Skull variation in populations of the Indian gerbil *Tatera indica* (Gerbillinae, Rodentia) sampled across its broad geographic range

Zainab Dashti[1], Hasan Alhaddad[1], Bader H. Alhajeri[1]

1 *Department of Biological Sciences, Kuwait University, Safat, 13060, Kuwait*

http://zoobank.org/5B7EB0CF-50AD-47F3-9727-57FEF48674F4

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

Citation: Dashti Z, Alhaddad H, Alhajeri BH (2022) Skull variation in populations of the Indian gerbil *Tatera indica* (Gerbillinae, Rodentia) sampled across its broad geographic range. Vertebrate Zoology 72 1077–1098. https://doi.org/10.3897/vz.72.e90474

Abstract

Populations of broadly distributed species often exhibit geographic structuring, which is sometimes reflected in phenotype. The monotypic Indian gerbil (*Tatera indica*) is an example of a widely distributed species, with its range encompassing much of Asia. This study aims to determine if *T. indica* populations exhibit marked variation in skull morphology—this structure is particularly adaptable and thus could be amenable to show such variation. Furthermore, the potential drivers of skull variation are examined, including the role of climate and geography. To achieve these goals, 21 linear measurements were measured on the skulls of 509 specimens, coming from 111 different localities, across this species wide range. The specimens were then assigned into one of four broad geographic groups (≈ populations) based on their geographic proximity, and the overall and the pairwise differences in the 21 skull measurements among these groups were assessed. Specimens from Pakistan significantly differed from those belonging to the West Iran, East Iran, and India populations, which in turn did not significantly differ from each other. Pairwise bioclimatic and geographic distances between the localities explained a significant, yet small amount of variation in the measurements. Thus, while the Pakistani *T. indica* population was distinct in skull measurements, both climatic and non-climatic spatial factors seem not to account largely for its distinctiveness (from the other populations).

Keywords

Adaptation, climate, geographic variation, morphometrics, populations, skull morphology

Introduction

The Indian gerbil *Tatera indica* (Hardwicke, 1807) is a large rat-like rodent, with a heavily constructed body (Kryštufek and Vohralík 2009). This species is monotypic, and belongs to the subfamily Gerbillinae Gray, 1825, which includes gerbils and jirds, with its closest relatives belonging to the genera *Desmodillus* Smith, 1834, *Gerbilliscus* Thomas, 1897, and *Gerbillurus* Shortridge, 1942 (Alhajeri et al. 2015). *T. indica*'s cranium is robust, deep, and dorsally curved, with a prominent forward projecting zygomatic plate (Kryštufek and Vohralík 2009; Vaughan et al. 2011). It has a flat, reduced braincase (Shehab et al. 2011), grooved orange upper incisors, and hypsodont lamellate molars (Aulagnier 2009). *T. indica* is mixivorous, with a diet consisting of plant material (i.e., leaves,

grass, and grains) and occasionally insects and small mammals (Shenbrot and Krasnov 2001; Vaughan et al. 2011; Akhtar et al. 2017).

Indian gerbil populations are found throughout the Middle East and Central Asia (Bates 1988; Yigit et al. 2001; Kryštufek et al. 2017). In a large part of this species' range, it inhabits deserts and other arid regions, but it also occurs in forests, shrublands, and grasslands, with its distribution being restricted by cold temperature (Aulagnier 2009; Kryštufek et al. 2017; Kryštufek and Vohralík 2009; Vaughan et al. 2011). Despite occurring in arid environments, *T. indica* does not carry any extreme adaptations to arid conditions as do other gerbils (Aulagnier 2009). Rather, it has typical mammalian adaptations to aridity, including the ability to reduce water loss in dry periods by increasing urine concentration (Goyal et al. 1988), and being nocturnal, only emerging from their burrows to forage during the relatively cool night (Goyal and Ghosh 1993).

Within a single species, geographically separated populations can show differences in morphology, including in the skull. When these conspecific populations occupy widely different habitats, then selective pressures could potentially direct this differentiation (i.e., adaptive), otherwise, morphological differentiation could be neutral (i.e., when spatially separated populations occupy similar habitats). The effect of spatial separation and climate on skull differentiation is a widely studied topic in rodents, both within populations (Monteiro et al. 2003; Tabatabaei Yazdi and Adriaens 2011; Cordero and Epps 2012; Renvoisé et al. 2012; Tabatabaei Yazdi et al. 2012; Nanova 2014; Quintela et al. 2016; Alhajeri 2019; Alhajeri 2021a) and between species (Renaud and Millien 2001; Croft et al. 2011; Martínez and Di Cola 2011; Alhajeri 2018; Alhajeri 2021c; Tabatabaei Yazdi and Alhajeri 2018; Yağcı and Gurbanov 2019; Álvarez et al. 2021).

While skull differentiation has been examined in *T. indica*, most studies so far have been geographically restricted to a single country such as Syria (e.g., Shehab et al. 2011) and Turkey (e.g., Yigit et al. 2001), but mostly within Iran (e.g., Mirshamsi et al. 2007; Ashrafzadeh et al. 2012; Khajeh et al. 2019). Most such studies report only minor differences (e.g., Mirshamsi et al. 2007). Thus, the main aim of the present study is to expand the scope of the comparisons of skull differences between *T. indica* populations from across its broad range, to see if these widely distributed populations show marked differentiation. Furthermore, we also examine the possible influence of climatic and non-climatic spatial factors in driving skull variation. More specifically, skull variation could be influenced by climatic spatial factors as a result of local adaptation to specific climatic conditions. This could happen directly, such as how body size is influenced by temperature in accordance with Bergmann's rule (see Alhajeri and Steppan 2016) or primary productivity (i.e., precipitation) in accordance with the resource availability hypothesis (see Alhajeri et al. 2020a). Alternatively, this could happen indirectly, by climate determining the type of plant resources that would be available in each habitat, and therefore influencing the availability of food items (i.e., the diet) available to each *T. indica* population. Of

course, these effects need not be on body size alone, and could influence overall cranial morphology (particularly shape) as well, as have been shown in other murids (see Martin et al. 2016). On the other hand, non-climatic spatial factors could drive skull differentiation by reducing gene flow between spatially distant populations, as predicted by the isolation-by-distance model (Wright 1943). This would lead to spatially structured clinal variation in skull morphology as has been shown to occur in other gerbils (e.g., Alhajeri 2019).

Materials and methods

Geographic sampling

The skulls of 509 *T. indica* specimens, sampled from a range extending from Kuwait to Nepal (Fig. 1), were photographed from the following museums: the United States National Museum of Natural History (USNM) in Washington D.C., the American Museum of Natural History (AMNH) in New York, the Field Museum of Natural History (FMNH) in Chicago, the Museum of Vertebrate Zoology (MVZ) at the University of California in Berkeley, and the Florida Museum of Natural History (UF) at the University of Florida. The USNM photographs were taken by ZD (n = 393), and those from AMNH, FMNH, MVZ, and UF by BHA (n = 116), both following the same standardized photography protocol outlined in previous studies (Alhajeri 2021b; Alhajeri 2022a; Alhajeri 2022b). Only specimens with completely erupted mandibular molars were sampled in this study.

Photographs were captured with a Nikon D3200 DSLR camera using a Micro-Nikkor lens (Nikon, Tokyo, Japan) at a 6016×4000-pixel resolution. The dorsal, ventral, and lateral views of the crania were photographed, along with the occlusal and lateral views of the mandibles. A scale bar was included in all photographs to convert pixels to millimeters. The left side of each skull was photographed. In a few specimens, this side was severely broken, and as such the right side was photographed and the photograph was reflected before taking the measurements. A total of 21 linear measurements were extracted from each photograph (Fig. 2; Table S1) using ImageJ 1.52 (Abramoff et al. 2004). Briefly, these are: upper incisors width (IW), molar row length (ML), M1 breadth (MB), basicranial width (BW), basioccipital length (BOL), bulla length (BL), bulla width (WOB), condyle breadth (CB), incisor depth (ID), rostrum depth (DR), bulla height (HB), cranial length (DL), cranial width (DW), interorbital width (WOI), nasals width (WN), nasals length (LN), lower incisors width (WI), diastema length (MDL), m1 height (FH), angular length (ALM), and condyloid length (CLM) (see Table S1 for a more detailed description of each measurement). These measurements are standard for mammalian skulls and are used in many similar studies on rodents (e.g., Alhajeri and Steppan 2018). All measurements were collected by ZD.
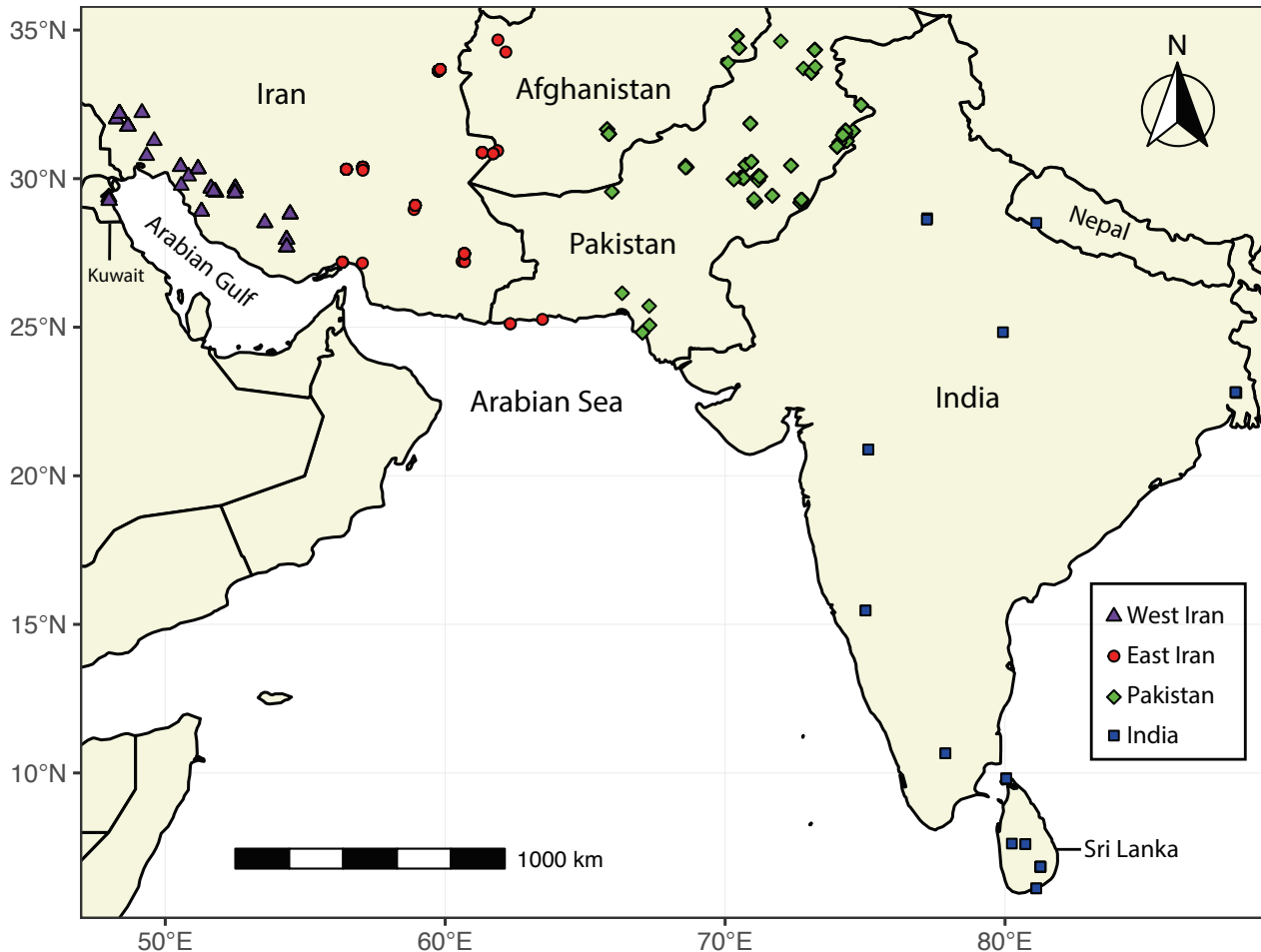
Figure 1. Map of specimen sampling localities. Sampling localities are divided into four geographic groups based on their proximity (see legend). The Y and X axis are the latitude and longitude, respectively. A geographic WGS84 projection is used. For more details, see the "Materials and methods" and Appendix 1. The full range of the species according to the International Union for Conservation of Nature (IUCN) Red List (Kryštufek et al. 2017) appears in Fig. S1.

Each measurement was taken twice in order to allow for the calculation of the repeatability of the measurements using the intraclass correlation coefficient (ICC) (Wolak et al. 2012). The ICC values based on the two trials were within an acceptable range (see results), and therefore the average of two trials was used in all subsequent analyses (Data S1). The ICC was calculated using an Analysis of Variance (ANOVA) based on the ICCest() function in the *ICC* package (Wolak et al. 2012) in R (R Core Team 2019). For all analyses in this study, the level of significance was set at $\alpha = 0.050$. Each specimen was assigned a latitude and a longitude value based on their respective online museum database.

In instances where geographic coordinates were not available in museum databases, specimens were assigned this information based on the most precise locality information available from these databases and/or specimen vial labels. This was done by converting geographic information into coordinates using the distance-measuring tool in Google Maps (Google 2019). The provided locality information is not always precise, and thus geographic coordinates are approximations and would be most useful for detecting broad-scale patterns, which are the

focus of this study. For specimens with no precise locality information (i.e., only a town name, a province, or a country was available), coordinates were based on the approximate geographic centroid (e.g., center of mass of a county) or the midpoint (e.g., midpoint of a bridge) of the most precise locality information available—this was estimated based on the geographic divisions of Google Maps. Specimens with no locality information or those for which coordinates could not be determined using the provided locality information were omitted (i.e., not part of the 509 specimens sampled in this study).

