

<sup>1</sup>Department of Biological Science, Florida State University, Tallahassee FL, USA; <sup>2</sup>Department of Biological Sciences, Kuwait University, Safat, Kuwait; <sup>3</sup>Eastern Virginia Medical School, Norfolk VA, USA

## Molecular systematics of gerbils and deomyines (Rodentia: Gerbillinae, Deomyinae) and a test of desert adaptation in the tympanic bulla

BADER H. ALHAJERI<sup>1,2</sup>, ONDREIA J. HUNT<sup>1,3</sup> and SCOTT J. STEPPAN<sup>1</sup>

### Abstract

Recent molecular studies in gerbils found multiple instances of discordance between molecular and morphological phylogenies. In this study, we analyse the largest molecular data set to date of gerbils and their sister group the deomyines to estimate their phylogenetic relationships. Maximum-likelihood and Bayesian analyses were largely concordant, and both generally had high levels of node support. For gerbils, the results were generally concordant with previous molecular phylogenies based on allozymes, chromosomes, DNA/DNA hybridization and DNA sequences, and discordant with morphological phylogenies. None of the traditional gerbil tribes and subtribes were monophyletic. In addition, paraphyly was found in the genera *Gerbillus*, *Gerbilliscus* and *Meriones* as well as in five subgenera within *Dipodillus*, *Gerbillurus* and *Meriones*. Short branches separating taxa in small clusters within *Dipodillus* and *Meriones* suggest synonymy. Within deomyines, all genera and subgenera were monophyletic; however, two species groups within *Acomys* appear to contain synonymous taxa. We also find support for the discordance between molecular and morphological phylogenies in gerbils being partly due to convergent adaptations to arid environments, primarily in the suite of traits associated with inflation of the tympanic bullae. Relative bullar size does appear to be a desert adaptation and is correlated with aridity independent of phylogeny. Further, it varies more strongly along bioclimatic clines than between binary habitat classifications (desert versus mesic).

**Key words:** Arid environments – geometric morphometrics – molecular phylogenetics – Muroidea – skull morphology

### Introduction

Muridae is the most diverse family of mammals, consisting of over 700 species and 150 genera (Musser and Carleton 2005) in four subfamilies: Murinae, Lophiomyiinae, Deomyinae and Gerbillinae (Jansa and Weksler 2004; Schenk et al. 2013). Murids are found throughout the Old World, living in most habitats from hyperarid deserts to hypermesic rainforests (Nowak 1999). Murid systematics have changed drastically in the past 20 years, as a consequence of molecular phylogenetic studies (Musser and Carleton 2005). While these studies unambiguously support the sister status of gerbils and deomyines (Agulnik and Silver 1996; Martin et al. 2000; Michaux and Catzeflis 2000; Michaux et al. 2001; Jansa and Weksler 2004; Stepan et al. 2004a; Chevret and Dobigny 2005; Schenk et al. 2013), relationships within these two clades are less understood, with great discordance between morphological and molecular phylogenies of gerbils (Musser and Carleton 2005). No molecular study has thoroughly sampled deomyines (e.g. Jansa and Weksler 2004; Stepan et al. 2004a), and we provide the first such sampling here.

Members of Deomyinae (Thomas 1888), a subfamily that consists of 42 species and 4 genera of spiny mice and relatives (Musser and Carleton 2005), are found mostly in Africa and part of the Middle East including the Arabian Peninsula and Turkey (Fig. 1); within this range, they occupy various habitats including grasslands, forests, savannahs and deserts (Nowak 1999). Deomyines vary in behaviour, with diurnal, crepuscular and fully nocturnal species and also vary in nesting habits with both burrowing and non-burrowing species (Nowak 1999). This group does not share a large suite of morphological traits and were only assembled recently based on molecular evidence (Musser and Carleton 2005). Before this assembly, *Deomys* was placed in Dendromurinae by Thomas (1896) and remained there until molecular phylogenetic analyses (DNA/DNA hybridization, Denys et al. 1995; cytochrome *b*, Verheyen et al. 1995; VWF and/

or LCAT, Michaux and Catzeflis 2000; Michaux et al. 2001) robustly placed this genus in the same clade as the ‘acomyines’ (*Acomys*, *Lophuromys* and *Uranomys*). Similar to *Deomys*, ‘acomyines’ were only recently isolated from murines based on molecular data (e.g. chromosomes, Viegas-Péquignot et al. 1986; allozymes, Bonhomme et al. 1985; DNA/DNA hybridization, Chevret et al. 1993; 12S, Hänni et al. 1995; cytochrome *b*, Verheyen et al. 1995) that strongly supported their monophyly and isolation from murines (Musser and Carleton 2005).

Gerbillinae (Gray 1825), an Old World subfamily that consists of 103 species and 16 genera of gerbils, jirds and relatives (Musser and Carleton 2005), have a much larger geographic distribution than deomyines, covering most of Africa and a larger extent of Asia which ranges from the Middle East to central Asia (Fig. 1), wherein they occupy mostly arid, unproductive, open regions including deserts, grasslands and savannahs (Nowak 1999). One of the most characteristic features of many gerbils is their largely inflated tympanic bullae (cranial chambers that house the auditory ossicles) which function in sound amplification (aids in detecting interspecific vocalizations and foot drumming, as well as sounds from predators) in open, mostly desert, habitats where sound dissipates quickly (Lay 1972). Members of this group share a large suite of morphological and behavioural traits; the majority are diurnal, have varying degrees of ricochet locomotion (associated elongated tails and narrow hind legs) and are mostly desert adapted, which is evident by their efficient burrowing, well-developed vision (large eyes), and efficient water conservation (Nowak 1999). Unlike deomyines, gerbils have long been grouped together and defined based on derived skeletal, dental and male genital characters (e.g. similar tympanic bulla morphology and mastoid pneumatization; shared dental formula, enamel patterns and overall teeth morphology; small or absent coronoid process; reviewed in Lay 1972; Petter 1973; Carleton and Musser 1984; Pavlinov et al. 1990; Pavlinov 2001, 2008).

Despite sharing a large suite of morphological synapomorphies, relationships among gerbils are highly debated, with the two most inclusive morphological phylogenies (all 16 genera, Pavlinov 2001; 14 genera, Tong 1989) having low resolution at

Corresponding author: Bader H. Alhajeri (bader.alhajeri@ku.edu.kw)

Contributing authors: Ondreia J. Hunt (ondreia221@hotmail.com), Scott J. Stepan (stepan@bio.fsu.edu)

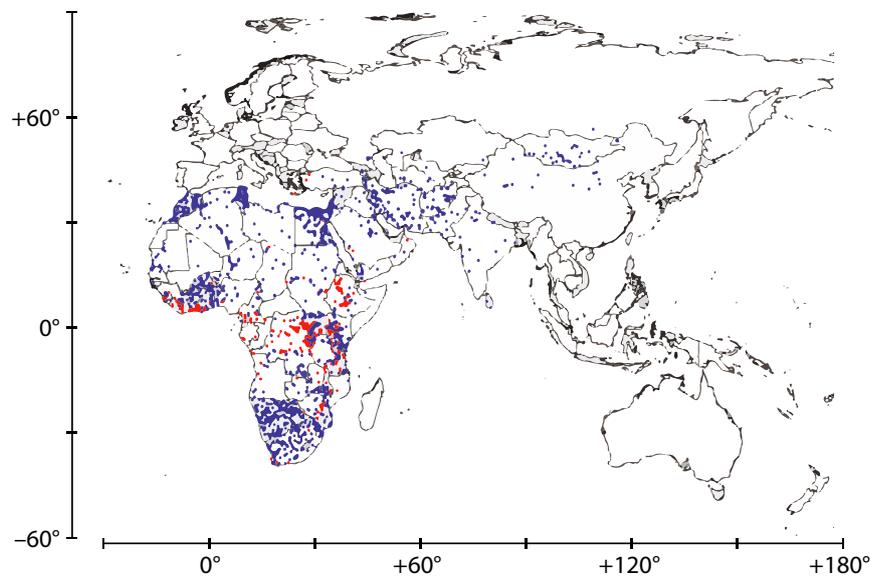


Fig. 1. Geographic distribution of gerbils and deomyines. Blue points = gerbils; red points = deomyines. Points indicate all species occurrence records from the Global Biodiversity Information Facility (GBIF 2013) data portal (<http://data.gbif.org/>). Individuals from all gerbil and deomyine species available in GBIF are mapped.

deep nodes and disagreements in a great deal of the systematic inferences (Musser and Carleton 2005). No one molecular phylogeny includes all extant gerbil genera, and most of the early phylogenies suffered from low generic sampling (e.g. the two studies with the largest sampled genera: seven genera using DNA/DNA hybridization, Chevret 1994; 6 genera using chromosome data, Benazzou 1984), low overlap in sampled species, lack of outgroups, and/or small data sets, making comparisons between studies difficult (Chevret and Dobigny 2005). Chevret and Dobigny (2005) sequenced 29 species in 11 recognized genera for cytochrome *b* and 12S and, based on this data set, suggested major taxonomic revisions at the tribal, subtribal, generic and subgeneric levels. The concordance between this DNA sequence phylogeny and previous genic phylogenies based on non-sequence data (e.g. DNA/DNA hybridization and chromosomes) is greater than the concordance between previous morphological phylogenies, which led Chevret and Dobigny (2005) to hypothesize that most of the characters used in morphological phylogenies are homoplastic and evolved as a consequence of strong constraints imposed by arid environments.

Similar to Chevret and Dobigny's (2005) study, most molecular phylogenies of gerbils were based on one or two mitochondrial genes. Moreover, most incorporated fewer gerbil species and genera and were more narrowly focused in taxonomic (i.e. relationships within genera) and geographic scale than Chevret and Dobigny's (2005) study. These studies include an investigation of the phylogenetic relationships within *Gerbilliscus* (Colangelo et al. 2007) and their morphological variation (Colangelo et al. 2010). In these studies, Colangelo and colleagues recognized three major clades within *Gerbilliscus* that correspond to major geographic subdivisions, and supported the synonymy of the *Gerbillurus* and *Gerbilliscus* (Colangelo et al. 2007). Similarly, Abiadh et al. (2010) examined the phylogenetic relationships within *Gerbillus* and found evidence for the synonymy of *Dipodillus* and *Gerbillus*. Ito et al. (2010) estimated the phylogenetic relationships of gerbils at the subfamily level and found evidence for the paraphyly of *Meriones* and a synonymy between two *Meriones* species.

In this study, we use the supermatrix approach to combine new sequences from multiple loci with most published DNA

sequences for gerbils and deomyines to provide the most comprehensive estimation of the systematics of the group to date. We perform both maximum-likelihood and Bayesian analyses on 13 loci, including nuclear and mitochondrial protein-coding genes, mitochondrial ribosomal RNA-coding genes, and introns, of 57 gerbil species from 14 genera, as well as 25 deomyine species from all four genera. This phylogeny was used for two main purposes: (1) to provide the basis for a taxonomic revision of these two subfamilies, and (2) to test the prediction that discordance between molecular and morphological phylogenies is a consequence of rapid morphological convergent adaptation in gerbils to aridity (e.g. Chevret and Dobigny 2005). We test this prediction using a fossil-calibrated chronogram along with a geometric morphometric (Zelditch et al. 2004) data set that includes the crania of most gerbils (including species lacking sequence data). This expanded morphological data set includes both mesic and desert species and was used to test the association between cranial (and specifically tympanic bulla) morphology and aridity indices, after correcting for correlations due to phylogenetic relatedness using phylogenetic analyses of variance and phylogenetic generalized least squares (PGLS). We focused on the tympanic bulla because it is the most widely studied potential mammalian desert adaptive character and is widely used in traditional gerbil systematics (see Discussion).

Hypertrophied tympanic bullae have long been attributed to desert adaptation in mammals (e.g. Lay 1972); however, no study used either phylogenetic correction or continuous bioclimatic variables to test this hypothesis. Gerbils are particularly suited for this study because they exhibit a great range of variation in tympanic bulla morphology. In a comparison of 13 gerbil species, Lay (1972) found that species living in more arid environments have increased auditory sensitivity as a consequence of increased anatomical specializations in the middle and inner ear anatomy (e.g. of the auditory bulla). This increased auditory sensitivity that accompanies the enlargement of the bullae was hypothesized to be an adaptation for predator avoidance in open xeric habitats where sound dissipates quickly and early detection is necessary to improve chances of escaping more effective predators (Lay 1972; Webster and Webster 1975). The use of phylogenetic correction in our study allows for isolating the effect of

phylogeny from adaptation in interspecific comparisons of bullar morphology.

Geometric morphometrics facilitates the statistical separation of size from shape variation better than traditional morphometrics when shape changes are complex. In this study, we use geometric morphometrics to estimate the overall size of the tympanic bulla and cranium, but do not investigate shape variation at this point. Because geometric morphometrics enables the estimation of a size measure that is mathematically independent of shape (Zelditch et al. 2004), this size measure may be a precise estimate of size, regardless of the choice of landmarks or arbitrary points outlining the structure of interest.

We also employ both discrete and continuous climatic variables to examine whether desert adaptation occurs in a continuous or binary fashion, the latter of which suggests two adaptive optima representing the bullar condition of mesic versus desert species.

## Material and Methods

### Phylogenetic sampling

A total of 95 species were included, representing all gerbil (57) and deomyine (25) species with published sequence data in addition to our newly sequenced species, and 13 outgroups. We sampled 14 of 16 gerbil and all 4 deomyine genera, totalling 55% of gerbil and 59% of deomyine species, respectively (Musser and Carleton 2005). The unsampled genera (*Ammodillus* and *Microdillus*) are monotypic and rare (found only in Somalia). Outgroups were sampled to maximize the inclusion of murid and cricetid fossil calibrations from Schenk et al. (2013). We included the monotypic Lophiomyiinae, the sister group of gerbils and deomyines. All taxonomy followed Musser and Carleton (2005).

