

## Annoying neighbors: Multi-scale distribution determinants of two sympatric sibling species of birds

Alban GUILLAUMET<sup>1, 2\*</sup>, Guillaume LEOTARD<sup>3</sup>

<sup>1</sup> EPHE, CEFE-CNRS UMR 5175, 1919 route de Mende, 34293 Montpellier cedex 5, France

<sup>2</sup> Current address: Office of Research/PIPES, 200 W. Kawili Street, Army Reserve Building A, Hilo, Hawaii, 96720, USA

<sup>3</sup> CEFE-CNRS, Département Interaction, Ecologie et Sociétés, 1919 route de Mende, 34293 Montpellier cedex 5, France

**Abstract** We tested the role of interspecific competition in driving species distribution at multiple spatial scales using two sibling species of *Galerida* larks (*G. cristata* and *G. theklae*) in Morocco (sympatry), Balearic islands (*G. theklae* only) and Israel (*G. cristata* only). We first investigated regional-scale determinants by contrasting allopatric *versus* sympatric patterns in five distinct habitat types. We next focused on a single habitat used by both species, the coastal sand dunes. Dune quadrats were established along the Moroccan coast and completed by a quadrat in the nearest distinct landscape habitat. Poisson regressions were used to model *Galerida* counts together with ecological predictors as concerns the climate, topography, vegetation structure and soil granulometry. At the local scale, both species preferred grey dunes over white sand dunes, and both were negatively affected by the abundance of the congeneric species in the dune. However, we found that *G. theklae* tended to replace *G. cristata* in more arid sand dunes, even if the transition was not strictly clinal. Instead, the transition occurred when the surrounding landscape changed from coastal wetlands to bathas (grasslands with shrubs), highlighting the importance of habitat composition at the landscape scale. The fact that *G. cristata* used bathas in allopatry, but not in sympatry, suggested that the competitive environment contributed to determine sand dune occupancy. We suggest that landscape-level effects may be pivotal in explaining species distribution not only at the local scale, by affecting the pool of potential immigrants, but also at the regional scale, by contributing to species' range limit [*Current Zoology* 61 (1): 10–22, 2015].

**Keywords** Allopatry, Ecological release, *Galerida* lark, Interspecific competition, Landscape, Range limit

Unraveling the key factors driving species distribution has always been a fundamental biogeography issue. Yet, climate change being perceived as one of the largest forthcoming threats to biodiversity, the development of accurate models of species distribution under future climate scenarios is becoming an urgent conservation issue and a priority research area (Pressey et al., 2007; Thuiller et al., 2008). Despite being popular and sometimes actually supported by the data (Driscoll et al., 2010), deterministic (niche-based) approaches are repeatedly criticized because the species-environment relationship is fitted without explicit consideration of the spatial context (e.g. Dormann, 2007). In addition to evident effects of intrinsic habitat quality, the probability of presence (or abundance) of a species in a given patch generally depends on the species abundance in surrounding areas, because of species' limited dispersal abilities (spatial autocorrelation, see e.g. Selmi et al., 2003; De Marco et al., 2008). In turn, the connectivity of surrounding habitats will also determine the number of individuals that can disperse into the patch (e.g. Ren-

jifo, 2001; Mortelliti and Boitani, 2009).

The importance of spatial context may be even more critical when considering biotic interactions such as interspecific competition (e.g. Apps et al., 2006; Araujo and Luoto, 2007). Community ecologists have traditionally emphasized the importance of local and regional scales as structuring forces on local communities (Spiesman and Cumming, 2008). The local scale corresponds to the spatial extent of foraging activities. For instance, the complexity of microhabitat topography strongly altered the outcome of competition in two Pacific Island geckos (Petren and Case, 1998). At the regional scale extinction and speciation rates and geographic dispersal determine regional species pools (Buckley and Roughgarden, 2005; Spiesman and Cumming, 2008). The existence of 'rules' of community assembly, such as nonrandom patterns of co-occurrence of species of different size guilds (Adams, 2007) suggests that interspecific competition can play a role in determining extinction and dispersal rates.

However, the composition and arrangement of envi-

---

Received Jan. 24, 2014; accepted May 11, 2014.

\* Corresponding author. E-mail: albang@hawaii.edu

© 2015 *Current Zoology*

ronmental features at the intermediate landscape scale can also influence the outcome of competition, in a way that may not be predicted based on an assessment of local or regional conditions alone (Amarasekare, 2003; Leibold et al., 2004; Spiesman and Cumming, 2008). Coexistence at the landscape scale might be achieved through trade-offs in performance within particular habitats or through a competition-colonization trade-off (Amarasekare, 2003; Kneitel and Chase, 2004). For instance, Spiesman and Cumming (2008) who studied sandhill ants community structure found that habitat generalists tended to replace habitat specialists at sites with a high proportion of habitat matrix in the surrounding landscape; Rodriguez et al. (2007) showed superior competitors became less frequent as forest patches were more isolated in a guild of passerines.

Hence, considering structuring forces at multiple spatial scales should provide the context necessary for a more thorough understanding of community organization (Spiesman and Cumming, 2008). However, the general importance of landscape heterogeneity for patterns of species coexistence remain under appreciated because of the paucity of studies that have examined this issue in nature (Oliver et al., 2009, Driscoll et al., 2010; but see also e.g. Delgado et al., 2013, Fisher et al., 2013), and multiple scales are rarely considered simultaneously (Spiesman and Cumming, 2008). In this paper, we used two sibling species of *Galerida* larks (*G. cristata* and *G. theklae*) as a model system and surrogate hypotheses (sensu Araujo and Luoto, 2007), rather than fully controlled experiments, to examine the influence of interspecific competition across multiple spatial scales. Hence, our study does not aim at formally demonstrating the existence of interspecific competition between the two species, but instead examines whether patterns of distribution and abundance are consistent with a negative impact of their interaction at the spatial scales investigated.

The existence of interspecific competition between *G. cristata* and *G. theklae* has been suggested using a variety of phenotypic and acoustic clues. In particular, morphological shifts were observed where they established secondary contact (Guillaumet et al., 2008a), and both larks responded to congeneric territorial calls with an aggressive response as intense as when responding to the intrusion of a conspecific, which also suggested that competition is symmetric (Laiolo, 2013). Current evidence based on ecological data is more limited. Both species apparently favor different types of habitats that segregate along a gradient of increasing ruggedness (corresponding to increased landscape complexity, from

plains to hills, and increased particle size, from clay to rock boulders); typically, cultivations for *G. cristata* and, at the rugged-end of the gradient, grasslands with shrubs (hereafter, bathas) for *G. theklae* (Guillaumet et al., 2010; Laiolo, 2012). Since no species appears dominant over the other, it might be that habitat segregation shown by *Galerida* species essentially responds to real habitat preferences and not to interspecific density-dependent habitat shifts (Laiolo, 2013). Contradicting this view, however, Guillaumet et al. (2010) showed that their habitat overlap in southern France was reduced during the presumably more limiting winter season, while in the Ebro valley, local densities of the two larks were negatively correlated even after controlling for habitat factors (Laiolo, 2013).

Theoretical work has shown that spatial segregation of species might ultimately be the consequence of the interaction between an Allee effect and interspecific competition (Ferdy and Molofsky, 2002). Comparing the patterns of habitat use in sympatry and in allopatry may therefore provide critical information. Laiolo (2012) studied sympatric and allopatric localities in Spain for both the crested and Thekla larks and reported an apparent lack of competitive release in allopatry. However, since testing ecological release was not the aim of the study, no efforts were made to ensure that the same habitat types were sampled in similar proportions everywhere, and that critical components of habitat selection, such as ruggedness of habitats, were used to compare allopatric and sympatric localities (Laiolo P., pers. comm). In this paper, we therefore tested the 'ecological release' hypothesis by comparing the patterns of habitat use in Morocco (*cristata* + *theklae*), the Balearic Islands (*theklae* only) and Israel (*cristata* only) using a standardized protocol while controlling for residual environmental variance.

In addition, the apparent lack of hierarchy between species in resource exploitation (Laiolo, 2013) calls into question the mechanisms underlying occupancy in habitats shared by both species. In southern France, a previous study showed that *G. cristata* is dominant in large patches of vineyards in the 'Languedoc' plains, but not in small vineyards that are enclosed in bathas in nearby hilly 'Corbieres', which are dominated by *G. theklae* (Guillaumet et al., 2010). In this paper, we therefore tested the hypothesis that habitat occupancy in a given habitat is influenced by the surrounding habitat matrix (e.g., Spiesman and Cumming, 2008). For this analysis, we focused on the coastal sand dunes because unlike other habitats, coastal sand dunes are essentially surrounded by water and other sand dunes. As a result,

it is possible to characterize the habitat matrix that hypothetically influences sand dune occupancy by sampling the single inland habitat directly adjacent to the dune.

In summary, we first tested the 'ecological release' hypothesis in *Galerida* larks by contrasting allopatric versus sympatric patterns in five distinct habitat types, predicting an increase in the number of habitats occupied and / or within-habitat density compensation in allopatry. This is our 'Regional' analysis. We next tested the hypothesis that landscape features, including the presence or absence of competitors in the landscape, contributed to determine patch occupancy patterns in coastal sand dunes. This is our 'Landscape' analysis. Both analyses accounted for intrinsic features of habitats at the local scale. Combined with previous knowledge on these species gained from phylogenetic, phenotypic and acoustic studies, these data may offer one of the most comprehensive frameworks on the evolutionary history of two remarkably similar sympatric bird species.

## 1 Materials and Methods

We first present elements of methods common to both the 'Regional' and 'Landscape' analyses before detailing specific aspects of each analysis.

### 1.1 Sampling

The sampling unit (called a quadrat) is approximately equivalent to a 150×150 m station. It was classified as of one of the following habitats: D = sand dune, B = batha (grasslands with shrubs), C = crop, A = anthropic (e.g., parking lot), F = fallow land, S = saltflat. While some quadrats consisted of entirely homogeneous habitat, others were a mixture of one dominant and one or several other habitat types. In the latter case the dominant habitat type was used to classify the quadrat. A sampling session corresponds to a 20 minutes period where the observer covers the quadrat and counts the *Galerida*, followed by the recording in the field of up to 36 primary ecological variables (concerning the topography, vegetation structure, granulometry and floristics of the quadrat), including critical components of *cristata* - *theklae* habitat segregation in sympatry such as steepness of terrain and proportion of rocks in soil (see Appendix S1 and Guillaumet et al., 2010). Sampling was performed during the 2008 breeding season, from February 18 to April 30, in Morocco (*cristata* + *theklae*), the Balearic Islands (*theklae* only; Majorca, Ibiza and Formentera) and Israel (*cristata* only). Because in Morocco *G. cristata* is parapatric and apparently competing with another species of *Galerida* towards the interior part of the country, namely *G. macrorhyncha* (see Guillaumet et al., 2010), sampling was performed in the

*cristata* part of the range to avoid confusing the effects of *theklae* and *macrorhyncha* on *cristata*.

### 1.2 Species identification

Determination was made in the field using 8×32 binoculars and a 20 × 60 spotting scope. Despite the strong experience of the observers, not all individuals could be safely identified in Morocco because these *Galerida* are among the most similar sympatric bird species in the region and intraspecific variability is important (Guillaumet et al., 2005). Of course, observation conditions were also a factor. Out of a total of 432 individuals recorded during our sampling sessions in Morocco, 324 (75.0%) were identified with a degree of confidence estimated to be > 99% (ID = 99), 77 (17.8%) were identified with a degree of confidence between 51 and 99% (ID = 51), and 31 (7.2%) were left unidentified (ID = 50). We ensured that unidentified birds did not introduce biases in our analyses using methods detailed separately for each analysis (below). A total of 199 *G. cristata* was sampled in Israel, and 67 *G. theklae* were sampled in Balearic Islands.

### 1.3 The 'Regional' analysis

#### 1.3.1 Data

The ecological release was studied in five main habitats, namely crop, dune, batha, fallow and anthropic. Sampling was performed in the 'Mediterranean' region, as defined here by an aridity Index (as calculated using formula given in Guillaumet et al., 2010) < -3.89, in Morocco, the Balearic Islands and Israel. For two countries characterized by a steep aridity gradient (Morocco and Israel), and three habitats (crop, dune, batha), sampling was also performed independently in the "Desertic" region as defined here by an aridity Index > -3.70. An average of 8.8 quadrats were performed country<sup>-1</sup> region<sup>-1</sup> habitat<sup>-1</sup> (Appendix S2.1).

