

Multi-locus phylogeny of the *Crocidura poensis* species complex (Mammalia, Eulipotyphla): influences of the paleoclimate on its diversification and evolution

Running title: Phylogeography of *Crocidura poensis* complex

Violaine Nicolas¹, François Jacquet¹, Rainer Hutterer², Adam Konečný³, Stephane Kan Kouassi⁴, Lies Durnez^{5,6}, Aude Lalis¹, Marc Colyn⁷, Christiane Denys¹

¹ Institut de Systématique, Evolution, Biodiversité, ISYEB - UMR 7205 - CNRS, MNHN, UPMC, EPHE, Sorbonne Universités, Paris, France

² Zoologisches Forschungsmuseum Alexander Koenig, Bonn, Germany

³ Department of Botany and Zoology, Masaryk University, Brno, Czech Republic

⁴ Département Technique et Technologique, Institut Pasteur de Côte d'Ivoire, Abidjan, Côte d'Ivoire

⁵ Department of Biology, University of Antwerp, Antwerp, Belgium

⁶ Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium

⁷ ECOBIO UMR 6553, CNRS-Université de Rennes 1, Station Biologique de Paimpont, Paimont, France

Corresponding author: Violaine Nicolas violaine.colin@mnhn.fr

Acknowledgements

This work was supported by the ‘Service de Systematique Moleculaire’ of the MNHN (UMS 2700, Paris, France), the network ‘Bibliotheque du Vivant’ funded by the CNRS, the MNHN, the INRA and the CEA (Genoscope) and the ‘Action Transversale du Muséum: DNA Barcode et gestion durable des collections’ (MNHN). Field campaigns were funded by EU-DGVIII BIOFAC and ECOFAC programs, WHO-Ebola Research Project in Côte d'Ivoire, WWF Gabon, PGRR-GFA Terra Systems Project in Guinea, INCO-DEV Threat Control in Guinea and Côte d'Ivoire, ANR Biodiversité IFORA ANR-06-BDIV-0014, Czech Science Foundation project P506-10-0983, Flemish Interuniversity Council, the Directorate-General for Development Cooperation (Brussels, Belgium), WCS-RAP 2010 Survey of the Lékoumou Department, Republic of Congo. We thank all African authorities for their assistance and permit autorisations. We are grateful to the curators of the following museum for allowing us to use their collections: FMNH, Chicago; IVB ASCR, Brno; CBGP, Montferrier-sur-Lez; Université de Cocody-Abidjan, Abidjan; ZFMK, Bonn; University of Antwerp; IRSNB, Bruxelles; RMCA, Tervuren; CSRS, Abidjan. We are grateful to all collectors and particularly B. Kadjo, J. Decher, S. M. Goodman, M. H. Huhndorf, R. H. Pine, members of the Malawi Mammal Expedition of the FMNH, R. Šumbera, P. Vogel, R.W. Norris, M. Barej, J. Nopper, E. Fichet-Calvet, E. Lecompte, A. D. Missoup, V. Aniskine, A. Ondzie, H. Bohoussou, F. Kourouma, B. Soropogui, M. Sylla, H. Leirs. We thanks S. Quéroil for her participation in sequencing work, and P. Barrière for his help in morphological identification of the specimens.

Abstract

Aim. This study aims to reconstruct the evolutionary history of the *Crocidura poensis* species complex and to identify factors driving diversification within it. We tested whether: 1) there is a pattern of allopatric differentiation coincident with the location of hypothesized Pleistocene forest refugia, 2) sister taxa are separated by broad rivers, 3) sister taxa occupy adjacent but distinct habitat.

Location. Sub-Saharan African forests and adjacent savannah.

Taxon. Shrews.

Methods. Analyses were based on 247 specimens collected from across the distribution of the species complex. We reconstructed the phylogeny (Bayesian and maximum likelihood methods) and assessed historical biogeography of this taxonomic group using a combination of mitochondrial and nuclear markers. We mapped the genetic diversity, and estimated the divergence times by a relaxed clock model. Informed by multilocus species delimitation methods, we discussed possible taxonomic implications.

Results. This complex is composed of nine major genetic lineages (proposed species). The earliest split within this complex occurred after 2.0-2.4 Ma, which corresponds to a period of increased aridity and/or extreme environmental variability. Most other divergence events occurred after the Early-Middle Pleistocene Transition (1.2-0.8 Ma). Divergent selection across ecological gradients could explain diversification within the West African lineage. In Central Africa the observed phylogeographic pattern fits the Pleistocene refuge hypothesis and supports the existence of multiple small, rather than a few large forest refugia during glacial maxima. Large rivers, like the Congo and Sanaga Rivers, are important barriers to gene flow for several lineages, but probably were not the primary cause of differentiation.

Main conclusions. Both geographic isolation in distinct forest refugia and divergent selection along ecological gradients could explain Pleistocene diversification within this complex.

Keywords

Africa, diversification, ecotone, Pleistocene, phylogeography, taxonomy, tropical forest, forest refugia.

Introduction

The African Guineo-Congolian rainforest constitutes the second largest block of tropical forest on earth, and it hosts a remarkable biodiversity (Myers *et al.*, 2000). The evolutionary processes responsible for this high species richness have fascinated biologists for decades and remain a highly debated topic today. Two main hypotheses for divergence and speciation in rainforests have been proposed and adapted for Africa: the first one emphasizing the role of geographic isolation (Pleistocene forest refugia, riverine barriers), and the second one emphasizing the role of divergent selection along ecological gradients (Mayr & O'Hara, 1986; Haffer, 1997; Smith *et al.*, 1997; Moritz *et al.*, 2000; Plana, 2004). The Pleistocene forest refuge hypothesis postulates that Pleistocene climatic oscillations caused phases of contraction and expansion of different vegetation types. Severe and long-lasting dry and cold periods reduced forests to isolated refugia allowing allopatric diversification of forest taxa. The riverine barrier hypothesis argues that rivers acted as physical barriers to species dispersion, promoting diversification. The gradient model of diversification suggests that strong environmental gradients result in adaptive divergence and speciation for taxa tolerant of a broad range of habitats. While controversy still remains over which diversification process is the most prevalent for African forests, each of these models has been demonstrated to underlie speciation and population differentiation across a variety of taxa. Ecology is of course central to the processes of divergence and speciation (Orr & Smith, 1998) but it plays a contrasting role in the two speciation hypotheses. In the former model speciation occurs through geographic isolation, and niche conservatism (Wiens *et al.*, 2010) is expected over the phylogeny. In contrast according to the ecological gradients hypothesis, the ecological divergence and adaptation to new niches is a fundamental prerequisite (Couvreur *et al.*, 2011).

Molecular studies of diversity within species and relationships among species can be used to test alternative hypotheses about speciation mechanisms. The predictions of the three major models (refugia, riverine and gradient model) were reviewed by Moritz *et al.* (2000) and Anthony *et al.* (2007). Very briefly, the refugia model predicts that recently evolved taxa occur in adjacent refugia, and that there should be evidence from intraspecific gene trees for range expansions and secondary contact between refugia. The riverine model predicts that sister taxa should occur across major rivers rather than being segregated within fluvial areas. The gradient model predicts that sister taxa should occupy distinct but adjacent habitats. Small terrestrial mammals, owing to their low dispersal ability, are excellent models for testing patterns of pan-African biogeography. There have been several large-scale studies on sub-Saharan rodents (Nicolas *et al.*, 2008b; Missoup *et al.*, 2012; Olayemi *et al.*, 2012; Colangelo *et al.*, 2013; Bohoussou *et al.*, 2015; Bryja *et al.*, 2017), but only one study was devoted to the pan-African biogeography of shrews (Jacquet *et al.*, 2015). The genus *Crocidura* is hyper-diversified in Africa, with 109 recognised species (Happold & Happold, 2013). The systematics of this genus is difficult because of the high morphological similarity between species. This has led to the creation of taxonomic units called “species complex” within which the number of species and their geographical distribution are highly debated. In this study we will focus on the species complex *Crocidura poensis* (Fraser, 1843) which is distributed throughout the Guineo-Congolian rainforest and adjacent savannah. Given its large distribution in tropical Africa and the fact that it harbours both forest specialists and open habitat species it can be considered as a good model to test alternative scenarios of diversification. This complex includes both the dark forest and swamp shrew species *C. poensis*, *C. buettikoferi* Jentink, 1888, *C. nigrofusca* Matschie, 1895, *C. turba* Dollman 1910, *C. batesi* Dollman, 1915, *C. nigeriae* Dollman, 1915, *C. wimmeri* Heim de Balsac & Aellen, 1958, *C. grandiceps* Hutterer, 1983, *C. longipes* Hutterer and Happold, 1983, *C. fingui* Ceriaco *et al.*, 2015 and several species paler and greyer in colour, inhabiting the Sahelo-Sudanian and the Guinean savannah: *C. foxi* Dollman, 1915 and *C. theresae* Heim de Balsac, 1968 (Hutterer & Happold, 1983; Hutterer, 2005). Several attempts have been made to resolve taxonomic issues within this complex. The first steps were made using morphological data (Heim de Balsac & Meester, 1977; Hutterer & Joger, 1982; Hutterer & Happold, 1983; Brosset, 1988; Grubb *et al.*, 1998). Nevertheless, the morphology of shrews is known to be highly dependent on their environment (Rychlik *et al.*, 2006) and a taxonomy based exclusively on such information can lead to the grouping of phylogenetically unrelated forms (Jacquet *et al.*, 2013). Cytogenetic and allozyme studies (Meylan, 1971; Meylan & Vogel, 1982; Maddalena, 1990; Schlitter *et al.*, 1999), and several recent molecular phylogenies (Quérouil *et al.*, 2005; Dubey *et al.*, 2008; Jacquet *et al.*, 2012; Ceriaco *et al.*, 2015) shed new light on the systematics of this group. No molecular study including all the species of the complex was carried out, and previous sampling schemes were not suitable (few specimens and specimens from few localities) to sort apart intra- versus inter-specific divergence. The validity of several species still needs to be assessed as well as the geographical distribution of most species (Hutterer, 2005; Happold & Happold, 2013). A robust and reliable taxonomic framework is essential to understand the evolutionary history of this complex but also for other fields of science such as epidemiology (*C. theresae* is a reservoir of a distinct hantavirus; Klempa *et al.*, 2007) or conservation (Nicolas *et al.* 2009). Here, we present the first comprehensive phylogeographical survey of the *C. poensis* species complex. We used multilocus datasets to resolve phylogenetic relationships between species and to generate a dated phylogeny. This study allowed us to test the validity of the current species forming the *C. poensis* complex, and to assess the relative roles that candidate Pleistocene refugia, major rivers, and environmental gradients may have played in driving the evolutionary diversification of this species complex. To this aim we tested the predictions of

