Patterns of Genetic Diversity and Gene Flow Associated With an Aridity Gradient in Populations of Common Mole-rats, *Cryptomys hottentotus hottentotus*

Hana N. Merchant [ID] [1,2,*], Anastasia Ivanova[2], Daniel W. Hart[3], Cristina García[1], Nigel C. Bennett[4], Steven J. Portugal[1,5], Chris G. Faulkes[2]

[1]Department of Biological Sciences, School of Life and Environmental Sciences, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK

[2]School of Biological and Behavioural Sciences, Queen Mary University of London, London, UK

[3]Department of Zoology and Entomology, University of Pretoria, Pretoria, Gauteng, South Africa

[4]Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Pretoria, Gauteng, South Africa

[5]Department of Biology, University of Oxford, Oxford OX1 3SZ, UK

*Corresponding author: E-mail: Hana.Merchant.2020@live.rhul.ac.uk.

Abstract

Genetic adaptation is the change of a population toward a phenotype that best fits the present ecological conditions of the environment it inhabits. As environmental conditions change, allele frequencies shift, resulting in different populations of the same species possessing genetic variation and divergent phenotypes. Cooperatively breeding common mole-rats (*Cryptomys hottentotus hottentotus*) inhabit environments along an aridity gradient in South Africa, which provides an opportunity for local genetic adaptations to occur. Using one mitochondrial gene (cytochrome b) and 3,540 SNP loci across the whole genome, we determined the phylogenetic relationship, population structure and genetic diversity of five populations of *C. h. hottentotus* located along an aridity gradient. Mitochondrial data identified population-specific clades that were less distinct in the two mesic populations, potentially indicating historical or recent gene flow, or the retention of ancestral haplotypes. Arid and semi-arid populations formed a distinct cluster from the non-arid populations. Genetic diversity and gene flow were higher in arid-dwelling individuals, suggesting greater connectivity and interactions between colonies in arid regions in comparison to mesic ones. Using an Aridity Index, we determined that isolation by environment, rather than isolation by geographical distance, best explains the genetic distance between the populations. Further analyses using target loci may determine if there are differing underlying genetic adaptations among populations of *C. h. hottentotus*. These analyses could help unravel population differences in response to environmental factors within a subspecies of bathyergid mole-rat and determine the adaptive capacity of this small nonmigratory subterranean rodent species in response to aridification in the face of climate change.

Key words: arid-dwelling mammals, Bathyergidae, environmental adaptation, genetic adaptation, genome-wide variation, landscape genomics.

Significance

This study identifies population-level genomic differences that occur in association with extreme arid conditions and highlights that some populations may be more at risk in response to climate change and aridification than others as a result of changes in allelic composition and genetic diversity. The potential for loss of distributional range in this species could thus be high, and therefore further investigations of their genetics and specific adaptations may influence conservation policies to ensure the survival of this and related species.

Introduction

Genetic adaptation is affected by population-level genetic processes such as genetic drift, natural selection, random mutations, and gene flow (Cortázar-Chinarro et al. 2017). The variation in selective pressures between habitats influences the occurrences and effects of these processes resulting in genetic variation (McCallum et al. 2014) and divergent phenotypes (Orr 2005; Hoffmann and Willi 2008). Certain environments require extreme and more specialized adaptations, and thus can exert a stronger influence on the genetic differentiation of populations than actual geographical (physical) distance (Wang and Bradburd 2014); often the more extreme the environment, the greater the adaptive drive. One example of an extreme environment is xeric biomes, which include deserts and arid regions (Harris et al. 1998). Arid environments are typically characterized by extreme water scarcity, high temperatures, and minimal vegetation, all of which present challenges that influence the adaptive strategies of the organisms living in these conditions (Harris et al. 1998).

Certain genomic regions interface more with the environment, and thus will be more relevant regarding adaptation to extreme desert conditions (Rocha et al. 2021). Previously identified genes linked to mammalian desert adaptations include those associated with energy storage, low energy expenditure, adaptive tolerance to starvation, water retention, stress responses (such as oxidative damage and heat stress), radiation tolerance, and metabolism of toxins from the diet (Rocha et al. 2021). The use of genome-wide single nucleotide polymorphisms (SNPs) has successfully located these regions within the genome that may be associated with environmental stressors, specifically those associated with arid environments (Rocha et al. 2021). Further analyses using genome-wide studies may be used to examine population structure and genetic variability. Landscape genetic analysis can shed light on intraspecific genetic variability, and the adaptive capacity of various populations. Such approaches are becoming particularly critical in light of global rapidly changing environments, and provide insight into population connectivity, habitat structure, dispersal patterns, and biodiversity (Drake et al. 2022). For example, desert-dwelling cactus mice (*Peromyscus eremicus*) are a prime example of a species known to display heat tolerance and dehydration adaptations. Using whole genome data, Tigano et al. (2020) identified sites across the genome associated with selective sweeps. Functional annotations of candidate genes revealed signatures of selection at sites with genes related to cellular mechanisms used to cope with thermal and hyperosmotic stress. Additionally, population structure analyses identified distinct clusters of populations isolated by geographical barriers, regardless of geographic proximity, driven by the limited dispersal ability of *P. eremicus* (Tigano et al. 2020). Moreover, a study on greater gliders (*Petauroides volans*) different

geographic locations identified candidate markers based on genome-wide SNPs. Geographically isolated populations showed low genetic diversity, and candidate markers indicated the detection of regions of adaptation to temperature using genotype–environment association (Knipler et al. 2023). Such explorations enable a deeper understanding of population genetics, coupled with local adaptations in response to an aridity gradient. These studies highlight the impact of environmental and geographical features on genetic diversity and variation. Nevo and Beiles (1989) compared genetic diversity in species of crickets, land snails, lizards, mice, and pocket gophers that inhabit the desert, Mediterranean, and steppe regions of Israel. This study found that on Mount Carmel, also termed "the Evolution Canyon", genetic diversity and heterozygosity increased toward ecologically heterogeneous environments. This study also found evidence of increased genetic diversity in Mediterranean species thriving in and near the harsher desert environment (Nevo and Beiles 1989), whilst genetic diversity was low in the species existing solely in the arid habitat. The result from this study necessitates further investigation into the population genetics and structure of such species whose range spans different biomes, from arid to mesic, to further understand how aridity influences genetic diversity. A deeper understanding of intraspecific genetic variation has numerous practical applications. Genetic diversity, for example, can be a reliable indicator of translocation success in reintroduced individuals (Scott et al. 2020). Scott et al. (2020) found that Mojave Desert tortoises (*Gopherus agassizii*) with higher individual heterozygosity survived at much higher rates than those with lower individual heterozygosity when reintroduced into wild populations.

African mole-rats (Bathyergidae) are subterranean rodents found across sub-Saharan Africa (Faulkes and Bennett 2021). Mole-rats are an ideal clade for comparative analyses because they occupy a broad distribution crossing many climatic regions, coupled with variations in food availability and soil type. The ability of common mole-rats and other African mole-rat species to persist and thrive in arid regions of Africa is due to the adaptive benefits of group living which increases the energy allocations dedicated to foraging and locating stochastically distributed food sources (Jarvis et al. 1994). The aridity food distribution hypothesis (AFDH) posits that larger numbers of individuals per colony are found in arid regions to increase the chances of finding stochastically distributed food resources to secure sufficient energy allocations of the colony dedicated to foraging (Jarvis et al. 1994). The scarcity of food is thought to limit the number of individuals that can exist in these arid environments, and thus reduce genetic diversity and connectivity between colonies in arid populations (Visser et al. 2018). Two of the three social genera, *Fukomys* and *Cryptomys*, share common ancestry around 10 to 12 million years ago. Throughout much of the range of *Fukomys*, speciation and diversity have been shaped by landscape evolution driven by physical, ecological,

and climatic changes linked to the formation of the African Rift Valley. Cladogenesis is associated with significant bouts of volcanic activity (Faulkes et al. 2010, 2017), whilst changes in drainage patterns within the major river systems of the Zambezian region in South-Central Africa are particularly important for structuring present populations and potential speciation events in some *Fukomys* populations (Van Daele et al. 2004, 2007, 2013). In comparison, the genus *Cryptomys* has been restricted to the less geologically active southern African region, diversifying almost entirely across South Africa (Faulkes et al. 2017). Within the genus, common mole-rats (*Cryptomys hottentotus hottentotus*) are unusual in being found along an environmental cline in South Africa, ranging from mesic to arid habitats. This allows for within-species comparison, between different populations in widely different environments. Population-level differences in morphology have recently been identified in response to an aridity gradient (Merchant et al. 2024), thus indicating that aridity is acting as a selective pressure on populations of *C. h. hottentotus*. Further exploration into genetic adaptation in response to aridity is required to determine the level of adaptation occurring at the population level.

Using microsatellite genetic markers, Bishop et al. (2004, 2007) explored, for the first time, *C. h. hottentotus* colony structure and relatedness in both arid and mesic populations, revealing that colonies are dominated by outbreeding family groups consisting of a breeding queen and siblings with sometimes non-related adult individuals. Paternity was found to be primarily attributed to the dominant male in the colony; however, the presence of extra-colony paternity by floating males was shown at both sites by the occurrence of offspring whose paternity could not be assigned to the resident breeding male. Although the study sites differed considerably in area, the results suggested that movement between colonies by individual mole-rats was greater at the mesic (Somerset West) site (Bishop et al. 2004, 2007).

