Evolutionary history of the thicket rats (genus **Grammomys**) mirrors the evolution of African forests since late Miocene

Original Article

Evolutionary history of the thicket rats (genus *Grammomys*) mirrors the evolution of African forests since late Miocene

Josef Bryja[1,2*], Radim Šumbera[3], Julian C. Kerbis Peterhans[4,5], Tatiana Aghová[1,2], Anna Bryjová[1], Ondřej Mikula[1,6], Violaine Nicolas[7], Christiane Denys[7] and Erik Verheyen[8,9]

[1] Institute of Vertebrate Biology of the Czech Academy of Sciences, 603 65, Brno, Czech Republic

[2] Department of Botany and Zoology, Faculty of Science, Masaryk University, 611 37, Brno, Czech Republic

[3] Department of Zoology, Faculty of Science, University of South Bohemia, 370 05, České Budějovice, Czech Republic

[4] College of Professional Studies, Roosevelt University, 60605, Chicago IL, USA

[5] Field Museum of Natural History, 60605, Chicago, IL, USA

[6] Institute of Animal Physiology and Genetics of the Czech Academy of Sciences, 602 00, Brno, Czech Republic

[7] Institute of Systematics and Evolution, UMR7205 CNRS-MNHN-EPHE-Sorbonne Universités, 75005, Paris, France

[8] Royal Belgian Institute for Natural Sciences, Operational Direction Taxonomy and Phylogeny, 1000, Brussels, Belgium

[9] Evolutionary Ecology Group, Biology Department, University of Antwerp, 2020, Antwerp, Belgium

***Correspondence**: Josef Bryja, Institute of Vertebrate Biology of the Czech Academy of Sciences, Research Facility Studenec, Studenec 122, 675 02 Koněšín, Czech Republic; E-mail: bryja@brno.cas.cz

Running head: Evolution of climbing rats and African forests

Keywords: Arvicanthini, coastal forests, late Miocene, lowland forests, mountain forests, phylogeography, Plio-Pleistocene climate changes, Rodentia, tropical Africa

ABSTRACT

Aim *Grammomys* are mostly arboreal rodents occurring in forests, woodlands and thickets throughout sub-Saharan Africa. We investigated whether the divergence events within the genus follow the existing evolutionary scenario for the development of African forests since the late Miocene.

Location Sub-Saharan African forests and woodlands.

Methods We inferred the molecular phylogeny of *Grammomys* using Bayesian and maximum likelihood methods and DNA sequences of 351 specimens collected from across the distribution of the genus. We mapped the genetic diversity, estimated the divergence times by a relaxed clock model and compared evolution of the genus with forest history.

Results Phylogenetic analysis confirms the monophyly of *Grammomys* and reveals five main *Grammomys* lineages with mainly parapatric distributions: (1) the poensis group in Guineo-Congolese forests; (2) the selousi group with a distribution mainly in coastal forests of southern and eastern Africa; (3) the dolichurus group restricted to the easternmost part of South Africa; (4) the macmillani group in the northern part of eastern and Central Africa with one isolated species in Guinean forests; and (5) the surdaster group, widely distributed in eastern Africa south of the equator. Every group contains well supported sublineages suggesting the existence of undescribed species. The earliest split within the genus (groups 1 versus 2-5) occurred in the late Miocene, and coincides with the formation of the Rift Valley which resulted in the east-west division of the initially pan-African forest. The subsequent separation between groups (2 versus 3-5) also dates to the end of the Miocene and suggests the split between *Grammomys* from

coastal to upland forests in eastern Africa followed by a single dispersal event into western Africa during the Pleistocene.

Conclusions The evolutionary history of the genus *Grammomys* reflects closely the accepted scenario of major historical changes in the distribution of tropical African forests since the late Miocene.

INTRODUCTION

Tropical forests in Africa contain rich biodiversity. For example, the Eastern Arc Mountains support ca 3300 km^2 of forest that harbours 211 endemic or nearly endemic vertebrate species (Rovero *et al*., 2014) whereas the Albertine Rift mountains host the largest suite of endemic mammals on the continent (Plumptre *et al.*, 2007). However, biological diversity is not equally distributed across the African tropics (e.g. de Klerk *et al*., 2002), but knowledge of its distribution is crucial in prioritizing conservation activity.

A recent study of forest composition in tropical Africa identified six floristic clusters associated with particular environmental conditions (Fayolle *et al*., 2014; Fig. 1). The origin of these forest types is the outcome of a complex evolutionary history that started from a single continuous equatorial forest that covered sub-Saharan Africa during the period of humid climate of the Early and Middle Miocene (Plana, 2004). By the Late Miocene, tectonic uplift created the Rift Valley and split the pan-African rainforest into the Guineo-Congolese forests in western and Central Africa and the forests situated east of the rift. The rift formation combined with declining global temperatures and changes in monsoon winds resulted in an arid climate that caused the disappearance of forests along the slope of the rift mountains, hence creating the so-called "arid corridor" that periodically connected the northern (Sudanian and Somalian) and southern (Zambezian) savannas (Bobe, 2006). However, some old mountain ranges (e.g. Albertine Rift and Eastern Arc mountains) served as long-term forest refugia allowing the evolution of species-rich communities (e.g. Loader *et al*., 2014). Throughout this period, West (=Guinean) and Central (=Congolese) African forests continued to exist as a single unit

that underwent periodic fragmentation during the Pleistocene (Maley, 1996). Since the Middle Pleistocene, the forested mountain chains in eastern Africa also underwent fragmentation, as suggested by increasing proportions of C4 vegetation, most likely indicating the origin of the current tropical grasslands around these mountains (Cerling, 1992).

Based on the concept of phylogenetic niche conservatism (Wiens & Donoghue, 2004), this study proposes to use a phylogeographic approach for forest-dwelling mammals to investigate the evolutionary history and past connections among African forests. Phylogeographic patterns for widely distributed taxa with specific ecological requirements can be used to test alternative hypotheses of African forest evolution. Although an increasing number of studies have used this approach on sub-Saharan vertebrates (e.g. Huntley & Voelker, 2016), so far few studies have targeted widespread taxa living in various forest types (for a rare example see Couvreur *et al.*, 2008). It is in this context that we have used DNA sequences to infer for the first time the phylogeny of thicket rats of the genus *Grammomys.* These partly arboreal rodents, belonging to the tribe Arvicanthini (Ducroz *et al.*, 2001, Lecompte *et al.*, 2008, Missoup *et al.*, 2016), occur in a variety of forests and woodlands in sub-Saharan Africa. Although 11 to 14 *Grammomys* species are currently recognized, the monophyly of the genus remains uncertain and its taxonomic sampling incomplete (Musser & Carleton, 2005). Because these climbing rats are widely distributed in sub-Saharan forests and woodlands, they may represent a suitable model group to trace the evolutionary histories of the forested habitats in which they occur. Moreover, the fact that they represent a genus originating

during the radiation of Arvicanthini ca 8 Ma (Ducroz *et al*., 2001) provides an opportunity to study their evolutionary history since the Late Miocene, a crucial era for the development of African forests.

