Family Wide Molecular Adaptations to Underground Life in African Mole-Rats Revealed by Phylogenomic Analysis

Kalina T.J. Davies,*,[1] Nigel C. Bennett,[2] Georgia Tsagkogeorga,[1] Stephen J. Rossiter,*,[1] and Christopher G. Faulkes*,[1]

[1]School of Biological & Chemical Sciences, Queen Mary University of London, London, United Kingdom
[2]Department of Zoology & Entomology, University of Pretoria, Pretoria, South Africa

*Corresponding author: E-mail: kalinadavies@gmail.com; s.j.rossiter@qmul.ac.uk; c.g.faulkes@qmul.ac.uk.

Associate editor: Katja Nowick

Abstract

During their evolutionary radiation, mammals have colonized diverse habitats. Arguably the subterranean niche is the most inhospitable of these, characterized by reduced oxygen, elevated carbon dioxide, absence of light, scarcity of food, and a substrate that is energetically costly to burrow through. Of all lineages to have transitioned to a subterranean niche, African mole-rats are one of the most successful. Much of their ecological success can be attributed to a diet of plant storage organs, which has allowed them to colonize climatically varied habitats across sub-Saharan Africa, and has probably contributed to the evolution of their diverse social systems. Yet despite their many remarkable phenotypic specializations, little is known about molecular adaptations underlying these traits. To address this, we sequenced the transcriptomes of seven mole-rat taxa, including three solitary species, and combined new sequences with existing genomic data sets. Alignments of more than 13,000 protein-coding genes encompassed, for the first time, all six genera and the full spectrum of ecological and social variation in the clade. We detected positive selection within the mole-rat clade and along ancestral branches in approximately 700 genes including loci associated with tumorigenesis, aging, morphological development, and sociality. By combining these results with gene ontology annotation and protein–protein networks, we identified several clusters of functionally related genes. This family wide analysis of molecular evolution in mole-rats has identified a suite of positively selected genes, deepening our understanding of the extreme phenotypic traits exhibited by this group.

Key words: adaptive evolution, positive selection, subterranean, mole-rats.

Introduction

During their evolutionary radiation, mammals have colonized a wide range of extreme ecological niches, including the air, sea, and underground. The subterranean environment presents its own unique set of physiological challenges such as low oxygen (hypoxia), high carbon dioxide (CO_2) levels (hypercapnia), high relative humidity, and little or no exposure to sunlight (Nevo 1979; de Vries et al. 2008). Yet despite these harsh environmental conditions, several mammalian lineages have made the transition from a habitat above ground to a predominantly underground niche (Nevo 1979). Of these, one of the most successful groups is the African mole-rats (Family: Bathyergidae), comprising six genera and approximately 30 recognized species (e.g., [Van Daele et al. 2007; Faulkes and Bennett 2013], and see supplementary fig, S1, Supplementary Material online) with several other taxa yet to be described (e.g., [Ingram et al. 2004; Van Daele et al. 2007; Faulkes et al. 2010, 2011]). Similar adaptations are also seen in two other unrelated speciose mammalian groups, notably the afrotherian golden moles (family Chrysochloridae) and the laurasiatherian moles and their allies (family Talpidae) although not all members of the latter are strictly subterranean (Wilson and Reeder 2005).

African mole-rats rarely emerge above ground and show a suite of phenotypic adaptations to cope with the environmental constraints of their subterranean niche (Nevo 1979; de Vries et al. 2008). Morphologically, they possess extrabuccal incisors for digging, and elongated bodies with shortened limbs for efficient tunneling (Bennett and Faulkes 2000). In addition, as a consequence of living in complete darkness, they show reduced eyes (Nikitina and Kidson 2014) although they do possess at least some functional visual pigment genes (Zhao, Ru, et al. 2009; Emerling and Springer 2014) and can perceive light (Kott et al. 2010; Nikitina and Kidson 2014). Indeed, mole-rats exhibit some degree of circadian locomotory rhythmicity (Oosthuizen et al. 2003). Although behavioral studies suggest these rhythms are sometimes weak and highly variable both within and between species, but can be entrained with light (Riccio and Goldman 2000; Oosthuizen et al. 2003; de Vries et al. 2008). Mole-rats also show reductions in their auditory systems; they lack external ear pinnae, and show reduced auditory sensitivity and limited high frequency hearing capabilities (Heffner and Heffner 1993). Apart from vocalizations, some species also utilize seismic signals (drumming) for communication (Lacey et al. 1991; Narins et al. 1992; Yosida et al. 2007). Similarly, olfaction is hypothesized to be important for foraging; a recent study found evidence that positive selection may have acted within certain olfactory receptor families, and thus suggests olfaction may be enhanced in mole-rats (Stathopoulos et al. 2014).

Open Access

Some of the most extreme conditions that subterranean species must overcome are the predominantly hypoxic and hypercapnic conditions (Roper et al. 2001; Shams et al. 2005). In the case of mole-rats, these are likely exacerbated by the fact that many individuals can often share confined burrow systems (Brett 1991). One possible adaptation for elevated CO_2 levels in naked mole-rats appears to be their relative acid insensitivity via the nociceptor sodium channel NaV1.7 (Park et al. 2008; Smith et al. 2011). These conditions are similar to those experienced during hibernation; and this similar selection pressure may have led to convergent functional amino acid changes in NaV1.7 in other hibernating mammals (Liu et al. 2013).

Another putative adaptation to life underground is the loose and elastic skin of mole-rats that is particularly evident in the naked mole-rat (Daly and Buffenstein 1998). Loose skin in subterranean animals is hypothesized to aid movement through the substrate, but could also impact on thermoregulation and moisture loss. Recently, it has been speculated that the specialized skin of naked mole-rats may be due to the secretion of a high molecular mass form of hyaluronan (HMM-HA) (Tian et al. 2013). Furthermore, secretion of HMM-HA may have also conveyed cancer resistance to naked mole-rats via early contact inhibition (Seluanov et al. 2009; Liang et al. 2010; Tian et al. 2013). An additional possible consequence of the naked mole-rats' cancer resistance is their extreme longevity—with some individuals living greater than 30 years—approximately five times that predicted for their body size. Furthermore, naked mole-rats display negligible senescence into old age (Buffenstein and Jarvis 2002; Buffenstein 2008). For these reasons, mole-rats are recognized as an emerging model for studying aging (e.g., [Azpurua et al. 2013; Rodriguez et al. 2014]).

The ecological success of mole-rats can in part be attributed to their diet of geophytes (the storage organs of plants), which has allowed them to occupy a range of climatically varied habitats across sub-Saharan Africa ([Faulkes et al. 2004; Faulkes and Bennett 2013], supplementary fig. S1, Supplementary Material online). However, the need to find ephemeral food items might also have helped drive the cooperative breeding seen in many mole-rat species, and has led to the generation of the aridity food distribution hypothesis as an ultimate explanation for the evolution of sociality in mole-rats (as reviewed by Faulkes and Bennett [2013]). In fact social structure differs markedly across the family, from solitary genera, such as *Bathyergus*, *Georychus*, and *Heliophobius*, to highly social species, such as *Heterocephalus glaber* and *Fukomys damarensis* (Faulkes and Bennett 2013). The two latter species are often referred to as eusocial, that is, exhibiting a reproductive division of labor, overlapping generations and cooperative care of young (Bennett and Faulkes 2000). The social state of the ancestral mole-rat is unclear; however, given the mole-rat phylogeny the switch between solitary and a social/cooperative lifestyle must have occurred several times (Faulkes and Bennett 2013).

Study Aims

The morphological, physiological and life-history traits displayed by mole-rats make them a valuable group for comparative evolutionary studies. Yet, while the genomes of two naked mole-rat individuals and one Damaraland mole-rat (*F. damarensis*) have recently been sequenced (Kim et al. 2011; Fang, Seim, et al. 2014; Keane et al. 2014), very little is known regarding molecular adaptations across the family, with only a small number of candidate gene studies (e.g., [Edrey et al. 2012; Davies et al. 2014; Stathopoulos et al. 2014; Faulkes et al. 2015]). Here, to identify genes that may have been targets of selection across African mole-rats, we undertook transcriptome sequencing of seven phylogenetically divergent species. In combination with the naked mole-rat genomic data, our transcriptomes encompass all six recognized genera (Faulkes and Bennett 2013), as well as the full spectrum of social and ecological variation. We constructed gene alignments based on approximately 14,000 one-to-one human guinea pig orthologs, and searched for the presence of either positive selection along the ancestral mole-rat branch or divergent selection, that is, different nonsynonymous-to-synonymous rate ratios (ω), between the mole-rat clade and outgroup taxa. Specifically, we predicted that due to the constraints of their subterranean lifestyle, genes relating to morphological development, sensory perception, hypoxia/hypercapnia, cancer susceptibility, aging, pigmentation, and social behavior, would be under positive selection in the African mole-rats. Finally, we attempted to investigate the evolution of sociality across the mole-rat family by screening candidate "sociality" genes for episodic diversifying selection pressures, that is, to search for the presence of positive selection acting on sites along specific lineages. It was predicted that if these genes were associated with mole-rat social behavior they would show evidence of positive selection along branches that represented transitions between solitary and social life-histories.

Results

Taxonomic Representation

We sequenced RNA from seven mole-rat taxa with diverse ecologies and life-histories, in terms of social structure (i.e., eusocial, social, and solitary species), body mass, and longevity (see table 1 and fig. 1 and supplementary fig. S1, Supplementary Material online, for more details). We combined our transcriptomes with the publically available naked mole-rat genome, so that our sampling encompassed all six currently recognized bathyergid genera. For comparative purposes, we collected transcriptome sequence data from a phylogenetically divergent subterranean rodent the East African root rat (*Tachyoryctes splendens*) from the family Spalacidae.

Protein-Coding Gene Ortholog Identification

Reciprocal BLAST searches using human proteins as queries recovered an average of 11,722 candidate genes across the newly assembled transcriptomes (see table 1). Similar searches using guinea pig (*Cavia porcellus*) proteins recovered an average of 10,631 candidate genes (results not shown).

Table 1. Locality, Life-History, and Data Set Information for the Mole-Rat (Bathyergidae) and *Tachyoryctes* (Spalacidae) Species Included in This Study.

Species	Locality	Sociality (group size)	Body Mass (g)	Max. Lifespan (yr)	Number of Orthologs Recovered	Number of One-To-One Orthologs Recovered
Bathyergus suillus	Darling, Western Cape Province; SA	Solitary (1)	93	> 6	11,928	10,119
Georychus capensis	Darling, Western Cape Province; SA	Solitary (1)	181	11.2	11,084	9,384
Cryptomys h. natalensis	Glengarry; SA	Social (8)	132.5	11	12,132	10,261
Cryptomys h. pretoriae	Lynnwood Road, Pretoria; SA	Social (12)	132.5	11	12,449	10,570
Cryptomys h. mahali	Patryshoek, Pretoria; SA	Social (NA)	132.5	11	11,769	10,000
Fukomys damarensis	Black Rock, Hotazel; SA	Eusocial (41)	165.0 (M), 141.5 (F)	11.9	12,121	10,257
Heliophobius emini	Mlali; Tanzania	Solitary (1)	160	> 7.5	9,243	7,844
Heterocephalus glaber	(Keane et al. 2014)—colony established by V. Gorbunova (University of Rochester, USA); founder animals from colony of J.U. Jarvis (University of Cape Town, SA)	Eusocial (≤300)	35	31	16,625	13,741
Tachyoryctes splendens	Makundi Village, Kilimanjiro; Tanzania	Solitary (1)	220	NA	13,050	10,998

Note.—Number of orthologs recovered refers to the number of rodent transcripts recovered from reciprocal BLASTs using 23,393 human proteins; number of one-to-one orthologs recovered refers to the number of these matching the 14,873 human guinea pig one-to-one orthologs. Body mass and lifespan values were obtained from the AnAge database (Tacutu et al. 2013) and (Faulkes et al. 1997; Sherman and Jarvis 2002; Bennett and Jarvis 2004; Weigl 2005; Schmidt et al. 2013). SA, South Africa; yrs, years; g, grams; NA, not available; M, male; F, female.

Based on the higher number of hits, sequences extracted with human proteins were used for further analyses. Our novel sequences were combined with coding sequences from *H. glaber* (Keane et al. 2014), identified using the same methods as above. Candidate genes were then filtered by one-to-one human–guinea pig orthologous protein-coding genes to give multifasta files representing approximately 14,000 genes with at least one mole-rat representative (see fig. 2 for work-flow).

Phylogenomic Reconstruction

An alignment of 3,999 concatenated genes was partitioned and used to reconstruct maximum likelihood (ML) phylogenies based on either gene or both gene and codon position together. Both analyses recovered the division between mouse-like rodents (Myomorpha) and guinea pig-related rodents (Hystricomorpha) (see fig. 1 and supplementary fig. S2*a*, Supplementary Material online). This relationship was also recovered by a Bayesian analysis of a subsampled data set of 60,783 amino acids (see supplementary fig. S2*b*, Supplementary Material online). African mole-rats were recovered as monophyletic, with the naked mole-rat, *H. glaber*, and the silvery mole-rat, *H. emini* occupying more basal positions within the clade. The recently split genera of *Cryptomys* and *Fukomys* were recovered as a monophyletic group. *Georychus capensis* and *Bathyergus suillus* were recovered as sister taxa. All nodes were recovered with maximum statistical support of 100 bootstraps (see fig. 1 and supplementary fig. S2, Supplementary Material online).

