University of South Bohemia

Faculty of Science

Phylogeny of Grey-bellied Pygmy Mouse (*Mus triton*) complex

Master thesis

Jarmila Krásová

Supervisor: doc. Radim Šumbera, Ph.D.

Consultant: doc. Josef Bryja, Ph.D.

České Budějovice 2014

Krásová, J., 2014: Phylogeny of Grey-bellied Pygmy Mouse (*Mus triton*) complex. Mgr. Thesis, in English. – 48 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Abstract:

The Grey-bellied Pygmy Mouse (*Mus triton*) has been for a long time considered as a single species, although validity of the single species status was questioned. In order to revise current taxonomy of *M. triton*, I analyzed sequences of one mitochondrial (cytochrome b) and two nuclear genes (IRBP and Intron 7 of the β-fibrinogen) from specimens collected across the most of its known distributional range. Four well-supported phylogroups at species level, differentiated during the Plio-Pleistocene, were evidenced. Divergence dating suggests that the diversification of *"triton"* species complex was likely caused by Plio-Pleistocene climatic oscillations together with highly diverse topography of Eastern Africa.

Prohlašuji, že svoji diplomovou práci jsem vypracovala samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. V platném znění souhlasím se zveřejněním své diplomové práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

V Českých Budějovicích, 12. 11. 2014

Acknowledgements

"The best teachers are those who show you where to look, but don't tell you what to see."

Alexandra K. Trenfor

Foremost, I would like to express my sincere gratitude to my supervisor Radim Šumbera for his patience, motivation, enthusiasm, and immense knowledge. My sincere thanks also go to my consultant, Josef Bryja. Both of them have my great admiration for being very hardworking, excellent scientists. I would also like to extend my thanks to Vláďa Mazoch for providing an introduction to phylogenetic methods and sharing his personal experience from field work in Eastern Africa, both aspects have been extremely beneficial for my work. My sincere thanks also go to Táňa Aghová and Hanka Patzenhauerová-Konvičková for helping me with laboratory part of my work. Last but not least, I would like to thank my husband. He was always there cheering me up and stood by me through the good times and bad.

Content:

Introduction

African Pygmy mice of the subgenus *Nannomys* (formerly *Leggada*) are small-sized murine rodents that form the only native lineage of the genus *Mus* in the African continent. With 18 recognized species (Musser and Carleton, 2005; Britton-Davidian et al., 2012), it is the most species-rich subgenus of the genus *Mus* representing about 50 % of genus richness. Majority of *Nannomys* species are savanna-dwelling rodents (Veyrunes et al., 2005), but some of them have been found in forests, swampy habitats and in agriculture fields (e.g. Fichet-Calvet et al., 2009; Kaleme et al., 2007; Kan Kouassi et al., 2008; Lamb et al., 2014; Mulungu et al., 2008).

History of the *Nannomys* lineage has started in the late Miocene after the ancestor of the subgenus *Nannomys* had crossed the Arabian land-bridge from Asia to Africa (Britton-Davidian et al., 2012) between 6.8 and 7.8 Mya (Chevret et al., 2005; Lecompte et al., 2008; Veyrunes et al., 2005). According to the most recent study of Lamb et al. (2014), the first *Nannomys* radiation took place in the late Miocene around 6.33 Mya. Oldest *Nannomys* fossils were found in Tugen Hills, Kenya (4.5 Mya; Winkler, 2002) and in Omo valley and Hadar, Ethiopia (5-2.5 Mya; Catzeflis and Denys, 1992; Denys, 1999). Subsequent extensive radiations occurred mainly during Pliocene and Pleistocene (Chevret et al., 2005; Lamb et al., 2014; Veyrunes et al., 2005). Authors of these three studies suggested different times of the radiation events. Chevret et al. (2005) dated the first Pliocene radiation to 5.0 Mya and later radiation events to 4 - 3 Mya. According to Veyrunes et al. (2005), the first *Nannomys* radiation took place 4.0 Mya and later radiations 3.2 – 1.4 Mya. Lamb et al. (2014) estimated the first radiation to 4.9 Mya and the later diversifications between 3.5 and 2.5 Mya.

The Pliocene and Pleistocene radiation events in *Nannomys* were probably related to global climatic changes (e.g. Bonnefille, 2010; deMenocal, 2004; Gasse, 2006; Potts, 2013; Trauth et al., 2007; 2009), which are known to be a driving force of speciation in many other small African mammals (e.g. Bryja et al., 2014; Colangelo et al., 2013; Demos et al., 2014; Dobigny et al., 2013; Nicolas et al., 2011). According to Lamb et al. (2014) the whole subgenus has likely diversified in mountain areas, because its basal lineages, *Mus callewaerti* from Mbala Nondolo village, Angola, and *M. triton* from Kahuzi-Biega, Albertine Rift, DRC, are montane species. The most intensive radiation of the *Nannomys* dated to the late Pliocene 3.5 – 2.5 Mya, (Lamb et al., 2014) was probably affected by a strong shift from warm and mesic to cold and arid climate with the peak of aridity near 2.8 Mya (deMenocal, 2004). This shift is well documented in a fossil record of mammalian fauna from the Omo basin (Ethiopia),

where remarkable decrease of the number of woodland and forest species and increase of the number of grassland species occurred between 3.6 and 2.4 Mya (Gasse, 2006). Latter, the oscillations in African climate continued with main aridity peaks near 1.7 and 1.0 Mya (deMenocal, 2004). Periodic increases of aridity caused spreading of drier habitats such as open canopy woodlands and savannahs which is well documented by increases of shrubs, heath and grasses pollen (Bonnefille, 1983). According to Pleistocene forest refuge hypothesis, forest fragmentation during glacial minima led to the isolation and subsequent diversification of forest associated taxa (Haffer, 1982 In: Nicolas et al., 2011). Expanding of open habitats together with very variable climate characterized by changing of long wet and dry periods led to extensive dispersion and adaptation followed by radiation in many plants and animals (Potts, 2013). Nevertheless, the dispersion could be restricted by geographic barriers emerging throughout the tectonic history of East African rift-system continuing in Pliocene and Pleistocene (Chorowicz, 2005).

Radiations of the *Nannomys* described above led to a high cryptic diversity of the subgenus (e.g. Veyrunes et al., 2005). Recently, Lamb et al. (2014) proposed existence of three new *Nannomys* species. If we consider relatively poor sampling in this study in some areas of *Nannomys* distribution, we can assume that the number of species within this group can be even higher. In my study I focus on genetic variability in one *Nannomys* species, the Grey-bellied Pygmy mouse (*Mus triton*) living mainly in mountainous areas of Central and Eastern Africa. Although it has been known for a long time that taxonomy of *M. triton* is complicated (see Chimimba and Bennett, 2005), until now there has been neither phylogenetic nor even morphological taxonomic study carried out on this species. Lamb et al. (2014) is the only published molecular-phylogenetic study, where *M. triton* has been included. Based on a single cytochrome b sequence from a single individual from D. R. Congo, the authors found that *M. triton* has a moderately supported relationship to *M. callewaerti* from Angola. Both these species form a sister lineage to all other *Nannomys.*

The Grey-bellied Pygmy mouse is one of the largest *Nannomys* species with the mean body mass of 12.1 g (range 8-18, N=27; Britton-Davidian et al., 2012). Similarly to all species of the genus, *M. triton* has mainly nocturnal activity. Its diet consists primarily of insects, but plant material is also consumed. In Malawi, reproduction takes place mainly during the wet season in between April and July (Hanney, 1965) and the highest mortality rate occurs during cold dry seasons between July and September (Happold and Happold, 1989).

Area of distribution of this species covers Democratic Republic of the Congo, south Uganda, Tanzania, South Ethiopia, Northern Mozambique, Zambia and Central and Northeast Angola (http://www.mol.org/; Dieterlen and Agwanda, 2008; Olson et al., 2001; Figure 1) Although its southernmost recent occurrence is recorded in Tete District, Mozambique, Avery (1991) discovered this species in upper Pleistocene deposits in South Africa (Border Cave, the Lebombo Mountain), which is about 1000 kilometers from the border of the current distribution.

Mus triton occupies wide range of habitats including grasslands, scrublands, swamps and cultivated areas from 1000 to 3500 m (Dieterlen and Happold, 2013). In Malawi the species prefers either grasslands, such as in Zomba Plateau, or *Phillipia* swamps and bracken, as in Nyika Plateau and Mulanje Mountains (Chitaukali et al., 2001; Hanney 1965). In Mount Kilimanjaro in Tanzania, *M. triton* is the most common in disturbed forests, gardens and young pine plantations (Mulungu et al., 2008). In Southwestern Uganda the species was found only in grassland habitats in altitudes from 1280 to 2350 m. The only known low attitude area (around 500 m) inhabited by *M. triton* is in Kikwit in southwestern DRC (Leirs et al., 1999).

Mus triton was originally described as *Leggada triton* from Mt Elgon, Uganda (Thomas 1909). It was defined as a large species (head and body length 81 mm, tail length 51 mm) with dark pelage, below greyish. Upper sides of feet are dark brown. Thomas (1909) also noted that the skull is markedly larger (23 mm) than that of any other known *Leggada* species by which it can be readily distinguishable from its allies. On the front feet, the third and fourth fingers are the longest, while on the hind feet the first and fifth toes are short, the remaining three much longer and all about the same length (Skinner and Chimimba, 2005). In some morphometric traits there are conspicuously large differences between populations of *Mus triton*. For example, specimens collected in Angola (Hill and Carter, 1941), northern Zimbabwe (Ansell, 1957 in Hanney, 1965) and Malawi (Hanney, 1965) significantly differ in the length of diastema. Furthermore, specimens from Tanzania and the Democratic republic of Congo differ in chromosome number (Robbins and Baker, 1978). These facts indicate that the taxa may represent a species complex, rather than a single species (c.f. Lavrenchenko, 2000).

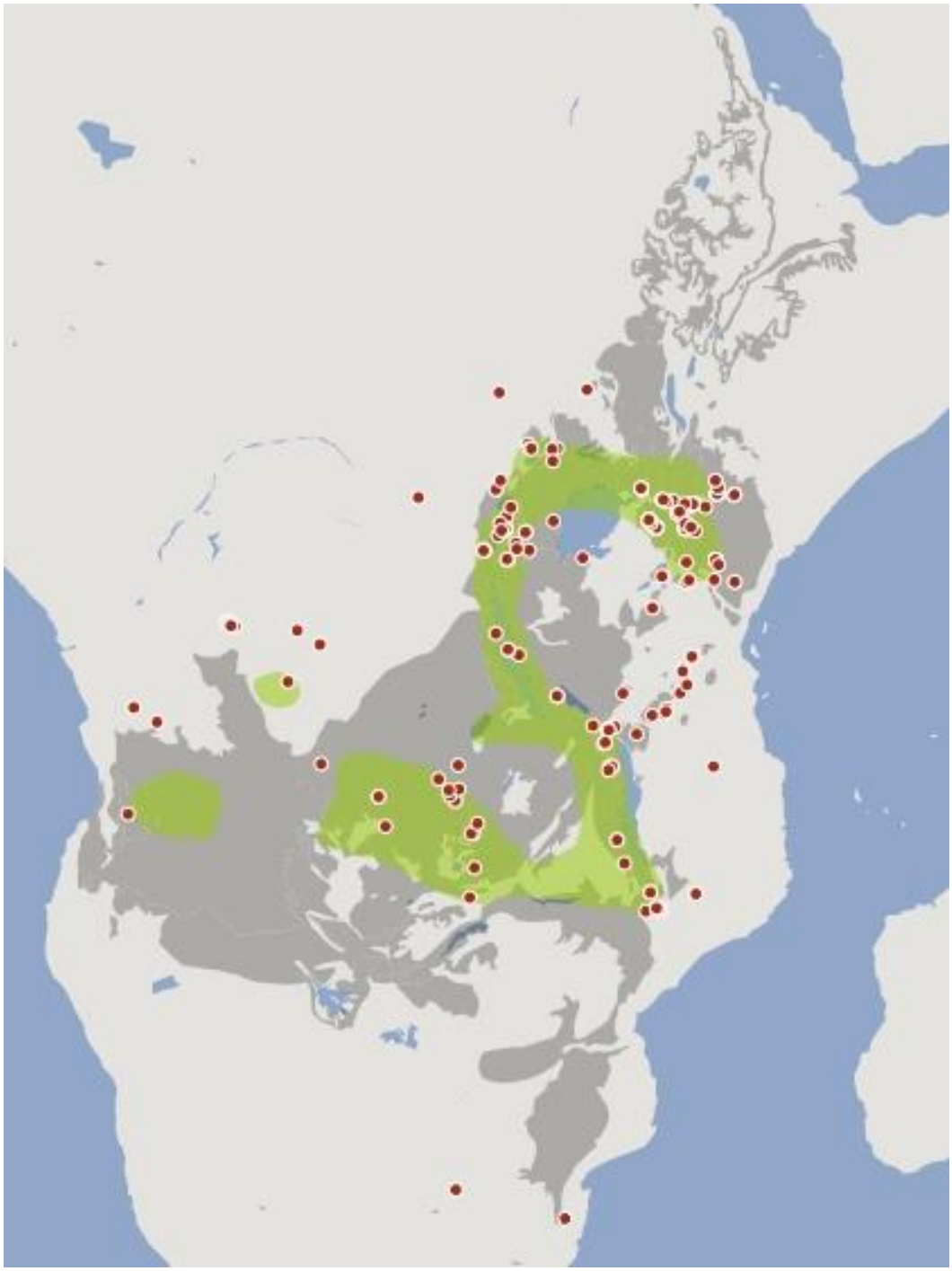


Figure 1. Distribution of *Mus triton* according to Map of life (http://www.mol.org/). Green area – species distribution from IUCN (Dieterlen and Agwanda, 2008). Grey area – regional checklist from WWF (Olson et al., 2001), Red dots represents geotagged points obtained from Global Biodiversity Information Facility (GBIF.org) data portal. Two points in South Africa are positions of the fossil findings (Avery, 1991).

The Grey-bellied Pygmy mouse inhabits mainly montane areas in the Eastern and Central Africa, i.e. areas which are known for a high rate of endemism likely formed by dramatic climatic and tectonic history of the region (see above). Due to its basal position within *Nannomys* subgenus (Lamb et al., 2014) and its relatively large area of distribution it could be intriguing to reconstruct its phylogeography and historical biogeography. In this study I uncover genetic variability of *Mus triton* using a combination of mitochondrial marker (cytochrome b) and two nuclear markers (Intron 7 of the β-fibrinogen and Interphotoreceptor retinoid-binding protein) on individuals collected from most of its distributional range. Main aims of the study are: 1) to revise taxonomy of *M. triton* complex and infer its evolutionary history; 2) to reconstruct its phylogeography and identify historical processes leading to its present distribution.

Methods

Sampling

Sequences of *M. triton* used in this study have been obtained either directly from specimens collected during field expeditions to Eastern Africa (between 2005 – 2013) or downloaded from publicly accessible databases. A total of 101 individuals was captured at 27 different localities (in Ethiopia, Kenya, Tanzania, Zambia and Malawi), using snap traps and Shermann traps. Spleen tissues were preserved in 96% ethanol and stored at -20°C until DNA extraction. In total, 19 sequences from Kenya, Rwanda, DRC, Tanzania, Malawi and Angola (Table 1) were downloaded from GenBank (www.ncbi.nlm.nih.gov/genbank) or African Rodentia databases (http://projects.biodiversity.be/africanrodentia/). Please see Appendix 1 for further reference.