Latitude and longitude values of the localities were obtained in decimal degrees to the nearest four decimal places (precision ~11 meters at the equator). A geographic coordinate system with the 1984 World Geodetic System (WGS84) datum (EPSG:4326) was assigned to the latitude and longitude values for subsequent analyses (Wasser 2017).

The resulting geographic coordinates were used in order to divide the 111 localities into four geographic groups based on their geographic proximity—this was mostly geographic breaks along longitude, which is more variable than the latitude in the geographic distribution of
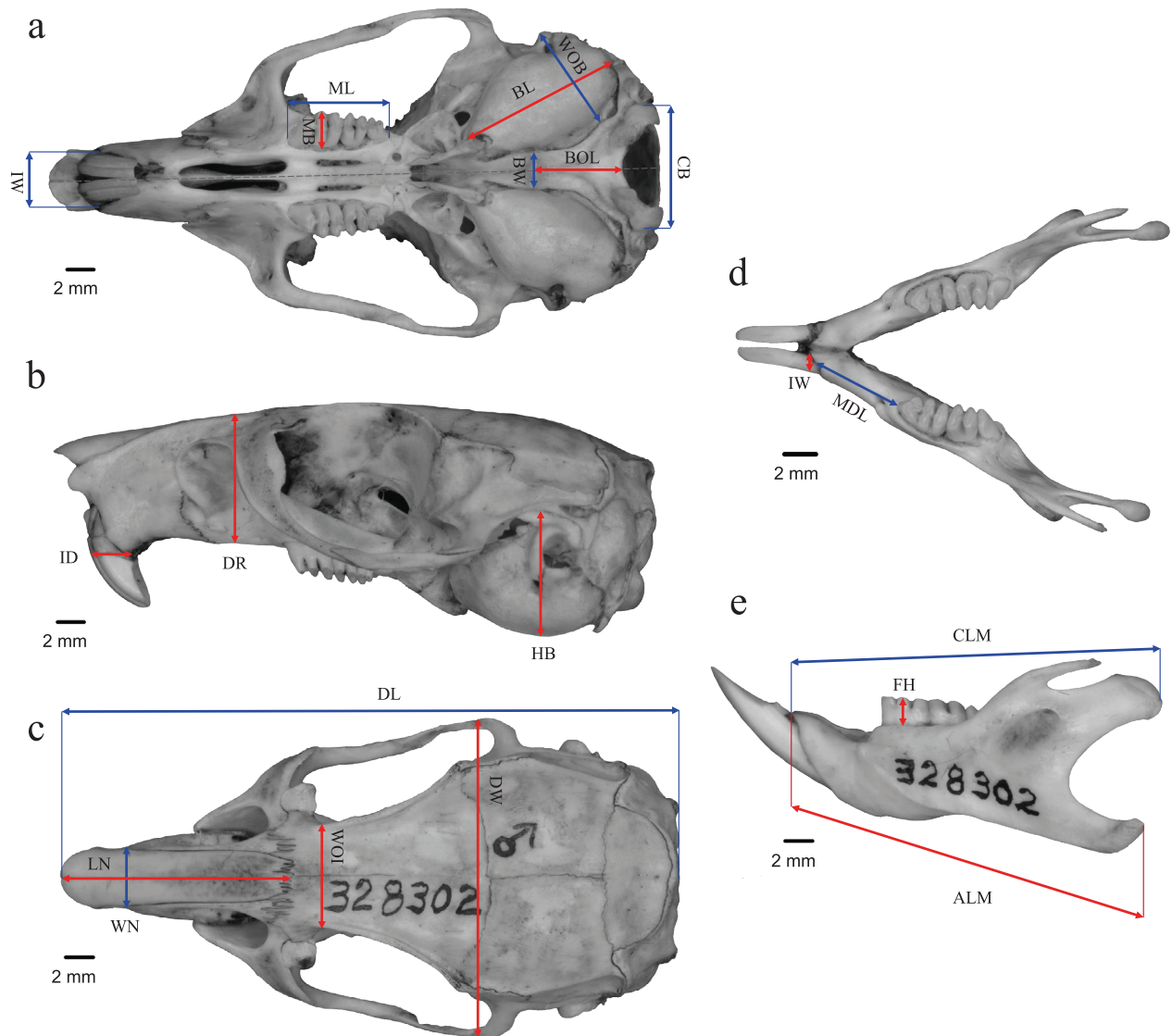
Figure 2. Extracted linear measurements. The 21 measurements collected in this study, shown on the (a) ventral cranium, (b) lateral cranium, (c) dorsal cranium, (d) occlusal mandible, and (e) lateral mandible, of USNM 328302 from Esfedan, Iran. Acronyms used in this figure match the measurement descriptions in Table S1. Scale bars are also indicated for each view. Colors in this figure are irrelevant and used to more easily distinguish the measurements.

this species (Fig. 1; Data S1; Appendix 1). The geographic groups were, from west to east, "West Iran" (mostly specimens along or west of the Zagros Mountains, including specimens from Kuwait), "East Iran" (mostly specimens on the eastern part of Iran, including some on the western borders of each of Afghanistan and Pakistan), "Pakistan" (mostly specimens in Pakistan, along with a few specimens in the eastern border of Afghanistan), and "India" (specimens from India, Sri Lanka, and Nepal) (Fig. 1). The full range of *T. indica* according to the International Union for Conservation of Nature (IUCN) Red List (Kryštufek et al. 2017) appears in Fig. S1.

The sample includes 255 females, 247 males, and 7 specimens of unknown sex (Data S1). Most specimens come from Iran (233) and Pakistan (198); the rest of the countries were much less sampled: Afghanistan had 43 specimens, Sri Lanka had 15 specimens, India had 12, Kuwait had seven, and Nepal had one. (Data S1). The

most sampled geographic group was "Pakistan" (235), followed by "West Iran" (126), "East Iran" (120), and India (28), respectively (Data S1). Out of the 111 localities, the most sampled was "Changa Manga Forest Park" (31), followed by "Andimeshk, 30 km S" (21), and "Kandahar International Airport" (21) (Data S1; Appendix 1).

Morphometric analysis

The first set of analyses were conducted to determine if *T. indica* populations (as represented by the four geographic groups; see above) show marked skull differentiation. Firstly, morphometric variation was summarized in the form of descriptive statistics (mean, standard deviation, minimum, maximum, and number of specimens) both overall and divided by geographic groups. Descriptive statistics were calculated using the stat.desc() function in

the *pastecs* library (Grosjean and Ibanez 2018) in R. Differences in each of the 21 morphometric measurements among the geographic groups were visualized using multipaneled boxplots using the ggplot() function in the R package *ggplot2* (Wickham 2016).

Subsequently, all 21 measurements were natural log-transformed to reduce the effect of outliers and to linearize their relationships for subsequent statistical analyses. The effect of group membership, sex, and their interaction on overall morphology was then examined with a permutational MANOVA (PERMANOVA) (Anderson 2001). PERMANOVA is a nonparametric test based on Euclidean distance matrices (distinguishes differences in centroids) that does not assume normality. The p-values for each model term were determined using 1000 permutations, with the F ratio calculated using a pseudo-F statistic (Anderson et al. 2008). Because the groups are unbalanced, a Type-II sums of squares (SS) PERMANOVA was conducted with the adonis.II() function in the R library RVAideMemoire (Hervé 2019). Since PERMANOVA detected no significant sexual dimorphism, and no interaction between sex and the geographic group factor was detected (see Results), males and females were combined in all subsequent analyses. All specimens with missing measurement data and/or are of unknown sex were removed prior to the PERMANOVA.

While PERMANOVA is robust to unbalanced designs (Anderson and Walsh 2013), significant p-values could be driven by unequal multivariate dispersions among groups. As such, the variance to each group's median was calculated using PERMDISP (a test of homogeneity of multivariate variances), which was carried out using the betadisper() function in the R package *vegan* (Oksanen et al. 2019). The significance of the differences in group dispersions was tested using an unrestricted permutation test using the permutest() function in *vegan*. The PERMDISP results indicate heterogeneity in dispersion of geographic groups but not sex (see Results), which indicates that the results of the PERMANOVA are not conservative for the geographic group factor. Thus, for the geographic group factor, the PERMANOVA results were verified both visually based on principal component analysis (PCA) scatterplots (see below) and via post-hoc pairwise PERMANOVA comparisons; the latter also serves to determine which groups are significantly different from each other. Pairwise PERMANOVAs were carried out using the adonis.pair() function in the *EcolUtils* R library (Salazar 2018), based on Euclidean distances with p-values computed using 1000 permutations. In these pairwise analyses, Holm's (1979) correction was used to correct for multiple comparisons, which maintains a table wide error rate of 5%. For all PERMANOVAs, the amount of the total variation explained by each model term was computed as the ratio of the SS of each factor to the total SS of the model (i.e., R^2).

To determine which measurements differed the most between the geographic groups, univariate tests on individual logged measurements were conducted. Both parametric ANOVA assumptions were not satisfied for most measurements. Normality of model residuals was tested using the Shapiro-Wilk test as implemented in the shapiro.test() function in the R *base* package and homogeneity of variance across groups were tested using Levene's test as implemented in the leveneTest() function in the in the *car* R package (Fox and Weisberg 2011). Since the parametric ANOVA assumptions were not satisfied, a nonparametric Kruskal-Wallis test was used. This test is rank-based and does not make any assumptions about the distribution of the data (Kruskal and Wallis 1952). Because 21 related Kruskal-Wallis tests were conducted (for each measurement), in addition to reporting the raw p values, Holm-corrected p-values were also computed and used to determine significance. The Kruskal-Wallis test was run using the kruskal.test() function and Holm-correction using the p.adjust() function, both in the R *base* package. The effect size for each Kruskal-Wallis test was computed as eta-squared (η^2) based on the H-statistic, carried out using the kruskal_effsize() function in the *rstatix* R package (Kassambara 2019). Finally, post-hoc Scheffé tests were used to detect the pairs of geographic groups which were significantly different in each of the logged measurements (Scheffé 1999). This technique adjusts the level of significance for numerous comparisons of unequal sample sizes. The pairwise comparison effect size was based on Hedges' g (Hedges 1981), which corrects for small sample size. Scheffé tests were conducted using the ScheffeTest() function in the R package *DescTools* (Signorell 2018), while the Hedges' g was computed using the hedg_g() function in the R package *esvis* (Anderson 2018). All specimens with missing measurement data were removed prior to the relevant Kruskal-Wallis or ScheffeTest and their associated tests of effect size.

The affinities of the four geographic groups based on overall morphology were visualized using PCA scatterplots of the first three principal component axes, with specimens color-coded based on group membership. For these scatterplots, the PCA was applied to the 21 logged measurements using the pca() function in the *pcaMethods* R package (Stacklies et al. 2007). The PCA was conducted on the standardized data (i.e., scaled to unit variance and mean centered) using the "SVDimpute" algorithm (Troyanskaya et al. 2001), which allows for the estimation of missing data.

Morphometric data preprocessing

The above analyses showed that the variation in the 21 skull measurements is geographically structured (see Results). Furthermore, similar results were found for most of the univariate tests of the 21 examined measurements, and their loadings onto the first principal component seem to suggest that the reason for this concordance is their potentially high correlation with each other, by virtue of being associated with overall size (see Results). The subsequent analyses test the possible influence of climatic and non-climatic spatial factors in driving this geographic skull variation. To achieve this goal, the morphometric data need to

first be processed to be suitable for redundancy analysis (RDA) (as described in the section: Association between morphometric variation with climate and other spatial factors). This data preprocessing, along with its rationale is described in detail in Table S2. Briefly, the *mice* package (Van Buuren and Groothuis-Oudshoorn 2011) was used to impute missing data, and the RDA was conducted both on the logged variables and on size-corrected log variables derived by "shearing" (see Humphries et al. 1981).

Extraction of climate variables

To determine the effect of climatic spatial factors (i.e., climatic adaptation) in driving geographic skull variation, a proxy for overall climate need to be used for the RDA (see section: Association between morphometric variation with climate and other spatial factors). Climate was based on 19 bioclimatic variables downloaded as raster files from WorldClim (version 2) (Fick and Hijmans 2017), at a spatial resolution of 2.5 min (~4.5 km at the equator) (Table S3; Data S1). The extracted raster files were processed using the R package *raster* (Hijmans 2019). The CRS of the bioclimatic raster files matched those for the locality information (WGS84, unprojected latitude/longitude), and so no (re)projections were necessary. Since the raster cells are already ~4.5 km each, the raw values of the raster (climate) cells where the geographic coordinates of the localities fall in were obtained and used in subsequent analyses. The mean values based on neighboring cells were not used (i.e., a circular buffer of a given radius) since the resolution of the raster cells was already sufficiently coarse, and the home range of gerbils is often lower than these grid cells. As such, uncertainty in the locality positions (see above) was considered in the computation of these climatic variables by using these relatively coarse grid cells.

The climatic variables were standardized by transforming them to a mean of zero and unit variance using the scale() function in the R *base* package to account for different units. Preliminary analyses indicated high multicollinearity among the standardized climatic variables. Variance inflation factors (VIFs) were computed using the R function vif() from the *usdm* package (Naimi et al. 2014). VIF values ranged from ~3 (BIO15) to ~ 3095503 (BIO7) (see Results). Such a high level of correlation among the variables could be an issue if they are used as explanatory variables for regression models (i.e., RDA) as they could bias the results. Furthermore, high VIFs may indicate variables that are functionally related to each other, as is expected of the bioclimatic variables, which are all based on the same precipitation or temperature dataset. As such, the function vifstep() from the *usdm* package was used to reduce the number of variables via a stepwise procedure (Naimi et al. 2014). In this process, the function calculates the VIF of all the variables, then excludes the variable with the highest VIF score, and then recalculates the VIF scores for the remaining variables, and it repeats the process until a specified VIF thresh-

old of < 20 is achieved for the remaining variables. This process led to the retention of BIO2, BIO3, BIO5, BIO8, BIO9, BIO13, BIO14, BIO15, BIO18, and BIO19, with VIFs that now range from ~2 (BIO15) to ~13 (BIO8). Only these variables were used in the subsequent RDA analyses (see section: Association between morphometric variation with climate and other spatial factors).

