Journal of Animal Ecology 2002
71, 1015–1029

Peak morphological diversity in an ecotone unveiled in the chukar partridge by a novel Estimator in a Dependent Sample (EDS)

SALIT KARK*†, TAPAN MUKERJI‡, URIEL N. SAFRIEL†§, IMANUEL NOY-MEIR¶, RACHEL NISSANI** and ARIEL DARVASI†

**Department of Biological Sciences, Stanford University, Stanford, CA 94305–5020, USA; †Department of Geophysics, Stanford University, Stanford, CA 94305, USA; ‡Department of Evolution, Systematics and Ecology, The Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel; §Jacob Blaustein Institute for Desert Research, Sede Boqer Campus, Ben-Gurion University of the Negev, 84990, Israel; ¶Department of Agricultural Botany, The Faculty of Agricultural, Food and Environmental Quality Sciences, The Hebrew University of Jerusalem, Rehovot 76100, Israel, **Clinical Genetics Unit, Kaplan Hospital, Rehovot 76100, Israel*

Summary

1. Areas of environmental transition (i.e. ecotones) have recently been shown to play an important role in the maintenance of genetic diversity, divergence and in speciation processes. We test the hypothesis that ecotone populations maintain high phenotypic diversity compared to other populations across the distribution range.
2. Focusing on the chukar partridge (*Alectoris chukar* Gray), we study trends in morphological diversity across a steep ecotone within the species native range in Israel and Sinai. Using 35 traits and 23 ratios between traits, we apply a novel weighted average statistic that we term 'Estimator in a Dependent Sample' (EDS). This estimator enables us to compare levels of diversity across populations using multiple-correlated traits and is especially useful when sample sizes are small.
3. We provide a program for calculating the EDS and a bootstrapping procedure to describe its confidence interval and standard deviation. This estimator can be applied widely in a range of studies using multiple-correlated traits in evolutionary biology, ecology, morphology, behaviour, palaeontology, developmental biology and genetics.
4. Our results indicate that within-population diversity peaks in chukar populations located in the Mediterranean-desert ecotone in Israel. However, had we not included the ecotone region in our study, we would have drawn different conclusions regarding patterns of morphological diversity across the range. We suggest that ecotones should be given higher priority in future research and conservation planning, potentially serving as within-species diversity hotspots.

Key-words: *Alectoris chukar*, biodiversity hotspots, ecotone, Estimator in a Dependent Sample (EDS), morphological diversity.

Journal of Animal Ecology (2002) **71**, 1015–1029

Introduction

The recent increase in the rate of habitat loss, alien species' invasion, native species' extinction and the resulting decline in both among and within-species biodiversity are requiring scientists and decision makers to prioritize conservation efforts and budgets (Risser 1995; Mace *et al*. 2000; Myers *et al*. 2000; Balmford *et al*. 2001). A central approach at the community level focuses attention towards areas that maintain high species diversity, i.e. diversity hotpsots (Myers *et al*. 2000). At the within-species level, a similar approach could be taken, emphasizing regions that sustain populations with especially high levels of within-species genetic and phenotypic diversity.

Correspondence: Dr Salit Kark, Department of Evolution, Systematics and Ecology, The Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel. Tel: 972-2-6585075. Fax: 972-2-6584741. E-mail: salit@mail.ls.huji.ac.il

Detecting such areas with high potential for generating and maintaining diversity within-species ranges is an important first step in achieving this goal (Myers *et al*. 2000; Smith *et al*. 2001). Populations located in transitional environments and along steep ecological gradients have been suggested recently as important centres for generation and maintenance of genetic diversity (Kark *et al*. 1999) and morphological divergence (Kark *et al*. 1999; Smith *et al*. 2001). These ecological transition zones (i.e. ecotones) have been proposed as playing a key role in speciation processes (Schilthuizen 2000). In areas of climatic and environmental change, many species approach their continuous distribution limits (Safriel *et al*. 1994; Danin, Arbel & Levy 1998). A recent study focusing on patterns of genetic (allozyme) diversity in a bird species, the chukar partridge, has shown that genetic diversity peaks in the Mediterranean-desert ecotone region in Israel (Kark *et al*. 1999; Kark *et al*. 2001). In this area, the species reaches the edge of its continuous distribution.

In this study we test the hypothesis that ecotone populations maintain especially high morphological diversity, in addition to their high genetic diversity. More specifically, using a novel Estimator in a Dependent Sample (EDS), we test the hypothesis that within-population morphological diversity peaks in the Mediterranean-desert ecotone of Israel compared with other areas across the distribution range. We focus on the chukar partridge (*Alectoris chukar*), which is distributed along the steep environmental gradient in Israel and in Sinai, Egypt. The region comprises a narrow land bridge between Europe, Asia, and Africa. Steep climatic and ecological gradients occur in the region within relatively short distances (Bitan & Rubin 1991; Danin *et al*. 1998). The region offers a unique opportunity to test our research hypothesis comparing levels of diversity across a species range within geographically proximate populations that are potentially connected by dispersal and gene flow, yet experience extremely different environments and ecological conditions.

A range of studies in ecology, evolution, morphology, behaviour and various other fields (e.g. palaeontology, developmental biology and genetics) use multiple-correlated traits (or loci) to estimate diversity across populations or groups. Common estimates of variation within populations, including the variance, standard deviation (SD) and coefficient of variation (CV) have a skewed distribution. Therefore, comparing them across populations using a single trait often requires much larger sample sizes than available (in the order of hundreds per population) in order to detect statistically significant differences between populations (Zar 1999). This is often difficult to achieve in studies of natural populations, and especially when focusing on vertebrates or on endangered species. In addition, when multiple statistical tests are repeated for many traits, the *P*-value should be corrected for multiple comparisons (e.g. using the sequential Bonferroni correction; Rice 1989). This results in the fact that it is more difficult to reject the null hypothesis in every single test. Thus studies that use many correlated traits may lack the power to detect differences in levels of within-population diversity across study populations. In addition, multiple traits may be correlated due to allometry (Soulé & Cuzin-Roudy 1982) and should not be treated independently.

Developing scientific approaches that enable us to compare diversity quantitatively is a crucial step in understanding spatial diversity patterns. At the morphological level, several approaches complement each other and may assist in achieving this goal. The first approach has recently been receiving substantial attention (e.g. Schneider *et al*. 1999) and focuses on comparing levels of morphological divergence across regions, where divergence is 'the evolution of increasing differences between lineages in one or more characters' (Futuyma 1997). A second direction focuses on comparing within-population diversity across populations or groups. When following this latter approach, the variance, coefficient of variation or other estimate of variability, rather than mean trait values, are compared across populations. Such approaches are somewhat equivalent to those used at the genetic level, where studies often include a sample of multiple loci (e.g. allozyme, microsatellite) used to assess genome-wide variation. While the two approaches (i.e. calculating divergence and diversity) are often used interchangeably at the within-species level, they each deal with a different aspect of biodiversity, both being important. Statistical methods for the study of morphological divergence using multiple traits and for reducing the dimensionality of a set of data while trying to preserve the structure, have received much attention (e.g. principal components analysis (PCA)). However, detecting divergence was not our goal in this study. Our goal here is to compare levels of within-population diversity across a species range using information from all traits measured taking into account their correlation and to come up with one diversity estimate per population (rather than to compare trends in single traits). To achieve this goal, we developed and applied a statistical estimator for within-population morphological diversity in multiple-correlated traits. It is estimated in terms of variance, standard deviation (SD), coefficient of variation (CV) or any other estimate of variation within a group. This simple statistic is especially useful when data from many, often correlated traits are available, or when sample size is too small to detect significant trends across populations if single traits are analysed. We provide two versions of the statistic, which we call Estimator in a Dependent Sample (EDS), the first using the absolute value of the correlation coefficient (r) between traits and the second using the coefficient of determination (r^2). We use the EDS to test our research hypothesis that morphological diversity in the chukar partridge peaks in the ecotone region in Israel.