### DNA extraction and sequencing

The final supermatrix used in the phylogenetic analyses is a concatenation of two mitochondrial ribosomal genes [1019 base pairs (bp) of 12S rRNA; 513 bp of 16S rRNA], eight nuclear protein-coding genes [148 bp of exon 3, 253 bp of intron 2 and 101 bp of exon 4 of acid phosphatase five (ACP5); 754 bp of exon 2, 220 bp of intron 2 and 79 bp of exon 3 of arginine vasopressin receptor 2 (AVPR2); 2388 bp of exon 11 of breast cancer 1 (BRCA1); 1214 bp of intron 3 of benzodiazepine receptor gene (BZRP); 878 bp of exon 10 of growth hormone receptor (GHR); 1122 bp of exon 1 of interphotoreceptor retinoid binding protein (IRBP); 2025 bp of the single exon of recombination activation gene 1 (RAG1); and 1264 bp of exon 28 of von Willebrand factor (VWF)], and three mitochondrial protein-coding genes [324 bp of cytochrome *c* oxidase I (COI); 684 bp of cytochrome *c* oxidase II (COII); and 1140 bp of cytochrome *b* (CYTB)], for a total of 14 126 sites. Most sequences were generated in our laboratory, either new here (27 ACP5, 10 BZRP, 26 CYTB, 5 GHR, 2 IRBP and 5 RAG1) (Appendix 1) or published previously (Steppan et al. 2004a, 2005; Rowe et al. 2008; Schenk et al. 2013), but were supplemented by sequences downloaded from GenBank (mostly published in Michaux et al. 2001; Jansa and Weksler 2004; Chevret and Dobigny 2005; Colangelo et al. 2007; Bösel et al. 2009; Ito et al. 2010; Ndiaye et al. 2014). Prior to inclusion of published sequences, multiple sequences per species and at least one sequence for all available species were aligned and analysed using maximum likelihood to generate gene trees. When possible for the concatenated analyses, only one sequence per gene per species was retained, using the criteria of monophyletic species in gene trees, maximum sequence length, and no apparent sequencing errors. Taxonomic identifications of newly sequenced samples and of published sequences were also assessed by reference to systematic authorities and alpha systematic studies (e.g. Musser and Carleton 2005; Yazdi and Adriaens 2013). Because there is no universal objective criterion for concatenating sequences of different genes from different individuals, we simplified this issue by selecting only one sequence for each gene in each species. This procedure assumes that species have been correctly identified. Inclusion of all available species provides an opportunity to detect misidentifica-

tions should they exist. Specimen locality information is listed in Appendix 2.

For newly generated sequences, total genomic DNA was extracted from vouchered museum tissues using standard phenol–chloroform–isoamyl alcohol extraction procedure as described by Sambrook et al. (1989). Polymerase chain reactions (PCRs) included 10× GoTaq buffer (Promega, Madison, WI, USA), 1 unit of GoTaq polymerase, 10 μM of forward and reverse primers, 0.15 mM of dNTPs, 3 mM of MgCl<sub>2</sub>, 0.2 μg BSA, 20–25 ng of DNA template, and ddH<sub>2</sub>O to a total volume of 25 μl. A negative control without template DNA was included in all PCRs to test for DNA contamination of reagents. Amplification and sequencing were completed with primer sequences under reaction conditions described previously depending on specific taxa (Jansa and Voss 2000; Adkins et al. 2001; DeBry and Seshadri 2001; Steppan et al. 2004a,b, 2005; Rowe et al. 2008; Schenk et al. 2013). ACP5 was amplified using a combination of the primer sequences 120FWD, 139FWD, 545REV, 564REV and S223 (DeBry and Seshadri 2001; Rowe et al. 2008). All BZRP amplifications used the primers S221 and S222 (Rowe et al. 2008). CYTB was amplified using a combination of the primer sequences S199, P484 and P485 (Rowe et al. 2008). GHR was amplified with the primers GHREXON10 and GHREND (Adkins et al. 2001). The IRBP region was amplified using a combination of the primers 119A2, B2 and 878F (Jansa and Voss 2000; Weksler 2003). RAG1 was amplified using a combination of the primers S70, S142, S73, S278 and S279 (Steppan et al. 2004b; Schenk et al. 2013). A summary of the used primers is given in Table S1.

PCRs were visualized on 1% agarose gels with ethidium bromide, and successful reactions were prepared by enzymatic digestion with EXOSAP-IT (Affymetrix, Cleveland, OH, USA). Both the 5' and 3' directions of the sequences were generated using the aforementioned primers at the FSU core facilities or at the DNA Analysis Facility on Science Hill at Yale University. SEQUENCHER 4.7 (Gene Codes Corporation, Ann Arbor, MI, USA) was used to assemble single sequence reads into a contiguous sequence with heterozygous sites scored as polymorphic. The 10 protein-coding sequences were realigned manually using MESQUITE 2.75 (Maddison and Maddison 2010) using the codon structure as a guide by consolidating indels and resulted in an unambiguous alignment. The remaining three sequences were aligned using MUSCLE (Edgar 2004). Species were represented in the concatenated data matrix by one to 11 loci (Appendix 1). New sequences have been deposited in GenBank under accession numbers KR088975–KR089049, and the matrix and trees to TreeBase under submission identification number 16150.

### Phylogenetic analysis

Phylogenetic estimation was conducted using maximum likelihood (ML) and Bayesian inference (BI) using RAXML 7.6.3 (Stamatakis 2006) and MRBAYES 3.2.2 (Huelsenbeck and Ronquist 2001), respectively; both run on the CIPRES Science Gateway (Miller et al. 2010). MODELTEST 3.1 (Posada and Crandall 1998) was used to estimate the best-fit DNA substitution model for each locus separately, for each data-type partition separately (e.g. codon positions, introns), and for the concatenated data using the Akaike Information Criterion (AIC; Akaike 1974). The GTR+Γ+I model was used in all the phylogenetic analyses because it was the best-fit model for all the data-type partitions, the concatenated data, and all loci except for BRCA1 and RAG1 (where GTR+Γ was the preferred model), BZRP and GHR (where HKY+Γ fit the data best); ACP5 (where K81uf+I fit the data best), and IRBP (where TrN+Γ+I fit the data best). In cases where the best-fit model was not implemented in RAXML, MRBAYES and/or BEAST analyses, the most similar available model was used instead (in all cases, GTR+Γ+I).

For the concatenated data, we conducted phylogenetic analyses using three partitioning schemes: (1) no partition, (2) 13 partitions corresponding to loci and (3) eight partitions corresponding to across-gene codon position and data type (i.e. mitochondrial ribosomal, nuclear introns, three mitochondrial codon positions, and three nuclear codon positions). Trees estimated using these alternative partitioning schemes were similar (data not shown), and only trees estimated using the eight-partition (across-gene codon position and data type) scheme are shown because it fit the data best in a previous phylogenetic study of muroids (Schenk et al. 2013). Parameter values among all partitions were unlinked.

RAXML searches were run 10 times from different random starting trees for individual loci and for the concatenation to escape local optima (Morrison 2007). The resulting trees for the multiple searches looked indistinguishable and the results of only one search are presented here. Clade support for the ML trees was determined using nonparametric bootstrapping (BS) as implemented in RAXML in CIPRES using rapid bootstrap inferences each optimized with ML, which resulted in 100–1000 replicated bootstrap searches.

For the MRBAYES analyses, a flat Dirichlet prior was applied on all trees and the GTR+ $\Gamma$ +I DNA substitution model for all partitions with clade support determined using Bayesian posterior probabilities (PP). Metropolis-coupled Markov chain Monte Carlo (MC<sup>3</sup>) was run independently four times, using different random starting trees, for 36 million generations each, sampling every 5000 generations from the posterior distribution. All the trees where the standard deviation of the split frequencies was >0.01 were discarded as burn-in generations (first 5–11% of the MC<sup>3</sup> chains); convergence and stationary of the post-burn-in trees was confirmed by evaluating all parameter values in the MC<sup>3</sup> chains in TRACER 1.5 (Rambaut and Drummond 2005), and no additional trees were excluded as burn-in. All four independent MC<sup>3</sup> chains had a post-burn-in effective sample size (ESS) of >600 for all parameters, and the combined MC<sup>3</sup> chain had an ESS > 3000 for each parameter. The post-burn-in trees from the four independent runs were combined manually and summarized with TREEANNOTATOR 1.8.0 (Drummond and Rambaut 2007) using the maximum-clade credibility tree criterion.

### Divergence-time analysis

Divergence times were estimated simultaneously with topology and branch lengths using an uncorrelated lognormal relaxed-clock model in BEAST 1.7.5 (Drummond and Rambaut 2007) using CIPRES. A GTR+ $\Gamma$ +I DNA substitution model was applied for all eight partitions (as in the MRBAYES analyses above) with clade support determined using Bayesian PP. The MC<sup>3</sup> chain was run for 50 million generations, sampling every 5000 generations from the posterior distribution. TRACER was used to determine appropriate burn-in based on convergence and stationary leading to the exclusion of the first 10% of the of the MC<sup>3</sup> chain as burn-in. Of 119 parameters, 113 had ESS > 200 and only 4 had ESS < 100, and the post-burn-in trees were summarized using TREEANNOTATOR. Five fossil calibrations were used to calibrate the chronogram (Table 1), all of which were used previously (see Schenk et al. 2013 and references therein for justification). Lognormal prior distributions were applied to all calibrations with means and standard deviations chosen to construct 95% confidence intervals (for the origination of the taxon based on first occurrence and stratigraphic sampling) spanning 95% Marshall indices (Marshall 1994) as reported by the Paleobiology Database (PDB 2013). The BI chronogram was used in subsequent comparative analyses.

### Morphological data collection

Morphological data were collected from 1 to 10 crania of 102 species of muroids (78 gerbils, 13 deomyines, 1 lophiomyine and 10 outgroups) for a total of 429 specimens (Table S2). We use four or more individuals where possible to reduce error in estimating the species mean. Although four specimens is too few to quantify intraspecific variation (Cardini and Elton 2007; Cardini et al. 2015), at a supraspecific and the suprageneric

Table 1. Calibration-point distributions including estimates for the BEAST analysis

Node	Clade	SD	Offset	5%	95%
5	<i>Acomys</i>	1.93	5.26	5.3	29.05
3	<i>Apodemus</i>	0.48	4.85	5.3	7.06
4	Gerbillinae	1.25	15.87	15.9	23.70
2	Murinae	0.89	9.8	9.8	14.05
1	Sigmodontini	1.41	4.80	4.9	14.93

All BEAST calibrations were assigned lognormal prior distributions. Node numbers correspond to those in Fig. 4. The ages are in million years before present. SD = standard deviation

level, it should be sufficient to characterize mean sizes for species, and any error from small sample sizes is likely to be merely a small addition of noise when examining broad-scale among-species patterns.

We sampled all the available gerbil species and most of the deomyine species in the visited institutions. At least four individuals were measured for every taxon, except for nine species that were too rare. All 16 of the gerbil genera were sampled except for the rare Somali pygmy gerbil, *Microdillus peeli*, which was not available; all four deomyine genera were sampled. The morphological sample includes 75% of gerbil species and 31% of deomyine species (Musser and Carleton 2005).

Skulls were examined of voucher specimens from the American Museum of Natural History (AMNH), the Field Museum (FMNH), the Museum of Vertebrate Zoology (MVZ), the Smithsonian Institution National Museum of Natural History (USNM), the Florida Museum of Natural History (UF), and the Laboratorio de Citogenética Mamíferos, Facultad de Medicina, Universidad de Chile (LCM). Wild-caught adults were chosen, as determined by basioccipital–basisphenoid epiphyseal fusion following Robertson and Shadle (1954) and Samuels (2009), the complete eruption 3rd molars assessed by reaching the occlusal surface (Steppan 1997) and the examination of the size associated skins.

Outlines of the ventral cranium and tympanic bulla semilandmarks were digitized on photographs taken with a Nikon D3200 digital SLR camera using a Nikon 40 mm f/2.8G AF-S DX Micro-Nikkor Lens (Nikon, Tokyo, Japan) at 24 megapixels in a standardized manner. All photographs included a scale bar. The left side of the cranium was digitized unless it was damaged; in these instances, the right side was digitized and reflected. Semilandmarks of those curves were digitized using TpsDig 2.16 (Rohlf 2010). Although projection of 3D skull shape onto 2D photographic images will cause some distortion of distance, the effect is relatively minor for cranial outlines where landmarks are close to the midplane (Zelditch et al. 2004). Moreover, since we use the semilandmarks to estimate only size and not shape, this distortion, if present, will be unlikely to significantly affect the estimate of cranial and bullar size. We confirmed this by testing the association between the centroid size of the tympanic bulla and skull obtained using semilandmarks, with more common linear estimates of tympanic bulla and cranial size extracted from Alhajer (2014); we found a strong positive correlation (see Results).

