This item is the archived peer-reviewed author-version of:

Diversity and evolution of African Grass Rats (Muridae: Arvicanthis) : from radiation in East Africa to repeated colonization of northwestern and southeastern savannas

Reference:

Bryja Josef, Colangelo Paolo, Lavrenchenko Leonid A., Meheretu Yonas, Šumbera Radim, Bryjová Anna, Verheyen Erik K., Leirs Herwig, Castiglia Riccardo.- Diversity and evolution of African Grass Rats (Muridae: Arvicanthis) : from radiation in East Africa to repeated colonization of northwestern and southeastern savannas
Journal of zoological systematics and evolutionary research - ISSN 0947-5745 - 57:4(2019), p. 970-988
Full text (Publisher's DOI): <https://doi.org/10.1111/JZS.12290>
To cite this reference: <https://hdl.handle.net/10067/1601280151162165141>

Title: Diversity and evolution of African Grass Rats (Muridae: *Arvicanthis*) - from radiation in East Africa to repeated colonization of north-western and south-eastern savannahs

Short running title: Grass rats in sub-Saharan savannas

Josef Bryja^{1,2}, Paolo Colangelo³, Leonid A. Lavrenchenko⁴, Yonas Meheretu⁵, Radim Šumbera⁶, Anna Bryjová¹, Erik Verheyen^{7,8}, Herwig Leirs⁸, Riccardo Castiglia⁹

¹ Institute of Vertebrate Biology of the Czech Academy of Sciences, Brno, Czech Republic

² Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic

³ National Research Council, Institute of Agro-environmental and Forest Biology (CNR-IBAF), Rome, Italy

⁴ A.N.Severtsov Institute of Ecology and Evolution of the Russian Academy of Sciences, Moscow, Russia

⁵ Department of Biology and Institute of Mountain Research and Development, Mekelle University, Mekelle, Tigray, Ethiopia

⁶ Department of Zoology, Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

⁷ Royal Belgian Institute for Natural Sciences, Operational Direction Taxonomy and Phylogeny, Brussels, Belgium

⁸ Evolutionary Ecology Group, Biology Department, University of Antwerp, Antwerp, Belgium

⁹ Department of Biology and Biotechnology "Charles Darwin", Sapienza University of Rome, Rome, Italy

Prepared for: Journal of Zoological Systematics and Evolutionary Research as Original article

***Corresponding author:** 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

Word count: 10598 words incl. References, 7 Figures, 2 Tables

Keywords: taxonomy, Ethiopia, biogeography, reticulate evolution, tropical Africa

36 **Abstract**

37 African grass rats of the genus *Arvicanthis* Lesson, 1842, are one of the most important groups of
38 rodents in sub-Saharan Africa. They are abundant in a variety of open habitats, they are major
39 agricultural pests and they became a popular model in physiological research because of their diurnal
40 activity. Despite this importance, information about their taxonomy and distribution is
41 unsatisfactory, especially in Eastern Africa. In this study, we collected the most comprehensive
42 multilocus DNA dataset to date across the geographic and taxonomic range of the genus (229
43 genotyped specimens from 130 localities in 16 countries belonging to all currently recognized
44 species). We reconstructed phylogenetic relationships, mapped the distribution of major genetic
45 clades, and used the combination of cytogenetic, nuclear and mitochondrial markers for species
46 delimitations and taxonomic suggestions. The genus is composed of two major evolutionary groups,
47 called here the ANSORGEI and NILOTICUS groups. The former contains four presumed species, while
48 the latter is more diverse and we recognized nine species. Most relationships among species are not
49 resolved, which suggests a rapid radiation (dated to early-middle Pleistocene). Further, there is an
50 indication of reticulate evolution in Ethiopia, i.e. the region of the highest *Arvicanthis* diversity. The
51 distribution of genetic diversity suggests diversification in Eastern Africa, followed by repeated
52 dispersals to the west (Sudano-Guinean savannahs), and to the south (Masai steppe). We propose
53 nomenclatural changes for Ethiopian taxa and provide suggestions for future steps towards solving
54 remaining taxonomic questions in the genus.

55

Introduction

Despite their practical importance for human, *e.g.* as vectors of pathogens or agricultural pests, the general knowledge of rodents in sub-Saharan Africa is relatively limited. Surprisingly, this is often true also for abundant taxa, where even the basic information about their taxonomy and distribution is missing. One of the most successful groups of African murine rodents in open habitats is the genus *Arvicanthis* Lesson, 1842, which has a wide distribution ranging along the entire Nile Valley south to the Zambezi River, and from Senegal in the west to the Horn of Africa in the east. These rodents occur in a variety of ecosystems, from the lowland savannahs and dry coastal areas to the high altitude grasslands of the Ethiopian highlands. They are major agricultural pests, as their reproductive success (especially after annual dry-season depletion) makes them more successful than other competing rodents (Sicard, Diarra & Cooper, 1999). Because most species are diurnal and easy to breed, *Arvicanthis* became a popular model also in physiological research, especially in studies of circadian rhythms, vision and endocrinology (*e.g.* Castillo-Ruiz, Indic & Schwartz, 2018; Giancesini, Clesse, Tosini, Hicks & Laurent, 2015; Langel, Smale, Esquiva & Hannibal, 2015).

The genus *Arvicanthis* shows a remarkable degree of diversity and constitutes a typical historical example of the taxonomic issues caused by the subsequent activities of lumpers and splitters during the last century. For example, Allen (1939) listed 37 taxa within the genus, while Honaki, Kinman & Koepl (1982) considered the genus monotypic with the single species *A. niloticus*. The genus has been later on the subject of intense taxonomic, systematic and evolutionary studies. Numerous papers over three last decades have summarised the information so far collected through morphometric data (Baskevich & Lavrenchenko, 2000; Bekele, Capanna, Corti, Marcus & Schlitter, 1993; Corti & Fadda, 1996; Fadda & Corti, 2001) and karyotype analyses (Castiglia et al., 2003; Castiglia, Bekele, Makundi, Oguge & Corti, 2006; Civitelli, Castiglia, Codja & Capanna, 1995; Corti, Civitelli, Bekele, Castiglia & Capanna, 1995; Corti, Civitelli, Castiglia, Bekele & Capanna, 1996; Corti et al., 2005; Volobouev et al., 2002). Especially karyotypes were found very useful for understanding the evolution of these rodents (see review in Castiglia et al., 2006) helping significantly to delimit currently recognized species (*e.g.* Volobouev et al., 2002). In the most recent compendium of the Handbook of the Mammals of the World (Denys et al. 2017), seven species are recognized. However, the review of described karyotypes suggests cryptic diversity in the genus (Castiglia et al., 2006) and the high numbers of synonyms included in some taxa reflects a complex situation (Musser & Carleton, 2005). The definition and distribution limits of several species clearly require further investigation, and additional species may prove to be present especially in central and eastern Africa (Happold, 2013).

In contrast to cytogenetic and morphometric approaches, the use of DNA sequence data in phylogenetic studies of *Arvicanthis* is relatively scarce (Abdel Rahman Ahmed et al., 2008; Dobigny et al., 2011, 2013; Ducroz, Volobouev & Granjon, 1998, 2001). These studies are limited in number of markers, based mainly on mitochondrial sequences (mostly cytochrome *b*, *CYTB*), except Dobigny et al. (2013), who used one nuclear gene with very limited variability. They are also geographically biased to west-central Africa and only one clade, *A. niloticus* sensu lato, is analysed in wider geographical context (Dobigny et al., 2013). Other studies either addressed the taxonomic diversity in geographically limited regions (Abdel Rahman Ahmed et al., 2008 in Sudan; Dobigny et al., 2011 in northern Cameroon), or attempted to derive phylogenetic relationships for the genus with a very limited number of specimens collected throughout the distribution range of this genus (Ducroz et al., 1998, 2001). Available results suggest that most genetic diversity is located in Eastern Africa, which is in agreement with cytogenetic data (Castiglia et al., 2006). However, no detailed study based on DNA sequences from Eastern Africa have been published yet. As a result of relative morphological uniformity, general lack of DNA data, and especially complete absence of morphological studies of genotyped individuals, the species in the genus are delimited only vaguely, there have been numerous confusions and misunderstandings (e.g. misuse of species names; see Table 1), and the distributional limits for most taxa are not well defined.

For this study, we collected the most comprehensive multilocus dataset of DNA sequences to date across the geographic range of the genus with special focus on undersampled region of eastern Africa. With these data in hand, we had following aims. (1) We used *CYTB*-barcoded samples for realistic assessment of the distribution of main mitochondrial genetic clades, possibly corresponding to different species. (2) Using *CYTB* and six nuclear markers, we reconstructed phylogenetic relationships in the genus. (3) In areas of sympatric occurrence of multiple genetic clades (especially in Ethiopia that seems to be the region of the highest *Arvicanthis* diversity), the combination of nuclear and mitochondrial markers together with available karyotypes was employed for species delimitation, taxonomic suggestions, and definitions for future systematic research. (4) Data on pan-African distribution of the genetic diversity of the genus were used to propose evolutionary hypotheses for the Plio-Pleistocene changes of open habitats in sub-Saharan Africa.

Materials and methods

Sampling

This study is based on 229 specimens of *Arvicanthis* genotyped at *CYTB*, 184 of them were newly genotyped and 45 sequences were downloaded from GenBank (Appendix S1). Five additional individuals from Castiglia et al. (2006) were unambiguously assigned to particular clades based on

their karyotype and geographic distribution. Collected material sufficiently covers the known distribution of the genus (130 localities from 16 countries; Appendix S1) and it represents the most comprehensive available material of all currently recognized *Arvicanthis* species (sensu Denys et al., 2017). The tissue samples were stored in 96% ethanol or SO 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).

Genotyping

DNA was extracted by commercial kits and all samples were genotyped at mitochondrial gene for cytochrome *b* (*CYTB*) using the protocol from Bryja et al. (2014). Preliminary phylogenetic analyses of *CYTB* showed division of the genus into two main clades (see also Dobigny et al., 2013). Hereafter we will call these two major groups as the NILOTICUS and ANSORGEI groups according to two characteristic species within them (Dobigny et al., 2013; Volobouev et al. 2002). Because several mitochondrial lineages in the NILOTICUS group seem to have overlapping distribution in eastern Africa, we employed nuclear markers to assess whether they might represent separate gene pools (i.e. reproductively isolated species) and to reconstruct their nuclear phylogeny. We selected individuals representing all but one mitochondrial lineages from the NILOTICUS group (we were not able to get nuclear sequences of one genetic lineage; in total 32 individuals + three outgroups from the ANSORGEI group) and sequenced them at two nuclear exons (*IRBP*, *RAG*) and four nuclear introns (*WLS*, *DHCR*, *TRPV*, *SMO*). For more details on used markers, including primer sequences and PCR protocols, see Appendix S2.

