

Genetic differentiation between insular and continental populations of migratory and resident warblers, the Great Reed Warbler *Acrocephalus arundinaceus* and the Cetti's Warbler *Cettia cetti*

Francesco Ceresa<sup>1\*</sup>, Eduardo J. Belda<sup>2</sup>, Laura Kvist<sup>3</sup>, Mira Kajanus<sup>3</sup>, Juan S. Monrós<sup>1</sup>

<sup>1</sup>Institute Cavanilles of Biodiversity and Evolutionary Biology, University of Valencia, C/Catedrático José Beltrán 2, ES-46980 Paterna (Valencia), Spain; <sup>2</sup>IGIC, Universidad Politécnica de Valencia, C/Paranimf 1, ES-46730 Gandía (Valencia), Spain; <sup>3</sup>Department of Ecology and Genetics, University of Oulu, PO Box 3000, FI-90014 Oulu, Finland.

---

\* Correspondence author. Email: francesco.ceresa01@gmail.com

1 **Abstract**

2

3 Island populations are frequently smaller than continental populations, have lower genetic diversity,  
4 are more inbred and show genetic differentiation from the mainland. However, sufficient numbers of  
5 immigrants may reduce the differentiation of insular populations and moderate the effects of genetic  
6 drift. In this study, we compared insular and continental populations of Cetti's Warbler *Cettia cetti*  
7 and Great Reed Warbler *Acrocephalus arundinaceus*. We assessed the degree of genetic  
8 differentiation between insular and continental birds, their demographic history and genetic diversity.  
9 We compared the results, taking into account the differences in migration strategy and morphology  
10 of the two warblers. We found slightly lower genetic diversity in the insular populations than in the  
11 continental birds, possibly because of the lower population size or reduced immigration. The genetic  
12 differentiation between island and mainland birds was low, but higher in the Cetti's Warbler than in  
13 the Great Reed Warbler, suggesting differences in the species' capability of crossing the sea. We found  
14 evidence for a past bottleneck in both the insular and continental populations of the Cetti's Warbler,  
15 while for the Great Reed Warbler we found no signs of past population reductions. High dispersal  
16 capability of the Great Reed Warbler may have allowed high gene flow, which may explain the  
17 observed interspecific differences in the demographic history.

18

19 **Keywords** Dispersal; Dispersal barrier; Gene flow; Insularity; Microsatellites

## 20 **Introduction**

21

22 Compared to mainland populations, island populations are often smaller, have a lower genetic  
23 diversity and are more inbred (Frankham et al. 2010). These characteristics imply a higher extinction  
24 risk, in fact most of the recently documented extinctions were of insular species, even though they  
25 represent a minority of all species (Frankham et al. 2010). In species with sufficient dispersal  
26 capability, island populations may receive immigrant individuals from the mainland. If such  
27 immigration leads to sufficient gene flow, it may reduce genetic differentiation between insular and  
28 mainland populations and moderate effects of genetic drift (Madsen et al. 1999; Hogg et al. 2006;  
29 Ortego et al. 2008; Frankham et al. 2010). Birds have generally high dispersal ability (Koenig et al.  
30 1996), but there are large differences among species, apparently related to characteristics such as  
31 migration strategy, population size and habitat preference (Paradis et al. 1998; Sutherland et al. 2000).  
32 The environmental characteristics of areas, which separate suitable breeding habitats, also influence  
33 the rate of gene flow (Frankham et al. 2010). Dispersal between islands and mainland imply crossing  
34 a more or less wide water body, which may represent for some bird species a more effective dispersal  
35 barrier than a similarly wide extension of land (Hodges and Krementz 1996; Matchans et al. 1996;  
36 Ceresa et al. 2015).

37 Knowledge of the degree of connectivity between insular and continental bird populations may be  
38 important for conservation purposes (e.g., Agudo et al. 2010) and for obtaining information about  
39 species' capability to disperse and cross geographical barriers. Furthermore, comparing species with  
40 different ecological and morphological characteristics may help to better understand the general  
41 dispersal patterns of birds.

42 In this study, we compare insular and continental populations of two passerine birds, the Cetti's  
43 Warbler *Cettia cetti* and the Great Reed Warbler *Acrocephalus arundinaceus*, based on their  
44 population genetic structure, genetic diversity and demographic history. The Cetti's Warbler breeds

45 in a wide range from Western Europe and North-Western Africa to Central Asia. European  
46 populations are mostly resident and breed in a variety of moist habitats such as marshes, lake- and  
47 riversides, reedbeds, and sometimes also in dry scrublands (Kennerley and Pearson 2010). The Great  
48 Reed Warbler is widely distributed across Europe, breeding mainly in inundated reedbeds (Kennerley  
49 and Pearson 2010). It migrates over long distances and overwinters in sub-Saharan Africa (Kennerley  
50 and Pearson 2010).

51 Both species breed on the island of Mallorca, which is one of the Balearic Islands located about  
52 150 km from the Spanish and 250 km from the Algerian coast. The Great Reed Warbler is presumed  
53 to have higher dispersal capability than the Cetti's Warbler, given its longer migration distance and  
54 more pointed wings (Kennerley and Pearson 2010), which allow higher flight efficiency (Norberg et  
55 al. 1990). We therefore expect a weaker impact of insularity in the Mallorcan populations of the Great  
56 Reed Warbler than in the Cetti's Warbler.