Genome-Wide Screens for Changes in Selection Pressure in Mole-Rats

Detection of Positive Selection on Mole-Rat Branches

For each gene, we fitted branch-site models to our data to test for the presence of lineage-specific positive selection on either the ancestral branch leading to all mole-rats in the alignment, or, in the cases when only a single mole-rat species was present, on its terminal branch. In total, 513 genes were shown to be under significant positive selection (likelihood ratio test [LRT]: uncorrected $P \leq 0.05$) with foreground $\omega \geq 1.00$ (see supplementary table S1, Supplementary Material online). Genes under positive selection included several relating to tumor suppression and disease (e.g., *BRCA1* and *APLP2*), *GHRHR* the receptor of growth hormone releasing hormone, *TYRP1* involved in the production of melanin (Slominski et al. 2004) and *arginine vasopressin receptor 1A* (*AVPR1A*) which has documented roles in rodent social behavior and pair-bond formation (Lim et al. 2004). Following false discovery rate (FDR) correction, at 0.05, for the multiple comparisons made, we found statistical evidence of positive selection in 31 genes, with 22 of these along the ancestral branch of mole-rats and nine on a terminal mole-rat branch (see supplementary table S1, Supplementary Material online).

Detection of Divergent Selection between the Mole-Rat Clade and Outgroups

To test for the presence of divergent selection pressures acting on the mole-rats compared with other taxa, we ran clade models (Bielawski and Yang 2004). Of the genes tested with these models, 7,232 revealed evidence of significant divergent selection between mole-rats and the outgroup clade (LRT: uncorrected $P \leq 0.05$). Genes found to be under significant divergent selection between members of the mole-rat clade and outgroup taxa included approximately 50 genes associated with hypoxia and decreased oxygen levels, for example, *EGLN3*, *hypoxia-inducible factor 1-alpha* (*HIF1A*), *HIF1AN*, and *HMOX1*. In the majority of cases, these hypoxia-related genes were found to be under purifying selection in both mole-rats and outgroups (i.e., both foreground and background $\omega_2 < 1.00$). Seven thousand and fourteen of the genes under
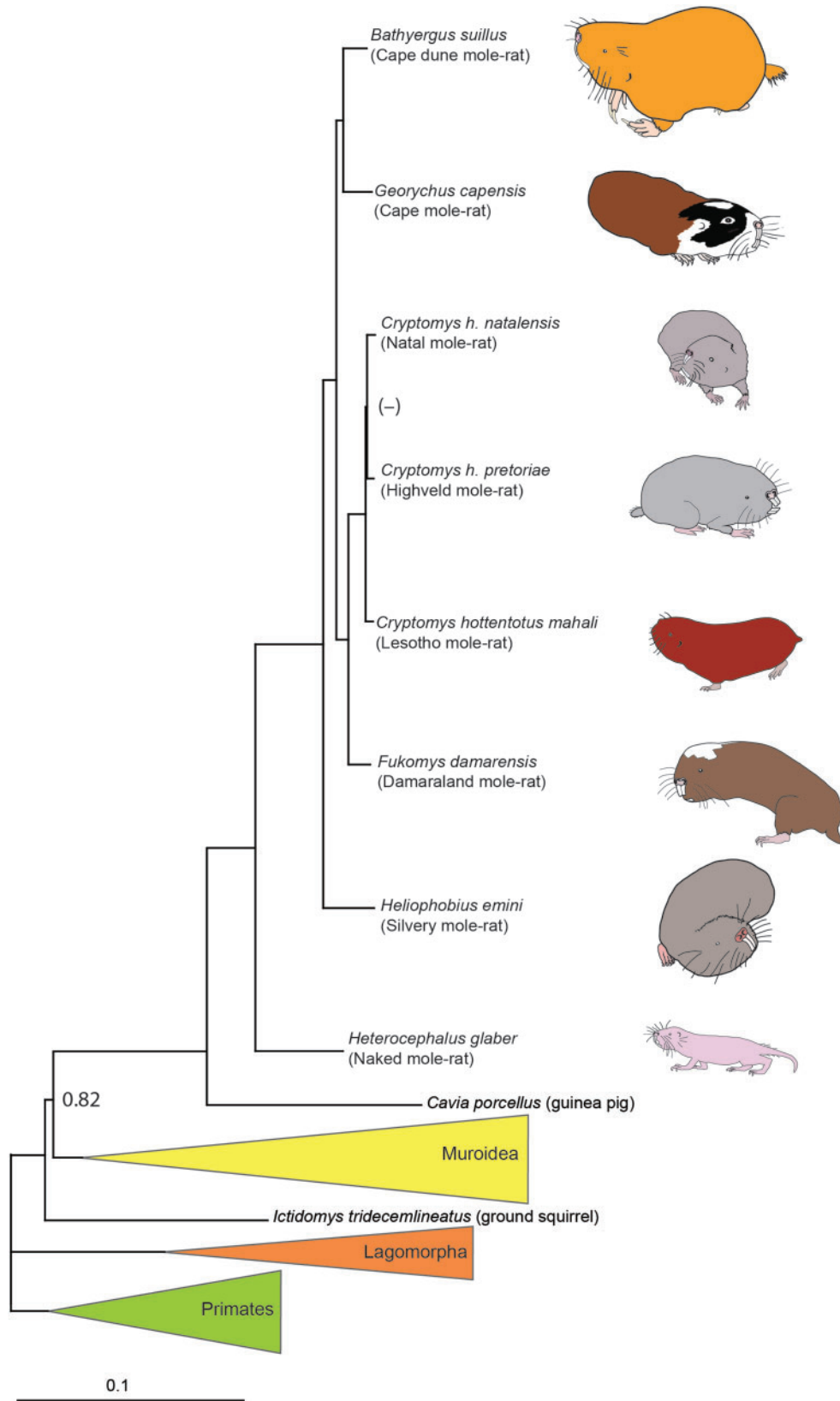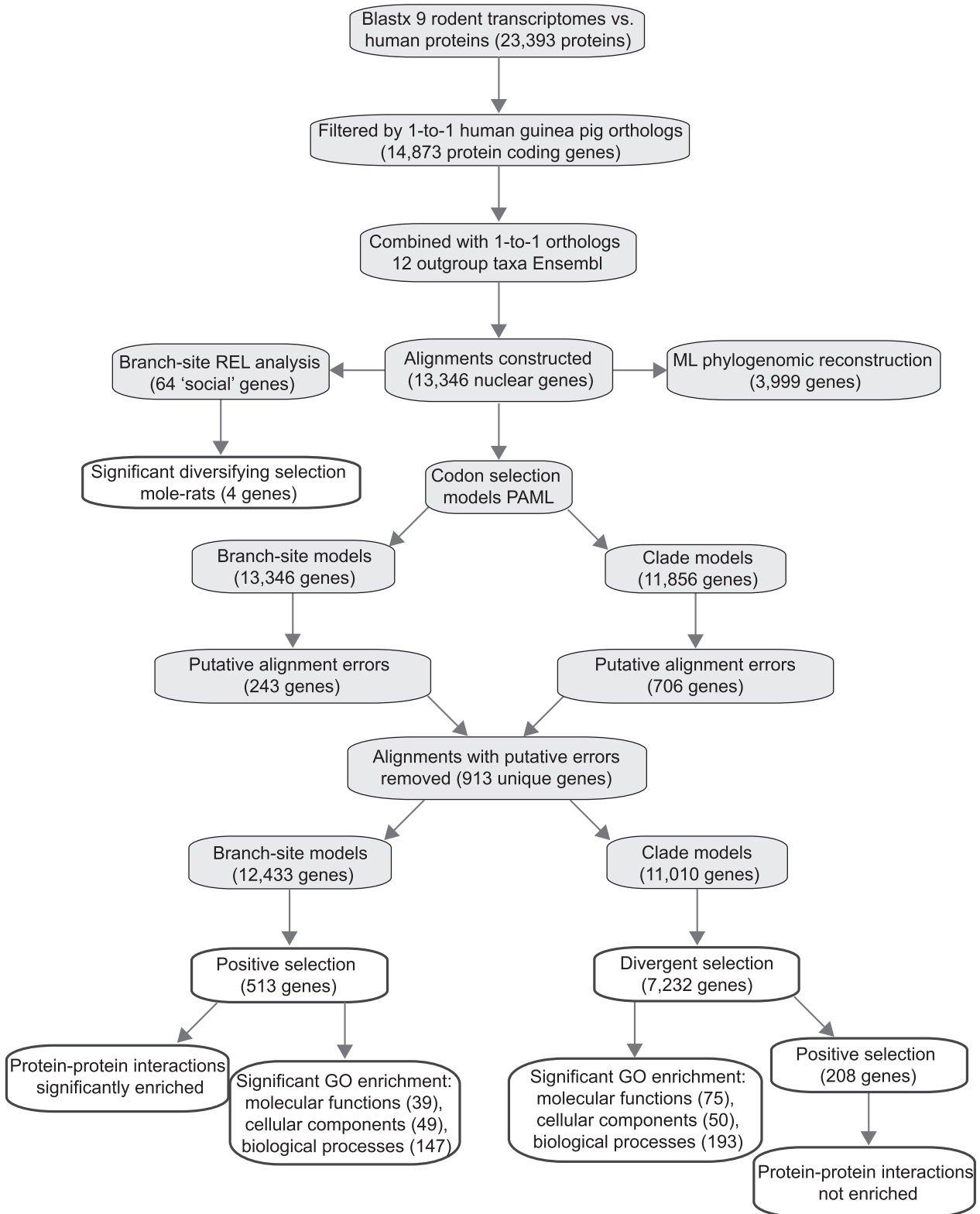
Fig. 1. ML phylogeny generated by analysis of a concatenated data set of 3,999 genes, partitioned by gene. All nodes were recovered with bootstrap support of 100%. The identical topology was recovered by analysis of the same data set partitioned by gene and codon position (first + second and third position). Bayesian analyses of the 172 most informative gene partitions (60,783 amino acids), recovered the same topology with the exception of the relationship among the three *Cryptomys hottentotus* subspecies, all nodes received a BPP support of 1.00 unless otherwise shown. Line drawings represent each mole-rat species included in the study; images are scaled to represent relative body sizes (also see table 1 for values). Outgroup taxa are represented by colored clades as follows: green, Hominidae, Hylobatidae, and Galagidae; orange, *Ochotona princeps* and *Oryctolagus cuniculus*; yellow, Muridae, Spalacidae, and Heteromyidae.

```
┌─────────────────────────────┐
│ Blastx 9 rodent transcriptomes vs. │
│ human proteins (23,393 proteins)   │
└─────────────────────────────┘
               │
               ▼
┌─────────────────────────────┐
│ Filtered by 1-to-1 human guinea pig orthologs │
│ (14,873 protein coding genes)                 │
└─────────────────────────────┘
               │
               ▼
┌─────────────────────────────┐
│ Combined with 1-to-1 orthologs │
│ 12 outgroup taxa Ensembl       │
└─────────────────────────────┘
               │
               ▼
```

Branch-site REL analysis ← Alignments constructed → ML phylogenomic reconstruction
(64 'social' genes)   (13,346 nuclear genes)   (3,999 genes)

Significant diversifying selection   Codon selection
mole-rats (4 genes)   models PAML

Branch-site models   Clade models
(13,346 genes)   (11,856 genes)

Putative alignment errors   Putative alignment errors
(243 genes)   (706 genes)

Alignments with putative errors
removed (913 unique genes)

Branch-site models   Clade models
(12,433 genes)   (11,010 genes)

Positive selection   Divergent selection
(513 genes)   (7,232 genes)

Protein-protein interactions   Significant GO enrichment: Significant GO enrichment: Positive selection
significantly enriched   molecular functions (39),   molecular functions (75),   (208 genes)
                                cellular components (49),   cellular components (50),
                                biological processes (147)   biological processes (193)
                                                                                            Protein-protein interactions
                                                                                            not enriched

divergent selection remained significant after FDR correction at 0.05. On closer inspection 208 genes (189 after FDR at 0.05) were found to be under significant positive selection in mole-rats and not the outgroup taxa, that is, foreground $\omega_2 > 1.00$ and background $\omega_2 < 1.00$ (LRT: uncorrected $P \le 0.05$; see supplementary table S2, Supplementary Material online). These genes included *APCS* which may play a role in Alzheimer's disease (Botto et al. 1997), *ATRN* which may

play a role in obesity (He et al. 2001) and *THRA*, one of the receptors of thyroid hormone.

Gene Ontology Enrichment Analyses

Gene ontology (GO) enrichment analyses performed on genes identified as either 1) under positive selection on the ancestral mole-rat branch or 2) under divergent selection in the mole-rat clade, revealed several significant patterns based on Fisher's exact test. We report results from both gene sets across the three GO domains—molecular function (MF), cellular component (CC), and biological process (BP).

GO Enrichment Analysis of Genes under Positive Selection on Mole-Rat Branches

We found 39 MF, 49 CC, and 147 BP terms to be significantly enriched (see supplementary tables S3–S5, Supplementary Material online, respectively, for full results). Enriched MF terms included several related to receptor activity (e.g., "neuropeptide receptor activity"), hormone binding (e.g., "growth factor binding"), and enzymatic activity (e.g., "metallopeptidase activity"). Enriched CC terms included a wide variety of components including "platelet alpha granule," "photoreceptor inner segment," and "chromatoid body."