In this study, we use targeted genome-wide SNPs, baited to target ultra-conserved elements (UCEs). UCEs are highly conserved regions of the genome, and 481 UCE sites are shared among extremely divergent taxa (Bejerano et al. 2004). The flanking DNA regions of UCEs are highly variable, and the SNP data from these variable regions are useful for reconstructing population-level evolutionary histories (Smith et al. 2014). The role of UCEs is currently not fully known, however, they have been associated with gene regulation and organismal development (Sandelin et al. 2004; Pennachhio et al. 2006). UCEs are believed to be functional and important parts of the genome and are almost universally conserved across evolutionary distant taxa. Gene knockouts of UCE loci in mice produced viable, fertile offspring however, and thus, their role remains cryptic (Ahituv et al. 2007).

We tested hypotheses to infer the evolutionary processes and ecological drivers that underpin genetic variation across the distribution of *C. h. hottentotus* and determine population differences in genetic diversity and gene flow. Using mitochondrial DNA (cytochrome b) and genome-wide SNP data for 39 individuals from 29 colonies across five *C. h. hottentotus* populations varying in Aridity Index (two arid, one semi-arid, and two non-arid), we determined if genetic diversity is associated with aridity, and identified differences in population structure and movement patterns in response to aridity. We predicted that the evolutionary relationships of the different populations would indicate an arid–mesic split. We predicted that genetic differentiation would be linked to aridity, and that genetic diversity and gene flow would be higher in non-arid populations due to the greater prevalence of food resources, and thus a larger effective population size in the environment. Arid populations were predicted to be under stronger selection pressures due to the harsh environmental conditions and consequently have lower genetic diversity and increased directional selection for genes associated with arid adaption. Additionally, we predicted that the genetic structure of the populations would reveal separate clusters for the arid and non-arid populations and genetic distance was related to biome type (arid, semi-arid, and non-arid/mesic).

Results

Evolutionary Relationships and Mitochondrial Variation Along an Aridity Gradient

Cryptomys h. hottentotus mtDNA-cytochrome b gene (1,140 bp) sequences were obtained for 39 individuals representing five populations. The mitochondrial cytochrome b data demonstrated 69 sites recorded with SNPs. Overall mtDNA-cytochrome b nucleotide diversity (π) across all populations was low (0.014). Within the biomes, π was 0.0064 for arid populations, 0.0022 for semi-arid, and 0.0042 for mesic populations. The number of haplotypes was the highest in arid areas (9), and the lowest in semi-arid (5). Tajima's D statistic returned a negative value of −0.22, indicating a higher-than-expected number of low-frequency polymorphisms.

In total, 20 haplotypes were identified, and 19 out of 20 haplotypes (95%) were private (i.e. restricted to that population). These include all haplotypes from Somerset West, Darling, and Klawer. One haplotype (supplementary table S1, Supplementary Material online, Haplotype 1) was present in two different populations, Steinkopf and No Heep. Non-synonymous mutations which lead to a change in the protein sequence were found in 9 out of 20 haplotypes.

The hierarchical (nested) analysis of molecular variance (AMOVA) for cytochrome b (Table 1) revealed very low genetic variation (3.3%) and the highest fixation index

($\Phi ST = 0.97$) within populations. High genetic variation (32.7%) and a high fixation index ($\Phi SC = 0.91$) were found between populations. Both these results were statistically significant ($P < 0.001$). The highest differences (64.0%) and the lowest fixation index ($\Phi CT = 0.64$) are found among the biome groups; however, this result was not statistically significant ($P = 0.07$).

The Unrooted Minimum Spanning Network (Fig. 1) identified 20 cytochrome b haplotypes overall. Haplotypes that belong to individuals from the same populations clustered together and exhibited a lower degree of genetic differentiation from each other (most often one to two mutations), as indicated by numbers in parentheses along the branches. Similarly, the differentiation between arid-dwelling populations was lower compared to the mesic populations. This

finding is further supported by the maximum parsimony and maximum likelihood trees of the observed haplotypes (supplementary fig. S1a and b, Supplementary Material online). The private nature of all haplotypes except for Haplotype 1 is evident, given the geographical clustering by population (Fig. 1).

Genetic Diversity and Structure Patterns Among Populations as a Function of an Aridity Gradient

Of the genome-wide UCE data identified, 3,540 diploid nuclear loci for 39 individuals across five populations were found. In total, four out of five populations (all except Klawer) showed significant deviations from Hardy–Weinberg equilibrium (HWE) using a Chi-Squared test,

Table 1 Nested analysis of molecular variance (AMOVA) among mtDNA-cytochrome b sequences across six populations of *C. h. hottentotus*

Source of variation	Degrees of freedom	Variation (%)	Fixation indices	P value
Among biomes	2	64.0	0.63984	=0.066[NS]
Among populations	3	32.7	0.90742	**<0.001***
Within populations	30	3.3	0.96666	**<0.001***

Biomes are defined by the aridity categories based on the Aridity Index for each population (arid, semi-arid, or mesic). Significant results are in bold and marked with an asterisk (*). Nonsignificant results are denoted by NS.
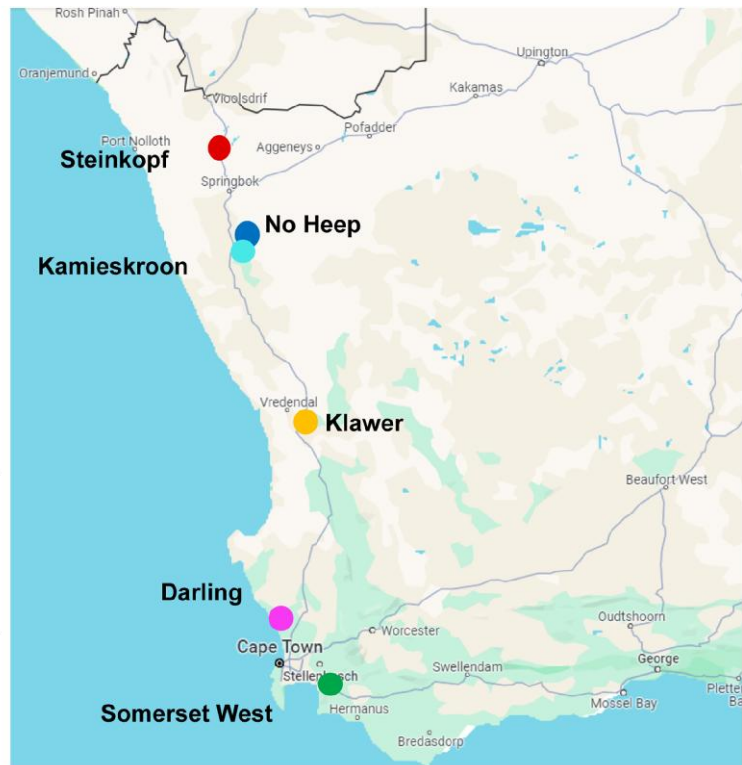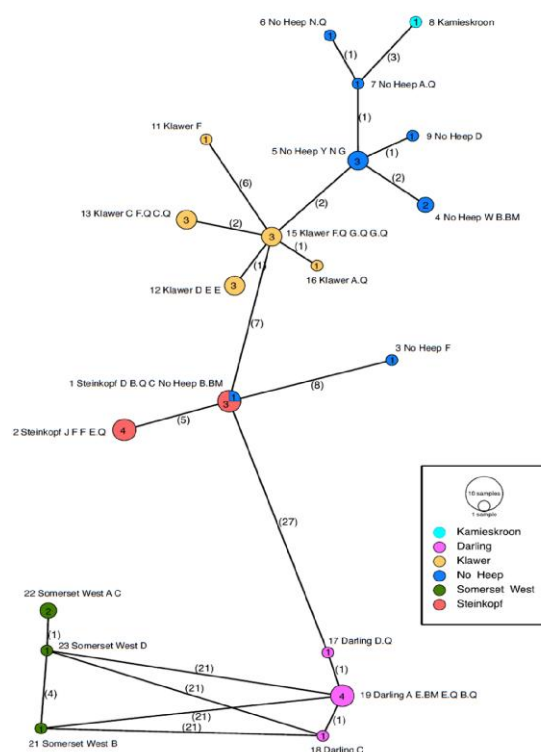

Fig. 1. Minimum Spanning Network for six populations of *C. h. hottentotus* haplotypes. Circles represent haplotype, and size and numbers within the circles indicate the number of individuals that exhibit each respective haplotype. Populations are color-coded. Branch numbers indicate the number of mutations between each haplotype. Branch length is non-informative, and the network is unrooted. Geographical distribution of the six populations of *C. h. hottentotus* is presented on a map of South Africa. See supplementary table S1, Supplementary Material online for haplotype naming system.