Extraction of spatial variables

To determine the effect of non-climatic spatial factors (i.e., geographic distance) in driving geographic skull variation, an index of geographic structure is needed for the RDA (see section: Association between morphometric variation with climate and other spatial factors). This was based on a set of geographic variables which were calculated as follows.

First, pairwise geographic distances were computed between the latitude and longitude values of the localities. Distance calculations between localities are inaccurate in unprojected data if the Euclidean distance formula is used, especially at increasingly large spatial scales. As an alternative to projecting the geographic coordinates to an equidistant projection, the rdist.earth() R function in the *fields* package (Nychka et al. 2017) was used to calculate the great-circle distance between the localities using the haversine formula (Inman 1835). This formula considers the Earth's curvature and allows for a more accurate calculation of distances on unprojected (WGS84) geographic coordinate data at large spatial scales.

The spatial relationships among the localities were then summarized using distance-based Moran's Eigenvector Maps (dbMEMs; Borcard and Legendre 2002; Legendre and Legendre 2012) based on the aforementioned spatial distance matrix. These dbMEMs are orthogonal eigenvectors; MEMs with high Moran's *I* values describe broad spatial structures involving the whole dataset, while those with lower Moran's *I* values describe more fine spatial structures involving limited numbers of localities (Dray et al. 2012). These dbMEM eigenvectors were computed from the geographic distance matrix using the R function dbmem() in the *adespatial* package (Dray et al. 2019). Only MEMs corresponding to positive autocorrelation were retained for subsequent analyses (see below). The resulting positive MEMs are more powerful descriptors of spatial relationships between the localities than the raw longitude and latitude values. These MEM variables were both used as predictors to study spatial patterns of morphometric variation as well as covariables (i.e., spatial filters) to correct for spatial autocorrelation (see below; henceforth, referred to as "spatial variables"). By design, all MEM variables are orthogonal to each other (for all variables, VIF = 1), and so no process of variable reduction was conducted on them at this stage (i.e., as was done for the climatic variables). As such, all these variables were used in the subsequent RDA analyses (see section: Association between morphometric variation with climate and other spatial factors).

Association between morphometric variation with climate and other spatial factors

The processed morphometric data (see section: Morphometric data preprocessing), along with the climatic (see section: Extraction of climate variables) and spatial variables (see section: Extraction of spatial variables) were used to quantify the influence of climatic and non-climatic spatial factors (i.e., geographic distance and/or climatic adaptation) in driving this geographic skull variation. This was done by variation partitioning (Borcard et al. 1992) by means of partial RDA, which determines the amount of morphological variation explained by the various unique and combined fractions of the (1) climatic variables and the (2) spatial variables.

Prior to running this analysis, the forward.sel() function in *adespatial* was used to perform a forward variable selection procedure to identify the most parsimonious combinations of explanatory variables in each of the climatic and spatial datasets (the variables that are the strongest predictors of the response variables) (and to prevent model overfitting) by permutation of residuals under the reduced model (1000 permutations) (Dray et al. 2019). The forward selection procedure was conducted separately for each set of explanatory variables after confirming that the global test that includes all variables was significant for each dataset. A double-stopping criterion was used to select the most parsimonious model for each dataset, where the selection procedure stopped either when the significance alpha level of 0.05 was reached (i.e., the candidate variable is non-significant), or when the adjusted coefficient of multiple determination (R^2_{adj}) value of the reduced model reached that of the global model that contains all the potential explanatory variables (i.e., the candidate variable brings the total model's R^2_{adj} value over that of the global model) (Blanchet et al. 2008).

The reduced sets of climatic and spatial variables were then subjected to variation partitioning by means of partial RDA in order to determine the unique and combined effects of both the spatial and the climatic variables on skull morphology. The varpart() function in *vegan* was used to partition variation, with a model that uses a matrix of the logged skull measurements (including the imputed data) as the response variables and the climatic and the spatial variables as two separate sets of explanatory variables. As such, this function partitions the variation in the morphological measurements into components accounted for by either the spatial or the climatic variables, and their combination—this partitioning is based on partial RDA, with the R^2_{adj} from the RDA used to assess the partitions explained by each subset of variables and their combinations (Liu 1997). The significance of all the testable fractions in the variance partitioning procedure was tested by the corresponding RDA or partial RDA models—the p-values were calculated via permutation tests with 1000 permutations. The response of the morphometric variation to climate variables were interpreted using RDA

plots including climate data as explanatory variables and spatial variables as covariables.

The exact same analyses, starting with forward selection, were then repeated on the size-corrected morphometric variables, to determine the effect of size in driving the pattern detected in the previous analyses.

Results

Morphometric variation among geographic groups

The two trials of the morphometric measurements were highly repeatable, all of which had an ICC score > 0.71. This is greater than the average repeatability of other morphological studies (ICC = 0.65; Wolak et al. 2012). The descriptive statistics of the measurements (overall and separated by geographic groups) are shown in Table 1, which are in turn depicted in the form of boxplots in Fig. 3.

According to the PERMANOVA, a significant effect on the ln-transformed skull measurements was found for the geographic group factor ($p < 0.050$), but not for the sex factor ($p = 0.166$), nor the interaction of geographic group with sex ($p = 0.652$; Table 2). The insignificant interaction effect indicated that the effect of geographic group on the skull measurements does not depend on sex. A moderate effect was detected for the geographic group factor on the skull measurements ($R^2 = 0.106$; Table 2), indicating that 10.6% of the variation in the skull measurements can be attributed to geographic group membership. The post-hoc pairwise PERMANOVAs for the geographic groups indicated a significant difference in the skull measurements between Pakistan and each of: West Iran ($R^2 = 0.073$, $p < 0.050$), East Iran ($R^2 = 0.087$, $p < 0.050$), and India ($R^2 = 0.088$, $p < 0.050$; Table 3), all of which remained significant after Holm correction. All other pairwise comparisons were not significant (Table 3).

The PERMDISP analysis indicated homogeneous multivariate dispersions in the logged skull measurements for sex ($p = 0.060$) but not for the geographic group factor ($p < 0.050$). This indicated that for the geographic group factor, the significant PERMANOVA result detected earlier is not only explained by differences in medians of the geographic groups but is also driven by differences in group multivariate dispersions. However, the pairwise PERMANOVAs and the results of the PCA analyses seem to suggest substantial differences in the logged skull measurements between some of the geographic groups (and not just in their dispersions).

The Kruskal-Wallis tests found significant differences between the geographic groups in all the measurements ($p < 0.050$ for all measurements; Table 4). The effect size was highest for the M1 breadth ($\eta^2 = 0.352$) followed by molar row length ($\eta^2 = 0.212$; Table 4). The other measure-

Table 1. Descriptive statistics of measurements (in mm). Measurements are presented as mean ± standard deviation, minimum–maximum, and number of specimens. Values are visualized in Figure 3.

Measurement	Overall	West Iran	East Iran	Pakistan	India
IW	3.85 ± 0.38 2.67–5.11, n = 503	3.93 ± 0.39 2.93–4.86, n = 124	4.00 ± 0.40 2.67–4.86, n = 119	3.72 ± 0.30 2.67–4.45, n = 232	3.94 ± 0.42 3.16–5.11, n = 28
ML	7.16 ± 0.53 5.97–9.15, n = 500	7.42 ± 0.56 6.33–8.55, n = 122	7.36 ± 0.51 6.43–9.15, n = 117	6.89 ± 0.38 5.97–8.05, n = 233	7.51 ± 0.51 6.41–8.25, n = 28
MB	2.74 ± 0.24 2.17–3.50, n = 499	2.89 ± 0.20 2.51–3.50, n = 122	2.86 ± 0.22 2.28–3.32, n = 117	2.59 ± 0.17 2.17–3.14, n = 232	2.82 ± 0.25 2.30–3.32, n = 28
BW	2.10 ± 0.36 1.00–3.83, n = 466	2.07 ± 0.37 1.35–2.92, n = 103	2.22 ± 0.37 1.22–3.27, n = 105	2.01 ± 0.30 1.00–2.79, n = 231	2.43 ± 0.43 1.92–3.83, n = 27
BOL	5.30 ± 0.59 3.34–6.94, n = 443	5.28 ± 0.78 3.34–6.69, n = 98	5.57 ± 0.61 4.00–6.94, n = 94	5.20 ± 0.46 3.52–6.26, n = 225	5.24 ± 0.35 4.38–5.96, n = 26
BL	12.00 ± 0.88 9.43–15.28, n = 443	11.99 ± 0.97 10.02–14.18, n = 97	12.49 ± 1.03 10.03–15.28, n = 92	11.83 ± 0.68 9.43–13.35, n = 228	11.81 ± 0.84 9.96–13.33, n = 26
WOB	7.65 ± 0.58 6.12–10.54, n = 441	7.86 ± 0.59 6.53–9.36, n = 96	7.87 ± 0.70 6.52–10.54, n = 92	7.54 ± 0.45 6.28–8.96, n = 227	7.15 ± 0.55 6.12–8.53, n = 26
CB	8.56 ± 0.62 6.57–10.38, n = 418	8.92 ± 0.66 6.57– 10.32, n = 90	8.72 ± 0.69 7.36–10.26, n = 86	8.30 ± 0.42 6.86–9.76, n = 220	9.07 ± 0.54 8.14–10.38, n = 22
ID	2.61 ± 0.31 1.55–3.59, n = 495	2.68 ± 0.38 1.74–3.59, n = 120	2.76 ± 0.27 2.12–3.37, n = 119	2.48 ± 0.24 1.55–3.29, n = 228	2.74 ± 0.27 2.08–3.50, n = 28
DR	9.47 ± 0.97 6.66–12.98, n = 501	9.64 ± 1.08 7.36–11.76, n = 123	9.87 ± 1.04 7.57–12.69, n = 118	9.07 ± 0.63 6.66–10.86, n = 232	10.30 ± 1.02 8.29–12.98, n = 28
HB	9.23 ± 0.80 5.83–11.49, n = 452	9.35 ± 0.94 6.74–11.49, n = 102	9.54 ± 0.92 5.83–11.15, n = 96	9.05 ± 0.61 6.70–10.58, n = 227	9.10 ± 0.82 6.83–11.24, n = 27
DL	42.33 ± 3.50 30.26–52.23, n = 401	43.17 ± 4.06 34.95–50.92, n = 81	43.90 ± 3.51 34.40–50.20, n = 83	41.15 ± 2.67 30.26–47.26, n = 211	44.30 ± 4.12 34.91–52.23, n = 26
DW	21.80 ± 1.88 15.90–28.53, n = 458	22.14 ± 2.16 17.57–26.63, n = 106	22.80 ± 1.96 18.00–27.38, n = 105	21.11 ± 1.40 15.90–24.11, n = 222	22.23 ± 1.75 19.96–28.53, n = 25
WOI	7.14 ± 0.57 5.61–9.78, n = 501	7.28 ± 0.55 6.21–9.05, n = 123	7.36 ± 0.50 5.83–8.52, n = 116	6.87± 0.45 5.61–8.12, n = 234	7.79 ± 0.76 6.10–9.78, n = 28
WN	4.38 ± 0.42 3.12–6.06, n = 501	4.48 ± 0.44 3.60–5.63, n = 121	4.51 ± 0.41 3.50–5.73, n = 119	4.22 ± 0.32 3.12–5.03, n = 233	4.78 ± 0.52 3.76–6.06, n = 28
LN	16.38 ± 1.92 9.71–21.86, n = 465	17.02 ± 2.18 12.25–21.86, n = 103	16.93 ± 1.94 12.16–21.36, n = 116	15.65 ± 1.48 9.71–18.62, n = 218	17.34 ± 1.93 14.41–21.31, n = 28
WI	1.01 ± 0.14 0.62–1.75, n = 503	1.03 ± 0.18 0.62–1.75, n = 125	1.07 ± 0.14 0.75–1.41, n = 119	0.97 ± 0.12 0.66–1.36, n = 232	0.99 ± 0.12 0.76–1.30, n = 27
MDL	5.43 ± 0.59 2.47–7.60, n = 506	5.54 ± 0.62 4.27–7.13, n = 126	5.57 ± 0.59 4.16–7.60, n = 120	5.29 ± 0.54 2.47–6.64, n = 233	5.45 ± 0.58 4.36–7.06, n = 27
FH	1.56 ± 0.23 0.92–2.48, n = 502	1.60 ± 0.26 0.92–2.33, n = 124	1.63 ± 0.24 1.18–2.48, n = 120	1.50 ± 0.20 0.93–2.13, n = 231	1.57 ± 0.24 1.21–2.47, n = 27
CLM	20.49 ± 2.05 13.89–27.18, n = 481	21.08 ± 2.28 15.81–25.41, n = 114	21.36 ± 2.11 14.25–27.18, n = 112	19.66 ± 1.54 13.89–22.97, n = 228	21.32 ± 1.88 18.30–26.60, n = 27
ALM	20.35 ± 2.21 14.23–26.96, n = 458	20.98 ± 2.50 15.26–26.59, n = 113	21.38 ± 2.25 14.51–26.41, n = 111	19.34 ± 1.50 14.23–22.84, n = 209	21.34 ± 2.07 17.50–26.96, n = 25
Notes: Measurement acronyms are described in Table S1. All values are untransformed.					

ments had lower effect sizes. The results of the post-hoc Scheffé tests resembled those of the associated omnibus Kruskal-Wallis tests as well as the pairwise PERMANO-VAs mentioned above, both in terms of significance and effect size magnitude (i.e., based on Hedges' *g*) (omnibus tests are combined tests, as opposed to individual pairwise tests). Overall, the general pattern was that the Pakistan group was significantly different from the other three geographic groups (West Iran, East Iran, and India) in most of the measurements (and with larger effect sizes, based on Hedges' *g*), while for most measurements, the other three geographic groups did not significantly differ from each other (and had much lower effect sizes, based on Hedges' *g*) (Table 4). For the two measurements with

the highest η^2 (M1 breadth and molar row length), the post-hoc Scheffé tests indicated that Pakistan group significantly differed from all other geographic groups ($g = 1.084–1.651$, all $p < 0.050$), while the other three geographic groups did not significantly differ from each other in these two measurements (all $p \geq 0.344$; Table 4).