Many of the BP terms found to be significantly enriched can be summarized into broad biological categories including aging, sensory perception, development of morphological structures, reproduction, and immunity (see table 2 for a summary of representative terms and genes). Of the seven genes under selection associated with aging, many encode proteins that either interact with telomerase or act as tumor suppressors (Zhou and Lu 2001; Salomoni and Pandolfi 2002; Cornet et al. 2003). Genes associated with enrichment of sensory perception terms encode proteins linked to vision and hearing (see table 2). Among the 12 vision genes identified was *RP1*, mutations in which cause autosomal dominant retinitis pigmentosa—whereas the 11 genes associated with inner ear development terms included *SLC4A7*, disruption of which leads to cochlear and retinal degeneration in mice (Lopez et al. 2005), and *MYO3A* which is associated with nonsyndromic hearing loss (DFNB30). Additional categories of interest relating to sensory perception include "mechanoreceptor differentiation," which included the *NTF3* gene – a member of the neurotrophin family that controls survival of mammalian neurons.

Enriched terms relating to morphogenesis and/or development encompassed broad functions, some of which could relate to known adaptations in mole-rats, such as tooth development. Additionally, over ten terms associated with the circulatory system and/or blood were recorded (see table 2), genes from these categories included *COL4A1*, the C-terminal NC1 domain of which is known as arresten and has regulatory roles in angiogenesis, tumor growth, and expression of *HIF1A* (Sudhakar et al. 2005).

Genes found in the reproduction categories include *ZP4*, which encodes a glycoprotein expressed in the zona pellucida and which may act as a sperm receptor, as well as *SYT6*, which is involved in acrosomal exocytosis (Michauta et al. 2001). Significantly enriched terms associated with immune response were well represented (see table 2). Three genes

with roles in the innate immune system and antibacterial activity were shown to be related to the host–symbiont interactions terms; for example, *PGLYRP4* which encodes a peptidoglycan recognition protein (Dziarski and Gupta 2010).

GO Enrichment Analysis of Genes under Divergent Selection in Mole-Rats Compared with Outgroup Taxa

A total of 75 MFs, 50 CCs, and 193 BPs were found to be significantly enriched based on Fisher's exact test ($P < 0.05$, supplementary tables S6–S8, Supplementary Material online, respectively). Enriched MF terms included "retinoic acid receptor binding" and "oxygen binding," while several CC terms of interest included the telomere cap ("telomere cap complex" and "nuclear telomere cap complex") as well as the "photoreceptor outer segment."

BP terms could be grouped into a number of functionally related categories, including telomeres, responses to stress and blood (see table 2). Positively selected genes associated with enrichment of telomere terms encode proteins with roles in telomere stability and preventing chromosome end-to-end fusions (e.g., [van Steensel et al. 1998]). Genes linked to significant stress response terms (e.g., "response to gamma radiation" and "response to xenobiotic stimulus"; also see table 2) included *GSTO1*, which encodes a protein that might play a role in the metabolism of carcinogens, and has been linked to several neurodegenerative diseases (Li et al. 2003). Genes associated with salt/osmotic stress response include *AVP* that appears to show antidiuretic activity, and *SLC2A4*, which functions as an insulin-regulated facilitative glucose transporter and is implicated in insulin resistance (Huang and Czech 2007). Approximately ten terms relating to blood coagulation were enriched (see table 2); with the associated genes including a number of serpins (e.g., *SERPINC1*, *SERPINE1*, and *SERPINE2*). Numerous metabolic and catabolic processes were shown to be significantly enriched: for example, "androgen metabolic process," "estrogen metabolic process," and "amine catabolic process" (see supplementary table S8, Supplementary Material online). Additional enriched terms of interest include "regulation of keratinocyte differentiation" and also several terms relating to the phosphatidylinositol 3-kinase cascade (see supplementary table S8, Supplementary Material online).

Visualization of Protein–Protein Interaction Networks and Functional Clusters

Molecular adaptation may involve positive selection acting on a single gene or, in some cases, multiple genes that share similar functions and/or belong to common pathways (e.g., [Kosiol et al. 2008]). Therefore, we tested for evidence of "functional" clustering among the protein products of genes found to be under positive selection (based on branch-site or clade models) using protein–protein interaction networks and enriched GO BP terms.

In total, we identified 507 protein products in the STRING database (Franceschini et al. 2013) that were encoded by the 513 positively selected genes identified with branch-site models. These proteins displayed significant enrichment in interactions compared with the background 12,316 proteins,

Table 2. Summary of Some of the Uniting Biological Themes within BP GO Terms Found to be Significantly Enriched (Fisher's exact test, $P < 0.05$).[a]

Enriched Terms Recovered from Analyses of Genes under Positive Selection in Mole-Rats Identified with Branch-Site Models

Category	Representative GO Terms	Associated Genes:
Aging [3]	"Multicellular organismal aging," "determination of adult life-span," and "regulation of telomerase activity"	CTSV, GHRHR, INPP5D, LRRK2, PARM1, PML, and PINX1
Circadian rhythm [2]	"Circadian behavior" and "rhythmic behavior"	NAGLU, HCRTR2, and NPY2R
Circulatory system [> 10]	e.g., "Angiogenesis," "artery development," "artery morphogenesis," and "cardiovascular system development"	> 50 (e.g., ANGPTL3, COL4A1, PDGFRB, RYR1, and THSD7A)
Development of mineralized structures [5]	"Odontogenesis of dentin-containing tooth," "positive regulation of osteoblast differentiation," "regulation of bone resorption," "biomineral tissue development," and "regulation of bone remodeling"	18 (e.g., AHSG, EDAR, FGF23, ODAM, and LAMA5)
Immune response [>40]	e.g., "Response to virus," "defense response to bacterium," "response to wounding," "epidermis development," "regulation of natural killer cell activation," and "regulation of lymphocyte differentiation and tumor necrosis factor production"	> 80 (e.g., CALM3, CD38, HPX, and TESPA1)
Inner ear development [2]	"Inner ear receptor cell development" and "inner ear development"	DUOX2, FZD2, FZD6,GLI2, GLI3, JAG1, MYO3A, NAGLU, PDGFRB, SLC4A7, and SPRY2
Host–Symbiont interactions [3]	"Growth of symbiont in host," "growth involved in symbiotic interaction," and "growth of symbiont involved in interaction with host"	CTSG, PGLYRP4, and TIRAP
Pigmentation [3]	"Melanocyte differentiation," "pigmentation," and "pigment cell differentiation"	EDAR, GLI3, HPS1, LYST, TYRP1, USP13, and VPS33B
Reproduction [> 10]	e.g., "Single fertilization," "sperm motility," "spermatogenesis," "acrosome reaction," and "sexual reproduction"	> 40 (e.g., AVPR1A, SYT6, TDRD5, TDRD6, and ZP4)
Skin development [3]	"Collagen fibril organization," "hemidesmosome assembly," and "epidermis development"	ABCA12, COL3A1, CST6, DDR2, DPT, EDAR, FRAS1, FOXC2, FZD6, GLI2, JAG1, KRT14, KRT25, LAMA3, LAMA5, LAMB3, LATS2, MSX2, NTF3, and RYR1
Vision [6]	"Detection of visible light," "detection of light stimulus," "phototransduction, visible light," "cellular response to light stimulus," "retinoid metabolic process," and "phototransduction"	ASIC2, APOA4, APOB, CALM3, CDC25A, CYP1A1, NFATC4, TMEM161A, RDH11, RP1, TTR, and PLB1

Enriched Terms Recovered from Analyses of Genes under Divergent Selection in Mole-Rats Identified with Clade Models

Category	Representative GO Terms:	Associated Genes:
Telomeres [2]	"Regulation of telomerase activity" and "telomere maintenance via telomerase"	ACD, GREM1, MEN1, PARM1, PIF1, PINX1, PML, POT1, PPARG, RAD50, RFC1, TERF1, TERF2, TERF2IP, TINF2, TNKS, and WRAP53
Cell division and DNA modification/repair [> 20]	e.g., "Meiotic chromosome segregation," "meiotic cell cycle," "spindle assembly involved in mitosis," "centrosome duplication," "mismatch repair," and "establishment of chromosome localization"	> 250 (e.g., BRCA1, BRCA2, MEIOB, RGCC, and XRCC2)
Immune response [> 20]	e.g., "Inflammatory response," "humoral immune response," "immunoglobulin mediated immune response," and "adaptive immune response"	> 500 (e.g., BAD, PGLYRP1, SOCS6, TNFRSF1B, and TOLLIP)
Response to stress/stimulus [> 10]	e.g., "Response to salt stress," "cellular response to osmotic stress," "cellular response to xenobiotic stimulus," and "hyperosmotic response"	> 150 (e.g., AQP4, AVP, BRCA2, REN, TP63, TRPV4, SLC12A6, and SLC2A4)
Blood coagulation [> 10]	e.g., "Negative regulation of angiogenesis," "heme metabolic process," "regulation of blood coagulation," and "fibrin clot formation"	> 100 (e.g., AGT, F11, HPSE, HMOX1, and VASH1)
Pigmentation	"Pigment metabolic process," "cellular pigmentation," "pigment granule localization," "melanosome localization," "pigment biosynthetic process," and "pigmentation"	> 75 (e.g., ATRN, MITF, MLPH, OCA2, and PMEL)

[a]Numbers in parentheses denote the number of enriched GO terms assigned to each of the broadly defined biologically related "categories."

with a total of 350 interactions observed compared with 270 expected ($P = 2.53 \times 10^{-6}$). These observed interactions involved 240 proteins, and consisted of one large subnetwork, two small subnetworks, and several isolated triplets and pairs.

Notable clusters included those centered on BRCA1, POLQ, and APOB; AVPR1A, and GHSR; COL4A1, COL4A2, and ITGA1; EBNA1BP2, NOC4L, and WDR46 (see supplementary fig. S3, Supplementary Material online). The three proteins
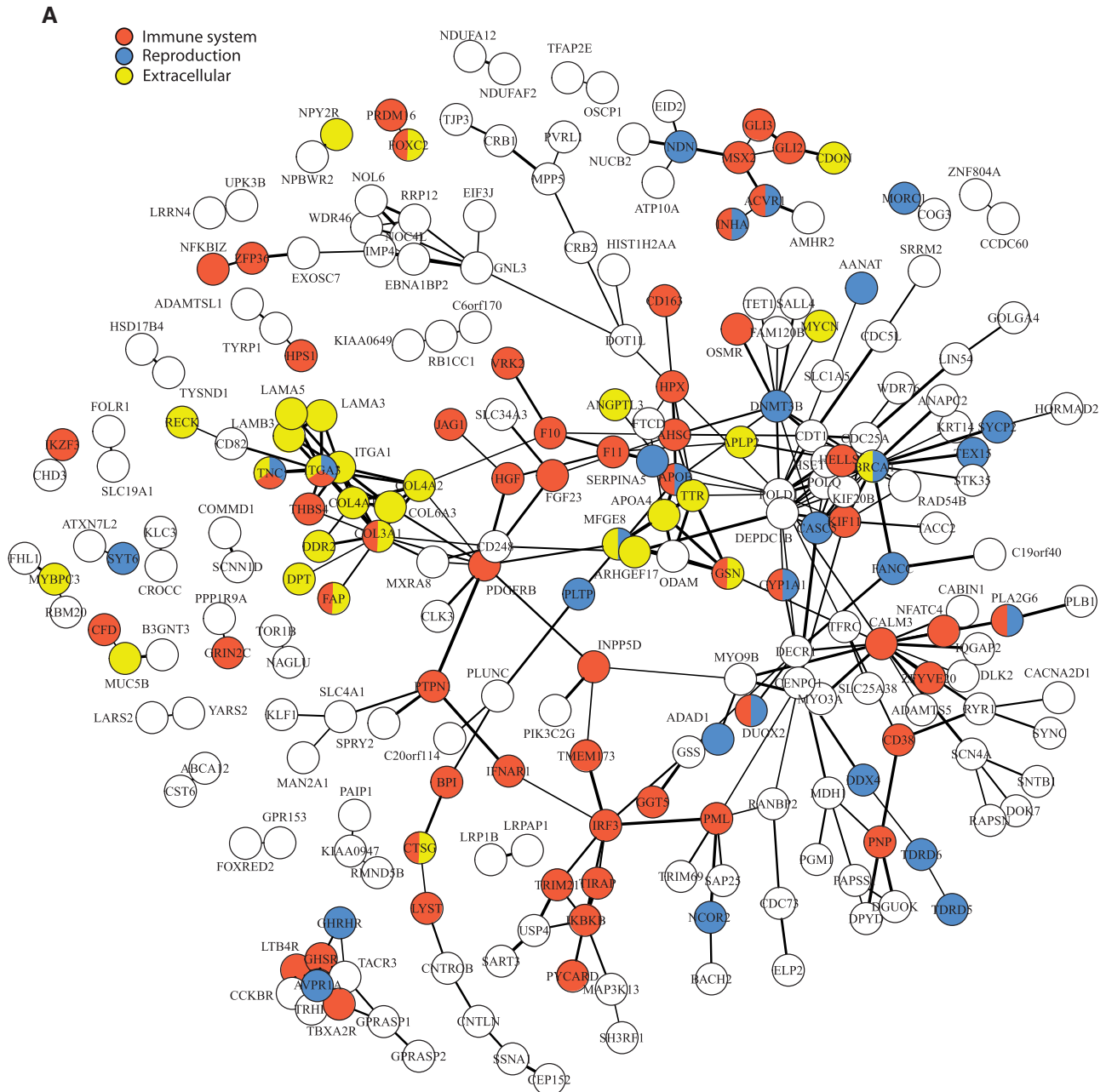
Fig. 3. Protein–protein interactions visualized as a network. Connection thickness corresponds to the confidence of the interaction, with a thicker line indicating stronger evidence. Evidence is based on combined interaction scores of the following lines of evidence: neighborhood, fusion, co-occurrence, homology, coexpression, experimental knowledge, and text mining. (A) Genes have been classified into three broad categories with nodes colored as follows: red, "immune system" (e.g., "response to virus," "regulation of natural killer cell activation"); blue, "reproduction" (e.g., "fertilization," "sexual reproduction"); and yellow, "extracellular" (e.g., "extracellular matrix disassembly," "biological adhesion"). (B) Genes have been classified into three broad categories with nodes colored as follows: red, "metabolism" (e.g., "hormone metabolic process," "fat-soluble vitamin metabolic process"); blue, "cell cycle" (e.g., "DNA methylation," "centriole replication"); and yellow, "cell" (e.g., "cell motility," "localization of cell"). (C) Genes have been classified into four broad categories with nodes colored as follows: red, "aging" (e.g., "determination of adult lifespan," "multicellular organismal aging"); blue, "morphology" (e.g., "biomineral tissue development," "anatomical structure regression"); yellow, "sensory" (e.g., "phototransduction," "inner ear receptor cell development"); and gray, "symbionts and host" (e.g., "growth involved in symbiotic interaction," "growth of symbiont in host").

with the greatest number of interactions, based on all sources of evidence, were BRCA1, POLD1, and F11 with 19, 17, and 16 interactions, respectively.

