Title

Aridity shapes adaptive genomic divergence and population connectivity in a Southern African rodent

Abstract

Elucidating the drivers of evolution in dry environments is central to understanding how organisms respond to climate change. While research on the genomics of adaptation is growing, aridity-driven intraspecific divergence remains poorly quantified. Here, we address this gap by using genomic data from 230 individuals of the arid-adapted four-striped mouse *Rhabdomys bechuanae*, sampled across an aridity gradient in southern Africa, a region facing increasing aridification. Combining these data with palaeoclimatic reconstructions and present-day aridity indices, we investigate, from a spatio-temporal perspective, how intraspecific genetic variation relates to aridity. Inference of past effective population size revealed a sharp decline around 15,000 years BP, coinciding with regional aridification and likely reflecting reduced connectivity during dry periods. Population structure followed a pattern of isolation by distance and mirrored the aridity gradient. Genotype-Environment Association analyses identified SNPs and genes significantly associated with aridity and genetically differentiated among populations, with functions related to water and energy conservation - as expected under arid conditions - as well as neurotransmission. These findings highlight the underappreciated role for neurological processes in coping with water and resource scarcity. More broadly, our integrative population genomics approach strongly suggests that aridity shapes population connectivity and adaptation, with implications for climate resilience.

Keywords

adaptation, aridification, effective population size, genotype-environment association, population genomics, *Rhabdomys bechuanae*

Teaser text

How do organisms adapt to increasingly arid environments? Using population genomic data from a desert-occurring rodent sampled along an aridity gradient in Southern Africa, we investigate how aridity shapes intraspecific population structure and adaptive divergence. We uncover historical reduction in connectivity linked to past aridification episodes, and identify genetic signals of adaptation to water and energy scarcity, highlighting neurological pathways as an important but often overlooked facet of aridity adaptation. Our findings shed new light on the evolutionary impact of aridification and demonstrate the value of integrative genomics for understanding climate resilience in wild species.

Introduction

Climate change has increasingly been associated with local extinctions (Wiens, 2016), shifts in phenology (Menzel et al., 2006) and alteration of species' geographic distributions (Chen et al., 2011). These responses tend to be particularly pronounced in harsh environments, such as arid regions, where species already face several ecological and physiological constraints (Easterling et al., 2000). With projections indicating a rise in the intensity and frequency of drought events in the coming decades (Naumann et al., 2018; Engelbrecht et al., 2024), it is essential to better understand how species respond to both temporal and spatial variation in aridity. By combining newly generated genomic data with available palaeoclimatic reconstructions and contemporary measures of aridity, this study aims to shed light on the role of aridification on the demographic history, spatial population structure and adaptive responses of a non-model rodent species. These insights are key for predicting resilience of natural populations under future climate scenarios and for guiding conservation strategies (Skelly et al., 2007).

At the individual level, the primary challenges of life in arid environments include maintaining body temperature and conserving water (Willmer et al., 2009). These challenges are especially evident for species that rely on evaporative water loss mechanisms, such as sweating, to dissipate heat, especially in mammals (Rymer et al., 2016). The physiological demands imposed by such environmental conditions, particularly when experienced over multiple generations, often exceed the limits of phenotypic plasticity in non-desert-adapted species. Consequently, it is widely assumed that strong selection has acted on complex physiological traits related to energy metabolism, water retention and thermoregulation in species inhabiting arid environments (Willmer et al., 2009; Rymer et al., 2016; Rocha et al., 2021).

Genomic studies have identified recurrent genetic patterns among mammals living in deserts (Rocha et al., 2021). Evidence of adaptation to arid environments has included elevated allele frequency differentiation (Malaspinas et al., 2016; Rocha et al., 2023), reduced genetic diversity (Malaspinas et al., 2016; Kim et al., 2016; Ababaikeri et al., 2020; Tigano et al., 2020), increased non-synonymous-to-synonymous substitution rate ratios (Cheng et al., 2023) and divergent gene expression patterns (Bittner et al., 2021) in genomic regions involved in key functions for desert adaptation. These genomic regions often relate to key physiological traits relevant to the survival of both mammalian specialist and non-specialist species in deserts, including DNA repair, protein synthesis and degradation (Tigano et al., 2020, Cheng et al., 2023), thyroid hormone levels (Malaspinas et al., 2016), water retention (Bridges and James, 1982), lipid (Cheng et al., 2023) and carbohydrate metabolism (Wu et al., 2014; MacManes & Eisen, 2014; Marra et al., 2014; Malaspinas et al., 2016; Kim et al., 2016). However, most of these studies have relied on qualitative comparisons between desert and non-desert species or populations, and relatively few have explicitly explored quantitative variation along continuous environmental gradients. Several other studies on the genomics of local adaptation have demonstrated the power of Genotype-Environment Association (GEA) analyses to identify polymorphisms associated with continuous

environmental variation (Coop et al., 2010; Frichot et al., 2013; Gautier, 2015). In this context, Rocha et al. (2021) highlighted the use of the Aridity Index (AI), known to be a good predictor of certain aridity-adapted phenotypes, as a powerful, composite and quantitative environmental variable to inform genomic studies of adaptation to arid conditions.

In this study, we applied a GEA approach to investigate the genomic signatures of adaptation to arid conditions in *Rhabdomys bechuanae*, a four-striped mouse species from Southern Africa. Adapted to arid environments, *R. bechuanae* provides a unique opportunity to study adaptation in a system where strong environmental gradients exist within a single species due to its distribution along a pronounced west-to-east aridity gradient. Its range overlaps with hot arid and hyper-arid summer rainfall regions of Southern Africa (Ganem et al., 2020). Divergence from its closest relatives within the *Rhabdomys* species complex occurred approximately 3.1-4.3 million years ago (du Toit et al., 2012), and this diversification has been associated with environmental niche differentiation (Meynard et al., 2012). It likely reflects a long-term trend towards increasing aridity and seasonality in Southern Africa since the late Miocene, coupled with a shift from closed subtropical woodland to sparse, shrubby vegetation (Dupont et al., 2013). Currently, *R. bechuanae* thrives predominantly in sparsely vegetated areas and nests in bushes. Phenotypic traits suggest adaptation to aridity; these traits include morphological (longer tail length related to potential differences in locomotion and/or thermoregulation) and behavioural divergence (selection of different habitats) from more mesic congeners (Ganem et al., 2020; Dufour et al., 2019), as well as enhanced osmoregulatory ability, as suggested by differences in blood metabolite concentrations between populations of *R. bechuanae* and of its sister species *R. dilectus,* occurring syntopically in semi-arid environments (Keilani et al., 2025). Although *R. bechuanae* is considered one of the most arid-adapted species of its genus, it still occupies a wide environmental gradient, providing an ideal system to investigate whether the aridity gradient imposes divergent selection pressures on key physiological traits. We therefore hypothesised that this aridity gradient would impose divergent selective pressures on key physiological traits related to adaptation to aridity, leading to detectable intraspecific population differentiation and genetic variation at loci associated with aridity-related functions.

To investigate genomic signatures of evolution under arid conditions in *R. bechuanae*, we generated a genome-wide SNP dataset using restriction-site associated DNA sequencing (RAD-seq; Baird et al., 2008). Our overarching aim was to evaluate how the temporal and spatial variation in aridity has shaped the population history and structure of *R. bechuanae*, by integrating inferences of historical changes in effective population size with GEA analyses.

Using an extensive sample of individuals from multiple localities spanning the species' aridity gradient, our first objective was to reconstruct past changes in effective population size in relation to palaeoclimatically inferred shifts in aridity. We expected that historical episodes of increased aridity would be associated with population declines, either due to demographic bottlenecks (Benning et al., 2023; Ventura-Rojas et

al., 2024), or via reductions in genetic diversity caused by habitat fragmentation and decreased population connectivity. Conversely, wetter periods were expected to correspond with population expansions or increased gene flow.

In addition to temporal variation, spatial variation in aridity may also influence current population structure and genetic diversity (Gomulkiewicz & Houle, 2009; Sarabia et al., 2021; Rocha et al., 2023). Therefore, our second objective was to characterise the current population structure within *R. bechuanae* and assess its relationship with spatial variation in aridity across the species' range. We predicted that population structure would correlate with the degree of aridity, reflecting its impact on gene flow and differentiation.

Our third objective was to assess the genomic basis of adaptation to aridity in *R. bechuanae* using a GEA approach with the Aridity Index as a quantitative predictor. This approach enabled us to identify loci associated with spatial variation in arid conditions across the species' range and to identify potential biological functions that may underpin survival in arid environments. We hypothesised that site frequency variation would concentrate in genomic regions linked to key functions for survival in deserts such as DNA repair, protein modification, osmoregulation, and carbohydrate/fat metabolism (Rocha et al., 2021; Cheng et al., 2023).

Material & Methods

Samples, aridity index and RAD-seq data production

Individual samples of four-striped mice of the genus *Rhabdomys* were previously collected in their natural environment between 2007 and 2019 by us or collaborators (Ganem et al., 2020). Population-level samples from *R. bechuanae* were collected specifically for the present study in 2018 and 2019 (details for trapping procedures in Ganem et al., 2020). We aimed to avoid related individuals as much as possible, by selecting mice trapped at least 100 m apart, except for breeding pairs (a male and a female) that could be trapped in the same nest. This distance was based on data obtained for *R. pumilio* indicating minimal relatedness (R=0 between males and R=0.06 between females) at this distance (Solmsen et al., 2011).

To study fine-scale intraspecific structure, as well as perform GEA and historical effective population size inferences in *R. bechuanae*, we selected a large number of individuals from this species, sampled in 11 different localities at least 20 km distant from each other and along the aridity gradient (*N*=230 individual samples in total, 19.82 +/- 4.72 individuals per locality, **Table S1**). Species identification for all sampled individuals in the *Rhabdomys* species complex composed of cryptic species had previously been obtained via COI sequencing or RFLP genotyping (Ganem et al., 2020 and procedure therein; **Table S1**).

To assess the degree of aridity for each *R. bechuanae* locality, we computed the Aridity Index (AI). AI is typically defined as the ratio between precipitation (*P*) and potential evapotranspiration (*PET*) in a given environment (*World Atlas of Desertification*, 1992). For each *R. bechuanae* locality, the Aridity Index (AI) was retrieved based on data computed using the 30-year average of *P*/*PET* (1981-2010):

$$\mathrm{AI} = \frac{\sum_{i=1}^{30}\left(\frac{P_i}{PET_i}\right)}{30}$$

where *i* denotes the ith year.

The Aridity index was computed on a 0.5° global grid, using data from Version 3 of the Global Aridity Index and Potential Evapotranspiration Database (Zomer et al., 2022).

DNA extractions for RAD-sequencing were performed from tail tissues stored in 98% alcohol, with a 96-Well Plate Animal Genomic DNA Miniprep Kit (BioBasic Inc.). Since samples from other *Rhabdomys* species and subspecies had also been collected for a separate study, we proceeded to RAD-seq library preparation and sequencing for *R. bechuanae* samples analysed in this study concurrently with these other *Rhabdomys* samples (*N* = 401 in total). Preparation of the libraries for RAD sequencing was performed following the protocol designed by Etter et al. (2011) and modified as detailed in Cruaud et al. (2014). Briefly, DNAs were digested with the 8-cutter restriction enzyme SbfI and ligated to a modified Illumina P1 adapter containing individual multiplex identifiers (MIDs) 5-6 bp long. Then, 20 pools of libraries were prepared based on equimolar mixes of 22 to 29 digested and tagged DNAs per pool (global *Rhabdomys* sample set including 230 *R. bechuanae* individuals). Each pool was sheared by sonication using an S220 ultra-sonicator (Covaris, Inc.) and size-selected for 300–600 bp by agarose gel excision. Different P2 adapters including pool multiplex identifiers (MIDs) were then ligated to different pools intended to be sequenced on the same sequencing lane. The DNA fragments containing both adapters (P1 and P2) were PCR enriched during 18 cycles. These 20 pools of RAD libraries were sequenced on a total of four S1 lanes and two SP lanes of an Illumina NovaSeq 6000, using a 2x150 bp paired-end protocol, at MGX-Montpellier GenomiX Facility (Montpellier, France).

RAD-seq raw data processing and filtering

Following standard Illumina processing and quality filtering, duplicate reads resulting from PCR amplification were discarded using the program *clone_filter* implemented in the *Stacks* v. 2.66 software (Catchen et al., 2013). The resulting paired-end reads were demultiplexed and quality filtered using the *process_radtags* pipeline, also implemented within *Stacks*. Sequence reads were aligned to the reference genome using *BOWTIE* (version 2.4.5, Langmead et al. 2009). The reference genome used was the one of the

closely related species *R. pumilio* (2.3 Gb in length; GCA_030674055.1; Richardson et al., 2023). Before mapping, putative repeated elements were hard-masked (i.e., repeated sequences were replaced by stretch of "N") from the reference genome using the mouse repeated element Repbase Update (v 27.06) database using *Repeatmasker* v. 4.1.3. (Smit et al., 2015). Reads mapping to more than one reference sequence were discarded, and the maximum number of mismatches allowed was three. The resulting files were converted to .bam files and sorted using *SAMtools* v. 1.18 (Li et al., 2009).