Over the past decades we have collected material of *Grammomys* rats from a large part of their distribution for molecular sampling. We inferred for the first time the phylogeny of the genus that we used together with estimated divergence dates as a proxy for the evolutionary histories of the different forest types in tropical Africa in which they occur. Lastly, based on observed diversity, we identified the geographic areas and genetic clades in which future taxonomic studies are most likely to result in discoveries of new *Grammomys* species.

MATERIALS AND METHODS

Sampling

The study is based on 351 specimens of *Grammomys* genotyped for at least one genetic marker (Table S1 in Appendix S1). The tissue samples were stored in 96% ethanol, DMSO or liquid nitrogen until DNA extraction. All fieldwork complied with legal regulations in the respective African countries and sampling was carried out in accordance with local legislation (see Acknowledgements). In total, the analysed dataset includes genetic information on specimens collected from 170 localities in 18 African countries (Fig. 1).

DNA sequencing

We collected the sequences for mitochondrial markers, either the cytochrome *b* gene (*CYTB*, 334 new sequences and 11 from GenBank), the 16S rRNA gene (*16S*, 164 new sequences) or both, for all 351 specimens. For 112 selected specimens we also obtained sequences of the nuclear gene for interphotoreceptor binding protein (*IRBP*, 110 new sequences and two from GenBank) to match detected mitochondrial diversity as far as possible with sequences from a nuclear locus (Table S1 in Appendix S1). Primers and PCR protocols for DNA from fresh material are detailed in Table S1 in Appendix S2. PCR products were Sanger sequenced from both sides in a commercial laboratory. Genetic data obtained from fresh material were complemented by eight museum samples (mostly dry skins) (Appendix S1) pyrosequenced on GS Junior using the *CYTB* mini-barcode protocol (Galan *et al*., 2012). This approach was used for samples from geographical areas that are difficult to access today or from the type localities of *G. dryas* and *G. poensis* (see more details in Bryja *et al*., 2014a)

Phylogenetic reconstructions within* Grammomys *and genetic distances

Sequences of *CYTB*, *16S* and *IRBP* were edited and aligned in SeqScape 2.5 (Applied Biosystems), producing final alignments of 1140, 575 and 1261 bp, respectively. We first reconstructed the mitochondrial phylogeny using the concatenated *CYTB* and *16S* dataset, because preliminary separate analyses of these two loci provided very similar topologies (not shown). We performed the final phylogenetic analyses with a reduced mtDNA dataset of 157 specimens (155 sequences of *CYTB* and 115 of *16S*) (Appendix S1), representing the main mtDNA lineages identified by preliminary analyses (not shown). The remaining 194 specimens (identical and/or shorter sequences from the same

or neighbouring localities) were unambiguously assigned to particular lineages by neighbour-joining analysis (bootstrap support > 90%; not shown) in MEGA 6.06 (Tamura *et al*., 2013). These data were used to increase the precision with which we mapped the geographical distribution of phylogenetic clades and assigned type material to particular genetic groups. To assess the monophyly of *Grammomys* reliably, we used as outgroups 24 mitochondrial sequences of 13 genera within the tribe Arvicanthini (*sensu* Lecompte *et al*., 2008), eight sequences of species from other tribes of Murinae and one species of the subfamily Gerbillinae (Table S2 in Appendix S1). We used PARTITIONFINDER 1.0.1 (Lanfear *et al*., 2012) to detect partitions and the most suitable substitution models simultaneously. Using the Bayesian information criterion (BIC), the best scheme supported four partitions (Table S2 in Appendix S2).

Mitochondrial phylogeny was analysed by maximum likelihood (ML) and Bayesian inference (BI) approaches. ML analysis was performed using RAXML 8.0 (Stamatakis, 2014). Because simpler models are not available in RAXML, the GTR+G model (option -m GTRGAMMA) was selected for the four partitions (option -q). The robustness of the nodes was evaluated by the default bootstrap procedure with 1,000 replications (option -# 1000). Bayesian analysis of evolutionary relationships was performed in MRBAYES 3.2.1 (Ronquist & Huelsenbeck, 2003). Three heated and one cold chain were employed in a partitioned analysis, and runs were initiated from random trees. Two independent runs were conducted with 5 million generations each and trees and parameters were sampled every 1000 generations. Convergence was checked using TRACER 1.5 (Rambaut & Drummond, 2007). For each run, the first 25% of sampled trees were discarded as burn-

179 in. Bayesian posterior probabilities (PP) were used to assess branch support of the
180 Markov chain Monte Carlo (MCMC) tree.
181
182 The number of base substitutions per site of *CYTB* averaging over all sequence pairs
183 between and within groups was calculated as uncorrected *p*-distance as well as using the
184 Kimura 2-parameter (K2P) model. The groups were defined on the basis of phylogenetic
185 analysis (see below and Fig. 2). This analysis was conducted in MEGA 6.06 and involved
186 155 *CYTB* sequences representing 28 mitochondrial lineages.
187
188 For the phylogenetic analyses of 101 retained nuclear *IRBP* sequences from all but one of
189 the mitochondrial lineages (m6 was missing because no *IRBP* sequence was obtained),
190 heterozygous sequences were phased using FASTPHASE (Scheet & Stephens, 2006)
191 implemented in DNASP 5.10 (Librado & Rozas, 2009). Using PARTITIONFINDER 1.0.1
192 and BIC, the best scheme supported two partitions (Table S2 in Appendix S2).
193 Phylogenetic analyses were performed in RAXML and MRBAYES as described above.
194
195 ***Dated phylogeny of Arvicanthini***
196 The ML and BI analyses of the concatenated mitochondrial dataset resulted in different
197 phylogenetic positions for the poensis group (see below). The ML tree suggests that the
198 poensis group represents a separate lineage within Arvicanthini, and does not belong to
199 *Grammomys*. As the basal divergences within this tribe were poorly supported (not
200 shown), we attempted to increase their degree of support by adding more mitochondrial
201 and nuclear sequences. The enhanced dataset contained four mitochondrial (*CYTB*,