Mitochondrial phylogeny and sequence divergence

the first step, we reconstructed mitochondrial phylogeny by Bayesian inference (BI) and maximum likelihood (ML) approaches. Sequences of *CYTB* were edited and aligned in Geneious 9.0.5 (Biomatters, Ltd.) producing a final alignment of 1140 bp. The Findmodel web application (<http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>) was used to identify the most appropriate substitution model. The Akaike information criterion (AIC), compared among 12 biologically relevant substitution models, revealed that the model best fitting the ingroup data was GTR + G. As outgroups, we used seven sequences of other murid rodents, including three from the tribe Arvicanthini, one from Otomyini, one from Murini, one from Rattini and one *Acomys* sequence from the sister subfamily Deomyinae (see Appendix S2). BI analysis (partitioned by codon position) was performed in MrBayes 3.2.6 (Ronquist & Huelsenbeck, 2003). Three heated and one cold chain were employed, and runs were initiated from random trees. Two independent runs were conducted with 10 million generations each and trees and parameters were sampled every 1000 generations.

Convergence was checked using Tracer 1.7 (Rambaut, Drummond, Xie, Baele & Suchard, 2014). For each run, the first 25% of sampled trees were discarded as burn-in. Bayesian posterior probabilities were used to assess branch support of the MCMC tree. ML analysis was performed using RAxML 8.2.8 (Stamatakis, 2014). We used GTRCAT model as suggested by authors of the programme. The robustness of the nodes was evaluated by the default bootstrap procedure with 1,000 replications. All phylogenetic analyses were run on CIPRES Science Gateway (Miller, Pfeiffer & Schwartz, 2010) and trees were edited in FigTree 1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>).

Genetic p -distances among mitochondrial clades (i.e. the number of base differences per *CYTB* site from averaging over all sequence pairs between groups) were calculated in MEGA 10.0.4 (Kumar, Stecher, Li, Knyaz & Tamura, 2018) using the same dataset as for mitochondrial phylogeny, i.e. 121 sequences, and pairwise deletion option. Standard error estimate(s) were obtained by a bootstrap procedure (1000 replicates).

Phylogeny reconstruction using nuclear markers

The number of genetic partitions in the concatenated alignment of six nuclear loci and the most suitable nucleotide substitution models were simultaneously estimated in PartitionFinder 2 (Lanfear, Frandsen, Wright, Senfeld & Calcott, 2016). Single gene trees were calculated by BI in MrBayes as described above using the best substitution models. Partitioned analyses of concatenated nuclear dataset (3926 bp) were performed in MrBayes (BI analysis specifying the best models for all four partitions identified by PartitionFinder; see Appendix S2) and RAxML (ML analysis with GTRCAT model for all partitions) using the specifications described above. In MrBayes, the fifth partition was defined for six parsimony informative indels encoded as standard morphological data (i.e. 1 or 0) and with substitution model specified by Nst=1 rates=gamma coding=variable.

The species tree based on sequences of nuclear markers was calculated under the fully Bayesian framework using the multi-species coalescent model implemented in StarBEAST 2 (Ogilvie, Bouckaert & Drummond, 2017), an extension of BEAST 2.5.0 (Bouckaert et al., 2014). Alignments for each of six genes were imported into BEAUti 2.5.0, where separate and unlinked substitution and tree models, but the same strict clock rate for all partitions, were defined. We set up the lognormal priors for birthRate.t:Species (with $M=4.0$ and $S=1.25$) and popMean ($M=-5.0$ and $S=1.2$), the Yule speciation model and linear Pop function with constant root. Two independent runs were carried out for 20×10^6 generations with sampling every 2000 generations in BEAST. The resulting parameter and tree files from the two runs were combined using LogCombiner 2.5.0 (Bouckaert et al., 2014) and examined for convergence and effective sample sizes (ESS) in Tracer v1.7 (Rambaut et al., 2018; all

parameters having ESS > 200) and a maximum credible tree was found in TreeAnnotator 2.5.0 (Bouckaert et al., 2014).

Congruent topologies of mitochondrial and nuclear genes (see Results) suggest that past hybridization likely played a role in the evolution of *Arvicanthis*. We therefore reconstructed the Bayesian species network (reflecting both incomplete lineage sorting and admixture) using the newly developed approach SpeciesNetworks (Zhang, Ogilvie, Drummond & Stadler, 2017), implemented in BEAST 2.5.0 (Bouckaert et al., 2014). We used the same dataset, substitution and tree models as well as set up for clock rate as in StarBEAST analysis described above. In general, we followed the tutorial for SpeciesNetworks for defining parameters of the analysis in BEAUTI, however, we modified some priors as follows. We used exponential prior for netDivRate.t:Species (i.e. speciation rate minus hybridization rate) with mean 1.0. The exponential prior was used for the origin time of the species network (originTime.t:Species) with mean 0.001. The prior for popMean.t:Species was set up as gamma with shape (alpha) 2.0 and scale (beta) 0.005, which corresponds to distribution with a mean of 0.01. The analysis in BEAST was run with 20 million iterations, sampled every 2000th. All ESS were > 200 after checking in Tracer. The individual networks were explored in IcyTree, which is a web application for visualizing phylogenies, including phylogenetic networks (Vaughan, 2017).

Divergence dating

The temporal frame of divergences in the genus was analysed in StarBEAST 2. For this analysis, we used 35 individuals sequenced for both nuclear and mitochondrial loci (Appendix S1). The species tree shape was modelled by the birth-death process using an uninformative prior for the net diversification rate and a weakly informative prior, Beta(2,2), for the fraction of extinct species. Because we do not expect differences in substitution rates among lineages within the complex, we assumed a strict molecular clock but partition-specific clock rates (it is reasonable to assume e.g. much higher substitution rate in the mitochondrial marker). Because the variability of nuclear markers was relatively low, we defined more informative exponential prior (with mean = 0.002) for the clock rate in all nuclear genes, based on genetic distances between the two major groups in the genus. The fossil record of *Arvicanthis* does not allow the placing of fossils on the particular sites of the phylogenetic tree of the genus, so it is not possible to calibrate the molecular clock by ingroup fossils. We therefore performed a secondary calibration using the time to most recent common ancestor (TMRCA) of the NILOTICUS and ANSORGEI clades estimated by Aghová et al. (2018). In their study, the authors used the refined set of nine controlled fossil calibrations for the family Muridae and reconstructed a dated phylogeny of the family using a comprehensive multilocus dataset. They included also three species from the NILOTICUS group and two from the ANSORGEI group and

estimated TMRCA of all taxa at 4.45 Mya (with 95% HPD 3.65-5.23) . We therefore specified a normal prior for TMRCA of the two major *Arvicanthis* groups with mean 4.45 and standard deviation = 0.5, providing the distribution corresponding to primary divergence estimates (i.e. median 4.45 Mya, 95% HPD 3.63-5.27). The MCMC analysis in StarBEAST and analysis of outputs were performed as described above (but with 50 million iterations per run, sampled every 5000th, and 25% burnin) and the final figure was edited using R package 'strap' (Bell & Lloyd, 2014).

Results

Mitochondrial phylogeny and karyotypic variability

Both BI and ML analyses provided very similar topology of mitochondrial tree, with the genus *Arvicanthis* highly supported and split into two major clades, the ANSORGEI and NILOTICUS groups (Fig. 1). These two groups have partially overlapping distributions across the range of the genus in sub-Saharan Africa (see Fig. 2 for central-east Africa and Volobouev et al., 2002, for west-central Africa), but significantly differ by DNA sequences as well as by karyotypes (with the former having NFa = 72-78, while the latter NFa = 60-64; Fig. 1). Within each group we defined several usually highly supported clades that we will consider below as separate "species" (but see Discussion for taxonomic issues). The ANSORGEI group contains four species, while the NILOTICUS group is more diverse and we recognized nine species (Fig. 1).

In the ANSORGEI group, there are two species in eastern Africa: (a) *A. nairobae* is the only species in the group, whose monophyly at mtDNA is not significantly supported. It is highly structured with three distinct geographically separated clades (with unresolved relationships) distributed along the Albertine rift from Uganda to southern Tanzania, between Nairobi and north-eastern Tanzania, and in southern coastal Kenya, respectively (Fig. 2B). (b) *A. sp. "ANI-6"* was found only in the Great Rift Valley (GRV) in Ethiopia and is separated into two distinct mitochondrial clades in north-south direction. In addition, there is one species in west-central Africa (*A. rufinus*, see Fig. 2B for central African localities) and one limited to western Africa, west of the Niger River (*A. ansorgei*, not shown). The former one is sister to all other taxa in the ANSORGEI group, while the latter is closely related to Ethiopian endemic *A. sp. "ANI-6"*.

All nine species in the NILOTICUS group can be found in eastern Africa (Fig. 2A), while only one (*A. niloticus* "C2-C4"; named according to Dobigny et al., 2013) is also widespread in Sudanian savannas up to western Senegal. Except Ethiopia, the distribution of most species is roughly parapatric. Eastern border of the distribution of *A. niloticus* "C2-C4", represented in eastern Africa only by the clade C2, seems to be limited by GRV in southern Ethiopia and Kenya. On the other hand, *A. somalicus* is the

typical species of arid Somali-Masai savanna in Ethiopia and Kenya, east of GRV. *Arvicanthis neumanni* was recorded in the so-called Masai steppe in Tanzania, while its mitochondrial sister clade *A. sp. "Masai Mara"* only in two localities in GRV in Tanzania. Three distinct haplogroups of *A. niloticus* s. str. are distributed along the Nile River in Egypt and Sudan, in northern Ethiopia, and it is the only *Arvicanthis* species confirmed from Yemen in the Arabian Peninsula. Besides these five species, we distinguished four additional taxa; all of them are endemic to Ethiopia (Fig. 3). There are two species living in Afroalpine grasslands in the highest mountains of Ethiopia: *A. blicki* in Bale and Arsi Mts. (and Debre Sina, see below) and *A. abyssinicus* in Semien Mts and three additional highest mountain ranges in north-eastern part of Ethiopian highlands. Monophyly of *A. blicki* is not supported at *CYTB*, but the two species clearly differ at nuclear markers (see below). Two additional Ethiopian clades are *A. sp. "Menangesha"* (incorrectly called "*abyssinicus*" in previous genetic papers; see Table 1 and Discussion) from four localities around Addis Ababa, and *A. sp. "Metahara"* from lowland savannahs in the northern part of GRV (Fig. 3).

Mean uncorrected genetic distances at *CYTB* among the putative species (genetic clades) were > 8% and in most cases even > 10% (Table 2). There were two exceptions for the pairs *A. abyssinicus* - *A. blicki* ($5.1 \pm 0.5\%$) and *A. sp. "ANI-6"* - *A. ansorgei* ($6.9 \pm 0.6\%$). We also calculated p-distances in three species with pronounced intraspecific structure, i.e. between three clades of *A. niloticus* "C2-C4" ($4.1\text{--}4.7\%$), three clades of *A. nairobae* ($5.8\text{--}7.4\%$), and three clades of *A. rufinus* ($5.1\text{--}5.8\%$).