The reference-based *Gstacks* pipeline was applied to cluster reads from all available samples (*N*=401) into RAD loci using the Marukilow model, minimum mapping quality of 40, and alpha thresholds (for mean and variance) of 0.05 for discovering single nucleotide polymorphisms (SNPs). Using the catalogue of loci generated with *Gstacks*, the *populations* program was then used on *R. bechuanae* samples only, to proceed to SNP and sample filtering steps and generate SNP datasets adapted to each type of downstream analyses: dataset (A) for population structure analyses/genetic differentiation estimates, dataset (B) for inferences of historical changes in effective population size and dataset (C) for GEA analyses. Some SNP and sample filtering steps were applied similarly to all three datasets: a minimum Phred genotype quality score of 40, a minimum read depth per genotype of three, a minimum 10X and maximum 30X coverage (based on empirical distributions), and a minimum percentage of the samples genotyped of 80 %. We also retained in each dataset only the first SNP of each locus to minimise linkage disequilibrium bias in downstream analyses, and only individuals with less than 50% of missing data. The minor variant count filter (--mac) was only applied (mac=3) to datasets A and C. Finally, a downsampling procedure was applied to datasets A and B. Unbalanced sample sizes can induce some bias in the inference of population structure (Puechmaille, 2016) and demographic history (McLaughlin & Winker, 2020), notable through overestimating admixture with the *sNMF* method (Wang, 2022). We downsampled each locality up to a maximum of 19 samples, with only the samples with the least missing data being included for each locality, as missing or low frequency variants can lead to imprecisions and biases in population genetic inferences (Wright et al., 2019; Crawford & Lazzaro, 2012). Several different parameter sets were tested for each filter parameter, partly based on empirical distributions, but intermediate stringency was ultimately chosen as a compromise between SNP numbers *vs*. coverage. The sequence order of the different SNP and sample filtering steps, their application to each analytical dataset and their effect on the number of final samples and SNP retained are provided in **Table S2**.

Inferences of historical changes in effective population size

To evaluate historical events that could have significantly impacted current population genetic diversity and structure, and test whether historical variations in effective population size (N_e) coincide with known past variations in the degree of aridity, we estimated historical changes in N_e using *Stairway Plot 2* v.2.1.2, which conducts multi-epoch coalescent inference to assess changes in population size over time (Liu & Fu, 2015; Liu & Fu, 2020). We run this analysis at two different sample scales: 1) at the *R.*

bechuanae species level, pooling individual samples from all localities, and 2) at the locality level: within *R. bechuanae*, analysing two localities representing the two extremes of the Aridity Index across the sampled distribution.

The Python program *easySFS* (https://github.com/isaacovercast/easySFS) was first used to generate the folded site frequency spectrum (SFS) formatted file with VCF files and a sample list file (v0.0.1; Coffman et al., 2015). Since demographic inference from the SFS is particularly dependent on the presence of rare alleles (i.e., singletons and doubletons; Gutenkunst et al., 2009), we used the SNP data set obtained without the *-mac* filter (dataset B). In addition to the species-level analysis with N_{max} = 19 per locality (see above), we also carried out an analysis using a *R. bechuanae* pooled sample with N = 1 per locality. By reducing to one the sample size per locality, the goal was to remove any possible signal of population structure among localities, and therefore disentangle true population size changes from effects of population structure within *R. bechuanae* on estimations of N_e over time, as suggested by Chikhi et al. (2010, 2018).

In all cases, the samples with the least missing data were included, and then projected down, to cover the optimal number of SNPs. For all analyses, the mutation rate was set to 5.7×10^{-9} per site per generation (based on the germline mutation rate found in *Mus musculus* in Milholland et al., 2017), with a generation time of one year.

We considered the entire allele frequency spectrum with and without singletons. Ultimately, 67 % of sites were used to create an SFS training set and train the model, and the remaining sites were used to create an SFS testing set and test the goodness of fit of the trained model. The goodness of fit of the trained models was tested using ¼, ½, ¾ or n-2 “breakpoints”, which define the boundaries of each epoch, and the best-fit model was used to produce the final inference. Two hundred simulations were carried out for each estimation.

Population structure analyses

To describe the structure of genetic diversity within *R. bechuanae*, we first computed differentiation statistics (pairwise F_{ST}, Weir and Cockerham, 1983) for all sampled localities with sufficient sample sizes using the *vcftools* program v.0.1.16 (Danecek et al., 2011). To test isolation-by-distance and isolation-by-environment within *R. bechuanae*, simple and partial Mantel tests (R *ncf* package v. 1.3-2, Bjornstad & Bjornstad, 2016) were conducted to test for correlations between genetic differentiation among all *R. bechuanae* localities (using $F_{ST}/(1- F_{ST})$ following Rousset, 1997) and geographic distance, as well as environmental distance (using the Aridity Index). To characterise population structure within *R. bechuanae*, we performed a principal component analysis (PCA) and a sparse non-negative matrix factorisation (*sNMF)* analysis using the R package *LEA* v.3.14 (Frichot & François, 2015). For *sNMF* analyses, we ran 1,500 repetitions for each value of *K* ranging from 1 to 8 and chose the best *K* by evaluating cross-entropy values. Additionally, a neighbour-joining tree was

generated using SNP data of these same samples with the R package *ape* (v. 5.7-1, Paradis & Schliep, 2019).

Genotype-Environment Association analyses

To explore signals of genetic adaptation to aridity, we conducted Genotype–Environment Association (GEA) analyses. Using extensions of F_{ST}-based models, the rationale was that environmental variables distinguishing differentiated populations should be associated with allele frequency differences at loci subjected to the selective constraints they impose (Coop et al., 2010). To test for an association between allele frequency differences between all *R. bechuanae* localities and the environmental variable of interest, the Aridity Index, we used the program *BayPass* v.2.4, which implements a method accounting for the neutral correlation of allele frequencies across populations to detect loci associated with covariables (Gautier, 2015). We ran *BayPass* under the standard covariate model, which uses an Importance Sampling (IS) approximation to identify SNPs associated with the Aridity Index, and corrects for population structure by using the scaled covariance matrix Ω of population allele frequencies. Three runs (with different starting seeds) were performed to test for convergence. To determine the level of association of each SNP with the Aridity Index and identify SNPs significantly associated, we computed the empirical Bayesian p-value (eBPis) and Bayes factors (*BF*) expressed in deciban units (dB). Associations between each SNP and Aridity Index were judged from Bayes Factors under a Markov Chain Monte Carlo (MCMC) sampling method (BFmc). Three thresholds of *BF* (or degrees of association) were considered, following Jeffreys, 1961: 10 dB < *BF* < 15 dB for “strong” evidence; 15 dB < *BF* < 20 dB for “very strong” evidence; *BF* > 20 dB for “decisive” evidence.

Additionally, *XtX* statistics (Coop et al., 2010) were outputted by the *BayPass* core model to identify SNPs with significant genetic differentiation among *R. bechuanae* populations. We chose *XtX* over the more traditional F_{ST} statistics because *XtX* explicitly accounts for population structure through Ω. *P*-values associated with the *XtX* calibrated estimator were not distributed as expected under the null hypothesis (**Figure S1**). Therefore, *p*-value thresholds were not used to identify outliers. Instead, *XtX* statistics were calibrated through the simulation and analysis of a pseudo-observed dataset as described in Gautier (2015), simulated with 44,900 SNPs.

To construct an accurate gene annotation file, we mapped the *Mus musculus* GRCm39 assembly annotation (Mouse Genome Sequencing Consortium, 2002) onto the *R. pumilio* genome using the default settings of *liftoff* software version 1.6.3 (Shumate & Salzberg, 2021), which was specifically designed for accurate mapping of gene annotation files. SNPs with a *BF* greater than 10 dB, as well as those that both had a *BF* greater than 10 dB and a significant *XtX*, were mapped onto the annotation. We used this annotation to identify genes containing SNPs significantly associated with the Aridity Index (outlier ‘association’ genes) and among those, genes containing SNPs both significantly associated and differentiated among populations (outlier

‘association/differentiation’ genes). For each gene, we first retrieved a full description from the Alliance of Genome Resources (Alliance of Genome Resources Consortium, 2024), which was then summarised to obtain a meaningful set of functional information. We performed functional enrichment analyses on these two lists of outlier genes using *Metascape*, an online tool for gene function annotation analysis. The enrichment was initially considered statistically significant at $p < 0.05$. P-values were then corrected for false discovery rates according to the Benjamini-Hochberg method. Only enrichments exhibiting a corrected $p_{adj} < 0.05$ were considered significant, with enrichments on the verge of significance also reported in this article.

Results

RAD-seq and genotype dataset

Across all 230 *R. bechuanae* samples analysed, an average of 2,316,057 sequence reads per sample were mapped against the *R. pumilio* reference genome, after using *process_radtags* to filter for quality and *clone-filter* to mask potential PCR duplicates (see **Table S3** for details). The alignment rate to the repeat-masked *R. pumilio* genome averaged 55.96 % across *R. bechuanae* samples. The *Gstacks* analysis resulted in an average raw number of RAD loci identified per sample of 151,936 ± 1,779. RAD loci were genotyped with a mean per-sample coverage of 10.94 ± 0.21 X and a mean of 227.88 ± 0.09 base pairs per locus.

The SNP and sample filtering procedures, used to generate the three different *R. bechuanae* datasets for population structure analyses, inferences of past effective population sizes and GEA analyses, resulted in 44,900 (population structure, dataset A; GEA: dataset C), and 147,761 (inferences of past effective population sizes: dataset B) final sets of SNPs (**Table S2**). After SNP and genotype quality filtering, missing data were on average 9.56 ± 0.68 % per sample (**Figure S2**). A total of 154 samples were kept after individual missing data filtering (< 50%) and downsampling for analysis of population structure and historical changes in population effective size, while 218 samples were kept after individual missing data filtering for the GEA (**Tables S2** and **S3**).

Genetic diversity and population structure in time and space

Historical changes in effective population size in *R. bechuanae*

Aridity Index (AI) estimations validated our sampling of *R. bechuanae* along an aridity gradient in South Africa: AI of sampled localities ranged from 0.0322 (Klein Pella, classified as hyper-arid) to 0.2457 (Sandveld, classified as semi-arid) (**Table S4**).

We detected no significant differences in the temporal patterns of N_e associated with contrasting contemporary aridity levels among localities at the semi-arid (Sandveld) or hyper-arid (Klein Pella) extremes along the aridity gradient, nor any evidence of recent population decline (**Figure S3**). We therefore analysed historical changes in N_e for *R. bechuanae* as a whole, pooling data from all sampled localities and with N_{max} = 19 individuals per locality. Results from the *Stairway Plot 2* analysis performed incorporating singletons (**Figure 1**, solid line) closely matched those obtained without singletons (**Figure S4**). In both cases, the demographic trajectory of *R. bechuanae* showed a steep increase in N_e beginning around 140 thousand years ago (kya), followed by a stabilisation around 100 kya, with a median N_e estimate of 4,434,261 individuals reached around 60 kya. Around 30 kya - shortly before the Last Glacial Maximum-LGM - N_e began to decline gradually, before undergoing a sharp drop from 15 kya to the present. In contrast, the analysis using a *R. bechuanae* pooled sample with N = 1 per locality revealed a stable N_e during of approximately 4 x 10^6 throughout the same period (**Figure 1**, dashed line). This discrepancy suggests that the recent sharp decline in N_e observed when treating the species as a whole may reflect reduced population connectivity during that time. Notably, we found that this period of sharp N_e decline coincides with repeated and localised arid episodes that occurred 15 – 11, 7–5 and 1.5–0.5 kya in the central summer rainfall zones of Southern Africa (Chevalier & Chase, 2016), suggesting that past aridification events may have caused reduced population connectivity, higher genetic differentiation, and whole population decline.

Contemporary *R. bechuanae* population structure

The analysis of contemporary population structure on all retained *R. bechuanae* individuals revealed population structure despite low level of intraspecific differentiation within this species (**Figure 2**, **Tables S5 and S6**). Pairwise F_{ST} estimates among *R. bechuanae* sampling localities were relatively low, spanning from 0.010 to 0.058 (**Table S5**). Combined, the Principal Component Analysis (PCA) (**Figure S5**), the *sNMF* analysis (**Figure 2A**) and the neighbour-joining tree (**Figure 2C**) did not indicate a clear-cut separation of several sampling localities or groups of localities from each other.

The *sNMF* analysis yielded similarly low cross-entropy values for K = 3 and 4 clusters (**Figure S6**), suggesting that these two inferences of population structure are almost equally likely (**Figure 2A**). However, the minimal cross-entropy value was found at K = 3, indicating that *R. bechuanae* samples were primarily assigned to three main genetic clusters (**Figure 2**, **Table S6**). A first *R. bechuanae* genetic cluster grouped most localities from a region corresponding to the Southern Kalahari (Lake Naute, Mariental, Tswalu, Molopo, Kolomela, Sandveld), which ranged from a hyper-arid to a semi-arid climate according to the estimated Aridity Index. Another cluster was more prevalent in the Gariep and Tussen die Riviere semi-arid localities, and a third cluster was mostly found in the hyper-arid Klein Pella locality (**Figure 2B**, **Table S6**). Genetic differentiation between these groups of localities was also visible on the neighbour-joining tree (**Figure 2C**). Most individuals from localities located centrally in our sampling scheme

(Benfontein, Kolomela Mine) did not have ancestry coefficients > 85 % from one specific cluster and were instead composed of mixed continuous ancestries (**Figure 2C, Table S6**), as also suggested by the PCA plot (**Figure S5**).

Consistent with the gradual pattern of genetic ancestries observed across the sampled localities (**Figure 2**), simple Mantel tests using F_{ST} distances among *R. bechuanae* localities (**Table S5**) showed significant positive correlations between genetic distance and geographic distance (p = 0.038), indicating isolation by distance, as well as between genetic distance and the environmental variable of interest, the Aridity Index (p = 0.037), indicating isolation by environment (**Figure 3**). Since we found a very strong correlation between geographic and environmental distances (p <0.001) and no significant partial Mantel tests (**Table S7**), these results suggest that patterns of isolation by distance and by environment are closely intertwined, and that genetic differentiation may be primarily shaped by the aridity gradient, which closely mirrors the geographic distance gradient.

