Phylogeography and cryptic diversity of the solitary-dwelling silvery mole-rat, genus *Heliophobius* (Family: Bathyergidae)

C.G. Faulkes[1], N.C. Bennett[2,3,], F.P.D. Cotterill[4], W. Stanley[5], G. F. Mgode[6] & E. Verheyen[7]

[1]Queen Mary University of London, School of Biological & Chemical Sciences, Mile End Road, London E1 4NS UK

[2]Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Pretoria 0002, South Africa

[3]Department of Zoology, King Saud University, Riyadh, Saudi Arabia

[4]AEON, Department of Geological Sciences & Department of Molecular and Cell Biology, University of Cape Town, Rondebosch 7701, South Africa

[5]Department of Zoology, Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, IL 60605 USA

[6]Pest Management Centre, Sokoine University of Agriculture, PO Box 3110, Chuo Kikuu, Morogoro, Tanzania

[7]Royal Belgian Institute of Natural Sciences, Brussels, Belgium

Keywords: *Heliophobius*, silvery mole-rat, Bathyergidae, African mole-rats, phylogeography, taxonomy, Rift Valley, tectonics

Corresponding author: Chris G. Faulkes, Queen Mary University of London, School of Biological & Chemical Sciences, Mile End Road, London E1 4NS. e-mail c.g.faulkes@qmul.ac.uk

Running title: Phylogeography and cryptic diversity of the silvery mole-rat

Abstract

Alongside the eusocial naked mole-rat, *Heterocephalus glaber*, *Heliophobius argenteocinereus* is the second oldest lineage within the African mole-rat family Bathyergidae, and phylogenetically intermediate between the East African *H. glaber* and the South African genera *Bathyergus* and *Georychus*. Across its geographic range *H. argenteocinereus* is widely distributed on both sides of the East African Rift System (EARS), and is a key taxon for understanding the phylogeographic patterns of divergence of the family as a whole. Phylogenetic analysis of 62 mitochondrial cyt *b* sequences, representing 48 distinct haplotypes from 26 geographic locations across the range of *Heliophobius*, consistently and robustly resolved six genetically divergent clades that we recognize as distinct evolutionary species. Early species descriptions of *Heliophobius* were synonymised into a monotypic taxonomy that recognized only *H. argentocinereus*. These synonyms constitute available names for these rediscovered cryptic lineages, for which combined morphological and genetic evidence for topotypical populations endorses the recognition of six to eight distinct taxa. Bayesian estimates of divergence times using the fossil *Proheliophobius* as a calibration for the molecular clock suggest that the adaptive radiation of the genus began in the early Miocene, and that cladogenesis, represented in the six extant species, reflects a strident signature of tectonic activity that forged the principal graben in the EARS.

Introduction

African mole-rats of the genus *Heliophobius* are subterranean Hystricomorph rodents that constitute the most widespread and potentially biodiverse of the solitary dwelling taxa of the family Bathyergidae. As a family, African mole-rats are distributed throughout sub-Saharan Africa, and are renowned for their range of social and reproductive strategies that reach a pinnacle in the naked mole-rat, *Heterocephalus glaber* (for recent reviews see O'Riain & Faulkes, 2008; Faulkes & Bennett, 2007, 2009). Molecular phylogenies (based on both nuclear and mitochondrial genes) have consistently resolved six distinct clades, corresponding to the six genera currently recognised within the family, and these generic divisions are also supported by morphological characteristics (Allard & Honeycutt 1992; Faulkes *et al*., 1997a; Walton *et al*., 2000; Huchon & Douzery, 2001; Faulkes *et al*., 2004; Ingram *et al*., 2004).

In all molecular phylogenies the eusocial naked mole-rat lineage is basal, followed by lineages giving rise to three extant genera characterised by a strictly solitary lifestyle (*Heliophobius*, *Bathyergus*, and *Georychus*, respectively). In general, these solitary mole-rats are larger in body size and restricted primarily to mesic regions, where precipitation generally exceeds 400 mm per annum. Of these, *Heliophobius* has by far the widest distributional range, occurring in the sandy soils of savannas and woodlands of southern Kenya, throughout Tanzania, parts of eastern and the south-eastern Democratic Republic of Congo (DRC), through eastern Zambia and Malawi and into central Mozambique (Fig. 1). These areas are characterized by a high annual rainfall, which on average exceeds 900 mm (Bennett & Faulkes, 2000). In contrast, the social genera *Cryptomys* and *Fukomys* share a more recent common ancestor than the solitary-dwelling mole-rat genera, and the ranges of some of these social species, together with *Heterocephalus*, extend into areas where rainfall may be very low (sometimes less than 200 mm per annum), and also sporadic and unpredictable.

Adaptations to this arid environment and the resulting patterns of food distribution are thought to drive social evolution in these taxa (Jarvis & Bennett, 1990, 1991, 1993; Jarvis *et al*., 1994; Faulkes *et al*., 1997a; Bennett & Faulkes, 2000; Faulkes & Bennett, 2007, 2009; O'Riain & Faulkes, 2008, but see Burda *et al*., 2000 for counter arguments).

Phylogeographic reconstructions reveal a southward dispersal of ancestral forms from Central and East Africa (*Heterocephalus* and *Heliophobius*) into southern Africa (*Bathyergus* and *Georychus*) coupled with either one or more evolutionary losses of sociality or the retention of an ancestral solitary state. The group containing the reciprocally monophyletic clades of *Cryptomys* and *Fukomys* then appears to have either convergently evolved sociality, or possibly retained an ancestral social lifestlye. *Fukomys* particularly has undergone an extensive radiation in southern and Central/West Africa, to produce the most speciose genus within the family with the widest geographical distribution (Faulkes *et al*., 2004; Ingram *et al*., 2004; Van Daele et al., 2004; 2007a,b). Two reasons emphasise why *Heliophobius* is thus a key taxon for understanding the phylogeographic patterns that have underpinned the adaptive radiation of the family: (i) as living representatives of an early divergence phylogenetically intermediate between East African *Heterocephalus* and the South African genera *Bathyergus* and *Georychus*, (ii) the genus *Heliophobius* is widely distributed on both sides of a major geographic and ecological boundary, the East African Rift System (EARS), which forms part of the African Rift Valley. In contrast, the two other extant bathyergid genera distributed in the region of the EARS (approx. 8ºN to 17ºS) are confined either to the east (*Heterocephalus*) or to the west (*Fukomys*) of the EARS, although some extralimital and possibly disjunct populations of *Fukomys* are known to be extant in Tanzania (Faulkes *et al*., 2004, 2010).

To date, relatively little is known about intra-generic phylogenetic relationships of *Heliophobius*, because encompassing molecular genetic studies of the Bathyergidae have only

included a limited number of population samples (e.g. Faulkes *et al*., 2004; Ingram *et al*., 2004). However, in the early literature, a number of *Heliophobius* taxa were described, but their taxonomic status is subsequently unclear. Reviewing earlier literature, Ellerman (1940) listed a total of nine described taxa in *Heliophobius*. With the exception of only *H. spalax* (Thomas, 1910), these were all relegated to subspecies of *H. argenteocinereus* on criteria of pelage colour, qualitative differences in skull morphology, and body size, as follows: *H. a. argenteocinereus* Peters, 1852, *H. a. angonicus* Thomas, 1917, *H. a. robustus* Thomas, 1906, *H. a. marungensis* Noack, 1887, *H. a. emini* Noack, 1894, *H. a. kapiti* (Heller, 1909), *H. a. albifrons* (Gray, 1864) and *H. a. mottoulei* (Schouteden, 1913). Of these, the provenance of *H. albifrons* remains unclear, being vouched for by a specimen collected by Captain Speke in "East Africa", which according to Swynnerton (1945), can probably be restricted to the coast of Tanganyika Territory. More recently, Honeycutt *et al*. (1991) re-examined the holotype of *H. spalax* and concluded that the characters used to differentiate it from *H. argenteocinereus* can be attributed solely to age variation. In terms of chromosomal evolution and karyotypic variation, information is limited for *Heliophobius*. Scharff and coworkers (2001) reported minor differences in the karyotype of Zambian populations of *Heliophobius* (2n = 62) compared with the karyotype of Kenyan *Heliophobius* (2n = 60; George 1979), although a micrograph of the chromosomal spread was not published. While Ellerman (1940) and later Honeycutt *et al*. (1991) were thus in favour of taxonomic simplification, the recent availability of DNA sequence data argues against this, because it has revealed that some populations of *Heliophobius* are highly divergent (up to 17.8% sequence difference at the mitochondrial cytochrome *b* (cyt *b*) gene; Faulkes *et al*., 2004).

Using molecular genetic techniques, the aims of this study were (i) to reconstruct the phylogenetic relationships within *Heliophobius*; (ii) to resolve the taxonomic uncertainty represented in described taxa across the geographic range of the genus; (iii) account for the

phylogeographic origins and evaluate the role of tectonic activity (related to evolution of the African Rift Valley) in diversification of these poorly known bathyergids.

Methods

Sampling, PCR and sequencing

Samples for DNA analysis (total n=125) were either collected from locations across the range of *Heliophobius* between 1987 and 2008 (n=42), or obtained from museum specimens originally collected between 1912–1948 (The Royal Museum for Central Africa (RMCA), Tervuren, Belgium; n=18), 1964–1992 (Harrison Institute (HZM), Sevenoaks, Kent, UK; n=8), 1908–1997 (The Transvaal Museum (TM), Pretoria, South Africa; n=26) and 1953-1970 (Natural History Museum of Zimbabwe (NMZB), Bulawayo, Zimbabwe; n=10), 1905–2003 (Field Museum of Chicago (FMC), Chicago, USA; n=21). Full sample details are listed in Table 1 and their respective localities displayed in Fig. 1, together with the localities of published holotypes.

Tissue (muscle or skin biopsies from freshly collected animals) was fixed in 95% ethanol and then stored at –20ºC prior to DNA extraction. For the museum samples, small pieces of dried skin (approximately 5 mm square) were removed as unobtrusively as possible, e.g. from the abdominal area where the specimen was sutured, or from inside the mouth. Genomic DNA was extracted from the samples using a standard protocol (Faulkes *et al*., 1997a), and PCR amplification of the entire cyt *b* gene carried out using primers L14724 and H15915 (Irwin *et al*., 1991) and protocols previously described for African mole-rats by Faulkes *et al*., (1997b). These were used in conjunction with additional primers designed to amplify smaller fragments from within the cyt *b* gene for the museum samples, where DNA degradation had occurred: L14724 with MRH4HA (H15143: 5'-GAACGATATYTGYCCTCATGGWAG-3') and MRH3HA (H15165: 5'-

AATCACAGTAGCTCCTCAGA-3'); MRL4HA (L15163: 5'-CTWCCATGACARATATCGTTC-3') or MRL3HA (L15184: 5'-TGAGGAGCTACTGTGATT-3') with MRH5HA (H15497: 5'-GTGATGTTCTCAGGGTGTCC-3'); MRL5HA (L15516: 5'-GGACACCCTGACAACTACAC-3') with H15915 (in all cases numbers refer to the position of the 3' end of the primer relative to the human mitochondrial DNA map; Anderson *et al.,* 1981). A negative control was used in all PCR reactions to guard against contamination of museum samples. Sequencing was carried out in both directions using combinations of primers to obtain complementary partially or fully overlapping strands. An Amersham MegaBACE 1000 automated sequencer was used to separate fragments produced using the Dynamic ET Terminator Sequencing kit (Amersham, UK).