Table 1. Total number of sequences, number of haplotypes, sequence length, number of specimens obtained during field work and sequences downloaded from online databases.

	cyt b	**IRBP**	**β-fibint 7**
Total sequences	119	70	31
Number of haplotypes	45	42	24
Sequence length (bp)	1096	1049	696
Captured individuals	100	70	31
African Rodentia	15	0	0
GenBank	4	0	0

Amplification and sequencing

DNA extraction was performed using the Qiagen method (DNeasy tissue kit). Amplifications of cytochrome b (cyt b), Interphotoreceptor retinoid-binding protein (IRBP) and Intron 7 of the β-fibrinogen (β-fibint 7) were performed by polymerase chain reaction (PCR) using procedures, thermal profiles and primers listed below.

Cytochrome b (cyt b)

Mitochondrial sequences were isolated using L14723 and H15915 primers (Irwin et al., 1991; Table 2). Amplifications were performed by the following procedure: a first cycle of an initial denaturation at 94°C for 2 min, then 35 cycles with denaturation at 94°C for 1 min, annealing at 52°C for 1 min, and extension at 72°C for 1 min. These steps were followed by a 10 min extension at 72°C.

Intron 7 of the β-fibrinogen (β-fibint 7)

β-fibint 7 sequence was obtained by PCR using primers fgb-I7L and fgb-I7U (Wickliffe et al., 2003; Table 2). The first denaturation cycle started at 94°C for 40 seconds, then 35 cycles with denaturation at 94°C for 40 seconds, annealing at 53°C for 45 seconds, and extension at 72°C for 90 seconds (Wickliffe et al., 2003).

Interphotoreceptor retinoid-binding protein (IRBP)

Amplified DNA of nuclear IRBP was obtained by primers IRBP217 (Stanhope et al., 1992; Table 2). The PCR procedure was as follows: an initial denaturation of 3 min at 94°C, then 30 cycles with denaturation at 94°C for 60 seconds, annealing at 55°C for 60 seconds, and extension at 70°C for 2 minutes. The final extension lasted for 10 min at 72°C.Each PCR reaction used 25 – 250 ng of template DNA, 100 ng of each primer, 0.1 mM of each dNTP and 1 U of Taq polymerase in a final volume of 25 μl. Purified products were sequenced by Macrogen Europe by using the same primers as for PCR reaction.

Table 2. Regions sequenced, length of aligned sequences before and after removal of ambiguous sites, and primers used for three loci used in this study. L = forward primer, H = reverse primer; bp = base pairs.

	Region	**Primers**	**Reference**
cyt b	Entire locus (1140 bp)	L: 5′-ACCAATGACATGAAAAATCATCGTT-3′ H: 5′-TCTCCATTTCTGGTTTACAAGAC-3′	Irwin et al., 1999
IRBP	Part of the exon (1276 bp)	L: 5′-ATGGCCAAGGTCCTCTTGGATAACTACTGCTT-3′ H: 5′-CGCAGGATGATGAGGTGCTCCGTGTCCTG-3′	Stanhope et al., 1992
β-fibint 7	Entire intron (cca 750 bp)	L: 5′-ACCCCAGTAGTATCTGCCATTCGGATT-3′ H: 5′-GGGGAGAACAGAACCATGACCATCCAC-3′	Wickliffe et al., 2003

Phylogenetic reconstructions

Sequences of cyt b, IRBP and β-fibint 7 were examined using the SeqScape (version 2.6, Applied Biosystems) and BioEdit 7.0.9.0 (Hall, 1999). Further processing of sequences was conducting using Geneious 7.0.6. Haplotype dataset was created using DnaSP 5.10.01 (Librado and Rozas, 2009). Only long, high quality long sequences after removing of unambiguous sites were used in final alignments (Table 1, Appendix 1). Shorter sequences of cyt b were assigned to particular mitochondrial group by conducting simple Neighbor Joining tree in MEGA 6.04. This information were used to determine distribution of genetic groups. Complications during amplification of nuclear markers did not allow me to obtain sequences from all individuals (see Table 1, Appendix 1).

Evolutionary relationships among haplotypes were estimated by constructing phylogenetic trees using Maximum Likelihood and Bayesian Markov chain Monte Carlo phylogenetic analyses.

The FindModel web application (http://www.hiv.lanl.gov/content/sequence /findmodel /findmodel.html) was used to identify the most appropriate substitution model for each gene. The Akaike information criterion (AIC) compared among 12 model tests with initial tree constructed using Weighbor. General Time Reversible plus Gamma (GTR + G) was selected as a model with the highest AIC for cyt b and IRBP. The best fitting model for β-fibint 7 was Hasegawa-Kishino-Yano plus gamma (HKY + G). Sequences of three Murinae species downloaded from GenBank (http://www.ncbi.nlm.nih.gov/genbank/) were included to analysis as outgroups for each gene (Appendix 2).

Maximum Likelihood analysis was performed using PhymML 3.0 online web application (Guindon et al., 2010) with 1000 bootstrap replications and BIONJ starting tree. Bayesian analysis was performed by MrBayes 3.2.3 on the Cipress Scientific Gateway clusters (Miller et al., 2010). Three independent runs were performed with 10 million generations. Trees and parameters were saved every 1000 generation with first 25% of calculated trees discarded as a burn-in. Convergence was checked using TRACER v1.5 (Rambaut and Drummond, 2007). Bayesian posterior probabilities were used to assess branch support of the Bayesian tree. Phylogenetic trees were displayed using FigTree v1.4 (Rambaut, 2012).

Concatenated tree

A total of 37 sequences of cyt b, IRBP and β-fibint 7 were used to analyze concatenated dataset. Alignments of three genes were joined using FaBox 1.4.1 (Villesen, 2007). ML analysis was performed on the PhyML 3.0 (Guindon et al., 2010) with 500 bootstrap replications and BIONJ starting tree. Bayesian analysis was performed by MrBayes 3.2.3 on XSEDE with the same parameters as for single gene analyses (see above).

Network analysis

To visualize relationships among mitochondrial haplotypes I used program SPLITSTREE 4.10 (Huson and Bryant, 2006) using uncorrected p distances. Network was drawn using the equal angle method. Bootstrap (1000 replications) was used to calculate branch support. To further visualization of relationships among mitochondrial haplotypes, I constructed median-joining network in program NETWORK 4.6.1.1 (Bandelt et al., 1999).

Genetic distances and genetic variability

Genetic variability within cyt b dataset was assessed using DNAsp 5.10.01 (Librado and Rozas, 2009) to calculate haplotype diversity (Hd). Distribution of Kimura-2 parameter (K2P) corrected genetic distances on cyt b were calculate in MEGA 6.04 to evaluate possible sister species since the mean value between sister species of rodents is 7.3 % (sensu Baker and Bradley, 2006).

Divergence dating

Divergence times were estimated only for more complete cyt b dataset. Bayesian inference was performed using BEAST (Drummond and Rambaut, 2007) on XSEDE 1.8.0 clusters of Cipres Science Gateway (Miller et al., 2010) for cyt b. Input XLM file was prepared using BEAUti (Drummond et al., 2013). The three calibration points were derived from paleontological data on related murine rodents: origin of extant *Mus* (Jacobs and Flynn, 2005), origin of extant *Apodemus* species (Schenk et al., 2003) and the *Arvicanthis*/*Otomys* lineage split derived from the earliest records of *Otomys* (5 Mya; Denys, 1999) an the earliest records of genera *Aethomys*, *Arvicanthis* and *Lemniscomys* (6.08-6.12 Mya; Ambrose et al., 2007) compared to the next relevant sample where are these three and related genera absent except for *Aethomys* (9.5-10.50 Mya; Mein et al., 2004). Both standard deviations and offsets were exactly entered to set 5% and 95% quantiles to the fossil derived minimum and maximum ages (Table 3). Specimens from genus *Praomys* were included to avoid disproportionate evolution rate between used groups. All outgroup sequences are reported in detail in the Appendix.

Bayesian analyses were carried out using the GTR substitution model using coalescent model with constant size as the tree prior. Divergence times and their credibility intervals were estimated using relaxed clock model with substitution rates drawn from lognormal distribution. The MCMC simulation ran for 25 million generations, sampling every 2500 generations. Seven independent runs were performed from which the three runs with the highest effective sample sizes (ESS) were selected. Convergence and mixing were assessed in TRACER v1.5 (Rambaut and Drummond, 2007). Selected runs were combined in LogCombiner 1.8.0 (Drummond et al., 2013) and 10% of the trees were discarded as burn-in. The maximum clade credibility tree was calculated by TreeAnnotator 1.8.0 (Drummond et al., 2013) and the tree viewed in FigTree v.1.4.0 (Rambaut, 2012).

Table 3. Details on offsets, quantiles (in Mya) and standard deviations (S.D.), of the three calibration points used in dating of divergence.

Calibration point	**Offset**	**5% quantile**	**95% quantile**	**S.D.**
Origin of *Mus*	7.00	7.30	10.38	0.74
Origin of *Apodemus*	4.89	5.30	7.30	0.54
Arvicanthis/*Otomys* split	5.81	6.08	9.54	0.80

Ecoregions mapping

Habitat characteristics were obtained by overlaying gps coordinates with WWF terrestrial ecoregions of the world dataset (Olson et al., 2001), using R statistical software (R Core Team, 2014; Appendix 1).

Results

Phylogeny and distribution of genetic clades

Concatenated trees of the three genes (cyt b, IRBP and β-fibint 7) were well resolved with very similar topology in both Maximum Likelihood and Bayesian analyses and highlight the existence of several well-supported clades within *Mus triton* (Figure 2**Chyba! Nenalezen zdroj odkazů.**).

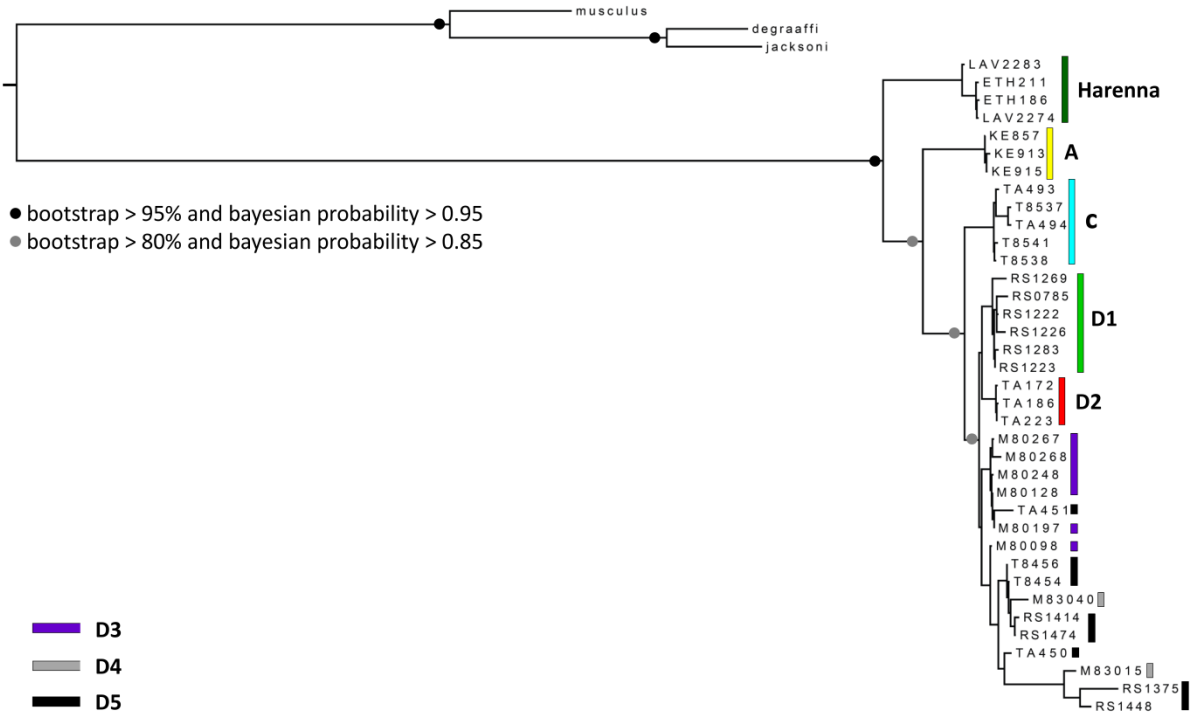


Figure 2. Phylogenetic relationships between 37 sequences of concatenated dataset of cyt b, IRBP and β-fibint 7 recovered by Maximum Likelihood analysis. The main clades are identified on the right border. Tip values indicate names of specimens.

Generated dataset of cyt b contained 119 sequences. In total, 69 cytochrome b sequences (55 new sequences, 11 sequences from GenBank and 3 sequences from African Rodentia) consisting of 1096 sites were used to perform. Additional 50 sequences were assigned to each mitochondrial clade to explore geographical range of particular phylogenetic groups (see Appendix 1).

Both Maximum Likelihood and Bayesian phylogenetic analyses of cyt b provided well resolved trees with five strongly supported groups (Figure 3). First split delimiting lineage called hereafter Harenna (according to the name of the single locality in Bale Mountains in Ethiopia, where is distributed) is strongly supported (99% for both ML bootstrap support (BS) and Bayesian posterior probability (BP). The second group (group A) with 98/99% (BS/BP) nodal support is represented by specimens from Kenyan Highlands (Mt Elgon, Nyahururu, Aberdares and Nanyuki) and northern part of Albertine Rift (Kigali, Rwanda and Kahuzi Biega, DRC). Group B (79% both BS and BP) corresponds to individuals from one lowland area in south-west DRC (Mbwambala near Kikwit). Groups C and D form a strongly supported monophyletic cluster; group C (node support 99% for both BS and BP) comprises individuals from Eastern Arc Mountains in Tanzania (Lulanzi, Mt Ngozi, Mufindi and Ukaguru), while the group D is widespread in Zambia, Tanzania, Malawi and Angola (see Figure 3). The latter consist of several well supported subgroups (node supports between 75 to 99% for both analyses). Subgroup D1 is distributed in Northern and Northwestern Province of Zambia (Nchila, Mutinodo, Kapyshia Hot Springs) and Angola (Huambo), D2 in Tanzania (Gerodom, Magangwe, Kyelele and Kalebe). Subgroups D3 and D4 are present in Malawi (D3 in Mpalaganga estate, Mulanje Mts, Ndawambe, Ntchisi, Perekezi Forest, Viphya Plateau and D4 in Nyika and Chelinda). The last subgroup D5 is distributed in southwestern Tanzania (Mbizi Hill Forest Reserve) and northeastern Zambia (Uningi Pans, Chishimba Falls, Kapuma Falls, Lumangwe Falls and Ntumbachushi Falls, see Figure 6).

Maximum Likelihood tree of both nuclear genes provided similar topology. The Harenna group again formed a sister group to all remaining lineages. Group A was supported as a sister clade to groups (C+D) (Figure 4Figure 5). The reciprocal monophyly of mitochondrial groups C and D, however, was not supported in any of two genes.