To assess levels of clustering of proteins with shared functions, we grouped the 147 significantly enriched GO BPs terms into ten functionally related categories (e.g., "sensory," "reproduction," and "metabolism") for clarity.

These categories were then mapped onto relevant proteins in the above network (see fig. 3). As might be expected given the adaptive evolution of immunity in mammals, around one-quarter of the proteins in the network were associated with GO terms concerning immunity; these included several proteins clustered around IRF3 (see fig. 3A). We also recorded clustering among approximately 30 proteins
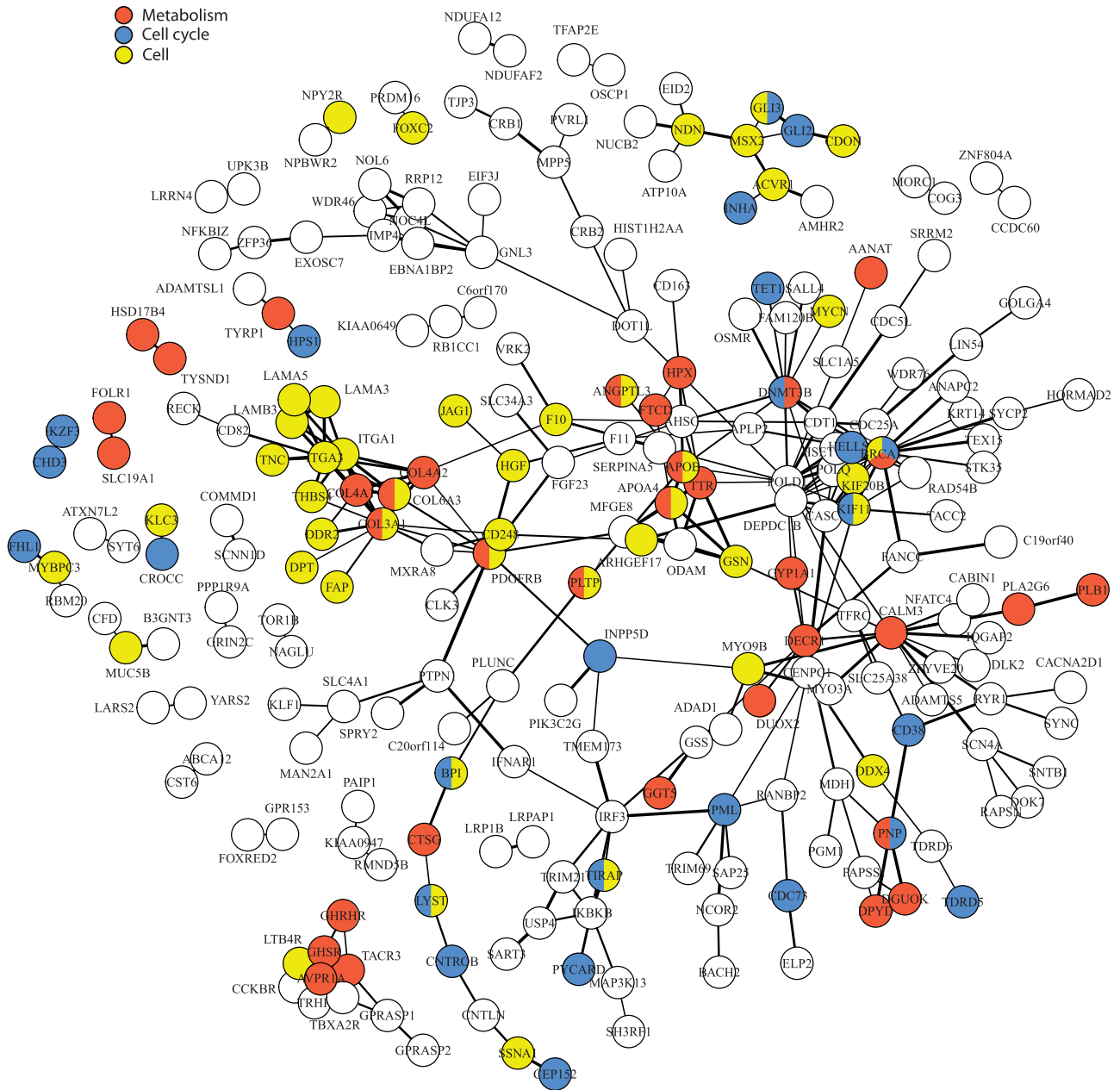
Fig. 3. Continued.

belonging to the "extracellular interactions" category (e.g., "extracellular matrix disassembly") that included multiple collagens and laminins, whereas approximately 30 proteins encoded by genes associated with reproduction did not display obvious clustering (see fig. 3A). Clustering among collagens and laminins was also seen in the category relating to "cell" (e.g., "cell motility"), however, little, if any, clustering was evident among proteins encoded by genes associated with either "metabolism" (e.g., "fat-soluble vitamin metabolic process") or "cell cycle" (e.g., "centriole replication") (see fig. 3B). Over 100 proteins, encoded by genes associated with BP related to "morphological structures" were present in the network. Many of these were found in numerous subnetworks, such as those centered on AVPR1A or MSX2

(see fig. 3C). Although approximately 40 proteins, encoded by genes related to "sensory/response to stimulus" (e.g., "phototransduction") were included in the network, little spatial association was shown. The categories "aging" and "symbionts and host" were represented by only three and two proteins, respectively, therefore, network clustering was not possible (see fig. 3C).

We also tested for interactions among the protein products of the 208 genes under positive selection in the mole-rat clade compared with the background clade. However, when these interactions were compared with those between the background 10,913 proteins, no significant enrichment was found; with a total of 49 observed interactions compared with 40.9 expected ($P = 0.119$).
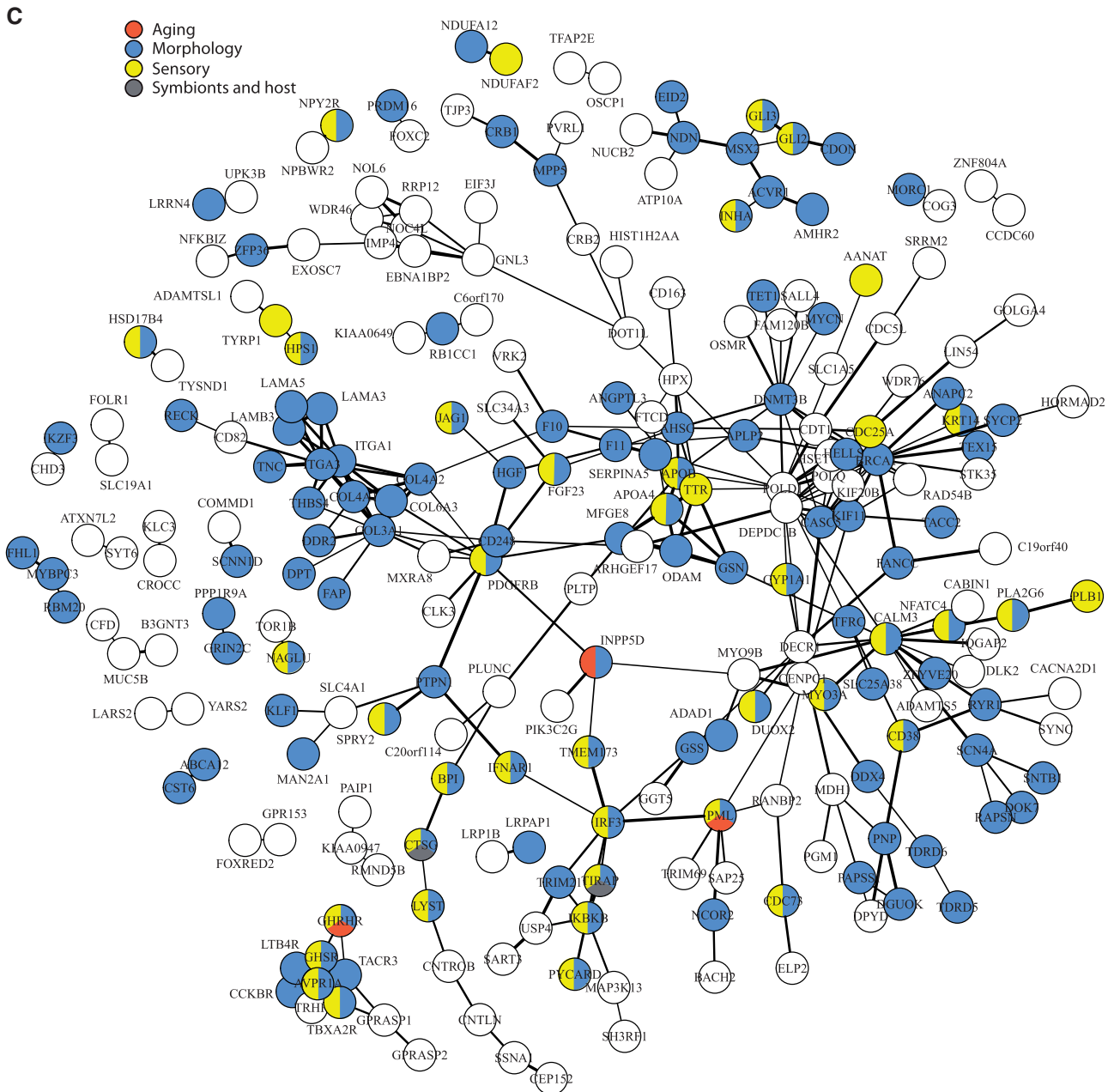
Fig. 3. Continued.

HYPHY Analysis of Genes Relating to Social Behavior

Given the mole-rat species tree, several alternative hypotheses involving multiple gains or losses of solitary/social dwelling are equally supported (see supplementary fig. S4, Supplementary Material online; [Faulkes and Bennett 2013]). We attempted to investigate the evolution of sociality across the family by screening 58 genes thought to play a role in mammalian social behavior for differential selection pressures across key branches. Tests based on 30 genes revealed branches showing evidence of episodic positive selection, reduced to 15 after correcting for multiple tests (Holm $P \leq 0.05$, see supplementary table S9, Supplementary Material online). Alignments with significant LRTs were screened by eye, and seven genes were rejected as potential false positive due to alignment/

sequencing errors. Within the mole-rats, we identified three genes (*NMUR2*, *CHRNB2*, and *IL1B*) as potentially being under positive selection along the ancestral branch, and one gene (*MECP2*) along the *F. damarensis* branch (see supplementary table S9 and fig. S5, Supplementary Material online). We also identified four genes (*DLG4*, *IRAK1*, *PPT1*, and *PTEN*) showing signatures of positive selection in outgroup Hystricognathi.

Discussion

We sequenced and assembled mixed-tissue transcriptomes from seven African mole-rat species and one spalacid rodent mole, and combined these with existing data to undertake a genome-wide study of molecular evolution spanning this highly specialized and ecologically diverse group. We

predicted that mole-rats would show Darwinian selection in genes underpinning their highly specialized morphological, physiological, and sensory adaptations. By ensuring our newly sequenced focal taxa included solitary species (*H. emini*, *B. suillus*, and *G. capensis*) in addition to the social and eusocial species, we conducted the first examination of whether mole-rats have undergone selection acting on sociality-related genes associated with the evolution of their diverse social systems.