57

## 58 **Methods**

59

### 60 **Sampling and DNA extraction**

61 We obtained blood samples from 35 Cetti's Warblers and 44 Great Reed Warblers, mist-netted during  
62 the breeding season of 2012-2013 (18 May – 30 July of 2012 and 9 May – 5 July of 2013) at four  
63 marshland areas. For the Cetti's Warbler, we obtained samples from S'Albufera de Mallorca (39°47'  
64 N - 3°06' E; extension ~1700 hectares) in Spain and Larache (35°08' N - 6°05' W; ~3600 ha) in  
65 Morocco (Fig 1). For the Great Reed Warbler, we obtained samples from the Pego-Oliva Natural Park  
66 (38°51' N - 0°03' W; ~1250 ha) and Villafranca de los Caballeros (39°27' N - 3°19' W; ~200 ha), both  
67 located in the Spanish mainland, S'Albufera de Mallorca and Larache (Fig 1). Sampling took place  
68 late enough to exclude any possible wintering/migrating Cetti's Warblers, while during May both  
69 local and migratory Great Reed Warblers can occur at the same time in our study areas (Kennerley

70 and Pearson 2010; Clement 2017). However, all Great Reed Warblers sampled in May were breeding  
71 birds according to the development of the incubation patch or the cloacal protuberance (in females  
72 and males, respectively; Svensson 1992). Therefore, occurrence of non-local individuals in our  
73 sample is extremely unlikely. From Larache, we were able to sample only two Great Reed Warblers,  
74 thus we did not include this locality in site-specific statistical analyses (see below). Sample sizes of  
75 each site are given in Table 1. Blood samples (5-15  $\mu$ l) were obtained by puncturing the brachial vein  
76 and stored in 96% ethanol. DNA was extracted using UltraClean™ Blood Spin™ Kit (MoBio  
77 Laboratories) according to the manufacturer's protocol.

78

#### 79 DNA amplification

80 We amplified 13 polymorphic microsatellite loci from the Cetti's Warblers: Cu $\mu$ 28 (Gibbs et al. 1999,  
81 *Catharus ustulatus*), Gf05 (Petren 1998, *Geospiza fortis*), FhU2 (Ellegren 1992, *Ficedula hypoleuca*),  
82 Ppi2 (Martínez et al. 1999, *Pica pica*), Zl54 (Frentiu et al. 2003, *Zosterops lateralis*), Pca3 (Dawson  
83 et al. 2000, *Cyanistes caeruleus*), Hru6 (Primmer et al. 1995, *Hirundo rustica*), Pdo $\mu$ 1 (Neumann and  
84 Wetton 1996, *Passer domesticus*), Aar4 and Aar8 (Hansson et al. 2000, *Acrocephalus arundinaceus*),  
85 Esc3 and Esc6 (Hanotte et al. 1994, *Emberiza schoeniclus*) and Ase64 (Richardson et al. 2000,  
86 *Acrocephalus sechellensis*). From the Great Reed Warblers, we amplified 12 polymorphic loci:  
87 Cu $\mu$ 28, FhU2, Pdo5 (Griffith et al. 1999, *P. domesticus*), Ppi2, Zl54, Pca3, Aar4, Aar5 (Hansson et  
88 al. 2000, *Acrocephalus arundinaceus*), Aar8, Esc6, Ase34 (Richardson et al. 2000, *A. sechellensis*)  
89 and Ase64. We performed the Polymerase Chain Reaction (PCR) in 10- $\mu$ l volumes using 1  $\mu$ l of DNA  
90 template, 1  $\mu$ l of PCR-buffer, 1  $\mu$ l of dNTPs (2 mM), 0.8  $\mu$ l of primers, 0.06  $\mu$ l of DNA-polymerase  
91 (Biotools) and 0.5  $\mu$ l of MgCl<sub>2</sub> (0.4  $\mu$ l for HrU6, 0.6  $\mu$ l for Esc3). The PCR procedure consisted of  
92 an initial denaturation for 5 min (94°C), 35 cycles of denaturation for 30 s (94°C), annealing for 30 s  
93 (50°C) and synthesis for 45 s (72°C), and a final synthesis for 7 min (72°C). Annealing temperatures  
94 for the primers were 48°C for Pdo5 and Gf05, 52°C for Ppi2 and Ase64, 54°C for Cu $\mu$ 28 and ZL54,

95 and a touchdown from 50° to 45°C for Esc6, Hru6 and Esc3, from 63°C to 53°C for Aar5, FhU2,  
96 Pca3 and Pdoμ1, from 54°C to 48°C for Aar4 and Ase34, and from 60°C to 50°C for Aar8. We ran  
97 the PCR products on an ABI PRISM 3730 DNA Analyzer (Applied Biosystems) and scored the alleles  
98 with GeneMapper 4.0. We checked the data with the program MICROCHECKER 2.2.3 (van  
99 Oosterhout et al. 2004) for possible genotyping errors (null alleles, scoring errors due to stuttering,  
100 large allele dropouts).

101

## 102 Statistical analysis

103 We used GENEPOP 4.2 (Raymond and Rousset 1995; Rousset 2008) to assess possible deviations  
104 from the Hardy-Weinberg and linkage equilibrium separately for each sampling site and for the total  
105 sample of both species. We calculated the expected heterozygosity ( $H_e$ ) using Arlequin 3.5.1  
106 (Excoffier and Lischer 2010), and inbreeding coefficient ( $F_{IS}$ ) and allelic richness ( $A$ ) with FSTAT  
107 2.9.3 (Goudet 1995).

