Journal of Ornithology
https://doi.org/10.1007/s10336-021-01885-x

ORIGINAL ARTICLE

Disentangling the taxonomic status and phylogeographic structure of Marmora's (*Curruca sarda*) and Balearic Warbler (*Curruca balearica*): a genetic multi-marker approach

Davide Nespoli[1] · Irene Pellegrino[2] · Marco Galaverni[3] · Romolo Caniglia[4] · Joseph Sunyer[5] · Chiara Mengoni[4] · Ettore Randi[4,6] · Andrea Galimberti[7] · Diego Rubolini[1] · Fernando Spina[8] · Gabriel Gargallo[9] · Mattia Brambilla[10,11]

Received: 28 September 2020 / Revised: 14 March 2021 / Accepted: 5 April 2021

Abstract

Marmora's Warbler (*Curruca sarda*) and Balearic Warbler (*C. balearica*) are allopatric sibling species and were recently split mostly based on morphological and ethological characteristics. Here we provide the first phylogenetic and phylogeographic analyses of this species complex to support the taxonomic status of *C. sarda* and *C. balearica* in light of integrative taxonomy. We sampled the two taxa in most of their breeding ranges and we sequenced three mitochondrial and one nuclear gene region. All *C. balearica* individuals had private haplotypes for the four markers and formed monophyletic clades. Genetic distances between the two taxa were comparable with those found between other species belonging to the *Curruca* genus. Furthermore, most of the genetic variance was expressed at the interspecific level, rather than between different populations within taxa or between individuals within populations. Our results strongly support the current taxonomic status of these two warblers as distinct species.

Keywords DNA barcoding · Endemic species · Integrative taxonomy · Mediterranean

Communicated by M. Wink.

✉ Irene Pellegrino
irene.pellegrino@uniupo.it

[1] Dipartimento di Bioscienze, University of Milan, via Celoria 26, 20133 Milan, Italy

[2] Dipartimento di Scienze e Innovazione Tecnologica, Università del Piemonte Orientale, Viale Teresa Michel 11, 15121 Alessandria, Italy

[3] WWF Italia Onlus, Via Po 25/c, 00198 Rome, Italy

[4] Area per la Genetica della Conservazione (BIO-CGE), Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), via Cà Fornacetta 9, 40064 Ozzano dell'Emilia, BO, Italy

[5] Servei d'Agents de Medi Ambient, Conselleria d'Agricultura, Medi Ambient i Territori. Govern de les Illes Balears, Gremi Corredors 10, Polígon Son Rossinyol, 07009 Palma, Spain

[6] Section of Biology and Environmental Science, Department of Chemistry and Bioscience, Aalborg University, Fredrik Bajers Vej 7H, 9220 Aalborg Øst, Denmark

[7] ZooPlantLab, Dipartimento di Biotecnologie e Bioscienze, Università degli Studi di Milano-Bicocca, P.za della Scienza 2, 20126 Milano, Italy

[8] Area per l'Avifauna Migratrice, Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), via Cà Fornacetta 9, 40064 Ozzano dell'Emilia, BO, Italy

[9] Catalan Ornithological Institute (ICO), Plaça Leonardo da Vinci 4-5, 08019 Barcelona, Spain

[10] Sezione di Zoologia dei Vertebrati, Museo Delle Scienze, Corso del Lavoro e della Scienza 3, 38123 Trento, Italy

[11] Settore Biodiversità e Aree Protette, Fondazione Lombardia per l'Ambiente, Largo 10 Luglio 1976 1, 20822 Seveso, MB, Italy

Published online: 18 April 2021

Zusammenfassung

Taxonomischer Status und phylogeografische Struktur von Sarden- ***(Curruca sarda)*** **und Balearengrasmücke** ***(Curruca balearica)*****: Aufklärung durch genetische Analysen mehrerer Loci**

Sardengrasmücke (*Curruca sarda*) und Balearengrasmücke (*Curruca balearica*) sind allopatrische Zwillingsarten, welche kürzlich vor allem aufgrund morphologischer und ethologischer Merkmale voneinander abgetrennt wurden. Hier präsentieren wir die erste phylogenetische und phylogeografische Analyse dieses Artkomplexes, um den taxonomischen Status von *C. sarda* and *C. balearica* im Sinne der integrativen Taxonomie zu bestätigen. Wir beprobten beide Taxa über den Großteil ihrer Brutverbreitung und sequenzierten drei mitochondriale sowie eine nukleäre Genregion. Alle Individuen von *C. balearica* wiesen exklusive Haplotypen für die vier Marker auf und bildeten monophyletische Kladen. Die genetischen Abstände zwischen den beiden Taxa waren mit den bei anderen Arten aus der Gattung *Curruca* festgestellten vergleichbar. Weiterhin spielte sich die genetische Varianz überwiegend auf der interspezifischen Ebene ab, weniger zwischen verschiedenen Populationen innerhalb der Taxa oder zwischen Individuen innerhalb von Populationen. Unsere Ergebnisse liefern daher eine klare Bestätigung des derzeitigen taxonomischen Status dieser beiden Grasmücken als eigenständige Arten.

Introduction

The Mediterranean basin is one of the world's best explored areas in terms of ornithological research and it is considered a main biodiversity hotspot for the Western Palearctic (Myers et al. 2000). In fact, it acted as a center of origin and diversification for many taxa (e. g. Perktas et al. 2011; Pellegrino et al. 2015; Kakhki et al. 2018) due to glaciations and salinity crisis events (Taberlet et al. 1998). Since the advent of the Linnean classification, taxonomic research on Mediterranean bird species has usually been based on the analysis of morphological differences among populations. Such traditional approach might, however, have led to underestimating the actual degree of diversification occurring between isolated populations across the Mediterranean region. This divergence can be better evaluated in terms of genetic divergence and the development of reproductive isolation barriers, which in oscine birds are usually mediated by vocalizations and niche differentiation (Mason et al. 2017).