Sensory, Morphological, and Physiological Adaptations for Life Underground

Consistent with the expectation that mole-rats would show signatures of adaptive evolution associated with shifts in sensory modality—a consequence of their transition to a niche characterized by low light and ambient noise—we found Darwinian selection acting on several "sensory perception" genes in this group. These include *RP1*, which is required for the differentiation of photoreceptor cells, and *SLC24A2*, a component of the visual transduction cascade. On the other hand, *rhodopsin*, *RHO*, was seen to be under purifying selection in the naked mole-rat, in agreement with previous results ([Zhao, Ru, et al. 2009]; but see [Fang, Seim, et al. 2014]). We found three genes associated with circadian rhythms to be under positive selection in mole-rats: *NAGLU*, deficiency of which alters circadian behaviors (Heldermon et al. 2007); *HCRTR2*, which may have roles in the regulation of feeding and sleep/wakefulness (Kilduff and Peyron 2000), and *NPY2R*—which encodes a protein with diverse activities including the modulation of circadian rhythms (Soscia and Harrington 2005) (see table 2). Although mole-rats are able to maintain circadian rhythmicity despite living in dark underground burrows of near-constant temperature (Riccio and Goldman 2000), previous work showed that the receptors of melatonin—a key regulator of circadian rhythms and body temperature (Cagnacci et al. 1992)—are nonfunctional in naked mole-rats (Kim et al. 2011). More research is needed to determine the specific roles of *NAGLU*, *HCRTR2*, and *NPY2R* in mole-rats, as well as to elucidate the molecular basis of their circadian pathways in general.

We also found several other genes associated with sensory perception under selection in mole-rats, which may help to explain their sensory adaptations (see Introduction). Hearing or deafness genes identified included *JAG1*, *HGF*, *MYO3A*, *TNC*, and *SLC4A7* (Tsai et al. 2001; Schultz et al. 2009; Zhao et al. 2013), the latter of which encodes a sodium bicarbonate cotransporter that functions in maintaining the intracellular pH of both ears and eyes (Bok et al. 2003). Such cases of selection in hearing and vision genes contributed to the observed enrichment in inner ear development and vision GO categories. Among loci involved in chemosensory perception, we found positive selection in two olfactory receptor genes (*OR13G1* and *OR5J2*) and one bitter taste receptor (*TAS2R41*), whereas divergent selection was seen in several other taste receptors including *TAS1R1* and *TAS1R3*, which together form heterodimers responsible for sensing umami

(Xiaodong et al. 2002). Previous studies have suggested olfaction and bitter perception may be enhanced in mole-rats (Hayakawa et al. 2014; Stathopoulos et al. 2014), but have disagreed about the number of taste receptors detected (Kim et al. 2011; Hayakawa et al. 2014), possibly reflecting differences in the gene annotation methods employed and highlighting the need for careful in-depth analyses when studying large complex gene families. Overall, despite our findings, coverage of sensory genes was rather limited in our alignments, and thus the inclusion of sensory organ-specific transcriptome data would increase our ability to investigate the molecular evolution of sensory genes. Data obtained from DNA sequencing as opposed to RNA-Seq would also be needed to test for gene loss-of-function, which can arise from relaxed selection and has been widely reported in sensory genes (e.g., [Zhao, Rossiter, et al. 2009; Emerling and Springer 2014]).

Cases of enrichment stemming from selection acting on related genes were also seen in terms associated with pigmentation and melanosome localization (see table 2). Although naked mole-rats lack fur, they do show countershading of the skin (Braude et al. 2001), that has been hypothesized to play roles in thermoregulation, camouflage, or protection against ultraviolet (UV) light (Braude et al. 2001). Interestingly the melanin-producing cells (melanocytes) are found within the dermis of naked mole-rats (Daly and Buffenstein 1998), whereas in other mole-rat species they occur in the epidermis. Given that the melanophores of cold-blooded vertebrates, such as reptiles and amphibians, also occur in the dermis, some have speculated that the unusual melanocyte distribution in naked mole-rats may relate to their poikilothermic-like thermoregulation (Daly and Buffenstein 1998). Given the infrequent exposure to UV light of subterranean mammals, it might be expected that genes involved in cellular resistance to UV light may be under relaxation in mole-rats, indeed, previous studies have shown that naked mole-rat-derived fibroblasts are sensitive to UV light (Salmon et al. 2008). Although we did not find enrichment in GO terms relating to a response UV, we did record positive selection in a number of genes associated with the repair of DNA damage caused by factors such as UV light, such as *CDC25A*, *NFATC4*, *PML*, *POLD1*, and *TMEM161A*. An earlier naked mole-rat genome study suggested that a single amino acid substitution at position 397 in the *Hairless* (*Hr*) gene may be associated with hairlessness in the naked mole-rat (Kim et al. 2011); however, we found the C397W substitution is shared by *Cryptomys h. mahali* and the guinea pig. It thus seems unlikely that this substitution is solely responsible for the hairlessness of the naked mole-rat, as also suggested by a recent study of several hystricognath rodents (Delsuc and Tilak 2015).

Evidence of positive selection was also found in genes associated with tooth development including *ODAM*, *EDAR*, and *BMP4*. The first of these functions in enamel mineralization (Kestler et al. 2008) and has been specifically linked to incisor development in laboratory rats (Park et al. 2007); it might thus have a role in the continuous tooth replacement exhibited by some mole-rat species (Gomes Rodrigues et al. 2011).

Despite the fact that mole-rats live in low oxygen environments, we found little evidence of positive selection or GO enrichment in known hypoxia genes, although some divergent selection was detected. Of 65 genes annotated as being involved with "cellular response to hypoxia," we found divergent selection in 47, while of 67 genes listed as being involved with "cellular response to decreased oxygen levels," we found divergent selection in 48. These genes included *HIF1A*, which encodes the alpha subunit of HIF-1, and its inhibitor, *HIF1AN*. Many other genes under divergent but not positive selection were associated with oxygen levels but not explicitly hypoxic conditions (*n* = 59; data not shown).

Genes Associated with Tumorigenesis and DNA Repair

Evidence from captive and wild mole-rats has highlighted both their natural resistance to cancer and resilience to experimentally induced tumorigenesis (Seluanov et al. 2009; Liang et al. 2010; Tian et al. 2013). Genes associated with cancer that showed $\omega > 1$ in mole-rats included the breast cancer 1, early onset gene (*BRCA1*) as well as *BAP1* and *CNTROB*, which interact with *BRCA1* and *BRCA2*, respectively. Interestingly, *BRCA1* was previously found to be under positive selection in the naked mole-rat, although the reported positively selected sites (PSS) varies among studies, specifically in the Zinc Finger domain (Morgan et al. 2013), residues 430–470 (Keane et al. 2014) and the 3′-end of the protein (this study). These discrepancies may arise from differences in data filtering, alignment and models used. Other genes related to cancer and also under positive selection included *CHD5*, a putative tumor suppressor gene (Gorringe et al. 2008), *SERPINA5*, with possible roles in tumor growth and development (Jing et al. 2014), *SCAI*, which suppresses cancer cell invasion, and *COL4A2* (*α2 chain of type IV collagen*), one of the most common basement membrane collagens that has been implicated in the inhibition of tumor growth ([Kamphaus et al. 2000], also see [Kim et al. 2011]).

Other genes under selection could relate to the naked mole-rat's negligible senescence and, for its body size, extreme longevity (see table 1). These included several genes relating to the stability of chromosomes and telomerase activity, as well as genes encoding proteins related to growth, notably *GHRHR* and *GHSR* (*growth hormone secretagogue receptor*). Such findings warrant further investigation, but are intriguing given the documented role growth hormone plays in longevity and body size, especially given that the *GHR* gene is highly conserved across mole-rats (Davies et al. 2014). Mole-rats, and the naked mole-rat in particular, have also previously been studied in relation to levels of amyloid beta (Aβ) peptides (Edrey et al. 2013), the accumulation of which are thought to be associated with Alzheimer's disease. Here, we found positive selection in several genes related to Aβ, specifically *APCS* (*amyloid P component, serum*), *APLP2* (*amyloid beta (A4) precursor-like protein 2*) and *GSN* (*gelsolin*). The gene *APLP2* is also thought to play a role in spatial learning (Weyer et al. 2011).

Mole-Rat Systematics and the Evolution of Sociality

Phylogenomic reconstructions undertaken by this study strongly supported the recovery of Myomorpha as the sister-group of Hystricomorpha. Intraordinal relationships within Rodentia, for example, those between the mouse-related (Myomorpha), squirrel-related (Sciuromorpha), and guinea pig-related (Hystricomorpha) clades, have long been debated (e.g., [Huchon et al. 2002; Blanga-Kanfi et al. 2009]). This is perhaps unsurprising given that Rodentia is the largest mammalian order, in terms of species, and also that the large effective population size and short generation times displayed by most rodents are associated with rapid evolutionary rates. Our findings based on phylogenomic analysis agrees with several multimarker studies that used phylogenetic information from six nuclear (Blanga-Kanfi et al. 2009), 11 nuclear plus mitochondrial genes (Fabre et al. 2012), and genome-wide distributions of indels and short interspersed elements (SINEs) (Churakov et al. 2010). On the other hand, Meredith et al. (2011) recovered Sciuromorpha as the sister-group of Hystricomorpha based on a supermatrix of 26 gene fragments. Within the mole-rats themselves, the currently recognized interspecific relationships were recovered with high support, with the naked mole-rat basal with respect to the other species (Faulkes et al. 2004; Ingram et al. 2004).

The phylogenetic relationships among mole-rat species support several alternative hypotheses concerning the origins of sociality, which all involves multiple gains or losses. We found limited evidence of episodic positive selection acting on four candidate sociality genes in mole-rats; these signatures were detected on the ancestral mole-rat branch and on the eusocial *F. damarensis* branch. Therefore, this distribution precluded clear inferences regarding the switches in sociality, or the direction of the trait evolution within the family. However, several genes detected as being under positive selection, either by this analysis or by branch-site and clade models are nonetheless noteworthy. For example, positive selection was seen to be acting on *AVPR1A*. The expression of *AVPR1A* has previously been linked to social behavior in mammals (e.g., as reviewed by Donaldson and Young [2008]), with experiments that increased expression of *AVPR1A* in the brains of solitary meadow voles seeing an increase in partner preference (Lim et al. 2004). Furthermore, we recorded divergent selection in *arginine vasopressin* (*AVP*), which encodes the protein that is cleaved to form arginine-vasopressin, and which was previously reported as being under positive selection in the naked mole-rat (Kim et al. 2011). Aside from its association with social behavior, AVP's main biological function is fluid regulation, which may also be important to mole-rats given their diets and exploitation of arid habitats in some species. We also detected positive selection in *CD38*, which may have a role in the release of the hormone oxytocin (Jin et al. 2007; Lopatina et al. 2013). Both oxytocin and its receptor have regulatory roles in a wide range of behaviors such as parturition and lactation (Donaldson and Young 2008), and has more recently been linked to disorders such as autism (e.g., [Jacob et al. 2007]). Finally, we detected episodic positive selection on the eusocial *F. damarensis* branch in *MECP2*,

mutations in which can cause Rett syndrome, and autistic-like behavior in humans, but which have been associated with enhanced prosocial behavior in mice (Pearson et al. 2012).

It is important to note that by focusing on coding sequences, we were unable to rule out cases of molecular adaptation operating in other parts of the genome, such as noncoding regulatory regions, or epigenetic changes. For example, in the context of sociality, it was previously shown that a microsatellite repeat expansion in the promoter region of *AVPR1A* is associated with monogamy in prairie voles (Young et al. 1999), and silencing genes involved in methylation in honeybees can dramatically alter caste differentiation (Kucharski et al. 2008). Such features have hitherto been rarely examined in detail in mole-rats, but offer considerable opportunities for future research.

Comparison with Previous Genome Studies of Subterranean Mammals

Branch-site and clade models implemented in this study revealed, respectively, 513 (4.13%) and 208 (1.67%) positively selected genes in mole-rats after filtering, which decreased to 31 (0.25%) and 189 (1.72%) genes, respectively, after correcting for multiple tests. While such filtering aimed to remove false positives, it will likely have led to the removal of some genuine cases of positive selection. Furthermore, it is widely known that tests of selection based on ω (i.e., d_N/d_S) can lack power in highly conserved genes, and are sensitive to alignment errors (as discussed in Yang and dos Reis [2011]). As a consequence, direct comparison of selection results across this and previous genome-wide studies (Kim et al. 2011; Fang, Seim, et al. 2014) are not straightforward. Although overall numbers of positively selected genes reported here are broadly similar to those documented before, the overlap in gene identity is surprisingly low. Specifically we found positive selection in just 11 genes (*ADAMTS7*, *BPIFB1*, *BTF3*, *CEACAM16*, *COL3A1*, *COL4A2*, *FOLR1*, *NTN5*, *SCARF1*, *SLC30A5*, and *SYCP2*) of the approximately 140 previously reported to be under positive selection in the naked mole-rat (Kim et al. 2011) and just three genes (*INO80D*, *TMEM19*, and *TTR*) of those estimated as having $\omega > 1$ in *H. glaber* and *F. damarensis* (Fang, Seim, et al. 2014). Additionally, Kim et al. (2011) found *TERF1* to be under positive selection in the naked mole-rat, whereas our clade models suggest $\omega > 1$ in both mole-rats and outgroup species, implying that changes in this gene might not relate specifically to mole-rat biology. Comparing our results with those from the subterranean blind mole-rat, *Spalax galili* (Spalacidae) (Fang, Nevo, et al. 2014) revealed four positively selected genes in common (*ATRN*, *MDGA1*, *TMEM216*, and *TMEM180*). The first of these has been implicated in wide functions, including obesity (also see Results), pigmentation and, along with *MDGA1* and *TMEM216*, in aspects of the nervous system (e.g., [Kuramoto et al. 2001]). Finally, with respect to studies of bathyergid mole-rats, variation in gene lists could also reflect other differences in methodological approaches, such as ortholog prediction, representation and alignment methods, model and branch specification, and taxon sampling. In particular both earlier studies included fewer mole-rat and outgroup species, but with the addition of the phylogenetically more distant dog (*Canis lupus familiaris*) (Kim et al. 2011; Fang, Seim, et al. 2014), while we restricted our analyses to members of the Euarchontoglires.