Phylogeny based on nuclear markers

Bayesian analysis based on six concatenated nuclear loci provided relatively well-resolved tree, clearly separating ANSORGEI and NILOTICUS groups (Fig. 4). All species identified by mtDNA within the NILOTICUS group were supported by PP > 0.97 also at nuclear markers (and bootstrap support of ML analysis > 86 in all but one species), but their mutual relationships were different from the ones inferred from the mtDNA tree (compare topologies at Figs. 1 and 4). The most important difference was in splitting the Afroalpine mitochondrial clade "*blicki/abyssinicus*" into two very distinct groups; one of them (*A. blicki*) clustering individuals from Bale + Arsi Mts. (east of GRV) with those from Debre Sina (west of GRV; Fig. 3), while the second included *A. abyssinicus* from the highest mountains in northwestern part of Ethiopian highlands. Diagnostic indels in introns grouped together (1) *A. niloticus* "C2-C4" with *A. sp. "Menangesha"* and *A. blicki*, and (2) *A. niloticus* s. str. with *A. sp. "Metahara"*. It is also interesting that Tanzanian *A. neumanni* was sister to all seven remaining species, which all were found in Ethiopia.

However, the individual nuclear genes were not very informative, when analysed separately, and individual gene trees provided slightly different topologies (Appendix S3), which can be the result of incomplete lineage sorting and/or past admixture. We therefore calculated Bayesian species tree in *BEAST (taking into account incomplete lineage sorting) and Bayesian species network (reflecting both incomplete lineage sorting and admixture) (Fig. 5). The resolution of the output was rather weak (reflecting low number of markers and their low variability), but some relationships were consistently supported in both analyses. In contrast to the concatenated tree (Fig. 4), both species tree and species network clustered *A. blicki* and *A. abyssinicus*, with the population from Debre Sina being closer to *A. abyssinicus*. Their closest sister is *A. sp. "Menangesha"*; all three species have the same N_{Fa}=64 (Fig. 1). Further, the sister relationship of *A. niloticus* s.str. and *A. sp. "Metahara"* was repeatedly confirmed. On the contrary, the phylogenetic position of *A. niloticus* "C2-C4", *A. somalicus* and *A. neumanni* strongly depended on the used method of analysis and the nuclear markers were unable to identify their sister clades. The sister relationship of *A. neumanni* and *A. somalicus* (considered to be synonyms by some authors, e.g. Musser & Carleton, 2005) was not confirmed by any of the used analyses and markers.

Divergence dating

The results of divergence dating are summarized at Fig. 6. The posterior median of TMRCA of all species of *Arvicanthis* (4.2 Mya) is slightly younger than the prior is (4.45 Mya). The species tree based on the combination of nuclear and mitochondrial data again supports two major *Arvicanthis* group, but the topology of the NILOTICUS group is not well resolved. Most divergences leading to current species diversity occurred simultaneously in early Pleistocene (1.5-2 Mya). The only exception is the clade of Ethiopian Afroalpine endemics (0.4-0.7 Mya), but these estimates should be taken with caution, as the evolution of Ethiopian taxa was likely affected by reticulation (see Discussion), which is not assumed by multi-species coalescent model in StarBEAST.

Discussion

Multilocus phylogenetic analysis performed in this study identified two major genetic groups within the genus (the NILOTICUS and ANSORGEI groups), thus confirming previous studies based on DNA sequences (Dobigny et al., 2013) and karyotypes (Castiglia et al., 2006). Both groups are widely distributed across sub-Saharan Africa and the species from these groups might occur in sympatry, e.g. in the inner delta of the Niger River (Volobouev et al., 2002), northern Cameroon (Dobigny et al., 2011), or southern Ethiopia (this study). The delimitation of species within the two main groups, however, is not yet satisfactorily done, especially in eastern Africa.

Relatively low diversity and well-resolved taxonomy of grass rats in West Africa

The taxonomic situation of West African taxa is relatively clear thanks to intensive research in last decades (e.g. Dobigny et al., 2011, 2013; Fadda & Corti, 2001; Volobouev et al., 2002; see Fig. 7). Based on combined evidence from cytogenetics, mitochondrial sequences and geometric morphometrics it was proposed that three species occupy open habitats of west-central Africa. There is one species from the NILOTICUS group (*A. niloticus* "C2-C4" sensu Dobigny et al., 2013; ANI-1a and ANI-1b cytotypes sensu Volobouev et al., 2002), widespread in dry savannahs of Sahelian and northern Sudanian domains (Fig. 7A), and two species from the ANSORGEI group (Fig. 7B). The first species *A. ansorgei* (= ANI-3 cytotype), with the holotype from Gunnal in Guinea Bissau (Volobouev et al., 2002), occupies bush and woodland savannas of the Sudanian domain in southern Senegal, Gambia, Guinea, Mali, Burkina-Faso, northern Benin and south-western Niger (Fig. 7B based on Denys et al., 2009; Dobigny, Nomao & Gautun, 2002; Fadda & Corti, 2001; Konečný, Koubek & Bryja, 2010; Volobouev et al., 2002). This species is parapatric with *A. niloticus* "C2-C4", with overlapping populations in the inner delta of the Niger River in Mali (Volobouev et al., 2002). The only record east of the Niger River was reported by Fadda & Corti (2001) from one locality in southern Chad, but it was based on two specimens identified by a discriminant function from geometric morphometric data and it should be taken with caution, awaiting genetic confirmation (not shown in Fig. 7B). The second species is *A. rufinus* (with the karyotype ANI-4 and likely also ANI-2; see Castiglia et al., 2006; Ducroz et al., 1998; and Dobigny et al., 2011 for taxonomic discussion), distributed in the most humid zones of the Sudanian-Guinean region and confirmed (cyto-)genetically from Guinea (C. Denys et al., pers. comm) through Benin and northern Cameroon to CAR (Dobigny et al., 2011; Volobouev et al., 2002). Based on the discriminant function derived from geometric morphometry of skulls, the species was also reported from Sierra Leone, Ghana (where is also the type locality) and western Nigeria (Fadda & Corti, 2001). This species exhibits intraspecific variation with two mitochondrial lineages (both with ANI-2 cytotype) east of the Niger River and one lineage (ANI-4 cytotype) west of the Niger River (Fig. 1; Dobigny et al., 2011). The study of possible reproductive barriers between the two cytotypes (i.e. ANI-2 and ANI-4 sensu Volobouev et al., 2002) would require new sampling, especially from Nigeria.

Species limits in the ANSORGEI group in East Africa

The taxonomic situation in eastern Africa is much more complicated and, unfortunately, the new multilocus DNA data analysed in this paper sheds light on some groups only. The less speciose clade ANSORGEI is represented in eastern Africa by two major groups that we propose to represent two distinct species. The first one is *A. nairobae*, the taxon described from Nairobi in Kenya. Castiglia et al. (2006) reported a very distinct karyotype ($2n = 62$, $NFa = 78$) from the type locality and they

assigned it to this taxon. An identical karyotype has been found previously at two localities in north-eastern Tanzania (Castiglia et al., 2003) and all these specimens form a strongly supported monophyletic group at mitochondrial DNA (Fig. 1). Its sister subclade grouped individuals from Uganda (Ducroz et al., 1998) and our new samples from western and southern Tanzania. The third, newly reported, mitochondrial clade clusters four individuals from two localities in coastal southern Kenya (Figs. 1 and 2B). The relationships of these three groups are not satisfactorily solved at mitochondrial DNA (and we were not able to amplify the nuclear markers). However, considering the geographic proximity and genetic divergences at *CYTB*, we propose them conspecific and suggest the name *A. nairobae* for all of them. Unfortunately, the type material of *A. nairobae* was not included in morphometric study of Fadda & Corti (2001), so it is difficult to identify, which morphogroup in their study might represent this species and what is the real extent of its distribution. However, the distribution of their "*A. sp1*" significantly overlaps with genetically identified *A. nairobae* in our study and it is the only *Arvicanthis* species in western Tanzania. It is also morphologically well distinct taxon in East Africa, suggesting it belongs to the ANSORGEI group (Fadda & Corti, 2001). We therefore assigned their "*A. sp1*" to *A. nairobae* at Fig. 7B. Awaiting further study, the species seems to be distributed in miombo habitats along the Albertine rift and along eastern part of Kenya/Tanzania border (with the type locality, Nairobi, laying at the margin of its distribution; Fig. 7B).

The second East African species from the ANSORGEI group is called here *A. sp. "ANI-6"*, based on the karyotype described by Corti et al. (2005) from Ethiopia (we included some of their specimens in our phylogenetic analysis). This species very likely includes also the populations with the karyotype ANI-6a (Orlov, Baskevich & Bulatova, 1992; see Castiglia et al., 2006). Based on mitochondrial DNA analyses, it is a significantly supported sister taxon to West African *A. ansorgei* and its distribution is limited to Ethiopian lowlands, especially in GRV (Fig. 2B). It has clear north-south intraspecific structure, separating mtDNA haplotypes into two subgroups. A similar phylogeographic pattern was found also in other rodents living in Ethiopian Somali-Masai savannahs, e.g. gerbils (Aghová et al., 2017) or spiny mice (Aghová et al., under review). Intraspecific divergence can be a result of a vicariance event caused by the African rift lakes forming important barriers to gene flow, especially during humid periods of Pleistocene (Trauth et al., 2010; see also discussion in Aghová et al., 2017).