For the PCA based on the 21 measurements, the first principal component (PC1) by far explained most of the variation (61.7%), while PC2 and PC3 explained only 6.7% and 4.2% of the variation, respectively (Fig. 4; Fig. S2). The scatterplot based on the first two principal components showed some division of the Pakistan geographic group from all other geographic groups, which in turn did not separate markedly from each other on these two PC
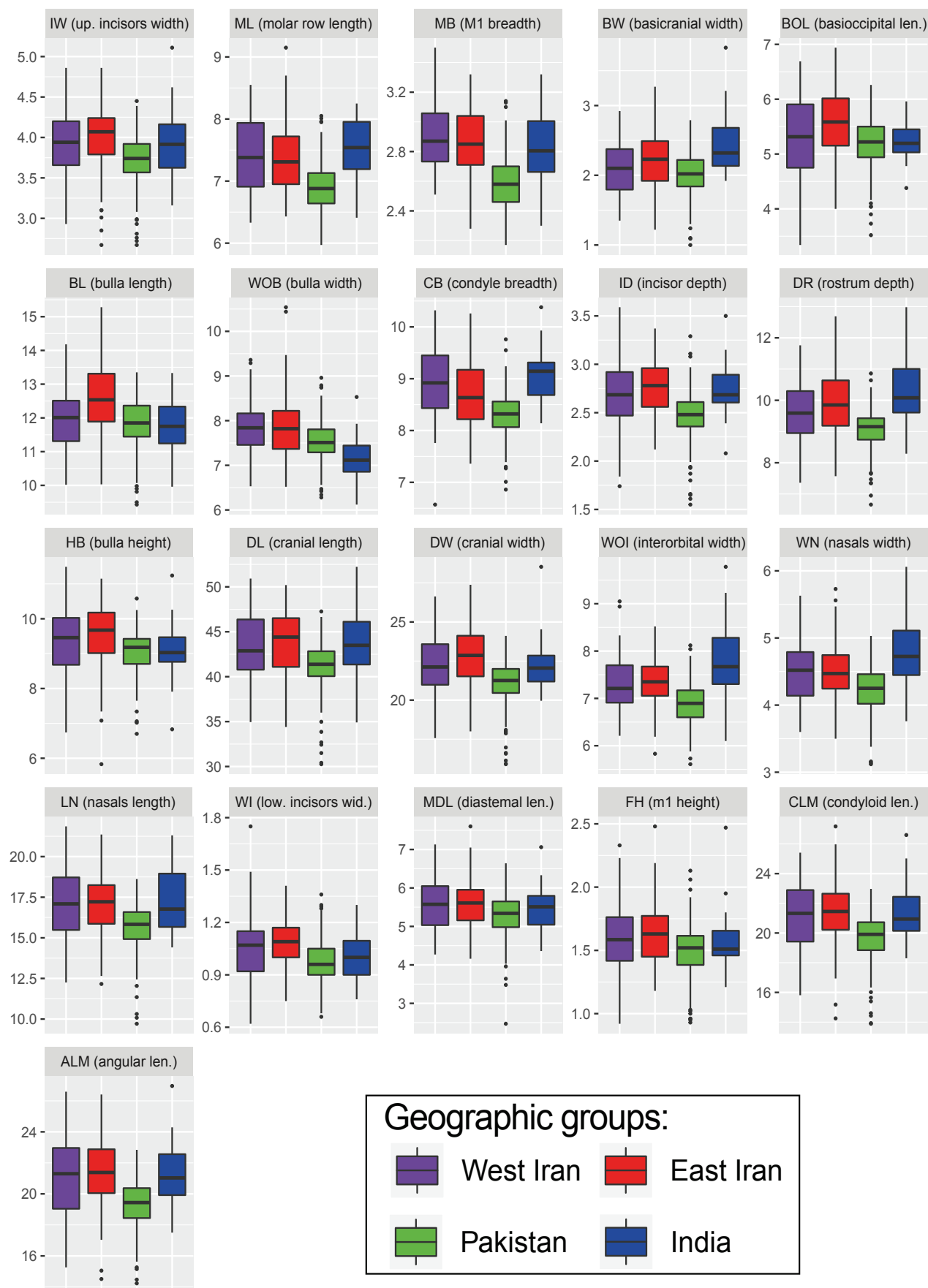
Figure 3. Boxplots of the 21 measurements, separated by geographic group. Inner box lines are medians (50% quantile), the lower and upper hinges are the 25th and 75th quantiles, the upper and lower points of the whiskers represent the minimum and maximum, and the outer points are outliers. Values appear in Table 1. Measurement acronyms are described in Table S1. All values are untransformed. The plot was generated using *ggplot2*.

Table 2. Multifactorial PERMANOVA table. The effect of geographic group ("Group"), sex ("Sex"), and their interaction on the logged measurements is examined.

Terms	DF	SS	MS	*F*	*R*²	*p*
Group	3	7.25	2.41	11.93	0.1064	**< 0.001**
Sex	1	0.32	0.32	1.60	0.0047	0.166
Group × Sex	3	0.43	0.14	0.71	0.0063	0.652
Residuals	297	60.18	0.20		0.8835	
Total	304	68.11			1.0012	

Notes: df = degrees of freedom; SS = Type-II sums of squares; MS = mean squares; *F* = pseudo-*F* statistic; *R*² = amount of the total variation explained by each model term (term SS/total SS); *p* = *p*-value based on 1000 permutations. Type-II SS are non-additive, as such the "Total" SS is not exactly the sum of all the model term SS, and the *R*² values for the terms do not add up to exactly one (as is the case for Type-I SS). Statistically significant factors (*p* < 0.05) are highlighted in bold.

Table 3. Pairwise PERMANOVA tables. The effect of the geographic group factor on the logged measurements is examined.

Group	SS	MS	F	*R*²	*p*	Holm-*p*
EI vs. IN	0.67	0.67	2.88	0.0343	0.058	0.176
EI vs. PK	3.66	3.66	21.44	0.0870	< 0.001	**0.005**
EI vs. WI	0.28	0.28	1.05	0.0083	0.300	0.300
IN vs. PK	2.62	2.62	17.10	0.0885	< 0.001	**0.005**
IN vs. WI	0.75	0.75	2.53	0.0323	0.095	0.191
PK vs. WI	3.32	3.32	17.44	0.0734	< 0.001	**0.005**

Notes: All column labels correspond to those in Table 2. *p* = raw *p*-value based on 1000 permutations; Holm-*p* = *p*-value corrected for multiple comparisons using Holm's (1979) method. Geographic groups: WI = West Iran; EI = East Iran; PK = Pakistan; IN = India. Statistically significant factors (Holm-*p* < 0.05) are highlighted in bold.

axes (Fig. 4). More specifically, specimens in the Pakistan group are marked by large PC1 scores and small PC2 scores, while those from the three other groups (West Iran, East Iran, and India) are marked by small PC1 scores and large PC2 scores (Fig. 4). This result confirmed the previous results detected by each of the pairwise PERMANOVAs and the Scheffé tests. Further examination of the loadings of each of the 21 logged measurements highly indicates that PC1 is a size-vector, as all the loadings had the same sign (negative) and similar magnitudes, which was not the case for PC2 nor PC3 (Table 5). The scatterplot of PC2 vs. PC3 also displayed some separation of the Pakistan group from the other geographic groups (Fig. S2), but not as much as the PC1 vs. PC2 plot (Fig. 4), and this separation is mostly driven by the aforementioned PC2 axis, and not the PC3 axis. Looking more closely at the loadings, it seems positive PC1 scores are associated with small size, while positive PC2 scores are associated with large ML, MB, and CB vs. small WOB, and positive PC3 scores are associated with small BL, WOB, and CB (Table 5).

The similar results that were found for most of the 21 examined measurements for the Kruskal-Wallis and the Scheffé tests (Table 4), and the fact that PC1 likely represents skull size, suggesting that the cause of this concordance among the measurements is that they are highly associated with the overall size of the skulls. However, pairwise correlation analyses on the 21 measurements seem to suggest that they were only moderately correlated (Pearson's *r* = 0.220–0.970; mean *r* = 0.628; see supplementary file 2: Data S2); size does not seem to over-

whelm the variation in the measurements (i.e., *r* was not > 0.800 for all measurements). This implies that while these measurements were indeed associated with size, there are different degrees of residual shape variation retained in each of these measurements.

Association between morphometric variation with climate and other spatial factors

The most parsimonious combination of variables chosen by the forward selection procedure (for the variation partitioning RDA analysis) for the climatic dataset were (in order of inclusion in the model): BIO9 (Mean Temperature of Driest Quarter), BIO3 (Isothermality), BIO5 (Max Temperature of Warmest Month), BIO15 (Precipitation Seasonality), BIO8 (Mean Temperature of Wettest Quarter), and BIO2 (Mean Diurnal Range); and for the spatial dataset were: MEM1, MEM6, MEM5, MEM19, MEM7, MEM3, MEM4, MEM9, MEM15, MEM17, and MEM2 (Table S4). The *vegan* function vif.cca() confirmed that all remaining explanatory variables in both datasets had VIFs ≤ 10, indicating an acceptable level of multicollinearity has been reached. For the size-corrected variables, the forward selection procedure led to the retention of the following sets of variables (in order of inclusion in the model): BIO8 (Mean Temperature of Wettest Quarter), BIO3 (Isothermality), BIO9 (Mean Temperature of Driest Quarter), BIO5 (Max Temperature of Warmest Month), BIO13 (Precipitation of Wet-

Table 4. Univariate statistical analyses. The analyses are of the differences in each of the 21 measurements among the four geographic groups. Results of both the omnibus Kruskal-Wallis and the pairwise Scheffé tests are shown.