#### 1.3.2 Handling unidentified individuals

To avoid that unidentified birds introduce biases in our analyses, by underestimating the actual number of *Galerida* in sympatric localities, we re-affected unidentified individuals (ID = 50) to either species as follows : 1) for each habitat and region, we estimated *pt*, the actual *G. theklae* proportion, as the frequency of *G. theklae* among the individuals identified with a degree of confidence > 99%; and 2) each unidentified individual was ascribed to *G. theklae* with probability *pt* and to *G. cristata* with probability 1- *pt*. Several runs were conducted which resulted in different labelling of unidentified individuals, but significance levels and conclusions were almost identical, so we only present the results of one such run. Including only those individuals identified with a degree of confidence > 99% also yielded very

similar results (not shown).

### 1.3.3 Testing of the 'Ecological release' hypothesis

Two approaches were developed to test for an ecological release while accounting for possible between-countries ecological differences that could have affected *Galerida* counts. As a preliminary step for both, we conducted a Principal Component Analysis (PCA) on 24 environmental factors + aridity (see Appendix S1). In the first approach, one-tailed Wilcoxon rank-sum tests were used to compare species abundance in allopatry and sympatry. Results were regarded as potentially inconclusive when allopatric and sympatric quadrats occupied different portions of the PCA habitat space, as estimated visually (see Discussion).

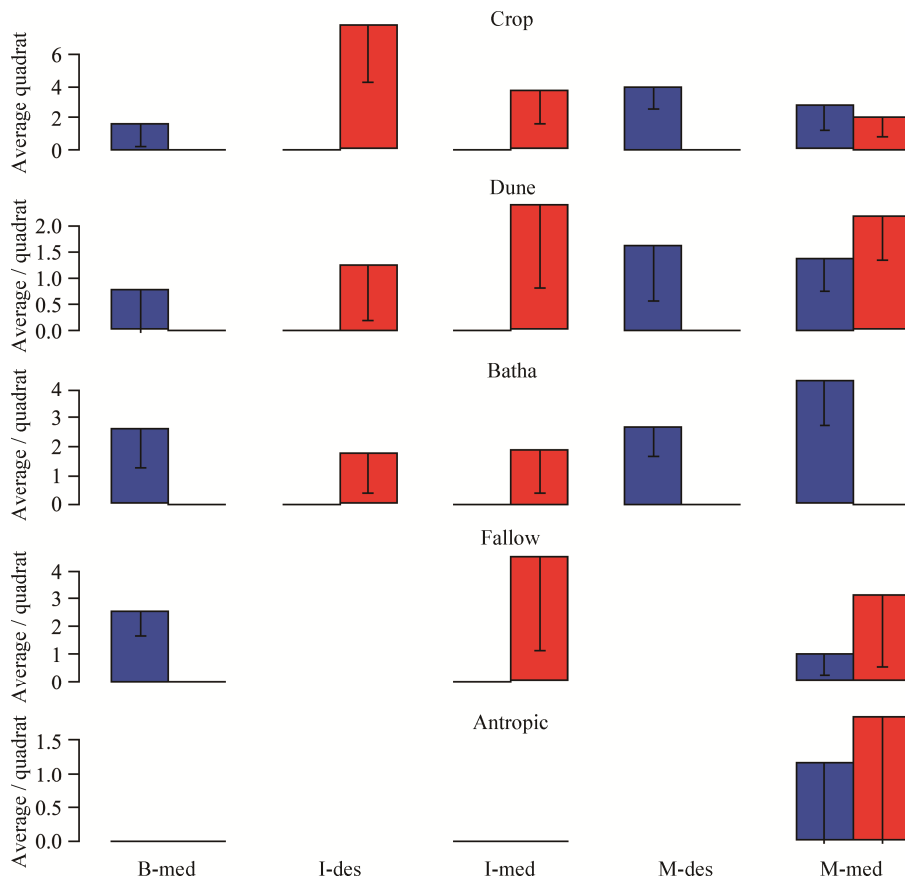
In the second approach, which was applied only to habitats with sufficient sample size (crop, dune and batha), allopatric and sympatric data were combined. We developed habitat-specific Poisson regression models using *G. theklae* counts as our response (alternatively, *G. cristata*), *G. cristata* counts as our 'competition' factor (alternatively, *G. theklae*), while simultaneously control-

ling for environmental factors such as climate and ruggedness of habitats by incorporating the first four axes of the PCA as covariates. Models without competition factors and the same models augmented with competition factors were compared using Akaike Information Criterion corrected for small sample size (AICc; QAICc, the quasi-likelihood modification of AICc yielded similar results so we only present the former). When 'competition' coefficients were unreliable (due to a cell in the cross-classification with only zero observation), we also performed a linear model for comparison; square-root and log(x+1) transformation of count data were used to ensure residuals were approximately normally distributed.

### 1.4 The 'Landscape' analysis

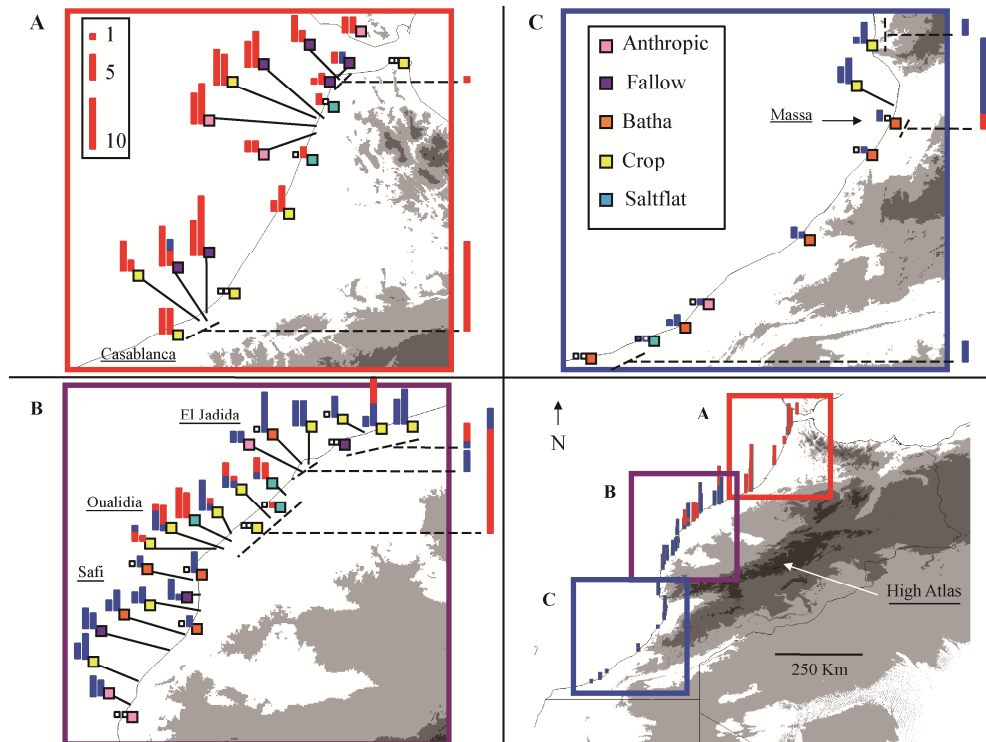
#### 1.4.1 Data

Dune quadrats were established at a fairly regular spacing (given accessibility) along the Atlantic coastline of Morocco, from the Gibraltar Strait (including one station in Spain and one on the Mediterranean side of N Morocco) to the northern Sahara margins (Fig. 2). A



**Fig. 1** Habitat-specific histograms of *Galerida* abundance ('Regional' analysis)

Legend for x-axis: B = Balearic Islands (strictly Mediterranean, *theklae* only), I = Israel (*cristata* only), M = Morocco (sympatry); med = Mediterranean region, des = desertic region. Color legend: blue = *G. theklae*, red = *G. cristata*. In sympatry, abundance values were corrected to account for unidentified individuals (see text for details and Appendix S2.2 for histograms based on raw values). Histograms show the mean abundance and approximate lower bound of the 95 % confidence interval around the estimated mean (lack of bar indicates lower bound  $\leq 0$ ).



**Fig. 2** Histograms of *Galerida* abundance mapped onto a map of Morocco ('Landscape' analysis)

The study zone (bottom right) is divided into the regions A, B and C which are detailed in the rest of the figure. Within each region, the paired histograms show the approximate geographic position and *Galerida* abundance for the dune (left) and the landscape (right) quadrats (ID = 99). Histogram size is proportional to the actual abundance (see region A for illustration), with *G. cristata* in red and *G. theklae* in blue (stacked charts). Empty squares are quadrats with no *Galerida* seen. The dominant habitat of the landscape quadrat is symbolized by a filled square: pink = Anthropogenic, purple = fallow, orange = batha, yellow = crop, blue = saltflat. Histograms outside the three regions (on the right) represent for comparison voucher specimens that were obtained during previous studies within 10 km of the coastline (approximately in the dotted zone area). These birds were identified in the hand by a combination of phenotypic traits and also, for most of them, by their mtDNA haplotype (Guillaumet et al. 2008b). We note however that they were sampled without any standardized protocol (sampling effort between habitat types was not balanced).

sampling session was performed in the nearest distinct (inland) habitat in the vicinity of the dune quadrat. Overall, 51 such paired quadrats were performed. The average ( $\pm SE$ ) distance between the center of one dune quadrat and its nearest neighbor dune quadrat was  $14.91 \pm 2.70$  km, while the average distance between the dune quadrat and its landscape quadrat was  $0.67 \pm 0.12$  km. For the 'landscape' analysis, we included only individuals identified with a probability  $> 99\%$  (ID = 99), but exploratory analyses using ID = 51 yielded similar results (not shown). Note that dune and landscape quadrats were also used for the 'Regional' analysis; the difference is that the surrounding landscape of these quadrats was not accounted for in the 'Regional' analysis.

#### 1.4.2 Response and explanatory factors

Model selection was performed for each species independently. The number of *G. cristata* (or *G. theklae*) recorded within a dune quadrat was the response variable. Poisson regressions were used to model the relationship between the response and explanatory variables. We did not consider the possible interactions between

factors in this paper both for simplicity and due to small sample size.

Explanatory factors were grouped into several classes. The 'Habitat preferences' class (DUN) consisted of thirty-two ecological variables associated with the characteristics of the dune quadrat, plus the aridity index (see Appendix S1 for details). To avoid collinearity, we conducted a PCA on these 33 variables; the first four axes, called PC1-4.DUN, were used to characterize the DUN class in subsequent analyses. The 'Landscape' (LAN) class consisted of the same 32 ecological variables measured this time in the landscape quadrat, plus the abundance of the focal species (and not of the competitor) in the landscape. A PCA was performed on this set of 33 variables, and the first four axes (PC1-4.LAN) were used to characterize the LAN class.

Properly quantifying the influence of one species on the distribution of the other was a bit more difficult. In a first series of run, the potential effect of competition was accounted for by two distinct classes of one single variable each, namely the abundance of the other *Ga-*

*lerida* species (the competing species) in the dune (COMP.DUN class) and the abundance of the competing species in the paired landscape (COMPLAN class). While this allowed us to separate possible competition effects at the local and landscape scales (we therefore called this run 'separate'), we were also interested in estimating the global effects of competition. In a second run (called 'global'), we therefore used a single class for competition (called COMP), where the two competition variables were subjected to a PCA, and summarized by the first axis (PC1.COMP).

It was important to recognize that our estimation of competition effects, as described above, might be too conservative; hence we refer to them as 'lower bound'. This was so because the first axis of the landscape class (called PC1.LAN) was actually a ruggedness axis (see results). While previous studies showed that both species are strongly separated along this ruggedness axis (see Introduction), our 'Regional' analysis revealed that *G. cristata* occupied rugged habitat in Israel, while *G. theklae* occupied non-rugged habitats, such as coastal sand dunes in Morocco, wherein the crested lark is rare (see results). Accordingly, the ruggedness of landscape, as measured by PC1.LAN, could be regarded as a measure of the competitive environment for both species. Therefore, we repeated all analyses above with the difference that PC1.LAN was excluded from the 'LAN' class and transferred to the relevant competition class. We call this series of run 'upper bound' because we cannot exclude that part of the information conveyed by PC1.LAN actually corresponded to auto-ecological processes. Conversely, the 'lower bound' run measure the residual effect of the abundance of the competing species in the landscape that is not predicted by the ruggedness of the habitat.

In summary, we conducted four series of run resulting from the combination of either: i) 'global' or 'separate' measure of competition; and ii) 'lower bound' versus 'upper bound' estimate of competition effects. The 'DUN' class of factors was always PC1.DUN + PC2.DUN + PC3.DUN + PC4.DUN. The 'LAN' class was PC1.LAN + PC2.LAN + PC3.LAN + PC4.LAN (- PC1.LAN for 'upper bound' run). For the 'global' run, the 'COMP' class was PC1.COMP (+ PC1.LAN for 'upper bound' run). For the 'separate' run, 'COMP' was divided up into two classes, namely 'COMP.DUN' = abundance of competing species in dune and 'COMP.LAN' = abundance of competing species in landscape (+ PC1.LAN for 'upper bound run').