Genus Arvicanthis in Ethiopian highlands

Ethiopian highlands are inhabited by *Arvicanthis* populations with two very divergent mitochondrial haplogroups, called here *A. sp. "Menangesha"* and *A. abyssinicus/blicki* (Fig. 1). The concatenated nuclear dataset clearly separated *A. abyssinicus* and *A. blicki* (Fig. 4), suggesting that three highland species of *Arvicanthis* can occur in Ethiopia (Fig. 3). *Arvicanthis abyssinicus* was described from

Semien Mts., so we assigned this name to the only genetic clade of the genus captured in these mountains during several recent expeditions (Appendix S1). It is a high-elevation specialist, occupying Afroalpine habitats at 3600-3900 m a.s.l. However, there are two reasons to cast doubts on the assignment of this species name to these populations. First, the type locality (Entschetqab, Simien Mts.) was reported from elevation about 3250 m a.s.l. ("ca. 10,000 French feet"; Rüppell, 1842), which is slightly lower than known altitudinal range of this genetic clade (besides Semien Mts. the species is reported from three additional highest mountain ranges west of GRV, always above 3600 m a.s.l.; Fig. 3). Second, the type material of *A. abyssinicus* is unfortunately damaged to the extent that it cannot be used for any multivariate analysis of skull morphology (O. Mikula, pers. comm.; see also Yalden, Largen & Kock, 1976). *Arvicanthis blicki*, morphologically very distinct species (Fadda & Corti, 2001) was considered to be distributed only in Bale and Arsi Mts., east of GRV (Denys et al. 2017), but the concatenated nuclear tree suggests that it is distributed also in Debre Sina, west of GRV (Figs. 3 and 4). It seems very likely that GRV was repeatedly crossed by highland specialists during glacial periods, because similar biogeographic pattern, i.e. genetically very close populations on both sides of GRV, was observed in other taxa of mountain rodents (Bryja et al., 2018; Lavrenchenko, Verheyen, Verheyen, Hulselmans & Leirs, 2007) and frogs (Freilich, Tollis & Boissinot, 2014). Finally, *A. sp. "Menangesha"* is the species that we reported from relatively limited area in central Ethiopia based on DNA genotyping and karyotypes (Fig. 3, Appendix S1). In recent literature, this species was incorrectly reported under the name "*A. abyssinicus*" (Table 1). Most localities reported by Yalden et al. (1976) and Yalden, Largen, Kock & Hillman (1996) as "*A. abyssinicus*" from the elevation of 2000-3400 m a.s.l. from the western part of Ethiopian plateau very likely represent *A. sp. "Menangesha"*. It usually has a mid-dorsal dark stripe and rather dark agouti fur on the venter (Dollman, 1911; Yalden et al., 1976; J. Bryja and R. Castiglia, pers. obs.), but the detailed phenotypic analysis of genotyped individuals is required to evaluate this interpretation.

The evolutionary history of *Arvicanthis* in Ethiopian highlands, however, was probably complex and there are indications of past reticulate processes. Species tree and species network based on nuclear loci (Fig. 5) both suggest that all Ethiopian highland populations form a monophyletic group, where *A. abyssinicus* is closely related to *A. blicki* from Debre Sina. Close relationship between *A. blicki* east of GRV (Bale and Arsi Mts.) and *A. sp. "Menangesha"* west of GRV is documented by concatenated nuclear phylogeny, shared insertion in DHCR gene (Fig. 4), allozymes (Capula, Civitelli, Corti, Bekele & Capanna, 1997) and by the same inversion on chromosome 3 (Castiglia et al., 2006). On the other hand, they are very distinctive for body size, pelage coloration, diploid chromosome number and skull morphology (Castiglia et al., 2006; Fadda & Corti 2001; Yalden et al., 1976). Although we are unable to resolve the phylogenetic relationships of *Arvicanthis* in Ethiopian highlands with the

present dataset, we can propose an evolutionary scenario that could be tested in future work by genomic approaches. This scenario must include: (1) the gene flow across GRV, probably during glacial periods, and cross-breeding of ancestors of *A. sp. "Menangesha"* and *A. blicki*, which is suggested by shared traits at DHCR gene and inversion on chromosome 3; (2) rapid karyotype changes (Castiglia et al., 2006) and evolution of very distinct phenotype of *A. blicki* (Fadda & Corti, 2001) in Afroalpine conditions of the Bale/Arsi mountains, probably in response to strong environmental selection (see similar examples e.g. in Šumbera et al., 2018; Kostin et al., 2018); (3) hybridization of ancestors of *A. blicki* with the ancestors of *A. abyssinicus*, which resulted in the (adaptive?) introgression of mitochondrial DNA of *A. blicki* into *A. abyssinicus*, potentially allowing *A. abyssinicus* to colonize the highest mountains in north-western Ethiopian plateau. The population in Debre Sina may represent one of intermediate populations of hybrid origin between these two taxa, and its newly reported karyotype ($2n = 62$, $NFa = 64$) suggest that the hybridization predated the karyotypic changes of *A. blicki* in Bale and Arsi Mts. It is also possible that Ethiopian endemic *Arvicanthis* sp. "Menangesha", able to live also at relatively low altitudes (see Yalden et al., 1976), potentially hybridized with *A. niloticus* "C2-C4" in southwestern Ethiopia and adopted its DHCR gene (Fig. 4) and mtDNA (Fig. 1), or vice versa. Albeit this scenario seems to be quite complicated, the reticulate speciation processes are common in Ethiopian highlands, at least in rodents (Bryja et al., 2018; Kostin et al. 2018; Lavrenchenko et al., 2004) and amphibians (Freilich et al., 2014), i.e. groups, where both nuclear and mitochondrial markers have been studied so far. For example, two distinct Afroalpine species exist in *Stenocephalemys* (*S. albocaudata* and *S. sp. "A"*) and their distribution is the same as that of *A. blicki* and *A. abyssinicus*, respectively (compare Fig. 3 with Bryja et al., 2018). Interestingly, the two Afroalpine *Stenocephalemys* species have very similar mtDNA, but are clearly distinct at nuclear markers, which is also in agreement with the situation in *Arvicanthis*.

Diversity of the NILOTICUS group in Sudanian and Somali-Masai savannahs

There have been numerous confusions in using the name *A. niloticus* (see e.g. Musser & Carleton 2005 for a review). This species was described from Egypt and in the most comprehensive genetic work of Dobigny et al. (2013) the name was used for widespread populations with ANI-1a and ANI-1b cytotypes from Senegal to eastern Sudan, harbouring four distinct mitochondrial haplogroups C1-C4. We revealed similar structure at mitochondrial DNA, when using more samples especially from eastern Africa, but our nuclear markers show that the easternmost clade C1 is very distinct and is not even a sister species to C2-C4 (in contrast, C2-C4 individuals form a well-supported monophyletic group at nuclear markers despite large geographic distances among them; Fig. 4). This supports previous views, e.g. based on skull morphology (Fadda & Corti, 2001), that populations living along the Nile valley, in northern Ethiopia and Yemen represent the same species (see also taxonomic

discussion in Musser & Carleton, 2005), here called *A. niloticus* s. str., distinct from Sudano-Sahelian taxon, i.e. *A. sp. "C2-C4"* (sensu Dobigny et al., 2013). The samples from Egypt belong to the mitochondrial clade C1 and the name "*niloticus*" should be therefore used for the species distributed along the Nile River in Egypt and Sudan, and in northern Ethiopia and Yemen (Fig. 7A), with frequently used name "*dembeensis*" being its younger synonym. The species *A. niloticus* "C2-C4", widely distributed in the belt of Sudanian savanna from Senegal to southern Ethiopia and western Kenya (Fig. 7A), should thus receive another name. It is important to note that Fadda & Corti (1998) found conspicuous variation in skull morphology in the potential contact zone of C1 (i.e. *A. niloticus* s. str.) and C2 (called "*testicularis*" in their work) in Sudan and South Sudan, where future studies should investigate the extent of reproductive isolation between these two taxa and their phenotypic differences.

Within the NILOTICUS group, there are three distinct mitochondrial clades that potentially co-occur in the central and northern part of GRV in Ethiopia - *A. sp. "Metahara"*, *A. somalicus*, and *A. niloticus* s. str. (Figs. 1 and 3). Also based on nuclear markers, they are clearly distinct, even if we did not analyse the latter one from the central rift, but only from northern Ethiopia. In Ethiopia, *A. niloticus* s. str. prefers higher altitudes (Koka and Dera Dilfekar in GRV, cca 1700 m a.s.l.), which is also true in northern Ethiopia (up to 2700 m a.s.l.; Meheretu et al., 2012, 2014; J. Bryja et al., unpubl. data). Two remaining species, however, are syntopic at low elevation (≤ 1000 m a.s.l.) in the southern part of the Afar triangle, e.g. in the Awash National Park (Corbet & Yalden, 1972; Demeter, 1983; this study). *Arvicanthis somalicus*, described from northern Somaliland (Dollman, 1911) is an arid-adapted species with small body size and it is probably the only species of *Arvicanthis* living in dry conditions of the Somali region in south-eastern Ethiopia, Somalia and eastern Kenya (Fig. 7A). The species has been considered as a synonym of *A. neumanni*, another small-bodied species living in arid habitats of the Masai steppe in Tanzania. Both nuclear and mitochondrial markers suggest that the two taxa are genetically very distinct (Figs. 1, 4, 5), which agrees with earlier findings that they also differ significantly in the shape of the skull (Fadda & Corti, 2001) and karyotype (Castiglia et al., 2006). Hence, it seems they should be considered as two different species (see also suggestions in Musser & Carleton, 2005 and Denys et al., 2017; for their distribution see Fig. 7A). All genetically identified individuals of *A. sp. "Metahara"* were collected from the central and northern part of GRV (Fig. 3) and this species was recently captured also near Dire Dawa further eastward (C. Denys, unpubl. data). It is likely that previous records of larger *Arvicanthis* from this area (Corbet & Yalden, 1972; Demeter, 1983; some of them reported as *A. dembeensis* by Yalden et al., 1976) represent this species. Its precise distributional limits and phylogenetic relationships with *A. niloticus* s. str. should be further

investigated, but it can be anticipated that they are relatively close relatives as suggested by shared insertion at SMO gene (Fig. 4) and sister relationship in the species tree (Fig. 5).

The mitochondrial lineage *A. sp. "Masai-Mara"* is the only one, where we were unable to sequence nuclear markers and its specific status thus remains obscure. This lineage is only known from two localities in the so-called Serengeti ecosystem, which hosts several endemic species of small mammals, e.g. *Mastomys pernanus* (Denys et al., 2017). They represent relatively large-bodied specimens and even if they are sister with *A. neumanni* based on mitochondrial DNA (Fig. 1), they have different karyotype (ancestral ANI-1a; Ducroz, 1998). The name *A. muansae* Matschie, 1911, described from Mwanza on the south coast of Lake Victoria, is a possible candidate for relatively large individuals from this region, but more detailed taxonomic study using new material is required to solve this question. Fadda & Corti (2001) identified a morphologically distinct group of individuals along the Lake Victoria (called "*A. sp.*" in their paper) that might represent this species, but unfortunately, none of these individuals was sequenced. They were morphologically most similar to their "*A. sp2*" and "*A. testicularis*" that we included in *A. niloticus* "C2-C4" based on significant overlap with the distribution ranges of our genotyped individuals (Fig. 7). The genetic and phenotypic analyses of populations from northern Tanzania, western Kenya, Uganda, northern DRC and South Sudan is now necessary to solve the remaining taxonomic issues in the NILOTICUS group.

Taxonomic implications

Using the combination of nuclear and mitochondrial DNA markers with available cytogenetic and morphological data, we provide the most comprehensive description of the diversity of *Arvicanthis* in sub-Saharan Africa (Fig. 7). Most recent compendia of rodent diversity (Denys et al., 2017; Monadjem, Taylor, Denys & Cotterill, 2015) recognize seven *Arvicanthis* species. Our results show that the species diversity within the genus is much higher and at least some of the revealed genetic groups should receive a species name. We do not aim to perform a comprehensive integrative taxonomic revision of the genus in this paper, as our study mainly used genetic data. For example, we have insufficient information to interpret the taxonomic implications of our findings for *A. niloticus* "C2-C4". The numerous *Arvicanthis* descriptions originate from Kenya and Uganda (see Allen, 1939), from where the substantial morphological variability was reported (Fadda & Corti, 2001) and where the taxon *A. sp. "Masai Mara"* could occur. It will be necessary to perform combined genetic and morphological analysis from this region to answer the question of how many species exist there and how they are related to *A. niloticus* "C2-C4" in central and western Africa. Similarly, there are at least four descriptions from "Anglo-Egyptian Sudan" (Allen, 1939), i.e. current Sudan and South Sudan, which is a region of potential contact zone of *A. niloticus* s. str. and *A. niloticus* "C2-C4".

