History of the Crested Lark in the Mediterranean region as revealed by mtDNA sequences and morphology

Alban Guillaumet a,*, Jean-Marc Pons b, Bernard Godelle a, Pierre-Andre Crochet c

a Laboratoire “Génome, Populations, Interactions, Adaptation”, CNRS UMR 5171, Université Montpellier II, C. C. 63, 34095 Montpellier cedex 5, France
b Origine, Structure et Evolution de la biodiversité, UMR 5202, C. P. 51, 55 rue Buffon & Service de Systématique moléculaire, IFR 101 CNRS, 43 rue Cuvier, 75005 Paris, France
c CEFE, CNRS UMR 5175, 1919 route de Mende, 34293 Montpellier cedex 5, France

Received 4 August 2005; revised 18 December 2005; accepted 4 January 2006
Available online 13 February 2006

Abstract

The Crested Lark has a very complex taxonomy, partly as a result of a strong variation in plumage ground color seemingly linked with environmental factors. However, large variations in body size and bill shape further complicate the situation in the Maghreb. In this paper, we first present a set of hypotheses to explain patterns of morphological variation around the Mediterranean Sea. A phylogeographical analysis covering all major biogeographical areas in the species’ range is then performed to test these scenarios. Three mtDNA groups with distinct geographical distribution were identified. The randonii clade (= G. (c.) randonii) is endemic from central Maghreb and is phylogenetically basal relative to cristata and senegallensis. These two latter groups are much more widespread. The cristata clade is found in NW Morocco, throughout Europe and W Asia and in NE Africa, while senegallensis regroups the populations sampled in the Western Sub-Saharan Africa and in NE Maghreb (E Algeria, Tunisia). A combination of genetic and paleoenvironmental evidences supports a scenario of allopatric differentiation of these two lineages outside the Maghreb, with subsequent range expansion leading to their secondary presence in the Maghreb. However, the alternative hypothesis of differentiation in two, or even three separate Maghreb refuges cannot be completely dismissed with the present data. Interestingly, the Sahara desert and the Gibraltar Strait did not act as permanent barriers to dispersal in this species. In addition, the populations in the Maghreb are consistently longer-billed than their closest relatives, suggesting a role for natural selection or phenotypic plasticity.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Crested Lark; Galerida cristata; Phylogeography; Cytochrome b; Morphology; Maghreb; Sahara; Gibraltar Strait

1. Introduction

Phylogeographical studies have brought many insights into the Quaternary history of species in various regions of the world, but especially in North America and in Europe. Similar studies in other parts of the world are currently few, and documenting history of biotas in these poorly known regions provide a considerable challenge (Hewitt, 2001). One of such region is the Maghreb in North Africa. The Maghreb is limited to the West by the Atlantic Ocean, to the North by the Mediterranean Sea, and to the East and South by the Lybian and Saharan deserts, respectively. The Maghreb is actually separated from the Iberian Peninsula by a gap of 14 km of open sea. Despite this short distance, recent phylogeographical studies revealed that the Gibraltar Strait often constitutes a major barrier to gene flow, even in flying species such as birds (in the Pied Flycatcher Ficedula hypoleuca, Saetre et al., 2001; in the Blue Tit Parus caeruleus, Salzburger et al., 2002; but not in the Bearded Vulture Gypaetus barbatus, Godoy et al., 2004), and bats (in the Greater Mouse-eared bat Myotis myotis, Castella et al., 2000; in the ‘austriacus’ group of long-eared bats of genus Plecotus, Juste et al., 2004). The role of the Sahara desert in promoting genetic and morphological divergences in North-African taxa is a question of great biogeographical
interest that is still poorly understood (Douady et al., 2003). The Sahara currently constitute a strong barrier to dispersal for organisms that are not adapted to live in the most arid parts of the deserts. Nevertheless, changes in aridity level and extent of the desert have taken place repeatedly in the past. Early phases of aridification trace back to the mid-Miocene, and may have triggered the diversification of some species through vicariance from that time (Douady et al., 2003). Latter, the Sahara experienced several phases of contractions during the Pleistocene, as a response to climatic oscillations (Jolly et al., 1998a; Tiedemann et al., 1989). Dispersal and genetic exchange have probably been possible during wetter phases between populations currently separated by a wide gap of desert. However, relevant phylogeographic studies are lacking to test this hypothesis (see Franck et al., 2001 for an exception).

As well as the restrictions to gene flow between the Maghreb and the other regions discussed above, it is also possible that some barriers between populations occur within the Maghreb. Paleoenvironmental data and climatic reconstructions show that the Maghreb underwent drastic ecological changes during Pleistocene climatic oscillations (Jolly et al., 1998a). These fluctuations could have led to allopatric differentiation in separate Maghreb refuges, although this phenomenon is currently poorly understood (Cosson et al., 2005). While some studies reveal clear patterns of East-West vicariance in the Maghreb (in the land snail Helix aspersa: Guiller et al., 2001; in the greater white-toothed shrew: Cosson et al., 2005), others do not (in the Woodmouse Apodemus sylvaticus: Libois et al., 2001; in the Chaffinch Fringilla coelebs: Griswold and Baker, 2002). Clearly, more phylogeographical studies of North-African taxa are needed to search and identify concordant patterns among species, and unravel the influence of taxonomy or different life-histories on current distribution of genetic diversity within North-African species or species complexes.