these three alternative models of diversification: 1) we determined whether there is a pattern of allopatric differentiation coincident with the location of hypothesized Pleistocene forest refugia; 2) we tested whether sister taxa are separated by broad rivers, and 3) we tested whether sister taxa occupy adjacent but distinct habitats.

Material and methods

DNA extraction and amplification

DNA was extracted from liver, kidney or intercostal tissues using the NucleoSpin96 tissue kit (Machery-Nagel). For all individuals (247 ingroup and 22 outgroup specimens) we sequenced two mitochondrial (mtDNA) markers, the cytochrome b (Cytb) gene and the 16s rRNA gene (16S; see Table S1 in Appendix S1 in Supporting Information). For 114 selected specimens (92 ingroup and 22 outgroup specimens) we sequenced the mtDNA marker cytochrome oxidase I (Co1) and three nuclear (nDNA) markers, one exon, Breast Cancer 1 (BRCA1) and two introns, Signal Transducer and Activator of Transcription 5A (STAT5A) and Histone Deacetylase 2 (HDAC2) following the protocol described in Jacquet *et al.* (2013). These 92 ingroup specimens were selected in order to match detected mitochondrial diversity as far as possible.

Purification of PCR products and cycle-sequencing reactions of both DNA strands were performed at the Genoscope (Evry, France). Sequences were deposited in GenBank (see Table S1 in Appendix S1).

Phylogenetic analyses

Sequences were aligned using ClustalW (Thompson *et al.*, 1994) and edited manually. The final alignments were of 635, 1063, 538, 476, 562 and 682 bp for the Co1, Cytb, 16S, BRCA1, STAT5A and HDAC2 genes, respectively.

We first reconstructed a mitochondrial phylogeny using the concatenated Cytb and 16S dataset (the Co1 gene was not used in this analysis because it was available for only 92 ingroup specimens, while the Cytb and 16S genes were available for 247 ingroup specimens). We also constructed a nuclear phylogeny using the concatenated BRCA1, STAT5A and HDAC2 sequences. Phylogenies were analysed by maximum likelihood (ML) and Bayesian inference (BI) approaches. A detailed version of the phylogenetic methods is available within Appendix S2.

Species delimitation

On the basis of the lineages obtained by the phylogenetic analyses, two species delimitation scenarios were tested: a scenario with 9 candidate species (lineages I to IX) and a scenario with 19 candidate species (I-1 to I-11, II to IX). We tested the distinctiveness of their gene pools using the combined mtDNA + nDNA dataset (Cytb, Co1, 16S, BRCA1, STAT5A and HDAC2; 92 individuals) and the nDNA dataset alone (BRCA1, STAT5A and HDAC2; 92 individuals). For species delimitation analyses all the nuclear genes were phased, and the mitochondrial genes were combined into a single linked partition.

Bayesian species delimitation was conducted using the program BPP and the joint species delimitation and species-tree estimation model (Yang, 2015). To verify results across species delimitation methods, we also used the program STACEY (Jones, 2017) implemented in BEAST 2.4.2 (Bouckaert *et al.*, 2014). A detailed version of the species delimitation methods is available within Appendix S2.

Morphological identification of specimens was carried out by Violaine Nicolas and confirmed by Rainer Hutterer. It was based on both external and cranial morphology. Our specimens

were compared to type specimens and relevant literature (original species descriptions and the following review papers: Heim de Balsac & Meester, 1977; Hutterer & Joger, 1982; Hutterer & Happold, 1983; Brosset, 1988; Grubb et al., 1998).

Divergence time estimates and biogeographic analyses

Divergence times were estimated from the combined supermatrix of mtDNA+nDNA sequences. The dataset for divergence dating included one sequence for each of the 19 mtDNA lineages (Table S1 in Appendix S1). We used BEAST V1.8.4 (Drummond *et al.*, 2012) assuming a Yule model of speciation and an uncorrelated log-normal distribution molecular clock as tree priors. To calibrate the phylogeny we used two strategies. First we used a fossil constraint: the oldest record of an African *Crocidura* was dated at 6 Ma (Mein & Pickford, 2006). Second we used three secondary calibration points derived from the study of Dubey et al. (2008): the split between the Eurasian *Suncus* and *Crocidura* at 6.8 Myr ago; within *Crocidura*, the split between the Afrotropical clade (including the West Palaearctic *C. russula* and *C. pachyura*) and the other clades which occurred 6.2 Myr ago, and the split between the strictly Afrotropical clade and the latter West Palaearctic species which occurred 5.9 Myr ago. A detailed version of divergence time analyses is available in Appendix S2. Biogeographical reconstructions were performed using the phylogenetic tree generated by BEAST with outgroups removed, and the use of BioGeoBEARS (Matzke, 2013) in RASP 4.02 (Yu *et al.*, 2015). Six models of geographic range evolution were compared in a likelihood framework: Dispersal-Extinction Cladogenesis model (DEC); DEC +J model which adds founder-event speciation with long-distance dispersal to the DEC framework; Dispersal Vicariance Analysis (DIVA); DIVA with long-distance dispersal (DIVA +J); Bayesian inference of historical biogeography for discrete areas (BayArea); BayArea with long distance dispersal (BayArea +J). Model fit was assessed using the Akaike information criterion and likelihood-ratio tests. Each model allows for a different subset of biogeographical possibilities, such as dispersal, vicariance and extinction (Matzke, 2013). As terminal biogeographical units we used four areas corresponding to different ecosystems (Happold & Lock, 2013; Sayre *et al.*, 2013): a –Guinean Rainforest, b –Congolian Rainforest, c –Sudanian Savannah, d- Zambezian woodland. The number of maximum areas was set to four. Dispersal rate between adjacent areas was fixed to 1 (a-c, c-b, a-b, b-d, c-d), whereas dispersal of 0.7 was specified for long-distance dispersal (a-d). To define in which biogeographical unit a given lineage is present we followed Happold & Happold (2013) and modified it according to our own results (Table 1). Lineage III was considered to be present only in the Zambezian woodland given that we never captured it in West Central Africa despite extensive sampling in this area. We followed Hutterer & Happold (1983) and Happold & Happold (2013) in considering that lineages IV and VII are present in both Guinean and Congolian Rainforests. Several specimens of these two lineages from Nigeria are available as vouchers at the BMNH and the USNM and there is no doubt about their identification. Even if we could not include any specimens from Nigeria in our molecular study, we therefore consider that these two lineages are present in the Congolian Rainforest zone.

Results

Phylogenetic results

Cytb + 16S DNA

Of the 1601 included characters 542 are variable and 437 are parsimony informative. The *C. poensis* complex is monophyletic and is the sister lineage of *C. hildegardeae* (Fig. 1). Within the *C. poensis* complex nine distinct lineages are recovered (I to IX). All these lineages, except lineage I, are highly supported (PP = 1, BP > 95). Lineages I to III form a distinct

Central-East African clade and they exhibit allopatric distributions: lineage I is restricted to central African forests and the island of Bioko, lineage II is restricted to the island of Príncipe and lineage III is restricted to the East African rift (Fig. 2A). Lineages IV to VIII form a distinct West African clade, and most of them have overlapping geographical distributions in Guinea, Liberia and Ivory Coast. Lineage IX is present in Benin and western Cameroon and occurs in sympatry with lineage VII in Benin.