Detailed analysis of the type material and genomic analysis of potential contact zone, especially in Sudan, is required to delimit correctly these taxa.

On the other hand, our sampling and genotyping allow to revise at least the suitable names for Ethiopian species. Based on Allen (1939), following taxa were named from Ethiopia:

Mus abyssinicus Rüppel, 1842 - described from Entschetqab, Simien province, ca. 10,000 French feet (3250 m a.s.l.). We use this name for the high-elevation taxon *A. abyssinicus*, recorded in Afroalpine habitats in four highest mountains in north-western part of Ethiopian highlands (Fig. 3). It is the only species of *Arvicanthis* captured in Semien Mts. during our recent expeditions. "*Mus rufidorsalis* Heuglin, 1877" from "grassy plain in Simien and Wogara" was synonymized with *A. abyssinicus* (Allen, 1939). The name "*abyssinicus*" has been frequently incorrectly used for other Ethiopian taxa, especially *A. sp.* "Menangesha" (see below).

Arvicanthis abyssinicus blicki Frick, 1914 was described from South Chilalo Mts., Hora Mt., 9000 ft (2743 m a.s.l.) and it represents another high-elevation specialist, *A. blicki*, which is also morphologically very distinct (Fadda & Corti, 2001). It lives in Afroalpine habitats of Bale and Arsi Mts., east of GRV, but another population (affected by hybridization with *A. abyssinicus*?) was confirmed from Debre Sina, west of GRV (Fig. 3). The understanding of possible reticulate evolution of highland Ethiopian taxa would require the use of genomic approaches.

Arvicanthis abyssinicus fluvicinctus Osgood, 1936 was described from Bichana, Gojjam (close to Debre Marcos) and it was synonymized with "*abyssinicus*" in Musser & Carleton (2005), corresponding to *A. sp.* "Menangesha" in our study.

Arvicanthis abyssinicus mearnsi Frick, 1914 was described from "Sadi Malka, Hawash River, 2800 ft (853 m a.s.l.)", close to the current Awash NP, where two *Arvicanthis* species co-occur (*A. sp.* "Metahara" and *A. somalicus*). The name is reported as synonym of "*niloticus*" in Musser & Carleton (2005), but it is the most suitable name for *A. sp.* "Metahara", i.e. bigger lowland species living (at least) in northern part of GRV and at the southern margin of the Afar triangle. The hindfoot length in the type series is 27-29 mm, while all individuals of genotyped *A. somalicus*, i.e. its sympatric taxon, in our collection had it shorter than 24.5 mm.

Arvicanthis abyssinicus raffertyi Frick, 1914 was described from Gardula, south of Arba Minch from the elevation of 4000 ft (ca 1220 m a.s.l.). This is the area of sympatry between *A. sp.* "ANI-6" and *A. niloticus* "C2-C4". Our unpublished data suggest that the two species differ by the length of hindfoot (albeit overlapping), being significantly longer for *A. niloticus* "C2-C4". The measurements provided for the type series of "*raffertyi*" correspond to *A. sp.* "ANI-6" (most of them 27.0-28.5 mm, while frequently above 29.0 mm in *A. niloticus* "C2-C4"). Based on the description, "*raffertyi*" also differs from *A. abyssinicus zaphiri* Dollman, 1911 (synonymized with *A.*

niloticus, see below) by the lack of the broad brain-case (Frick, 1914). To sum up, we propose the name *A. raffertyi* for A. sp. "ANI-6", which is genetically the most distinct among Ethiopian *Arvicanthis*, but has been overlooked in previous studies (Fadda & Corti reported it as "A. sp.3", morphologically similar to *A. ansorgei*, which correspond to genetics as well). It should be noted that Yalden et al. (1976) mentioned that the altitude of the type locality (Gardula) has an approximate elevation of 2500 m a.s.l. (and not 1200 m as reported by Frick, 1914), but at the same time it is possible that the type material was collected in the lowlands "near Gardula".

Arvicanthis abyssinicus saturatus Dollman, 1911 from Didessa River near Guma is the oldest available name for *A. sp. "Menangesha"*, reported in most previous works (except the original description) as *A. abyssinicus*. The description by Dollman (1911) perfectly fits this taxon, i.e. general colour dark blackish, backs of feet dark brown, toes black and dorsal stripe (usually) well marked. In addition, the type locality is not far from our genotyped material of this taxon (Fig. 3) and the holotype was assigned to this taxon in morphometric analysis of skulls (Fadda & Corti, 2001). Based on the analysis of 3D geometric morphometry (Fadda & Corti, 2001) the species is reported from additional localities at 1600-2880 m a.s.l., mostly west of GRV. Occurrence at two localities east of GRV should be confirmed genetically as the species identification is based only on discriminant function. The name "*fluvicinctus*" is the younger synonym of this taxon (see above).

Arvicanthis abyssinicus zaphiri Dollman, 1911 was described from Guma, 2200 ft (650 m a.s.l.) and was synonymized with *A. niloticus* by Musser & Carleton (2005). Based on the geographic location of the type locality, it will rather belong to *A. niloticus* "C2-C4", even if the hindfoot (31 mm) is relatively long and rather similar to *A. niloticus* s. str. (our unpubl. data).

Meriones lacernatus Rüppel, 1842 was described from grassy plains around Lake Dembea (= Lake Tana). Allen (1939) mentioned "*dembeensis* Rüppel, 1842" and "*pelliceus* Thomas, 1928" described from the same area as its synonyms and both "*dembeensis*" and "*pelliceus*" are considered synonyms to *A. niloticus* by Musser & Carleton (2005) (even if they do not list "*lacernatus*"). The genotyped material from around Lake Tana confirms that all these taxa very likely belong to *A. niloticus* s. str.

Mus ochropus Heuglin, 1877 was described from the Bogos country, Abyssinia, which in fact now belong to Eritrea (=Keren). The species is listed as synonym of *A. niloticus* by Musser and Carleton (2005), and it geographically corresponds to *A. niloticus* s. str.

Based on this information, we propose that Ethiopia is inhabited by one species from the ANSORGEI group (*A. raffertyi* = *A. sp. "ANI-6"*) and seven species from the NILOTICUS group. Four of them were described from Ethiopia and represent the endemic species for the country: *A. abyssinicus*, *A. blicki*, *A. saturatus* (= *A. sp. "Menangesha"*) and *A. mearnsi* (= *A. sp. "Metahara"*). Three additional

Ethiopian species are more widespread (Fig. 7): *A. niloticus* (= *A. niloticus* s. str. described from Egypt), *A. niloticus* "C2-C4" (the naming of this taxon requires additional taxonomic work, see above) and *A. somalicus* (described from Somaliland and widely distributed in Somali-Masai savannah).

Further cryptic diversity?

Even if we have provided a significant amount of new data illuminating the taxonomic and evolutionary diversity of *Arvicanthis*, especially in Eastern Africa, there are still unanswered questions concerning the systematics of the genus. Besides those highlighted in previous text (e.g. reticulate evolution in Ethiopia or species limits of *A. niloticus* "C2-C4"), the previous cytogenetic research suggested that diversity of *Arvicanthis* can be even higher than revealed by the DNA markers in our study (see review in Castiglia et al., 2006). Some populations with very distinct karyotypes were included in our study. For example, the ANI-5 (sensu Castiglia et al., 2006) clearly represents an example of chromosomal fusions within *A. niloticus* "C2-C4" (Fig. 1). The ANI-5 population was described from the margin of the distribution range of this species, hence it is possible that peripatric evolution in marginal populations has facilitated the fixation of the chromosomal mutations; see similar examples e.g. in the rodent genera *Rhabdomys* (Castiglia et al., 2012) or *Taterillus* (Dobigny et al., 2015). Two other reported karyotypes (ANI-7 and ANI-8 sensu Castiglia et al., 2006), are worth of further work, which might reveal new (species) diversity in the ANSORGEI group. The ANI-7 (2n=56, Nfa=78 described from the Gambela area by Bulatova et al., 2002) originates from the westernmost karyotyped locality in Ethiopia and may represent a missing link between *A. sp.* "ANI-6" and non-Ethiopian taxa of the ANSORGEI group. Whether it represents just intraspecific variation within *A. sp.* "ANI-6" (as suggested by morphological study of Fadda & Corti, 2006, who identified part of specimens from Gambela as *A. sp.* 3 = *A. sp.* "ANI-6"; see Table 1 and Fig. 7B) or it is a separate taxon should be further researched. Similarly, the karyotype ANI-8 (2n=44, NF=72, Nfa not known) described by Capanna & Civitelli (1988) from Afgoi in Somalia stands that the genus is prone to karyotypic divergence. In this case, the divergence seems relatively huge and may result in reproductive isolation between the Afgoi population and geographically closest taxa *A. sp.* "ANI-6" or *A. nairobae*. Considering the cradle of the genus in East Africa, the occurrence of yet undetected species in poorly sampled areas is definitely possible.

Acknowledgements

This study was supported by the Czech Science Foundation (project no. 18-17398S) and the Russian Foundation for Basic Research (project no. 18-04-00563-a). Samples utilised in the study have been lawfully acquired and were collected prior to The Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from Their Utilization to the Convention on

Biological Diversity has been in effect. We are indebted to many local authorities for providing permits to carry out the research, especially the Ethiopian Wildlife Conservation Authority (EWCA), Government of Ethiopia and the Oromia Forest and Wildlife Enterprise (OFWE) in Ethiopia, the Sokoine University of Agriculture in Morogoro (Tanzania), the Kenyan Forest Service and the Kenyan Wildlife Service (Kenya). For help during the field work and providing additional samples we acknowledge E. Capanna, A. Bekele, N Oguge, R. Makundi, C. Denys, P. Benda, S. Gryssels, J. Šklíba, T. Vlasatá, E. Hrouzková, S. Šafarčíková, M. Lövy, O. Mikula, A. Konečný, A. Hánová, L. Cuypers, A. Ribas, J. Krásová, J. Sádlová, D. Frynta, J. Votýpka, W. T. Stanley, E. C. Craig, C. Sabuni, A. Katakweba, A. Massawe, K. Welegerima, A.A. Warshavsky, Yu. F. Ivlev, D. Yu. Alexandrov, D. S. Kostin, A. A. Martynov, A. R. Gromov, Z. Tomass and all local collaborators. For help with genotyping we acknowledge T. Aghová and A. Hánová. Constructive comments during the data analysis and writing were provided by O. Mikula and L. Granjon.