### 1.4.3 Model selection

An Information-theoretic approach was used to

compare alternative models. Following Burnham and Anderson (2001), we used QAICc, the quasi-likelihood modification to AICc of account for possible overdispersion of count data; note that QAICc reduces to AICc whenever the overdispersion coefficient  $c \leq 1$ . We built all eight ('global' run) or sixteen ('separate') possible models containing between zero and three or four classes of factors. The models were ranked by rescaling the QAICc values in comparison with the model with the lowest QAICc, and we calculated QAICc weights. Top models with  $\Delta_i < 2$  were considered the most parsimonious, and the confidence set on the best model was obtained by summing QAICc weights of the models sequentially, starting at the best model, until we reached 95 %. Assessment of the relative importance of the class  $j$  was obtained by calculating  $w_{+}(j)$ , the sum of QAICc weights of models containing the class  $j$ .

To ensure that regressions were unbiased, residual spatial autocorrelation was investigated by drawing a correlogram in which Moran's spatial autocorrelation coefficient  $I$  (Moran, 1950) was plotted against classes of equidistant stations (ten classes of 10 km). A correlogram was considered significant if at least one of its Moran's  $I$  values was significant at the 5 % level, after Bonferroni correction (i.e.  $P < 0.005$ ). No spatial autocorrelation could be detected in the residuals of the best models, indicating that the regressions were unbiased: all  $P > 0.048$ .

All statistical analyses were performed using R version 2.13.1 (© The R Foundation for Statistical Computing).

## 2 Results

### 2.1 The 'Regional' analysis

Model selection based on combined allopatry plus sympatry data yielded evidence of ecological release (as indicated by AICc weight of competition model  $> 0.95$  and negative coefficient) for *G. cristata* in crop, dune and batha and for *G. theklae* in dune (Table 1). Wilcoxon rank-sum tests comparing sympatric and allopatric values yielded consistent results for *G. cristata*, while suggesting that ecological release was most pronounced in bathas and desertic region (Table 2). No ecological release was detected for *G. cristata* in fallow lands and anthropic habitats, but note small sample size (Appendix S2.1). For *G. theklae*, Wilcoxon rank-sum tests did not yield evidence of ecological release in dunes while conversely, they suggested a significant ecological release in fallow lands (Table 2).

**Table 1** Test of ecological release ('Regional' analysis) - Habitat-specific Poisson regression results for habitats with sufficient sample size

Species	Habitat	Model	AICc.e	AICc.ec	w.ec	Estimate	SE	z	P-value
<i>theklae</i>	Crop	Poisson	158.874	158.089	0.597	-0.116	0.064	-1.809	0.070
<i>cristata</i>	Crop	Poisson	243.159	215.184	1.000	-0.315	0.071	-4.450	<0.001
<i>theklae</i>	Dune	Poisson	236.131	224.614	0.997	-0.237	0.071	-3.344	0.001
<i>cristata</i>	Dune	Poisson	267.686	253.007	0.999	-0.251	0.070	-3.558	<0.001
<i>theklae</i>	Batha	Poisson	135.298	135.298	0.500	/	/	/	/
<i>cristata</i>	Batha	Poisson	142.696	104.888	1.000	-17.194	<b>2358.858</b>	<b>-0.007</b>	<b>0.994</b>
<i>cristata</i>	Batha	Gaussian	74.406	65.414	0.989	-0.447	0.130	-3.436	0.002

AICc.e = AIC corrected for small sample size (AICc) for a model where 'Species' counts are explained by environmental and climatic factors only; AICc.ec = AICc of the same model augmented by competitor counts; w.ec = Akaike weight of the 'competition' model; Estimate of the competitor coefficient is given together with associated standard error (SE), z- and P-values. Lack of estimate for *theklae* in batha results from a lack of *cristata* in bathas within *theklae* range, while values in bold correspond to unreasonably high SE resulting from a cell in the cross-classification with only zero observation (*cristata* in bathas were never associated with *theklae*, resulting in numerically null fitted rates; see Fig. 1 for illustrations); in the latter case, we also performed a linear model on transformed data (to ensure residuals were approximately normally distributed) for comparison (Model = Gaussian).

**Table 2** Test of ecological release ('Regional' analysis): Wilcoxon rank-sum tests

Habitat	<i>theklae</i> - Med	<i>cristata</i> - Med	<i>cristata</i> - Des
Crop	0.824	0.068	<i>0.005</i>
Dune	0.744	0.306	<i>0.016</i>
Batha	0.893	<i>0.013</i>	<i>0.001</i>
Fallow	<i>0.009</i>	0.263	/
Anthropic	0.945	0.968	/

P-values of one-tailed tests testing H1: counts in allopatry are greater than in sympatry; values in italic are significant; Med = Mediterranean and Des = Desertic; for instance, *theklae*-Med tests whether *theklae* counts are greater in Balearic Islands (strictly Mediterranean) than in the Mediterranean region of Morocco, while *cristata*-Des tests whether *cristata* counts are greater in desertic region of Israel than in desertic region of Morocco.

## 2.2 The 'Landscape' analysis

For *G. cristata*, the 'global' and 'lower bound' analysis indicates that all three classes of factor, namely dune, landscape and competition, play an important role in driving the patterns of *G. cristata* abundance in the dune: all  $w_+(j) \geq 0.99$  (Table 3). At the local scale, we found that *G. cristata* was more abundant in less arid environment (see also Fig. 2), in grey (fixed) as opposed to white sand dunes, and in dunes with less vegetation cover (Table 4). At the landscape scale, *G. cristata* was more abundant when the surrounding landscape was less rugged and, to a lesser extent, when the landscape had finer-grained soils, notably near saltflats (see Appendix S3.2) and when the landscape had a higher tree cover. Finally, *G. cristata* was more abundant when *G. theklae* was less abundant. The 'separate' and 'lower bound' analysis suggested that residual competition effects (after habitat factors were partialled out) were

stronger in the dune than in the landscape:  $w_+(\text{COMP.DUN}) = 0.97$  versus  $w_+(\text{COMP.LAN}) = 0.31$ , although they were similar (both  $\geq 0.96$ ) when the ruggedness of the landscape was regarded as a measure of the competitive environment ('upper bound'). In the latter case, the importance of the landscape class dropped from  $w_+(\text{LAN}) = 1.00$  to 0.24 (Table 3).

The results for *G. theklae* differed in a series of points. First, a larger array of models was necessary to build the 95 % confidence set; second,  $w_+$  ranking yielded the competition class > landscape > dune, although all classes were necessary to build the 95 % confidence set (Table 3). Third, *G. theklae* was positively, and not negatively, influenced by environmental aridity, and it was negatively, rather than positively, influenced by fine-grained soils in the landscape (Table 4). Fourth, including the ruggedness of the landscape as a measure of the competitive environment did not induce an increase of  $w_+(\text{COMP.LAN})$ , which actually dropped from 0.23 (Table 3) to 0.20 (not shown).

## 3 Discussion

### 3.1 Ecological release in allopatry

Ecological release was detected for *G. cristata* in the three main habitats sampled (dune, crop and batha). Since bathas constitute a stronghold for *G. theklae* wherein it occurs (e.g. Guillaumet et al., 2010), we will now examine this habitat in more details, as it certainly provides the strongest case for ecological release. Although our results suggest the existence of release in bathas independently in both Mediterranean and desertic regions (Table 2), we observe that Mediterranean bathas in Morocco and Israel do not occupy the same

**Table 3** Information-theoretic statistics of the 'Landscape' analysis

Species	Global	Bound	c	D	L	C	C.D	C.L	K	QLL	QAICc	ΔQAICc	w	Σw
<i>cristata</i>	TRUE	lower	1.00	1	1	1	/	/	10	-48.25	121.99	0.00	0.98	0.98
				0.99	1.00	1.00	/	/						
<i>cristata</i>	TRUE	upper	1.00	1	0	1	/	/	7	-51.50	119.61	0.00	0.76	0.76
				1	1	1	/	/	10	-48.25	121.99	2.38	0.23	1.00
				1.00	0.24	1.00	/	/						
<i>cristata</i>	FALSE	lower	1.00	1	1	/	1	0	10	-48.70	122.90	0.00	0.67	0.67
				1	1	/	1	1	11	-47.95	124.66	1.77	0.28	0.95
				0.98	1.00	/	0.97	0.31						
<i>cristata</i>	FALSE	upper	1.00	1	0	/	1	1	8	-51.41	122.24	0.00	0.74	0.74
				1	1	/	1	1	11	-47.95	124.66	2.42	0.22	0.96
				1.00	0.24	/	0.96	1.00						
<i>theklae</i>	TRUE	lower	1.60	0	0	1	/	/	3	-48.06	102.64	0.00	0.42	0.42
				0	1	1	/	/	7	-43.08	102.76	0.12	0.40	0.81
				1	1	1	/	/	11	-38.11	104.99	2.35	0.13	0.94
				1	0	1	/	/	7	-45.64	107.88	5.24	0.03	0.97
				0.17	0.55	0.97	/	/						
<i>theklae</i>	FALSE	lower	1.58	0	0	/	1	0	3	-47.16	100.82	0.00	0.34	0.34
				0	1	/	1	0	7	-42.28	101.16	0.34	0.29	0.63
				0	0	/	1	1	4	-47.10	103.07	2.25	0.11	0.74
				1	1	/	1	0	11	-37.21	103.19	2.37	0.11	0.85
				0	1	/	1	1	8	-42.28	103.98	3.16	0.07	0.92
				1	0	/	1	0	7	-44.49	105.57	4.75	0.03	0.95
				1	1	/	1	1	12	-37.16	106.53	5.70	0.02	0.97
0.17	0.50	/	0.98	0.23										

For each species, we conducted four series of runs resulting from the combination of either: i) 'global' or 'separate' (global = FALSE) measure of competition; and ii) 'lower bound' versus 'upper bound' estimate of competition effects. Legend: classes of factors (class included = 1, 0 otherwise): D = DUN (ecological factors associated with dune habitat), L = LAN (ecological factors associated with the landscape adjacent to the dune), C = COMP (interspecific competition); the latter class was divided up into two classes for the 'separate' analyses, namely C.D = COMP.DUN and C.L = COMPLAN, the separate competition effects due to the presence of the competing species in the dune and the landscape, respectively. For each run, we give the 95 % confidence set, i.e. the set of models necessary to reach a cumulative QAICc weight ( $\Sigma w$ ) of 0.95; K = number of parameters, QLL = Log-Likelihood divided by c (the coefficient of overdispersion, identical for all models in a same run). In the last line of each run are given  $w_+(j)$ , the sum of QAICc weights of models containing the class  $j$ . Since competition-related  $w_+$  were not higher for *G. theklae* using 'upper bound' estimates, we only present the results obtained using 'lower bound' runs.

fraction of the ecological space (Appendix S2.3.3), making the conclusions of the corresponding Wilcoxon rank-sum test questionable. Nevertheless, a global test which accounted for country-specific as well as regional (due e.g. to environmental aridity) habitat differences also indicated release for *G. cristata* in bathas in the absence of *G. theklae* (Table 1). The release is such that in desertic region, where habitats of both countries were similar (Appendix S2.3.3), the abundance of *G. cristata* in Israeli bathas was not statistically distinguishable from the abundance of *G. theklae* in Moroccan bathas (Wilcoxon rank-sum test:  $P = 0.287$ ; see also Fig. 1). Our results are consistent with those published by Shkedy and Safriel (1992) who studied the niche breadth of

*G. cristata* in the Neguev desert of Israel. They reported that *G. cristata* was not only present on stony hillsides, but that it further spent most of its time foraging in the central portions of the rocky slopes, i.e. in the most rugged fraction of the available habitat.