Lineage I (PP=1, BP = 68) exhibits the highest genetic diversity (see Fig. S1 in Appendix S3 for uncollapsed nodes), with eleven supported sublineages (I-1 to I-11; Fig. 1). It presents a strong phylogeographic signal: most of these eleven sublineages are geographically restricted to one or several geographically close localities, and they exhibit allopatric or parapatric distributions (Fig. 2B). Some genetic variability is recovered in other lineages but few nodes are supported and no strong phylogeographic signal is apparent.

nDNA

Of the 1720 included characters 264 are variable and 137 are parsimony informative. Based on nuclear data, only lineages V and IX are recovered among the nine mtDNA lineages (Fig. S2 in Appendix S3). The mtDNA lineage IV is paraphyletic in the nuclear tree and forms a clade with the mtDNA lineage VIII. The two mtDNA lineages VI and VII form a unique lineage based on nuclear data. The mtDNA lineage I is not monophyletic in the nuclear tree and is split in two lineages: the first grouping specimens from Bioko Island, Mt Kupe and Yabassi (ie mtDNA sublineages I-1, I-2 and I-4), and the second grouping specimens from all other localities. The latter is paraphyletic and forms a monophyletic group with the mitochondrial lineages II and III.

Species delimitation

When considering nine candidate species in BPP analyses (lineages I to IX), all analyses support the existence of nine species with high posterior probability, and the posterior probability of each delimited species is also high (PP > 0.95; Table S1 in Appendix S3; Fig. 3). When considering 19 candidate species (lineages I-1 to I-11, and II to IX), BPP most often recognize the 17-species delimitation scenario as the most probable one (lineages I-2 and I-4 are collapsed in one single species, as well as lineages I-8 and I-9), but the posterior probability of this model is low (< 0.59), and several lineages have posterior probabilities below 0.95 preventing their recognition as distinct species.

STACEY results based on nine candidate species indicate high support for nine clusters. These nine clusters were recovered 99% of the time for the combined mtDNA+nDNA dataset, and 82% of the time for the nDNA dataset. The species tree based on the combined mtDNA+nDNA dataset (Fig. S3 in Appendix S3), indicates that species I, II and III are sister species, as well as species VI and VII, and species IV and VIII. Results for the 19 candidate species analyses are inconclusive. For the combined mtDNA+nDNA dataset 18 clusters are recovered 46 to 63% of the time (lineages I-2 and I-4 are collapsed in one single species). For the nDNA dataset 17 clusters are recovered 20% of the time (I-2 and I-4; and I-8 and I-9 collapsed), 16 clusters are recovered 15% of the time (I-2 and I-4; I-8 and I-9; and I-11 and II collapsed) and 18 clusters are recovered 10% of the time (I-2 and I-4 collapsed).

Historical biogeography and divergence dating

Median dates of the nodes are slightly older with the secondary calibrations compared to the fossil calibration, even if the confidence intervals are largely overlapping (Table S2 in Appendix S3). Only the tree obtained with the three secondary calibration points is presented in figure 3 (for the whole tree including calibration points and outgroups see Fig. S4 in Appendix S3). This tree was used for the biogeographic reconstruction. Among the six

models of geographic-range evolution compared in a likelihood framework in BioGeoBEARS, the DIVA model was chosen because of its best likelihood and AICc associated scores (Table S3 in Appendix S3). Divergence time estimates indicate a Pleistocene origin of the *C. poensis* complex with a median age of 2.0-2.4 Ma. According to our ancestral area reconstruction the MRCA of the *C. poensis* complex was probably largely distributed in tropical Africa (areas now corresponding to the Guinean Rainforest, Sudanian savannah and Congolian Rainforest). It then diverged in two lineages through a vicariance event, the first one corresponding to lineages I-II-III whose ancestor lived in the Congolian Rainforest, and the second one corresponding to lineages IV-V-VI-VII-VIII-IX whose ancestor lived in the Guinean Rainforest and Sudanian Savanna. Twelve diversification events occurred within the first lineage: a first one about 1.0-1.2 Ma separating Príncipe Island specimens from mainland specimens, and eleven events after 0.7-0.8 Ma leading to lineage III and sublineages I-1 to I-11. The distribution of lineages III and I-11, endemic to Zambezian Woodland, could be explained by two vicariant events. All other sublineages are present in the Congolian Rainforest. Within the second lineage five diversification events occurred: a vicariance event between the Guinean Rainforest lineages (IV-V-VI-VII-VIII) and the Sudanian Savannah lineage (IX) about 1.9-2.2 Ma, and four events between 1.2-1.4 and 0.7-0.8 Ma leading to lineages IV to VIII. Two dispersal events from the Guinean to the Congolian Rainforest explain the actual distribution of lineages IV and VII in both areas (these two lineages are present in Nigeria even if we could not include any specimen from this country in our molecular analysis). Lineages V, VI and VIII co-occur in the Guinean Rainforest.

Discussion

Systematic revision of the *C. poensis* complex

Our molecular analyses confirm the monophyly of the so-called *C. poensis* complex, and its sister clade relationship with *C. hildegardae*. BPP and STACEY analyses consistently support the recognition of nine species within this complex, corresponding to lineages I to IX.

According to our morphological identification, lineages II, III, V, VII, VIII and IX corresponds to *C. fingu*, *C. turba*, *C. theresae*, *C. grandiceps*, *C. wimmeri* and *C. foxi*, respectively (Table S1 in Appendix S1 in Supporting Information, Fig. 3).

Lineage VI groups several specimens captured in W Guinea which are morphologically close to *C. longipes*, a species only known from two localities in NW Nigeria. Given the high geographical distance between Nigeria and Guinea we provisionally name lineage VI *C. cf. longipes*.

Lineage IV is composed of all specimens morphologically described as *C. buettikoferi* and of the West African specimens morphologically identified as *C. poensis* and *C. nigeriae*. Heim de Balsac (1974) noticed the great phenotypic resemblance between *C. buettikoferi*, *C. poensis* and *C. nigeriae* occurring in sympatry in West Africa and questioned their taxonomic status. *Crocidura buettikoferi* was described from Robertsport in Liberia, and we obtained molecular data from Zياما (Guinea) and Mt Nimba (Liberia), less than 300 km from the type locality. *Crocidura poensis* was described from Bioko Island, and our molecular data clearly show the distinctiveness between Bioko (lineage I) and West African *C. poensis* specimens (lineage IV). We propose to restrict the distribution of *C. poensis* to West-Central Africa and to rename all the specimens of the lineage IV captured in West Africa as *C. buettikoferi*. *Crocidura nigeriae* was described from Nigeria. Owing to the lack of data from Nigeria, we

could not assess the real status of *C. nigeriae*, and we simply consider West African records as misidentifications.

Lineage I is composed of specimens morphologically identified as *C. poensis*, *C. batesi* and *C. nigrofuscus* from Central African rainforests. In this lineage a high genetic variability is observed and this genetic variability is not partitioned according to morphological identification, except for *C. nigrofuscus*. *Crocidura nigrofuscus* (lineage I-11) is the only lineage captured on the left bank of the Congo River. This river possibly played a role in its diversification. The specimens morphologically identified as *C. poensis* and *C. batesi* do not cluster in two separate lineages: 10 lineages are observed and most of them are found in only one or several geographically restricted localities. Our BPP and STACEY analyses show that these lineages cannot be considered as distinct species, and support the recognition of only one species in West Central Africa. Several studies highlighted the difficulty to distinguish *C. poensis* and *C. batesi* based on morphological data and also raised doubts on the validity of these species (Hutterer, 2005; Happold & Happold, 2013). More genetic data at the genomic scale, as well as quantified morphometric data including genetically identified specimens are necessary to further assess their taxonomical status. Several recent studies showed the power of morphometric geometric approaches to discriminate cryptic shrew species and to associate a genetic lineage to a type specimen (Jacquet *et al.*, 2013; Jacquet *et al.*, 2014).

To conclude the combination of geographical, morphological and molecular data allowed us to consider *C. theresae*, *C. foxi*, *C. wimmeri*, *C. buettikoferi*, *C. grandiceps*, *C. cf. longipes*, *C. turba* and *C. fingu* as distinct species. The status of *C. nigeriae*, as well as the taxonomy of the West Central African species (*C. poensis*, *C. batesi* and *C. nigrofuscus*) still need to be confirmed by additional data. In the rest of this study, we will temporarily give the name *C. poensis* to all specimens of lineage I.

Evolutionary history of the *C. poensis* complex

All diversification events within the *C. poensis* species complex occurred during the Pleistocene. During this period three major transitions or climate events have influenced the African climate (Maslin *et al.*, 2014). 1) The intensification of Northern Hemisphere Glaciation (iNHG, 3.2-2.5 Ma). It was proposed that in response to the iNHG increase aridity occurred in tropical Africa at around 2.8 Ma (deMenocal, 2004); though this has been disputed by Trauth *et al.* (2009). Recent advances tend to indicate a highly variable and changing climatic regime at this period. 2) The development of the Walker Circulation (DWC, 2.0-1.7 Ma), which coincides with an increase in aridity and climatic variability on the continent after 1.9-1.5 Ma (Trauth *et al.*, 2009). 3) The Early-Middle Pleistocene Transition (EMPT, 1.2-0.8 Ma). The EMPT is the marked prolongation and intensification of glacial-interglacial climate cycles. Before the EMPT, global climate conditions appear to have responded primarily to the obliquity orbital with glacial-interglacial cycles with a mean period of 41 kyrs. After the EMPT glacial-interglacial cycles occur with a much longer mean quasi-periodicity of ~100 kyrs. The EMPT, in addition to marking a change in periodicity, also marks a dramatic sharpening of the contrast between warm and cold periods (Maslin *et al.*, 2014).