202 *COI+COII+ATPase8*, *16S*, *12S*) and five nuclear markers (*IRBP*, *RAG1*, *GHR*, *BRCA1*,
203 *AP5*). In total, this multi-locus dataset included 34 species of Arvicanthini (sensu
204 Lecompte *et al*., 2008) comprising 14 genera. The genus *Grammomys* was represented by
205 sequences of representatives of the five groups that were identified by the mitochondrial
206 phylogeny. As outgroups, we used representatives of six other tribes of Murinae (Table
207 S3 in Appendix S1). The total length of the concatenated dataset was 9458 bp with 46%
208 missing data. We performed analyses in RAxML and MrBayes using the partitioned
209 datasets (Table S2 in Appendix S2) as described above.
210
211 The same dataset was used to estimate the times to most recent common ancestors
212 (TMRCAs) of the clades that were identified by earlier analyses. We used a relaxed clock
213 model with branch rates drawn from an uncorrelated lognormal distribution in Beast
214 1.8.2 (Drummond et al. 2012). Calibration of the molecular clock was based on four
215 fossil taxa. Three represent the oldest records of three Arvicanthine genera
216 (*Lemniscomys*, *Arvicanthis*, *Aethomys*) from the Lemudong´o locality 1, Kenya (Manthi,
217 2007; 6.12-6.08 Ma), for which we used exponential priors with mean = 1.0 and offset =
218 6.1 for TMRCA of these genera. The fourth calibration point was represented by the
219 *Mus*/*Arvicanthis* split (Kimura *et al*., 2015; 11.1 Ma), for which we set an exponential
220 prior with mean 1.0 and offset 11.1. For more details see Table S4 in Appendix S2. For
221 divergence dating analysis we used the partitioned multi-locus dataset (Table S2 in
222 Appendix S2) with priors set to the Yule speciation process, and we constrained the tree
223 topology based on the results of the previous ML analysis. We used a linked partition
224 tree, and unlinked clock and site models. The MCMC simulations were run twice with 20

million iterations, with genealogies and model parameters sampled every 1000 iterations. The outputs from BEAST were analysed as described above, following the removal of 25% trees as burn-in. All phylogenetic analyses were run on CIPRES Science Gateway (Miller *et al.*, 2010).

***Species tree and dating of divergences within* Grammomys**

We used the concatenated mitochondrial sequences (*CYTB* + *16S*) and unphased nuclear IRBP genes of the genus *Grammomys* to obtain a dated species tree under the fully Bayesian framework implemented in the *BEAST package (Heled & Drummond, 2010), an extension of BEAST 1.8.2 (Drummond *et al.*, 2012). Alignments for mitochondrial and nuclear genes were given separate and unlinked substitution, clock and tree models (the latter was linked for two mitochondrial markers). The monophyly of the five main lineages was constrained and the tree was calibrated (relaxed log-normal clock, secondary calibration) using the TMRCAs of the main *Grammomys* lineages estimated from the primary divergence date analysis of Arvicanthini (Table S4 in Appendix S2). Two independent runs were carried out for 20 million generations with sampling every 2000 generations in BEAST. The resulting parameter and tree files from the two runs were examined for convergence in TRACER 1.5 and combined in LOGCOMBINER 1.8.2 (Drummond *et al*., 2012) after removing 10% burn-in. A maximum clade credibility tree was calculated in TREEANNOTATOR 1.8.2 (Drummond *et al*., 2012).

Biogeographical analysis

The dispersal-extinction-cladogenesis model of LAGRANGE (DEC model; Ree & Smith, 2008) estimates geographic range evolution using a phylogenetic tree with branch lengths scaled to time, geographic (habitat) areas for all tips, and an adjacent matrix of plausibly connected areas. We used the optimization on multiple trees (i.e. Bayes-Lagrange or S-DEC model) implemented in the in RASP 3.1 software (Yu *et al*., 2015) to take into account topological uncertainty. RASP computes the likelihood values of all possible ancestral distributions in LAGRANGE and, relying on a composite Akaike weight, it summarizes the biogeographic reconstructions across trees.

Using the distribution data for particular lineages (Fig. 3), we assigned the distribution of tips on the species tree to six main forest types defined by Fayolle *et al*. (2014; see Fig. 1B). In S-DEC analysis, the maximum number of current and ancestral ranges was set at two (as currently no lineage occurs in more than two main forest types) and all six areas were allowed to be mutually connected in the past. For background phylogenetic information we used 18000 trees from the species tree analysis in *BEAST. The probability of ancestral areas was plotted in the form of pie-charts along the species tree.

RESULTS

Phylogenetic analysis of the mitochondrial dataset and distribution of genetic variability

The topology of mitochondrial *Grammomys* trees was similar in ML and BI analyses, except for the position of the poensis group (see below). Based on the topology and statistical support for the branches of the inferred tree we defined five main genetic

groups within the genus (Fig. 2; for the tree with tip labels and outgroups see Appendix S3). These groups have largely parapatric distribution ranges with up to three groups partially overlapping in north-eastern Tanzania and south-eastern Kenya (Fig. 1). The group names are based on the ongoing taxonomic revision of the genus (J. Bryja *et al*., unpublished data).

(1) The **poensis group** includes specimens from Guineo-Congolese forests on the north bank of the Congo River, including montane forests of the Cameroon volcanic line (Fig. 1). In BI analysis the poensis group formed a sister clade to the remaining *Grammomys* taxa (Fig. 2), but in ML topology it formed a deeply divergent lineage with unresolved relationships to other genera of Arvicanthini. The group can be subdivided into four lineages (p1-p4; Fig. 2) with parapatric distributions. The most distinct populations (= p1) are found in Gabon, isolated by the river Ogooué (Fig. 3A). The lineage p2 may correspond to *G. kuru* (Thomas & Wroughton, 1907), described from north-eastern Democratic Republic of the Congo (DRC). *Grammomys poensis* was described from Bioko Island and corresponds to lineage p4 (Eisentraut, 1965).

(2) The **selousi group** is named after a recently described species, *G. selousi* Denys *et al.*, 2011, from south-eastern Tanzania, for which *CYTB* sequence of type material was included in the analysis. The group is subdivided into five lineages with allopatric or parapatric distribution ranges within a narrow belt along the East African coast (se1-se5; Figs 2 & 3A) and appears to prefer lowland forests, e.g. coastal forests inhabited by se4 and se5 (but the latter also occurs in the Usambara Mts and hills of south-eastern Kenya;

293 Fig. 3A). The only lineage within this group that is restricted to highlands is se1 in the
294 Southern Rift Mountains (SRM) of southern Tanzania and northern Malawi. The South
295 African lineage se3 may represent *G. cometes* (Thomas & Wroughton, 1908).
296
297 3) The **dolichurus group** occurs south of the Zambezi (Fig. 3B). Our sample size was
298 too small for detailed analysis of internal genetic structure, but the three lineages seem to
299 correspond to populations distributed along a north-south trajectory (not shown).
300
301 4) The **macmillani group** is composed of eight highly divergent genetic lineages (m1-
302 m8; Figs 2 & 3A). Based on mostly non-overlapping distributions, three lineages can be
303 assigned to earlier species descriptions, although comparisons with type material are
304 required to confirm our current taxonomic interpretation. The m4 lineage is probably *G.*
305 *macmillani* (Wroughton, 1907) described from Wouida, north of Lake Turkana in
306 Ethiopia); m1 corresponds to *G. dryas* (Thomas, 1907) described from the Ruwenzori
307 Mts in Uganda, and m3 to *G. buntingi* (Thomas, 1911), which is the only *Grammomys*
308 species occurring west of the Dahomey gap. Furthermore, m5 may represent *G. gazellae*
309 (Thomas, 1910), a taxon described from South Sudan and synonymised with *G.*
310 *macmillani* (Hutterer & Dieterlen 1984).
311
312 5) The **surdaster group** is named after *G. surdaster* (Thomas & Wroughton, 1908), a
313 synonym of *G. dolichurus* (Musser & Carleton, 2005). However, if the dolichurus group
314 is an exclusively southern African clade (see above), we recommend applying the name
315 surdaster to populations north of the Zambezi as has been suggested by Musser &