108 To investigate the genetic population structure within our samples, we used the Bayesian  
109 program STRUCTURE 2.3.4 (Pritchard et al. 2000; Falush et al. 2003), which allows estimation of  
110 the most probable number of distinct genetic clusters ( $K$ ) in the data set. We adopted a model with  
111 population admixture and correlated allele frequencies (Falush et al. 2003) and carried out the analysis  
112 both with and without prior spatial information of sampling sites (i.e. sampling locality). We  
113 performed ten independent runs for each value of  $K$  between 1 and 10, with a burn-in period of 50000  
114 iterations and 500000 Markov chain Monte Carlo (MCMC) replications. Starting from the  
115 STRUCTURE results, we also calculated the *ad hoc* statistics  $\Delta K$ , which estimates the second order  
116 rate of change of  $K$ -values between the consecutive numbers of genetic clusters. The highest value  
117 obtained is inferred as the best estimator of the actual number of clusters (Evanno et al. 2005).  
118 Although the program STRUCTURE is widely used to investigate the genetic structure of wild  
119 populations, using and comparing more statistical methods is recommended, especially in case of

120 weak population structuring or complex phylogeographic patterns (e.g., Frosch et al. 2014, Kraus et  
121 al. 2013, 2016, Wang et al. 2017). In order to obtain further insights about population structuring, we  
122 carried out also the discriminant analysis of principal components (DAPC), using package *adegenet*  
123 (Jombart 2008) version 2.1.0 in R 3.3.2 (R Core Team 2016). This method identifies genetic clusters  
124 through the K-means clustering algorithm, using the Bayesian Information Criterion (BIC). To obtain  
125 more reliable assignments of individuals to clusters, we used the *optim.a.score* function with 25  
126 simulations to determine the optimal number of principal components to be used in the DAPC.  
127 Furthermore, to estimate the genetic differentiation within and among sampling locations, we  
128 performed an analysis of molecular variance (AMOVA) with the program Arlequin, which was used  
129 also to calculate the pairwise  $F_{ST}$  values between the sampling sites of the Great Reed Warbler. To  
130 obtain further details about gene flow, the program MIGRATE-N 3.6.11 (Beerli and Felsenstein 1999,  
131 2001) was used to perform a maximum likelihood estimation of migration rates among sampling sites  
132 and among the genetic clusters identified by STRUCTURE and DAPC. We applied the stepwise  
133 mutation model, ran 10 short and one long chain, recorded every 100 steps of 50 000 or 500 000  
134 genealogies, respectively and discarded 10000 trees from each chain as a burn-in.

135 To investigate the demographic history of our insular and continental samples, we looked for  
136 past bottlenecks using the program BOTTLENECK 1.2.02 (Cornuet et al. 1996; Piry et al. 1999),  
137 which tests for heterozygosity excess caused by a recent reduction of the effective population size  
138 (Piry et al. 1999). We used the Wilcoxon test under the two-phase mutation model with 95% single-  
139 step mutations. We also estimated the Garza-Williamson index that compares the mean ratio of the  
140 number of alleles to the range in allele size (Garza and Williamson 2001) with Arlequin, in order to  
141 search for signs of population bottlenecks in a more remote past. This index decreases in a  
142 bottlenecked population. For the Cetti's Warbler, we found weak population structuring between  
143 samples from Mallorca and Larache and evidence for bottlenecks (see Results). Therefore, we used  
144 the program DIY ABC 2.0.3 (Cornuet et al. 2008) to estimate the timing and severity of population

145 size reductions, as well as to further assess their occurrence. DIY ABC allows to compare different  
146 competing historical/demographic scenarios, and for both Mallorca and Larache, we contrasted three  
147 alternative scenarios: the first with no population size changes, the second describing population  
148 growth and the third including a bottleneck (hereafter, scenario 1, 2 and 3, respectively). We used the  
149 default range of priors for the time of population size changes (10-10000 generations) and set priors  
150 for the effective population sizes to 10-10000 for Mallorca and 10-100000 for Larache. We applied  
151 the default Generalized Stepwise Mutation model (Estoup et al. 2002) and four default summary  
152 statistics. With these settings, a total of 3000000 simulated data sets were calculated (1000000 per  
153 scenario). Among them, the 10000 sets calculated for the most supported scenario and closest to the  
154 observed data according to the summary statistics were used for parameter estimations.

155

## 156 **Results**

157

### 158 Cetti's Warbler

159 Using MICROCHECKER, we found possible null alleles in loci Esc6 and Hru6. Therefore, we  
160 excluded these loci from calculations of the Hardy-Weinberg equilibrium, linkage disequilibrium and  
161 inbreeding coefficient, and also from BOTTLENECK analysis, while for the other analyses we used  
162 all 13 loci. We found no evidence of large allele dropouts or scoring errors in the data set. The total  
163 sample was not in Hardy-Weinberg equilibrium ( $\chi^2_{40} = 92.95$ ,  $p < 0.001$ ), nor was Larache ( $\chi^2_{22} =$   
164  $49.05$ ,  $p < 0.001$ ) or Mallorca ( $\chi^2_{18} = 43.90$ ,  $p < 0.001$ ). Linkage disequilibrium was detected in loci  
165 FhU2 - Pca3 and FhU2 - Pdo $\mu$ 1 from both sampling sites and the total sample and in Pca3 - Pdo $\mu$ 1  
166 from Mallorca and the total sample. After Bonferroni correction for multiple testing, linkage  
167 disequilibrium only between FhU2 and Pca3 remained significant (adjusted significant p-values:  
168 Mallorca,  $p < 0.0033$ ; Larache and total sample  $p < 0.0009$ ). Both FhU2 and Pca3 were anyway  
169 included in the analyses. Expected heterozygosity and allelic richness were higher in Larache than in