Ethical note

Sampling of wild populations focused where possible around agricultural areas, where mole-rats are considered pests. For Tanzanian animals, collection permits were issued by the Sokoine University of Agriculture and respective District Authorities. Export permits were obtained from the Wildlife Department (Ministry of Natural Resources and Tourism, Tanzania); and a Zoosanitary/Veterinary permit from the Ministry of Livestock Development and Fisheries. Upon capture, animals were temporarily stored in covered plastic containers with sand, nest material and food. On the evening of the day of capture, individuals were euthanased with an overdose of chloroform, and determination of sex and body weight, followed by tissue collection for DNA analysis was carried out *post mortem*.

Analysis of mitochondrial DNA sequences

Sequences were aligned manually using previous mole-rat cyt *b* alignments for reference (Faulkes *et al*., 2004) using MacClade version 3 (Madison & Madison, 1992). Phylogenetic relationships were investigated using standard parsimony (MP), maximum likelihood (ML) and Bayesian approaches. Firstly, Modeltest version 3.7 (Posada & Crandall, 1998) was used to establish the evolutionary model. Parameters were then used in subsequent model-based analyses. Maximum likelihood and parsimony analyses were undertaken with PAUP* version 4.0b8 (Swofford, 2001). For all the phylogenetic analyses of the full dataset, a single representative of each of the separate haplotypes obtained from the different geographical locations were included, together with the sequence for the naked mole-rat (*H. glaber*) as an outgroup (Faulkes *et al*., 1997a). Because the complete data set was large, for MP we used the heuristic search option both with all characters having an equal weighting, and with characters weighted *a posteriori* according to their rescaled consistency index. In addition, after initial exploration of the data, a reduced dataset (n=12) containing representatives of the main clades was analysed using an exhaustive search option (that considers every possible tree). Samples with the most complete sequence information were chosen for this analysis, which is only possible for up to 12 samples in PAUP*. Gaps were treated as missing data. Bootstrap analysis using these parsimony criteria was conducted with 100 replicates of the full dataset and 1000 replicates of the reduced dataset of 12 taxa. ML was conducted using the heuristic search option and the parameters obtained from Modeltest. Corrected and uncorrected measures of genetic distance between haplotypes were also determined with MEGA5 (Tamura & Nei, 1993; Tamura *et al*., 2007). Bayesian phylogenetic analysis was undertaken using BEAST version 1.6.1 (Drummond & Rambaut, 2007). The Hasegawa, Kishino and Yano model of DNA substitution with a gamma correction and considering the proportion of invariant sites (HKY+G+I; Hasegawa *et al*., 1985), with empirical base frequencies, a relaxed molecular clock (Drummond *et al*., 2006) and a Yule tree prior (the

most suitable for interspecies comparisons), was implemented. A total of twenty-three independent chains were run before a final run of 30,000,000 generations, sampling output every 3,000 generations, with priors for all parameters optimised. The first 3,000 trees (10%) were discarded during burn-in. Mixing and convergence of MCMC chains were evaluated using Tracer v1.5 (Drummond & Rambaut, 2007) to ensure sufficient iterations and sampling were performed before samples from the posterior distribution of trees were summarized using TreeAnnotator v1.6.1, and trees drawn using FigTree version 1.3.1 (Drummond & Rambaut, 2007).

Molecular clock calibration and estimation of divergence times

Evidence for rate heterogeneity along lineages was investigated using a likelihood ratio test with and without a molecular clock enforced during maximum likelihood analysis in PAUP* (Muse & Weir, 1992). A fossil-based calibration was then used to estimate the divergence time of lineages under a relaxed molecular clock model (Drummond *et al*., 2006). The minimum fossil age of 19–20 million years ago (Mya), was used to constrain the initial divergence of the *Heliophobius* lineage within the bathyergid family tree. This date is based on the occurrence of the early Miocene fossil *Proheliophobius*, thought to be ancestral to the genus *Heliophobius* (Lavocat, 1973). Divergence times among the lineages were estimated in the Bayesian phylogenetic analysis performed with BEAST (v1.6.1), by setting the time to most recent common ancestor (tmrca) of the ingroup (*Heliophobius*) as a prior with a lognormal distribution with the following settings: mean = 1.0, log (stdev) = 0.8, offset = 19.5 Mya. Initial runs using BEAST were used to estimate the approximate substitution rate over the tree, and the settings used for priors for the relevant parameters were as follows: ucld.stdev: mean = 1.0, log (stdev) = 0.8, offset = 0, initial value = 0.333; ucld.mean: mean =

0.008, log (stdev) = 0.8, offset = 0, initial value = 0.008; yule.birthRate: = mean = 1.0, log (stdev) = 0.8, offset = 0, initial value = 1.0.

Results

From the total of 125 individuals of *Heliophobius* from which DNA extraction was attempted, 62 amplified successfully. These represent 48 distinct haplotypes from 26 geographic locations, 42 of these sequences are previously unpublished, while the remaining 20 have variously been included in other phylogenetic studies of the Bathyergidae (Faulkes *et al*., 2004, 2010; Table 1, Fig. 1). Despite repeated attempts, many of the DNA extractions from museum skin snips failed to amplify the target cyt *b* locus. Those that were successful tended to be the more recently collected samples although three were 80 years old (range from 1931-1992), and the amount of sequence obtained ranged from 23–85% (263–969 bp) of the entire cyt *b* gene (1140 bp).

Maximum parsimony analysis of the 48 distinct *cyt b* haplotypes using the heuristic search option in PAUP* produced 28,662,607 rearrangements. From the total of 1140 characters, 668 were constant and 472 variable, of which 340 were parsimony informative and 132 uninformative. An initial consensus cladogram from the 1000 trees retained (according to the parameters set in PAUP*) had a tree length of 1013 steps. In a second round of parsimony analysis, the data were resampled by bootstrapping (with 100 replications) both with unweighted characters, and with 233 characters re-weighted *a posteriori* with their respective rescaled consistency index (907 characters retained a weighting of one). The bootstrap 50% Majority Rule consensus tree is displayed in Fig. 2a. Five main clades with clear geographic structure (labeled and colour coded in Figs 2, 3 and 4 and colour coded on Fig. 1) were recovered with bootstrap support of 95, 100, 100, 55 and 56% respectively, as follows: Clade 1 contains sequences from museum samples collected in Central Africa (DRC

and Rwanda) west of the Albertine Rift, designated as *H. robustus/mottoulei* (from Kisandje) and *H. mottoulei* (from Kiambe); Clade 2 constitutes sequences from samples east of the EARS, with a clear geographical separation of sub-clades 2a (Kenya) and 2b (Tanzania); Clade 3 also contains sequences from samples east of the Albertine Rift in Tanzania; Clade 4 comprises sequences from samples west of the Albertine Rift in Zambia and Malawi, and the single sample from Mozambique which is east of this southern section of the Albertine Rift; Finally, Clade 5 contains sequences from samples west of the Albertine Rift in Zambia including museum material designated as *H. marungensis*. With the exception of the single sample from Mount Namuli in Mozambique (3765), these clades show clear geographic structuring with respect to the EARS, with Clades 2 and 3 to the east, and Clades 1, 4 and 5 to the west. A reduced data set of 11 ingroup taxa representing each of these clades was selected with the outgroup for an exhaustive parsimony search. Fig. 3 shows the cladogram representing the single most parsimonious tree (length = 827 steps), together with bootstrap values following 100 replications of the branch and bound search option with characters both unweighted, and re-weighted *a posteriori* with their respective rescaled consistency indices. Weaker nodes (Clades 4 and 5 in Fig. 2) increased in bootstrap support following re-weighting, implying that any uncertainty was due to homoplasy in the data set, and all five clades were supported at 99-100% (Fig. 3). The topology of this single most parsimonious tree is identical to the parsimony tree containing the full dataset shown in Fig. 2a.

Following hierarchical likelihood ratio tests (hLRT) of different models of evolution using Modeltest, that found to best fit the cyt *b* data corresponded to the Tamura and Nei with a gamma correction that considers the proportion of invariant sites model of sequence evolution (TrN+G+I; Tamura & Nei, 1993). The parameters thus chosen for maximum likelihood analysis were unequal base frequencies (A, 0.3333; C, 0.2721; G, 0.1042; T, 0.2904), proportion of invariable sites (I) = 0.4679 and, for variable sites (G), a gamma

distribution shape parameter of 1.2305. Akaike Information Criterion (AIC) tests with Modeltest produced a slightly different result and found a transversion model with a gamma correction that considers the proportion of invariant sites (TVM+I+G) to be the best fit to the data, with unequal base frequencies (A, 0.3255; C, 0.2705; G, 0.1149; T, 0.2891), proportion of invariable sites (I) = 0.4586 and, for variable sites (G), a gamma distribution shape parameter of 1.1996. Using the heuristic search option in PAUP*, both models produced maximum likelihood trees with the same topology, the likelihood score of the two best trees retained from 66570 rearrangements implementing the AIC model was 6159.43. The topology of these two trees (not illustrated separately) was identical at all main nodes, differing only slightly in the placement of Nyika samples 3638 and 3602 relative to one another within their clade. Overall, the maximum likelihood (ML) phylogeny (mirrored in Fig. 2b) was very similar to the maximum parsimony (MP) analysis (Fig. 2a), recovering a monophyletic Clade 2 (with subclades 2a and 2b), but with Mbizi samples within subclade 2b in the ML tree. Clade 3 and two distinctive subclades were also recovered, with small differences only in the branching order at the terminal nodes. The most significant difference in the ML tree was that the monophyletic Clade 5 is now within a paraphyletic Clade 4, rather than a separate, more basal divergence. Clade 1 lineages were still basal within the ingroup taxa, but not monophyletic, with *H. robustus/mottoulei* (from Kisandji) diverging first, followed by a lineage with the two *H. mottoulei* sequences (from Kiambe), then two reciprocally monophyletic groups of Clade 2, and Clades 3, 4 and 5. Constraining the topology of the exhaustive parsimony tree shown in Fig. 3 to reflect the maximum likelihood Topology 2 (Fig. 4) resulted in an increase in tree length of just 8 steps (to 835).

The likelihood ratio test revealed a significant deviation from the null model (P = 0.000012) with (tree score = 5486.9) and without (tree score = 5437.8) a molecular clock

enforced during maximum likelihood analysis in PAUP*. Thus for the subsequent analysis of divergence times using BEAST, a relaxed molecular clock was assumed.

