Title

Fossils know it best: using a new set of fossil calibrations to improve the temporal phylogenetic framework of murid rodents (Rodentia: Myomorpha: Muroidea: Muridae)

Authors

Tatiana Aghová^{a,b,c,*}, Yuri Kimura^d, Josef Bryja^{a,c}, Gauthier Dobigny^{e,f}, Laurent Granjon^e, Gael J. Kergoat^g

Authors' affiliations

^a *Institute of Vertebrate Biology of the Czech Academy of Sciences, Květná 8, 603 65, Brno, Czech Republic*

^b *Department of Zoology, National Museum, Václavské náměstí 68, 115 79, Prague, Czech Republic*

^c *Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, 611 37, Brno, Czech Republic*

^d *Department of Geology and Paleontology, National Museum of Nature and Science, 4-1-1 Amakubo, Tsukuba, 305-0005 Ibaraki, Japan*

^e *CBGP, IRD, CIRAD, INRA, Montpellier SupAgro, Univ. Montpellier, Montpellier, France*

^f *Ecole Polytechnique d'Abomey-Calavi, Abomey-Calavi University, 01BP2009, Cotonou, Benin*

^g *CBGP, INRA, CIRAD, IRD, Montpellier SupAgro, Univ. Montpellier, Montpellier, France*

*Corresponding author: tatiana.aghova@gmail.com

Running title: Dated phylogeny of Muridae

Abstract

Murid rodents (Rodentia: Myomorpha: Muroidea: Muridae) represent the most diverse and abundant mammalian group. In this study, we reconstruct a dated phylogeny of the family using a multilocus dataset (six nuclear and nine mitochondrial gene fragments) encompassing 160 species representing 82 distinct murid genera from four extant subfamilies (Deomyinae, Gerbillinae, Lophiomyinae, and Murinae). In comparison with previous studies on murid or muroid rodents, our work stands out for the implementation of multiple fossil constraints within the Muridae thanks to a thorough review of the fossil record. Before being assigned to specific nodes of the phylogeny, all potential fossil constraints were carefully assessed; they were also subjected to several cross-validation analyses. The resulting phylogeny is consistent with previous phylogenetic studies on murids, and recovers the monophyly of all sampled murid subfamilies and tribes. Based on nine controlled fossil calibrations, our inferred temporal timeframe indicates that the murid family likely originated in the course of the Early Miocene, 23.0-16.0 million years ago (Ma), and that most major lineages (i.e. tribes) have started diversifying *ca.* 10 Ma. Historical biogeography analyses support the Paleotropical origin for the family, with an initial internal split (vicariance event) followed by subsequent migrations between Afrotropical and Indomalayan lineages. During the course of their diversification, the biogeographic pattern of murids is marked by several dispersal events toward the Australasian and the Palearctic regions, mostly from the Indomalaya. The Afrotropical region was also secondarily colonized at least three times from the Indomalaya, indicating that the latter region has acted as a major centre of diversification for the family.

Keywords

Fossils, historical biogeography, molecular dating, Muridae, *Mus*, *Rattus*

1. Introduction

With about 150 genera and more than 730 recognized species, Muridae is the most diverse family of mammals (Musser and Carleton, 2005). Collectively murids have colonized highly distinct ecological niches, adapting to a wide array of environments ranging from warm (deserts or tropical forests) to cold habitats (high altitude mountain ranges, tundra; Vaughan et al., 2011). Life habits in murids are also diverse, as the family encompasses amphibious, arboreal, fossorial, or terrestrial taxa (Michaux et al., 2007; Musser and Carleton, 2005).

All murid species are native to the Old World (Musser and Carleton, 2005), but some species (especially the black rat *Rattus rattus* Linnaeus, the Norway rat *Rattus norvegicus* (Berkenhout) and the house mouse *Mus musculus* Linnaeus) now have a worldwide distribution due to commensalism and dissemination by humans. Murid species diversity is especially high in the Australasian and Indomalayan regions which accommodate half of the species diversity of the family (Rowe et al., 2016a). Second to that is the species diversity in the Afrotropical region (more than 200 species; Musser and Carleton, 2005). By contrast, there are much less native murid taxa in the Palearctic region (e.g. *Apodemus* Kaup, *Diplothrix* Thomas, or *Tokudaia* Kuroda).

The history of murid systematics is complex and convoluted with numerous changes occurring in the past sixty years (see Table 1 for a summary). Simpson (1945) divided representatives of family Muridae (as currently understood) into two separate families: Cricetidae (with subfamilies Gerbillinae, Lophiomyinae and others) and Muridae (subfamilies Murinae and Otomyinae). Chaline et al. (1977) considered “murid” rodents to belong to four families: Cricetidae (including Lophiomyinae), Gerbillidae, Muridae (exclusively Murinae) and Nesomyidae (including Otomyinae). Lavocat (1978) simplified this classification by recognizing only two families: Muridae (Murinae) and Nesomyidae (in which he included Gerbillinae, Lophiomyinae and Otomyinae). Another major change was later made by Carleton

and Musser (1984), who defined family Muridae in the broad sense with no less than 14 subfamilies (including Gerbillinae, Lophiomyinae, Murinae and Otomyinae). Following the introduction of molecular systematics, changes in the classification of family Muridae continued at a fast rate. Using molecular phylogenetics Chevret et al. (1993a) demonstrated that *Acomys* I. Geoffroy is not a member of the subfamily Murinae but belongs to a separate monophyletic clade including *Deomys* Thomas, *Lophuromys* Peters and *Uranomys* Dollman. All four genera were assigned to the subfamily Deomyinae, which is closely related to the Gerbillinae. In another study, Chevret et al. (1993b) showed that Otomyinae are closely allied to the tribe Arvicanthini, thus unequivocally constituting a subset of the subfamily Murinae at the tribe level (Ducroz et al., 2001; Jansa and Weksler, 2004). Jansa and Weksler (2004) also strongly suggested that Lophiomyinae belonged to the Muridae. Only part of these proposals was followed by Musser and Carleton (2005) who recognized the following five subfamilies in the family Muridae: Deomyinae, Gerbillinae, Leimacomyinae, Murinae and Otomyinae. Nowadays the most consensual classification agrees on five subfamilies: Deomyinae (four genera and *ca.* 42 species), Gerbillinae (16 genera and *ca.* 103 species), Leimacomyinae (only one species, possibly extinct; Kingdon, 2015), Lophiomyinae (only one species) and Murinae (129 genera and *ca.* 584 species; see the review of Granjon and Montgelard, 2012).

Musser and Carleton's (2005) comprehensive catalogue listed 730 species in the family Muridae. Estimates of species diversity in this family are very likely not definitive, as new murid taxa are being regularly described (e.g. Carleton et al., 2015; Esselstyn et al., 2015; Missoupe et al., 2016; Mortelliti et al., 2016; Rowe et al., 2016a). Expected and ongoing raise in species number can be accounted for by an increased focus on poorly known regions with high levels of endemism, especially in tropical Asia and Africa. It is also linked with the development of integrative taxonomy studies, where molecular genetic approaches are able to

detect taxa and geographical regions with high cryptic diversity (e.g. Bryja et al., 2014, 2017; Ndiaye et al., 2016).

Because of the high species richness of the family, determining the precise timing of its radiation is of particular paleobiogeographic interest. Several dated estimates for the age of Muridae are available owing to studies either focusing on the order Rodentia (Adkins et al., 2001, 2003; Fabre et al., 2012; Montgelard et al., 2008), on the superfamily Muroidea (Muroidea; Schenk et al., 2013 ; Steppan et al., 2004) or on various murid subsets (e.g. Bryja et al., 2014; Chevret and Dobigny, 2005; Dobigny et al., 2013; Fabre et al., 2013; Pagès et al., 2016; Rowe et al., 2008, 2011, 2016b). However, no clear consensus could be reached for the age of the family Muridae. Indeed, age estimates derived from all aforementioned studies are far from being congruent, likely because their datasets have not been designed for this particular purpose. In addition, all these studies used very diverse dating procedures, some of them relying on substitution rate calibrations (e.g. Arbogast et al., 2001; Nicolas et al., 2008) whereas others used fixed ages (e.g. the putative *Mus/Rattus* split at 12 Ma; Steppan et al., 2004), very distant fossil constrains (Adkins et al., 2003; Fabre et al., 2012; Montgelard et al., 2008) or primary calibrations using various fossil constraints within or outside the family Muridae (e.g. Bryja et al., 2014; Pagès et al., 2016; Rowe et al., 2016b; Schenk et al., 2013).

For fossil-based calibrations of molecular clocks, it is crucial: (i) to properly assign and place fossils on the tree, and (ii) to correctly estimate the age of fossil-bearing formations (Parham et al., 2012; Sauquet et al., 2012). Unfortunately the fossil record of oldest murids is quite fragmentary and mostly consists of isolated teeth and mandible remains, thus sometimes making taxonomic identification difficult. The earliest representatives for the family Muridae include the tribe Myocricetodontini with genera such as †*Myocricetodon* Lavocat, †*Dakkamys* Jaeger and †*Mellalomys* Jaeger (Jacobs and Flynn, 2005; Lazzari et al., 2011). Extinct members of the genus †*Potwarmus* Lindsey could be considered as a stem group of the subfamily

Murinae based on detailed analyses of dental morphology (Lazzari et al., 2011; López Antoñanzas, 2009; Wessels, 2009). The earliest unequivocal representative of the subfamily Murinae is the genus †*Antemus* Jacobs (Jacobs and Downs, 1994; Jacobs and Flynn, 2005; Kimura et al., 2015). †*Antemus* possesses a new cusp (anterostyle, also known as t1), which is a synapomorphy of Murinae. The earliest record of †*Antemus chinjiensis* is dated at 13.8 Ma (Jacobs et al., 1990) based on specimens from the locality YGSP 491, Chinji Formation in the Potwar Plateau, Pakistan (Jacobs, 1977). In the fossil record of the Potwar Plateau, two more derived fossil genera are of particular interest: †*Karnimata* Jacobs and †*Progonomys* Schaub. Based on the relative position of the anterostyle to the lingual anterocone on M1, Jacobs (1978) hypothesized that †*Karnimata* is related to *Rattus* and that †*Progonomys* is a member of the lineage including *Mus*. Hence, their first stratigraphic occurrence has been used to define the widely used *Mus/Rattus* calibration (ca. 12 My; Jacobs and Downs, 1994). However, in 2015, Kimura et al. revisited these fossils from a paleontological perspective and proved this calibration point to be controversial. They showed that †*Karnimata* is a member of the Arvicanthini-Millardini-Otomyini clade rather than a member of the lineage encompassing the genus *Rattus* and its relatives (i.e. tribe Rattini). Therefore, they demonstrated that the continuous fossil record of the murine rodents from the Potwar Plateau actually provides a minimum age for the most recent common ancestor of the lineages leading to *Arvicanthis* Lesson and *Mus* (= *Mus/Arvicanthis* split).

Recent progresses in divergence dating analyses lead us to revisit results previously obtained by favouring the implementation of a new set of well-justified primary fossil calibrations within a Bayesian framework. In comparison to previous studies (listed above), our study can be considered as medium-sized in terms of taxonomic sampling, and essentially focused on the family Muridae. But our study stands out for rigorous evaluation of the fossil data for this highly diverse mammalian family. The present study has four main objectives: (i)

to design a comprehensive multi-marker molecular dataset for the family Muridae, (ii) to review the murid fossil record in order to identify reliable and suitable primary fossil calibrations, (iii) to provide a reliable estimate of the timing of diversification of the family using multiple fossil calibrations, and (iv) to lean on the resulting dated phylogeny to reconstruct the biogeographic history of the family using up-to-date analytical approaches.