For *G. theklae* however, evidence of ecological release was not clear-cut, and we believe that no definitive conclusion can be reached based on present data. Although Poisson regression suggested ecological release in dune (Table 1), no such effect was found using a Wilcoxon rank-sum test (Table 2). We note however that dunes in Balearic Islands differ strikingly from Mediterranean dunes in Morocco and Israel along the first axis of the PCA, which is negatively related to vegeta-



**Table 4** Overdispersion-adjusted Poisson regression coefficients (Estimate) for models predicting the abundance of 'Species' (*G. cristata* or *G. theklae*) in the dune ('Landscape' analysis)

Species	Factor (F)	Class	signature var. (sv)	r(sv, F)	Estimate	SE	t	Interpretation
<i>cristata</i>	PC1.DUN	DUN	biovolume	0.87	-0.38	0.15	-2.59	- vegetation cover
	PC2.DUN	DUN	rock	0.81	0.03	0.14	0.21	/
	PC3.DUN	DUN	aridity	0.82	-0.47	0.31	-1.53	- aridity
	PC4.DUN	DUN	grey	0.81	0.51	0.17	2.93	+ grey dunes
	PC1.LAN	LAN	pBatha	-0.92	1.85	0.58	3.20	- ruggedness
	PC2.LAN	LAN	pAnthr	0.73	0.04	0.18	0.21	/
	PC3.LAN	LAN	cv4	0.82	0.14	0.12	1.17	+ tree cover
	PC4.LAN	LAN	clay	0.68	0.34	0.22	1.55	+ fine-grained soils
	the_AB_D	COMP.DUN	/	/	-0.44	0.20	-2.27	- competition DUN
	the_AB_L	COMP.LAN	/	/	-0.17	0.15	-1.15	- competition LAN
<i>theklae</i>	PC1.DUN	DUN	biovolume	0.87	-0.29	0.25	-1.18	- vegetation cover
	PC2.DUN	DUN	rock	0.81	0.03	0.18	0.19	/
	PC3.DUN	DUN	aridity	0.82	0.24	0.23	1.04	+ aridity
	PC4.DUN	DUN	grey	0.81	0.45	0.19	2.31	+ grey dunes
	PC1.LAN	LAN	pBatha	0.92	-0.45	0.23	-1.93	- ruggedness
	PC2.LAN	LAN	pAnthr	0.73	-0.33	0.21	-1.58	- Anthropic
	PC3.LAN	LAN	cv4	0.81	0.53	0.23	2.36	+ tree cover
	PC4.LAN	LAN	clay	0.60	-0.15	0.19	-0.77	- fine-grained soils
	cri_AB_D	COMP.DUN	/	/	-0.55	0.24	-2.24	- competition DUN
	cri_AB_L	COMP.LAN	/	/	0.05	0.16	0.33	/

The models presented here correspond to four classes of factors ('separate') + 'lower bound' measures of competition. All classes of factors were included. 'Signature' variable (sv) is a single raw variable strongly correlated to the explanatory Factor (F) resulting from a within-class Principal Component Analysis; r (sv, F) is the corresponding Pearson's correlation coefficient;  $t = \text{Estimate} / SE$ . The 'Interpretation' column gives biological interpretation; e.g. '- aridity' indicates that the presence of *G. cristata* in the dune is negatively influenced by environmental aridity; no interpretation is given when Estimate  $\sim 0$ ; the\_AB\_D is the abundance of *G. theklae* in the dune, cri\_AB\_L the abundance of *G. cristata* in the landscape quadrat. See Appendix S1 for summary of variables, Appendices S3.1–3 for details on Principal Components Analyses, and Appendix S3.4 for results of the 'global' analysis.

tion cover (Appendix S2.3.2). Hence, a possible explanation for a lack of ecological release based on Wilcoxon rank-sum test is that dunes in Balearic Islands, characterized by higher vegetation cover, were intrinsically less suitable for *G. theklae* than Moroccan dunes (see Table 3 for evidence of a negative effect of vegetation cover in sand dunes), an effect which was accounted for in Poisson regression (Table 1). Similarly, although Wilcoxon rank-sum test detected a significant release in fallow lands for *G. theklae*, we cannot exclude that fallow lands may have been intrinsically more favorable to *G. theklae* in Balearic Islands given that they do occupy a different portion of the ecological space (see Appendix S2.3.4).

Testing for competition release in natural systems is a difficult task, as it is impossible to compare habitats that differ only by the absence or presence of the competitor. While some factors can be anticipated to have an impact

and be statistically controlled for, as was done in this paper (Table 1), we emphasize that other factors could actually play a role. For instance, we observed during the course of the study that both *Galerida* seemed positively influenced by the diversity of herbaceous vegetation in the dune, possibly reflecting a more perennial source of seeds and insect preys, but this factor was not initially quantified and so was not accounted for. Similarly, the amount of water irrigation in desertic crops was not quantified, and water supply undoubtedly contributed to the unexpectedly high abundance of *G. cristata* in Israeli desertic crops (Fig. 1). Since desertic crops in Israel were essentially sampled along the Arava valley, while desertic bathas holding *G. cristata* were essentially sampled in a different area in the Neguev (e.g., Sde Boker), we stress the competition release detected in Israeli bathas should not be an artifact of differential irrigation systems in nearby crops (see also

Shkedy and Safriel, 1992, for evidence that *G. cristata* readily uses bathas in Israel even in the absence of large nearby crops).

### 3.2 Determinants of patch occupancy in sympatry

As could be expected, intrinsic features of the dune were important to explain occupancy patterns, although noticeably less so for *G. theklae* (Table 3). For instance, both species preferred grey (fixed) dunes over white sand dunes, where they were frequently seen feeding within sparse herbaceous vegetation, and both (although especially *G. cristata*) were rather intolerant to the presence of dense vegetation cover (Table 4). However, both species responded in opposite ways to aridity gradients (PC3.DUN in Table 4), as there was a clear trend for *G. theklae* to replace *G. cristata* in more arid environment (towards south of the Moroccan coastline, Fig. 2). Despite this, aridity alone, or in combination with dune habitat factors, cannot entirely explain observed patterns for two reasons. First, aridity was included as a covariate in the 'habitat preference' (DUN) class, which was clearly insufficient to explain the abundance patterns of both species (Table 3). Second, species' distribution did not entirely correspond to a clinal transition between *G. theklae*-dominated patches in the south and *G. cristata*-dominated patches in the north. Rather, the distribution was a more complex mosaic in the center of the gradient, with *G. theklae* apparently dominating in the Safi and El Jadida areas, but not in the Oualidia region sandwiched in between them (Fig. 2B).

The spatial context, as measured by the landscape class of factors, also had a strong apparent influence (Table 3). For instance, we found that the abundance of *G. cristata* in the dune diminished when the dune was surrounded by rugged landscape (bathas), but increased when surrounding habitats were characterized by fine-grained soils, such as coastal wetlands and associated saltmarsh vegetation (Fig. 2B, Table 4 and Appendix S.3.2). Hence, our results highlight the importance of habitat connectivity on species' distribution (see also e.g. Jaquiere et al., 2008; Driscoll et al., 2010).

Finally, interspecific competition appeared to have a strong influence at the local scale, where the abundance of each species in the dune negatively impacted the abundance of the other, even after habitat effects had been accounted for (Tables 3–4). We only detected a small, if any, effect of competition at the landscape scale for the Thekla lark: the relative importance of the corresponding class (COMPLAN) was weak ( $w_{+} = 0.23$ , Table 3), it did not increase when the ruggedness of the landscape was included as a measure of the competitive environ-

ment, and the coefficient estimate for crested lark abundance ( $cri\_AB\_L$ ) was  $\sim 0$  (Table 4). For the crested lark, conversely, the coefficient estimate for Thekla lark abundance ( $the\_AB\_L$ ) was negative ( $-0.17 \pm 0.15$ , Table 4), and the relative importance of the COMPLAN class was strongly boosted when ruggedness of the landscape was included as a measure of the competitive environment (Table 3).

### 3.3 A key role of the competitive environment?

We found that *G. cristata* were more frequent in the dune when the surrounding landscape was less rugged (Fig. 2B, Table 4). While it could be that bathas are intrinsically less suitable for *G. cristata*, we demonstrated in our 'Regional' analysis that *G. cristata* readily used bathas in the absence of *G. theklae* (Tables 1–2, Fig. 1). This suggests that avoidance of rugged landscape in Morocco, which partly determined *G. cristata* abundance in coastal sand dunes, might be more the consequence of a competitive exclusion by *G. theklae* than simple auto-ecological processes. In other words, the apparent exclusion of *G. cristata* from bathas by *G. theklae* seemed to have carry-over effects in other habitats such as sand dunes.

Comparing our results with those from a similar study in southern France (Guillaumet et al., 2010) provided additional support to the hypothesis that the surrounding landscape ('competitive environment') plays a prominent role in determining actual patch occupancy. This is because only *G. cristata* was recorded in coastal sand dunes in southern France, even if *G. theklae* was common in nearby bathas situated less than 10 km from the coast [Guillaumet et al. (2010): Fig. I-B in their Appendix S1]. Critically, sand dunes in southern France were adjacent to coastal lagoons. This pattern is consistent with our results from Morocco, where *G. cristata* and *G. theklae* responded in opposite ways to the fourth axis of the landscape PCA, positively correlated with fine-grained soils and the presence of saltflats (Table 4 and Appendices S3.2 and S3.3); hence, sand dunes adjacent to coastal wetlands were dominated by *G. cristata*, even though *G. theklae* was regionally dominant (Fig. 2B).

Somewhat complicating the picture, however, Guillaumet et al. (2010) reported that *G. theklae* was dominant in small vineyards enclosed in bathas, while *G. cristata* was dominant in large patches of vineyards. Hence, we expected *G. theklae* to be favoured by a rugged landscape. However, we failed to evidence such a positive effect and actually found the opposite pattern, as the abundance of *G. theklae* in the dune was, like *G.*

*cristata*, negatively affected by the ruggedness of the landscape (Table 4; note however that the size effect, as measured by the absolute value of 'Estimate', was four times larger for *G. cristata*). Part of this apparent contradiction may stem from the fact that rugged landscape along the Moroccan coast increased in frequency towards the most arid-end of the gradient (Fig. 2), where productivity is probably not enough to support large *Galerida* numbers. Further work will be required to better understand the response of *G. theklae* and *G. cristata* to rugged landscape.

### 3.4 Habitat partitioning driven by a trade-off in species performance?

This study, together with previous studies on the same two species (Guillaumet et al., 2012; Laiolo, 2013) clearly demonstrate that a series of factors acting at multiple spatial scales contribute, and likely interact, to determine patch occupancy and partitioning in these two interacting sibling species of larks. In this section, we raise and discuss the hypothesis that a trade-off in species performance mediated by environmental aridity and/or habitat productivity is a major determinant of species composition at the landscape scale which, in turn, contributes (together with other landscape and local habitat factors) to determine local patch occupancy in sympatry. Specifically, we suggest that the populations of *G. cristata* we sampled in Morocco might be better adapted to mesic and human-modified productive habitats, while being less competitive than *G. theklae*, but also *G. macrorhynca* (see Guillaumet et al., 2010), in less productive natural or arid habitats, for two reasons. First, genetic-based inferences demonstrated the ancient presence of *G. theklae* in the Maghreb (Maghreb-Ethiopian split dated at ~ 2.8 million years ago: Guillaumet et al., 2008a). Conversely, the presence of *G. cristata* in Morocco is much more recent and was promoted by the conversion of forest to agriculture in the northwest, more mesic, part of the country (Guillaumet et al., 2005; Guillaumet et al., 2008a). In fact, the colonization of the whole Eurasia (and secondarily, Morocco from Spain) by *G. cristata* was likely facilitated by the development of human agriculture that started around ten thousand years ago (Guillaumet et al., 2008a). Second, larks inhabiting mesic environments tend to have a different set of adaptations as compared to species living in more arid environments, including increased basal metabolic rate and evaporative water loss but also increased clutch size, number of clutches and parental effort (Tieleman et al., 2003, 2004), factors that could influence the outcome of interspecific com-

petition.

Our hypothesis of a trade-off in species performance could explain the finding that *G. cristata* dominates *G. theklae* in the mesic-end of the gradient, both in northern Morocco (Fig. 2A) and southern France (Guillaumet et al., 2010), but also possibly around coastal wetlands towards the arid-end in Morocco (Fig. 2B; although higher productivity in surrounding habitats remain to be demonstrated). While conversely, *G. theklae* seems 'relegated' into rugged, arid and semi-arid habitats (see Palmer, 2003, for an example where the outcome of competition depended on patch productivity in some African ants). Interestingly, the known southern range limit of *G. cristata* in the Massa area also coincides with an abrupt decrease in the abundance of *G. theklae* in both the dune and surrounding landscape (Fig. 2C), perhaps signaling a sharp decline in the productivity. This hypothesis would also help explaining why the ecological release we detected in *G. cristata* was stronger in desertic regions (Table 2). The fact that we failed to detect a general increase of *G. theklae* in mesic regions whereby *G. cristata* was absent (Fig. 1) suggests that *G. theklae* might be a specialist of rugged/arid environments. We note however that habitats in Balearic Islands tended to differ significantly from their Moroccan counterparts (Appendix S2.3), suggesting that our statistical power to detect ecological release may have been low.