Bayesian phylogenetic analysis performed with BEAST also consistently resolved the five main clades described above, with an overall topology very similar to the ML analysis. Fig. 4 displays the tree with the greatest likelihood, having a mean likelihood score of -6236. All the nodes supporting the five principal clades had high Bayesian posterior probability values (BPP), and it is noteworthy that the lowest of these at Node D (84.3%) reflects the different topology of this part of the tree revealed in the MP analysis (Figs 2a and 3), where bootstrap support was also low (54% in MP analysis of the full dataset and 58% when characters were unweighted in the exhaustive MP tree).

In terms of the magnitude of genetic differences within and among clades, while within population mitochondrial cyt *b* sequence differences were generally low, between geographic areas genetic divergence was high, with uncorrected p-distances of up to 17.7%, rising to 25.1% following Tamura–Nei + gamma correction (TN+G; Table 2) between Clades 1 and 2. Within Clade 2, the higher mean within clade differences were due to subclades 2a and 2b that distinguish the Kenyan and Tanzanian populations. Between these the mean ± standard error of the mean (SEM) percentage difference was 10.8 ± 0.8 for the uncorrected p-distance, rising to 13.2 ± 1.2% for the TN+G corrected genetic distance.

Molecular clock estimates of divergence times over the Bayesian tree, together with their upper and lower bound 95% highest posterior density (HPD) intervals are summarised in Fig. 4. Divergence of the ingroup clade (Node A) is constrained to the early Miocene (19–20 Mya) by the *Proheliophobius* fossil-based calibration set as the prior, and corresponds to a period when formation of the Kenya Rift was initiated (Ebinger 1993). Extant populations in DRC/Rwanda, which represent the ambiguously labelled *H. robustus/mottoulei* from Kisandji

and *H. mottoulei*, represent a basal lineage derived from the common ancestor of the genus at Node A.

Divergences at Nodes B and D in Figure 4 coincide with the commencement and progressive formation of the Western Rift, and are dated at 15.1 and 11.6 Mya respectively. It is the placement of Clade 3 with respect to Nodes B and D that gives rise to the important difference between the MP topology (Figs 2a and 3) and ML/Bayesian topology (Figs 2b and 4), i.e. the relative timings and patterns of Neogene vicariance and cladogenesis due to rifting, represented in the Usango trough and Kilombero graben of central Tanzania. The nodes B and D have overlapping 95% HPD intervals (lower Node B 12.2, upper Node D 15.0 Mya), and the possible range of divergence times are well within the period of increasing tectonic activity in the Western Rift. Within Clade 3 there was a clear separation between the more northern localities around the Ukaguru Mountains and Liwale in the south. The latter population was characterised by two distinct clades, which have populated north-east Tanzania and appear to have recently become sympatric, although within the region there was no pattern with respect to their colony locations.

Node C also represents an important geographical divergence between Clades 2a and 2b (samples from Kenya and Tanzania respectively) and dated at 10.6 Mya (9.0–13.2). The extant populations that form these two subclades are separated by the northern edge of the Eastern Arc Mountains (Taita Hills, Pare and Usambara Mountains). The geographically distinct Mbizi population – isolated on the Ufipa Plateau – is also divergent within Clade 2b, separating from the Usambara populations at 9.6 Mya.

Node E defines a well-supported division between Clades 4 and 5. Within Clade 4, the single sample sequenced from Mount Namuli in Mozambique is anomalous, being the only lineage clearly situated east of the Rift Valley. The divergence at Node F between the single Mount Namuli sequence and the two from Rumphi in Malawi is dated at 2.1 Mya, with

uncorrected p-distances comparatively low at 4.5 and 4.8%. The other samples in Clade 4 were variously distributed in eastern Zambia, including the Fort Jameson sample from the Harrison Institute which was labelled as *H. argenteocinereus angonicus*. Clade 5 includes all the RMCA samples designated as *H. marungensis*, together with a Transvaal Museum sample collected from the Nyika Plateau.

Discussion

This study represents the most extensive molecular phylogenetic reconstruction to date for *Heliophobius*, and provides strong evidence for deep divergences within the genus. This cladogenesis is apparently driven by tectonic activity that formed the rugged relief of the African Rift Valley. Moreover, these results support the validity of early taxonomic nomenclature (e.g. Ellerman, 1940). The main clades resolved by the different phylogenetic analyses were strongly supported by the congruence in branching order and topology recovered by the different methods, although some of the deeper nodes were not equally well supported by all methods of analyses employed (e.g. Clade 1 in Fig. 2b, and the position of Clade 3 in Figs 2a and 3 versus 2b and 4). The inclusion of museum skin specimens greatly extended the sampling range beyond more recently collected and better-preserved material. Despite the difficulties in extraction and the shorter DNA sequences obtained from dried skin snips, these genotypes reveal a robust phylogenetic signal with well supported clades (e.g. Clades 1, 4 and 5; Figs 2, 3 and 4).

Phylogeographic Events

A pervasive difficulty that undermines reconstructions of phylogeographic history from DNA sequence data is the reliability of molecular clock calibration (Kumar, 2005). In the case of the Bathyergidae, previous studies have employed two main calibrations. Huchon and

Douzery (2001) used the occurrence of the first caviomorph fossil in the Tinguiririan of South America dated at 31-37 Mya (Wyss *et al*., 1993) to estimate divergence times of all the Hystricognath families. This gave a value of 40 or 48 Mya for the basal node of the bathyergid phylogeny. A local-clock calibration was used by Ingram *et al*. (2004) and Van Daele *et al*. (2007b) based on a minimum timing of 19–20 Mya for the divergence of the *Heliophobius* clade within the family, derived from the occurrence of the fossil *Proheliophobius leakeyi* (Lavocat, 1973). This fossil is interpreted as a possible close outgroup to extant bathyergids, and a number of specimens were described by Lavocat from East Africa (including Rusinga, Songhor and Napak) all within the distributional range of extant *Heliophobius*. The similarity of the palate in *Proheliophobius* to that of extant *Heliophobius* (broad and scarcely extending behind the third molar M3) prompted the naming of the fossil genus. Furthermore, Early Miocene *Heterocephalus* fossils have also been recorded in deposits at Napak in Uganda (Bishop 1962; Bishop *et al*., 1969), suggesting that *Heterocephalus* and *Proheliophobius* may have coexisted temporally (and questioning the latter as an earlier outgroup of the former). Thus, there is a strong case for using the age of *Proheliophobius* to constrain the first appearance of *Heliophobius*.

The East African Rift System (EARS) that divides extant populations of *Heliophobius* is comprised of two major lineaments in East Africa, with a third south-eastern branch in the Mozambique Channel (see Chorowicz, 2005 for recent review). The eastern branch of the EARS (Kenyan or Gregory Rift) ramifies from the Afar triangle in the north, through the main Ethiopian Rift, across the central Kenya Rift to culminate in the North-Tanzanian divergence (the latter includes Mt. Kilimanjaro, the Taita Hills, and the Usambara and Pare mountains). The initiation of this eastern branch of the EARS is estimated at 23 Mya with volcanism continuing to 11 Mya, and was followed by a period of shoulder uplift and graben basin development which continued until 5.3 Mya. Tectonic activity continues today in parts

of the Kenyan Rift (Smith & Mosely, 1993). The western branch of the EARS (Albertine Rift) contains the rift valley lakes and runs from Lake Albert (also known as Lake Mobutu) in the north, to Lake Malawi (also known as Lake Nyasa) in the south, encompassing lakes Edward, Kivu, Tanganyika and Rukwa. Formation of the Albertine Rift started at approximately 12 Mya in the northern section, to 7 Mya in the south (Ebinger, 1989; Ebinger *et al*., 1993) with rapid, ongoing uplift of the western flank (Ruwenzori Mountains) since the Pliocene (Bauer *et al*., 2010). It is estimated that Lakes Tanganyika and Nyasa attained deep-water conditions at 6–12 Mya (Cohen *et al*., 1993) and 4.5 Mya (Delvaux, 1995; Delvaux *et al*., 1998), respectively.

The Bayesian estimates of divergence times of the major clades resolved in our phylogeny, provide strong evidence that tectonic activity associated with the development of the EARS has played a significant role in cladogenesis in *Heliophobius*. The populations in DRC (blue circles on Fig. 1) consistently form a distinct basal clade (Clade 1, Figs 2, 3 and 4) isolated in Katanga, west of the Albertine Rift in the southern Congo basin (Fig. 1). At the time of the proposed initial radiation of the *Heliophobius* clade (19–20 Mya), volcanism was underway along the Kenyan Rift, but significant rift formation did not apparently commence along the Kenyan Rift, or in the Western Rift, until 11–12 Mya, with major tectonic activity continuing to 5–7 Mya. This coincides with the initial divergence of Clade 2 at Node B and the separation of Clade 3 from Clade 4 and 5 at Node D in Fig. 4, dated at approximately 15.1 and 11.6 Mya respectively (with 95% HPD covering a period from 18.4 to 7.7 Mya). All populations within Clades 2 and 3 occur east of the Western Rift, and the split at Node C between sub-clades 2a (east of the main Kenya Rift) and 2b (mainly eastern Tanzania) also occurred within this time frame at around 10.6 Mya. As with Clade 1, Clades 4 and 5 show a clear geographical pattern and include populations west of the Western Rift, situated in DRC, Rwanda, Zambia and Malawi. Within Clade 5 there is one exception to this geographical

trend, the single sample collected from Mount Namuli in Mozambique (south-east of Lake Nyasa). The divergence at Node F (2.1 Mya with 95% HPD of 1.0 to 3.1 Mya) between the Mount Namuli lineage and the sister group from Rumphi in Malawi potentially just pre-dates the period of increased tectonic activity at the Mbeya Triple Junction between Lake Rukwa and Lake Nyasa which commenced from 2.5 Myr. This region has previously been implicated as a possible conduit for east-west dispersal of mole-rats of the genus *Fukomys* (Faulkes *et al*. 2010). It is also interesting that the range of this relatively young Clade 4 of *Heliophobius* exhibits close spatio-temporal congruence with early Pleistocene isolation of Yellow baboons, *Papio cynocephalus* in eastern Zambia (Sithaldeen *et al*. 2009; Keller et al. 2010). Molecular dating estimates of vicariance in *Heliophobius* and *Papio* provides new temporal biogeographical context on the recognized role of the Luangwa valley as a refugium for endemic mammals (Ansell 1978; Cotterill 2003).

An earlier study further supports the estimated timings of divergences of Clades B, C and D. Faulkes *et al*. (2004) employed the Huchon and Douzery (2001) molecular clock calibration value of 40 or 48 Mya for the basal node of the bathyergid phylogeny, which overlap with the estimates obtained here, Node B: 10.8–13.0 versus 12.2–18.4, Node C: 8.1–9.7 versus 9.0–13.2 and Node D: 6.3–7.6 versus 7.7–15.0 (upper and lower estimates from Faulkes *et al*., 2004 versus upper and lower 95% HPD from this study). Thus, the molecular and geologically-based chronology is consistent with the scenario of a widespread distribution of ancestral lineages prior to Miocene rifting, and vicariance of these populations was then caused by increasing tectonic activity in the EARS. The restriction of the ancestral clade (designated the *H. mottoulie* complex) in Katanga is interesting. Their localized range coincides with the Upemba (=Kamalondo) graben, in a region of high floristic diversity and endemism (Malaisse 1997; Linder 2001) that supports a high endemism of aquatic and terrestrial vertebrates (Broadley & Cotterill 2004; Cotterill 2006).