Authors' Contributions

JB, RS, PC, EV, HL, RC conceived and designed the study; JB, LAL, YM, RS, RC collected important part of samples; AB, YM performed laboratory analysis; JB analysed data; JB wrote the first version of the manuscript. All authors contributed to the final version of the paper. All authors read and approved the final manuscript.

References

- Abdel Rahman Ahmed, E.H., Ducroz, J.-F., Mitchell, A., Lamb, J.M., Contrafatto, G., Denys, C. & Taylor, P.J. (2008). Phylogeny and historical demography of economically important rodents of the genus *Arvicanthis* (Mammalia: Muridae) from the Nile Valley: of mice and men. *Biological Journal of the Linnean Society*, 93, 641-655. doi: 10.1111/j.1095-8312.2007.00895.x
- Aghová, T., Šumbera, R., Piálek, L., Mikula, O., McDonough, M.M., Lavrenchenko, L.A., Meheretu, Y., Mbau, J.S. & Bryja, J. (2017). Multilocus phylogeny of East African gerbils (Rodentia, *Gerbilliscus*) illuminates the history of the Somali-Masai savanna. *Journal of Biogeography*, 44, 2295-2307. doi: 10.1111/jbi.13017
- Aghová, T., Kimura, Y., Bryja, J., Dobigny, G., Granjon, L. & Kergoat, G.J. (in press). Fossils know it best: using a new set of fossil calibrations to improve the temporal phylogenetic framework of murid rodents (Rodentia: Muridae). *Molecular Phylogenetics and Evolution*, 128, 98-111. doi: 10.1016/j.ympev.2018.07.017
- Aghová, T., Palupčíková, K., Šumbera, R., Frynta, D., Lavrenchenko, L.A., Meheretu, Y., Sádlová, J., Votýpka, J., Mbau, J.S., Modrý, D. & Bryja, J. (under review). Multiple radiations of spiny mice (Rodentia: *Acomys*) in dry open habitats of Afro-Arabia: evidence from multi-locus phylogeny. *BMC Evolutionary Biology*, under review.
- Allen, G.M. (1939). A checklist of African mammals. *Bulletin of the Museum of Comparative Zoology at Harvard College*, 83, 1-763.
- Baskevich, M.I. & Lavrenchenko, L.A. (2000). Review of karyological studies and the problems of systematics of Ethiopian *Arvicanthis* Lesson, 1842 (Rodentia: Muridae). In G. Rheinwald (Ed.), *Isolated Vertebrate Communities in the Tropics* (pp. 209-215). *Bonner Zoologische Monographien* 46, Bonn, Germany.
- Bekele, A., Capanna, E., Corti, M., Marcus, L.F. & Schlitter, D.A. (1993). Systematics and geographic variation of Ethiopian *Arvicanthis* (Rodentia, Muridae). *Journal of Zoology*, 230, 117-134. doi: 10.1111/j.1469-7998.1993.tb02676.x
- Bell, M.A. & Lloyd, G.T. (2017, August 5). strap: Stratigraphic Tree Analysis for Palaeontology. R package version 1.4. Retrieved from <https://cran.r-project.org/package=strap>
- Bouckaert, R.R., Heled J., Kuehnert, D., Vaughan, T.G., Wu, C.-H., Xie, D., Suchard M.A., Rambaut, A. & Drummond, A.J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10, e1003537. doi: 10.1371/journal.pcbi.1003537
- Bryja, J., Mikula, O., Šumbera, R., Meheretu, Y., Aghová, T., Lavrenchenko, L.A., Mazoch, V., Ogue, N., Mbau, J.S., Welegerima, K., Amundala, N., Colyn, M., Leirs, H. & Verheyen, E. (2014). Pan-African phylogeny of *Mus* (subgenus *Nannomys*) reveals one of the most successful mammal

- radiations in Africa. *BMC Evolutionary Biology*, 14, article no. 256. doi: 10.1186/s12862-014-0256-2
- Bryja, J., Kostin, D., Meheretu, Y., Šumbera, R., Bryjová, A., Kasso, M., Mikula, O. & Lavrenchenko, L.A. (2018). Reticulate Pleistocene evolution of Ethiopian rodent genus along remarkable altitudinal gradient. *Molecular Phylogenetics and Evolution*, 118, 75-87. doi: 10.1016/j.ympev.2017.09.020
- Bulatova, N.Sh., Lavrenchenko, L.A., Orlov, V. & Milishnikov, A. (2002). Notes on chromosomal identification of rodent species in western Ethiopia. *Mammalia*, 66, 128-132.
- Bulatova, N.Sh. & Lavrenchenko, L.A. (2005). Possible karyological affinities of small mammals from north of the Ethiopian plateau. In B.A. Huber, B.J. Sinclair & K.H. Lampe (Eds.), *African Biodiversity: Molecules, Organisms, Ecosystems* (pp. 315-319). Springer Science + Business Media, Inc., New York, USA.
- Capanna, E. & Civitelli, M.V. (1988). A cytotaxonomic approach of the systematics of *Arvicanthis niloticus* (Desmarest 1822) (Mammalia, Rodentia). *Tropical Zoology*, 1, 29-37. doi: 10.1080/03946975.1988.10539404
- Capula, M., Civitelli, M.V., Corti, M., Bekele, A. & Capanna, E. (1997). Genetic divergence in the genus *Arvicanthis* (Rodentia: Murinae). *Biochemical Systematic and Ecology*, 25, 403-409. doi: 10.1016/S0305-1978(96)00109-3
- Castiglia, R., Corti, M., Tesha, P., Scanzani, A., Fadda, C., Capanna, E. & Verheyen, W. (2003). Cytogenetics of the genus *Arvicanthis* (Rodentia, Muridae). 3. Comparative cytogenetics of *A. neumanni* and *A. nairobae* from Tanzania. *Genetica*, 118, 33-39. doi: 10.1023/A:1022903112274
- Castiglia, R., Bekele, A., Makundi, R., Ouge, N. & Corti, M. (2006). Chromosomal diversity in the genus *Arvicanthis* from east Africa: a taxonomic and phylogenetic evaluation. *Journal of Zoological Systematic and Evolutionary Research*, 44, 223-235. doi: 10.1111/j.1439-0469.2006.00356.x
- Castiglia, R., Solano, E., Makundi, R.H., Huselmans, J., Verheyen, E. & Colangelo, P. (2012). Rapid chromosomal evolution in mesic four-striped grass rat *Rhabdomys dilectus* (Rodentia, Muridae) revealed by mtDNA phylogeographic analysis. *Journal of Zoological Systematic and Evolutionary Research*, 50, 162-172. doi: 10.1111/j.1439-0469.2011.00627.x
- Castillo-Ruiz, A., Indic, P. & Schwartz, W. J. (2018). Time management in a co-housed social rodent species (*Arvicanthis niloticus*). *Scientific Reports*, 8, article no. 1202. doi: 10.1038/s41598-018-19365-3
- Civitelli, M.V., Castiglia, R., Codja, J.C. & Capanna, E. (1995). Cytogenetics of the genus *Arvicanthis* (Rodentia, Murinae). 1. *Arvicanthis niloticus* from Republic of Benin (West Africa). *Zeitschrift für Säugetierkunde*, 60, 215-225.
- Corbet, G.B. & Yalden, D.W. (1972). Recent records of mammals (other than bats) from Ethiopia. *Bulletin of the British Museum (Natural History), Zoology*, 22, 213-252.

- 738 Corti, M. & Fadda, C. (1996). Systematics of *Arvicanthis* (Rodentia, Muridae) from the Horn of Africa:
739 a geometric morphometrics evaluation. *Italian Journal of Zoology*, 63, 185-192. doi:
740 10.1080/11250009609356128
- 741 Corti, M., Civitelli, M.V., Bekele, A., Castiglia, R. & Capanna, E. (1995). The chromosomes of three
742 endemic rodents of the Bale mountains, South Ethiopia. *Rendiconti Lincei Scienze Fisiche e*
743 *Naturali*, 6, 157-164.
- 744 Corti, M., Civitelli, M.V., Castiglia, R., Bekele, A., & Capanna, E. (1996). Cytogenetics of the genus
745 *Arvicanthis* (Rodentia, Muridae). 2. The chromosomes of three species from Ethiopia: *A.*
746 *abyssinicus*, *A. dembeensis* and *A. blicki*. *Zeitschrift für Säugetierkunde*, 61, 339-351.
- 747 Corti, M., Castiglia, R., Colangelo, P., Capanna, E., Beolchini, F., Bekele A., Ouge, N., Makundi, R.H.,
748 Sichilima, A.M., Leirs, H., Verheyen, W. & Verhagen, R. (2005). Cytotaxonomy of rodent species
749 from Ethiopia, Kenya, Tanzania and Zambia. *Belgian Journal of Zoology*, 135, 197-216.
- 750 Demeter, A. (1983). Taxonomical notes on *Arvicanthis* (Mammalia: Muridae) from the Ethiopian Rift
751 Valley. *Annales du Musée Royal de 'Afrique Centrale (Série 8vo), Sciences Zoologiques, Tervuren*,
752 237, 129-136.
- 753 Denys, C., Lalis, A., Aniskin, V., Kourouma, F., Soropogui, B., Sylla, O., Doré, A., Koulemou, K.,
754 Beavogui, Z.B., Sylla, M., Camara, A., Camara, A.B., Camara, A.C., Kouassi Kan, S., Volobouev, V.,
755 Camara, C., Koivogui, L. & Bernard, A.K. (2009). New data on the taxonomy and distribution of
756 Rodentia (Mammalia) from the western and coastal regions of Guinea West Africa. *Italian Journal*
757 *of Zoology*, 76, 111-128. doi: 10.1080/11250000802616817
- 758 Denys, C., Taylor, P.J. & Aplin, K.P. (2017). Family Muridae. In D.E. Wilson, T.E. Jr. Lacher & R.A.
759 Mittermeier (Eds.), *Handbook of the Mammals of the World, Volume 7: Rodents II*. Barcelona:
760 Lynx Edicions.
- 761 Dobigny, G., Nomao, A. & Gautun, J.C. (2002). A cytotaxonomic survey of Rodents from Niger:
762 Implications for systematics, biodiversity and biogeography. *Mammalia*, 66, 495-523. doi:
763 10.1515/mamm.2002.66.4.495
- 764 Dobigny, G., Tatard, C., Kane, M., Gauthier, P., Brouat, C., Ba, K. & Duplantier, J.M. (2011). A
765 cytotaxonomic and DNA-based survey of rodents from Northern Cameroon and Western Chad.
766 *Mammalian Biology*, 76, 417-427. doi:10.1016/j.mambio.2010.10.002
- 767 Dobigny, G., Tatard, C., Gauthier, P., Ba, K., Duplantier, J.M., Granjon, L. & Kergoat, G. J. (2013).
768 Mitochondrial and nuclear genes-based phylogeography of *Arvicanthis niloticus* (Murinae) and
769 sub-Saharan open habitats Pleistocene history. *PLoS ONE*, 8, article no. 77815. doi:
770 10.1371/journal.pone.0077815