Table 2 Genetic diversity results for five populations of *C. h. hottentotus*

Population	HWE		LD	Heterozygosity			A_r	Private alleles	F_{is}
	Chi-Squared	MCMC	r_D^2	H_o	H_e	uH_e			
Steinkopf	0.029	0.008	0.136	0.215	0.197	0.21	1.191	25	−0.082
No Heep	0.03	0.014	0.03	0.215	0.21	0.221	1.205	132	−0.018
Klawer	0.092	0.011	0.122	0.186	0.22	0.237	1.209	109	0.148
Darling	0.028	0.01	0.238	0.129	0.137	0.147	1.13	57	0.037
Somerset West	0.028	0.012	0.074	0.121	0.107	0.115	1.104	283	−0.121

For each population, the values are listed for Hardy–Weinberg equilibrium (HWE) significance results for a Chi-Squared test, and MCMC permutation test, linkage disequilibrium (LD) r_D^2 values, observed heterozygosity (H_o), expected heterozygosity (H_e), unbiased expected heterozygosity (uH_e), allelic richness (A_r), number of private alleles, inbreeding coefficient (F_{is}), and number of individuals per population (n).

and all five populations showed significant deviation from HWE using a MC permutation test (Table 2). The total fraction of locus pairs with positive r_D^2 across populations and linked loci was 0.54, which was higher than expected under random association. Deviations from linkage equilibrium were significant for all populations (values higher than 0), and linkage disequilibrium was highest for the mesic population, Darling (la = 265.375, $r_D^2 = 0.238$, $P < 0.05$), and lowest in the arid population, No Heep (la = 64.207, $r_D^2 = 0.03$, $P < 0.05$). For all populations, the unbiased heterozygosity (uH_e) ranged from 0.147 to 0.237. The inbreeding coefficient (F_{is}) values were above zero in Klawer and Darling, indicating that these populations have lower individual variation, and thus a deficiency of heterozygotes. Rarefied allele richness (A_r) varied from 1.104 to 1.209, and the number of private alleles ranged from 25 to 283. All results are summarized in Table 2. The models identified that genetic diversity measured as A_r and heterozygosity decreased significantly with an increase in Aridity Index (AI) (supplementary fig. S2, Supplementary Material online), and the number of private alleles exhibited a quadratic relationship with AI, revealing a higher number of private alleles at extreme aridity.

Aridity as a Driver Underpinning Genetic Differentiation Patterns

The AMOVA showed that molecular variance was greatest among samples (74.6%) compared to populations (8.86%) and colonies (3.13%) (Table 3). Genetic differentiation (F_{st}) under Nei for the data was 0.268 and found to be significantly different ($P = 0.001$) from the F_{st} value expected under HWE. F_{st} values between the populations ranged from 0.041 to 0.252 (supplementary table S2, Supplementary Material online). Steinkopf and No Heep were the most genetically similar populations (0.041). The isolation models suggest a marginal positive correlation between genetic distances and geographical distances (F = 31.77, beta = 0.7835, df = 2, 17, $P = 0.048$) and aridity index (F = 31.77, beta = 6.2386, df = 2, 17, $P < 0.001$). A second model testing isolation by environment using the residuals from a model of genetic distance against geographical

Table 3 Hierarchical analysis of molecular variance (AMOVA) for SNPs across five populations of *C. h. hottentotus*

Test	Variance (%)	Standard deviation	P value
Variations between samples	74.594	4.428	**0.001***
Variations between colony	2.651	−0.111	0.514[NS]
Variations between population	7.506	1.048	0.151[NS]

Population and within-population colony assignments are predetermined based on provenance of collected samples. Significant results are in bold and marked with an asterisk (*). Nonsignificant results are denoted by NS.

distance identified a significant correlation (F = 12.16, beta = 0.37, df = 1, 18, $P = 0.0026$). Both models suggest a marginal correlation between genetic distance and geographical distance, but that isolation by the environment (IBE) explains the genetic differences between the populations (Fig. 2).

Aridity Gradient Shapes the Distribution of the Genetic Variation Among Populations

All clustering analyses (STRUCTURE, PCA, DAPC) concurred to show two main genetic clusters that encompassed (i) the two mesic populations, and (ii) the semi-arid and two arid populations. The STRUCTURE analysis showed two clusters (K = 2) with a mesic group encompassing the two mesic populations (Somerset West and Darling), and an arid group containing the two arid populations (Steinkopf and No Heep) and the intermediate, semi-arid population (Klawer). The assignment probability on an individual basis is listed in supplementary table S3, Supplementary Material online, and DeltaK values estimated using the Evanno methods (supplementary fig. S3, Supplementary Material online) from STRUCTURE results are shown in Fig. 3a and c.

The PCA results demonstrated the presence of two clusters, which is in line with the STRUCTURE results (Fig. 3b), and the DAPC results showed significant distinction between all five populations (supplementary fig. S4, Supplementary Material online). The results of the spatial
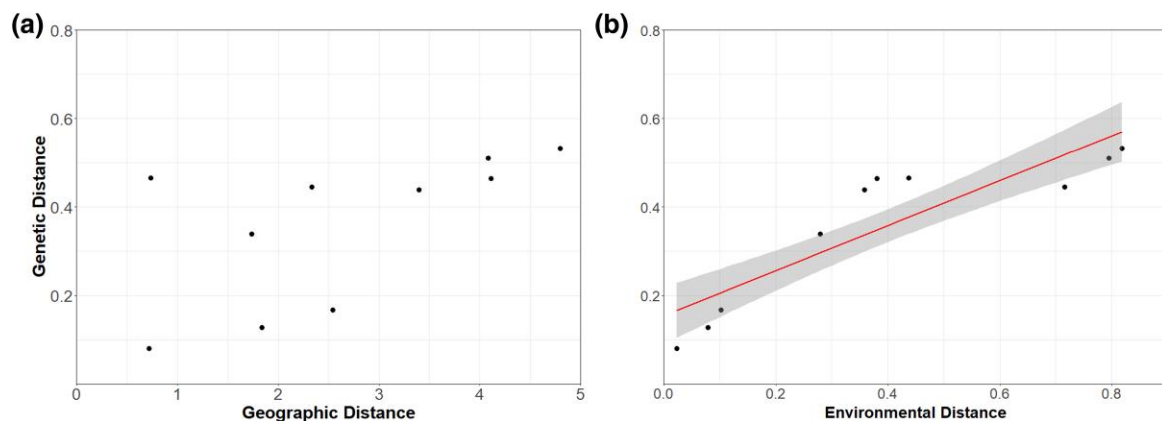
Fig. 2. Scatter plot for a) isolation by distance shows a relationship between the genetic (F_{st}) and geographic distances of the population pairs ($P = 0.048$), and b) isolation by environment shows a significant and linear relationship between the genetic (F_{st}) and environmental distance (Aridity Index) of the population pairs ($P < 0.001$). The linear relationship is depicted along with 95% confidence intervals.

PCA (sPCA) and Moran's I test supported the cluster analyses results and demonstrated the existence of a significant genetic structuring ($P = 0.048$) across the five populations. The sPCA did not detect a substructure within any of the groups.

Discussion

Overall, the results from this study suggest that genetic diversity and population structure correlate with the environmental Aridity Index. The results demonstrate evidence of genetic differentiation between arid and mesic populations, represented in both the mitochondrial and nuclear data. Higher levels of genetic diversity are detectable in arid populations when compared to mesic populations based on the nuclear data.

Evolutionary Relationships and Mitochondrial Variation Along an Aridity Gradient

A negative Tajima's D implies that a historic rapid population expansion (Dangalle et al. 2014) or a positive selection of certain alleles, known as selective sweep (De Jong et al. 2011), might have taken place. This is particularly evident in mesic-dwelling populations, suggesting that historically, *C. h. hottentotus* expanded from the north to the south and that mesic-dwelling individuals may have derived adaptations to mesic conditions, and arid populations retained ancestral haplotypes linked with arid adaption.

Genetic Diversity and Structure Patterns Among Populations as a Function of an Aridity Gradient

HWE was not present in any of the populations apart from Klawer under the Chi-Squared test, thus suggesting the nuclear SNPs may be under selection, and that gene flow and genetic drift may be present, or that nonrandom mating is occurring, due to the presence of colonies within the populations (Abramovs et al. 2020). Indeed, the lack of

HWE is further supported by the higher-than-expected presence of linkage disequilibrium found in all populations (Slatkin 2008). Due to the different ecological requirements in arid and non-arid environments, allele frequencies may be under different selection pressures, resulting in different genotypes (Orr 2005; Hoffmann and Willi 2008). However, due to the relatively small population sample sizes used in this study, there may be an increased effect of genetic drift. The smaller sample size of representative genomes may result in a random skew in allele frequencies, thus presenting with lower genetic variation than is present in the natural systems (Nei and Tajima 1981) and the fixation of certain alleles with no beneficial qualities in a population's adaptive response to its surrounding environment.