170 Mallorca (Table 1a).

171 In the analysis without geographical information of the sampling sites, program STRUCTURE  
172 attributed the highest likelihood to  $K = 1$ , while by adding geographical information (LOCPRIOR  
173 model) the highest likelihood for  $K = 2$  was obtained, also supported by  $\Delta K$ . Using the LOCPRIOR  
174 model, the assignment of individuals to two clusters corresponding to Mallorca and Larache is clearly  
175 seen in the bar plot produced by STRUCTURE. Without spatial information the difference between  
176 the two sites was less clear (Fig 3a). These results indicate weak population genetic structuring  
177 between the samples from Mallorca and Larache. The DAPC confirmed the occurrence of two genetic  
178 clusters (Fig. 4a), but in the continental sample some individuals were assigned to the ‘insular cluster’,  
179 in a clearer way than depicted by STRUCTURE analysis (Fig 5a); it is anyway possible that the  
180 DAPC is overconfident when assigning individuals to clusters (see Frosch et al. 2014). According to  
181 the AMOVA analysis, genetic differentiation between the two sampling sites was significant ( $F_{ST} =$   
182 0.064,  $p < 0.001$ ). The migration rate calculated with MIGRATE-N was higher from the continental  
183 to the insular sampling site ( $M = 7.643$ ; CI 95% 7.299 - 8.012) than in the opposite direction ( $M =$   
184 7.344; CI 95% 7.011 - 7.675), but the 95% confidence intervals of the two values overlapped widely.

185 We found signs of recent population reduction in Mallorca (Wilcoxon test, heterozygosity  
186 excess,  $p = 0.01$ ; shifted allele frequency class mode), but not in Larache (Wilcoxon test,  
187 heterozygosity excess,  $p = 0.91$ ; normal L-shaped allele frequency distribution). For both sampling  
188 sites, the Garza-Williamson index was lower than 0.68 (Table 1a), which indicates past bottlenecks  
189 (Garza and Williamson 2001). Past bottlenecks were indicated also by the DIY ABC analysis for both  
190 Mallorca and Larache. In Mallorca, scenario 3 (bottleneck) obtained the highest support according to  
191 the posterior probabilities calculated through both the direct (0.456) and the logistic approach (0.467),  
192 while the second best scenario described a constant population size (scenario 1; direct approach:  
193 0.406; logistic approach: 0.427). According to scenario 3, the bottleneck took place between 5060 (CI  
194 95% 979 - 9510) and 1900 (CI 95% 560 - 6360) generations ago. This would translate to

195 approximately 18700 - 7000 years ago, assuming a generation length of 3.7 years (BirdLife  
196 International 2016). The effective population size was 6820 (CI 95% 2580 - 9730) before the  
197 bottleneck, 1330 (CI 95% 194 - 3460) during and 3480 (CI 95% 712 - 8580) after it. The model  
198 checking performed for scenario 3 showed consistency between the posterior distributions and the  
199 observed data for two of the four summary statistics, while significant deviations ( $0.01 > p > 0.001$ )  
200 were found in the mean M index (Garza and Williamson 2001) and in the mean allele size variance  
201 across loci. For Larache, the bottleneck scenario (scenario 3) was strongly supported according to the  
202 posterior probabilities calculated through the direct and the logistic approach (0.946 and 0.973,  
203 respectively). The estimated bottleneck timing was between 3630 (CI 95% 638 - 8550) and 693 (CI  
204 95% 28 - 3030) generations ago, i.e. approximately between 13400 and 2500 years ago. The effective  
205 population size was 73900 (CI 95% 34800 - 98200) before the bottleneck, 1320 (CI 95% 279 - 3000)  
206 during and 40800 (CI 95% 4150 - 93100) after it. The model checking supported also the scenario 3,  
207 there were no significant deviations between the posterior distributions and the observed data of the  
208 four summary statistics.

209

210 Great Reed Warbler

211 MICROCHECKER did not detect null loci, large allele dropouts or scoring errors in the data set.  
212 There were no deviations from the Hardy-Weinberg equilibrium. Linkage disequilibrium was found  
213 between loci Cuu28 - ZI54, Pdo5 - Pca3, Cuu28 - Aar8 and Aar8 - Ase34 in Pego-Oliva, ZI54 - Aar8  
214 in Mallorca, ZI54 - Esc6 in Villafranca and Aar8 - Ase34 from combined data. However, after  
215 Bonferroni correction for multiple testing, none of these linkage disequilibria resulted to be  
216 significant (adjusted significant p-values: Mallorca,  $p < 0.0009$ ; Pego-Oliva, Villafranca and total  
217 sample,  $p < 0.0008$ ). All loci were used in the analyses. Expected heterozygosity and allelic richness  
218 were at very similar levels in Villafranca and Pego-Oliva, while lower in Mallorca (Table 1b).

219 Program STRUCTURE attributed the highest likelihood to  $K = 1$ , both with and without the

220 LOCPRIOR model. Consistently, the STRUCTURE bar plots of assignment likelihoods obtained for  
221  $K = 2$  (Fig 3b) showed almost no differences among individuals. It is not possible to calculate  $\Delta K$  for  
222  $K = 1$ , because  $\Delta K$  is based on the rates of change between the previous and the next  $K$ . In our case,  
223 we could only calculate  $\Delta K$  for  $K = 2-9$ , and for these values we did not observe any clear peaks of  
224 this statistics. More clarity was obtained with DAPC, which supported the occurrence of two distinct  
225 genetic clusters (Fig 4b) and highlighted the difference between the two Spanish continental sites,  
226 while the Mallorcan sample included birds from both clusters in similar proportions (Fig 5b). With  
227 the AMOVA analysis, we found a significant but low genetic differentiation among sampling sites  
228 ( $F_{ST} = 0.052$ ,  $p < 0.001$ ). Comparing the genetic differentiation between the pairs of sampling sites,  
229 we found significant though low differentiation in all comparisons (Mallorca – Pego-Oliva:  $F_{ST} =$   
230  $0.017$ ,  $p = 0.036$ ; Mallorca – Villafranca:  $F_{ST} = 0.063$ ,  $p = 0.018$ ; Pego-Oliva – Villafranca:  $F_{ST} =$   
231  $0.078$ ,  $p < 0.001$ ), and, consistently with the DAPC results, the highest differentiation was found  
232 between Pego-Oliva and Villafranca. The migration rates calculated between sampling sites using  
233 MIGRATE-N were lower from Mallorca to the continental sites than in the opposite direction and  
234 between the two continental sites (Table 2). We also merged sampling sites based on the DAPC  
235 clusters (cluster 1: Pego-Oliva + Mallorca; cluster 2: Villafranca + Larache; Fig. 5b); the migration  
236 rate was higher from cluster 2 to cluster 1 ( $M = 4.677$ ; CI 95% 4.291 - 5.085) than in the opposite  
237 direction ( $M = 1.758$ ; CI 95% 1.608 - 1.920).