Materials and methods

CHUKAR DISTRIBUTION AND STUDY AREA

The northern more mesic regions of Israel are characterized by Mediterranean climate, with over 450 mm mean annual rainfall (up to 1000 mm), precipitated largely during the winter. The southern Negev Desert and Arava Valley, although geographically close to the above regions, are ecologically very different. They are arid, with a mean annual rainfall of ≤ 100 mm (Bitan & Rubin 1991; Israel Meteorological Service, unpublished data). A sharp climatic gradient occurs at the Mediterranean-desert ecotone of the southern Shefela-northern Negev in Israel (Kadmon & Danin 1997; Kark 1999). In this area of transition between Mediterranean and desert ecosystems, rainfall decreases from over 450 to less than 100 mm within several dozen km, and between-years rainfall variability is high (Bitan & Rubin 1991).

Native to Israel and the region, the chukar generally inhabits the mesic and semi-arid areas and has large populations in Mediterranean and steppe environments (Shirihai 1996). The Mediterranean-arid ecotone in the northern Negev is the edge of the chukar continuous distribution (Kark *et al*. 1999). In the central and southern Negev and in the Sinai desert, where annual rainfall is < 100 mm and is highly variable among the years, chukar density decreases, distribution becomes discontinuous, and local populations become small and patchy (Pinshow, Degen & Alkon 1983; Degen, Pinshow & Shaw 1984; Shirihai 1996). This area comprises the extreme periphery of the chukar global distribution range. Chukars do not possess physiological adaptations to heat stress (Carmi-Winkler, Degen & Pinshow 1987; Frumkin 1983; Kam 1986), as opposed to the partly sympatric sand partridge (*Ammoperdix heyi*), which is well adapted to the desert environment (Degen, Pinshow & Alkon 1983). A main limiting factor for chukars in the desert is their ability to forage long enough without risking their heat balance. Extremely high temperatures limit foraging activity to levels insufficient for their energy demands (Carmi-Winkler *et al*. 1987). Therefore, in these marginal habitats the species occurs discontinuously, and is generally limited to temporary food and water patches (Shirihai 1996). These habitat patches must be rich enough to suffice for the birds' energetic needs in the short available daily foraging time, which is limited to the early morning (Degen *et al*. 1984; Carmi-Winkler *et al*. 1987). They must provide sufficient water during the hot and dry summer months (Carmi-Winkler *et al*. 1987; Degen *et al*. 1984). An isolated population is found in the southern Sinai mountain region and is probably a relict from the glacial periods of the Upper Pleistocene when chukar distribution extended more continuously into areas that are today arid deserts in the Negev and the Sinai Peninsula (Nissani 1974; Yom-Tov & Tchernov 1988).

STUDY POPULATIONS AND DATA SET

Populations were studied along the climatic gradient ranging from the mesic Mediterranean areas in the Upper Galilee through the Mediterranean-desert ecotone (Northern Negev area), south to the isolated population of Mount Sinai region at the very extreme periphery of the species range. The populations (listed from north to south) used for this study include: Upper Galilee, Jerusalem Mountains, Northern Negev, Negev Highlands and Southern Sinai Mountains (Table 1). The Northern Negev population is located in the Mediterranean-desert ecotone. Sample size in each population ranged from 26 to 36 individuals (sexes pooled) for a total of 158 chukars (Table 1). In this work we use extensive morphological data collected by one of us (RN; Nissani 1974) from populations of chukars sampled in Israel and Sinai between 1971 and 1973. Thirty-five morphological traits were measured in all individuals and 23 ratios between traits were calculated by Nissani (1974). The early study tested a different hypothesis (Nissani 1974), and aimed to examine whether Bergmann's and Allen's ecogeographical rules are supported across the gradient

Table 1. Description of chukar study regions (with acronym used in Figures), location of population studied, environmental description (based on Danin *et al*. 1998), and sample size for males and females

Region (acronym)	Location	Environment	Latitude & longitude	Sample size	
				Males	Females
Eastern Upper Galilee (UG)	Yiftach	Mediterranean maquis; Localized deciduous orchards	33°06′N 35°33′E	11	18
Jerusalem Mountains (JM)	Ness Harim	Mediterranean maquis; Localized deciduous orchards	31°44′N 35°03′E	16	18
Northern Negev (NN)	Mishmar Hanegev	Open semisteppe bathas with croplands	31°23′N 34°41′E	17	16
Negev Highlands (NH)	Sede Boqer	Arid steppes with some orchard oases	30°52′N 34°47′E	11	25
Southern Sinai Mountains (SM)	Saint Catherine	Desert vegetation of hyper-arid mountains limited to wadis with small orchard oases	28°33′N 33°57′E	13	13

(Begon, Harper & Townsend 1996). Therefore, Nissani compared population means for the different traits measured. However, in this study we are not interested in comparing change in means but rather we examine within-population diversity patterns across the ecotone in Israel and Sinai. In order to test this, we digitized the original hand-written data sheets and use the raw measurement data rather than the analyses done by Nissani. Only adult birds were included. For detailed description of traits, ratios and measurement methodology see Appendix 1.

CALCULATION OF THE ESTIMATOR IN A DEPENDENT SAMPLE (EDS)

When estimating diversity in a sample of traits that are statistically dependent, observations need to be weighted by their degree of independence. For example, when two traits are studied, their values may share a certain component that is reflected in the correlation between them. Each trait has a shared portion with the other trait and an unshared portion. In the case of two traits, when calculating the degree of variation in a population using both traits, we should count the shared and unshared portions only once each. Alternatively, if one considers the traits independently, the shared portion will be counted twice and thus will be over-represented compared to the unshared portion. As an extreme example, two traits that are 100% correlated will then be counted twice as though they were independent. In order to avoid this in a study that includes two or more traits, we propose that the correlation between each two traits in the sample can be calculated to derive weights for a weighted average statistic. The weighting parameter is the degree of independence of each trait, and is estimated by: $w_i^{abs} = \frac{1}{2} + \sum_{j=1}^{N}\left(1 - \left|\frac{r_{ij}}{2}\right|\right)$ for trait i where r_{ij} is the correlation coefficient (Pearson, Spearman or other) between traits i and j. The sum over j runs from 1 to the number of traits (N).

The weighted average is then:

$$\mathrm{EDS}_{\mathrm{abs}} = \frac{1}{\sum_{i=1}^{N} w_i^{abs}} \sum_{i=1}^{N} a_i w_i^{abs}$$

where N is the number of traits, a is a variability measure (e.g. coefficient of variation, standard deviation), and EDS stands for Estimator in a Dependent Sample. The weighted sum of the a_i values in the numerator is normalized by the sum of the weights in the denominator. While it may seem intuitive to want to weigh with the inverse of the correlation, doing so can cause the weights to blow up for 0 correlation.

Alternatively, the coefficient of determination, r^2, can be used instead of the absolute value of r giving:

$$w_i^{sq} = \frac{1}{2} + \sum_{j=1}^{N}\left(1 - \frac{r_{ij}^2}{2}\right) \text{ and}$$

$$\mathrm{EDS}_{\mathrm{sq}} = \frac{1}{\sum_{i=1}^{N} w_i^{sq}} \sum_{i=1}^{N} a_i w_i^{sq}.$$

Accordingly, for each trait i the unshared portion of a_i (which is $1 - |r_{ij}|$) with respect to j is counted once, while the shared portion of $|r_{ij}|$ is taken as the average of the a_i and a_j values, i.e. half is taken from trait i and half from trait j. Thus for trait i we have $(1 - |r_{ij}|) + (|r_{ij}|/2) = 1 - (|r_{ij}|/2)$. To take care of the specific case where $i = j$ (i.e. when trait i is weighed against itself and $r_{ii} = 1$) we add the constant $^1/_2$. This is because ii is counted once while ij occurs twice (as ij and as ji). Thus for a simple example of three traits (1, 2, 3, for instance) with two being perfectly correlated and one completely uncorrelated ($r_{12} = r_{21} = 1$; $r_{31} = r_{32} = 0$; and $r_{11} = r_{22} = r_{33} = 1$), the weights would be:

$$w_1 = w_2 = \frac{1}{2} + [(1 - 1/2) + (1 - 1/2) + 1] = 2.5$$
$$w_3 = \frac{1}{2} + [1 + 1 + (1 - 1/2)] = 3.$$

The EDS gives a weighted estimate of within-population diversity, using information based on all individuals and all traits measured within a population, and on their correlation matrix. The estimator is calculated for each population separately. This estimator may be useful in various analyses using multiple-correlated phenotypic traits, especially when the required emphasis is not on the absolute numerical values for each trait, but rather on the relative comparison of estimates from different species, populations or groups (e.g. sexes, social classes and age groups) or when testing for spatial and temporal trends in phenotypic variability. In order to avoid substantial complication in the calculation of the statistic, the equation corrects for the correlation between each pair of traits i, j at a time rather than for simultaneous correlations in a multidimensional space. This is perhaps a partial compensation, yet is better than not compensating at all. The fewer traits used, the higher is the degree of compensation. The full compensation will require data that are not available on the multidimensional correlation space. Under certain assumptions this can be achieved, but the calculation becomes very complicated and cumbersome.