Taxonomic implications

The clades resolved in this molecular phylogeny provide strong support for at least some of the nine *Heliophobius* holotypes described in the literature of the late nineteenth and early twentieth Centuries. However, in the absence of sequence information from original holotype specimens (a general problem due to age of samples and access), naming of particular phylogenetic clades must proceed with caution.

Although material was obtained from the holotype of *H. mottoulei* Schouteden, 1913 (RMCA-1222), collected in 1912 by Mottoule, extracted DNA failed to amplify under standard PCR protocols. However, two other samples collected and labelled as *H. mottoulei* in 1931 from Kiambe (DRC), were successfully sequenced, and together with a single sequence from a sample from Kisandje (DRC) with the skull labelled as *H. robustus* and the skin as *H. mottoulei* (RMCA 13487, form the basal Clade 1 (Figs 2, 3 and 4). It seems unlikely that *robustus* occurs in central Katanga, as the type locality is some distance southeast of Bangweulu (Fig. 1). In light of the type locality of *mottoulei* at Kilongwe near the Upemba Swamps, Lualaba river, these Kiambe specimens more likely represent an undescribed lineage – here designated *H. sp* "Kiambi" - isolated north of the Luvua river from populations in the Kamalondo (=Upemba) graben to the south. Besides its deep divergence (Late Miocene) from topotypical *H. mottoulei*, this hypothesis is underwritten by evidence for the marked vicariance of narrow range endemics of genus *Fukomys* across neighbouring northeast Zambia revealed by Van Daele *et al*., (2007a,b). These populations of mole-rats exhibit marked isolation in relation to trunk rivers and wetlands. Notably, the type locality of *H. robustus* is congruent with the East Bangweulu clade of the *F. whytei* complex (Van Daele *et al*., 2007a,b). Nevertheless, new specimens of *Heliophobius* from northeast Zambia are required to re-evaluate the status of *H. robustus* and test this hypothesis. This endorses

genotyping of new topotypical collections (Mpika plateau in the Chambeshi-Bangweulu catchment) of *H. robustus*. Bootstrap support of 100% and percentage sequence differences of 10.9 to 17.7 % (uncorrected p-distances; Table 2) with respect to the other clades provide strong support for considering Clade 1 as comprising at least two evolutionarily significant lineages. With regard to the taxonomic status, both *H. mottoulei* and *H. robustus* are large bodied species (adult female *H. mottoulei* head and body length = 250 mm; Fig. 5), but Schoutenden (1913) describes distinctive differences (presumably with reference to type of *robustus*, but this is not clear). These include the shape of the nasals, the zygomatic arch and coronoid process of the lower jaw. Differences in the morphology of the molars were also noted. Together with sequence differences of 8.7% (uncorrected p-distance) reported here, there may be a case for considering these as distinct taxa. Although the locality from which the Kisandjie sample sequenced here was collected is in close proximity to the type locality of *H. mottoulei*, the aforementioned confusion in labelling of the museum samples of the former endorses further sampling to resolve the taxonomy of these Katanga *Heliophobius*.

The reciprocally monophyletic groups (2a and 2b) within Clade 2 constitute two evolutionarily significant lineages. The Kenyan populations derive from the type locality of *H. kapiti* (Heller, 1909) and therefore almost certainly represent this species. They are sufficiently different both genetically (see above) and geographically from the sub-clade 2b, being divided by the northern arm of the Eastern Arc Mountains, for *H. kapiti* to be considered as a separate taxon from the populations in 2b. It is possible that sub-clade 2b represents *H. spalax* (Thomas, 1910), the type locality being Taveta, just north of the sub-clade 2b populations in the Usambara Mountains. Node C in Fig. 4 is constrained at around 10.6 Myr, significantly later than the formation of the Taita Hills and Usambara/Pare Mountains as a barrier (Bauernhofer *et al*., 2008). Ellerman (1940) also considered *H. spalax* to be a separate species to *H. argenteocinereus*, and Thomas (1910) distinguishes *H. spalax*

from *H. emini* (Noack, 1894), on the basis of the structure of the palate and other characteristic of the skull. Our results indicate that if sub-clade 2b is *H. spalax*, it may have made secondary contact and is now sympatric with Clade 3, which we putatively designate as *H. emini* on the basis of proximity of our sampling to the type locality (Fig. 1). The Bayesian phylogeny in Fig. 4 clearly suggests that Clades 2b and 3 have divergent histories, and samples karyotyped from Morogoro (Clade 3) have a different karyotype (2n = 62; Deuve *et al*., 2008), compared with sub-clade 2a (2n = 60; George, 1979). The validity of the other possible East African species, *H. albifrons* (Gray, 1864) remains questionable due to the uncertain provenance of the holotype and a lack of information (see above, as the type locality of the single specimen brought back to London by Captain Speke is simply cited as "East Africa" Swynnerton, 1940).

The two remaining groups clearly and robustly resolved in our phylogeny, Clades 4 and 5, should also be considered as evolutionarily significant lineages, although the mean uncorrected genetic difference between the two groups is on average lower than the among the other clades at 6.3 ± 0.7 % (Table 2). Within Clade 4, assessing the likely identity of the populations is problematic and the group potentially could include two taxa (*H. angonicus* and *H. argenteocinereus*), with the Fort Jameson sample (QM3601/HZM 1.10803) identified by the collector (W.F.H. Ansell) as *H. a. angonicus* Thomas, 1917. Thomas also designated material from Petuake (a locality we also sampled) as representing *angonicus*. The single sample from Mt. Namuli in Mozambique is of significance, being the only lineage within both Clades 4 and 5 that is located east of the southern Malawi branch of the EARS. The holotype that also represents the genotype - *H. argenteocinereus* Peters, 1852 - is from Tete on the Lower Zambezi River in Mozambique, a locality midway between Mt. Namuli and the type locality of *H. angonicus* and other Clade 4 populations in eastern Zambia (Fig. 1). Thus it is difficult to say whether the samples from Rumphi and/or Mt. Namuli correspond to *H.*

argenteocinereus and the other samples to *H. angonicus*. The fact that overall mean uncorrected p-distance within Clade 4 are low at 2.8 ± 0.4% might argue for sub-specific status i.e. *H. a. argenteocinereus* and *H. a. angonicus*. In terms of morphology, Thomas (1917) distinguishes *angonicus* from *argenteocinereus* on the grounds of greater development of the postorbital projections of the skull and the presence of a vertical ridge on the occipital plane (absent in all eight examples of *argenteocinereus* that were examined). A recent study by Barciova *et al*. (2009) investigated skull shape variation from populations in the Luano Valley, Zambia and the vicinity of Blantyre, Malawi (equivalent to the sampling range in Clade 4). Cranial shape variation was large and appeared to be influenced mainly by the type of habitat, and the latitudinal gradient, providing some evidence for local adaptation. In any case, fresh specimens for DNA analysis from Mozambique are a prerequisite to resolve these issues (PCR of museum samples was attempted but they failed to sequence). Scharff *et al*. (2001) reported a 2n = 62 karyotype of *Heliophobius* from the Luano Valley, near Petuake in Clade 4, corresponding to the Morogoro Clade 3 populations (Deauve *et al*., 2008), but differentiating these karyotypically from *H. kapiti* from Kenya.

Finally, five samples from the RMCA labelled by the collector as *H. marungensis* (Noack, 1887) collected in Tembwe (DRC) and Dubie (Rwanda) form Clade 5, together with a single sample obtained from the Transvaal Museum collected from the Nyika Plateau. Although these localities represent a wide geographical spread and potentially overlap the ranges of populations in Clades 1 and 4 (Fig. 1), the robustness of support for Clade 5 and the magnitude of genetic difference versus Clade 4 (mean uncorrected p-distance = 6.3 ± 0.7 %) argues for their specific status as *H. marungensis*. Further, Noack (1887) argues strongly for the distinctiveness of *marungensis* from *argenteocinereus* in terms of dental anatomy. He notes that *marungensis* has grooved incisors compared with smooth in *argenteocinereus*, and while the latter may have a remarkable molar dentition of 6/6 in full complement, the

marungensis specimens he described had only four molars in the upper jaw and up to five in the lower jaw, with no room for a sixth. While they overlap in geographical range, the taxa in Clades 4 and 5 are notably smaller in body size than those of Clade 1 (*mottoulei* and *robustus*), with head and body lengths of up to 182 mm (males) and 179 mm (females) in *H. argenteocinereus* (Roberts, 1913), 154 mm in *H. angonicus* (Thomas, 1917) and approx 130 mm in *H. marungensis* (Noak, 1887; Fig. 5).

Conclusions

This study has unambiguously identified six major evolutionary lineages of *Heliophobius*, whose deep temporal isolation and allopatry corroborate available species descriptions. As a result, we would suggest that there is a good case for recognising a *robustus*/*mottuolei* species complex (Clade 1), *H. kapiti* (Clade 2a), *H. spalax* (tentatively Clade 2b), *H. emini* (Clade 3), an *argenteocinereus*/*angonicus* complex (Clade 4), and *H. marungensis* (Clade 5). Molecular dating reveals that four of these clades have been allopatric for many millions of years, but interestingly, others appear to be potentially in sympatry due to secondary contact, e.g. Clades 2b and 3, or possibly diverging *in situ*, e.g. Clades 1, 4 and 5. At present it is unclear whether local adaptations have lead to ecological specialisations and reproductive isolation, or whether hybridisation is occurring between any of these lineages. Large body size differences between Clade 1 and Clades 4 and 5 may well preclude interbreeding. Further work utilising nuclear genetic markers including microsatellites would help to resolve this issue, and more sampling is especially needed in the DRC and Mozambique, together with a comprehensive analysis of skull shape and morphometrics. The conservation status of these populations of *Heliophobius* is unclear, although mole-rats are popular targets for bushmeat; this calls for reappraisal of the conservation status of these cryptic species. This evidence for the

previously overlooked cryptic genetic divergence of these interesting animals further endorses a reappraisal of the richness of mammal diversity of tropical East Africa.

Acknowledgements

Thanks to Tom Hutchinson and Eirini Daskalaki for various assistance in the lab, Laurence Gay and Heike Brinkman for translations of early literature from French and German respectively, Gus Mzumara for collection of the Rumphi samples and Julian Kerbis at the Field Museum in Chicago for the sample from Mt. Namuli, Mozambique. Samples and information were also kindly provided by Teresa Kearney at the Transvaal Museum in Pretoria, and David Harrison and Paul Bates at the Harrison Institute in Kent, UK. Michel Louette and Wim Wendelin at the Royal Museum for Central Africa in Tervuren were of great help in supplying samples and facilitating a visit by CGF in 2005 funded by a SYNTHESYS grant (BE-TAF-289). This work was also funded by grants from the National Research Foundation and University of Pretoria South Africa (to NCB) and the ERANDA and Bay Foundations (FC). Finally the authors are indebted to Dr Sarita Maree and an anonymous reviewer, whose comments greatly improved the final version of the manuscript.