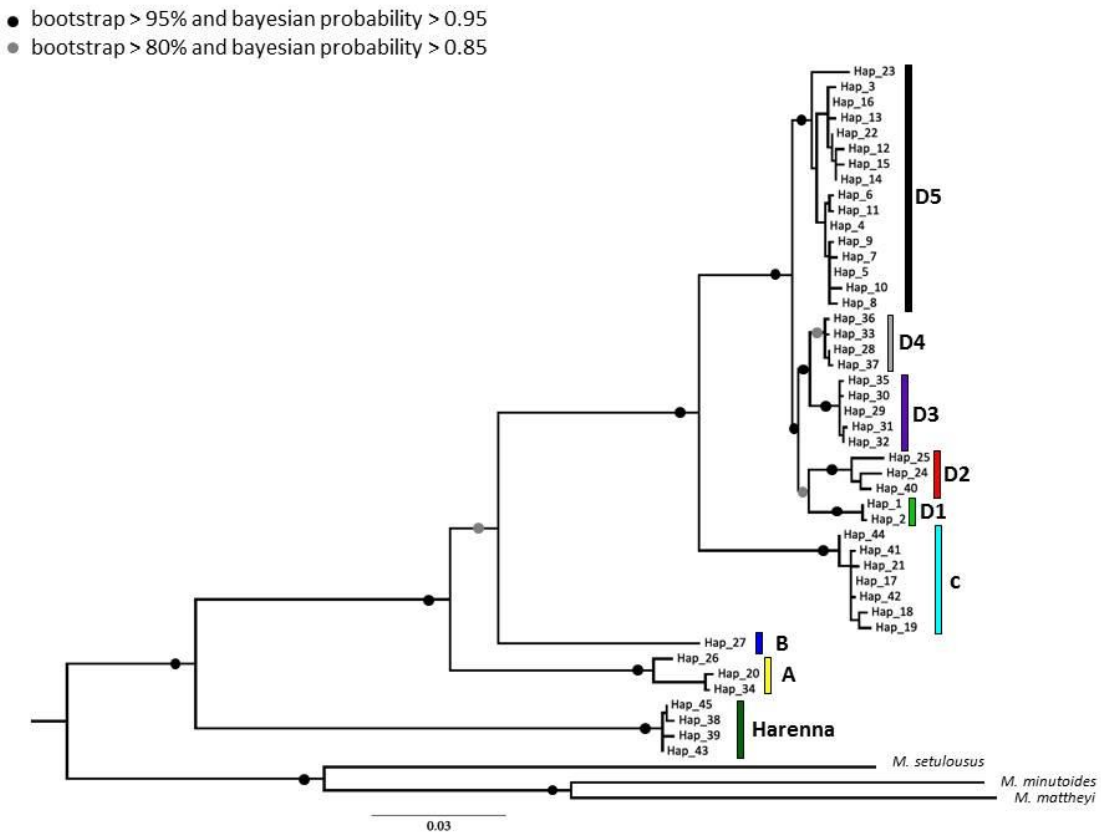


Figure 3. ML Phylogenetic tree based on analysis of 1086 nucleotides of the mitochondrial cyt b. The main clades are identified on the right border. Tip values indicate names of haplotypes.

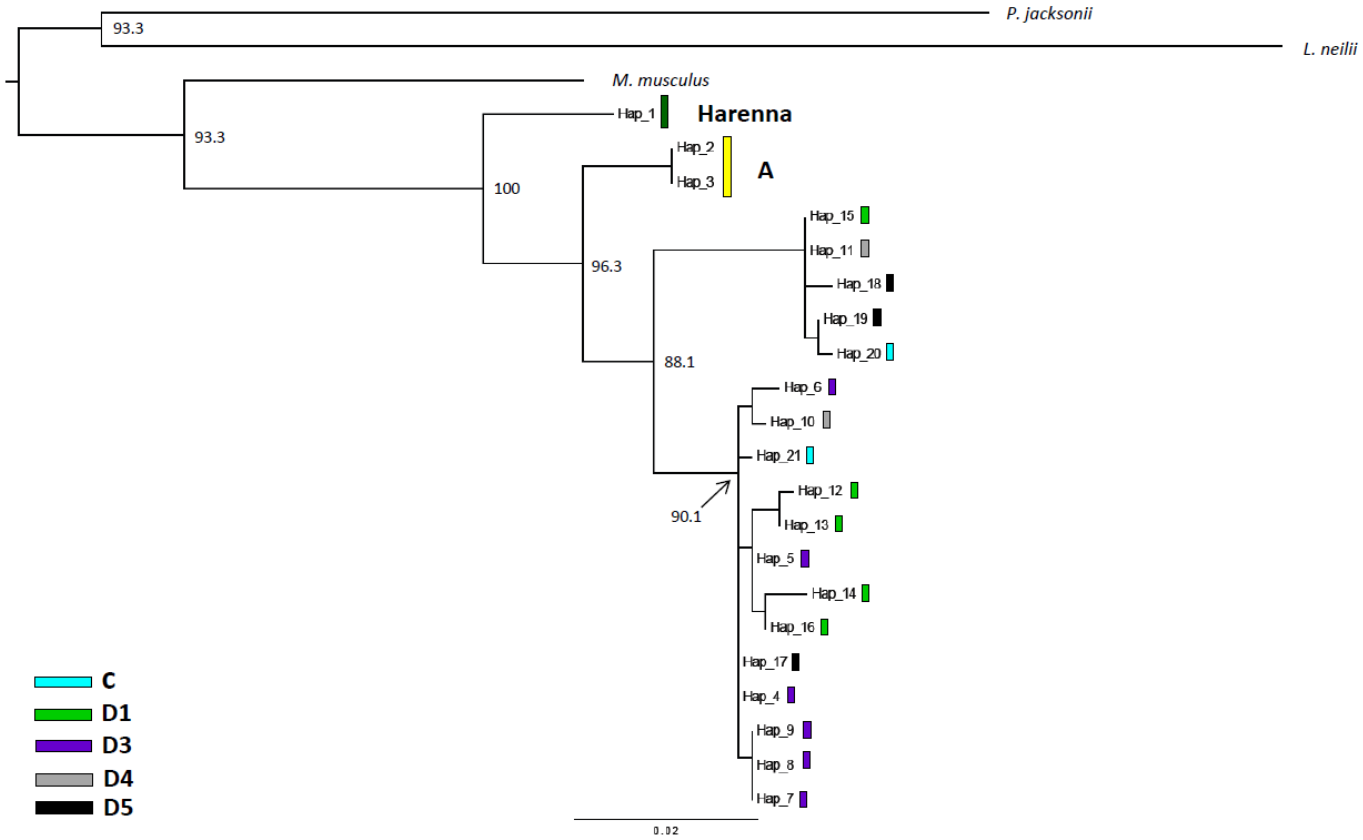


Figure 4. Results of ML based on analysis of the 700 nucleotides of the β-Fib7 nuclear gene. Numbers above differentiated branches indicate bootstrap scores (1000 replications) [%].

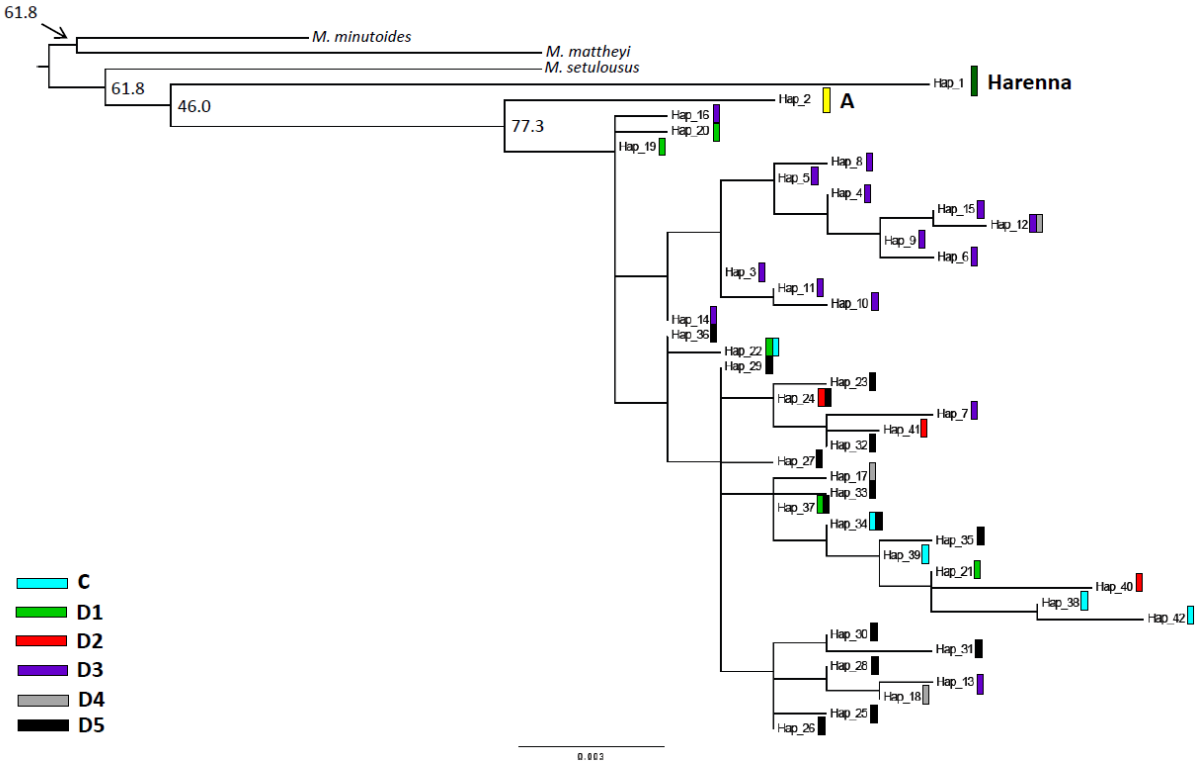


Figure 5. Results of ML based on analysis of 1049 nucleotides the nucleotides of the IRBP nuclear gene. Numbers above differentiated branches indicate bootstrap scores (1000 replications) [%].

Haplotype networks

Both haplotype networks (Figure 7a, b) connecting all 45 haplotypes of cyt b show the structure of *M. triton* complex confirming the presence of five well-differentiated phylogroups. Overall haplotype diversity (Hd) is high (0.997 ± 0.008) with the lowest H observed within subclades D1 and D2 (Table 6).

Genetic distances and divergence dating

K2P distances between five main cyt b mitochondrial groups (Harenna, A, B, C and D) range from 6.2 to 16.6% (Table 4). The greatest distance was between "Harenna" and remaining clades followed by group A. Differences between C and D were relatively low. K2-P distances among subgroups D1-D5 were quite low from 1.1 to 2.9% (Table 5).

Molecular clock estimated the split between the Harenna lineage from the ancestor of the remaining lineages to occur at the end of Miocene at 5.37 with 95% of the highest posterior density (HDP) 4.38-6.52. (Figure 9). The split between lineage A and phylogroup B+C+D occurred according to my analysis during Pliocene at 3.33 Mya (HDP 2.53-4.12) followed by the split of B and C+D at 2.7 Mya (2.01-3.43). Divergence between C and D took place at 1.65 (1.17-2.23) Mya. All divergences within group D are younger than 1 Mya.

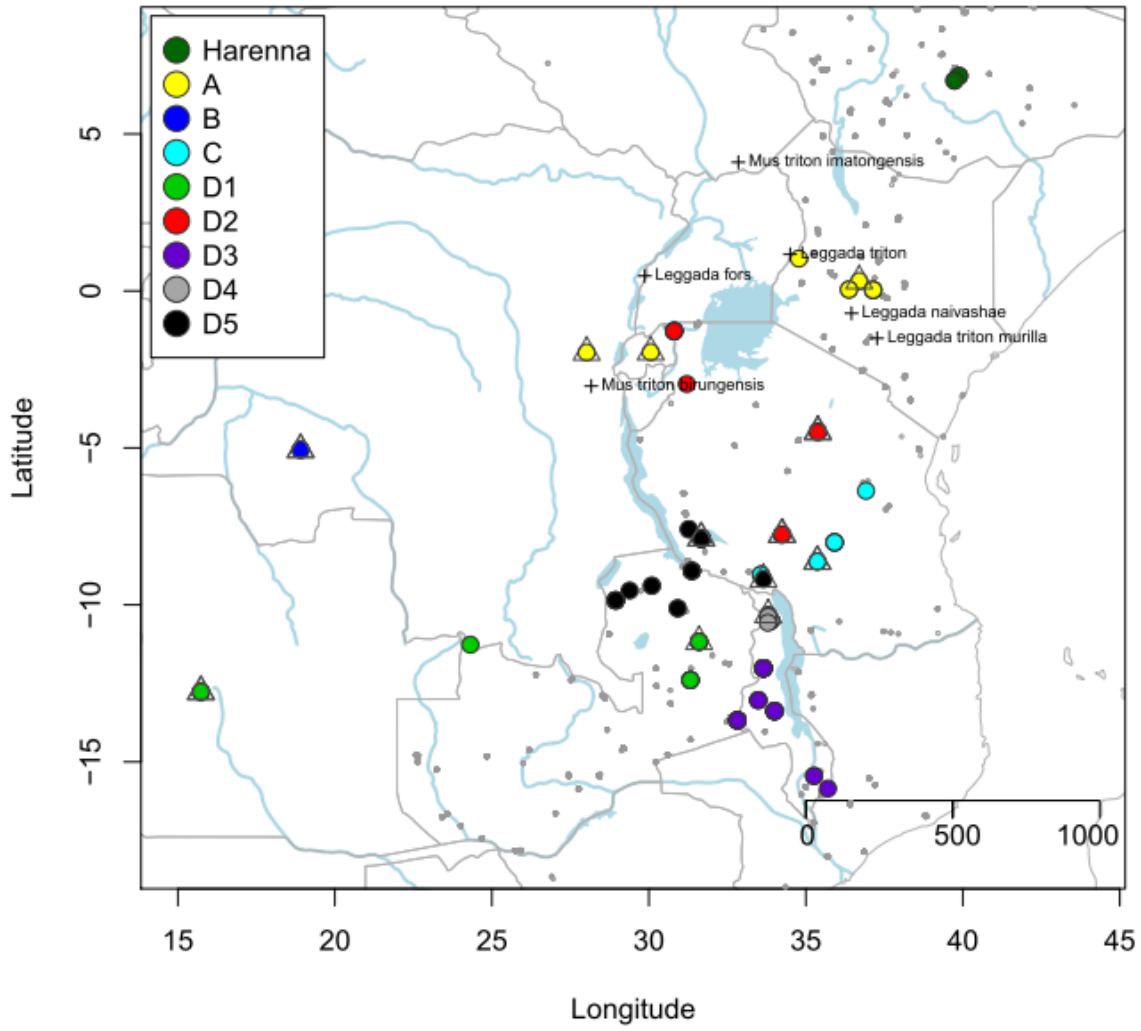


Figure 6. Geographical distribution of specimens identified as Mus (Nannomys) triton in our analysis. Group Harenna – dark green dots, group A – yellow dots, group B – dark blue dots, group C – light blue dots, group D1 – light green dots, group D2 – red dots, group D3 – dark purple dots, group D4 – grey dots, group D5 – black dots. Dots without triangles = samples collected during our fieldworks, dots with triangles = sequences obtained from African Rodentia database. Black crosses – original descriptions of Mus triton (sub)species. Grey dots on the background = all trapping localities