The Crested Lark (Galerida cristata) is a very good model to address these questions. This species has one of the most complex subspecific taxonomy within birds, with over 60 races described across the species range (Del Hoyo et al., 2004). In general, differences among races are slight, but notable exceptions do exist. In Europe, Middle East, North-Eastern Africa, and sub-Saharan Africa, the subspecies are short-billed and differ primarily in intensity of streaking, and plumage ground color (with variation often matching that of the substrate, see e.g., Cramp, 1988). In the Maghreb however, large variations in bill size and shape also contribute to the delimitation of races. This study aims at understanding the origin of morphological variation in the Crested Lark around the Mediterranean Sea with a special attention paid on Northern Africa. It should also provide valuable information on the history of past fragmentation and recent colonization in a North-African bird.

A recent phylogeographical analysis of the Crested Lark in Morocco (Guillaumet et al., 2005) identified two well-differentiated evolutionary units: Galerida (c.) cristata, which regroups three small-billed subspecies and lives in the NW of the country, and G. (c.) randonii, which regroups two large-billed subspecies, and replaces the former in more arid regions. These two semi-distinct races have probably diverged about 1 My ago. They have reached reciprocal monophyly in mtDNA and show evidence of strong reproductive isolation in their narrow contact zone, although hybridization is apparently still occurring. However, due to the restricted geographic coverage of this study, the question of the biogeographical origin of these taxa remained unresolved. In particular, northernmost populations of Galerida (c.) cristata in Morocco are often synonymized with the similarly small-billed populations of the “European group” of the Crested Lark (see e.g., Cramp, 1988), suggesting that Galerida (c.) cristata might be a (recent) immigrant from Western Europe. Some alternative sources are also possible. For instance, populations south of the Sahara are also short-billed (e.g., subspecies senegallensis or alexanderi in the “Sahel group,” see our Fig. 1). We can thus raise the hypothesis that populations North and South of the Sahara are currently connected, or that they were connected in the recent past.

Tunisia is another critical place for a better understanding of polymorphism in bill morphology in the Crested Lark. The populations in this country are traditionally divided into two subspecies—carthaginis along the Mediterranean coast and arenicola in desert areas (see e.g., Del Hoyo et al., 2004)—which are mainly distinguished on the basis of their coloration (paler in arenicola). Both subspecies have similar bill length but are clearly intermediate between Galerida (c.) cristata (small-billed) and G. (c.) randonii (large-billed) (see Cramp, 1988).

Several biogeographical hypotheses may account for intermediate bill size in Tunisia. First, Tunisian populations might represent the endpoint of a clinal variation of G. (c.) randonii or of G. (c.) cristata populations (Fig. 1). Second, long-billed G. (c.) randonii and short-billed G. (c.) cristata populations may have diverged in allopatry and two distinct expansion pathways could have resulted in two distinct contact zones with strong reproductive isolation or recent contact in the West (Morocco) producing a bimodal distribution of bill size, but more or less complete mixing in the East (Tunisia) producing a population with intermediate bill size. Third, East-West vicariance in the Maghreb might have lead to the formation of a third, morphologically intermediate, lineage in Tunisia. Last, the geographic pattern of morphological variation of the Crested Lark around the Mediterranean Sea might correspond to a ring species (or an “embedded” ring species, sensu Patten et al., 2004), because (i) there is no major gap in the current distribution of the species around the Mediterranean Sea; (ii) the two most distinctive forms meet in Morocco, where they show evidence of reproductive isolation (Guillaumet et al., 2005); (iii) a chain of intermediate forms seems to connect them through a fairly continuous geographic ring (Fig. 1; see Irwin and Irwin, 2002 for the criteria a ring species is expected to fulfil).
These hypotheses lead to different predictions about the genetic composition of the populations in Tunisia. The “clinal” model predicts that only cristata (or randonii) haplotypes will be found in Tunisia, the “secondary contact” scenario instead predicts a mixture of cristata and randonii haplotypes, while the “new lineage” model predicts that original haplotypes will be discovered. Last, the “ring-species” model predicts that geographically adjacent populations should be genetically more closely related to each other than to more distant populations around the Mediterranean basin, except at the end of the ring in Morocco.

In this study, we first detail the patterns of morphological variation that allowed us to generate candidate biogeographical scenarios for the distribution of the Crested Lark complex in North-Africa. We then use a phylogeography of the Crested Lark based on mitochondrial DNA to test these scenarios. Combination of morphological and genetical data is a powerful tool to decipher the role of history and selection in shaping morphological variation, in this case variation in bill size in the Crested Lark. As few phylogeographical studies have been conducted in Africa, this study will also bring new insights into the footprints left by the Quaternary climatic oscillations in North Africa, and the history of avian populations currently separated by the Sahara.