References

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.

Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P. Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Standen, R. & Young, I.G. (1981). Sequence and organization of the human mitochondrial genome. *Nature (Lond.)* **290,** 457–465.

Ansell, W. F. H. (1978). The Mammals of Zambia. Chilanga, Zambia: National Parks and Wildlife Service.

Bauer, F.U., Glasmacher, U.A., Ring, U. Schumann, A. & Nagudi, B. (2010). Thermal and exhumation history of the central Rwenzori Mountains, Western Rift of the East African Rift System, Uganda. *Int. J. Earth Sci. (Geol Rundsch)* **99,** 1575–1597.

Bauernhofer, A.H., Hauzenberger, C.A., Wallbrecher, Hoinkes, E.G., Muhongo, S. & Mathu, E.M. (2008). Pan-African deformation in SE Kenya and NE Tanzania: geotectonic implications for the development of the north-central Mozambique belt. *Afr. J. Sci. Tech., Ser. B.* **9,** 50–71.

Barciova, L., Sumbera, R. & Burda, H. (2009). Variation in the digging apparatus of the subterranean silvery mole-rat, Heliophobius argenteocinereus (Rodentia, Bathyergidae): the role of ecology and geography. *Biol. J. Linn. Soc.* **97,** 822–831.

Bennett, N.C. & Faulkes, C.G. (2000). *African Mole-Rats: Ecology And Eusociality*. Cambridge: Cambridge University Press.

Bishop, W.W. (1962). The mammalian fauna and geomorphological relations of the Napak volcanics, Karamoja. *Uganda Geological Survey, Records 1957-1958*, pp 1–18.

Bishop, W.W., Miller, J.A. & Fitch, F.J. (1969). New potassium-argon age determinations relevant to the Miocene fossil mammal sequence in East Africa. *Am. J. Sci.* **267**, 669–699.

Broadley, D. G. & F. P. D. Cotterill (2004). The reptiles of southeast Katanga, an overlooked 'hot spot'. *Afr. J. Herp.* **53,** 35–61.

Burda, H., Honeycutt, R.H., Begall, S., Löcker-Grutjen, O. & Scharff, A. (2000). Are naked and common mole-rats eusocial and if so, why? *Behav. Ecol. Sociobiol.* **47**, 293–303.

Chorowicz, J. (2005). The East African rift system. *J. Afr. Earth Sci.* **43,** 379–410.

Cohen, A.S., Soreghan, M.J. & Scholz, C.A. (1993). Estimating the age of formation of lakes-an example from Lake Tanganyika, East-African Rift System. *Geology* **21,** 511–514.

Cotterill, F.P.D. (2003). Geomorphological influences on vicariant evolution of some African mammals in the Zambezi Basin: some lessons for conservation. In *Proceedings of the ecology and conservation of mini-antelope: an international symposium on duiker and dwarf antelope in Africa*: 11–58. Plowman, A. B. (Ed.). Fürth: Filander Verlag.

Cotterill, F.P.D. (2006). Taxonomic status and conservation importance of the avifauna of Katanga (southeast Congo Basin) and its environs. *The Ostrich* **77,** 1–21.

Delvaux D. (1998). Denudation history of the Malawi and Rukwa Rift flanks (East African Rift System) from apatite fission track thermochronology. *J. Afr. Earth Sci.* **26,** 363–365.

Delvaux D. (1995). Age of Lake Malawi (Nyasa) and water level fluctuations. Musée royal de l'Afrique Centrale (Tervuren). *Département de Géologie et Minéralogie rapport annuel.* 1993(1994), 99–108.

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.

Drummond, A.J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214.

Drummond, A.J., Ho, S.Y.W, Phillips, M.J. & Rambaut, A. (2006). Relaxed Phylogenetics and Dating with Confidence. *PLoS Biol.* **4,** e88.

Ebinger, C.J., Deino ,A.L., Tesha, A.L., Becker, T. & Ring, U. (1993). Tectonic Controls on Rift Basin Morphology: Evolution of the Northern Malawi (Nyasa) Rift. *J. Geophys. Res.* **17**, 821–836.

Ebinger, C.J. (1989) Tectonic development of the western branch of the east African rift system. *Geol. Soc. Am. Bull.* **1**, 885–903.

Ellerman, J.R. (1940). *The Families And Genera Of Living Rodents. Vol 1.* Trustees of the British Musuem (Natural History), London.

Faulkes, C.G., Mgode, G.F., Le Comber, S.C. & Bennett, N.C. (2010). Cladogenesis and endemism in Tanzanian mole-rats, genus *Fukomys* (Family: Bathyergidae): A role for tectonics? *Biol. J. Linn. Soc.* **100,** 337–352.

Faulkes, C.G. & Bennett, N.C. (2009). Reproductive skew in African mole-rats: Behavioural and physiological mechanisms to maintain high skew. In *Reproductive Skew in Vertebrates: Proximate and Ultimate Causes*: 369–396. Hager, R. & Jones, C.B. (Eds). Cambridge: Cambridge University Press.

Faulkes, C.G. & Bennett, N.C. (2007). African mole-rats: Behavioral and Ecological Diversity. In *Rodent Societies: An Ecological and Evolutionary Perspective*: 427–437. Wolff, J. & Sherman, P.W. (Eds). Chicago: The University of Chicago Press.

Faulkes, C.G. & Bennett, N.C. (2001). Family values: Group dynamics and social control of reproduction in African mole-rats. *Trends Ecol. Evol.* **16**, 184–190.

Faulkes, C.G., Bennett, N.C., Bruford, M.W., O'Brien, H.P., Aguilar, G.H., & Jarvis, J.U.M. (1997a). Ecological constraints drive social evolution in the African mole-rats. *Proc. Roy. Soc. Lond. Ser. B.* **264**, 1619–1627.

Faulkes, C.G., Abbott, D.H., O'Brien, H.P., Lau, L., Roy, M.R., Wayne, R.K. & Bruford, M.W. (1997b). Micro- and macro-geographic genetic structure of colonies of naked mole-rats, *Heterocephalus glaber*. *Mol. Ecol.* **6**, 615–628.

Filippucci, M.G., Burda, H., Nevo, E. & Kocka, J. (1994). Allozyme divergence and systematics of common mole-rats *(Cryptomys*, Bathyergidae, Rodentia) from Zambia. *Z. Säugetierkd* **59,** 42–51.

George, W. (1979). Conservatism in the karyotypes of two African mole rats (Rodentia, Bathyergidae). *Z. Säugetierkd* **44**, 278–285.

Gray, J.E. (1864). Notes on the species of sand-moles (*Georychus*). *Proc. Zool. Soc. Lond.* **9,** 123–125.

Hasegawa M., Kishino H. & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–74.

Heller, E. (1909). A new rodent of the genus *Georychus*. *Smithsonian Miscellaneous Collections* **52,** 269–270.

Honeycutt, R.L., Allard, M.W., Edwards, S.V. & Schlitter, D.A. (1991). Systematics and evolution of the family Bathyergidae. In *The Biology of the Naked Mole-Rat*: 45–65. Sherman, P.W., Jarvis, J.U.M. & Alexander, R.D. (Eds). Princeton: Princeton University Press.

Huchon, D. & Douzery, E.J.P. (2001). From the Old World to the New World: A molecular chronicle of the phylogeny and biogeography of hystricognath rodents. *Mol. Phylogenet. Evol.* **20**, 238–251.

Ingram, C., 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.

Irwin, D.M., Kocher, T.D. & Wilson, A.C. (1991). Evolution of the cytochrome-*b* gene of mammals. *J. Mol. Evol.* **32**, 128–144.

Jarvis, J.U.M. & Bennett, N.C. (1990). The evolutionary history, population biology and social structure of African mole-rats: Family Bathyergidae. In: *Evolution of Subterranean Mammals at the Organismal and Molecular Levels*: 97–128. Nevo, E. & Reig, O.A. (Eds). New York: Wiley Liss.

Jarvis, J.U.M. & Bennett, N.C. (1991). Ecology and Behavior of the Family Bathyergidae. In *The Biology of the Naked Mole-Rat*: 66–96. Sherman, P.W., Jarvis, J.U.M. & Alexander, R.D. (Eds). Princeton: Princeton University Press.

Jarvis, J.U.M., O'Riain, M.J., Bennett, N.C. & Sherman, P.W. (1994). Mammalian eusociality: a family affair. *Trends in Ecology and Evolution* **9,** 47–51.

Keller, C., Roos, C., Groeneveld, L.F., Fischer, J. & Zinner, D. (2010). Introgressive hybridization in southern African baboons shapes patterns of mtDNA variation *Am. J. Phys. Anthropol.* **142,** 125–136

Kumar, S. (2005). Molecular clocks: four decades of evolution. *Nat. Rev. Genet.* **6**, 654–62.

Lavocat, R. (1973). Les rongeurs du Miocene d'Afrique Orientale. I. Miocene Inferieur. *Memoires et Travaux de L'ephe, Institut de Montpellier*, **1**, 1–284.

Linder, H.P. (2001). Plant diversity and endemism in sub-Saharan tropical Africa. *J. Biogeogr.* **28,** 169–182.

Madison, W.P. & Madison, D.R. (1992). *MacClade: Analysis of Phylogeny and Character Evolution*, Version 3. Sunderland, MA: Sinauer Associates.

Malaisse, F. (1997). Se Nourrir en Forêt Claire Africaine: Approache écologique et nutrionnelle. Gembloux Presses: Agronomiques de Gembloux.

Muse, S.V. & Weir B.S. (1992). Testing for equality of evolutionary rates. *Genetics* **132,** 269–276.

Noack, T.H. (1887). *Heliophobius marungensis var.* von *argenteocinereus* N. *Zool. Jahrb. Syst.* **2,** 223–226.

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

O'Riain, M.J. & Faulkes, C.G. (2008). African mole-rats: Eusociality, relatedness and ecological constraints. In *Ecology of Social Evolution*: 205–220. Heinze, J. & J. Korb, J. Berlin: Springer-Verlag.

Peters, W.H.C. (1852). Naturwissenschafteliche Reise nach Mossambique. *Zoologie I. Saugetiere*. Georg Reimer, Berlin.

Posada, D. & Crandall, K.A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Roberts, A. (1913). The collection of mammals in the Transvaal Museum registered up to the 31st March, 1913, with descriptions of new species. *Ann. Trans. Mus.* **4,** 91.

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

Schoutenden, H. (1913). Un rat-taupe nouveau du Congo. *Rev. Zool. Afr.* **2,** 203.

Sithaldeen, R., Bishop, J.M. & Ackermann, R.R. (2009). Mitochondrial DNA analysis reveals Plio-Pleistocene diversification within the chacma baboon. *Mol. Phylogenet. Evol.* 53, 1042–8.

Smith, M. & Mosley, P., (1993). Crustal heterogeneity and basement influence on the development of the Kenya Rift, East Africa. *Tectonics* **12,** 591–606.

Swofford D.L. (2001). *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sunderland, MA: Sinauer Associates.