A promising avenue for future research may be to compare physiological and life-history traits of *G. theklae* and *G. cristata*, both in allopatry and sympatry, in a series of different habitats spread along productivity, aridity, and ruggedness gradients. Aggressiveness towards the competing species (using e.g. playback experiment) and patterns of morphological differentiation (to account for possible phenotypic shifts, see e.g. Guillaumet et al., 2008a) could also be examined for possible variation along these gradients. Of course, other mechanisms such as competition-colonization trade-off (see review by Kneitel and Chase, 2004) and priority effects (e.g., Geange and Stier, 2009) may also contribute to observed patterns in these two species and could be investigated as well.

### 3.5 A role for landscape structure in determining range limits?

While *G. cristata* readily occupies the arid-end of the habitat gradient in allopatry, it was essentially absent from desertic environment in Morocco (Fig. 1; but see also Fig. 2), despite the fact that desertic quadrats in Morocco were slightly less arid than their Israeli coun-

terparts (not shown). We conclude that the distribution of *G. cristata* in Morocco is truncated towards the mesic-end at least partly because of its interaction with two competing *Galerida* species, namely *G. macrorhynca* towards the interior (Guillaumet et al., 2010) and *G. theklae* along the littoral, where *G. macrorhynca* is absent (this study). Concerning the latter, we further suggest that the southern range limit of *G. cristata* in Morocco coincides with a transition from a spatially heterogeneous to a spatially homogeneous competitive environment (Amarasekare, 2003) mediated by environmental aridity and landscape structure. As we have shown, the transition between *G. theklae*-dominated patches in the south and *G. cristata*-dominated patches in the north is not strictly clinal, rejecting the possibility that environmental aridity is the only driver. Instead, the transition south of Oualidia occurs when the landscape changes from coastal wetlands to rocky cliffs and bathas (Fig. 2B). Accordingly, we propose that landscape-level effects may be pivotal in explaining species distribution at both the local scale (patch occupancy, see Table 3) and regional scale (range limit). Mirroring the southern range limit of *G. cristata* in Morocco, we suggest that landscape-level effects may also be critical in determining the northern range limit of *G. theklae*. In France, the range limit of *G. theklae* coincides with a transition from a heterogeneous landscape (open bathas in the Corbieres massif alongside the littoral plain), favorable to both *G. cristata* and *G. theklae*, to a more homogeneous landscape (large-scale vineyards and coastal lagoons) occupied only by *G. cristata* (Gonin, 2008, Guillaumet et al., 2010, LPO Aude and A. Guillaumet, unpublished data).

**Acknowledgements** Roger Prodon and Finn Kjellberg supported the project, which was partly financed by GDRI BIOM, Roger Prodon suggested making the paired landscape quadrat, and Paola Laiolo kindly answered our questions concerning her previous work on ecological aspects of coexistence in *Galerida* larks. Two anonymous referees and Dr. Zhi-Yun JIA helped improving an earlier version of this manuscript.

## References

- Adams DC, 2007. Organization of *Plethodon* salamander communities: Guild-based community assembly. *Ecology* 88: 1292–1299.
- Amarasekare P, 2003. Competitive coexistence in spatially structured environments: A synthesis. *Ecology Letters* 6: 1109–1122.
- Apps CD, McLellan BN, Woods JG, 2006. Landscape partitioning and spatial inferences of competition between black and grizzly bears. *Ecography* 29: 561–572.
- Araujo MB, Luoto M, 2007. The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography* 16: 743–753.
- Buckley LB, Roughgarden J, 2005. Lizard habitat partitioning on islands: The interaction of local and landscape scales. *Journal of Biogeography* 32: 2113–2121.
- Burnham KP, Anderson DR, 2001. Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildlife Research* 28: 111–119.
- Delgado MP, Sanza MA, Morales MB, Traba J, Rivera D, 2013. Habitat selection and coexistence in wintering passerine steppe birds. *Journal of Ornithology* 154: 469–479.
- De Marco P, Diniz-Filho JAF, Bini LM, 2008. Spatial analysis improves species distribution modelling during range expansion. *Biology Letters* 4: 577–580.
- Dormann CF, 2007. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography* 16: 129–138.
- Driscoll DA, Kirkpatrick JB, McQuillan PB, Bonham KJ, 2010. Classic metapopulations are rare among common beetle species from a naturally fragmented landscape. *Journal of Animal Ecology* 79: 294–303.
- Ferdy J-B, Molofsky J, 2002. Allee Effect, Spatial Structure and Species Coexistence. *Journal of theoretical Biology* 217: 413–424.
- Fisher JT, Anholt B, Bradbury S, Wheatley M, Volpe JP, 2013. Spatial segregation of sympatric marten and fishers: The influence of landscapes and species-scapes. *Ecography* 36: 240–248.
- Geange SW, Stier AC, 2009. Order of arrival affects competition in two reef fishes. *Ecology* 90: 2868–2878.
- Gonin J, 2008. Le Cochevis de Thékla *Galerida theklae* en France: répartition, effectifs et sélection de l’habitat. Ecole Pratique des Hautes Etudes, unpublished thesis.
- Guillaumet A, Crochet PA, Godelle B, 2005. Phenotypic variation in *Galerida* larks in Morocco: The role of history and natural selection. *Molecular Ecology* 14: 3809–3821.
- Guillaumet A, Crochet PA, Pons JM, 2008a. Climate-driven diversification in two widespread *Galerida* larks. *BMC Evolutionary Biology* 8: 32.
- Guillaumet A, Ferdy JB, Desmarais E, Godelle B, Crochet PA, 2008b. Testing Bergmann’s rule in the presence of potentially confounding factors: A case study with three species of *Galerida* larks in Morocco. *Journal of Biogeography* 35: 579–591.
- Guillaumet A, Gonin J, Prodon R, Crochet PA, 2010. The geographic and seasonal dimensions of habitat use in *Galerida* larks: implications for species coexistence and range limits. *Ecography* 33: 961–970.
- Jaquiere J, Guelat J, Broquet T, Berset-Brandli L, Pellegrini E et al., 2008. Habitat-quality effects on metapopulation dynamics in greater white-toothed shrews *Crocidura russula*. *Ecology* 89: 2777–2785.
- Kneitel JM, Chase J, 2004. Trade-offs in community ecology: Linking spatial scales and species coexistence. *Ecology Letters* 7: 69–80.
- Laiolo P, 2012. Interspecific interactions drive cultural co-evolution and acoustic convergence in syntopic species. *Journal of Animal Ecology* 81: 594–604.
- Laiolo P, 2013. From inter-specific behavioural interactions to

- species distribution patterns along gradients of habitat heterogeneity. *Oecologia* 171: 207–215.
- Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM et al., 2004. The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters* 7: 601–613.
- Moran PAP, 1950. Notes on continuous stochastic phenomena. *Biometrika* 37: 17–23.
- Mortelliti A, Boitani L, 2009. Distribution and coexistence of shrews in patchy landscapes: A field test of multiple hypotheses. *Acta Oecologica* 35: 797–804.
- Oliver M, Luque-Larena JJ, Lambin X, 2009. Do rabbits eat voles? Apparent competition, habitat heterogeneity and large-scale coexistence under mink predation. *Ecology Letters* 12: 1201–1209.
- Palmer TM, 2003. Spatial habitat heterogeneity influences competition and coexistence in an African acacia ant guild. *Ecology* 84: 2843–2855.
- Petren K, Case TJ, 1998. Habitat structure determines competition intensity and invasion success in gecko lizards. *Proceedings of the National Academy of Sciences USA* 95: 11739–11744.
- Pressey RL, Cabeza M, Watts ME, Cowling RM, Wilson KA, 2007. Conservation planning in a changing world. *Trends in Ecology & Evolution* 22: 583–592.
- Renjifo LM, 2001. Effect of natural and anthropogenic landscape matrices on the abundance of subandean bird species. *Ecological Applications* 11: 14–31.
- Rodriguez A, Jansson G, Andren H, 2007. Composition of an avian guild in spatially structured habitats supports a competition-colonization trade-off. *Proceedings of the Royal Society of London B* 274: 1403–1411.
- Selmi S, Boulinier T, Faivre B, 2003. Distribution and abundance patterns of a newly colonizing species in Tunisian oases: the common blackbird *Turdus merula*. *Ibis* 145: 681–688.
- Shkedy Y, Safriel UN, 1992. Niche breadth of two lark species in the desert and the size of their geographical ranges. *Ornis Scandinavica* 23: 89–95.
- Spiesman BJ, Cumming GS, 2008. Communities in context: The influences of multiscale environmental variation on local ant community structure. *Landscape Ecology* 23: 313–325.
- Thuiller W, Albert C, Araujo MB, Berry PM, Cabeza M et al., 2008. Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology Evolution and Systematics* 9: 137–152.
- Tieleman BI, Williams JB, Bloomer P, 2003. Adaptation of metabolism and evaporative water loss along an aridity gradient. *Proceedings of the Royal Society of London B* 270: 207–214.

## SUPPORTING INFORMATION

## Appendix S1 Variables

	Definition	Name	Reg	DUN	LAN
	abundance of <i>G. cristata</i>	<b>cri_AB</b>	1		0.5
	abundance of <i>G. theklae</i>	<b>the_AB</b>	1		0.5
	value of the aridity index	<b>aridity</b>	1	1	
	proportion of bathas	<b>pBatha</b>		1	1
	proportion of anthropic habitats	<b>pAnthr</b>		1	1
	proportion of crops	<b>pCrop</b>		1	1
	proportion of fallow lands	<b>pFallow</b>		1	1
	proportion of saltflats	<b>pSaltflat</b>		1	1
	proportion of beach	<b>beach</b>		1	1
	proportion of white dune	<b>white</b>		1	1
	proportion of grey (fixed) dune	<b>grey</b>		1	1
	altitudinal amplitude	<b>AA</b>	1	1	1
	an index of the average slope ranging from 0 for flat grounds to 3 for slope > 25°	<b>slope</b>	1	1	1
	minimum slope	<b>min</b>	1	1	1
	maximum slope	<b>max</b>	1	1	1
	relative cover of low (< 25 cm) grass-like vegetation	<b>Gl</b>			
	relative cover of high grass-like vegetation	<b>Gh</b>			
	relative cover of rush-like vegetation	<b>J</b>			
	relative cover of reed-like vegetation	<b>R</b>			
	relative cover of low (< 25 cm) herbaceous (non-grass) vegetation	<b>Hl</b>			
	relative cover of high herbaceous (non-grass) vegetation	<b>Hh</b>			
	relative cover of low (< 25 cm) succulent vegetation	<b>Sul</b>			
	relative cover of high succulent vegetation	<b>Suh</b>			
	relative cover of chamephytes (shrubs < 25 cm)	<b>Ch</b>			
	relative cover of shrubs	<b>Sh</b>			
	relative cover of trees	<b>tree</b>	1	1	1
	relative cover of herbaceous-like vegetation (Gl+Gh+Hl+Hh+Sul)	<b>herb</b>	1	1	1
	relative cover of shrub-like vegetation (Ch+Sh+Suh)	<b>shrub</b>	1	1	1
	heterogeneity of (primary) vegetation types as measured by the index of diversity ( $1-\Sigma p_i^2$ )	<b>hetv</b>	1	1	1
	relative cover of the vegetation layers from 0 to 0.25 m	<b>cv0.25</b>	1	1	1
	relative cover of the vegetation layers from 0.25 to 0.5 m	<b>cv0.5</b>	1	1	1
	relative cover of the vegetation layers from 0.5 to 1 m	<b>cv1</b>	1	1	1
	relative cover of the vegetation layers from 1 to 2 m	<b>cv2</b>	1	1	1
	relative cover of the vegetation layers from 2 to 4 m	<b>cv4</b>	1	1	1
	relative cover of the vegetation layers from 4 to 8 m	<b>cv8</b>	1	1	1
	vegetation cover index = $\text{Ln}(0.25 \times \text{cv}0.25 + 0.25 \times \text{cv}0.5 + 0.5 \times \text{cv}1 + \text{cv}2 + 2 \times \text{cv}4 + 4 \times \text{cv}8)$	<b>biovolume</b>	1	1	1
	proportion of blocks larger than 10 m in bare ground	<b>Dm</b>	1	1	1
	proportion of rocks between 1 and 10 m in bare ground	<b>m</b>	1	1	1
	proportion of rocks between 10 and 100 cm in bare ground	<b>dm</b>	1	1	1
	proportion of rocks between 1 and 10 cm in bare ground	<b>cm</b>	1	1	1
	proportion of (smaller) gravels in bare ground	<b>gravel</b>	1	1	1
	proportion of sand in bare ground	<b>sand</b>	1	1	1
	proportion of finer-grained soils in bare ground	<b>clay</b>	1	1	1
	the relative cover of rocks in bare ground ( $\Sigma$ of all granulometry classes but sand and clay)	<b>rock</b>	1	1	1
	heterogeneity of (primary) soil granulometry types as measured by ( $1-\Sigma p_i^2$ )	<b>hetg</b>	1	1	1