2. Material and Methods

2.1. Taxon sampling

For this study, new DNA sequences were generated for five murid species (*Acomys cf. cineraceus*, *Acomys subspinosus* (Waterhouse), *Acomys wilsoni* Thomas, *Arvicanthis niloticus* (Desmarest), *Arvicanthis neumanni* (Matschie) see Appendix A). Though we largely relied on GenBank data for this work, it is worth underlining that our research group generated thousands of murid sequences (all deposited in GenBank) in the past 15 years (we used some of these sequences for 44 species included in this study). In total, our dataset (Appendix A) encompasses 160 murid species representing 82 of the 151 known murid genera. All four extant subfamilies (if considering the Togo mouse from Leimacomysinae to be extinct) of Muridae are included. For the largest subfamily Murinae, we included representatives of all 10 tribes that have been defined by Lecompte et al. (2008): Apodemini, Arvicanthini, Hydromyini, Malacomysini, Millardini, Murini, Otomyini, Phloeomyini, Praomyini and Rattini. As outgroup taxa, we selected five species of the family Cricetidae (from subfamilies Arvicolinae, Cricetinae, Neotomyinae and Tylomyinae), which constitutes the sister group of Muridae (Fabre et al., 2012). Finally, the tree was rooted using *Calomyscus baluchi* Thomas, a representative of the more distant family Calomyscidae (Fabre et al., 2012). All species names followed Musser and Carleton (2005) and Monadjem et al. (2015).

2.2. DNA extraction, sequencing and molecular matrix

DNA was extracted using a Qiagen[®] DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. Two nuclear gene fragments were targeted using the following combinations of polymerase chain reactions (PCR) primers: IRBP217 and IRBP1531 (Stanhope et al., 1992) for the fragment of the 'interphotoreceptor retinoid binding' (IRBP) gene; RAG1F1705 and RAG1R2951 (Teeling et al., 2000) for a fragment of the 'recombination activating gene 1' (RAG1) gene. For PCR protocols, see Bryja et al. (2017) and Teeling et al. (2000), respectively. PCR products were Sanger sequenced in both directions using the BigDye[®] Terminator chemistry (Thermo Fisher Scientific) either in the Institute of Vertebrate Biology on an 'Applied Biosystems[®] 3130xl Genetic Analyzer', or commercially through the GATC Biotech company (Konstanz, Germany). New sequences were deposited in GenBank under accession numbers KY634246 to KY634250.

The newly generated sequences were further combined with data from GenBank. The resulting matrix (see Appendix A) encompasses the following six nuclear and nine mitochondrial gene fragments: 'acid phosphatase 5' (AP5), BRCA1, intronic portion of 'Peripheral benzodiazapine receptor variant' (BZRP), 'growth hormone receptor' (GHR), IRBP and RAG1, for the nuclear genes, and '12S ribosomal RNA' (12S), '16S ribosomal RNA' (16S), 'ATP synthase 8' (ATPase8), 'cytochrome c oxidase I' (COI), 'cytochrome oxidase II' (COII), 'cytochrome *b*' (Cytb), 'Aspartic acid transfer RNA' (tRNA-Asp), 'Lysine transfer RNA' (tRNA-Lys), 'Serine transfer RNA' (tRNA-Ser), for the mitochondrial genes. For nine taxa (*Acomys* cf. *cineraceus*, *Acomys wilsoni* Thomas, *Aethomys chrysophilus* (de Winton), *Aethomys hindei* (Thomas), *Aethomys kaiseri* (Noack), *Aethomys silindensis* Roberts, *Arvicanthis nairobae* J.A. Allen, *Arvicanthis neumanni* and *Thallomys paedulus* (Sundevall), gene fragments were concatenated from two individuals to minimize the amount of missing data. For all protein-coding genes, we used Mesquite 3.2 (Maddison and Maddison, 2007) to

check the coding frame for possible errors or stop codons. The sequences of several markers (i.e. 12S, 16S, intronic portion of BZRP, tRNA-Asp, tRNA-Lys and tRNA-Ser) were variable in length; their alignment was accomplished using MUSCLE (Edgar, 2004) with default settings.

2.3. Phylogenetic analyses

Phylogenetic analyses were conducted using both Bayesian inference (BI) and maximum likelihood (ML). Analyses were performed on the online computer cluster CIPRES Science Gateway (Miller et al., 2010; www.phylo.org) and on the high performance computing (HPC) cluster hosted in the Centre de Biologie pour la Gestion des Populations (CBGP) in Montferrier-sur-Lez, France. For both phylogenetic analytical approaches, we carried out partitioned analyses to improve phylogenetic accuracy (Nylander et al., 2004). The molecular dataset was divided *a priori* into 33 partitions: we used three partitions for each of the protein-coding genes (AP5, ATPase8, BRCA1, COI, COII, Cytb, GHR, IRBP and RAG1) and one partition for each of the rRNA-tRNA genes (12S, 16S, tRNA-Asp, tRNA-Lys and tRNA-Ser) as well as the BZRP intronic portion. The best partitioning scheme and substitution models were determined with PartitionFinder 1.1.1 (Lanfear et al., 2014) using a greedy heuristic algorithm; because of the risk of over-parameterization associated with the high number of specified partitions, the ‘unlinked branch lengths’ option was chosen over the ‘linked branch lengths’ option. The Bayesian information criterion (BIC) was also preferentially used to compare partitioning schemes and substitution models following the recommendation of Ripplinger and Sullivan (2008).

PartitionFinder (based on BIC) identified the same three partitions for both BI and ML analyses: two partitions are associated with a Generalized-Time-Reversible (GTR + Γ +I) model

and one partition is associated with a Hasegawa-Kishino-Yano (HKY + Γ +I) model (see Table 2).

Bayesian inference analyses were carried out using MrBayes v3.2.6 (Ronquist et al., 2012b). Two independent runs with four MCMC (one cold and three incrementally heated chains) were conducted: they ran for 50 million generations, with trees sampled every 1,000 generations. A conservative 25% burn-in was applied after checking for stability on the log-likelihood curves and the split-frequencies of the runs. Support of nodes for MrBayes analyses was provided by clade posterior probabilities (PP) as directly estimated from the majority-rule consensus topology. Following Erixon et al. (2003), nodes supported by $PP \geq 0.95$ were considered strongly supported.

Maximum likelihood analyses were performed using RAxML v8.2.8 (Stamatakis, 2014). Because this software does not allow simpler substitution models, we used three partitions with a General Time Reversible (GTR + Γ +I) model (see Table 2). The best ML tree was obtained using heuristic searches with 100 random addition replicates. Clade support was then assessed using a non-parametric bootstrap procedure with 1,000 replicates. Following Hillis and Bull (1993), nodes supported by bootstrap values (BV) ≥ 70 were considered strongly supported.

2.4. Evaluation of suitable fossil calibrations

Following the recommended criteria of Parham et al. (2012) for fossil calibrations, we rigorously compiled a list of potential candidates from the paleontological literature and eventually retained 18 candidate fossils (see Table 3). The candidate fossils possess the information for the collection site, unique identification number, and the state of preservation along with justification for the age of the fossil (i.e., age of fossil-bearing formation and stratigraphic level, preferably with an absolute age by radiometric dating and/or reliable relative age estimates, for example, by magnetostratigraphy).

In the next step, diagnostic morphological characters were reassessed to determine whether they could be reliably used as minimum age constraints in our phylogeny, either as crown or stem calibrations. Seven fossils were discarded following this step (see the ‘Results’ section).

For the remaining 11 fossils, we used the cross-validation procedure developed by Near and Sanderson (2004) and Near et al. (2005). The following approach was used: (i) we identified potential inconsistencies within the 11 remaining fossil calibrations, and (ii) we explored the impact of the inclusion of each of these fossils on our divergence time estimates. Each of the 11 fossil constraints was enforced at a time in a specific Bayesian relaxed-clock (BRC) analysis to estimate the ages of the remaining nodes (see also section 2.5). First, the sum of the squared differences between the molecular and fossil age estimates (SS) was calculated (for more details see Near and Sanderson, 2004). All calibration points were then ranked based on the magnitude of its SS score; here the fossil with the greatest SS score is assumed to be the most inconsistent with respect to all other fossils in the analysis (Near and Sanderson, 2004). Second, we calculated the average squared deviation, s , for all fossil calibrations in the analysis. Following the method of Near et al. (2005), we removed the fossil with the greatest SS score and recalculated s with the remaining fossil calibration points. This process was pursued until only the two fossil calibration points with the lowest and second lowest magnitudes of SS remained (Near and Sanderson, 2004). The rationale behind this procedure is to assess whether calibration points are approximately equally informative and accurate (Near et al., 2005): if it is the case the magnitude of s should only decrease by a small fraction whenever a fossil calibration is removed.

2.5. Bayesian relaxed-clock analyses

Although divergence time dating is now a well-established cornerstone of evolutionary biology, there is still no widely accepted objective methodology for converting data from the fossil record to calibration information of use in molecular phylogenies (Drummond and Bouckaert, 2015). In the last few years, several methodological approaches to better implement fossil calibrations have been developed, for instance allowing one to directly include fossil lineages in phylogenies ('total-evidence dating'; Pyron, 2010; Ronquist et al. 2012a) or to account for information on the density of the fossil record ('fossilized birth-death (FBD) process'; Stadler, 2010; Heath et al., 2014). However, for our study, a 'total evidence dating' approach was not applicable since it would have required the coding of a morphological matrix for both fossils and extant taxa, which is problematic given the fragmentary nature of muroid fossils. The use of the FBD methodology was also not envisioned because the fossil record of muroid rodents is too sparse. Instead, we relied on a node-dating approach in which fossil information is enforced on specific nodes through the use of parametric distributions.

Following our assessment of the murid fossil record and the results of cross-validation analyses, nine fossil calibrations were finally retained for the dating procedure (for more information, see Tables 3 and 4, and Appendix B). Five of them were defined based on fossil material collected in the Siwalik Group of Pakistan (†*Antemus chinjiensis*, †*Karnimata darwini* Jacobs, †cf. *Karnimata* sp., †*Mus* sp. and †*Abudhabia pakistanensis* Flynn and Jacobs). Three additional accepted fossils originate from 6.1 Ma fossils discovered in the Lemudong'o locality in Kenya (†*Aethomys* sp., †*Arvicanthis* sp. and †*Gerbilliscus* sp.), and the last retained fossil calibration constraint is defined by the 9.6 Ma fossil of †*Parapodemus lungdunensis* Schaub. Priors for fossil constraints were defined by using either uniform or lognormal statistical distributions in two separate analyses. Statistical distributions were bounded by the minimum ages provided by the fossil constraints and a conservative maximum age (ca. 25 Ma) for the root derived from the study of Schenk et al. (2013; see Table 4). In a preliminary way (see

section 2.4 of the Material and Methods), BRC analyses were also conducted using one fossil constraint at a time to carry out cross-validation analyses.

Bayesian relaxed-clock analyses were conducted with BEAST v1.8.4 (Drummond et al., 2012) using uncorrelated lognormal relaxed clocks (Drummond et al., 2006). To limit the risk of over-parameterization: (i) we used three clock models (based on PartitionFinder results, Table 2); and (ii) we enforced a guide tree that corresponds to the topology with the best clade support (this topology corresponds to the topology obtained with MrBayes; see the ‘Results’ section). For the tree speciation model, a birth death process (Gernhard, 2008) was used in order to better account for extinct and missing lineages.