Heterozygosity was found to increase with aridity, i.e. as conditions become more arid, higher levels of heterozygosity are observed compared to mesic populations. We can thus infer that populations are variable and that habitat qualities such as aridity do play a role. Higher heterozygosity levels infer the presence of greater degrees of admixture in the arid region (Boca, et al. 2020). Greater admixture indicates more movement and greater occurrences of disperser events in arid regions between populations and colonies and thus, more gene flow. This pattern, however, is not consistent with the known examples of philopatry and dispersal differences between arid and mesic areas. The AFDH (Jarvis et al. 1994; Faulkes et al. 1997a; Spinks et al. 2000a; Spinks et al. 2000b; Jacobs et al. 2022) suggests that arid-dwelling *C. h. hottentotus* are more constrained in dispersal due to inhabiting a lower food-density environment. Therefore, movement has a greater energetic cost, thus, arid populations have significantly lower immigration and emigration rates and stronger social cohesion (Spinks et al. 2000a; Spinks et al. 2000b), in addition to being more xenophobic (Spinks 1998). Combining these factors is more likely to curtail the gene flow between
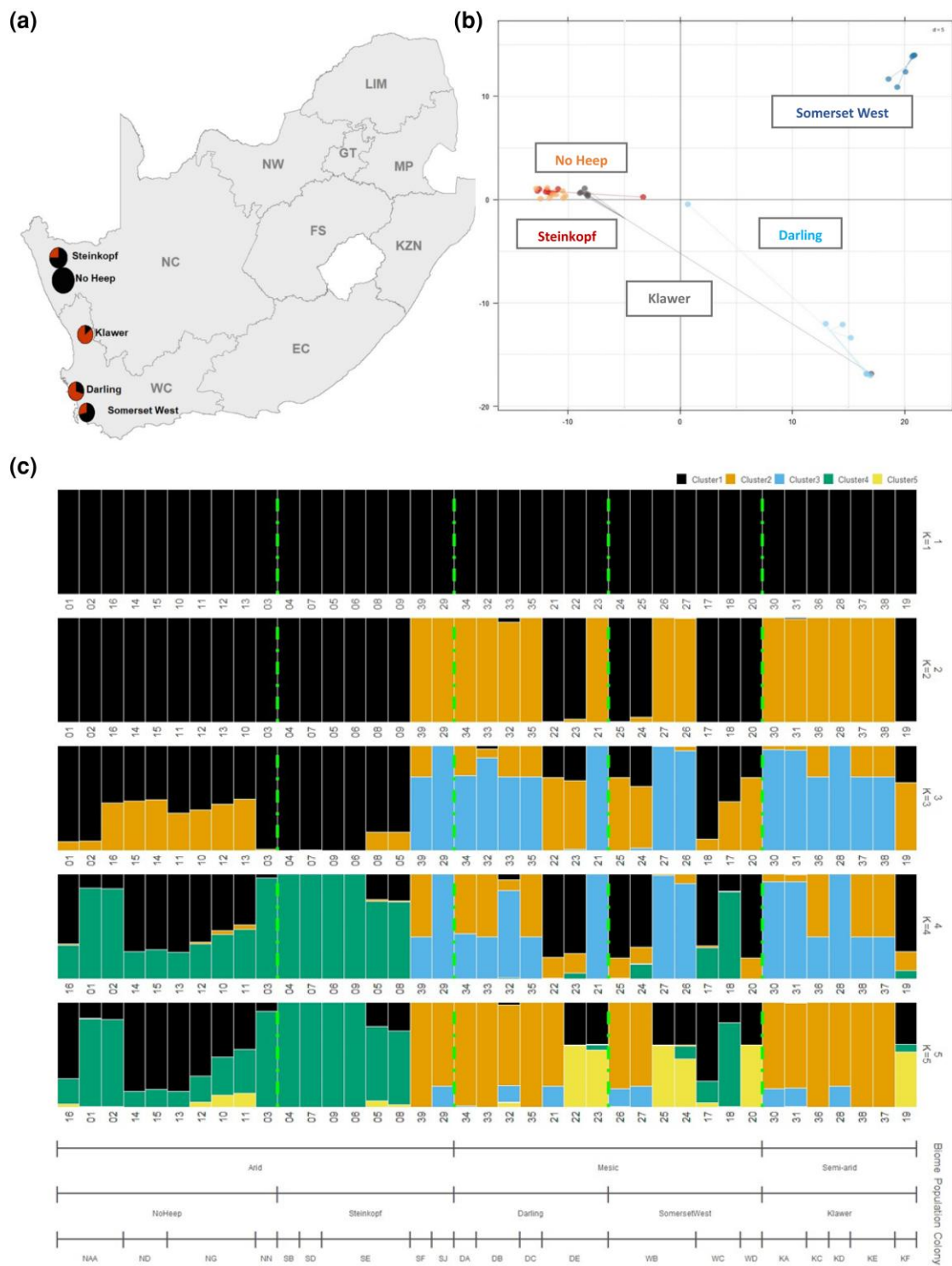
Fig. 3. a) Geographical distribution of common mole-rat populations across the Northern Cape (NC) and Western Cape (WC) of South Africa sequenced in this study. Pie charts show the proportion of K = 1 and K = 2 present in each population according to STRUCTURE analysis. Number of individuals is represented by the size of pie charts. b) Principal component analysis of 3,540 SNP loci for 39 individuals of *C. h. hottentotus* across five populations, Darling, Somerset West, Klawer, No Heep, and Steinkopf. c) Bar plots showing the STRUCTURE results for the first five values of K. Cluster (K) colors are listed in the top right corner. Samples are grouped by biome (arid, mesic, and semi-arid), then by population (Steinkopf, No Heep, Klawer, Darling, and Somerset West) within that, and further separated by colony. Dashed lines separate populations. A box highlights the K value preferentially selected by STRUCTURE analyses (K = 2).

populations, rather than promote it. Conversely, the AFDH states that ecological constraints in arid habitats promote the evolution of cooperative foraging and larger social groups. Indeed, these same arid conditions have been shown to drive arid-dwelling populations to dig longer burrow systems to compensate for low geophyte density (Spinks et al. 2000b) and arid populations are predicted to have increased numbers of individuals per colony, though this is yet to be demonstrated significantly (Spinks and Plagányi 1999; Spinks et al. 2000a). The greater colony size may increase genetic diversity, and larger burrow systems may potentially increase the number of encounters between the colonies and populations, providing opportunities for gene flow. Additionally, whilst there may be a case of reduced dispersal in arid populations, following rainfall events, individuals may disperse further than in the mesic populations, creating pathways for admixture (Lovegrove 1991; Spinks and Plagányi 1999). The connectivity of colonies in arid regions may be higher than expected, and there may be an effect on the distribution of colonies and individuals clustering around food sources of high geophyte abundance, thus leading to a greater chance of interactions.

The observed heterozygosity (H_o) was higher than the expected (H_e) for Steinkopf, No Heep, and Somerset West. However, when compared to the coefficient, uH_e, which takes into account the small effective population sizes, the H_o was higher for Steinkopf and Somerset West only. Higher than expected levels of heterozygosity again implies that these populations are breaking the assumptions of HWE, and can indicate an isolate-breaking effect, which is the mixing of two previously isolated populations (Namipashaki et al. 2015). As Steinkopf and Somerset West are populations found at the most distal ends of the environmental continuum, they are likely experiencing the extremes of this environmental cline and, therefore, may be under greater levels of selection than the other populations (Spinks et al. 2000b). These populations also may represent range edge populations, and thus have the potential for greater movement into new areas, and higher levels of admixture with populations of neighboring subspecies compared to the other populations (No Heep, Klawer, and Darling) which demonstrate lower-than-expected levels of heterozygosity. Individuals from No Heep presented with little variation in heterozygosity compared to all other populations, potentially explained by genetic isolation due to the surrounding mountain ranges around No Heep. Darling and Somerset West populations (mesic dwelling) are located in a currently densely populated area with Cape Town in between, which might act as an anthropogenic barrier preventing admixture and gene flow. The arid regions of the Northern Cape are much more sparsely populated and thus have fewer barriers to dispersal and movement. The precise distributional ranges of the populations and more

broadly the *Cryptomys* subspecies are not fully understood, and the presence of hybrid zones is unknown, but likely (Finn 2022).

The number of private alleles increased with both extreme high and extreme low AI. These results suggest that private alleles are accumulating in extreme arid and mesic environments, but less so in intermediate, semi-arid conditions. High numbers of unique loci in environmental extremes indicates selection for environmental adaptation, and thus these loci may be associated with thermal tolerance or osmoregulatory gene regions. A higher number of private alleles can also indicate past or recent genetic isolation in No Heep and Somerset West. A_r decreased significantly with an increase in AI, and thus was significantly lower in the mesic populations. Low A_r can indicate the presence of past genetic bottlenecks or reductions in population size and can suggest that these populations are particularly at risk of rapidly changing climates. Additionally, mole-rats have long generation lengths, so the individuals have a slower turnover and thus the movement of alleles within a population is much slower than other similarly sized rodents (Can et al. 2022). Different levels of linkage disequilibrium observed between the populations indicate the potential for populations to be independent, however the values did not differ greatly, so this remains uncertain. Linkage disequilibrium can also infer the presence of past bottleneck events occurring in the populations (Slatkin 2008) and may influence the heterozygosity results. Both of these findings indicate that the results from the program STRUCTURE need to be carefully considered (Pritchard et al. 2000), however the STRUCTURE analysis corroborated with PCA and DAPC and as a consequence we can infer that the STRUCTURE results can be considered in our determination of clusters.

Aridity Gradient Shapes the Distribution of the Genetic Variation Among Populations

The isolation by environment model was significant implying that genetic distinction is potentially linked to environmental factors, namely Aridity Index (AI). The residual variation in genetic distance against geographic distance was explained by aridity index, thus isolating the impact of AI once the impact of distance and autocorrelation had been accounted for, further supporting the isolation by environment model. Genetic distance is virtually identical at Steinkopf and No Heep but differs more between Somerset West and Darling. These differences in genetic distance are reflected in the differences in the AI at these sites, thus supporting the idea that aridity may influence genetic distinction. However, as isolation by geographic distance was on the threshold of significance ($P = 0.048$), it is worth considering that geographical distance may be impacting genetic distance. Furthermore, the small sample sizes in this study may be reducing the impact of both factors. As the sample sites were located in a linear

transect, geographic distance contributes to differences in aridity, so it remains difficult to tease apart the geographic and environmental impacts.

The results from the AMOVA indicate that genetic variance was greatest between individuals and lowest between colonies. The dispersal strategies of mole-rats naturally lead to high levels of relatedness between colonies. Dispersers move from a colony and migrate, likely to the next available colony, suggesting that colony proximity plays a role in genetic diversity (Greeff and Bennett 2000). This importance of colony proximity can lead to high levels of genetic homogeneity between colonies, with neighboring colonies containing relatives of each other (Spinks et al. 1999). Genetic variation between the populations was lowest indicating that population structures are not highly distinct and potentially not entirely independent. The high variation between individuals, however, suggests that relatively high levels of diversity are present in the species as a whole, with the presence of varied alleles that may or may not be linked to habitat adaptation. The high number of individuals from each population with genetically similar origins may again be due to the nontraditional, nonrandom mating system of the social common mole-rats.