As mentioned, for calculating EDS, various variability measures, including the variance, standard deviation and the coefficient of variation can be used. We prefer the coefficient of variation when comparing natural populations, as it is independent of absolute differences in trait means (Soulé & Cuzin-Roudy 1982) detected commonly along climatic and ecological gradients (Begon *et al*. 1996). Due to a trend of decreasing body size of chukars across the geographical gradient studied here (Shirihai 1996), we use the coefficient of variation to calculate EDS. Both $\mathrm{EDS}_{\mathrm{abs}}$ and $\mathrm{EDS}_{\mathrm{sq}}$ were calculated for the 35 traits and for the 23 ratios. We used the Pearson correlation coefficient for the calculation of the EDS. Males and females were separated in the analysis due to differences in life histories and social structure (Alkon 1974), which may affect their

morphological diversity patterns. Thus, the EDS is separately obtained for birds of each sex within each population, representing the weighted morphological diversity value over all traits or ratios. Although the analyses for each of the two sexes and for the traits and ratios were performed separately, these are probably not independent and so should not be treated as such.

BOOTSTRAPPING

In order to obtain statistics for the variation of the EDS statistic for each population and sex, including confidence intervals and standard deviations, we used bootstrapping techniques (Efron & Tibshirani 1998). The bootstrap procedure enabled us to compute distributions of the EDS statistics and 95% confidence intervals. The bootstrap method performs data-based simulation, which allows assessment of the accuracy of complicated estimation procedures using the power of computation (Efron & Tibshirani 1998). This is especially useful for sparse data. We used bootstrap in a non-parametric mode, which avoids restrictive and sometimes erroneous assumptions about the form of the underlying populations (Efron & Tibshirani 1998). We computed the EDS statistic for each bootstrap sample. We used for each bootstrap run a similar data structure as in the original data set (i.e. similar population sizes were drawn with replacement). This was performed for each of the two EDS types in both traits and ratios for each of the two sexes, giving a total of eight bootstrapping procedures. Typically 1200 bootstrap samples were drawn. In each run, we checked for missing values to ensure that the bootstrap EDS statistic could be computed with the drawn sample. Occasionally some of the samples had to be rejected because of missing data, which did not allow computation of the statistic. Overall, after rejection, the bootstrap sample size was larger than 1000 in all cases. The results of the bootstrap runs were then used to draw the distribution and to compute confidence intervals and ratio tests on the EDS statistic (Efron & Tibshirani 1998).

STATISTICAL SOFTWARE

The EDS calculations and the bootstrapping procedures are provided in a MATLAB program that we wrote for this work (Kark & Mukerji 2002), available in Appendix 2 and at URL http://www.stanford.edu/group/Mooney/salit/#Download the EDS. The program includes two files: calculationEDS.m and bootEDS.m; both should be downloaded.

HYPOTHESIS TESTING

To test the research hypothesis that the ecotone population of the Northern Negev originating from Mishmar Hanegev (NN) has higher EDS than any of the other populations sampled in the study, we followed several approaches. First, we applied a one-tailed Scheffé's multiple contrasts test (Miller 1991; Zar 1999), using the standard deviations obtained in the bootstrapping (Efron & Tibshirani 1998). Our null hypothesis was:

$$\mu_{NN} - {}^{1}/_{4}\,(\mu_{UG} + \mu_{JM} + \mu_{NH} + \mu_{SM}) = 0$$

(see Table 1 for population acronyms). In this test the sum of the coefficients of μ is 1 (Zar 1999). We tested this hypothesis for each of the two EDS types (EDS_{abs} and EDS_{sq}, see above) in both traits and ratios for males and females separately. This gave us a total of eight comparisons. Due to multiple comparisons, we applied a sequential Bonferroni correction (Rice 1989) to the calculation of the statistical significance (*P*-value) of the contrasts.

A second approach for testing the significance of the differences between the EDS of the ecotone vs. all other populations used the proportion between the EDS values. We computed the statistic R, which represents the ratio between EDS_{NN} (ecotone) and EDS_{OTHER} (all other populations). If indeed the ecotone has significantly higher levels of diversity, as suggested by our research hypothesis, R is predicted to be greater than 1. So we tested the null hypotheses $H_0 : R \leq 1$. This was performed using the values from the bootstrap calculations to compute a one-tailed lower confidence interval at the $P = 0{\cdot}05$ level.

In addition, we examined the proportion of the total bootstrap runs in which the ecotone population had the highest EDS level. The null expectation is that the ecotone population will have the highest EDS levels 20% of the runs (being one out of five populations) in each of the cases (male and female traits and ratios).

Results

In all eight cases, including traits and ratios in males and females for both EDS estimates (EDS_{abs} and EDS_{sq}), within-population morphological diversity, as estimated by the EDS value, peaked in the Northern Negev ecotone population compared to the other four populations studied across the distribution range (Fig. 1). Scheffé's multiple contrasts test was significant after a sequential Bonferroni correction in all eight cases ($P < < 0{\cdot}0001$), rejecting the null hypothesis (see Methods above), and suggesting significantly higher EDS levels in the ecotone population compared to all other populations in the study.

We used the EDS data from the bootstrapping to examine the ratio between EDS of the ecotone and the other populations combined, testing the null hypothesis $H_0 : R \leq 1$. The one-tailed ($P = 0{\cdot}05$) confidence intervals for both male ratios and traits (using EDS_{abs}) were larger than 1 (1·01 and 1·07, respectively). Hence the null hypothesis of no differences between ecotone and other populations can be rejected at the 0·05 level.