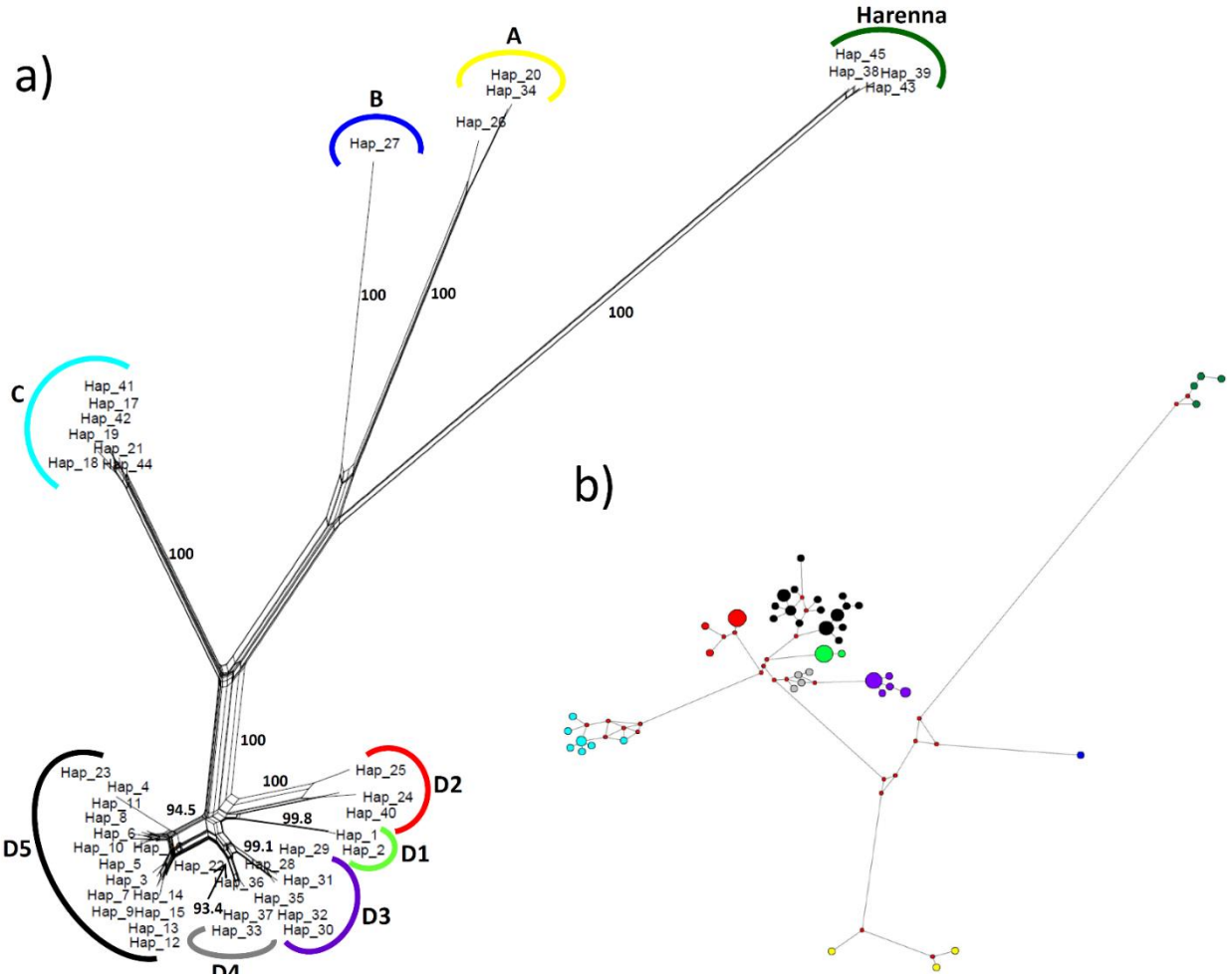


Figure 7. Phylogenetic networks constructed using Splits Tree 4.0 (a) and NETWORK (b). Taxon labels represent haplotype definition and highlight colors represents clades (dark green for Harenna, yellow for A, dark blue for B, light blue for C, light green for D 1, red dots group D2, dark purple for group D3, grey for D4, black forD5). Uncorrected p distances were used and the networks were drawn using the equal angle method. Bootstrap support values (1000 replicates) are shown on selected branches.

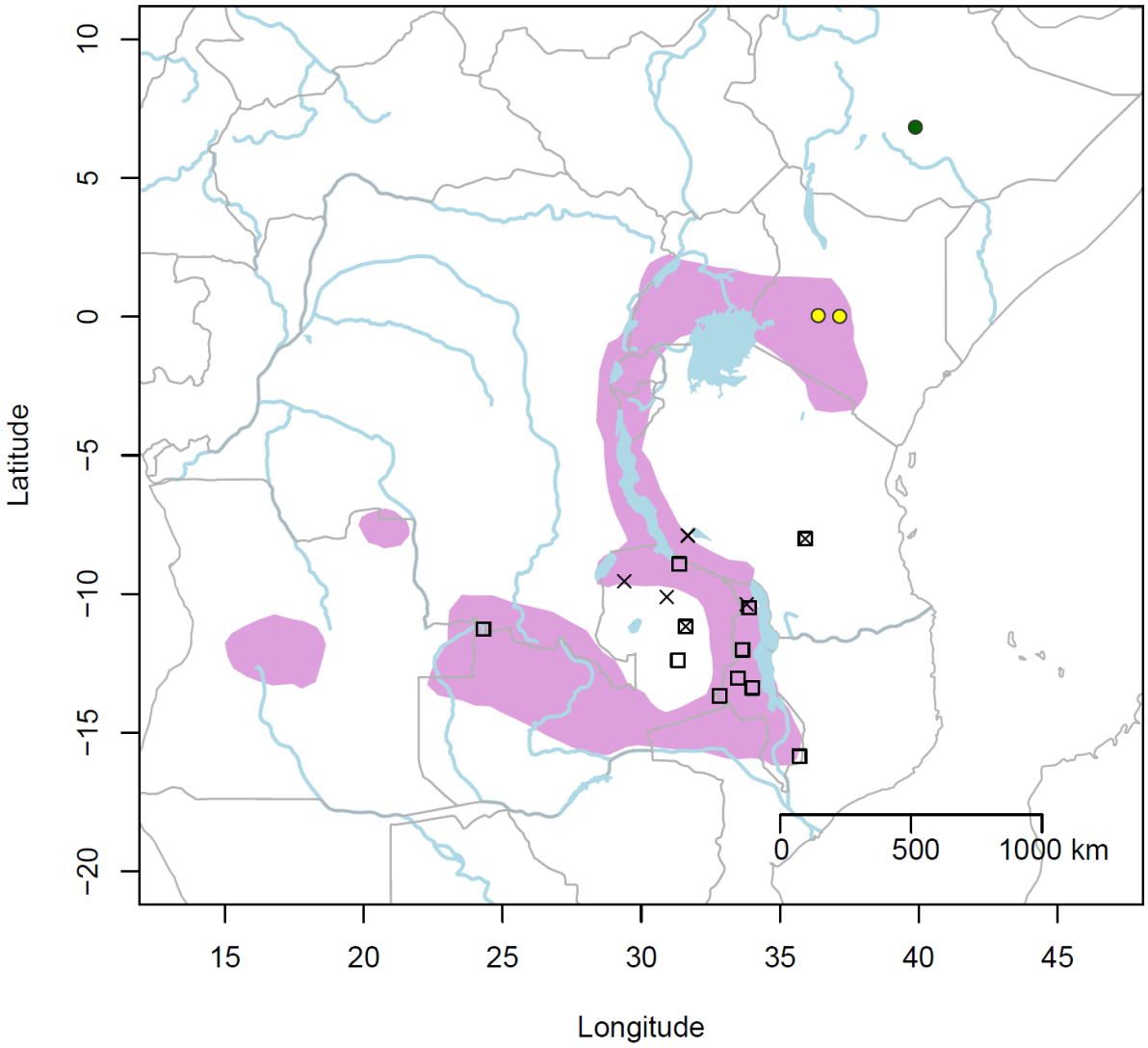


Figure 8. Geographical distribution of Mus (Nannomys) triton groups based on β-Fib7. Group Harenna – dark green dots, group A – yellow dots, group "C" – black crosses, group "D" black squares. Geographic range according to IUCN is marked pink.

Table 4. Average Kimura 2-parameter genetic distances within (diagonal) and between mitochondrial groups of *Mus triton*.

Group	**A**	**B**	**C**	**D**	**Harenna**
A	0.021				
B	0.096	0.000			
C	0.126	0.106	0.003		
D	0.124	0.107	0.062	0.018	
Harenna	0.164	0.151	0.166	0.158	0.002

Table 5. Average Kimura 2-parameter genetic distances within (diagonal) and between mitochondrial subgroups of *Mus triton*.

Clade	**D1**	**D2**	**D3**	**D4**	**D5**
D1	0.007				
D2	0.029	0.004			
D3	0.027	0.028	0.002		
D4	0.026	0.023	0.011	0.001	
D5	0.029	0.027	0.021	0.18	0.006

Table 6. Genetic diversity for cyt b sequences in *Mus triton*. Number of individuals (N), number of haplotypes (H), haplotype diversity (Hd), standard deviations (SD) and nucleotide diversity (π) are shown.

mtDNA groups	**N**	**H**	**Hd**
Harenna	4	4	1
A	3	3	1
B	1	1	x
C	8	7	0.964
D1	10	2	0.286
D2	10	3	0.464
D3	10	5	0.756
D4	4	4	1
D5	19	16	0.957
Total	69	45	0.977

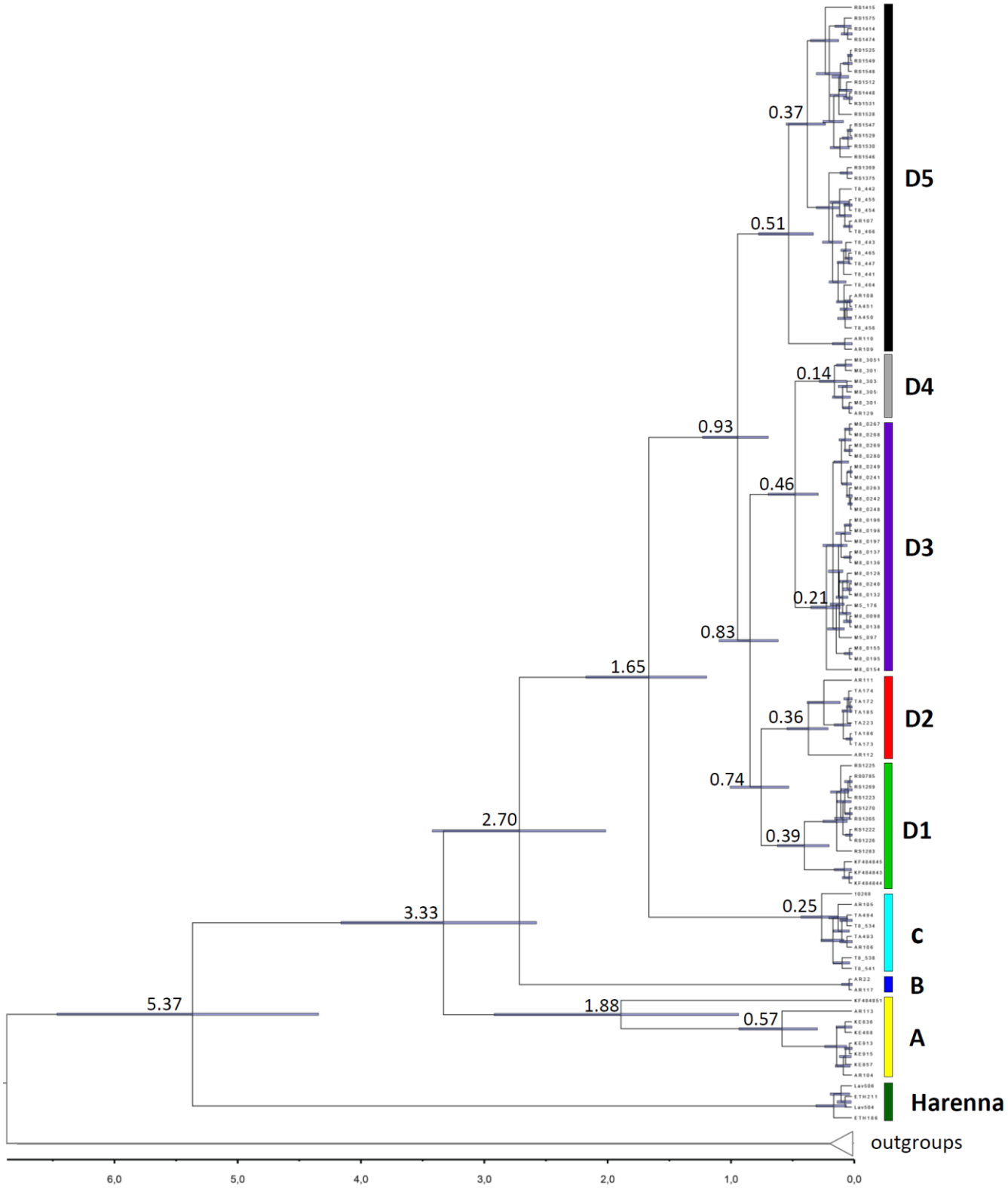


Figure 9. Chronogram illustrates divergence of major *M. triton* phylogroups based on cytochrome b. Relaxed clock model with uncorrelated lognormal distribution was used in BEAST 1.8.0. The scale is in million years. Horizontal dars represent 95% highest posterior densities (HPDs).

Discussion

In this study I present results of a phylogenetic investigation on the Grey-bellied Pygmy mouse (*Mus triton*) complex using three genetic markers. Phylogenetic analyses based on mitochondrial (cyt b) and two nuclear markers (IRBP and β-fibint 7*)* revealed a clear genetic structure within *M. triton*, previously considered as a single species (Dieterlen and Happold, 2013). According to the mitochondrial DNA, four genetically-well distinct monophyletic lineages with a clear geographic pattern exist within the "*triton*" complex (Figure 3). Further geographical structuring can be found also within the mitochondrial clade D. Both nuclear markers confirmed groups A and "Harenna" as distinct basal lineages but failed to recognize subtle pattern in more recent radiations of the groups C and D. These two groups form one, low supported group in case of IRBP and two moderately supported groups in case of β-fibint 7.

Incongruity between mitochondrial and nuclear results

Mitochondrial DNA (mtDNA) has been widely used over the last 25 years as a main genetic marker for phylogeny in many mammal species. The wide use of mtDNA has following practical reasons. It is easy to amplify and it is non-recombining molecule. In addition, a huge body of mitochondrial sequences has been accumulated over the time in online databases. The main differences to nuclear DNA are inheritance through the maternal lineage only, haploid genome, higher mutation rate, and smaller effective population size (Ballard and Whitlock, 2004).