The population statistic results are further supported by the cluster analyses and mitochondrial data, all of which suggest the presence of distinct mesic and arid clusters. The intermediate population, Klawer, share alleles with both arid and mesic clusters, but shares greater genetic similarities with the arid populations, which corresponds with the support for isolation by environment, as the AI for the site Klawer, is closer to the AI for the two arid populations, than the two mesic. These analyses demonstrate the multiple influences, proximity and environmental similarities, that effect gene flow and connectivity. Cluster analyses can be deceptive with populations that are not truly independent, for example, those found on a continuum or along an environmental cline (Lombaert et al. 2018). Thus, multiple cluster analyses were used in order to strengthen the results. Additionally, the known provenance of the individuals used in this study, and the colonial nature of this species, can also provide support for the results from these analyses.

Limitations and Future Work

The populations used in this study form a linear continuum and therefore the genetic diversity exhibited in our results reflects diversity along one dimension. As such, future work would benefit from the sampling of populations that vary in longitude as well as latitude. Further work on population structure and dispersal, investigating populations closer in proximity, is needed. A follow-up study using increased numbers per population from many colonies would provide a better representation of the natural populations that exist in the wild. This present study provides a baseline against which to unravel different colony roles, such as queens and breeding males to provide an understanding of the individual colony structures in each population, which may infer and explain the movements, or lack thereof between colonies in each population as well as further understand how they differ in the different habitat types. Finally, the next steps from this study would be to investigate variation in target loci for environmental and thermal adaptation to determine specific environmental-based differences between the populations.

Nonrandom mating can raise the observed frequency of homozygous alleles in the overall population to higher than expected under HWE, known as the Wahlund effect (Sinnock 1975). The mating system of social mole-rats creates an effect of subpopulations, or colonies, in which only certain individuals, the queen and breeding males, will mate with each other and produce offspring, forming colonies of closely related sibling and parent family groups (Bennett and Faulkes 2000). It is thus possible that incest avoidance is operational in common mole-rats and the evidence of inbreeding may simply result from sampling family members. The population inbreeding and heterozygosity values did not differ greatly when only one individual per colony was used. Nevertheless, the use of F_{is} can be misleading when considering colonial species that have nontraditional mating systems, as present in common mole-rats, confounding the inbreeding coefficient. Often, in the case of African mole-rats, the dispersers are (potential) breeding males, so these are likely to be the more genetically distinct individuals in a colony and increase genetic diversity (Torrents-Ticó et al. 2018). "Floater males" are also common in mole-rat populations, and thus paternities in colonies may not be attributed to any colony residents (Bishop et al. 2004). As breeding individuals were not included extensively in this study and may be important in further studies to gain a more in-depth picture of genetic diversity of common mole-rats. Interestingly, the Klawer population had two colonies, F and G, with two queens each. The presence of multiple reproductive females in one colony is rare for *C. h. hottentotus*. Multiple queens per colony have been previously encountered in mesic sites, but has, as of yet, never been formally reported in arid sites (Spinks et al. 2000a). Klawer is classified semi-arid and in colony G both queens had the same haplotype, which is insufficient to establish the matrilineal relatedness. In colony F, the haplotypes were different, suggesting that the two queens come from different maternal lineages, highlighting the importance of exploring breeding and nonbreeding individuals.

Conclusions

Overall, our results demonstrate a relationship between environment and genetic diversity and its distribution across

populations at a regional scale. This study shows that mole-rats in arid environments have increased genetic diversity and greater connectivity between populations than mesic populations. Genetic diversity may be greater in arid populations as a consequence of resource distribution and dispersal patterns, as well as human settlement distribution. Mesic populations are therefore potentially more at risk in a changing climate due to their low genetic diversity and low initial allelic composition. The potential for loss of distributional range for this species could thus be high, and therefore further investigation of their population genetics, and specific adaptations may influence conservation policies that need to be put in place to ensure the survival of this species and the various potential morphotypes that exist.

Methods

Study Species

Bathyergid mole-rats vary in social strategies from solitary to truly social or eusocial, living in colonies of closely related family groups (Bennett et al. 1990). Populations of the co-operatively breeding common mole-rats (C. h. hottentotus) present an ideal model system for the study of mammalian population genetics as the distributional range of this subspecies spans from the arid succulent Karoo in the Northern Cape through to the mesic Fynbos of the Western Cape (Spinks et al. 2000a).

Study Site

The sites were specifically selected to represent an aridity gradient based on an Aridity Index (AI), a numerical indicator of the degree of dryness of the climate at a given location (UNEP, 1992). The AI for the study populations was calculated from climatic data (ranging from the years 1981 to 2020) retrieved from the ERA5-Land of the European Centre for Medium-Range Weather Forecasts, created by the Copernicus Climate Change Service (Muñoz-Sabater et al. 2021) with a spatial resolution of 0.1° by 0.1°. Monthly averaged temperature (T_{air} in °C), total precipitation (t_p in m), and dew point temperature (d2m in °C) were used. These combined data were used to calculate the annual aridity index (AI) (equation (1)). Where t_p is

directly obtained from ERA5-Land and potential evapotranspiration (PET) is calculated from the Romanenko estimation (equation (2)) (Romanenko 1961). For equation (2), relative humidity (RH) was calculated from ERA5-Land d2m (equation (3)).

$$AI = \frac{t_p}{PET} \qquad (1)$$

$$PET = 0.00006 \times (100 - RH) \times (25 + T_{air})^2 \qquad (2)$$

$$RH = 100 \times 10^{7.591386\left(\frac{d2m}{d2m + 240.7263} - \frac{T_{air}}{T_{air} + 240.7263}\right)} \qquad (3)$$

Aridity classifications and corresponding AI values, as outlined by UNESCO (1979) and UNEP (1992), state that where PET is greater than t_p, the climate is considered to be arid (Colantoni et al. 2015). AI values at each of the five sites used in this study are listed in Table 4 and have been selected to range in AI across different aridity classifications. The two municipalities of Steinkopf and No Heep are considered arid populations, Klawer a semi-arid/intermediate, and the municipalities of Darling and Somerset West mesic populations.

Sample Collection

Samples of common mole-rats were collected from 39 individuals across five sites, from between seven and four colonies per site (supplementary table S1, Supplementary Material online). Mole-rat individuals were trapped using Hickman live traps (Hickman 1979) inserted into tunnels located under mounds and were baited with sweet potato. All sites have been previously documented as having common mole-rats (Spinks et al. 2000a; Visser et al. 2019; Hart et al. 2023). The animals used in this study were maintained in captivity for between 22 and 28 days prior to being euthanized using an overdose of isoflurane in line with strict veterinary procedures. Individuals were collected from multiple colonies within each of the five populations (two arid, one semi-arid, and two non-arid). The Animal Use and Care Committee of the University of Pretoria evaluated and approved experimental protocol (ethics clearance no. NAS016/2021) and DAFF section 20 approval (SDAH-Epi-21031811071). Tissues were

Table 4 List of five sites, sample size for both the mitochondrial dataset (cyt-b), and whole genome dataset (WG) with collection coordinates for individuals of common mole-rats, recent Aridity Index values (from the year 2020) for each site based on climate data taken from ERA5-Land dataset to two decimal places, and aridity classifications are listed

Site	Sample size: cyt-b/WG	Latitude	Longitude	Aridity index	Aridity classification
Steinkopf	7/8	−29.34531	17.7872	0.04	Arid
No Heep	9 (+1 Kamieskroon)/10	−30.04253	17.95852	0.07	Arid
Klawer	11/7	−31.7013476	18.7446117	0.11	Semi-arid
Darling	6/7	−33.406573	18.417538	0.42	Mesic
Somerset West	4/7	−34.035613	18.799499	0.86	Mesic

then harvested by cutting open the abdomen and chest cavity biceps femoris muscle tissue was extracted and fixed in 70% ethanol and stored at −20 °C prior to extraction.

DNA Extraction

Extractions were first carried out using the QIAGEN Dneasy Blood and Tissue Kit as specified by the manufacturer. The QIAGEN Kit extraction method yielded low DNA concentration values, thus a second set of extractions were carried out on the samples using a standard salt/chloroform extraction protocol as outlined by Müllenbach et al. (1989). A full list of analyses performed, and sample information is given in supplementary table S4, Supplementary Material online.

Mitochondrial DNA Sequencing

PCR amplification of the entire cytochrome b gene (1,140 bp) was carried out using primers (forward: L14724, reverse: H15915) described previously in Faulkes et al. (1997a, 1997b). Sanger sequencing was carried out in both directions using the Eurofins Genomics Value Read provided by Eurofins MWG Operon Inc. (Eurofins Scientific, Luxembourg, Luxembourg) to obtain complimentary and partially overlapping strands for 39 individuals.

Sequence alignment was performed manually in Mesquite version 3.70 (Maddison and Maddison 2021). Sequences of other African mole-rat species were obtained from GenBank (Clark et al. 2016), and used as references for sequence alignment, namely *Heterocephalus glaber* accession no. MT8453841 (Zemlemerova et al. 2021), *Fukomys damarensis* accession no. AF012223.1 (Faulkes et al. 1997a), and *C. h. hottentotus* accession no. MH186559.1 (Visser et al. 2019). One additional *C. h. hottentotus* sequence was added to the database from Kamieskroon (30°13′09.0″S 17°56′20.1″E), a site located near No Heep. This sequence was obtained from an old sample collected in 2016 for a previous study held at Queen Mary University of London.