### Morphometric processing and analyses

Coordinates marking the outline of the ventral outer edge of the cranium and the tympanic bulla were digitized following a method modified from Momtazi et al. (2008). For the tympanic bulla, the outline was estimated from points digitized along the structure in a clockwise fashion starting with the anterior–medial most of the junction between tympanic bulla and pterygoid process (Fig. 2). The points were then resampled along the curve by length, resulting in 200 semilandmarks. The same process was repeated for the ventral view of the left side of the cranium starting from

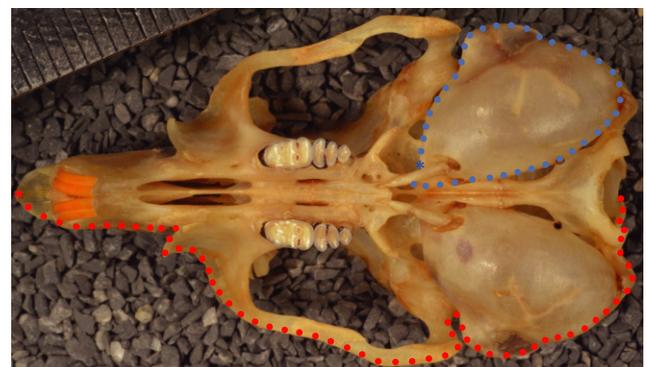


Fig. 2. Positions of the cranial semilandmarks used in the geometric morphometric analyses of morphological variation (blue = bulla, red = cranium) shown on the ventral view of Shaw's jird (*Meriones shawi*; USNM 474187). Not all semilandmarks are shown (bulla = 200; skull = 400). Cranial semilandmarks are shown on the right side but were digitized on the left.

the anterior most region of the midsagittal plane, between incisors/nasal bones (Fig. 2), with the points resampled to give 400 semilandmarks.

A generalized procrustes analysis (GPA; Rohlf and Slice 1990) was conducted on the data set to obtain centroid size (CS) with shape coordinates discarded. GPA translates the coordinate configurations to a common centroid by scaling them to unit centroid size (CS) and rotating them in order to minimize the sum of squared distances between the corresponding landmarks as well as account for the effects of translation (Zelditch et al. 2004). Semilandmarks were slid along their tangent directions using the procrustes distance criterion (Adams and Otárola-Castillo 2013). Centroid sizes (CS), outputted from a GPA on each specimen, were averaged to estimate species average size (Claude 2008) as implemented in the Geomorph library (Adams and Otárola-Castillo 2013) in R (R Development Core Team 2013). GPA analyses were conducted separately for the cranium and bulla data sets. CS were log-transformed prior to subsequent comparative analyses to meet the assumptions of normality of these analyses. Three size measures were analysed: (1) log bulla CS, (2) log cranium CS and (3) relative bulla CS [= log (bulla CS/cranium CS)].

### Extraction of habitat categories and continuous environmental data

Species were categorized as 'desert' species if they were classified by IUCN (2013) as '8. Desert' in their habitat classification scheme, and 'mesic' for any other classification (e.g. '1. Forest', '2. Savanna'; Figure S1; Table S3). IUCN relies on experts to classify the habitat of each species based on its distribution. This category included species that live in '8.1 Hot', '8.2 Temperate', and '8.3 Cold' deserts. In instances where species live in more than one habitat type (e.g. a desert and a mesic region), the species were classified as desert if their range included a desert. Therefore, desert species were not those that live exclusively in deserts, but rather those whose habitat classification included a desert, and mesic species were those whose habitat classification did not include a desert. Analyses were also conducted on a classification based on Shenbrot et al. (1999), and results of the two schemes were compared (Table S3). Shenbrot et al. (1999) considered desert rodents to be species with >50% of their range in an arid region while excluding species restricted to mesic refugia such as oases and species with exclusive subterranean lifestyle such as naked mole rats. By implication, all other species not found in this list were considered to be mesic species.

Bioclimatic variables were obtained for the habitat of each species from WORLDCLIM 2.5-min geographic information system (GIS) layer (Hijmans et al. 2005) using DIVA-GIS 7.5 (Hijmans et al. 2012) by cross-referencing geographic localities (range distribution maps) of each species (polygon shape files from IUCN) with the WORLDCLIM 2.5-min database. We used the average value of each grid cell observation (range = 2–1 318 798 grid cells) that fell within the boundary of the polygon range file. Mean annual temperature and mean annual precipitation were extracted for comparative analyses in addition to average temperature and precipitation of the driest quarters which were used to calculate the aridity index (Tables S3 and S4) following the method of de Martonne (1942), which is also known as the de Martonne–Gottman aridity index and the Pinna Combinative Index (Zambakas 1992):

$$AI = \frac{P}{T + 10} + \frac{12p}{t + 10}$$

where  $P$  is the mean annual rainfall in millimetres,  $T$  is the temperature in degrees Celsius,  $p$  is the average rainfall of the driest month in millimetres, and  $t$  is the average temperature of the driest month in degrees Celsius. This index was expanded from the formula that was introduced by de Martonne (1927) where the temperature and the rainfall of the driest months were not accounted for. In de Martonne's (1942) aridity index, temperature was used as a proxy for potential evapotranspiration which was not available from DIVA-GIS and otherwise might be preferred (Maliva and Missimer 2012). While evaporation was calculated as a function of temperature alone, in reality it is related to many

other factors including the amount of moisture in the soil, the type of soil, wind velocity, atmospheric pressure and plant cover (Walton 1969). Because DIVA-GIS does not output the average temperature of the driest month ( $t$ ), the aridity index was modified to use quarterly data (i.e.  $t$  = average temperature of the driest quarter;  $p$  = average rainfall of the driest quarter) following Eckert et al. (2010). Aridity here captures water availability as a function of temperature and precipitation and is arguably the best measure of the abiotic variables that desert organisms adapt to. This index is unitless, with lower values indicating increased aridity (Baltas 2007).

To meet the assumptions of normality of subsequent analyses, mean annual rainfall was log-transformed. The aridity index was also log-transformed, but only after the calculation of the index from the raw data (e.g. Arroyo et al. 2006). In addition, since the raw aridity index scores are usually positive values with a larger value indicating a more mesic region, the negative raw aridity index scores were dropped from the analyses (a value of 0 is the lowest possible aridity index score in de Martonne (1942) method). Negative values are a known problem in cold regions, where they result from very low temperatures, regardless of moisture content.

Species were coded as Sub-Saharan African or Eurasian (Palearctic, including arid North Africa, and the New World) following the biogeographic divisions of Schenk et al. (2013). All extant gerbils and deomyines were found to live in either Africa or Asia, with a few species spanning both. In these instances, species were unambiguously assigned to either Africa or Asia based on the degree of spanning either continent. Historical biogeographic reconstruction was conducted with Mesquite under maximum parsimony.

A species accumulation curve for the chronogram was visualized using a lineage-through-time (LTT) plot, constructed using the Ape library (Paradis et al. 2004) in R.

### Comparative analyses

Two trees were used for the phylogenetic comparative analyses, the BI chronogram (1) with species missing morphological data pruned off, and (2) the chronogram, also with species missing morphological data pruned off, but further with additional species that have morphological data grafted onto their closest relatives based on the taxonomy of Carleton and Musser (2005) (Figure S1). Grafting and pruning were conducted in using the Ape library.

The correlation between cranial and bullar size with habitat and climate was tested by conducting the following: (1) phylogenetic analysis of variance (PHYANOVA) of CS versus habitat (Table S3) with 1000 Brownian motion simulations following the method of Garland et al. (1993); and (2) PGLS analysis of CS versus mean annual precipitation, mean annual temperature and the aridity index (separately) following the method of Freckleton et al. (2002). PGLS and phylogenetic independent contrasts are equivalent methods when the same permutation model is used (Adams and Collyer 2015). We also conducted an ancestral state reconstruction analysis (ACE) on the relative bulla size on the pruned tree (tree 1) to determine how many times taxa evolved qualitatively enlarged bullae. PHYANOVA and ACE were conducted in the Phytools library (Revell 2012) and PGLS was conducted in the Caper library (Orme et al. 2013) in R.

## Results

### Phylogenetic analysis

Maximum likelihood analysis of individual loci led to a single ML tree each (Figure S2). Although largely in agreement, all individual locus trees had at least one area of incongruity with the single concatenated data set ML tree (Figure S3) in the estimated relationships among genera (Text S1). Most of the aforementioned incongruities in the relationships among genera in loci trees occur in nodes with low BS values and/or short branches and represent a shift in the placement of the clades by one or a

few nodes. Most differences between the ML and the BI trees occurred within genera (e.g. the exact placement of *Gerbillus latastei*). The only among-genus difference concerned the monophyly of *Meriones* where the ML and the BEAST BI analyses found the genus to be paraphyletic while the MRBAYES BI analysis found it to be monophyletic, but support was low in both cases. PP values from the BEAST and the MRBAYES analyses were highly similar, and the PP values in the following results are based on MRBAYES only (Fig. 3).

Most clades in the concatenated data analyses were well supported (76% of nodes  $\geq 85\%$  BS,  $84\% \geq 0.95$  PP) including all three of the polytypic subfamilies and both families (PP, 1.0; BS, 100%), and relationships among subfamilies. Most nodes with low PP and BS scores coincide with short branches within genera (e.g. within *Gerbillus*, *Meriones*, *Acomys* and *Lophuromys*) and/or in regions with most conflict in the individual locus trees (e.g. within the *Meriones* + *Psammomys* + *Brachionomys* + *Rhombomys* clade). Although the placement of *Taterillus* showed the most incongruence in the individual loci trees, its placement as a sister to Gerbillini I clade (Fig. 4; Figure S3) was strongly supported in the concatenated data set (PP, 1.0; BS, 100%).

Within Deomyinae, both polytypic genera were monophyletic with mixed support (*Acomys*, PP, 0.85; BS, 53%; *Lophuromys*, PP, 0.51; BS, 73%). Deomyinae consisted of a basal split between the monotypic genus *Uranomys* and the remaining three genera (PP, 1.0; BS, 100%), followed by a split between *Acomys* and *Lophuromys* + *Deomys* (PP, 1.0; BS, 93%). The split between the monotypic *Deomys* and *Lophuromys* was weakly supported (Fig. 3; Figure S3). Within *Acomys*, there was weak support for the basal split between the subgenus *Subacomys*, consisting of only *Acomys subspinosus*, from the rest of the sampled species which form the subgenus *Acomys* (PP, 0.85; BS, 53%). The monophyly of the subgenus *Acomys* and its exact placement within the genus is yet to be determined as *Acomys louisaie*, which forms the monotypic subgenus *Peracomys*, was not sampled. Within *Lophuromys*, based on the only sampled member of the subgenus *Kivumys* (*Lophuromys woosnami*), the basal split between *Kivumys* and the subgenus *Lophuromys* (all other sampled species) had strong support (monophyly of subgenus *Lophuromys* PP > 0.95; BS, 100%).

*Acomys minous* (two specimens, 'GroupA' and 'GroupB'), *Acomys cilicicus*, *Acomys nesiotis* and *Acomys cahirinus* form a well-supported clade (PP > 0.95; BS, 98%) and all are connected

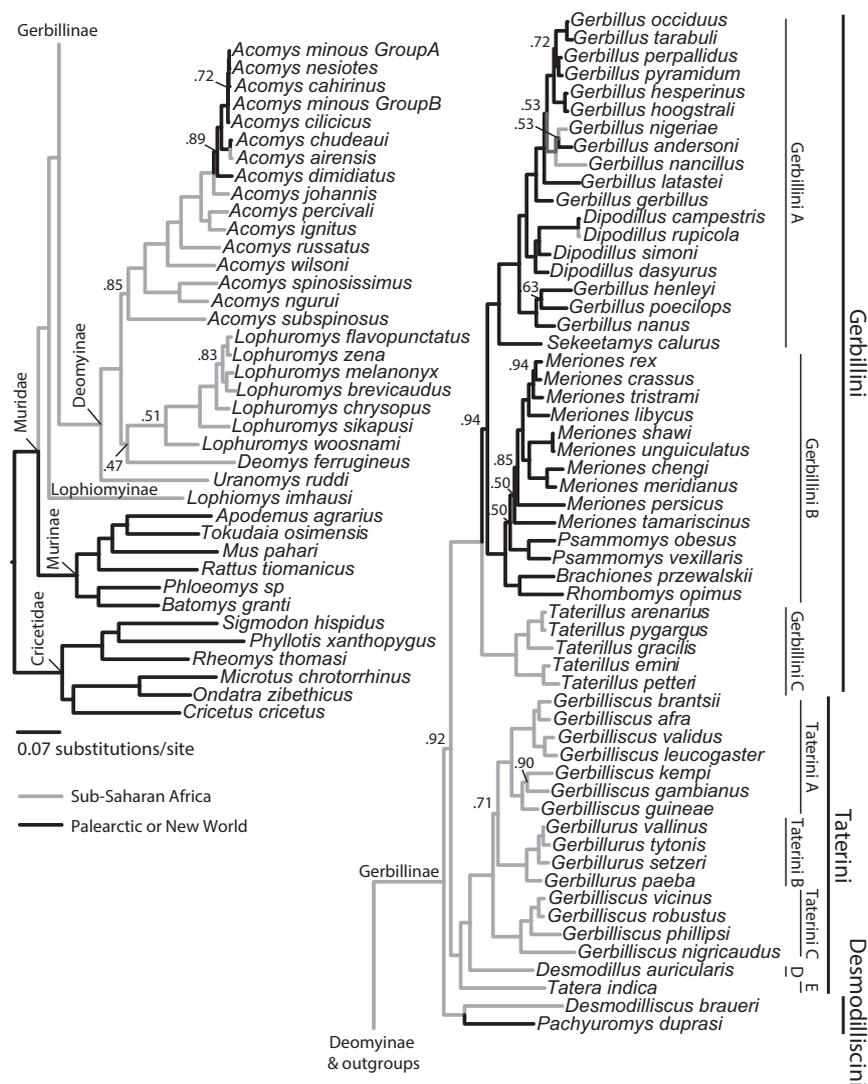


Fig. 3. BI phylogram of the concatenated data set from the MRBAYES analysis. PP values are indicated on the nodes. All other nodes (not annotated) are strongly supported (PP > 0.95). Proposed tribal and subtribal designations for gerbils based on Chevret and Dobigny (2005) and this study are also indicated. Branches are coded as black or grey corresponding to reconstructed geographic distributions using maximum parsimony.