The separation between the West African (*C. buettikoferi*, *C. theresae*, *C. grandiceps*, *C. wimmeri*, *C. cf. longipes* and *C. foxi*) and the Central-East African (*C. poensis*, *C. fingu* and *C. turba*) lineages is dated just after 2.0-2.4 Ma, which corresponds to the end of the iNHG period. The increased aridity and/or extreme environmental variability at this period could

have favored the diversification within the *C. poensis* species complex through a vicariance event. It is interesting to note that at more or less the same time period diversification events between West Africa and Central-East Africa were also recorded in several others forest-adapted small mammals (Nicolas *et al.*, 2006; Nicolas *et al.*, 2010; Hassanin *et al.*, 2015) and birds (Fuchs & Bowie, 2015; Fuchs *et al.*, 2017; Huntley & Voelker, 2016).

In the West African lineage all speciation events occurred between 2.2-1.9 and 0.8-0.7 Ma. *Crocidura foxi* is the unique species inhabiting Sudanian Savannah and it probably diverged from other species through a vicariant event. The ancestor of all other species of the West African lineage probably lived in the Guinean Rainforest, and the occurrence of *C. grandiceps* and *C. buettikoferi* in Nigeria (Hutterer & Happold, 1983; Happold & Happold, 2013) would be the result of two dispersal events. The Upper Guinean (corresponding to our Guinean biogeographical unit) and Lower Guinean forests (corresponding to the western part of our Congolian biogeographical unit) are separated by a broad savannah corridor, also known as the Dahomey Gap (Benin, Togo, and SE of Ghana). The Dahomey gap may be as old as 3 Myr (Robert & Chamley, 1987), but it did not persist throughout the last 3 Myr: forest expansion completely filled the gap several times (Dupont *et al.*, 2000; Salzmann & Hoelzmann, 2005). These successive phases of forest expansion and contraction could have favoured the dispersion of *C. grandiceps* and *C. buettikoferi* from Upper Guinea to Nigeria. How can we explain the diversification of the five species broadly sympatric in the Guinean Rainforest (*C. buettikoferi*, *C. wimmeri*, *C. theresae*, *C. grandiceps* and *Crocidura cf. longipes*)? A similar pattern, with several highly overlapping sister lineages, was previously observed in other Upper Guinean small mammals (Nicolas *et al.*, 2008a; Jacquet *et al.*, 2014; Bohoussou *et al.*, 2015). The Pleistocene refuge hypothesis was used to explain this pattern. According to this hypothesis tropical forests contracted into smaller fragments during the more arid phases of the Pleistocene (glacial cycles), leading to the isolation and allopatric diversification of forest-associated taxa. Diversification in several microrefugia of Upper Guinean forests, intensive genetic drift in small populations and secondary contact of already differentiated populations would thus explain the observed pattern. However, an intriguing observation suggest that the explanation could be more complex for the species of the *C. poensis* complex. The five species broadly sympatric in Guinea, Liberia and Ivory Coast do not have exactly the same habitat requirements (Table 1). *Crocidura buettikoferi* can be found in a wide range of habitats such as grassland within the rainforest, forests, forests relicts in derived savanna, fallows and cocoa plantations. *Crocidura grandiceps* is considered as a forest species. *Crocidura theresae* is mostly captured in open habitats such as grasslands and fields, and only occasionally in forest relicts. *Crocidura wimmeri* is only found in Eastern Ivory Coast and Western Ghana in wet, swampy forests near the coast. *Crocidura cf. longipes* is only found in W Coastal Guinea in small gallery forests isolated in savannahs. These differences in habitat preferences between species and their broadly sympatric geographical areas suggest that speciation could have occurred along an ecological gradient from forest to savannah, congruently with the model of divergent selection across ecological gradients. The same holds true for the parapatric West African species, *Crocidura foxi*, inhabiting Sudanian Savannah. According to this model, ecotones can promote parapatric speciation through disruptive selection (Smith *et al.*, 1997; Moritz *et al.*, 2000). Under the ecological gradient hypothesis, selection can drive phenotypic divergence despite high gene flow, but genetic divergence is expected to increase across the ecotone over time. Multiple vertebrate taxa exhibit a pattern of selection-driven phenotypic or genetic divergence across the forest-savannah ecotone, lending support to the ecological gradient hypothesis (Smith *et al.*, 1997; Smith *et al.*, 2005; Freedman *et al.*, 2010; Smith *et al.*, 2011; Mitchell *et al.*, 2015; Zhen *et al.*, 2017). Diversification events between West African species seem to have occurred

between 1.4 and 0.8 Ma, i.e. at the EMPT. According to Cowling et al. (2008), at this time, long and cold glacial periods were characterized by a heterogeneous vegetation landscape. This heterogeneous vegetation landscape could have favoured ecological speciation. Coupled phenotypic and genome wide approaches allowing the detection of signatures of natural selection are necessary to test this hypothesis.

Within the Central-East lineage, *C. fingsi*, endemic to Príncipe Island, is the first species to diverge (about 1.0-1.2 Ma). Príncipe Island originated 31 Ma (Lee et al., 1994) and it was never connected to the mainland. Thus, colonization of Príncipe Island most probably occurred by oceanic rafting via the Congo River aided by the Congo current (Measey et al., 2007).

All other diversification events within the Central-East lineage occurred between 0.8-0.7 and 0.2 Ma, i.e. after the EMPT.

Our results indicate that gene flow between Bioko Island and the continent is not ongoing, and suggest a recent (Late Pleistocene, around 0.2 Ma) population divergence from the mainland. Bioko is currently separated from mainland Africa by approximately 30 km of shallow sea. The volcanic peaks that comprise Bioko Island formed 1 Ma (Marzoli et al., 2000) and since then, cycles of rising and retreating sea levels resulted in several periods of isolation and connectivity between Bioko and the adjacent continent. The most recent periods of high connectivity between Bioko and the continent coincide with the last glacial maximum (~21 ka) and glacial period preceding the Last Interglacial (~120 ka) (Bell et al., 2017). It seems likely that the population from Bioko remained genetically isolated throughout these recent cycles of geographic connectivity. The Bioko population of *C. poensis* is most closely related to mainland populations in southwestern Cameroon, similar to the pattern recovered in African forest geckos (Leaché & Fujita, 2010) and the amphibian *Hyperolius ocellatus* (Bell et al., 2017). These results are consistent with a phenomenon of diversification linked to vicariance due to marine incursions.

A high number (12) of genetic clades were recovered in Central-East Africa. All of them are restricted to the forest habitat and most of them have allopatric distributions (Table 1, Fig. 2). The phylogeographic pattern observed seems to fit the Pleistocene refuge hypothesis. Based on reconstructions of paleovegetation and current patterns of species richness and endemism, several Pleistocene forest refugia have been proposed for the last glacial maximum, but they vary in their location and extent (Maley, 1996; Anhof et al., 2006; Cowling et al., 2008). Our results agree with the hypothesis of Maley (1996) that several distinct forest refugia occurred on the West Central African Atlantic coast (N Sanaga -Mount Cameroon- corresponding to lineage I-2, S Sanaga -Campo Na'an- corresponding to lineage I-6, N Ogooue -Mont de Cristal- corresponding to lineage I-7, S Ogooue -Chaillu mountains- corresponding to lineage I-5, S Ogooue -Monts Doudou- corresponding to lineage I-8, Coastal Congo -Mayombe mountains- corresponding to lineage I-9). Lineages I-7, I-5 and I-8 have parapatric distributions, which could be explained by secondary contact after population expansion. Our results also tend to support the view of Maley (1996) and Anhof et al. (2006) suggesting the existence of distinct forest refugia south of the Congo River (lineage I-11) and in the Albertine rift Mountains (lineage III). We also found distinct lineages in areas never proposed as refugial areas such as SE Cameroon and SW CAR (lineage I-3) or SW Congo (lineage I-10). The distribution of *C. poensis* spans three major rivers (Congo, Ogooue and Sanaga) that are barriers to dispersal in many mammals across the Guineo-Congolian forest (Quérouil et al., 2003; Telfer et al., 2003; Anthony et al., 2007; Bohoussou et al., 2015). Distinct lineages were found on each bank of the Sanaga River (I-2 and I-6) and on each bank of the Congo River (I-10 and I-11). Lineages I-2 and I-6 are not sister-lineages, suggesting that the Sanaga River probably did not drive the diversification between these two lineages. The node

supporting the sister relationship between lineages I-10 and I-11 is not supported ($PP < 0.56$). Lineage I-8 seems to be restricted to the left bank of the Ogooue River. Lineages I-5 and I-7 are present on both banks of the Ogooue River. Although phylogeographic studies have been widely used to support either the refuge hypothesis or the riverine barrier hypothesis, these interpretations can be subjective (Portik *et al.*, 2017). Given the geographic arrangement of proposed forest refugia and rivers it is often not possible to decide whether the observed genetic pattern is better explained by forest refugia or riverine barriers in central Africa. An improved method for testing the location and number of refugia involves inferring areas of habitat stability through environmental niche modelling and coupling these results with phylogeographic and demographic inferences. The identification of areas of species-specific habitat stability can provide strong evidence for the impact of forest refugia on diversification patterns, but this approach has been rarely used for African taxa (but see Faye *et al.*, 2016; Portik *et al.*, 2017). An important consideration is that the locations of areas of habitat stability may vary greatly between species due to differences in their ecological traits. Because the geographic positions of forest refugia directly guide predictions about spatial genetic diversity and gene flow, the identification of species-specific climatically stable regions should ideally be used to inform phylogeographic hypotheses and evaluate the relevance of classically defined refugia. In the future it would be very interesting to try to identify these areas of climatic stability for the *C. poensis* complex through ecological niche modelling. However we consider that the number of sampling localities in which *C. poensis* was correctly identified is presently too low to perform such analyses. Genomic data could also provide the necessary signal to distinguish between demographic models derived from forest refugia and riverine barriers (Portik *et al.*, 2017).

