Refugia, colonization and diversification of an arid-adapted bird: coincident patterns between genetic data and ecological niche modelling

RAFAEL BARRIENTOS,*¶ 1 LAURA KVIST,†¶ ANDRÉS BARBOSA,* FRANCISCO VALERA,* FARES KHOURY,‡ SARA VARELA§ and EULALIA MORENO*

*Estación Experimental de Zonas Áridas (EEZA-CSIC), Ctra. de Sacramento s/n, La Cañada de San Urbano, E-04120 Almería, Spain, †Department of Biology, University of Oulu, POB 3000, FIN-90014 Oulu, Finland, ‡Department of Biology and Biotechnology, American University of Madaba, Madaba, Jordan, §Department of Ecology, Faculty of Science, Charles University, Vinicated, 7, 128 44 Praha 2, Praha, Czech Republic

Abstract

Phylogeographical studies are common in boreal and temperate species from the Palaeartic, but scarce in arid-adapted species. We used nuclear and mitochondrial markers to investigate phylogeography and to estimate chronology of colonization events of the trumpeter finch Bucanetes githagineus, an arid-adapted bird. We used 271 samples from 16 populations, most of which were fresh samples but including some museum specimens. Microsatellite data showed no clear grouping according to the sampling locations. Microsatellite and mitochondrial data showed the clearest differentiation between Maghreb and Canary Islands and between Maghreb and Western Sahara. Mitochondrial data suggest differentiation between different Maghreb populations and among Maghreb and Near East populations, between Iberian Peninsula and Canary Islands, as well as between Western Sahara and Maghreb. Our coalescence analyses indicate that the trumpeter finch colonized North Africa during the humid Marine Isotope Stage 5 (MIS5) period of the Sahara region 125 000 years ago. We constructed an ecological niche model (ENM) to estimate the geographical distribution of climatically suitable habitats for the trumpeter finch. We tested whether changes in the species range in relation to glacial–interglacial cycles could be responsible for observed patterns of genetic diversity and structure. Modelling results matched with those from genetic data as the species’ potential range increases in interglacial scenarios (in the present climatic scenario and during MIS5) and decreases in glacial climates (during the last glacial maximum, LGM, 21 000 years ago). Our results suggest that the trumpeter finch responded to Pleistocene climatic changes by expanding and contracting its range.

Keywords: Bucanetes githagineus, climate change, Mediterranean basin, microsatellites, mitochondrial DNA, range expansion

Received 7 October 2012; revision received 5 November 2013; accepted 6 November 2013

Introduction

The impact of Quaternary climatic oscillations on population processes of a plethora of boreal and temperate taxa has been studied in recent decades (reviewed in Taberlet et al. 1998; Hewitt 1999, 2004; Bennet & Provan 2008). In contrast, phylogeographical studies focused on arid-adapted animal species from southern regions (e.g. North Africa or Middle East) have remained scarce, with only a handful of published examples involving snails (Guiller
et al. 2001), mammals (e.g. Cosson et al. 2005; Nicolas et al. 2009; Ben Faleh et al. 2011), birds (e.g. Idaghdour et al. 2004; Guillamet et al. 2006, 2008; García et al. 2008; Förstchler et al. 2010) and, mainly, reptiles (e.g. Brown et al. 2002; Carranza et al. 2004, 2006; Rato et al. 2007, 2010), with none of them including niche modelling (see below). In this region, the role of successive expansions and contractions of the Sahara Desert in structuring populations (reviewed in Drake et al. 2011; see also Besnard et al. 2007; Guillamet et al. 2006, 2008; Illera et al. 2008) is still debated and poorly understood. Even though the Sahara Desert may form an effective dispersal barrier at present, this has not always been the case. Biogeographical and paleohydrological studies show that the Sahara has been covered by dense river networks, for example, during Marine Isotope Stage 5 (hereafter, MIS5; 130 000–75 000 years ago) and at the beginning of the Holocene (11 000–8000 years ago), likely enabling several water-dependent species to cross the Sahara (Drake et al. 2011). Nevertheless, after the late Pleistocene, the Sahara became increasingly dry (Kröpelin et al. 2008). Moreover, the ongoing climate warming, combined with an increase in human pressure via, for example, agriculture and animal husbandry, can increase desertification (Le Houérou 2000). This may cause shifts in range not only in temperate species (Parmesan & Yohe 2003; Root et al. 2003), but also in species from semi-arid environments, which could be particularly at risk (Brown et al. 1997; Sanz 2002). Although these latter organisms are adapted to warm conditions, slight increases in aridity can cause their habitats to become unsuitable, for instance, due to the loss of the scarce water sources as a consequence of prolonged droughts, especially if this occurs during the breeding season (Allen & Saunders 2002).

The circum-Mediterranean region north of the Sahara is an acknowledged biodiversity hotspot (Myers et al. 2000), suggesting a complex evolutionary history for the region. Indeed, different colonization or radiation patterns have been described in North Africa: (i) westward expansion from Asia or the Middle East (Franck et al. 2001; Arnaiz-Villena et al. 2008; Illera et al. 2008), (ii) eastward from western or southwestern Maghreb (Carranza et al. 2004, 2006; Guillamet et al. 2008) or even from the Canary Islands (Illera et al. 2011) and (iii) colonization from Europe (Guillamet et al. 2006). These patterns suggest multiple refugia or origins, colonization routes and radiations for the species from the region.

The trumpeter finch (Bucanetes githagineus) is a small passerine (ca. 21 g) distributed throughout the semi-arid regions of the Western Palaearctic, from northwestern India throughout the Middle East to Western Sahara and the Canary Islands (Cramp & Perrins 1994), where it occupies semi-arid landscapes (Barrientos et al. 2009a). Based on plumage coloration and morphology, four subspecies have been recognized (Cramp & Perrins 1994; see below). The species is expanding at the northwestern part of its range and has recently colonized Europe (Carrillo et al. 2007a; Barrientos et al. 2009b). Contrary to other species, whose expansions are assumed to have been human-mediated (e.g. Carranza et al. 2004; Cosson et al. 2005), the current range of the trumpeter finch seems to be a consequence of recent, natural expansion (Serra et al. 2006; Carrillo et al. 2007a). However, in some areas of the Near East, it is possible that the species is expanding in the most arid regions tracking human settlements that provide birds with water ad libitum (Cramp & Perrins 1994; Khoury & Al-Shamlih 2006).

The information available to explore past climatic conditions is continuously increasing (e.g. Waltari et al. 2007), and there is a recent tendency to integrate molecular ecology and modelling of past ranges to study the evolutionary histories of species (reviewed in Waltari et al. 2007; see also Hope et al. 2011; Rebelo et al. 2012). As no wide-scale land cover data with good resolution are yet available to reconstruct habitat traits in the past millennia, the best approach to date is to generate envelopes of climatic suitability, which assume that there is niche conservatism over time (Peterson et al. 1999).