Following Tieleman et al. (2003), we used a synthetic measure of the climate corresponding to the aridity index (*aridity*) as calculated in Guillaumet et al. (2010). Basic climatic data were extracted from a database at a 10-minute resolution (Hijmans et al., 2005). Covers were estimated visually by comparison with a reference chart which allows a reliability of  $\pm 10\%$  (Prodon and Lebreton, 1981). The relative cover of eleven biological types (Gl  $\rightarrow$  Tree) of plants was estimated within a radius of about 50 meters from the quadrat center. The relative cover for vegetation structure (cv0.25  $\rightarrow$  cv8) and granulometry (Dm  $\rightarrow$  clay) were estimated on the entire quadrat. The subset of variables included into the 'Regional' analyses are indicated by a 1 in the column 'Reg'. All variables were primary factors (measured in the field) except aridity, herb, shrub, hetv, biovolume, rock and hetg which were synthetic factors. All variables were measured in both the dune and the landscape (except aridity, dune only); a suffix is added to the name of the variable in the text to indicate where it has actually been measured. For instance, the abundance of *G. cristata* in the dune (resp. in the landscape) would be noted cri\_AB\_D (resp cri\_AB\_L, see e.g. Table 4). DUN and LAN show the explanatory factors used in the corresponding class of factors for the 'Landscape' analysis; a value of 0.5 indicates that only the focal species was used for LAN (e.g. if the response variable is cri\_AB\_D, then the value of cri\_AB in the column LAN (i.e. cri\_AB\_L) would be 1, while the \_AB\_L would be 0).

## Additional references

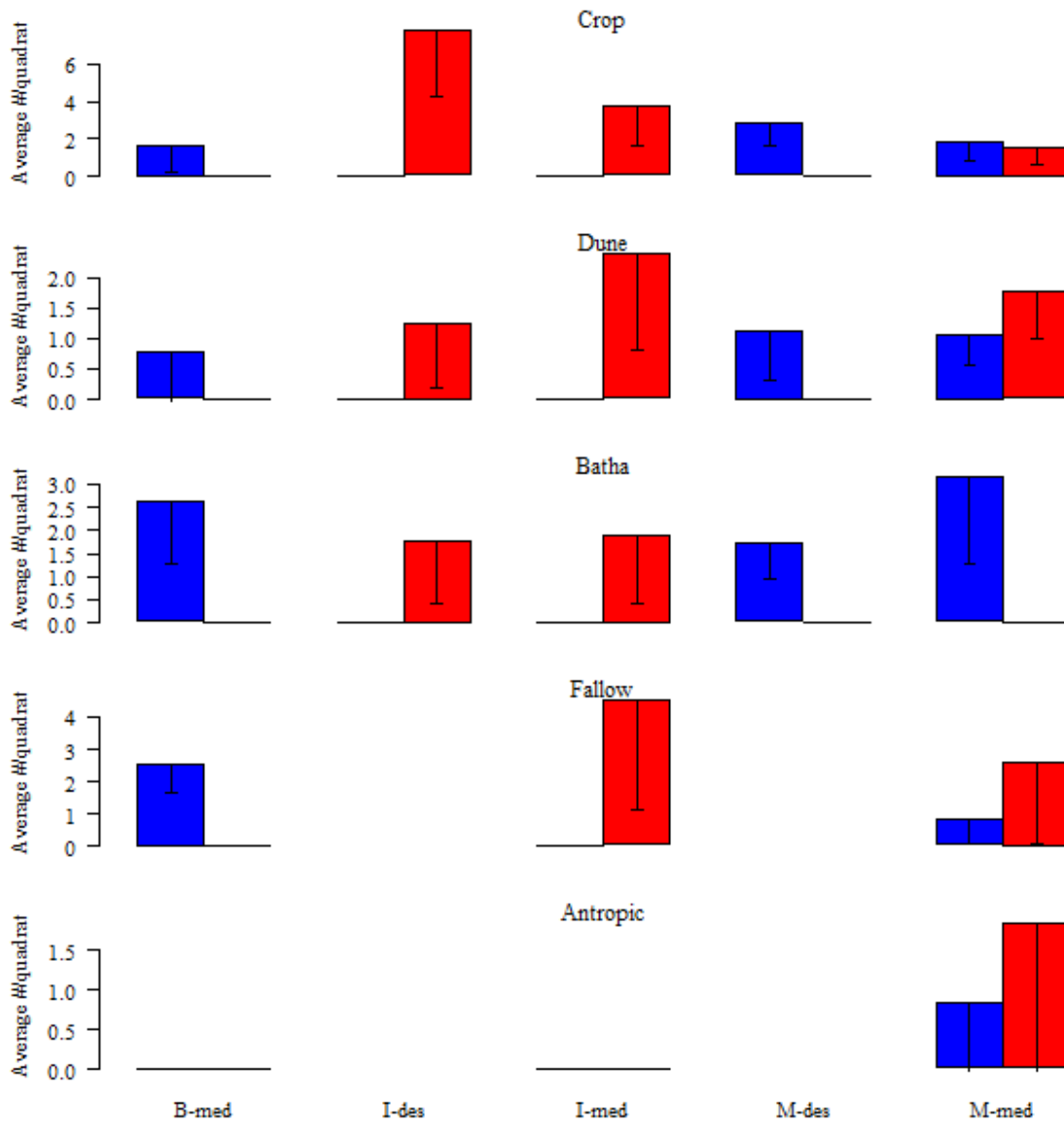
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A, 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.
- Prodon R, Lebreton JD, 1981. Breeding avifauna of a Mediterranean succession: the holm oak and cork oak series in the eastern Pyrenees, 1. Analysis and modelling of the structure gradient. *Oikos* 37: 21–38.

## Appendix S2 Regional analysis

**S2.1) Sample size (number of quadrats) for each of the five main habitats (rows) and five regions (columns) defined by country (B = Balearic islands, I = Israel and M = Morocco) and aridity class (med = Mediterranean  $\leftrightarrow -4.99 < \text{aridity index} < -3.89$ ; des = Desertic  $\leftrightarrow -3.70 < \text{aridity index} < -1.26$ )**

	B-med	I-des	I-med	M-des	M-med
<b>Crop</b>	9	8	8	6	20
<b>Dune</b>	9	8	8	8	43
<b>Batha</b>	8	9	11	14	7
<b>Fallow</b>	9	0	8	0	10
<b>Anthropic</b>	6	0	5	0	6

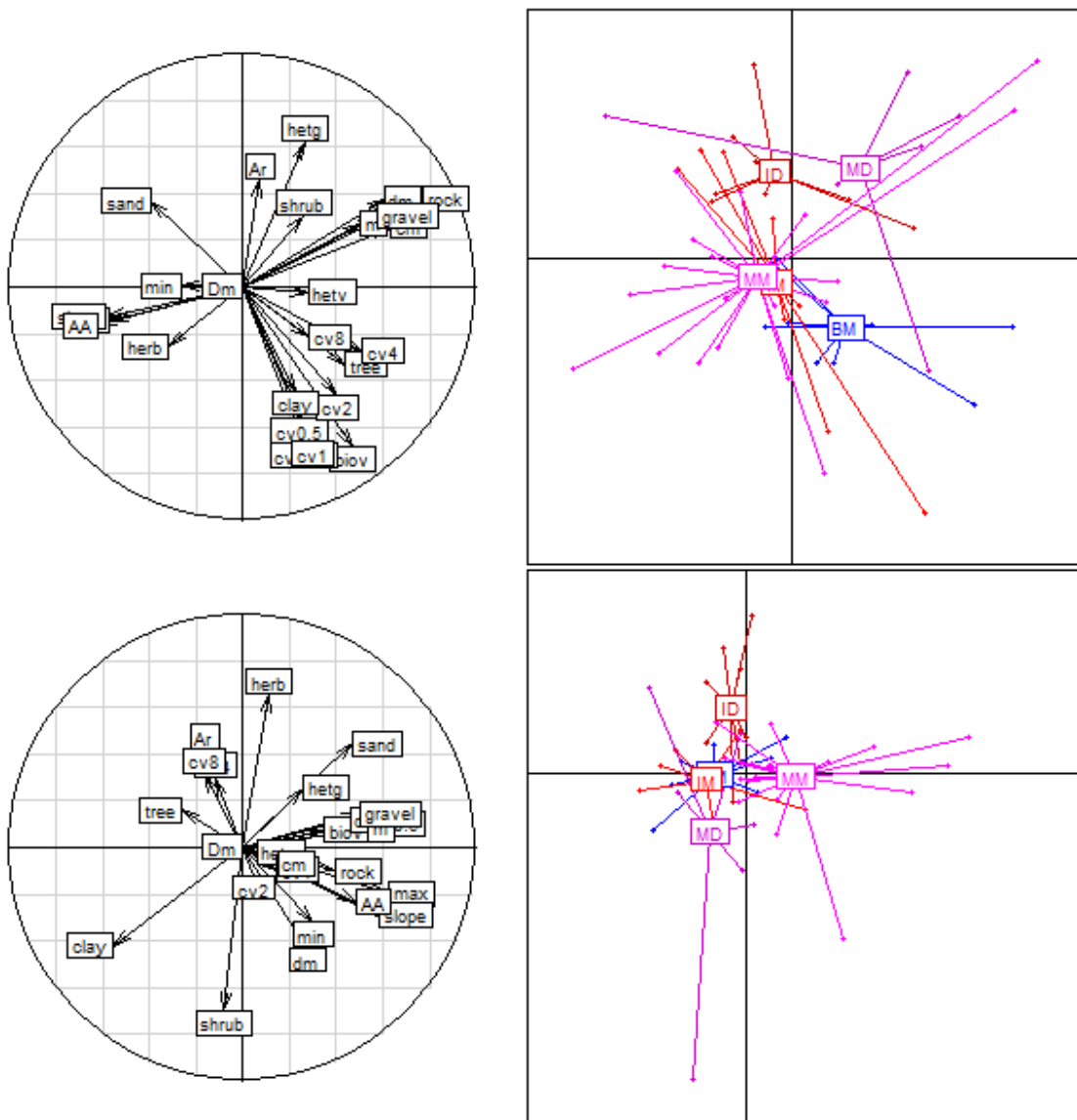
S2.2) Habitat-specific histograms of *Galerida* abundance - Same as Fig. 1, except that we only kept the birds identified with a level of confidence > 99 % in sympatry (Morocco); hence, birds in other categories (ID = 51 and ID = 50) were not re-ascribed to either species