Tables

Table 1: Geographical distribution and habitat requirements of the main lineages recovered in the phylogenetic analysis. Data based on Hutterer & Happold (1983), Nicolas *et al.* (2009); Happold & Happold (2013) and the present study (see Suppl Table S1). The four biogeographical units (Guinean Rainforest, Congolian Rainforest, Sudanian Savannah, Zambezian woodland) were defined based on Happold & Lock (2013) and Sayre *et al.* (2013).

Lineage	Distribution	Habitat
I-1	Congolian Rainforest (Bioko Island)	Forest
I-2	Congolian Rainforest (SW Cameroon)	Forest
I-3	Congolian Rainforest (E Cameroon, SW Central African Republic))	Forest, fallow
I-4	Congolian Rainforest (SW Cameroon)	Forest
I-5	Congolian Rainforest (Gabon, NW Congo Republic)	Forest
I-6	Congolian Rainforest (S Cameroon)	Forest
I-7	Congolian Rainforest (N Gabon)	Forest
I-8	Congolian Rainforest (S Gabon)	Forest
I-9	Congolian Rainforest (SE Gabon, SW Congo Republic)	Forest
I-10	Congolian Rainforest (E Congo Republic)	Forest/savanna mosaïc
I-11	Zambezian woodland (S Democratic Republic of Congo, N Zambia)	Disturbed secondary forest, ecotone forest/grassland
II	Congolian Rainforest (Príncipe Island)	Old banana plantations, houses, near forest, rocky outcrops

III	Zambeian woodland (East Africa)	Fields, bushlands and grasslands
IV	Guinean and Congolian Rainforest (Guinea to Nigeria)	Grassland within the rainforest, forest, forests relicts in derived savanna, fallows and plantations
V	Guinean Rainforest (Guinea to W Ghana)	Mostly in open habitats such as grasslands and fields, only occasionally in forest relicts
VI	Guinean Rainforest (W coastal Guinea)	Small gallery forests isolated in savannahs
VII	Guinean and Congolian Rainforests (Guinea to Nigeria)	Forest
VIII	Guinean Rainforest (E Ivory Coast and W Ghana)	Wet, swampy forests near the coast
IX	Sudanian Savannah (Senegal to Sudan)	Grassy areas, secondary bush, rocky grasslands recently burned, swamp areas, cultivated areas

Figures legends

Figure 1: Phylogenetic tree estimated from Maximum Likelihood analysis of Cytb and 16S sequences. Populations/species at tips have been collapsed into triangles where the apex represents the deepest split within the group. Values at nodes represent posterior probabilities (Bayesian analysis) and bootstrap support (Maximum Likelihood analysis), respectively. A zoom on lineage I is shown in the upper framed window.

Figure 2: Map of sampling points showing the distribution of the main phylogenetic lineages identified (differentiated by colours) on the basis of mtDNA analyses (16S + Cytb). To improve clarity, geographically close localities were grouped together on the map. CAR = Central African Republic, DRC = Democratic Republic of Congo.

Figure 3: Chronogram of the *C. poensis* complex with ancestral areas reconstructions indicated, and results of molecular species delimitation via BPP and STACEY based on nine candidate species. To improve clarity outgroup taxa are not represented (but see Fig. S4 in Appendix S3 for the tree including outgroup taxa). Numbers above branches represent posterior probabilities (60 million generations per run, two runs). Terminal taxa are colour coded according to the biogeographical areas in which they occur: a – Guinean Rainforest, b – Congolian Rainforest, c – Sudanian Savannah, d- Zambeian woodland. Colors at nodes indicate most likely ancestral areas recovered by BioGeoBEARS under the DIVA model. Dispersion and vicariance events are indicated at nodes by the letters D and V, respectively. The tree displayed here is the maximum lineage credibility tree from BEAST analyses (3 mitochondrial – Co1, Cytb, 16S -; and 3 nuclear genes - BRCA1, STAT5A, HDAC2 -). Node numbers refer to Table S2 in Appendix S3.