Genetic basis of adaptation to aridity

Genotype-Environment Association analysis and genomic scan for differentiation

To take advantage of the full aridity gradient across *R. bechuanae*'s area of distribution, we performed the GEA analysis at the locality level using *BayPass* (by pooling individual genotype information from each of the 11 *R. bechuanae* localities). Given the above correlations, we expected stringent conditions for the GEA analysis. Indeed, since *BayPass* explicitly standardises population allele frequencies for their correlation structure - here induced by the neutral population structure but also the isolation by environment signal - we expected a high rate of false negatives when testing for an association with the Aridity Index.

Overall, despite these stringent conditions, 108 SNPs were found to be significantly associated with the Aridity Index (significance threshold of 10 dB), distributed across all genomic scaffolds (**Figure 4A**, **Table S8**). Ten SNPs displayed "decisive" evidence (Bayes Factor (*BF*) > 20 dB), 26 "very strong" evidence (15 dB < *BF* < 20 dB) and 72 SNPs displayed "strong" evidence (10 dB < *BF* < 15 dB). Among these, 68 were mapped within an annotated coding region of the *R. pumilio* genome, corresponding to 63 outlier 'association' genes in total.

The population differentiation analysis using *XtX* statistics simulated *XtX* values of 44,900 neutral SNPs to infer a 5% *XtX* significance threshold of 14.4. Across the three runs of the core model, which gave similar results, a total of 1,682 variants were identified as *XtX* outliers (**Figure 4B**). Of these, 65 SNPs displayed "strong evidence" of association with aridity, as well as a significant *XtX* value. Among them, 45 were

mapped within an annotated coding region of the *R. pumilio* genome, corresponding to 41 outlier 'association/differentiation' genes.

Functional analyses

All 63 genes associated with aridity, as well as the subset of 41 genes that were both associated with aridity and showed significant genetic differentiation across *R. bechuanae* populations, had identified homologues in the *M. musculus* genome, according to *Metascape* analyses (for a complete list of outlier genes with functional details, see **Table S8**).

The 41 'association/differentiation' outlier genes corresponded to twelve broad functional physiological pathways (**Table S8**), reflecting diverse metabolic functions. In **Figure 4**, we present one representative gene from each pathway, selected for showing the strongest signal of association with the Aridity Index. Among the most promising candidate genes were those involved in nervous system functions (e.g., *ADGRB1*, *WHAMM,* see **Table S8**). Notably, Gene Ontology (GO) terms related to neurotransmission showed the strongest degree of enrichment ($p_{adj} < 0.01$) among genes associated with the Aridity Index, with twelve GO terms significantly enriched overall **(Figure 5, Table S9**). These results suggest that genes involved in neuronal processes are those responding the most strongly to selection associated with varying degrees of aridity.

Several genes were also implicated in maintaining cellular integrity and stability of genetic or protein information under water-limited conditions. These included genes associated with osmoregulation (e.g., *EFCC1*), cellular response to stress (e.g., *SLC37A3*), and protein metabolism (e.g., *TMEM44*). Both gene sets – those associated with aridity and those both associated and differentiated – contained genes involved in the regulation of transcription (e.g., *ESRP2*, see **Table S8**). Relevant GO terms were significantly enriched among genes associated with aridity (**Figure 5**), and approached significance ($p_{adj} = 0.063$) among associated/differentiated genes **(Table S9**). Furthermore, GO terms related to cytoskeleton regulation and GTPase activity were also near the threshold for enrichment significance in the associated/differentiated gene set ($p_{adj} = 0.063$) (**Table S9**), with genes such as *WHAMM* or *EEFSEC* among top candidates (**Figure 4, Table S8**).

Finally, several genes with potential relevance for energy conservation under arid conditions were also identified. These included genes involved in lipid metabolism (e.g., *SDR9C7*), carbohydrate metabolism (e.g., *GNAQ*), immunity (e.g., *STING1*), and sweet taste perception (*TAS1R2*, see **Table S8**).

Discussion

In this study, we generated RAD-seq data to investigate patterns of intraspecific genetic diversity and responses to aridity in the African four-striped mouse, *Rhabdomys bechuanae*, a rodent species broadly distributed across a gradient of aridity - from semi-arid to hyper-arid summer rainfall regions of Southern Africa. By developing an original integrative approach, combining genomic data with fine-scale contemporary and historical aridity data available for the Southern African region and leveraging a natural aridity gradient to investigate signatures of aridity-driven intraspecific divergence using GEA analyses, our study provides significant insights into population-level responses to increased aridity in a homeothermic rodent species.

Our results reveal a recent decline in effective population size at the species level. Complementary analyses suggest this decline may reflect increasing population differentiation due to reduced connectivity among populations during Holocene aridification events. Additionally, we found that genetic diversity within the species is shaped by spatial structure of the environment, closely linked to variation in aridity across the *R. bechuanae* distribution range. We detected allele frequency shifts associated with aridity levels, enabling identification of candidate genes and functions potentially involved in adaptation to arid environments. As predicted, we identified candidate genomic regions associated with classical desert survival phenotypes, including DNA repair/protein modification, osmoregulation and carbohydrate/fat metabolism (Rocha et al., 2021; Cheng et al., 2023). However, remarkably, the strongest functional enrichment signal emerged for genes involved in neuronal transmission functions, bringing novel insights into the role of neurological processes in coping with osmotic stress and resource scarcity. Together, these findings support the hypothesis that genetic variation in *R. bechuanae* reflects both its demographic history, shaped by past climatic fluctuations, and the selective pressures imposed by spatial variation in aridity across its distribution range (Hereford & Winn, 2008; Kolbe et al., 2012).

Population structure and historical changes in effective population size in response to aridification

Although genetic differentiation amongst *R. bechuanae* localities was relatively low overall, our analyses revealed significant genetic clustering (three clusters) with the *sNMF* approach, along with significant correlation between genetic and geographical distances. This intraspecific structure was also supported by both PCA and neighbour-joining analyses. The observed pattern of isolation by distance provided evidence of restricted gene flow across the species' distribution range, which aligns with the patchiness of favourable habitat available for *R. bechuanae* in the arid and semi-arid regions of Southern Africa (Dufour et al., 2015; Ganem et al., 2020). Such habitat patchiness constrains dispersal, thereby promoting the genetic differentiation observed despite the species' relatively continuous regional distribution. Our findings support this

hypothesis through the detection of significant isolation by environment associated with the Aridity Index variable, although this variable showed strong correlation with geographic distance. The population structure analysis identified genetic clusters that correspond to populations at the opposing ends of the aridity gradient (Klein Pella representing the arid extreme, and Sandveld the semi-arid extreme), both of which have relatively peripheral positions in the species' range. This spatial arrangement suggests that the observed genetic structure could reflect adaptation to distinct environmental conditions along the aridity gradient, which may have imposed divergent selective pressures on *R. bechuanae* populations, supporting the hypothesis that aridity is a key driver of population structure in this species.

Historical demographic events, such as range expansions, population bottlenecks or divergence due to local adaptation or neutral processes, may have shaped patterns of population structure and diversity in *R. bechuanae* and may be partly related to past climatic events. This motivated us to infer historical changes in effective population size and correlate results with known past climatic events. We found that past variation in effective population size (N_e) was not specific to 'arid' or 'semi-arid' localities ; rather, a recent decline in N_e was inferred at the species level from 15 thousand years ago, which corresponds to a time interval during which multiple dry episodes took place in the *R. bechuanae* distribution range (15 – 11, 7–5 and 1.5–0.5 thousand years before present) according to palaeoclimatic reconstructions from central Southern Africa (Chevalier & Chase, 2016). During this time, the Southern Kalahari experienced a gradual increase in temperature (Lukich & Ecker, 2022) that may have contributed to this pattern. This observation upholds our prediction of a decrease in effective population size during dry periods, and suggests that biome changes induced by the last deglaciation in Southern Africa (Zhao et al., 2024) may have driven *R. bechuanae* population history.

Signals of reduced effective population size in demographic analyses do not necessarily reveal overall reduction in population size but can also reflect population structuring events (Loog, 2021). Although a previous study on the closely-related species *R. pumilio* showed the impact of heat and food availability on short and long-term population dynamics (Nater et al., 2018), we tested whether population fragmentation could also contribute to the observed reduced effective population size. Interestingly, when analysis was restricted to one sample per locality, the population decline was not detected, suggesting that the demographic signal observed at species level reflects partly population fragmentation. This result suggests that periods of increased aridity may have promoted genetic differentiation among *R. bechuanae* populations by reducing connectivity between them. Specifically, arid conditions and their impact on the vegetation may alter the distribution of nesting sites, concentrating them in scattered patches of appropriate vegetation (Dufour et al., 2015), and lack of favourable habitat (for food and shelter) may create geographic barriers that impede dispersal and gene flow between populations (Ota, 1998). This would align with our prediction that decreased population connectivity during dry episodes can drive population differentiation.

The patterns of genetic diversity emerging from aridification remain relatively unexplored in the literature; previous studies on vertebrate models report a link between past changes in aridity and genetic diversification (Garcia et al., 2011), or changes in connectivity (Merchant et al., 2024). Overall, our results support the hypothesis that past climatic fluctuations in the Southern Kalahari and central South Africa may have been important factors shaping effective population size and population structure of *R. bechuanae,* although other neutral and adaptive factors could also be involved. Historical periods of increased aridity could not only have reduced population connectivity but may have also created opportunities for local adaptation if geographically separated populations experienced different environmental pressures along the aridity gradient. This dual effect of aridity, both fragmenting populations and creating divergent selective environments, provides a comprehensive explanation for the observed genetic structure in *R. bechuanae*.

Adaptive genomic variation along an aridity gradient

This study enabled the identification of several SNPs associated with the Aridity Index variable, despite the very stringent conditions of the *BayPass* analysis due to significant correlations between genetic, geographic and environmental (Aridity Index) distances. While these conditions reduced the overall power of the analysis, they also strengthened the reliability of the identified SNPs. Additional approaches, such as genome-wide association studies based on well characterised physiological phenotypes (Keilani et al., 2025), could complement our findings by revealing other genomic regions underlying key traits for adaptation to aridity. It is unlikely our inferences were biased by small sample sizes in some localities, as GEA studies using Bayes' factor are not particularly sensitive to sample sizes (Forester et al., 2018). Refining our candidate list of outlier SNPs by focusing on SNPs located within coding regions and exhibiting significant genetic differentiation among populations allowed us to identify the most promising outlier SNPs and candidate genomic regions.

Our hypothesis of differential selective pressures acting on *R. bechuanae* populations along the aridity gradient was supported by the identification of outlier SNPs located within coding regions of genes associated with important biological functions for adaptation to an arid or desert lifestyle. We identified genes involved in energetic balance (e.g., lipid metabolism, immunity) and osmotic balance (e.g., osmoregulation, neuronal processes, body fluids regulation), as well as cellular responses to dehydration (e.g., protein metabolism, regulation of transcription and cytoskeleton). *Rhabdomys bechuanae* face environmental constraints, which can involve high temperatures and resource scarcity. Even in semi-arid conditions, these challenges cause a strain on homeostasis, as revealed in *R. bechuanae* populations from semi-arid locations by blood markers of osmoregulation, immunity and energy levels (Keilani et al., 2025). More arid conditions could cause cellular stress via thermal stress, and exacerbate the negative effects of dehydration due to lack of water and rapid water loss (e.g., respiratory water loss, evaporative cooling). Overall, consistent with our predictions, our results suggest that various physiological mechanisms may enable *R. bechuanae* to

cope with harsh conditions and inhabit drier environments. We synthesised previous findings from other species aligning with our results in **Table S8**, and discuss in more detail the functions of interest for adaptation to arid conditions in light of this literature review.

In numerous studies carried out on mammalian species, the same genes, or those involved in similar functions as reported in our study, are highlighted as containing SNPs varying with an aridity climatic index (e.g., Moisture index, Flori et al., 2019), differentially expressed in arid conditions (e.g., water restriction, McManes & Eisen, 2014), or showing signs of selection in specialist organisms (e.g., Cheng et al., 2023). Many of these gene lists identified from previous studies and in our study intersect, suggesting a genetic basis for adaptation to aridity partly shared across various mammals. Studies on the basis of adaptation to xeric environments show convergent changes in tissue composition, gene expression as well as coding sequences, with significant overlap mostly among related species (Rocha et al., 2021, Chalopin et al., 2025).

Our findings align with key shared mechanisms of adaptation to low resource availability observed across mammals, especially those involved in maintaining energetic balance, such as lipid metabolism (Wu et al.; 2014; Kim et al., 2016; Sugden et al., 2018; Chebii et al., 2021; Rocha et al., 2021), and immune response (Vermeulen et al., 2015; McDade et al., 2016; Brusch et al., 2019). Interestingly, acquiring energy in a food-scarce environment may also involve diet shifts in response to arid food sources, such as desert plants and insects. These shifts can be facilitated by mutations in taste receptor genes, such as the sweet taste receptor identified in our study or in the numbat *Myrmecobius fasciatus* (Peel et al., 2022), and bitter taste receptors in the *Peromyscus eremicus* cactus mouse, where reduced sensitivity may broaden dietary options by increasing palatability in resource-limited habitats (Tigano et al., 2020).

Given the scarcity of water in deserts, osmotic balance maintenance imposes strong selective pressures on osmoregulation mechanisms. In *R. bechuanae*, we identified genes containing significantly associated outlier SNPs involved in osmoregulatory, many of which were also reported in other desert rodents (e.g., *Allactaga sibirica*, *Dipus sagitta*, *Meriones meridianus* and *Phodopus roborovskii* : Cheng et al., 2023, **Table S8**).

