Species limits and phylogeographic structure in two genera of solitary African mole-rats *Georychus* and *Heliophobius*

M. Uhrová[a,1],*, O. Mikula[a,b,1], N.C. Bennett[c], P. Van Daele[a], L. Piálek[a], J. Bryja[b], J.H. Visser[d,e], B.J ansen van Vuuren[d], R. Šumbera[a]

[a]Department of Zoology, Faculty of Sciences, University of South Bohemia, Czech Republic
[b]Institute of Vertebrate Biology, Czech Academy of Sciences, Czech Republic
[c]Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, 0002, South Africa
[d]Centre for Ecological Genomics and Wildlife Conservation, Department of Zoology, University of Johannesburg, Auckland Park, Johannesburg, South Africa
[e]Department of Conservation and Marine Sciences, Cape Peninsula University of Technology, P.O. Box 652, Cape Town 8000, South Africa

*Corresponding author. uhrovm00@jcu.cz; uhrova4@gmail.com

[1]MU and OM contributed equally to this work.

Highlights

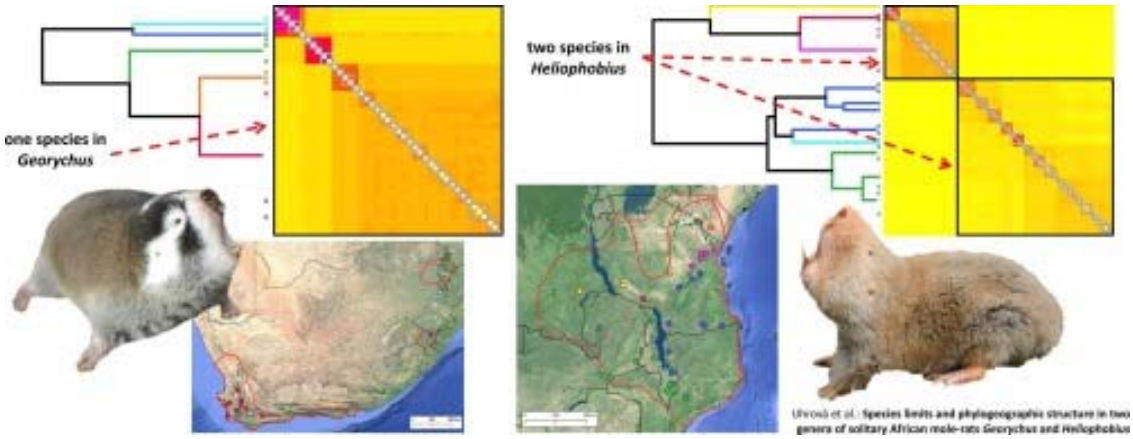
- Solitary African mole-rats, *Georychus* and *Heliophobius*.
- Species delimitation by multispecies coalescent model and clustering on co-ancestry matrices.
- Comparison of Sanger and ddRAD sequence data.
- One species of *Georychus* and two of *Heliophobius*.
- Intrageneric structure was shaped by Pleistocene climatic oscillations.

Abstract

African mole-rats (Bathyergidae) are an intensively studied family of subterranean rodents including three highly social and three solitary genera. Although their phylogenetic interrelations are clear, genetic diversity and the number of species within each genus is much less certain. Among the solitary genera, *Heliophobius* and *Georychus* were for a long time considered as monotypic, but molecular studies demonstrated strong phylogeographic structure within each genus and proposed that they represent complexes of cryptic species. The present study re-evaluates their internal genetic/phylogenetic structure using a combination of methodological approaches. We generated datasets of one mitochondrial and six specifically selected nuclear markers as well as of a large number of double digest restriction site associated (ddRAD) loci and then applied species delimitation analyses based on the multispecies coalescent model or clustering on co-ancestry matrices. The population structure was largely congruent across all analyses, but the methods differed in their resolution scale when determining distinct gene pools. While the multispecies coalescent model distinguished five *Georychus* and between eleven to thirteen *Heliophobius* gene pools in both Sanger sequenced and ddRAD loci, two clustering algorithms revealed significantly finer or coarser structure in ddRAD based co-ancestry matrices. Tens of clusters were distinguished by fineRADstructure and one (in *Georychus*) or two clusters (in *Heliophobius*)

by Infomap. The divergence dating of the bathyergid phylogeny estimated that diversification within both genera coincided with the onset of the Pleistocene and was likely driven by repeated large-scale climatic changes. Based on this updated genetic evidence, we suggest recognizing one species of *Georychus* and two species of *Heliophobius*, corresponding to a northern and southern major lineage, separated by the Eastern Arc Mountains. Yet, the final taxonomic revision should await integrated evidence stemming from e.g. morphological, ecological, or behavioral datasets.

Graphical abstract



Keywords: Bathyergideae; ddRAD; Molecular phylogeny; Diversification; Population structure; Pleistocene climatic fluctuations

1. Introduction

Although the description and cataloguing of biodiversity are now commonly based on genetic data, best practices for the inclusion of large multi-locus data are yet not firmly established in the field. The inclusion of reduced representation of genomic data (including RAD sequences) can shed a brighter light on the current differentiation and evolutionary past of studied species, but its potential is yet to be fully explored. This is also true for the description of small mammal diversity that still relies heavily on the analyses of a handful of mitochondrial and nuclear loci.

The African mole-rats (Bathyergidae) is a case in hand. These animals are subterranean hystricognathous rodents distributed throughout savannah habitats in sub-Saharan Africa. The family includes six genera with remarkable differences in their social lifestyles. Sister to all other genera is the social genus *Heterocephalus*, followed by the solitary *Heliophobius*. Next, the ancestor of two social genera, *Fukomys* and *Cryptomys,* diverged from the ancestor of two solitary genera, *Georychus* and *Bathyergus* (e.g. Allard and Honeycutt, 1992, Faulkes et al., 2004, Davies et al., 2015, Patterson and Upham, 2014, Visser et al., 2020).

In contrast to the highly supported generic classification and phylogeny, the intrageneric diversity remains largely unresolved. Among other reasons, this is due to limitations of hitherto available data and due to inconsistent application of species delimitation criteria. The majority of taxonomic revisions, including more recent studies and species descriptions, were based on morphological data (GIPPOLITI and AMORI, 2011, Van Daele et al., 2013),

karyotypes (Burda et al., 1999) or just one or a few genetic markers (e.g. Faulkes et al., 2004, Faulkes et al., 2011, Faulkes et al., 2017, Visser et al., 2020). Morphological approaches are limited by the paucity of diagnostic characters as a result of the strong selection imposed by the subterranean environment (Stein, 2000). Karyotypic differences were shown to be of little significance for reproductive isolation, especially in the case of Robertsonian fusions (Maputla et al., 2011, Horn et al., 2012); these are common in mole-rats (Nevo et al. 1986). Finally, single unlinked loci are differently affected by stochastic population genetic processes, and may not reflect genome-wide differentiation (Avise et al., 1987, Degnan and Rosenberg, 2009, Nichols, 2001). Of particular importance is that mitochondrial and nuclear loci are prone to mutual phylogenetic discordance (Toews and Brelsford, 2012). Due to its different mode of inheritance, mitochondrial DNA (mtDNA) may introgress into different nuclear backgrounds, blurring otherwise clear differentiation patterns (e.g. Runck et al., 2009, Furman et al., 2014, Good et al., 2015, Ait Belkacem et al., 2016).

On the other hand, multi-locus genetic analyses can reveal substantial differentiation even between populations that are very similar in most other respects (e.g. Hulva et al., 2004, Boratynski et al., 2012). In African mole-rat genera, substantial phylogenetic diversity was revealed in the solitary *Heliophobius* and *Georychus* (Faulkes et al., 2011, Visser et al., 2018) that were traditionally considered monotypic (Bennett and Faulkes, 2000, Happold, 2013, Wilson et al., 2016). Notably, both of these genera have fragmented distributions spanning diverse environmental conditions, which can trigger speciation processes. *Heliophobius* has a large but fragmented distribution in East Africa where it occupies various savanna habitats from grassland to miombo woodland (Burda, 2001, Šumbera et al., 2007). *Georychus* is largely confined to the South African south-western coastal regions, which are dominated by fynbos vegetation, but it has two isolated populations in the grasslands of KwaZulu-Natal and Mpumalanga provinces (Bronner, 1990, Visser et al., 2018).

In *Heliophobius*, Zambian and Malawian populations have a different karyotype then the Kenyan one (Scharff et al., 2001, Šumbera et al., 2007, Deuve et al., 2008), and analysis of a geographically comprehensive sample of mitochondrial and nuclear genes defined up to nine distinct lineages (Faulkes et al., 2011, Bryja et al., 2018). This led Faulkes et al. (2011) to distinguish at least six species, a classification followed by some taxonomic compendia (e.g. Monadjem et al., 2015). Bryja et al. (2018), on the contrary, suggested keeping the monotypic status of the genus until new data are collected. In *Georychus*, similarly, allozyme data and mitochondrial sequences revealed remarkable differentiation between disjunct populations from the Western Cape, KwaZulu-Natal, Mpumalanga provinces, and within the Western Cape itself (Honeycutt et al., 1987, Nevo et al., 1987, Janecek et al., 1992, Faulkes et al., 2004). More recently, studies covering the entire geographic range, and using mitochondrial and nuclear sequences, identified five lineages within the genus and proposed species status for them (Visser et al., 2018, Visser et al., 2019, Visser et al., 2020).

Moreover, published estimates of divergence dates vary substantially. The most recent common ancestor of extant genera lived about 45 million years ago (Ma) according to a phylogeny calibrated by a tentative phylogenetic placement of the extinct *Proheliophobius* (Visser et al., 2019, Visser et al., 2020), but only 31.2 Ma according to a phylogeny of the ctenohystrican rodents calibrated using multiple (non-bathyergid) node constraints (Upham and Patterson, 2015). An even younger estimate (21.7 Ma) was obtained by tip dating based on a set of bathyergid and other related fossils (Bryja et al., 2018). Consequently, the same phylogeographic lineages (=putative species) were estimated to be of Miocene (Visser et al.,

2019, Visser et al., 2020) or Plio-Pleistocene origin (Bryja et al., 2018), profoundly changing conclusions about factors underlying their differentiation.

Here, the genetic diversity of *Heliophobius* and *Georychus* is for the first time analyzed based on genome-wide nuclear data, generated by double-digest restriction-site-associated DNA (ddRAD) sequencing. We also include Sanger sequenced loci: mitochondrial cytochrome *b* (*CYTB*) gene and an enlarged set of nuclear markers. Using these complementary sequencing strategies, we addressed the following questions: (1) how many distinct lineages (putative species) are present within the genera; (2) what are their phylogenetic relationships, and (3) how old are putative species. In addition, our work joins a discussion of species delimitation studies using comparable Sanger and ddRAD sequence data. We took care to allow for the direct comparability between the analyses of both data sets, and examined whether they give similar results.

2. Material and methods

2.1. Overview of material

The material analyzed in this study includes 43 *Georychus capensis*, and 101 *Heliophobius argenteocinereus* specimens (Fig. 1, Fig. 2 and Supplementary file 1). *Georychus* samples, from 13 localities in South Africa, were collected as described in Visser et al. (2018), while *Heliophobius* samples were collected from 32 localities in Kenya, Malawi, Mozambique, Tanzania and Zambia (Fig. 2 and Supplementary file 1). The tissues from *Heliophobius* were taken immediately after the animals were sacrificed, while the samples of *Georychus* were obtained from frozen carcasses. In addition, representatives of thirteen other species of mole-rats and related families were included in the divergence dating analyses (Supplementary file 1). Tissue samples (spleen, liver, muscle, or toe) were taken and stored in 96% ethanol until DNA extraction. DNA was isolated using the commercial DNeasy Blood & Tissue Kit (Qiagen) kit according to the enclosed protocol. The extractions quality was verified on a 1% agarose gel, and the DNA concentration was fluorometrically measured by The Qubit® 2.0 Fluorometer using dsDNA BR Assay Kit (Qubit).

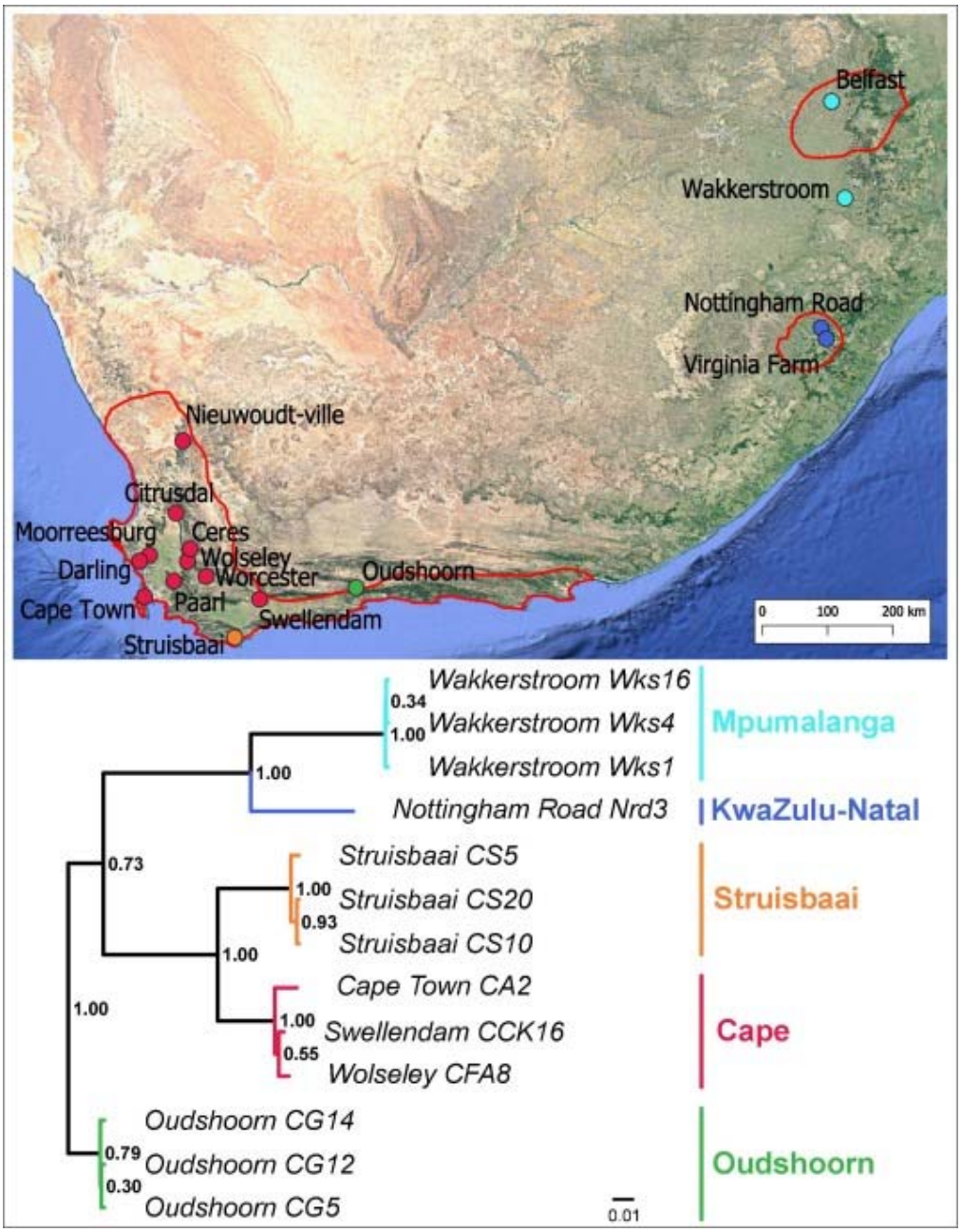


Fig. 1. Sampling sites and *CYTB* tree of *Georychus*. mPTP lineages are marked by different colors and the map shows their geographical distribution. Red outlines show estimates of distributional ranges according to IUCN Red List (IUCN 2020) and Bennett and Faulkes (2000), the latter being used for distribution in KwaZulu-Natal and Mpumalanga provinces. The tree was outgroup rooted, but the outgroups were pruned out for the display.

