt, W. L., Borchsenius, F., Sandel, B., Kissling, W. D., & Svenning, J. C. (2015). Late Cenozoic climate and the phylogenetic structure of regional conifer floras world-wide. *Global Ecology and Biogeography*, 24, 1136–1148.
- Elith, J., Phillips, S. J., Hastie, T., Dudik, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17, 43–57.
- Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61, 1–10.
- Fjeldsø, 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 and Conservation*, 6, 325–346.
- Fouquet, A., Cassini, C., Haddad, C. F. B., Pech, N., & Rodrigues, M. T. (2014). Species delimitation, patterns of diversification and historical biogeography of a Neotropical frog genus; *Adenomera* (Anura, Leptodactylidae). *Journal of Biogeography*, 41, 855–870.
- Gehrke, B., & Linder, H. P. (2011). Time, space and ecology: Why some clades have more species than others. *Journal of Biogeography*, 38, 1948–1962.
- Godoy, F. L., Tabor, K., Burgess, N. D., Mbilinyi, B. P., Kashaigili, J. J., & Steininger, M. K. (2012). Deforestation and CO₂ emissions in coastal Tanzania from 1990 to 2007. *Environmental Conservation*, 39, 62–71.
- González-Orozco, C. E., Mishler, B. D., Miller, J. T., Laffan, S. W., Knerr, N., Unmack, P., ... Gruber, B. (2015). Assessing biodiversity and endemism using phylogenetic methods across multiple taxonomic groups. *Ecology and Evolution*, 5, 5177–5192.
- Graham, C. H., VanDerWal, J., Phillips, S. J., Moritz, C., & Williams, S. E. (2010). Dynamic refugia and species persistence: Tracking spatial shifts in habitat through time. *Ecography*, 33, 1062–1069.
- Gudde, R. M., Joy, J. B., & Mooers, A. O. (2013). Imperilled phylogenetic endemism of Malagasy lemuriformes. *Diversity and Distributions*, 19, 664–675.
- Habel, J. C., Borghesio, L., Newmark, W. D., Day, J. J., Lens, L., Husemann, M., & Ulrich, W. (2015). Evolution along the Great Rift Valley: Phenotypic and genetic differentiation of East African white-eyes (Aves, Zosteropidae). *Ecology and Evolution*, 5, 4849–4862.
- Haffer, J. (1969). Speciation in Amazonian forest birds. *Science*, 165, 131–137.
- Hall, J., Burgess, N. D., Lovett, J., Mbilinyi, B., & Gerau, R. E. (2009). Conservation implications of deforestation across an elevational gradient in the Eastern Arc Mountains, Tanzania. *Biological Conservation*, 142, 2510–2521.
- Hamilton, A. C. (1976). The significance of patterns of distribution shown by forest plants and animals in tropical Africa for the reconstruction of upper Pleistocene palaeo-environments: A review. *Palaeoecology of Africa*, 9, 63–97.
- Harper, E. B., Measey, G. J., Patrick, D. A., Menegon, M., & Vonesh, J. R. (2010). *Field guide to the amphibians of the Eastern Arc Mountains and Coastal Forests of Tanzania and Kenya*. Nairobi, Kenya: Camerapix International.
- Hijmans, R. J. (2012). Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null model. *Ecology*, 93, 679–688.
- Hijmans, R. J., & van Etten, J. (2012). *Raster: Geographic analysis and modeling with raster data*. R package version 2.0-12. Retrieved from <http://CRAN.R-project.org/package=raster>
- 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.
- IUCN (2016). *The IUCN red list of threatened species*. Version 2016-3. Retrieved from <http://www.iucnredlist.org> Accessed December 1, 2016.