In the present work, we used a phylogeographical approach complemented with ecological niche modelling (hereafter ENM) to trace the population history and colonization routes of the trumpeter finch and to match chronology of phylogeographical events within the species to global climatic changes (e.g. Hope et al. 2011; Rebelo et al. 2012). Specifically, we aimed to (i) determine the climatically suitable areas for the species during both glacial and interglacial periods, identifying where populations might have persisted in spite of Pleistocene climatic changes, and to track the colonization routes used by the species from ancient refuges; (ii) estimate dates of colonization and splitting events; and (iii) assess the population genetic structure of the current populations. For this purpose, we included 16 populations, totalling 271 birds (both fresh and museum specimens) across the species range. We used microsatellites and mitochondrial data to track recent range shifts and historical colonizations and diversification (Avise 2000). The ENM also allows us to identify which climatic variables were important in delimiting the range of our study species.

Material and methods

Study species

The trumpeter finch range is divided into three main regions with several smaller, patchy sites (Fig. 1). As most of their range has been poorly studied (Cramp & Perrins 1994; Del Hoyo et al. 2010), the knowledge on
connectivity among populations is still scarce (but see Barrientos et al. 2009b). Four subspecies are tentatively recognized based on plumage coloration and morphology (Cramp & Perrins 1994): B. g. zedlitzi in Mauritania, Maghreb, Libya and the Iberian Peninsula, perhaps also in northern Niger and northern Chad; B. g. githagineus in central and southern Nile Valley (Egypt), southeastern Egypt and northcentral and northeastern Sudan; B. g. crassirostris in Arabia, Sinai (Egypt), southcentral Turkey, Syria, Israel, Jordan, Iraq, Kuwait, Iran, Afghanistan, Uzbekistan, western Pakistan and northwestern India; and, B. g. amantun in the Canary Islands. The northern birds belonging to B. g. zedlitzi and B. g. crassirostris subspecies are larger than southern B. g. githagineus and B. g. amantun subspecies (Cramp & Perrins 1994). Colour variation in the northern subspecies is clinal, whereas it is more heterogeneous in the southern subspecies. The northern and southern subspecies are suggested to be mixed across their ranges in central Egypt (Cramp & Perrins 1994).

The species was first documented to breed in Europe in 1971 (García 1972). Its breeding range is expanding in southeastern Iberia (Carrillo et al. 2007a; Barrientos et al. 2009b), with several records outside of the breeding season in the western Mediterranean region (Spain, Llabrés 2010 or Italy, Gildi 1997), and even in North European countries such as Finland (Birdlife Finland 2008). This presence of individuals outside the current species range during the nonbreeding season precedes the establishment of new breeding sites in the periphery of its range (Barrientos et al. 2009b). Although the information is less detailed in Near East populations, it seems that the species began an expansion process there around 1950s (Cramp & Perrins 1994; Serra et al. 2006).

Sampling and laboratory protocols
We sampled 271 birds from 16 populations (Tunisia and Algeria were finally combined, see below; Fig. 1): four populations from the Iberian Peninsula (Gorafe, Cabo de Gata, Monnegre and Tabernas), four from Canary Islands (Lanzarote, Fuerteventura, Tenerife and La Gomera), three from Maghreb (Morocco, Algeria and Tunisia), one from south Sahara (Western Sahara), one from Egypt, two from the Near East (Jordan and Israel) and one from the Middle East (Iran). Samples include both fresh and museum samples. Fresh samples were obtained mainly by mist-netting, but also occasionally by sampling a single chick from every nest (see Barrientos et al. 2009b). Detailed information of sampling is given in Table S1 in Supporting Information and in Fig. 1.

DNA was extracted with the standard phenol–chloroform method (Sambrook & Russell 2001). Museum samples were extracted in a separate room using a fume hood with UV-light exposure prior to the extraction. Seven microsatellite loci – Lox1, Lox2 and Lox8 (Piertney et al. 1998), Ppi2 (Martínez et al. 1999), Pcc6 (Bensch et al. 1997), Pdo5 (Griffith et al. 1999) and Pk12 (GenBank Accession no. AF041466) – were amplified using the procedure detailed by Barrientos et al. (2009b). Microsatellites were run on ABI3730 and scored with GENEMAPPER v. 3.7 (Applied Biosystems, Foster City, CA, USA). Part of the mitochondrial control region was amplified using primers TrfinchL20 and passeriformesH830 (Kvist et al. 2011; Table S2, Supporting Information), which amplify most of domains I and II of the control region. The museum samples were amplified in two or three fragments using primers TrfinchL20 and TrfinchH465.
Microsatellite analyses

Locus Pk12 did not amplify in the Western Sahara samples, and therefore, this locus was removed from analyses where the Western Sahara was involved. We did not get good PCR products from some populations with museum samples, and therefore, these were excluded (see populations included in Table 1). We first used Microchek v. 2.2.3 (van Oosterhout et al. 2004) to check for null-alleles, large allele dropouts and scoring errors. The loci that showed evidence of null-alleles were removed from estimation of inbreeding coefficients, \( F_{IS} \), calculated with Genepop (Raymond & Rousset 1995; Rousset 2008). The same program was used to calculate linkage disequilibrium for each pair of loci in each population. All loci were used to estimate basic diversity values, observed and expected heterozygosities (with ARLEQUIN v. 3.1; Excoffier et al. 2005) and allelic richness (with FSTAT v. 2.9.3; Goudet 2001) from each population sampled with at least 10 individuals.

In addition, we calculated the \( M \)-ratio (Garza & Williamson 2001) with Arlequin to detect possible bottlenecks or founder events. These demographic events can produce gaps in the size distribution of microsatellite alleles, which can be quantified as the \( M \)-ratio (or Garza–Williamson index), the mean ratio of the number of observed alleles to all the potential repeats within the allele size range across all loci. When seven or more loci are analysed, \( M \)-ratios below 0.68 suggest a reduction in population size (Garza & Williamson 2001; Excoffier et al. 2005). Population bottlenecks can cause genetic signatures in the distributions of allele sizes and generate an excess in expected heterozygosities as rare alleles that are lost from a population contribute little to

### Table 1 Sample sizes for microsatellite data, observed (Ho) and expected (He) heterozygosities, allelic richness (\( A \)), \( M \)-ratio (or Garza–Williamson index), results from testing for bottlenecks with the program Bottleneck (Wilcoxon test for heterozygote excess and mode shift test) and \( F_{IS} \) (over four loci) estimated from the microsatellite data. Values in regular text are estimated with six loci, and in italics with seven loci