2. Materials and methods

2.1. Samples

A total of 160 Crested Lark specimens were included in this study. Sampling was designed to cover most of the distribution of the Crested Lark with a special attention paid on the Mediterranean basin.

Moroccan specimens were collected under license for the present project (see Guillaumet et al., 2005). Most of the additional specimens consists of museum skins conserved at the French Muséum National d’Histoire Naturelle in Paris. A few blood or muscle samples have been sent directly to us by contributors (see names in Acknowledgements). Last, we included in our analysis previously published cyt b sequence of G. cristata from Morocco (AY769746, AY769748, AY769749), and Saudi Arabia (GenBank AY165151).

We obtained complete morphological measurements for 117 specimens and partial cyt b sequences for 65 specimens. Sample size for each data set and country is given in Fig. 2 (see Appendix 1 for additional information including voucher/specimen number).

2.2. Morphology

Whenever possible, the specimens used in genetic analysis were measured for six morphological variables (juveniles were not measured, and some specimens had missing data for one or more variables). Measurements were done by the same author (A.G.) using a caliper (to the nearest 0.1 mm) or a ruler. These variables are: bill length (BL) from bill tip to distal edge of nostrils, bill depth (BD) at distal edge of nostrils, bill width (BW) at distal edge of nostrils on the upper mandible, tarsus length (TarL), tail length (TailL) using the T(R) method, and wing length (WL) using the maximum length method (Svensson, 1992).

To derive the candidate scenarios presented in Section 1, specimens were grouped into six populations: Europe (EU), Asia (AS), Eastern Maghreb (East Algeria and Tunisia,
A. Guillaumet et al. / Molecular Phylogenetics and Evolution 39 (2006) 645–656

EM), Western Sahel (from Senegal to W Chad, WS), and in Morocco: G. (c.) cristata (MO), and G. (c.) randonii (MR) (see Guillaumet et al., 2005, for Moroccan specimens). Because of a slight size dimorphism in Galerida larks, and to present a single value for each population, bill length values were calculated as the mean of the males and females values per population (Table 1).

Genetic analysis revealed that each of our six populations is characterized by a single or a largely dominant mtDNA haplotype. Three genetic groups were thus defined on the basis of dominant haplotypes (see Section 3). Morphological variation was then assessed as follows. First, we performed a principal component analysis (PCA) and a discriminant function analysis (DA) on the six morphological variables. The results of these two analyses being very comparable, we only present the results of the PCA. Differences among the genetic groups were assessed using two-way ANOVAs and Tukey tests performed on the first and second principal components of the PCA (PC1 and PC2, taken as dependent variables), while “sex” and “group” were treated as categorical variables. Within some genetic groups represented in several regions, differences among regions were assessed using the same two-way ANOVAs and Tukey tests as described above. In all cases, normality of the distributions was checked before analysis by a Shapiro–Wilk test.

Last, the morphological similarity among taxa or regions was assessed by constructing a morphological tree. Its topology was then compared to the phylogeny derived from the cyt b, to assess whether morphology accurately reflects history between populations. The morphological tree was based on Mahalanobis (squared) distances as implemented in the discriminant function module of STATISTICA (version 5 © StatSoft, Inc). Mahalanobis distance is appropriate here because it accounts for the correlation between variables. The resulting matrix was then used to reconstruct a tree with R version 2.0.1 (© The R Foundation for Statistical Computing). Seven reconstruction methods were tested, and five of them (including UPGMA) yielded the same topology, so we only present the results obtained with the UPGMA method. The R software was also used for all other statistical treatments.

Table 1

<table>
<thead>
<tr>
<th>Group</th>
<th>Region</th>
<th>n</th>
<th>Bill length</th>
<th>Bill depth</th>
<th>Bill width</th>
<th>Tarsus length</th>
<th>Tail length</th>
<th>Wing length</th>
<th>PC1</th>
<th>PC2</th>
</tr>
</thead>
<tbody>
<tr>
<td>crisata</td>
<td>Europe</td>
<td>7–10</td>
<td>12.0</td>
<td>5.1</td>
<td>4.4</td>
<td>24.9</td>
<td>58.8</td>
<td>101.3</td>
<td>−1.3</td>
<td>−0.3</td>
</tr>
<tr>
<td></td>
<td>Asia</td>
<td>14–17</td>
<td>13.0</td>
<td>5.1</td>
<td>4.4</td>
<td>25.2</td>
<td>65.6</td>
<td>107.0</td>
<td>−0.5</td>
<td>−1.0</td>
</tr>
<tr>
<td></td>
<td>Morocco (MO)</td>
<td>24–26</td>
<td>13.0</td>
<td>5.4</td>
<td>4.4</td>
<td>25.5</td>
<td>63.6</td>
<td>104.2</td>
<td>−0.6</td>
<td>−0.4</td>
</tr>
<tr>
<td>senegallensis</td>
<td>Maghreb</td>
<td>37–52</td>
<td>14.0</td>
<td>5.5</td>
<td>4.7</td>
<td>25.9</td>
<td>62.8</td>
<td>105.2</td>
<td>−0.1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>W Sahel</td>
<td>9–19</td>
<td>11.9</td>
<td>5.6</td>
<td>4.7</td>
<td>23.7</td>
<td>57.0</td>
<td>99.2</td>
<td>−1.2</td>
<td>1.9</td>
</tr>
<tr>
<td>randonii</td>
<td>Morocco (MR)</td>
<td>24–25</td>
<td>16.3</td>
<td>5.9</td>
<td>4.9</td>
<td>26.7</td>
<td>67.5</td>
<td>109.7</td>
<td>1.0</td>
<td>0.2</td>
</tr>
</tbody>
</table>