316 Carleton (2005). The surdaster group is sister to the macmillani group in all
317 mitochondrial trees. Both groups have largely parapatric distribution ranges with a
318 relatively narrow overlap in northern Tanzania and in the Albertine Rift. The surdaster
319 group is widespread in the eastern African highlands between the equator and the
320 Zambezi River (except for a single locality in central Mozambique; Fig. 1), and may also
321 occur in Angola and southern DRC as suggested by su5 from the Kikwit area in south-
322 western DRC (see also the distribution map in Monadjem *et al.* 2015 under the name *G.*
323 *dolichurus*). The group can be divided into 10 well supported mitochondrial lineages with
324 mostly parapatric distribution ranges (su1-su10; Figs 2 & 3B). The relations among them
325 are unresolved, although in most topologies su1 is sister to all the other lineages and su5-
326 su7 and su8-su10 are monophyletic clades.
327
328 ***Genetic distances***
329 Genetic distances for *CYTB* within and among mitochondrial lineages of *Grammomys* are
330 summarized in Table S3 in Appendix S2. Uncorrected *p*-distances (and similarly K2P-
331 corrected distances) among lineages belonging to different groups were high and ranged
332 from 8.4% (m5 × su2) to 18.7% (p2 × se5). The genetic distances among lineages within
333 each group ranged between 6 and 12% (Table 1), except for the surdaster group, in which
334 11 of 45 lineage pairs differed by less than 5% (Appendix S2).
335
336 ***Analysis of nuclear IRBP gene***
337 The phylogenetic analysis of phased *IRBP* sequences provided a less resolved tree (Fig.
338 S1 in Appendix S2). Of five major mitochondrial clades, only two (poensis and selousi)

were reliably recovered by *IRBP*. The poensis group formed a clade with the genus *Thallomys* exclusive of the other *Grammomys* clades. In the selousi group, only se1 and se3 were significantly supported. In the macmillani group, the geographically adjacent m1 and m2 clades from the Albertine Rift Mts differed substantially in *IRBP* sequences, while m3 from western Africa was significantly supported as the sister taxon of m5 from Central Africa. There was no obvious structure in the surdaster group, and specimens assigned to different mitochondrial lineages often had very similar or identical *IRBP* sequences (Fig. S1 in Appendix S2).

***Monophyly and phylogenetic position of* Grammomys**

The multi-locus ML and BI phylogenies yielded very similar topologies that validated the Arvicanthini tribe (Fig. S2 in Appendix S2). All *Grammomys* representatives clustered in a monophyletic clade, but with low support for the placement of the poensis group. Sister groups that diverged successively were *Thallomys* and *Aethomys*, though the nodes were weakly supported. Surprisingly, *Grammomys* was reconstructed as distantly related to *Thamnomys*, a genus that historically has been thought to be closely affiliated to it (Musser & Carleton, 2005). *Thamnomys* diverged at the beginning of the Arvicanthini radiation, and appears to be the sister genus of *Oenomys*. The remaining arvicanthine genera formed three well supported clades: (1) *Hybomys* + *Stochomys*, (2) *Desmomys* + *Rhabdomys*, and (3) *Arvicanthis* + *Pelomys* + *Lemniscomys*; and two lineages with long and unresolved branches (*Dasymys* and *Micaelamys*).

***Divergence dating within Arvicanthini and species tree of* Grammomys**

The time of divergence between *Grammomys* and its sister genus *Thallomys* was estimated as Late Miocene (median TMRCA= 8.83 Ma; Fig. S2 in Appendix S2). Soon after their split, the poensis group diverged from the rest of the genus (TMRCA of *Grammomys* = 8.21 Ma). The selousi group then separated (6.58 Ma) from the three remaining groups, which diverged from each other in the Pliocene. Based on secondary calibration of the species tree, TMRCAs of lineages within the five main *Grammomys* groups are mostly Pleistocene in age, i.e. < 2.5 Ma (Fig. 4).

Biogeographical analysis

The most probable scenario of the S-DEC model proposed the continuous distribution of ancestral *Grammomys* in the Late Miocene forests that covered eastern and Central Africa, followed by a vicariance event that separated the Central (the poensis group) and East African groups (Fig. 4). The poensis group subsequently diverged by vicariance to p1 (Wet Central Africa) and remaining lineages (Moist Central Africa), from where the lineage p4 dispersed into West Africa (Nigeria). In East Africa, the ancestors of the selousi group dispersed to coastal forests in the Late Miocene, but lineage se1 remained in the uplands and split by vicariance from the rest of the group. The ancestral areas of both the macmillani and surdaster groups are clearly situated in the East African mountain forests. From there, a single dispersal event to wet-moist West African forests followed by diversification occurred in the m3 lineage (Fig. 4).

DISCUSSION

Deep divergence in Grammomys *and the fragmentation of Miocene forests*

385 The multi-locus phylogeny of Arvicanthini supports the monophyly of *Grammomys*. The
386 > 8 Ma divergence between the poensis group and the remaining lineages makes it one of
387 the oldest intrageneric divergences among African murids (assuming that the poensis
388 group remains in the genus *Grammomys*, which could be re-evaluated using the data
389 presented here). This finding thus fits the model of fragmentation of the African Miocene
390 forest into the current Guineo-Congolese forests and coastal and mountain forests in East
391 Africa at this time (Lovett, 1993; Plana, 2004). The formation of the Rift Valley and the
392 decline in global temperatures during the Late Miocene resulted in greater rainfall
393 seasonality, and the spread of grassy vegetation and fragmentation of forests situated east
394 of the rift (Bobe, 2006). An increasing number of studies have shown that the genetic
395 diversification between animal and plant taxa occurring in both the central and eastern
396 African forests started during the Late Miocene. For example, the splits between
397 Congolese and eastern African species of the plant genera *Uvariodendron* and *Monodora*
398 are dated to ca 8.4 Ma (Couvreur *et al*., 2008). Similarly, the contraction and
399 fragmentation of the Pan-African forest at this time played a key role in the
400 diversification of some groups of African chameleons (Tolley *et al*., 2013). Additionally,
401 two rodent lineages, endemic to montane forests of East Africa (the denniae group of
402 *Hylomyscus* and *Praomys delectorum*), split from their sister lineages living mostly in
403 Guineo-Congolese forests at the beginning of the Praomyini radiation dated to the end of
404 the Miocene (Demos *et al*., 2014; Lecompte *et al*., 2005; Missoup *et al*., 2012).
405
406 ***Palaeoendemism in coastal forests of East Africa***