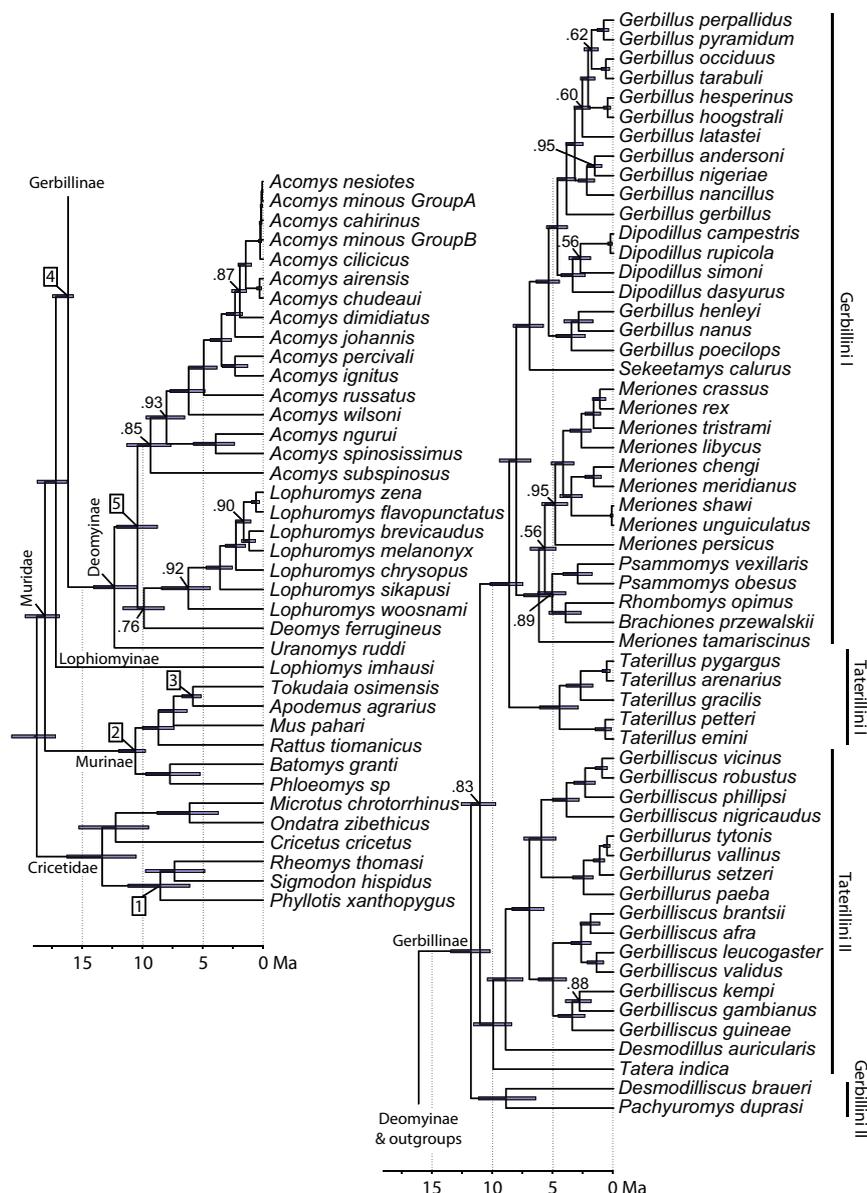


Fig. 4. Time-calibrated ultrametric chronogram from the BEAST analysis of the concatenated data set. PP values are indicated on the nodes. All other nodes (not annotated) are strongly supported (PP > 0.95). Node bars denote the 95% highest posterior densities. Nodes constrained in the analysis based on fossil calibrations are indicated with numbers inside squares that correspond with the fossils described in Table 1. Traditional tribal designations for gerbils based on the morphological phylogenies of Pavlinov et al. (1990) and Pavlinov (2001) are also indicated (Pavlinov's Gerbillini and Taterillini each form a clade that includes both the I and II clades).

by very short branches suggesting that they might collectively represent a single species, as noted by Barome et al. (2001). *Acomys chudeaui* and *Acomys airensis* are also potentially the same species, because they are similarly separated by very short branches (Fig. 3 and Figure S3).

Within Gerbillinae, both tribes Taterillini (*Tatera*, *Taterillus*, *Gerbilliscus*, *Gerbillurus* and *Desmodillus*) and Gerbillini (all other genera) were paraphyletic with Taterillini nested within Gerbillini. Within Taterillini, the two subtribes Gerbillurina (*Gerbillurus* and *Desmodillus*) and Taterillina (*Tatera*, *Gerbilliscus* and *Taterillus*) were also paraphyletic (Fig. 3 and Figure S3) with multiple clades that conflicted with monophyly strongly supported (PP, 1.00; BS, 100%). *Gerbillus*, *Gerbilliscus* and *Meriones* were paraphyletic: the first two strongly supported (PP, 1.00; BS, 100%): *Gerbillus* with respect to *Dipodillus*, and *Gerbilliscus* with respect to *Gerbillurus*. Paraphyly of *Meriones* with

respect to *Brachiones* + *Rhombomys* + *Psammomys* was weakly supported in the ML tree (Figure S3) and BI chronogram (Fig. 4) but was monophyletic in the BI tree (Fig. 3).

Gerbillinae consisted of a basal split between a highly supported clade that consists of the two monotypic genera *Pachyuromys* and *Desmodilliscus* (PP, 1.00; BS, 99%) and the rest of Gerbillinae, followed by a split between *Tatera* + *Desmodillus* + *Gerbilliscus* + *Gerbillurus* and the remaining genera. The first clade consisted of a basal split between the monotypic *Tatera* and *Desmodillus* + *Gerbilliscus* + *Gerbillurus* followed by a split between the monotypic *Desmodillus* and *Gerbilliscus* + *Gerbillurus*, the latter forming a paraphyletic group.

The second clade consisted of a basal split between *Taterillus* and *Psammomys* + *Rhombomys* + *Brachiones* + *Meriones* + *Sekeetamys* + *Gerbillus* + *Dipodillus* followed by a split between *Psammomys* + *Rhombomys* + *Brachiones* + *Meriones* and *Seke-*

*tamys* + *Gerbillus* + *Dipodillus*. The relationships among *Psammomys*, *Rhombomys*, *Brachiones* and *Meriones* were ambiguous and were incongruent between the ML tree and the BI chronogram versus the BI tree. In the ML tree and BI chronogram, *Meriones tamariscinus* was sister to *Psammomys* + *Rhombomys* + *Brachiones* + *Meriones* with *Psammomys* and *Rhombomys* + *Brachiones* as sister clades. In the BI tree, *Rhombomys* + *Brachiones* were sister to *Psammomys* plus a monophyletic *Meriones*. Because *Meriones* was monophyletic for both individual gene trees (CYTB and COII), its paraphyly in ML and the BI chronogram might be an artefact arising from a very short internal branch (Ito et al. 2010). *Meriones tamariscinus* is the type species, and so resolution of the taxonomy will require additional data from nuclear genes. The monotypic *Sekeetamys* was sister to a paraphyletic *Gerbillus* with *Dipodillus* nested within.

Within *Dipodillus*, both subgenera are paraphyletic; subgenus *Dipodillus* (represented in the sample by *Dipodillus simoni*) was nested within subgenus *Petteromys* (all other sampled *Dipodillus*). The extremely short branches separating *Dipodillus rupicola* and *Dipodillus campestris* indicated that these two species may be synonymous. *Gerbillurus* consisted of a basal split between the monotypic subgenus *Progerbillurus* (*Gerbillurus paeba*) and the other two subgenera where the polytypic subgenus *Gerbillurus* was paraphyletic having *Gerbillurus tytonis* (subgenus *Paratatera*) nested within it. Within *Gerbillus*, both subgenera were monophyletic with the genus consisting of a basal split between subgenus *Hendecapleura* (represented in the sample by *Gerbillus henleyi*, *Gerbillus nanus* and *Gerbillus poecilops*) and *Dipodillus* + *Gerbillus* (the latter consisting of all other sampled *Gerbillus*) that in turn form two monophyletic groups. Despite the ambiguity of the relationships among genera close to *Meriones*, in all instances, subgenus *Meriones* (consisting of only *M. tamariscinus*) was clearly separated from all other *Meriones* representing the basal split within the genus or within a larger clade that subsumes the genus. However, the other two sampled subgenera were paraphyletic with *Meriones rex* (which along with *Meriones persicus* form subgenus *Parameriones*) nested within *Pallasiomys* (all other sampled *Meriones*); *M. persicus* was however divergent from other *Meriones*. The very short branches separating *Meriones shawi* and *Meriones unguiculatus* also suggested that these two species may be synonymous.

### Association of cranial size with climate

Regression analyses between the centroid size of the cranium versus traditional linear estimates of cranial size (skull length and width) as well as between the centroid size of the bulla versus traditional linear estimates of bullar size (bullar length and

width) were all significantly positive (all  $p$  values < 0.000001 and all  $R^2 > 0.90$ ; Table S5), indicating a strong positive association. Similarly, an estimate of bullar volume based on the formula of an elliptical cone was also significantly positively correlated with bullar CS, indicating that CS extracted using semilandmarks in this study may faithfully estimate cranial and bullar size.

Species classified by IUCN as occurring in desert habitats lived in regions that were significantly more arid (low aridity index scores) received significantly less mean annual precipitation and had slightly colder temperatures (by about 1°C) than species classified as mesic (data not shown; see Alhajeri 2014). Similar results were found for the comparative analyses using both the pruned and the grafted (expanded) trees (Figure S1), as well as when using the full data set or when restricting analyses to gerbils; therefore, unless otherwise specified (in cases of discordance in significance), all results will refer to the full data set, grafted tree analyses.

PHYANOVA indicated that cranial CS did not significantly differ between desert and mesic species ( $F = 0.059$ ,  $p = 0.931$ ; Fig. 5a). PGLS showed that cranial CS was significantly positively although weakly correlated with the aridity index (Coef = 0.026,  $R^2 = 0.063$ ,  $p = 0.007$ ; Fig. 5b) and significantly negatively correlated with mean annual temperature (Coef = -0.002,  $R^2 = 0.047$ ,  $p = 0.0167$ ; Fig. 5d), indicating that cranial size was greater in more mesic and colder regions. PGLS indicated that cranial CS was not significantly correlated with mean annual precipitation (Coef = 0.015,  $R^2 = 0.002$ ,  $p = 0.371$ ; Fig. 5c).

PHYANOVA indicated that desert species had significantly greater bullar CS than mesic species ( $F = 14.58$ ,  $p = 0.032$ ; Fig. 6a); however, this result was not significant in the pruned full data set analysis ( $F = 11.60$ ,  $p = 0.113$ ) or in both the grafted ( $F = 2.49$ ,  $p = 0.169$ ) and the pruned ( $F = 0.22$ ,  $p = 0.699$ ) gerbil only analyses. PGLS indicated that bullar CS was significantly negatively correlated with the aridity index (Coef = -0.029,  $R^2 = 0.071$ ,  $p = 0.004$ ; Fig. 6b) and mean annual temperature (Coef = -0.003,  $R^2 = 0.143$ ,  $p < 0.0001$ ; Fig. 6d) indicating that bullae were larger in more arid and colder environments. PGLS indicated that bullar CS was not significantly correlated with mean annual precipitation (Coef = -0.021,  $R^2 = 0.004$ ,  $p = 0.236$ ; Fig. 6c). PHYANOVA indicated that desert species had significantly greater relative bullar CS than mesic species ( $F = 35.88$ ,  $p = 0.001$ ; Fig. 7a). PGLS indicated that relative bullar CS was significantly negatively correlated with the mean annual precipitation (Coef = -0.036,  $R^2 = 0.107$ ,  $p = 0.0005$ ; Fig. 7c) and mean annual temperature (Coef = -0.001,  $R^2 = 0.065$ ,  $p = 0.0059$ ; Fig. 7d), indicating that relative bullar size was greater in environments that receive

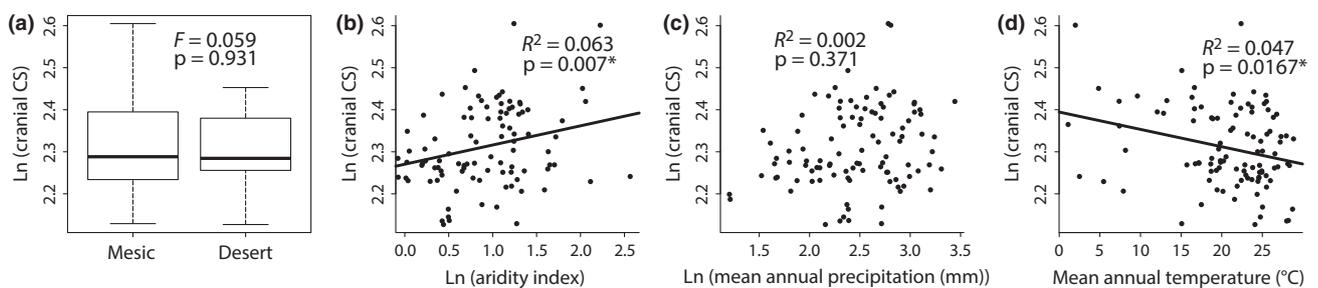


Fig. 5. Association between log cranial CS and (a) habitat, (b) aridity index, (c) mean annual precipitation and (d) mean annual temperature. The association between morphological variables with the binary habitat was tested using PHYANOVA, whereas the association between morphological variables and the three continuous environmental variables were tested using phylogenetic generalized least squares (PGLS); both on the grafted phylogeny. A best-fit line is shown in significant regressions.  $R^2$  values, as in the text, are adjusted values.