n corresponds to the number of specimens measured for each variable. Mean of males and females mean values are given. PC1, first axis of the PCA; PC2, second axis of the PCA.
2.3. Genetics

DNA extraction and sequencing was performed as follows. For fresh tissue samples, the protocol described in Guillaumet et al. (2005) was applied (protocol I). Briefly, total DNA was isolated using Sigma GenElute mammalian genomic DNA kit. A cyt b fragment of 1100 bp was amplified by the polymerase chain reaction using the primers A (chicken position L-14995) and F (H-16065) (Helbig and Seibold, 1999). Reactions were performed in 50 μl volumes using 5 μl of 10× buffer, 2.5 mM MgCl₂, 0.2 mM of dNTPs, 20 pmol of each primer, and 1 U of Taq DNA polymerase (Promega), at an annealing temperature of 52°C. A 291 bp fragment was then sequenced using primer F as sequencing primer. Amplification products were read with a Pharmacia LKB A.L.F automatic DNA sequencer following recommended procedures.

Because museum bird specimens generally contain minute quantities of DNA, a specific procedure (protocol II) adapted from Mundy et al. (1997) was applied. DNA was extracted from museum specimens (toe pad) using the CTAB procedure (Winneppeninckx et al., 1993). A 291 bp fragment of the cyt b was then amplified using primers cyt-H’ (L-15728; 5’-CAC ATT AAA CCA GAA TAC TTC CTC TT-3’) and F. The primer cyt-H’ corresponds roughly to the primer mt-H of Helbig et al. (1995), which we modified to better match Galerida sequences. The amplifications were performed in a final volume of 50 μl. Cycling conditions were 92°C for 40 s, 52°C for 40 s, and 72°C for 60 s for 45 cycles. After purification [QiaQuick PCR Purification Kit, Qiagen (Holden, Germany)], direct sequencing was performed with the same primers was performed on an automated sequencer following the supplier’s procedures (Beckmann Coulter, Fullerton, CA, USA).

Three major groups that differ by a minimum of three diagnostic substitutions were identified with the 291 bp alignment (see Section 3). To assess the phylogenetic relationships between these groups, we obtained longer sequences for a subset of specimens of each group. We obtained (from the 3’ end) 504 bp in at least four specimens of each group, and 715 bp in at least two specimens of each group. For fresh specimens, this was done by sequencing with both primers A and F. For museum specimens, the cyt b was divided into three contiguous and overlapping fragments: A’C, BD, and H’F. New internal primers were specifically designed to amplify and sequence these fragments: cyt-A’ (L 15246, 5’-ATC GGC CGA GGA NTC TAC TAC-3’), cyt-B (L 15413; 5’-GAA TGA GCA TGA GGC GGA TTC TCA g-3’), cyt-D (H15897; 5’-ATT TGT GAG AGG GGA CGG AAG GTT ATT GA-3’), and cyt-C (H15540; 5’-CCT GTT TCG TGG AGG GTG AGG TG-3’). The 715 bp results from concatenation of the three fragments (BD was sequenced with D and A’C was sequenced with C). Because statistical support was not improved with 715 bp (when compared with the 504 bp alignment), we only presents the results obtained with the latter.

For fresh specimens, we are confident that our cyt b sequences are of mitochondrial origin because no stop codon were found except at the end of the cyt b gene, and the base composition and the pattern of base substitution were typical of active mitochondrial genes (Guillaumet et al., 2005). To ensure that the fragments amplified from museum specimens were also of mitochondrial origin, we compared the sequences obtained in three specimens (two G. (c.) cristata and one G. (c.) randonii sampled in 2002 in Morocco): (i) using pectoral muscle (protocol I), and (ii) using pieces of toe pad, after the specimens were prepared at the Paris museum (protocol II).