238 We found no signs of recent population size reductions from BOTTLENECK analysis. Garza-  
239 Williamson indexes were higher than or almost equal to 0.68 (Table 1b), thus there were no  
240 indications for past bottlenecks.

241

## 242 **Discussion**

243

244 Cetti's Warbler

245 The lower expected heterozygosity and allelic richness found in Mallorca may be a consequence of  
246 fewer arrivals of immigrant individuals and of the lower population numbers in this island site than  
247 in the continental breeding site. Yet, the standard deviations of diversity estimates from Mallorca and  
248 Larache overlapped widely (Table 1a), thus the difference in genetic diversity between the two  
249 sampling sites is not strong.

250 The weak population structuring and the significant but low genetic differentiation suggest that  
251 the large distance (~ 960 km) and the sea limit gene flow between the two sampling sites, but do not  
252 prevent it. This suggests that Cetti's Warbler has a good dispersal ability, consistently with information  
253 from a mark-recapture study of the British population (Robinson et al. 2007). The recent colonization  
254 of Britain obviously shows the capability of this species to cross the sea during dispersal and our  
255 results indicate that the species is able to fly over stretches of sea notably wider than the English  
256 Channel. The slightly lower migration rate from Mallorca to Larache than in the opposite direction  
257 may indicate a lower tendency of the insular individuals to disperse.

258 The demographic history of the Cetti's Warbler was partly similar in the two sampling areas,  
259 although the bottleneck was more recent and severe in Larache than in Mallorca. The results suggest  
260 that the Western Mediterranean population of this species was reduced for a relatively long period  
261 (approx. 10000 years). The bottleneck period estimated for Mallorca started during the cold and dry  
262 period MIS2, included the Younger Dryas (approx. 13000 – 12000 yr BP; also a dry and cool period)  
263 and ended in a transition period to a wetter and less continental climate (Fletcher and Sánchez Goñi  
264 2008). A long period of severe climatic conditions in the Western Mediterranean, with the dominance  
265 of semi-desertic vegetation (Fletcher and Sánchez Goñi 2008), may explain the observed population  
266 reduction. In the case of the Larache population, the estimated bottleneck time started shortly before  
267 the Younger Dryas, but also included wetter periods with less continental climate; a possible  
268 relationship between the population size reduction and climate is therefore not clearly identifiable.

269

270 Great Reed Warbler

271 As observed for the Cetti's Warbler, the lower expected heterozygosity and allelic richness found in  
272 Mallorca may be explained by limited immigration or smaller population size in the Mallorca Island  
273 than in the continental breeding sites. However, also in this case the standard deviations of diversity  
274 estimates overlapped widely (Table 1b). Overall, our results of population differentiation indicate  
275 high levels of gene flow among sampling sites that is nevertheless limited, possibly due to long  
276 distances between sampling sites and high philopatry of this species (e.g., Procházka and Reif 2000;  
277 Hansson et al. 2002; Mátrai et al. 2012; Koleček et al. 2015). The different results obtained through  
278 STRUCTURE analysis (K=1) and DAPC (K=2) are probably due to the different assumptions of the  
279 analysis methods (Jombart et al. 2010), which can lead to identification of different numbers of  
280 distinct genetic clusters (Jombart et al. 2010; Kraus et al. 2013). In any case, the results of the two  
281 analysis methods are compatible; in fact, likelihood values calculated by STRUCTURE often fail in  
282 identifying the real number of genetic clusters (Evanno et al. 2005), and it was not possible to use the  
283 more reliable *ad hoc* statistics  $\Delta K$  to compare  $K = 1$  and  $K = 2$ . The strong connection among our  
284 sampling sites is consistent with the available information from mark-recapture data, which have  
285 revealed cases of long distance dispersal for up to more than 100 km (Hansson et al. 2002; Mátrai et  
286 al. 2012). Our results suggest that, at least in the case of our study population, the sea does not  
287 represent any more effective dispersal barrier than a similarly wide extension of land. In fact, Pego-  
288 Oliva is approximately equally distant from Mallorca and from Villafranca (~ 290 km), but the  
289 differentiation is even slightly higher between the two continental sites than between the island and  
290 mainland sites. The lower migration rates from the insular to the continental sites than in the opposite  
291 direction suggest a lower tendency to long-distance dispersal in the Mallorcan breeding population.  
292 The demographic history was similar in Mallorca and in the continental breeding areas, i.e. no  
293 population bottlenecks were detected.