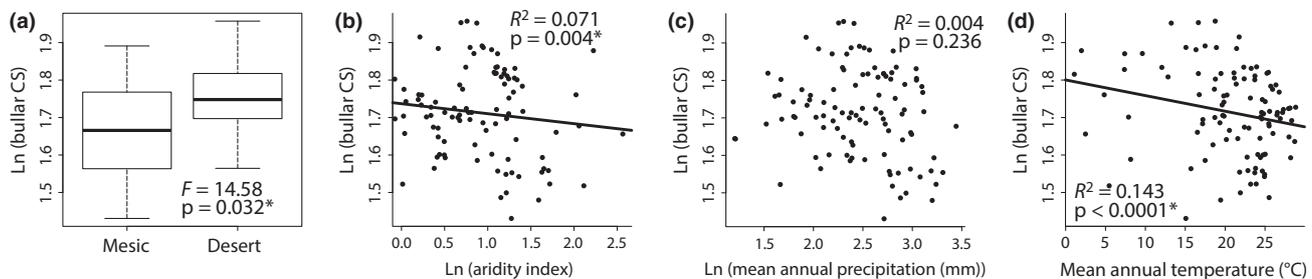


Fig. 6. Association between log bullar CS and (a) habitat, (b) aridity index, (c) mean annual precipitation and (d) mean annual temperature. See Fig. 5 legend for information.

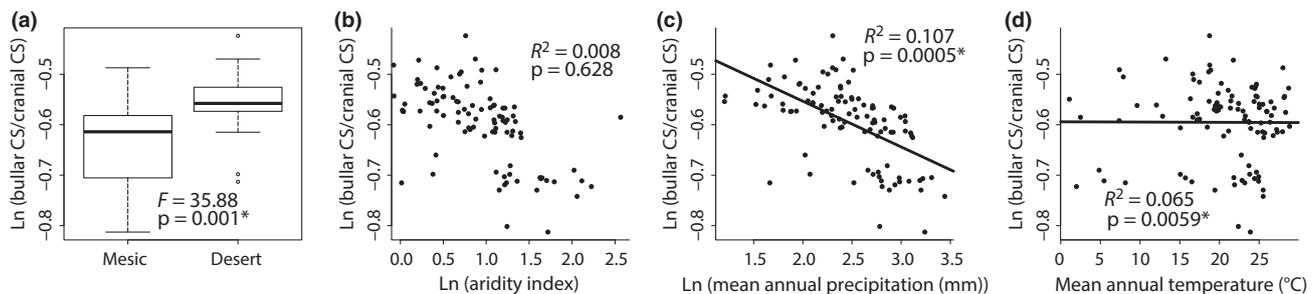


Fig. 7. Association between relative bullar CS (log (bullar CS/cranial CS)) and (a) habitat, (b) aridity index, (c) mean annual precipitation and (d) mean annual temperature. In (d), the line is horizontal despite the significant p value because the trend line is non-phylogenetic, whereas the p value is from the comparative analysis. See Fig. 5 legend for information.

less rainfall and are colder. PGLS indicated that relative bullar CS was not significantly correlated with aridity index scores (Coef = 0.003,  $R^2 = 0.008$ ,  $p = 0.628$ ; Fig. 7b).

We interpreted the aforementioned results cautiously and conservatively because multiple testing may have inflated type I errors.

## Discussion

### Gerbillinae and Deomyinae systematics

For Gerbillinae, our results are highly congruent with previous molecular phylogenies (e.g. Chevret and Dobigny 2005; Steppan et al. 2005; Schenk et al. 2013). However, our greater sampling in species and genes has resulted in increased confidence and detail. At the same time, the results are strongly incongruent with phylogenies and classifications based on morphology (e.g. Tong 1989; Pavlinov et al. 1990), with both traditional tribes Taterillini and Gerbillini being reciprocally paraphyletic. Even within Taterillini, the two subtribes Gerbillurina and Taterillina were also paraphyletic (Fig. 3; Figure S3). Furthermore, we found two to three of the largest genera to be paraphyletic (*Gerbillus*, *Gerbilliscus* and possibly *Meriones*) and five paraphyletic subgenera (within *Dipodillus*, *Gerbillurus* and *Meriones*). Given the concordance between molecular phylogenies based on allozymes (Benzazzou 1984), DNA/DNA hybridization (Chevret 1994), mtDNA sequences (Chevret and Dobigny 2005; Ito et al. 2010) and combined mtDNA and nuDNA sequences (Steppan et al. 2005; Schenk et al. 2013; this study), major revisions are needed within all taxonomic levels (tribal, subtribal, generic and subgeneric) to the current morphologically based taxonomy of gerbils. We therefore support the taxonomic divisions of Chevret and Dobigny (2005) that we follow for the rest of the discussion, and propose formalization.

We found three major groups for gerbils in our phylogeny, corresponding to Chevret and Dobigny's (2005) tribes: (1) a

basal lineage (their tribe I) consisting of the two monotypic genera *Pachyuromys* + *Desmodilliscus*; the name Desmodilliscini (Pavlinov, 1982), would apply; Desmodilliscini is sister to two clades that consist of (2) (their tribe II) *Tatera* + *Desmodillus* + *Gerbilliscus* + *Gerbillurus* and (3) (their tribe III) consisting of *Taterillus* + *Rhombomys* + *Brachiones* + *Psammomys* + *Meriones* + *Sekeetamys* + *Gerbillus* + *Dipodillus* (Chevret and Dobigny 2005 did not sample *Brachiones*). The available names for these two more diverse clades are Taterini (new rank, name proposed by Chevret and Dobigny (2005), referring to the most recent common ancestor of *Tatera* and *Gerbilliscus* and all of its descendants) and Gerbillini (Gray 1825), respectively. The status of *Ammodillus* and *Microdillus* remains uncertain. Furthermore, we recovered the same two clades within our Gerbillini that Chevret and Dobigny (2005) found: (their subtribe a) *Sekeetamys* + *Gerbillus* + *Dipodillus* (they considered *Dipodillus* as a synonym of *Gerbillus*) and (their subtribe b) *Psammomys* + *Rhombomys* + *Brachiones* + *Meriones*. We find that the closest relative of *Brachiones* within the sample is *Rhombomys* and together they form a relatively strongly supported clade (PP, 0.95; BS, 76%). Many of the subtribes proposed by Chevret and Dobigny (2005) and others would be monotypic and therefore largely redundant, and so we do not recommend any names for subtribes here.

We find strong support for the monophyly of a clade that consists of all *Gerbilliscus* and *Gerbillurus* (PP, 1.0; BS, 100%) that is consistent with previous morphological (Pavlinov et al. 1990; Pavlinov 2001) and molecular studies (Qumsiyeh et al. 1991; Chevret 1994; Chevret and Dobigny 2005). However, relationships among the three principal clades, *Gerbilliscus robustus* + *G. vicinus* + *G. phillipsi* + *G. nigricaudus*, *Gerbillurus*, and all other *Gerbilliscus* species are unsettled, with ML and BI trees yielding conflicting, well-supported resolutions, but both containing a paraphyletic *Gerbilliscus*. Chevret and Dobigny (2005) considered *Gerbillurus* to be a synonym of *Gerbilliscus*.

Alternatively, *Gerbillurus* could be retained as a genus by splitting *Gerbilliscus* into two genera. Because the type species of *Gerbilliscus* (*Gerbilliscus*) – *G. boehmi* – has not been sequenced, we recommend that taxonomic resolution wait for a comprehensive morphological and molecular analysis that includes *G. boehmi*.

Our results also indicate that *Dipodillus* is nested within *Gerbillus*; however, unlike the situation with *Gerbilliscus*, all analyses show strong support for *Gerbillus* (*Gerbillus*) being more closely related to *Dipodillus* than to *Gerbillus* (*Hendecapleura*) (PP, 1.00; BS, 97%). Here again, *Dipodillus* could be returned to a subgenus of *Gerbillus*. Given the morphological diversity in this group, and the monophyly of the subgenera, we recommend elevating *Hendecapleura* to its own genus; either solution is consistent with the phylogeny. The situation within *Meriones* is more tenuous because support for paraphyly is weak and limited to ML and the BEAST analyses, and further, the two polytypic subgenera (*Parameriones* and *Pallasiomys*) are clearly paraphyletic with respect to each other. Therefore, the monophyly of *Meriones* and consequently whether or not to synonymize *Brachiones* + *Rhombomys* + *Psammomys* with *Meriones* or split *M. tamariscinus* cannot yet be determined.

The data strongly support the monophyly of Deomyinae (PP, 1.0; BS, 100%) as a clade that consists of the two diverse monophyletic sister genera *Acomys* and *Lophuromys* and two monotypic genera *Uranomys* and *Deomys* as in previous molecular phylogenies based on DNA/DNA hybridization (Denys et al. 1995), mtDNA cytochrome *b* sequences (Verheyen et al. 1995), nuDNA VWF and/or LCAT sequences (Michaux and Catzeflis 2000; Michaux et al. 2001), a combination of DNA/DNA hybridization, mtDNA and nuDNA sequences (Chevret et al. 2001), and nuDNA BRCA1, GHR, IRBP and RAG1 (Schenk et al. 2013). The monophyly of Deomyinae and its isolation from Murinae is also supported by some morphological traits including pelage texture, palatal and molar occlusal patterns, and reproductive biology (Musser and Carleton 2005). Unlike Gerbillinae, the present results support the monophyly of all of the deomyine genera and all of the sampled deomyine subgenera. Replicating the findings from Barome et al. (2001) who analysed the same CYTB data, we find two groups of species that might need some synonymization: (1) *A. minous GroupA* + *A. nesiotus* + *A. cahirinus* + *A. minous GroupB* + *A. cilicicus* and (2) *A. chudeaui* + *A. airensis*, both consisting of sequences connected by very short branches. The former grouping was noted by Barome et al. (2001) and also is concordant with evidence of near-identical karyotypes (Zima et al. 2001) and breeding experiments (Frynta and Sadlova 1998), as discussed in Musser and Carleton (2005), indicating the need for further sampling and suggesting possible synonymy.

### Divergence and biogeography of gerbils and deomyines

In a previous study, Chevret and Dobigny (2005) found that gerbils split from deomyines at around 17 Myr before the present. We estimated a similar date, and the credibility interval of the gerbil–deomyine split was reduced to 15.9–17.6 Myr before the present (Fig. 4). The oldest available fossils of *Gerbillus* and *Meriones* are both 2 Myr old (Chevret and Dobigny 2005). These fossils are consistent with our tree, when *Dipodillus* is included in *Gerbillus* and *Brachiones* + *Rhombomys* + *Psammomys* is included in *Meriones* (4.5–6.5 and 4.8–7.2, respectively; Fig. 4).

The biogeographic reconstruction was in agreement with the one conducted for muroid rodents in general, in Schenk et al. (2013) which indicates that Africa was the point of origin for the

clade that includes gerbils, deomyines and lophiomyines (Fig. 3; Figure S6). The biogeographic reconstruction was also unambiguous in detecting the first transition from Africa to Eurasia occurring in a gerbil clade that consists of *Sekeetamys* + *Gerbillus* + *Dipodillus* + *Meriones* + *Psammomys* + *Rhombomys* + *Brachiones* (Fig. 3) at 6.9–9.5 Myr before the present (Fig. 4); apart from an independent colonization of Eurasia by *Pachyuromys duprasi* and *Tatera indica*, and the possible recolonization by *Gerbillus andersoni*, this was the only colonization of Eurasia in gerbils. The second potential colonization of Eurasia in the aforementioned clade occurred in a deomyine clade that consists of eight species of *Acomys* (Fig. 3) at 1.4–2.6 Myr before the present (Fig. 4); apart from an independent colonization of Eurasia by *Acomys russatus*, this was the only colonization of Eurasia in deomyines. The rest of the gerbils and deomyines retained their ancestral distribution in Africa, or recolonized Africa from Eurasia.

The LTT plot did not reveal any extraordinary bursts in diversification in our tree in general or coincident with any of the biogeographic transitions (Figure S7).

### Desert adaptation of the tympanic bulla

Gerbillinae is the most diverse subfamily of mammals inhabiting the Palaearctic Desert, which extends from Asia to Africa, with over 46 species in 8 genera being endemic to this arid region (Lay 1972). Chevret and Dobigny (2005) suggested that the discordance between morphological and molecular phylogenies in Gerbillinae is a result of convergence as a consequence of strong selection in arid environments, including in traits located at the posterior part of the skull, which along with teeth characters were the primary basis of gerbil classification. For example, features of the middle ear were considered derived character states in the morphological phylogeny of Tong (1989) and used to determine the relationships among genera. Even earlier, subgeneric relationships within gerbils have been based on tympanic bulla morphology in addition to dental characters (e.g. Petter 1973). These traits have been used to group genera (e.g. *Gerbillus* and *Gerbillurus*) that previous molecular phylogenies and this study find to be part of divergent lineages. Lay (1972) argued that the increase in bullar hypertrophy in gerbils, which he also found to be associated with arid environments, is a useful adaptation for predator avoidance in open habitats. Pavlinov and Rogovin (2000) similarly argued that tympanic bulla hypertrophy in several specialized desert rodents was useful in increasing auditory sensitivity at low frequencies. However, these studies did not apply phylogenetic correction and therefore may have inaccurately estimated the strength of the association between bullar morphology and aridity when testing for desert adaptation.