Swynnerton, G.H. (1945). A revision of the type localities of mammals occurring in the Tanganyika Territory. *Proc. Zool. Soc. Lond.* **115,** 49–84.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24,** 1596–1599.

Tamura, K. & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10,** 512–526.

Thomas, O. (1910). List of mammals from Mount Kilimanjaro, obtained by Mr. Robin Kemp, and presented to the British Museum by Mr. C.D. Rudd. *Ann. Mag. Nat. Hist.* **8**, 303–316.

Thomas, O. (1906). Mr O. Thomas on new African mammals. *Ann. Mag. Nat. Hist.* **7,** 179.

Thomas, O. (1917). A new *Heliophobius* from North-eastern Rhodesia. *Ann. Mag. Nat. Hist* **20,** 314–315.

Van Daele, P.A.A.G., Faulkes, C.G., Verheyen, E. & Adrians, D. (2007a). African mole-rats (Bathyergidae): A complex radiation in Afrotropical soils. In *Subterranean Rodents: News*

from Underground: 357–373 Begall, S, Burda, H. & Schleich, C.E. (Eds). Heidelberg: Springer-Verlag.

Van Daele, P.A.A.G., Verheyen, E., Brunain, M. & Adriaens, D. (2007b). Cytochrome b sequence analysis reveals differential molecular evolution in African mole-rats of the chromosomally hyperdivers genus Fukomys (Bathyergidae, Rodentia) from the Zambezian region. *Mol. Phylogenet. Evol.* **45,** 142–157.

Van Daele, P.A.A.G., Dammann, P., Meier, J.L., Kawalika, M., Van De Woestijne, C. & Burda, H. (2004). Chromosomal diversity in mole-rats of the genus Cryptomys (Rodentia : Bathyergidae) from the Zambezian region: with descriptions of new karyotypes. *J. Zool. (Lond.)* **264,** 317–326.

Walton, A.H., Nedbal, M.A. & Honeycutt, R.L. (2000). Evidence from Intron 1 of the nuclear transthyretin (prealbumin) gene for the phylogeny of African mole-rats (Bathyergidae). *Mol. Phylogenet. Evol.* **16**, 467–474.

Figure Legends

Figure 1: Map showing relative locations of the sampling sites, with yellow, blue, red and green colours corresponding to the coloured phylogenetic clades on Figs 2, 3 and 4. White squares denote locations of holotypes as labelled (excluding *H. albifrons* where the type locality is unknown), and * the sites of the fossil taxon *Proheliophobius*, used to calibrate the molecular clock. Dotted lines approximately delineate the western (Albertine Rift) and eastern (Kenya or Gregory Rift) branches of the East African Rift System.

Figure 2: Phylogenetic relationships of 48 *Heliophobius* mtDNA cyt *b* haplotypes and outgroup species *Heterocephalus glaber*. Bootstrap support indices (> 50%) are given above each node. The five main clades are labeled (1–5) at the relevant nodes. (a) Consensus cladogram obtained from maximum parsimony analysis after weighting informative sites with their rescaled consistency indices in a heuristic search (Tree length = 527.5 steps; Consistency Index 0.774, Homoplasy Index 0.226); (b) Consensus cladogram obtained from maximum likelihood analysis using a (TVM+I+G) model of sequence evolution (see text), tree score: –ln L = 6159.43.

Figure 3: The single most parsimonious cladogram reflecting evolutionary relationships of 11 *Heliophobius* cyt *b* haplotypes representing the main clades of the full sample set and the outgroup *Heterocephalus glaber* (see Table 1 for specimen details) generated by an exhaustive search in PAUP* (Tree Length = 827.0, Consistency Index 0.667, Homoplasy Index 0.333). Bootstrap support (> 50%) from 100 branch-and-bound replicates after weighting sites with the RCI (in bold), and without weighting, respectively are indicated at each node.

Fig. 4: Maximum clade credibility tree inferred using BEAST with an uncorrelated lognormal relaxed molecular clock model. Blue bars spanning nodes correspond to ages for the lower and upper bounds of the 95% highest posterior density (HPD) intervals (also shown in inset). Respective colours and labelled nodes of the five main clades correspond to Figs 1, 2 and 3 as discussed in the main text. Numbers at respective nodes refer to corresponding posterior probability support (%; for clarity, values are only given for principal nodes). Additional geological annotations are based on Ebinger (1989), Ebinger *et al*. (1993), Smith & Mosely

(1993) and Bauer *et al*. (2010); MTJ corresponds to the period of increased tectonic activity at the Mbeya Triple Junction.

Figure 5: Museum skins photographed to scale from the RMCA (a) *Heliophobius mottoulei* holotype (RMCA-1222) from Kilongwe, (b) "*H. robustus*" juvenile from Kisandji (RMCA-13487 – skull labelled as *robustus* and skin as *mottoulei*) and (c) *H. marungensis* from Dubie (RMCA-11262).

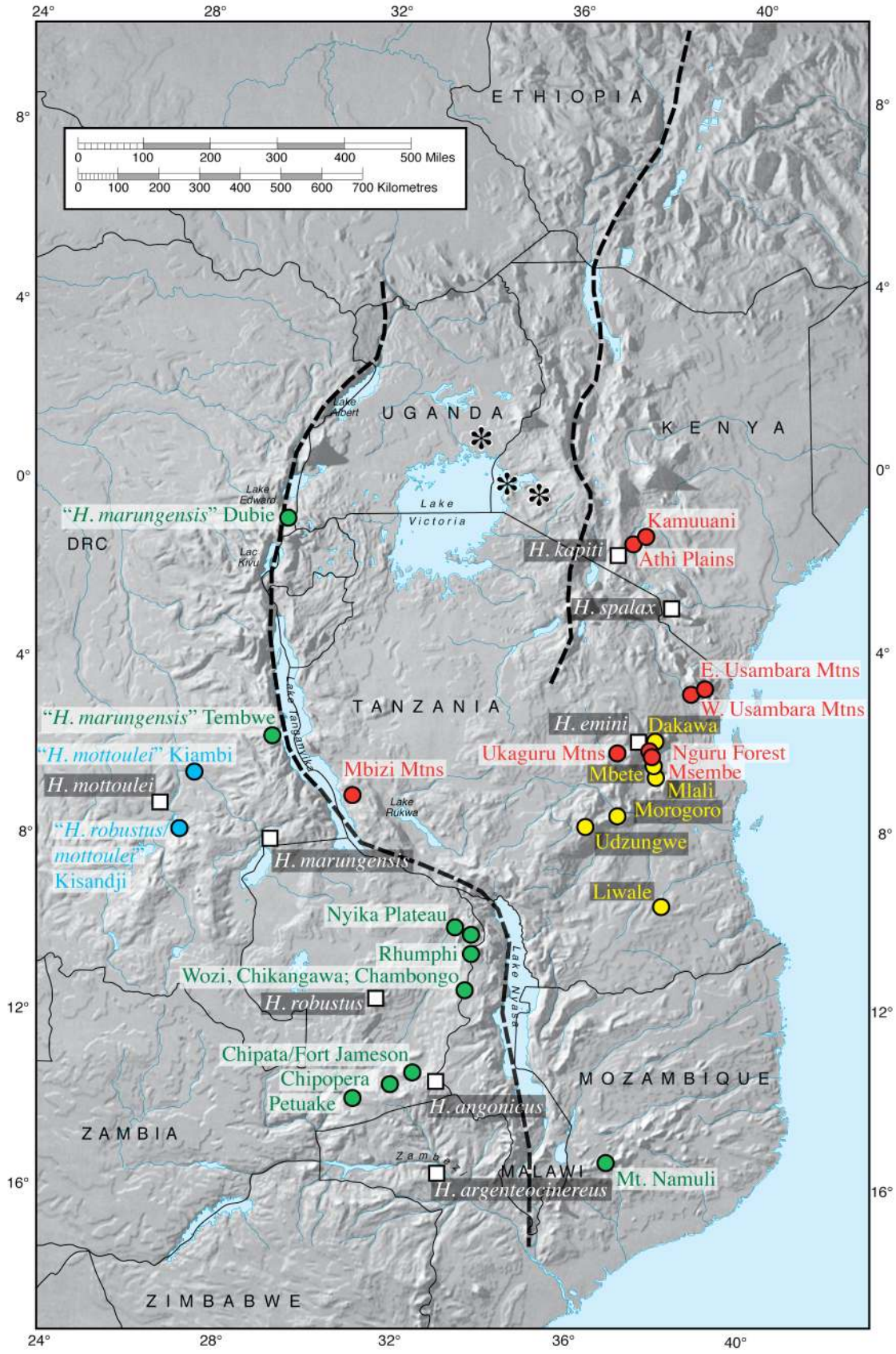
24°
28°
32°
36°
40°
8°
4°
0°
12°
16°
ETHIOPIA
0 100 200 300 400 500 Miles
0 100 200 300 400 500 600 700 Kilometres
Lake Albert
UGANDA
KENYA
Lake Edward
Lake Victoria
"H. marungensis" Dubie
DRC
Lac Kivu
Kamuuani
H. kapiti
Athi Plains
H. spalax
E. Usambara Mtns
W. Usambara Mtns
TANZANIA
Lake Tanganyika
"H. marungensis" Tembwe
H. emini
Dakawa
"H. mottoulei" Kiambi
Ukaguru Mtns
Nguru Forest
Mbizi Mtns
Mbete
Msembe
H. mottoulei
Mlali
Lake Rukwa
Morogoro
"H. robustus/ mottoulei" Kisandji
Udzungwe
H. marungensis
Liwale
Nyika Plateau
Rhumphi
Lake Nyasa
Wozi, Chikangawa; Chambongo
H. robustus
Chipata/Fort Jameson
Chipopera
MOZAMBIQUE
Petuake
H. angonicus
ZAMBIA
Zambezi
MALAWI
Mt. Namuli
H. argenteocinereus
ZIMBABWE