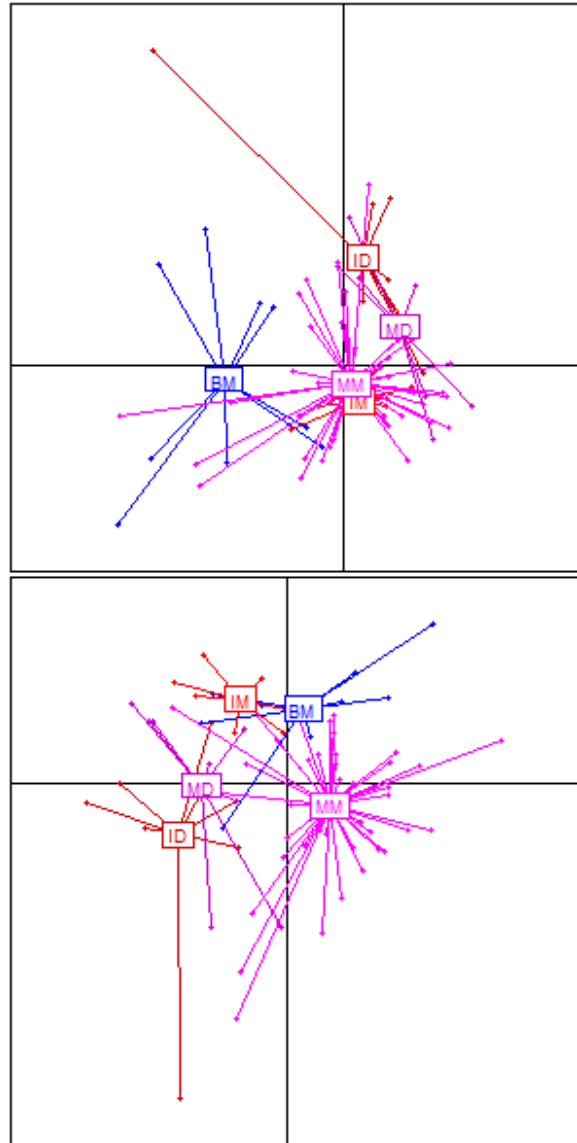
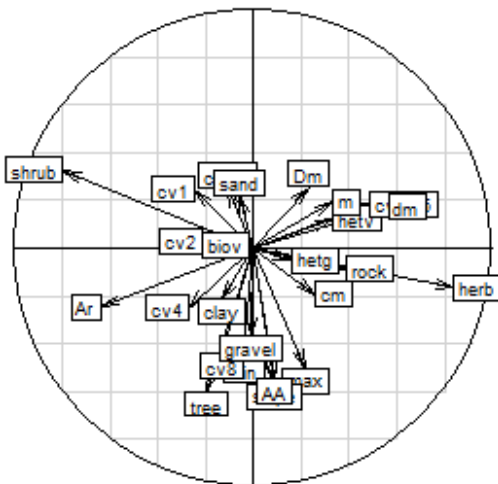
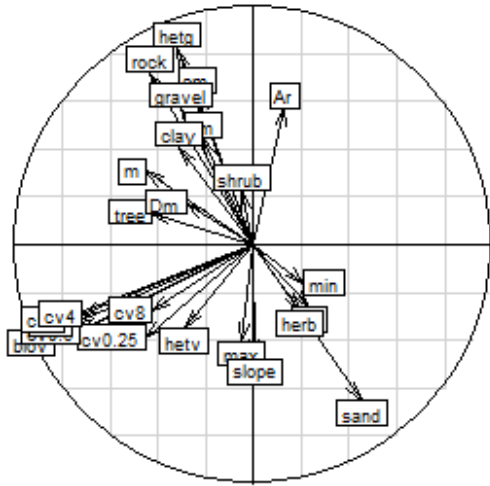


**S2.3) Habitat-specific Principal Component analysis (PCA) based on 24 environmental factors plus aridity. First row: right = First plane (x-axis = first axis called PC1, y-axis = second axis called PC2) and corresponding correlation circle (left); second row: right = PC3 (x-axis) + PC4 (y-axis) and corresponding correlation circle (left); Legend: BM = Balearic Islands - Mediterranean; ID = Israel - Desertic; IM = Israel - Mediterranean; MD = Morocco - Desertic; MM = Morocco - Mediterranean; see Appendix S1 for variable definitions.**

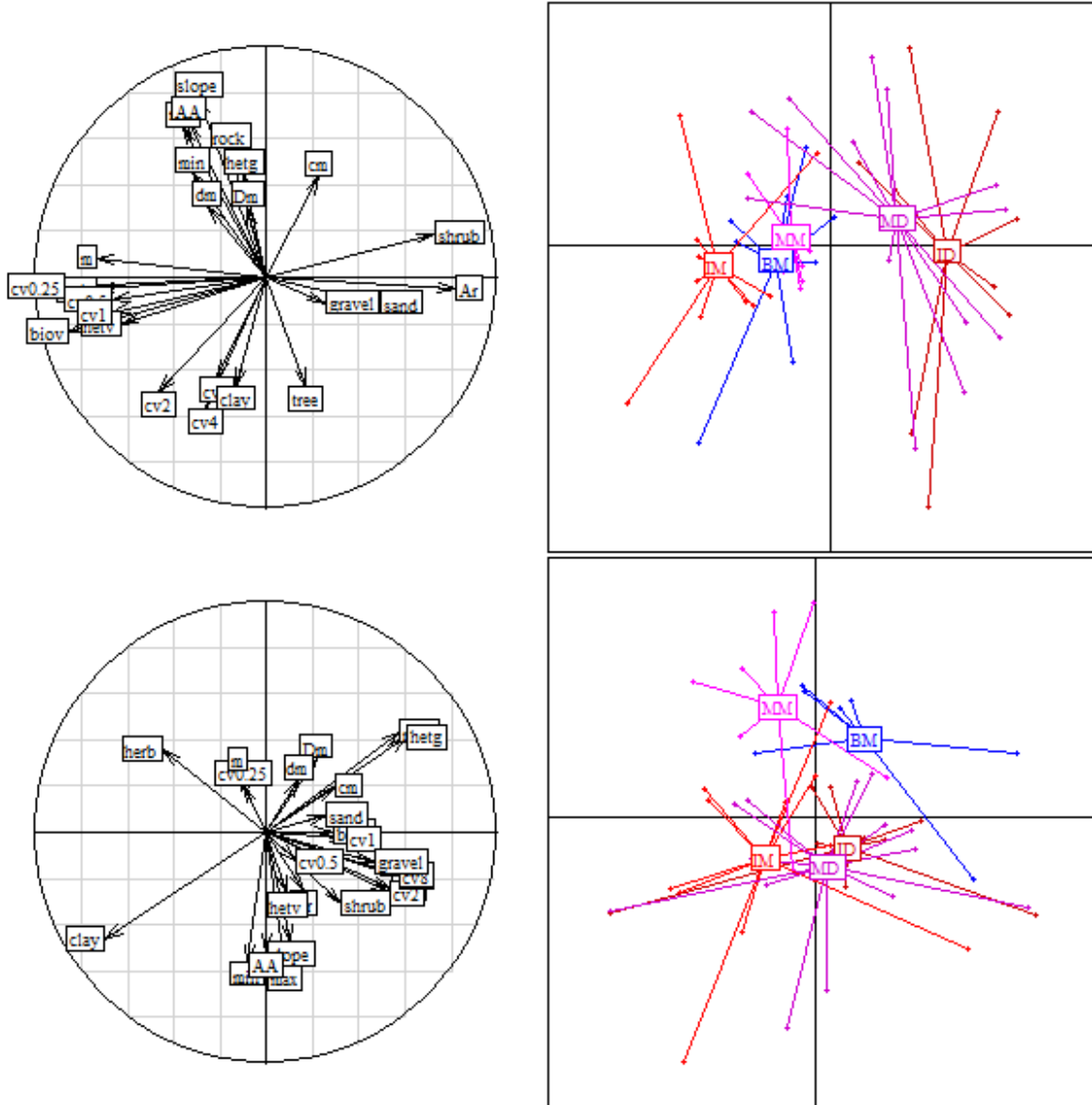
**S2.3.1) Crop.** Four retained axes explain a total of 58.6 % of total variance (19.7, 15.8, 13.5 and 9.6 % of total variance for each individual axis).



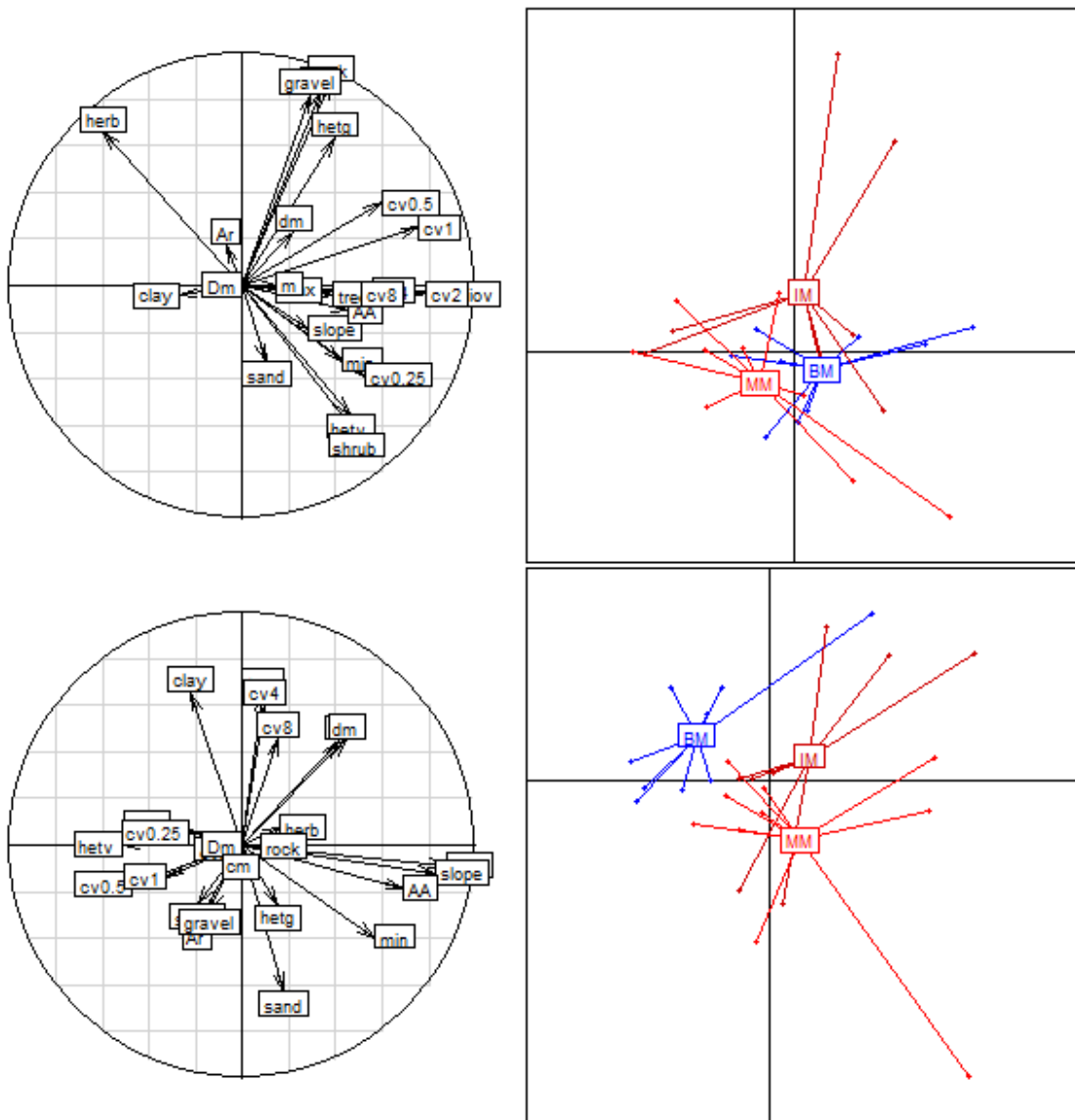
S2.3.2) **Dune.** Four retained axes explain a total of 60.8 % of total variance (19.3, 18.9, 12.7 and 9.9 % of total variance for each individual axis).



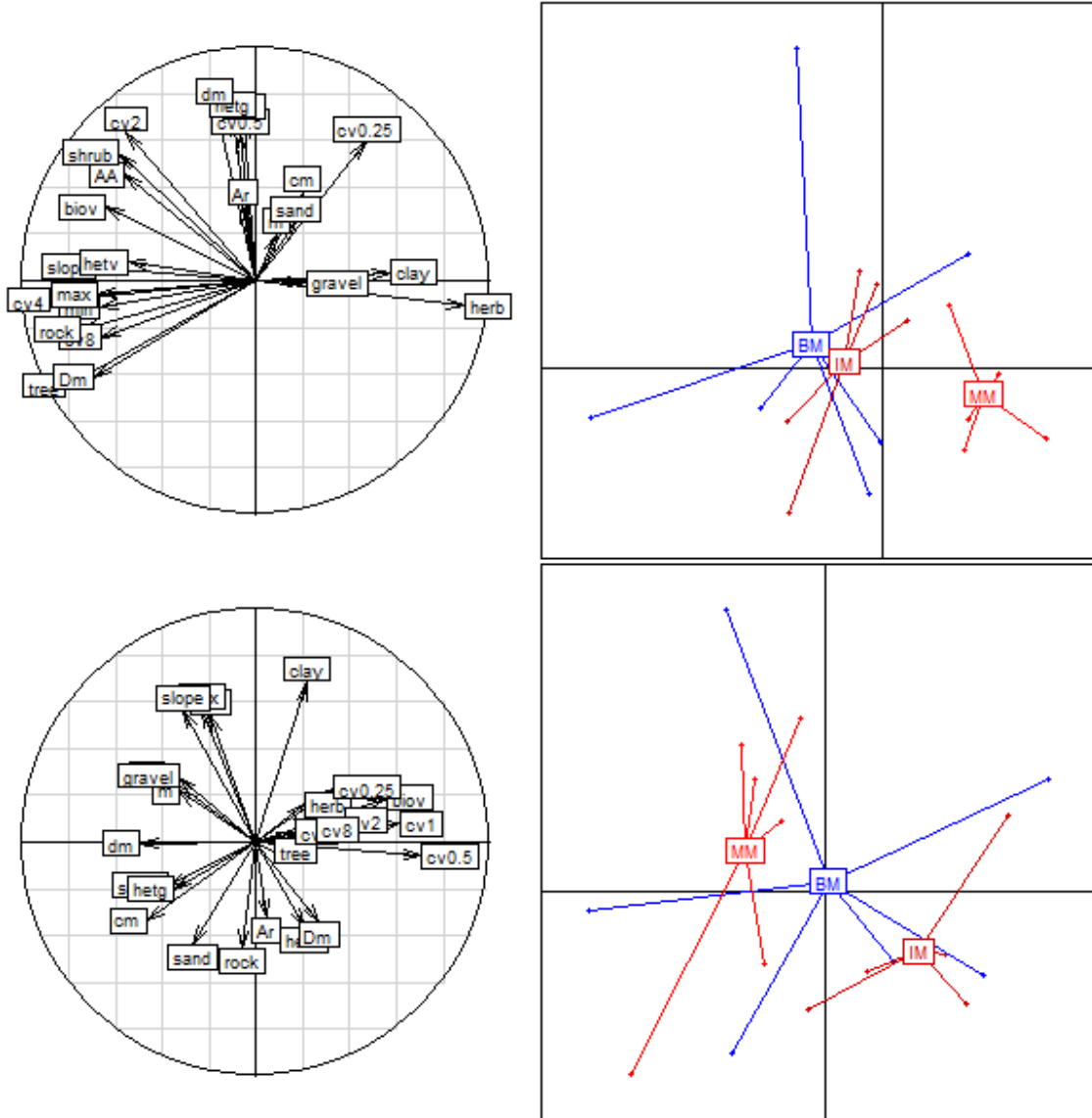
S2.3.3) Batha. Four retained axes explain a total of 65.8 % of total variance (25.5, 15.9, 14.0 and 10.4 % of total variance for each individual axis).