*S*everal possible explanations of the observed topological incongruences between mtDNA and nuclear DNA (nuDNA) can be drawn. First, mtDNA has higher evolutionary rate and four times larger effective population size than nuDNA. This leads to longer coalescent times and less clear structure observed by nuDNA. Second, the selective mitochondrial sweep on maternally inherited loci can generate these incongruences. This phenomenon was already suggested in various animal species as insects (Shaw, 2002), birds (Irwin et al., 2009) and also in rodents (Kennedy and Nachman, 1998). Third, introgression of mitochondrial DNA during hybridization is another explanation for differences between mtDNA and nuDNA. Introgressive hybridization has been already suggested in other African mammals as antelopes (Masembe et al., 2006), mice of genera *Praomys* (Bryja et al., 2014) *Mastomys* (Lecompte et al., 2003).

IRBP was successfully used in muroid rodents phylogenies (e.g. Jansa and Weksler, 2004, Weksler, 2003), but the data provided convincing evidence mainly on and above generic level. Altogether with the fact, that in fast evolving mammals (rodents) the substitution rate of mtDNA can be 100 times higher than the substitution rate of IRBP (Nabholz et al., 2008), the most possible solution is that in case of *M. triton* the nuclear IRBP was not able to accumulate sufficient number of changes in comparison with the mitochondrial cyt b.

"*Mus triton*" lineages

Until now, genetic diversity of *M. triton complex* was not investigated by any type of taxonomical approach. My study thus provides the first results about "triton" group diversity. Importantly, these results were assessed on material collected across the significant portion of the geographical distribution of the taxa.

Geographical distribution of "triton" lineages

Group "Harenna"

This basal lineage, supported by both mtDNA and nuDNA (Table 4Table 5, and Figure 2) phylogenies, forms a very distant phylogroup (see Table 4 for distances and Figure 3). This lineage has also different karyotype. Aniskin et al., (1998) reported 34 chromosomes in *M. triton* from Harenna, whereas M. triton from Kenya has only 32 chromosomes (Veyrunes unpubl. results). This lineage is probably endemic to Harenna forest in the Bale Mountains, Ethiopia, the site well-known for high species diversity and high rate of endemism of small mammals (e.g. Lavrenchenko, 2000, Yalden and Largen, 1996). Several morphological differences between the "Harenna" and the other lineages were also reported. Individuals from Harenna forest appear to be also heavier (weight = 16.25 g, N=2; our unpublished data) compared to specimens from the rest of the "triton" distributional range. In addition, dorsal pelage of the "Harenna" specimens is olive-brown (Yalden, 1988) differing from darkish-brown pelage of the other "*triton"* species (Dieterlen and Happold, 2013).

Group A

This highly supported, genetically distinct lineage (K2P distances to other *triton* groups ranging from 9.6 to 16.4%) is distributed in Kenya highlands and also in the Albertine Rift. This group was also supported by both nuclear genes. Available cytochrom b sequences from Kahuzi-Biega, eastern DRC (GenBank, accession number KF484851) published in Lamb et al. (2014) and Kigali Rwanda (African Rodentia, accession code RW5368) also belong to this

group, as well as herein published sequences of the specimens collected on Mt. Elgon, Nyahururu and Nanyuki localities in Kenya. All specimens of this group collected in the three localities of Kenyan highlands were trapped mainly in the moist, shrubby habitats and at the edges of cultivated fields.

Group B

This group is represented by specimens collected only at two nearby localities in Kikwit and Mbwambala, DR Congo. Nevertheless this lineage could be distributed much more widely. It seems that mice of this lineage are common around the Kikwit city, as the "*Mus triton*" was the most abundant rodent species collected during Ebola project (27.9% from 1170 specimens, Leirs et al. 1999). Unfortunately no other sequences apart from cyt b are available. These specimens were trapped in moist habitats such as in disturbed secondary forest, fallow fields and vegetation along watercourses. Interestingly, the altitude of Kikwit is very low (around 500 m a.s.l) compared to the altitudes where specimens of other *triton* lineages were collected (usually above 1000 m a. s. l; see Appendix 1).

Clades C and D

The last phylogroup (supported by all three markers) occurs in relatively large are area extending from Eastern Arc Mountains across Zambia to Angolan highlands and southern Malawi. Specimens from Angola identified as *M. callewaerti* (Thomas, 1925) (GenBank accession numbers KF484843, KF484845, KF484844) by Lamb et al. (2014) also belong to group C+D in my study. High mitochondrial divergence splits this lineage into two well supported mitochondrial clades C and D (K2P distance = 6.2%). The later lineage is further divided into five subclades. Clades C and D have distributional overlap near Mbeya, southwest Tanzania (Figure 6). The lineage D could be in contact also with the lineage A in north-western Tanzania, although sympatric occurrence has not been recorded. Across its distribution range, the C+D phylogroup occurs in a wide range of habitats. It should be noted, that mitochondrial lineage C seems to be confined to Eastern Arc, which is known for a high level of endemism (see Burgess et al. 2007), in contrast to the lineage D with an affinity to more open forests, woodlands and forest grassland mosaics (see Appendix 3).

Temporal diversification of "*Mus triton*" complex

Analysis of mitochondrial and nuclear markers confirms basal divergence of the lineage found only in the isolated Harenna Forest. According to cyt b the split of this lineage occurred at 5.37 MY (mean value, see Figure 9). The first *Nannomys* radiation probably occurred in

Ethiopia (or Eastern Africa) soon after the ancestor of all *Nannomys* crossed the Red Sea (between 6.8 and 7.8 Mya). This time frame is supported by the oldest fossils belonging to *Nannomys* (see Introduction). According to Lamb et al., (2014), the first radiation occurred around 6.33 Mya. The split of the "Harenna" lineage may belong to this early radiation of *Nannomys* on the African continent. Nowadays, the Harenna forest is very unique habitat in Ethiopia and in fact in the whole Africa. It lies on the Southern slope of the Bale Mountains. The Bale Mountains form the largest continuous area above 3000 m in Africa. This mountain block is characteristic by a distinctive endemic fauna and flora, resulting from the combination of a relatively large area, geographic isolation, and climatic history (Kingdon, 1989). The southern slopes of the Bale Mountains represent a continuous range of natural vegetation ranging from 1500 to 4400 m a.s.l. The largest part of this area is covered by montane evergreen forest (1500 to 3250 m), which is isolated from similar forest types by Rift Valley to the west, afroalpine Sanetti Plateau to the North and arid lowlands on the other sides. We may assume that ancestor of the "Harenna" lineage adapted to conditions of the montane evergreen forest or its ecotone (It should be mentioned that both two specimens collected by Šumbera and Bryja were collected in forest clearings). The moist climate at this period probably allowed the ancestor of remaining lineages *of "triton"* complex to spread its distribution to the southern montane areas in Kenya. During the Upper Miocene (8 to 6 Mya) there was an aridification trend which is documented by transition from forest to grassland followed by an increase in the biomass of C_4 plants (Cerling et al., 1997). Later (5 to 3 Mya), the increase of aridity continued as it is possible to see in the fossil record from Omo valley, Ethiopia represented by increasing dominance of grazing mammal species (Bobe, Behrensmeyer and Chapman, 2002). We may suppose that the lineage from the Harenna forest was isolated by arid conditions in areas around the Bale massif in the beginning of this aridification process. Because of the rainfall from the Indian Ocean and also the presence of large glaciers during Pleistocene which supported montane forest zone (Osmaston et al 2005), humid climate persisted even during regular arid-humid cycles during Pliocene and Pleistocene. Genetic divergence between the "Harenna" and remaining lineages of the "triton" group indicates that the Harenna Forest was isolated from other forest blocks for a long period of time (It should be noted that other *Nannomys* species were collected in areas around Harenna forest, but this lineage of the "triton" complex was not found elsewhere outside the forest, Šumbera pers. communication).

The split between lineage A and phylogroup containing B+C+D clades is estimated to occur between 4.2-2.5 Mya (Mean 3.33 Mya, see Figure 9). Again this event could be caused

by general aridification occurring in Eastern Africa during this period. Present discontinuous distribution of the lineage A in Kenyan Highlands and Albertine rift may reflect habitat unsuitability for this lineage or alternatively lack of sampling in sites between both mountainous ranges. I suggest that the gap is natural, because between both mountainous areas there is drier lowland area called Uganda gap. Nevertheless, it is well known that this gap did not form impenetrable barrier for small mammals as is already known for montane forest small mammals of the genera *Hylomyscus* and *Sylvisorex*. (Demos et al., 2014). Reduction of montane and lowland forests occurred during xeric phases of the Plio-Pleistocene climatic oscillations. Three main peaks of aridity at 2.8, 1.7 and 1.0 Mya (deMenocal, 2004) were followed by mesic periods. Presumably, these mesic periods allowed migration of forest-dwelling species by corridors which were probably covered with suitable habitat. Thus, *Mus triton* could pass along forests from Kenyan highlands to Albertine rift during humid climate between two arid periods at 2.8 and 1.7 at 1.88 as resulted from BEAST analysis in this study. Although associated with woodland/forest habitats *Mus triton* seems to be present mainly in dense vegetation on clearings, edges of the forest, along water courses, etc. (Šumbera, Mazoch pers. communication), so I expect that humid window for dispersal could have been opened for a longer time period when compared to *Hylomyscus* or *Sylvisorex*.

Geographical distribution of the lineage B from Kikwit is far from other "*triton*" specimens described in this study. Missing data from southern Congo do not allow to confirm if the population is truly isolated or alternatively, if there is a lack of sampling in this area. I assume, that the ancestor of this lineage crossed the Congo basin after splitting from C+D clades at approx. 2.7 Mya. It is clear, that large rivers could complicate crossing the Congo basin. Nevertheless, even the largest river here, the Congo River, is not necessarily insuperable barrier. It is known that this river did not act as barrier to gene flow in another rodent species, *Lemniscomys striatus* which is even more confined to dry habitats (Nicolas et al., 2008). Nevertheless, due to the lack of sequences in the area, any colonisation scenarios explaining current distribution of the lineage B around Kikwit are speculative. Interestingly, area around Kikwit was previously suggested as a place with a pronounced local endemism (Colangelo et al., 2005)

The split of lineage containing the clades C+D took place around 2.7 Mya, i.e. during the onset of climatic changes caused by expansions of glaciers and the first aridity peak at end of Pliocene (deMenocal 1995). Although trapping was also conducted in northern parts of East African Mountains in Tanzania and Kenya (this study) no "*triton*" was among the trapped

specimens. Due to an influence of the Indian Ocean, the forest habitats of East African Mountains experienced long-term stability leading to the persistence of forest species, reducing rates of extinction, presence of relict lineages, etc. (for review see Burgess et al. 2007). These factors probably contributed to establishing and persistence of the clade C. During further expansion of the clade to the Malawian Southern Rift Montane area, it had to cross Makambako gap which is an area of a dry habitat (Burgess et al., 2007). It probably happened during an expansion of favourable habitats in a more humid period. The split of the clades C and D occurred at 1.65 Mya. Again this time frame correlates with one of the suggested aridity peaks at 1.7 Mya (deMenocal, 1995, 2004) which again caused contraction of forest habitats. The clade C remained in Eastern Arc Mountains, whereas the clade D is more confined to drier open canopy woodland habitats.

No other important split – at least among extant taxa - occurred between 1.7 Mya and the next aridity peak at approx. 1.0 Mya (see – Beast Figure 9). After 1.0 Ma African climate variability shifted towards the amplitude of 100 kyr cycles synchronous with the high-latitude ice cycles (deMenocal 2004). The highest posterior densities values (95%) of main splits within D clade roughly correspond to these 100 kyr cycles (0.93, 0.84, and 0.74) suggesting connection between climate fluctuations and segregation of mtDNA subclades within D lineage.

The most recent mitochondrial divisions occurred in the southern part of the "*triton*" complex distribution. Several hypothesis can be drawn from geographic confinement of the D subclades. The first noticeable feature of the clade D phylogeny is geographic separation of the clade D2 (red dots in the Figure 6) on Tanzania Plateau (area in Tanzania defined by Eastern Arc Mountains on the east, Lake Tanganyika on the west and in the south by Lake Rukwa area).

A similar pattern of distribution of one of the mtDNA lineages of *Mastomys natalensis* was observed by Colangelo et al. (2013). Authors stress out importance of fault systems of the Tanganyika-Rukwa-Malawi (TRM) segment of EARS as an area of the current secondary contact zone. In this area important tectonic activity and faulting was recorded up to the mid-Pleistocene, accompanied by significant water level fluctuations. For example Lake Rukwa flowed into Lake Tanganyika during the last humid period dated (Delvaux et al., 1998). TRM forms an important barrier even for larger mammals (review Lorenzen 2012). During periods of increased humidity, both Rukwa and Malawi rifts could act as a barriers subdividing clades of *M. triton*, in this case for clade D2. Formidable barrier to *M. triton* dispersal to the other

areas outside Tanzania plateau could also have been provided by a proposed ‘forest circle’ covering the EAM as well as the southern and Albertine Rift during humid Pleistocene periods (please see Colangelo et al. 2013, Supplementary Figures for further reference).

The clear geographic pattern between clades D1, D5 on one side and D3, D4 (confined to specific conditions of Nyika plateau) on the other side could be also related to Pleistocene climatic events. It is known that at this period was formed do called "arid corridor" allowing migration of arid fauna between dry eastern Africa and dry south-western Africa. At this time more mesic habitats such as miombo woodlands were contracted along this expanding arid habitats. Grubb et al. (1999) proposed that in periods of extreme aridity, drier vegetation might have replaced miombo woodlands completely. There are still some remains of this corridor especially along Luangwa valley. In the low altitudes along Zambezi and Luangwa rivers, a typical vegetation type is Mopane woodland (see Appendix 3 for distribution of Mopane). It seems that this type of woodland is impenetrable for species adapted for more humid habitats such as “*triton*”. Indeed, intensive trapping in mopane did not confirm occurrence of this specie here (Mazoch and Šumbera pers. Com.).

Clade D5 is restricted to SRM in south Tanzania and northeast Zambia (black dots on Figure 6). This clade has a recent zone secondary contact with the clade C near Mbeya, Tanzania and with the clade D2 in Mbizi forest, Tanzania. As previously noted, humid periods could have restricted clade D5 from access to Tanzanian plateau by “forest circle” and fluctuating water levels in Lake Rukwa area (TRM). On the east side the contact with D3 and D4 clades could have been prohibited by expanding Mopane woodlands during phases of aridification. So the last remaining hypotheses should target historical geographic isolation of D5 from D1 clade. Clade D5 is enclosed by water barriers, consisted by Luapula River and Lake Bangweulu, surrounded by vast wetlands on the south-west and it is never crossing Chambeshi River on the east. On the other side D1 does not seems leave Luangwa escarpment. Moreover, all three rivers have complicated recent (> 1 Mya) drainage evolution, including reversion of flow direction (see Stankiewicz 2006).

Recently announced publication of detailed map of global ecological land units of the world by the U.S. Geological Survey (USGS) and Esri. Pattern of Hot Moist (Figure 10) bioclimate surprisingly accurately predicts distribution of *M. triton* groups, including Harenna lineage, even South African fossil sites. The map is also showing belt of Hot Moist climate, which stretches from central Zambia to west Angola. This belt corresponds to confirmed GPS locations of clade D1 specimens and helps with explanation of this rather unusual distribution.

Figure 10. map of global ecological land units of the world by the U.S. Geological Survey (USGS) and Esri. Pattern of Hot Moist climate is marked with green color.