Meas.	Kruskal-Wallis test	Scheffé test					
		WI–EI	WI–PK	WI–IN	EI–PK	EI–IN	PK–IN
IW	$\eta^2 = 0.115$, **$p < 0.050$**	$g = 0.185$, $p = 0.526$	$g = 0.632$, **$p < 0.050$**	$g = 0.033$, $p = 0.999$	$g = 0.843$, **$p < 0.050$**	$g = 0.147$, $p = 0.899$	$g = 0.709$, **$p < 0.050$**
ML	$\eta^2 = 0.212$, **$p < 0.050$**	$g = 0.106$, $p = 0.855$	$g = 1.161$, **$p < 0.050$**	$g = 0.162$, $p = 0.832$	$g = 1.084$, **$p < 0.050$**	$g = 0.287$, $p = 0.540$	$g = 1.550$, **$p < 0.050$**
MB	$\eta^2 = 0.352$, **$p < 0.050$**	$g = 0.169$, $p = 0.555$	$g = 1.651$, **$p < 0.050$**	$g = 0.345$, $p = 0.344$	$g = 1.379$, **$p < 0.050$**	$g = 0.160$, $p = 0.833$	$g = 1.245$, **$p < 0.050$**
BW	$\eta^2 = 0.062$, **$p < 0.050$**	$g = 0.391$, **$p < 0.050$**	$g = 0.176$, $p = 0.690$	$g = 0.922$, **$p < 0.050$**	$g = 0.622$, **$p < 0.050$**	$g = 0.549$, $p = 0.106$	$g = 1.293$, **$p < 0.050$**
BOL	$\eta^2 = 0.041$, **$p < 0.050$**	$g = 0.405$, **$p < 0.050$**	$g = 0.136$, $p = 0.944$	$g = 0.051$, $p = 0.999$	$g = 0.707$, **$p < 0.050$**	$g = 0.563$, $p = 0.167$	$g = 0.096$, $p = 0.976$
BL	$\eta^2 = 0.059$, **$p < 0.050$**	$g = 0.502$, **$p < 0.050$**	$g = 0.194$, $p = 0.616$	$g = 0.181$, $p = 0.853$	$g = 0.805$, **$p < 0.050$**	$g = 0.673$, **$p < 0.050$**	$g = 0.025$, $p = 0.998$
WOB	$\eta^2 = 0.084$, **$p < 0.050$**	$g = 0.018$, $p = 0.999$	$g = 0.616$, **$p < 0.050$**	$g = 1.184$, **$p < 0.050$**	$g = 0.592$, **$p < 0.050$**	$g = 1.047$, **$p < 0.050$**	$g = 0.828$, **$p < 0.050$**
CB	$\eta^2 = 0.164$, **$p < 0.050$**	$g = 0.298$, $p = 0.117$	$g = 1.184$, **$p < 0.050$**	$g = 0.222$, $p = 0.735$	$g = 0.778$, **$p < 0.050$**	$g = 0.519$, $p = 0.071$	$g = 1.747$, **$p < 0.050$**
ID	$\eta^2 = 0.172$, **$p < 0.050$**	$g = 0.221$, $p = 0.180$	$g = 0.684$, **$p < 0.050$**	$g = 0.156$, $p = 0.744$	$g = 1.101$, **$p < 0.050$**	$g = 0.061$, $p = 0.995$	$g = 1.053$, **$p < 0.050$**
DR	$\eta^2 = 0.157$, **$p < 0.050$**	$g = 0.209$, $p = 0.280$	$g = 0.699$, **$p < 0.050$**	$g = 0.614$, **$p < 0.050$**	$g = 0.996$, **$p < 0.050$**	$g = 0.417$, $p = 0.169$	$g = 1.800$, **$p < 0.050$**
HB	$\eta^2 = 0.064$, **$p < 0.050$**	$g = 0.201$, $p = 0.460$	$g = 0.396$, **$p < 0.050$**	$g = 0.267$, $p = 0.604$	$g = 0.660$, **$p < 0.050$**	$g = 0.480$, $p = 0.125$	$g = 0.076$, $p = 0.997$
DL	$\eta^2 = 0.095$, **$p < 0.050$**	$g = 0.192$, $p = 0.543$	$g = 0.623$, **$p < 0.050$**	$g = 0.277$, $p = 0.540$	$g = 0.918$, **$p < 0.050$**	$g = 0.109$, $p = 0.976$	$g = 1.098$, **$p < 0.050$**
DW	$\eta^2 = 0.122$, **$p < 0.050$**	$g = 0.317$, $p = 0.064$	$g = 0.599$, **$p < 0.050$**	$g = 0.043$, $p = 0.989$	$g = 1.042$, **$p < 0.050$**	$g = 0.293$, $p = 0.622$	$g = 0.771$, **$p < 0.050$**
WOI	$\eta^2 = 0.198$, **$p < 0.050$**	$g = 0.137$, $p = 0.732$	$g = 0.831$, **$p < 0.050$**	$g = 0.848$, **$p < 0.050$**	$g = 1.021$, **$p < 0.050$**	$g = 0.774$, **$p < 0.050$**	$g = 1.852$, **$p < 0.050$**
WN	$\eta^2 = 0.126$, **$p < 0.050$**	$g = 0.057$, $p = 0.961$	$g = 0.702$, **$p < 0.050$**	$g = 0.634$, **$p < 0.050$**	$g = 0.797$, **$p < 0.050$**	$g = 0.611$, **$p < 0.050$**	$g = 1.563$, **$p < 0.050$**
LN	$\eta^2 = 0.108$, **$p < 0.050$**	$g = 0.043$, $p = 0.996$	$g = 0.776$, **$p < 0.050$**	$g = 0.145$, $p = 0.864$	$g = 0.772$, **$p < 0.050$**	$g = 0.206$, $p = 0.789$	$g = 1.093$, **$p < 0.050$**
WI	$\eta^2 = 0.078$, **$p < 0.050$**	$g = 0.215$, $p = 0.200$	$g = 0.400$, **$p < 0.050$**	$g = 0.238$, $p = 0.763$	$g = 0.721$, **$p < 0.050$**	$g = 0.534$, $p = 0.134$	$g = 0.137$, $p = 0.956$
MDL	$\eta^2 = 0.034$, **$p < 0.050$**	$g = 0.050$, $p = 0.976$	$g = 0.426$, **$p < 0.050$**	$g = 0.141$, $p = 0.940$	$g = 0.492$, **$p < 0.050$**	$g = 0.199$, $p = 0.847$	$g = 0.287$, $p = 0.621$
FH	$\eta^2 = 0.043$, **$p < 0.050$**	$g = 0.133$, $p = 0.671$	$g = 0.451$, **$p < 0.050$**	$g = 0.105$, $p = 0.979$	$g = 0.625$, **$p < 0.050$**	$g = 0.250$, $p = 0.706$	$g = 0.359$, $p = 0.478$
CLM	$\eta^2 = 0.140$, **$p < 0.050$**	$g = 0.124$, $p = 0.750$	$g = 0.774$, **$p < 0.050$**	$g = 0.106$, $p = 0.931$	$g = 0.965$, **$p < 0.050$**	$g = 0.019$, $p = 1.000$	$g = 1.044$, **$p < 0.050$**
ALM	$\eta^2 = 0.163$, **$p < 0.050$**	$g = 0.163$, $p = 0.520$	$g = 0.857$, **$p < 0.050$**	$g = 0.146$, $p = 0.849$	$g = 1.134$, **$p < 0.050$**	$g = 0.015$, $p = 1.000$	$g = 1.268$, **$p < 0.050$**

Notes: p = raw p-value. For the Kruskal-Wallis tests, all p-values remained significant at $p < 0.050$ after correcting for multiple comparisons using Holm's (1979) method. η^2 = eta-squared; g = Hedges' g. Both the η^2 and the g values are the absolute values of the effect sizes. Measurement acronyms are described in Table S1. Geographic groups: WI = West Iran; EI = East Iran; PK = Pakistan; IN = India. Statistically significant factors ($p < 0.05$) are highlighted in bold.

test Month), BIO18 (Precipitation of Warmest Quarter), BIO2 (Mean Diurnal Range), and BIO15 (Precipitation Seasonality) for the climatic dataset, and MEM1, MEM5, MEM4, MEM3, MEM11, MEM2, MEM19, MEM12, MEM8, MEM10, MEM7, MEM13, MEM16, MEM18, MEM15, and MEM17 (Table S4) for the spatial dataset, all of which had VIFs ≤ 12.

The variation partitioning by RDA analysis indicated that all fractions were significant at $p < 0.05$ for both the non-sheared and the sheared data (Table 6). The shared influence of spatial and climatic variables explained 25.0% of the total variation ($R^2_{adj} = 0.250$; Table 6). The "pure" variation explained by spatial variables (spatial variation that does not correspond to climatic variation) was similar to that of climatic variables (i.e., non-spatial climatic variation) ($R^2_{adj} = 0.085$ vs. 0.087, respectively; Table 6). Spatially structured climatic variation and the converse (unexplained spatial variation) were also similar ($R^2_{adj} = 0.164$ vs. 0.162, respectively), while the variation explained by joint effect of spatial and climate variables
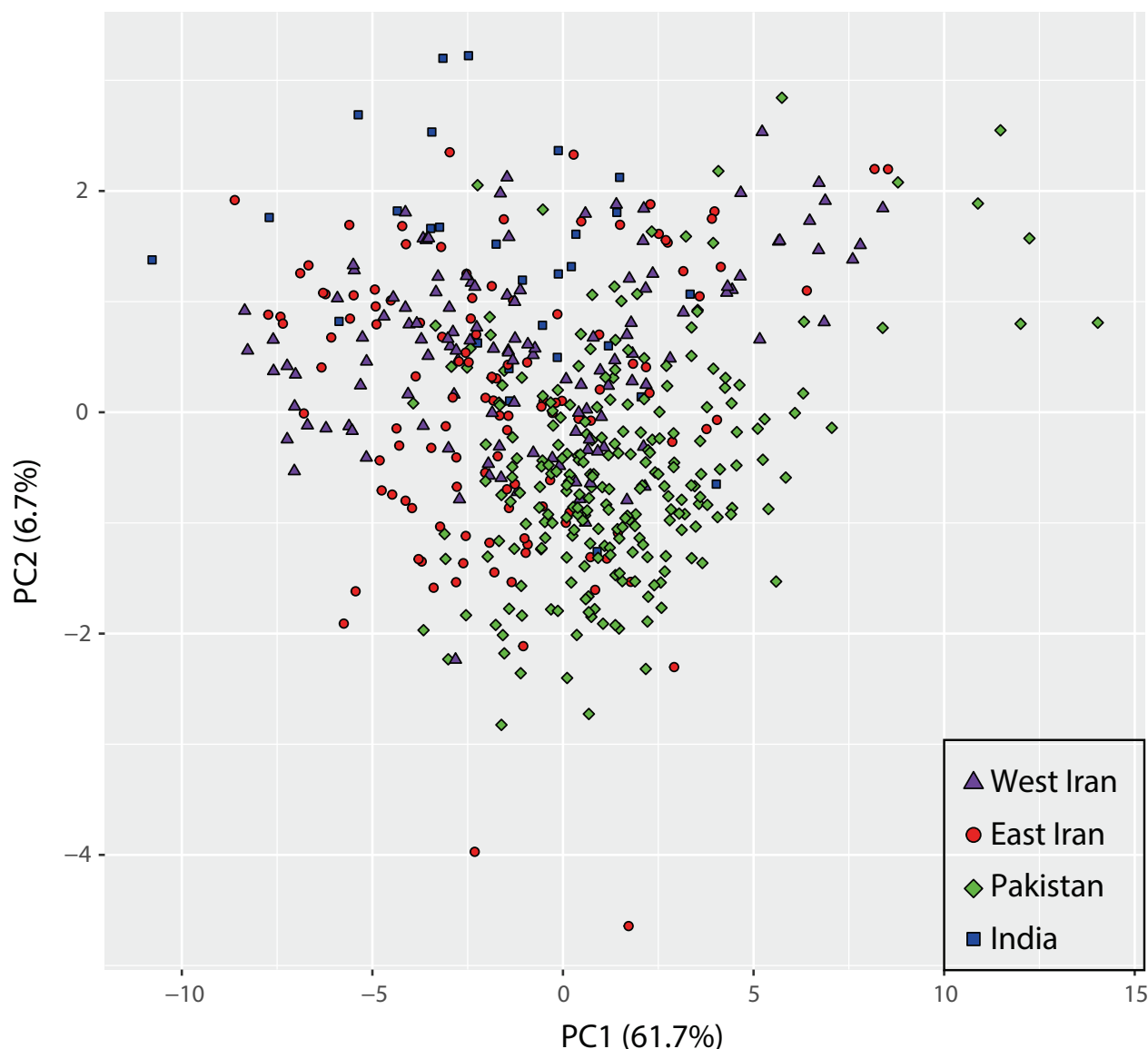
Figure 4. Scatterplot of the first two principal components of PCA conducted on the 21 logged measurements. Specimens are divided into the four geographic groups (see legend). The amount of explained variance by each PC axis is indicated in parentheses. The plot was generated using *ggplot2*.

was 7.7% (R^2_{adj} = 0.077; Table 6). Similar results were found after size correction, but with smaller effect sizes (R^2_{adj}). The shared influence of spatial and climatic variables explained 14.9% of total variation after shearing (R^2_{adj} = 0.149; Table 6). Of this variation, the variation explained by spatial variables was 11.9% and dropped to 6.5% after partialing out the effects of climatic variables, and the variation explained by climatic variables was 8.4% and dropped to 3% after accounting for spatial autocorrelation (Table 6). The variation explained by joint effect of spatial and climate variables (after shearing) was 5.5% (R^2_{adj} = 0.055; Table 6).

The RDA plots of the climate variables after removing the effect of the spatial variables (showing the response of the morphometric variation to climate variables) show no clear separation of the geographic groups on the first two RDA axes both for the non-sheared data (Fig. S3a) and for the sheared data (Fig. S3c). The amount of vari-

ation accounted for by the first two RDA axes and their significance are indicated in the Fig. S3 legend. When not considering the geographic groups, and examining the specimens as a whole, the strongest separation among the specimens (main axes of variation for climate, represented by the longest vectors) seems to be based on BIO5 (Max Temperature of Warmest Month), followed by BIO9 (Mean Temperature of Driest Quarter), and BIO8 (Mean Temperature of Wettest Quarter), respectively, for the non-sheared data (Fig. S3b), while for the sheared data, the separation seems to be greatest for BIO5 (Max Temperature of Warmest Month) followed by BIO18 (Precipitation of Warmest Quarter) (Fig. S3d). For the most part, the directions of this climatic variation seem to be neither strongly correlated nor opposed (Fig. S3b, d). However, among the variables, specimens with increasing BIO9 (Mean Temperature of Driest Quarter) have decreasing BIO8 (Mean Temperature of Wettest Quarter)

Table 5. PCA loadings. Loadings describe how much each of the 21 logged measurements contributes to each PC axis.

Measurement	PC1	PC2	PC3
IW	−0.196	0.126	0.184
ML	−0.218	**0.352**	−0.135
MB	−0.156	**0.586**	−0.256
BW	−0.215	0.047	0.167
BOL	−0.204	−0.268	−0.163
BL	−0.207	−0.282	**−0.371**
WOB	−0.158	**−0.335**	**−0.424**
CB	−0.159	**0.313**	**−0.339**
ID	−0.243	0.003	0.033
DR	−0.268	0.037	0.050
HB	−0.202	−0.268	−0.218
DL	−0.227	−0.059	−0.096
DW	−0.253	−0.059	−0.005
WOI	−0.230	0.179	−0.100
WN	−0.243	0.070	0.113
LN	−0.240	−0.051	0.142
WI	−0.209	−0.124	0.291
MDL	−0.223	−0.094	0.242
FH	−0.191	−0.020	**0.343**
CLM	−0.254	−0.040	0.099
ALM	−0.231	0.022	0.128

Notes: Measurement acronyms are described in Table S1. To improve interpretation, measurments with PC loadings above an arbitrarily large cut-off value of ≥ |0.3| are highlighted in bold. Only PC1–3 are shown.

(for the non-sheared data; Fig. S3b), while specimens with increasing BIO5 (Max Temperature of Warmest Month) have decreasing BIO18 (Precipitation of Warmest Quarter) (for the sheared data; Fig. S3d). BIO5 (Max Temperature of Warmest Month) seems to be the strongest contributor to RDA axis 1 for the non-sheared data (Fig. S3b), while BIO8 (Mean Temperature of Wettest Quarter) seems to be the strongest contributor to RDA axis 1 for the sheared data (Fig. S3d). The other climatic variables seem to more strongly contribute to RDA axis 2 (Fig. S3b, d).

Discussion

The results of each of the omnibus and the pairwise PER-MANOVAs, along with the omnibus Kruskal-Wallis and the pairwise Scheffé tests, in addition to the PCA, all support the idea that the considered geographic groups differ from each other in skull morphology. The Pakistan group was the most distinct among the geographic groups (Figs.4, S2; Table 3), and the most discriminating of the measurements were M1 breadth and molar row length (Table 4). Perhaps more interestingly, the variation partitioning RDA analysis indicated equal (significant, but of low effect size) contribution of each of climatic and non-climatic spatial factors in explaining the examined skull variation, with size-correction further reducing the

effect size (Table 6). Taken together, the results could perhaps indicate that the Pakistan *T. indica* population is the most differentiated among the examined, and that its differentiation cannot be explained largely by climate, geography, nor size.