BEAST .xml files were modified to implement the path-sampling procedure for Bayes factor (B_F) estimation following the recommendations of Baele et al. (2013). Out of the two calibrations, the calibration procedure with lognormal prior has the best harmonic mean (-208117.74 *versus* -208262.58 for the procedure with a uniform prior) and is recovered as the best-fit calibration procedure with a statistically significant B_F of 289.68 ($B_F > 10$, Kass and Raftery, 1995). The final analysis (with nine verified fossil constraints and lognormal prior distribution for calibration constraints) was carried out by two independent runs each with 50 million generations and trees sampled every 5,000 generations. We used a conservative burn-in of 12.5 million generations per run. Post burn-in trees from both analyses were further combined using the LogCombiner module of BEAST. Convergence of runs was assessed graphically under Tracer v.1.6 and by examining the ESS of parameters.

2.6. Historical biogeography

Ancestral biogeography was reconstructed using the R package ‘BioGeoBEARS’ (Matzke, 2013). Data for species’ ranges were obtained from the International Union for Conservation of Nature website (<https://www.iucn.org/>). Five major biogeographic areas were

defined on the basis of Olson et al. (2001): A, West Palearctic (from Western Europe to the Ural Mountains, including North Africa); B, East Palearctic (from the Urals to Japan); C, Indomalaya (from Afghanistan through the Indian subcontinent and Southeast Asia to lowland southern China, and through Indonesia as far as Java, Bali, and Borneo, west of the Wallace line); D, Australasia (Australia, New-Zealand, Papua-New-Guinea and neighbouring small islands); E, Afrotropics (Africa, northern part excluded). Dispersal rate between adjacent areas was fixed to 1 (A-B; B-C), whereas the dispersal of 0.7 (A-E; C-D) and 0.3 (B-D; B-E) was specified for long-distance dispersal or whenever a geographical barrier had to be crossed. Dispersal was disallowed between geographical areas separated by two or more areas (A-D; D-E). Six models of geographic range evolution were compared in a likelihood framework: (i) Dispersal-Extinction Cladogenesis model (DEC) similar to Lagrange (Ree and Smith, 2008), which parameterizes dispersal and extinction; (ii) DEC +J model (Matzke, 2013; 2014), which adds founder-event speciation with long-distance dispersal (cladogenesis, where daughter lineage is allowed to jump to a new range outside the range of the ancestor; Matzke, 2013) to the DEC framework; (iii) Dispersal Vicariance Analysis (DIVA; Ronquist, 1997); (iv) DIVA with long-distance dispersal (DIVA +J; Matzke, 2013); (v) Bayesian inference of historical biogeography for discrete areas (BayArea; Landis et al., 2013); and (vi) BayArea with long-distance dispersal (BayArea +J; Matzke, 2013). Model fit was assessed using the Akaike information criterion (AIC) and likelihood-ratio tests (LRT).

3. Results

3.1. Phylogeny of Muridae

Our multilocus dataset representing all major lineages of the family Muridae is 10,482 bp long with 42.5% missing data. Both BI and ML analyses yield similar topologies (see Fig. 1 for the topology inferred under BI, and Fig. S1 in Appendix D for the best-fit ML tree), as indicated

by a high proportion of shared nodes (160 out of 162). BI and ML analyses differ only in the position of *Pelomys fallax* (Peters) and *Zyzomys argurus* (Thomas), but their placements are not significantly supported in either analysis. Clade support is moderate to high on average; if considering the number of nodes that are supported by $PP \geq 0.95$ or $BV \geq 70\%$, BI analyses yield a slightly more robust topology (135 well-supported nodes) compared to the ML tree (122 well-supported nodes).

Phylogenetic analyses confirm the monophyly of the family Muridae, of all its four constituent subfamilies, as well as of the previously defined tribes of the subfamily Murinae (Fig. 1). On the contrary, the phylogenetic position of some genera (e.g. *Acomys* I. Geoffroy, *Dasymys* Peters, *Golunda* Gray, *Melomys* Thomas, *Micaelamys* Ellerman, *Pelomys* Peters, *Oenomys* Thomas and *Otomys* F. Cuvier) within particular tribes was only partly supported.

3.2. Evaluation of suitable fossil calibrations

We summarized all fossils considered in this study in Table 3 and Appendix B regarding taxonomic information and specification for prior settings (see also Figure 2 for their respective positions within the tree). Five out of 18 preselected fossils (i.e. †*Parapelomys robertsi* Jacobs, †*Potwarmus primitivus*, †*Preacomys kikiae* Geraads, †cf. *Progonomys* sp. Schaub and †aff. *Stenocephalomys* Frick) were excluded from further analyses because the scarcity of paleontological interpretation about their phylogenetic relationships impeded assigning them to specific nodes of the phylogeny (Appendix B). We also excluded fossils of *Acomys* and *Lemniscomys* Trouessart from the Lemudong'O locality, Kenya (Manthi, 2007), because first upper molars, which possess the most diagnostic characters in the murine dentition, are not described from the locality (Table 3; see more details also in Appendix B). The two-step cross-validation procedure resulted in a further reduction of the fossil set of possible calibration points. Specifically, we excluded two fossils: one is a 2.4 Ma fossil identified as the genus

Gerbillus Desmarest, while the other one corresponds to a 6.1 Ma fossil identified as the genus *Mastomys* Thomas (Appendix C). The rationale is that (i) these two fossil calibrations exhibited the largest magnitude of SS, and that (ii) their removal also resulted in a very high (fivefold) decrease in s (see Appendix C for more details). As a result of the latter series of selection steps, nine fossils were finally retained for divergence dating (see Figure 2 for their position on the tree, Table 4 for specification of priors, and Appendix B for more details on all considered fossils).

3.3 Historical biogeography and divergence dating

Among the six models of geographic-range evolution compared in a likelihood framework in BioGeoBEARS, the Dispersal-Extinction Cladogenesis model with founder-event speciation (DEC +J) was chosen because of its best likelihood and AICc associated scores (lnL=-117.1, AICc=240.4; Table 5).

The dated tree resulting from the BRC analyses is shown in Figure 2 while dating estimates for all internal nodes are provided in Table 6, and results of ancestral distribution reconstructions are presented in Figure 3. The most recent common ancestor (MRCA) of Muridae originated during the early Miocene (median age of 19.3 Ma; 95% highest posterior density (HPD): 17.06-21.92 Ma) in the Afrotropical and Indomalayan bioregions.

Three subfamilies (Deomyinae, Gerbillinae and Lophiomyinae) belonging to the same clade started their diversification in the Afrotropics. Within this clade, a first split occurred *ca.* 18.6 Ma (95% HPD: 16.35-21.11 Ma) between the Lophiomyinae and the clade encompassing the Deomyinae and Gerbillinae. Deomyinae started their diversification *ca.* 14.6 Ma (95% HPD: 12.68-16.82 Ma) while Gerbillinae started theirs *ca.* 12.2 Ma (95% HPD: 10.46-14.16 Ma). In Deomyinae and Gerbillinae, several lineages were able to colonize the Palearctic region

from the Afrotropics (in our dataset, this concerns *Acomys russatus* (Wagner) and *Gerbillus gerbillus* Olivier).

The subfamily Murinae originated in the Indomalayan region during the middle Miocene (median age of 14.2 Ma; 95% HPD: 12.70-16.07 Ma); the corresponding basal split separated the Phloeomyini and all remaining murines ('core murines' *sensu* Steppan et al. 2005). The next major split occurred in the Indomalaya between Rattini and the remaining tribes of Murinae (median age of 12.3 Ma; 95% HPD: 11.28-13.62 Ma), with an origin of Rattini estimated at 10.4 Ma (95% HPD: 9.23-11.93 Ma). The ancestral area of Rattini was also inferred to be the Indomalaya; during the course of their diversification, a few taxa colonized Australasia (e.g. *Bunomys andrewsi* (J.A. Allen), *Melasmothrix naso* Miller and Hollister, *Paruromys dominator* Thomas and *Rattus leucopus* (Gray)) as well as the West and East Palearctic (e.g. *Micromys minutus* (Pallas) and *Diplothrix legata* (Thomas)). The Hydromyini split from the remaining Murinae at *ca.* 11.9 Ma (95% HPD: 10.95-13.10 Ma); although basal lineages of this tribe are currently found in the Indomalaya (e.g. *Archboldomys luzonensis* Musser, *Apomys datae* (Meyer), *Apomys hylocoetes* Mearns, *Chrotomys gonzalesi* Rickart and Heaney, *Chiropodomys gliroides* (Blyth) and *Rhynchomys isarogensis* Musser and Freeman), a specific and diverse lineage of Hydromyini also colonized and radiated in the Australasia *ca.* 8.1 Ma (95% HPD: 7.22-9.09 Ma). The clade gathering Apodemini, Malacomyini, Murini and Praomyini likely originated in the Afrotropics, with several lineages secondarily colonizing the Indomalaya and the West and East Palearctic. The split between Malacomyini (which remained in the Afrotropics) and Apodemini (which dispersed and differentiated mainly in the West and East Palearctic) is estimated at *ca.* 10.2 Ma (95% HPD: 9.33-11.39 Ma). Murini started to diversify in the Indomalaya at *ca.* 7.2 Ma (95% HPD: 6.24-9.29 Ma). The intense radiation (51 extant species, Monadjem et al., 2015) of Praomyini occurred in the Afrotropics (median age of 6.8 Ma for the MRCA of Praomyini; 95% HPD: 6.06-7.77 Ma). The Indomalayan Millardini

split from the predominantly Afrotropical Arvicanthini + Otomyini tribes at *ca.* 10.8 Ma (95% HPD: 9.88-12.04 Ma). The respective first diversifications within Arvicanthini, Millardini and Otomyini are estimated at *ca.* 8.8 Ma (95% HPD: 7.92-9.72 Ma), 9.2 Ma (95% HPD: 7.72-10.75 Ma) and 4.9 Ma (95% HPD: 4.12-5.98 Ma), respectively. The position of Asian *Golunda* within Arvicanthini is not resolved (Fig. 1); the dispersal to Indomalaya of the lineage leading to the extant *Golunda* species at *ca.* 8.5 Ma (95% HPD: 7.33-9.57 Ma; as suggested in Fig. 3) should therefore be taken with caution.

4. Discussion

4.1. Selection of taxa and molecular markers

Our sampling of 160 species from 82 genera represents 22% of known murid species diversity and more than half of the generic diversity of the family Muridae. When one compares our sampling effort to previous studies (Table 7), only the study of Fabre et al. (2012) relied on a better sampling for the family Muridae (302 species from 105 genera, i.e. about 41% of known species diversity). In the study of Schenk et al. (2013), 18% of murid species are included. The number of sampled murid species is also lower in Lecompte et al. (2008) and Rowe et al. (2008) because their studies focussed on specific tribes and subfamilies.

4.2. Calibration of molecular clock and divergence dating

Using the classical *Mus/Rattus* calibration as prior for divergence dating often lead to an underestimation of the age of the subfamily Murinae, inferring median ages that are generally comprised between 13.3 to 12.0 Ma. Only the most recent studies (e.g. Rowe et al., 2016b) used the correct *Mus/Arvicanthis* calibration with a prior median age of 11.1 Ma (as suggested by Kimura et al., 2015). This resulted in the estimation of Murinae age of *ca.* 14.0 Ma (Rowe et al., 2016b), which is consistent with our study (Table 6).