- Jetz, W., Rahbek, C., & Colwell, R. K. (2004). The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecology Letters*, 7, 1180–1191.
- Kent, P. E., Hunt, J. A., & Johnstone, D. W. (1971). *Geophysics of coastal Tanzania*. National Library of Australia: Institute of Geological Sciences, Geophysical papers, 101 pp.
- Kissling, W. D., Blach-Overgaaard, A., Zwaan, R. E., & Wagner, P. (2016). Historical colonization and dispersal limitation supplement climate and topography in shaping species richness of African lizards (Reptilia: Agamidae). *Scientific Reports*, 6, 34014.
- Kissling, W. D., & Carl, G. (2008). Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*, 17, 59–71.
- Kissling, W. D., Eiserhardt, W. L., Baker, W. J., Borchsenius, F., Couvreur, T. L. P., Balslev, H., & Svenning, J.-C. (2012). Cenozoic imprints on the phylogenetic structure of palm species assemblages worldwide. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 7379–7384.
- Kolbert, E. (2014). *The sixth extinction*. New York, NY: Henry Holt and Company.
- Kramer-Schadt, S., Niedballa, J., Pilgrim, J. D., Schröder, B., Lindenborn, J., Reinfelder, V., ... Wilting, A. (2013). The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions*, 19, 1366–1379.
- Laity, T., Laffan, S. W., Gonzalez-Orozco, C. E., Faith, D. P., Rosauer, D. F., Byrne, M., ... Newport, K. (2015). Phylodiversity to inform conservation policy: An Australian example. *Science of the Total Environment*, 534, 131–143.
- Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29, 1695–1701.
- 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.
- Linder, H. P., de Klerk, H. M., Born, J., Burgess, N. D., Fjeldsø, J., & Rahbek, C. (2012). The partitioning of Africa: Statistically defined biogeographical regions in sub-Saharan Africa. *Journal of Biogeography*, 39, 1189–1205.
- Loader, S. P., Ceccarelli, F. S., Menegon, M., Howell, K. M., Kassahun, R., Mengistu, A. A., ... Gower, D. J. (2014). Persistence and stability of Eastern Afromontane forests: Evidence from brevipitid frogs. *Journal of Biogeography*, 41, 1781–1792.
- Lorenzen, E. D., Heller, R., & Siegmund, H. R. (2012). Comparative phylogeography of African savannah ungulates. *Molecular Ecology*, 21, 3656–3670.
- Maslin, M. A., Brierley, C. M., Milner, A. M., Shultz, S., Trauth, M. H., & Wilson, K. E. (2014). East African climate pulses and early human evolution. *Quaternary Science Reviews*, 101, 1–17.
- Mayr, E., & O'Hara, R. J. (1986). The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. *Evolution*, 40, 55–67.
- Merow, C., Smith, M. J., Edwards, T. C., Guisan, A., McMahon, S. M., Normand, S., ... Elith, J. (2014). What do we gain from simplicity versus complexity in species distribution models? *Ecography*, 37, 1267–1281.
- Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, 36, 1058–1069.
- Moreau, R. E. (1933). Pleistocene climatic changes and the distribution of life in East Africa. *Journal of Ecology*, 21, 415–435.
- Mumbi, C. T., Marchant, R., Hooghiemstra, H., & Wooller, M. J. (2008). Late Quaternary vegetation reconstruction from the Eastern Arc Mountains, Tanzania. *Quaternary Research*, 69, 326–341.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403, 853–858.
- Mynhardt, S., Maree, S., Pelser, I., Bennett, N. C., Bronner, G. N., Wilson, J. W., & Bloomer, P. (2015). Phylogeography of a morphologically cryptic golden mole assemblage from South-Eastern Africa. *PLoS One*, 10, e0144995.
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259.
- Pollock, L. J., Rosauer, D. F., Thornhill, A. H., Kujala, H., Crisp, M. D., Miller, J. T., & McCarthy, M. A. (2015). Phylogenetic diversity meets conservation policy: Small areas are key to preserving eucalypt lineages. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 370, 20140007. <https://doi.org/10.1098/rstb.2014.0007>
- Poynton, J. C., Loader, S. P., Conradie, W., Rödel, M.-O., & Liedtke, H. C. (2016). Designation and description of a neotype of *Sclerophrys maculata* (Hallowell, 1854), and reinstatement of *S. pusilla* (Mertens, 1937) (Amphibia: Anura: Bufonidae). *Zootaxa*, 4098, 73–94.