S2.3.4) Fallow. Four retained axes explain a total of 67.3 % of total variance (22.3, 18.1, 15.2 and 11.8 % of total variance for each individual axis).



S2.3.5) Anthropogenic. Four retained axes explain a total of 69.9 % of total variance (30.0, 16.6, 12.6 and 10.7 % of total variance for each individual axis).



### Appendix S3 Landscape analysis

#### S3.1) Column coordinates of the Principal Component analysis (PCA) performed on the habitat preference class (DUN), based on 32 environmental factors plus aridity

variable	r(PC1.DUN, var)	r(PC2.DUN, var)	r(PC3.DUN, var)	r(PC4.DUN, var)
aridity	-0.15	-0.24	<b>0.82</b>	0.23
beach	-0.08	0.29	-0.20	-0.38
white	-0.24	-0.21	-0.25	<b>-0.61</b>
grey	0.28	0.03	0.21	<b>0.81</b>
pSaltflat	-0.04	0.02	0.45	-0.22
pFallow	0.00	0.00	0.00	0.00
pAnthr	-0.11	0.05	0.08	-0.06
pBatha	-0.14	0.35	-0.10	0.23
pCrop	0.03	0.21	-0.10	-0.32
AA	0.19	-0.39	-0.14	0.55
min	0.13	-0.36	0.06	0.51
max	0.47	-0.10	-0.19	0.30
slope	0.47	-0.29	-0.25	0.58
hetv	0.38	0.17	-0.18	-0.18
herb	-0.13	0.28	<b>-0.83</b>	0.09
shrub	0.02	-0.24	<b>0.87</b>	-0.04
tree	0.49	-0.21	-0.27	-0.14
biovolume	<b>0.87</b>	0.19	-0.11	-0.14
cv0.25	0.46	0.32	-0.24	0.09
cv0.5	<b>0.72</b>	0.31	0.24	-0.02
cv1	<b>0.80</b>	0.27	0.27	-0.16
cv2	<b>0.81</b>	0.22	0.11	-0.15
cv4	<b>0.80</b>	0.00	-0.11	-0.15
cv8	0.52	-0.16	-0.25	-0.12
rock	-0.23	<b>0.81</b>	-0.07	0.31
hetg	-0.20	<b>0.79</b>	0.25	-0.06
Dm	0.01	0.13	0.04	0.05
m	-0.15	0.37	-0.09	0.27
dm	-0.32	0.59	-0.14	0.13
cm	-0.15	<b>0.72</b>	-0.06	0.28
gravel	-0.03	0.57	0.06	0.26
sand	-0.19	<b>-0.64</b>	-0.39	0.05
clay	0.30	0.32	0.44	-0.18

Four retained axes explain a total of 50.1 % of total variance (16.1, 14.0, 10.7 and 9.3 % of total variance for each individual axis).

|Pearson's correlation coefficients| > 0.6 are highlighted in bold. See Appendix S1 for variables definition.

**S3.2) Column coordinates of the Principal Component analysis (PCA) performed for *G. cristata* for the landscape class (LAN), based on 32 environmental factors plus the abundance of the focal species (*G. cristata*) in the landscape**

variable	r(PC1.LAN, var)	r(PC2.LAN, var)	r(PC3.LAN, var)	r(PC4.LAN, var)
beach	-0.05	0.49	0.15	-0.21
white	-0.09	0.14	-0.25	0.26
grey	0.19	0.06	-0.02	-0.18
pSaltflat	0.20	0.00	-0.25	0.49
pFallow	0.32	-0.01	-0.18	-0.01
pAnthr	-0.02	<b>0.73</b>	0.34	-0.05
pBatha	<b>-0.92</b>	-0.09	-0.11	0.12
pCrop	0.40	-0.45	0.22	-0.43
AA	-0.60	-0.26	0.06	-0.39
min	<b>-0.71</b>	-0.32	0.04	-0.02
max	<b>-0.77</b>	-0.32	0.01	-0.19
slope	<b>-0.71</b>	-0.39	0.00	-0.29
hetv	-0.15	-0.10	0.30	0.23
herb	<b>0.62</b>	0.09	0.08	-0.42
shrub	<b>-0.67</b>	-0.14	-0.12	0.40
tree	-0.09	0.56	0.57	-0.12
biovolume	0.08	<b>-0.73</b>	0.49	0.16
cv0.25	0.27	<b>-0.72</b>	0.23	0.02
cv0.5	0.05	<b>-0.72</b>	0.29	-0.14
cv1	-0.02	<b>-0.61</b>	0.27	0.09
cv2	-0.20	-0.38	0.54	0.14
cv4	0.03	0.18	<b>0.82</b>	0.14
cv8	0.03	0.30	<b>0.66</b>	0.29
rock	<b>-0.88</b>	0.22	0.07	0.00
hetg	<b>-0.77</b>	0.23	0.05	0.08
Dm	-0.45	0.04	-0.21	-0.08
m	<b>-0.72</b>	0.00	-0.20	-0.03
dm	<b>-0.89</b>	-0.07	-0.08	0.10
cm	-0.58	0.40	0.17	-0.18
gravel	-0.47	0.27	0.52	0.26
sand	0.16	0.06	0.09	<b>-0.72</b>
clay	0.42	-0.19	-0.13	<b>0.68</b>
cri_AB_L	0.41	0.13	0.13	0.30

Four retained axes explain a total of 54.8 % of total variance (24.1, 13.3, 9.4 and 8.0 % of total variance for each individual axis). | Pearson's correlation coefficients | > 0.6 are highlighted in bold. See Appendix S1 for variables definition.

**S3.3) Column coordinates of the Principal Component analysis (PCA) performed for *G. theklae* for the landscape class (LAN), based on 32 environmental factors plus the abundance of the focal species (*G. theklae*) in the landscape**

variable	r(PC1.LAN, var)	r(PC2.LAN, var)	r(PC3.LAN, var)	r(PC4.LAN, var)
beach	0.04	0.49	0.17	-0.12
white	0.08	0.15	-0.24	0.35
grey	-0.19	0.06	-0.02	-0.07
pSaltflat	-0.21	0.01	-0.23	0.56
pFallow	-0.32	-0.01	-0.19	0.00
pAnthr	0.02	<b>0.73</b>	0.35	-0.04
pBatha	<b>0.92</b>	-0.10	-0.12	0.06
pCrop	-0.40	-0.45	0.22	-0.47
AA	0.60	-0.26	0.06	-0.22
min	<b>0.71</b>	-0.33	0.03	0.01
max	<b>0.76</b>	-0.32	0.02	-0.06
slope	<b>0.70</b>	-0.39	0.00	-0.17
hetv	0.16	-0.11	0.29	0.15
herb	-0.60	0.08	0.06	-0.59
shrub	<b>0.66</b>	-0.13	-0.10	0.56
tree	0.09	0.56	0.59	-0.08
biovolume	-0.08	<b>-0.73</b>	0.48	0.09
cv0.25	-0.27	<b>-0.72</b>	0.22	-0.05
cv0.5	-0.06	<b>-0.72</b>	0.30	-0.12
cv1	0.01	-0.60	0.28	0.13
cv2	0.19	-0.38	0.54	0.21
cv4	-0.02	0.17	<b>0.81</b>	0.07
cv8	-0.02	0.29	<b>0.65</b>	0.24
rock	<b>0.89</b>	0.20	0.05	-0.13
hetg	<b>0.79</b>	0.21	0.03	-0.03
Dm	0.46	0.03	-0.23	-0.25
m	<b>0.73</b>	-0.02	-0.24	-0.23
dm	<b>0.90</b>	-0.08	-0.10	0.01
cm	0.57	0.40	0.18	-0.13
gravel	0.48	0.26	0.51	0.23
sand	-0.17	0.06	0.11	-0.56
clay	-0.42	-0.19	-0.13	0.60
the_AB_L	0.22	-0.11	-0.08	-0.54

Four retained axes explain a total of 54.8 % of total variance (23.8, 13.3, 9.3 and 8.4 % of total variance for each individual axis).

| Pearson's correlation coefficients | > 0.6 are highlighted in bold. See Appendix S1 for variables included.



**S3.4) Overdispersion-adjusted Poisson regression coefficients (Estimate) for models predicting the abundance of 'Species' (*G. cristata* or *G. theklae*) in the dune**

Species	Factor (F)	Class	signature var. (sv)	r(sv, F)	Estimate	SE	t	Interpretation
<i>cristata</i>	PC1.DUN	DUN	biovolume	0.87	-0.39	0.15	-2.62	- vegetation cover
	PC2.DUN	DUN	rock	0.81	0.02	0.14	0.12	/
	PC3.DUN	DUN	aridity	0.82	-0.51	0.30	-1.69	- aridity
	PC4.DUN	DUN	grey	0.81	0.50	0.17	2.89	+ grey dunes
	PC1.LAN	LAN	pBatha	-0.92	1.80	0.58	3.11	- ruggedness
	PC2.LAN	LAN	pAnthr	0.73	0.04	0.18	0.21	/
	PC3.LAN	LAN	cv4	0.82	0.13	0.12	1.09	+ tree cover
	PC4.LAN	LAN	clay	0.68	0.33	0.22	1.49	+ fine-grained soils
	PC1.COMP	COMP	the_AB_D	0.85	-0.92	0.31	-2.91	- competition
	<i>theklae</i>	PC1.DUN	DUN	biovolume	0.87	-0.26	0.25	-1.04
PC2.DUN		DUN	rock	0.81	0.01	0.18	0.06	/
PC3.DUN		DUN	aridity	0.82	0.32	0.23	1.39	+ aridity
PC4.DUN		DUN	grey	0.81	0.42	0.19	2.16	+ grey dunes
PC1.LAN		LAN	pBatha	0.92	-0.48	0.24	-2.00	- ruggedness
PC2.LAN		LAN	pAnthr	0.73	-0.31	0.19	-1.59	- Anthropic
PC3.LAN		LAN	cv4	0.81	0.57	0.23	2.46	+ tree cover
PC4.LAN		LAN	clay	0.60	-0.19	0.19	-0.97	- fine-grained soils
PC1.COMP		COMP	cri_AB_D	-0.94	0.97	0.39	2.46	- competition

The models presented here correspond to three classes of factors ('global') + 'lower bound' measures of competition. All classes of factors were included. 'Signature' variable (sv) is a raw variable strongly correlated to the explanatory Factor (F) resulting from a within-class Principal Component Analysis; r(sv, F) is the corresponding Pearson's correlation coefficient; t = Estimate / SE. The 'Interpretation' column gives biological interpretation; e.g. '- aridity' indicates that the presence of *G. cristata* in the dune is negatively influenced by environmental aridity; no interpretation is given when Estimate ~ 0. See Appendix S1 for summary of variables, Appendices S3.1–3 for details on Principal Components Analyses, and Table 4 for results of the 'separate' analysis.