Other fossils frequently used for molecular clock calibration are from the genus †*Parapodemus*. Recent studies used these fossils in two ways: the first occurrence of †*Parapodemus* sp. (Martín-Suárez and Mein, 1998) in the late Miocene in Europe (Lungu, 1981; see Appendix B) was used to calibrate the MRCA of ‘Apodemurini’ (representing the split between Apodemini+Malacomyini and Murini+Praomyini; Fabre et al., 2013; Rowe et al., 2011). Another calibration point is based on the discovery of †*Parapodemus pasquierae* Aguilar and Michaux, from ‘Lo Fournas 6’ site (Roussillon, France). Authors postulated that the latter species and the smaller †*Parapodemus lugdunensis* co-occurred during the same time period ‘MN10’ (Aguilar et al., 1999; Montuire et al., 2006), dated approximately at 9.7 Ma (Mein, 2003). Michaux et al. (2002) considered the differences between these two species as representative of the split between the large *Apodemus mystacinus* Danford and Alston and all smaller *Apodemus* species from the subgenus *Sylvaemus*, but they used a younger age of 7.0 Ma as a prior for their divergence. Numerous authors followed this calibration (e.g. Bryja et al., 2014; Fabre et al., 2013; Lecompte et al., 2008; Schenk et al., 2013) even if there is no clear rationale for it. The estimated dates of MRCA of Apodemini range from 7.5 Ma (Rowe et al. 2016b; Schenk et al., 2013) to 9.6 Ma (Bryja et al., 2014; Michaux et al., 2002; Lecompte et al., 2008). In our study, we conservatively used the †*Parapodemus lugdunensis* fossil as a stem constraint for Apodemini and this placement resulted in an estimation of their MRCA at 9.0 Ma (95% HPD: 7.92-10.13 Ma).

There are several localities in Africa (e.g. Lukeino Formation, Winkler, 2002; Lemudong’o, Ambrose et al., 2007; Manthi, 2007) where fossil representatives of *Arvicanthis* were identified. These fossils were used for molecular clock calibration in several studies with a prior MRCA for the genus at 6.1 Ma (e.g. Fabre et al., 2013; Rowe et al., 2011). The fossils from Lemudong’o were also used, in a less conservative way, by Bryja et al. (2014): based on early records of †*Otomys* sp. (ca. 5.0 Ma; Denys, 1990), †*Aethomys* sp., †*Arvicanthis* sp.,

†*Lemniscomys* sp. (from Lemudong'o = 6.1 Ma; Ambrose et al., 2007; Manthi et al., 2007) and other relevant samples where these and related genera were absent (9.50-10.50 Ma; Mein et al. 2004), they set the split between Arvicanthini and Otomyini within 6.08-9.54 Ma. This calibration resulted in an estimation of the MRCA for the tribe Arvicanthini at 7.8 Ma (Bryja et al., 2014), which is about 1 million year younger than our own estimate (Table 6). Rowe et al. (2016b) used as minimum age 8.7 Ma for the split between *Arvicanthis* and *Otomys* based on the study of Kimura et al. (2015) that set minimum and maximum ages for locality Y388, where †*Karnimata darwini* was found. This calibration resulted in an estimated MRCA of Arvicanthini at 8.5 Ma. In our study, we instead used the age of an older locality (Y182; median age of 9.2 Ma) where †*Karnimata darwini* was also found (Jacobs, 1978; Kimura et al., 2015), in order to set a crown calibration for the Arvicanthini/Millardini/Otomyini clade. This placement resulted in an estimation of their MRCA at 8.8 Ma (95% HPD: 7.92-9.72 Ma) (Table 6).

4.3. Historical biogeography with focus on faunal exchanges between the Afrotropics and the Indomalaya

Our study could not resolve the origin of murid rodents, but it was either in the Afrotropics or in the Indomalaya. Our inferred ancestral tropical range for the MRCA of murids is consistent with the fact that most extant murid taxa are still distributed in warm and moist tropical areas. During the Early Miocene (23.0-16.0 Ma), the rotation of Africa and Arabia, and finally the collision with Eurasia formed a landbridge between Africa and Eurasia (the so-called 'Gomphotherium landbridge'; Rögl, 1999). During this time period, early murids colonized both geographical regions. The subsequent reopening of the Mediterranean-Indo-Pacific seaway ('Indo-Pacific recurrence'; Rögl, 1999) separated Africa from Eurasia again, thus giving rise to the main clades of Afrotropics and Indomalaya rodents. Three subfamilies,

Deomyinae, Gerbillinae and Lophiomyinae, then likely diversified in the Afrotropics (Chevret and Dobigny, 2005; Ndiaye et al., 2016; Schenk et al., 2013, this study; Figure 3). This hypothesis is supported by paleontological records since the oldest fossils tied to these subfamilies were found in the Afrotropics (e.g. late Miocene *Acomys*, *Gerbilliscus*, *Lophiomy*s and †*Preacomys* from East Africa; Winkler et al. 2010 and references therein). The subfamily Murinae started to diversify in Indomalaya, most probably in Southeast Asia, where we can also find the hitherto highest phylogenetic diversity, including the oldest offshoots of this clade (e.g. the ancestor of Phloeomyini probably lived in the Philippines, those of Rattini and Hydromyini in South-east Asia, etc.; Fabre et al., 2013).

During the Middle Miocene (16.0-11.6 Ma), the Mediterranean-Indo-Pacific seaway closed again at the beginning of the Serravallian *ca.* 13.8 Ma ('Parathethys Salinity Crisis'; Rögl, 1999), co-incidentally with a global cooling that caused vegetation shifts and a general aridification (Prista et al., 2015). The newly formed landbridge (Rögl, 1999) allowed repeated dispersals of murine rodents from Asia to both Africa and Eurasia. Murine fossil records provide clear evidence for connections between the Indomalaya, the Palearctic, and the Afrotropics. Among them, there are two conspicuous examples: (i) †*Progonomys* was recorded in many Indomalayan Middle Miocene localities (Jacobs and Flynn, 2005) as well as in the Palearctic region (Algeria: Wessels, 2009; China: Qiu et al., 2004; Egypt: Heissig, 1982; France: Mein et al., 1993; Spain: Weerd, 1976); and (ii) the oldest records of †*Parapelomys* spp. were found synchronously in Africa (8.5 Ma; Chorora, Ethiopia; Geraads, 2001) and in Pakistan (*ca.* 8.0 Ma; Jacobs and Flynn, 2005). During this period, representatives of several murine tribes occurred in the Afrotropics (Arvicanthini, Malacomyini, Otomyini and Praomyini) and the Indomalaya (Millardini, Murini, Rattini, and basal lineages of Hydromyini).

The last faunal interchange of murid taxa between Africa, Asia and Western Palearctic (Benammi et al., 1996; Sabatier, 1982; Sen, 1977, 1983; Winkler, 2002) is coincident with

Messinian Salinity Crisis *ca.* 6 Ma during the Late Miocene (Hsü et al., 1973, 1978). During this period of global sea level depression (Haq et al., 1987) Africa and Arabia were reconnected through Neguev-Sinai landbridge ('Levantine corridor', Fernandes et al., 2006) and landbridge in the Bab-el-Mandeb (Bosworth et al., 2005). In murids, evidence to support this faunal exchange can be found in the African subgenus *Nannomys* (genus *Mus*), which colonized Afrotropics and started there its radiation *ca.* 5.2 Ma (Bryja et al., 2014). A possible example for an opposite west-to-east migration is the genus *Golunda*, which belongs to the Arvicanthini tribe. In a predominantly Afrotropical clade, *Golunda* is the only genus that occurs in the Indomalaya, probably since the end of Miocene (Ducroz et al., 2001; Fig. 3). However, one should be cautious with this scenario since the position of *Golunda* within Arvicanthini is not well supported (Fig. 1). Africa-to-Asia dispersals at the Miocene/Pliocene boundary have been also recorded in other taxa, such as rodents (e.g. *Myomyscus yemeni* (Sanborn and Hoogstraal); our unpubl. data), reptiles (e.g. *Varanus yemenensis*: Böhme et al., 2003, Portik and Papenfuss 2012; *Hemidactylus* geckos: Šmíd et al. 2013; *Echis* vipers: Pook et al. 2009) and hamadryas baboons (Winney et al. 2004).

5. Conclusion and perspectives

In this study, we provided an improved multilocus dated phylogeny for the highly speciose family Muridae. Both our dating and historical biogeography analyses suggest that the family originated during the Early Miocene, and subsequently gave rise to four extant subfamilies: three in the Afrotropical region (Deomyinae, Gerbillinae and Lophiomyinae) and one in the Indomalaya (Murinae). Our study also supports a dynamic biogeographic scenario in which repeated colonisation events occurred in the Australasian (Hydromyini, Rattini), Afrotropical (Malacomyini, Praomyini, Arvicanthini, Otomyini) and Palearctic (Apodemini) regions. One of the strong aspects of this study lies in the assessment and treatment of fossil

data (Appendix B); such data is likely to be useful for further studies investigating the timing of diversification of rodents, or even mammals in general. For an easy access to all corresponding fossil records, we have made data available on the Date-a-Clade Website (<http://palaeo.gly.bris.ac.uk/fossilrecord2/dateaclade/index.html>), Paleobiology Database (<http://fossilworks.org/>) and TimeTree Database (<http://timetree.org>).

Acknowledgments

This study was supported by two projects of the Czech Science Foundation, no. 15-20229S, the Ministry of Culture of the Czech Republic (DKRVO 2017/15, National Museum, 00023272) and JSPS KAKENHI JP15H06884 (Grant-in-Aid for Young Scientists Start-up). Most analyses were run on CBGP HPC computational platform, while a minor part was performed on CIPRES Gateway. We thank Alexandre Dehne Garcia for his help on the HPC CBGP cluster and Nick J. Matzke for introduction into BioGeoBEARS analysis. We would also like to thank Alisa Winkler and Christiane Denys for discussion about some paleontological aspects, as well as Arame Ndiaye and Pascal Chevret for providing additional sequences. We are grateful to Vincent Lazzari, Fredrick K. Manthi, Pierre Mein, Rajeev Pantanik and Sevkett Sen who provided permission for the reproduction of the figures in Appendix B.

Authors' contributions: GJK, GD, LG, JB conceived the ideas; TA collected genetic data (including new genotyping); YK, TA collected and analysed paleontological data; TA, GJK performed phylogenetic analyses; TA, JB, GJK wrote the first version of the manuscript that was then implemented by all authors.

References

- Adkins, R.M., Gelke, E.L., Rowe, D., Honeycutt, R.L., 2001. Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. *Mol. Biol. Evol.* 18, 777–91.
- Adkins, R.M., Walton, A.H., Honeycutt, R.L., 2003. Higher-level systematics of rodents and divergence time estimates based on two congruent nuclear genes. *Mol. Phylogenet. Evol.* 26, 409–420.
- Aguilar, J.P., Escarguel, G., Michaux, J., 1999. A succession of Miocene rodent assemblages from fissure fillings in southern France: Palaeoenvironmental interpretation and comparison with Spain. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 145, 215–230.
- Ameur, A., 1984. Découverte de nouveaux Rongeurs dans la formation miocène de Bou Hanifia (Algérie occidentale). *Géobios*, 17, 167-175.
- Ambrose, S.H., Nyamai, C.M., Mathu, E.M., Williams, M.A.J., 2007. Geology, Geochemistry, and Stratigraphy of the Lemudong’o Formation, Kenya Rift Valley. *Kirtlandia, Clevel. Museum Nat. Hist.* 53–64.
- Arbogast, B. S., Browne, R. A., Weigl, P. D., 2001. Evolutionary genetics and Pleistocene biogeography of North American tree squirrels (*Tamiasciurus*). *J. Mammal.* 82, 302-319.
- Baele, G., Li, W.L.S., Drummond, A.J., Suchard, M. A., Lemey, P., 2013. Accurate model selection of relaxed molecular clocks in bayesian phylogenetics. *Mol. Biol. Evol.* 30, 239–43.
- Benammi, M., Calvo, M., Prévot, M., Jaeger, J.-J., 1996. Magnetostratigraphy and paleontology of Ait Kandoula Basin (High Atlas, Morocco) and the African-European late Miocene terrestrial fauna exchanges. *Earth Planet. Sci. Lett.* 145, 15-29.
- Bernor, R.L., Brunet M., Ginsburg, L., Mein, P., Pickford, M., Rögl, F., Sen, S., Steininger F., Thomas, H., 1987. A consideration of some major topics concerning old world Miocene

595 mammalian chronology, migrations and paleogeography. *Geobios*. 20, 431-439.