<table>
<thead>
<tr>
<th>Population (Region)</th>
<th>( N )</th>
<th>Ho (SD)</th>
<th>He (SD)</th>
<th>( A )</th>
<th>( M )-ratio (SD)</th>
<th>Wilcoxon test, Mode shift</th>
<th>( F_{IS} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gorafe (Iberian Peninsula)</td>
<td>26</td>
<td>0.5400 (0.2631)</td>
<td>0.7242 (0.3035)</td>
<td>2.88</td>
<td>0.708 (0.245)</td>
<td>NS, normal distribution</td>
<td>0.0541</td>
</tr>
<tr>
<td>Tabernas (Iberian Peninsula)</td>
<td>51</td>
<td>0.5853 (0.2675)</td>
<td>0.7608 (0.2614)</td>
<td>3.01</td>
<td>0.804 (0.169)</td>
<td>NS, normal distribution</td>
<td>0.0627</td>
</tr>
<tr>
<td>Cabo de Gata (Iberian Peninsula)</td>
<td>22</td>
<td>0.6057 (0.2839)</td>
<td>0.7534 (0.3027)</td>
<td>3.03</td>
<td>0.722 (0.238)</td>
<td>NS, normal distribution</td>
<td>–0.0139</td>
</tr>
<tr>
<td>Monnegre (Iberian Peninsula)</td>
<td>16</td>
<td>0.5433 (0.3362)</td>
<td>0.7357 (0.3014)</td>
<td>2.96</td>
<td>0.616 (0.291)</td>
<td>NS, normal distribution</td>
<td>0.1097</td>
</tr>
<tr>
<td>Lanzarote (Canary Islands)</td>
<td>25</td>
<td>0.6754 (0.2680)</td>
<td>0.7842 (0.2333)</td>
<td>3.07</td>
<td>0.654 (0.259)</td>
<td>( P &lt; 0.05 ), normal distribution</td>
<td>–0.0490</td>
</tr>
<tr>
<td>Fuerteventura (Canary Islands)</td>
<td>26</td>
<td>0.6951 (0.2540)</td>
<td>0.7803 (0.2325)</td>
<td>3.05</td>
<td>0.680 (0.251)</td>
<td>( P &lt; 0.05 ), normal distribution</td>
<td>–0.0201</td>
</tr>
<tr>
<td>Jordan (Near East)</td>
<td>28</td>
<td>0.5747 (0.2501)</td>
<td>0.7599 (0.2519)</td>
<td>3.12</td>
<td>0.756 (0.246)</td>
<td>( P &lt; 0.01 ), normal distribution</td>
<td>0.1098</td>
</tr>
<tr>
<td>Morocco &amp; Tunisia (Maghreb)</td>
<td>36</td>
<td>0.6731 (0.2596)</td>
<td>0.7749 (0.3045)</td>
<td>3.11</td>
<td>0.776 (0.167)</td>
<td>( P = 0.07 ), normal distribution</td>
<td>–0.0029</td>
</tr>
<tr>
<td>Western Sahara (Western Sahara)</td>
<td>10</td>
<td>0.7913 (0.2509)</td>
<td>0.7912 (0.2148)</td>
<td>3.02</td>
<td>0.573 (0.344)</td>
<td>( P &lt; 0.01 ), shifted distribution</td>
<td>–0.2159</td>
</tr>
</tbody>
</table>

© 2013 John Wiley & Sons Ltd
overall heterozygosity (Luikart & Cornuet 1998; Garza & Williamson 2001). We used BOTTLENECK v. 1.2.01 (Piry et al. 1999) to look for these signatures using the two-phase mutation model (with 70% of stepwise mutations) for microsatellites and examining the overall distribution of allele frequency classes (‘mode shift’ test) and using a Wilcoxon test to detect signs of excess of heterozygotes.

To measure genetic differentiation between populations, we calculated Jost’s $D_{st}$ (Jost 2008) with GENALEX 6.5 (Peakall & Smouse 2012). $D_{st}$ is a relative measure of genetic differentiation, which has been shown to be less biased than $F_{st}$ for measuring differentiation, especially when the number of sampled individuals varies between populations (Gerlach et al. 2010). Significance for the $D_{st}$-values was obtained by 999 permutations. In addition, Arlequin (Excoffier et al. 2005) was used to estimate pairwise $F_{ST}$-values between the sampling sites for those sites from which at least seven individuals were sampled. The statistical significance was obtained by permuting the samples by 1000 permutations, and $P$-values were corrected with the Bonferroni step-down correction (Holm 1979). We chose only $F$-statistics and not $R$-statistics due to the relatively low sample size from some of the populations (Slatkin 1995; Gaggiotti et al. 1999). To test for isolation by distance (coordinates are shown in Table S3, Supporting Information), the pairwise distances between sampling sites, using Slatkin’s linear $F_{ST}$, were correlated with logarithms of geographical distances using the matrix correlation test, the Mantel test, in Arlequin. Significance of correlation was tested with 1000 permutations. In addition, we used the molecular variance analysis (AMOVA) implemented in Arlequin to test for the best grouping of the populations. For this, we combined the samples from Iberia (Gorafe, Tabernas, Cabo de Gata and Monnegre) into one, as well as samples from Canary Islands (Fuerteventura and Lanzarote). We constructed nine different groups using three hierarchical levels, within populations, between populations and among groups. The nine groups were formed with the intention of testing all the possible connections between the populations taking into account the geographical locations of the populations. Significance was tested with 1000 permutations.

Population structure was further studied using STRUCTURE v. 2.2 (Pritchard et al. 2000; see also Falush et al. 2003). This program infers the number of populations ($K$) using a Markov chain Monte Carlo approach. We used no prior information of the sampling locations and assumed a model with population admixture and correlated allele frequencies within populations (Falush et al. 2003). We used 10 000 iterations as a burn-in and collected data for 100 000 iterations for a series of ten independent runs for each value of $K$ between 1 and 10. The likelihood of the data and subsequent log probabilities for the different numbers of subpopulations were calculated for each $K$. We also computed the posterior probabilities according to Bayes’ Rule for each $K$ (the Bayes factor) following the guidelines in the Structure manual. In addition, we applied the ad hoc method by Evanno et al. (2005), which is based on estimating $\Delta K$ between consecutive numbers of populations. The highest $\Delta K$ can be inferred as the best estimator of the actual $K$, and this method should detect the highest level of population structure even if several hierarchical levels exist (Evanno et al. 2005).

**Mitochondrial analyses**

The mitochondrial sequences were aligned by eye with Bioedit (Hall 1999). We constructed two data sets: one containing 576 bp long sequences from 175 individuals and one containing 194 bp from 187 individuals (due to the failure to sequence the whole fragment from twelve individuals, eight of them from Egypt). Diversity estimates including nucleotide diversity ($\pi$), theta (0), number of haplotypes and haplotype diversity ($h$) were calculated with DNASP v. 5.0 for both data sets (Librado & Rozas 2009). The following applies to the 576 bp data if not stated otherwise. We estimated the ‘expansion parameters’, Tajima’s $D$ and Fu’s $F$ and calculated raggedness index ($r$) and $\tau = 2\mu t$ (where $\mu$ is the mutation rate per generation and $t$ time in generations) from mismatch distributions to detect possible changes in population size with DnaSP. The best substitution model for the data was estimated with Modelgenerator (Keane et al. 2006) according to Bayesian information criterion. This substitution model was used in analyses with Arlequin.

Pairwise $\theta_{ST}$-values between the study populations (sampled for more than seven individuals) were estimated with Arlequin using Tamura–Nei distance, and isolation by distance tested as described for microsatellites. AMOVA was performed as above, but using 16 different combinations instead of nine because the number of populations was larger in this data set. For this, we combined the samples from Iberia into one (Gorafe, Tabernas, Cabo de Gata and Monnegre), as well as samples from Near East (Jordan and Israel), Maghreb (Morocco, Tunisia and Algeria) and Canary Islands (Fuerteventura, Lanzarote, Tenerife and La Gomera). Egypt is represented by only two individuals.