At the cellular level, dehydration-induced changes, such as cytoskeletal rearrangements (Burg et al., 2007), may explain the observed association between aridity and genes regulating the cytoskeleton (e.g., GO terms: filopodium assembly), suggesting possible adaptations to hyperosmotic stress in the striped mouse. However, dehydration affects cells beyond cytoskeletal changes: it includes protein denaturation, transcription/translation suppression, altered adhesion, increased DNA breaks and protein oxidation, cell cycle arrest, and apoptosis (Kampinga, 1993; Lamitina, et al., 2006; Burg et al., 2007). Consistent with these responses, SNPs associated with aridity were found in genes involved in protein synthesis and modification, DNA repair, and cellular cycle regulation. Genes containing outlier SNPs were also enriched for GO terms related to transcription regulation. These findings align with previous results from selective sweeps studies in arid-adapted rodents (e.g., *Peromyscus eremicus* : Tigano

et al., 2020), and meta-analyses of metazoan genomic and transcriptomic data showing selection on translation-related genes in harsh environments (Porcelli et al., 2015). Indeed, cellular responses to hyperosmotic stress commonly involve regulating protein translation and degradation (Lamitina et al., 2006). Overall, our results suggest strong selection on genes mitigating thermal or hyperosmotic stress and restoring cellular function after acute stress (Kampinga, 1993).

Water balance in mammals is regulated by brain-body communication mediated by central nervous system structures that exhibit marked plasticity in response to environmental cues in arid-adapted species (Rabhi et al., 1993; Djazouli Alim et al., 2019). This may help explain the strong enrichment of GO terms related to nervous system functions, such as nervous system development and neurotransmitter regulation, which emerged as the most significant signal in our study. Although neurological functions are occasionally noted in aridity adaptation studies (e.g., Kim et al., 2015; Yang et al., 2016; Flori et al., 2019; Blumstein & MacManes, 2024), they are rarely emphasised as key survival factors under arid conditions. When mentioned, they are typically discussed in the context of selective brain cooling in large, heat-adapted mammals (Hetem et al., 2012; Strauss et al., 2017). Neurotransmitters, essential for maintaining homeostasis (Mittal et al., 2017), regulate blood flow – a critical function in arid environments where dehydration can restrict organ perfusion (Trangmar & Gonzalez-Alonso, 2017). Our identification of genes involved in heart rate regulation and vasculature development - also differentially expressed in other desert-adapted rodents such as *Abrothrix olivacea* (Giorello et al., 2018), supports this. Additionally, *R. bechuanae* exhibit lower blood concentrations than mesic-dwelling congeners (Keilani et al., 2025), a condition possibly alleviated by improved blood flow. Together, these results support the involvement of a previously underappreciated functional category – neurological pathways - in the genomic response to aridity, offering new directions for research in the evolutionary physiology of aridity adaptation.

Conclusion

This study sheds new light on how historical climatic fluctuations and environmental gradients shaped population structure and genetically based responses to aridity in *Rhabdomys bechuanae*, a rodent species inhabiting some of the most arid regions of Southern Africa. By integrating genomic data with fine-scale environmental information, we uncovered signatures of population differentiation and polygenic adaptation along an aridity gradient, involving functions linked to osmoregulation, energetic balance, and cellular stress responses. The list of candidate genes and biological functions provides a valuable foundation for further research on the molecular mechanisms of adaptation to arid environments. While we focused on coding regions to identify genes and functions associated with the Aridity Index, future work incorporating transcriptomic and non-coding genomic data will be essential to assess the role of gene expression regulation in this species' response to increasing aridity and to disentangle the relative contributions of coding versus regulatory variation. Moreover, although this study

revealed genetically based signatures of adaptation, avoidance behaviours and plastic physiological responses likely play a critical role in coping withenvironmental stress imposed by seasonal variation in aridity (Blumstein & MacManes, 2023). Future physiological and transcriptomic studies manipulating water and food availability, will be key to elucidating the interplay between phenotypic plasticity and local adaptation in shaping species' responses to arid conditions.

References

Ababaikeri, Buweihailiqiemu, Shamshidin Abduriyim, Yilamujiang Tohetahong, Tayerjan Mamat, Adil Ahmat, and Mahmut Halik. “Whole-Genome Sequencing of Tarim Red Deer (Cervus Elaphus Yarkandensis) Reveals Demographic History and Adaptations to an Arid-Desert Environment.” Frontiers in Zoology 17, no. 1 (October 16, 2020). https://doi.org/10.1186/s12983-020-00379-5.

Baird, Nathan A., Paul D. Etter, Tressa S. Atwood, Mark C. Currey, Anthony L. Shiver, Zachary A. Lewis, Eric U. Selker, William A. Cresko, and Eric A. Johnson. “Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers.” PLoS ONE 3, no. 10 (October 13, 2008): e3376. https://doi.org/10.1371/journal.pone.0003376.

Benning, John W., Alexai Faulkner, and David A. Moeller. “Rapid Evolution during Climate Change: Demographic and Genetic Constraints on Adaptation to Severe Drought.” Proceedings of the Royal Society B: Biological Sciences 290, no. 1998 (May 10, 2023). https://doi.org/10.1098/rspb.2023.0336.

Bittner, Noëlle K. J., Katya L. Mack, and Michael W. Nachman. “Gene Expression Plasticity and Desert Adaptation in House Mice*.” Evolution 75, no. 6 (January 23, 2021): 1477–91. https://doi.org/10.1111/evo.14172.

Bjornstad, O. N., & Bjornstad, M. O. N. (2016). Package ‘ncf’. *Spatial nonparametric covariance functions*, *5*, 1-1.

Blumstein, D. M., & MacManes, M. D. (2023). When the tap runs dry: The physiological effects of acute experimental dehydration in Peromyscus eremicus. Journal of Experimental Biology, 226(23). https://doi.org/10.1242/jeb.246386

Blumstein, D., & MacManes, M. (2024). The multi-tissue gene expression and physiological responses of water deprived Peromyscus eremicus. *BMC genomics*, *25*(1), 770.

Bridges, Te, and Nv James. “The Hypothalamo□Neurohypophysial system of native australian desert rodents. The vasopressin and oxytocin contents of hypothalamus and posterior pituitary of Notomys alexis and Pseudomys australis compared with those of the laboratory rat and mouse in different states of water balance.” Australian Journal of Experimental Biology and Medical Science 60, no. 3 (June 1982): 265–83. https://doi.org/10.1038/icb.1982.31.

Brusch, George A., IV, Keith Christian, Gregory P. Brown, Richard Shine, and Dale F. DeNardo. “Dehydration Enhances Innate Immunity in a Semiaquatic Snake from the Wet□dry Tropics.” Journal of Experimental Zoology Part A: Ecological and Integrative Physiology 331, no. 4 (March 14, 2019): 245–52. https://doi.org/10.1002/jez.2260.

Burg, Maurice B., Joan D. Ferraris, and Natalia I. Dmitrieva. “Cellular Response to Hyperosmotic Stresses.” Physiological Reviews 87, no. 4 (October 2007): 1441–74. https://doi.org/10.1152/physrev.00056.2006.

Catchen, Julian, Paul A. Hohenlohe, Susan Bassham, Angel Amores, and William A. Cresko. “Stacks: An Analysis Tool Set for Population Genomics.” Molecular Ecology 22, no. 11 (May 24, 2013): 3124–40. https://doi.org/10.1111/mec.12354.

Chalopin, Domitille, Carine Rey, Jeremy Ganofsky, Juliana Blin, Pascale Chevret, Marion Mouginot, Bastien Boussau, Sophie Pantalacci and Marie Sémon (2025). Convergent transcriptomic and genomic adaptation in xeric rodents.

Chebii, Vivien J., Emmanuel A. Mpolya, Samuel O. Oyola, Antoinette Kotze, Jean-Baka Domelevo Entfellner, and J. Musembi Mutuku. “Genome Scan for Variable Genes Involved in Environmental Adaptations of Nubian Ibex.” Journal of Molecular Evolution 89, no. 7 (June 17, 2021): 448–57. https://doi.org/10.1007/s00239-021-10015-3.

Chen, I-Ching, Jane K. Hill, Ralf Ohlemüller, David B. Roy, and Chris D. Thomas. “Rapid Range Shifts of Species Associated with High Levels of Climate Warming.” Science 333, no. 6045 (August 19, 2011): 1024–26. https://doi.org/10.1126/science.1206432.

Cheng, Jilong, Xingwen Peng, Hong Li, Anderson Feijó, Lin Xia, Georgy I. Shenbrot, Deyan Ge, Zhixin Wen, Dehua Wang, and Qisen Yang. “Similar Adaptative Mechanism but Divergent Demographic History of Four Sympatric Desert Rodents in Eurasian Inland.” Communications Biology 6, no. 1 (January 12, 2023). https://doi.org/10.1038/s42003-023-04415-y.

Chevalier, Manuel, and Brian M. Chase. “Determining the Drivers of Long□term Aridity Variability: A Southern African Case Study.” Journal of Quaternary Science 31, no. 2 (February 2016): 143–51. https://doi.org/10.1002/jqs.2850.

Chikhi, L., Rodríguez, W., Grusea, S., Santos, P., Boitard, S., & Mazet, O. (2018). The IICR (inverse instantaneous coalescence rate) as a summary of genomic diversity: insights into demographic inference and model choice. Heredity, 120(1), 13-24.

Chikhi, L., Sousa, V. C., Luisi, P., Goossens, B., & Beaumont, M. A. (2010). The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. Genetics, 186(3), 983-995.

Coop, Graham, David Witonsky, Anna Di Rienzo, and Jonathan K Pritchard. “Using Environmental Correlations to Identify Loci Underlying Local Adaptation.” Genetics 185, no. 4 (August 1, 2010): 1411–23. https://doi.org/10.1534/genetics.110.114819.

Crawford, J. E., & Lazzaro, B. P. (2012). Assessing the accuracy and power of population genetic inference from low-pass next-generation sequencing data. Frontiers in genetics, 3, 66.

Cruaud, Astrid, Mathieu Gautier, Maxime Galan, Julien Foucaud, Laure Sauné, Gwenaëlle Genson, Emeric Dubois, Sabine Nidelet, Thierry Deuve, and Jean-Yves Rasplus. “Empirical Assessment of RAD Sequencing for Interspecific Phylogeny.” Molecular Biology and Evolution 31, no. 5 (February 3, 2014): 1272–74. https://doi.org/10.1093/molbev/msu063.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. Bioinformatics, 27(15), 2156-2158.

Djazouli Alim, Fatma Zohra, Elena V. Romanova, Yea-Ling Tay, Ahmad Yamin bin Abdul Rahman, Kok-Gan Chan, Kar-Wai Hong, Mark Rogers, et al. “Seasonal Adaptations of the Hypothalamo-Neurohypophyseal System of the Dromedary Camel.” PLOS ONE 14, no. 6 (June 18, 2019): e0216679. https://doi.org/10.1371/journal.pone.0216679.

Dufour, Claire M. S., Christine Meynard, Johan Watson, Camille Rioux, Simon Benhamou, Julie Perez, Jurie J. du Plessis, Nico Avenant, Neville Pillay, and Guila Ganem. “Space Use Variation

in Co-Occurring Sister Species: Response to Environmental Variation or Competition?" PLOS ONE 10, no. 2 (February 18, 2015): e0117750. https://doi.org/10.1371/journal.pone.0117750.

Dufour, C. M. S., Pillay, N., Avenant, N., Watson, J., Loire, E., & Ganem, G. (2019). Habitat characteristics and species interference influence space use and nest-site occupancy: Implications for social variation in two sister species. Oikos, 128, 503–516. https://doi.org/10.1111/oik.05357

Dupont, L. M., Rommerskirchen, F., Mollenhauer, G., & Schefuß, E. (2013). Miocene to Pliocene changes in South African hydrology and vegetation in relation to the expansion of C4 plants. *Earth and Planetary Science Letters*, *375*, 408-417.

Du Toit, Nina, Bettine Jansen van Vuuren, Sonja Matthee, and Conrad A. Matthee. "Biome Specificity of Distinct Genetic Lineages within the Four-Striped Mouse Rhabdomys Pumilio (Rodentia: Muridae) from Southern Africa with Implications for Taxonomy." Molecular Phylogenetics and Evolution 65, no. 1 (October 2012): 75–86. https://doi.org/10.1016/j.ympev.2012.05.036.

Easterling, David R., Gerald A. Meehl, Camille Parmesan, Stanley A. Changnon, Thomas R. Karl, and Linda O. Mearns. "Climate Extremes: Observations, Modeling, and Impacts." Science 289, no. 5487 (September 22, 2000): 2068–74. https://doi.org/10.1126/science.289.5487.2068.

Engelbrecht, F.A., Steinkopf, J., Padavatan, J., Midgley, G.F. (2024). Projections of Future Climate Change in Southern Africa and the Potential for Regional Tipping Points. In: von Maltitz, G.P., *et al.* Sustainability of Southern African Ecosystems under Global Change. Ecological Studies, vol 248. Springer, Cham. https://doi.org/10.1007/978-3-031-10948-5_7.

Etter, Paul D., Susan Bassham, Paul A. Hohenlohe, Eric A. Johnson, and William A. Cresko. "SNP Discovery and Genotyping for Evolutionary Genetics Using RAD Sequencing." In Methods in Molecular Biology, 157–78. Totowa, NJ: Humana Press, 2011. http://dx.doi.org/10.1007/978-1-61779-228-1_9.

Flori, L., Moazami□Goudarzi, K., Alary, V., Araba, A., Boujenane, I., Boushaba, N., ... & Gautier, M. (2019). A genomic map of climate adaptation in Mediterranean cattle breeds. Molecular ecology, 28(5), 1009-1029.