The coastal forests of East Africa were recognised as a distinct phytogeographical unit by White (1983) and, more recently, by Fayolle *et al*. (2014). They exhibit a patchy distribution extending from southern Somalia to the Limpopo River in southern Mozambique and represent endangered centres of biodiversity. There is evidence that most of the coastal forest endemics, including mammals, are palaeoendemics (Burgess *et al*., 1998). Phylogenetic reconstruction of *Grammomys* revealed the split of the selousi group from other East African *Grammomys* ca 6.5 Ma (Fig. 4), indicating a Late Miocene separation of coastal and highland forests in eastern Africa (Fig. 6). This is concordant with the divergence time (ca 6.5 Ma) proposed by Mikula *et al*. (2016) between the genus *Beamys* (a rodent typical of African coastal forests), and its sister genus *Cricetomys* (widespread in various African forests). The *Grammomys* lineage se3 from east coastal South Africa suggests a historical connection between coastal forests in East Africa and those further south, which has not been reported before. Species inhabiting these coastal forests are able to reach higher altitude forests (possibly via riverine gallery forests) as suggested by the presence of se2 in the Mulanje Mts, se5 in the Usambara Mts and the observation that *Beamys* occurs in coastal forests as well as in the Southern Rift Mountains (SRM) (Happold, 2013). The clear north-south structuring within the selousi group reflects the fragmented nature of coastal forests; this separation may be maintained by large rivers (e.g. Rufiji, Zambezi, Limpopo) as observed for other lowland species (Bartáková *et al*., 2015; McDonough *et al.*, 2015). Alternative hypotheses of divergence within coastal forests include climatic changes in the Plio-Pleistocene or increases in sea level, shrinking suitable habitats into isolated fragments situated at higher elevations (Burgess *et al*., 1998).

Evolution of the eastern Afromontane biodiversity hotspot during Plio-Pleistocene climatic oscillations

A reversal of the cooling trend occurred in the Early Pliocene. This represented the warmest period over the last 5 Myr, leading to the suggestion that East African forests may have expanded at this time, especially at higher elevations (Feakins & deMenocal, 2010). More continuous forest cover probably facilitated the dispersion of the dolichurus group in south-eastern Africa during that period. However, after 3.5 Ma temperatures decreased and the Plio-Pleistocene aridification events linked with significant expansion of grass-dominated ecosystems in East Africa generated more diverse mosaic environments (Bobe, 2006). Within the genus *Grammomys*, these environmental changes are reflected by intensive radiations that occurred in the eastern Afromontane hotspot, especially in the Eastern Arc Mountains and Southern Rift Mountains (EAM + SRM; the surdaster group) and the Kenyan Highlands and Albertine Rift Mountains (KH+ARM; the macmillani group) (Fig. 5). The overlap in the distribution ranges of mammal species occurring in the main blocks of the Afromontane region (i.e. EAM+SRM versus KH+ARM) is generally very low (e.g. Carleton *et al*., 2015), suggesting that the faunas of the EAM+SRM and the KH+ARM pursued long-term independent evolutionary trajectories. The distribution ranges for the macmillani and surdaster groups reported in this study appear to agree with this scenario (Fig. 1).

Demos et al. (2014) provided evidence of repeated Pleistocene connections between small mammal taxa inhabiting forests of the Albertine Rift Mts and the Kenyan

Highlands. This explains the sister-group relationship between two lineages restricted to high elevations of the Albertine Rift Mts (i.e. palaeoendemics m1 + m2) and the rest of the macmillani group, the geographic origin of which is presumed to be in the Kenyan highlands. It can be argued that during one of the humid Pleistocene periods, lineage m4 from the Kenyan highlands colonized the southern Kenyan and northern Tanzanian mountains (e.g. the volcanoes in the Rift Valley inhabited by m7 and m8). Subsequently, the lineage leading to m5 appears to have descended from high, humid montane forest to drier, forested savanna habitats. We hypothesize that an increased ability to colonize drier habitats may have allowed *Grammomys* to colonize relatively large areas at the interface between the Guineo-Congolese forests and the Sudanian savanna, and consequently, the Guinean forests-savanna mosaic of West Africa (m3; see below).

The diversification events within the surdaster group may also be linked to Pleistocene climatic changes. There is increasing evidence that, during humid periods within the last 2 Myr, the currently fragmented mountain forests of the EAM and SRM were repeatedly united, allowing the periodic exchange of forest-dependent faunas. However, it is unlikely that a single spatio-temporal scenario applies for all faunal components, as even species with presumably similar ecological requirements may have different responses to the same environmental changes (Carleton & Stanley, 2012). For example, phylogenetic reconstructions of the forest-dependent rodent *Praomys delectorum* revealed two distinct lineages corresponding to the Usambara Mts in the north and Nguru Mts in the south, which are separated by the wide savanna belt in north-eastern Tanzania (Bryja *et al*., 2014b). However both sides of this belt are inhabited by a single mitochondrial

476 *Grammomys* lineage (su10; Fig. 3). Such conflicting patterns may be due to a lower
477 dependency of *Grammomys* on the prevailing ecological conditions in humid montane
478 forests. This would have allowed them to colonize both miombo woodlands (lineage su4)
479 and savanna-forest mosaics on the south-eastern edge of the Congolese forests (su5-su7).
480 Such distribution patterns have not been observed in previously studied forest specialists
481 restricted to the EAM and SRM (e.g. Bryja *et al*., 2014b; Lawson, 2010; Loader *et al*.,
482 2014; Tolley *et al*., 2011).
483
484 ***Long-distance dispersal along the northern edge of the Congo Basin***
485 In order to explain similarities between eastern and western African montane forests and
486 grasslands, many authors have assumed that, during climatic changes and especially
487 during colder periods, the mountain floras and faunas must have extended to the
488 lowlands, which facilitated dispersal between mountain massifs (White, 1981). The zones
489 characterized by the mosaic of forest and savanna north of the Congo basin are among
490 the least known areas of Africa. However, our results concerning the distribution of
491 *Grammomys* m5 suggest that there is a clear biogeographical connection between Uganda
492 (+ westernmost Kenya) and Central Africa (north-eastern DRC, CAR, South Sudan).
493 This link is not only indicated by this study, but also by earlier studies which revealed
494 that identical genetic lineages of other rodents occur in this forest/savanna mosaic, e.g.
495 *Mus* cf. *bufo* (Bryja *et al*., 2014a), or *Aethomys hindei* (Monadjem *et al*., 2015). The
496 biogeographic scenario suggests that, during humid phases, the Pleistocene lowland
497 forests of the Congo Basin extended further north than they do today. This situation may
498 have allowed the ancestors of *Grammomys* m3+m5 from eastern Africa to disperse along

the northern margin of the Congolese forest and colonize north-eastern DRC, CAR and South Sudan (Fig. 5). It seems plausible that, after the northern edge of the lowland forests in the Congo Basin receded, some populations persisted in the resulting relict forests in forest-savanna mosaics (i.e. *G.* m5 in CAR), montane areas (probably *G. aridulus* in Jebel Marra region in Sudan; Fig. 1) or adapted to new environments, where *Grammomys* mice were previously absent (*G. buntingi* = m3 in West Africa).