To explore the colonization history in more detail, we used DIYABC v. 1.0.4.45 (Cornuet et al. 2008) to perform coalescence analyses. This program utilizes a coalescent-based approximate Bayesian computation algorithm to infer the population history by looking backwards in time to examine genealogy of alleles until the time at which the most recent common ancestor (hereafter
TMRCA) is reached. We carried out preliminary tests by building numerous different combinations of splitting and admixture along the scenarios starting from the six sampling regions (Fig. 1), that is, the Canary Islands, the Western Sahara, the Iberian Peninsula, Maghreb, the Near East (including Egypt, which did not deviate from the Near East population based on pairwise ΦST-values) and Iran at time 0 (present). We used prior historical information and geographical proximities of populations for building evolutionarily realistic scenarios. In addition, we performed a simplified run by combining the populations into three based on pairwise ΦST-values and geography: (i) Near East (including Egypt) and Middle East combined, (ii) Maghreb and Iberia combined and (iii) Western Sahara and Canary Islands combined. All three possible cladograms of a nonrooted tree were included in the scenarios. After the preliminary tests, we used altogether seven different historical scenarios for the final analysis (Fig. 2 shows the three best supported scenarios). We used a minimum substitution rate of $6.8 \times 10^{-4}$ and maximum of $1.72 \times 10^{-7}$ per generation based on conservative substitution rate estimates of 2–5% per Myr for the control region (Sato et al. 1999; Päckert et al. 2007) and a generation length ($G$) of 2.9 years for the trumpeter finch. Generation length was obtained from capture-recapture data in the Tabernas population (E. Moreno, unpublished data) and calculated using the equation $G = A + s/1-s$, where $A$ is the age at first reproduction and $s$ is the survival rate (Lande et al. 2003). Modelgenerator suggested that the Tamura–Nei substitution model was the best. However, the program crashed while trying to change the substitution model setting. Thus, we used Kimura two-parameter distance as an alternative. Priors used for effective population sizes were 100–1 000 000 and for splitting times from 1 to 5000 up to 5000 to 100 000, depending on the population. We simulated 7 000 000 data sets, which were compared to the observed data to choose the scenario that best explained the data by estimating posterior probabilities for each scenario. We checked for the fit of observed and simulated summary statistics under the seven scenarios as recommended in the DIYABC manual for the following statistics: (i) for one sample statistics, we used the number of haplotypes and segregating sites, mean number of pairwise differences and its variance, Tajima’s $D$, private segregating sites, mean numbers of the rarest nucleotide at segregating sites and its variance; and (ii) for two samples, we used the numbers of haplotypes and segregating sites, mean of pairwise differences and ΦST. In addition, we evaluated the confidence in scenario choice by estimating type I and II errors as suggested in the DIYABC manual. The scenario that best explained the data was used to estimate divergence times and effective population sizes. We used TCS v. 1.21 (Clement et al. 2000) for constructing the haplotype network.

**Mitochondrial and microsatellite combined analyses**

We also combined the microsatellite and mitochondrial data (576 bp fragment) for coalescence analysis with DIYABC (Cornuet et al. 2008). We used the same large-scale sample populations as above, except that we did not have microsatellite data from Iran and Egypt, so these sites were excluded (i.e. there were five populations: Iberia, Maghreb, Near East, Western Sahara and Canary Islands). We assumed a mutation rate of $1 \times 10^{-5}$ for all the microsatellite loci (Kruglyak et al. 1998), otherwise the run conditions were as above. Prior distributions for effective population sizes and splitting times were as above, with slight modifications between the two runs performed. The simulated data were again compared to the observed data to choose the scenario that best explains the data. Posterior probabilities were estimated for each scenario, and the three best were chosen for a final run with 3 000 000 simulations (Fig. S1, Supporting Information). The scenario that best explained the data was used to estimate divergence times and effective population sizes.

![Fig. 2](image-url)  
Fig. 2 Historical scenarios constructed for the coalescence analyses with mitochondrial (576 bp fragment) data.
Ecological niche modelling

We selected 197 species occurrences from the georeferenced range map of the trumpeter finch from Fig. 1 (Table S4, Supporting Information; see also Fig. 5d). We used the 19 bioclimatic variables from WorldClim (version 1.4, see http://www.worldclim.org/ for variable descriptions). The WorldClim variables for the present interglacial scenario are the maps obtained by statistical interpolation of weather stations, with 30-arcsec (ca. 1-km) resolution (see Hijmans et al. 2005). For the Last Glacial Maximum (hereafter, LGM) cold scenario (21 000 years ago), we used the predictions of two general circulation models: the ‘Community Climate System Model’ (hereafter CCSM) and the ‘Model for Interdisciplinary Research on Climate’ (hereafter MIROC; version 3.2; downloaded from worldclim.org).

We modelled the current range of the trumpeter finch using the maximum entropy algorithm, MAXENT 3.3.3 (Phillips et al. 2006; Phillips & Dudík 2008), with the R library ‘dismo’ and the default settings (Hijmans et al. 2011). MAXENT calculates the climatic suitability of the species based on the species occurrence records and a set of background points, calibrating a model to discriminate between the species niche and the background conditions (used vs. available; Elith et al. 2006; Phillips et al. 2006). Model parameters were calibrated from 500 replicate model runs, where species records were randomly divided (80–20%) into 157 points for training the model and 40 points for testing the model predictions, plus 10 157 background points. Variable importance was determined from the percentage contribution of each variable to the model predictions. Overall model performance was evaluated using the area under the receiving operator characteristics curve (AUC), which ranges from 0.5 (randomness) to 1 (perfect prediction). The logistic output of MAXENT consists of a grid map with each cell having an index of suitability between 0 and 1. Low values indicate that conditions are climatically unsuitable for the species to occur, whereas high values indicate that conditions are suitable. We calibrated the model using the current range of the species, and we projected it into the two LGM cold climatic scenarios (21 000 years ago; that is, one of the most demanding scenarios for several species; see Introduction). We used a consensus approach to map the range of the species in the LGM. We projected the model in both general circulation models (CCSM and MIROC) and multiplied both maps. Thus, the LGM map shows the areas with high levels of climatic suitability for the species using both climatic models. We did not project the models to MIS5 (Last Interglacial Maximum, 126 000 years ago) because the currently available resolution of the climatic models for this period is still poor. Alternatively, we used the current climate conditions as a proxy of the MIS5 following Varela et al. (2009), because MIS5 matches the TMRCA (see below) and we wanted to explore climatic suitability for the trumpeter finch at that time.

Results

Microsatellite analyses

The highest expected heterozygosities were found in Jordan, Western Sahara, Fuerteventura and Lanzarote and the highest observed heterozygosities in Western Sahara, Lanzarote, Fuerteventura and Morocco. Allelic richness was the highest in Jordan and Morocco. All four populations from the Iberian Peninsula had diversity estimates somewhat lower than the other populations. Inbreeding coefficients estimated over four loci (Lox 2, Lox 8 and Pk12 were removed due to possible null-alleles) resulting in nonsignificant values for all populations (Table 1). No linkage disequilibrium was detected.

Bottleneck tests suggest a past bottleneck by revealing significant heterozygote excess in Lanzarote, Fuerteventura, Jordan and Western Sahara, which also showed a shifted allele size class distribution. M-ratio was low (below 0.68) in Monnegre, Lanzarote and Western Sahara (Table 1).