596 Blanga-Kanfi, S., Miranda, H., Penn, O., Pupko, T., DeBry, R.W., Huchon, D., 2009. Rodent
597 phylogeny revised: analysis of six nuclear genes from all major rodent clades. *BMC Evol.*
598 *Biol.* 9, 71.

599 Bosworth, W., Huchon, P., McClay, K., 2005. The Red Sea and Gulf of Aden Basins. *J. Afr.*
600 *Earth Sci.* 43, 334–378.

601 Böhme, W., 2003. Checklist of the living monitor lizards of the world (family Varanidae). *Zool.*
602 *Verhandel.* 341, 1–43.

603 Bryja, J., Aghova, T., Bryjova, A., Denys, C., Verheyen, E., 2017. Evolutionary history of the
604 thicket rats (genus *Grammomys*) mirrors the evolution of African forests since late
605 Miocene. *J. Biogeogr.* 44, 182–194.

606 Bryja, J., Mikula, O.E., Umbera, R., Meheretu, Y., Aghová, T., Lavrenchenko, L. A, Mazoch,
607 V., Ogue, N., Mbau, J.S., Welegerima, K., Amundala, N., Colyn, M., Leirs, H.,
608 Verheyen, E., 2014. Pan-African phylogeny of *Mus* (subgenus *Nannomys*) reveals one of
609 the most successful mammal radiations in Africa. *BMC Evol. Biol.* 14, 256.

610 Campisano, C.J., Feibel, C.S., 2008. Depositional environments and stratigraphic summary of
611 the Pliocene Hadar formation at Hadar, Afar Depression, Ethiopia. in: Quade, J., Wynn,
612 J.G. (Eds.), *The Geology of Early Humans in the Horn of Africa*. *Geol. Soc. Am.*, vol.
613 446, pp. 179-201.

614 Carleton M.D., Musser G.G., 1984. Muroid rodents. in: Anderson S., Jones J.K., Jr., (Eds.),
615 *Orders and families of recent mammals of the world*. Wiley, New York, pp. 289–379.

616 Chaline, J., 1977. Les rongeurs au pilori? Analyse critique des reconstitutions paléoécologiques
617 en microthériodontologie. *Approche écologique de l’Homme fossile*. Paris, 77-81.

618 Cheema, I. U., S.Mahmood Raza, L. J. Flynn, A. R. Rapjar, Tomida, Y., 2000. Miocene
619 smallmammals from Jalalpur, Pakistan, and their biochronologic implications. *Bull. Natl.*

620 Sci. Mus., Tokyo, C 26, 57–77.

621 Chevret, P., Denys, C., Jaeger, J.J., Michaux, J., Catzefflis, F.M., 1993a. Molecular evidence
622 that the spiny mouse (*Acomys*) is more closely related to gerbils (Gerbillinae) than to true
623 mice (Murinae). Proc. Natl. Acad. Sci. U.S.A. 90, 3433–3436.

624 Chevret, P., Denys, C., Jaeger, J.-J., Michaux, J., Catzefflis, F., 1993b. Molecular and
625 palaeontological aspects of the tempo and mode of evolution in *Otomys* (Otomyinae:
626 Muridae: Mammalia). Biochem. Syst. Ecol. 21, 123–131.

627 Chevret, P., Dobigny, G., 2005. Systematics and evolution of the subfamily Gerbillinae
628 (Mammalia, Rodentia, Muridae). Mol. Phylogenet. Evol. 35, 674–688.

629 Daxner-Höck, G., 2003. *Cricetodon meini* and other rodents from the Mühlbach and Grund,
630 Lower Austria (Middle Miocene, late MN5). Ann. Nat.hist. Mus. Wien (A) 104, 267–291.

631 DeBry, R.W., 2003. Identifying conflicting signal in a multigene analysis reveals a highly
632 resolved tree: the phylogeny of Rodentia (Mammalia). Syst. Biol. 52, 604–617.

633 Deino, A.L., Ambrose, S.H., 2007. ⁴⁰Ar/³⁹Ar dating of the Lemudong'o late Miocene fossil
634 assemblages, southern Kenya rift. Kirtlandia. 56, 65–71.

635 Denys, C., 1990. The oldest *Acomys* (Rodentia, Muridae) from the lower Pliocene of South
636 Africa and the problem of its murid affinities. Paleontogr. Abt. A, 210, 79–91.

637 Denys, C., Michaux, J., Catzefflis, F., Ducroq, S., Chevret, P., 1995. Morphological and
638 molecular data against the monophyly of Dendromurinae (Muridae: Rodentia). Bönner
639 Zool. Beitr. 45, 173–190.

640 Dobigny, G., Tatard, C., Gauthier, P., Ba, K., Duplantier, J., Granjon, L., Kergoat, G.J., 2013.
641 Mitochondrial and Nuclear Genes-Based Phylogeography of *Arvicanthis niloticus*
642 (Murinae) and Sub-Saharan Open Habitats Pleistocene History. PLoS One 8.

643 Donoghue, P. C., Benton, M. J., 2007. Rocks and clocks: calibrating the Tree of Life using
644 fossils and molecules. Trends Ecol. Evol. 22, 424–431.

645 Drummond, A.J., Bouckaert, R.R., 2015. Bayesian evolutionary analysis with BEAST 2.
646 Cambridge University press.

647 Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and
648 dating with confidence. PLoS Biol. 4, e88.

649 Drummond, A. J., Suchard, M. A., Xie, D., Rambaut, A. 2012. Bayesian phylogenetics with
650 BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969-1973.

651 Ducroz, J.F., Volobouev, V., Granjon, L., 2001. An Assessment of the Systematics of
652 Arvicanthine Rodents Using Mitochondrial DNA Sequences: Evolutionary and
653 Biogeographical Implications J. Mammal. Evol. 8, 173-206.

654 Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high
655 throughput. Nucleic Acids Res. 32, 1792–7.

656 Erixon, P., Svennblad, B., Britton, T., Oxelman, B., 2003. Reliability of Bayesian posterior
657 probabilities and bootstrap frequencies in phylogenetics. Syst. Biol. 52, 665–673.

658 Fabre, P.H., Hautier, L., Dimitrov, D., Douzery, E.J., 2012. A glimpse on the pattern of rodent
659 diversification: a phylogenetic approach. BMC evolutionary biology, 12(1), 88.

660 Fabre, P.H., Pagès, M., Musser, G.G., Fitriana, Y.S., Fjeldså, J., Jennings, A., Jönsson, K. A.,
661 Kennedy, J., Michaux, J., Semiadi, G., Supriatna, N., Helgen, K.M., 2013. A new genus
662 of rodent from wallacea (Rodentia: Muridae: Murinae: Rattini), and its implication for
663 biogeography and Indo-Pacific Rattini systematics. Zool. J. Linn. Soc. 169, 408–447.

664 Fernandes, C.A., Rohling, E.J., Siddall, M., 2006. Absence of post-Miocene Red Sea land
665 bridges: Biogeographic implications. J. Biogeogr. 33, 961–966.

666 Flynn, L.J., Jacobs, L.L., 1999. Late Miocene small mammal faunal dynamics: the crossroads
667 of the Arabian peninsula. in: Whybrow, P.J., Hill, A., (Eds.), Fossil vertebrates of Arabia.
668 Yale University Press, New Haven, pp. 412–419.

669 Gavryushkina, A., Heath, T. A., Ksepka, D. T., Stadler, T., Welch, D., Drummond, A. J., 2017.

670 Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst.*
671 *Biol.* 66, 57-73.

672 Geraads, D., 2001. Rongeurs du Miocene superieur de Chorora (Ethiopie): Dendromuridae,
673 Muridae et conclusions. *Palaeovertebrata*. 30, 89–109.

674 Geraads, D., Alemseged, Z., Bellon, H., 2002. The late Miocene mammalian fauna of Chorora,
675 Awash basin, Ethiopia: systematics, biochronology and the 40K–40Ar ages of the
676 associated volcanics. *Tertiary Research*. 21, 113–127.

677 Gernhard, T., 2008. The conditioned reconstructed process. *J. Theor. Biol.* 253, 769–778.

678 Granjon, L., Colangelo, P., Tatard, C., Colyn, M., Dobigny, G., Nicolas, V., 2012. Intrageneric
679 relationships within *Gerbilliscus* (Rodentia, Muridae, Gerbillinae), with characterization
680 of an additional West African species. *Zootaxa*. 25, 1–25.

681 Granjon, L., Montgelard, C., 2012. The Input of DNA Sequences to Animal Systematics:
682 Rodents as Study Cases. in: Munshi, A. (Ed.), *DNA Sequencing - Methods and*
683 *Applications*, InTech, pp. 104-140.

684 Haq, B.U., Hardenbol, J., Vail, P.R., 1987. Chronology of fluctuating sea levels since the
685 Triassic. *Science*. 235, 1156-1167.

686 Heath, T. A., Huelsenbeck, J.P., Stadler, T., 2014. The fossilized birth-death process for
687 coherent calibration of divergence-time estimates. *Proc. Natl. Acad. Sci. U.S.A.* 111,
688 E2957-66.

689 Heissig, K., 1982. Kleinsäuger aus einer obermiozänen (Vallesium) Karstfüllung Ägyptens.
690 *Mitt. Bayer. Staatssamml. Paläontol. Hist. Geol.* 22, 97-101.

691 Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing
692 confidence in phylogenetic analysis. *Syst. Biol.* 42, 182–192.

693 Hsü, K.J., 1978. Stratigraphy of the lacustrine sedimentation in the Black Sea. *Initial Reports*
694 *of the DSDP*. 42B, 509–524.

695 Hsü, K.J., Cita, M.B., Ryan, W.B.F., 1973. The origin of the Mediterranean evaporites. Initial
696 Reports of the DSDP. 42, 1203–1231.

697 Jacobs, L. L. 1977. A new genus of murid rodent from the miocene of Pakistan and comments
698 on the origin of the Muridae. *PaleoBios*. 25, 1–11

699 Jacobs, L. L., 1978. Fossil rodents (Rhizomyidae and Muridae) from Neogene Siwalk deposits,
700 Pakistan. *Mus. North. Ariz. Bull.* 52, 1–103.

701 Jacobs, L.L., Downs, W.R., 1994. The evolution of murine rodents in Asia. *Natl. Sci. Museum*
702 *Monogr.* 8, 149–156.

703 Jacobs, L. L., Flynn, L.J., Downs, W.R., 1989. Neogene Rodentia of Southern Asia. in: Black,
704 C.C., Dawson, M.R. (Eds.), *Papers on fossil rodents in honor of Albert Elmer Wood*.
705 *Natural History Museum, Los Angeles County, Science Series No. 33*, pp. 157–177.

706 Jacobs, L., Flynn, L., Downs, W., Barry, J., 1990. Quo vadis Antemus? The Siwalik muroid
707 record. in: Lindsay, E.H, Fahlbusch, V., Mein, P. (Eds.), *European Neogene Mammal*
708 *Chronology*. Plenum Press. New York. pp. 573–586.