CONCLUSION

This is the first phylogenetic study of *Grammomys* rodents that includes samples from most of its distribution area in sub-Saharan Africa. Our results suggest that the genus is monophyletic and unrelated to *Thamnomys*, and that its intrageneric divergences are among the oldest in African murids (> 8 Ma). The majority of the five detected clades have parapatric distribution ranges, and the times of divergence estimated among these clades agree with accepted scenarios for the evolutionary history of the African forests since the Late Miocene. The distribution of these lineages does not agree with the current taxonomy. Our results suggest that a revision of this genus will lead to discoveries of new species, especially in highland and coastal forests in East Africa. ~~Finally, since the discovery of four *Plasmodium* parasites in *Grammomys* from the Democratic Republic of Congo (Vincke & Lips, 1948), no new rodent *Plasmodium* isolates have been obtained (Keeling & Rayner, 2015). We suggest that the taxonomic diversity reported for thicket rats might imply a significant underestimation of *Plasmodium* diversity. New surveys may lead to a better understanding of the origin and evolutionary history of these malaria causing blood parasites in rodents and other mammals.~~

ACKNOWLEDGEMENTS

This study was supported by the Czech Science Foundation, project no. 14-36098G. The French ANR – IFORA project and the INCO-DEV Project – TREATCONTROL allowed work in West-Central Africa. Fieldwork in DR Congo was supported by the Belgian Directorate for Development Cooperation (DGD), and the Flemish Inter-University Council – University Development Cooperation (VLIR-UOS). For help during the field work we acknowledge V. Mazoch, H. Konvičková, J. Šklíba, M. Lövy, G. Mhamphi, F. Sedláček, S. Šafarčíková, A. Konečný, S. Gambalemoke Mbalitini, B. Kadjo, F. Kourouma, M. Sylla, D. Mory, E. Kemming, A.D. Missoup, R. Cornette, S. Moulin, E. Lecompte, A. Lalis and all other local field collaborators. The assistance of R. Makundi, A. Massawe, C. Sabuni, J. Mbau, W.N. Chitaukali, M. Colyn, B. Dudu Akaibe, L. Koivogui, C. Camara, and the late W. Verheyen with project logistics and collection of samples is highly appreciated. H. Konvičková and L. Piálek helped with genotyping, F. Jacquet, C. Sabuni and J. Goüy de Bellocq provided unpublished sequences. For permission to carry out the research and to collect specimens we are obliged to the National Research Council and Forestry Department in Malawi, the Uganda National Council for Science and Technology, the Kenyan Forest Service and the Kenyan Wildlife Service, COSTECH Tanzania, the Guinean and Cameroonian Ministries of Water and Forest Management, the Zambian Wildlife Authority, Sokoine University of Agriculture in Morogoro (Tanzania), and the 'Centre de surveillance de la biodiversité' in Kisangani (DR Congo). We also thank the SYNTHESYS programme (BE-TAF-5113 to OM and FR-TAF-5799 to JB) and museum curators that allowed us to study collections in their

545 care: W. Wendelen (RMCA), R. Hutterer (ZFMK), J. Phelps (FMNH), C. Conroy
546 (MVZ), D. Lunde (USNM), and V. Volpato (SMF). B. van Vuuren, J. Masters, P. Linder
547 and two anonymous reviewers provided very useful comments on previous version of the
548 manuscript.

549

550

REFERENCES

Bartáková, V., Reichard, M., Blažek, R., Polačik, M. & Bryja, J. (2015) Terrestrial fishes: rivers are barriers to gene flow in annual fishes from the African savanna. *Journal of Biogeography*, **42**, 1832-1844.

Bobe, R. (2006) The evolution of arid ecosystems in eastern Africa. *Journal of Arid Environments*, **66**, 564-584.

Bryja, J., Mikula, O., Šumbera, R., Meheretu, Y., Aghová, T., Lavrenchenko, L.A., Mazoch, V., Oguge, N., Mbau, J.S., Welegerima, K., Amundala, N., Colyn, M., Leirs, H. & Verheyen, E. (2014a) Pan-African phylogeny of *Mus* (subgenus *Nannomys*) reveals one of the most successful mammal radiations in Africa. *BMC Evolutionary Biology*, **14**, e256.

Bryja, J., Mikula, O., Patzenhauerová, H., Oguge, N., Šumbera, R. & Verheyen, E. (2014b) The role of dispersal and vicariance in the Pleistocene history of an East African mountain rodent, *Praomys delectorum*. *Journal of Biogeography*, **41**, 196-208.

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.

Carleton, M.D. & Stanley, W.T. (2012) Species limits within the *Praomys delectorum* group (Rodentia: Muridae: Murinae) of East Africa: a morphometric reassessment and biogeographical implications. *Zoological Journal of the Linnean Society*, **165**, 420–469.

573 Carleton, M.D., Banasiak, R.A. & Stanley, W.T. (2015) A new species of the rodent
574 genus *Hylomyscus* from Angola, with a distributional summary of the *H. anselli*
575 species group (Muridae: Murinae: Praomyini). *Zootaxa*, **4040**, 101-128.
576 Cerling, T.E. (1992) Development of grasslands and savannas in East Africa during the
577 Neogene. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **97**, 241-247.
578 Couvreur, T.L.P., Chatrou, L.W., Sosef, M.S.M. & Richardson, J.E. (2008) Molecular
579 phylogenetics reveal multiple tertiary vicariance origins of the African rain forest
580 trees. *BMC Biology*, **6**, e54.
581 Demos, T.C., Kerbis Peterhans, J.C., Agwanda, B. & Hickerson, M.J. (2014) Uncovering
582 cryptic diversity and refugial persistence among small mammal lineages across the
583 Eastern Afromontane biodiversity hotspot. *Molecular Phylogenetics and Evolution*,
584 **71**, 41-54.
585 Denys, C., Lalis, A., Lecompte, E., Cornette, R., Moulin, S., Makundi, R., Machang'u,
586 R., Volobouev, V. & Aniskine V. (2011) Notes on a faunal survey in Kingu Pira
587 (South Tanzania) with new karyotypes of several small mammals and description of a
588 new Murid species (Mammalia, Rodentia). *Zoosystema*, **33**, 5-47.
589 Ducroz, J.F., Volobouev, V. & Granjon, L. (2001) An assessment of the systematics of
590 arvicanthine rodents using mitochondrial DNA sequences: Evolutionary and
591 Biogeographical implications. *Journal of Mammalian Evolution*, **8**, 173-206.
592 Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics
593 with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969-1973.
594 Fayolle, A., Swaine, M.D., Bastin, J.-F., Bourland, N., Comiskey, J.A., Dauby, G.,
595 Doucet, J.-L., Gillet, J.-F., Gourlet-Fleury, S., Hardy, O.J., Kirunda, B., Kouamé, F.N.