Pairwise $D_{olv}$-values (Table 2), as well as $F_{ST}$-values (not shown), within the Iberian populations and within the Canary Islands were mostly nonsignificant (but Górafe significantly differed from Tabernas) and close to zero. Significant differentiation based on $D_{olv}$-values was observed between comparisons of most of the Iberian populations and the Canary Island populations. The Moroccan population differed significantly from two or three of the Iberian populations (depending on the number of loci used), from both Canary Island populations and from Western Sahara. Interestingly, differentiation was low across vast geographical distances, between Western Sahara and Jordan and between Morocco and Jordan. Isolation by distance was not detected ($r = 0.034, P > 0.5$). The hierarchical variance analysis resulted in the highest among-group variation when the Canary Island populations were considered as one group and all the others were combined. This grouping resulted in the highest and significant $F_{CT}$-value (Table S5, Supporting Information).

Structure gave the highest log probability for three clusters with a Bayes factor of 1.00 when data from six loci were used, but the application of the Evanno’s method reduced the most likely number to two. When data from seven loci were used (excluding Western Sahara), the highest probability was obtained for six clusters (Bayes factor 1.00), whereas the Evanno’s method still resulted in $K = 2$. Results from six loci are
Table 2  Pairwise $D_{est}$ values from six or seven (in italics) microsatellite loci above and $\Phi_{ST}$-values from mitochondrial data (576 bp) below the diagonal for populations including at least 7 samples. Significant $D_{est}$ and $\Phi_{ST}$-values ($P < 0.05$) are shown in bold. Region abbreviations are as follows: Iber. Pen., Iberian Peninsula; Canar. Is., Canary Islands; West. Sah., Western Sahara.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gorafe</td>
<td>Iber. Pen.</td>
<td>0.0334</td>
<td>0.0049</td>
<td>0.0581</td>
<td>0.0958</td>
<td>0.1006</td>
<td>0.0929</td>
<td>0.0599</td>
<td>—</td>
<td>—</td>
<td>0.2300</td>
</tr>
<tr>
<td>Tabernas</td>
<td>Iber. Pen.</td>
<td>0.0051</td>
<td>0.0000</td>
<td>0.0293</td>
<td>0.0619</td>
<td>0.0628</td>
<td>0.0819</td>
<td>0.0757</td>
<td>—</td>
<td>—</td>
<td>0.1684</td>
</tr>
<tr>
<td>Cabo de Gata</td>
<td>Iber. Pen.</td>
<td>0.0031</td>
<td>0.0000</td>
<td>0.0238</td>
<td>0.0687</td>
<td>0.0717</td>
<td>0.0597</td>
<td>0.0359</td>
<td>—</td>
<td>—</td>
<td>0.1761</td>
</tr>
<tr>
<td>Monnegre</td>
<td>Iber. Pen.</td>
<td>0.0557</td>
<td>0.0000</td>
<td>0.0282</td>
<td>0.1215</td>
<td>0.1138</td>
<td>0.0740</td>
<td>0.0802</td>
<td>—</td>
<td>—</td>
<td>0.1597</td>
</tr>
<tr>
<td>Lanzarote</td>
<td>Canar. Is</td>
<td>0.1545</td>
<td>0.1059</td>
<td>0.0989</td>
<td>0.0140</td>
<td>0.0007</td>
<td>0.1021</td>
<td>0.0802</td>
<td>—</td>
<td>—</td>
<td>0.2362</td>
</tr>
<tr>
<td>Fuerteventura</td>
<td>Canar. Is</td>
<td>0.2735</td>
<td>0.2060</td>
<td>0.2246</td>
<td>0.0985</td>
<td>0.0138</td>
<td>0.0908</td>
<td>0.1100</td>
<td>—</td>
<td>—</td>
<td>0.1978</td>
</tr>
<tr>
<td>Jordan</td>
<td>Near East</td>
<td>0.3414</td>
<td>0.2704</td>
<td>0.2662</td>
<td>0.1442</td>
<td>0.0466</td>
<td>0.0653</td>
<td>0.0757</td>
<td>0.0154</td>
<td>0.0054</td>
<td>0.1602</td>
</tr>
<tr>
<td>Morocco</td>
<td>Maghreb</td>
<td>0.0160</td>
<td>0.0583</td>
<td>0.0594</td>
<td>0.1315</td>
<td>0.2351</td>
<td>0.3383</td>
<td>0.4030</td>
<td>—</td>
<td>—</td>
<td>0.0160</td>
</tr>
<tr>
<td>Tunisia &amp; Algeria</td>
<td>Maghreb</td>
<td>0.0170</td>
<td>0.0265</td>
<td>0.0044</td>
<td>0.0145</td>
<td>0.0037</td>
<td>0.1020</td>
<td>0.1724</td>
<td>0.0707</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Western Sahara</td>
<td>West. Sah.</td>
<td>0.3397</td>
<td>0.2498</td>
<td>0.2965</td>
<td>0.1239</td>
<td>0.0008</td>
<td>0.0218</td>
<td>0.3998</td>
<td>0.1866</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
shown in Fig. 3. When the results suggested three separated clusters, most of the Iberian individuals as well as those from the Canary Islands fell into their own groups, but the third group included birds from Maghreb, Western Sahara and Near East. The number of populations inferred using the Evanno’s approach for the six loci data set (two) showed no clear grouping according to the sampling locations. When the data set of seven loci was used, the samples from the Canary Islands grouped together, and others were mixed.

Mitochondrial analyses

We obtained 576 bp of the mitochondrial control region from 175 individuals. In addition, we sequenced 194 bp from twelve additional individuals (eight from Egypt, two from Tunisia, one from Algeria and one from Israel) from which we failed to sequence the whole 576-bp fragment. The sequences from the 175 individuals resulted in 76 different haplotypes with a mean nucleotide diversity of 0.00468. Haplotype network is shown in Fig. 4.
highest nucleotide diversity was found in Iran (0.01042, \(n = 3\)) and the Near East (0.00436, \(n = 31\); Jordan 0.00419, \(n = 27\) and Israel 0.00550, \(n = 4\)), and the lowest from Egypt (0.00174, \(n = 2\)) and the Western Sahara (0.00197, \(n = 10\); Table 3). The nucleotide diversities estimated from the 194-bp fragment from the populations with increased sample sizes were 0.00619 for Israel (\(n = 5\)) and 0.00309 for Egypt (\(n = 10\)); in Tunisia and Algeria, all haplotypes were identical (\(n = 10\)).

The parsimony network (Fig. 4) shows three main haplotypes found in most populations, but with different frequencies. These three haplotypes form centres of three connected starlike networks, indicating past population expansions. However, even though in many cases, \(\theta\)-values of populations were much larger than nucleotide diversities (resulting in negative Tajima’s \(D\)-values typical for expanded populations; Table 3), Tajima’s \(D\) were significant only for the Iberian Gorafe and the Moroccan population. Fu’s \(F\)-values, which are more sensitive to past expansions than Tajima’s \(D\), were significant for all Iberian and Maghreb populations, Lanzarote and Jordan. Smooth, unimodal mismatch distributions from Tabernas, Cabo de Gata, Monnegre, Jordan, Morocco and Western Sahara indicated past population expansions, although the raggedness indexes were not significant.