Analysis of Evolutionary Relationships and Mitochondrial Variation Along an Aridity Gradient

The mitochondrial data consisted of sequences from 39 samples (supplementary table S4, Supplementary Material online and one additional sample from Kamieskroon), and one reference sequence for *F. damarensis* as an outgroup. Phylogenetic trees were constructed using Molecular Evolutionary Genetic Analysis version 11 (MEGA11; Tamura et al. 2021; Stecher et al. 2020). The Maximum Parsimony (MP) method for analysis of taxa (Nei and Kumar 2000) with the Subtree-Pruning-Regrafting (SPR) algorithm (Nei and Kumar 2000) and the Maximum Likelihood (ML) method using Hasegawa–Kishino–Yano model (Hasegawa et al. 1985) were used. To estimate the robustness of the tree topology, bootstrap analysis was conducted with 1,000 replicates of the dataset (Felsenstein 1985).

A Minimum Spanning Network (Bandelt et al. 1999) analysis was used to determine the haplotypes present in the mtDNA full sequence dataset containing 39 sequences. PopART version 1.7 (Leigh and Bryant 2015) was used to tag sequences by population, which were then grouped into three biomes according to respective Aridity Index (Steinkopf, No Heep, Kamieskroon as arid, Klawer as semi-arid, Darling and Somerset West as mesic). A hierarchical (nested) AMOVA (Excoffier et al. 1992) was performed to calculate Φ-statistics and identify geographical patterns of genetic differentiation between haplotypes. Significance was calculated for 1,000 permutations. Values for nucleotide diversity (π), defined as average pairwise differences per site for all sequences, and Tajima's *D* statistic for population neutrality (Tajima 1989) were obtained, and a map of the geographical distribution of haplotypes was drawn from haplotypes data in PopART.

UCE Targeted Sequence Capture

Probes were designed and synthesized by Daicel Arbor Biosciences, Ann Arbor, MI, USA. Genome libraries were prepared using the NEBNext Ultra II FS DNA Library Prep Kit for Illumina TruSeq (New England BioLabs, Inc., Ipswich, MA, USA), as per the manufacturer's recommended protocol for fragmentation system DNA library preparations with inputs ≥ 100 ng. Fragment sizes of between 300 and 500 bp were required and therefore, Size Selection protocol using AMPure XP Beads was included, and specific conditions within the protocol were selected based on this fragment size range as per the Library Prep Kit protocol, and as requested by Daicel Arbor Biosciences, Ann Arbor, MI, USA. Libraries were pooled in equimolar proportions into six pools prior to shipping to Daicel Arbor Biosciences. Total DNA was quantified via a spectrofluorimetric assay, and the mass of molecules per library was quantified via a quantitative PCR assay. Captures were performed following the myBaits v5.02 protocol (Daicel Arbor Biosciences, Ann Arbor, MI, USA) using myBaits UCE Tetrapods 5Kv1 baits with an overnight hybridization and washes at 65 °C. Post-capture, the reactions were amplified for ten cycles and were quantified again with a spectrofluorimetric assay. The captures were pooled in approximately equimolar ratios (supplementary table S5, Supplementary Material online). Samples were sequenced on the Illumina NovaSeq 6000 platform on a partial S4 PE150 lane to approximately 1 Gbp per library. Demultiplexed fastq data were returned with successful sequence reads for 39 samples (supplementary table S4, Supplementary Material online).

Bioinformatics for UCE SNP Loci

Using the Python package, PHYLUCE Version 1.7.2 (Faircloth 2016), through the Linux distribution, Ubuntu version 22.04.1 (Sobell 2015), UCE data were processed, contigs assembled and loci aligned as per the protocol outlined by Faircloth (2016). Sequences were trimmed of adaptor contamination

and low-quality bases using Illumiprocessor. Cleaned reads were inspected with fastQC to ensure only high-quality reads remained. Contigs from each individual were subsequently assembled from the cleaned reads to a reference set of UCE loci using the bwa-meme algorithm for each sample against the composite reference. Alignment included cleaning, validating, and marking duplicates that were then removed in the resulting Binary Alignment/Map (BAM) file. The reference genome used was from *F. damarensis* since no genomes have been sequenced within the genus *Cryptomys*. Extracted contigs were taken for each UCE locus, and depth and quality were assessed. The mapped reads were indexed. Using the probes to create a bed file, through bedtools to restrict to target regions only. Depth values were extracted, and sequences filtered to only select SNPs for each of the UCE loci. Sequences were filtered using VCFtools, creating a variant call format (VCF) file. The VCF file was filtered to retain a minimum genotype depth of 5 and a maximum genotype depth of 15 (mean depth for 50% and 80% of individuals) for SNPs successfully genotyped in ≥80% of the individuals. The parameters created a complete matrix with a minimum genotype quality (GQ) of 20. The SNPs were then pruned to only one SNP per locus to prevent extreme effects of linkage disequilibrium and to provide a set of independent markers to compete the VCF file.

Genetic Diversity and Structure Analyses Among Populations Along an Aridity Gradient

Calculations and statistical analyses were performed using the statistical software R version 4.2.2 (R Core Team 2021). We compared population diversity among population using eight estimates of genetic diversity and genetic structure (Table 2) available in *adegenet* (Jombart 2008) and *hierfstat* (Goudet 2005) R packages. We test for deviations from the HWE for each population by both a Chi-Squared test and Monte Carlo permutation test with 1,000 replicates. Linkage disequilibrium values (r_D^2) were also determined for each population using 1,000 replicates, and for each pair of loci, to create an average value for our UCE loci. We used estimates of expected and observed heterozygosity (H_e and H_o, respectively) and unbiased expected heterozygosity (uH_e) to metrics of population-level genetic diversity, uH_e is defined by equation (4). We also estimated rarefied allele richness (A_r), and the number of private alleles unique to each site as proxies of genetic diversity. The inbreeding coefficient (F_{is}), calculated using equation (5), measured the proportion of the variance for each individual within a subpopulation and it is typically used as a proxy of gene flow among subpopulations, with high values of F_{is} implying low gene flow levels. Linear mixed effects models with binomial distributions (heterozygosity and F_{is}) and general linear models (allelic richness, private allele number) were used to determine

differences between the populations for each of the parameters and this was carried out on the full dataset and a subset of one individual per colony present in the study to explore the effects of family groups.

$$uH_e = H_e \times \frac{2 \times \text{Population size}}{2 \times \text{Population size} - 1} \quad (4)$$

$$F_{is} = \frac{1 - \text{mean}(H_o)}{\text{mean}(uH_e)} \quad (5)$$

Analysis of Genetic Structure in Response to Aridity

An AMOVA was carried out based on 999 replicates to assess the variation in genetic diversity both among the populations and within populations, through variation between individuals (samples), and variation between the different colonies found within a population. Genetic differentiation across the populations (F_{st}) was assessed using the R package *hierfstat* (Goudet 2005). F_{st} was calculated under a Nei distribution (Nei 1978). We tested for isolation by distance (IBD) and IBE (Wang et al. 2013), in our study system through models using the pairwise Nei's genetic distances, geographic distances based on longitude and latitude for each site, and a distance matrix of Aridity Index for each site (environmental distance) (Table 4). Using the R package *spaMM*, isolation patterns were first tested using a linear model using matrices for genetic distance as the dependent variable, and geographic distance and environmental distance as dependent variables. A second test was carried out by extracting residuals from a model of genetic distance against geographical distance, and modeling this against environmental distance.

Cluster Analysis of Populations Along an Aridity Gradient

We test for clustering patterns and whether aridity underlies observed patterns by applying an equilibrium-based model (STRUCTURE) and nonequilibrium-based clustering models (PCA and DAPC). The VCF file was converted into STRUCTURE file format using PGDSpider version 2.1.1.5 (Lischer and Excoffier 2012). STRUCTURE version 2.3.4 software (Pritchard et al. 2000) was first used to carry out clustering analyses. STRUCTURE uses Bayesian model-based methods that presume HWE and linkage disequilibrium, analyzing the individual-level multilocus genotypes and assigning individuals to clusters based on their assignment probability. We set a burn-in period of 50,000 iterations and 200,000 iterations with Markov Chain Monte Carlo (MCMC). An admixture model was assumed, and geographic location was considered in the analysis. K values ranged from 1 to 21, and ten runs per K value. We applied the Evanno method (Evanno et al. 2005) to select the most likely K value and as implemented and visualized using the package *pophelper* in R (Francis 2017).

STRUCTURE analyses can be confounded by populations that are not truly independent, for example, with populations found along an environmental cline. To strengthen the results, further clustering analyses were carried out using the *adegenet* R package, to combine different methodological approaches and compare results from equilibrium-based models (STRUCTURE) and nonequilibrium-based models (PCA and DAPC). A principal component analysis (PCA) was used to confirm the clustering of individuals into distinctive groups. A PCA synthesizes the multilocus data to create a multivariate space where individuals are mapped according to their similarities and determine principal components that best explain the data. Finally, a distance-based discriminant analysis of principal components (DAPC) was used. DAPC build genetic clusters from genetic data by combining the alleles. Neither the PCA or the DAPC require assumptions of HWE or linkage disequilibrium. The most likely K value for the DAPC was selected using the Bayesian information criterion (BIC) and the optimal K value identified through an accompanying decrease in the BIC.

Due to the evident spatial patterns in the PCA, a spatial PCA (sPCA) was carried out using *adegenet*. A sPCA uses Moran's I to incorporate spatial autocorrelation between samples so that genetic structure can be modeled without a priori population assignment and allows for visualizing population structure in space. This analysis method, similar to a PCA, does not assume HWE or linkage disequilibrium. Longitude and latitude for each colony within a population were used, and unique individual coordinates were created with the "jitter" function in R to develop unique x and y values. sPCA eigenvalues were tested for global and local structures, which correspond to positive spatial autocorrelation between individuals, such as patches, clines and intermediates, and strong genetic differences between neighbors, using 1,000 Monte Carlo tests replicates.