It has been proposed that diversification in many African small mammals was related to climatic oscillations during the Plio-Pleistocene (Rambau et al. 2003; Russo et al. 2006; Russo et al. 2010, Taylor et al. 2009). History of the *Mus triton* species complex was obviously shaped by these changes in climatic conditions as well as by the presence of various barriers. Since *M. triton* complex consists of higher altitude dwelling species living mainly in moist habitats, its various populations remain isolated by reduction of suitable habitats during the Pleistocene glacial = aridification phases. These dry and grassy habitats unsuitable for the "triton" vanished with the onsets of moist interglacial periods, when forest habitats replaced some drier habitats (deMenocal, 1995, 2004). Woodland bridges between previously isolated montane islands have favored further migrations of *M. triton* complex.

Taxonomic implications for "*Mus triton*" complex

My findings correspond with assumptions about high cryptic diversity within *Nannomys* subgenus (Lamb et al., 2014, Veyrunes et al., 2005) and also within *Mus triton* (XXX). Nodal support (based on mtDNA and partially also nuclear markers) together with genetic distances justifies division of "*M. triton"* into several species. Separate species status for "*Nannomys triton*" from Harenna Forest was already suggested by Lavrenchenko (2000). Distinct status of this lineage is supported by high mitochondrial and nuclear divergences. Further data on morphometry, genetics and ecology are needed to properly define taxonomic status of this lineage. Lineage A is also well supported by mitochondrial and nuclear markers. The geographic position of the type locality of *Mus triton* (Mt. Elgon, Kenya; Thomas, 1909) and absence of other lineages in Kenyan highlands supports denomination of lineage A as *M. triton* sensu stricto. Apart from Kenyan highlands this species occurs also in Albertine Rift. Specimens from Kikwit (DRC) forms a distinct mitochondrial lineage (B) but due to the absence of nuclear data it is difficult to convincingly decide about its taxonomic status. Two remaining lineages C and D have the largest geographic distribution. Although genetic differences on cytochrome b exceeded five percent margin, an appropriate threshold for separation of monophyletic populations into separate species (according to Baker and Bradley 2006), I advise to treat both groups as a single species, because the lineages does not form monophyletic clusters when nuDNA analysis is employed. I suggest to name lineage (C+D) as *Mus callewaerti* since the Angolan specimens reported in Lamb et al. (2014) belongs to group D and were described as *M. callewaerti* (Thomas, 1925). Further detailed study on morphology and comparing with type material are necessary for confirming validity of this name for the lineage containing clades C+D.

References

Ambrose, S. H., Bell, C. J., Bernor, R. L., Boisserie, J. R., Garcia, N., 2007. The paleoecology and paleogeographic context of Lemudong'o locality 1 A late Miocene terrestrial fossil site in southern Kenya. The Cleveland Museum of Natural History 56, 38–52.

Aniskin, V. M., Lavrenchenko, L. A., Varshavskii, A. A., Milishnikov, A. N., 1998. Karyotypes and Cytogenetic Differentiation of Two African Mouse Species of genus *Mus* (Rodentia, Muridae). Russian journal of genetics 34(1), 80-85.

Ansell, W.F.H., 1957. Some mammals from Northern Rhodesia, Annals and Magazine of Natural History (12) 10, 529-551. In: Hanney, P., 1965. The Muridae of Malawi (Africa: Nyasaland). Journal of Zoology 146, 577-633.

Avery, D.M. 1991. Late Quaternary incidence of some micromammalian species in Natal. Durban Museum Novitates 16, 1-11.

Baker, R. J., Bradley, R. D., 2006. Speciation in mammals and the genetic species concept. Journal of Mammalogy 87(4), 643-662.

Ballard, J. W. O., Whitlock, M. C., 2004. The incomplete natural history of mitochondria. Molecular ecology *13*(4), 729-744.

Bandelt H.,J., Forster P., Röhl A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16, 37-48.

Bonnefille, R., 2010. Cenozoic vegetation, climate changes and hominid evolution in tropical Africa. Global and Planetary Change 72, 390-411.

Bonnefille, R., 1983. Evidence for a cooler and drier climate in the Ethiopian Uplands towards 2.5 Myr ago. Nature 303, 487-491.

Britton-Davidian, J., Robinson, T.J., Veyrunes, F., 2012. Systematics and evolution of the African pygmy mice, subgenus *Nannomys*: A review. Acta Oecologica 42, 41-49.

Bryja, J., Mikula, O., Patzenhauerová, H., Oguge, N.O., Šumbera, R., Verheyen, E., 2014. 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., Butynski, T. M., Cordeiro, N. J., Doggart, N. H., Fjeldså, J., Howell, K. M., Stuart, S. N., 2007. The biological importance of the Eastern Arc Mountains of Tanzania and Kenya. Biological conservation 134(2), 209-231.

Catzeflis, F.M., Denys, C., 1992. The African *Nannomys* (Muridae): An early offshoot from the *Mus* lineage – evidence from scnDNA hybridization experiments and compared morphology. Israel Journal of Zoology 38, 219-231.

Chevret, P., Veyrunes, F., Britton-Davidian, J., 2005. Molecular phylogeny of the genus Mus (Rodentia: Murinae) based on mitochondrial and nuclear data. Biological Journal of the Linnean Society 84, 417-427.

Chimimba, C.T., Bennet, N., 2005: Order: Rodentia. In: Skinner, J.D., Chimimba, C.T. (Eds.), The mammals of the southern African. Cambridge Univ. Press, 142.

Chitaukali, W. N., Burda, H., and Kock, D., 2001. On small mammals of the Nyika Plateau, Malawi. African small mammals, 415-426.

Chorowicz, J., 2005. The East African rift system. Journal of African Earth Sciences 43(1), 379-410.

Cerling, T. E., Harris, J. M., MacFadden, B. J., Leakey, M. G., Quade, J., Eisenmann, V., Ehleringer, J. R., 1997. Global vegetation change through the Miocene/Pliocene boundary. Nature 389(6647), 153-158.

Colangelo, P., Corti, M., Verheyen, E., Annesi, F., Oguge, N., Makundi, R. H., Verheyen, W., 2005. Mitochondrial phylogeny reveals differential modes of chromosomal evolution in the genus< i> Tatera</i>(Rodentia: Gerbillinae) in Africa. Molecular phylogenetics and evolution 35(3), 556-568.

Colangelo, P., Verheyen, E., Leirs, H., Tatard, 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(4), 901-916.

Delvaux, D., Kervyn, F., Vittori, E., Kajara, R. S. A., Kilembe, E., 1998. Late Quaternary tectonic activity and lake level change in the Rukwa Rift Basin. *Journal of* African Earth Sciences 26(3), 397-421.

DeMenocal, P. B., 1995. Plio-Pleistocene African climate. Science (New York, NY) 270(5233), 53-59.

deMenocal, P.B., 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. Earth and Planetary Science Letters 220, 3-24.

Demos, T.C., Kerbis Peterhans, J.C., Agwanda, B., Hickerson, M.J., 2014. Uncovering cryptic diversity and refugial persistence among small mammal lineages across the Eastern Afromontane biodiversity hotspot. Molecular Phylogenetics and Evolution 71, 41-54.

Denys, C., 1999. Of mice and men. Evolution in the East and South Africa during Plio-Pleistocene times. In: Bromage, T.G., Schrenk, F, (Eds.), African Biogeography, Climate Change, and Human Evolution. Oxford University Press, New York, 226-252.

Dieterlen, F. and Agwanda, B., 2008. *Mus triton*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on 04 December 2014.

Dieterlen, F. and Happold, D.C.D., 2013. Mammals of Africa. Volume III: Rodents, Hares and Rabbits. Bloomsbury Publishing, London, 497-498.

Dobigny, G., Tatard, C., Gauthier, P., Ba, K., Duplantier, J. M., Granjon, L., Kergoat, G. J., 2013. Correction: Mitochondrial and Nuclear Genes-Based Phylogeography of *Arvicanthis niloticus* (Murinae) and Sub-Saharan Open Habitats Pleistocene History. PloS one 8, 1-16.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology *7*, 214.

Drummond, A. J., Rambaut, A., Suchard, M., 2013. BEAST v1. 8.0 2002-2013. Bayesian Evolutionary Analysis Sampling Trees. Available at <http://Beast.Bio.Ed.Ac.Uk/>.

Fichet-Calvet, E., Lecompte, E., Veyrunes, F., Barrière, P., Nicolas, V., Koulémou, K., 2009. Diversity and dynamics in a community of small mammals in coastal Guinea, West Africa. Belgian Journal of Zoology 139, 93-102.

Gasse, F., 2006. Climate and hydrological changes in tropical Africa during the past million years. Comptes rendus PALEVOL *5*(1), 35-43.

Grubb, P., Sandrock, O., Kullmer, O., Kaiser T.M., Schrenk, F., 1999. Relationships between eastern and southern mammal faunas. African biogeography, climate change, and human evolution (Bromage GT and Schrenk F, eds) Oxford University Press, Oxford, United Kingdom, 253-267.

Guindon, S., Gascuel, O., 2003. PhyML :"A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood." Systematic Biology 52(5), 696-704.

Haffer, J., 1982. General aspects of the refuge theory. Biological diversification in the tropics (ed. By G.T. Prance), pp. 6.-24. Colombia University Press, New York. In: Nicolas, V., Missoup, A. D., Denys, C., Kerbis Peterhans, J., Katuala, P., Couloux, A., Colyn, M. (2011). The roles of rivers and Pleistocene refugia in shaping genetic diversity in Praomys misonnei in tropical Africa. Journal of Biogeography 38(1), 191-207.

Hall, T. A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic acids symposium series 41, 95-98.

Happold, D.C.D. and Happold, M., 1989. Biogeography of montane small mammals in Malawi, Central Africa. Journal of Biogeography 16, 353 – 367.

Hanney, P., 1965. The Muridae of Malawi (Africa: Nyasaland). Journal of Zoology 146, 577-633.

Hill, J. E., Carter, T. D., 1941. The mammals of Angola, Africa. Bulletin of the American Museum of Natural History 78, 85.

Huson, D. H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Molecular biology and evolution 23(2), 254-267.

Irwin, D.M., Kocher, T.D., Wilson, A.C., 1991. Evolution of the cytochrome b gene of mammals. Journal of Molecular Evolution 32, 128–144.

Irwin, D. E., Brelsford, A., Toews, D. P., MacDonald, C., Phinney, M., 2009. Extensive hybridization in a contact zone between MacGillivray's warblers Oporornistolmiei and mourning warblers O. philadelphia detected using molecular and morphological analyses. Journal of Avian Biology 40(5), 539-552.

Jacobs, L. L., Flynn L. J., 2005. Of mice… again: the Siwalik rodent record, murine distribution, and molecular clocks. Interpreting the Past: Essays on Human, Primate, and Mammal Evolution in Honor of David Pilbeam, 63-80.

Jansa, S. A., Weksler, M., 2004. Phylogeny of muroid rodents: relationships within and among major lineages as determined by IRBP gene sequences.Molecular phylogenetics and evolution 31(1), 256-276.

Kaleme, P.K., Bates, J., Peterhans, J.K., Jacques, M.M., Ndara, B.R., 2007. Small mammal diversity and habitat requirements in the Kahuzi-Biega National Park and surrounding areas, eastern Democratic Republic of Congo. Integrative Zoology 2, 239-246.

Kan Kouassi, S., Nicolas, V., Aniskine, V., Lalis, A., Coloux, A., Colyn, M., Dosso, M., Koivogui, L., Verheyen, E., Akoua-Koffi, C., Denys, C., 2008. Taxonomy and biogeography of the African Pygmy mice, Subgenus *Nannomys* (Rodentia, Murinae, *Mus*) in Ivory Coast and Guinea (West Africa). Mammalia 72, 237-252.

Kennedy, P., Nachman, M. W., 1998. Deleterious mutations at the mitochondrial ND3 gene in South American marsh rats (Holochilus). Genetics 150(1), 359-368.

Kingdon, J., 1989. Island Africa: The Evolution of Africa's Rare Animals and Plants.

Lamb, J., Dows, S., Eiseb, S., Taylor, P.J., 2014. Increased geographic sampling reveals considerable new genetic diversity in the morphologically conservative African Pygmy Mice (Genus *Mus*; Subgenus *Nannomys*). Mammalian Biology 79, 24-35.

Lavrenchenko, L.A., 2000. The mammals of the isolated Harenna Forest (southern Ethiopia): structure and history of the fauna. Bonner Zoologische Monographien 46, 223-231.

Lecompte, E., Denys, C., Granjon, L., Volobouev, V., 2003. Integrative systematics: contributions to Mastomys phylogeny and evolution. Aciar Monograph series 96, 536-540.

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. Evolutionary Biology 8, 199.

Leirs, H., Mills, J. N., Krebs, J. W., Childs, J. E., Akaibe, D., Woollen, N., Ksiazek, T. G., 1999. Search for the Ebola virus reservoir in Kikwit, Democratic Republic of the Congo: reflections on a vertebrate collection. Journal of Infectious Diseases 179 (Supplement 1), 155-163.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11), 1451-1452.

Lorenzen, E.D., Heller, R., Siegismund, H.R., 2012. Comparative phylogeography of African savannah ungulates. Molecular ecology 21(15), 3656–3670.

Maley, J., 1996. The African rain forest – main characteristics of changes in vegetation and climate from the Upper Cretaceous to Quaternary. Proceedings of the Royal Society of Edinburgh 104B, 31-73.

Map of Life [online]. Available at: http://mol.org/. Downloaded on 04 December 2014.

Masembe, C., Muwanika, V. B., Nyakaana, S., Arctander, P., Siegismund, H. R. (2006). Three genetically divergent lineages of the Oryx in eastern Africa: Evidence for an ancient introgressive hybridization. Conservation genetics *7*(4), 551-562.

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

Mulungu, L.S., Makudi, R.H., Massawe, A.W., Machang'u, R.S., Mbije, N.E., 2008. Diversity and distribution of rodent and shrew species associated with variations in altitude on Mount Kilimanjaro, Tanzania. Mammalia 72, 178-185.

Musser, G.G., Carleton, M.D., 2005. Superfamily Muroidea. Mammal Species of the World a Taxonomic and Geographic Reference. D.E. Wilson and D.M. Reeder eds. John Hopkins University Press, Baltimore, 894-1531

Mein, P., Pickford M., Senut, B., 2004. Late Miocene micromammals from the Harasib karst deposits, Namibia. Part 2b—Cricetomyidae, Dendromuridae and Muridae, with an addendum on the Myocricetodontinae. Communs. Geol. Surv Namibia 13, 43–61.

Nabholz, B., Glémin, S., Galtier, N., 2008. Strong variation of mitochondrial mutation rate across mammals – the longetivity hypothesis. Mol. Biol. Evol. 25(1), 120-130.

Nicolas, V., Mboumba, J-F., Verheyen, E., Denys, C., Lecompte, E., Olayemi, A., Missoup, A.D., Katuala, P., Colyn, M., 2008. Phylogeographic structure and regional history of *Lemniscomys striatus* (Rodentia: Muridae) in tropical Africa. Journal of Biogeography 35, 2074-2089.

Nicolas, V., Missoup, A. D., Denys, C., Kerbis Peterhans, J., Katuala, P., Couloux, A., Colyn, M., 2011. The roles of rivers and Pleistocene refugia in shaping genetic diversity in Praomys misonnei in tropical Africa. Journal of Biogeography 38(1), 191-207.