Because multiple hits were detected in our data set, major groups and alternative potential evolutionary paths among the haplotypes sampled were identified using a median-joining network (Bandelt et al., 1999), with the parameter ε set to 1. Net divergence between genetic groups (Da, Nei, 1987) and tests of neutrality performed on each group separately (Tajima’s D, 1989; Fu and Li’s D* and F*, 1993) were performed with DnaSP (Rozas and Rozas, 1999). Divergence was estimated with the longer alignment available (715 bp), while neutrality tests were performed with the 291 bp data set. Phylogenetic relationships between haplotypes was inferred using neighbor-joining (NJ), maximum parsimony (MP), and maximum likelihood (ML) methods using Phylowin (Galtier et al., 1996). One sequence of Galerida magnirostris (GenBank AY165169) and two sequences of Galerida theklae (GenBank AY769740, AY769741) were used as outgroups. Pairwise genetic distances were calculated with Kimura’s two parameters model (1980). For ML, a transition over transversion ratio of 3.4 (corresponding to the estimated value for the genus Galerida) was applied. The tree building ML algorithm in Phylowin is that of the FASTDNAML program (Olsen et al., 1994). Bootstrap tests with 5000 replicates (500 for ML) were performed to assess the robustness of the clades (i.e., bootstrap support: BS). We also performed maximum likelihood analysis with a more sophisticated model as implemented in PHYML (Guindon and Gascuel, 2003). We used the TN93 model of nucleotides substitution (Tamura and Nei, 1993) and a discrete gamma model with eight categories (500 bootstrap replications). The gamma shape parameter and the transition/transversion ratio were estimated from the data. However, we obtained identical results with this method and thus only present the results obtained with Phylowin.

3. Results

3.1. Genetics

We are confident that the sequences obtained from museum specimens cannot be attributed to contamination because old DNA (museum skins) was manipulated in a separate laboratory where no fresh samples had been handled, because the negative extraction controls did not yield any band when used as PCR templates, and because original
haplotypes were found in museum specimens, including one geographically restricted lineage that is phylogenetically nested within other sequences of Crested Lark, but was not found in fresh specimens (see below). Furthermore, the sequences obtained in three specimens using both pectoral muscle and museum skin were perfectly identical.

In a first step, 291 bp of the cyt b gene were aligned for 64 of the 65 sequenced specimens. Fourteen variable sites defining eight haplotypes were detected in the Crested Lark (Table 2). A median-joining network was reconstructed with these eight haplotypes. The genetic variation in the Crested Lark is subdivided into three major groups that differ by a minimum of three diagnostic substitutions (Fig. 3 and Table 2). Each mtDNA group is characterized by a common haplotype sampled throughout the range of the group, and by zero to three rarer haplotypes present in single individuals. The first group (hereafter called the randonii group) corresponds to a single haplotype (randonii) found only in G. (c.) randonii from SE Morocco (n = 18). The second group (cristata group) regroups three haplotypes. The most frequent (cristata) is found in NW Morocco (n = 16), but also in Europe [Portugal (1), France (1)], Western Asia [Turkey (2), Iran (6), Iraq (1), Kazakhstan (1), Saudi Arabia (1), Israel (1)], and NE Africa [Egypt (1), E Chad (1)]. Two rarer haplotypes were found, respectively, in Iran (1) and Morocco (1). The specimen from Yugoslavia with a 153 bp fragment could be unambiguously ascribed to the cristata group because it did not differ from the haplotype cristata, while it differed from the haplotypes of other groups by a minimum of two diagnostic mutations (not shown). The third group is called the senegallensis group because senegallensis Muller 1776 is the older available name for it. It groups four haplotypes. The most frequent is found in Tunisia (n = 5), Algeria (1), Mali (1), Mauritania (1), and Senegal (2). The three other haplotypes were found in single specimens from Chad (W), Mali, and Tunisia. These three groups are thus found in distinct parts of the range of the Crested Lark complex. We did not find areas of overlap between these groups (Fig. 4), except for a narrow contact zone between the cristata and randonii groups in Morocco (see Guillaumet et al., 2005).

3.2. Phylogeny

Sequences have been deposited in GenBank under Accession Nos. DQ028951–DQ028957. For the randonii group, neutrality tests were not performed as no polymorphism was detected. For the cristata and senegallensis groups, the statistics were consistently negative, indicating a tendency for an excess of singletons (not shown). This tendency reached significance for D* and F* in the cristata group (D* = −2.75, P < 0.05; F* = −2.85, P < 0.05). For all other tests, D in both groups, D* and F* in the senegallensis

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>GenBank</th>
<th>Country</th>
<th>Nucleotide position</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>randonii</td>
<td>AY769749</td>
<td>Morocco</td>
<td>T A T G T A C T A C T G C</td>
<td>18</td>
</tr>
<tr>
<td>Tu6</td>
<td>DQ028955</td>
<td>Tunisia</td>
<td>C . C A C G . . . A G . A T</td>
<td>1</td>
</tr>
<tr>
<td>Ch1</td>
<td>DQ028954</td>
<td>Chad</td>
<td>C C C A C A G . . . G . A T</td>
<td>1</td>
</tr>
<tr>
<td>Ma2</td>
<td>DQ028956</td>
<td>Mali</td>
<td>C . C A C G . . . . A T</td>
<td>1</td>
</tr>
</tbody>
</table>

The haplotype cristata (*1) was sampled in Morocco, Portugal, France, Turkey, Iran, Iraq, Kazakhstan, Saudi Arabia, Israel, Egypt, Chad; the haplotype senegallensis (*2) was sampled in Tunisia, Algeria, Mauritania, Senegal, Mali. n corresponds to sample size.
phylogenetic status of the group *G. (c.) randonii* (R) appears to be the sister taxon to the monophyletic group containing the remaining *G. (c.) cristata* (C) and *G. (c.) senegallensis* (S). However, birds from the Eastern Maghreb are larger than their sub-Saharan counterparts in all dimensions (they are in particular much longer-billed), except bill depth and bill width (see Table 1). Accordingly, the former have significantly larger value for PC1 but the reverse is true for PC2 (Table 4).