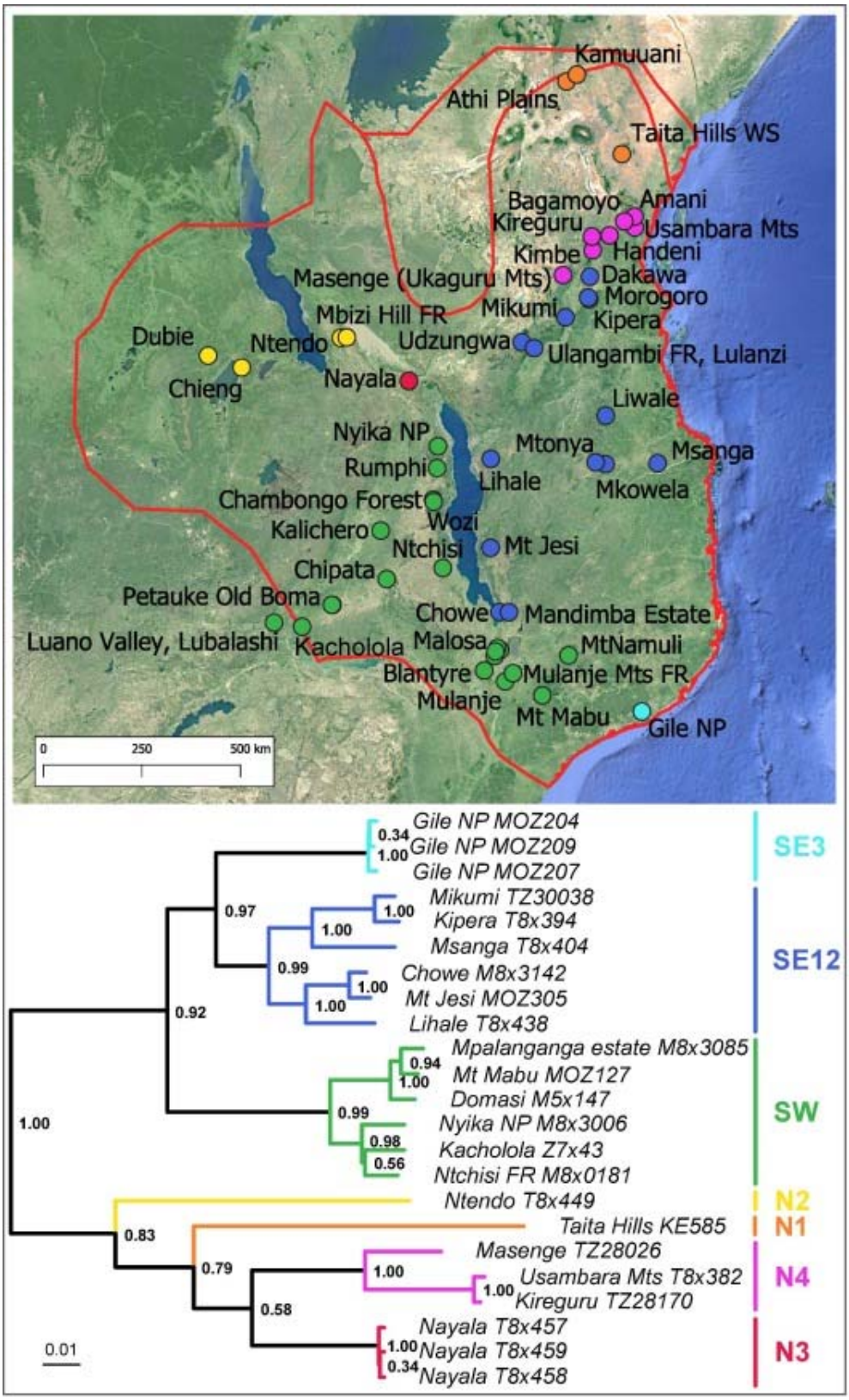


Fig. 2. Sampling sites and *CYTB* tree of *Heliophobius*. mPTP lineages are marked by different colors and the map shows their geographical distribution. Red outlines show estimates of distributional ranges according to IUCN Red List (IUCN 2020). The tree was outgroup rooted, but the outgroups were pruned out for the display.

2.2. Sanger sequencing

The mitochondrial gene for *CYTB* and six nuclear loci: recombination activating gene 1 (*RAG1*), intron 7 of the gene for β-fibrinogen (*FGB*), dehydrocholesterol reductase 24–7 (*DHCR*), NAD synthetase 1 (*NADSYN*), smoothened, frizzled class receptor 9 (*SMO*), and transient receptor potential cation channel, subfamily V, member 4 (*TRPV*) were amplified by polymerase chain reaction using the primers and protocols specified in Supplementary file 2. The newly generated sequences are available in GenBank under accession numbers OK506463-OK506476 (*CYTB*), OK506477-OK506523 (*RAG1*), OK506524-OK506570 (*FGB*), OK524003-OK524048 (*DHCR*), OK524049-OK524082 (*NADSYN*), OK524083-OK524129 (*SMO*) and OK524130-OK524176 (*TRPV*) (Supplementary file 1).

The *CYTB* data set consisted of 13 *Georychus* and 23 *Heliophobius* sequences, representing localities from across their distribution ranges and covering the previously described *CYTB* lineages (Visser et al., 2019, Bryja et al., 2018). For *Heliophobius*, we analyzed also 25 additional *CYTB* sequences to obtain a clearer picture of the *CYTB* lineage distribution (nine of them from localities included in ddRAD but not in the Sanger data, three of them were newly obtained). Two of these additional *CYTB* sequences were 136 bp long mini-barcodes designed by Galan et al. (2012) and obtained by amplicon sequencing on Illumina MiSeq platform (Illumina, San Diego, CA, USA). These sequences were obtained from skin samples of museum voucher specimens provided by Livingstone Museum (Livingstone, Zambia) and Royal Museum for Central Africa (Tervuren, Belgium). The complete list of included specimens is listed in Supplementary file 1. For *Georychus*, we analyzed also three other published *CYTB* sequences from localities not included in our main data set.

The nuclear sequences were obtained from the same 13 *Georychus* and 23 *Heliophobius* specimens preselected based on their *CYTB* haplotypes. They were computationally phased by PHASE (Stephens et al., 2001), a Bayesian method that assumes haplotypes are present in Hardy-Weinberg equilibrium frequencies. Even though the haplotype reconstruction provided us with two alleles per individual, we randomly sampled and retained just one of them to reduce the spatial non-independence of our sequence data. (By definition, two alleles from the same individual originate from nearby individuals one generation before.)

2.3. ddRAD sequencing

Material comprising 100 specimens of *Heliophobius* and 43 specimens of *Georychus* was processed into one final library for ddRAD sequencing (Peterson et al., 2012). Library preparation and basic bioinformatic processing followed procedures detailed in Piálek et al. (2019). The library was sequenced on an Illumina HiSeq 2000/2500 150P/E in the EMBL Genomics Core Facility (Heidelberg, Germany).

In the pre-processing step, ddRAD sequences were demultiplexed and trimmed of adaptors, barcodes and overhangs of restriction sites. The ddRAD loci were assembled separately in *Georychus* and *Heliophobius* using the published genome sequence of *Fukomys damarensis* (GenBank accession GCA_000743615.1) as a mapping reference. The procedure was conducted in STACKS v. 1.19 (Catchen et al., 2013). Finally, the assembled ddRAD loci were filtered to retain only those with higher than 90% sequencing success across individuals. This threshold was chosen to minimize the porportion of missing data, while retaining large number of loci. At the same time this threshold did not introduce any apparent ascertainment bias as the mean genetic distance, the number of SNPs and the number of distinct alleles was

similar whether the threshold was 50% or 90%. The final dataset contained 76,710 loci in *Georychus* and 59,973 loci in *Heliophobius*, whose alignments are available in the form of partitioned nexus files in Dryad repository associated with this paper (https://doi.org/10.5061/dryad.3n5tb2rgs0). The ddRAD sequencing produces phased sequences. For analyses that parallel those of Sanger sequenced loci we retained just one sequence per individual per locus, for the others we used both sequences.

2.4. Analysis of mitochondrial cytochrome *b* gene

We started by reconstructing the *CYTB* trees for the individuals included in our data sets. In the analysis of *Georychus*, three sequences of *Heliophobius* were used as outgroups for *post hoc* rooting, and vice versa. The trees were inferred as unrooted with branch lengths in substitution units using Bayesian inference as implemented in MrBayes 3.2.6 (Ronquist et al., 2012). For the calculation of the phylogenetic likelihood, we used a HKY + G nucleotide substitution model (Hasegawa et al., 1985, Yang, 1994) and a 12–3 scheme of partitioning according to the codon position. Four independent Markov Chain Monte Carlo (MCMC) simulations were conducted to test for convergence, which was accomplished visually in Tracer 1.7 (Rambaut et al., 2018) and numerically using the package coda (Plummer et al., 2006) for R (R Core Team, 2020). Convergence statistics included the potential scale reduction factor (Gelman & Rubin, 1992) and the effective sample size. After discarding the first 10% of trees as burn-in, the four posterior samples were pooled, the trees rooted, and outgroups removed. The pooled sample was represented by the Maximum Clade Credibility (MCC) tree (Drummond & Bouckaert 2015, p. 162) with the mean common ancestor node heights (Drummond & Bouckaert 2015, p. 92). The MCC tree was calculated in R using packages ape (Paradis and Schliep, 2019), phangorn (Schliep, 2011) and custom functions available at https://github.com/onmikula/mcctree_mrbayes.

The *CYTB* tree was partitioned into putative species using the model of multi-rate Poisson tree processes, implemented in the software mPTP (Kapli et al., 2017). The model assumes that the tree can be partitioned into *k* monophyletic lineages connected by the backbone of ancestral branches, and that every lineage as well as the ancestral backbone can be characterized by its own exponential distribution of branch lengths. The lineages can be interpreted as putative species. Average pairwise Kimura two-parameter (K2P, Kimura, 1980) distances were calculated between them for easier comparison of differentiation in these and other rodent genera. The sequences not included in the CYTB tree were classified into the lineages using the nearest neighbor criterion as applied to uncorrected genetic distances between these lineages and the included sequences.

2.5. Phylogeny and species delimitation using Sanger sequenced nuclear loci

Allelic variation of the Sanger sequenced loci was first assessed through the construction of haplotype networks using the TCS algorithm (Templeton et al., 1992) implemented in PopART v1.7 (Leigh and Bryant, 2015). For each pair of alleles, the algorithm calculates the probability that they are derived from each other, following the most parsimonious evolutionary scenario. The largest number of mutational differences still associated with a probability higher than 0.95 is taken as a criterion for inclusion of the inter-allelic links in the resulting network (Clement et al., 2000).

Species delimitation was inferred in the Bayesian framework using the multispecies coalescent (MSC) model (Rannala & Yang, 2003) as implemented in BPP v. 4.3.0 (Flouri et

al., 2018). This program uses reversible jump MCMC (Green, 1995) to sample possible species delimitations and their species trees. BPP requires the specification of candidate species, which are never split during the analysis but can be repeatedly merged and separated. Here, we used individuals as the candidate species, with each represented by a single sequence at every locus. The posterior sample of species trees was summarized by MCC tree using the mean common ancestry node heights. The tree has the candidate species at its tips. In the MSC framework, merging the candidate species is equivalent to setting their divergence time to zero and hence, all candidate species are present in all sampled trees in spite of merging. Given the presence of mergers, the MCC species tree is annotated by two sets of posterior probabilities (PPs). One set relates to the clades (species tree topology), while the other relates to the species (mergers of species tree tips). The maximum credibility species delimitation was obtained by taking species in the order of decreasing PP until all candidate species are classified. We use a term MSC-species for these BPP-delimited units to distinguish them from taxonomic species. If there were any conflicting classifications, they would be skipped and commented on separately. The program uses inverse gamma (IG) priors for population size parameters (θs) and the root divergence time of the species tree (τ_0), while other divergence times are derived from τ_0 under flat Dirichlet prior. For θs we used IG($\alpha = 3$, $\beta = 0.006$) in both genera, while for τ_0 we used IG($\alpha = 3$, $\beta = 0.015$) in *Georychus* and IG($\alpha = 3$, $\beta = 0.025$) in *Heliophobius*. The choice of priors was motivated by a preliminary analysis considering the maximum observed genetic distance (a rough surrogate of $2\tau_0$) and the average distance between individuals from the same *CYTB* lineage (relating to the value used for θ). We used species model prior 3, which gives the same prior probability to any possible number of species (Yang & Rannala, 2014). In both genera, we used the HKY + G nucleotide substitution model and assumed a strict clock, while the clock rate was allowed to vary across loci as multiples of the mean rate. The MCMC convergence was assessed by the same statistics as in *CYTB* tree, based on four independent runs of the analyses.

Even when highly supported and included in the same maximum credibility solution, MSC-species can differ widely in the degree of their genealogical distinctiveness. In particular, very young species with large population sizes may not be reciprocally monophyletic in a large proportion of loci and their genetic make-up may therefore be very similar, in spite of detectable signal of divergence. We therefore quantified the genetic distinctiveness of every MSC-species by the genealogical divergence index (*gdi*, Jackson et al., 2017, Leaché et al., 2019). BPP was rerun with the same settings, but the species tree and species delimitation was fixed to their maximum credibility solutions and the estimated parameters (θs and τs) were used for the calculation according to Eq. 7 in Leaché et al. (2019). The index is scaled to equal zero if the species is completely intermixed with othersin all gene trees and unity if all gene copies belonging to the species coalesce within its branch of the species tree. Its value is not available for MSC-species represented by just one sequence, making the corresponding θ unidentifiable.

2.6. Analysis of ddRAD sequenced nuclear loci

A subset of 100 ddRAD loci from the Sanger-sequenced individuals was subjected to the MSC based analyses in BPP with the same settings as used for the analysis of Sanger sequenced data. The loci were selected at random using a custom R routine. The number of 100 loci was considered sufficiently large to make the inference reliable while keeping computation time reasonably short. All input files of these and other BPP analyses are available in this paper's Dryad repository.

Genealogical information present in all ddRAD sequences was summarized in the co-ancestry matrix (Lawson et al., 2012). Given *N* individuals in the data set, the co-ancestry matrix is of dimension *N*x*N* and every row represents a co-ancestry profile of a single individual. For each of the other individuals, the profile contains the number of loci at which that particular individual is more similar to the profiled individual than the remaining individuals in the data set (Malinsky et al., 2018). These counts approximate how frequently the profiled individual is associated with others via its first-order cross-coalescence. The co-ancestry matrix was estimated in RADpainter (Malinsky et al., 2018).

Clusters of similar co-ancestry profiles suggest some individuals are genealogically more similar to each other than to the rest, and hence some population structure is present. Along the same line of reasoning, some level of population structure can be identified with the presence of distinct species. We used two clustering methods (and the eponymic software) to identify populations or species in the co-ancestry matrix. The first is fineSTRUCTURE (Lawson et al., 2012), a Bayesian method based on multinomial likelihood, which employs MCMC to sample the posterior distribution of possible partitions. The other is Infomap (Rosvall & Bergstrom, 2008), which treats the matrix as a graph (Mikula et al., in prep.) and uses the Louvain clustering algorithm (Blondel et al., 2008) to find the partition with the minimum description length in the sense of information theory (Rosvall & Bergstrom, 2008). Both methods can be used iteratively to merge delimited populations successively in the order of the minimum loss of likelihood and the minimum increase of description length, respectively. As a Bayesian analysis, fineSTRUCTURE outputs posterior probabilities of populations as a measure of their statistical support. The calculation of co-ancestry matrix, as described above, implicitly assumes independence of loci. A bias caused by linkage disequilibrium may be alleviated by multiplication of the matrix by factor *c*, which makes the profiles flatter. The factor was estimated as a part of co-ancestry matrix estimation. Finally, we checked for the impact of missing data by a visual comparison of the co-ancestry and missingness matrix, the latter being calculated in R from binary records of missing data.

2.7. Divergence dating

Ancestral ages of *Georychus* and *Heliophobius* lineages were estimated in a broader phylogenetic context using a combination of node calibration and tip dating of 21 taxa representing the diversity of mole-rats (Bathyergidae) and their closest living relatives, dassie rats (Petromuridae) and cane rats (Thryonomidae). According to the fossil record (Winkler et al., 2010, Barbière and Marivaux, 2015) and divergence dating of so-called Phiomorph rodents (Sallam & Seiffert, 2020) and Ctenohystricans (Upham and Patterson, 2015), these families have their origins in the Late Eocene to Early Miocene. Given the sparsity of fossil records and the uncertainty around the phylogenetic placement of fossil forms, we used just a few broad node constraints. First, we assumed an Eocene origin of the concerned speciation-extinction process (uniform density from 34 to 56 Ma before present). Then, we assumed an Oligocene origin of lineages referred to as Heterocephalinae and Bathyerginae, i.e. the lineages leading to the extant *Heterocephalus* and to the other five extant genera, respectively. We did not constrain them to be in a sister relationship, however. This was accomplished by putting the same uniform calibration density within Oligocene boundaries (23.0–34 Ma) on the origin of branches leading immediately to the most recent common ancestor of the particular group, i.e. on the origins of stem Heterocephalinae and stem Bathyerginae. Tip dating information comes from the inclusion of 42 fossil taxa that can be associated with some of the internal lineages, at least with stem Bathyergidae or stem Petromuridae + Thryonomidae (see López Antoñanzas and Sen, 2005, Denys, 2011, Mein

and Pickford, 2006, Mein and Pickford, 2008, Winkler et al., 2010, and Sallam & Seiffert, 2020 for the reference). Taken together, both extinct and extant taxa were classified into 13 clades, being constrained as monophyletic (Bathyerginae being the exception). Molecular data from the extant taxa are from five nuclear markers (*DHCR*, *FGB*, *RAG1*, *SMO*, *TRPV*) either sequenced or retrieved from GenBank. The complete list of included taxa, their classification, specimens representing them and reference to original data is contained in Supplementary file 1. The input file of the divergence dating analysis is available in this paper's Dryad repository.