Modelgenerator suggested the Tamura–Nei substitution model with invariable sites as the best model, with a fraction of invariable sites of 0.79. This model applied to pairwise \(\Phi_{ST}\) calculations showed that Iberian subpopulations are not differentiated from each other, nor are the Canary Island and Near East subpopulations as all pairwise comparisons within these groups resulted in nonsignificant \(\Phi_{ST}\)-values between populations. In the Maghreb group, Morocco was significantly different from pooled Tunisia and Algeria. Significant \(\Phi_{ST}\)-values were also found between Iberian populations and most populations from Canary Islands, Jordan and Western Sahara. Two of the Iberian populations were significantly differentiated from Israel, Iran and Morocco. Notably, the Western Saharan population was significantly different from the Maghreb populations, but was not so from

Fig. 4 Haplotype network of the mitochondrial control region sequences based on 576 bp from 175 individuals. Each population is marked with a different colour and abbreviations as shown in the figure. When there were more than one haplotype in the same population, their numbers are shown. Branching point of an outgroup Carduelis hornemanni, GenBank Accession no EU400528, is marked with an arrow.

© 2013 John Wiley & Sons Ltd
the Canary Island populations, and Egypt was not differentiated from Near East populations (Table 2). Contrary to microsatellite data, mitochondrial data show a significant isolation-by-distance pattern (r = 0.239, P < 0.05).

The hierarchical molecular analyses with fifteen different groupings partitioned the largest amount of variance between groups when Iberia was combined with Maghreb, Canary Islands with Western Sahara, Near East with Egypt and Iran treated separately (21.3%). This grouping also resulted in the highest FCT-value and a low FSC-value (Table S6, Supporting Information).

**Coalescence analyses**

When only the 576 bp mitochondrial data were used in DIYABC analyses, the best of seven historical scenarios was scenario 1 in Fig. 2. This scenario included the first split between Iberia and Maghreb at t1, the second split between Western Sahara and Canary Islands at t2 and the third at t3 including a simultaneous split of four branches leading to Iran, to the Near East, to the Western Sahara-Canary Islands and to Maghreb-Iberia. Posterior probability using logistic regression resulted in 0.53 for this scenario, and the second best (scenario 2 in Fig. 2) had a posterior probability of 0.15. Comparison of summary statistics of the observed data with the summary statistics obtained according to each of the seven simulated scenarios did not lend clear support to any of the scenarios; most of the statistics were within the ranges of the simulated ones. Type I error (the true scenario did not have the highest probability) was 0.148, and the mean of type II errors (the false scenario is not rejected) was 0.070. The best scenario for the combined microsatellite and mitochondrial data was quite similar to that obtained only with mitochondrial data (scenario 1 in Fig. S1, Supporting Information; logistic regression 0.43; t3, i.e. TMRCA, was 125 000 years ago with combined data), but the estimated error rates remained high with 0.312 for type I and 0.295 for type II errors. Posterior estimates for splitting times and effective population sizes from mitochondrial analyses are shown in Table 4.

**Ecological niche modelling and climate change**

Our model showed a high power of discrimination between presences and background, with an AUC for the calibrating data set of 0.87 (i.e. 87% of the records were correctly predicted). Thus, our model was highly accurate and its prediction was significantly greater than the random one (AUC = 0.5; one-tailed Wilcoxon signed rank test; P = 0.002). Moreover, our model correctly predicted the validation data set, AUC = 0.80, which indicates that our model did not overfit.

Our analysis indicated that climatic suitability is directly proportional to the temperature in summer
(BIO9; importance of 28%), the precipitation seasonality (BIO15; 15%) and the winter mean temperature (BIO11; 13%), with the climatic suitability being highest at around 18°C in this last variable. Thus, our model showed that the species climatic requirements are related to warm and dry conditions.

The present potential range based on climatic variables (Fig. 5a) is basically consistent with the observed current range (Fig. 1). The map showed that the most suitable conditions are concentrated in the Middle East and the northwestern Africa. It is outstanding that the large gap present in the trumpeter finch range in inland Egypt and Libya (Fig. 1) was also detected with ENM (Fig. 5a), suggesting that there are no suitable climatic conditions for the species to survive in those areas. Climatic suitability is intermediate to high in some areas where the species has not been recorded like Turkey, coastal Egypt and Libya, East Iran or the central Iberian Peninsula.

The estimated climatic suitability for the trumpeter finch during the LGM underwent a general range contraction, with the main area of suitable habitat identified in Iran, Iraq, Syria, Jordan, Saudi Arabia, southeastern Morocco, Algeria and Niger (Fig. 5b). These refugia were likely isolated, with at least the Maghreb population isolated from the easternmost nuclei. Comparing the model predictions between glacial and interglacial scenarios, we observed that, globally, the climatic suitability for the species had increased (Fig. 5c). Both area expansion and an increase in local climatic suitability have occurred after the glacial maximum, suggesting that the species could have experienced an expansion associated with this climatic warming. Finally, we created a map to identify the areas that showed climatically suitable conditions for the species in both glacial and interglacial scenarios (Fig. 5d). We consider that those areas could have been the climatic refugia for the species during the

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne Iberia</td>
<td>$9.73 \times 10^5 (1.95 \times 10^5 - 9.87 \times 10^5)$</td>
</tr>
<tr>
<td>Ne Maghreb</td>
<td>$3.83 \times 10^5 (1.09 \times 10^5 - 9.79 \times 10^5)$</td>
</tr>
<tr>
<td>Ne Canary Islands</td>
<td>$3.44 \times 10^5 (1.28 \times 10^5 - 9.87 \times 10^5)$</td>
</tr>
<tr>
<td>Ne Western Sahara</td>
<td>$1.46 \times 10^5 (6.07 \times 10^5 - 9.56 \times 10^5)$</td>
</tr>
<tr>
<td>Ne Near East</td>
<td>$6.54 \times 10^5 (2.65 \times 10^5 - 9.71 \times 10^5)$</td>
</tr>
<tr>
<td>Ne Iran</td>
<td>$4.23 \times 10^5 (7.94 \times 10^5 - 9.76 \times 10^5)$</td>
</tr>
<tr>
<td>t1 (split Iberia/Maghreb)</td>
<td>$4.83 \times 10^3 (1.48 \times 10^3 - 9.98 \times 10^3)$</td>
</tr>
<tr>
<td>t2 (split C-Islands/W-Sahara)</td>
<td>$5.53 \times 10^3 (1.43 \times 10^3 - 1.83 \times 10^3)$</td>
</tr>
<tr>
<td>t3 (TMRCA of the study populations: split Iberia+Maghreb/Iran/N-East/C-Islands=W-Sahara)</td>
<td>$8.67 \times 10^3 (2.81 \times 10^3 - 1.4 \times 10^4)$</td>
</tr>
</tbody>
</table>

Fig. 5 Species climatic suitability in the present interglacial scenario (a), and during the LGM (b). Increment in climatic suitability from glacial to interglacial periods (c), obtained from subtracting habitat suitability in LGM (a) from the present scenario (b). Climatic suitability during glacial-interglacial cycles (d). Red areas indicate the locations that showed suitable conditions for the species during both glacial and interglacial scenarios. Black dots show the current species occurrences used in our analyses (see Table S4 in Supporting Information).
Pleistocene as climatic conditions were more stable there. These areas are located in Iran, Jordan, north Arabia and southwestern Maghreb (Morocco). The high number of localities where the species is currently present, but that are placed outside of these potential refugia (black dots from Fig. 5d and Table S4, Supporting Information), highlights that the species has expanded its geographical range since the LGM.