The distinctiveness of the Pakistani *T. indica* population found here could not be detected in the earlier *T. indica* morphometric studies because of their more restrictive geographic scale. For example, Shehab et al. (2011) found similarities in craniodental measurements between Syrian, southern Turkish, and Iranian *T. indica* populations, except with respect to size (the prior two populations were larger than the latter) (also see Yigit et al. 2001). Similarly, within Iranian populations, most prior studies found only slight skull differences (e.g., Mirshamsi et al. 2007; Ashrafzadeh et al. 2012; Khajeh et al. 2019). By largely expanding the geographic scale of the compared Indian gerbil populations (Fig. 1), the present study was able to detect more pronounced intraspecific differences (Figs. 3, 4, S2; Tables 3, 4).

The distinctiveness of the Pakistani *T. indica* skulls detected in this study was mostly a result of differences in their dentition, as captured by M1 breadth and the molar row length (Table 4). Dentition, particularly that of the molars, is largely associated with diet, as numerous previous rodent studies have demonstrated (Renaud et al. 2005; Samuels 2009; Croft et al. 2011; Martin et al. 2016; Verde Arregoitia et al. 2017). This is mostly a consequence of dental characteristics being related to mastication and all other forms of oral food processing. Since *T. indica* have

Table 6. Variation partitioning by RDA analysis. Summary of variation partitioning by (partial) RDA partitions explained by each subset of variables and their combinations. In all instances the response variables are the 21 morphometric measurements (logged, including imputed missing data; top = not sheared, bottom = sheared). The different combinations of spatial and climatic variables (explanatory variables) are indicated in each row.

Terms	DF	F	p	R^2_{adj}
Before size correction				
Spatial	11	9.94	< 0.050	0.162
Climate	6	17.63	< 0.050	0.164
Spatial | climate	11	6.19	< 0.050	0.085
Climate | spatial	6	10.64	< 0.050	0.087
Spatial + climate	17	10.94	< 0.050	0.250
Joint effect [b]				0.077
After size-correction				
Spatial	16	5.31	< 0.050	0.119
Climate	8	6.83	< 0.050	0.084
Spatial | climate	16	3.39	< 0.050	0.065
Climate | spatial	8	3.14	< 0.050	0.030
Spatial + climate	24	4.71	< 0.050	0.149
Joint effect [b]				0.055

Notes: "Spatial" = fraction explained by spatial variables (RDA model that only includes spatial explanatory variables); "Climate" = fraction explained by climate variables (RDA model that only includes climatic explanatory variables); "Spatial | climate" = fraction explained by spatial variables after partialing out the effects of climatic variables (partial RDA model that includes spatial explanatory variables and climatic conditioning variables); "Climate | spatial" = fraction explained by climatic variables after partialing out the effects of spatial variables (partial RDA model that includes climatic explanatory variables and spatial conditioning variables); "Spatial + climate" = total explained by both spatial and climatic variables (RDA model that includes both spatial and climatic explanatory variables); and "Joint effect [b]" = joint effect (common fraction) of spatial and climate variables (non-testable component calculated as the difference "Climate" minus "Climate | spatial" or "Spatial" minus "Spatial | climate"; this fraction arises because explanatory variables in different sets are correlated). Geographic variables were extracted from dbMEM analysis (positive MEM variables), while the climatic variables are standardized (to a mean of zero and unit variance) and based on WorldClim (version 2). F = pseudo-F based of permutation tests, R^2_{adj} = adjusted coefficient of multiple determination. All p-values are significant. For the un-sheared dataset, spatial variables are based on (in order of inclusion in the model): MEM1, MEM6, MEM5, MEM19, MEM7, MEM3, MEM4, MEM9, MEM15, MEM17, and MEM2, while the climatic variables are based on: BIO9, BIO3, BIO5, BIO15, BIO8, and BIO2. For the sheared dataset, spatial variables are based on (in order of inclusion in the model): MEM1, MEM5, MEM4, MEM3, MEM11, MEM2, MEM19, MEM12, MEM8, MEM10, MEM7, MEM13, while the climatic variables are based on BIO8, BIO3, BIO9, BIO5, BIO13, BIO18, BIO2, and BIO15. Statistical significance was based on 1000 permutations.

a broad diet that includes both plant material and animal matter (Shenbrot and Krasnov 2001; Vaughan et al. 2011; Akhtar et al. 2017), it stands to reason that different populations, potentially with access to different nutritional sources, and in different proportions, can undergo selective pressures (or at least developmental plasticity) that would drive their differentiation from each other.

While minor, this study found that non-climatic spatial factors did explain some of the skull variation (Table 6; Fig. S2). Potentially, this could be explained by the isolation-by-distance model (Wright 1943), whereby the further the populations are from each other, the more limited the dispersal between them will be, and consequently the more differentiation. The association between skull morphometric and geographic distance is not a novel observation, and has been shown to occur in numerous previous rodent studies (Muñoz-Muñoz et al. 2011; Cordero and Epps 2012; Quintela et al. 2016; Alhajeri 2018; Alhajeri 2019). Other non-climatic spatial factors unexamined in this study could also explain the detected association with skull variation, such as environmental heterogeneity and geographic barriers (Quintela et al. 2016).

Similarly, this study also found a significant, but minor, influence of climatic spatial factors on geographic skull variation, potentially explained by climatic adaptation (Table 6; Fig. S2). Perhaps a noteworthy observation was that, for the for the morphometric variables not corrected for size, most of the retained climatic variables are temperature-related, whereas for the variables corrected for size, more precipitation-related variables were retained (Table S4). These same variables also seem to be associated with their respective RDA axes (Fig. S3). Climate has previously been shown to influence the size of *T. indica* (Khajeh et al. 2019). The morphology of other rodents, including that of the skull, has been similarly demonstrated to be influenced by climate (Monteiro et al. 2003; Renaud et al. 2005; Pergams and Lawler 2009; Wolf et al. 2009; Tabatabaei Yazdi and Adriaens 2011; Cordero and Epps 2012; Alhajeri et al. 2020b; Alhajeri et al. 2020a). There are many potential explanations for this association between climate and morphology, including the association of climate with resource availability (Wolf et al. 2009; Cordero and Epps 2012; Alhajeri 2018; Magnus et al. 2018) and climatic adaptation (e.g., to aridity) (Magnus et al. 2017; Alhajeri and Steppan 2018; Khajeh et al. 2019). More spe-

cifically, the aforementioned temperature-related climatic variables, that seem to be associated with the morphometric variables not corrected for size (which may still retain some residual size/allometric information), may partly reflect minor clinal variation in size associated with Bergmann's rule (see Alhajeri and Steppan 2016). On the other hand, the precipitation-related variables that are apparently associated with the morphometric variables corrected for size (which presumably should be less associated with size, and more with shape), may capture spatial (geographic) dietary variation of the different *T. indica* populations, as a result of having access to different food items in habitats with different precipitation regimes. Different cranial morphologies would be more adapted to different diets.

Conclusion

T. indica specimens from Pakistan significantly differed in skull morphology from those belonging to the West Iran, East Iran, and India populations, which in turn did not significantly differ from each other. This differentiation was found to be driven to a similar (but small) degree by both climatic and non-climatic spatial factors. Thus, the distinctiveness of the Pakistani *T. indica* population could be driven by unexamined factors. Potential factors that could be examined in future studies include environmental heterogeneity and dispersal barriers. The taxonomic implications of the results of the present study are that *T. indica* could hide cryptic diversity, if only at the subspecific level. This could be explored further and in more detail with a larger dataset that incorporates more specimens from, or from near the Pakistan group of localities, as well as by including other localities that are unexamined in the present study. Characters other than the cranial measurments examined herein, and other phenotypic characters can also be explored to further elucidate the aforementioned patterns.

Data availability

All data generated in this study are deposited in the supporting information (see supplementary file 1: Data S1).

Author contributions

BHA conceived and designed the study. ZD and BHA photographed the specimens. ZD collected and analyzed the data. ZD, BHA, and HA wrote the manuscript. All authors read and approved the final version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

References

Abramoff MD, Magalhaes PJ, Ram SJ (2004) Image Processing with ImageJ. Biophotonics International 11: 36–42.

Akhtar M, Naz S, Iqbal M, Azeem W, Beg MA (2017) Effect of Seasonal Swings and Age Specific Variations on Body Weight of Indian Gerbille (*Tatera indica*). Pakistan Journal of Zoology 49: 1–4.

Alhajeri BH (2018) Craniomandibular Variation in the Taxonomically Problematic Gerbil Genus *Gerbillus* (Gerbillinae, Rodentia): Assessing the Influence of Climate, Geography, Phylogeny, and Size. Journal of Mammalian Evolution 25: 261–276. https://doi.org/10.1007/s10914-016-9377-2

Alhajeri BH (2019) Cranial variation in geographically widespread dwarf gerbil *Gerbillus nanus* (Gerbillinae, Rodentia) populations: Isolation by distance versus adaptation to local environments. Journal of Zoological Systematics and Evolutionary Research 57: 191–203. https://doi.org/10.1111/jzs.12247

Alhajeri BH (2021a) A Geometric Morphometric Analysis of Geographic Mandibular Variation in the Dwarf Gerbil *Gerbillus nanus* (Gerbillinae, Rodentia). Journal of Mammalian Evolution 28: 469–480. https://doi.org/10.1007/s10914-020-09530-9

Alhajeri BH (2021b) A morphometric comparison of the cranial shapes of Asian dwarf hamsters (*Phodopus*, Cricetinae, Rodentia). Zoologischer Anzeiger 292: 184–196. https://doi.org/10.1016/j.jcz.2021.04.001

Alhajeri BH (2021c) Cranial variation in allactagine jerboas (Allactaginae, Dipodidae, Rodentia): a geometric morphometric study. Zoological Research 42: 182–194. https://doi.org/10.24272/j.issn.2095-8137.2020.302

Alhajeri BH (2022a) *Desmodilliscus braueri* crania compared to *Pachyuromys duprasi* (Desmodilliscini, Gerbillinae, Rodentia). Mammalia 86: 77–87. https://doi.org/10.1515/mammalia-2021-0036

Alhajeri BH (2022b) Geometric differences between the crania of Australian hopping mice (*Notomys*, Murinae, Rodentia). Australian Mammalogy 44: 24–38. https://doi.org/10.1071/AM20067

Alhajeri BH, Steppan SJ (2016) Association between climate and body size in rodents: A phylogenetic test of Bergmann's rule. Mammalian Biology 81: 219–225. https://doi.org/http://dx.doi.org/10.1016/j.mambio.2015.12.001

Alhajeri BH, Steppan SJ (2018) A phylogenetic test of adaptation to deserts and aridity in skull and dental morphology across rodents.

Journal of Mammalogy 99: 1197–1216. https://doi.org/10.1093/jmammal/gyy099

Alhajeri BH, Hunt OJ, Steppan SJ (2015) Molecular systematics of gerbils and deomyines (Rodentia: Gerbillinae, Deomyinae) and a test of desert adaptation in the tympanic bulla. Journal of Zoological Systematics and Evolutionary Research 53: 312–330. https://doi.org/10.1111/jzs.12102

Alhajeri BH, Porto LM V, Maestri R (2020a) Habitat productivity is a poor predictor of body size in rodents. Current Zoology 66: 135–143. https://doi.org/10.1093/cz/zoz037

Alhajeri BH, Fourcade Y, Upham NS, Alhaddad H (2020b) A global test of Allen's rule in rodents. Global Ecology and Biogeography 29: 2248–2260. https://doi.org/10.1111/geb.13198

Álvarez A, Ercoli MD, Olivares AI, De Santi NA, Verzi DH (2021) Evolutionary Patterns of Mandible Shape Diversification of Caviomorph Rodents. Journal of Mammalian Evolution 28: 47–58. https://doi.org/10.1007/s10914-020-09511-y

Anderson D (2018) esvis: Visualization and Estimation of Effect Sizes. Available from: https://cran.r-project.org/package=esvis

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecology 26: 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Anderson MJ, Walsh DCI (2013) PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? Ecological Monographs 83: 557–574. https://doi.org/10.1890/12-2010.1

Anderson MJ, Gorley RN, Clarke RK (2008) Permanova+ for Primer: guide to software and statisticl methods. PRIMER-E. Plymouth, UK.

Ashrafzadeh MR, Shahi T, Karami M, Darvish J (2012) Intraspecific variations within *Tatera indica* Hardwicke, 1807 (Rodentia: Muridae) populations in Hormozgan province, Iran. Journal Of Natural Environment (Iranian Journal Of Natural Resources) 64. Available from: https://www.sid.ir/en/journal/ViewPaper.aspx?ID=261522.

Aulagnier S (2009) Mammals of Europe, North Africa and the Middle East. A. & C. Black, London.

Bates PJJ (1988) Systematics and zoogeography of *Tatera* (Rodentia: Gerbillinae) of north-east Africa and Asia. Bonner zoologische Beiträge 33: 265–303.