The time-calibrated species tree was inferred using Bayesian implementation of the fossilized birth–death model (Heath et al., 2014) available in BEAST 2 (Bouckaert et al., 2014). The fossilized birth–death process was assumed constant and uninformative priors were put on its parameters. As morphology of most of the fossils highly deviate from that of extant species, we did not assume they were ancestors of modern taxa and set the removal probability parameter of the model to 0.75. For all loci, we assumed a strict clock and HKY + G substitution model. The analysis was run four times to check for convergence, and the pooled posterior sample of trees was represented by the MCC tree with mean common ancestor heights. The tree was visualized in R using package MCMCtreeR (Puttick & Title, 2019).

2.8. Species distribution modelling

Species distribution models were fit separately for each genus using MaxEnt (Phillips et al., 2006) as implemented in the R package maxnet (Phillips et al., 2017). The background area for each genus was selected *ad hoc* to form an overlapping convex hull around our presence records with presumably suitable habitats selected among ecoregions of Olson et al. (2011). The presence data comprise all georeferenced sequence records and all georeferenced records from museum databases accessed via vertnet (www.vertnet.org) and African Mammalia (http://projects.biodiversity.be/africanmammalia) web sites. As predictors, we used three climatic curves (average monthly precipitation, minimum and maximum temperature as the elementary descriptors of a climate), downloaded from WorldClim database (Fick & Hijmans, 2017). To gain insight into past distributional changes, we also downloaded the paleoclimatic layers for the mid-Holocene and the last glacial maximum (LGM, Gent et al., 2011). Both presence records and predictor values were sub-sampled to 0.5° resolution, which means the background was discretized into quadrate cells with a size of 0.5°×0.5°. MaxEnt predictions were expressed in the form of relative occurrence rates (RORs), i.e. relative probabilities of occurrence in particular cells. These were scaled relative to the uniform prior expectation (1/no. of background cells) so that the values > 1 indicate that the model brought extra-evidence for the species' presence. When plotted, the color scales were adjusted to enhance the interpretability of the maps. In particular, the range of colors was reduced to reflect the strength of evidence provided by the model, and the scales were made only as contrasting as the predicted RORs.

The predictor curves were treated as described in Mikula (2021), including their superimposition and projection to natural cubic splines. The superimposition was based on present precipitation curves only, but the superimposed present curves served as a reference for superimposition of the past ones, and the same curve shifts were automatically applied also to the corresponding temperature curves. The number of knots (and hence complexity) of cubic splines was selected in a preliminary analysis using AICc (Burnham and Anderson, 2002, Warren and Seifert, 2011) as a measure of model fit. For precipitation, minimum temperature and maximum temperature curves, respectively, the procedure suggested 4, 9,

and 8 knots in *Georychus*, while it was 12, 8, and 8 knots in *Heliophobius*. When represented as splines, the curves were used without any further transformation (MaxEnt's "linear feature"), and lasso regularization was performed as a part of model fit. R functions necessary for the handling of climatic curves are available at https://github.com/onmikula/maxent_tools. The estimates of suitable climatic conditions were obtained as the ROR-weighted means of the original climatic curves. As the curves entered the model time-shifted, the averaging was done separately for each group of identically shifted curves.

3. Results

3.1. Structure of mitochondrial DNA

The *CYTB* tree for *Georychus* was estimated to contain five mPTP-delimited lineages, each restricted to a specific geographical area (Fig. 1). These were labelled according to the names of regions or sampling sites where they occur, namely as Cape (multiple sites in the Western Cape Province), Struisbaai (a site at the very tip of the Western Cape), Oudshoorn (a site further east in the Western Cape), KwaZulu-Natal, and Mpumalanga (these are represented by single collection sites). While the sister relationships between the Cape and Struisbaai, as well as between KwaZulu-Natal and Mpumalanga, are well supported (PP = 1.00), the position of Oudshoorn is not well resolved (in the MCC tree it is sister to the rest, but with only PP = 0.73). Average K2P distances between lineages (Table 1) vary from 6.1% (Cape-Struisbaai) to 14.1% (Cape-Mpumalanga).

Table 1. Average K2P distances between *CYTB* lineages of *Georychus* expressed as percentage of sequence divergence.

	Cape	**Struisbaai**	**Oudshoorn**	**KwaZulu-Natal**
Struisbaai	6.1			
Oudshoorn	8.4	9.0		
KwaZulu-Natal	13.2	13.3	10.8	
Mpumalanga	14.1	13.0	11.6	9.2

In *Heliophobius*, the *CYTB* tree contained seven mPTP-delimited lineages with a strong geographic pattern (Fig. 2). Two large clades are apparent in the tree: the northern clade with four lineages (N1, N2, N3, N4) and the southern one with three lineages (SE12, SE3, SW). The letters N, SE and SW refer to north, southeast, and southwest, respectively. The numbering in labels was chosen for the sake of compatibility with Bryja et al. (2018). All lineages represented by more than one individual were found to be monophyletic with PP $\geq$ 0.99, but deeper phylogenetic relationships were sometimes resolved with more limited support. The northern and southern clades are monophyletic with PP = 0.83 and 0.92, respectively, SE12 is sister to SE3 with PP = 0.97, and N3 to N4 with PP = 0.58. The northern lineages were present in southern Kenya (N1), the Eastern Arc Mountains (N4), and highlands along the Rukwa Rift (N2, N3). The southern lineages occupy the oceanic side of the Eastern Arc Mountains, highlands rimming the eastern coast of Lake Malawi and lowlands in southern Tanzania (SE12), the Indian Ocean coast in northern Mozambique (SE3) and highlands of Malawi and northeastern Zambia (SW). The northern and southern clades were found most closely to each other on the opposite sides of the Eastern Arc Mountains. Average K2P distances between lineages (Table 2) vary from 7.3% (SE12-SE3) and 7.4% (N3-N4) to 15.7% (N1-SW).

Table 2. Average K2P distances between *CYTB* lineages of *Heliophobius* expressed as percentage of sequence divergence.

	N1	N2	N3	N4	SE12	SE3
N2	13.8					
N3	11.7	10.9				
N4	10.9	11.3	7.4			
SE12	14.8	13.2	12.4	13.1		
SE3	14.8	13.6	12.3	11.4	7.3	
SW	15.7	14.0	13.4	14.2	9.0	9.4

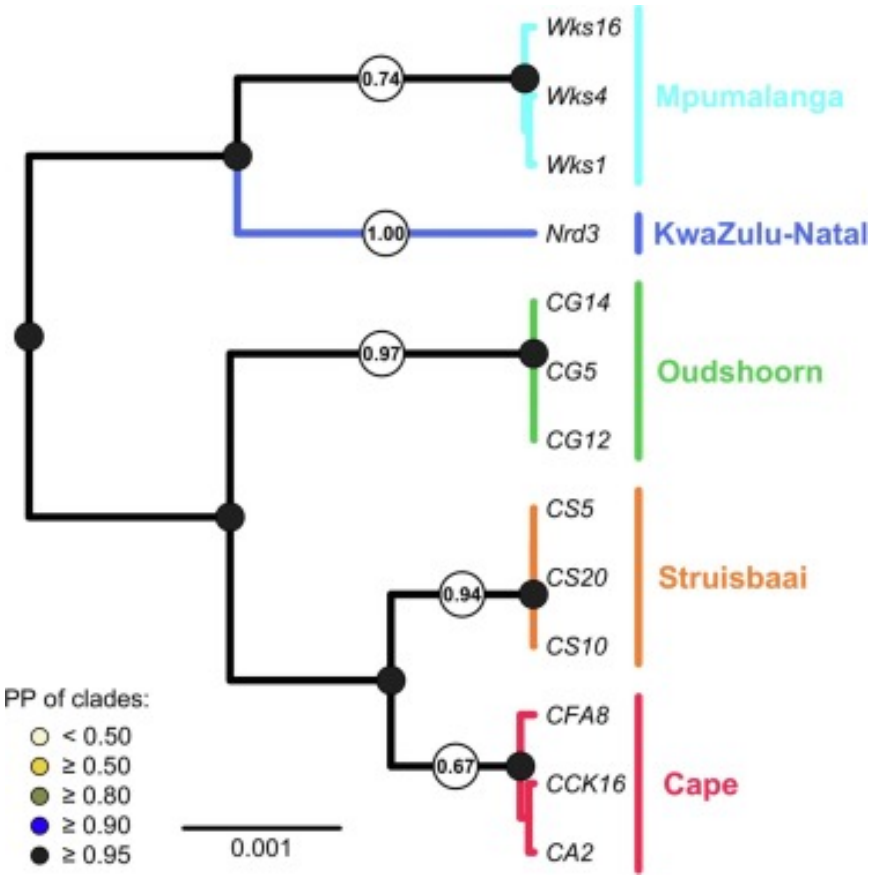


Fig. 3. Species tree of *Georychus* inferred in BPP from ddRAD loci. Maximum credibility MSC-species are marked by sidelines, labels and different colors and their PPs are shown as numbers on branches supporting them. PPs of the clades are indicated by color dots as specified in the legend.

3.2. Bayesian delimitation of MSC-species

In *Georychus*, analysis of 100 ddRAD loci in BPP identified five MSC-species that matched the *CYTB* lineages (Fig. 3). While support for both the Cape and Mpumalanga groups were low (PP = 0.67 and 0.74 respectively), support was high for Struisbaai (PP = 0.94), Oudshoorn (PP = 0.97), and KwaZulu-Natal (PP = 1.00). These MSC-species form two main clades in the species tree, separated by a divergence of $\tau_0 = 0.0032$. The southwestern clade

consists of MSC-species Cape, Struisbaai and Oudshoorn. The divergence time of Cape + Struisbaai and Oudshoorn is $\tau = 0.0019$, while between the Cape and Struisbaai it is $\tau = 0.0009$. The eastern clade has two MSC-species KwaZulu-Natal and Mpumalanga, whose divergence is also $\tau = 0.0019$. All these clades were supported with PP = 1.00. In terms of genealogical distinctiveness (Figure S.1, Supplementary file 3), Mpumalanga, Oudshoorn and Struisbaai were clearly separated (gdi $\geq$ 0.81), while separation of Cape was inconsipcuous (*gdi* = 0.55). The index could not be calculated for KwaZulu-Natal as it consists of a single individual only.

The analysis of six Sanger loci (Figure S.2, Supplementary file 3) shows one important difference: an individual from Swellendam forms its own MSC-species (PP = 0.64), sister to Struisbaai collapsed with the rest of Cape, a merger included in the maximum credible species delimitation, but having PP = 0.32 only. The remaining three MSC-species were Oudshoorn (PP = 0.93), KwaZulu-Natal (PP = 1.00) and Mpumalanga (PP = 0.91). The species tree topology was congruent with that of ddRAD-based phylogeny, but divergences were markedly deeper ($\tau_0 = 0.0081$). The *gdi* was high in Mpumalanga (0.96) and Oudshoorn (0.96), but much lower in Cape + Struisbaai (0.61).

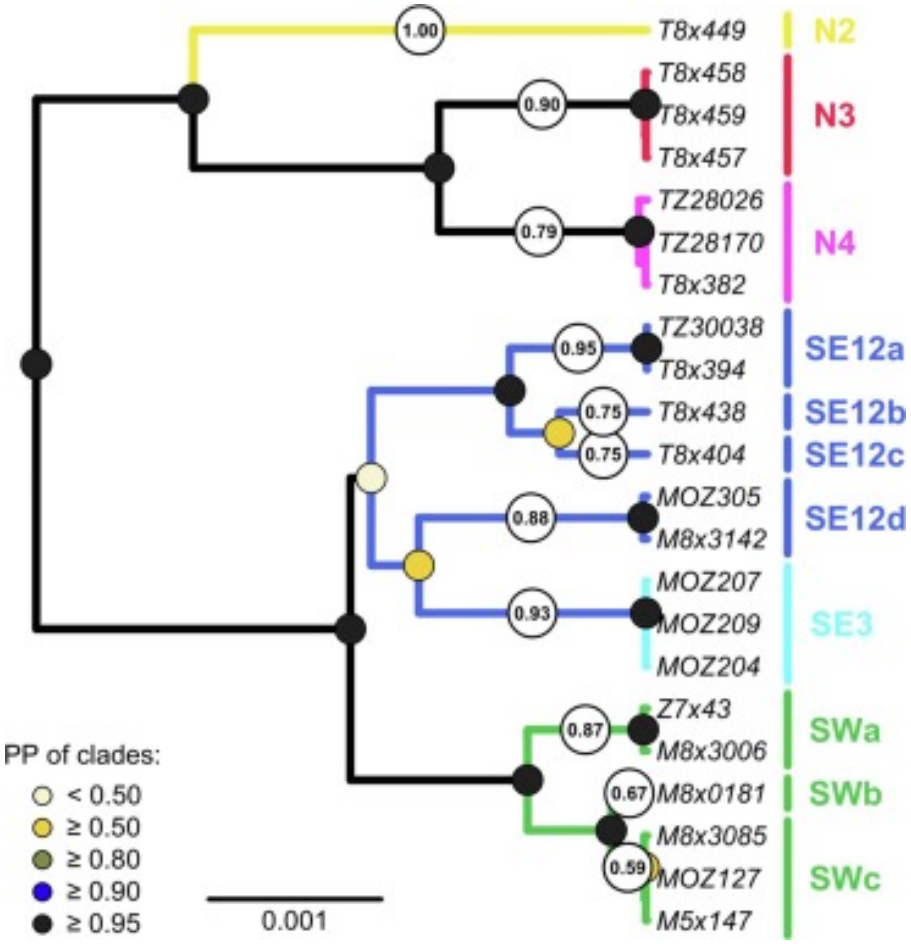


Fig. 4. Species tree of *Heliophobius* inferred in BPP from ddRAD loci. Maximum credibility MSC-species are marked by sidelines, labels and different colors and their PPs are shown as numbers on branches supporting them. PPs of the clades are indicated by color dots as specified in the legend.

In *Heliophobius*, analysis of 100 ddRAD loci recovered eleven MSC-species matching all *CYTB* lineages except for SE12 and SW, which went out fragmented into smaller units (Fig. 4). The MSC-species are labelled in accord with the corresponding *CYTB* lineages, the smaller units being marked by letters 'a' to 'd'. N1 lineage from *CYTB* tree is not present in the tree due to lack of data. PP of the MSC-species varied from 0.59 to 1.00. It was low (<0.80) in MSC-species SE12b, SE12c and SWb, which are represented by single individuals and do not match any separate *CYTB* lineage, but also in SWc, which is a loose association of three individuals from geographically distant sites in Malawi and northern Mozambique. The species tree has its root at $\tau_0 = 0.0035$ and it is structured into the northern and the southern clade, both supported with PP = 1.00. The northern clade is structured more deeply, with basal divergence time $\tau = 0.0026$, compared to $\tau = 0.0017$ in the southern clade. The internal topology of the northern clade is also fully resolved (PP = 1.00), unlike to the southern clade, where monophyly of SE clade is doubtful (PP = 0.41) and so are some of relationships within this group of MSC-species. Genealogical distinctiveness (*gdi*) of MSC-species varied from 0.37 to 0.89, but only in four cases exceeded 0.70, the conventinal sigificance level of Jackson et al. (2017). These were N4 (0.72), N3 (0.81), SE12d (0.81) and SE3 (0.89). For the complete *gdi* results, see Figure S.3 (Supplementary file 3).

Six Sanger loci analysis gave a similar result with 13 MSC-species (Figure S.4, Supplementary file 3). Here, the N1 lineage was present and found as an unequivocally supported MSC-species (PP = 1.00). Other N lineages matched the previous results, but with much lower support for N3 (PP = 0.53). In the southern clade, the support was lower for SE12d (PP = 0.58) and SE3 (PP = 0.79) and species delimitation was not resolved within SW subclade. PPs of all its MSC-species were ≤ 0.50 and SWa was split to single individuals from mutually distant sites of Nyika (10.6°S, 33.8°E) and Kacholola (14.8°S, 30.6°E). The species tree's topology was less resolved with a doubtful monophyly of the northern clade (PP = 0.67) and the placement of N1 uncertain (PP = 0.47). Compared to its ddRAD-based counterpart, the Sanger-based species tree of *Heliophobius* was deeper ($\tau_0 = 0.0056$). The values of *gdi* were high in all six MSC-species, for which they could be calculated: SE12a (0.70), SWc (0.75), SE12d (0.83), SE3 (0.87), N3 (0.86) and N4 (0.90).

3.3. Nuclear haplotype networks

In *Georychus*, haplotype networks of six Sanger sequenced loci generally show KwaZulu-Natal and Mpumalanga populations separated from the Cape and Struisbaai populations (Figure S.5, Supplementary file 3). Oudshoorn is generally closer to the Cape + Struisbaai, although in SMO it is just between the two major groups and in FGB it shares alleles with Mpumalanga.