We support the hypothesis that convergent adaptation to aridity led to discordance between molecular and morphological phylogenies in the morphology of the tympanic bulla. This conclusion was partially anticipated by Pavlinov (2008) who argued that bullar hypertrophy evolved twice based on differences in the compartmentalization, once in his Taterillini (*Gerbillus* and *Desmodillia*) and once in his Gerbillini (*Pachyuromys*), although in our tree *Gerbillus* and *Desmodillia* are not sister taxa. The results indicate that even after accounting for phylogenetic relationships, there is a trend of more arid and warmer environments to be associated with significantly smaller cranial size (Fig. 5), whereas both absolute and relative size of the tympanic bulla are significantly larger in more arid and in colder environments (Fig. 6, although the correlation is weak). The ancestral state reconstruction analysis suggests that enlargement of the bullae occurred multiple times and to varying degrees (the

greatest hypertrophy being in the *Pachyuromys* + *Desmodilliscus* clade, within *Gerbillurus*, *Desmodillus*, and within *Meriones*; Figures S4 and S5). There does not appear to be a discrete qualitative 'shift' in bullar hypertrophy within gerbils, although there does appear to have been a shift to moderately enlarged bullae on the branch leading to Gerbillinae (Figure S5). Interestingly, the species cluster with the most hypertrophied bullae do not reside in the most arid deserts (compared with the rest of the species in or sample).

Considering that morphology of the tympanic bulla is widely used in morphological phylogenies of gerbils (Tong 1989; Pavlinov et al. 1990; Pavlinov 2008) that serve as the traditional basis for the taxonomy of the group, convergence in this trait along with other cranial structures that increase fitness in deserts would explain part of the discordance between molecular and morphological phylogenies and the frequent paraphyly at tribal, subtribal, generic and subgeneric levels. Although we focus on the bulla, other cranial characters common in traditional morphological phylogenies of rodents are also expected to experience convergent adaptation to aridity (e.g. Alhajer 2014). Rodents might be especially prone to adaptive convergence because of their small size, short dispersal distances and short generation times, enabling them to rapidly respond to environmental pressures (Samuels 2009).

The bony auditory bulla (including both tympanic and mastoid portions) is part of the middle ear and surrounds three small bones that function in sound transmission by transferring sound wave energy from the membrane of the tympanum (eardrum) to the oval window (Pavlinov 2008). Bullar hypertrophy in desert rodents is a consequence of pneumatization (formation air cavities) leading to increased volume of this structure as well as the occasional formation of new bony structures in the mastoid (Pavlinov 2008). Overall bullar hypertrophy with limited shape variation is a consequence of changes in shape and/or size of the tympanic membrane and the malleus (hammer-shaped middle-ear bone that transmits sound vibrations from the eardrum to the incus), whereas the mastoid bulla's function is to reinforce the bullar walls and semicircular canals (Lay 1972; Pavlinov 1988, 2008; Pavlinov et al. 1990). This implies that bullar hypertrophy is correlated with the enlargement of the eardrum and/or the malleus, which in turn function in sound amplification. The modification of the external bulla morphology provides the middle ear transmitting structures with the needed ear volume to function as well as to fine-tune sound wave energy transfer in the middle ear through the emergence of new bony elements that aid in sound detection, all of which are useful to sound wave transmission in arid air (Pavlinov 2008).

Our confirmation of an association between bullar hypertrophy and environmental condition, rejecting the null hypothesis that the association was a consequence of phylogenetic relatedness, highlights the need for further study of the function and ecological importance of this unusual feature of the understudied auditory system.

## Acknowledgements

We are grateful to all museum curators, collectors and collection managers for tissue loans used in this project and access to their collections: American Museum of Natural History (Darrin Lunde and Eileen Westwig); California Academy of Sciences (John Dumbacher and Maureen Flannery); Carnegie Museum of Natural History (Sue McLaren and John R. Wible); Field Museum of Natural History (Lawrence R. Heaney, William Stanley, John Phelps, Julian C. Kerbis-Peterhans and Bruce D. Patterson); Laboratorio de Citogenética Mamíferos, Universidad de Chile (Ángel Spotorno and Laura Walker); Museum of Vertebrate Zoology, Berkeley (James L. Patton and Chris Conroy); Texas Tech University

(Jorge Salazar-Bravo); United States National Museum of Natural History (Kristofer Helgen, Darrin Lunde, Helen Kafka and Michael D. Carleton); University of Florida Museum of Natural History (Candace McCaffery and David Reed); and Ron Adkins (University of Tennessee, Memphis). We particularly thank Darcy Ogada at the National Museum of Kenya and Julian Kerbis-Peterhans for the *Lophiomys* tissue. J. Schenk provided helpful input regarding research design. The manuscript benefitted from constructive comments by Andea Cardini and two anonymous reviewers. Financial support for this work was provided by a doctoral dissertation fellowship from Kuwait University to BHA (to the Florida State University) and a grant from the National Science Foundation to SJS (DEB-0841447).

## Data archive

GenBank, TreeBASE and Supporting Information.

## References

- Abiadh A, Chetoui M, Lamine-Cheniti T, Capanna E, Colangelo P (2010) Molecular phylogenetics of the genus *Gerbillus* (Rodentia, Gerbillinae): implications for systematics, taxonomy and chromosomal evolution. *Mol Phylogenet Evol* **56**:513–518.
- Adams DC, Collyer ML (2015) Permutation tests for phylogenetic comparative analyses of high-dimensional shape data: what you shuffle matters. *Evolution* **69**:823–829.
- Adams DC, Otárola-Castillo E (2013) Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol Evol* **4**:393–399.
- Adkins RM, Gelke EL, Rowe D, Honeycutt RL (2001) Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. *Mol Biol Evol* **18**:777–791.
- Agulnik SI, Silver LM (1996) The Cairo spiny mouse *Acomys cahirinus* shows a strong affinity to the Mongolian gerbil *Meriones unguiculatus*. *Mol Biol Evol* **13**:3–6.
- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* **19**:716–723.
- Alhajer BH (2014) Adaptation, diversification, and desert ecology of the most diverse order of mammals (Mammalia, Rodentia). Ph.D. dissertation, Florida State University, Tallahassee, FL.
- Arroyo MTK, Chacon P, Cavieres L (2006) Relationship between seed bank expression, adult longevity and aridity in species of *Chaetanthera* (Asteraceae) in central Chile. *Ann Bot* **98**:591–600.
- Baltas E (2007) Spatial distribution of climatic indices in northern Greece. *Meteorol Appl* **14**:69–78.
- Barome PO, Lymberakis P, Monnerot M, Gautun J-C (2001) Cytochrome b sequences reveal *Acomys minous* (Rodentia, Muridae) paraphyly and answer the question about the ancestral karyotype of *Acomys dimidiatus*. *Mol Phylogenet Evol* **18**:37–46.
- Benazzou T (1984) Contribution à l'étude chromosomique et de la diversification biochimique des Gerbillidés (Rongeurs). Ph.D. Thesis, Université de Paris XI, Paris, France.
- Bonhomme F, Iskandar D, Thaler L, Pette F (1985) Electromorphs and phylogeny in muroid rodents. In: Lockett WP, Hartenberger JL (eds), *Evolutionary Relationships Among Rodents*. Plenum, New York, pp 671–683.
- Böselt I, Römppler H, Hermsdorf T, Thor D, Busch W, Schulz A, Schöneberg T (2009) Involvement of the V2 vasopressin receptor in adaptation to limited water supply. *PLoS One* **4**:e5573.
- Cardini A, Elton S (2007) Sample size and sampling error in geometric morphometric studies of size and shape. *Zoomorphology* **126**:121–134.
- Cardini A, Seetah K, Barker G (2015) How many specimens do I need? Sampling error in geometric morphometrics: testing the sensitivity of means and variances in simple randomized selection experiments. *Zoomorphology* 1–15.
- Carleton M, Musser G (1984) Muroid rodents. In: Anderson S, Jones J (eds), *Orders and Families of Recent Mammals of the World*. John Wiley and Sons, New York, pp 289–379.
- Carleton M, Musser G (2005) Order Rodentia. Volume 2 of *Mammal Species of the World: a Taxonomic and Geographic Reference*, 3rd edn. Johns Hopkins University Press, Baltimore, MD.

- Chevret P (1994) Etude évolutive des Murinae (Rongeurs, Mammifères) africains par hybridation ADN/ADN: comparaisons avec les approches morphologiques et paléontologiques. Ph.D. Thesis, Université de Montpellier II, France.
- Chevret P, Dobigny G (2005) Systematics and evolution of the subfamily Gerbillinae (Mammalia, Rodentia, Muridae). *Mol Phylogenet Evol* **35**:674–688.
- Chevret P, Denys C, Jaeger JJ, Michaux J, Catzeflis FM (1993) Molecular evidence that the spiny mouse (*Acomys*) is more closely related to gerbils (Gerbillinae) than to true mice (Murinae). *Proc Natl Acad Sci* **90**:3433–3436.
- Chevret P, Catzeflis F, Michaux JR (2001) Acomyinae: new molecular evidences for a muroid taxon (Rodentia: Muridae). In: Denys C, Granjon L, Poulet A (eds), *African Small Mammals*. IRD Editions, Paris, pp 114–125.
- Claude J (2008) *Morphometrics with R*. Springer, New York.
- Colangelo P, Granjon L, Taylor PJ, Corti M (2007) Evolutionary systematics in African gerbilline rodents of the genus *Gerbilliscus*: inference from mitochondrial genes. *Mol Phylogenet Evol* **42**:797–806.
- Colangelo P, Castiglia R, Franchini P, Solano E (2010) Pattern of shape variation in the eastern African gerbils of the genus *Gerbilliscus* (Rodentia, Muridae): environmental correlations and implication for taxonomy and systematics. *Mamm Biol* **75**:302–310.
- de Martonne E (1942) Nouvelle carte mondiale de l'indice d'aridité. *Ann Geogr* **51**:242–250.
- DeBry RW, Seshadri S (2001) Nuclear intron sequences for phylogenetics of closely related mammals: an example using the phylogeny of *Mus*. *J Mammal* **82**:280–288.
- Denys C, Michaux J, Catzeflis F, Ducrocq S, Chevret P (1995) Morphological and molecular data against the monophyly of Dendromurinae (Muridae: Rodentia). *Bonn Zool Beitr* **45**:173–190.
- Drummond A, Rambaut A (2007) BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**:1–8.
- Eckert AJ, van Heerwaarden J, Wegrzyn JL, Nelson CD, Ross-Ibarra J, González-Martínez SC, Neale DB (2010) Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). *Genetics* **185**:969–982.
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**:1792–1797.
- Freckleton RP, Harvey PH, Pagel M (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *Am Nat* **160**:712–726.
- Frynta D, Sadlova J (1998) Hybridisation experiments in spiny-mice from the Eastern Mediterranean: systematic position of *Acomys cilicicus* revisited. *Z Säuget*, **63**(Suppl):18.
- Garland T, Dickerman AW, Janis CM, Jones JA (1993) Phylogenetic analysis of covariance by computer simulation. *Syst Biol* **42**:265–292.
- GBIF (2013) GBIF data portal. Available at <http://www.gbif.net/>.
- Gray JE (1825) Outline of an attempt at the disposition of the Mammalia into tribes and families with a list of the genera apparently appertaining to each tribe. *Ann Philos ns Ser 2* **10**:337–344.
- Hänni C, Laudet V, Barriel V, Catzeflis FM (1995) Evolutionary relationships of *Acomys* and other Murids (Rodentia, Mammalia) based on complete 12S rRNA mitochondrial DNA sequences. *Isr J Zool* **41**:131–146.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Clim* **25**:1965–1978.
- Hijmans RJ, Guarino L, Mather P (2012) DIVA-GIS Version 7.5. A geographic information system for the analysis of species distribution data. Available at <http://www.diva-gis.org/>.
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: bayesian inference of phylogenetic trees. *Bioinformatics* **17**:754–755.
- Ito M, Jiang W, Sato JJ, Zhen Q, Jiao W, Goto K, Sato H, Ishiwata K, Oku Y, Chai JJ, Kamiya H (2010) Molecular phylogeny of the subfamily Gerbillinae (Muridae, Rodentia) with emphasis on species living in the Xinjiang-Uygur autonomous region of China and based on the mitochondrial cytochrome b and cytochrome c oxidase subunit II genes. *Zoolog Sci* **27**:269–278.
- IUCN (2013) IUCN red list of threatened species. Version 2013.1. Available at <http://www.iucnredlist.org/>.
- Jansa S, Voss R (2000) Phylogenetic studies on didelphid marsupials I. Introduction and preliminary results from nuclear IRBP gene sequences. *J Mamm Evol* **7**:43–77.
- Jansa S, Weksler M (2004) Phylogeny of muroid rodents: relationships within and among major lineages as determined by IRBP gene sequences. *Mol Phylogenet Evol* **31**:256–276.
- Lay DM (1972) The anatomy, physiology, functional significance and evolution of specialized hearing organs of Gerbillinae rodents. *J Morphol* **138**:41–120.
- Maddison WP, Maddison DR (2010) Mesquite v2.75. A modular system for evolutionary analysis. Available at <http://mesquiteproject.org/>.
- Maliva R, Missimer T (2012) Aridity and drought. In: Maliva R, Missimer T (Eds.), *Arid Lands Water Evaluation and Management*. Springer, Berlin/Heidelberg, New York, pp. 21–39.
- Marshall CR (1994) Confidence intervals on stratigraphic ranges; partial relaxation of the assumption of randomly distributed fossil horizons. *Paleobiology* **20**:459–469.
- Martin Y, Gerlach G, Schlotterer C, Meyer A (2000) Molecular phylogeny of European muroid rodents based on complete cytochrome b sequences. *Mol Phylogenet Evol* **16**:37–47.
- de Martonne E (1927) Regions of interior basin drainage. *Geogr Rev* **17**:397–414.
- Michaux J, Catzeflis F (2000) The bushlike radiation of muroid rodents is exemplified by the molecular phylogeny of the LCAT nuclear gene. *Mol Phylogenet Evol* **17**:280–293.
- Michaux J, Reyes A, Catzeflis F (2001) Evolutionary history of the most speciose mammals: molecular phylogeny of muroid rodents. *Mol Biol Evol* **18**:2017–2031.
- Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES science gateway for inference of large phylogenetic trees. *Gatew Comput Environ Work* 1–8.
- Momtazi F, Darvish J, Ghassemzadeh F, Moghimi A (2008) Elliptic Fourier analysis on the tympanic bullae in three *Meriones* species (Rodentia, Mammalia): its application in biosystematics. *Acta Zool Cracoviensia Ser A Vertebr* **51**:49–58.
- Morrison DA (2007) Increasing the efficiency of searches for the maximum likelihood tree in a phylogenetic analysis of up to 150 nucleotide sequences. *Syst Biol* **56**:988–1010.
- Musser GG, Carleton MD (2005) Superfamily Muroidea. In: Wilson DE, Reeder DM (eds), *Mammal Species of the World*, 3rd edn. The Johns Hopkins University Press, Baltimore, MD, pp 894–1531.
- Ndiaye A, Hima K, Dobigny G, Sow A, Dalecky A, Ba K, Thiam M (2014) Integrative study of a poorly known Sahelian rodent species, *Gerbillus nancillus* (Rodentia, Gerbillinae). *Zool Scr* **41**:11–28.
- Nowak RM (1999) *Walker's Mammals of the World*, 6th edn. The Johns Hopkins University Press, Baltimore, MD.
- Orme D, Freckleton R, Thomas G, Petzoldt T, Fritz S, Isaac N, Pearse W (2013) CAPER: Comparative analyses of phylogenetics and evolution in R. R Package v0.5.2. Available at <http://cran.r-project.org/web/packages/caper/>.
- Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**:289–290.
- Pavlinov IY (1988) Evolution of the mastoid portion of auditory bulla of the specialized desert rodents. *Zool Zhurn* **67**:739–750.
- Pavlinov IY (2001) Current concepts of Gerbillid phylogeny and classification. In: Denys C, Granjon L, Poulet A (eds), *African Small Mammals*. Editions IRD, Paris, pp 141–150.
- Pavlinov IY (2008) A review of phylogeny and classification of Gerbillinae (Mammalia: Rodentia). *Zoologicheskoe Issledovanie* **9**:1–68.
- Pavlinov IY, Rogovin KA (2000) Relation between size of pinna and of auditory bulla in specialized desert rodents. *J Gen Biol* **61**:87–101.
- Pavlinov IY, Dubrovsky YA, Rossolimo OL, Potapova EG (1990) *Gerbils of the World*. Nauka, Moscow.
- PDB (2013) The Paleobiology Database. Available at <http://fossilworks.org/>.
- Petter F (1973) Tendances évolutives dans le genre *Gerbillus* (Rodentia, Gerbillidés). *Mammalia* **37**:631–636.
- Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**:817–818.
- Qumsiyeh MB, Hamilton MJ, Dempster ER, Baker RJ (1991) Cytogenetics and systematics of the rodent genus *Gerbillurus*. *J. Mammal*. **72**:89–96.