**3.3. Morphology**

Normality of the data was not rejected for any variable in any taxon or region after correction for multiple tests (not shown). The first principal component of the PCA represents 59.9% of the total variance. All variables are positively and strongly correlated to PC1 ($r > 0.67$) so that it is a size axis. In contrast, PC2 is a shape axis: birds with proportionately deeper and wider bill obtain the maximum PC2 scores (PC2 is positively correlated with bill depth $r = 0.58$ and bill width $r = 0.56$, but negatively with tail length $r = -0.41$ and tarsus length $r = -0.35$).

Morphological differences among genetic groups are strong and significant for all synthetic variables, and all but one comparisons are significant (the *senegallensis* and *randonii* groups do not significantly differ for PC2, see Table 3). However, no group is perfectly diagnosable in the first plane of the PCA (Fig. 6). Discriminant function performs better (not shown), but still retains some overlaps as exemplified by the *senegallensis* group: longer-billed specimens (e.g., Tu34, BL = 16.9) are misclassified as *randonii*, while specimens with shortest bill (e.g., Ch1, BL = 9.8) are misclassified as *cristata*.

We also found morphological differences among regions within mtDNA groups (Table 4; see also Fig. 6). In the *cristata* group, European birds are significantly smaller than those sampled in Asia or Morocco (PC1), and Asian birds have a proportionally weaker bill than the remaining populations (PC2). In the *senegallensis* group, birds from the Eastern Maghreb are larger than their sub-Saharan counterparts in all dimensions (they are in particular much longer-billed), except bill depth and bill width (see Table 1). The morphological tree (Fig. 7) is only partially congruent with the mtDNA tree (Fig. 5): the three regions that harbor a *cristata* haplotype cluster in a single group; however, birds of the *senegallensis* group from E Maghreb cluster with the *randonii* group and not with their closest relatives from W Sahel.
4. Discussion

4.1. A Mediterranean origin for G. (c.) cristata in Morocco

Two semi-species in the Crested Lark complex have been recently identified in Morocco: G. (c.) cristata lives in the NW of the country, and is replaced in more arid regions by the much larger G. (c.) randonii (see Guillaumet et al., 2005). One of the aim of this study was to unravel the extent of their actual distribution, and also to address the question of the origin of G. (c.) cristata in Morocco (see Section 1).

Galerida (c.) randonii turned out to be endemic of the central Maghreb, with its main repartition in the High Plateaux of Morocco and probably Algeria (this form was initially described from the High Plateaux of Algeria—see Loche (1858)—and some of our samples in Morocco are less than 50 km of the boundary). Sampling in some regions of West Africa is currently sparse (see Fig. 2), but we are confident that this lineage is effectively restricted to the
Maghreb. Beyond its mtDNA, *G. (c.) randonii* is also characterized by a specific morphology (very large size, proportionately long bill, see Guillaumet et al., 2005; Cramp, 1988), and we have shown that even the most similar populations from Eastern Maghreb (particularly, the subspecies *arenicola*) belong to a distinct lineage (see Fig. 7 and Section 4.2), while populations in the “islands” of the Sahara or in sub-Saharan Africa are all characterized by distinctively shorter size and bill (see e.g., Cramp, 1988; Keith et al., 1992; also Table 1).

The fact that the *G. (c.) cristata* specimens from Morocco shared haplotypes with the three European specimens (from Portugal, Southern France, and Yugoslavia) but had no haplotype from the Western sub-Saharan lineage (Fig. 4 and Table 2) demonstrates a Mediterranean origin of *G. (c.) cristata* in Morocco. Further, this shows that the Gibraltar Strait did not constitute a permanent barrier to dispersal in this species, in agreement with the results found in the Honeybee *Apis mellifera* (Franck et al., 2001), and in the Bearded Vulture *G. barbatus* (Godoy et al., 2004), but not in many other flying species of bats and birds (e.g., Castella et al., 2000; see Section 1).

Two lines of arguments suggest that Morocco was colonized from Europe, and not the reverse. First, the *cristata*
clade is distributed worldwide and reaches the edge of its distribution in Morocco (Fig. 4). Second, *G. (c.) cristata* in Morocco is essentially found in man-modified areas, such as low-intensity cereal crops that developed to the detriment of sclerophyllous forest (see Benabid, 2000). Forested areas are clearly avoided by *G. (c.) cristata*, while natural steppes, on the other hand, are inhabited by *G. (c.) randonii* (personal observations). Taken together, these elements suggest that the colonization of North Africa by *cristata* may have coincided with recent anthropogenic modifications of the landscape. However, additional data using e.g., hypervariable polymorphic markers will be required to confirm this inference (see e.g., Caujape-Castells and Jansen, 2003).