Forester, Brenna R., Jesse R. Lasky, Helene H. Wagner, and Dean L. Urban. "Comparing Methods for Detecting Multilocus Adaptation with Multivariate Genotype–Environment Associations." Molecular Ecology 27, no. 9 (April 23, 2018): 2215–33. https://doi.org/10.1111/mec.14584.

Frichot, Eric, Sean D. Schoville, Guillaume Bouchard, and Olivier François. "Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models." Molecular Biology and Evolution 30, no. 7 (March 29, 2013): 1687–99. https://doi.org/10.1093/molbev/mst063.

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. Methods in ecology and evolution, 6(8), 925-929.

Ganem, Guila, Claire Marie Soleil Dufour, Nico L. Avenant, Pierre Caminade, Seth J. Eiseb, Christelle Tougard, and Neville Pillay. "An Update on the Distribution and Diversification of Rhabdomys Sp. (Muridae, Rodentia)." Journal of Vertebrate Biology 69, no. 2 (July 7, 2020): 1. https://doi.org/10.25225/jvb.20013.

Garcia, J. T., Alda, F., Terraube, J., Mougeot, F., Sternalski, A., Bretagnolle, V., & Arroyo, B. (2011). Demographic history, genetic structure and gene flow in a steppe-associated raptor species. BMC evolutionary biology, 11, 1-11.

Gautier, Mathieu. “Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates.” Genetics 201, no. 4 (October 19, 2015): 1555–79. https://doi.org/10.1534/genetics.115.181453.

Giorello, F. M., Feijoo, M., D'Elía, G., Naya, D. E., Valdez, L., Opazo, J. C., & Lessa, E. P. (2018). An association between differential expression and genetic divergence in the Patagonian olive mouse (Abrothrix olivacea). Molecular Ecology, 27(16), 3274-3286.

Gomulkiewicz, Richard, and David Houle. “Demographic and Genetic Constraints on Evolution.” The American Naturalist 174, no. 6 (December 2009): E218–29. https://doi.org/10.1086/645086.

Gutenkunst, Ryan N., Ryan D. Hernandez, Scott H. Williamson, and Carlos D. Bustamante. “Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data.” PLoS Genetics 5, no. 10 (October 23, 2009): e1000695. https://doi.org/10.1371/journal.pgen.1000695.

Hereford, Joe, and Alice A. Winn. “Limits to Local Adaptation in Six Populations of the Annual plantDiodia Teres.” New Phytologist 178, no. 4 (March 31, 2008): 888–96. https://doi.org/10.1111/j.1469-8137.2008.02405.x.

Hetem, R. S., Strauss, W. M., Fick, L. G., Maloney, S. K., Meyer, L. C., Fuller, A., ... & Mitchell, D. (2012). Selective brain cooling in Arabian oryx (Oryx leucoryx): a physiological mechanism for coping with aridity?. *Journal of Experimental Biology*, *215*(22), 3917-3924.

Jeffreys, Harold. The theory of J EFFREYS , H. (1961). Theory of Probability, 3rd ed. Oxford Classic Texts in the Physical Sciences. Oxford Univ. Press, Oxford. MR1647885.

Kampinga, Harm H. “Thermotolerance in Mammalian Cells Protein Denaturation and Aggregation, and Stress Proteins.” Journal of Cell Science 104, no. 1 (January 1, 1993): 11–17. https://doi.org/10.1242/jcs.104.1.11.

Keilani, H., Avenant, N., Caminade, P., Pillay, N., & Ganem, G. (2025). Negative impact of mild arid conditions in natural rodent populations revealed using markers of physiological condition in natura. Peer Community Journal, 5.

Kim, E-S, A R Elbeltagy, A M Aboul-Naga, B Rischkowsky, B Sayre, J M Mwacharo, and M F Rothschild. “Multiple Genomic Signatures of Selection in Goats and Sheep Indigenous to a Hot Arid Environment.” Heredity 116, no. 3 (November 11, 2015): 255–64. https://doi.org/10.1038/hdy.2015.94.

Kolbe, Jason J., Manuel Leal, Thomas W. Schoener, David A. Spiller, and Jonathan B. Losos. “Founder Effects Persist Despite Adaptive Differentiation: A Field Experiment with Lizards.” Science 335, no. 6072 (March 2, 2012): 1086–89. https://doi.org/10.1126/science.1209566.

Lamitina, Todd, Chunyi George Huang, and Kevin Strange. “Genome-Wide RNAi Screening Identifies Protein Damage as a Regulator of Osmoprotective Gene Expression.” Proceedings of the National Academy of Sciences 103, no. 32 (August 8, 2006): 12173–78. https://doi.org/10.1073/pnas.0602987103

Langmead, Ben, Cole Trapnell, Mihai Pop, and Steven L Salzberg. “Ultrafast and Memory-Efficient Alignment of Short DNA Sequences to the Human Genome.” Genome Biology 10, no. 3 (2009): R25. https://doi.org/10.1186/gb-2009-10-3-r25

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and 1000 Genome Project Data Processing Subgroup, The Sequence alignment/map (SAM) format and SAMtools, *Bioinformatics* (2009) 25(16) 2078-9

Liu, Xiaoming, and Yun-Xin Fu. “Exploring Population Size Changes Using SNP Frequency Spectra.” Nature Genetics 47, no. 5 (April 6, 2015): 555–59. https://doi.org/10.1038/ng.3254

Liu, Xiaoming, and Yun-Xin Fu. “Stairway Plot 2: demographic history inference with folded SNP frequency spectra.” Genome Biology 280, no. 21 (November 17, 2020): 1474-760X. https://doi.org/10.1186/s13059-020-02196-9

Loog, L. (2021). Sometimes hidden but always there: the assumptions underlying genetic inference of demographic histories. Philosophical Transactions of the Royal Society B, 376(1816), 20190719.

Lukich, Vasilija, and Michaela Ecker. “Pleistocene Environments in the Southern Kalahari of South Africa.” Quaternary International 614 (March 2022): 50–58. https://doi.org/10.1016/j.quaint.2021.03.008.

MacManes, Matthew D., and Michael B. Eisen. “Characterization of the Transcriptome, Nucleotide Sequence Polymorphism, and Natural Selection in the Desert Adapted mousePeromyscus Eremicus.” PeerJ 2 (October 28, 2014): e642. https://doi.org/10.7717/peerj.642.

Malaspinas, Anna-Sapfo, Michael C. Westaway, Craig Muller, Vitor C. Sousa, Oscar Lao, Isabel Alves, Anders Bergström, et al. “A Genomic History of Aboriginal Australia.” Nature 538, no. 7624 (September 21, 2016): 207–14. https://doi.org/10.1038/nature18299.

Marra, Nicholas J., Andrea Romero, and J. Andrew DeWoody. “Natural Selection and the Genetic Basis of Osmoregulation in Heteromyid Rodents as Revealed byRNA□seq.” Molecular Ecology 23, no. 11 (May 16, 2014): 2699–2711. https://doi.org/10.1111/mec.12764.

McDade, Thomas W., Alexander V. Georgiev, and Christopher W. Kuzawa. “Trade-Offs between Acquired and Innate Immune Defenses in Humans.” Evolution, Medicine, and Public Health 2016, no. 1 (2016): 1–16. https://doi.org/10.1093/emph/eov033.

McLaughlin, J. F., & Winker, K. (2020). An empirical examination of sample size effects on population demographic estimates in birds using single nucleotide polymorphism (SNP) data. PeerJ, 8, e9939.

Menzel, Annette, Tim H. Sparks, Nicole Estrella, Elisabeth Koch, Anto Aasa, Rein Ahas, Kerstin Alm□kübler, et al. “European Phenological Response to Climate Change Matches the Warming Pattern.” Global Change Biology 12, no. 10 (August 21, 2006): 1969–76. https://doi.org/10.1111/j.1365-2486.2006.01193.x.

Merchant, H. N., Ivanova, A., Hart, D. W., García, C., Bennett, N. C., Portugal, S. J., & Faulkes, C. G. (2024). Patterns of genetic diversity and gene flow associated with an aridity gradient in populations of common mole-rats, Cryptomys hottentotus hottentotus. Genome Biology and Evolution, 16(7), evae144.

Meynard, Christine N., Neville Pillay, Manon Perrigault, Pierre Caminade, and Guila Ganem. "Evidence of Environmental Niche Differentiation in the Striped Mouse (Rhabdomys Sp.): Inference from Its Current Distribution in Southern Africa." Ecology and Evolution 2, no. 5 (April 12, 2012): 1008–23. https://doi.org/10.1002/ece3.219.

Milholland, Brandon, Xiao Dong, Lei Zhang, Xiaoxiao Hao, Yousin Suh, and Jan Vijg. "Differences between Germline and Somatic Mutation Rates in Humans and Mice." Nature Communications 8, no. 1 (May 9, 2017). https://doi.org/10.1038/ncomms15183.

Mittal, R., Debs, L. H., Patel, A. P., Nguyen, D., Patel, K., O'Connor, G., ... & Liu, X. Z. (2017). Neurotransmitters: The critical modulators regulating gut–brain axis. *Journal of cellular physiology*, *232*(9), 2359-2372.

Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520–562 (2002). https://doi.org/10.1038/nature01262

Nater, Chloé R., Koen J. van Benthem, Cindy I. Canale, Carsten Schradin, and Arpat Ozgul. "Density Feedbacks Mediate Effects of Environmental Change on Population Dynamics of a Semidesert Rodent." Journal of Animal Ecology 87, no. 6 (August 21, 2018): 1534–46. https://doi.org/10.1111/1365-2656.12888.

Naumann, G., Alfieri, L., Wyser, K., Mentaschi, L., Betts, R. A., Carrao, H., ... & Feyen, L. (2018). Global changes in drought conditions under different levels of warming. *Geophysical Research Letters*, *45*(7), 3285-3296.

Ota, H. (1998). Geographic patterns of endemism and speciation in amphibians and reptiles of the Ryukyu Archipelago, Japan, with special reference to their paleogeographical implications. Researches on Population Ecology, 40, 189-204.

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics, 35(3), 526-528.

Peel, E., Silver, L., Brandies, P., Hayakawa, T., Belov, K., & Hogg, C. J. (2022). Genome assembly of the numbat (Myrmecobius fasciatus), the only termitivorous marsupial. Gigabyte, 2022, gigabyte47.

Porcelli, D, R K Butlin, K J Gaston, D Joly, and R R Snook. "The Environmental Genomics of Metazoan Thermal Adaptation." Heredity 114, no. 5 (March 4, 2015): 502–14. https://doi.org/10.1038/hdy.2014.119.

Puechmaille, Sebastien J. "The Program Structure Does Not Reliably Recover the Correct Population Structure When Sampling Is Uneven: Subsampling and New Estimators Alleviate the Problem." Molecular Ecology Resources 16, no. 3 (March 2, 2016): 608–27. https://doi.org/10.1111/1755-0998.12512.

Rabhi, M., Ugrumov, M., Polenova, O., Bengelloun, W., & Calas, A. (1993). Modifications morphofonctionnelles du système hypothalamo-hypophysaire vasopressinergique après déshydratation chronique chez la Mérione (Meriones shawi). *Ann Endocrinol*, *54*, 15.

Richardson, Rose, Charles Y. Feigin, Beatriz Bano-Otalora, Matthew R. Johnson, Annette E. Allen, Jongbeom Park, Richard J. McDowell, et al. "The Genomic Basis of Temporal Niche Evolution in a Diurnal Rodent." Current Biology 33, no. 15 (August 2023): 3289-3298.e6. https://doi.org/10.1016/j.cub.2023.06.068.

Rocha, Joana L., Raquel Godinho, José C. Brito, and Rasmus Nielsen. “Life in Deserts: The Genetic Basis of Mammalian Desert Adaptation.” Trends in Ecology & Evolution 36, no. 7 (July 2021): 637–50. https://doi.org/10.1016/j.tree.2021.03.007.

Rocha, Joana L., Pedro Silva, Nuno Santos, Mónia Nakamura, Sandra Afonso, Abdeljebbar Qninba, Zbyszek Boratynski, et al. “North African Fox Genomes Show Signatures of Repeated Introgression and Adaptation to Life in Deserts.” Nature Ecology & Evolution 7, no. 8 (June 12, 2023): 1267–86. https://doi.org/10.1038/s41559-023-02094-w

Rousset, François. “Genetic Differentiation and Estimation of Gene Flow from F-Statistics Under Isolation by Distance.” Genetics 145, no. 4 (April 1, 1997): 1219–28. https://doi.org/10.1093/genetics/145.4.1219.

Rymer, Tasmin L., Neville Pillay, and Carsten Schradin. “Resilience to Droughts in Mammals: A Conceptual Framework for Estimating Vulnerability of a Single Species.” The Quarterly Review of Biology 91, no. 2 (June 2016): 133–76. https://doi.org/10.1086/686810.

Sarabia, Carlos, Bridgett vonHoldt, Juan C. Larrasoaña, Vicente Uríos, and Jennifer A. Leonard. “Pleistocene Climate Fluctuations Drove Demographic History of African Golden Wolves (Canis Lupaster).” Molecular Ecology 30, no. 23 (January 16, 2021): 6101–20. https://doi.org/10.1111/mec.15784.

Shumate, A., & Salzberg, S. L. (2021). Liftoff: accurate mapping of gene annotations. Bioinformatics, 37(12), 1639-1643.

Skelly, David K., Liana N. Joseph, Hugh P. Possingham, L. Kealoha Freidenburg, Thomas J. Farrugia, Michael T. Kinnison, and Andrew P. Hendry. “Evolutionary Responses to Climate Change.” Conservation Biology 21, no. 5 (August 14, 2007): 1353–55. https://doi.org/10.1111/j.1523-1739.2007.00764.x.