- 771 Dobigny, G., Aniskin, V., Granjon, L., Cornette, R. & Volobouev, V. (2005). Recent radiation in West
772 African *Taterillus* (Rodentia, Gerbillinae): the concerted role of chromosome and climatic changes.
773 *Heredity*, 95, 358-368. doi: 10.1038/sj.hdy.6800730
- 774 Dollman, G. (1911). On *Arvicanthis abyssinicus* and allied East-African species, with descriptions of
775 four new forms. *Annals and Magazine of Natural History*, 8(45), 334-353. doi:
776 10.1080/00222931108693036
- 777 Ducroz, J.F. (1998). *Contribution des approches cytogénétique et moléculaire à l'étude systématique*
778 *et évolutive des rongeurs de la "division" Arvicanthis*. Doctoral thesis, Muséum National d'Histoire
779 Naturelle, Paris, France.
- 780 Ducroz, J.F., Volobouev, V. & Granjon, L. (1998). A molecular perspective on the systematics and
781 evolution of the genus *Arvicanthis* (Rodentia, Muridae): Inferences from complete cytochrome b
782 gene sequences. *Molecular Phylogenetics and Evolution*, 10, 104-117. doi:
783 10.1006/mpev.1997.0477
- 784 Ducroz, J.F., Volobouev, V. & Granjon, L. (2001). An assessment of the systematics of arvicanthine
785 rodents using mitochondrial DNA sequences: evolutionary and biogeographical implications.
786 *Journal of Mammalian Evolution*, 8, 173-206. doi: 10.1023/A:1012277012303
- 787 Fadda, C. & Corti, M. (2001). Three dimensional geometric morphometrics of *Arvicanthis*:
788 implications for systematics and taxonomy. *Journal of Zoological Systematics and Evolutionary*
789 *Research*, 29, 585-696. doi: 10.1046/j.1439-0469.2001.00169.x
- 790 Freilich, X., Tollis, M. & Boissinot, S. (2014). Hiding in the highlands: Evolution of a frog species
791 complex of the genus *Ptychadena* in the Ethiopian highlands. *Molecular Phylogenetics and*
792 *Evolution*, 71, 157-169. doi: 10.1016/j.ympev.2013.11.01
- 793 Frick, C. (1914). A new genus and some new species and subspecies of Abyssinian rodents. *Annals of*
794 *the Carnegie Museum*, 9(1914-1915), 7-28.
- 795 Giancesini, C., Clesse, D., Tosini, G., Hicks, D. & Laurent V. (2015). Unique regulation of the melatonin
796 synthetic pathway in the retina of diurnal female *Arvicanthis ansorgei* (Rodentia). *Endocrinology*,
797 156, 3292-3308. doi: 10.1210/EN.2015-1267
- 798 Happold, D. (ed.) (2013). *Mammals of Africa: vol. 3, Rodents, hares and rabbits*. Bloomsbury
799 Publishing, London.
- 800 Honaki, J.H., Kinman, K.E. & Koeppl, J.W. (1982). *Mammal species of the world. A taxonomic and*
801 *geographic reference*. Allen Press and Association of Systematic Collection. Lawrence.
- 802 Konečný A., Koubek, P. & Bryja, J. (2010). Indications of higher diversity and abundance of small
803 rodents in human-influenced Sudanian savanna than in the Niokolo Koba National Park (Senegal).
804 *African Journal of Ecology*, 48, 718-726. doi: 10.1111/j.1365-2028.2009.01169.x

- 805 Kostin, D.S., Kasso, M., Komarova, V.A., Martynov, A.A., Gromov, A.R., Alexandrov, D.Y., Bekele, A.,
806 Zewdie, C., Bryja, J. & Lavrenchenko, L.A. (2018). Taxonomic and genetic diversity of rodents from
807 the Arsi Mountains (Ethiopia). *Mammalia*, in press. doi: 10.1515/mammalia-2017-0135
- 808 Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura K. (2018). MEGA X: Molecular Evolutionary Genetics
809 Analysis across computing platforms. *Molecular Biology and Evolution*, 35, 1547-1549. doi:
810 10.1093/molbev/msy096
- 811 Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2016). PartitionFinder 2: new
812 methods for selecting partitioned models of evolution for molecular and morphological
813 phylogenetic analyses. *Molecular Biology and Evolution*, 34, 772-773. doi:
814 10.1093/molbev/msw260
- 815 Langel, J. L., Smale, L., Esquivá, G. & Hannibal, J. (2015). Central melanopsin projections in the diurnal
816 rodent, *Arvicanthis niloticus*. *Frontiers in Neuroanatomy*, 9, article no. 93. doi:
817 10.3389/fnana.2015.00093
- 818 Lavrenchenko, L.A., Verheyen, W.N., Verheyen, E., Hulselmans, J., Leirs, H. (2007). Morphometric and
819 genetic study of Ethiopian *Lophuromys flavopunctatus* Thomas, 1888 species complex with
820 description of three new 70-chromosomal species (Muridae, Rodentia). *Bulletin de l'Institut Royal*
821 *des Sciences Naturelles de Belgique, Biologie*, 77, 77-117.
- 822 Lavrenchenko, L., Verheyen, E., Potapov, S., Lebedev, V., Bulatova, N., Aniskin, V., Verheyen, W. &
823 Ryskov, A., 2004. Divergent and reticulate processes in evolution of Ethiopian *Lophuromys*
824 *flavopunctatus* species complex: evidence from mitochondrial and nuclear DNA differentiation
825 patterns. *Biological Journal of the Linnean Society*, 83, 301-316. doi: 10.1111/j.1095-
826 8312.2004.00390.x
- 827 Meheretu, Y., Čížková, D., Těšíková, J., Welegerima, K., Tomas, Z., Kidane, D., Girma, K., Schmidt-
828 Chanasit, J., Bryja, J., Günther, S., Bryjová, A., Leirs, H. & Goüy de Bellocq J. (2012). High diversity
829 of RNA viruses in rodents, Ethiopia. *Emerging Infectious Diseases*, 18, 2047-2050. doi:
830 10.3201/eid1812.120596
- 831 Meheretu, Y., Sluydts, V., Welegerima, K., Bauer, H., Teferi, M., Yirga, G., Mulungu, L., Haile, M.,
832 Nyssen, J., Deckers, J., Makundi, R. & Leirs, H. (2014). Rodent abundance, stone bund density and
833 its effects on crop damage in the Tigray highlands, Ethiopia. *Crop Protection*, 55, 61-67. doi:
834 10.1016/j.cropro.2013.10.016
- 835 Miller, M.A., Pfeiffer, W. & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of
836 large phylogenetic trees. In *Gateway Computing Environments Workshop (GCE)*. IEEE (pp. 1-8).
- 837 Monadjem, A., Taylor, P.J., Denys, C. & Cotterill, F.P. (2015). Rodents of sub-Saharan Africa: a
838 biogeographic and taxonomic synthesis. De Gruyter, Berlin.

- 839 Musser, G.G. & Carleton, M.D. (2005). Superfamily Muroidea. In D.E. Wilson & D.M. Reeder (Eds.),
840 *Mammal species of the world: a taxonomic and geographic reference*, 3rd edition. Johns Hopkins
841 University Press, Baltimore.
- 842 Ogilvie, H.A., Bouckaert, R.R. & Drummond, A.J. (2017). StarBEAST2 brings faster species tree
843 inference and accurate estimates of substitution rates. *Molecular Biology and Evolution*, 34, 2101-
844 2114. doi: 10.1093/molbev/msx126
- 845 Orlov, V.N., Baskevich, M.I. & Bulatova, N.S. (1992). Chromosomal sets of rats of the genus
846 *Arvicanthis* from Ethiopia. *Zoologicheskii Zhurnal*, 71, 103-112.
- 847 Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018). Posterior summarisation in
848 Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, syy032. doi:10.1093/sysbio/syy032
- 849 Ronquist, F. & Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed
850 models. *Bioinformatics*, 19, 1572-1574. doi: 10.1093/bioinformatics/btg180
- 851 Rüppel, E. (1842). Säugethiere aus der Ordnung der Nager, beobachtet im nordöstlichen Africa.
852 *Museum Senckenbergianum: Abhandlungen aus dem Gebiete der beschreibenden*
853 *Naturgeschichte*, 3 (1833-1845), 91-116.
- 854 Sicard, B., Diarra, W. & Cooper, H.M. (1999). Ecophysiology and chronobiology applied to rodent pest
855 management in semi-arid agricultural areas in Sub-Saharan West Africa. In G.R Singleton, L.A.
856 Hinds, H. Leirs & Z. Zhang (Eds.), *Ecologically-based management of rodent pests* (pp. 409-440).
857 ACIAR monographs no. 59, Canberra.
- 858 Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
859 phylogenies. *Bioinformatics*, 30, 1312-1313. doi: 10.1093/bioinformatics/btu033
- 860 Šumbera, R., Krásová, J., Lavrenchenko, L.A., Mengistu, S., Bekele, A., Mikula, O. & Bryja, J. (2018).
861 Ethiopian highlands as a cradle of the African fossorial root-rats (genus *Tachyoryctes*), the genetic
862 evidence. *Molecular Phylogenetics and Evolution*, 126, 105-115. doi:
863 10.1016/j.ympev.2018.04.003
- 864 Trauth, M.H., Maslin, M.A., Deino, A.L., Junginger, A., Lesoloyia, M., Odada, E.O., Olago, D.O., Olaka,
865 L.A., Strecker, M.R. & Tiedemann, R. (2010). Human evolution in a variable environment: the
866 amplifier lakes of Eastern Africa. *Quaternary Science Reviews*, 29, 2981-2988. doi:
867 10.1016/j.quascirev.2010.07.007
- 868 Vaughan, T.G. (2017). IcyTree: rapid browser-based visualization for phylogenetic trees and
869 networks. *Bioinformatics*, 33, 2392-2394. doi: 10.1093/bioinformatics/btx155.
- 870 Volobouev, V.T., Ducroz, J.-F., Aniskin, V.M., Britton-Davidian, J., Castiglia, R., Dobigny, G., Granjon,
871 L., Lombard, M., Corti, M., Sicard, B. & Capanna, E. (2002). Chromosomal characterization of
872 *Arvicanthis* species (Rodentia, Murinae) from western and central Africa: implications for
873 taxonomy. *Cytogenetic and Genome Research*, 96, 250-260. doi: 10.1159/000063041

- 874 Yalden, D.W., Largen, M.J. & Kock, D. (1976). Catalogue of mammals of Ethiopia. 2. Insectivora and
875 Rodentia. *Monitore Zoologico Italiano*, suppl. 8, 1-118.
- 876 Yalden, D.W., Largen, M.J., Kock, D. & Hillman, J.C. (1996). Catalogue of the mammals of Ethiopia
877 and Eritrea. Revised checklist, zoogeography and conservation. *Tropical Zoology*, 9, 73-164. doi:
878 10.1080/03946975.1996.10539304
- 879 Zhang, C., Ogilvie, H.A., Drummond, A.J. & Stadler T. (2017). Bayesian inference of species networks
880 from multilocus sequence data. *Molecular Biology and Evolution*, 35, 504-517. doi:
881 10.1093/molbev/msx307

For Review Only

Figure legends

Fig. 1 Mitochondrial BI tree of *Arvicanthis* based on the alignment of 1140 bp of *CYTB*. The colour of squares indicate posterior probability from MrBayes, circles indicate statistical support from 1000 bootstraps in ML analysis (see legend for range of values represented by each colour). Only one outgroup is shown. The colour bars indicate separate species discussed in the paper. Information about karyotypes, where available, was placed on the tree.