Haplotype networks of the same loci in *Heliophobius* separate the northern and southern lineage, although even here we can see exceptions. When mitochondrial lineages are mapped to the networks (see the colors), the two major clades are fully separable on three nuclear loci (*FGB, SMO, NAD SYN*), but not on the others (Figure S.6, Supplementary file 3).

3.4. Clustering on co-ancestry matrix

In the co-ancestry matrix calculated from all ddRAD sequenced loci, fineSTRUCTURE delimited 31 populations of *Georychus* (all with PP = 1.00), most often with one or two members (grey boundaries in the heat map, Fig. 5). Close examination of these populations reveals they always comprised individuals from the same locality, sampled up to three km

from each other, but the nearest spatial neighbors were not always found in the same population. In contrast, Infomap merged all individuals into a single population (black boundary around the heat map). Visual examination of the heat map suggests the presence of several clusters that are consistent with MSC-species inferred by BPP. For instance, the most pronounced co-ancestry peak is observed in the upper left corner of the heat map (see the square of magenta color) and all individuals involved were from Mpumalanga, which is corroborated by the match with BPP tree. At the same time, however, the co-ancestry signal of variable strength is apparent also between these diagonal blocks. For instance, the Oudshoorn block shows a uniform signal of co-ancestry both with KwaZulu-Natal + Mpumalanga and Struisbaai + Cape blocks (note the wide orange stripes). Hierarchical fineSTRUCTURE clustering shows that populations with the same *CYTB* type also had a similar nuclear genome (see the coloring of branches in Figure S.7, Supplementary file 3).

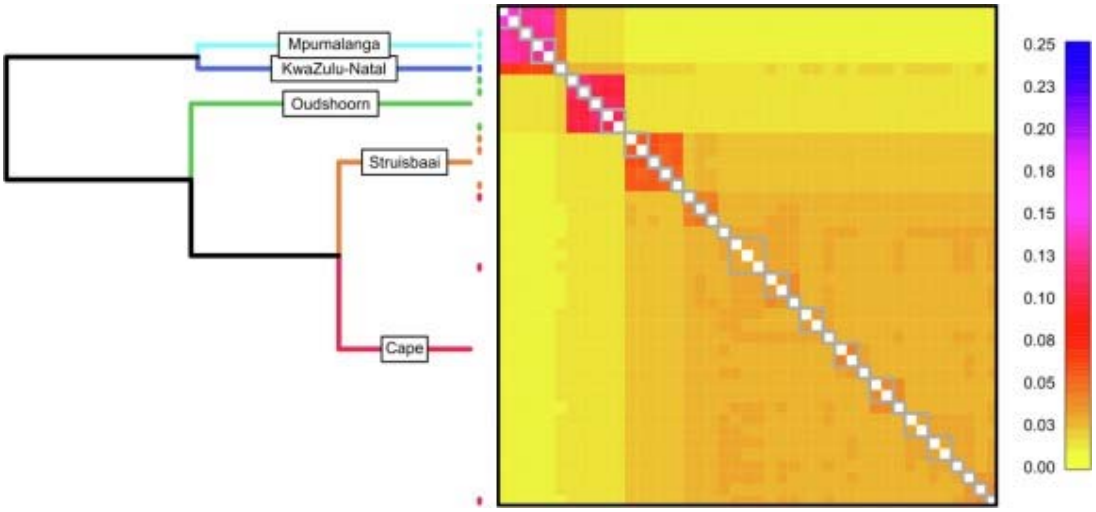


Fig. 5. Co-ancestry matrix of *Georychus* shown as a heat map. The scale indicates proportion of the nearest neighbor loci. fineSTRUCTURE clusters are shown in grey and Infomap cluster in black. On the left, there is the ddRAD based species tree with individual colored tips (short vertical bars) matching the corresponding rows of the heat map (co-ancestry profiles).

Fig. 6 shows the co-ancestry matrix, population clustering and BPP species tree of *Heliophobius*, fineSTRUCTURE delimited 53 populations (all with PP = 1.00). They often had just one or two members and were typically confined to a single locality or two localities up to five kilometres apart. In two cases, however, populations contained individuals sampled tens of kilometres from each other (not shown). Also, the nearest neighbors were not always found in the same population. Infomap delimited just two populations congruent with the northern and southern clades in BPP species tree. Also visually, the presence of these two clusters is the most pronounced feature of the heat map as they mutually share just a uniformly low background co-ancestry (note the large yellow rectangles). Within these clusters, rich patterns of co-ancestry are apparent, including conspicuous diagonal blocks, larger than fineSTRUCTURE populations, which show various proportions of mutual co-ancestry and roughly correspond to BPP's MSC-species.

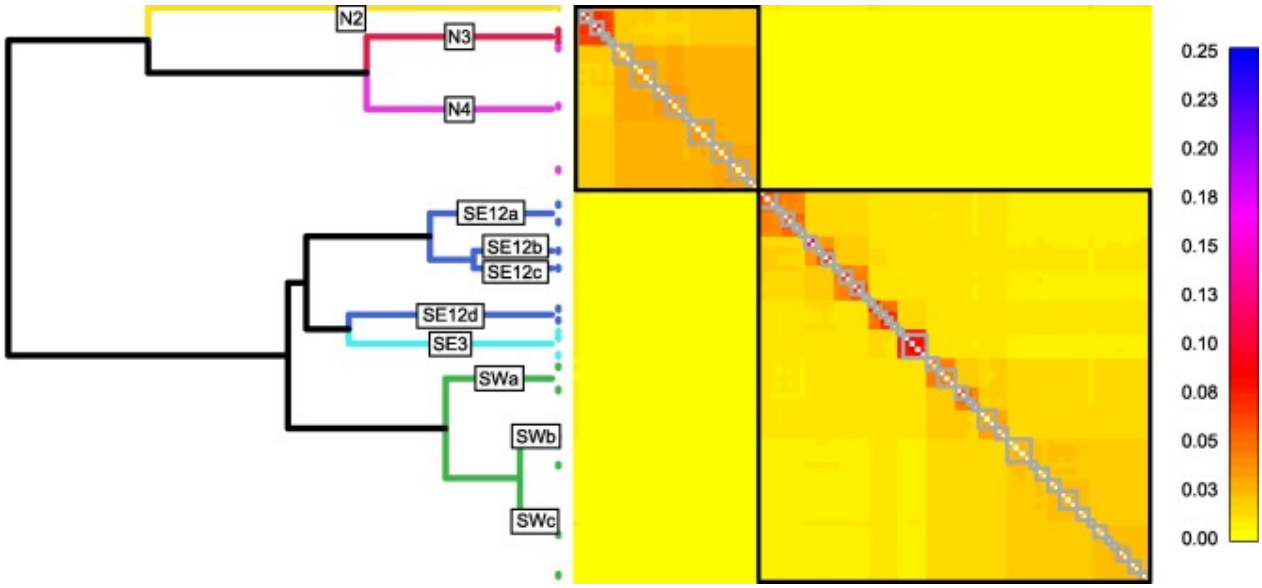


Fig. 6. Co-ancestry matrix of *Heliophobius* shown as a heat map. The scale indicates proportion of the nearest neighbor loci. fineSTRUCTURE clusters are shown in grey and Infomap cluster in black. On the left, there is the ddRAD based species tree with individual colored tips (short vertical bars) matching the corresponding rows of the heat map (co-ancestry profiles).

With just two populations, a hierarchical Infomap solution would be trivial. Hierarchical fineSTRUCTURE clustering (Figure S.8, Supplementary file 3) is characterized by unambiguously deep distinction between the northern and southern Infomap populations and a series of less conspicuous but clear differences within them. Visual check does not suggest missingness to introduce bias into the population clustering results (see Figure S.8).

3.5. Divergence dating

The diversification of extant bathyergid lineages (Fig. 7) began in middle Oligocene when *Heterocephalus* split from the other genera, an event dated to 29.02 Ma, with 95% highest posterior density (HPD) interval 27.04–31.01. *Heliophobius* clade separated in middle Miocene, 13.37 (11.18–15.49) Ma, and the clades of *Bathyergus* + *Georychus* and *Cryptomys* + *Fukomys* diverged in the same period, 10.63 (8.94–12.51) Ma. Splits of *Georychus* and *Bathyergus* as well as of *Cryptomys* and *Fukomys* date to the end of Miocene 6.96 (5.47–8.43) Ma and 5.85 (4.47–7.41) Ma, respectively.

The intrageneric diversification of the studied genera occurred during Pleistocene. The first split within genus *Georychus* (lineages in the Western Cape from the rest) occurred at 2.20 (1.47–2.91) Ma followed by the splits between clades Cape + Struisbaai and Oudshoorn 1.96 (1.21–2.73) Ma, Mpumalanga and KwaZulu-Natal 1.33 (0.66–2.00) Ma and Cape and Struisbaai 0.27 (0.01–0.62) Ma. The N and SE + SW clades within genus *Heliophobius* diverged 2.31 (1.47–3.12) Ma and clades SE and SW split in 1.08 (0.59–1.66) Ma.

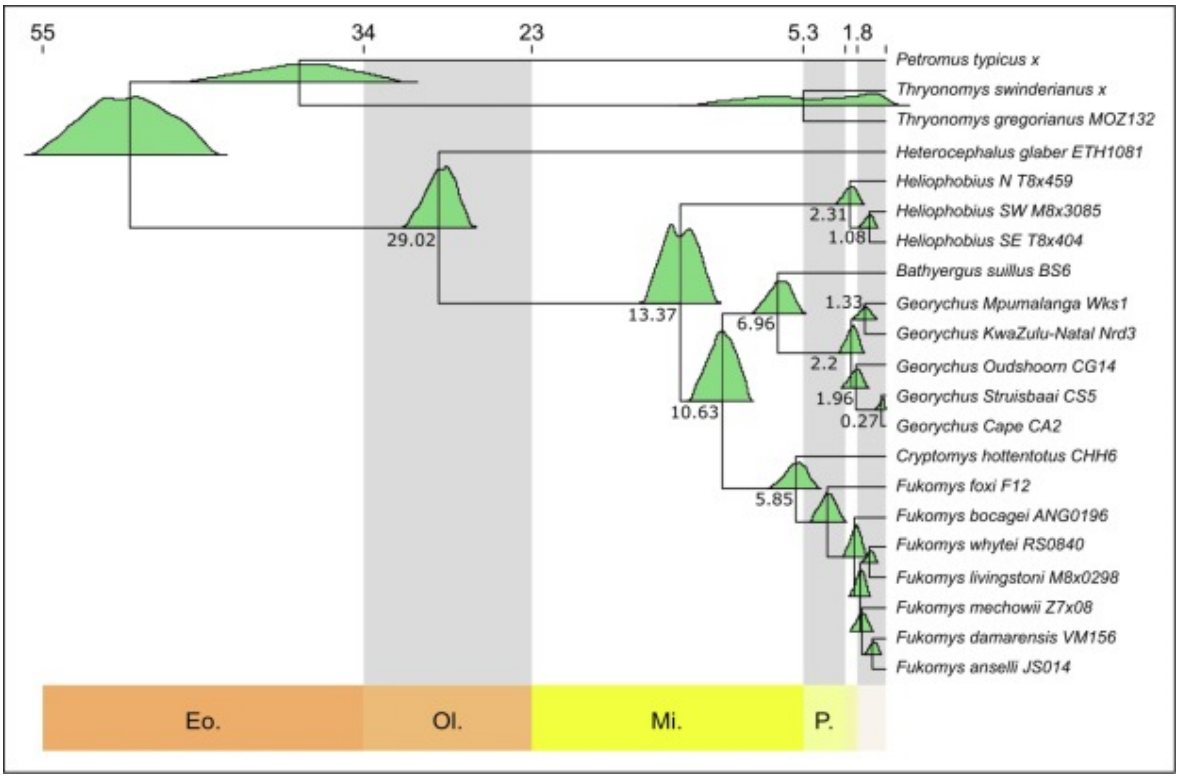


Fig. 7. Time-calibrated phylogeny of bathyergids and their closest relatives. Fossil species were pruned off the tree for the display. Time scale is annotated with names of geological periods (Eo. = Eocene, Ol. = Oligocene, Mi. = Miocene, P. = Pliocene) and their time limits (in Ma). Posterior distributions of ages are shown in green at the particular nodes. In addition, median ages are given for the nodes discussed in the text.

3.6. Models of present and past distributions

The present distribution model of *Georychus* shows the genus is confined to discontinuous and rather sharply delimited areas with suitable climate (Fig. 8). One such an area corresponds to the coastal region in the Western Cape, and some distance from the coast it is quickly replaced by an inhospitable environment. Towards the east, it continues only in a very narrow coastal belt. On the eastern side, there are two poorly connected areas with scaled RORs > 1, encompassing the sampled localities in KwaZulu-Natal and Mpumalanga provinces. The prediction for the LGM shows a very similar pattern. The suitable areas are on the same places as in the present, but they are larger, their margins are a bit less contrasting against the background, and the eastern part of the predicted distribution is continuous. Fig. 9 shows the ROR-weighted climatic curves (in green) on the top of the curves observed at presence locations (in dark yellow) and the frequency distribution density of background values (grey violin plots). There are two precipitation regimes, one with May-to-August and the other with December-January rainfall peak, but both supporting *Georychus* by providing substantially more moisture than present in the rest of southern Africa. The geographical arrangement of the regimes is shown in Figure S.9 (Supplementary file 3). While the Cape region has its rainy season from May to August and its monthly precipitations vary around 100 mm, in the eastern part of *Georychus* distributions, precipitation is mainly concentrated to December and January and its monthly values may reach 150 mm. Temperature curves show *Georychus* prefers colder parts of the region. The rainy season is cold in the Cape region (winter rainfall), but warm in the east (summer or year-round rainfall).

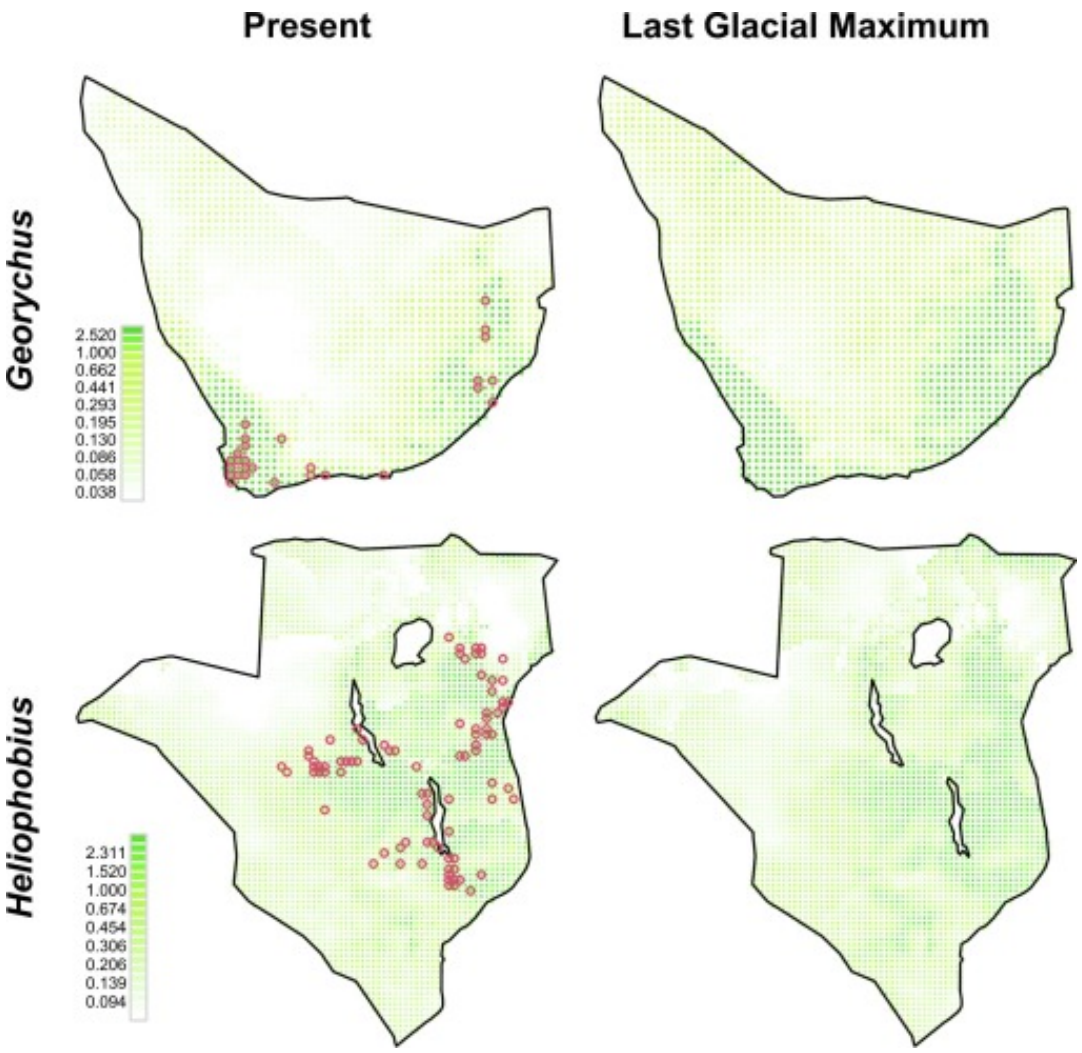


Fig. 8. Species distribution models of *Georychus* and *Heliophobius* for present and the Last Glacial Maximum conditions. Coloration indicates scaled RORs, red dots mark cells with presence records.