4.2. Medium-sized populations in Tunisia belong to a distinct mitochondrial lineage

Analysis of variation in bill size in the Crested Lark (as traditionally defined) showed that, in Morocco, short-billed and long-billed populations corresponded to two semi-species (Guillaumet et al., 2005). Populations that inhabit Tunisia were hard to assign to either of these taxa since they have an intermediate bill length (on average), while extreme specimens are similar to either *G. (c.) cristata* or *G. (c.) randonii* (Fig. 6). We raised four hypotheses that might explain such an intermediate bill length (see Section 1).

Among the six specimens sequenced from Tunisia, three had a *G. (c.) cristata*-like bill length (BL = 13.0–14.4, compare with data in Table 1 of Guillaumet et al., 2005), and three had a *G. (c.) randonii*-like bill length (BL = 15.7–16.9). However, none had a *cristata* or *randonii* haplotype, which invalidates two of our scenarios: the clinal and the hybridization hypotheses.

Under the ring-species model, we expected that large-billed populations in Eastern and Southern Morocco were descended from medium-sized populations in Algeria and Tunisia (Fig. 1). This prediction is not met: *senegallensis* populations in the E Maghreb are genetically more closely related to *cristata* populations in NW Morocco than to neighboring *randonii* populations in E Morocco, and this receives a strong bootstrap support (80–94%, Figs. 4 and 5B). Another expectation of the ring-species model is that differentiation occurred as a result of isolation by distance (Irwin and Irwin, 2002). However, at least two of the three groups in the Crested Lark complex are monophyletic in their mtDNA (the *randonii* and the *cristata* clades, see Fig. 5; see also Guillaumet et al., 2005), which is generally achieved through periods of geographic isolation. Accordingly, the transition between *cristata* and *senegallensis* in North Africa is coincident with a gap in the species distribution in Lybia, suggesting that Lybian desert might have constituted (and still constitute) a long term barrier to gene flow (Fig. 4; but note low sample size in this region). We conclude that the Crested Lark complex is not a good candidate for true ring speciation. In the future, other predictions of the ring-species model could be tested to definitively eliminate this possibility. For instance, if speciation occurred in a ring, we would expect that no breaks in gene flow exists between *senegallensis* and *randonii* populations meeting in Algeria.

The discovery of a distinct mitochondrial lineage in Tunisia and in sub-Saharan Africa (Figs. 3 and 4) supports the hypothesis that the Tunisian populations with intermediate bill size constitute a separate evolutionary unit. However, the phylogenetic relationships of haplotypes that constitute this assemblage is unclear (Fig. 5). While distance method suggests they are monophyletic, maximum parsimony, and maximum likelihood methods instead suggest they are paraphyletic relative to *cristata*, even if bootstrap support for this is weak (<50). This conflict could be due to the recent divergence of the *cristata* and *senegallensis* groups, since net divergence between them is only about 0.5%. Whatever the actual relationships of the *cristata* and *senegallensis* haplotypes, they are clearly much more closely related than either is to *randonii*. Despite this pattern of historical relationships, *senegallensis* populations in Tunisia are morphologically more similar to *G. (c.) randonii* than to their relatives in sub-Saharan Africa (Fig. 7). The intermediate bill length of the populations in Tunisia (Fig. 1) is thus the result of a recent change in morphology, the possible reasons of which are discussed below.

4.3. Morphological variation: combined effects of history and selection

Results presented in this paper reveal that part of morphological variation can be explained by the history of the populations. First, the region with the greater amount of morphological variation (the Maghreb, Fig. 1) is also the single region where the three mtDNA lineages are meeting (Fig. 4). In addition, we find a correlation between genetic and morphological divergences, as evidenced by the three populations of the *cristata* clade that cluster in a single morphological group (Fig. 7).

However, history alone seems not sufficient to account for all the variation. In particular, the populations in the Maghreb are consistently larger, and particularly, longer-billed than their closest relatives (*randonii* clade—which is endemic of the Maghreb—vs all other groups and populations; Europe vs Morocco in the *cristata* clade; Maghreb vs Sahel in *senegallensis*; see Table 1). Natural selection could have favored larger size and longer bill in the Maghreb. For instance, competition with the smaller but otherwise similar Thelka Lark could result in character displacement (*G. theklae* inhabits the Maghreb, Iberian Peninsula, and southernmost France, but is lacking elsewhere throughout the range of the Crested Lark included in the present study, see Cramp, 1988). Alternatively, longer bill and more generally, larger body size could result from ecoclimatic reasons (e.g., Bergmann’s rule, Ashton, 2002 or James’ modified version of this rule, see James, 1970). Finally, phenotypic plasticity could also produce larger body size and longer
bills in some environments, as demonstrated in birds by James (1983). Clearly, more work will be required to determine the role of selection and plasticity in the observed pattern of morphological variation.