Smit, AFA, Hubley, R & Green, P. RepeatMasker Open-4.0. 2013-2015 http://www.repeatmasker.org.

Solmsen, N., J. Johannesen, and C. Schradin. “Highly Asymmetric Fine-Scale Genetic Structure between Sexes of African Striped Mice and Indication for Condition Dependent Alternative Male Dispersal Tactics.” Molecular Ecology 20, no. 8 (March 3, 2011): 1624–34. https://doi.org/10.1111/j.1365-294x.2011.05042.x.

Strauss, W. M., Hetem, R. S., Mitchell, D., Maloney, S. K., O'Brien, H. D., Meyer, L. C., & Fuller, A. (2017). Body water conservation through selective brain cooling by the carotid rete: a physiological feature for surviving climate change?. *Conservation Physiology*, *5*(1), cow078.

Sugden, Lauren Alpert, Elizabeth G. Atkinson, Annie P. Fischer, Stephen Rong, Brenna M. Henn, and Sohini Ramachandran. “Localization of Adaptive Variants in Human Genomes Using Averaged One-Dependence Estimation.” Nature Communications 9, no. 1 (February 19, 2018). https://doi.org/10.1038/s41467-018-03100-7.

The Alliance of Genome Resources Consortium , Updates to the Alliance of Genome Resources central infrastructure, *Genetics*, Volume 227, Issue 1, May 2024, iyae049, https://doi.org/10.1093/genetics/iyae049

Tigano, Anna, Jocelyn P. Colella, and Matthew D. MacManes. “Comparative and Population Genomics Approaches Reveal the Basis of Adaptation to Deserts in a Small Rodent.” Molecular Ecology 29, no. 7 (March 23, 2020): 1300–1314. https://doi.org/10.1111/mec.15401.

Trangmar, S. J., & González-Alonso, J. (2017). New insights into the impact of dehydration on blood flow and metabolism during exercise. *Exercise and sport sciences reviews*, *45*(3), 146-153.

Ventura□Rojas, Perla D., Alberto González□Romero, Claudia E. Moreno, and Vinicio J. Sosa. “Effect of Rainfall, Temperature and Climate Change on the Ecology of the Rodents of Arid Zones: A Review.” Mammal Review, June 26, 2024. https://doi.org/10.1111/mam.12372.

Vermeulen, Anke, Wendt Müller, Kevin D. Matson, B. Irene Tieleman, Lieven Bervoets, and Marcel Eens. “Sources of Variation in Innate Immunity in Great Tit Nestlings Living along a Metal Pollution Gradient: An Individual-Based Approach.” Science of The Total Environment 508 (March 2015): 297–306. https://doi.org/10.1016/j.scitotenv.2014.11.095.

Wang, Jinliang. “Fast and accurate population admixture inference from genotype data from a few microsatellites to millions of SNPs.” Heredity 129, no.2 (August 1st, 2022) : 1365-2540. https://doi.org/10.1038/s41437-022-00535-z

Wiens, John J. “Climate-Related Local Extinctions Are Already Widespread among Plant and Animal Species.” PLOS Biology 14, no. 12 (December 8, 2016): e2001104. https://doi.org/10.1371/journal.pbio.2001104.

Willmer, Pat, Graham Stone, and Ian Johnston. Environmental Physiology of Animals. John Wiley & Sons, 2009.

Wright, B. R., Grueber, C. E., Lott, M. J., Belov, K., Johnson, R. N., & Hogg, C. J. (2019). Impact of reduced-representation sequencing protocols on detecting population structure in a threatened marsupial. Molecular Biology Reports, 46(5), 5575-5580.

Wu, Huiguang, Xuanmin Guang, Mohamed B. Al-Fageeh, Junwei Cao, Shengkai Pan, Huanmin Zhou, Li Zhang, et al. “Camelid Genomes Reveal Evolution and Adaptation to Desert Environments.” Nature Communications 5, no. 1 (October 21, 2014). https://doi.org/10.1038/ncomms6188.

Yang, J. I., Li, W. R., Lv, F. H., He, S. G., Tian, S. L., Peng, W. F., ... & Liu, M. J. (2016). Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Molecular biology and evolution*, *33*(10), 2576-2592.

Zhao, X., Lu, X., Meadows, M., Dupont, L., Mao, L., Xu, Y., ... & Zabel, M. (2024). Biome variability in southernmost Africa since the last deglaciation recorded in marine sediments. Quaternary Science Reviews, 343, 108912.

Zomer, Robert J., Jianchu Xu, and Antonio Trabucco. “Version 3 of the Global Aridity Index and Potential Evapotranspiration Database.” Scientific Data 9, no. 1 (July 15, 2022). https://doi.org/10.1038/s41597-022-01493-1

Figures

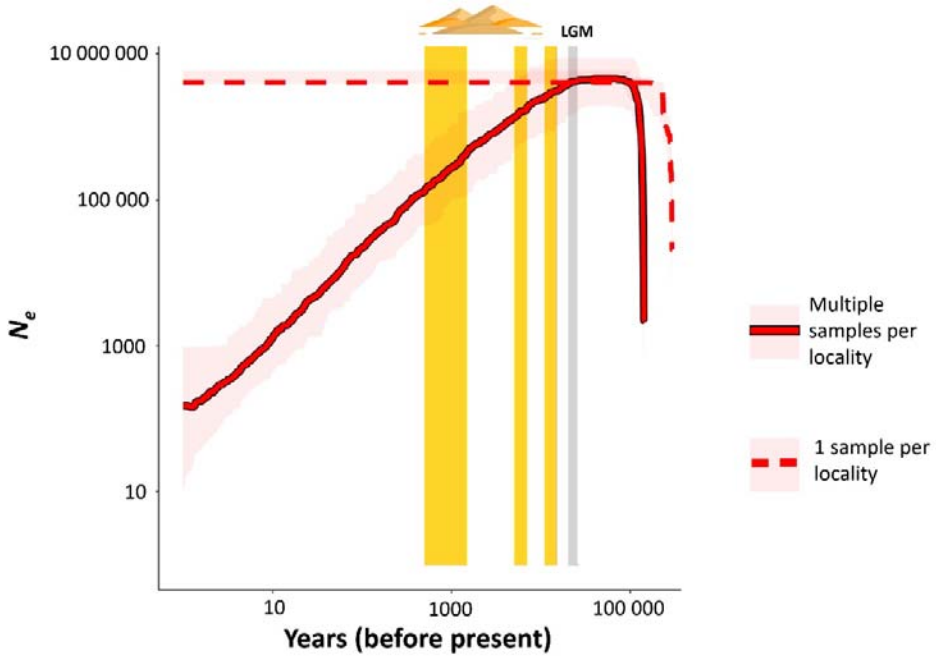


Figure 1: Historical changes in effective population size in *R. bechuanae*.

Stairway Plot 2 analysis (147,761 SNPs, with singletons) showing historical changes in effective population size (N_e) over the past 130,000 years based on N = 154 individuals (multiple samples per locality, solid line) and N = 11 individuals (one sample per locality, dashed line). The red line indicates the estimated median N_e, and the shaded red area represents the 95% confidence interval. The grey rectangle marks the period of the Last Glacial Maximum (LGM), while the yellow rectangles denote recent arid episodes in the Summer rainfall region of Southern Africa (Chevalier & Chase, 2016). Estimates were generated using a mutation rate of 5.7×10^{-9} mutations per site per year and a generation time of one year.

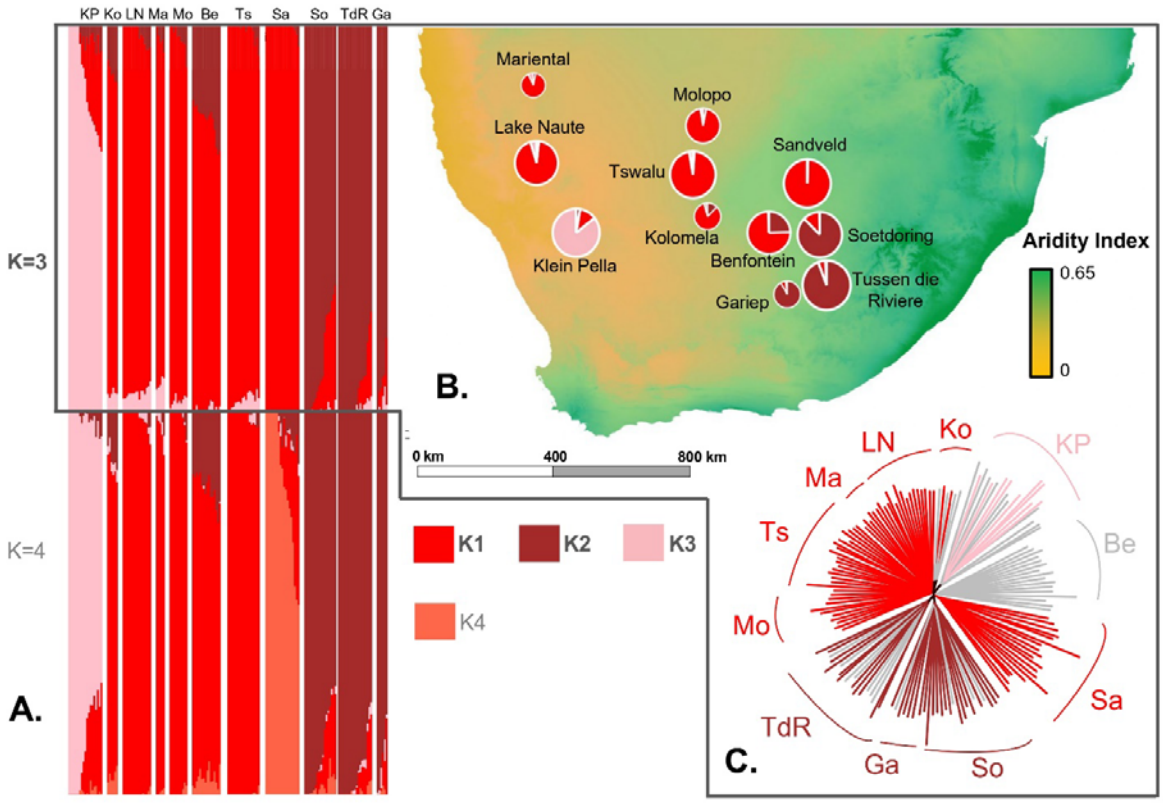


Figure 2: Population structure among *R. bechuanae* samples.

A. *sNMF* results from K=3 to K=4 for *R. bechuanae* samples in each locality. Abbreviations of sampling localities are : Klein Pella (KP); Kolomela Mine (Ko); Lake Naute (LN); Mariental (Ma); Molopo (Mo); Benfontein (Be); Tswalu (Ts); Sandveld (Sa); Soetdoring (So); TdR (Tussen die Riviere); Ga (Gariep Dam). **B.** Map showing the distributions and genetic compositions of *R. bechuanae* sampling localities. Pie diagrams indicate the frequencies of genetic clusters in each locality. Base map: World Topographic Map Esri Standard, Aridity Index layer was computed from a 0.5° global grid, using data from Version 3 of the Global Aridity Index and Potential Evapotranspiration Database (Zomer et al., 2022). **C.** Neighbour-joining tree built using *ape.* Individuals with an ancestry coefficient > 0.85 for one cluster are coloured according to the code colour of the corresponding genetic cluster used in panel A. *R. bechuanae* individuals with ancestry coefficients < 0.85 for any genetic cluster are coloured in grey. All analyses were performed using *N*=152 individuals and 44,900 SNPs.

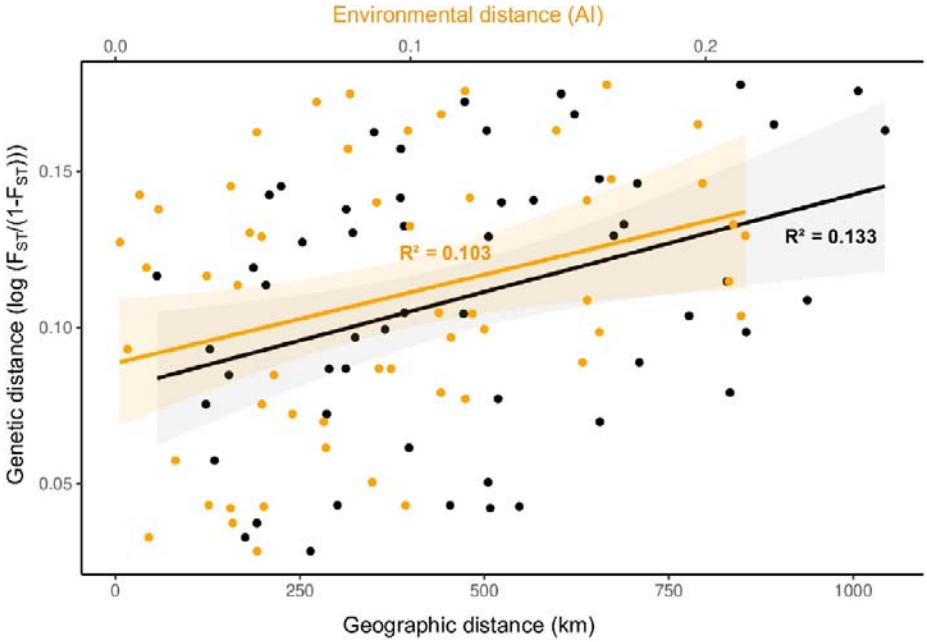


Figure 3: **Correlations between genetic distance and geographic/Aridity Index distances**.

Correlation plots and correlation coefficient R^2 estimates between genetic distance (F_{ST}) and geographic distance (in km) matrices (black), and between environmental distance and geographic distance matrices (orange). Environmental distance is represented by Euclidean distance in terms of the Aridity Index.