Beck MW (2019) ggord: Ordination Plots with ggplot2. Available from: https://github.com/fawda123/ggord

Blanchet FG, Legendre P, Borcard D (2008) Forward selection of explanatory variables. Ecology 89: 2623–2632. https://doi.org/10.1890/07-0986.1

Borcard D, Legendre P (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling 153: 51–68. https://doi.org/10.1016/S0304-3800(01)00501-4

Borcard D, Legendre P, Drapeau P (1992) Partialling out the Spatial Component of Ecological Variation. Ecology 73: 1045–1055. https://doi.org/10.2307/1940179

Van Buuren S, Groothuis-Oudshoorn K (2011) mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software 45. https://doi.org/10.18637/jss.v045.i03

Cordero GA, Epps CW (2012) From Desert to Rainforest: Phenotypic Variation in Functionally Important Traits of Bushy-Tailed Woodrats (Neotoma cinerea) Across Two Climatic Extremes. Journal of Mammalian Evolution 19: 135–153. https://doi.org/10.1007/s10914-012-9187-0

Croft DA, Niemi K, Franco A (2011) Incisor morphology reflects diet in caviomorph rodents. Journal of Mammalogy 92: 871–879. https://doi.org/10.1644/10-mamm-a-226.1

Dray S, Bauman D, Blanchet G, Borcard D, Clappe S, Guenard G, Jombart T, Larocque G, Legendre P, Madi N, Wagner H (2019) adespatial: Multivariate Multiscale Spatial Analysis. Available from: https://cran.r-project.org/package=adespatial

Dray S, Pélissier R, Couteron P, Fortin M-J, Legendre P, Peres-Neto PR, Bellier E, Bivand R, Blanchet FG, De Cáceres M, Dufour A-B, Heegaard E, Jombart T, Munoz F, Oksanen J, Thioulouse J, Wagner HH (2012) Community ecology in the age of multivariate multiscale spatial analysis. Ecological Monographs 82: 257–275. https://doi.org/10.1890/11-1183.1

Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37: 4302–4315. https://doi.org/10.1002/joc.5086

Fox J, Weisberg S (2011) An R Companion to Applied Regression. Second Edition. A. & C. Black, London.

Google (2019) Google maps. Available from: google.com/maps.

Goyal SP, Ghosh PK (1993) Burrow structure of two gerbil species of Thar desert, India. Acta Theriologica 38: 453–456. https://doi.org/10.4098/AT.arch.93-37

Goyal SP, Ghosh PK, Sasidharan TO, Chand P (1988) Body water relations in two species of gerbil (*Tatera indica indica* and *Meriones hurrianae*) of the Indian desert. Comparative Physiology B 158: 127–134. https://doi.org/10.1007/BF00692736

Grosjean P, Ibanez F (2018) pastecs: Package for Analysis of Space-Time Ecological Series. Available from: https://cran.r-project.org/package=pastecs

Hedges L V (1981) Distribution Theory for Glass's Estimator of Effect size and Related Estimators. Journal of Educational Statistics 6: 107–128. https://doi.org/10.3102/10769986006002107

Hervé M (2019) RVAideMemoire: Testing and Plotting Procedures for Biostatistics. Available from: https://cran.r-project.org/package=RVAideMemoire

Hijmans RJ (2019) raster: Geographic Data Analysis and Modeling. Available from: http://cran.r-project.org/package=raster

Holm S (1979) A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics 6: 65–70. https://doi.org/10.2307/4615733

Humphries JM, Bookstein FL, Chernoff B, Smith GR, Elder RL, Poss SG (1981) Multivariate Discrimination by Shape in Relation to Size. Systematic Zoology 30: 291–308. https://doi.org/10.2307/2413251

Inman J (1835) Navigation and Nautical Astronomy for the Use of British Seamen. 3rd Edition. C. and J.Rivington, London.

Kassambara A (2019) rstatix: Pipe-Friendly Framework for Basic Statistical Tests. Available from: https://rpkgs.datanovia.com/rstatix.

Khajeh A, Mohammadi Z, Ghorbani F, Meshkani J (2019) Variation of the auditory system in the Indian Gerbil, *Tatera indica* Hardwicke, 1807, (Muridae, Rodentia) from the east of Iran. Journal of Asia-Pacific Biodiversity 12: 139–143. https://doi.org/10.1016/j.japb.2018.12.003

Kruskal WH, Wallis WA (1952) Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association 47: 583–621. https://doi.org/10.2307/2280779

Kryštufek B, Vohralík V (2009) Mammals of Turkey and Cyprus, Rodentia II: Cricetinae, Muridae, Spalacidae, Calomyscidae, Capromyidae, Hystricidae, Castoridae. Science and Research Centre of the Republic of Slovenia, Ljubljana.

Kryštufek B, Shenbrot G, Sozen M, Molur S (2017) *Tatera indica*. Available from: http://www.iucnredlist.org/details/21514/0

Legendre P, Legendre L (2012) Numerical Ecology. 3rd Edition. Elsevier, Amsterdam.

Liu Q (1997) Variation Partitioning By Partial Redundancy Analysis (RDA). Environmetrics 8: 75–85. https://doi.org/10.1002/(SICI)-1099-095X(199703)8:2<75::AID-ENV250>3.0.CO;2-N

Magnus LZ, Machado RF, Cáceres N (2017) Comparative ecogeographical variation in skull size and shape of two species of woolly opossums (genus *Caluromys*). Zoologischer Anzeiger 267: 139–150. https://doi.org/10.1016/j.jcz.2017.03.003

Magnus LZ, Machado RF, Cáceres N (2018) Ecogeography of South-American Rodentia and Lagomorpha (Mammalia, Glires): Roles of size, environment, and geography on skull shape. Zoologischer Anzeiger 277: 33–41. https://doi.org/10.1016/j.jcz.2018.06.002

Martin SA, Alhajeri BH, Steppan SJ (2016) Dietary adaptations in the teeth of murine rodents (Muridae): a test of biomechanical predictions. Biological Journal of the Linnean Society 119: 766–784. https://doi.org/10.1111/bij.12822

Martínez JJ, Di Cola V (2011) Geographic distribution and phenetic skull variation in two close species of *Graomys* (Rodentia, Cricetidae, Sigmodontinae). Zoologischer Anzeiger 250: 175–194. https://doi.org/10.1016/j.jcz.2011.03.001

Mirshamsi O, Darvish J, Kayvanfar N (2007) A preliminary study on Indian Gerbils, *Tatera indica* Hardwicke, 1807 at population level in eastern and southern parts of Iran (Rodentia: Muridae). Iranian Journal of Animal Biosystematics 3: 49–61.

Monteiro LR, Duarte LC, dos Reis SF (2003) Environmental correlates of geographical variation in skull and mandible shape of the punaré rat *Thrichomys apereoides* (Rodentia: Echimyidae). Journal of Zoology 261: 47–57. https://doi.org/10.1017/S0952836903003893

Muñoz-Muñoz F, Sans-Fuentes MA, López-Fuster MJ, Ventura J (2011) Evolutionary modularity of the mouse mandible: dissecting the effect of chromosomal reorganizations and isolation by distance in a Robertsonian system of *Mus musculus domesticus*. Journal of Evolutionary Biology 24: 1763–1776. https://doi.org/10.1111/j.1420-9101.2011.02312.x

Naimi B, Hamm N, Groen T, Skidmore A, Toxopeus A (2014) Where is positional uncertainty a problem for species distribution modelling. Ecography 37: 191–203.

Nanova O (2014) Geographical variation in the cranial measurements of the midday jird *Meriones meridianus* (Rodentia: Muridae) and its taxonomic implications. Journal of Zoological Systematics and Evolutionary Research 52: 75–85. https://doi.org/10.1111/jzs.12032

Nychka D, Furrer R, Paige J, Sain S (2017) fields: Tools for spatial data. Available from: https://cran.r-project.org/package=fields.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2019) Package 'vegan.' Available from: https://github.com/vegandevs/vegan.

Pergams ORW, Lawler JJ (2009) Recent and Widespread Rapid Morphological Change in Rodents. PLOS ONE 4: e6452. https://doi.org/10.1371/journal.pone.0006452

Quintela FM, Fornel R, Freitas TRO (2016) Geographic variation in skull shape of the water rat *Scapteromys tumidus* (Cricetidae, Sigmodontinae): isolation-by-distance plus environmental and geographic barrier effects? Anais da Academia Brasileira de Ciências 29: 451–466. https://doi.org/10.1590/0001-3765201620140631.

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: http://www.r-project.org.

Renaud S, Millien V (2001) Intra- and interspecific morphological variation in the field mouse species *Apodemus argenteus* and *A. speciosus* in the Japanese archipelago: the role of insular isolation and biogeographic gradients. Biological Journal of the Linnean Society 74: 557–569. https://doi.org/10.1111/j.1095-8312.2001.tb01413.x

Renaud S, Michaux J, Schmidt DN, Aguilar J-P, Mein P, Auffray J-C (2005) Morphological evolution, ecological diversification and climate change in rodents. Proceedings of the Royal Society B 272: 609–17. https://doi.org/10.1098/rspb.2004.2992

Renvoisé E, Montuire S, Richard Y, Quéré J-P, Gerber S, Cucchi T, Chateau-Smith C, Tougard C (2012) Microevolutionary relationships between phylogeographical history, climate change and morphological variability in the common vole (*Microtus arvalis*) across France. Journal of Biogeography 39: 698–712. https://doi.org/10.1111/j.1365-2699.2011.02611.x

Salazar G (2018) EcolUtils: Utilities for community ecology analysis. Available from: https://github.com/GuillemSalazar/EcolUtils.

Samuels JX (2009) Cranial morphology and dietary habits of rodents. Zoological Journal of the Linnean Society 156: 864–888. https://doi.org/10.1111/j.1096-3642.2009.00502.x

Scheffé H (1999) The Analysis of Variance. Wiley, New York.

Shehab AH, Al-Ahmad SA, Samara FF (2011) Morphology and Distribution of the Indian Gerbil, *Tatera indica* (Hardwicke, 1807), in Syria. Zoology in the Middle East 52: 3–10. https://doi.org/10.1080/09397140.2011.10638473

Shenbrot G, Krasnov B (2001) Geographic variation in the role of gerbils and jirds (Gerbillinae) in rodent communities across the Great Palaearctic Desert Belt. In: Denys C, Granjon L, Poulet A (Eds) African Small Mammals/Petit mammifères africains. IRD Editions, Paris, 511–529.

Signorell A (2018) DescTools: Tools for Descriptive Statistics. Available from: https://cran.r-project.org/package=DescTools.

Stacklies W, Redestig H, Scholz M, Walther D, Selbig J (2007) pcaMethods – A bioconductor package providing PCA methods for incomplete data. Bioinformatics 23: 1164–1167. https://doi.org/10.1093/bioinformatics/btm069

Tabatabaei Yazdi F, Adriaens D (2011) Patterns of skull shape variation in *Meriones persicus* (Rodentia: Muridae) in relation to geoclimatical conditions. Iranian Journal of Animal Biosystematics 7: 129–142.

Tabatabaei Yazdi F, Alhajeri BH (2018) Sexual dimorphism, allometry, and interspecific variation in the cranial morphology of seven *Meriones* species (Gerbillinae, Rodentia). Hystrix 29: 162–167. https://doi.org/10.4404/hystrix-00018-2017

Tabatabaei Yazdi F, Adriaens D, Darvish J (2012) Geographic pattern of cranial differentiation in the Asian Midday Jird *Meriones meridianus* (Rodentia: Muridae: Gerbillinae) and its taxonomic implications. Journal of Zoological Systematics and Evolutionary Research 50: 157–164. https://doi.org/10.1111/j.1439-0469.2011.00642.x

Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, Botstein D, Altman RB (2001) Missing value estimation methods for DNA microarrays. Bioinformatics 17: 520–525. https://doi.org/10.1093/bioinformatics/17.6.520

Vaughan TA, Ryan JM, Czaplewski NJ (2011) Mammology. Jones and Bartlett's, Burlington, MA.

Verde Arregoitia LD, Fisher DO, Schweizer M (2017) Morphology captures diet and locomotor types in rodents. Royal Society Open Science 4: 160957. https://doi.org/10.1098/rsos.160957

Wasser L (2017) GIS With R: Projected vs Geographic Coordinate Reference Systems. Available from: https://www.earthdatascience.org/courses/earth-analytics/spatial-data-r/geographic-vs-projected-coordinate-reference-systems-UTM.

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.

Wolak ME, Fairbairn DJ, Paulsen YR (2012) Guidelines for estimating repeatability. Methods in Ecology and Evolution 3: 129–137. https://doi.org/10.1111/j.2041-210X.2011.00125.x

Wolf M, Friggens M, Salazar-Bravo J (2009) Does weather shape rodents? Climate related changes in morphology of two heteromyid species. Naturwissenschaften 96: 93–101. https://doi.org/10.1007/s00114-008-0456-y

Wright S (1943) Isolation by Distance. Genetics 28: 114–138. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1209196.

Yağcı T, Gurbanov RR (2019) The impact of the different habitats on skull variation in the fossorial subterranean rodents (Rodentia: Spalacidae) from Middle Anatolia. Ecological Research 34: 802–812. https://doi.org/10.1111/1440-1703.12048

Yigit N, Çolak E, Verimli R (2001) A Study on the Distribution, Morphology and Karyology of *Tatera indica* (Hardwicke, 1807) (Mammalia: Rodentia) in Turkey. Turkish Journal of Zoology 25: 67–70.

Appendix 1

Geographic localities of the 509 *T. indica* specimens sampled in this study. The voucher number of each specimen is indicated. In instances where museum databases and/or vial labels had spelling mistakes or outdated locality information (i.e., the name of a county changed), locality information was unified, with the original spelling or locality information between square brackets.

Afghanistan

Kandahar, Herat Highway [Herat-Kandahar Rd], 5 mi S (FMNH 103249). Zaranj [Qala-i-Kang], 10 mi S (FMNH 103283, 103284). Jalalabad Airport (USNM 602945, 602946, 602948, 602969). Jalalabad, 25 mi N (FMNH 103253, 103254, 103255, 103256, 103257, 103258, 103259). Kandahar International Airport [Air Field] (USNM 600267, 600268, 600269, 600270, 600271, 602166, 602168, 602169, 602258, 602419, 602420, 602421, 602422, 602423, 602424, 602425, 602427, 602430, 602987, 602988, 603015). Kandahar, 4 mi N (FMNH 103260, 103262, 103268, 103273, 103274, 103275, 103278, 103279).

India

Delhi (FMNH 28986). Dharwad [Dharwar] (FMNH 83067, 83068). Khajuraho (AMNH 250026). Kurumbapatti (FMNH 83065, 83066). Mollasimla, 2 mi E Singur F S Dist Hoogly (FMNH 104617). New Delhi (USNM 319045). Parola [E Khandesh Parola] (FMNH 83063, 83064). Singur (USNM 355835, 355836).