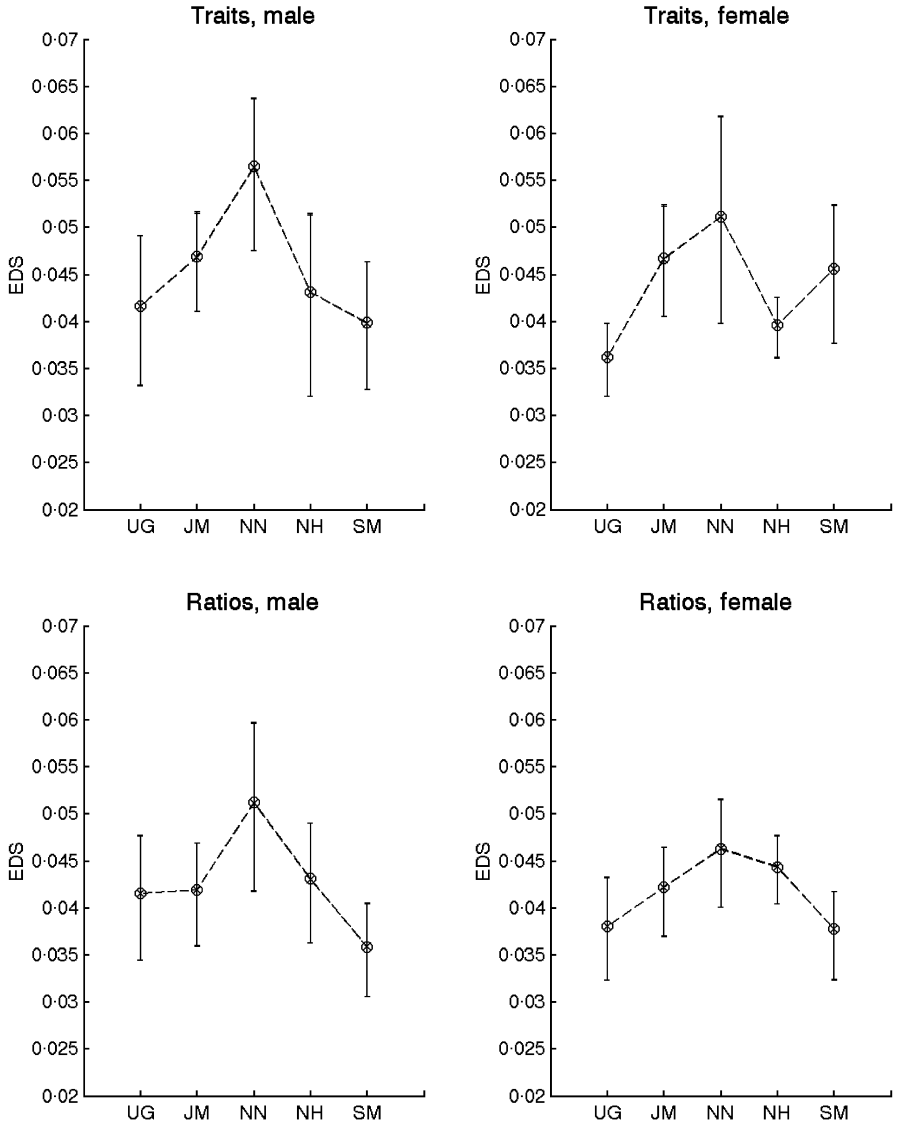

Fig. 1. Trends in within-population morphological diversity in chukar (*Alectoris chukar*) populations in Israel and Sinai, as estimated based on 35 traits or 23 ratios using the Estimator in a Dependent Sample (EDS) calculated with the coefficient of variation. Population acronyms are shown in Table 1. Populations are shown from north to south. Note the high levels of diversity in the NN population, which originated from the ecotone region (see Table 1 for details on location). Mean EDS in sampled population and 95% confidence intervals calculated using the bootstrap procedure are shown for males and females, traits and ratios, and for the two forms of EDS, which were very similar (where O – EDS_{abs} and X – EDS_{sq}).

However, for female ratios and traits, the corresponding one-tailed confidence intervals were 0·95 and 0·91, respectively, indicating that the null cannot be rejected at the 0·05 level following this approach.

For both EDS types (EDS_{abs} and EDS_{sq}), the proportion of bootstrap runs in which the ecotone population had the highest EDS value out of all five populations studied was much higher than the expected 20%. For the male traits and ratios, the proportions for EDS_{abs} were 96·2% and 87·7%, respectively. For female traits and ratios, they were somewhat lower at 67·2% and 66·2%, respectively, yet still higher than the 20% predicted by the null hypothesis. The distribution of the bootstrap results is shown in Figs 2 and 3 for EDS_{abs}. Results of all tests using EDS_{sq} were very similar (not shown).

Discussion

SPATIAL TRENDS IN MORPHOLOGICAL DIVERSITY ACROSS THE RANGE

Results from the chukar case study support the hypothesis predicting peak levels of within-population diversity in the ecotone. These results are consistent with the analyses of genetic (allozyme) diversity across the chukar range in the same region, which show a hump-shaped trend with peak diversity at the ecotone and genetic structuring of populations despite substantial levels of gene flow between populations (Kark 1999; Kark *et al*. 2001). If the ecotone region would not have been included in the study, we may have concluded that there are no significant changes in within-population

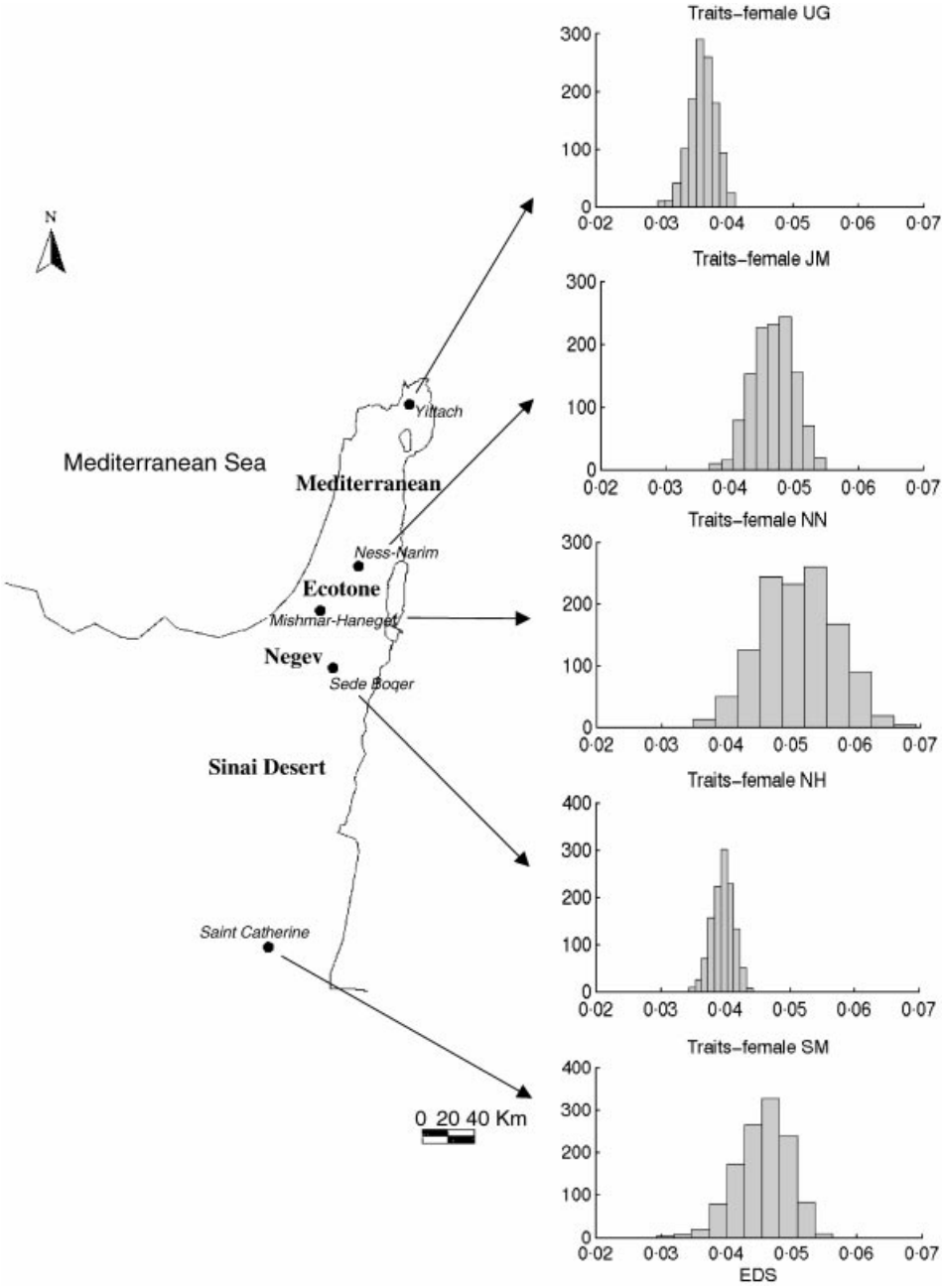

Fig. 2. Map of study area and location of populations with histograms showing the distribution of the frequency of the EDS_{abs} values in bootstrap runs of female traits across the distribution range. Detail for the EDS_{abs} bootstrap run distribution in male traits is shown in Fig. 3.

diversity across the chukar distribution range in Israel and Sinai. These ecotone regions are often small in size, and are intermediate between ecosystems or ecological communities, and thus do not fall into many ecosystem or community-orientated research agendas (e.g. study of Mediterranean ecosystems). Yet sampling and including gradients and areas of transition in research plans may be an important component in revealing spatial diversity patterns across species ranges, especially in the search for potential 'diversity hotspots'. While we cannot separate the genetic and environmental components affecting the phenotypic diversity detected in this study, at least some of the traits analysed have been shown to have high heritability in birds (Boag & van Noordwijk 1987; Smith *et al*. 1997) and are correlated with fitness, flight, movement and feeding ecology (Nissani 1974; Smith 1993; Smith *et al*. 1997).