However, it is important to acknowledge that while our study included wide phylogenetic sampling of the mole-rats, data are still missing from several key members of the Hystricomorpha suborder. In particular, large sequence data sets are currently unavailable for the sister taxa of African mole-rats: the cane rats (Thryonomyidae) and the dassie rat (Petromuridae) (e.g., [Blanga-Kanfi et al. 2009]). Given that many hystricomorphs members share certain life-history traits (e.g., longer gestation time, increased lifespans and sociality) (e.g., [Shadle 1948; Rathbun and Rathbun 2006; Tacutu et al. 2013]), we cannot rule out the possibility that some instances of positive selection detected on the ancestral mole-rat lineage in fact reflect adaptations deeper in the ancestry of this group. On the other hand, many of the ecological and anatomical traits exhibited by mole-rats and discussed in this study (e.g., obligate subterranean niche and eusociality) are thought to have evolved following the divergence of this monophyletic clade from their closest relatives. As such, our integrated approach combining selection analyses, tests for GO term enrichment and protein–protein interaction networks provides compelling evidence for widespread molecular adaptations underpinning the evolution of the African mole-rat family.

Materials and Methods

Species Representation and Data sets

We studied eight African mole-rat species (family Bathyergidae), representing all six genera. For seven of these species, transcriptomes were generated from specimens collected in South Africa, while coding sequences for *H. glaber* were obtained from the online Naked Mole-Rat Genome Resource (Keane et al. 2014). Additionally, we generated transcriptome data for the East African root rat *T. splendens* (Spalacidae) (see table 1). African mole-rats and the spalacid, *T. splendens* are distantly related having diverged approximately 70 Ma (Fabre et al. 2012; Fang, Nevo, et al. 2014). All transcriptomes were based on pooled samples of multiple tissue types (brain, liver, lung, skeletal muscle, heart, and kidney). RNA was isolated by BGI (Hong Kong), which also performed paired-end sequencing, with a read length of either 2×90 or 2×100 bp, on an Illumina HiSeq2000 platform (Illumina, San Diego). Following sequencing, reads containing adaptors, unknown nucleotides greater than 5% and low quality reads (i.e., $> 20\%$ bases quality scores < 10) were removed. Clean reads were deposited in the GenBank Sequence Read Archive (SRA) under SRA Study accession SRP061925 (see supplementary tables S10 and S11, Supplementary Material online, for more information). De novo transcriptome assembly was performed with Trinity (release February 25, 2013) with default settings (Grabherr et al. 2011).

De novo transcriptome assemblies can include erroneous chimeric transcripts. To reduce the number of *trans*-self-chimeras, created by orthologous reads being assembled in the incorrect direction, we cut rodent contigs in which multiple high-scoring segment pairs (HSPs) were detected in both minus and plus directions, following the method of Yang and Smith (2013). Briefly, each rodent assembly was used as the query in BLASTx searches with an *e* value cutoff of 0.01 against a database of 23,393 human proteins downloaded from Ensembl 75 (Flicek et al. 2014). Only HSPs that met the default parameters of identity $\geq$ 30% and length $\geq$ 100 bp were considered. Initially self-chimeras were searched for, and, where undetected, an additional check was made for multigene chimeras. Detected putative chimeras meeting the above criteria were cut into segments and retained if greater than 100 bp. Following the removal of putative *trans*-self-chimeras *Heliophobius emini* had the fewest contigs ($n = 105,685$) and *Cryptomys h. mahali* the most ($n = 320,282$; see supplementary table S11, Supplementary Material online). Note this method did not remove *cis*-chimeras (Yang and Smith 2013).

Protein-Coding Gene Ortholog Identification

Human and guinea pig protein sequences were downloaded from Ensembl 75, and filtered by the longest sequence per gene. In total, 23,393 human and 14,933 guinea pig sequences were used as tBLASTn queries against each rodent RNA-seq database, the top hit was kept with an *e* value cutoff of $<1e^{-6}$. Reciprocal BLASTs were performed using BLASTx, with rodent transcripts as queries against human and guinea pig protein databases, again only retaining the top hit and *e* value $< 1e^{-6}$. Percentage coverage of each rodent hit against Ensembl proteins was calculated with the perl script analyze_blastPlus_topHit_coverage.pl available in Trinity utils. Transcripts that met the reciprocal BLAST searches were kept as candidates for further analysis. Matches were filtered by the 14,873 one-to-one human guinea pig orthologous protein-coding genes downloaded from Ensembl 75.

Alignment Construction

Candidate mole-rat transcripts were combined with one-to-one canonical coding sequences downloaded via the Ensembl-API. Sequences downloaded from Ensembl with greater than 50% missing data were removed. In total, multifasta files consisted of a maximum of 21 species from the superorder Euarchontoglires; these included eight mole-rats, six outgroup rodents, two lagomorphs, and five primates. All sequences containing internal stop codons, likely to be either alternative isoforms or pseudogenes, were removed. Genes containing the amino acid selenocysteine (encoded by TGA) were retained. As the inclusion of short sequences can increase the likelihood of alignment errors all sequences less than 150 nt were removed. Multifasta files were aligned with GUIDANCE (Penn et al. 2010), using the prank algorithm (Löytynoja and Goldman 2005) with codons enforced and ten bootstraps. Following alignment, sequences with a

GUIDANCE score below the default of 0.6 were removed and the alignment process repeated. Columns were removed if they scored below the GUIDANCE score default of 0.93, or if they contained greater than 50% gaps using a perl script which kept codons intact. Finally, sequences less than 50 codons were removed from alignments. Alignments were retained if $\geq$100 codons in length and contained at least one mole-rat sequence, resulting in a final data set of 13,346 nuclear genes.

Phylogenomic Reconstruction

We constructed ML phylogenetic trees with RAxML (Stamatakis 2006) using a concatenated alignment of 3,999 genes (representing 1,754,367 codons) that each contained sequence information for all eight mole-rat taxa. Outgroup species number varied from 2 to 13 per gene. The concatenated alignment was analyzed under the general time reversible + GAMMA + I substitution model, with a partitioned analysis whereby model parameters were estimated independently for each gene partition. Nodal support for the recovered tree topology was calculated with 100 standard bootstraps. We repeated the above ML analysis with an additional partition based on codon position, whereby the codons of each gene were partitioned into first + second and third codon position.

Bayesian phylogenomic reconstruction was performed under the CAT site-heterogeneous mixture model using PhyloBayes 3.3f (Lartillot and Philippe 2004; Lartillot et al. 2009). Due to computational limitations caused by the model's complexity, the initial data set of 3,999 genes were subsampled so as to include only the 172 most informative gene partitions (60,783 amino acids) with MAREv0.1.2-rc (Meyer et al. 2011; Misof et al. 2013). Two independent Monte Carlo Markov Chains (MCMCs) were run in parallel on the reduced data set for approximately 70,000 cycles, starting from a random tree topology and saving a point every cycle. Convergence of the two chains was checked using the program bpcomp, and Bayesian posterior probabilities (BPPs) were obtained from the MCMC sampled trees after discarding the first 5,000 points as the burn-in.

Genome-Wide Screens for Changes in Selection Pressure in Mole-Rats
Alignment Error Detection and Removal

Alignment errors may lead to erroneous detection of positive selection (Fletcher and Yang 2010; Markova-Raina and Petrov 2011; Jordan and Goldman 2012). Thus to identify gene alignments containing possible errors, we filtered the data sets based on inferences of high levels of aggregated PSS based on the rationale that these are more likely to be due to alignment errors than genuine (also see [Harrison et al. 2014]). We ran branch-site and clade models, see below, across all alignments, and the resulting rst files were parsed to provide summaries of the number and location of sites inferred to be under positive selection. We used the Bayes Empirical Bayes estimates of model parameters, and considered sites to be positively selected if the foreground $\omega > 1.00$ and

proportion of sites in this category was > 0.00. We then filtered alignments in which the median interval distance among sites under selection was ≤ 10 amino acids. This threshold was based on examination of sample alignments generated from a number of taxonomic groups including rodents, cetaceans, and bats (data not shown). In total, 243 alignments were detected using the branch-site models and 706 alignments were detected with the clade models with a median interval distance ≤ 10, representing a total of 913 unique genes that were removed from both data sets. We visually inspected the 913 rejected alignments and deemed that 424 had obvious alignment/sequencing errors accounting for all potential PSS. Of the remaining 489 alignments, many had alignment errors accounting for some, but not all, potentially PSS, in total 75 of these alignments passed the branch-site LRT and 216 passed the clade model LRT, prior to FDR correction.

Detection of Positive Selection on Mole-Rat Branches

To identify genes that may have been under selection during the early evolution of mole-rats we set the ancestral mole-rat branch as the foreground branch and ran branch-site models (Zhang et al. 2005) with codeml in PAMLv4.4 (Yang 2007). This allowed us to test for evidence of positive selection on the ancestral mole-rat branch or, in the cases where there was only a single mole-rat species in the alignment, the respective terminal branch. In the branch-site test, a branch is designated as the foreground, and site-wise estimates of ω (the ratio of the number of nonsynonymous substitutions per nonsynonymous site [d_N] to the number of synonymous substitutions per synonymous site [d_S]) are compared with estimates across the remaining background branches in the tree under model A. This model is then compared with the null model A using an LRT, and significance of the model fit assessed by chi-squared test with one degree of freedom (df). For specifying the relationship among the mole-rats, we used the tree topology recovered in this study, which is consistent with that of Faulkes et al. (2004). We also applied the FDR (Benjamini and Hochberg 1995) and Holm (Holm 1979) correction to the P values calculated above to account for the multiple tests performed with a significance level of 0.05. In total, we tested 12,433 genes that had passed the above filtering based on PSS median interval; of these, 1,423 contained a single mole-rat species, and, of these, 1,356 were represented by *H. glaber*.

Detection of Divergent Selection between the Mole-Rat Clade and Outgroups

We also tested for evidence of divergent selection pressures acting across the mole-rat clade compared with the outgroup species. All branches of the mole-rat clade, including, the ancestral branch were set as the foreground clade and the remaining branches in the tree as the background. We implemented the clade model C (Bielawski and Yang 2004), in which the estimated averaged ω of the foreground clade is compared with that of the averaged ω of the background clade. Values estimated by each clade model were then compared with model M1a (nearly neutral) via a LRT with three df. We again applied the FDR and Holm correction method to the calculated P values.

GO Enrichment Analyses

We used topGO (Alexa and Rahnenfuhrer 2010) in R v.2.15 (R Development Core Team 2012) to assess enrichment in GO terms in the genes identified as either 1) under positive selection on the ancestral mole-rat branch or 2) under divergent selection in the mole-rat clade. We downloaded the GO term accession codes and associated GO domains for each of the 12,433 and 11,011 genes tested in the branch-site and clade models respectively from Ensembl. This information was then divided into three GO maps based on the three GO domains: 1) MF, 2) CC, and 3) BP. In each test our genes of interest were those defined by an uncorrected P value < 0.05, resulting in 510 and 7,232 genes for the branch-site and clade models, respectively. We implemented Fisher's exact test to identify overrepresented gene sets, which were then carried out using the classic, elim, and weight algorithms.

Visualization of Protein–Protein Interaction Networks and Functional Clusters

We tested for the presence of enriched protein–protein interactions in our samples of the protein products encoded by the identified positively selected genes using the STRING (v9.1) database (Franceschini et al. 2013) available at http://string-db.org/ (last accessed October 28, 2014). We tested all genes identified by branch-site models, and clade models with a foreground $\omega \geq 1.00$ and background $\omega < 1.00$, filtered by median interval of PSS > 10 and that passed the LRT with an uncorrected P value ≤ 0.05. We used the human Ensembl gene identifiers to represent the protein products of the genes and calculated enrichment in connections compared with the custom background of all tested gene products (again filtered by median interval between PSS > 10). We retrieved combined interaction scores based on the following lines of evidence: neighborhood, fusion, co-occurrence, homology, coexpression, experimental knowledge and text mining, and visualized networks with igraph v0.7.0 (Csardi and Nepusz 2006).

Additionally, we summarized the results of the selection tests, network analysis, and GO enrichment analyses by color coding the products of the positively selected genes based on broad-scale GO terms in the network. For each of our enriched GO terms, we retrieved all genes found to be significant with our selection tests, based on the annotations from topGO (Alexa and Rahnenfuhrer 2010) in R v.2.15 (R Development Core Team 2012). We grouped the GO terms into functional categories hypothesized to be important during mole-rat evolution: "immune system" (e.g., "response to virus"), "reproduction" (e.g., "fertilization"), "extracellular" (e.g., "biological adhesion"), "metabolism" (e.g., "hormone metabolic process"), "cell cycle" (e.g., "centriole replication"), "cell" (e.g., "cell motility"), "aging" (e.g., "determination of adult lifespan"), "morphology" (e.g., "biomineral tissue development"), "sensory" (e.g., "phototransduction"), and

"symbionts and host" (e.g., "growth involved in symbiotic interaction").