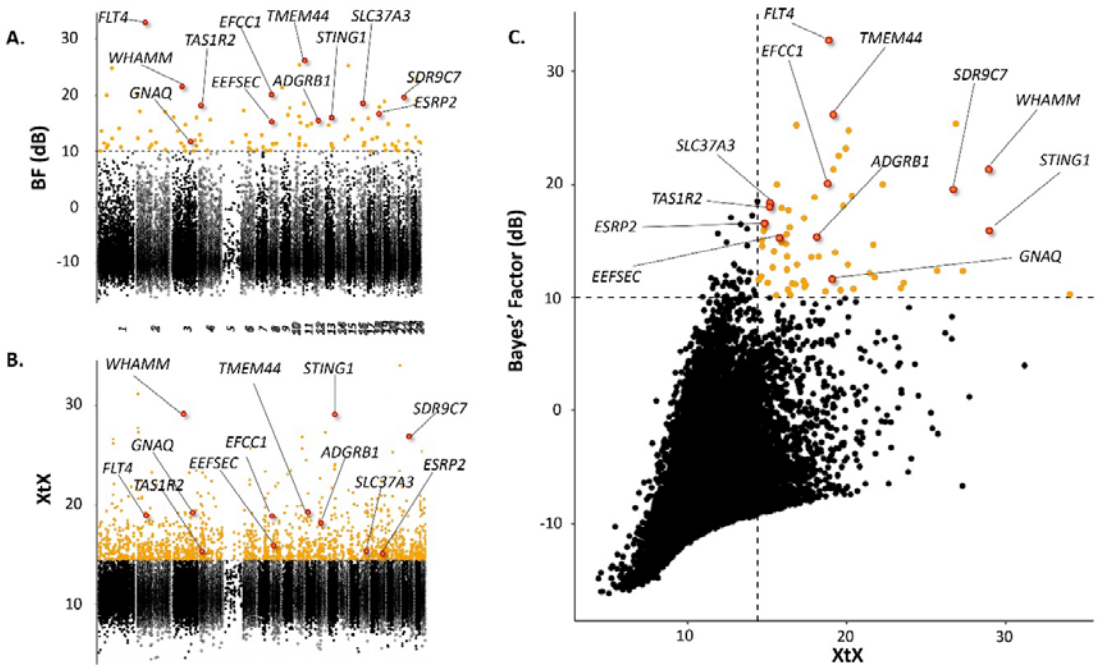


Figure 4: **Genotype-Environment Associations and genetic differentiation in *R. bechuanae***.

A. Manhattan plot of Bayes Factor (*BF*, in decibans) values from *BayPass* (standard covariate model), showing associations between 44,900 SNPs and the Aridity Index across 24 chromosome-level scaffolds (x-axis : genomic position). The dashed line marks the *BF*=10 ("strong" level of association) significance threshold. Outlier SNPs are in orange ; others alternate black/white fill colour by scaffold). **B.** Manhattan plot of genetic differentiation (*XtX* values) from *BayPass* (core model). The dashed line indicates the 5% POD significance threshold (*XtX* = 14.4). Colour scheme as in A. **C**. Joint plot of *BF* and *XtX* values. Thresholds are shown as dashed lines; SNPs significant for both metrics are in orange. In all plots, we show the names of some genes containing significantly associated and differentiated SNPs, involved in functions relevant to the response to arid conditions (for each function, the gene containing the SNP with the highest BF was highlighted; SNPs within these genes in red). All analyses were performed using *N*=218 individuals and 44,900 SNPs.

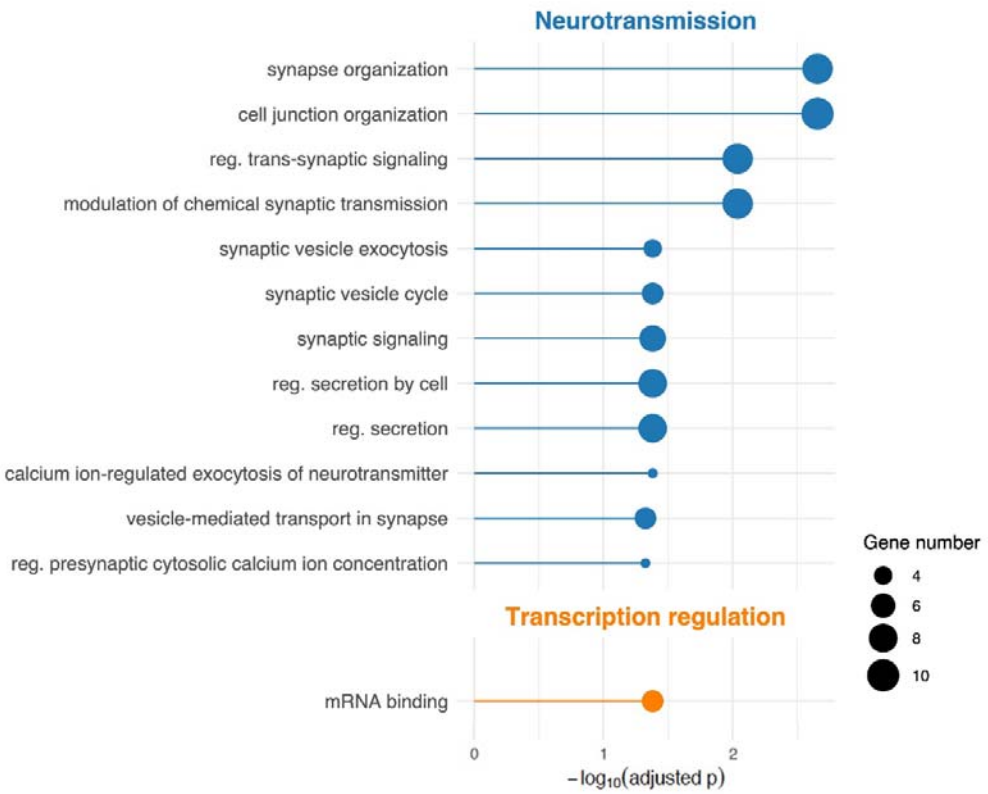
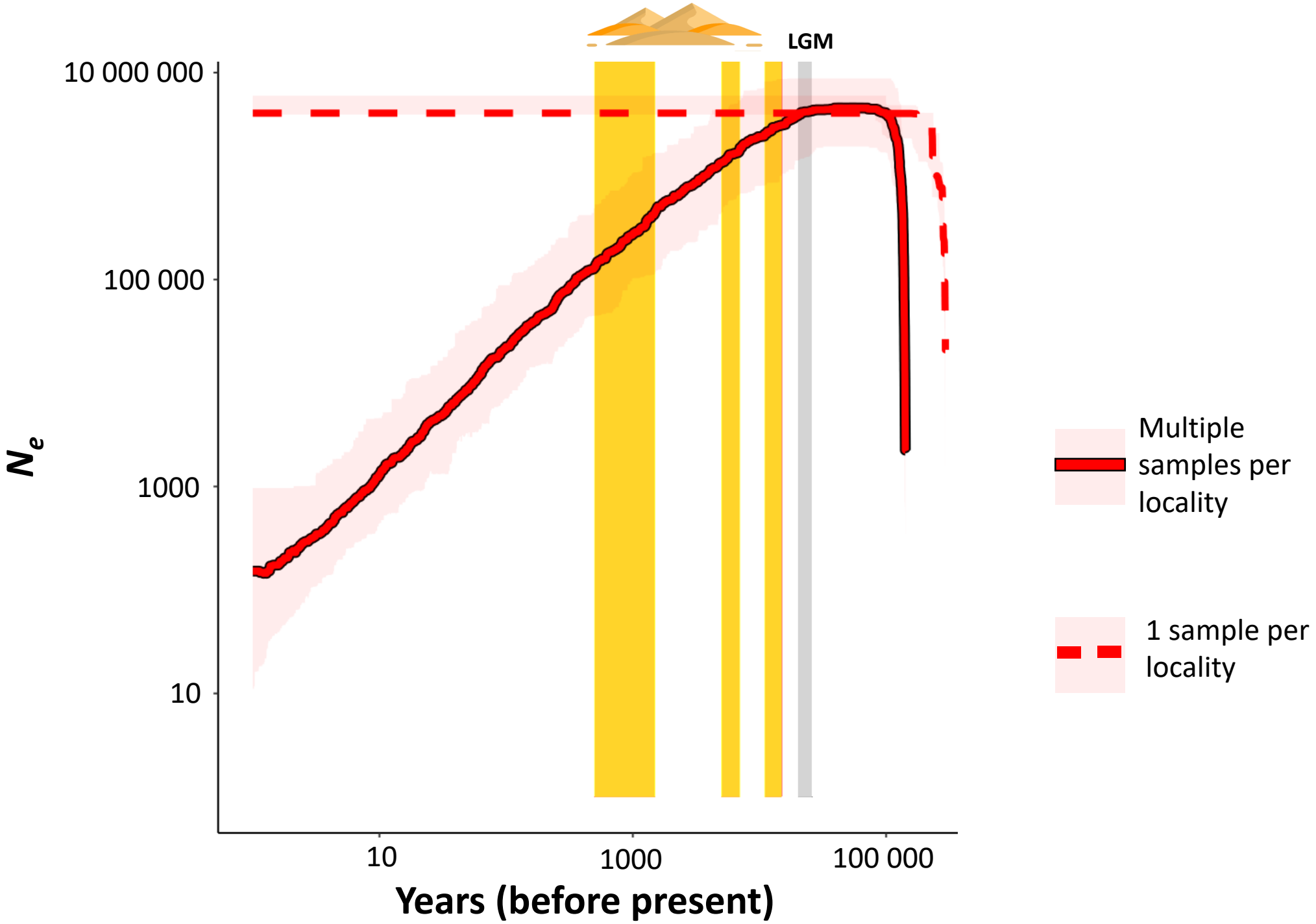
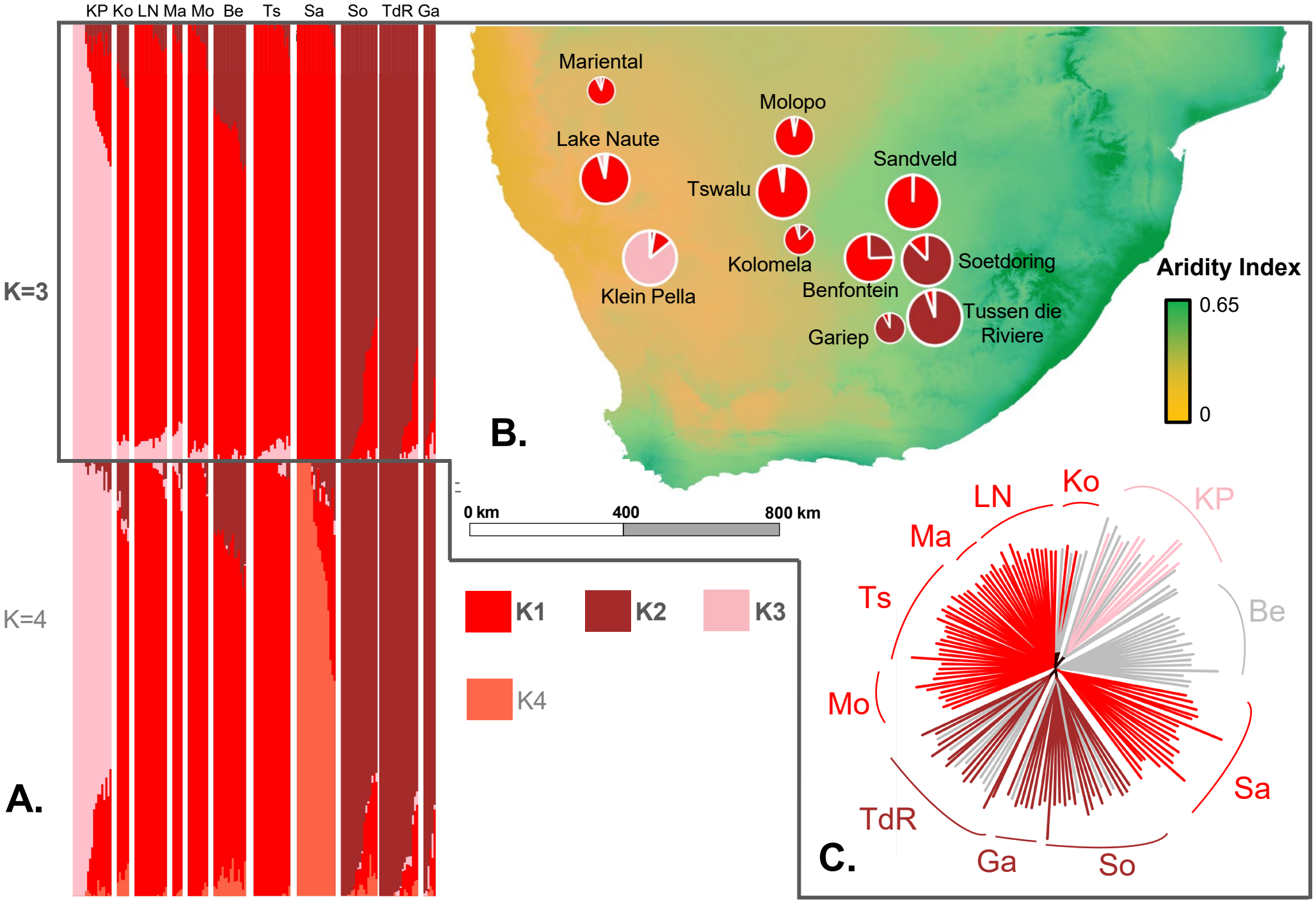
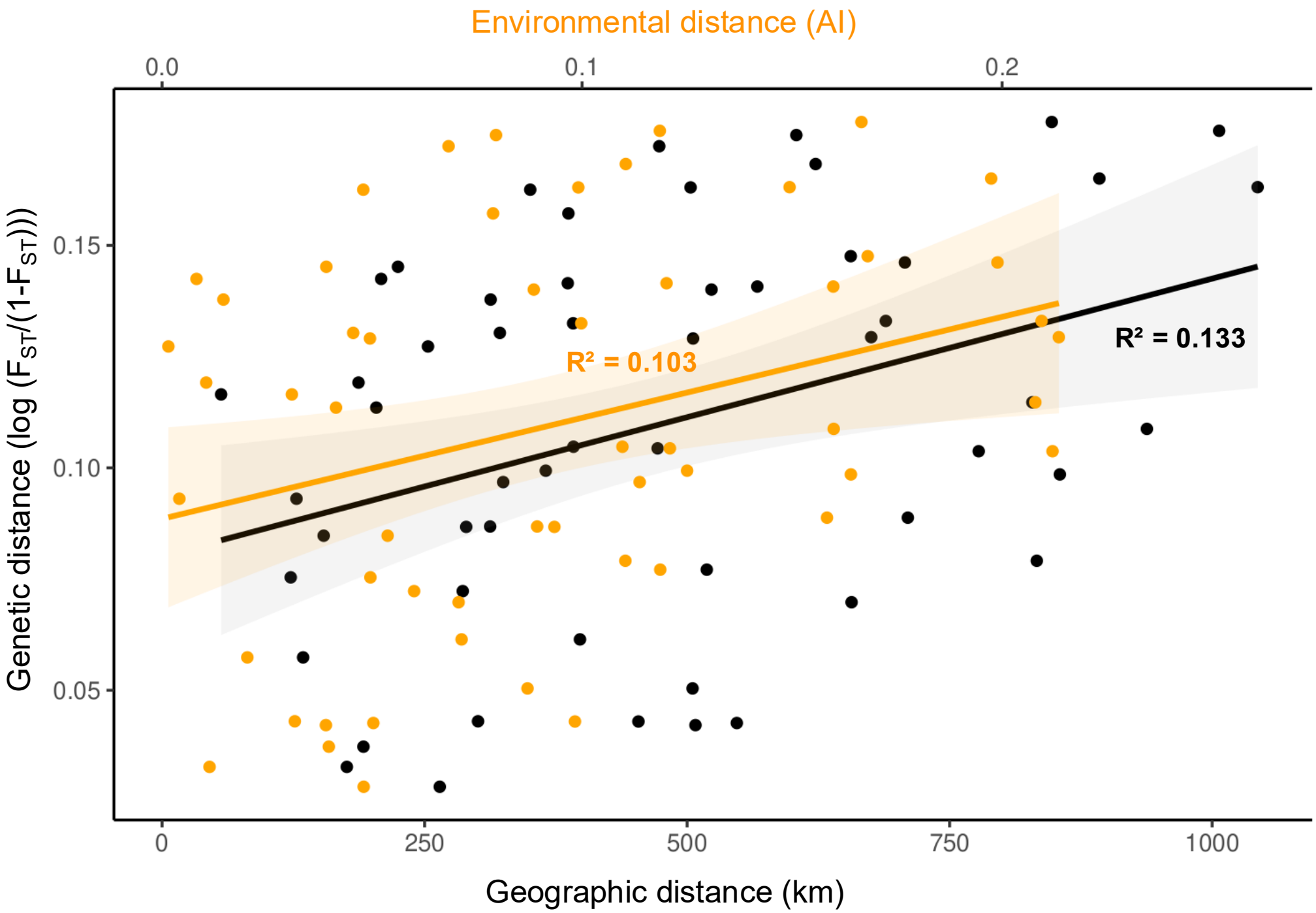