709 Jacobs L.L., Flynn J.J., 2005. Of mice ... again: the Siwalik rodent record, murine distribution,
710 and molecular clocks. in: Leiberman. D.E., Smith, R.J., Kelley, J., (Eds.), *Interpreting the*
711 *past: essays on human, primate, and mammal evolution in honor of David Pilbeam*. Brill
712 *Academic Publishers Inc, Boston*. pp. 63–80.

713 Jaeger, J.-J., 1977. Les rongeurs du Miocène moyen et supérieur du Maghreb. *Palaeovertebrata*.
714 8, 3–166.

715 Jansa, S. A, Weksler, M., 2004. Phylogeny of muroid rodents: relationships within and among
716 major lineages as determined by IRBP gene sequences. *Mol. Phylogenet. Evol.* 31, 256–
717 76.

718 Kass, R. E., Raftery, A. E., 1995. Bayes factors. *J. Am. Stat. Assoc.* 90, 773-795.

719 Kimbel, W., Walter, R., Johanson, D., Reed, K., Aronson, J., Assefa, Z., Marean, C.W., Eck,

720 G., Bobe, R., Hovers, E., Rak, Y., Vondra, C., Yemane, T., York, D., Chen, Y., Evensen,
721 N., Smith, P., 1996. Late Pliocene Homo and Oldowan tools from the Hadar formation
722 (Kada Hadar member), Ethiopia. *J. Hum. Evol.* 31, 549-561.

723 Kimura, Y., Hawkins, M.T.R., McDonough, M.M., Jacobs, L.L., Flynn, L.J., 2015. Corrected
724 placement of *Mus-Rattus* fossil calibration forces precision in the molecular tree of
725 rodents. *Nat. Publ. Gr.* 1–9.

726 Kimura, Y., Jacobs, L.L., Cerling, T.E., Uno, K.T., Ferguson, K.M., Flynn, L.J., Patnaik, R.,
727 2013. Fossil Mice and Rats Show Isotopic Evidence of Niche Partitioning and Change in
728 Dental Ecomorphology Related to Dietary Shift in Late Miocene of Pakistan. *PLoS One*
729 8, 1–10.

730 Kingdon, J., 2015. *The Kingdon field guide to African mammals*. Bloomsbury Publishing,
731 London.

732 Lavocat, R., 1978. Rodentia and Lagomorpha. in: Maglio, V.J., Cooke, H.B.S. (Eds.), *Evolution*
733 *of African mammals*. Harvard University Press, Cambridge. pp. 68-89.

734 Landis, M.J., Matzke, N.J., Moore, B.R., Huelsenbeck, J.P., 2013. Bayesian analysis of
735 biogeography when the number of areas is large. *Syst. Biol.* 62, 789–804.

736 Lanfear, R., Calcott, B., Kainer, D., Mayer, C., Stamatakis, A., 2014. Selecting optimal
737 partitioning schemes for phylogenomic datasets. *BMC Evol. Biol.* 14, 82.

738 Lazzari, V., Tafforeau, P., Michaux, J., 2011. When homologous cusps display non-
739 homologous wear facets: An occlusal reorganization ensures functional continuity during
740 dental evolution of Murinae (Rodentia, Mammalia). *Arch. Oral Biol.* 56, 194–204.

741 Lecompte, E., Aplin, K., Denys, C., Catzeflis, F., Chades, M., Chevret, P., 2008. Phylogeny
742 and biogeography of African Murinae based on mitochondrial and nuclear gene sequences,
743 with a new tribal classification of the subfamily. *BMC Evol. Biol.* 8, 199.

744 Lindsay, E.H., 1988. Cricetid Rodents from Siwalik Deposits near Chinji Village. Part 1:

Megacricetodontinae, Myocricetodontinae and Dendromurinae. *Palaeovertebrata*. 18, 95-154.

López Antoñanzas, R., 2009. First potwarmus from the miocene of Saudi Arabia and the early phylogeny of murines (Rodentia: Muroidea). *Zool. J. Linn. Soc.* 156, 664–679.

Lungu, A. N., 1981. Hipparion Fauna of Middle Sarmatian of Moldavia (Insectivora, Lagomorpha and Rodentia). State Tiraspol Institut Pedagogiki Publ. 1557, 1-137.

Maddison, W. P. and Maddison, D.R., 2017. Mesquite: a modular system for evolutionary analysis. Version 3.2 <http://mesquiteproject.org>

Macholán, M., 2012, (Ed.), Evolution of the house mouse (No. 3). Cambridge University Press.

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

Martín-Suárez, E., Mein, P., 1998. Revision of the genera *Parapodemus*, *Apodemus* *Rhagamys* and *Rhagapodemus* (Rodentia, Mammalia). *Geobios*. 31, 87–97.

Matzke, N.J., 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing.

Berkeley Plan. J. 26, 217–220.

Matzke, N.J., 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.* 63, 951–970.

Mein, P., Martín-Suárez, E., Agustí, J. 1993. *Progonomys* Schaub, 1938 and *Huerzelerimys* gen. nov.(Rodentia); their evolution in Western Europe. *Scripta Geol.* 103, 41-64.

Mein, P., 2003. On Neogene rodents of Eurasia: distributions and migrations - in: Reumer, J.W.F., Wessels, W. (Eds.), *Distribution and migration of Tertiary mammals in Eurasia*. A volume in honour of Hans de Bruijn. *Natuurmuseum Rotterdam*. pp. 407-418.

Mein, P., Pickford, M., Senut, B., 2004. Late Miocene micromammals from the Harasib karst deposits, Namibia. Part 2b-Cricetomyidae, Dendromuridae and Muridae, with an

770 addendum on the Myocricetodontinae. *Communs. Geol. Surv. Namibia*. 13, 43–63.

771 Michaux, J. R., Chevret, P., Filippucci, M. G., Macholán, M., 2002. Phylogeny of the genus
772 *Apodemus* with a special emphasis on the subgenus *Sylvaemus* using the nuclear IRBP
773 gene and two mitochondrial markers: cytochrome b and 12S rRNA. *Mol. Phylogenet.*
774 *Evol.* 23, 123-136.

775 Michaux, J., Chevret, P., Renaud, S., 2007. Morphological diversity of Old World rats and mice
776 (Rodentia, Muridae) mandible in relation with phylogeny and adaptation. *J. Zool. Syst.*
777 *Evol. Res.* 45 (3), 263-279.

778 Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for
779 inference of large phylogenetic trees, in: 2010 Gateway Computing Environments
780 Workshop, GCE 2010. pp. 1–8.

781 Monadjem, A., Taylor, P. J., Denys, C., Cotterill, F. P., 2015. Rodents of sub-Saharan Africa:
782 a biogeographic and taxonomic synthesis. Walter de Gruyter GmbH,
783 Berlin/Munich/Boston.

784 Montgelard, C., Forty, E., Arnal, V., Matthee, C. A., 2008. Suprafamilial relationships among
785 Rodentia and the phylogenetic effect of removing fast-evolving nucleotides in
786 mitochondrial, exon and intron fragments. *BMC Evol. Biol.* 8, 321.

787 Montuire, S., Maridet, O., Legendre, S., 2006. Late Miocene–Early Pliocene temperature
788 estimates in Europe using rodents. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 238, 247-262.

789 Munthe, J., 1987. Small-mammal fossils from the Pliocene Sahabi Formation of Libya. in.
790 Boaz, N.T., El-Arnauti, A., Gaziry, A.W, de Heinzelin, J., Boaz, D.D, (Eds.), *Neogene*
791 *Paleontology and Geology of Sahabi*, Alan R. Liss. New York. pp. 135–144.

792 Musser, G.G., Carleton M.D., 2005. Superfamily Muroidea, in: Wilson, D.E., Reeder, D.A.M.
793 (Eds.), *Mammal Species of the World. A Taxonomic and Geographic Reference*. The
794 Johns Hopkins University Press. Baltimore. pp. 894–1531.

795 Ndiaye, A., Chevret, P., Dobigny, G., Granjon, L., 2016. Evolutionary systematics and
796 biogeography of the arid habitat-adapted rodent genus *Gerbillus* (Rodentia, Muridae): a
797 mostly Plio-Pleistocene African history. *J. Zool. Syst. Evol. Res.* 54, 299-317.

798 Near, T.J., Meylan, P. A., Shaffer, H.B., 2005. Assessing concordance of fossil calibration
799 points in molecular clock studies: an example using turtles. *Am. Nat.* 165, 137–146.

800 Near, T.J., Sanderson, M.J., 2004. Assessing the quality of molecular divergence time estimates
801 by fossil calibrations and fossil-based model selection. *Philos. Trans. R. Soc. Lond. B.*
802 *Biol. Sci.* 359, 1477–1483.

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

806 Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian
807 Phylogenetic Analysis of Combined Data. *Syst. Biol.* 53, 47–67.

808 Ogg, J. G., Smith, A. G., 2004. Geologic Time Scale 2004. in: Gradstein, F. M., Ogg, J. G.,
809 Smith, A. G.A (Eds.), Cambridge University Press. pp. 63–86.

810 Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N.,
811 Underwood, E.C., D’amico, J. A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J.,
812 Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P.,
813 Kassem, K.R., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth.
814 *Bioscience* 51, 933.

815 Pagès, M., Fabre, P.-H., Chaval, Y., Mortelliti, A., Nicolas, V., Wells, K., Michaux, J.R.,
816 Lazzari, V., 2016. Molecular phylogeny of Southeast Asian arboreal murine rodents. *Zool.*
817 *Scr.* 45, 349-364.

818 Parham, J.F., Donoghue, P.C.J., Bell, C.J., Calway, T.D., Head, J.J., Holroyd, P. A., Inoue,
819 J.G., Irmis, R.B., Joyce, W.G., Ksepka, D.T., Patané, J.S.L., Smith, N.D., Tarver, J.E.,

820 Van Tuinen, M., Yang, Z., Angielczyk, K.D., Greenwood, J.M., Hipsley, C. A., Jacobs,
821 L., Makovicky, P.J., Müller, J., Smith, K.T., Theodor, J.M., Warnock, R.C.M., Benton,
822 M.J., 2012. Best practices for justifying fossil calibrations. *Syst. Biol.* 61, 346–359.

823 Pook, C. E., Joger, U., Stümpel, N., Wüster, W., 2009. When continents collide: phylogeny,
824 historical biogeography and systematics of the medically important viper genus *Echis*
825 (*Squamata: Serpentes: Viperidae*). *Mol. Phylogenet. Evol.* 53, 792-807.

826 Portik, D. M., Papenfuss, T. J., 2012. Monitors cross the Red Sea: the biogeographic history
827 of *Varanus yemenensis*. *Mol. Phylogenet. Evol.* 62, 561-565.

828 Prista, G.A., Agostinho, R.J., Cachão, M.A., 2015. Observing the past to better understand the
829 future: A synthesis of the Neogene climate in Europe and its perspectives on present
830 climate change. *Open Geosci.* 7, 65–83.

831 Pyron, R.A., 2010. A likelihood method for assessing molecular divergence time estimates and
832 the placement of fossil calibrations. *Syst. Biol.* 59, 185–194.

833 Qiu, Z., Zheng, S., Zhang, Z., 2004. Murids from the late Miocene Bahe formation, Lantian,
834 Shaanxi. *Vertebrata Pal. Asiatica*, 42, 67-76.

835 Ree, R.H., Smith, S. A., 2008. Maximum likelihood inference of geographic range evolution
836 by dispersal, local extinction, and cladogenesis. *Syst. Biol.* 57, 4–14.