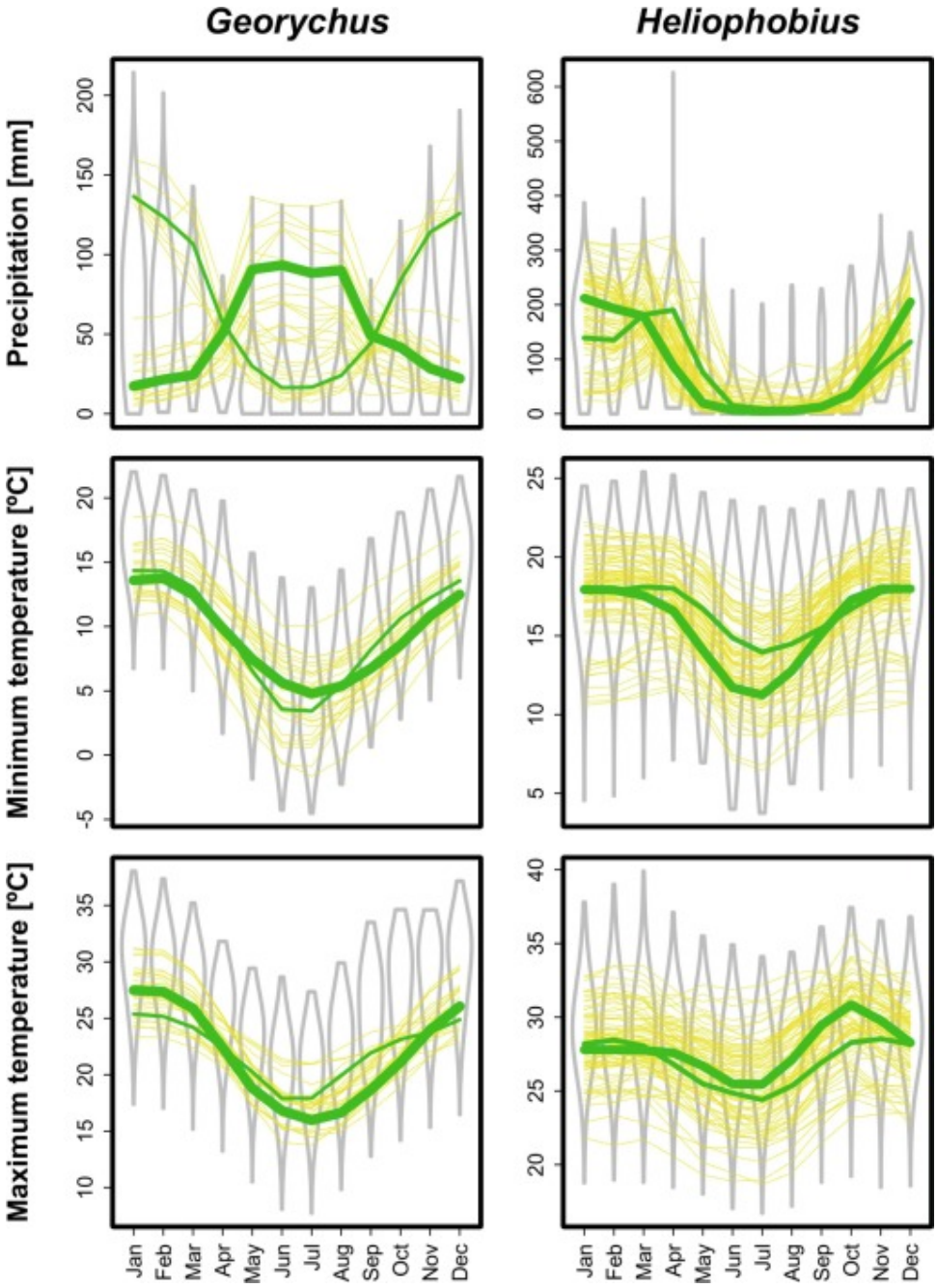


Fig. 9. Climatic curves predictive of *Georychus* and *Heliophobius* presence. Green lines show ROR-weighted means of the curves for particular climatic regimes (identified with distinct shifts in curve superimposition). The width of the lines reflect sum of RORs assigned to the regime. Dark yellow lines are the input climatic curves at the presence locations. Violin plots show kernel density estimates of probability distributions of particular climatic variables across the relevant background.

In general, the predicted present distribution of *Heliophobius* (Fig. 8) is rather complex in its shape, although at its northern margin (in southern Kenya) predicted RORs drop extremely sharply from high to almost zero values. The core of the predicted distribution includes the wider surroundings of Tanzanian montane circle and the Lake Malawi Rift, including lowlands in the southeastern Tanzania. The predicted LGM distribution shows no conspicuous differences from the present one. The average climatic curves are also shown for

two dominant precipitation regimes (Fig. 9), i.e. for the main areas, where precipitation curves were found shifted by the algorithm preparing MaxEnt predictors. While precipitation peaks in April along EAM, in most of the other sites with presence records the peak comes earlier, in January (Figure S.9 Supplementary file 3).

4. Discussion

Both solitary genera of African mole-rats studied here, namely *Georychus* and *Heliophobius*, were recently considered to include several cryptic species. In our study, we approached the issue using a large genomic (ddRAD) data set. Cluster analyses of the data highlighted strong local population structure, but also large-scale similarities within these populations. The most parsimonious descriptions of pairwise genetic similarities suggested the presence of a single species of *Georychus* and two species of *Heliophobius*. In addition, we conducted identical MSC-based species delimitation analyses with a subset of these genomic data and six Sanger sequenced loci. Both analyses detected phylogeographic structures similar to those described in previous studies, with the subset (one hundred) of ddRAD loci producing better supported and presumably more accurate results.

4.1. Species delimitation highly dependent on the method used

The results of MSC species delimitation analyses on six Sanger-sequenced and 100 randomly selected ddRAD loci were broadly congruent, but ddRAD-based analyses resulted in slightly less, but better supported, MSC-species. On the other hand, the number of clusters found in the same co-ancestry matrices differed by orders of magnitude, depending on the clustering algorithm used: 31 and 53 with fineSTRUCTURE compared to one and two with Infomap, within *Georychus* and *Heliophobius* respectively. Notably, MSC-species trees of both Sanger-sequenced and ddRAD loci and fineSTRUCTURE dendrogram obtained from the co-ancestry matrix have very similar topologies. The clustering and species delimitation analyses, therefore, detected the same underlying population structure, but differed in their resolution scale.

These differences are best explained by different assumptions made by the analyses. MSC models fully resolve gene trees as coming from discrete populations separated by a hierarchy of successive splits, which were not followed by any post-split gene flow. Although the nearest-neighbor approximation of pair-wise genealogical proximity producing the co-ancestry matrix omits part of the genealogical information, it allows for any kind of allele sharing across all phylogenetic scales. As a result, MSC based analyses tend to avoid fine scale clustering, which is not accountable without invoking ongoing gene flow (Zhang et al., 2011), but also coarse scale clustering, if it created clusters with recognizable traces of internal splits (Leaché et al., 2019). Accordingly, BPP returned a moderate number of MSC-species, often highly supported in terms of PP, although not always distinctive in terms of *gdi*. BPP is a fully Bayesian analysis jointly inferring both species and gene trees, and therefore its gene trees tend to reflect those genealogical patterns that can be explained without gene flow. On the contrary, the co-ancestry matrix may reflect a multitude of processes, as long as they left a trace in the distribution of genetic nearest-neighbors. It can contain a variety of features supporting both uniqueness of small groups of individuals or genealogical connections between them, depending on the chosen point of view (Hellenthal et al., 2014). In this respect, fineSTRUCTURE and Infomap are somewhat opposites. While fineSTRUCTURE looks for the smallest groups with specific co-ancestry profiles (Lawson et al., 2012), Infomap searches for clustering that makes the description as concise as possible

(Rosvall and Bergstrom, 2008). If two groups of individuals find a different number of the nearest-neighbor loci within each other, fineSTRUCTURE infers them as separate populations. The same mutual similarities, however, can easily cause Infomap to merge the groups together, provided they are strong enough to make the two-group description more complex. Interesting in this respect is to consider haplotype networks of Sanger sequenced loci (Figures S5 and S6). In spite of being few, they show instances of allele sharing which can ultimately cause the merging of the populations in Infomap.

The nuclear datasets differ in two important aspects: Sanger sequenced loci are few in number (six here), but they are specifically targeted and thus can be selected to be useful for particular research questions (e.g. Rodríguez-Prieto et al., 2014). In contrast, ddRAD loci are numerous (tens of thousands here), but their selection is governed only by the distribution of restriction sites across the genome. This can make them a random sample of a genome, but the ability to test this assumption is limited. In this study, it appears that ddRAD dataset can provide a more accurate picture of genetic variability of sequenced individuals and hence more accurate results. This interpretation follows from the fact that the delimitations based on ddRAD loci were simpler and less ambiguous than the delimitation based on Sanger sequenced loci. On the other hand, it implies that six Sanger sequenced loci were almost as effective as one hundred of ddRAD loci. In this case at least, but possibly in many others, analyses based on a handful of pre-specified loci may provide a quite accurate picture of MSC-species boundaries. Species delimitation studies seldom use Sanger and ddRAD sequence data in fully comparable fashion, although Gottscho et al. (2017) did so in their study of ringe-toed lizards (*Uma notata*). They applied Bayes factor delimitation to Sanger and ddRAD-derived SNP data from Leaché et al. (2014) and found in both datasets good support for the same species rich delimitation.

4.2. Genetic diversity of *Georychus*

All hitherto phylogeographic studies of *Georychus capensis* are largely in agreement with our results – all trees, dendrograms and networks have a similar topology regardless of the methods and datasets employed (targeted Sanger loci vs. restriction site dependent ddRAD loci, likelihood phylogenies vs. haplotype networks vs. clustering on co-ancestry matrix). What does differ is the number of clusters considered as distinct. The most recent studies revealed six lineages based on two mitochondrial markers (Visser et al., 2018), the follow-up analyses of *CYTB* (Visser et al., 2019, Visser et al., 2020) and three nuclear fragments (Visser et al., 2019) merged two of them, hence giving five genetic lineages. In this study, BPP confirmed the presence of five distinct MSC-species, but their genetic similarities in the co-ancestry matrix cause Infomap to collapse all of them into one gene pool and also *gdi* values show they are often not very distinctive.

4.3. Genetic diversity of *Heliophobius*

The first comprehensive study of *Heliophobius* reported strong phylogeographic structure across its distribution and proposed six main mitochondrial lineages as separate species (Faulkes et al., 2011). This study contained, however, several mistakes including overlaps in lineage distributions due to misplaced sampling sites and seemingly distinct phylogeographic lineages formed by chimeric museum sequences (see Bryja et al., 2018, Suppl. 2 for a detailled account). The following analyses of mitochondrial (*CYTB*) variation confirmed the presence of strong phylogeographic structure (Bryja et al., 2018, Visser et al., 2019). Bryja et al. (2018) delimited eight mitochondrial lineages informally as visually distinctive

phylogenetic clusters confined to a specific region, but then assessed whether they correspond to separate nuclear gene pools. They were designated as candidate species in a multispecies coalescent species delimitation analysis, and almost all of them were found distinctive, although their divergences were rather shallow. The only deep divergences were between members of northern and southern lineages (cf. Bryja et al., 2018, Fig. 3).

Our study shows similar results, but joins mPTP-based maximum likelihood delimitation of *CYTB* lineages and confronts them with nuclear MSC-species, whose inference was unaware of mPTP results. The obtained delimitations are roughly congruent, although not identical. In particular, the structure of the southeastern lineage differs on the mitochondrial (*CYTB*) and nuclear (species) tree. The clustering on co-ancestry matrix points towards the same conclusion as exploration of MSC-species delimitations in Bryja et al. (2018, Fig. 3): fineSTRUCTURE shows strong local structure (53 populations), but Infomap shows just a single really deep divergence, between the northern and southern populations. The strong local structure present within these major clusters is also supported by high *gdi* of MSC-species.

Two other points are worth mentioning here. Microsatellite analysis of Bryja et al. (2018) showed one southern population close to the presumed northern-southern contact zone to possess alleles otherwise found in northern populations. Whether it represents recent gene flow is an important unresolved question, but the same individuals are included in our ddRAD data set and show no apparent trace of admixture. Secondly, our sampling was rather comprehensive, but populations living southwestwards to the East African Rift System (EARS) in the Democratic Republic Congo is represented just by a single *CYTB* mini-barcode. This short sequence indicates the presence of N2 lineage, but extended sampling in this region is of prime importance for completing a revision of the genus.

4.4. Dating of intraspecific divergences in *Heliophobius* and *Georychus*

According to estimates using the fossilized birth–death model, recent lineages of both genera are of Pleistocene origin. This estimate is in agreement with estimated diversification times in many other savannah rodent taxa in the same regions. Examples from eastern Africa include genera *Mastomys* (Colangelo et al., 2013), *Mus* (subgenus *Nannomys*, Bryja et al., 2014), *Gerbilliscus* (McDonough et al., 2015), *Saccostomus* (Mikula et al., 2016), *Acomys* (Aghová et al., 2017, Petružela et al., 2018) and *Aethomys* (Mazoch et al., 2018). In southern Africa, it includes *Rhabdomys* (Rambau et al., 2003), *Micaelamys* (Russo et al., 2010) and *Otomys* (Edwards et al., 2011). The Pleistocene divergence age for extant lineages contradicts the Early to Middle Miocene age (19 Ma in *Heliophobius* and 16 Ma in *Georychus*) estimated by Visser et al. (2020). These significantly older estimates were obtained due to exclusive reliance on the fossil genus *Proheliophobius* (20.0 Ma) as an ancestor of extant *Heliophobius*, the fossil which is often used for dating of mole-rat divergences (e.g. Ingram et al., 2004; Faulkes et al., 2010). Paleontological literature, however, does not suggest such an assignment (Winkler et al., 2010, Lavocat, 1973) and also visual comparison of the skull and mandible of *Proheliophobius* (Lavocat, 1973, Fig. 12) and *Heliophobius* (Gomes Rodrigues et al., 2016, Fig. 1) makes the assumption of a close relationship less convincing. It is also doubtful that bathyergid genetic structure has remained stable for approx. 15 Ma, while other savannah living rodents underwent massive radiations about 10 Ma later, during climatic oscillations in the Plio-Pleistocene. As demonstrated by the present divergence dating, *Proheliophobius* most likely constitutes an extinct genus of incertae sedis among Bathyergidae.

The basal divergence in *Heliophobius* was 2.3 Ma old according to our analysis, a slightly younger date than that obtained by Bryja et al. (2018) also using the fossilized birth–death model. The difference can be due to the use of a more comprehensive set of fossils (42 instead of 14), including multiple species of the same fossil genus or alternatively could arise from the use of a strict clock and exclusively nuclear loci. If there are just few nuclear loci and the *CYTB* substitution rate is much higher, it can have a disproportionate overestimating effect on divergence times.

Beyond the issue of intrageneric divergence, our study also provides a new dating of divergences between bathyergid genera. Our estimates of divergence dates broadly agree with the estimates obtained in earlier ctenohystrican phylogenetic studies, which calibrated their reconstructions with multiple (non-bathyergid) fossil constraints. Although being based on node calibrations, not fossilized birth–death model, they place the basal split of the extant bathyergids in the Oligocene: 26.0 Ma (Fang et al., 2014), 26.2 Ma (Sallam et al., 2009) and 31.2 Ma (Upham and Patterson, 2015). This is also in accord with the fossil record (Mein and Pickford, 2008, Winkler et al., 2010, Sallam and Seiffert, 2020) and the record of Oligocene global cooling and retreat of forests (Anka and Séranne, 2004, Couvreur et al., 2021). This climate change promoted the evolution of plants with subterranean storage organs (geophytes) and this, in turn, created niches for subterranean rodents, which depend on geophytes as their staple food (Nevo, 1979). On the other hand, it should be noted that our estimate of 29.2 Ma is largely based on our calibration priors and can hardly be made more precise without additional Oligocene fossils and/or improved understanding of the fossilized traits and their evolution.