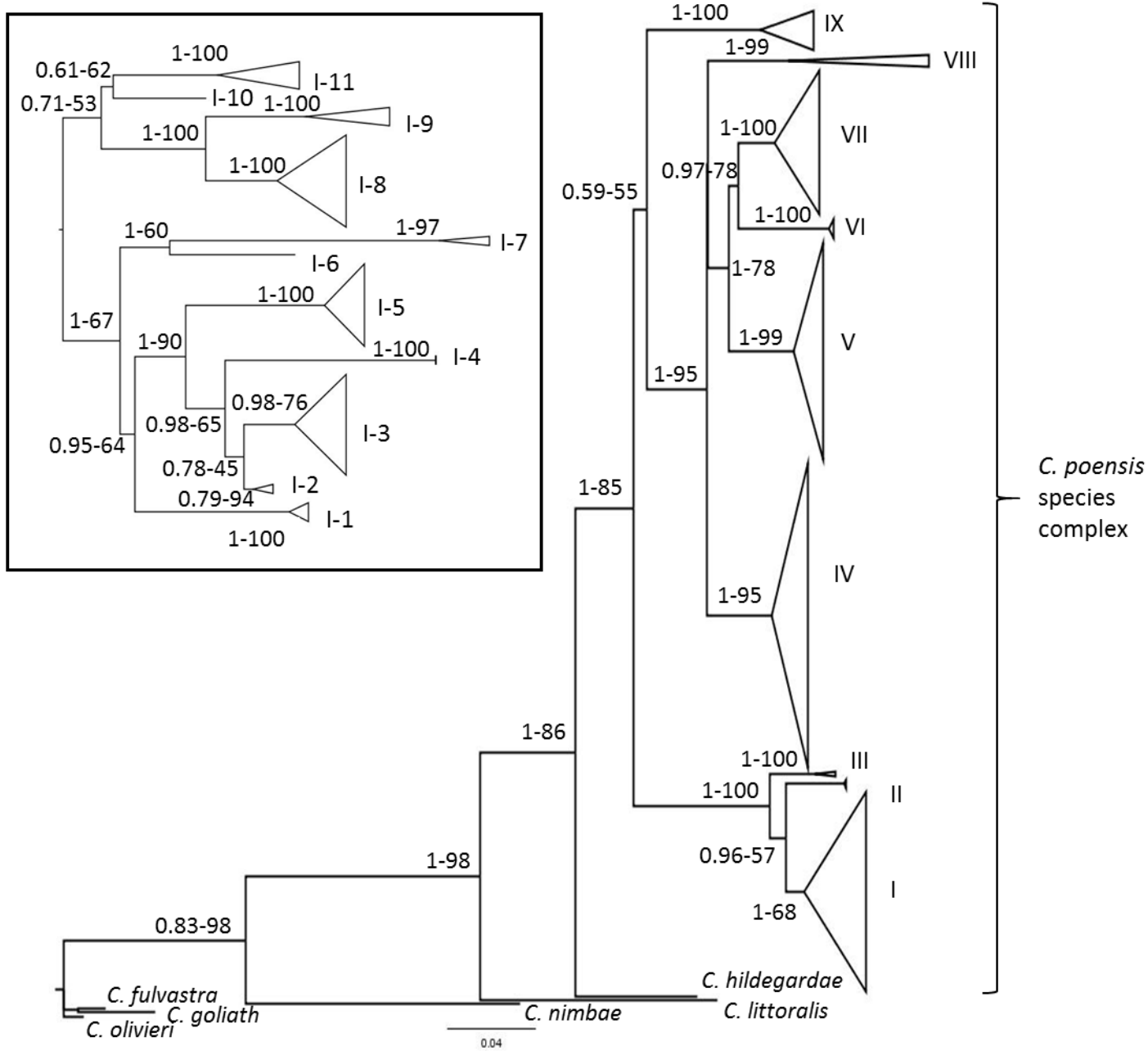


Figure 1

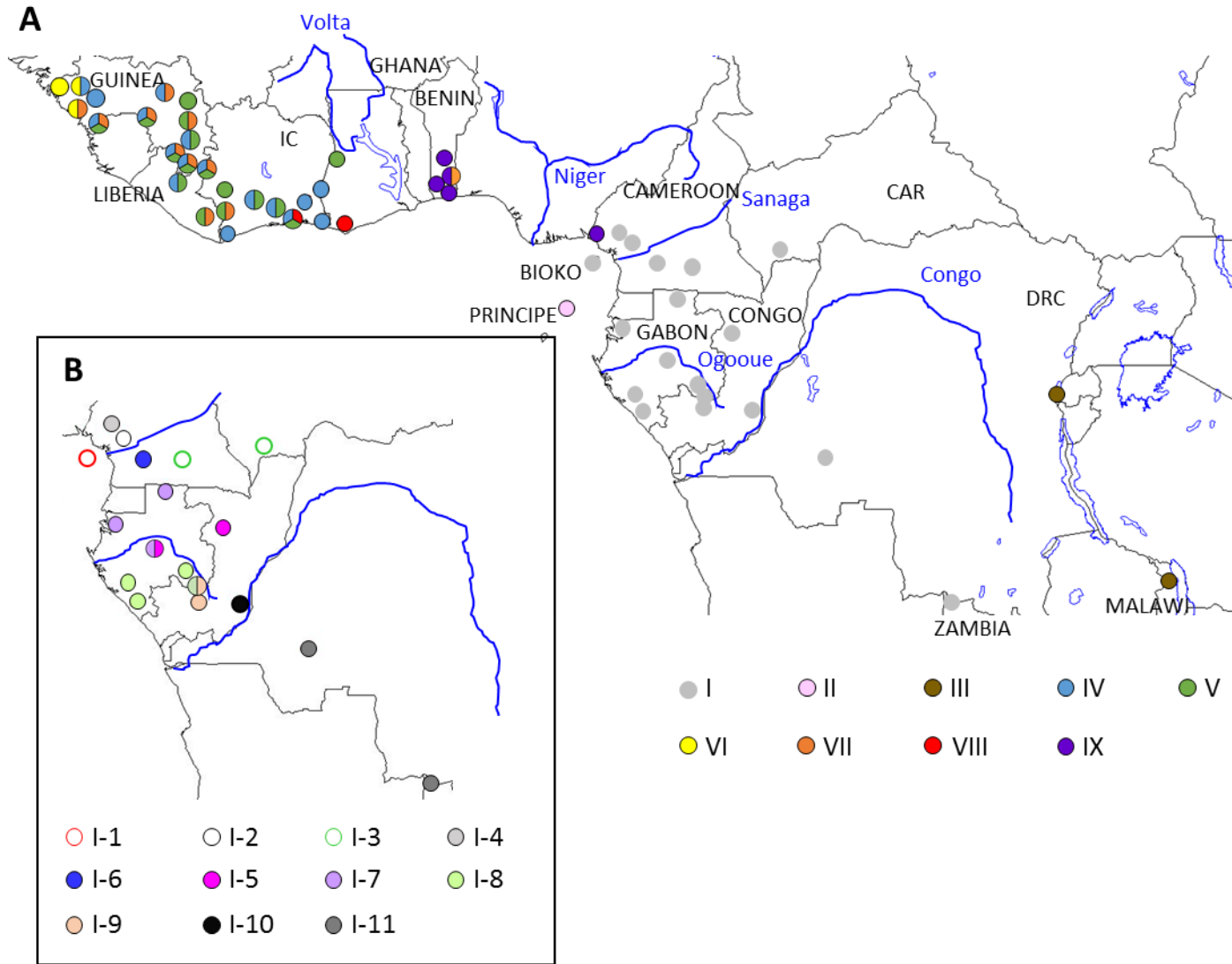


Figure 2

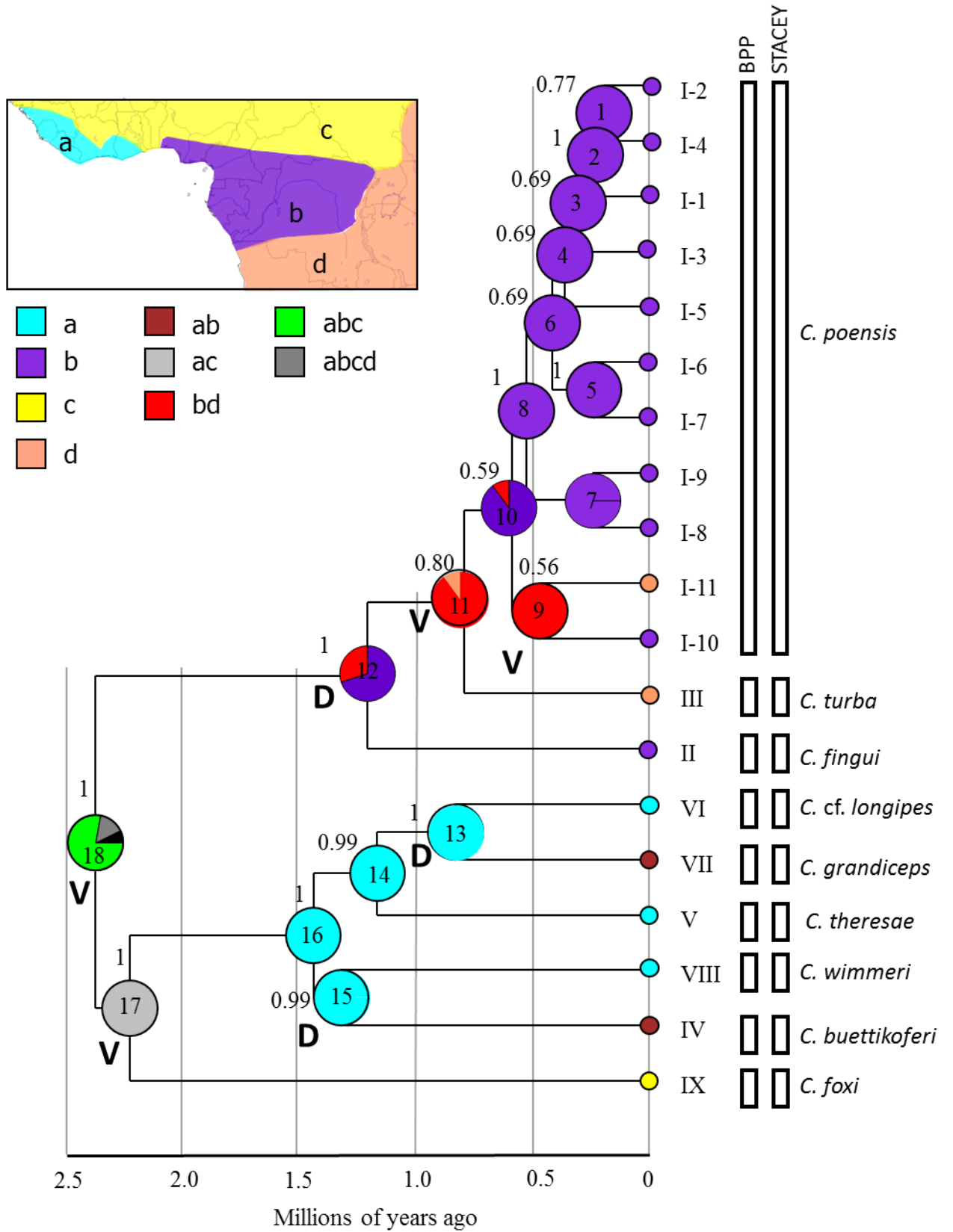


Figure 3

References

- Anhuf, D., P Ledru, M.P., Behling, H., Da Cruz JR, F.W., Cordeiro, R.C., Van Der Hammen, T., Karmann, I., Marengo, J.A., De Oliveira, P.E., Pessenda, L., Siffedine, A., Albuquerque, A.L. & Da Silva Dias, P.L. (2006) Paleo-environmental change in Amazonian and African rainforest during the LGM. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **239**, 510-527
- Anthony, N.M., Johnson-Bawe, M., Jeffery, K., Clifford, S.L., Abernethy, K.A., Tutin, C.E., Lahm, S.A., White, L.J., 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**, 20432-6
- Bell, R.C., Parra, J.L., Badjedjea, G., *et al.* (2017) Idiosyncratic responses to climate-driven forest fragmentation and marine incursions in reed frogs from Central Africa and the Gulf of Guinea Islands. *Molecular ecology*, **26**, 5223-5244
- Bohoussou, K.H., Cornette, R., Akpatou, B., Colyn, M., Kerbis Peterhans, J., Kennis, J., Šumbera, R., Verheyen, E., N'Goran, E., Katuala, P. & Nicolas, V. (2015) The phylogeography of the rodent genus *Malacomys* suggests multiple Afrotropical Pleistocene lowland forest refugia. *Journal of Biogeography*, **42**, 2049-2016
- Bouckaert, R., Heled, J., Kuhnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology*, **10**, e1003537
- Brosset, A. (1988) Le peuplement de mammifères insectivores des forêts du nord-est du Gabon. *Revue d'Ecologie, La Terre et la Vie*, **43**, 23-46
- Bryja, J., Šumbera, R., Kerbis Peterhans, J.C., Aghová, T., Bryjová, A., Mikula, O., Nicolas, V., Denys, C. & Verheyen, E. (2017) Evolutionary history of the thicket rats (genus *Grammomys*) mirrors the evolution of African forests since late Miocene. *Journal of Biogeography*, **44**, 182–194
- Ceríaco, L.M.P., Marques, M.P., Jacquet, F., Nicolas, V., Colyn, M., Denys, C., Sardinha, P.C. & Bastos-Silveira, C. (2015) Description of a new endemic species of shrew (Mammalia, Soricomorpha) from Príncipe Island (Gulf of Guinea). *Mammalia*, **79**, 325-341
- Colangelo, P., Verheyen, E., Leirs, H., Tatar, C., Denys, C., Dobigny, G., Duplantier, J.M., Brouat, C., Granjon, L. & Lecompte, E. (2013) A mitochondrial phylogeographic scenario for the most widespread African rodent, *Mastomys natalensis*. *Biological Journal of the Linnean Society*, **108**, 901–916
- Couvreur, T.L., Porter-Morgan, H., Wieringa, J.J. & Chatrou, L.W. (2011) Little ecological divergence associated with speciation in two African rain forest tree genera. *BMC Evolutionary Biology*, **11**, 296
- Cowling, S.A., Cox, P.M., Jones, C.D., Maslin, M.A., Peros, M. & Spall, S.A. (2008) Simulated glacial and interglacial vegetation across Africa: implications for species phylogenies and trans-African migration of plants and animals. *Global Change Biology*, **14**, 827-840
- deMenocal, P. (2004) African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters*, **220**, 3-24
- Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969-73
- Dubey, S., Salamin, N., Ruedi, M., Barriere, P., Colyn, M. & Vogel, P. (2008) Biogeographic origin and radiation of the Old World crocidurine shrews (Mammalia: Soricidae)

- inferred from mitochondrial and nuclear genes. *Molecular Phylogenetics and Evolution*, **48**, 953-963
- Dupont, L.M., Jahns, S., Marret, F. & Ning, S. (2000) Vegetation change in equatorial West Africa: time-slices for the last 150 ka. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **155**, 95-122
- Faye, A., Deblauwe, V., Mariac, C., Richard, D., Sonke, B., Vigouroux, Y. & Couvreur, T.L.P. (2016) Phylogeography of the genus *Podococcus* (Palmae/Arecaceae) in Central African rain forests: Climate stability predicts unique genetic diversity. *Molecular Phylogenetics and Evolution*, **105**, 126-138
- Freedman, A.H., Thomassen, H.A., Buermann, W. & Smith, T.B. (2010) Genomic signals of diversification along ecological gradients in a tropical lizard. *Molecular Ecology*, **19**, 3773-88
- Fuchs, J. & Bowie, R.C. (2015) Concordant genetic structure in two species of woodpecker distributed across the primary West African biogeographic barriers. *Molecular Phylogenetics and Evolution*, **88**, 64-74.
- Fuchs, J., Fjeldså, J. & Bowie, R.C.K. (2017) Diversification across major biogeographic breaks in the African Shining/Square-tailed Drongos complex (Passeriformes: Dicuridae). *Zoologica Scripta*, **46**, 27-41.
- Grubb, P., Jones, T.S., Davies, A.G., Edberg, E., Starin, E.D. & Hill, J.E. (1998) *Mammals of Ghana, Sierra Leone and the Gambia*. The Trendline Press, Cornwall.
- Haffer, J. (1997) Alternative models of vertebrate speciation in Amazonia: an overview. *Biodiversity and Conservation*, **6**, 451-476
- Happold, D. & Lock, J.M. (2013) The Biotic Zones of Africa. *Mammals of Africa Volume I* (ed. by J. Kingdon, D.C.D. Happold, M. Hoffmann, M. Happold and J. Kalina), pp. 57-74. Bloomsbury.