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

Author Contributions

Conceptualization, methodology, and planning, H.N.M., S.J.P., C.G.F., D.W.H., and N.C.B. Data collected by H.N.M., A.I., D.W.H., and N.C.B. Methodological development, H.N.M., C.G.F., C.G., D.W.H., S.J.P., and N.C.B. Data analysis was performed by H.N.M. and A.I. All authors contributed to the intellectual interpretation of the results. The initial draft of the paper was written by H.N.M. and A.I., with contributions and edits by all authors.

Conflict of Interest

We have no competing interests.

Data Availability

All data are available as an electronical supplementary file.

Ethical Approval

The Animal Use and Care Committee of the University of Pretoria evaluated and approved experimental protocol (ethics clearance no. NAS016/2021) and DAFF section 20 approval (SDAH-Epi-21031811071). Relevant provincial animal capture permits were obtained (Western Cape: CN44-87-13780; Northern Cape: FAUNA 0419/2021, FAUNA 042/2021; Gauteng: CPF6-0124). All methods were performed following the relevant guidelines and regulations. In addition, all experimental procedures were carried out under the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health.

Literature Cited

Abramovs N, Brass A, Tassabehji M. Hardy–Weinberg equilibrium in the large scale genomic sequencing era. Front Genet. 2020:11: 210. https://doi.org/10.3389/fgene.2020.00210.

Ahituv N, Zhu Y, Visel A, Holt A, Afzal V, Pennacchio LA, Rubin EM. Deletion of ultraconserved elements yields viable mice. PLoS Biol. 2007:5(9):e.234. https://doi.org/10.1371/journal.pbio.0050234.

Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 1999:16(1):37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036.

Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D. Ultraconserved elements in the human genome. Science. 2004:304(5675):1321–1325. https://doi.org/10.1126/science.1098119.

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

Bennett NC, Jarvis JUM, Wallace DB. The relative age structure and body masses of complete wild-captured colonies of two social mole-rats, the common mole-rat, *Cryptomys hottentotus hottentotus* and the Damaraland mole-rat, *Cryptomys damarensis*. J Zool. 1990:220(3):469–485. https://doi.org/10.1111/j.1469-7998.1990.tb04319.x.

Bishop JM, Jarvis JUM, Spinks AC, Bennett NC, O'Ryan C. Molecular insight into patterns of colony composition and paternity in the common mole-rat *Cryptomys hottentotus hottentotus*. Mol Ecol. 2004:13(5):1217–1229. https://doi.org/10.1111/j.1365-294X.2004.02131.x.

Bishop JM, O'Ryan C, Jarvis JU. Social common mole-rats enhance outbreeding via extra-pair mating. Biol Lett. 2007:3(2):176–179. https://doi.org/10.1098/rsbl.2006.0607.

Boca SM, Huang L, Rosenberg NA. On the heterozygosity of an admixed population. J Math Biol. 2020:81(6-7):1217–1250. https://doi.org/10.1007/s00285-020-01531-9.

Can E, Smith M, Boukens BJ, Coronel R, Buffenstein R, Riegler J. Naked mole-rats maintain cardiac function and body composition well into their fourth decade of life. Geroscience. 2022:44(2):731–746. https://doi.org/10.1007/s11357-022-00522-6.

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2016:44(D1):67–72. https://doi.org/10.1093/nar/gkv1276.

Colantoni A, Delfanti LMP, Cossio F, Baciotti B, Salvati L, Perini L, Lord R. Soil aridity under climate change and implications for agriculture in Italy. Appl Math Sci. 2015:9:2467–2475. https://doi.org/10.12988/ams.2015.52112.

Cortázar-Chinarro M, Lattenkamp EZ, Meyer-Lucht Y, Luquet E, Laurila A, Höglund J. Drift, selection, or migration? Processes affecting genetic differentiation and variation along a latitudinal gradient in an amphibian. BMC Evol Biol. 2017:17(1):1–14. https://doi.org/10.1186/s12862-017-1022-z.

Dangalle CD, Pallewatta N, Vogler AP. Inferring population history of tiger beetle species of Sri Lanka using mitochondrial DNA sequences. Ceylon J Sci. 2014:43(2):47–63. https://doi.org/10.4038/cjsbs.v43i2.7324.

De Jong MA, Wahlberg N, Van Eijk M, Brakefield PM, Zwaan BJ. Mitochondrial DNA signature for range-wide populations of *Bicyclus anynana* suggests a rapid expansion from recent refugia. PLoS One. 2011:6(6):e.21385. https://doi.org/10.1371/journal.pone.0021385.

Drake J, Lambin X, Sutherland C. The value of considering demographic contributions to connectivity: a review. Ecography. 2022:2022(6):e.05552. https://doi.org/10.1111/ecog.05552.

Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005:14(8):2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x.

Excoffier L, Smouse PE, Quattro J. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics. 1992:131(2):479–491. https://doi.org/10.1093/genetics/131.2.479.

Faircloth BC. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics. 2016:32(5):786–788. https://doi.org/10.1093/bioinformatics/btv646.

Faulkes CG, Abbott DH, O'Brien HP, Lau L, Roy MR, Wayne RK, Bruford MW. Micro- and macrogeographical genetic structure of colonies of naked mole-rats *Heterocephalus glaber*. Mol Ecol. 1997b:6(7):615–628. https://doi.org/10.1046/j.1365-294X.1997.00227.x.

Faulkes CG, Bennett NC. Social evolution in African mole-rats—a comparative overview. Adv Exp Med Biol. 2021:1319:1–33. https://doi.org/10.1007/978-3-030-65943-1_1.

Faulkes CG, Bennett NC, Bruford MW, O'Brien HP, Aguilar GH, Jarvis JU. Ecological constraints drive social evolution in the African mole-rats. Proc R Soc London Ser B Biol Sci. 1997a:264(1388):1619–1627. https://doi.org/10.1098/rspb.1997.0226.

Faulkes CG, Mgode GF, Archer EK, Bennett NC. Relic populations of *Fukomys* mole-rats in Tanzania: description of two new species *F. livingstoni* sp. nov. and *F. hanangensis* sp. nov. PeerJ. 2017:5:e.3214. https://doi.org/10.7717/peerj.3214.

Faulkes CG, Mgode GF, Le Comber SC, Bennett NC. Cladogenesis and endemism in Tanzanian mole-rats, genus Fukomys: (Rodentia Bathyergidae): a role for tectonics? Biol J Linn Soc. 2010:100(2):337–352. https://doi.org/10.1111/j.1095-8312.2010.01418.x.

Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 1985:39(4):783–791. https://doi.org/10.2307/2408678.

Finn KT. Sociality in African mole-rats: exploring how rainfall affects dispersal and genetic exchange in the Natal mole-rat (Cryptomys hottentotus natalensis) (doctoral dissertation). Pretoria: University of Pretoria; 2022.

Francis RM. pophelper: an R package and web app to analyse and visualize population structure. Mol Ecol Resour. 2017:17(1):27–32. https://doi.org/10.1111/1755-0998.12509.

Goudet J. Hierfstat, a package for R to compute and test hierarchical F-statistics. Mol Ecol Notes. 2005:5(1):184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x.

Greeff JM, Bennett NC. Causes and consequences of incest avoidance in the cooperatively breeding mole-rat, *Cryptomys darlingi* (Bathyergidae). Ecol Lett. 2000:3(4):318–328. https://doi.org/10.1046/j.1461-0248.2000.00162.x.

Harris DJ, Arnold EN, Thomas RH. Rapid speciation, morphological evolution, and adaptation to extreme environments in South African sand lizards (Meroles) as revealed by mitochondrial gene sequences. Mol Phylogenet Evol. 1998:10(1):37–48. https://doi.org/10.1006/mpev.1997.0463.

Hart DW, Bennett NC, Best C, van Jaarsveld B, Cheng H, Ivy CM, Kirby AM, Munro D, Sprenger RJ, Storey KB, et al. The relationship between hypoxia exposure and circulating cortisol levels in social and solitary African mole-rats: an initial report. Gen Comp Endocrinol. 2023:339:114294. https://doi.org/10.1016/j.ygcen.2023.114294.

Hasegawa M, Kishino H, Yano TA. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol. 1985:22(2):160–174. https://doi.org/10.1007/BF02101694.

Hickman GC. A live-trap and trapping technique fort fossorial 794 mammals. Af Zool. 1979:14(1):9–12. https://doi.org/10.1080/02541858.1979.11447641.

Hoffmann AA, Willi Y. Detecting genetic responses to environmental change. Nat Rev Genet. 2008:9(6):421–432. https://doi.org/10.1038/nrg2339.

Jacobs PJ, Hart DW, Merchant HN, Janse van Vuuren AK, Faulkes CG, Portugal SJ, Van Jaarsveld B, Bennett NC. Tissue oxidative ecology along an aridity gradient in a mammalian subterranean species. Antioxidants. 2022:11(11):e.2290. https://doi.org/10.3390/antiox11112290.

Jarvis JUM, O'Riain MJ, Bennett NC, Sherman PW. Mammalian eusociality: a family affair. Trends Ecol Evol. 1994:9(2):47–51. https://doi.org/10.1016/0169-5347(94)90267-4.

Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008:24(11):1403–1405. https://doi.org/10.1093/bioinformatics/btn129.