Olson, D.M. et al., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. BioScience, 51(11), p.933. Available at: http://bioscience.oxfordjournals.org/content/51/11/933.full. Downloaded on 04 December 2014.

Osmaston, H. A., Mitchell, W. A., Osmaston, J. A., 2005. Quaternary glaciation of the Bale Mountains, Ethiopia. Journal of Quaternary Science 20(6), 593-606.

Potts, R., 2013. Hominin evolution in settings of strong environmental variability. Quaternary Science Reviews 73, 1-13.

Rambaut, A., Drummond, A.J., 2007. TRACER v1.5. Available at: http://beast.bio.ed.ac.uk/software/tracer/.

Rambaut A., 2012. FigTree, version 1.4.0. Available at: http://tree.bio.ed.ac.uk/software/figtree/.

Robbins, L. W., Baker, R. J., 1978. Karyotypic data for African mammals, with a description of an in vivo bone marrow technique. Bulletin of Carnegie Museum of Natural History 6, 188-210.

Russo, I.R.M., Chimimba, C.T., Bloomer, P., 2006. Mitochondrial DNA differentiation between two species of Aethomys (Rodentia: Muridae) from southern Africa. Journal of Mammalogy 87(3), 545–553.

Russo, I., Chimimba, C., Bloomer, P., 2010. Bioregion heterogeneity correlates with extensive mitochondrial DNA diversity in the Namaqua rock mouse, Micaelamys namaquensis(Rodentia: Muridae) from southern Africa- evidence for a species complex. BMC Evolutionary Biology 10(1), 307.

Schenk, J. J., Rowe, K. C., Steppan, S. J., 2013. Ecological opportunity and incumbency in the diversification of repeated continental colonizations by muroid rodents. Systematic biology 62(6), 837-864.

Shaw, K. L., 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proceedings of the National Academy of Sciences 99(25), 16122-16127.

Stanhope, M. J., Czelusniak, J., Si, J. S., Nickerson, J., Goodman, M., 1992. A molecular perspective on mammalian evolution from the gene encoding interphotoreceptor retinoid binding protein, with convincing evidence for bat monophyly. Molecular phylogenetics and evolution 1(2), 148-160.

Stankiewicz, J., de Wit, M.J., 2006. A proposed drainage evolution model for Central Africa—Did the Congo flow east? Journal of African Earth Sciences 44(1), 75–84.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution 30, 2725-2729.

Taylor, P.J. et al., 2009. Speciation mirrors geomorphology and palaeoclimatic history in African laminate-toothed rats (Muridae: Otomyini) of the *Otomys denti* and *Otomys lacustris* species-complexes in the “Montane Circle” of East Africa. Biological Journal of the Linnean Society 96(4), 913–941.

Thomas, O., 1909. New African mammals. The annals and magazine of Natural History 4, 542-549.

Trauth, M.H., Larrasoaña, J.C., Mudelsee, M., 2009. Trends, rhythms and events in Plio-Pleistocene African climate. Quaternary Science Reviews 28, 399-411.

Trauth, M.H., Maslin, M.A., Deino, A.L., Strecker, M.R., Bergner, A.G.N, Dühnfort, M., 2007. High- and low-latitude forcing of Plio-Pleistocene East African climate and human evolution. Journal of Human Evolution 53, 475-486.

Veyrunes, F., Britton-Davidian, J., Robinson, T.J., Calvet, E., Denys, C., Chevret, P., 2005. Molecular phylogeny of the African pygmy mice, subgenus *Nannomys* (Rodentia, Murinae, *Mus*): Implications for chromosomal evolution. Molecular Phylogenetics and Evolution 36, 358-369.

Villesen, P., 2007, FaBox: an online fasta sequence toolbox, http://www.birc.au.dk/software/fabox.

Weksler, M. (2003). Phylogeny of Neotropical oryzomyine rodents (Muridae: Sigmodontinae) based on the nuclear IRBP exon. Molecular Phylogenetics and Evolution 29(2), 331-349.

Wickliffe, J.K., Hoffmann, F.G., Carroll, D.S., Dunina-Barkovskaya, Y.V., Bradley, R.D., Baker, R.J., 2003. Intron 7 (FGB-I7) of the fibrinogen, B beta polypeptide (FGB): a nuclear DNA phylogenetic marker for mammals. Occas. Pap. Mus. Texas Tech Univ. No. 219, 1–6.

Winkler, A.K., 2002. Neogene paleobiogeography and East African paleoenvironments: contributions from the Tungen Hills rodents to lagomorphs. J. Human Evol. 42 (1/2), 237-256.

Yalden, D. W., 1988. Small mammals of the Bale Mountains, Ethiopia. In: Lavrenchenko, L.A., 2000. The mammals of the isolated Harenna Forest (southern Ethiopia): structure and history of the fauna. Bonner Zoologische Monographien 46, 223-231.

Yalden, D. W., Largen, M. J., Kock, D., Hillman, J. C., 1996. Catalogue of the mammals of Ethiopia and Eritrea, 7. Revised checklist, zoogeography and conservation. Tropical Zoology 9(1), 73-164.

Appendix

Appendix 1. List of all specimen used in this study. TEOW – Habitat.identified from Terrestrial Ecoregions of the World database.

ID	Lineage	Database	Cyt b	IRBP	β-fib7	Locality	Longitude	Latitude	Alt (m)	TEOW
KF484851	A	(GB) KF484851				KahuziBiega (DRC)	28.019	-1.958	1230	Northeastern Congolian lowland forests
AR104	A	(AR) NKR28176	Hap_20			Aberdares (Kenya)	36.700	0.317	1825	N Acacia-Commiphora bushlands, thickets
AR113	A	(AR) RW5368	Hap_26			Kigali (Rwanda)	30.059	-1.950	1362	Victoria Basin forest-savanna mosaic
KE468	A		Hap_34			Mt Elgon NP (Kenya)	34.778	1.031	2839	Victoria Basin forest-savanna mosaic
KE836	A				Hap_2	Nyahururu (Kenya)	36.373	0.045	2288	East African montane forests
KE857	A			Hap_2	Hap_3	Nyahururu (Kenya)	36.373	0.045	2288	East African montane forests
KE913	A			Hap_2	Hap_3	Nyahururu (Kenya)	37.143	0.032	2288	East African montane forests
KE915	A			Hap_2	Hap_3	Nyahururu (Kenya)	37.143	0.032	2288	East African montane forests
KE923	A					Nanyuki farm (Kenya)	37.143	0.032	2026	East African montane forests
KE933	A					Nanyuki farm (Kenya)	37.143	0.032	2026	East African montane forests
AR117	B	(AR) 7	Hap_27			Mbwambala (DR Congo)	18.909	-5.058	448	Southern Congolian forest-savanna mosaic
AR22	B	(AR) 13				Mbwambala (DR Congo)	18.909	-5.058	448	Southern Congolian forest-savanna mosaic
AR105	C	(AR) 10866	Hap_21			Mufindi (Tanzania)	35.367	-8.617	1969	Eastern Arc forests
AR106	C	(AR) 10869	Hap_17			Mufindi (Tanzania)	35.367	-8.617	1969	Eastern Arc forests
T8_534	C		Hap_17	Hap_38		Ulangambi Forest Reserve (Tanzania)	35.914	-8.007	2068	Eastern Arc forests
T8_537	C			Hap_34	Hap_20	Ulangambi Forest Reserve (Tanzania)	35.914	-8.007	2068	Eastern Arc forests
T8_538	C		Hap_18	Hap_34	Hap_21	Ulangambi Forest Reserve (Tanzania)	35.914	-8.007	2068	Eastern Arc forests
T8_541	C		Hap_19	Hap_39	Hap_21	Ulangambi Forest Reserve (Tanzania)	35.914	-8.007	2068	Eastern Arc forests
TA493	C		Hap_41	Hap_42		Mt Ngozi, Poroto Range (Tanzania)	33.571	-9.035	1937	South Rift montane forest-grassland mosaic
TA494	C		Hap_42	Hap_22		Mt Ngozi, Poroto Range (Tanzania)	33.571	-9.035	1937	South Rift montane forest-grassland mosaic
10268	C		Hap_44			Mamiwa, Ukaguru (Tanzania)	36.917	-6.367	2177	Eastern Miombo woodlands
RS1283	D1	(GB) KF484843	Hap_2		Hap_16	Kapyshia Hot Springs (Zambia)	31.600	-11.170	1352	Central Zambezian Miombo woodlands
KF484845	D1	(GB) KF484844	Hap_2			Huambo (Angola)	15.734	-12.767	1683	Angolan Miombo woodlands
KF484843	D1	(GB) KF484845	Hap_2			Huambo (Angola)	15.734	-12.767	1683	Angolan Miombo woodlands
RS0785	D1			Hap_19	Hap_12	Nchila Wildlife Reserve (Zambia)	24.317	-11.267	1405	Central Zambezian Miombo woodlands
RS1222	D1		Hap_1		Hap_13	Mutinondo Wilderness (Zambia)	31.323	-12.392	1457	Central Zambezian Miombo woodlands
RS1223	D1		Hap_1	Hap_20		Mutinondo Wilderness (Zambia)	31.323	-12.392	1457	Central Zambezian Miombo woodlands
RS1225	D1		Hap_1			Mutinondo Wilderness (Zambia)	31.323	-12.392	1457	Central Zambezian Miombo woodlands
RS1226	D1		Hap_1	Hap_19		Mutinondo Wilderness (Zambia)	31.323	-12.392	1457	Central Zambezian Miombo woodlands
RS1265	D1					Kapyshia Hot Springs (Zambia)	31.600	-11.170	1352	Central Zambezian Miombo woodlands
RS1269	D1		Hap_1	Hap_21	Hap_15	Kapyshia Hot Springs (Zambia)	31.600	-11.170	1352	Central Zambezian Miombo woodlands
RS1270	D1		Hap_1	Hap_22		Kapyshia Hot Springs (Zambia)	31.600	-11.170	1352	Central Zambezian Miombo woodlands

KF484844	D1		Hap_2			Huambo (Angola)	15.734	-12.767	1683	Angolan Miombo woodlands
AR111	D2	(AR) 7372	Hap_24			Gerodom (Tanzania)	35.383	-4.467	1774	East African montane forests
AR111	D2	(AR) 7372				Gerodom (Tanzania)	35.383	-4.500	1774	East African montane forests
AR112	D2	(AR) 9308	Hap_25			Magangwe (Tanzania)	34.233	-7.767	1391	Central Zambezian Miombo woodlands
AR112	D2	(AR) 9308				Magangwe (Tanzania)	34.250	-7.750	1391	Central Zambezian Miombo woodlands
TA172	D2		Hap_40	Hap_37		Kyelele (Tanzania)	30.806	-1.278	1470	Victoria Basin forest-savanna mosaic
TA173	D2		Hap_40	Hap_24		Kyelele (Tanzania)	30.806	-1.278	1470	Victoria Basin forest-savanna mosaic
TA174	D2		Hap_40	Hap_40		Kyelele (Tanzania)	30.806	-1.278	1470	Victoria Basin forest-savanna mosaic
TA185	D2		Hap_40			Kyelele (Tanzania)	30.806	-1.278	1470	Victoria Basin forest-savanna mosaic
TA186	D2		Hap_40	Hap_24		Kyelele (Tanzania)	30.806	-1.278	1470	Victoria Basin forest-savanna mosaic
TA223	D2		Hap_40	Hap_24		Kelebe (Tanzania)	31.208	-2.964	1282	Central Zambezian Miombo woodlands
M5_097	D3		Hap_29			Mpalaganga estate (Malawi)	35.261	-15.447	960	South Malawi montane forest-grassland mosaic
M5_176	D3		Hap_35			Mpalaganga estate (Malawi)	35.261	-15.447	960	South Malawi montane forest-grassland
M8_0098	D3		Hap_29		Hap_4	Mulanje Mts (Malawi)	35.705	-15.849	2148	South Malawi montane forest-grassland
M8_0128	D3		Hap_29	Hap_3	Hap_4	Ndawambe (Malawi)	32.823	-13.675	1328	Southern Miombo woodlands
M8_0129	D3					Ndawambe (Malawi)	32.823	-13.675	1328	Southern Miombo woodlands
M8_0132	D3					Ndawambe (Malawi)	32.823	-13.675	1328	Southern Miombo woodlands
M8_0136	D3					Ndawambe (Malawi)	32.823	-13.675	1328	Southern Miombo woodlands
M8_0137	D3			Hap_4		Ndawambe (Malawi)	32.823	-13.675	1328	Southern Miombo woodlands
M8_0138	D3		Hap_29			Ndawambe (Malawi)	32.823	-13.675	1328	Southern Miombo woodlands
M8_0154	D3					Ndawambe (Malawi)	32.823	-13.675	1226	Southern Miombo woodlands
M8_0155	D3			Hap_5	Hap_5	Ndawambe (Malawi)	32.823	-13.675	1226	Southern Miombo woodlands
M8_0193	D3			Hap_6		Ntchisi (Malawi)	34.003	-13.381	1440	Central Zambezian Miombo woodlands
M8_0194	D3				Hap_6	Ntchisi (Malawi)	34.003	-13.381	1440	Central Zambezian Miombo woodlands
M8_0195	D3		Hap_29			Ntchisi (Malawi)	34.003	-13.381	1440	Central Zambezian Miombo woodlands
M8_0196	D3		Hap_30		Hap_4	Ntchisi (Malawi)	34.003	-13.381	1440	Central Zambezian Miombo woodlands
M8_0197	D3			Hap_7	Hap_4	Ntchisi (Malawi)	34.003	-13.381	1440	Central Zambezian Miombo woodlands
M8_0198	D3			Hap_8		Ntchisi (Malawi)	34.003	-13.381	1440	Central Zambezian Miombo woodlands
M8_0240	D3			Hap_9	Hap_4	Perekezi Forest Reserve (Malawi)	33.650	-12.017	1495	Southern Rift montane forest-grassland
M8_0241	D3			Hap_10	Hap_7	Perekezi Forest Reserve (Malawi)	33.650	-12.017	1495	Southern Rift montane forest-grassland
M8_0242	D3		Hap_31	Hap_11		Perekezi Forest Reserve (Malawi)	33.650	-12.017	1495	Southern Rift montane forest-grassland
M8_0248	D3		Hap_31	Hap_12	Hap_8	Perekezi Forest Reserve (Malawi)	33.650	-12.017	1495	Southern Rift montane forest-grassland
M8_0249	D3			Hap_13	Hap_8	Perekezi Forest Reserve (Malawi)	33.650	-12.017	1495	Southern Rift montane forest-grassland
M8_0262	D3					Perekezi forest reserve (Malawi)	33.650	-12.017	1495	Southern Rift montane forest-grassland
M8_0263	D3			Hap_12		Perekezi Forest Reserve (Malawi)	33.650	-12.017	1495	Southern Rift montane forest-grassland