- R Development Core Team (2013) R: A language and environment for statistical computing. Available at <http://cran.r-project.org/>.
- Rambaut A, Drummond AJ (2005) Tracer v1.5. Available at <http://beast.bio.ed.ac.uk/Tracer/>.
- Revell LJ (2012) Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* **3**:217–223.
- Robertson RA, Shadle AR (1954) Osteologic criteria of age in beavers. *J Mammal* **35**:197–203.
- Rohlf FJ (2010) TpsDig. Version 2.16. Department of Ecology and Evolution, State University of New York, Stony Brook. Available at <http://life.bio.sunysb.edu/morph/>.
- Rohlf FJ, Slice D (1990) Extensions of the procrustes method for the optimal superimposition of landmarks. *Syst Biol* **39**:40–59.
- Rowe KC, Reno ML, Richmond DM, Adkins RM, Steppan SJ (2008) Pliocene colonization and adaptive radiations in Australia and New Guinea (Sahul): multilocus systematics of the old endemic rodents (Muroidea: Murinae). *Mol Phylogenet Evol* **47**:84–101.
- Sambrook E, Fritsch F, Maniatis T (1989) *Molecular Cloning*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Samuels JX (2009) Cranial morphology and dietary habits of rodents. *Zool J Linn Soc* **156**:864–888.
- Schenk JJ, Rowe KC, Steppan SJ (2013) Ecological opportunity and incumbency in the diversification of repeated continental colonizations by murid rodents. *Syst Biol* **62**:837–864.
- Shenbrot GI, Krasnov BR, Rogovin KA (1999) *Spatial Ecology of Desert Rodent Communities*. Springer-Verlag, Berlin, Heidelberg.
- Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**:2688–2690.
- Steppan SJ (1997) Phylogenetic analysis of phenotypic covariance structure. I. Contrasting results from matrix correlation and common principal component analysis. *Evolution* **51**:571–586.
- Steppan S, Adkins R, Anderson J (2004a) Phylogeny and divergence-date estimates of rapid radiations in murid rodents based on multiple nuclear genes. *Syst Biol* **53**:533–553.
- Steppan SJ, Storz BL, Hoffmann RS (2004b) Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from c-myc and RAG1. *Mol Phylogenet Evol* **30**:703–719.
- Steppan SJ, Adkins RM, Spinks PQ, Hale C (2005) Multigene phylogeny of the Old World mice, Murinae, reveals distinct geographic lineages and the declining utility of mitochondrial genes compared to nuclear genes. *Mol Phylogenet Evol* **37**:370–388.
- Thomas O (1888) On a new and interesting annectant genus of Muridae, with remarks on the relations of the Old- and New-World members of the family. *Proc Zool Soc London* **1888**:130–134.
- Thomas O (1896) On the genera of rodents: an attempt to bring up to date the current arrangement of the order. *Proceedings of the Zoological Society of London*: 1012–1028.
- Tong H (1989) Origine et évolution des Gerbillidae (Mammalia, Rodentia) en Afrique du Nord. *Mém Soc Géol Fr* **155**:1–120.
- Verheyen E, Colyn M, Verheyen W (1995) The phylogeny of some African muroids (Rodentia) based upon partial mitochondrial cytochrome b gene sequences. *Belg J Zool* **125**:403–407.
- Viegas-Péquignot E, Petit D, Benazzou T, Prod'homme M, Lombard M, Hoffschir F (1986) Phylogénie chromosomique chez les Sciuridae, Gerbillidae et Muridae, et étude d'espèces appartenant à d'autres familles de Rongeurs. *Mammalia* **50**:164–202.
- Walton K (1969) *The Arid Zones*. Transaction Publishers, Piscataway, NJ.
- Webster D, Webster M (1975) Auditory systems of Heteromyidae: functional morphology and evolution of the middle ear. *J Morphol* **146**:343–376.
- Weksler M (2003) Phylogeny of Neotropical oryzomyine rodents (Muridae: Sigmodontinae) based on the nuclear IRBP exon. *Mol Phylogenet Evol* **29**:331–349.
- Yazdi FT, Adriaens D (2013) Cranial variation in *Meriones tristrami* (Rodentia: Muridae: Gerbillinae) and its morphological comparison with *Meriones persicus*, *Meriones vinogradovi* and *Meriones libycus*: a geometric morphometric study. *J Zool Syst Evol Res* **51**:239–251.
- Zambakas J (1992) *General Climatology*. Department of Geology, National & Kapodistrian University of Athens, Athens, Greece.
- Zelditch ML, Swiderski DL, Sheets HD, Fink WL (2004) *Geometric Morphometrics for Biologists: a Primer*. Elsevier Academic Press, London.
- Zima J, Macholan M, Pialek J, Slivkova L, Suchomelova E (2001) Chromosomal banding pattern in the Cyprus spiny mouse, *Acomys nesiotis*. *Folia Zool* **48**:149–152.

**Appendix 1 Sequences used in phylogenetic analyses. GenBank accession numbers are provided. Accession numbers in bold are newly acquired in this study**

	12S	16S	ACP5	AVPR2	BZRP	BRCA1	COI	COII	CYTB	GHR	IRBP	RAG1	VWF
<i>Acomys airensis</i> Thomas and Hinton, 1921									AJ012021				
<i>Acomys cahirinus</i> (E. Geoffroy, 1803)	HQ652130								FJ415482	FN984742	FN984743		FN984745
<i>Acomys chudeaui</i> Kollman, 1911									FJ415488				
<i>Acomys cilicicus</i> Spitzenberger, 1978									FJ415481				
<i>Acomys dimidiatus</i> (Cretzschmar, 1826)									Z96062				
<i>Acomys ignitus</i> Dollman, 1910						AY295008	DQ019086	DQ019086	JN247674	AY294923	KC953348	AY294951	
<i>Acomys johannis</i> Thomas, 1912									AJ010567				
<i>Acomys minous</i> Bate, 1906									AJ233951				
<i>Acomys minous</i> GroupA									GU046553				
<i>Acomys minous</i> Bate, 1906													
<i>Acomys minous</i> GroupB													
<i>Acomys nesiotis</i> Bate, 1903									AJ233952				
<i>Acomys ngurui</i> Verheyen et al., 2011									JX244274				
<i>Acomys perisvalli</i> Dollman, 1911			<b>KR088975</b>						<b>KR089012</b>				
<i>Acomys russatus</i> (Wagner, 1840)				FJ411185					Z96066	FMI62071	FMI62053		FMI62065
<i>Acomys spinosissimus</i> Peters, 1852									<b>KR089013</b>		AY326074		
<i>Acomys subspinosus</i> (Waterhouse, 1838)									AJ010557				
<i>Acomys wilsoni</i> Thomas, 1892								U18832					
<i>Apodemus agrarius</i> (Pallas, 1771)	AJ311130					EU349658	DQ019092	DQ019092	EU349733	DQ019054	AB096842	DQ023472	AB303284
<i>Batomys granti</i> Thomas, 1895									AY324459	AY294917	EU349838	AY241461	
<i>Brachiones przewalskii</i> (Büchler, 1889)									AB381915				

Appendix 1. (continued)

	12S	16S	ACP5	AVPR2	BZRP	BRCA1	COI	COII	CYTB	GHR	IRBP	RAG1	VWF
<i>Cricetus cricetus</i> (Linnaeus, 1758)	AY997834	AJ633740	<b>KR088976</b>			KC953168			<b>KR089014</b>	KC953253	AY277410	KC953488	AM000051
<i>Deomys ferrugineus</i> Thomas, 1888	AJ250350				<b>KR089003</b>	AY295007			EU349745	AY294922	KC953373	AY241460	AJ402716
<i>Desmodilliscus braueri</i> Wetstein, 1916	AJ851259								AJ851273		FN357289		
<i>Desmodilliscus auricularis</i> (Smith, 1834)	AJ851257		<b>KR088977</b>	FJ411197	<b>KR089004</b>	KC953171	DQ019081	DQ019081	<b>KR089015</b>	DQ019048	KC953374	KC953494	
<i>Dipodillus campestris</i> (Loche, 1867)	AJ851243		<b>KR088978</b>		<b>KR089005</b>				<b>KR089016</b>	<b>KR089038</b>		<b>KR089045</b>	
<i>Dipodillus dasyurus</i> (Wagner, 1842)			<b>KR088979</b>	FJ411203					<b>KR089017</b>	<b>KR089039</b>	FM162054	<b>KR089046</b>	FM162066
<i>Dipodillus rupicola</i> Granjon, Aniskin, Volobouev & Sicard, 2002									KF496284				
<i>Dipodillus simoni</i> Lataste, 1881			<b>KR088980</b>		<b>KR089006</b>				GU356579	<b>KR089040</b>		<b>KR089047</b>	
<i>Gerbilliscus affra</i> (Gray, 1830)	AJ430552								AJ430560				
<i>Gerbilliscus bramtsii</i> (Smith, 1836)		AM409236							AM409392				
<i>Gerbilliscus gambianus</i> (Thomas, 1910)	X84391	AM409230							AM409386				AJ402698
<i>Gerbilliscus guineae</i> (Thomas, 1910)	AJ430551	AM409223							AM409379				
<i>Gerbilliscus kempii</i> (Wroughton, 1906)		AM409227							AM409384				
<i>Gerbilliscus leucogaster</i> (Peters, 1852)	AJ851241	AJ878527	<b>KR088981</b>						AJ851260		<b>KR089043</b>		
<i>Gerbilliscus nigricaudus</i> (Peters, 1878)		AJ878518							T50217				
<i>Gerbilliscus phillipsi</i> (de Winton, 1898)						JF716030			JF704146				
<i>Gerbilliscus robustus</i> (Cretzschmar, 1826)		AM409222	<b>KR088982</b>		<b>KR089007</b>	AY295005	DQ019084	DQ019084	<b>KR089018</b>	<b>KR089041</b>	AY326113	KC953587	
<i>Gerbilliscus validus</i> (Bocage, 1890)		Z83919							Z96044				
<i>Gerbilliscus vicinus</i> (Peter, 1878)						JF716057			JF704165				
<i>Gerbilliscus paebe</i> (A. Smith, 1836)	AJ430591				<b>KR089008</b>				<b>KR089019</b>	KC953261	KC953376	KC953500	
<i>Gerbilliscus setzeri</i> (Schlitter, 1973)	AJ430592		<b>KR088983</b>						<b>KR089020</b>	<b>KR089042</b>	<b>KR089044</b>	<b>KR089048</b>	