Discussion

Trumpeter finch diversification

With mitochondrial data, we detected signs of past population expansions in populations from the Iberian Peninsula, the Near East, Morocco, and combined Tunisia and Algeria. These signs of expansion are likely of an older origin than the recent expansions recorded in the Near East and Iberia (López-Seoane 1861; Vayreda 1883; García 1972; Khoury & Al-Shamlih 2006). Values of \( \tau \) from the mismatch distributions (Table 3) indicate that these expansions had already occurred 220,000–68,000 years ago (estimated using divergence rates of 2.5%/Myr and \( \tau \)-values from 1.97 to 2.55; from lower range in Gabo de Gata, southeastern Iberian Peninsula, to upper range in Jordan, respectively), likely reflecting an expansion in the ancestral population. This period overlaps with the humid period during MIS5 (130,000–75,000 years ago) in the Sahara region. The TMRCA of trumpeter finch populations was estimated at about 125,000 years ago with the combined mitochondrial and microsatellite data set. These estimates indicate that the common ancestor existed shortly before the humidification of the Sahara, which facilitated the expansion and spread of the species across North Africa (see Drake et al. 2011 for other examples from this humid phase). High diversity and large \( \tau \)-values in populations from the Near and the Middle East indicate that the colonization occurred from east to west, but topologies of the phylogenetic trees (best coalescent scenarios, Fig. 2) or the maximum parsimony network (Fig. 4) do not lend support to any direction. Westward colonization would agree with results from Arnaiz-Villena et al. (2008), who suggest that Asia is the most likely origin for the ‘Arid-Zone’ Carduelini finches, a group including the trumpeter finch. According to these authors, the split between the trumpeter finch and its closely related sister-species, the Mongolian finch (B. mongolica), occurred about 6.5 Mya.

The present North African populations are differentiated in mitochondrial markers (pairwise \( F_{ST} \) values 0.3998 between Western Sahara and Morocco and 0.1866 between Western Sahara and Tunisia and Algeria combined) and in microsatellite loci (pairwise \( D_{est} \) between Western Sahara and Morocco 0.1602, but pairwise \( F_{ST} \) value between Western Sahara and Morocco was nonsignificant). This differentiation could have resulted from at least two scenarios. First, the species could once have been distributed across a wider range in northwestern Africa, when the region was moister. For instance, Guillaumet et al. (2006, 2008) found that steep increases in aridity in Africa are coincident with north–south vicariance events mediated by the Sahara in the Galerida larks, a group of arid-adapted birds. Similarly, populations of smooth snakes Macroprotodon sp. in southern Morocco and coastal Western Sahara, where habitats are more suitable than further inland, have been suggested to be relics from more mesic periods (Carranza et al. 2004). Saharan cycles of desert expansion and retraction may have also provided the conditions for population bottlenecks leading to radiation in West African Taterillus, rodents living in subdesert habitats (Dobigny et al. 2005). When the Sahara expanded after aridity increased, populations of the trumpeter finch became isolated. However, there are still remnants of shared ancestral polymorphism in these populations. For instance, we found a statistically significant divergence based on mitochondrial analyses between the geographically nearby populations of Morocco and Tunisia and Algeria combined, even though they share common haplotypes, supporting a common history followed by segregation. It is worth mentioning that this pattern of genetic differentiation between Morocco and Tunisia based on mitochondrial data has also been found in Dupont’s lark Chersophilus duponti (García et al. 2008), a passerine living in subdesert steppes, as well as in several other arid-adapted species (reviewed in Barata et al. 2008), suggesting a common process at a regional scale. A second possibility is that the species colonized northwestern Africa using two routes: one north of the Sahara and another south of the Sahara, which can also explain the observed differentiation in North Africa. The two colonization routes might meet somewhere between our sampling sites from the Western Sahara and Morocco. Mitochondrial differentiation is stronger between Morocco and the Western Sahara than between Morocco and samples from Tunisia and Algeria, suggesting a closer connection between the northernmost African populations. In addition, populations as distant as the Near East and the Western Sahara are not differentiated from each other, which may result from the Near East having served as a source for colonization in the Western Sahara (or vice versa), supporting this second explanation of colonization around the Sahara. A secondary contact of the two colonization routes would likely also result in elevated diversity values in populations located in the zone, which we did not detect. This does not completely rule
out the possibility of existence of a contact zone, because it is possible that our sampling sites are not located directly in the zone, as there is actually a 1000-km gap between the sampling sites in the Western Sahara and Morocco. In any case, the genetic structure we found does not support the four trumpeter finch subspecies recognized to date based on plumage coloration and morphology (Cramp & Perrins 1994).

Pairwise $F_{ST}$ and $\Phi_{ST}$-values (negative values for both) and AMOVA analyses support close connectivity of the Western Sahara and Canary Islands and differentiation between Morocco and the Canary Islands. On the other hand, $D_{est}$-values were higher between the Western Sahara and the Canary Islands than between Morocco and the Canary Islands. These islands, a volcanic archipelago that has never been connected to the continent (Illera et al. 2012), were likely colonized from a source located in the present northwestern Africa (although it is possible that colonization has also occurred the other way round, from the Canary Islands to continental Africa; see Illera et al. 2011). Our results give more support to colonization from a region around the Western Sahara than from Morocco ($F_{ST}$ and $\Phi_{ST}$-values; AMOVA and coalescence analyses), but not decisively ($D_{est}$-values). Coalescence time suggests that divergence of Canary Island from Western Sahara populations occurred about 5500 years ago. The increased aridification of the westernmost part of the Sahara region in the late Pleistocene–early Holocene could have been one of the driving forces for the colonization of the Canary Islands. At that time, the current Sahara Desert was either steppe at low elevation, or temperate xerophytic woods/scrub (Jolly et al. 1998), but since 5500 years ago aridification increased, resulting the desert we know today around 2700 years ago (Kröppelin et al. 2008). Aridification in semi-arid landscapes not only reduced water availability but also changed vegetation composition, likely affecting the trumpeter finch as well, even though it has been described as a ‘facultative drinker’ (as defined by Bartholomew 1972; MacMillen 1990) because it maximizes water intake when foraging on seeds (Carrillo et al. 2007b). Naturally, it is also possible that the Canary Islands were colonized by chance via jump dispersal without any habitat-related effects.

We found evidence of recent bottlenecks in Lanzarote and Fuerteventura. The heterozygote excess method implemented in Bottleneck should be able to detect a bottleneck of $N_e = 50$ for 25–250 generations (0.25–2.5 times $2N_e$; Cornuet & Luikart 1996) and $M$-ratios for a few hundred generations (Garza & Williamson 2001) after the initiation of the bottleneck. The bottlenecks detected from the Canary Islands cannot therefore reflect a founder event, but instead some other event that has reduced population sizes during historical times. The detection of a bottleneck also in the geographically nearby Western Sahara population may imply some common event for the region, although sample size is low and could constrain our findings.