Fig. 1

Fig. 2

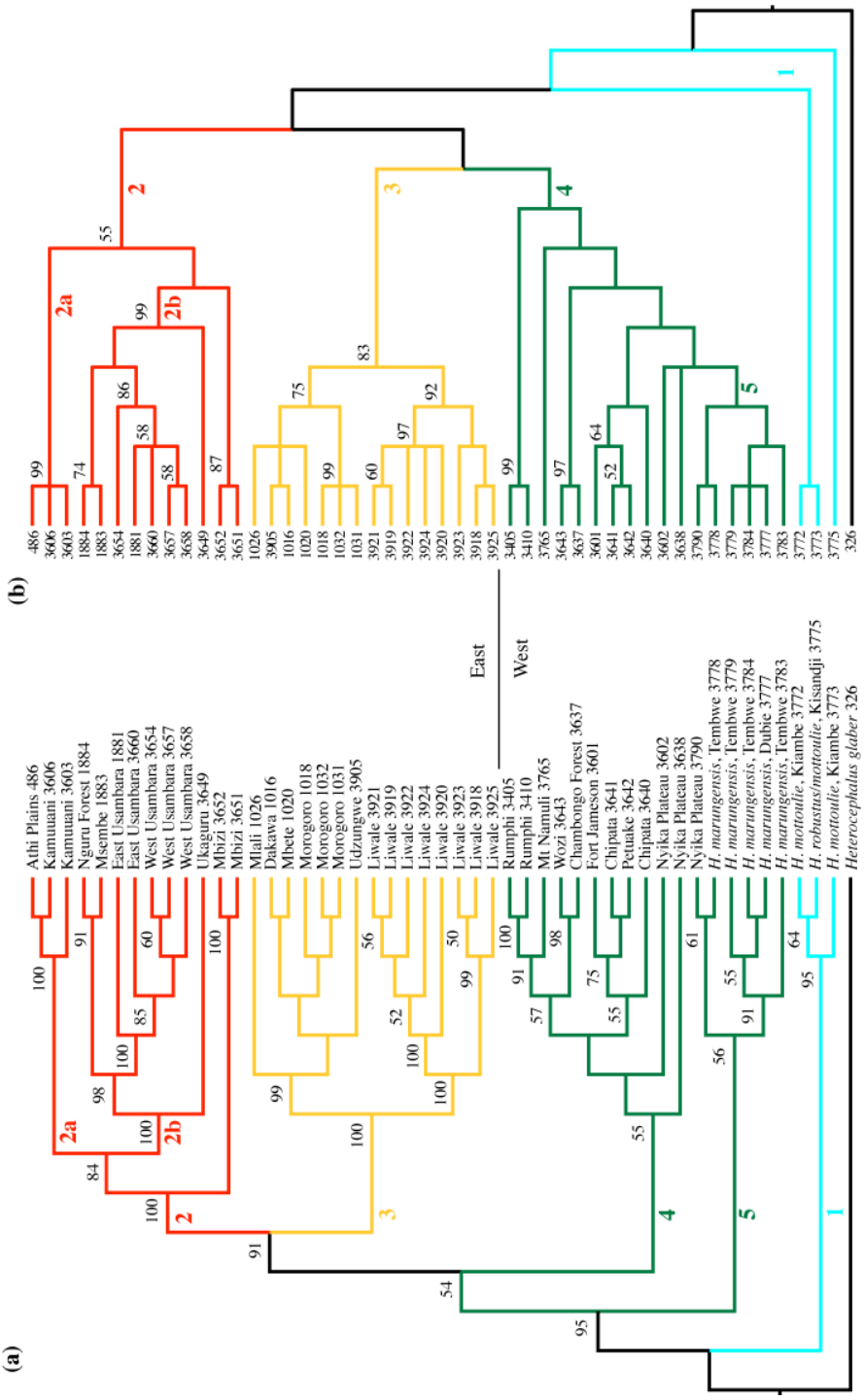
(a)
(b)
East
West
Athi Plains 486
Kamuuani 3606
Kamuuani 3603
Nguru Forest 1884
Msembe 1883
East Usambara 1881
East Usambara 3660
West Usambara 3654
West Usambara 3657
West Usambara 3658
Ukaguru 3649
Mbizi 3652
Mbizi 3651
Mlali 1026
Dakawa 1016
Mbete 1020
Morogoro 1018
Morogoro 1032
Morogoro 1031
Udzungwe 3905
Liwale 3921
Liwale 3919
Liwale 3922
Liwale 3924
Liwale 3920
Liwale 3923
Liwale 3918
Liwale 3925
Rumphi 3405
Rumphi 3410
Mt Namuli 3765
Wozi 3643
Chambongo Forest 3637
Fort Jameson 3601
Chipata 3641
Petuake 3642
Chipata 3640
Nyika Plateau 3602
Nyika Plateau 3638
Nyika Plateau 3790
H. marungensis, Tembwe 3778
H. marungensis, Tembwe 3779
H. marungensis, Tembwe 3784
H. marungensis, Dubie 3777
H. marungensis, Tembwe 3783
H. mottoulie, Kiambe 3772
H. robustus/mottoulie, Kisandji 3775
H. mottoulie, Kiambe 3773
Heterocephalus glaber 326
1
2
2a
2b
3
4
5

Fig. 3

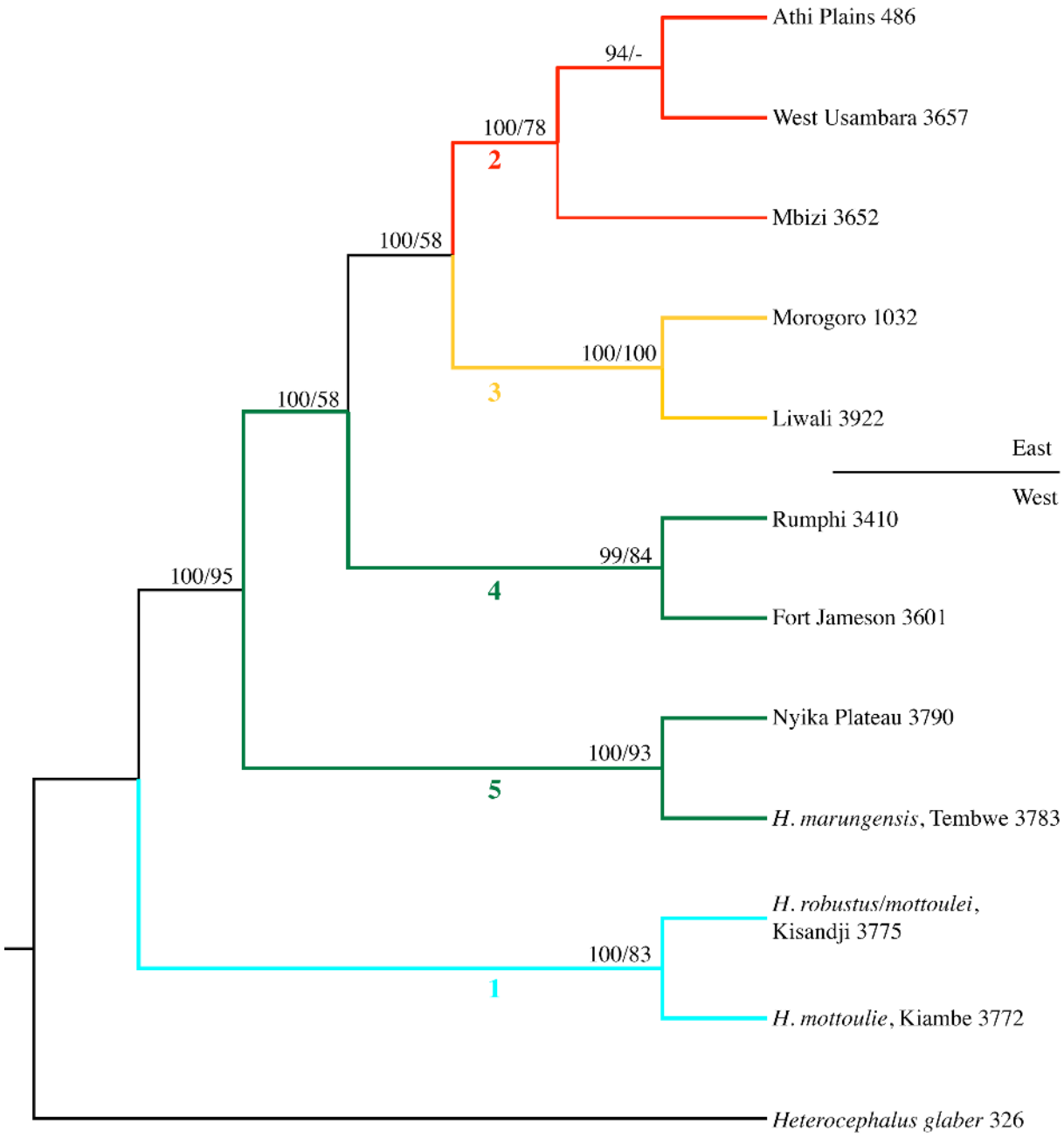
Athi Plains 486
94/-
West Usambara 3657
100/78
2
Mbizi 3652
100/58
Morogoro 1032
100/100
3
Liwali 3922
East
West
100/58
Rumphi 3410
99/84
100/95
4
Fort Jameson 3601
Nyika Plateau 3790
100/93
5
H. marungensis, Tembwe 3783
H. robustus/mottoulei, Kisandji 3775
100/83
1
H. mottoulie, Kiambe 3772
Heterocephalus glaber 326