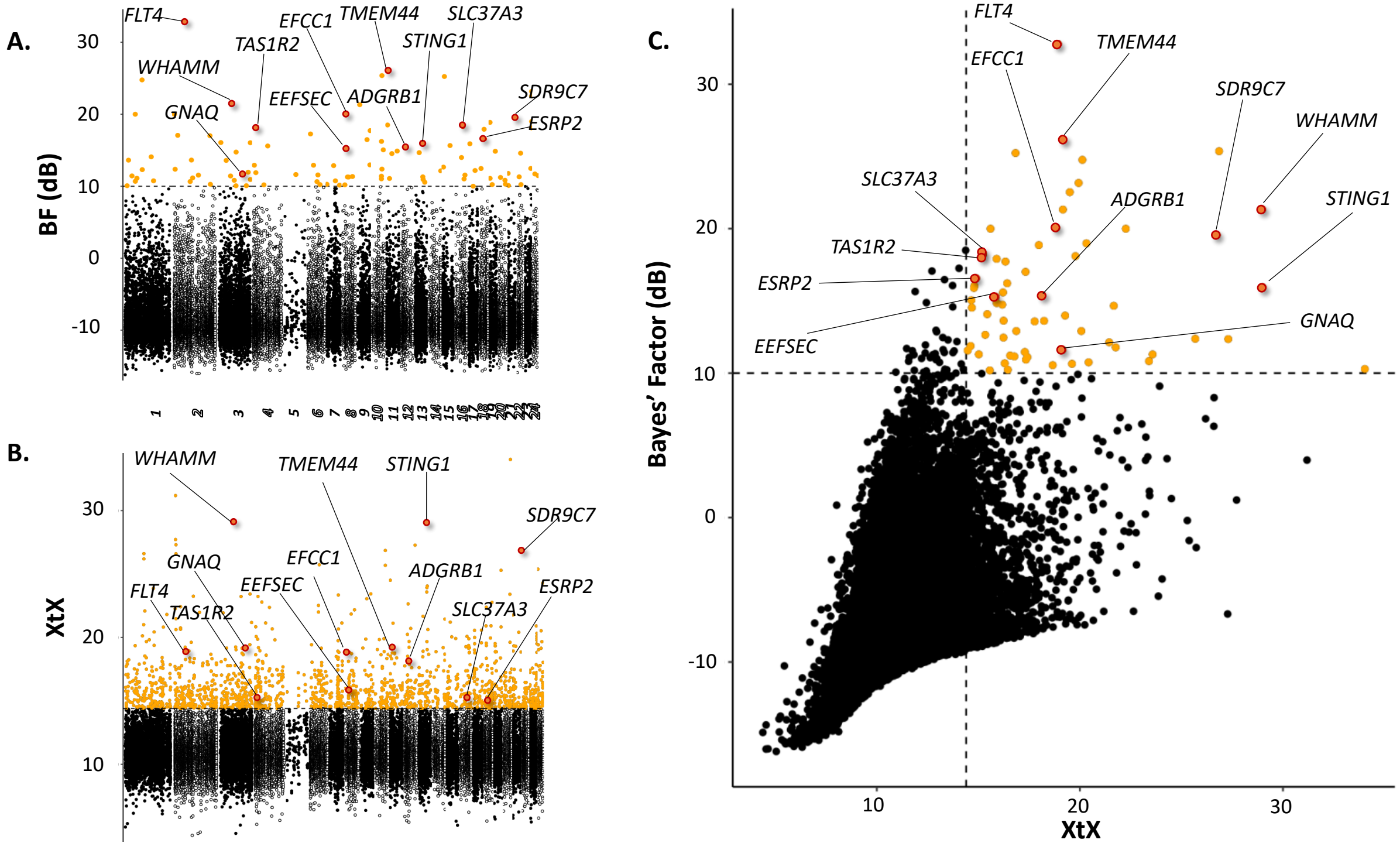
Figure 5: Functional enrichment among genes significantly associated with the Aridity Index.

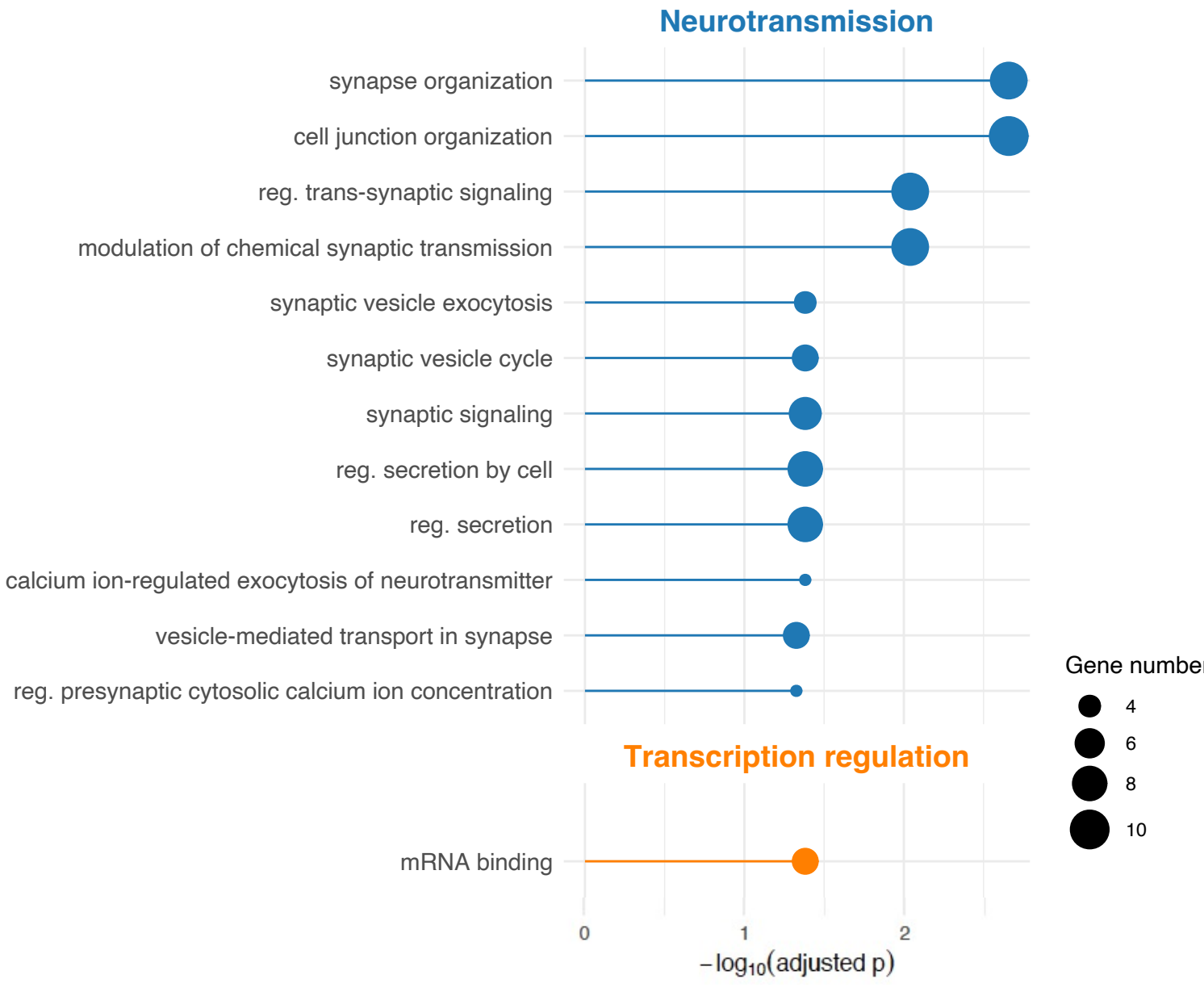
Abbreviated description of significantly enriched GO terms ($p_{adj} < 0.05$) are presented on the left, classified into more general physiological pathways, with level of significance after Benjamini-Hochberg correction reported on the x-axis. Circle size represents the number of outlier genes found to be associated with each GO term.


LGM
10 000 000
100 000
1000
10
N_e
10
1000
100 000
Years (before present)
Multiple samples per locality
1 sample per locality


KP Ko LN Ma Mo Be Ts Sa So TdR Ga
K=3
K=4
A.
Mariental
Molopo
Lake Naute
Sandveld
Tswalu
Kolomela
Soetdoring
Benfontein
Klein Pella
Tussen die Riviere
Gariep
B.
Aridity Index
0.65
0
0 km
400
800 km
K1
K2
K3
K4
C.
LN
Ko
KP
Ma
Ts
Be
Mo
Sa
TdR
Ga
So


Environmental distance (AI)
0.0
0.1
0.2
Genetic distance (log (F_ST/(1-F_ST)))
0.15
0.10
0.05
R² = 0.103
R² = 0.133
0
250
500
750
1000
Geographic distance (km)


A.
BF (dB)
FLT4
EFCC1
TMEM44
SLC37A3
TAS1R2
STING1
WHAMM
EEFSEC
ADGRB1
SDR9C7
GNAQ
ESRP2
B.
XtX
WHAMM
TMEM44
STING1
SDR9C7
GNAQ
EFCC1
ADGRB1
FLT4
EEFSEC
ESRP2
TAS1R2
SLC37A3
C.
Bayes' Factor (dB)
XtX
FLT4
TMEM44
EFCC1
SDR9C7
WHAMM
SLC37A3
ADGRB1
STING1
TAS1R2
ESRP2
GNAQ
EEFSEC


Neurotransmission
synapse organization
cell junction organization
reg. trans-synaptic signaling
modulation of chemical synaptic transmission
synaptic vesicle exocytosis
synaptic vesicle cycle
synaptic signaling
reg. secretion by cell
reg. secretion
calcium ion-regulated exocytosis of neurotransmitter
vesicle-mediated transport in synapse
reg. presynaptic cytosolic calcium ion concentration
Transcription regulation
mRNA binding
0
1
2
−log10(adjusted p)
Gene number
4
6
8
10