Knipler ML, Gracanin A, Mikac KM. Conservation genomics of an endangered arboreal mammal following the 2019–2020 Australian megafire. Sci Rep. 2023:13(1):e.480. https://doi.org/10.1038/s41598-023-27587-3.

Leigh JW, Bryant D. PopART: full-feature software for haplotype network construction. Methods Ecol Evol. 2015:6(9):1110–1116. https://doi.org/10.1111/2041-210X.12410.

Lischer HEL, Excoffier L. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics. 2012:28(2):298–299. https://doi.org/10.1093/bioinformatics/btr642.

Lombaert E, Guillemaud T, Deleury E. Biases of STRUCTURE software when exploring introduction routes of invasive species. Heredity (Edinb). 2018:120(6):485–499. https://doi.org/10.1038/s41437-017-0042-1.

Lovegrove BG. The evolution of eusociality in molerats (Bathyergidae): a question of risks, numbers, and costs. Behav

Ecol Sociobiol (Print). 1991:28(1):37–45. https://doi.org/10.1007/BF00172137.

Maddison WP, Maddison DR. 2021. [accessed 2023 Dec 2]. http://www.mesquiteproject.org.

McCallum KP, Guerin GR, Breed MF, Lowe AJ. Combining population genetics, species distribution modelling and field assessments to understand a species vulnerability to climate change. Austral Ecol. 2014:39(1):17–28. https://doi.org/10.1111/aec.12041.

Merchant HN, Portugal SJ, Bennett NC, Janse van Vuuren AK, Faulkes CG, Bowen J, Hart DW. New insights into morphological adaptation in common mole-rats (*Cryptomys hottentotus hottentotus*) along an aridity gradient. Ecol Evol. 2024:14(4):11301. https://doi.org/10.1002/ece3.11301.

Müllenbach R, Lagoda PJ, Welter C. An efficient salt-chloroform extraction of DNA from blood and tissues. Trends Genet. 1989:5(12):391–391.

Muñoz-Sabater J, Dutra E, Agustí-Panareda A, Albergel C, Arduini G, Balsamo G, Boussetta S, Choulga M, Harrigan S, Hersbach H, et al. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst Sci Data. 2021:13(9):4349–4383. https://doi.org/10.5194/essd-13-4349-2021.

Namipashaki A, Razaghi-Moghadam Z, Ansari-Pour N. The essentiality of reporting Hardy–Weinberg equilibrium calculations in population-based genetic association studies. Cell J (Yakhteh). 2015:17(2):187. 10.22074/cellj.2016.3711.

Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 1978:89(3):583–590. https://doi.org/10.1093/genetics/89.3.583.

Nei M, Kumar S. Molecular evolution and phylogenetics. Oxford: Oxford University Press; 2000.

Nei M, Tajima F. Genetic drift and estimation of effective population size. Genetics. 1981:98(3):625–640. https://doi.org/10.1093/genetics/98.3.625.

Nevo E, Beiles A. Genetic diversity in the desert: patterns and testable hypotheses. J Arid Environ. 1989:17(2):241–244. https://doi.org/10.1016/S0140-1963(18)30911-X.

Orr HA. The genetic theory of adaptation: a brief history. Nat Rev Genet. 2005:6(2):119–127. https://doi.org/10.1038/nrg1523.

Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD, et al. In vivo enhancer analysis of human conserved non-coding sequences. Nature. 2006:444(7118):499–502. https://doi.org/10.1038/nature05295.

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000:155(2):945–959. https://doi.org/10.1093/genetics/155.2.945.

R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021. URL https://www.R-project.org/.

Rocha JL, Godinho R, Brito JC, Nielsen R. Life in deserts: the genetic basis of mammalian desert adaptation. Trends Ecol Evol. 2021:36(7):637–650. https://doi.org/10.1016/j.tree.2021.03.007.

Romanenko V. Computation of the autumn soil moisture using a universal relationship for a large area. Proceedings of Ukrainian Hydrometeorological Research Institute. 1961:3:12–25. https://doi.org/10015209318.

Sandelin A, Bailey P, Bruce S, Engström PG, Klos JM, Wasserman WW, Ericson J, Lenhard B. Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. BMC Genomics. 2004:5(1):1–9. https://doi.org/10.1186/1471-2164-5-99.

Scott PA, Allison LJ, Field KJ, Averill-Murray RC, Shaffer BH. Individual heterozygosity predicts translocation success in threatened desert tortoises. Science. 2020:370(6520):1086–1089. https://doi.org/10.1126/science.abb0421.

Sinnock P. The Wahlund effect for the two-locus model. Am Nat. 1975:109(969):565–570. https://doi.org/10.1086/283027.

Slatkin M. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. Nat Rev Genet. 2008:9(6):477–485. https://doi.org/10.1038/nrg2361.

Smith BT, Harvey MG, Faircloth BC, Glenn TC, Brumfield RT. Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. Syst Biol. 2014:63(1):83–95. https://doi.org/10.1093/sysbio/syt061.

Sobell MG. A practical guide to Ubuntu Linux. Hoboken, NJ: Pearson Education; 2015.

Spinks AC. 1998. Sociality in the common mole-rat, *Cryptomys hottentotus hottentotus*: the effects of aridity [PhD thesis] University of Cape Town, Cape Town.

Spinks AC, Bennett NC, Jarvis JUM. Regulation of reproduction in female common mole-rats (*Cryptomys hottentotus hottentotus*): the effects of breeding season and reproductive status. J Zool. 1999:248(2):161–168. https://doi.org/10.1111/j.1469-7998.1999.tb01192.x.

Spinks AC, Bennett NC, Jarvis JU. A comparison of the ecology of two populations of the common mole-rat, *Cryptomys hottentotus hottentotus*: the effect of aridity on food, foraging and body mass. Oecologia. 2000a:125(3):341–349. https://doi.org/10.1007/s004420000460.

Spinks AC, Jarvis JU, Bennett NC. Comparative patterns of philopatry and dispersal in two common mole-rat populations: implications for the evolution of mole-rat sociality. J Animal Ecol. 2000b:69(2):224–234. https://doi.org/10.1046/j.1365-2656.2000.00388.x.

Spinks AC, Plaganyi EE. Reduced starvation risks and habitat constraints promote cooperation in the common mole-rat, *Cryptomys hottentotus hottentotus*: a computer-simulated foraging model. Oikos. 1999:85(3):435–444. https://doi.org/10.2307/3546693.

Stecher G, Tamura K, Kumar S. Molecular evolutionary genetics analysis (MEGA) for macOS. Mol Biol Evol. 2020:37(4):1237–1239. https://doi.org/10.1093/molbev/msz312.

Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989:123(3):585–595. https://doi.org/10.1093/genetics/123.3.585.

Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021:38(7):3022–3027. https://doi.org/10.1093/molbev/msab120.

Tigano A, Colella JP, MacManes MD. Comparative and population genomics approaches reveal the basis of adaptation to deserts in a small rodent. Mol Ecol. 2020:29(7):1300–1314. https://doi.org/10.1111/mec.15401.

Torrents-Ticó M, Bennett NC, Jarvis JU, Zöttl M. Sex differences in timing and context of dispersal in Damaraland mole-rats (*Fukomys damarensis*). J Zool. 2018:306(4):252–257. https://doi.org/10.1111/jzo.12602.

U.N.E.P. World atlas of desertification. 80. Kent: UNEP and Edward Arnold Publishers; 1992.

UNESCO. MAB Technical Notes No. UNESCO Press; 1979.

Van Daele PAAG, Blondé P, Stjernstedt R, Adriaens D. A new species of African Mole-rat (*Fukomys*, Bathyergidae, Rodentia) from the Zaire-Zambezi Watershed. Zootaxa. 2013:3636(1). https://doi.org/10.11646/zootaxa.3636.1.7.

Van Daele PAAG., Dammann P, Meier JL, Kawalika M, Van De Woestijne C, Burda H. Chromosomal diversity in mole-rats of the genus *Cryptomys* (Rodentia: Bathyergidae) from the Zambezian region: with descriptions of new karyotypes. J Zool. 2004:264(3):317–326. https://doi.org/10.1017/S0952836904005825.

Van Daele PAAG, Verheyen E, Brunain M, Adriaens D. 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. 2007:45(1):142–157. https://doi.org/10.1016/j.ympev.2007.04.008.

Visser JH, Bennett NC, Jansen van Vuuren B. Spatial genetic diversity in the Cape mole-rat, *Georychus capensis*: extreme isolation of populations in a subterranean environment. PLoS One. 2018:13(3):e0194165. https://doi.org/10.1371/journal.pone.0194165.

Visser JH, Bennett NC, van Vuuren BJ. Evolutionary and ecological patterns within the South African Bathyergidae: implications for taxonomy. Mol Phylogenet Evol. 2019:130:181–197. https://doi.org/10.1016/j.ympev.2018.10.017.

Wang IJ, Bradburd GS. Isolation by environment. Mol Ecol. 2014:23(23):5649–5662. https://doi.org/10.1111/mec.12938.

Wang IJ, Glor RE, Losos JB. Quantifying the roles of ecology and geography in spatial genetic divergence. Ecol Lett. 2013:16(2):175–182. https://doi.org/10.1111/ele.12025.

Zemlemerova ED, Kostin DS, Lebedev VS, Martynov AA, Gromov AR, Alexandrov DY, Lavrenchenko LA. Genetic diversity of the naked mole-rat (*Heterocephalus glaber*). J Zool Syst Evol Res. 2021:59(1):323–340. https://doi.org/10.1111/jzs.12423.

Associate editor: Emily Kopania