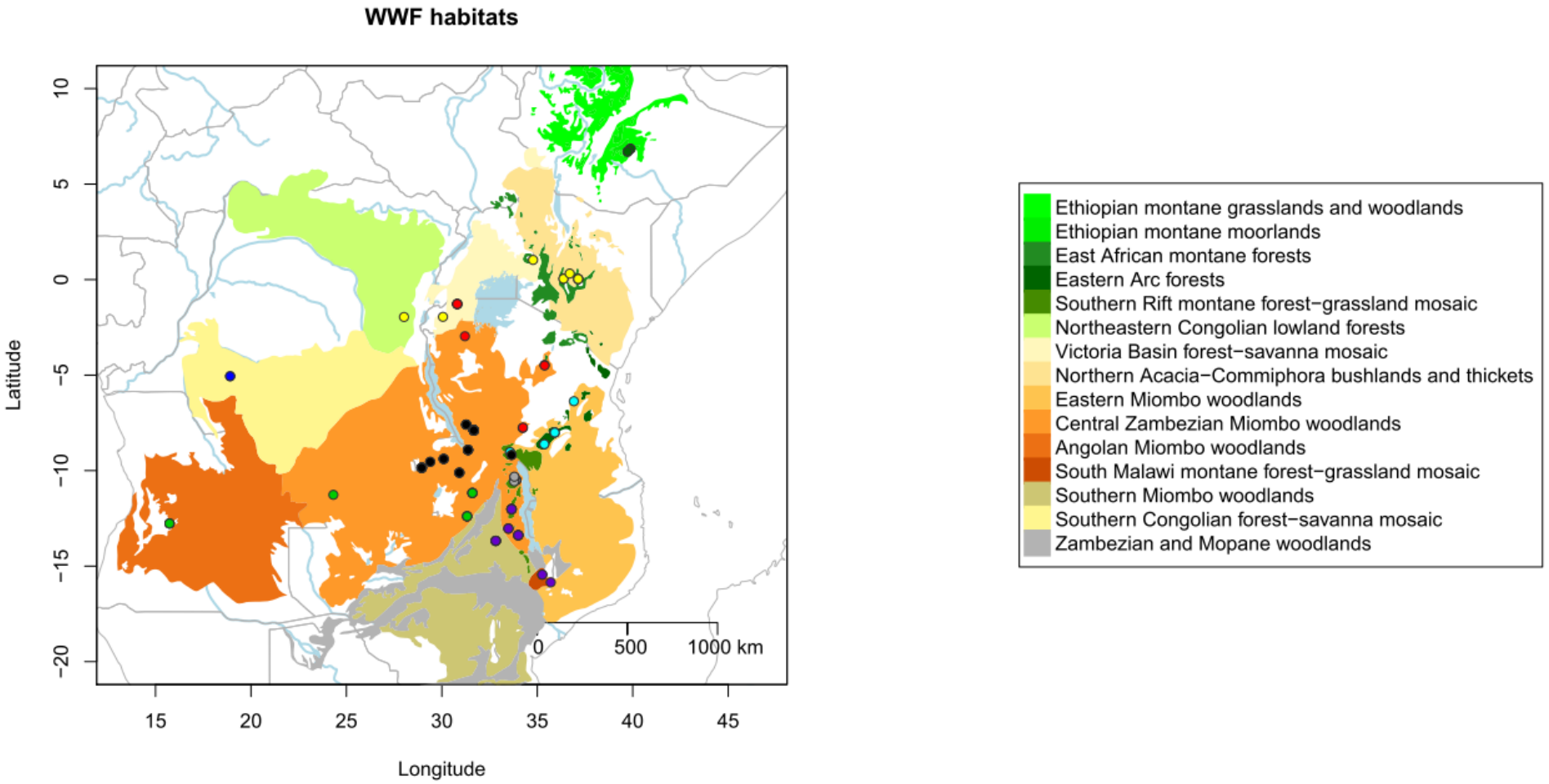
M8_0267	D3			Hap_14	Hap_9	Perekezi Forest Reserve (Malawi)	33.650	-12.017	1495	Southern Rift montane forest-grassland
M8_0268	D3			Hap_11	Hap_4	Perekezi Forest Reserve (Malawi)	33.650	-12.017	1495	Southern Rift montane forest-grassland
M8_0269	D3			Hap_10		Perekezi Forest Reserve (Malawi)	33.650	-12.017	1495	Southern Rift montane forest-grassland
M8_0277	D3			Hap_15		Viphya Plateau (Malawi)	33.485	-13.033	1740	Central Zambezian Miombo woodlands
M8_0278	D3					Viphya Plateau (Malawi)	33.485	-13.033	1740	Central Zambezian Miombo woodlands
M8_0280	D3		Hap_32			Viphya Plateau (Malawi)	33.485	-13.033	1740	Central Zambezian Miombo woodlands
AR129	D4	(AR) TM41758				Chelinda (Malawi)	33.800	-10.317	1436	Southern Rift montane forest-grassland
M8_3014	D4			Hap_16		Nyika NP (Malawi)	33.886	-10.496	2099	Southern Rift montane forest-grassland
M8_3015	D4			Hap_17		Nyika NP (Malawi)	33.886	-10.496	2099	Southern Rift montane forest-grassland
M8_3019	D4		Hap_36			Nyika NP (Malawi)	33.886	-10.496	2099	Southern Rift montane forest-grassland
M8_3026	D4			Hap_18	Hap_10	Nyika NP (Malawi)	33.886	-10.496	2099	Southern Rift montane forest-grassland
M8_3035	D4		Hap_37			Nyika NP (Malawi)	33.886	-10.496	2099	Southern Rift montane forest-grassland
M8_3040	D4			Hap_12	Hap_11	Nyika NP (Malawi)	33.799	-10.371	1667	Southern Rift montane forest-grassland
M8_3051	D4		Hap_33			Nyika NP (Malawi)	33.799	-10.371	1667	Southern Rift montane forest-grassland
M8_3052	D4					Nyika - Jallawe valley (Malawi)	33.799	-10.371	1667	Southern Rift montane forest-grassland
M8_3058	D4					Nyika - Jallae camp 2 (Malawi)	33.799	-10.371	2171	Southern Rift montane forest-grassland
M8_3067	D4					Nyika - air strip (Malawi)	33.790	-10.566	2350	Southern Rift montane forest-grassland
AR110	D5	(AR) 6186	Hap_23			Suma (Tanzania)	33.650	-9.183		Southern Rift montane forest-grassland
AR107	D5	(AR) 13023	Hap_16			Mbisi (Tanzania)	31.667	-7.867		Southern Rift montane forest-grassland
AR108	D5	(AR) 13025	Hap_22			Mbisi (Tanzania)	31.667	-7.867		Southern Rift montane forest-grassland
RS1368	D5			Hap_23	Hap_17	Uningi Pans (Zambia)	31.369	-8.917	1686	Central Zambezian Miombo woodlands
RS1369	D5			Hap_24		Uningi Pans (Zambia)	31.369	-8.917	1686	Central Zambezian Miombo woodlands
RS1370	D5			Hap_25		Uningi Pans (Zambia)	31.369	-8.917	1686	Central Zambezian Miombo woodlands
RS1375	D5		Hap_3	Hap_24		Uningi Pans (Zambia)	31.369	-8.917	1686	Central Zambezian Miombo woodlands
RS1414	D5			Hap_26	Hap_18	Chishimba Falls (Zambia)	30.918	-10.108	1320	Central Zambezian Miombo woodlands
RS1415	D5		Hap_4	Hap_27		Chishimba Falls (Zambia)	30.918	-10.108	1320	Central Zambezian Miombo woodlands
RS1447	D5			Hap_26		Kapuma Falls (Zambia)	30.095	-9.388	1368	Central Zambezian Miombo woodlands
RS1448	D5		Hap_5	Hap_28		Kapuma Falls (Zambia)	30.095	-9.388	1368	Central Zambezian Miombo woodlands
RS1474	D5		Hap_6	Hap_29	Hap_19	Lumangwe Falls (Zambia)	29.388	-9.541	1326	Central Zambezian Miombo woodlands
RS1525	D5		Hap_7	Hap_30		Ntumbachushi Falls (Zambia)	28.944	-9.853	1142	Central Zambezian Miombo woodlands
RS1528	D5		Hap_8	Hap_29		Ntumbachushi Falls (Zambia)	28.944	-9.853	1142	Central Zambezian Miombo woodlands
RS1530	D5			Hap_31		Ntumbachushi Falls (Zambia)	28.944	-9.853	1142	Central Zambezian Miombo woodlands
RS1546	D5		Hap_10	Hap_26		Ntumbachushi Falls (Zambia)	28.944	-9.853	1142	Central Zambezian Miombo woodlands
RS1548	D5		Hap_7	Hap_26		Ntumbachushi Falls (Zambia)	28.944	-9.853	1142	Central Zambezian Miombo woodlands
RS1549	D5		Hap_7	Hap_24		Ntumbachushi Falls (Zambia)	28.944	-9.853	1142	Central Zambezian Miombo woodlands
T8_441	D5		Hap_12	Hap_32		Mbizi Hill Forest Reserve (Tanzania)	31.679	-7.893	2203	Southern Rift montane forest-grassland

T8_442	D5	Hap_13	Hap_29		Mbizi Hill Forest Reserve (Tanzania)	31.679	-7.893	2203	Southern Rift montane forest-grassland
T8_443	D5	Hap_14	Hap_26		Mbizi Hill Forest Reserve (Tanzania)	31.679	-7.893	2203	Southern Rift montane forest-grassland
T8_447	D5	Hap_15	Hap_26		Mbizi Hill Forest Reserve (Tanzania)	31.679	-7.893	2203	Southern Rift montane forest-grassland
TA450	D5	Hap_22	Hap_41		Chala (Tanzania)	31.273	-7.590	953	Central Zambezian Miombo woodlands
TA451	D5	Hap_22			Chala (Tanzania)	31.273	-7.590	953	Central Zambezian Miombo woodlands
T8_454	D5	Hap_16	Hap_33	Hap_19	Mbizi Hill Forest Reserve (Tanzania)	31.679	-7.893	2203	Southern Rift montane forest-grassland
T8_455	D5	Hap_16	Hap_34		Mbizi Hill Forest Reserve (Tanzania)	31.679	-7.893	2203	Southern Rift montane forest-grassland
T8_456	D5		Hap_34	Hap_19	Mbizi Hill Forest Reserve (Tanzania)	31.679	-7.893	2203	Southern Rift montane forest-grassland
T8_464	D5		Hap_35		Mbizi Hill Forest Reserve (Tanzania)	31.679	-7.893	2203	Southern Rift montane forest-grassland
T8_465	D5		Hap_36		Mbizi Hill Forest Reserve (Tanzania)	31.679	-7.893	2203	Southern Rift montane forest-grassland
T8_466	D5	Hap_16	Hap_37		Mbizi Hill Forest Reserve (Tanzania)	31.679	-7.893	2203	Southern Rift montane forest-grassland
ETH0186	Harenna	Hap_38			Bale NP, Harrena forest (Ethiopia)	39.881	6.850	2338	Eth montane grasslands and woodlands
ETH0211	Harenna	Hap_39	Hap_1	Hap_1	Bale NP, Harrena forest (Ethiopia)	39.881	6.850	2338	Eth montane grasslands and woodlands
LAV2274	Harenna	Hap_43			Harrena forest (Ethiopia)	39.733	6.700	2541	Ethiopian montane moorlands
LAV2283	Harenna	Hap_45			Harrena forest (Ethiopia)	39.733	6.700	2541	Ethiopian montane moorlands

Appendix 2. Sequences used as outgroups in ML and Bayesian analysis of separate genes.

Gene	Species	GenBank Accession number
cyt b	*Mus minutoides*	AY057816
	Mus setulosus	AJ875083
	Mus mattheyii	AJ875069
IRBP	*Mus minutoides*	AJ875087
	Mus setulosus	AJ875088
	Mus mattheyii	AJ698889
β-fibint 7	*Mus musculus*	EF605469
	Praomys jacksoni	JN636419
	Praomys degraaffi	JN636367

Appendix 3. Habitats according to Olson et al. 2001 and positions of the mitochondrial lineages. Group Harenna – dark green dots, group A – yellow dots, group B – dark blue dots, group C – light blue dots, group D1 – light green dots, group D2 – red dots, group D3 – dark purple dots, group D4 – grey dots, group D5 – black dots.


WWF habitats
Latitude
Longitude
10
5
0
−5
−10
−15
−20
15
20
25
30
35
40
45
0
500
1000 km
Ethiopian montane grasslands and woodlands
Ethiopian montane moorlands
East African montane forests
Eastern Arc forests
Southern Rift montane forest−grassland mosaic
Northeastern Congolian lowland forests
Victoria Basin forest−savanna mosaic
Northern Acacia−Commiphora bushlands and thickets
Eastern Miombo woodlands
Central Zambezian Miombo woodlands
Angolan Miombo woodlands
South Malawi montane forest−grassland mosaic
Southern Miombo woodlands
Southern Congolian forest−savanna mosaic
Zambezian and Mopane woodlands

Appendix 4. Subspecies and synonyms of *M. triton*.

Form name	Locality	Reference
Leggada naivashae	Naivasha, Aberdare Mts (Kenya)	Heller (1910)
Leggada triton murilla	Machakos (Kenya)	Thomas (1910)
Leggada fors	Butagu Valley, Ruwenzori (DRC)	Thomas and Wroughton (1910)
Mus triton birungensis	Kivu region (DRC)	Lönnberg and Gyldenstolpe (1925)
Mus triton imatongensis	Imatong mountains (Sudan)	Setzer (1953)

Appendix 5. List of sequences used for the divergence dating analysis in BEAST. GenBank accession codes of cytochrome b sequences are shown.

Taxon	GenBank Accession number
Aethomys	AJ604526
Apodemus agrarius	AB096809
Apodemus argenteus	AB032848
Apodemus draco	AB109397
Apodemus flavicollis	AB032853
Apodemus gurkha	AB032852
Apodemus latronum	AB096834
Apodemus mystacinus	AJ748237
Apodemus speciosus	AB032849
Apodemus witherbyi	AB303228
Arvicanthis neumanni	EU349737
Arvicanthis niloticus	AF004569
Colomys	AF518372
Dasymys	AF141217
Golunda	AM408338
Grammomys	AM408345
Heimyscus	AF518332
Hylomyscus parvus	AF518329
Hylomyscus stella	AF518331
Lemniscomys rosalia	AF141209
Lemniscomys striatus	AF141210
Malacomys edwardsi	DQ022379
Malacomys longipes	DQ022380
Mastomys coucha	AF518334
Mastomys erythroleucus	AM409395
Mastomys pernanus	AF518343
Micaelamys	AF141215
Mus cookii	AB125767
Mus crociduroides	AJ698878

Mus musculus	V00711
Mus pahari	AY057814
Mus platythrix	AJ698880
Myomyscus brockmani	AF518352
Myomyscus yemeni	AF518356
Oenomys	EU349769
Otomys angoniensis	AM408343
Otomys denti	EU874449
Parotomys	AF141224
Stenocephalemys albipes	AF518346
Stenocephalemys albocaudata	AF518369
Stochomys	EU292149
Tokudaia	AB033703
Zelotomys	AF518375

Appendix 6. Sequences used as outgroups in ML and Bayesian analysis of concatenated dataset of cyt b, IRBP and β-fibint 7.

Gene	Species	GenBank Accession number
cyt b	*Mus musculus*	AB205311
	Praomys jacksoni	AF518361
	Praomys degraaffi	AF518359
IRBP	*Mus musculus*	JX457616
	Praomys jacksoni	KC953443
	Praomys degraaffi	DQ022410
β-fibint 7	*Mus musculus*	EF605469
	Praomys jacksoni	JN636419
	Praomys degraaffi	JN636367