HYPHY Analysis of Genes Relating to Social Behavior

We identified 64 candidate mammalian "social" genes, based on annotations with the following GO terms: "social behavior (GO:0035176)," "vocalization behavior (GO:0071625)," "grooming behavior (GO:0007625)," and "regulation of grooming behavior GO:2000821" using AmiGO2 version 2.14 (Carbon et al. 2009) (see supplementary table S12, Supplementary Material online). For these candidate genes, we combined novel coding sequences retrieved from our transcriptomes together with sequences obtained for available Hystricognathi mammals from GenBank (C. porcellus, Octodon degus, Chinchilla lanigera, F. damarensis (Fang, Seim, et al. 2014), and H. glaber (Kim et al. 2011). Alignments were constructed with GUIDANCE and prank as described above, codon sites with greater than 75% gaps were removed. We ran the Branch-site REL (Pond et al. 2011) model in HyPhy (Pond et al. 2005), to identify branches within the species tree containing a proportion of sites under episodic diversifying selection. Documented social structures (i.e., solitary, social + cooperative breeding and eusocial) (Faulkes and Bennett 2013) were mapped onto the species tree, and possible evolutionary scenarios explaining the distribution of social structure across the family were inferred. These scenarios were used to highlight branches we might expect to find to be under diversifying selection (see supplementary fig. S4, Supplementary Material online).

Supplementary Material

Acknowledgments

References

Alexa A, Rahnenfuhrer J. 2010. topGO: enrichment analysis for gene ontology. R package version 2.16.0.

Azpurua J, Ke Z, Chen IX, Zhang Q, Ermolenko DN, Zhang ZD, Gorbunova V, Seluanov A. 2013. Naked mole-rat has increased translational fidelity compared with the mouse, as well as a unique 28S ribosomal RNA cleavage. *Proc Natl Acad Sci U S A.* 110:17350–17355.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J Roy Stat Soc B Met.* 57:289–300.

Bennett NC, Faulkes CG. 2000. African mole-rats: ecology and eusociality. Cambridge University Press.

Bennett NC, Jarvis JUM. 2004. *Cryptomys damarensis. Mamm Species.* 756:1–5.

Bielawski JP, Yang ZH. 2004. A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *J Mol Evol.* 59:121–132.

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

Bok D, Galbraith G, Lopez I, Woodruff M, Nusinowitz S, BeltrandelRio H, Huang W, Zhao S, Geske R, Montgomery C, et al. 2003. Blindness and auditory impairment caused by loss of the sodium bicarbonate cotransporter NBC3. *Nat Genet.* 34:313–319.

Botto M, Hawkins PN, Bickerstaff MCM, Herbert J, Bygrave AE, Mcbride A, Hutchinson WL, Tennent GA, Walportz MJ, Pepys MB. 1997. Amyloid deposition is delayed in mice with targeted deletion of the *serum amyloid P component* gene. *Nat Med.* 3:855–859.

Braude S, Ciszek D, Berg NE, Shefferly N. 2001. The ontogeny and distribution of countershading in colonies of the naked mole-rat (*Heterocephalus glaber*). *J Zool.* 253:351–357.

Brett RA. 1991. The ecology of naked mole-rat colonies: burrowing, food, and limiting factors. In: Sherman PW, Jarvis JUM, Alexander RD, editors. The biology of the naked mole-rat. Princeton (NJ): Princeton University Press.

Buffenstein R. 2008. Negligible senescence in the longest living rodent, the naked mole-rat: insights from a successfully aging species. *J Comp Physiol B.* 178:439–445.

Buffenstein R, Jarvis JU. 2002. The naked mole rat—a new record for the oldest living rodent. *Sci Aging Knowledge Environ.* 2002:pe7.

Cagnacci A, Elliott JA, Yen SS. 1992. Melatonin: a major regulator of the circadian rhythm of core temperature in humans. *J Clin Endocrinol Metab.* 75:447–452.

Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, AmiGO Hub, Web Presence Working Group. 2009. AmiGO: online access to ontology and annotation data. *Bioinformatics* 25:288–289.

Churakov G, Sadasivuni MK, Rosenbloom KR, Huchon D, Brosius J, Schmitz J. 2010. Rodent evolution: back to the root. *Mol Biol Evol.* 27:1315–1326.

Cornet AM, Hanon E, Reiter ER, Bruyninx M, Nguyen VH, Hennuy BR, Hennen GP, Closset JL. 2003. Prostatic androgen repressed message-1 (PARM-1) may play a role in prostatic cell immortalisation. *Prostrate* 56:220–230.

Csardi G, Nepusz T. 2006. The igraph software package for complex network research. *InterJ Complex Syst.* 1695.

Daly TJM, Buffenstein R. 1998. Skin morphology and its role in thermoregulation in mole-rats, *Heterocephalus glaber* and *Cryptomys hottentotus. J Anat.* 193:495–502.

Davies KTJ, Tsagkogeorga G, Bennett NC, Dávalos LM, Faulkes CG, Rossiter SJ. 2014. Molecular evolution of growth hormone and insulin-like growth factor 1 receptors in long-lived, small-bodied mammals. *Gene* 549:228–236.

de Vries JL, Oosthuizen MK, Sichilima AM, Bennett NC. 2008. Circadian rhythms of locomotor activity in Ansell's mole-rat: are mole-rat's clocks ticking? *J Zool* 276:343–349.

Delsuc F, Tilak M-K. 2015. Naked but not hairless: the pitfalls of analyses of molecular adaptation based on few genome sequence comparisons. *Genome Biol Evol.* 7:768–774.

Donaldson ZR, Young LJ. 2008. Oxytocin, vasopressin, and the neurogenetics of sociality. *Science* 322:900–904.

Dziarski R, Gupta D. 2010. Mammalian peptidoglycan recognition proteins (PGRPs) in innate immunity. *Innate Immun.* 16:168–174.

Edrey YH, Casper D, Huchon D, Mele J, Gelfond JA, Kristan DM, Nevo E, Buffenstein R. 2012. Sustained high levels of neuregulin-1 in the longest-lived rodents; a key determinant of rodent longevity. *Aging Cell* 11:213–222.

Edrey YH, Medina DX, Gaczynska M, Osmulski PA, Oddo S, Caccamo A, Buffenstein R. 2013. Amyloid beta and the longest-lived rodent: the naked mole-rat as a model for natural protection from Alzheimer's disease. *Neurobiol Aging.* 34:2352–2360.

Emerling CA, Springer MS. 2014. Eyes underground: regression of visual protein networks in subterranean mammals. *Mol Phylogenet Evol.* 78:260–270.

Fabre P-H, Hautier L, Dimitrov D, Douzery EJP. 2012. A glimpse on the pattern of rodent diversification: a phylogenetic approach. *BMC Evol Biol.* 12:88.

Fang X, Nevo E, Han L, Levanon EY, Zhao J, Avivi A, Larkin D, Jiang X, Feranchuk S, Zhu Y, et al. 2014. Genome-wide adaptive complexes to underground stresses in blind mole rats *Spalax*. *Nat Commun.* 5:3966.

Fang X, Seim I, Huang Z, Gerashchenko MV, Xiong Z, Turanov AA, Zhu Y, Lobanov AV, Fan D, Yim SH, et al. 2014. Adaptations to a subterranean environment and longevity revealed by the analysis of mole rat genomes. *Cell Rep.* 8:1354–1364.

Faulkes CG, Bennett NC. 2013. Plasticity and constraints on social evolution in African mole-rats: ultimate and proximate factors. *Philos Trans R Soc B Biol Sci.* 368:20120347.

Faulkes CG, Bennett NC, Bruford MW, O'Brien HP, Aguilar GH, Jarvis JUM. 1997. Ecological constraints drive social evolution in the African mole-rats. *Proc R Soc Lond Ser B Biol Sci.* 264:1619–1627.

Faulkes CG, Bennett NC, Cotterill FPD, Stanley W, Mgode GF, Verheyen E. 2011. Phylogeography and cryptic diversity of the solitary-dwelling silvery mole-rat, genus *Heliophobius* (Family: Bathyergidae). *J Zool.* 285:324–338.

Faulkes CG, Davies KTJ, Rossiter SJ, Bennett NC. 2015. Molecular evolution of the *hyaluronan synthase* 2 gene in mammals: implications for adaptations to the subterranean niche and cancer resistance. *Biol Lett.* 11:20150185.

Faulkes CG, Mgode GF, Le Comber SC, Bennett NC. 2010. Cladogenesis and endemism in Tanzanian mole-rats, genus *Fukomys* (Family: Bathyergidae): a role for tectonics? *Biol J Linn Soc* 100:337–352.

Faulkes CG, Verheyen E, Verheyen W, Jarvis JUM, Bennett NC. 2004. Phylogeographical patterns of genetic divergence and speciation in African mole-rats (Family: Bathyergidae). *Mol Ecol.* 13:613–629.

Fletcher W, Yang Z. 2010. The effect of insertions, deletions, and alignment errors on the branch-site test of positive selection. *Mol Biol Evol.* 27:2257–2267.

Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, et al. 2014. Ensembl 2014. *Nucleic Acids Res.* 42:D749–D755.

Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, et al. 2013. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 41:D808–D815.

Gomes Rodrigues H, Marangoni P, Šumbera R, Tafforeau P, Wendelen W, Viriot L. 2011. Continuous dental replacement in a hyper-chisel tooth digging rodent. *Proc Natl Acad Sci U S A.* 108:17355–17359.

Gorringe KL, Choong DYH, Williams LH, Ramakrishna M, Sridhar A, Qiu W, Bearfoot JL, Campbell IG. 2008. Mutation and methylation analysis of the *Chromodomain-helicase-DNA binding* 5 gene in ovarian cancer. *Neoplasia* 10:1253–1258.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol.* 29:644–652.

Harrison PW, Jordan GE, Montgomery SH. 2014. SWAMP: sliding window alignment masker for PAML. *Evol Bioinform Online.* 10:197–204.

Hayakawa T, Suzuki-Hashido N, Matsui A, Go Y. 2014. Frequent expansions of the bitter taste receptor gene repertoire during evolution of mammals in the Euarchontoglires clade. *Mol Biol Evol.* 31:2018–2031.

He L, Gunn TM, Bouley DM, Lu X-Y, Watson SJ, Schlossman SF, Duke-Cohan JS, Barsh GS. 2001. A biochemical function for attractin in agouti-induced pigmentation and obesity. *Nat Genet.* 27:40–47.

Heffner RS, Heffner HE. 1993. Degenerate hearing and sound localization in naked mole rats (*Heterocephalus glaber*), with an overview of central auditory structures. *J Comp Neurol.* 331:418–433.

Heldermon CD, Hennig AK, Ohlemiller KK, Ogilvie JM, Herzog ED, Breidenbach A, Vogler C, Wozniak DF, Sands MS. 2007. Development of sensory, motor and behavioral deficits in the murine model of Sanfilippo Syndrome type B. *PLoS One* 2:e772.

Holm S. 1979. A simple sequentially rejective multiple test procedure. *Scand J Stat.* 6:65–70.

Huang S, Czech MP. 2007. The GLUT4 glucose transporter. *Cell Metab.* 5:237–252.

Huchon D, Madsen O, Sibbald MJJB, Ament K, Stanhope MJ, Catzeflis F, de Jong WW, Douzery EJP. 2002. Rodent phylogeny and a timescale for the evolution of Glires: evidence from an extensive taxon sampling using three nuclear genes. *Mol Biol Evol.* 19:1053–1065.

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

Jacob S, Brune CW, Carter CS, Leventhal BL, Lord C, Cook Jr. EH. 2007. Association of the *oxytocin receptor* gene (*OXTR*) in Caucasian children and adolescents with autism. *Neurosci Lett.* 417:6–9.

Jin D, Liu H-X, Hirai H, Torashima T, Nagai T, Lopatina O, Shnayder NA, Yamada K, Noda M, Seike T, et al. 2007. CD38 is critical for social behaviour by regulating oxytocin secretion. *Nature* 446:41–45.

Jing Y, Jia D, Wong CM, Oi-Lin Ng I, Zhang Z, Liu L, Wang Q, Zhao F, Li J, Yao M, et al. 2014. SERPINA5 inhibits tumor cell migration by modulating the fibronectin-integrin β1 signaling pathway in hepatocellular carcinoma. *Mol Oncol.* 8:366–377.

Jordan G, Goldman N. 2012. The effects of alignment error and alignment filtering on the sitewise detection of positive selection. *Mol Biol Evol.* 29:1125–1139.

Kamphaus GD, Colorado PC, Panka DJ, Hopfer H, Ramchandran R, Torre A, Maeshima Y, Mier JW, Sukhatme VP, Kalluri R. 2000. Canstatin, a novel matrix-derived inhibitor of angiogenesis and tumor growth. *J Biol Chem.* 275:1209–1215.

Keane M, Craig T, Alföldi J, Berlin AM, Johnson J, Seluanov A, Gorbunova V, Di Palma F, Lindblad-Toh K, Church GM, et al. 2014. The naked mole rat genome resource: facilitating analyses of cancer and longevity-related adaptations. *Bioinformatics* 30:3558–3560.

Kestler DP, Foster JS, Macy SD, Murphy CL, Weiss DT, Solomon A. 2008. Expression of odontogenic ameloblast-associated protein (ODAM) in dental and other epithelial neoplasms. *Mol Med.* 4:318–326.

Kilduff TS, Peyron C. 2000. The hypocretin/orexin ligand-receptor system: implications for sleep and sleep disorders. *Trends Neurosci.* 23:359–365.