294

295 Interspecific comparison

296 In both species, we observed slightly lower genetic diversity in the insular than in the continental  
297 breeding areas. A further pattern shared by both warblers is the similarity between the demographic  
298 history in the Mallorcan and the continental samples, although in the Cetti's Warbler the estimated  
299 bottleneck time was partly different between populations. The main interspecific difference is  
300 represented by the evidence for past bottlenecks found in the Cetti's Warbler but not in the Great Reed  
301 Warbler. This suggests that, in the case of the Great Reed Warbler, higher dispersal capability may  
302 have allowed constant arrival of individuals from other areas. Such capability may also have been  
303 advantageous for the Great Reed Warbler in facing climatic and environmental changes, e.g., by  
304 finding new breeding areas, and/or by migrating over long distances to find suitable wintering areas.  
305 The higher dispersal capability of the Great Reed Warbler is supported by the obtained migration  
306 rates among sampling sites, which were generally higher in this species than in the Cetti's Warbler.  
307 The interspecific comparison of the genetic differentiation and migration rates among sampling sites  
308 needs to be taken cautiously, given that most of Great Reed Warblers were sampled at sites separated  
309 by shorter distances than in the case of Cetti's Warblers. However, according to the STRUCTURE  
310 results, the two Great Reed Warblers from Larache were very similar to those sampled in the Spanish  
311 mainland and in Mallorca (Fig 3b), suggesting higher gene flow than in the Cetti's Warbler. A previous  
312 study carried out in the same area (including the Mallorca Island) on two other passerines also  
313 reported higher gene flow in the long-distance migratory species, the Eurasian Reed Warbler  
314 *Acrocephalus scirpaceus*, than in the mainly resident Moustached Warbler *Acrocephalus*  
315 *melanopogon* (Ceresa et al. 2015). Possibly, as hypothesized in Ceresa et al. (2015, 2016), also in our  
316 study, the higher gene flow in the long-distance migratory species may be partly explained by a higher  
317 capability of crossing the sea.

318

319 **Acknowledgements** - We acknowledge M. Rebassa, H. Rguibi-Idrissi, M. Marín, J. Gómez, Santi,

320 Luis, Kames and family and many others, who have helped us with the field work. We would also  
321 like to thank the authorities of the Marjal de Pego-Oliva Natural Park, the S'Albufera de Mallorca  
322 Natural Park, "Consejería de medio Ambiente y Desarrollo Rural de Castilla-La Mancha", the "Servei  
323 de Conservació de la Biodiversitat de la Generalitat Valenciana" and the "Haut Commissariat aux  
324 Eaux et Forêts et à la lutte contre la désertification de Rabat, Morocco" for providing the facilities to  
325 work in protected areas and for the relevant permits. We are grateful to M. Serra and the researchers  
326 and Ph.D. students of the Laboratory of Evolutionary Ecology (Institute Cavanilles of Biodiversity  
327 and Evolutionary Biology – University of Valencia), for providing the laboratory for DNA extraction.  
328 This study has been partly financed by Projects CGL2005-02041/BOS and CGL2010-21933-C02-02  
329 of the Spanish Ministry of Science and Innovation and by the University of Oulu. F. Ceresa was  
330 supported by an "Atraent talent" grant from the University of Valencia.

331

## 332 **References**

333

334 Agudo R, Rico C, Hiraldo F, Donázar JA (2011) Evidence of connectivity between continental and  
335 differentiated insular populations in a highly mobile species. *Divers Distrib* 17:1-12

336

337 Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective  
338 population numbers in two populations using a coalescent approach. *Genetics* 152: 763-773

339

340 Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective  
341 population sizes in n subpopulations by using a coalescent approach. *Proc Natl Acad Sci U S A* 98:  
342 4563-4568

343

344 BirdLife International (2016) *Cettia cetti*. The IUCN Red List of threatened species 2016:

345 e.T22714445A87542508. [http://dx.doi.org/10.2305/IUCN.UK.2016-](http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22714445A87542508.en)  
346 [3.RLTS.T22714445A87542508.en](http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22714445A87542508.en). Accessed 4 May 2017  
347  
348 Bonham PF, Robertson JCM (1975) The spread of Cetti's Warbler in north-west Europe. Br Birds  
349 68:393-408  
350  
351 Ceresa F, Belda EJ, Kvist L, Rguibi-Idrissi H, Monrós JS (2015) Does fragmentation of wetlands  
352 affect gene flow in sympatric *Acrocephalus* warblers with different migration strategies? J Avian Biol  
353 46:577-588  
354  
355 Ceresa F, Belda EJ, Monrós JS (2016) Similar dispersal patterns between two closely related birds  
356 with contrasting migration strategies. Popul Ecol 58:421427  
357  
358 Clement P (2017) Cetti's Warbler (*Cettia cetti*). In: del Hoyo J, Elliott A, Sargatal J, Christie DA, de  
359 Juana E (ed) Handbook of the Birds of the World Alive. Lynx Edicions, Barcelona.  
360 <http://www.hbw.com/node/58749>. Accessed 7 March 2017  
361  
362 Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent  
363 population bottlenecks from allele frequency data. Genetics 144:2001-2014  
364  
365 Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin JM, Balding DJ, Guillemaud T, Estoup A  
366 (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian  
367 computation. Bioinformatics 24:2713-2719  
368  
369 Dawson DA, Hanotte O, Greig C, Stewart IRK, Burke T (2000) Polymorphic microsatellites in the



370 blue tit *Parus caeruleus* and their cross-species utility in 20 songbird families. *Mol Ecol* 9:1941-1944  
371

372 Ellegren H (1992) Polymerase-chain-reaction (PCR) analysis of microsatellites - a new approach to  
373 studies of genetic relationships in birds. *Auk* 109:886-895  
374

375 Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their  
376 consequences for population genetics analysis. *Mol Ecol* 11:1591-1604  
377

378 Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the  
379 software STRUCTURE: a simulation study. *Mol Ecol* 14:2611-2620  
380

381 Excoffier L, Lischer H (2010) Arlequin suite ver 3.5: a new series of programs to perform population  
382 genetics analyses under Linux and Windows. *Mol Ecol Resour* 10:564-567  
383

384 Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus  
385 genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567-1587  
386

387 Fletcher WJ, Sánchez Goñi MF (2008) Orbital- and sub-orbital-scale climate impacts on vegetation  
388 of the western Mediterranean basin over the last 48,000 yr. *Quatern Res*  
389 doi:[10.1016/j.yqres.2008.07.002](https://doi.org/10.1016/j.yqres.2008.07.002)  
390