Appendix 1. (continued)

	12S	16S	ACP5	AVPR2	BZRP	BRCA1	COI	COII	CYTB	GHR	IRBP	RAG1	VWF
<i>Gerbillurus tytonis</i> (Bauer & Niethammer, 1960)	AJ430593		<b>KR088984</b>		<b>KR089009</b>				<b>KR089021</b>				
<i>Gerbillurus vulliamus</i> (Thomas, 1918)	U67291		<b>KR088985</b>			EU349643	GVU62574		<b>KR089022</b> KF496286	AF332022	KC953377	AY294948	
<i>Gerbillus andersoni</i> de Winton, 1902													
<i>Gerbillus gerbillus</i> (Olivier, 1801)	AJ851242		<b>KR088986</b>	FJ411204		EU349700	DQ019082	DQ019082	<b>KR089023</b> KF496235	DQ019049	EU349846	DQ023452	AJ402699
<i>Gerbillus henleyi</i> (de Winton, 1903)						JF716023							
<i>Gerbillus hesperinus</i> Cabrera, 1936									JN652803				
<i>Gerbillus hoogstrali</i> Lay, 1975									JN021420				
<i>Gerbillus latastei</i> Thomas & Trouessart, 1903									GU356550				
<i>Gerbillus nancillus</i> Thomas & Hinton, 1923									KF496279				
<i>Gerbillus nanus</i> Blanford, 1875	AJ851244		<b>KR088987</b>						<b>KR089024</b>	KC953262	KC953378	KC953501	
<i>Gerbillus nigeriae</i> Thomas & Hinton, 1920	X84381	AF141257							AJ430555		AM408333		
<i>Gerbillus occidentalis</i> Lay, 1975													
<i>Gerbillus perpallidus</i> Setzer, 1958									KF496282				
<i>Gerbillus poecilops</i> Yerbury & Thomas, 1895									JN652806				
<i>Gerbillus pyramidum</i> Geoffroy, 1825									JQ753064				
<i>Gerbillus tarabuli</i> Thomas, 1902	AJ851245								KF496283				
<i>Lophiomys inhausi</i> Milne-Edwards, 1867			<b>KR088988</b>						GU356571				
<i>Lophuromys brevicaudus</i> Osgood, 1936		AY828276							<b>KR089025</b>		KC953389	KC953514	
<i>Lophuromys chrysopus</i> Osgood, 1936		AY828304							AY828232				

Appendix 1. (continued)

	12S	16S	ACP5	AVPR2	BZRP	BRCA1	COI	COII	CYTB	GHR	IRBP	RAG1	VWF
<i>Lophuromys flavopunctatus</i> Thomas, 1888	U67294					AY295006	DQ019087	DQ019087	EU349754	AY294921	AY326091	AY294950	
<i>Lophuromys melanonyx</i> F. Petter, 1972		AY828279							AY828235				
<i>Lophuromys sikapusi</i> (Temminck, 1853)	AJ250349		<b>KR088989</b>						<b>KR089026</b>	KC953271	KC953390	KC953515	AJ402717
<i>Lophuromys woosnami</i> Thomas, 1906									DQ902807				
<i>Lophuromys zena</i> Dollman, 1909							AB381912		<b>KR089027</b>	KC953272	KC953391	KC953516	
<i>Meriones chengi</i> Wang, 1964									AB381900				
<i>Meriones crassus</i> Sundevall, 1842	AJ851251		<b>KR088990</b>						<b>KR089028</b>				
<i>Meriones libycus</i> Lichtenstein, 1823	AJ851250						AB381913		<b>KR089029</b>				
<i>Meriones meridianus</i> (Pallas, 1773)	AJ851252			FJ411223			AB381908		AJ851268				
<i>Meriones persicus</i> (Blanford, 1875)			<b>KR088991</b>						<b>KR089030</b>				
<i>Meriones rex</i> Yerbury & Thomas, 1895	AJ851248								AJ851265				
<i>Meriones shawi</i> (Duvernoy, 1842)				FJ411225		AY295004	DQ019083	DQ019083	<b>KR089031</b>	AF332021	KC953400	AY294947	
<i>Meriones tamariscinus</i> Pallas, 1773							AB381916		AB381904				
<i>Meriones tristrami</i> Thomas, 1892			<b>KR088992</b>						<b>KR089032</b>				
<i>Meriones unguiculatus</i> (Milne-Edwards, 1867)	AJ851249		<b>KR088993</b>	FJ411228			AB177842		<b>KR089033</b>	AF247184	AY326095		FN984757
<i>Microtus chrotorrhinus</i> (Miller, 1894)									AF163893	AM392383	AM919403		
<i>Mus pahari</i> Thomas, 1916	AB125793		EU349629						EU349767	KC953280	EU349864	EU349906	AB285474
<i>Ondatra zibethicus</i> (Linnaeus, 1766)			<b>KR088994</b>	FJ411232		AY295011			KC563206	AY294925	KC953427	AY294953	
<i>Pachyuromys duprasi</i> Lataste, 1880	AJ851258			FJ411234					AJ851274				
<i>Phloeomys sp.</i> <i>Phyllotis xanthopygus</i> (Waterhouse, 1837)			DQ023451 <b>KR088995</b>			EU349644 KC953208			DQ191484 <b>KR089034</b>	DQ019070 KC953314	AY326103 AY163632	DQ023480 KC953561	

Appendix 1. (continued)

	12S	16S	ACP5	AVPR2	BZRP	BRCA1	COI	COII	CYTB	GHR	IRBP	RAG1	VWF
<i>Psammomys obesus</i> Cretzschmar, 1828	AJ851253		<b>KR088996</b>		<b>KR089010</b>				<b>KR089035</b>	KC953314	FN357290	FN357290	
<i>Psammomys vexillaris</i> Thomas, 1925									AY934541				
<i>Rattus tiomanicus</i> (Miller, 1900)			<b>KR088997</b>						HM217391	KC953320	KC953449	KC953568	
<i>Rheomys thomasi</i> Dickey, 1928									<b>KR089036</b>	KC960491	KC953451	<b>KR089049</b>	
<i>Rhombomys opimus</i> (Lichtenstein, 1823)	AJ430590								AJ430556				
<i>Sekeetamys calurus</i> (Thomas, 1892)	AJ851246		<b>KR088998</b>						AJ851276				
<i>Sigmodon hispidus</i> Say & Ord, 1825	X89788		<b>KR088999</b>			AY295016			AF425200	AF540641	AY277479	AY241465	
<i>Tatera indica</i> (Hardwicke, 1807)	AJ430553	AM409239							AJ430563				
<i>Taterillus arenarius</i> Robbins, 1974	AJ851254								AJ851261		FN357288		
<i>Taterillus emini</i> (Thomas, 1892)		Z83921	<b>KR089000</b>		<b>KR089011</b>	KC953224	DQ019085	DQ019085	<b>KR089037</b>	DQ019050	KC953461	DQ023453	
<i>Taterillus gracilis</i> (Thomas, 1892)	AJ851256	AM409238							AJ851263				
<i>Taterillus petteri</i> Gautun, Tranier & Sicard, 1985)						JF716025							
<i>Taterillus pygargus</i> (F. Cuvier, 1838)	AJ851255								AJ851262				
<i>Tokudaia osimensis</i> (Abe, 1934)	AJ311133		EU349640			EU349659			AB033703	EU349828	EU349878	EU349918	
<i>Uranomys ruddi</i> Dollman, 1909	X84388	Z83922	<b>KR089001</b>			EU349642	DQ019088	DQ019088	HM635858	DQ019051	EU360812	DQ023454	AJ402714

## Appendix 2 List of specimens sequenced and localities

### Cricetidae

Arvicolinae: *Ondatra zibethicus* NA, Locality unknown.

Cricetinae: *Cricetus cricetus* MVZ 155880, Austria, 1 km NE Guntramsdorf.

Sigmodontinae: *Phyllotis xanthopygus* AK 13014, Argentina; *Rheomys thomasi* CN 101294, El Salvador, Ahuachapan, El Imposible, San Francisco Menendez; *Sigmodon hispidus* CN, Locality unknown.

### Muridae

Gerbillinae: *Desmodillus auricularis* RA 1, Namibia, Kanabeam, Karasburg District, 375 m. 28°07'17"S 17°33'32"E; *Dipodillus campestris* TK 40900, Tunisia, Mournagia Subgov't, Jebel Ain Es Seed; *Dipodillus dasyurus* TK 25570, Jordan, Azraq Oasis, lava hills S Azraq Ed Druz; *Dipodillus simoni* TK 40877, Tunisia, Gabes Subgov't, Gar'at El Mekky; *Gerbilliscus leucogaster* CAS 28615, Namibia, Hardap Region, Namib Rand Nature Reserve, Keerweder Farm, 24°56'37.7154"S, 16°4'4.0794"E; *Gerbilliscus robustus* FMNH 158105, Tanzania, Babati District, Tarangire National Park, near Engelhardt Bridge; *G. robustus* FMNH 151229, Tanzania, Kilimanjaro District, Gonja Forest Reserve, near Higililu River; *Gerbillurus paeba* RA 49, Namibia, Kanabeam, Karasburg District, 375 m. 28°07'17"S 17°33'32"E; *Gerbillurus setzeri* TK 32329, Namibia, Approximately 5 km N Gobabeb; *Gerbillurus tytonis* RA 08, Namibia, Kanabeam, Karasburg District, 375 m. 28°07'17"S 17°33'32"E; *Gerbillurus vallinus* H 675, Locality unknown; *Gerbillus gerbillus* CM 113822, Egypt, Giza Governorate, 50 km SW Giza (by road) on El Faiyum Rd. 29 42N, 30 58E; *Gerbillus gerbillus* CM 113823, Egypt, Giza Governorate, 50 km SW Giza (by road) on El Faiyum Rd. 29 42N, 30 58E; *Gerbillus nanus* MVZ 192048, Iran, Agricultural Station near Zahek, SE of Zabol; *Meriones crassus* TK 25637, Egypt, S Sinai, near El Tor; *Meriones libycus* MVZ 191969, Iran, 8 km SSW Kerman; *Meriones persicus* MVZ 191974, Iran, Zar Rud Bala, Aabshar-e Rayen, Kuhehazar, W of Rayen; *Meriones shawi* H 583, Locality unknown; *Meriones tristrami* TK 25568, Jordan, 7 Mi E Irbid; *Meriones unguiculatus* TK 20358, Pet trade; *Psammomys obesus* TK 40892, Tunisia, El Guettar Subgov't, Jebel Ank Mine, Gafsa Gov't; *Sekeetamys calurus* RA, Locality unknown; *Taterillus emini* CM 102330, Kenya, Eastern Province, Machakos District, Kathekani, 760 m, 02 37S, 38 09E.

Deomyinae: *Acomys perisvalli* H 767, Locality unknown; *Acomys spinosissimus* FMNH 196233, Malawi, 1.6 km SE (by air) Chinunka; *Deomys ferrugineus* FMNH 160403, Uganda, Kigezi, Bufumbira, Nteko Parish, edge of Bwindi-Impenetrable NP; *Lophuromys sikapusi* FMNH 137803, Uganda, Masaka, Bugala

Island, Kalangala, 1.0 km N, 0.5 km E; *Lophuromys sikapusi* FMNH 137798, Uganda, Bugala Island, Kalangala, 1.0 km N, 0.5 km E; *Lophuromys zena* FMNH 190640, Kenya, Nyeri Dist., Aberdare Range, 28.5 km W, 4.9 km N Nyeri; *Uranomys ruddi* CM 113723, Ghana, Greater Accra Region, Shai Hills Game Production Reserve Headquarter, 05 53N, 00 03E.

Lophiomyinae: *Lophiomyys imhausi* DO 111, Kenya.

Murinae: *Batomys granti* USNM 458948, Philippines, Luzon Island, Mt. Isarog, 1750 m; *Rattus tiomanicus* USNM 590331, Malaysia, Sarawak, Bintulu Division, Ulu Kakas, Bukit Sarang.

Abbreviations and acronyms (for institutions and collectors) are as follows. CAS, California Academy of Sciences; CM, Carnegie Museum of Natural History; FMNH, Field Museum of Natural History; MVZ, Museum of Vertebrate Zoology; TK, Museum of Texas Tech University; USNM, National Museum of Natural History; RA, Ron Adkins; NA, not available. Unlisted abbreviations are internal reference numbers.

### Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Data S1.** Summary of incongruities between gene trees and the concatenated data set tree.

**Figure S1.** Bayesian chronograms used for phylogenetic comparative analyses.

**Figure S2.** Maximum-likelihood phylograms of the 13 loci reconstructed with the GTR+ $\Gamma$ +I substitution model.

**Figure S3.** ML phylogram of the concatenated data set.

**Figure S4.** Maximum-likelihood ancestral state reconstruction of relative bulla size (bulla CS/skull CS) in the pruned tree (tree 1).

**Figure S5.** Tree that conveys the same information as Fig. S4 in the form of a 'traitgram' which is a projection of the phylogenetic tree.

**Figure S6.** Biogeographic regions used in ancestral state estimations of the biogeography of gerbils and deomyines.

**Figure S7.** Lineage-through-time plot indicating the diversification pattern observed in the chronogram in Fig. 4.

**Table S1.** Primers used in this study. For and Rev refer to forward and reverse, respectively.

**Table S2.** List of specimens digitized for morphological analyses. See Appendix 2 legend for museum abbreviations.

**Table S3.** Average of the raw bioclimatic variables encountered by species.

**Table S4.** Description and units of the bioclimatic variables from Table S2.

**Table S5.** Statistical summary of regression analyses between centroid size estimates from this study with distance based estimates of size from Alhajeri (2014).