- Happold, M. & Happold, D.C.D. (2013) *Mammals of Africa, volume IV: Hedgehogs, Shrews and Bats*. Bloomsbury Publishing, London.
- Hassanin, A., Khouider, S., Gembu, G.C., S, M.G., Kadjo, B., Nesi, N., Pourrut, X., Nakoune, E. & Bonillo, C. (2015) The comparative phylogeography of fruit bats of the tribe Scotonycterini (Chiroptera, Pteropodidae) reveals cryptic species diversity related to African Pleistocene forest refugia. *Comptes Rendus Biologies*, **338**, 197-211
- Heim de Balsac, H. (1974) Les insectivores de Lamto (Côte d'Ivoire). **38**, 637-646
- Heim de Balsac, H. & Meester, J. (1977) Order Insectivora. *The mammals of Africa: an identification manual* (ed. by J. Meester and H.W. Setzer), pp. 1-29. Smithsonian Press, Washington.
- Huntley, J.W. & Voelker, G. (2016) Cryptic diversity in Afro-tropical lowland forests: The systematics and biogeography of the avian genus *Bleda*. *Molecular Phylogenetics and Evolution*, **99**, 297-308
- Hutterer, R. & Joger, U. (1982) Kleinsäuger aus dem Hochland von Adamaoua, Kamerun. *Bonner Zoologische Beiträge*, **33**, 119-132
- Hutterer, R. & Happold, D.C.D. (1983) The shrews of Nigeria (Mammalia: Soricidae). *Bonner Zoologische Monographien*, **18**, 5-79
- Hutterer, R. (2005) Order Soricomorpha. *Mammal Species of the World: A Taxonomic and Geographic Reference, third ed.* (ed. by D.E. Wilson and D.M. Reeder), pp. 220-311. Johns Hopkins University Press, Baltimore.
- Jacquet, F., Nicolas, V., Bonillo, C., Cruaud, C. & Denys, C. (2012) Barcoding, molecular taxonomy and exploration of the diversity of shrews (Soricomorpha: Soricidae) on Mount Nimba (Guinea). *Zoological Journal of the Linnean Society*, **166**, 672-687
- Jacquet, F., Hutterer, R., Nicolas, V., Decher, J., Colyn, M., Couloux, A. & Denys, C. (2013) A new status for two African giant forest shrew *Crocidura goliath goliath* and

- Crocidura goliath nimbasilvanus* (Mammalia: Soricomorpha), based on molecular and geometric morphometric analyses. *African Zoology*, **48**, 13-29
- Jacquet, F., Nicolas, V., Colyn, M., Kadjo, B., Hutterer, R., Akpatou, B., Cruaud, C. & Denys, C. (2014) Forest refugia and riverine barriers promote diversification in the West African pygmy shrew (*Crocidura obscurior* complex, Soricomorpha). *Zoologica Scripta*, **43**, 131-148
- Jacquet, F., Denys, C., Verheyen, E., Bryja, J., Hutterer, R., Kerbis Peterhans, J.C., Stanley, W.T., Goodman, S.M., Couloux, A., Colyn, M. & Nicolas, V. (2015) Phylogeography and evolutionary history of the *Crocidura olivieri* complex (Mammalia, Soricomorpha): from a forest origin to broad ecological expansion across Africa. *BMC Evolutionary Biology*, **15**, 71
- Jones, G. (2017) Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *Journal of Mathematical Biology*, **74**, 447-467
- Klempa, B., Fichet-Calvet, E., Lecompte, E., Auste, B., Aniskin, V., Meisel, H., Barriere, P., Koivogui, L., ter Meulen, J. & Kruger, D.H. (2007) Novel hantavirus sequences in Shrew, Guinea. *Emerging Infectious Diseases*, **13**, 520-2
- Leaché, A.D. & Fujita, M.K. (2010) Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proceedings of the Royal Society Biological Sciences*, **277**, 3071-3077
- Lee, D.-C., Halliday, A.N., Fitton, J.G. & Poli, G. (1994) Isotopic variations with distance and time in the volcanic islands of the Cameroon line: evidence for a mantle plume origin. *Earth and Planetary Science Letters*, **123**, 119-138
- Maddalena, T. (1990) Systematics and biogeography of afrotropical and palaeartic shrews of the genus *Crocidura* (Insectivora : Soricidae) : An electrophoretic approach. *Vertebrates in the tropics* (ed. by G. Peters and R. Hutterer), pp. 297-308. Museum Alexander Koenig, Bonn.
- Maley, J. (1996) The African rain forest - main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh*, **104B**, 31-73
- Marzoli, A., Piccirillo, E.M., Renne, P.R., Bellieni, G., Iacumin, M., Nyobe, J.B. & Tongwa, A.T. (2000) The Cameroon volcanic line revisited: petrogenesis of continental basaltic magmas from lithospheric and asthenospheric mantle sources. *Journal of Petrology*, **41**, 87-109
- Maslin, M.A., Brierleya, 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
- Matzke, N. (2013) BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts. In. University of California, Berkeley
- Mayr, E. & O'Hara, E.J. (1986) The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. *Evolution*, **40**, 55-67
- Measey, J.G., Vences, M., Drewes, R.C., Chiari, Y., Melo, M. & Bourles, B. (2007) Freshwater paths across the ocean: molecular phylogeny of the frog *Ptychadena newtoni* gives insights into amphibian colonization of oceanic islands. *Journal of Biogeography*, **34**, 7-20
- Mein, P.R. & Pickford, M. (2006) Late Miocene micromammals from the Lukeino Formation (6.1 to 5.8 Ma), Kenya. *Bulletin Mensuel de la Societe Linneenne de Lyon*, **75**, 183-223
- Meylan, A. (1971) Chromosomes de Soricidés de Côte d'Ivoire (Mammalia, Insectivore). *Revue Suisse de Zoologie*, **78**, 603-613