In the ecotone region, where the edge of the continuous distribution range of the chukar occurs, fluctuating environmental conditions shift between more and less favourable for the species. Many other species reach the edge of their continuous distribution ranges in these areas (Danin *et al*. 1998; Kark 1999). Spatial and temporal fluctuations in climatic and environmental conditions may lead to higher levels of both genetic and morphological diversity and to the maintenance of within-population diversity in this region. Indeed, recent studies suggest that ecotones sustain high

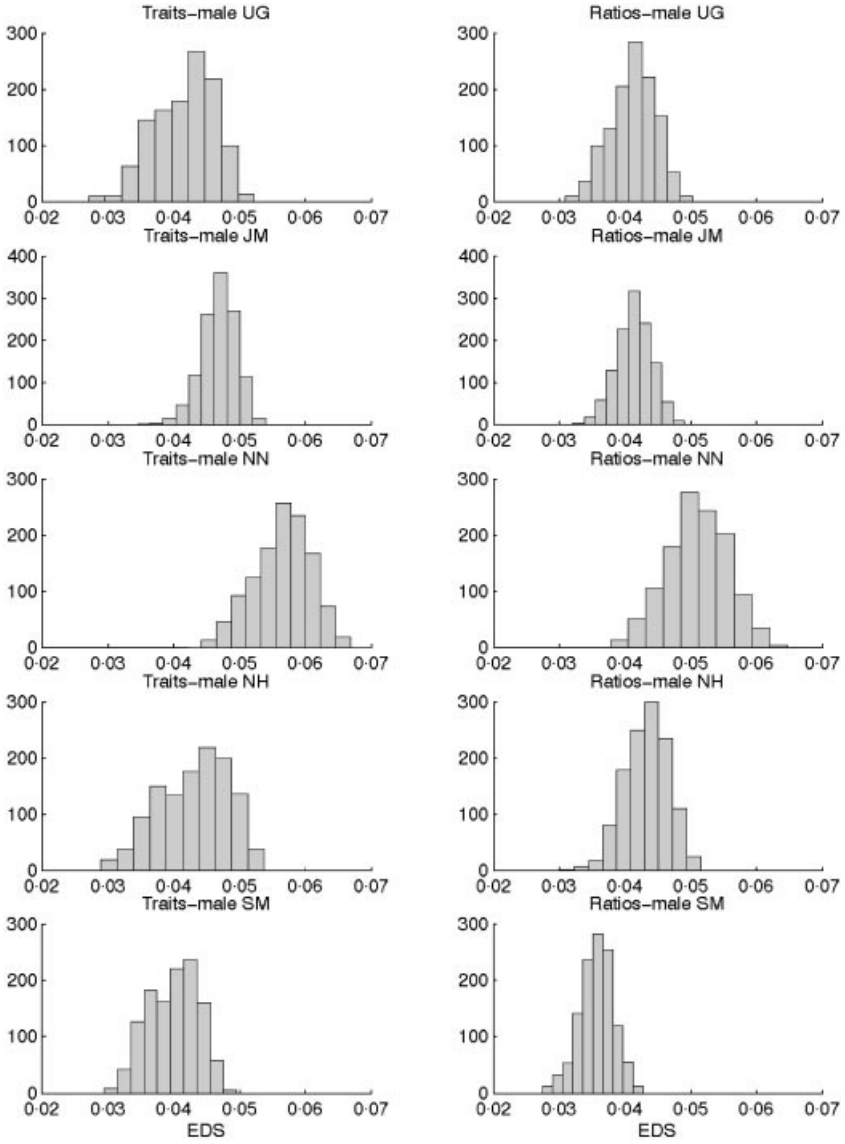

Fig. 3. Histograms showing the distribution of the frequency of EDS_{abs} values in bootstrap runs for male ratios and traits across the distribution range from north to south (top to bottom). Similar results were found for females. The ecotone population is NN (third from top). See Table 1 for population acronyms and Fig. 2 for locations.

morphological divergence in the face of high levels of gene flow (Schilthuizen 2000; Smith *et al*. 2001). Moreover, it has been shown recently that ecotones sustain not only high divergence but also high levels of genetic diversity within populations (Kark *et al*. 1999). A recent study focusing on comparison of developmental instability and environmental stress in the chukar across the same gradient in Israel suggests that the ecotone region is an area of shifts in the levels of environmental stress, as perceived by the birds and as reflected in their level and type of bilateral asymmetry (Kark 2001).

In this study we aimed to compare levels of diversity within populations. Other work using the same database compared population means aiming to examine Bergmann's ecogeographical rule. Nissani (1974), using all traits, found that ecotone populations did not show different means than other populations. She also suggested that while birds from Sinai morphologically differ from other populations, all Israeli populations, including those from the Negev desert, southern to the ecotone (see Fig. 2) should be included as a single subspecies. A recent study testing for temporal changes in phenotypic traits found that body mass (and marginally also tarsus length) declined significantly with decreasing latitude, but ambient temperature explained a much smaller fraction of the variation in body mass than latitude (Yom-Tov, Benjamini & Kark 2002). These studies further emphasize that trait means and their variances (or coefficient of variation) across species ranges may show very different spatial patterns.

OTHER APPROACHES

A more traditional approach that could potentially be used to calculate variation in a multivariate correlated

sample is the Mahalanobis D^2 distance. The Mahalanobis distance is used for classification and discrimination of different categories with multivariate attributes (Duda, Hart & Stork 2001). D^2 is given by: $(x–m)^T * C^{-1} * (x–m)$, where x is the sample point, m is the multivariate mean of the traits and C is the covariance matrix. It is a measure of the spread around the mean, normalized by the covariance. To estimate diversity in a multicorrelated sample, we could take the average of the Mahalanobis distance within each population or group. Calculation of the Mahalanobis distance requires inversion (by factorization) of the covariance matrix. While the Mahalanobis D^2 is potentially a useful estimator for calculating diversity, it has several important limitations for our application:

1. In the calculation of the D^2 the covariance matrix has to be inverted (or factorized). When the sample size is smaller than the number of traits (as is in the case presented here and in many cases when studying vertebrates and rare or endangered species, where sample sizes are relatively small) the covariance matrix can be nearly singular and badly conditioned. Therefore the inversion is inaccurate or in some cases not possible at all. When there are many traits, large sample sizes are required to compute the Mahalanobis distance. The EDS avoids this problem.

2. Mahalanobis D^2 is less flexible than EDS, as it always uses the square distance (while EDS can use the absolute value in addition to the square). In addition, it is very simple to choose in EDS between the variance, standard deviation, CV and other estimates of variation within populations.

3. Missing values may introduce a problem in calculating the covariance matrix and hence in calculating the Mahalanobis D^2. The missing values need to be filled in and this requires making assumptions about them. In the EDS calculation there is no need to fill in missing values as the inverse of the covariance matrix is not calculated. Hence the EDS is noncommittal about the missing values.