837 Reed, D. N., 2011. New murid (*Mammalia, Rodentia*) fossils from a late Pliocene (2.4 Ma)
838 locality, Hadar A. L. 894, Afar Region, Ethiopia. *J. Vert. Paleontol.* 31, 1326–1337.

839 Renaud, S., Michaux, J., Mein, P., Aguilar, J.P., Auffray, J.C., 1999. Patterns of size and shape
840 differentiation during the evolutionary radiation of the European Miocene murine rodents.
841 *Lethaia* 32, 61–71.

842 Ripplinger, J., Sullivan, J., 2008. Does choice in model selection affect maximum likelihood
843 analysis? *Syst. Biol.* 57, 76–85.

844 Ronquist, F., 1997. Dispersal-Vicariance Analysis: A New Approach to the Quantification of

845 Historical Biogeography. *Syst. Biol.* 46, 195.

846 Ronquist, F., Klopstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D.L., Rasnitsyn, A.P.,
847 2012a. A total-evidence approach to dating with fossils, applied to the early radiation of
848 the hymenoptera. *Syst. Biol.* 61, 973–999.

849 Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B.,
850 Liu, L., Suchard, M. A, Huelsenbeck, J.P., 2012b. MrBayes 3.2: efficient Bayesian
851 phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–
852 42.

853 Rowe, K.C., Achmadi, A.S., Esselstyn, J.A., 2016a. A new genus and species of omnivorous
854 rodent (Muridae: Murinae) from Sulawesi, nested within a clade of endemic carnivores. *J.*
855 *Mammal.* 97, 978-991.

856 Rowe, K.C., Achmadi, A.S., Esselstyn, J.A., 2016b. Repeated evolution of carnivory among
857 Indo-Australian rodents. *Evolution.* 70, 653-665.

858 Rowe, K.C., Aplin, K.P., Baverstock, P.R., Moritz, C., 2011. Recent and rapid speciation with
859 limited morphological disparity in the genus *rattus*. *Syst. Biol.* 60, 188–203.

860 Rowe, K.C., Reno, M.L., Richmond, D.M., Adkins, R.M., Stepan, S.J., 2008. Pliocene
861 colonization and adaptive radiations in Australia and New Guinea (Sahul): Multilocus
862 systematics of the old endemic rodents (Muroidea: Murinae). *Mol. Phylogenet. Evol.* 47,
863 84–101.

864 Rögl, F., 1999. Mediterranean and Paratethys. Facts and hypotheses of an Oligocene to
865 Miocene paleogeography (short overview). *Geol. Carpathica.* 50, 339–349.

866 Sabatier, M., 1982. Les rongeurs du site Pliocène à Hominidés de Hadar (Ethiopie).
867 *Palaeovertebrata.* 12. 1–56.

868 Sauquet, H., Ho, S.Y.W., Gandolfo, M.A., Jordan, G.J., Wilf, P., Cantrill, D.J., Bayly, M.J.,
869 Bromham, L., Brown, G.K., Carpenter, R.J., Lee, D.M., Murphy, D.J., Sniderman, J.M.K.,

870 Udovicic, F., 2012. Testing the impact of calibration on molecular divergence times using
871 a fossil-rich group: The case of nothofagus (Fagales). *Syst. Biol.* 61, 289–313.

872 Sen, S. 1977. La faune de rongeurs Pliocenes de Calta (Ankara, Turquie). *Bull. Mus. Natl. Hist.*
873 *Nat. Sci. Terre.* 465, 1–171.

874 Sen, S., 1983. Rongeurs et lagomorphes du gisement pliocène de Pul-e Charkhi, bassin de
875 Kabul, Afghanistan. *Bull. Mus. Natl. Hist. Nat.* 5, 33-74.

876 Schenk, J. J., Rowe, K. C., Steppan, S. J., 2013. Ecological opportunity and incumbency in the
877 diversification of repeated continental colonizations by muroid rodents. *Syst. Biol.* 62,
878 837-864.

879 Simpson, G. G., 1945. The principles of classification and a new classification of mammals.
880 *Bull. Am. Mus. Nat. Hist.* 85, 1–350.

881 Stadler, T., 2010. Sampling-through-time in birth-death trees. *J. Theor. Biol.* 267, 396–404.

882 Stamatakis, A., 2014. RAxML version 8 a tool for phylogenetic analysis and post-analysis of
883 large phylogenies. *Bioinformatics.* 30, 1312-1313.

884 Stanhope, M. J., Czelusniak, J., Si, J. S., Nickerson, J., Goodman, M., 1992. A molecular
885 perspective on mammalian evolution from the gene encoding interphotoreceptor retinoid
886 binding protein, with convincing evidence for bat monophyly. *Mol. Phylogenet. Evol.* 1,
887 148-160.

888 Steppan, S., Adkins, R., Anderson, J., 2004. Phylogeny and divergence-date estimates of rapid
889 radiations in muroid rodents based on multiple nuclear genes. *Syst. Biol.* 53, 533–553.

890 Steppan, S.J., Adkins, R.M., Spinks, P.Q., Hale, C., 2005. Multigene phylogeny of the Old
891 World mice, Murinae, reveals distinct geographic lineages and the declining utility of
892 mitochondrial genes compared to nuclear genes. *Mol. Phylogenet. Evol.* 37, 370–88.

893 Suwa, G., Beyene, Y., Nakaya, H., Bernor, R. L., Boissarie, J. R., Bibi, F., ... & Asfaw, B. 2015.
894 Newly discovered cercopithecoid, equid and other mammalian fossils from the Chorora

895 Formation, Ethiopia. *J. Anthropol. Sci.* 123, 19-39.

896 Šmíd, J., Carranza, S., Kratochvíl, L., Gvoždík, V., Nasher, A.K., Moravec, J., 2013. Out of
897 Arabia: a complex biogeographic history of multiple vicariance and dispersal events in the
898 gecko genus *Hemidactylus* (Reptilia: Gekkonidae). *PLoS One*. 8, e64018.

899 Teeling, E.C., Scally, M., Kao, D.J., Romagnoli, M.L., Springer, M.S., Stanhope, M.J., 2000.
900 Molecular evidence regarding the origin of echolocation and flight in bats. *Nature*, 403,
901 188–192.

902 Vangengeim, E.A., Lungu, A.N., Tesakov, A.S., 2006. Age of the Vallesian Lower Boundary
903 (Continental Miocene of Europe). *Stratigr. Geol. Correl.* 14, 655-667.

904 Vaughan, T.A., Ryan, J.M., Czaplewski, N.J., 2011. *Mammology*. 5th ed. Jones and Bartlett
905 Publishers, Sudbury.

906 Vrba, E. S. 1995. On the connections between paleo- climate and evolution. in: Vrba, E.S.,
907 Denton, G.H., Partridge, T.C., Burckle, L.H. (Eds.), *Paleoclimate and Evolution with*
908 *Emphasis on Human Origins*. Yale University Press. New Haven. pp. 24–48.

909 Weerd, A. D., 1976. Rodent faunas of the Mio-Pliocene continental sediments of the Teruel-
910 Alfambra region, Spain. Doctoral dissertation, Utrecht University.

911 Wessels, W., 2009. Miocene rodent evolution and migration. Muroidea from Pakistan, Turkey
912 and Northern Africa. *Geologica Ultraiectina* 307. Doctoral dissertation, Departement
913 Aardwetenschappen.

914 Winkler, A.J., 2002. Neogene paleobiogeography and East African paleoenvironments:
915 contributions from the Tugen Hills rodents and lagomorphs. *J. Hum. Evol.* 42, 237–256.

916 Winkler, A.J., Denys, C., Avery, D.M., 2010. Rodentia, in: Werdelin, L., Sanders, W.J. (Eds.),
917 *Cenozoic mammals of Africa*. Univ of California Press. pp. 261–304.

918 Winney, B. J., Hammond, R. L., Macasero, W., Flores, B., Boug, A., Biquand, V., ... & Bruford,
919 M. W. 2004. Crossing the Red Sea: phylogeography of the hamadryas baboon, *Papio*

hamadryas hamadryas. Mol. Ecol. 13, 2819-2827.

Figure titles

Figure 1. Molecular phylogeny of family Muridae based on Bayesian inference (BI) in MrBayes. Very similar topology was obtained also by maximum likelihood (ML) analysis in RAxML (see Appendix D). Red points show nodes supported in BI analysis (posterior probability $PP \geq 0.95$), blue points show high bootstrap support in ML analysis (bootstrap $BV \geq 0.70$). Violet nodes are supported by both analyses.

Figure 2. Divergence dating analysis of family Muridae. Nodes show medians of times to most recent common ancestor (MRCA), node bars indicate 95% HPD intervals. Latin numbers in yellow squares indicate positions of fossil constraints selected by multiple-step evaluation and used for final analysis (see Table 4 for more details).

Figure 3. Ancestral reconstruction for family Muridae with BioGeoBEARS (DEC+J; $d=0.008$; $e=0$; $j=0.0246$; $LnL=-117.11$). Five biogeographical areas are represented using different colours: A, West Palearctic (dark blue); B, East Palearctic (light blue); C, Indomalaya (green); D, Australasia (yellow); E, Afrotropics (red).

Tables:

Table 1: Brief summary of taxonomic changes in the family Muridae. The taxonomic position of particular lineages has changed significantly and they were either considered as separate families, or included in other families outside Muridae (names of families are in bold).

Table 2: PartitionFinder results showing optimum partitioning schemes and best fit models for each analysis (RAxML, MrBayes, BEAST). Settings: BIC, unlinked branch lengths, greedy algorithm.

Table 3: Overview of 18 candidate fossils with stratigraphic age, locality and relevant references. For more details see Appendix B.

Table 4: Overview of fossils finally selected for divergence dating, with parameters of uniform and lognormal prior distribution.

Table 5: Comparison of models used for BioGeoBEARS; likelihood scores (LnL), number of parameters (numparams), dispersal rate (d), extinction rate (e), free parameter controlling the relative probability of founder-event speciation events at cladogenesis (j), corrected Akaike Information Criterion (AICc), and AICc model weights.

Table 6: Results of divergence dating analysis. Time to the most common ancestor (MRCA) is shown as median in Ma, with 95% highest posterior density (HPD). Estimates from previous studies dealing with divergence dating of murid rodents are reviewed here for comparison.

Table 7: Comparison between previous relevant studies.

963 **Supplementary material**

964 **Appendix A.** List of taxa and genetic markers.

965 **Appendix B.** Description of considered fossils.

966 **Appendix C.** Results of cross-validation of fossil constraints.

967 **Appendix D.** Maximum likelihood phylogenetic tree.

Table 1: Brief summary of taxonomic changes in the family Muridae, The taxonomic position of particular lineages has changed significantly and they were either considered separate families or included in other families outside Muridae (name of families are in bold).

Simpson (1945)		Chaline et al. (1977)		Lavocat (1978)	
Cricetidae	Lophiomyinae	Cricetidae	Lophiomyinae	Nesomyidae	Lophiomyinae
	Gerbillinae	Gerbillidae			Gerbillinae
Muridae	Murinae	Muridae	Murinae		Otomyinae
	Otomyinae	Nesomyidae	Otomyinae	Muridae	
Carleton and Musser (1984)		Musser and Carleton (2005)		Granjon and Montgelard (2012)	
Muridae	Lophiomyinae	Cricetidae	Lophiomyinae	Muridae	Lophiomyinae
	Gerbillinae	Muridae	Otomyinae		Gerbillinae
	Murinae		Gerbillinae		Murinae
	Otomyinae		Murinae		Deomyinae
			Deomyinae		Leimacomyinae
			Leimacomyinae		

Table 2: PartitionFinder results showing optimum partitioning schemes and best fit models for each analysis (RAxML, MrBayes, BEAST). Settings: BIC, unlinked branch lengths, greedy algorithm.