4.5. The causes of intraspecific diversifications of *Heliophobius* and *Georychus*

Similar estimates (around 2.2 Ma) for the beginning of diversifications in extant lineages of both genera suggest a single large-scale factor playing a central role. The most obvious candidate is a change or repeated changes in global climate. The Quaternary (since 2.58 Ma) is characteristic of climate oscillations causing profound reorganization of ecosystems and leaving different kinds of recognizable traces. Based on a combination of ancient levels in EARS lakes (Trauth et al., 2007) and dust flows apparent in the adjacent marine sediments (deMenocal, 2004) it was concluded that three periods, 2.79–2.47, 1.89–1.69, and 1.12–0.92 Ma, were characterized by substantial climate variability, e.g. periods of frequent changes from humid to dry climates (Trauth et al., 2009, Potts, 2013). We assume that the climate fluctuations were the primary driver for the splitting of both genera; however, due to the credibility intervals being as long as 1.65 Ma, the diversifications could not be matched with precise climatic events.

In the case of *Heliophobius*, an inhabitant of the EARS, the ongoing formation of the rift system could also separate the ancestral populations. Although the main tectonic activity creating EARS had concluded by the Late Miocene, there are observations of peak rifting, subsidence and sedimentations in several basins during the Pleistocene (around 1–2 Ma; MacGregor, 2015), which could possibly affect ancestral populations of silvery mole-rat. However, in the case of *Georychus,* occuring in South Africa, there is no evidence for comparable topographical changes at that time that could represent the primary driver for its intrageneric splitting (Baby et al., 2020).

The *Heliophobius* distribution is characterized generally by mesic conditions with > 900 mm annual rainfall (Šumbera et al., 2007) but also precipitation seasonality (cf. Fig. 8), which

supports the occurence of geophytes. It inhabits mainly miombo woodlands and avoids humid forests where geophytes are absent. Therefore, the expansion of humid forests during the moister and warmer periods may be responsible for the persistent isolation of populations. The previously isolated populations can admix when climate becomes more mesic and seasonal, causing retreat of the forests, but the traces of admixture may be erased again in the subsequent episode of forest expansion. According to this model, the so-called East African "montane circle" may have represented such a barrier dividing ancestral *Heliophobius* population into the recent northern and southern lineages. This "montane circle" encompasses the Southern and Western Rifts of the EARS and the Eastern Arc Mountains (EAM) and hosts sites with the record of forest expansion in wet Pleistocene periods (Finch et al., 2014, Schüler et al., 2012). These repeated changes of habitats in diverse mountain ranges have been proposed generating the rich phylogeographic structure of many taxa, including bats (Fahr et al., 2002, Taylor et al., 2012), shrews (Demos et al., 2014, Stanley et al., 2015), and rodents (Taylor et al., 2009, Colangelo et al., 2013, Bryja et al., 2014, Krásová et al., 2019). Very similarly to the preceding scenario, the forests along the Southern Rift and higher level of Lake Malawi (Trauth et al., 2007) might have acted as a dispersal barrier for the populations of southeastern and southwestern lineages.

There is just a single fossil record of *Heliophobius* distribution, but it comes from the northern margin of the present distribution, namely from Isinya in southern Kenya (Brugal and Denys, 1989), a site dated to about 1 Ma (Durkee and Brown, 2014). The fossil demonstrates the ancient presence of silvery mole-rat in the region now occupied by a distinctive N1 lineage. Overall, our modelling of *Heliophobius* distribution during the LGM (∼21 thousand years ago) suggests its distribution remained remarkably stable throughout the major climatic oscillations. This implies that core populations within present day lineages have had a long history of persistence within the same region, not too distant from each other. Making this assumption, ecological or intrinsic barriers to gene flow have to be invoked to explain distinctiveness of the lineages (meaning MSC-species or Infomap clusters).

Similarly, the occurrence of *Georychus* depends on the availability of geophytes. Yet, except for the presence of vleis or rivers, no other habitat preferences are known, as environmental conditions vary remarkably across its distribution (Visser et al., 2017). The geographically widely distributed fossil records of *Georychus* (Pickford and Mein, 1988, Avery, 1991, Avery, 1998, Avery, 2001, Klein et al., 2007, Figure S.10) contrast with the current restricted and isolated populations and suggest that suitable environmental conditions were more widespread in the past. With respect to repeated changes of climate during Plio-Pleistocene, the distribution of *Georychus* probably underwent extensive retractions and expansions. Our MaxEnt model supports this view – the predicted present and LGM distributions are found in the same geographical domains and have comparably sharp boundaries, but the LGM distribution is more extensive. Especially in the Western Cape additional evidence suggests it served as a stable refugium. Firstly, the stable climate prevailed in the Cape region throughout the Quaternary (Cowling et al., 1997), which can explain an extremely high species diversity of plants in the region (Procheş et al., 2006). Secondly, at present, this is the place with the highest abundance of *Georychus* indicating habitat suitability for the species (Visser et al., 2017). Thirdly, local fynbos vegetation hosts the highest diversity and abundance of geophytes in the world (Cowling et al., 1997, Procheş et al., 2006), ensuring mole-rats sufficient food supply.

When it comes to the disjunct distribution of *Georychus*, either today or in the past, we hypothesize climatic fluctuations (Trauth et al., 2009; Potts et al., 2013; Quick et al., 2016),

to cause the retreat and consequently separation of *Georychus* populations during arid and cold periods. At the same time, however, many non-fossorial rodents, such as *Otomys irroratus*, *Myomyscus verreauxii*, *Rhabdomys pumilio*, *Gerbiliscus afra*, and *Acomys subspinosus*, are currently distributed along the whole southern and south-eastern coast, up to the grasslands of the eastern part of South Africa (Skinner and Chimimba, 2005, Monadjem et al., 2015). Therefore, we assume there is no unrecognized geomorphological barrier. The additional splitting of lineages could have been preserved due to the presence of the Breede River, separating the population in Oudshoorn, or by the Agulhas Plain, which was periodically flooded during the fluctuation of sea level in Pleistocene and isolated the population in Struisbaai (Visser et al., 2018).

Characteristics of the soil and food are the most important parameters influencing the occurence of subterranean mammals. Across its distribution, *Georychus* lives sympatricaly with several taxa of *Cryptomys* mole-rats (Monadjem et al., 2015). The nature of their coexistence can be illustrated in the Darling area, where three bathyergid species co-occur, but usually not in full syntopy. Whereas, solitary *B. suillus* is restricted to very loose sandy soils, *G. capensis* and the social *C. h. hottentotus* occur in more consolidated loamy and clayey soils. Here, the former occurs mainly in deep humid soils, whereas latter is located in drier stony substrates (Reichman and Jarvis 1989). There are also remarkable differences in food characteristics between both species. While *C. h. hottentotus* feeds almost exclusively on geophytes, the diet of *G. capensis* includes a high proportion of aboveground parts of plants (Robb et al., 2012). In the eastern (inland) part of its distribution, *Georychus* lives in higher rainfall areas with couch grass vegetation (J. Visser, unpublished data). On the basis of the limited, coarse data available, the relict populations of *Georychus* seem to be located mainly in sites with humid soils which at the same time show a lower availability of geophytes, both factors that would probably restrict colonization of such areas by *Cryptomys*. Here, *Georychus* may survive feeding mainly on aboveground parts of plants. It may be also possible that larger *Georychus* is competitively superior to *Cryptomys* in such areas, but unable to colonise drier sites, where social *Cryptomys* are more succesfull due to benefits of cooperative searching for food (Spinks and Plaganyi, 1999). To test the potential roles of habitat specialisation and/or competitive exclusion in comparison with other mole-rat species that co-occur with *G. capensis*, one requires detailed behavioural and ecological studies (knowledge of habitat requirements).

4.6. Taxonomic implications

Despite the power of molecular methods, well founded taxonomic proposals require integrative approach (Dayrat, 2005, Padial et al., 2010, Riedel et al., 2013), which puts genetic evidence into context of other data, including ecological requirements, behavioral traits, or reproductive isolation in contact zones. This kind of analysis together with the final proposal of the taxonomic status of various *Heliophobius* and *Georychus* lineages we defer to subsequent studies. These should demonstrate presence or absence of diagnostic morphological differences, ecological niche differentiation and reproductive isolation in zones of contact between putative species. Yet, based on the present genetic evidence, we propose the following taxonomic arrangement.

In the case of *Heliophobius*, we suggest two distinct species represented by the northern and southern lineages. Division of these two species is supported by mtDNA markers (Faulkes et al., 2011, Bryja et al., 2018), by MSC analysis of Sanger sequenced loci (Bryja et al., 2018; this study) and 100 ddRAD nuclear markers (this study) also by Infomap clustering algorithm

based on tens of thousands of ddRAD loci (this study). The karyotype from the Kenyan population is also slightly different from populations in Malawi and Zambia (2n = 60 and 62 respectively; Scharff et al., 2001, Šumbera et al., 2007, Deuve et al., 2008). The southern lineage should retain the name *H. argenteocinereus* (Peters, 1846) described from Tete (Mozambique, 16.14°S, 33.60°E), even though the locality likely represents a broader region of origin rather than actual sampling site of the type material. The oldest name available for the northern lineage is *H. kapiti* (Heller, 1909) from Potha, Kapiti Plains (Kenya, 1.57°S, 37.17°E). The even older name, *H. emini*, is not applicable for the northern lineage. It belongs to the taxon described from Simba Muënne (Noack, 1894), which according to Reichenow (1902), is synonymous with Simbamwenni. This is a locality at the outskirts of Morogoro (Tanzania, 6.80°S, 37.63°E), which is within the range of the southern lineage.

The split of both potential species is dated to around 2.31 Ma according to the fossilized-birth death analysis, and currently they are separated by the East African montane circle. Importantly, there is a potential contact zone between northern and southern lineages along the EAM. Without any obvious migration barrier, the admixture was suggested by the analysis of microsatellite data (Bryja et al., 2018) but it can be limited to a narrow hybrid zone (Barton and Hewitt, 1985, Macholán et al., 2007) and the apparent admixture pattern can also represent the ancestral polymorphism (cf. Lawson et al., 2018). A very similar situation with no real migration barrier, but the presence of distinct genetic clades was also found in the same region in co-distributed non-fossorial rodents, such as *Aethomys* (Mazoch et al., 2018), *Acomys* (Petružela et al., 2018), or *Lemniscomys* (Hánová et al., 2021).

The situation in the *Georychus* is less clear. Mitochondrial and nuclear gene sequences, allozymes, and restriction length polymorphisms suggest at least two distinct species from Western Cape and KwaZulu-Natal + Mpumalanga (Honeycutt et al., 1987, Nevo et al., 1987, Janecek et al., 1992, Faulkes et al., 2004, Visser et al., 2018, Visser et al., 2019, Visser et al., 2020; this study). The haplotype networks (this study) also show the evident separation of these two populations. Relatively deep historical splits around 2.20 Ma of these two potential species would also support their delimitation. On the other hand, the significant genetic similarities among the lineages are notable in a co-ancestry matrix based on ddRAD dataset, and as a result, the Infomap algorithm detected the sampled *Georychus* as one gene pool. Moreover, absence of any obvious geomorphological barrier between suggested species implies possible instability of the current genetic structure due to population expansion or retraction in the course of future climatic changes. Therefore, we propose keeping the current taxonomy with one species of this genus.

CRediT authorship contribution statement

M. Uhrová: Writing – original draft, Investigation, Formal analysis, Data curation, Visualization. **O. Mikula:** Writing – original draft, Formal analysis, Software, Data curation, Visualization. **N.C Bennett:** Writing – review & editing, Resources. **P. Van Daele:** Writing – review & editing, Resources. **L. Piálek:** Methodology, Investigation. **J. Bryja:** Writing – review & editing, Resources. **J.H. Visser:** Writing – review & editing, Resources. **B. Jansen van Vuuren:** Writing – review & editing, Resources. **R. Šumbera:** Conceptualization, Writing – original draft, Supervision, Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by the Czech Science Foundation project no. 31-20-10222S. We thank T. Aghová, W.N. Chitaukali, S. Gryseels, A. Hánová, A. S. Katakweba, A. Konečný, H. Konvičková, J. Krásová, M. Lövy, R. Makundi, J. Mbau, V. Mazoch, G. Mhamphi , J. Okrouhlík, C. Sabuni, F. Sedláček, J. Šklíba and R. Tsau for their help with field sampling and H. Burda and J. Zrzavý for useful comment on earlier version. We thank GRBC, the Technical Committee of the National Research Council of Malawi, Chancellor College, University of Malawi, the Kenyan Forest Service and the Kenyan Wildlife Service, Zambian Wildlife Authority, Sokoine University of Agriculture and Ministry of Tourism of Mozambique for permission to carry out our research. We are grateful to C. Mateke (Livingstone Museum) and W. Wendelen (Royal Museum for Central Africa) for access to the collection. All fieldwork complied with legal regulations in the respective countries. Access to the National Grid Infrastructure MetaCentrum provided under the programme CESNET LM2015042 is greatly appreciated and so is the access to CIPRES Science Gateway (https://www.phylo.org/). These facilities were used to run most of the phylogenetic analyses.

References

Aghov, 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. *J. Biogeogr*. 44, 2295–2307. https://doi.org/10.1111/jbi.2017.44.issue-1010.1111/jbi.13017.

Ait Belkacem, A., Gast, O., Stuckas, H., Canal, D., LoValvo, M., Giacalone, G., Päckert, M., 2016. North African hybrid sparrows (*Passer domesticus, P. hispaniolensis*) back from oblivion - ecological segregation and asymmetric mitochondrial introgression between parental species. *Ecol. Evol*. 6, 5190–5206. https://doi.org/10.1002/ece3.2274.

Allard, M.W., Honeycutt, R.L., 1992. Nucleotide sequence variation in the mitochondrial 12S rRNA gene and the phylogeny of African mole-rats (Rodentia: Bathyergidae). *Mol. Biol. Evol*. 9, 27–40. https://doi.org/10.1093/oxfordjournals.molbev.a040706.

Anka, Z., Séranne, M., 2004. Reconnaissance study of the ancient Zaire (Congo) deep-sea fan (ZaiAngo Project). *Mar. Geol*. 209, 223–244. https://doi.org/10.1016/j.margeo.2004.06.007.

Avery, D.M., 2001. The Plio-Pleistocene vegetation and climate of Sterkfontein and Swartkrans, South Africa, based on micromammals. *J. Hum. Evol*. 41, 113–132. https://doi.org/10.1006/jhev.2001.0483.

Avery, D.M., 1998. An assessment of the Lower Pleistocene micromammalian fauna from Swartkrans Members 1–3, Gauteng, South Africa. *Geobios* 31, 393–414. https://doi.org/10.1016/S0016-6995(98)80022-3.

Avery, D.M., 1991. Late Quaternary incidence of some micromammalian species in Natal. *Durban Mus. Novit*. 16, 1–11.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C., 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst*. 18, 489–522.

Baby, G., Guillocheau, F., Braun, J., Robin, C., Dall'Asta, M., 2020. Solid sedimentation rates history of the Southern African continental margins: implications for the uplift history of the South African Plateau. *Terra Nov*. 32, 53–65. https://doi.org/10.1111/ter.v32.110.1111/ter.12435.

Barbiére, F., Marivaux, L., 2015. Phylogeny and evolutionary history of hystricognathous rodents from the Old World during the Tertiary: new insights into the emergence of modern "phiomorph" families. In: Cox, P.G., Hautier, L. (Eds.), *Evolution of the Rodents – Advances in Phylogeny, Functional Morphology and Development.* University Press, Cambridge, pp. 87–138.

Barton, N.H., Hewitt, G.M., 1985. Analysis of hybrid zones. *Ann. Rev. Ecol. Syst.* 16, 113–148. https://doi.org/10.1146/annurev.es.16.110185.000553.

Bennett, N.C., Faulkes, C.G., 2000. *African Mole-rats: Ecology and Eusociality*. University Press, Cambridge.

Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* 2008, P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008.

Boratynski, Z., Brito, J.C., Mappes, T., 2012. The origin of two cryptic species of African desert jerboas (Dipodidae: *Jaculus). Biol. J. Linn. Soc*. 105, 435–445. https://doi.org/10.1111/j.1095-8312.2011.01791.x.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10 https://doi.org/10.1371/journal.pcbi.1003537 e1003537.

Bronner, G.N., 1990. New distribution records for four mammal species, with notes on their taxonomy and ecology. *Koedoe* 33, 1–7. https://doi.org/10.4102/koedoe.v33i2.435.

Brugal, J.P., Denys, C., 1989. Vertébrés du site acheuléen d'Isenya (Kenya, District de Kajiado) – implications paléoécologiques et paléobiogéographiques. *C. R. Acad. Sci. Biol. Paris* 308, 1503–1508.

Bryja, J., Konvičková, H., Bryjová, A., Mikula, O., Makundi, R., Chitaukali, W.N., Šumbera, R., 2018. Differentiation underground: range-wide multilocus genetic structure of the silvery mole-rat does not support current taxonomy based on mitochondrial sequences. *Mamm. Biol*. 93, 82–92. https://doi.org/10.1016/j.mambio.2018.08.006.

Bryja, J., Mikula, O., Šumbera, R., Meheretu, Y., Aghová, T., Lavrenchenko, L.A., Mazoch, V., Oguge, N., Mbau, J.S., Welegerima, K., Amundala, N., Colyn, M., Leirs, H., Verheyen, E., 2014. Pan-African phylogeny of Mus (subgenus *Nannomys*) reveals one of the most successful mammal radiations in Africa. *BMC Evol. Biol*. 14, 256. https://doi.org/10.1186/s12862-014-0256-2.