4.4. Phylogeographical implications

Paleontological records as well as climatic reconstruction show that many major ecological changes have occurred during the Pleistocene in the Maghreb, including variations in the relative extent of woodlands and steppe/desert associations (Jolly et al., 1998a). These fluctuations are suspected to have triggered allopatric differentiation in separate Maghreb refuges, resulting in strong phylogeographical structure in some animal taxa. For instance, East-West vicariance in the Maghreb is suggested by recent studies in the Land Snail Helix aspersa (divergence 3.3–1.2 My, Guillier et al., 2001), and in the Greater Whitetoothed Shrew Crocidura russula (divergence ~2.2 My, Cosson et al., 2005). Our analysis of phylogeography in the Crested Lark complex revealed that three distinct mitochondrial lineages actually share the Maghreb: cristata in the West, senegallensis in the East, and randomii in between them. They differ by a maximum of 2.5% sequence divergence in the cyt b, which would translate roughly into 1.3 My divergence according to the crude approximation of 2% per My in birds (but see Lovette, 2004). Hence, our data are compatible with a scenario of in situ differentiation during the Pleistocene. Under this first hypothesis, cristata and senegallensis would have subsequently expanded their range beyond the Maghreb, respectively, into Eurasia and sub-Saharan Africa. Alternatively, one or both of them might have actually differentiated ex situ. In this case, colonization of the Maghreb would be instead secondary.

Evidence for a recent range expansion is given by the star-like phylogeny of haplotypes (Fig. 3), resulting in negative Fu and Li’s D* and F* (significant for cristata, marginally so for senegallensis). At least in Europe, such an expansion has been documented over recent centuries, as the cristata clade expanded its range into various man-modified areas that simulate semi-desert habitats (Tucker and Heath, 1994). However, we have currently little data to determine whether cristata and senegallensis initially spread from the Maghreb, or whether they are recent invaders here. For cristata, we previously argued that an European origin was more likely (see Section 4.1). For senegallensis, we may note that Sahelian populations appear slightly more polymorphic and might harbor more ancient haplotypes than their relatives in N Africa (see Section 3 and Fig. 5), suggesting a Sahelian origin for this group.

One interesting finding of our study is the presence of senegallensis on both sides of the Sahara (Fig. 4). This indicates that the Sahara did not function as a permanent barrier to dispersal in this species. In the East of the continent, the cristata group has also been found on both sides of the Sahara. Its dispersal might be facilitated by the Nile corridor, where the species is continuously distributed (Fig. 4; see also Juste et al., 2004). However, unlike senegallensis, the presence of the haplogroup cristata in sub-Saharan Africa is currently evidenced by a single specimen (from E Chad). An error in the labeling of this Museum specimen being always possible, we consider this conclusion only tentative.

In the future, more intensive sampling in some regions could be performed to refine the history of the cristata and senegallensis lineages. First, it is currently unclear whether these two lineages actually differentiated in Africa, with secondary expansion of the cristata haplogroup into Eurasia, or whether cristata instead differentiated outside from Africa (e.g., in the Middle East) and only secondarily invaded Africa where it formed a suture zone with senegallensis in the Chad area (Fig. 4). Second, isolated populations in desert areas, such as that present around Illizi in central Sahara (Fig. 4), could be relics of a repartition that covered the whole Sahara in a recent past. In the mid-Holocene in particular (i.e., around 6000 BP), biome reconstruction from pollen and plant macrofossil data suggests that the Sahara desert was extremely reduced and replaced by the steppe at low-elevation sites and woodland scrub at higher elevations (Jolly et al., 1998b). Such a continuity of grasslands habitats from the West African savannah zone to the Mediterranean Sea may have greatly facilitated colonization of the Maghreb from the south, or vice-versa. This hypothesis is supported by one recent study in the Honeybee A. mellifera (Franck et al., 2001), which found very weak mtDNA differentiation between subspecies from northern and southern sides of the Sahara, and possibly also by a study on the Bearded Vulture G. barbatus (but sample size in North Africa is limited, Godoy et al., 2004). Furthermore, colonization of the Maghreb by the African fauna is attested by many records of species that got extinct before classical times, such as Burchell’s Zebra Equus burchelli (Dobson and Wright, 2000), or in more recent times, such as the Hartebeest Alcelaphus buselaphus (Aulagnier and Thévenot, 1986), but also by relic populations of Sahelian reptiles, such as the Puff Adder Bitis arietans (Carranza et al., 2004).

Acknowledgments

We thank M. Kaboli, R. Yosef, and S. Drovetski for providing us with some tissue samples, E. Desmarais, A. Tillier, C. Bonillo, and J. Lambourdière who kindly provided technical assistance with the laboratory work, and one anonymous referee for helpful suggestions on a previous version of this paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2006.01.002.
References


Carranza, S., Arnold, E.N., Wade, E., Fahd, S., 2004. Phylogeography of


Caujape-Castells, J., Jansen, R.K., 2003. The in


Grisswold, C.K., Baker, A.J., 2002. Time to the most recent common ances-


Jolly, D., Harrison, S.P., Damanti, B., Bonnefille, R., 1998a. Simulated climate and biomes of Africa during the late quaternary: compari-


A. Guillaumet et al. / Molecular Phylogenetics and Evolution 39 (2006) 645–656

656