Kim EB, Fang X, Fushan AA, Huang Z, Lobanov AV, Han L, Marino SM, Sun X, Turanov AA, Yang P, et al. 2011. Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature* 479:223–227.

Kosiol C, Vinař T, da Fonseca RR, Hubisz MJ, Bustamante CD, Nielsen R, Siepel A. 2008. Patterns of positive selection in six mammalian genomes. *PLoS Genet.* 4:e1000144.

Kott O, Šumbera R, Němec P. 2010. Light perception in two strictly subterranean rodents: life in the dark or blue? *PLoS One* 5:e11810.

Kucharski R, Maleszka J, Foret S, Maleszka R. 2008. Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 319:1827–1830.

Kuramoto T, Kitada K, Inui T, Sasaki Y, Ito K, Hase T, Kawagachi S, Ogawa Y, Nakao K, Barsh GS, et al. 2001. Attractin/mahogany/zitter plays a critical role in myelination of the central nervous system. *Proc Natl Acad Sci U S A.* 98:559–564.

Lacey EA, Pepper JW, Alexander RD, Braude SH, Sherman PW, Jarvis JUM. 1991. An ethogram for the naked mole-rat: non-vocal behaviours. In: Sherman PW, Jarvis JUM, Alexander RD, editors. The biology of the naked mole-rat. New York: Princeton University Press. p. 209–242.

Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25:2286–2288.

Lartillot N, Philippe H. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol Biol Evol. 21:1095–1109.

Li Y-J, Oliveira SA, Xu P, Martin ER, Stenger JE, Scherzer CR, Hauser MA, Scott WK, Small GW, Nance MA, et al. 2003. Glutathione S-transferase omega-1 modifies age-at-onset of Alzheimer disease and Parkinson disease. Hum Mol Genet. 12:3259–3267.

Liang S, Mele J, Wu Y, Buffenstein R, Hornsby PJ. 2010. Resistance to experimental tumorigenesis in cells of a long-lived mammal, the naked mole-rat (Heterocephalus glaber). Aging Cell 9:626–635.

Lim MM, Wang Z, Olazábal DE, Ren X, Terwilliger EF, Young LJ. 2004. Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. Nature 429:754–757.

Liu Z, Wang W, Zhang TZ, Li GH, He K, Huang JF, Jiang XL, Murphy RW, Shi P. 2013. Repeated functional convergent effects of NaV1.7 on acid insensitivity in hibernating mammals. Proc R Soc Lond Ser B Biol Sci. 281:20132950.

Lopatina O, Inzhutova A, Salmina AB, Higashida H. 2013. The roles of oxytocin and CD38 in social or parental behaviors. Front Neurosci. 6:182.

Lopez IA, Acuna D, Galbraith G, Bok D, Ishiyama A, Liu W, Kurtz I. 2005. Time course of auditory impairment in mice lacking the electroneutral sodium bicarbonate cotransporter NBC3 (slc4a7). Dev Brain Res. 160:63–77.

Löytynoja A, Goldman N. 2005. An algorithm for progressive multiple alignment of sequences with insertions. Proc Natl Acad Sci U S A. 102:10557–10562.

Markova-Raina P, Petrov D. 2011. High sensitivity to aligner and high rate of false positives in the estimates of positive selection in the 12 Drosophila genomes. Genome Res. 21:863–874.

Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simao TLL, Stadler T, et al. 2011. Impacts of the cretaceous terrestrial revolution and KPg extinction on mammal diversification. Science 334:521–524.

Meyer B, Meusemann K, Misof B. 2011. MARE: matrix reduction—a tool to select optimized data subsets from supermatrices for phylogenetic inference. Bonn (Germany): Zentrum für Molekulare Biodiversitätsforschung (zmb) am ZFMK.

Michauta M, De Blasa G, Tomesa CN, Yunesa R, Fukudab M, Mayorgaa LS. 2001. Synaptotagmin VI participates in the acrosome reaction of human spermatozoa. Dev Biol. 235:521–529.

Misof B, Meyer B, von Reumont BM, Kück P, Misof K, Meusemann K. 2013. Selecting informative subsets of sparse supermatrices increases the chance to find correct trees. BMC Bioinformatics 14:348.

Morgan CC, McCartney AM, Donoghue MT, Loughran NB, Spillane C, Teeling EC, O'Connell MJ. 2013. Molecular adaptation of telomere associated genes in mammals. BMC Evol Biol. 13:251.

Narins PM, Reichman OJ, Jarvis JUM, Lewis ER. 1992. Seismic signal transmission between burrows of the Cape mole-rat, Georychus capensis. J Comp Physiol A. 170:13–21.

Nevo E. 1979. Adaptive convergence and divergence of subterranean mammals. Annu Rev Ecol Syst. 10:269–308.

Nikitina NV, Kidson SH. 2014. Eye development in the Cape dune mole rat. Dev Genes Evol. 224:107–117.

Oosthuizen MK, Cooper HM, Bennett NC. 2003. Circadian rhythms of locomotor activity in solitary and social species of African mole-rats (family: Bathyergidae). J Biol Rhythms. 18:481–490.

Park JC, Park JT, Son HH, Kim HJ, Jeong MJ, Lee CS, Dey R, Cho MI. 2007. The amyloid protein APin is highly expressed during enamel mineralization and maturation in rat incisors. Eur J Oral Sci. 115:153-160.

Park TJ, Lu Y, Jüttner R, Smith ESJ, Hu J, Brand A, Wetzel C, Milenkovic N, Erdmann B, Heppenstall PA, et al. 2008. Selective inflammatory pain insensitivity in the African naked mole-rat (Heterocephalus glaber). PLoS Biol. 6:e13.

Pearson BL, Defensor EB, Pobbe RL, Yamamoto LH, Bolivar VJ, Blanchard DC, Blanchard RJ. 2012. Mecp2 truncation in male mice promotes affiliative social behavior. Behav Genet. 42:299–312.

Penn O, Privman E, Landan G, Graur D, Pupko T. 2010. An alignment confidence score capturing robustness to guide-tree uncertainty. Mol Biol Evol. 27:1759–1767.

Pond SLK, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21:676–679.

Pond SLK, Murrell B, Fourment M, Frost SDW, Delport W, Scheffler K. 2011. A random effects branch-site model for detecting episodic diversifying selection. Mol Biol Evol. 28:3033–3043.

R Development Core Team. 2012. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rathbun GB, Rathbun CD. 2006. Social monogamy in the noki or dassierat (Petromus typicus) in Namibia. Mammal Biol. (Zeitschrift für Säugetierkunde) 71:203–213.

Riccio AP, Goldman BD. 2000. Circadian rhythms of locomotor activity in naked mole-rats (Heterocephalus glaber). Physiol Behav. 71:1–13.

Rodriguez KA, Osmulski PA, Pierce A, Weintraub ST, Gaczynska M, Buffenstein R. 2014. A cytosolic protein factor fromthe naked-mole-rat activates proteasomes of other species and protects these from inhibition. Biochim Biophys Acta. 1842:2060–2072.

Roper TJ, Bennett NC, Conradt L, Molteno AJ. 2001. Environmental conditions in burrows of two species of African mole-rat, Georhychus capensis and Cryptomys damarensis. J Zool. 254:101–107.

Salmon AB, Sadighi Akha AA, Buffenstein R, Miller RA. 2008. Fibroblasts from naked mole-rats are resistant to multiple forms of cell injury, but sensitive to peroxide, ultraviolet light, and endoplasmic reticulum stress. J Gerontol A Biol Sci Med Sci. 63A:232–241.

Salomoni P, Pandolfi PP. 2002. The role of PML in tumor suppression. Cell Tissue Res. 108:165–170.

Schmidt CM, Jarvis JUM, Bennett NC. 2013. The long-lived queen: reproduction and longevity in female eusocial Damaraland mole-rats (Fukomys damarensis). Afr Zool. 48:193–196.

Schultz JM, Khan SN, Ahmed ZM, Riazuddin S, Waryah AM, Chhatre D, Starost MF, Ploplis B, Buckley S, Velásquez D, et al. 2009. Noncoding mutations of HGF are associated with nonsyndromic hearing loss, DFNB39. Am J Hum Genet. 85:25–39.

Seluanov A, Hine C, Azpurua J, Feigenson M, Bozzella M, Mao Z, Catania KC, Gorbunova V. 2009. Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat. Proc Natl Acad Sci U S A. 106:19352–19357.

Shadle AR. 1948. Gestation period in the porcupine, Erethizon dorsatum dorsatum. J Mammal. 29:162–164.

Shams I, Avivi A, Nevo E. 2005. Oxygen and carbon dioxide fluctuations in burrows of subterranean blind mole rats indicate tolerance to hypoxic-hypercapnic stresses. Comp Biochem Physiol A Mol Integr Physiol. 142:376–382.

Sherman PW, Jarvis JUM. 2002. Extraordinary life spans of naked mole-rats (Heterocephalus glaber). J Zool. 258:307–311.

Slominski A, Tobin DJ, Shibahara S, Wortsman J. 2004. Melanin pigmentation in mammalian skin and its hormonal regulation. Physiol Rev. 84:1155–1228.

Smith ESJ, Omerbašić D, Lechner SG, Anirudhan G, Lapatsina L, Lewin GR. 2011. The molecular basis of acid insensitivity in the African naked mole-rat. Science 334:1557–1560.

Soscia SJ, Harrington ME. 2005. Neuropeptide Y does not reset the circadian clock in NPY Y2−/− mice. Neurosci Lett. 373:175–178.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Stathopoulos S, Bishop JM, O'Ryan C. 2014. Genetic signatures for enhanced olfaction in the African mole-rats. PLoS One 9:e93336.

Sudhakar A, Nyberg P, Keshamouni VG, Mannam AP, Li J, Sugimoto H, Cosgrove D, Kalluri R. 2005. Human alpha1 type IV collagen NC1 domain exhibits distinct antiangiogenic activity mediated by alpha1-beta1 integrin. J Clin Investig. 115:2801–2810.

Tacutu R, Craig T, Budovsky A, Wuttke D, Lehmann G, Taranukha D, Costa J, Fraifeld VE, de Magalhaes JP. 2013. Human ageing genomic resources: integrated databases and tools for the biology and genetics of ageing. Nucleic Acids Res. 41:D102–D1033.

Tian X, Azpurua J, Hine C, Vaidya A, Myakishev-Rempel M, Ablaeva J, Mao Z, Nevo E, Gorbunova V, Seluanov A. 2013. High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. Nature 499:346–349.

Tsai H, Hardisty RE, Rhodes C, Kiernan AE, Roby P, Tymowska-Lalanne Z, Mburu P, Rastan S, Hunter AJ, Brown SD, et al. 2001. The mouse slalom mutant demonstrates a role for Jagged1 in neuroepithelial patterning in the organ of Corti. Hum Mol Genet. 10:507–512.

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

van Steensel B, Smogorzewska A, de Lange T. 1998. TRF2 protects human telomeres from end-to-end fusions. Cell 92:401–413.

Weigl R. 2005. Longevity of mammals in captivity; from the living collections of the world. Stuttgart: Kleine Senckenberg-Reihe Band 48.

Weyer SW, Klevanski M, Delekate A, Voikar V, Aydin D, Hick M, Filippov M, Drost N, Schaller KL, Saar M, et al. 2011. APP and APLP2 are essential at PNS and CNS synapses for transmission, spatial learning and LTP. EMBO J. 30:2266–2280.

Wilson DE, Reeder DM, editors. 2005. Mammal species of the world. A taxonomic and geographic reference Baltimore: Johns Hopkins University Press.

Xiaodong L, Staszewski L, Xu H, Durick K, Zoller M, Adler E. 2002. Human receptors for sweet and umami taste. Proc Natl Acad Sci U S A. 99:4692–4696.

Yang Y, Smith SA. 2013. Optimizing de novo assembly of short-read RNA-seq data for phylogenomics. BMC Genomics 14:328.

Yang ZH. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24:1586–1591.

Yang ZH, dos Reis M. 2011. Statistical properties of the branch-site test of positive selection. Mol Biol Evol. 28:1217–1228.

Yosida S, Kobayasi KI, Ikebuchi M, Ozaki R, Okanoya K. 2007. Antiphonal vocalization of a subterranean rodent, the naked mole-rat (Heterocephalus glaber). Ethology 113:703–710.

Young LJ, Nilsen R, Waymire KG, MacGregor GR, Insel TR. 1999. Increased affiliative response to vasopressin in mice expressing the V1a receptor from a monogamous vole. Nature 400:766–768.

Zhang JZ, Nielsen R, Yang ZH. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol. 22:2472–2479.

Zhao HB, Rossiter SJ, Teeling EC, Li CJ, Cotton JA, Zhang SY. 2009. The evolution of color vision in nocturnal mammals. Proc Natl Acad Sci U S A. 106:8980-8985.

Zhao HB, Ru BH, Teeling EC, Faulkes CG, Zhang SY, Rossiter SJ. 2009. Rhodopsin molecular evolution in mammals inhabiting low light environments. PLoS One 4:e8326.

Zhao Y, Zhao F, Zong L, Zhang P, Guan L, Zhang J, Wang D, Wang J, Chai W, Lan L, et al. 2013. Exome sequencing and linkage analysis identified Tenascin-C (TNC) as a novel causative gene in nonsyndromic hearing loss. PLoS One 8:e69549.

Zhou XZ, Lu KP. 2001. The Pin2/TRF1-interacting protein PinX1 is a potent telomerase inhibitor. Cell 107:347–359.