391 Frankham R, Ballou JD, Briscoe DA (2010) *Introduction to Conservation Genetics*, 2nd edn.  
392 Cambridge University Press, Cambridge  
393

394 Frentiu FD, Lange CL, Burke T, Owens IPF (2003) Isolation of microsatellite loci in the Capricorn

395 silvereeye, *Zosterops lateralis chlorocephalus* (Aves: Zosteropidae). *Mol Ecol Not* 3:462-464  
396

397 Frosch C, Kraus RH, Angst C, Allgöwer R, Michaux J, Teubner J, Nowak C (2014) The genetic legacy  
398 of multiple beaver reintroductions in Central Europe. *PloS one* 9: e97619  
399

400 Garza JC, Williamson EG (2001) Detection of reduction in population size using data from  
401 microsatellite loci. *Mol Ecol* 10:305-318  
402

403 Gibbs HL, Tabak LM, Hobson K (1999) Characterization of microsatellite DNA loci for a neotropical  
404 migrant songbird, the Swainson's thrush (*Catharus ustulatus*). *Mol Ecol* 8: 1551-1552  
405

406 Goudet J (1995) FSTAT: a computer program to calculate F-statistics. *J Hered* 86:485-486  
407

408 Griffith SC, Stewart IRK, Dawson DA, Owens IPF, Burke T (1999) Contrasting levels of extra-pair  
409 paternity in mainland and island populations of the house sparrow (*Passer domesticus*): is there an  
410 'island effect'? *Biol J Linn Soc* 68:303-316  
411

412 Hansson B, Bensch S, Hasselquist D, Lillandt BG, Wennerberg L, Von Schantz T (2000) Increase of  
413 genetic variation over time in a recently founded population of great reed warblers (*Acrocephalus*  
414 *arundinaceus*) revealed by microsatellites and DNS fingerprinting. *Mol Ecol* 9:1529-1538  
415

416 Hansson B, Bensch S, Hasselquist D, Nielsen B (2002) Restricted dispersal in a long-distance migrant  
417 bird with patchy distribution, the great reed warbler. *Oecologia* 130:536-542  
418

419 Hodges MF Jr, Krementz DG (1996) Neotropical migratory breeding bird communities in riparian

420 forests of different widths along the Altamaha River, Georgia. *Wilson Bull* 108:496-506

421

422 Hogg JT, Forbes SH, Steele BM, Luikart G (2006) Genetic rescue of an insular population of large  
423 mammals. *P Roy Soc Lond B Bio* 273:1491-1499

424

425 Jombart T (2008) Adegnet: a R package for the multivariate analysis of genetic markers.  
426 *Bioinformatics* 24:1403-1405

427

428 Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new  
429 method for the analysis of genetically structured populations. *BMC Genet* 11:94

430

431 Kennerley P, Pearson D (2010) Reed and Bush Warblers. Christopher Helm Publishers Ltd., London

432

433 Koenig WD, Van Vuren D, Hooge PN (1996) Detectability, philopatry, and the distribution of  
434 dispersal distances in vertebrates. *Trends Ecol Evol* 11:514-517

435

436 Koleček J, Jelínek V, Požgayová M, Trnka A, Baslerová P, Honza M, Procházka P (2015) Breeding  
437 success and brood parasitism affect return rate and dispersal distances in the great reed warbler. *Behav*  
438 *Ecol Sociobiol* 69:1845-1853

439

440 Kraus RH, Figuerola J, Klug K (2016) No genetic structure in a mixed flock of migratory and non-  
441 migratory Mallards. *J Ornithol* 157:919-922.

442

443 Kraus RH, Hooft P, Megens HJ, Tsvey A, Fokin SY, Ydenberg RC, Prins HH (2013) Global lack of  
444 flyway structure in a cosmopolitan bird revealed by a genome wide survey of single nucleotide

445 polymorphisms. *Mol Ecol* 22:41-55  
446  
447 Machtans CS, Villard MA, Hannon SJ (1996) Use of riparian buffer strips as movement corridors by  
448 forest birds. *Conserv Biol* 10:1366-1379  
449  
450 Madsen T, Shine R, Olsson M, Wittzell H (1999) Conservation biology: restoration of an inbred adder  
451 population. *Nature* 402:34-35  
452  
453 Martínez JG, Soler JJ, Soler M, Møller AP, Burke T (1999) Comparative population structure and  
454 gene flow of a brood parasite, the great spotted cuckoo (*Clamator glandarius*), and its primary host,  
455 the magpie (*Pica pica*). *Evolution* 53:269-278  
456  
457 Mátrai N, Gyurác J, Lenczl M, Hoffmann G, Bakonyi G, Mátics R (2012) Philopatry analysis of the  
458 great reed warbler (*Acrocephalus arundinaceus*) based on ringing data in Europe. *Biologia* 67:596-  
459 601  
460  
461 Neumann K, Wetton JH (1996) Highly polymorphic microsatellites in the house sparrow *Passer*  
462 *domesticus*. *Mol Ecol* 5:307-309  
463  
464 Norberg UM (1990) Vertebrate flight, mechanics, physiology, morphology, ecology and evolution.  
465 Springer Verlag, Berlin  
466  
467 Ortego J, Aparicio JM, Cordero PJ, Calabuig G (2008) Individual genetic diversity correlates with  
468 the size and spatial isolation of natal colonies in a bird metapopulation. *P Roy Soc Lond B Bio*  
469 275:2039-2047

470

471 Paradis E, Baillie SR, Sutherland WJ, Gregory RD (1998) Patterns of natal and breeding dispersal in  
472 birds. *J Anim Ecol* 67:518-536

473

474 Petren K (1998) Microsatellite primers from *Geospiza fortis* and cross-species amplification in  
475 Darwin's finches. *Mol Ecol* 7:1782-1784

476

477 Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent  
478 reductions in the effective population size using allele frequency data. *J Hered* 95:536-539

479

480 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus  
481 genotype data. *Genetics* 155:945-959

482

483 Procházka P, Reif J (2000) Analysis of ringing recoveries of great reed warblers (*Acrocephalus*  
484 *arundinaceus*) ringed or recovered in the Czech Republic and Slovakia. *Sylvia* 36:91–105

485

486 R Core Team (2016). R: A language and environment for statistical computing. R foundation for  
487 statistical computing, Vienna.