& Plumptre, A.J. (2014) Patterns of tree species composition across tropical African forests. *Journal of Biogeography*, **41**, 2320-2331.

Feakins, S.J. & deMenocal, P.B. (2010) Global and African regional climate during the Cenozoic. *Cenozoic Mammals of Africa* (ed. by L. Werdelin), pp. 45-55. University of California Press, Berkeley and London.

Galan, M., Pagès, M. & Cosson, J.F. (2012) Next-generation sequencing for rodent barcoding: species identification from fresh, degraded and environmental samples. *PLoS ONE*, **7**, e48374.

Happold, D.C.D. (2013) *Mammals of Africa, Volume III Rodents, hares and rabbits.* Bloomsbury Publishing, London and New York.

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

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. & Dieterlen, F. (1984) Zwei neue Arten der Gattung *Grammomys* aus Äthiopien und Kenia (Mammalia: Muridae). *Stuttgarter Beiträge zur Naturkunde, Serie A (Biologie)*, **374**, 1-18.

~~Keeling, P.J. & Rayner, J.C. (2015) The origins of malaria: there are more things in heaven and earth … *Parasitology*, **142**, S16-S25.~~

Kimura, Y., Hawkins, M.T.R., McDonough, M.M., Jacobs, L.L. & Flynn, L.J. (2015) Corrected placement of *Mus-Rattus* fossil calibration forces precision in the molecular tree of rodents. *Scientific Reports*, **5**, e14444.

De Klerk, H.M., Crowe, T.M., Fjeldsa, J. & Burgess, N.D. (2002) Biogeographical patterns of endemic terrestrial Afrotropical birds. *Diversity and Distributions*, **8**, 147-162.

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.

Lawson, L.P. (2010) The discordance of diversification: evolution in the tropical-montane frogs of the Eastern Arc Mountains of Tanzania. *Molecular Ecology*, **19**, 4046–4060.

Lecompte, E., Denys, C. & Granjon, L. (2005) Confrontation of morphological and molecular data: The *Praomys* group (Rodentia, Murinae) as a case of adaptive convergences and morphological stasis. *Molecular Phylogenetics and Evolution*, **37**, 899–919.

Lecompte, E., Aplin, K., Denys, C., Catzeflis, F., Chades, M. & Chevret, P. (2008) Phylogeny and biogeography of African Murinae based on mitochondrial and nuclear gene sequences, with a new tribal classification of the subfamily. *BMC Evolutionary Biology*, **8**, 199.

Librado, P. & Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Loader, S.P., Ceccarelli, F.S., Menegon, M., Howell, K.M., Kassahun, R., Mengistu, A.A., Saber, S.A., Gebresenbet, F., de Sá, R., Davenport, T.R.B., Larson, J.G., Müller, H., Wilkinson, M. & Gower, D.J. (2014) Persistence and stability of Eastern

Afromontane forests: evidence from brevicipitid frogs. *Journal of Biogeography*, **41**, 1781-1792.

Lovett, J.C. (1993) Eastern Arc moist forest flora. *Biogeography and Ecology of the Rain Forests of Eastern Africa* (ed. by J.C. Lovett and S.K. Wasser), pp. 33-55. Cambridge University Press, Cambridge.

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.

Manthi, F.K. (2007) A preliminary review of the rodent fauna from Lemudong'o, southwestern Kenya, and its implication to the late Miocene paleoenvironments. *Kirtlandia*, **56**, 92-105.

McDonough, M.M., Šumbera, R., Mazoch, V., Ferguson, A.W., Phillips, C.D. & Bryja, J. (2015) Multilocus phylogeography of a widespread savanna-woodland-adapted rodent reveals the influence of Pleistocene geomorphology and climate change in Africa's Zambezi region. *Molecular Ecology*, **24**, 5248-5266.

Mikula, O., Šumbera, R., Aghová, T., Mbau, J., Katakweba, A.A.S., Sabuni, C.A. & Bryja, J. (2016) Evolutionary history and species diversity of African Pouched Mice of the genus Saccostomus (Rodentia: Nesomyidae). *Zoologica Scripta*, in press.

Miller, M.A., Pfeiffer, W. & Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*, 14 Nov. 2010, New Orleans, LA, 1-8.

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.

Missoup, A.D., Nicolas, V., Eiseb, S., Chung, E.K. & Denys, C. (2016) Phylogenetic position of the endemic Mount Oku rat, *Lamottemys okuensis* (Rodentia: Muridae), based on molecular and morphological data. *Zoological Journal of the Linnean Society*, in press.

Monadjem, A., Taylor, P.J., Denys, C. & Cotterill, F.P.D. (2015) *Rodents of sub-Saharan Africa. A biogeographic and taxonomic synthesis*. Walter de Gruyter GmbH, Berlin/Munich/Boston.

Musser, G.G. & Carleton, M.D. (2005) Superfamily Muroidea. *Mammal Species of the World: A Taxonomic and Geographic Reference.* (ed. by D.E. Wilson and D.M. Reeder), pp. 894-1531. Johns Hopkins University Press, Baltimore.

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

Plumptre, A.J., Davenport, T.R.B., Behangana, M., Kityo, R., Eilu, G., Ssegawa, P., Ewango, C., Meirte, D., Kahindo, C., Herremans, M., Kerbis Petterhans, J., Pilgrim, J.D., Wilson, M., Languy, M. & Moyer, D. (2007) The biodiversity of the Albertine Rift. *Biological Conservation*, **134**, 178-194.