Fig. 2 Geographical distribution of genetic lineages within the two major *Arvicanthis* groups. (A) NILOTICUS group; (B) ANSORGEI group. Different species are marked by different symbols, whose colours and names correspond to Fig. 1. The symbols grouped by dashed lines belong to separate subclades within the species. The distribution of western African taxa is not shown, i.e. for *A. niloticus* "C2-C4" only the easternmost C2 clade is shown.

Fig. 3 Genetic diversity of the NILOTICUS group in Ethiopia. Different species are marked by different symbols, whose colours and names correspond to Fig. 1. The dashed lines indicate three separate subclades of *A. niloticus* s. str. The type localities of four taxa described from Ethiopia and proposed as Ethiopian endemic species are shown by arrows and corresponding names (see Taxonomic implications).

Fig. 4 Nuclear BI tree based on partitioned analysis of concatenated dataset of DNA sequences of two exons and four introns. Parsimony informative indels were also included and encoded as morphological characters. The numbers above branches indicate posterior probabilities estimated in MrBayes/bootstrap support in ML tree produced by partitioned analysis of the same dataset in RAxML (without indels). "-" means that topology was different in ML tree. Coloured squares indicate informative indels in particular introns: orange = 1 bp insertion in *DHCR*, blue = 3 bp insertion in *SMO*, light green = 3 bp insertion in *WLS*, pink = 2 bp deletion in *DHCR*, yellow = 8 bp insertion in *DHCR* + 38 bp deletion in *TRPV*, dark green = 22 bp deletion in *TRPV* + 1 bp insertion in *SMO*.

Fig. 5 (A) Species tree of the NILOTICUS group reconstructed by multi-species coalescent approach in *BEAST (Helled and Drummond 2010). The numbers above branches indicate posterior probabilities of particular clades. (B) The most probable Bayesian species network calculated in BEAST2 from nuclear multilocus data using the approach of Zhang et al. (2018). It is one of the 444 networks and its posterior probability is 0.02. We therefore summarized the topologies of 10 networks with highest posterior probabilities (in sum they had PP = 0.128) and the numbers above branches indicate in how

many of them the same group of species was clustered. The taxon *blicki* "DS" represent the high elevation population from Debre Sina, west of GRV (see Fig. 3).

Fig. 6 Divergence dating of the species tree inferred using a multi-species coalescent approach and the combination of both mitochondrial and nuclear markers in StarBEAST 2. The numbers in circles are medians of TMRCA of particular clades (with 95% HPD indicated by red bars). PP = posterior probability for each clade; note that the topology of the NILOTICUS group is not resolved, indicating rapid radiation around 1.5-2 Mya.

Fig. 7 Distribution of *Arvicanthis* taxa in sub-Saharan Africa based on DNA sequences (Ducroz et al. 1998, Abdel Rahman Ahmed et al. 2008, Dobigny et al. 2011, 2013, this study), cytogenetics (Volobouev et al. 2002, Castiglia et al. 2006) and geometric morphometrics (Fadda and Corti 2001). Particular species include localities of taxa as specified in Table 1. The proposed names for Ethiopian species are shown in the legend. See text for more details.

933 **List of Supporting Information**

934

935 **Appendix S1** List of used specimens with details on their karyotypes, DNA sequences and localities.

936

937 **Appendix S2** Additional details about the used methods (primers and PCR conditions, list of outgroup
938 sequences and used substitution models.

939

940 **Appendix S3** Bayesian gene trees of six nuclear markers.

941

For Review Only

942

943 **Tables**944 **Table 1** Species names of *Arvicanthis* used in previous studies and their correspondence with the names used in this study.

Group	Species (this study)	Ducroz et al. (1998)	Ducroz et al. (2001)	Fadda & Corti (2001)	Volobouev et al. (2002)	Castiglia et al. (2006)	Abdel Rahman Ahmed et al. (2008)	Dobigny et al. (2011)	Dobigny et al. (2013)	Proposed name	Type locality
ANSORGEI	<i>ansorgei</i>	sp 1-7	sp 3	ANI-3	ANI-3	-	ansorgei	-	ansorgei	<i>ansorgei</i> Thomas, 1910	Gunnal, Guinea- Bissau
	<i>nairobae</i>	nairobae, sp 11	sp 2	sp1	-	nairobae	-	-	-	<i>nairobae</i> J.A.Allen, 1909	Nairobi, Kenya
	sp. "ANI-6"	-	-	sp3	-	ANI-6, ANI- 6a	-	-	-	<i>raffertyi</i> Frick, 1914	Gardula, Ethiopia
	<i>rufinus</i>	sp 8-10	sp 1	ANI-2, ANI-4	ANI-2, ANI-4	-	rufinus	ANI-2, ANI-4, rufinus	ANI-2, cf. rufinus	<i>rufinus</i> (Temminck, 1853)	Elmina, Ghana
NILOTICUS	<i>abyssinicus</i>	-	-	-	-	-	-	-	-	<i>abyssinicus</i> (Rüppel, 1842)	Entschetqab, Simien Mts., Ethiopia
	<i>blicki</i>	-	-	blicki	-	blicki	-	-	-	<i>blicki</i> Frick, 1914	Chilalo Mts., Ethiopia
	<i>niloticus</i> s. str.	dembeensis, niloticus	niloticus	niloticus	-	niloticus	dembeensi, niloticus	niloticus	niloticus C1	<i>niloticus</i> (É. Geoffroy, 1803)	Egypt
	sp. "Masai Mara"	-	-	sp. (?)	-	niloticus	-	-	cf. niloticus	unresolved	-
	<i>neumanni</i>	somalicus	neumanni	sp4	-	neumanni	neumanni	somalicus	neumanni	<i>neumanni</i>	Burungo,

										(Matschie, 1894)	Irangi, Tanzania
	<i>niloticus</i> "C2-C4"	niloticus	-	ANI-1, sp. (?), sp2, testicularis	ANI-1a, ANI- 1b	niloticus, ANI-5	niloticus	niloticus	niloticus C2, C3, C4	unresolved	-
	<i>somalicus</i>	-	-	somalicus	-	somalicus, cf. somalicus	-	-	-	<i>somalicus</i> Thomas, 1903	Shuk, Northern Somaliland
	<i>sp.</i> "Menangesha"	abyssinicus	abyssinicus	abyssinicus	-	abyssinicus	abyssinicus	-	abyssinicus	<i>saturatus</i> Dollman, 1911	Didessa River, Ethiopia
	<i>sp.</i> "Metahara"	-	-	-	-	niloticus	-	-	-	<i>mearnsi</i> Frick, 1914	Sadi Malka, Ethiopia
not sequenced	ANI-7	-	-	-	-	ANI-7 (Gambella, Ethiopia)	-	-	-	unresolved	-
	ANI-8	-	-	-	-	ANI-8 (Somalia)	-	-	-	unresolved	-

945

946

947

Table 2 Mean genetic distances between groups calculated as uncorrected p -distances in MEGA using the same dataset as for mitochondrial phylogeny, i.e. 121 sequences. Standard error estimates are shown above the diagonal and were obtained by a bootstrap procedure (1000 replicates). Shaded cells indicate genetic distances within the NILOTICUS and ANSORGEI groups, respectively.

Species	<i>niloticus</i> "C2-C4"	<i>niloticus</i> s. str.	sp. "Menangesha"	sp. "Metahara"	<i>somalicus</i>	<i>neumanni</i>	sp. "Masai Mara"	<i>abyssinicus</i>	<i>blicki</i>	sp. "ANI-6"	<i>ansorgei</i>	<i>nairobae</i>	<i>rufinus</i>
<i>niloticus</i> "C2-C4"	-	0.0066	0.0071	0.0080	0.0075	0.0076	0.0074	0.0082	0.0074	0.0085	0.0088	0.0081	0.0080
<i>niloticus</i> s. str.	0.085	-	0.0073	0.0088	0.0088	0.0078	0.0079	0.0081	0.0075	0.0080	0.0082	0.0084	0.0087
sp. "Menangesha"	0.096	0.086	-	0.0083	0.0079	0.0077	0.0076	0.0085	0.0075	0.0090	0.0092	0.0085	0.0087
sp. "Metahara"	0.101	0.114	0.109	-	0.0082	0.0081	0.0084	0.0088	0.0082	0.0091	0.0088	0.0085	0.0083
<i>somalicus</i>	0.101	0.107	0.101	0.100	-	0.0077	0.0076	0.0085	0.0080	0.0094	0.0094	0.0086	0.0090
<i>neumanni</i>	0.106	0.105	0.099	0.108	0.105	-	0.0074	0.0084	0.0074	0.0089	0.0089	0.0086	0.0080
sp. "Masai Mara"	0.095	0.103	0.102	0.103	0.097	0.091	-	0.0082	0.0075	0.0088	0.0090	0.0079	0.0081
<i>abyssinicus</i>	0.097	0.106	0.116	0.111	0.100	0.105	0.105	-	0.0050	0.0089	0.0095	0.0085	0.0086
<i>blicki</i>	0.101	0.111	0.115	0.116	0.103	0.105	0.104	0.051	-	0.0082	0.0087	0.0082	0.0082
sp. "ANI-6"	0.137	0.127	0.142	0.144	0.141	0.145	0.133	0.136	0.137	-	0.0060	0.0061	0.0081
<i>ansorgei</i>	0.132	0.135	0.144	0.137	0.142	0.147	0.135	0.143	0.141	0.069	-	0.0064	0.0079
<i>nairobae</i>	0.140	0.133	0.147	0.143	0.135	0.141	0.129	0.140	0.141	0.081	0.094	-	0.0074
<i>rufinus</i>	0.132	0.131	0.139	0.137	0.137	0.133	0.123	0.132	0.131	0.117	0.125	0.113	-

Fig. 2

Fig. 3

(A) Species tree from *BEAST

(B) Species network

3.0E-4

Fig. 7