Burda, H., 2001. Determinants of the distribution and radiation of African mole-rats (Bathyergidae, Rodentia): ecology or geography? In: Denys, C., Granjon, L., Poulet, A. (Eds.), *African Small Mammals*. IRD Editions, Collection Colloques et Seminaries, Paris, pp. 263–277.

Burda, H., Zima, J., Scharff, A., Macholan, M., Kawalika, M., 1999. The karyotypes of *Cryptomys anselli* sp. nova and Cryptomys kafuensis sp. nova: new species of the common mole-rat from Zambia (Rodentia, Bathyergidae). *Z. Säugetierk.* 64, 36–50.

Burnham, K.P., Anderson, D.R., 2002. *Model Selection and Multimodel Inference: a Practical Information-theoretical Approach*, 2nd. ed. Springer, New York.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol*. 22, 3124–3140.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9, 1657–1659. https://doi.org/10.1046/j.1365-294X.2000.01020.x.

Colangelo, P., Verheyen, E., Leirs, H., Tatard, C., Denys, C., Dobigny, G., Duplantier, J. M., Brouat, C., Granjon, L., Lecompte, E., 2013. A mitochondrial phylogeographic scenario for the most widespread African rodent, *Mastomys natalensis*. *Biol. J. Linn. Soc.* 108, 901–916. https://doi.org/10.1111/bij.2013.108.issue-410.1111/bij.12013.

Couvreur, T.L.P., Dauby, G., Blach-Overgaard, A., Deblauwe, V., Dessein, S., Droissart, V., Hardy, O.J., Harris, D.J., Janssens, S.B., Ley, A.C., Mackinder, B.A., Sonké, B., Sosef, M.S.M., Stévart, T., Svenning, J.C., Wieringa, J.J., Faye, A., Missoup, A.D., Tolley, K.A., Nicolas, V., Ntie, S., Fluteau, F., Robin, C., Guillocheau, F., Barboni, D., Sepulchre, P., 2021. Tectonics, climate and the diversification of the tropical African terrestrial flora and fauna. *Biol. Rev.* 96, 16–51. https://doi.org/10.1111/brv.v96.110.1111/brv.12644.

Cowling, R., Richardson, D., Mustart, P., 1997. Fynbos. In: Cowling, R., Richardson, D., Pierce, S. (Eds.), *Vegetation of Southern Africa*. University Press, Cambridge, pp. 99–130.

Davies, K.T.J., Bennett, N.C., Tsagkogeorga, G., Rossiter, S.J., Faulkes, C.G., 2015. Family wide molecular adaptations to underground life in African mole-rats revealed by phylogenomic analysis. *Mol. Biol. Evol*. 32, 3089–3107. https://doi.org/10.1093/molbev/msv175.

Dayrat, B., 2005. Towards integrative taxonomy. *Biol. J. Linn. Soc*. 85, 407–415. https://doi.org/10.1111/j.1095-8312.2005.00503.x.

Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol* 24, 332–340.

deMenocal, P.B., 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth Planet. Sci. Lett.* 220, 3–24. https://doi.org/10.1016/S0012-821X(04)00003-2.

Demos, T.C., Kerbis Peterhans, J.C., Agwanda, B., Hickerson, M.J., 2014. Uncovering cryptic diversity and refugial persistence among small mammal lineages across the Eastern Afromontane biodiversity hotspot. *Mol. Phylogenet. Evol*. 71, 41–54. https://doi.org/10.1016/j.ympev.2013.10.014.

Denys, C., 2011. Rodents. In: Harrison, T. (Ed.), *Paleontology and Geology of Laetoli: Human Evolution in Context. Volume 2.* Springer, New York, pp. 15–53.

Deuve, J.L., Bennett, N.C., Britton-Davidian, J., Robinson, T.J., 2008. Chromosomal phylogeny and evolution of the African mole-rats (Bathyergidae). *Chromosome Res.* 16, 57–74. https://doi.org/10.1007/s10577-007-1200-8.

Drummond, A.J., Bouckaert, R.R., 2015. *Bayesian evolutionary analysis with BEAST, Bayesian Evolutionary Analysis with BEAST*. Cambridge University Press. https://doi.org/10.1017/CBO9781139095112.

Durkee, H., Brown, F.H., 2014. Correlation of volcanic ash layers between the Early Pleistocene Acheulean sites of Isinya, Kariandusi, and Olorgesailie, Kenya. *J. Archeol. Sci.* 49, 510–517. https://doi.org/10.1016/j.jas.2014.06.006.

Edwards, S., Claude, J., Jansen van Vuuren, B., Matthee, C.A., 2011. Evolutionary history of the Karoo bush rat, *Myotomys unisulcatus* (Rodentia: Muridae): disconcordance between morphology and genetics. *Biol. J. Linn. Soc.* 102, 510–526. https://doi.org/10.1111/j.1095-8312.2010.01583.x.

Fahr, J., Vierhaus, H., Hutterer, R., Kock, D., 2002. A revision of the *Rhinolophus maclaudi* species group with the description of a new species from West Africa (Chiroptera: Rhinolophidae). *Myotis* 40, 95–126.

Fang, X., Seim, I., Huang, Z., Gerashchenko, M., Xiong, Z., Turanov, A., Zhu, Y., Lobanov, A., Fan, D., Yim, S., Yao, X., Ma, S., Yang, L., Lee, S.G., Kim, E., Bronson, R., ˇSumbera, R., Buffenstein, R., Zhou, X., Krogh, A., Park, T., Zhang, G., Wang, J., Gladyshev, V., 2014. Adaptations to a subterranean environment and longevity revealed by the analysis of mole rat genomes. *Cell Rep. 8*, 1354–1364. https://doi.org/10.1016/j.celrep.2014.07.030.

Faulkes, C.G., Bennett, N.C., Cotterill, F.P.D., Stanley, W.T., Mgode, G.F., Verheyen, E., Kitchener, A., 2011. Phylogeography and cryptic diversity of the solitary-dwelling silvery mole-rat, genus *Heliophobius* (family: Bathyergidae). *J. Zool.* 285, 324–338. https://doi.org/10.1111/jzo.2011.285.issue-410.1111/j.1469-7998.2011.00863.x.

Faulkes, C.G., Mgode, G.F., Archer, E.K., Bennett, N.C., 2017. Relic populations of *Fukomys* mole-rats in Tanzania: description of two new species *F. livingstoni* sp. nov. and *F. hanangensis* sp. nov. *PeerJ* 5, e3214. https://doi.org/10.7717/peerj.3214.

Faulkes, C., Moodle, G., Le Comber, S., Bennett, N., 2010. Cladogenesis and endemism in Tanzanian mole-rats, genus *Fukomys*: (Rodentia Bathyergidae): a role for tectonics? *Biol. J. Linn. Soc.* 100, 337–352. https://doi.org/10.1111/j.1095-8312.2010.01418.x.

Faulkes, C.G., Verheyen, E., Verheyen, W., Jarvis, J.U.M., Bennett, N.C., 2004. Phylogeographical patterns of genetic divergence and speciation in African mole-rats (Family: Bathyergidae*). Mol. Ecol*. 13, 613–629. https://doi.org/10.1046/j.1365-294X.2004.02099.x.

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas*. Int. J. Climatol*. 37, 4302–4315. https://doi.org/10.1002/joc.2017.37.issue-1210.1002/joc.5086.

Finch, J., Wooller, M., Marchant, R., 2014. Tracing long-term tropical montane ecosystem change in the Eastern Arc Mountains of Tanzania. *J. Quat. Sci.* 29, 269–278. https://doi.org/10.1002/jqs.2699.

Flouri, T., Jiao, X., Rannala, B., Yang, Z., 2018. Species tree inference with BPP using genomic sequences and the multispecies coalescent*. Mol. Biol. Evol*. 35, 2585–2593. https://doi.org/10.1093/molbev/msy147.

Furman, A., Çoraman, E., Çelik, Y.E., Postawa, T., Bachanek, J., Ruedi, M., 2014. Cytonuclear discordance and the species status of *Myotis myotis* and *Myotis blythii* (Chiroptera). *Zool. Scr*. 43, 549–561. https://doi.org/10.1111/zsc.12076.

Galan, M., Pagès, M., Cosson, J.F., Kolokotronis, S.O., 2012. Next-generation sequencing for rodent barcoding: species identification from fresh, degraded and environmental samples. *PLoS One* 7, e48374. https://doi.org/10.1371/journal.pone.0048374.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–472. https://doi.org/10.1214/ss/1177011136.

Gent, P.R., Danabasoglu, G., Donner, L.J., Holland, M.M., Hunke, E.C., Jayne, S.R., Lawrence, D.M., Neale, R.B., Rasch, P.J., Vertenstein, M., Worley, P.H., Yang, Z.L., Zhang, M., 2011. The community climate system model version 4. *J. Clim*. 24, 4973–4991. https://doi.org/10.1175/2011JCLI4083.1.

Gippoliti, S., Amori, G., 2011. A new species of mole-rat (Rodentia, Bathyergidae) from the Horn of Africa. *Zootaxa* 2918 (1), 39. https://doi.org/10.11646/zootaxa.2918.110.11646/zootaxa.2918.1.4.

Gomes Rodrigues, H., Šumbera, R., Hautier, L., 2016. Life in burrows channelled the morphological evolution of the skull in rodents: the case of African mole-rats (Bathyergidae, Rodentia). J. *Mammal. Evol*. 23, 175–189. https://doi.org/10.1007/s10914-015-9305-x.

Good, J.M., Vanderpool, D., Keeble, S., Bi, K., 2015. Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. *Evolution* 69, 1961–1972. https://doi.org/10.1111/evo.2015.69.issue-810.1111/evo.12712.

Gottscho, A.D., Wood, D.A., Vandergast, A.G., Lemos-Espinal, J., Gatesy, J., Reeder, T.W., 2017. Lineage diversification of fringe-toed lizards (Phrynosomatidae: *Uma notata* complex) in the Colorado Desert: delimiting species in the presence of gene flow. *Mol. Phylogenet. Evol.* 106, 103–117.

Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82, 711–732. https://doi.org/10.1093/biomet/82.4.711.

Hánová, A., Konečný, A., Nicolas, V., Denys, C., Granjon, L., Lavrenchenko, L.A., Šumbera, R., Mikula, O., Bryja, J., 2021. Multilocus phylogeny of African striped grass mice (*Lemniscomys*): stripe pattern only partly reflects evolutionary relationships. *Mol. Phylogenet. Evol*. 155, 107007. https://doi.org/10.1016/j.ympev.2020.107007.

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

Hasegawa, M., Kishino, H., Yano, T.A., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol*. 22, 160–174. https://doi.org/10.1007/BF02101694.

Heath, T.A., Huelsenbeck, J.P., Stadler, T., 2014. The fossilized birth-death process for coherent calibration of divergence-time estimates. *PNAS* 111, E2957–E2966. https://doi.org/10.1073/pnas.1319091111.

Hellenthal, G., Busby, G.B.J., Band, G., Wilson, J.F., Capelli, C., Falush, D., Myers, S., 2014. A genetic atlas of human admixture history. *Science* 343, 747–751. https://doi.org/10.1126/science:1243518.

Heller, E., 1909. A new rodent of the genus *Georychus. Smithsonian Misc. Coll*. 52, 469–470.

Honeycutt, R.L., Edwards, S.V., Nelson, K., Nevo, E., 1987. Mitochondrial DNA variation and the phylogeny of African mole-rats (Rodentia: Bathyergidae). *Syst. Zool.* 36, 280. https://doi.org/10.2307/2413067.

Horn, A., Basset, P., Yannic, G., Banaszek, A., Borodin, P.M., Bulatova, N.S., Jadwiszczak, K., Jones, R.M., Polyakov, A.V., Ratkiewicz, M., Searle, J.B., Shchipanov, N.A., Zima, J., Hausser, J., 2012. Chromosomal rearrangements do not seem to affect the gene flow in hybrid zones between karyotypic races of the common shrew, (*Sorex araneus). Evolution* 66, 882–889. https://doi.org/10.1111/j.1558-5646.2011.01478.x.

Hulva, P., Horáček, I., Strelkov, P.P., Benda, P., 2004. Molecular architecture of *Pipistrellus pipistrellus/Pipistrellus pygmaeus* complex (Chiroptera : Vespertilionidae): further cryptic species and Mediterranean origin of the divergence. *Mol. Phylogenet. Evol.* 32, 1023–1035. https://doi.org/10.1016/j.ympev.2004.04.007.

IUCN 2020. The IUCN Red List of Threatened Species. Version 2020-3. https://www.iucnredlist.org.

Ingram, C.M., Burda, H., Honeycutt, R.L., 2004. Molecular phylogenetics and taxonomy of the African mole-rats, genus Cryptomys and the new genus *Coetomys* Gray, 1864. *Mol Phylogenet Evol*. 31, 997–1014.

Jackson, N.D., Carstens, B.C., Morales, A.E., O'Meara, B.C., 2017. Species delimitation with gene flow. *Syst. Biol*. 66, 799–812. https://doi.org/10.1093/sysbio/syw117.

Janecek, L.L., Honeycutt, R.L., Rautenbach, I.L., Erasmus, B.H., Rei, S., Schlitter, D.A., 1992. Allozyme variation and systematics of African mole-rats (Rodentia: Bathyergidae). Biochem. *Syst. Ecol*. 20, 401–416. https://doi.org/10.1016/0305-1978(92)90081-N.

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., Flouri, T., 2017. Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics* 33, 1630–1638. https://doi.org/10.1093/bioinformatics/btx025.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–120. https://doi.org/10.1007/BF01731581.

Klein, R.G., Avery, G., Cruz-Uribe, K., Steele, T.E., 2007. The mammalian fauna associated with an archaic hominin skullcap and later Acheulean artifacts at Elandsfontein, Western Cape Province, South Africa. *J. Hum. Evol*. 52, 164–186. https://doi.org/10.1016/j.jhevol.2006.08.006.

Krásová, J., Mikula, O., Mazoch, V., Bryja, J., Říčan, O., Šumbera, R., 2019. Evolution of the Grey-bellied pygmy mouse group: highly structured molecular diversity with predictable geographic ranges but morphological crypsis. *Mol. Phylogenet. Evol.* 130, 143–155. https://doi.org/10.1016/j.ympev.2018.10.016.

Lawson, D.J., Hellenthal, G., Myers, S., Falush, D., Copenhaver, G.P., 2012. Inference of population structure using dense haplotype data. *PLoS Genet*. 8, e1002453. https://doi.org/10.1371/journal.pgen.1002453.

Lawson, D.J., Van Dorp, L., Falush, D., 2018. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat Comm.* 9, 3258. https://doi.org/10.1038/s41467-018-05257-7.

Leaché, A.D., Fujita, M.K., Minin, V.N., Bouckaert, R.R., 2014. Species delimitation using genome-wide SNP data. *Syst. Biol*. 63, 534–542. https://doi.org/10.1016/j.ympev.2016.09.008.

Leaché, A.D., Zhu, T., Rannala, B., Yang, Z., 2019. The spectre of too many species. *Syst. Biol*. https://doi.org/10.1093/sysbio/syy051.

Leigh, J.W., Bryant, D., 2015. Popart: full-feature software for haplotype network construction. *Methods Ecol. Evol*. 6, 1110–1116. https://doi.org/10.1111/2041-210X.12410.

López Antoňanzas, R., Sen, S., 2005. New species of *Paraphiomys* (Rodentia, Thryonomyidae) from the Lower Miocene of As-Sarrar, Saudi Arabia. *Palaeontology* 48, 223–233. https://doi.org/10.1111/j.1475-4983.2005.00445.x.

MacGregor, D., 2015. History of the development of the East African Rift System: a series of interpreted maps through time. *J. African Earth Sci.* 101, 232–252. https://doi.org/10.1016/j.jafrearsci.2014.09.016.

Macholán, M., Munclinger, P., Šugerková, M., Dufková, P., Bímová, B., Božíková, E., Zima, J., Piálek, J., 2007. Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution* 61, 746–771. https://doi.org/10.1111/evo.2007.61.issue-410.1111/j.1558-5646.2007.00065.x.

Malinsky, M., Trucchi, E., Lawson, D.J., Falush, D., 2018. RADpainter and fineRAD structure: population Inference from RADseq Data. *Mol. Biol. Evol.* 35, 1284–1290. https://doi.org/10.1093/molbev/msy023.

Maputla, N.W., Dempster, E.R., Raman, J., Ferguson, J.W.H., 2011. Strong hybrid viability between two widely divergent chromosomal forms of the pouched mouse. *J. Zool.* 285, 180–187. https://doi.org/10.1111/j.1469-7998.2011.00825.x.