EDS VS. PCA

A question that emerges is why there was a need to use this novel EDS rather than simply apply a PCA procedure to analyse the variation within populations and compare it across populations. Our goal in this study was to obtain and compare *across* populations a single measure of variance over all variables (traits or ratios in our case) corrected for the covariance common to all pairs of variables. One may suggest that a similar final result could alternatively be obtained by measuring the variance of the first (at least) principal component in a PCA of variable values. In such a case the values would be log-transformed and standardized. There would be two possibilities for applying a PCA: one could run (i) a separate PCA of the individuals in each population or alternatively or (ii) an overall PCA of individuals in all populations combined and then project each population onto the new global principle component variables and calculate variances. Both approaches have certain problems. The problem in the first approach is that intervariable correlations may differ for different populations, thus the principal components (and specifically the first principal component (PC)) in each population may represent different combinations of variables. If we run a PCA separately on each population it is invalid to compare the variances of the PCs between populations, as they are different. In the case of the EDS the correlations may be different but the traits for each population are the same and are known. The second approach using the overall PCA avoids this problem and gives a common first axis of morphological variation, but it mixes between- and within-population variation in each PC. It is still possible to calculate, then, the within-variance of the first PC for each population, but the definition of this PC may be strongly influenced by between-population variation. When PCA is run for all populations combined, we assume that the correlations and relationships between the different variables are equal across populations. However, if there is one population that is different it may not be considered correctly. Moreover, when we project each population onto the overall PC axes, the variances are no longer uncorrelated because the overall PCs are not the PCs for each population. This, then, again presents the problem of comparing correlated variances, which the EDS attempts to solve naturally without first going through PCA. PCA simplifies the data to give new linearly independent components that replace the original traits. However, it is not clear what the components and the new discontinuous traits, rather than the original ones, represent. Is this actually biologically relevant? When we ran the PCA for both individual populations and using the PCs from all populations pooled, we found indeed that the PCs greatly vary with different new linear combinations of traits being created for different populations. Results for the variances were inconsistent across sexes, were not comparable and were difficult to interpret (not shown).

The EDS technique proposed and applied in this work does not share most of the above caveats and, especially, it does not create new traits that are difficult to interpret and that differ across populations. The bootstrap procedure enables us to describe confidence intervals and standard deviations, and thus to compare statistically diversity across populations, separating the within and between-populations components of variation. In addition, it enables us to compare the coefficient of variation and can be used for both traits and ratios.

THE ECOTONE: A HYBRID ZONE?

One could hypothesize that the high diversity found in chukar partridges in the Mediterranean-desert ecotone is solely a result of secondary contact of previously isolated populations in a hybrid zone. Re-contact of these populations could potentially have resulted in

high diversity in the contact area. Indeed, hybridization has been shown to be an important factor determining spatial patterns of diversity (e.g. Barton & Hewitt 1989; Hewitt 2001; and references therein). Thus secondary re-contact of formerly divergent and isolated populations in a hybrid zone with introgression seems a compelling simple theory to explain diversity trends in Israeli chukars across the ecotone. If this is the case, re-contact of formerly isolated populations in the region following the Quaternary glacials could have potentially led to hybridization and increased genetic diversity, yet morphological studies and palaeontological work suggest that this alone is not likely to explain the patterns seen, which are probably more complex. Chukar distribution in this region in the past 120 000 years has been contracting rather than expanding (Nissani 1974; Yom-Tov & Tchernov 1988). Distribution of chukars during the Upper Pleistocene was more continuous throughout the southern parts of Israel and Sinai, and included areas that are currently more arid, where chukars are not present or are very scarce (Shirihai 1996; Tchernov, pers. comm.). Thus, as suggested by Kark *et al*. (1999), the argument for historical re-contact and hybridization in the Mediterranean-desert ecotone causing increased diversity does not explain entirely the patterns seen at this time. Recent work across the same gradient at the mtDNA level applying a nested-clade approach may unveil the role of introgression and of historical factors in determining the high genetic diversity seen in the ecotone region (E. Randi, S. Kark & C. Tabarroni, unpublished). We suggest that a combination of factors may have been contributing to the patterns seen.

IMPLICATIONS FOR CONSERVATION

The study of phenotypic diversity, in addition to genetic diversity, is important for understanding patterns in components of biodiversity that are related to selection and fitness. Although we often seek to preserve the species' genetic diversity, the diversity in its morphological traits may mirror the changing environments where the species is found, and may reflect some important components of diversity related to fitness at the genetic level, undetected with standard population genetics techniques. Tools enabling practical and inexpensive, yet reliable and scientifically sound estimation of morphological diversity should be further developed and applied. The estimator presented in this work (EDS) may be a useful tool for monitoring levels of within-population morphological diversity based on measurements of multiple traits. It may be used to compare spatial and temporal trends in natural populations and in collection material. Further work using this statistic is required for evaluating its robustness and applicability for diverse areas of research.

Biodiversity has been defined as 'the diversity within species, between species and of ecosystems' (United Nations Conference on Environment and Development 1992). Patterns of species diversity across latitudinal, productivity and disturbance gradients have received wide scientific attention in many communities and ecosystems, both theoretically and empirically (e.g. Ricklefs & Schluter 1993; Gaston & Blackburn 2000). This has led to major advances in the prioritization of conservation efforts based on species endemism, richness and other diversity estimates (Myers *et al*. 2000; Olson & Dinerstein 1998). We argue here that similar approaches can be adopted and extended to the lower within-species level, in a search for areas with especially high genetic and morphological diversity. The advantage of the statistical method we use in this paper is that it is sufficiently powerful to detect differences across populations in their diversity when the sample size is small, as is often the case when working with vertebrates and rare or endangered species in natural environments.

We propose here that populations located in ecotone regions and transitional environments may be worth further focus as conservation targets, given their potential to maintain high genetic and morphological diversity. Ecotones may serve as potential centres for evolutionary novelty and for adaptive response to changing environments. Their conservation may be important for further survival of species and for maintenance of dynamic biodiversity processes, and may thus have much significance in the face of the current environmental changes.

Using a novel statistical estimator, which enables the estimation of diversity using multiple-correlated traits, we here show that ecotone populations, located across this area of environmental shift maintain especially high levels of morphological diversity, and may potentially serve as 'within-species diversity hotpots'. Having potentially important implications, the patterns revealed in this study should be further examined in additional populations and traits, as well as in other species and ecotones to test their generality.

Acknowledgements

We thank B. Efron's statistical consulting laboratory of the Department of Statistics at Stanford University for their helpful statistical advice, I. Milstein and O. Levi for mathematical consulting, P. Alkon, C. Conroy, M. Feldman, L. Hadly, U. Motro, K. O'Keefe, E. Randi, S. Shafir and M. van Tuinen for fruitful discussion and for comments on earlier drafts, and E. Tchernov for access to unpublished data and manuscripts. Our thanks to three anonymous referees for most helpful comments on the manuscript. Support for this research was granted to S.K. from the Rothschild Foundation, the Pontremoli and Rieger Research Funds through the Jewish National Fund (JNF), the Inter-University Ecology Fund founded by the JNF, Joe Alon Center for Regional Studies, the Mitrani Fund and the Blaustein International Center for Desert Studies of the Blaustein Institute for Desert Research, Ben-Gurion University of the Negev. T.M is supported by the Stanford Rock Physics Project.

References

Alkon, P.U. (1974) *Social behavior of a population of chukar partridge (Alectoris chukar) in Israel*. PhD thesis, Cornell University, Ithaca, New York.

Balmford, A., Moore, J.L., Brooks, T., Burgess, N., Hansen, L.A., Williams, P. & Rahbek, C. (2001) Conservation conflicts across Africa. *Science*, **291**, 2616–2619.

Barton, N.H. & Hewitt, G.M. (1989) Adaptation, speciation and hybrid zones. *Nature*, **341**, 497–503.

Begon, M., Harper, J.L. & Townsend, C.R. (1996) *Ecology*, 3rd edn. Blackwell Science, Oxford.

Bitan, A. & Rubin, S. (1991) *Climatic atlas of Israel for physical and environmental planning and design*. Ministry of Transport, Jerusalem.

Boag, P.T. & van Noordwijk, A.J. (1987) Quantitative genetics. *Avian genetics: a population and ecological approach* (eds F. Cooke & P.A. Buckley), pp. 45–78. Academic Press, London.

Carmi-Winkler, N., Degen, A.A. & Pinshow, B. (1987) Seasonal time–energy budgets of free-living chukars in the Negev Desert. *Condor*, **89**, 594–601.