- Meylan, A. & Vogel, P. (1982) Contribution a la cytotaxonomie des Soricidés (Mammalia, Insectivora) de l'Afrique occidentale. *Cytogenetic and Genome Research*, **34**, 83-92
- Missoup, A.D., Nicolas, V., Wendelen, W., Keming, E., Bilong Bilong, C.F., Couloux, A., Atanga, E., Hutterer, R. & Denys, C. (2012) Systematics and diversification of *Praomys* species (Rodentia: Muridae) endemic to the Cameroon Volcanic Line (West Central Africa). *Zoologica Scripta*, **41**, 327–345
- Mitchell, M.W., Locatelli, S., Ghobrial, L., Pokempner, A.A., Sesink Clee, P.R., Abwe, E.E., Nicholas, A., Nkambi, L., Anthony, N.M., Morgan, B.J., Fotso, R., Peeters, M., Hahn, B.H. & Gonder, M.K. (2015) The population genetics of wild chimpanzees in Cameroon and Nigeria suggests a positive role for selection in the evolution of chimpanzee subspecies. *BMC Evolutionary Biology*, **15**, 3
- Moritz, C., Patton, J.L., Schneider, C.J. & Smith, T.B. (2000) Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics*, **31**, 533-563
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858
- Nicolas, V., Quérrouil, S., Verheyen, E., Verheyen, W., Mboumba, J.F., Dillen, M. & Colyn, M. (2006) Mitochondrial phylogeny of African wood mice, genus *Hylomyscus* (Rodentia, Muridae): implications for their taxonomy and biogeography. *Molecular Phylogenetics and Evolution*, **38**, 779-93
- Nicolas, V., Bryja, J., Akpatou, B., Konecny, A., Lecompte, E., Colyn, M., Lalis, A., Couloux, A., Denys, C. & Granjon, L. (2008a) Comparative phylogeography of two sibling species of forest-dwelling rodent (*Praomys rostratus* and *P. tullbergi*) in West Africa: different reactions to past forest fragmentation. *Molecular Ecology*, **17**, 5118-5134
- Nicolas, V., Mboumba, J.F., Verheyen, E., Denys, C., Lecompte, E., Olayemi, A. & Colyn, M. (2008b) Phylogeographical structure and regional history of *Lemniscomys striatus* (Rodentia: Muridae) in tropical Africa. *Journal of Biogeography*, **35**, 2072-2089
- Nicolas, V., Barrière, P., Tapiero, A. & Colyn, M. (2009) Shrew species diversity and abundance in Ziama Biosphere Reserve, Guinea: comparison among primary forest, degraded forest and restoration plots. *Biodiversity and Conservation*, **18**, 2043-2061
- Nicolas, V., Akpatou, B., Wendelen, W., Kerbis Peterhans, J., Olayemi, A., Decher, J., Missoup, A.D., Denys, C., Barrière, P., Cruaud, C. & Colyn, M. (2010) Molecular and morphometric variation in two sibling species of the genus *Praomys* (Rodentia: Muridae): implications for biogeography. *Zoological Journal of the Linnean Society*, **160**, 397-419
- Olayemi, A., Nicolas, V., Hulselmans, J., Missoup, A.D., Fichet-Calvet, E., Amundala, D., Dudu, A., Dierckx, T., Wendelen, W., Leirs, H. & Verheyen, E. (2012) Taxonomy of the African Giant Pouched Rats (Nesomyidae: *Cricetomys*): molecular and craniometric evidence support an unexpected high species diversity. *Zoological Journal of the Linnean Society*, **165**, 700–719
- Orr, M.R. & Smith, T.B. (1998) Ecology and speciation. *Trends in Ecology and Evolution*, **13**, 502-506
- Plana, V. (2004) Mechanisms and tempo of evolution in the African Guineo–Congolian rainforest. *Philosophical Transactions of the Royal Society, Biological Sciences*, **359**, 1585-1594
- Portik, D.M., Leache, A.D., Rivera, D., Barej, M.F., Burger, M., Hirschfeld, M., Rodel, M.O., Blackburn, D.C. & Fujita, M.K. (2017) Evaluating mechanisms of diversification in a Guineo–Congolian tropical forest frog using demographic model selection. *Molecular Ecology*, **26**, 5245-5263

- Qu  rouil, 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**, 24-37
- Qu  rouil, S., Barri  re, P., Colyn, M., Hutterer, R., Dudu, A., Dillen, M. & Verheyen, E. (2005) A molecular insight into the systematics of African *Crocidura* (Crocidurinae, soricidae) using 16sRNA sequences. *Advances in the biology of shrews II* (ed. by J.F. Merritt, S. Churchfield, R. Hutterer and B.I. Sheftel), pp. 99-113. Special Publication of the International Society of Shrew Biologist 01, New York.
- Robert, C. & Chamley, H. (1987) Cenozoic evolution of continental humidity and paleoenvironment, deduced from kaolinite content of oceanic sediments. *Palaeogeography, Palaeoclimatology and Palaeoecology*, **60**, 171-187
- Rychlik, L., Ramalhinho, G. & Polly, P.D. (2006) Response to environmental factors and competition: skull, mandible and tooth shapes in Polish water shrews (*Neomys*, Soricidae, Mammalia). *Journal of Zoological Systematics and Evolutionary Research*, **44**, 339-351
- Salzmann, U. & Hoelzmann, P. (2005) The Dahomey Gap: an abrupt climatically induced rain forest fragmentation in West Africa during the late Holocene. *The Holocene*, **15**, 190-199
- Sayre, R., Comer, P., Hak, J., *et al.* (2013) *A new map of standardized terrestrial ecosystems of Africa*. Association of American Geographers, Washington, DC.
- Schlitter, D.A., Hutterer, R., Maddalena, T. & Robins, L.W. (1999) New karyotypes of shrews (Mammalia: soricidae) from Cameroon and Somalia. *Annals of Carnegie Museum*, **68**, 1-14
- 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
- Smith, T.B., Calsbeek, R., Wayne, R.K., Holder, K.H., Pires, D. & Bardeleben, C. (2005) Testing alternative mechanisms of evolutionary divergence in an African rain forest passerine bird. *Journal of evolutionary biology*, **18**, 257-68
- Smith, T.B., Thomassen, H.A., Freedman, A.H., Sehgal, R.N.M., Buermann, W., Saatchi, S., Pollinger, J., Mila, B., Pires, D., Valkinunas, G. & Wayne, R.K. (2011) Patterns of divergence in the olive sunbird *Cyanomitra olivacea* (Aves: Nectariniidae) across the African rainforest–savanna ecotone. *Biological Journal of the Linnean Society*, **103**, 821-835
- Telfer, P.T., Souquiere, 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**, 2019-24
- Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680
- Trauth, M.H., Larrasoana, J.C. & Mudelsee, M. (2009) Trends, rhythms and events in Plio-Pleistocene African climate. *Quaternary Science Reviews*, **28**, 399-411
- Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V., Damschen, E.I., Jonathan Davies, T., Grytnes, J.A., Harrison, S.P., Hawkins, B.A., Holt, R.D., McCain, C.M. & Stephens, P.R. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310-1324
- Yang, Z. (2015) The BPP program for species tree estimation and species delimitation *Current Zoology*, **61**, 854-865

- Yu, Y., Harris, A.J., Blair, C. & He, X. (2015) RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular phylogenetics and evolution*, **87**, 46-49
- Zhen, Y., Harrigan, R.J., Rugg, K.C., Anderson, E.C., Ng, T.C., Lao, S., Lohmueller, K.E. & Smith, T.B. (2017) Genomic divergence across ecological gradients in the Central African rainforest songbird (*Andropadus virens*). *Molecular Ecology*, **26**, 4966-4977

Biosketch

Violaine Nicolas is a researcher and curator of small mammals at the National Museum of Natural History (MNHN, Paris). Her interests include the biogeography, ecology and taxonomy of small mammals.

Author contributions: V.N. conceived the ideas; V.N., A.K., S.K., L.D., A.L., M.C. and C.D. collected the specimens; V.N., F.J. and R.H. identified the species; V.N. and F.J. performed the molecular lab work and analysed the data; V.N. drafted the manuscript, which was approved by all authors.

Supporting information

Appendix S1: Supplementary details of individuals.

Appendix S2: Additions to methodology: phylogenetic analyses, species delimitation analyses and divergence time analyses.

Appendix S3: Additions to results.