Mazoch, V., Mikula, O., Bryja, J., Konvičková, H., Russo, I.R., Verheyen, E., Šumbera, R., 2018. Phylogeography of a widespread sub-Saharan murid rodent *Aethomys chrysophilus*: the role of geographic barriers and paleoclimate in the Zambezian bioregion. *Mammalia* 82, 373–387. https://doi.org/10.1515/mammalia-2017-0001.

McDonough, M.M., Šumbera, R., Mazoch, V., Ferguson, A.W., Phillips, C.D., Bryja, J., 2015. Multilocus phylogeography of a widespread savanna-woodland-adapted rodent reveals the influence of Pleistocene geomorphology and climate change in Africa's Zambezi region. *Mol. Ecol.* 24, 5248–5266. https://doi.org/10.1111/mec.13374.

Mein, P., Pickford, M., 2006. Late Miocene micromammals from the Lukeino formation (6.1 to 5.8 Ma). *Kenya. Bull. Mens. Soc. Linn. Lyon* 75, 183–223. https://doi.org/10.3406/linly.2006.13628.

Mein, P., Pickford, P., 2008. Early miocene rodentia from the northern Sperrgebiet, Namibia. *Mem. Geol. Surv.* 20, 235–290.

Mikula, O., Šumbera, R., Aghová, T., Mbau, J.S., Katakweba, A.S., Sabuni, C.A., Bryja, J., 2016. Evolutionary history and species diversity of African pouched mice (Rodentia: Nesomyidae: *Saccostomus*). *Zool. Scr.* 45, 595–617. https://doi.org/10.1111/zsc.12179.

Mikula, O., 2021. Climatic curves as predictors in MaxEnt niche modeling. bioRxiv 2021.11.08.467713. https://doi.org/10.1101/2021.11.08.467713.

Mikula, O., Falush D., in prep. Species delimitation in the co-ancestry matrix as exemplified on the Ethiopian brush furred mice (Lophuromys). bioRxiv.

Monadjem, A., Taylor, P.J., Denys, C., Cotterill, F.P.D., 2015. *Rodents of sub-Saharan Africa: a Biogeographic and Taxonomic Synthesis*. Walter de Gruyter GmbH, Berlin/Munich/Boston.

Nevo, E., 1979. Adaptive convergence and divergence of subterraneean mammals. *Ann. Rev. Ecol. Syst.* 10, 269–308.

Nevo, E., Ben-Shlomo, R., Beiles, A., Jarvis, J.U.M., Hickman, G.C., 1987. Allozyme differentiation and systematics of the endemic subterranean mole rats of South Africa. Biochem. Syst. Ecol. 15, 489–502. https://doi.org/10.1016/0305-1978(87)90066-4.

Nevo, E., Capanna, E., Corti, M., Jarvis, J.U.M., Hickman, G.C., 1986. Karyotype differentiation in the endemic subterranean mole-rats of South Africa (Rodentia, Bathyergidae). *Z. Säugetierk.* 51, 36–49.

Nichols, R., 2001. Gene trees and species trees are not the same. *Trends Ecol. Evol.* 16, 358–364. https://doi.org/10.1016/S0169-5347(01)02203-0.

Noack, T.H., 1894. *Heliophobius emini* n. sp. *Zool. Jahrb. Syst.* 7, 559–562.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., Kassem, K.R., 2001. Terrestrial ecoregions of the world: a new map of life on Earth: a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience* 51, 933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.

Padial, J.M., Miralles, A., De la Riva, I., Vences, M., 2010. The integrative future of taxonomy. *Front. Zool.* 7, 16. https://doi.org/10.1186/1742-9994-7-16.

Paradis, E., Schliep, K., 2019. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528. https://doi.org/10.1093/bioinformatics/bty633.

Patterson, B.D., Upham, N.S., 2014. A newly recognized family from the Horn of Africa, the Heterocephalidae (Rodentia: Ctenohystrica). *Zool. J. Linn. Soc.* 172, 942–963. https://doi.org/10.1111/zoj.12201.

Peters, W.H.C., 1846. *Bericht Verhandl.* K. Preuss. Akad. Wiss. Berlin 11, 259.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* 7. https://doi.org/10.1371/journal.pone.0037135e37135.

Petružela, J., Šumbera, R., Aghová, T., Bryjová, A., Katakweba, A.S., Sabuni, C.A., Chitaukali, W.N., Bryja, J., 2018. Spiny mice of the Zambezian bioregion – phylogeny, biogeography and ecological differentiation within the *Acomys spinosissimus* complex. *Mamm. Biol.* 91, 79–90. https://doi.org/10.1016/j.mambio.2018.03.012.

Phillips, S.B., Aneja, V.P., Kang, D., Arya, S.P., 2006. Modelling and analysis of the atmospheric nitrogen deposition in North Carolina. *Int. J. Global Environ.* 6, 231–252. https://doi.org/10.1016/j.ecolmodel.2005.03.026.

Phillips, S.J., Anderson, R.P., Dudík, M., Schapire, R.E., Blair, M.E., 2017. Opening the black box: an open-source release of Maxent. *Ecography* 40, 887–893. https://doi.org/10.1111/ecog.2017.v40.i710.1111/ecog.03049.

Piálek, L., Burress, E., Dragová, K., Almirón, A., Casciotta, J., Říčan, O., 2019. Phylogenomics of pike cichlids (Cichlidae: *Crenicichla*) of the *C. mandelburgeri* species complex: rapid ecological speciation in the Iguazú River and high endemism in the Middle Paraná basin. *Hydrobiologia* 832, 355–375. https://doi.org/10.1007/s10750-018-3733-6.

Pickford, M., Mein, P., 1988. The discovery of fossiliferous Plio-Pleistocene cave fillings in Ngamiland, Botswana. *C. R. Acad. Sci. Paris* 307, 1681–1686.

Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: convergence diagnosis and output analysis for MCMC. *R News* 6, 7–11.

Potts, R., 2013. Hominin evolution in settings of strong environmental variability. *Quat. Sci. Rev.* 73, 1–13. https://doi.org/10.1016/j.quascirev.2013.04.003.

Procheş, Ş., Cowling, R.M., Goldblatt, P., Manning, J.C., Snijman, D.A., 2006. An overview of the Cape geophytes. *Biol. J. Linn. Soc.* 87, 27–43. https://doi.org/10.1111/j.1095-8312.2006.00557.x.

Puttick, M.N., 2019. MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior ages on trees. *Bioinformatics* 35, 5321–5322. https://doi.org/10.1093/bioinformatics/btz554.

Quick, L.J., Meadows, M.E., Bateman, M.D., Kirsten, K.L., Mäusbacher, R., Haberzettl, T., Chase, B.M., 2016. Vegetation and climate dynamics during the last glacial period in the fynbos-afrotemperate forest ecotone, southern Cape*, South Africa. Quat. Int.* 404, 136–149. https://doi.org/10.1016/j.quaint.2015.08.027.

R Core Team, 2020. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rambau, R.V., Robinson, T.J., Stanyon, R., 2003. Molecular genetics of *Rhabdomys pumilio* subspecies boundaries: mtDNA phylogeography and karyotypic analysis by fluorescence in situ hybridization. *Mol. Phylogenet. Evol.* 28, 564–575. https://doi.org/10.1016/S1055-7903(03)00058-7.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904.

Rannala, B., Yang, Z., 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* 164, 1645–1656.

Reichenow, A., 1902. Die Vögel Afrikas, Atlas. Neumann, Neudamm.

Reichman, O.J., Jarvis, J.U.M., 1989. The influence of three sympatric species of fossorial mole-rats (Bathyergidae) on vegetation. *J. Mamm.*, 70, 763–771. https://doi.org/10.2307/1381710.

Riedel, A., Sagata, K., Suhardjono, Y.R., Tänzler, R., Balke, M., 2013. Integrative taxonomy on the fast track - towards more sustainability in biodiversity research. *Front. Zool.* 10, 15. https://doi.org/10.1186/1742-9994-10-15.

Robb, G.N., Woodborne, S., Bennett, N.C., Adler, F.R., 2012. Subterranean sympatry: An investigation into diet using stable isotope analysis. *PLoS One* 7, e48572. https://doi.org/10.1371/journal.pone.0048572.

Rodríguez-Prieto, A., Igea, J., Castresana, J., Brosius, J., 2014. Development of rapidly evolving intron markers to estimate multilocus species trees of rodents. *PLoS One* 9, e96032. https://doi.org/10.1371/journal.pone.0096032.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. Mrbayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. https://doi.org/10.1093/sysbio/sys029.

Rosvall, M., Bergstrom, C.T., 2008. Maps of random walks on complex networks reveal community structure. *PNAS* 105, 1118–1123.

Runck, A.M., Matocq, M.D., Cook, J.A., 2009. Historic hybridization and persistence of a novel mito-nuclear combination in red-backed voles (genus Myodes). *BMC Evol. Biol.* 9, 1–15. https://doi.org/10.1186/1471-2148-9-114.

Russo, I.R., Chimimba, C.T., Bloomer, P., 2010. Bioregion heterogeneity correlates with extensive mitochondrial DNA diversity in the Namaqua rock mouse, *Micaelamys namaquensis* (Rodentia: Muridae) from southern Africa - evidence for a species complex. *BMC Evol. Biol.* 10, 307. https://doi.org/10.1186/1471-2148-10-307.

Sallam, H.M., Seiffert, E.R., 2020. Revision of Oligocene '*Paraphiomys*' and an origin for crown Thryonomyoidea (Rodentia: Hystricognathi: Phiomorpha) near the Oligocene-Miocene boundary in Africa. *Zool. J. Linn. Soc.* 190, 352–371. https://doi.org/10.1093/zoolinnean/zlz148.

Sallam, H.M., Seiffert, E.R., Steiper, M.E., Simons, E.L., 2009. Fossil and molecular evidence constrain scenarios for the early evolutionary and biogeographic history of hystricognathous rodents. *PNAS* 106, 16722–16727. https://doi.org/10.1073/pnas.0908702106.

Scharff, A., Macholan, M., Zima, J., Burda, H., 2001. A new karyotype of *Heliophobius argenteocinereus* (Bathyergidae, Rodentia) from Zambia with field notes on the species. *Mamm. Biol.* 66, 376–378.

Schliep, K.P., 2011. Phangorn: phylogenetic analysis in R. *Bioinformatics* 27, 592–593. https://doi.org/10.1093/bioinformatics/btq706.

Schüler, L., Hemp, A., Zech, W., Behling, H., 2012. Vegetation, climate and fire-dynamics in East Africa inferred from the Maundi crater pollen record from Mt Kilimanjaro during the last glacial–interglacial cycle. *Quat. Sci. Rev*. 39, 1–13. https://doi.org/10.1016/j.quascirev.2012.02.003.

Skinner, J.D., Chimimba, C.T., 2005. *The Mammals of the Southern African Sub-region.* University Press, Cambridge. https://doi.org/10.1017/CBO9781107340992.

Spinks, A.C., Plaganyi, E.E., 1999. Reduced starvation risks and habitat constraints promote cooperation in the common mole-rat, *Cryptomys hottentotus hottentotus*: a computer-simulated foraging model. *Oikos* 85, 435–444. https://doi.org/10.2307/3546693.

Stanley, W.T., Hutterer, R., Giarla, T.C., Esselstyn, J.A., 2015. Phylogeny, phylogeography and geographical variation in the *Crocidura monax* (Soricidae) species complex from the montane islands of Tanzania, with descriptions of three new species. *Zool. J. Linn. Soc.* 174, 185–215. https://doi.org/10.1111/zoj.2015.174.issue-110.1111/zoj.12230.

Stein, B.R., 2000. Morphology of subterranean rodents. In: Lacey, E.A., Patton, J.L., Cameron, G.N. (Eds.), *Life Underground: The Biology of Subterranean Rodents*. The University of Chicago Press, Chicago, pp. 19–61.

Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet*. 68, 978–989.

Šumbera, R., Chitaukali, W.N., Burda, H., 2007. Biology of the silvery mole-rat (*Heliophobius argenteocinereus*). Why study a neglected subterranean rodent species? In: Begall, S., Burda, H., Schleich, C.E. (Eds.), *Subterranean Rodents*. Springer, Berlin, Heidelberg, pp. 221–236.

Taylor, P.J., Maree, S., Van Sandwyk, J., Kerbis Peterhans, J.C., Stanley, W.T., Verheyen, E., Kaliba, P., Verheyen, W., Kaleme, P., Bennett, N.C., 2009. Speciation mirrors geomorphology and palaeoclimatic history in African laminate-toothed rats (Muridae: Otomyini) of the *Otomys denti* and *Otomys lacustris* species-complexes in the 'Montane Circle' of East Africa. *Biol. J. Linn. Soc*. 96, 913–941. https://doi.org/10.1111/j.1095-8312.2008.01153.x.

Taylor, P.J., Stoffberg, S., Monadjem, A., Schoeman, M.C., Bayliss, J., Cotterill, F.P.D., Janke, A., 2012. Four new bat species (*Rhinolophus hildebrandtii* complex) reflect Plio-Pleistocene divergence of dwarfs and giants across an Afromontane Archipelago. *PLoS One* 7, e41744. https://doi.org/10.1371/journal.pone.0041744.

Templeton, A.R., 1992. Cladistic approaches to identifying determinants of variability in multifactorial phenotypes and the evolutionary significance of variation in the human genome. CIBA Found. Symp. 259–283 https://doi.org/10.1002/9780470514887.ch14.

Toews, D.P.L., Brelsvord, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* 21, 3907–3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x.

Trauth, Martin H., Larrasoa˜na, Juan C., Mudelsee, Manfred, 2009. Trends, rhythms and events in Plio-Pleistocene African climate. *Quat. Sci. Rev*. 28, 399–411. https://doi.org/10.1016/j.quascirev.2008.11.003.

Trauth, M.H., Maslin, M.A., Deino, A.L., Strecker, M.R., Bergner, A.G.N., Dühnforth, M., 2007. High- and low-latitude forcing of Plio-Pleistocene East African climate and human evolution. *J. Hum. Evol*. 53, 475–486. https://doi.org/10.1016/j.jhevol.2006.12.009.

Van Daele, P., Blonde, P., Stjernstedt, R., Adriaens, D., 2013. A new species of African mole-rat (*Fukomys*, Bathyergidae, Rodentia) from the Zaire-Zambezi watershed. *Zootaxa* 3636, 171–189. https://doi.org/10.11646/zootaxa.3636.1.7.

Visser, J.H., Bennett, N.C., Jansen van Vuuren, B., 2019. Evolutionary and ecological patterns within the South African Bathyergidae: implications for taxonomy. *Mol. Phylogenet. Evol*. 130, 181–197. https://doi.org/10.1016/j.ympev.2018.10.017.

Visser, J.H., Bennett, N.C., Jansen van Vuuren, B., 2020. Phylogeny and biogeography of the African Bathyergidae: a review of patterns and processes. *PeerJ* 7, e7730. https://doi.org/10.7717/peerj.7730.

Visser, J.H., Bennett, N.C., Jansen van Vuuren, B., Chiang, T., 2018. Spatial genetic diversity in the Cape mole-rat, *Georychus capensis*: extreme isolation of populations in a subterranean environment. *PLoS One* 13, e0194165. https://doi.org/10.1371/journal.pone.0194165.

Upham, N.S., Patterson, B.D., 2015. Evolution of the caviomorph rodents: a complete phylogeny and timetree of living genera. In: Vassallo, A.I., Antenucci, D. (Eds.), *Biology of Caviomorph Rodents: Diversity and Evolution*. SAREM, Buenos Aires, pp. 63–150.

Visser, J.H., Bennett, N.C., Jansen van Vuuren, B., 2017. Distributional range, ecology and mating system of the Cape mole-rat (*Georychus capensis*) family Bathyergidae. *Can. J. Zool*. 95, 713–726.

Warren, Dan L., Seifert, Stephanie N., 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol. Appl.* 21, 335–342. https://doi.org/10.1890/10-1171.1.

Wilson, D.E., Lacher, T.E.J., Mittermeier, R.A., 2016. *Handbook of the Mammals of the World* - Volume 6 Lagomorphs and Rodents I. Lynx Edicions, Barcelona.

Winkler, A., Denis, C., Avery, D., 2010. Fossil rodents of Africa. In: Werdelin, L., Sanders, W.J. (Eds.), *Cenozoic Mammals of Africa*. University of California Press, Berkeley, pp. 263–304.

Yang, Z., 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol*. 39, 306–314. https://doi.org/10.1007/BF00160154.

Yang, Z., Rannala, B., 2014. Unguided species delimitation using DNA sequence data from multiple loci. *Mol. Biol. Evol*. 31 (12), 3125–3135. https://doi.org/10.1093/molbev/msu279.

Zhang, C., Zhang, D.X., Zhu, T., Yang, Z., 2011. Evaluation of a Bayesian coalescent method of species delimitation. Syst. Biol. 60, 747–761. https://doi.org/10.1093/sysbio/syr071.

Lavocat, R., 1973. Les rongeurs du Miocène d'Afrique orientale (Unpublished doctoral dissertation). Universit´e de Montpellier, Montpellier.