Danin, A., Arbel, A. & Levy, N. (1998) *The Carta atlas of the fauna and flora of the Holyland*, Carta, Jerusalem.

Degen, A.A., Pinshow, B. & Alkon, P.U. (1983) Summer water turnover weights in free-living chukars (*Alectoris chukar*) and sand partridges (*Ammoperdix heyi*) in the Negev desert. *Condor*, **85**, 333–337.

Degen, A.A., Pinshow, B. & Shaw, P.J. (1984) Must desert chukars (*Alectoris chukar sinaica*) drink water? Water influx and body mass changes in response to dietary water content. *Auk*, **101**, 47–52.

von den Driesch, A. (1976) *A guide to the measurement of animal bones from archaeological sites*. Peabody Museum of Archeology and Ethnology, Harvard University, Harvard, USA.

Duda, R., Hart, P. & Stork, D. (2001) *Pattern classification*. John Wiley & Sons, New York.

Efron, B. & Tibshirani, R. (1998) *An introduction to bootstrap*, 2nd edn. Chapman & Hall, Boca Raton.

Frumkin, R. (1983) Thermoregulation and acid base status in heat stressed desert phasianids: a comparison between the chukar partridge (*Alectoris chukar*) and the sand partridge (*Ammoperdix heyi*). MSc thesis, Ben-Gurion University of the Negev [in Hebrew, English abstract].

Futuyma, D. (1997) *Evolutionary biology*, 3rd edn. Sinauer Associates, Inc., Sunderland, MA.

Gaston, K.J. & Blackburn, T.M. (2000) *Pattern and process in macroecology*. Blackwell Science, Oxford.

Hewitt, G.M. (2001) Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology*, **10**, 537–549.

Kadmon, R. & Danin, A. (1997) Floristic variation in Israel: a GIS analysis. *Flora*, **192**, 341–345.

Kam, M. (1986) Energy budget and seasonal water of free chukar (*Alectoris chukar*) and sand partridges (*Ammoperdix heyi*) in the Negev Highlands. MSc thesis, Ben Gurion University [in Hebrew, English abstract].

Kark, S. (1999) *Within-population diversity in the distribution range: Partridges as a research model*. PhD Thesis, The Hebrew University, Jerusalem and http://www.stanford.edu/group/Mooney/salit.

Kark, S. (2001) Shifts in bilateral asymmetry within a distribution range: the case of the chukar partridge. *Evolution*, **55**, 2088–2096.

Kark, S., Alkon, P.U., Safriel, U.N. & Randi, E. (1999) Conservation priorities for chukar partridge in Israel based on genetic diversity across an ecological gradient. *Conservation Biology*, **13**, 542–552.

Kark, S. & Mukerji, T. (2002) *A MATLAB program for the calculation of the estimator in a dependent sample (EDS) in multiple-correlated traits*. Stanford University, *http://www.stanford.edu/group/Mooney/salit/#Download the EDS*.

Kark, S., Safriel, U.N., Tabarroni, C. & Randi, E. (2001) Relationship between heterozygosity and asymmetry: a test across the distribution range. *Heredity*, **86**, 119–127.

Mace, G.M., Balmford, A., Boitani, L., Cowlishaw, G., Dobson, A.P., Faith, D.P., Gaston, K.J., Humphries, C.J., Vane-Wright, R.I., Williams, P.H., Lawton, J.H., Margules, C.R., May, R.M., Nicholls, A.O., Possingham, H.P., Rahbek, C. & van Jaarsveld, A.S. (2000) It's time to work together and stop duplicating conservation efforts.... *Nature*, **405**, 393.

Miller, R.G. (1991) *Simultaneous statistical inference*, 2nd edn. McGraw-Hill, New York.

Myers, N., Mittermeier, R., Mittermeier, C., da Fonseca, G. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Nissani, R. (1974) Geographic variability of the chukar *Alectoris chukar* (Gray) and Bergmann's rule. MSc Thesis, The Hebrew University, Jerusalem [in Hebrew, English abstract].

Olson, D.M. & Dinerstein, E. (1998) The Global 200: a representation approach to conserving the Earth's most biologically valuable ecoregions. *Conservation Biology*, **12**, 502–515.

Pinshow, B., Degen, A.A. & Alkon, P.U. (1983) Water intake, existence energy and responses to dehydration in the sand partridge (*Ammoperdix heyi*) and the chukar (*Alectoris chukar*), two phasianids of the Negev desert. *Physiological Zoology*, **56**, 281–289.

Rice, W.R. (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Ricklefs, R.E. & Schluter, D. (1993) *Species diversity in ecological communities: historical and geographical perspectives*. University of Chicago Press, Chicago.

Risser, P.G. (1995) Biodiversity and ecosystem function. *Conservation Biology*, **9**, 742–746.

Safriel, U.N., Volis, S. & Kark, S. (1994) Core and peripheral populations, and global climate change. *Israel Journal of Plant Sciences*, **42**, 331–345.

Schilthuizen, M. (2000) Ecotone: speciation prone. *Trends in Ecology and Evolution*, **15**, 130–131.

Schneider, C.J., Smith, T.B., Larison, B. & Moritz, C. (1999) A test of alternative models of diversification in tropical rainforests: ecological gradients vs. rainforest refugia. *Proceedings of the National Academy of Sciences*, **96**, 13869–13873.

Shirihai, H. (1996) *The birds of Israel*. Academic Press, London.

Smith, T.B. (1993) Disruptive selection and the genetic basis of bill size polymorphism in the African finch *Pyrenestes*. *Nature*, **363**, 618–620.

Smith, T.B., Kark, S., Schneider, C.J., Wayne, R.K. & Moritz, C. (2001) Biodiversity hotspots and beyond: the need for conserving environmental transitions. *Trends in Ecology and Evolution*, **16**, 431.

Smith, T.B., Wayne, R.K., Girman, D.J. & Bruford, M.W. (1997) A role for ecotones in generating rainforest biodiversity. *Science*, **276**, 1855–1857.

Soulé, M.E. & Cuzin-Roudy, J. (1982) Allomeric variation: 2. Developmental instability of extreme phenotypes. *American Naturalist*, **120**, 765–786.

United Nations Conference on Environment and Development (1992) Earth Summit – convention on biological diversity, Rio de Janeiro.

Yom-Tov, Y., Benjamini, Y. & Kark, S. (2002) Bergmann's rule, global warming and body mass – are they related? The chukar partridge case. *Journal of Zoology (London)*, **257**, 449–455.

Yom-Tov, Y. & Tchernov, E. (1988) *The zoogeography of Israel*. Dr W. Junk Publishers, Dordrecht.

Zar, J.H. (1999) *Biostatistical analysis*, 4th edn. Prentice Hall, Inc., Upper Saddle River, NJ.

Received 22 March 2002; revision received 27 June 2002

Appendix 1

THE MORPHOLOGICAL TRAITS AND RATIOS USED, AND THEIR MEASUREMENT METHODS

The traits used for the calculation of the statistical estimator include: body mass, wing length, tarsus length – external measurement, third toe length, culmen length, lower mandible length, tail length, length of gonads, total humerus length, maximal width of humerus distal end, maximal width of humerus proximal end, total ulna length, total carpometacarpus length, width of carpometacarpus proximal end, number of colour bands on wing, total coracoid length, dorsal manubrial spine length, maximal sternum labri width, distance between nutrient foramen and top of sternal crest, total femur length, maximal width of femur proximal end, maximal width of femur distal end, total tibia length, maximal width of tibia distal end, distance between the tibia muscular canals, total tarsometatarsus length, maximal width of tarsometatarsus proximal end, maximal width of tarsometatarsus distal end, mandibula (ramus) length, upper jaw: premaxilla length, maximal upper jaw length (proximal end of premaxilla to proximal end of nasalia), nasalia length, maximal length of nasal aperture, maximal length of ventral part of premaxilla, and total skull length.