Rambaut, A. & Drummond, A.J. (2007) Tracer v1.5, Available from http://beast.bio.ed.ac.uk/Tracer

685 Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range
686 evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57**, 4-
687 14.
688 Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference
689 under mixed models. *Bioinformatics*, **19**, 1572-1574.
690 Rovero, F., Menegon, M., Fjeldså, J., Collett, L., Doggart, N., Leonard, C., Norton, G.,
691 Owen, N., Perkin, A., Spitale, D., Ahrends, A., Burgess, N. D. (2014) Targeted
692 vertebrate surveys enhance the faunal importance and improve explanatory models
693 within the Eastern Arc Mountains of Kenya and Tanzania. *Diversity and Distributions*,
694 **20**, 1438–1449.
695 Scheet, P. & Stephens, M. (2006) A fast and flexible statistical model for large-scale
696 population genotype data: applications to inferring missing genotypes and haplotypic
697 phase. *American Journal of Human Genetics*, **78**, 629-644.
698 Stamatakis, A. (2014) RAxML Version 8: A tool for phylogenetic analysis and
699 postanalysis of large phylogenies. *Bioinformatics*, **30**, 1312-1313.
700 Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013) MEGA6:
701 Molecular Evolutionary Genetics Analysis version 6.0. Mol*ecular Biology and*
702 *Evolution*, **30**, 2725-2729.
703 Tolley, K.A., Tilbury, C.R., Measey, G.J., Menegon, M., Branch, W.R. & Matthee, C.A.
704 (2011) Ancient forest fragmentation or recent radiation? Testing refugial speciation
705 models in chameleons within an African biodiversity hotspot. *Journal of*
706 *Biogeography*, **38**, 1748–1760.

Tolley, K.A., Townsend, T.M. & Vences, M. (2013). Large-scale phylogeny of chameleons suggests African origins and Eocene diversification. *Proceedings of the Royal Society of London, Series B*, **280(1759)**, e20130184.

~~Vincke, J.H. & Lips, M. (1948) Un nouveau *Plasmodium* d'un rongeur suavage du Congo, *Plasmodium berghei* n.sp. *Annales de la Societe Belge de Medecine Tropicale*, **28**, 97-194.~~

White, F. (1981) The history of the Afromontane archipelago and the scientific need for its conservation. *African Journal of Ecology*, **19**, 33-54.

White, F. (1983) *The vegetation of Africa*. UNESCO, Switzerland.

Wiens, J.J. & Donoghue, M.J. (2004) Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution*, **19,** 639-644

Yu, Y., Harris, A.J., Blair, C. & He, X.J. (2015) RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution*, **87**, 46-49.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Collecting localities and genetic data.

Appendix S2 Additions to phylogenetic analyses.

Appendix S3 Detailed Bayesian phylogeny of mtDNA.

DATA ACCESSIBILITY

New sequences used in phylogenetic analyses are available in GenBank under accession numbers KU723898-KU724057 and KU747156- KU747161 (*CYTB*), KU723674-KU723792 (*16S*), KU723651- KU723656 and KU723793-KU723897 (*IRBP*), KU723660- KU723673 (*RAG1*), KU723657- KU723659 (*BRCA1*) (see Appendix S1). Further details of specimens, including museum numbers, are specified in Appendix S1.

BIOSKETCH

Josef Bryja is head of the molecular ecology group at the Institute of Vertebrate Biology ASCR, and has a general interest in factors affecting the evolution of vertebrate populations. His specialities include phylogeography and speciation in Africa, conservation genetics and mechanisms of host-parasite co-evolution.

744 Authors' contributions: JB, RŠ, CD and EV conceived and designed the study, JB, RŠ,
745 JKP, CD, VN, TA and EV collected important part of samples, TA and AB genotyped
746 most samples, JB, OM and TA analysed data, and JB wrote the first draft of the
747 manuscript. All authors contributed to the final version of the paper.
748
749 Editor: Judith Masters

FIGURE LEGENDS

Figure 1 (A) Distribution of sampled *Grammomys* specimens in sub-Saharan Africa. The five main genetic groups of *Grammomys* are represented by different symbols (see key). Black stars show type localities of currently valid species (except *G. surdaster*, which is considered a junior synonym of *G. dolichurus*) mentioned in the text. Main mountain blocks mentioned in the text are schematically demarcated by dashed lines: KH = Kenyan Highlands, ARM = Albertine Rift Mountains, EAM = Eastern Arc Mountains, SRM = Southern Rift Mountains. **(B)** Distribution of main forest types in sub-Saharan Africa. The dots represent localities downloaded from Fayolle *et al.* (2014). They correspond to the six floristic clusters defined by the analysis of 1175 tree species in 455 sampling sites of tropical African forests.

Figure 2 Mitochondrial Bayesian tree of *Grammomys* based on concatenated alignment of 1140 bp of *CYTB* and 575 bp of *16S*. The circles indicate statistical support for nodes, specifically 1000 bootstraps in maximum likelihood analysis (BS)/posterior probability from Bayesian analysis (PP). Only values BS>75 and PP>0.95 are shown. More detailed version of the tree with precise values of statistical support, tip labels and outgroups is shown in Appendix S3.

Figure 3 Geographical distribution of genetic lineages within the five main *Grammomys* groups. Different groups are shown by different symbol shapes and different lineages by different symbol colours. The names of lineages correspond to those in Fig. 2 and putative species names for some are in parentheses (see text for more details). **(A)**

poensis (squares), selousi (circles) and macmillani (stars) groups; **(B)** dolichurus (stars) and surdaster (triangles) groups.

Figure 4 Ultrametric *Grammomys* **s**pecies tree from *BEAST. The pie-charts indicate the most probable ancestral areas of particular clades as estimated by S-DEC model in Bayes-Lagrange (Ree & Smith, 2008).

Figure 5 Schematic illustration of major evolutionary events in *Grammomys*. **(A)** The fragmentation of Late Miocene pan-African forest into the ancestors of current Guineo-Congolese forests (green) and East African montane and coastal forests (purple). **(B)** The split between *Grammomys* inhabiting montane (red) and coastal (yellow) forests in East Africa. **(C)** During the Pliocene the ancestors of the dolichurus (orange), surdaster (red) and macmillani (blue) groups split along a south-north trajectory. The long-term forest refugia for the surdaster and macmillani groups were probably located in the EAM + SRM for the former and in KH + ARM for the latter. **(D)** Pleistocene climatic cycles caused repeated fragmentations and expansions of forest habitats leading to diversification within all five main clades. One of the expansions of the macmillani clade involved the colonization of Guinean forests (m3 lineage) by the "northern route", i.e. north of the Congolese forests. Note that the ellipses at (A) and (B) show only schematically the positions of ancestral populations and do not indicate precise geographical locations.

796 **TABLES**

797 **Table 1** Minimum and maximum genetic distances (K2P-corrected and uncorrected p-
798 distances) among lineages in four main *Grammomys* groups. Genetic variation within the
799 dolichurus group was not analysed because of the low number of available sequences.

800

Groups	Min distance			Max distance		
	K2P-distance	p-distance	Lineages	K2P-distance	p-distance	Lineages
selousi	0.093	0.086	se2 x se3	0.127	0.114	se1 x se5
poensis	0.072	0.067	p3 x p4	0.106	0.097	p1 x p2
macmillani	0.064	0.061	m7 x m8	0.134	0.119	m2 x m6
surdaster	0.037	0.036	su5 x su9	0.108	0.098	su1 x su2

801