- Poynton, J. C., Loader, S. P., Sherratt, E., & Clarke, B. T. (2007). Amphibian diversity in East African biodiversity hotspots: Altitudinal and latitudinal patterns. *Biodiversity and Conservation*, 16, 1103–1118.
- Ricklefs, R. E., Bermingham, E., Dick, C. W., & Moritz, C. (2005). Phylogenetic perspectives on patterns of regional and local species richness. In C. Moritz, E. Bermingham (Eds.), *Tropical rainforest: Past, present, and future* (pp. 16–40). University of Chicago Press.
- Rosauer, D. F., Blom, M. P. K., Bourke, G., Catalano, S., Donnellan, S., Gillespie, G., ... Moritz, C. (2016). Phylogeography, hotspots and conservation priorities: An example from the Top End of Australia. *Biological Conservation*, 204, 83–93. <https://doi.org/10.1016/j.biocon.2016.05.002>
- Rosauer, D. F., & Jetz, W. (2015). Phylogenetic endemism in terrestrial mammals. *Global Ecology and Biogeography*, 24, 168–179.
- Rosauer, D. F., Catullo, R. A., Vanderwal, J., & Moussalli, A. (2015). Lineage range estimation method reveals fine-scale endemism linked to Pleistocene stability in Australian rainforest herpetofauna. *PLoS ONE*, 10, e0126274. <https://doi.org/10.1371/journal.pone.0126274.s001>
- Rosauer, D., Laffan, S. W., Crisp, M. D., Donnellan, S. C., & Cook, L. G. (2009). Phylogenetic endemism: A new approach for identifying geographical concentrations of evolutionary history. *Molecular Ecology*, 18, 4061–4072.
- Rovero, F., Menegon, M., Fjeldsø, J., Collett, L., Daggart, N., Leonard, C., ... Burgess, N. D. (2014). Targeted vertebrate surveys enhance the faunal importance and improve explanatory models within the Eastern Arc Mountains of Kenya and Tanzania. *Diversity and Distributions*, 20, 1438–1449.
- Sandel, B., Arge, L., Dalsgaard, B., Davies, R. G., Gaston, K. J., Sutherland, W. J., & Svenning, J.-C. (2011). The influence of late Quaternary climate-change velocity on species endemism. *Science*, 334, 660–664.
- Tolley, K. A., Tilbury, C. R., Measey, G. J., Menegon, M., Branch, W. R., & Matthee, C. A. (2011). Ancient forest fragmentation or recent radiation? Testing refugial speciation models in chameleons within an African biodiversity hotspot. *Journal of Biogeography*, 38, 1748–1760.
- Trauth, M. H., Maslin, M. A., Deino, A., & Strecker, M. R. (2005). Late Cenozoic moisture history of East Africa. *Science*, 309, 2051–2053.
- Tuanmu, M. N., & Jetz, W. (2014). A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Global Ecology and Biogeography*, 23, 1031–1045.
- Van Bocxlaer, I., Loader, S. P., Roelants, K., Biju, S. D., Menegon, M., & Bossuyt, F. (2010). Gradual adaptation toward a range-expansion phenotype initiated the global radiation of toads. *Science*, 327, 679–682.
- Weber, L. C., Vanderwal, J., Schmidt, S., McDonald, W. J. F., & Shoo, L. P. (2014). Patterns of rain forest plant endemism in subtropical Australia relate to stable mesic refugia and species dispersal limitations. *Journal of Biogeography*, 41, 222–238.
- Vences, M., Thomas, M., van der Meijden, A., Chiari, Y., & Vieites, D. (2005). Comparative performance of the 16S rRNA gene in DNA barcoding of amphibians. *Frontiers in Zoology*, 2, 5.

- Zimkus, B. M., Lawson, L. P., Barej, M. F., Barratt, C. D., Channing, A., Dash, K. M., ... Lötters, S. (2017). Leapfrogging into new territory: How Mascarene ridged frogs have diversified across Africa. *Molecular Phylogenetics and Evolution*, 106, 254–269.
- Zug, G. R., Vitt, L. J., & Caldwell, J. P. (2001). *Herpetology: an introductory biology of amphibians and reptiles*. New York: Academic Press.

BIOSKETCH

Christopher D. Barratt is a PhD student at the University of Basel, with interests in biogeography and evolution, principally across sub-Saharan Africa using amphibians as a model system.

SUPPORTING INFORMATION

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

How to cite this article: Barratt CD, Bwong BA, Onstein RE, et al. Environmental correlates of phylogenetic endemism in amphibians and the conservation of refugia in the Coastal Forests of Eastern Africa. *Divers Distrib*. 2017;00:1–13. <https://doi.org/10.1111/ddi.12582>