The split between Maghreb and Iberia was dated as the most recent splitting event, about 4800 years ago. This timing seems quite old, as the trumpeter finch was first detected in southwestern Europe in the middle of the 19th century (López-Seoane 1861; Vayreda 1883), and first breedings were recorded as late as in the second half of the 20th century (García 1972). However, it is also possible that the population in Iberia could have persisted for a longer time remaining undetected or, more likely, that differentiation could already have occurred in the North African continent. In addition, incomplete lineage sorting, which can be seen throughout the studied range, confuses the timing of population splits. The case of the trumpeter finch mirrors that of the Dupont’s lark, which remained unnoticed in Europe until the 19th century (Vieillot 1820), but whose Iberian and Moroccan populations diverged, based on genetic data, 24 000 years ago (García et al. 2008). Whether the recent range expansions are a consequence of climate change (Carrillo et al. 2007a), human-induced habitat changes (Khoury & Al-Shamlih 2006) in the source or receiving area, or have happened just by chance, remains unknown. However, it is likely that the colonization has occurred by many individuals or it is still on-going as no bottlenecks, indicating the reduction in population sizes through recent founder events were detected (see also Barrientos et al. 2009b).

Ecological niche modelling and its congruence with genetic data

According to our models, the species should have had a wide range during the warm scenarios and should decrease their potential area in the glacial cycles. During MIS5, the climatic conditions were similar to the present (Varela et al. 2009), which support the expansion of the most recent common ancestor across a range similar to the current one (see also Drake et al. 2011). Subsequently, the climate became colder during the glaciations, and the trumpeter finch became isolated in several refugia like Iran, Iraq, Syria, Jordan, Saudi Arabia, southeastern Morocco, Algeria and Niger during LGM. This fact can lead to differences in genetic structure between geographically close populations. For instance, eastern Maghreb (Tunisia), a region that did not retain suitable climatic conditions during LGM, is genetically different from southwestern Maghreb (Morocco) as has also been seen in some other arid-adapted species (see above). The lack of differentiation between most of the continental populations based on
microsatellites could be explained by a recovery of climatic suitability in more recent times, and the subsequent population expansion across Asia, Africa and Europe. The climatic suitability found with ENM for the present is intermediate to high in some areas where the species has not been recorded (e.g. most of Turkey, coastal Egypt and Libya or northwestern Iran). The absence of occurrence data in a priori climatically suitable areas is likely due to the fact that the species has been poorly searched in several regions or, alternatively, could be related to dispersal limitations or biotic interactions. Future field work should add information to this open question and give more clues about the variables (other than climate-related ones) that are affecting the observed geographical expansion. Climatic suitability is also high in the central Iberian Peninsula, where it is known that the species is currently absent (Manrique et al. 2003), but which can be considered a potential region for continuing the on-going expansion of the trumpeter finch (Carrillo et al. 2007a).

On the basis of our genetic data complemented with ENM, we can conclude that both recent expansion in border regions and ancient colonization of North Africa, including later colonization of the Canary Islands, seem to have followed, at least partially, climatic events.

Acknowledgements

We thank CM Carrillo, J Benzal, L García, G Delgado, D Concepción, M López-Darias, J Hernández, M Nogales, JM Gasent, J Vázquez and D Ferrández for their help. G López-Iborra, JC Illera and R Moreno-Opo provided some samples, as well as J Cabot and M Sempere from Estación Biológica de Doñana, Y Yom-Tov and T Shariv from Tel Aviv University, J Fjeldsa and JB Kristensen from Copenhagen Zoological Museum, E Pasquet from Paris Natural History Museum, A Gamauf from Wien Natural History Museum and M Adams from British Natural History Museum. M. Jennings provided information on species range, and K Koivula and four anonymous reviewers improved the first drafts with their constructive commentaries. S. Young reviewed the English, and I. Jiménez helped us with the CSIC repository and J.M. Álvarez-Martínez with figure editing. R Barrientos received financial assistance from an FPI fellowship from the Spanish Ministry of Education and Science, from JCCM-FSE 2007/2013 postdoctoral program and from Juan de la Cierva program from the Spanish Ministry of Competitiveness. S Varela is supported by the project: ‘Support of establishment of the Charles University’, registration number CZ.1.07/2.3.00/20130022, supported by The Education for Competitiveness Operational Programme (ECOP) funded by the ESF and the government budget of the Czech Republic. The Spanish Ministry of Education and Science through its Project REN2002-00169 and CGL2005-01771, the European Regional Development Fund and the University of Oulu provided the financial support for this study. The Regional Governments of Andalusia and Fuerteventura, the Haut Commissariat aux Eaux et Forêts et à la Lutte Contre la Désertification of Morocco and A Jáurez in Coto Las Lomillas (Tabernas) granted permits for the field work and bird trapping. The Government of Fuerteventura also provided accommodation in the Estación Biológica de la Oliva.

References


© 2013 John Wiley & Sons Ltd


Guillaumet A, Crochet PA, Fons JM (2008) Climate-driven diversification in two widespread Galerida larks. BMC Evolutionary Biology, 8, 32.


Kvist L, Ponníkas S, Belda EJ et al. (2011) Decline in population size has different effects to genetic structures in two subspecies of the reed bunting (Emberiza schoeniclus witherbyi and *E. s. lusitanica*) in Iberian Peninsula. *Journal of Ornithology*, **152**, 681–693.


Sanz JJ (2002) Climate change, birds: have their ecological consequences already been detected in the Mediterranean region? Ardeola, 49, 109–120.
Vayreda E (1883) Fauna Ornitológica de la Provincia de Gerona. Imprenta Paciano Torres, Gerona, Spain.

Author information

RB has a general interest in the habitat ecology, biogeography and conservation of vertebrates, including the trumpeter finch, which he studied during his thesis. Recently, he has studied the role of habitat fragmentation in the life history of vertebrates. LK is a docent in molecular ecology and evolution. Her major research interests are the phylogeography and molecular ecology of several Palaearctic bird species. AB is an evolutionary ecologist interested in the study of the effects of climate change on the immune system in extreme environments. FV is interested in host–parasite interactions, sexual selection and expansive processes of birds in relation to human activity and environmental factors.

FK is an ornithologist specialized in arid-adapted birds. SV works on species range shifts, community changes across time and Quaternary mammal extinction events. EM is interested in disentangling the role of morphology in explaining species distribution. She is the leader of the project funding this trumpeter finch study.

R.B. and L.K. designed and performed research and wrote the article. L.K. analysed genetic data. A.B., F.V. and E.M. designed research and contributed to writing the article. F.K. contributed to collecting samples. S.V. performed ENM analyses and contributed to writing the article.

Data accessibility

Microsatellite data, DNA sequences and their corresponding GenBank Accession nos. are deposited in the CSIC institutional repository (http://digital.csic.es/handle/10261/85523). Geographical coordinates of the locations used to model the ecological niche of the trumpeter finch from its current range are provided in Table S4 (Supporting Information) and the R-script for running the ecological niche models are available in GitHub (https://github.com/SaraVarela/Bucanetes).

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Samples used, including population of origin, tissue from which the DNA was extracted, source, code (with ring or museum label) and GenBank Accession nos.
Table S2 Primers used to amplify the mitochondrial control region of the trumpeter finch.
Table S3 Geographical coordinates used in the isolation-by-distance analyses.
Table S4 Geographical coordinates of the locations used to model the ecological niche of the trumpeter finch from its current range.
Table S5 AMOVA analyses with different groupings from the microsatellite data.
Table S6 AMOVA analyses with different groupings using mitochondrial DNA.

Fig. S1 Historical scenarios constructed for the coalescence analyses with combined microsatellite and mitochondrial (576 bp fragment) data.

© 2013 John Wiley & Sons Ltd