Iran

Bam, 55 km E (USNM 329171, 329202, 329203, 329204, 329205, 329206, 329207, 329208, 329209, 329211, 329212, 329219, 329220, 329221). Bandar Abbas, 5 km E (USNM 329200, 329201, 329214, 329215, 329216, 329217, 329230, 329231, 329233). Esfedan [Isfehdeh] (USNM 328292). Esfedan [Isfehdeh], 3 km SW (USNM 328279, 328282, 328284, 328285, 328287, 328288, 328300, 328301, 328302, 328303, 328304, 328305, 328306, 328307). Esfedan [Isfehdeh], 5 km NE (USNM 328291, 328308, 328309, 328310, 328311, 328312, 328313, 328314, 328315). Fahraj (FMNH 97180). Haji Abad [Hadji-Abad] (USNM 329173, 329175, 329192, 329193, 329194, 329195, 329196, 329197, 329198, 329199, 329213, 329222, 329223, 329224, 329225, 329229). Iranshahr (FMNH 97178). Iranshahr, 30 km N (USNM 328294, 328316, 328317, 328319, 328320, 328324, 328325, 328326, 328327, 328328, 328329, 328330, 328331, 328332, 328333). Iranshahr, 8 mi W (FMNH 97176). Kerman (MVZ 198834). Kerman, 8 km NNW (MVZ 191993, 191998, 191999, 192000, 192001, 192002, 192004, 192005, 192006, 192007, 192012, 192013, 192014, 192015, 192024). Minab (FMNH 111939). Qasr-e Shirin [Qasr-I-Shirin] (FMNH 97207). Zabol, 15 mi SW (FMNH 97159, 97161, 97162, 97164, 97165, 97166, 97167, 97168, 97169, 97170, 97175). Zabol, SE (MVZ 191994, 191995, 191996, 191997). Kohban Ward Gwadar Old City [Gwadar] (UF 30260, 30266). Pasni [Gwadar] (UF 30245). Ahram (FMNH 97194, 97195, 97196, 97197, 97198, 97199, 97200, 97201). Ahvaz [Ahwaz], 45 km N (USNM 350543, 350544, 350545, 350546, 350547, 350548, 350550, 350551, 350552, 350553, 350554, 350555, 350556). Ahvaz [Ahwaz], 53 mi SE (FMNH 111934). Andimeshk, 30 km S (USNM 350521, 350522, 350523, 350524, 350525, 350526, 350527, 350528, 350529, 350530, 350531, 350532, 350533, 350534, 350535, 350536, 350537, 350538, 350539, 350540, 350541). Behbahan [Beahbehan], 80 km SE (USNM 350578). Behbahan [Behbehan], 93 km ESE (USNM 350580, 350581, 350514, 350515, 350516, 350517, 350518, 350519). Beriz [Bariz] (USNM 329179, 329180, 329218, 329232, 329234, 329236, 329237, 329238, 329239). Darab, 11 km NW (USNM 369682, 369683, 369684, 369685, 369686, 369687, 369689, 369690, 369691, 369692, 369693). Gachsaran [Gach Saran], 20 mi NW (FMNH 111925, 111926, 111927, 111928, 111929, 111930). Gara [Sach Gara Gul] (USNM 297641, 297642). Jahrom (FMNH 111935, 111936, 111937, 111938). Kazerun, 10 km SE (USNM 350573, 350574). Kazerun, 7 km N (FMNH 97190, 97191, 97192, 97193). Lar, 4 km N (USNM 369708, 369709, 369710, 369711, 369713, 369714, 369715). Mansoor Abad [Mansorabad], 19 km S (USNM 369716, 369718, 369719, 369720). Mansoor Abad [Mansorabad], 2 km SE (USNM 369696, 369697, 369698, 369700, 369701, 369702, 369703, 369704, 369705). Pol Abgineh [Pol-i-Abgineh] (FMNH 97183, 97185, 97186). Pol Abgineh [Pol-i-Abgineh], 5 km ESE (FMNH 97182). Ramhor-

moz [Ram Hormuz] (FMNH 111932). Shush, 12 mi S (FMNH 97202, 97203, 97204). Taj Maleki (USNM 350579).

Kuwait

Kuwait City [Al Kuwait] (USNM 282345, 282346, 282347, 282348, 282349, 282350, 282351).

Nepal

Tikapur (USNM 290081).

Pakistan

Amandara (USNM 413595). Ayub National Park (USNM 326456, 326457, 326458, 326460, 326461, 326462, 326463, 326464, 326465, 327104, 327105, 327106, 327107, 327108, 327109, 353675, 353676, 353677, 353678, 353679). Bahawalpur [Bahawalapur], 3 mi N (USNM 353653, 353654, 353655, 353673). Bela, 8 km S (USNM 369463). Changa Manga Forest Park [Reserve] (USNM 369097, 369098, 369099, 369100, 369101, 369104, 369105, 369107, 369108, 369128, 369129, 369130, 369131, 411057, 411058, 411059, 411060, 411061, 411062, 411063, 411064, 411065, 411066, 411067, 411068, 413582, 413590, 413591, 413592, 413593, 413594). Charwa [Charwa Village] (USNM 326451, 326452, 326453, 326454, 326455, 353640, 353641, 353642, 353643, 353644). Daira [Daera] Din Panah (USNM 411075, 411076, 411077, 411078). Dak Bangalow [Bungalow] (USNM 353635, 353636, 353637, 353638, 353639). Dera Ghazi [Ghasi] Khan, 3 mi S (USNM 353669). Dera Ghazi Khan, 2 mi S (USNM 369136, 369137). Dera Ghazi Khan, 2 mi SSE (USNM 369134, 369135). Dera Ghazi Khan, 3 mi W (USNM 369133). Dera Ghazi Khan, 6 mi W (USNM 369132). Dera Ismail Khan (USNM 413585, 413587, 413588, 413589). Fort Abbas, 10 mi NW (USNM 369109, 369110, 369111, 369112, 369113, 369114, 369115, 369116, 369117, 369119, 369120, 369121, 369122, 369123, 369124, 369125, 369126, 369127). Fort Abbas, 10 mi W (USNM 353652, 353670, 353671, 353672). Gizri (AMNH 217325, 217326, 217327, 217328, 217329). Jahangir's [Jehangir's] Tomb (USNM 326447, 326448, 326449, 326450). Karachi (USNM 13666). Karachi, 70 km N (USNM 369464). Lahore (USNM 353624). Lahore District Ravi River (USNM 369154). Lahore, 14 mi NE (USNM 369093, 369094, 369095, 369096). Lahore, 25 mi SW (USNM 369087, 369088, 369089, 369090, 369091). Loralai [Loralai Town] (USNM 413604, 413605, 413606). Loralai, 2 mi N (USNM 413601). Loralai, 3 mi E (USNM 413602). Mansehra [Manshera] (USNM 369155, 369156, 369158). Mian Channu (USNM 353649, 353650, 353651). Mustafa Abad [Lulliani] HWY, 10 mi (USNM 353627). Mustafa Abad [Lulliani], 4 mi NW on Bari Doab (USNM 353628, 353629). Muzaffargarh (USNM 369140, 369141, 369142, 369143). Muzaffargarh, 1 mi N (USNM 369138, 369139, 369144). Muzaffargarh, 2 mi E on Multan Rd (USNM 369148, 369149, 369150). Muzaffargarh, 4 mi E (USNM 369147). Muzaffargarh, 5 mi S on Alipur [Ali Pur] Rd (USNM 369145). Muzaffargarh, 9 mi S (USNM 369146). Nushki, 3 mi W (USNM 411086, 411089). Panjnad Canal, 2 mi downstream of Panjnad Headworks (USNM 353674). Parachinar (USNM 353681). Ravi Rd (USNM 353623). Rawalpindi, 16 mi NE on Murree Rd (USNM 353680). Rawalpindi, 16 mi NW (USNM 353630). River Ravi Motorway Bridge [Ravi River Bridge] (USNM 353625, 353626). Sakhi Sarwar (USNM 411079, 411080, 411081, 411082, 411083, 411084). Shadan Lund (USNM 411069, 411070, 411071, 411072, 411073, 411074). Shahpur Kanjra [Shah Pur Village], along Bari Doab Canal (USNM 413583, 413584). Uch Sharif (USNM 353645, 353646, 353647, 353648, 353661, 353662, 353663, 353664, 353665, 353666, 353667, 353668). University of The Punjab [New University of The Panjab] (USNM 369151, USNM 369152).

Sri Lanka

Sri Lanka centroid (AMNH 150073, USNM 277235). Hambantota (FMNH 83069, 83070). Kankesanturai (AMNH 240846, 240847, 240848, 240849, 240850). Uva Province [Uva South, Handikema] (FMNH 92217, 92218, 92219, 92220, 92221). Wariyapola (AMNH 240852).

Supplementary material 1

Figure S1

Authors: Dashti Z, Alhaddad H, Alhajeri B (2022)
Data type: .jpg
Explanation note: The full range of *Tatera indica* according to the International Union for Conservation of Nature (IUCN) Red List (Kryštufek et al. 2017).
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/vz.72.e90474.suppl1

Supplementary material 2

Figure S2

Authors: Dashti Z, Alhaddad H, Alhajeri B (2022)
Data type: .pdf
Explanation note: Scatterplot of PC2 and PC3 of the PCA conducted on the 21 logged measurements. Specimens are divided into the four geographic groups (see legend). The amount of explained variance by each PC axis is indicated in parentheses. The plot was generated using *ggplot2*.
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/vz.72.e90474.suppl2

Supplementary material 3

Figure S3

Authors: Dashti Z, Alhaddad H, Alhajeri B (2022)
Data type: .pdf
Explanation note: RDA plots. The plots of the specimens and the measurements on the plane defined by the first two RDA axes shown for the partial RDA model that includes climatic explanatory variables and spatial conditioning variables ("climate | spatial") and non-sheared (a, b) and sheared (c, d) morphometric measurements as the response variables. The amount of variation explained by each RDA axis as proportion of the total variation explained by each RDA model (proportion explained of accumulated constrained eigenvalues) is as follows. For the non-sheared morphometric variables, RDA axis 1 accounts for 87.7% of the variation while axis 2 accounts for 7.7% of the variation. For the sheared morphometric variables, RDA axis 1 accounts for 44.1% of the variation while axis 2 accounts for 17.3% of the variation. The significance of each RDA axis was tested using forward tests for axes with 1000 permutations (non-sheared axis 1: $F = 56.00$, $p < 0.050$; non-sheared axis 2: $p = 0.183$; sheared axis 1: $F = 11.08$, $p < 0.050$; sheared axis 2: $F = 4.35$, $p < 0.050$; sheared axis 3: $p = 0.088$). For all four plots, RDA axis 1 is on the x-axis, while RDA axis 2 is on the y-axis, and specimen scores are weighted sums of measurements. In (a) and (c) the specimens are color coded based on their geographic group membership, while crosses (+) in each plot correspond to the 21 measurements. In (b) and (d) the arrows are the explanatory climatic variables after removing the effect of the spatial covariables, shown on the same plane as plots (a) and (c). See the main text for more information about the bioclimatic variables. Plots were generated using *ggord* library (Beck 2019).

Link: https://doi.org/vz.72.e90474.suppl3

Supplementary material 4

Table S1

Authors: Dashti Z, Alhaddad H, Alhajeri B (2022)
Data type: .docx
Explanation note: Descriptions of the morphometric measurements used in this study. The acronyms match those in Figure 1.
Link: https://doi.org/vz.72.e90474.suppl4

Supplementary material 5

Table S2

Authors: Dashti Z, Alhaddad H, Alhajeri B (2022)
Data type: .docx
Explanation note: Supplemental methods.
Link: https://doi.org/vz.72.e90474.suppl5

Supplementary material 6

Table S3

Authors: Dashti Z, Alhaddad H, Alhajeri B (2022)
Data type: .docx
Explanation note: Bioclimatic variables used as an index of climate in this study. These variables are based on monthly temperature and precipitation values, and show annual trends, seasonality, and extremes. The descriptions are from: https://www.worldclim.org.
Link: https://doi.org/vz.72.e90474.suppl6

Supplementary material 7

Table S4

Authors: Dashti Z, Alhaddad H, Alhajeri B (2022)
Data type: .docx
Explanation note: Summary of forward stepwise variable selection. The most parsimonious combinations of explanatory variables in each of the climatic and the spatial dataset are indicated in their order of entry into the model. Variable selection stopped either when the significance alpha level of 0.05 was reached or when the adjusted coefficient of multiple determination (R^2_{adj}) value of the reduced model reached that of the global model.
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/vz.72.e90474.suppl7

Supplementary material 8

Data S1

Authors: Dashti Z, Alhaddad H, Alhajeri B (2022)
Data type: .cvs
Explanation note: Data matrix used for all the analyses in this study. For each specimen, the catalog number is indicated in the first column, followed by its sex (F = female, M = male, blank = unknown), the country to which it belongs, the most precise collection locality available for the specimen (i.e. specific locality, county, state, province, etc.), the basis for the geographic coordinates (see "Materials and methods"), the latitude and longitude of the locality, and the geographic group of each locality based on proximity, mostly along the longitude (see "Materials and methods"). The locality information matches Appendix 1. The localities and their groups are shown in Figure 2. The data matrix also shows all 21 linear measurements (average of the two trials, in mm, untransformed [raw]). Measurement acronyms are described in Table S1 and visualized in Figure 1. The following columns show the 19 extracted bioclimatic variables (see Table S3 for a description) along with the measurements' PC1 used for "shearing" size correction (see "Materials and methods"). Blank cells indicate missing data.
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/vz.72.e90474.suppl8

Supplementary material 9

Data S2

Authors: Dashti Z, Alhaddad H, Alhajeri B (2022)
Data type: .docx
Explanation note: Pairwise Pearson's correlation coefficients (r) among the 21 logged measurements. Measurement acronyms are described in Table S1.
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/vz.72.e90474.suppl9