Calculated ratios were usually between traits and total mass and between measured trait and total reference trait length (or width) and include: maximal width of humerus distal end/total humerus length, maximal width of humerus proximal end/total humerus length, width of carpometacarpus proximal end/total carpometacarpus length, maximal sternum labri width/dorsal manubrial spine, distance between nutrient foramen and top of sternal crest/dorsal manubrial spine, maximal width of femur proximal end/total femur length, maximal width of femur distal end/total femur length, maximal width of tibia distal end/total tibia length, distance between the tibia muscular canals/total tibia length, maximal width of tarsometatarsus proximal end/total tarsometatarsus length, maximal width of tarsometatarsus distal end/total tarsometatarsus length, praemaxilla length/skull length, maximal length of ventral part of praemaxilla/maximal length of nasal aperture, maximal length of nasal aperture/nasalia length, mandibula (ramus) length/skull length, total carpometacarpus length/total humerus length, total ulna length/total humerus length, total tarsometatarsus length/total femur length, total femur length/total tibia, wing length/mass, tarsus length/mass, 3rd toe length/mass, lower mandible length/mass (Nissani 1974).

Measurements of mass, external measurements, and descriptive details were taken soon after sampling, in the field, whereas other measurements were taken in the laboratory. Wing length was measured as the minimal distance between the tip of the longest primary to the carpal joints of the folded wing. Tarsal length was taken from the base of the last complete scale before the divergence of the toes to the angle of the intertarsal joint. The length of the upper beak was measured as a straight line along the cord of the culmen, from its tip to the edge of the feathering at the base of the skull. The length of the lower beak was taken from the articulation of the maxillary to its tip. Length of mandible was taken from the articulation to its tip (Nissani 1974; Nissani and Tchernov, unpublished ms). Measurements were taken by the same person and using Mitutuyo callipers (Nissani 1974). For detailed description of traits and measurements see von den Driesch (1976).

Appendix 2

MATLAB SOFTWARE FOR THE CALCULATION OF THE ESTIMATOR IN A DEPENDENT SAMPLE (EDS) AND FOR BOOTSTRAPPING

File calculationEDS.m:

```
function [EDS_abs,EDS_sq,sd,ave]=calculationEDS (data)
%EDS diversity statistic (1st Version)
%[EDS_abs,EDS_sq,sd,ave] = calculationEDS (data)
%Computes a single diversity statistic (EDS) for data with multiple
%attributes, taking in to account the correlation between different
%attributes. The EDS is computed using both the absolute value as well
%as the square of the correlation (R-squared). Often the two EDS may be
%very similar. For full definition of EDS please see Methods.
%
%INPUT:
% data = data matrix (n-by-m) with n rows corresponding to individual
%        samples, and m-columns corresponding to the different attributes.
%        Missing values should be indicated as 0.
```

```
%The m-file computes correlation coefficients between different attribute
%pairs taking in to account possible missing values.
%
%OUTPUTS:
% EDS_abs = Scalar number, EDS calculated with absolute value of correlation
% EDS_sq = Scalar number, EDS calculated with correlation squared.
% sd      = standard deviation of data, m-length vector, corresponding to
%           m attributes in data.
% ave     = average of data, m-length vector for m attributes.
%
%See also bootEDS, corrcoef, std, mean
%Reference:

%Written by Tapan Mukerji and Salit Kark, 2002 All rights reserved.

nvar=size(data,2);
r=zeros(nvar,nvar);

for k=1:size(data,2)
   indexk=data(:,k)~=0;
   datak=data(indexk,k);
  sd(k)=std(datak); ave(k)=mean(datak);
  a(k)=std(datak)/mean(datak);
  for j=k:size(data,2)
      indexj=data(:,j)~=0;
      index=indexk & indexj;
      if std(data(index,k))~=0 & std(data(index,j))~=0
      rr=corrcoef(data(index,k),data(index,j));
      r(k,j)=rr(1,2);
      r(j,k)=r(k,j);
      end;
  end;
end;

%%%%%%% Faster computation of correlation coefficient matrix
%%%%%%% If there are no missing data, and all elements of the
%%%%%%% data matrix are valid, then the following line of code
%%%%%%% computes 'r' the correlation coefficient matrix much faster.
%%%%%%% In this case, the above 'for loop' may be commented out,
%%%%%%  using the line below instead.
%
% r = corrcoef(data);

paren=1-abs(r)/2;
denom=sum(sum(paren)+0·5);
nom=sum (a.*(sum(paren)+0·5));
EDS_abs=nom/denom;

paren=1-((r.*r)/2);
denom=sum(sum(paren)+0·5);
nom=sum (a.*(sum(paren)+0·5));
EDS_sq=nom/denom;
```

File bootEDS.m:

```
function [eds,eds_stats,eds_data,eds_bias,bootsam]=bootEDS(data,nboot);
%Bootstrap calculation of Esitmator in a Dependent Sample (EDS) diversity statistic.
%[eds,eds_stats,eds_bias,bootsam]=bootEDS(data,nboot);
%
```

```
%INPUTS:
% data = data matrix (n-by-m) with n rows corresponding to individual
%        samples, and m-columns corresponding to the different attributes.
%        Missing values should be indicated as 0.
% nboot = Desired number (e.g. 1000) of bootstrap sampling.
%         If data matrix contains individuals with many missing measurements
%         then some of the drawn bootstrap samples may have to be rejected.
%         In this case, the final outputs may have slightly fewer
%         than nboot results.
%
%The EDS calculations for each bootstrap sample are done within
%another m-file, calculationEDS, which is called by bootEDS.
%EDS is computed using both absolute value, as well as the square
%of the correlation coefficient (R-squared). The outputs contain results and
%statistics of both computations. For complete definition of the
%EDS statistic, please see the reference.
%The outputs return the raw 95 percentile confidence intervals,
%as well as bias corrected 95 percentile confidence intervals, based
%on the bootstrap results.
%
%OUTPUTS:
% eds = 2 columns; 1st column eds_abs bootstrap
%                  2nd column eds_sq bootstrap
% eds_stats = 2 rows: 1st row eds_abs [mean, std, lower_ci, upper_ci];
%                     2nd row eds_sq [mean, std, lower_ci, upper_ci];
% eds_data = 2 rows: [eds_absdata, lower_ci_corrected, upper_ci_corrected;
%                     eds_sqdata, lower_ci_corrected, upper_ci_corrected]
% eds_bias: bias for [eds_abs, eds_sq]
% bootsam: index of bootstrap samples
%
%See also: calculationEDS, bootstrp (statistics toolbox)
%Reference:

%Written by Tapan Mukerji and Salit Kark 2002 All rights reserved.
[n,m]=size(data);
bootsam=unidrnd(n,n,nboot);

h=waitbar(0,'Please wait; bootstrapping');
for k=1:nboot
[eds_abs(k), eds_sq(k)]=calculationEDS(data(bootsam(:,k),:));
waitbar(k/nboot,h);
end;
delete(h);

[eds_absdata,eds_sqdata]=calculationEDS(data);

edsabs_bias=nanmean(eds_abs)-eds_absdata;
edssq_bias=nanmean(eds_sq)-eds_sqdata;
eds_bias=[edsabs_bias, edssq_bias];

eds=[eds_abs(:), eds_sq(:)];

sum(isnan(eds))

eds=eds(~isnan(eds(:,1)),:);
nbootnan=size(eds,1);

f1=sum(eds(:,1)<eds_absdata)/nbootnan;
f2=sum(eds(:,2)<eds_sqdata)/nbootnan;
```

```
z0abs=norminv(f1); z0sq=norminv(f2);
plabs=normcdf(2*z0abs-1·96);
puabs=normcdf(2*z0abs+1·96);
plsq=normcdf(2*z0sq-1·96);
pusq=normcdf(2*z0sq+1·96);

eds_stats(:,1)=nanmean(eds)';
eds_stats(:,2)=nanstd(eds)';
eds_stats(:,3)=prctile(eds, 2·5)';
eds_stats(:,4)=prctile(eds, 97·5)';

eds_absdataci=prctile(eds(:,1),100*[plabs, puabs]);
eds_sqdataci=prctile(eds(:,2),100*[plsq, pusq]);

eds_data=[eds_absdata eds_absdataci;
          eds_sqdata eds_sqdataci];
```