488

489 Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetic software for exact tests  
490 and ecumenism. *J Hered* 86:248-249

491

492 Richardson DS, Jury FL, Dawson DA, Salgueiro P, Komdeur J, Burke T (2000) Fifty Seychelles  
493 warbler (*Acrocephalus sechellensis*) microsatellite loci polymorphic in Sylviidae species and their  
494 cross-species amplification in other passerine birds. *Mol Ecol* 9:2225-2230

495

496 Robinson RA, Freeman SN, Balmer DE, Grantham MJ (2007) Cetti's Warbler *Cettia cetti*: analysis  
497 of an expanding population. *Bird Study* 54:230-235

498

499 Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows  
500 and Linux. *Mol Ecol Resour* 8:103-106

501

502 Sutherland GD, Harestad AS, Price K, Lertzman KP (2000) Scaling of natal dispersal distances in  
503 terrestrial birds and mammals. *Conservation Ecology* 4:16 URL  
504 <http://www.consecol.org/vol4/iss1/art16>. Accessed 23 Oct 2015

505

506 Svensson L (1992) Identification guide to European passerines, 4th edn. Svensson, Stockholm

507

508 Van Oosterhout CV, Hutchinson WF, Wills DPM (2004) MICROCHECKER: software for identifying  
509 and correcting genotyping errors in microsatellite data. *Mol Ecol Not* 4:535: 538

510

511 Wang E, Van Wijk RE, Braun MS, Wink M (2017) Gene flow and genetic drift contribute to high  
512 genetic diversity with low phylogeographical structure in European Hoopoes (*Upupa epops*). *Mol*  
513 *Phylogenet Evol* 113:113-125

514 Table 1. Sample sizes (N), expected heterozygosity ( $H_e$ ), allelic richness (A), inbreeding coefficient  
 515 ( $F_{IS}$ ) and Garza-Williamson index (G-W) of (a) Cetti's Warbler and (b) Great Reed Warbler, calculated  
 516 from microsatellite data for each sampling site and for the total sample. Standard deviation are shown  
 517 in parentheses.

a)

Sampling site	N	$H_e$ (SD)	A (SD)	$F_{IS}$ (SD)	G-W
Mallorca	15	0.503 (0.271)	3.164 (2.113)	-0.317 (0.447)	0.650
Larache	20	0.543 (0.272)	4.134 (2.852)	-0.071 (0.348)	0.612
Total	35	0.544 (0.266)	3.995 (2.657)	-0.110 (0.378)	0.631

b)

Sampling site	N	$H_e$ (SD)	A (SD)	$F_{IS}$ (SD)	G-W
Pego-Oliva	20	0.440 (0.259)	3.340 (2.202)	-0.136 (0.264)	0.739
Mallorca	10	0.390 (0.297)	2.783 (1.981)	-0.004 (0.422)	0.690
Villafranca	12	0.439 (0.251)	3.213 (2.304)	0.009 (0.320)	0.672
Total	42	0.442 (0.244)	4.358 (3.562)	-0.033 (0.234)	0.701

Table 2. Migration rates (M) among Great Reed Warbler's sampling sites, calculated from microsatellite data.

Direction	M (95% CI)
Mallorca → Pego-Oliva	8.202 (7.503 - 8.943)
Villafranca → Pego-Oliva	11.854 (11.010 - 12.739)
Pego-Oliva → Mallorca	13.298 (12.165 - 14.499)
Villafranca → Mallorca	10.758 (9.742 - 11.841)
Pego-Oliva → Villafranca	9.811 (9.077 - 10.584)
Mallorca → Villafranca	5.888 (5.326 - 6.488)



518 **Figure legends**

519

520 Figure 1. A map showing the sampling locations: (1) Larache, (2) Villafranca de los Caballeros, (3)  
521 Pego-Oliva Natural Park and (4) S'Albufera de Mallorca.

522

523 Figure 2. Cetti's warbler STRUCTURE results: mean likelihood ( $\pm$ SD) obtained for each K value  
524 between 1 and 10, (a) without and (b) including prior spatial information of sampling sites, and  $\Delta K$   
525 statistics obtained (c) without and (d) including prior spatial information of sampling sites.

526

527 Figure 3. Bar plots of the STRUCTURE results based on microsatellite data of (a) Cetti's Warbler and  
528 (b) Great Reed Warbler for K = 2, obtained (1) without geographic information about the sampling  
529 sites and (2) by adding such information. Each column represents an individual and indicates its  
530 probability to belong to one of the two clusters.

531

532 Figure 4. Number of distinct genetic clusters (K) identified through the discriminant analysis of  
533 principal components (DAPC) for (a) the Cetti's Warbler and (b) the Great Reed Warbler. For both  
534 species, the Bayesian information criterion (BIC) supported the occurrence of two clusters.

535

536 Figure 5. Bar plots of the discriminant analysis of principal components (DAPC) results based on  
537 microsatellite data of (a) Cetti's Warbler and (b) Great Reed Warbler for K = 2. Each column  
538 represents an individual and indicates its probability to belong to one of the two clusters.

539

540

541

542