Fig. 4

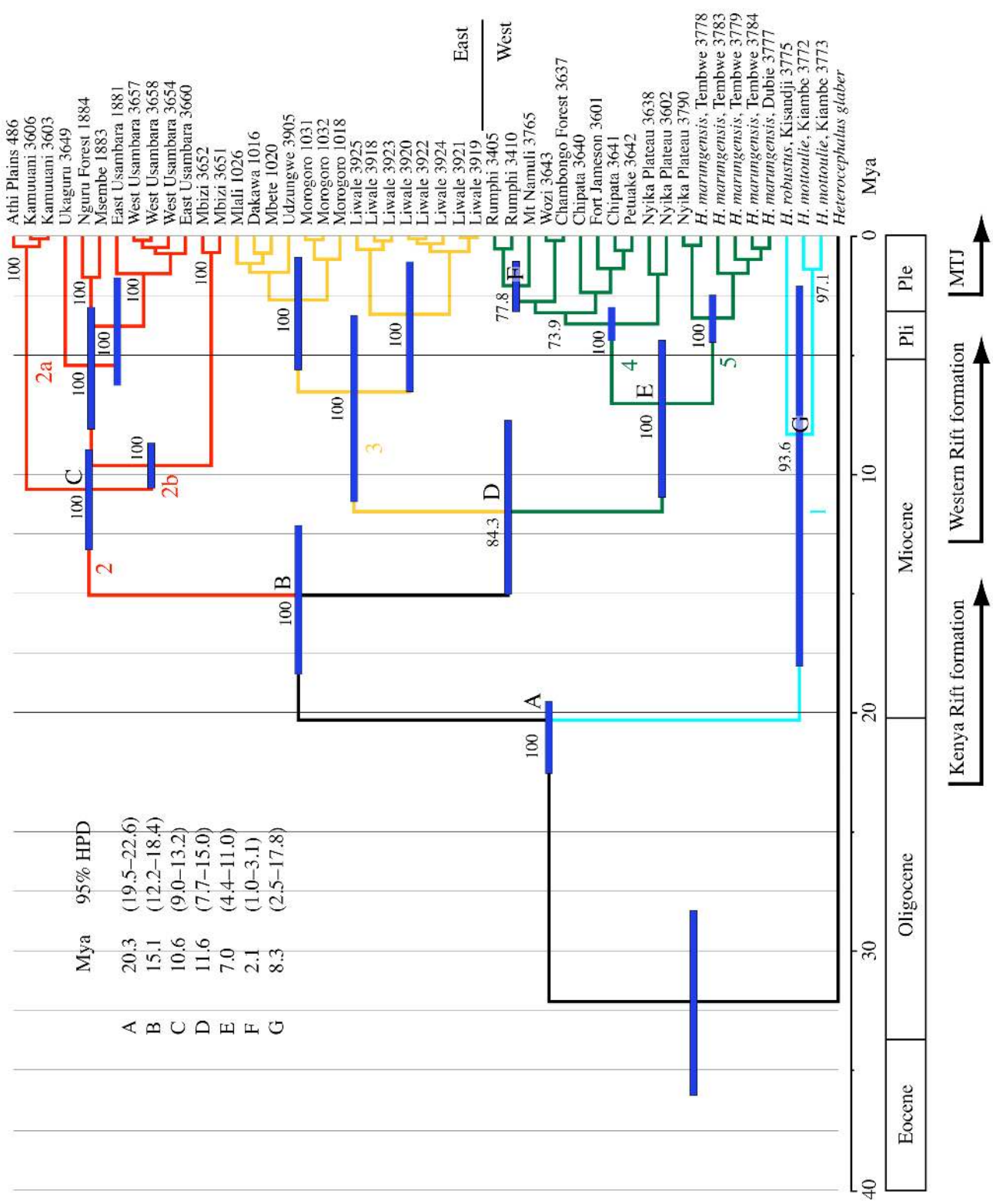
Athi Plains 486
Kamuuani 3606
Kamuuani 3603
Ukaguru 3649
Nguru Forest 1884
Msembe 1883
East Usambara 1881
West Usambara 3657
West Usambara 3658
West Usambara 3654
East Usambara 3660
Mbizi 3652
Mbizi 3651
Mlali 1026
Dakawa 1016
Mbete 1020
Udzungwe 3905
Morogoro 1031
Morogoro 1032
Morogoro 1018
Liwale 3925
Liwale 3918
Liwale 3923
Liwale 3920
Liwale 3922
Liwale 3924
Liwale 3921
Liwale 3919
Rumphi 3405
Rumphi 3410
Mt Namuli 3765
Wozi 3643
Chambongo Forest 3637
Chipata 3640
Fort Jameson 3601
Chipata 3641
Petuake 3642
Nyika Plateau 3638
Nyika Plateau 3602
Nyika Plateau 3790
H. marungensis, Tembwe 3778
H. marungensis, Tembwe 3783
H. marungensis, Tembwe 3779
H. marungensis, Tembwe 3784
H. marungensis, Dubie 3777
H. robustus, Kisandji 3775
H. mottoulie, Kiambe 3772
H. mottoulie, Kiambe 3773
Heterocephalus glaber
East
West
Mya
Mya
95% HPD
A 20.3 (19.5–22.6)
B 15.1 (12.2–18.4)
C 10.6 (9.0–13.2)
D 11.6 (7.7–15.0)
E 7.0 (4.4–11.0)
F 2.1 (1.0–3.1)
G 8.3 (2.5–17.8)
0
10
20
30
40
Eocene
Oligocene
Miocene
Pli
Ple
Kenya Rift formation
Western Rift formation
MTJ

Fig. 5

Table 1 – Number of individuals sequenced (*n*), number of haplotypes identified (*h*), the percentage of the complete *cyt* b gene (1140 bp) sequenced obtained for each sample together with their approximate geographical locations (also shown on Figure 1); individual identification numbers for all samples held at Queen Mary, University of London (prefixed with QM), together with respective museum reference numbers in parentheses (HZM: Harrison Institute; RMCA: Royal Museum of Central Africa; FMC: Field Museum of Natural History Chicago; NMZB, Natural History Museum of Zimbabwe; TM: Transvaal Museum) and GenBank accession numbers (further information on RMCA samples can be found at: http://projects.biodiversity.be/africanrodentia).

Country and capture location	Location Coordinates (degrees, minutes) Latitude, Longitude	*n*	*h*	%	Sample i.d./ (museum reference number)	GenBank accession numbers (where submitted)
Kenya, Athi Plains	1□43'S, 37□0'E	4	1	83.6 100 100 100	QM484 QM485 QM486 QM487	AY425923 AY425922 U87527 AY425921
Kenya, Kamuuani	1°33'S, 37°15'E	3	3	77.8 29.6 84.8	QM3603 (HZM 8.26807) QM3604 (HZM 7.26806) QM3606 (HZM 4.1331)	JN244276 JN244277 JN244278
Tanzania, Mlali	6°57'S, 37°40'E	7	1	98.6 100 96.2 100 100 100 63.0	QM1023 (RMCA 96.037-M-5586) QM1024 (RMCA 96.037-M-5587) QM1025 (RMCA 96.037-M-5588) QM1026 (RMCA 96.037-M-5629) QM1027 (RMCA 96.037-M-5670) QM1028 (RMCA 96.037-M-5671) QM1029 (RMCA 96.037-M-5680)	AY425938 AY425934 AY425927 AY425935 AY425939 AY425940 AY425941
Tanzania, Dakawa	6°26'S, 37°34'E	3	1	98.1 98.1 98.1	QM1016 (RMCA 96.036-M-5433) QM1017 (RMCA 96.036-M-5434) QM1021 (RMCA 96.036-M-5437)	AY425925 AY425924 AY425933
Tanzania, Morogoro	6□50'S, 37□39'E	3	3	86.6 98.0 100	QM1018 (RMCA 96.036-M-5435) QM1031 (RMCA 96.036-M-5449) QM1032 (RMCA 96.036-M-5458)	AY425926 AY425936 AY425937
Tanzania, Mbete	6°51'S, 37°40'E	1	1	95.8	QM1020 (RMCA 96.037-M-5585)	AY425932
Tanzania, Nguru Forest	6°39'S, 37°35'E	1	1	100	QM1884 (FMC 168421)	JN244280
Tanzania, Msembe	6°43'S, 37°37'E	1	1	100	QM1883 (FMC 166971)	JN244279
Tanzania, Amani (eastern Usambara Mountains)	4°59'S, 38°39'E	2	2	100 99	QM1881 (FMC 150371) QM3660 (FMC 150370)	AY425930 JN244281
Tanzania, Bagamoyo	5°6'S, 38°24'E	5	4	100 97.5 30.8 100 100	QM1882 (FMC 158272) QM3654 (FMC 147350) QM3656 (FMC 158272) QM3657 (FMC 158273) QM3658 (FMC 158274)	AY425931 JN244282 JN244283 JN244284 JN244285
Tanzania,	6°28'S, 36°50'E	1	1	97.7	QM3649 (FMC 166970)	JN244286

Ukaguru Mountains						
Tanzania, Udzungwe	8°9'S, 36°13'E	1	1	100	QM3905	JN244287
Tanzania, Mbizi Mountains	7°37', 31°8'	2	2	98.9 99.5	QM3651 (FMC 171473) QM3652 (FMC 171616)	JN244289 JN244288
Tanzania, Liwale	9°46'S, 37°57'E	8	8	94.3 98.1 85.4 5.7 99.1 99.1 90.3 76.9	QM3918 QM3919 QM3920 QM3921 QM3922 QM3923 QM3924 QM3925	JN244290 JN244291 JN244292 JN244293 JN244294 JN244295 JN244296 JN244297
Zambia, Fort Jameson	13°37'60S, 32°38'E	1	1	80.4	QM3601 (HM 2.10899)	JN244298
Zambia, Chipata	13°54'S, 32°10'E	2	2	34.4 24.6	QM3640 (NMZB 20227) QM3641 (NMZB 20226)	JN244299 JN244300
Zambia, Petuake	14°15'S, 31°19'E	1	1	32.8	QM2642 (NMZB 23013)	JN244301
Zambia, Nyika Plateau	10°28'S, 33°30'E	1	1	70.4 30.2 92.9	QM3602 (HZM 1.10803) QM3638 (NMZB 82584) QM3790 (TM 41757)	JN244302 JN244303 JN244304
Malawi, Wozi	11°49'S, 33°47'E	1	1	57.3	QM3643 (NMZB 83905)	JN244305
Malawi, Rumphi	11°1'S, 33°52'E	4	2	95.7 95.7 100 100	QM3404 QM3405 QM3410 QM3411	JN244306 JN244307
Malawi, Chambongo Forest	11°46'S, 33°46'E	1	1	60.0	QM3637 (NMZB 83907)	JN244308
Mozambique, Mt. Namuli	15°26'S, 37°3'E	1	1	67.6	QM3765 (FMC 183861)	JN244309
Democratic republic of Congo, Kiambe	7°8'S, 27°31'	2	2	64.6 35.8	QM3772 (RMCA 11307-M-) QM3773 (RMCA 11308-M-)	JN244310 JN244311
Democratic republic of Congo, Kisandji Mitwala	8°25'S, 27°12'E	1	1	31.6	QM3775 (RMCA 13487-M-)	JN244312
Democratic republic of Congo, Tembwe	6°17'S, 29°17'E	4	4	30.7 28.3 68.3 23.0	QM3779 (RMCA 8852-M-) QM3784 (RMCA 8845-M-) QM3783 (RMCA 8844-M-) QM3778 (RMCA 8851-M-)	JN244315 JN244316 JN244317 JN244314
Rwanda, Dubie	1°16'S, 29°17'E	1	1	30.5	QM3777 (RMCA 11262-M-)	JN244313

Table 2. Mean ± SEM cyt *b* genetic differences between the five *Heliophobius* clades (numbered 1–5 as in Figures 2, 3 and 4): Clade 1 = Central African samples (DRC and Rwanda) west of the Albertine Rift, designated as *H. robustus*/*mottoulei*; Clades 2 and 3 = Kenyan and Tanzanian samples, east of the Kenya Rift and Albertine Rift, respectively; Clade 4 = samples from Zambia and Malawi west of the Albertine Rift, and the single sample from Mozambique which is east of this part of the Rift; Clade 5 = samples west of the Albertine Rift in Zambia including museum material designated as *H. marungensis*. Above diagonal: uncorrected p-distances, below diagonal are TrN+G-corrected genetic distances (gamma-shape parameter = 1.230). Values along diagonal (in bold) depict within-clade uncorrected and TrN+G-corrected values (in parentheses).

	Clade 1	**Clade 2**	**Clade 3**	**Clade 4**	**Clade 5**
Clade 1	**4.7 ± 0.9 (5.5 ± 1.2)**	**17.7 ± 1.1**	**16.7 ± 1.2**	**16.7 ± 1.4**	**10.9 ± 0.8**
Clade 2	**25.1 ± 2.2**	**7.2 ± 0.5 (8.6 ± 0.6)**	**12.4 ± 0.7**	**12.2 ± 0.8**	**14.1 ± 1.3**
Clade 3	**22.8 ± 2.3**	**15.5 ± 1.1**	**2.4 ± 0.3 (2.5 ± 0.3)**	**8.0 ± 0.7**	**10.8 ± 1.2**
Clade 4	**23.7 ± 2.9**	**15.5 ± 1.0**	**9.3 ± 1.0**	**2.8 ± 0.4 (3.0 ± 0.4)**	**6.3 ± 0.7**
Clade 5	**15.5 ± 1.8**	**18.2 ± 1.2**	**13.4 ± 1.8**	**7.4 ± 0.9**	**3.2 ± 0.9 (3.5 ± 0.8)**