Analysis	N. of part	Best sub. model	Subsets
RAxML	3	GTR+I+G	12S, 16S, ATPase8_pos1-pos2, COII_pos1-pos2, COI_pos1-pos2, CYTB_pos1-pos2, tRNA-Asp, tRNA-Lys, tRNA-Ser
		GTR+I+G	ATPase8_pos3, COII_pos3, COI_pos3, CYTB_pos3
		GTR+I+G	AP5_pos1-pos3, BRCA_pos1-pos3, BZRP, GHR_pos1-pos3, IRBP_pos1-pos3, RAG_pos1-pos3
MrBayes and BEAST	3	GTR+I+G	12S, 16S, ATPase8_pos1-pos2, COII_pos1-pos2, COI_pos1-pos2, CYTB_pos1-pos2, tRNA-Asp, tRNA-Lys, tRNA-Ser
		HKY+I+G	ATPase8_pos3, COII_pos3, COI_pos3, CYTB_pos3
		GTR+I+G	AP5_pos1-pos3, BRCA_pos1-pos3, BZRP, GHR_pos1-pos3, IRBP_pos1-pos3, RAG_pos1-pos3

Table 3: Overview of 18 candidate fossils with stratigraphic age, locality and relevant references. For more details see Appendix B.

Fossil	Final analysis	Reason for excluding	Age (Ma)	Dating		Fossils	
				Method	References	Site	References
† <i>Potwarmus primitivus</i> (Wessels)	No	Equivocal placement	16.0	magnetostratigraphy	time scale of Ogg and Smith (2004)	Pakistan, Potwar Plateau, YGSP591	Lindsay (1988), Wessels (2009)
† <i>Antemus chinjiensis</i> Jacobs	Yes (1)		13.8	magnetostratigraphy	Jacobs and Flynn (2005)	Pakistan, Potwar Plateau, YGSP 491	Jacobs et al. (1989)
†cf. <i>Progonomys</i> sp. Schaub	No	Equivocal placement	11.6	magnetostratigraphy	time scale of Ogg and Smith (2004)	Pakistan, Potwar Plateau, YGSP 83, YGSP 504	Jacobs and Flynn (2005), Cheema et al. (2000), Kimura et al. (in prep.)
†cf. <i>Karnimata</i> sp. Jacobs	Yes (2)		11.2	magnetostratigraphy	Kimura et al. (2015) by time scale of Ogg and Smith (2004)	Pakistan, Siwalik Group, Nagri Formation, YGSP 791, YGSP 797,	Jacobs and Flynn (2005); Kimura et al. (2015)
† <i>Parapodemus lungdunensis</i> Schau	Yes (3)		9.6	magnetostratigraphy	Daxner-Höck (2003)	France, Dionay	Lungu (1981), Mein et al. (1993), Renaud et al. (1999)
† <i>Karnimata darwini</i> Jacobs	Yes (4)		9.2	magnetostratigraphy	Kimura et al. (2015) by time scale of Ogg and Smith (2004)	Pakistan, Siwalik Group, Dhok Pathan Formation, YGSP 182	Jacobs (1978); Kimura et al. (2015)
† <i>Abudhabia pakistanensis</i> Flynn and Jacobs	Yes (5)		8.7	magnetostratigraphy	Flynn and Jacobs (1999) by time scale of Ogg and Smith (2004)	Pakistan, Siwalik Group, Dhok Pathan Formation, YGSP387	Flynn and Jacobs (1999)
†aff. <i>Stenocephalomys</i> Frick	No	Equivocal placement	8.5	$^{40}\text{K}/^{40}\text{Ar}$	Geraads et al. (2002), Suwa et al. (2015)	Ethiopia, Chorora	Geraads (2001)
†cf. <i>Parapelomys</i> Jacobs	No	Equivocal placement	8.5	$^{40}\text{K}/^{40}\text{Ar}$	Geraads et al. (2002), Suwa et al. (2015)	Ethiopia, Chorora	Geraads (2001)

Table 3 (continued)

Fossil	Final analysis	Reason for excluding	Age (Ma)	Dating		Fossils	
				Method	References	Site	References
† <i>Preacomys kikiae</i> Geraads	No	Equivocal placement	8.5	$^{40}\text{K}/^{40}\text{Ar}$	Geraads et al. (2002), Suwa et al. (2015)	Ethiopia, Chorora	Geraads (2001)
† <i>Mus</i> sp. Linneaus	Yes (6)		8.0	magnetostratigraphy	Kimura et al. (2015) by time scale of Ogg and Smith (2004)	Pakistan, Siwalik Group, Dhok Pathan Formation, YGSP 547	Kimura et al. (2013; 2015)
† <i>Acomys</i> sp. I. Geoffroy	No	missing M1	6.1	$^{40}\text{Ar} / ^{39}\text{Ar}$	Deino and Ambrose (2007)	Kenya, Lemudong'o, locality 1	Manthi (2007)
† <i>Aethomys</i> sp. Thomas	Yes (7)		6.1	$^{40}\text{Ar} / ^{39}\text{Ar}$	Deino and Ambrose (2007)	Kenya, Lemudong'o, locality 1	Manthi (2007)
† <i>Arvicanthis</i> sp. Lesson	Yes (8)		6.1	$^{40}\text{Ar} / ^{39}\text{Ar}$	Deino and Ambrose (2007)	Kenya, Lemudong'o, locality 1	Manthi (2007)
† <i>Gerbilliscus</i> sp. (Thomas)	Yes (9)		6.1	$^{40}\text{Ar} / ^{39}\text{Ar}$	Deino and Ambrose (2007)	Kenya, Lemudong'o, locality 1	Manthi (2007)
† <i>Lemniscomys</i> sp. Trouessart	No	missing M1	6.1	$^{40}\text{Ar} / ^{39}\text{Ar}$	Deino and Ambrose (2007)	Kenya, Lemudong'o, locality 1	Manthi (2007)
† <i>Mastomys</i> sp. Thomas	No		6.1	$^{40}\text{Ar} / ^{39}\text{Ar}$	Deino and Ambrose (2007)	Kenya, Lemudong'o, locality 1	Manthi (2007)
† <i>Gerbillus</i> sp. Desmarest	No	cross-validation	2.4	geochronology (BKT-3 tephra), $^{40}\text{K}/^{40}\text{Ar}$, $^{40}\text{Ar} / ^{39}\text{Ar}$; sedimentology	Kimbel et al. (1996), Campisano and Feibel (2008)	Ethiopia, Hadar, AL894	Reed (2011)

Table 4: Overview of fossils finally selected for divergence dating, with parameters of uniform and lognormal prior distribution.

	Fossil	Position	Stem/Crown	Age (Ma)	Uniform distribution		Lognormal distribution		
					Min	Max	Offset	Log	Mean
1	† <i>Antemus chinjiensis</i>	crown Murinae	crown	13.8	13.24	25.29	13.24	1.0	3.2
2	†cf. <i>Karnimata</i> sp.	<i>Mus/Arvicanthis</i> split	crown	11.2	10.47	25.37	10.47	1.0	4.0
3	† <i>Parapodemus lugdunensis</i>	<i>Apodemus/Tokudaia</i> split	stem	9.6	8.93	25.41	8.93	1.0	4.5
4	† <i>Karnimata darwini</i>	TMRCA Millardini/Otomyini/Arvicanthini	crown	9.2	8.52	25.42	8.52	1.0	4.6
5	† <i>Abudhabia pakistanensis</i>	<i>Gerbilliscus/Desmodillus</i> split	crown	8.7	8.01	25.43	8.01	1.0	4.7
6	† <i>Mus</i> sp.	Murini	stem	8.0	7.29	25.45	7.29	1.0	4.9
7	† <i>Aethomys</i> sp.	<i>Aethomys</i>	stem	6.1	5.34	25.50	5.34	1.0	5.5
8	† <i>Arvicanthis</i> sp.	<i>Arvicanthis</i>	stem	6.1	5.34	25.50	5.34	1.0	5.5
9	† <i>Gerbilliscus</i> sp.	<i>Gerbilliscus</i>	stem	6.1	5.34	25.50	5.34	1.0	5.5

Table 5: Comparison of models used for BioGeoBEARS; likelihood scores (LnL), number of parameters (numparams), dispersal rate (d), extinction rate (e), free parameter controlling the relative probability of founder-event speciation events at cladogenesis (j), corrected Akaike Information Criterion (AICc), and AICc model weights.

	LnL	numparams	d	e	j	AICc	AICc_wt
DEC	-133.724	2	0.005	0.000	0.000	271.525	0.05181
DEC +J	-129.826	3	0.004	0.000	0.008	265.806	0.90415
DIVALIKE	-146.958	2	0.007	0.000	0.000	297.993	9.2665e-08
DIVALIKE +J	-132.856	3	0.004	0.000	0.012	271.866	0.04369
BAYAREALIKE	-178.311	2	0.005	0.036	0.000	360.698	2.2423e-21
BAYAREALIKE +J	-137.687	3	0.003	0.000	0.014	281.528	0.00035

Table 6: Results of divergence dating analysis. Time to the most common ancestor (TMRCA) is shown as median in Ma, with 95% highest posterior density (HPD). Estimates from previous studies dealing with divergence dating of murid rodents are reviewed here for comparison.

TMRCA	median	2.5%	97.5%	Steppan et al. (2004)	Lecompte et al. (2008)	Rowe et al. (2011)	Fabre et al. (2012)	Fabre et al. (2013)	Schenk et al. (2013)	Bryja et al. (2014)	Rowe et al. (2016b)
Muridae	19.3	17.06	21.92	22.5			~33-23		~21.0		
Lophiomyinae*	18.6	16.35	21.11								
Gerbillinae	12.2	10.46	14.16	9.3			~23-5		~10.0		
Deomyinae	14.6	12.68	16.82	13.1			~23-5		~13.0		
Murinae	14.2	12.70	16.07	12.0	12.3	13.3	~23-5		~14.5		~14.0
"core Murinae"	12.3	11.28	13.62	10.3	11.3			11.60	~12.5		~12.5
Phloeomyini	9.8	7.72	11.85		8.6				~10.0		~10.5
Rattini	10.4	9.23	11.93		9.7			8.70	~11.0		~11.0
Hydromyini	10.4	9.31	11.71		8.9				~11.0		~11.0
Malacomyini	3.4	2.35	4.60							4.4	
Apodemini	9.0	7.92	10.13		9.6				~7.5	9.5	~7.5
Murini	7.2	6.24	8.29		6.6	5.3			~6.0	7.4	~6.5
Praomyini	6.8	6.05	7.77		7.6				~5.0	6.8	~6.0
Millardini	9.2	7.72	10.75								
Otomyini	4.9	4.12	5.98						~6.0	3.8	~3.0
Arvicanthini	8.8	7.92	9.72		8.4	7.3			~7.5	7.8	~8.5

* offshoot from Gerbillinae + Deomyinae

Table 7: Comparison between previous relevant studies.

Author	Focus	Subfamilies	# genera	# species	Fossils in Muridae	# genetic markers
Fabre et al. (2012)	Rodentia	4	105	302	0	11 (8)*
Schenk et al. (2013)	Muroidea	4	85	136	4	4
our study	Muridae	4	82	160	9	15
Rowe et al. (2008)	Hydromyini (outgroup all Murinae + Gerbillinae+Deomyinae)	3	66	78	2	8
Lecompte et al. (2008)	Murinae	3	46	83	2	3

* 11 genetic markers for Rodentia, only 8 for Muridae