Recent studies on several bird taxa traditionally regarded as single 'biological' species have revealed the existence of morphologically similar (or nearly so) populations which are so highly diverging in terms of genetics and vocalizations that definitely belong to separate taxa (e.g. *Curruca subalpina, cantillans* and *iberiae*, Brambilla et al. 2008; Zuccon et al. 2020; *Muscicapa tyrrhenica,* Pons et al. 2016). This kind of research has recently led to similar results not only in birds but also in many other groups of vertebrates, including amphibians (e.g. North Italian tree frog *Hyla* sp., Dufresnes et al. 2018; *Pelodytes caucasicus-ibericus,* Díaz-Rodriguez et al. 2017), reptiles (e.g. *Tarentola mauritanica* complex, Rato et al. 2016) and mammals (e.g. *Myotis nattereri* complex, Juste et al. 2018; *Sciurus meridionalis,* Wauters et al. 2017), and to the identification of species and subspecies in invertebrate taxa, such as endangered crustaceans (e.g. Bernini et al. 2015) or insects (e.g. Galimberti et al. 2021). In this context, the application of advanced and integrative analytical methods is needed to shed light on the species complexes, but it is also useful in cases of apparently known and widespread taxa (Ancillotto et al. 2020; Zuccon et al. 2020; Pellegrino et al., 2017). The Mediterranean area is a biodiversity hotspot, but it is under threat from a high population density and the consequent human pressure on ecosystems, leading to biodiversity loss (Cutteload et al. 2008). On the islands, the Last Glacial Maximum, the distance from the continent and from other similar habitats, size and elevation (Steinbauer et al. 2013) modelled biodiversity and promoted phenotypic and genotypic differentiation, especially in animals with low dispersal abilities (such as reptilia, Escoriza 2020) but also in birds (Covas and Blondel, 1998). The Marmora's Warbler (*Curruca sarda*, Aves: Passeriformes: Sylviidae, formerly *Sylvia sarda*) is a Mediterranean endemic species with an insular distribution: it inhabits Sardinia, Corsica, Balearic Islands excluding Menorca and some smaller islands alongside Western Italy in the Tyrrhenian Sea. According to the 'traditional' taxonomy, it has been treated as a polytypic species with two allopatric subspecies, slightly differing in terms of plumage and bare-parts colouration: *Sylvia sarda balearica* (von Jordans, 1914) endemic to Balearic Islands, and the nominate *S. s. sarda* (Temminck, 1820), inhabiting the rest of the breeding range. Whereas a difference in vocalizations between the two subspecies was noted by von Jordans (1914), their taxonomic status had not been questioned until the early 2000s. Shirihai et al. (2001) proposed to elevate *S. s. sarda* and *S. s. balearica* to the rank of allospecies within the superspecies *S.* [*sarda*]. They based their taxonomic hypothesis on genetics, analyses of song, calls and morphometric divergence patterns, and secondarily, on plumage characters (less marked in juveniles and adult females) and bare-parts colouration. However, to date, no published genetic evidence about this taxonomic splitting and about the genetic structure of this species complex across the Mediterranean region has been provided.

Voelker and Light (2011), in a study about the influence of Palaeoclimatic events on the biogeographic distribution of *Sylvia* warblers, showed a phylogeny of the *Sylvia* genus based on ND2 (NADH dehydrogenase 2) and *cytb* (cytochrome b)

markers in which they a priori treated *S. sarda* and *S. balearica* as full species. On the basis of these results, Dickinson and Christidis (2014), Sangster et al. (2016), del Hoyo et al. (2016), and Shirihai and Svensson (2018) accepted them as two distinct species. Moreover, Dickinson and Christidis (2014) and more recently, Gill et al. (2020) recommended treating Marmora's and Balearic warblers as belonging to the *Curruca* rather than the *Sylvia* genus.

It seems, therefore, clear that a complete phylogeographic analysis is needed to definitely support the splitting of the two taxa and consequently to design appropriate management and conservation plans. Here, we aim to fill this knowledge gap in order to clarify the taxonomic status of the *Curruca* [*sarda*] species complex, integrating the already available information about morphology and vocalization with multi-marker genetic data and their geographical structuring.

Materials and methods

Field methods

Sample collection was carried out during 2013–2015 between May and June (which correspond to the peak of the breeding season; Brichetti and Fracasso 2010). Sampling sites covered 26 locations scattered within the breeding range of *C. sarda* and *C. balearica*, on the islands of Mallorca (Spain), Corsica (France), Sardinia and Elba (Italy) (Fig. 1). In order to obtain a sampling scheme balanced both over the whole breeding range of the species and within each island, we collected samples from localities as far apart as possible. Birds were trapped by mist-nets and biological samples were obtained by picking one or more feathers which were stored in 96% ethanol at − 20 °C until laboratory analysis. The birds were released immediately after sampling.

DNA extraction, amplification and sequencing

Total DNA was extracted from feathers using the DNeasy Blood and Tissue Kit (Qiagen), following the manufacturer's instructions, with the addition of 20 µl DTT (dithiothreitol) to the initial incubation step at 56 °C to accelerate the enzymatic digestion of membranes and protein structures. To examine the genetic differentiation and relationships within the species complex, we amplified and sequenced 2026 base pairs (bp) from three gene regions of the mitochondrial (mt) DNA and a nuclear (n) gene region, in order to verify potential mito-nuclear discordance phenomena, which may reflect incomplete lineage sorting or other kinds of processes which may not be highlighted by mtDNA alone (Ghorbani et al.

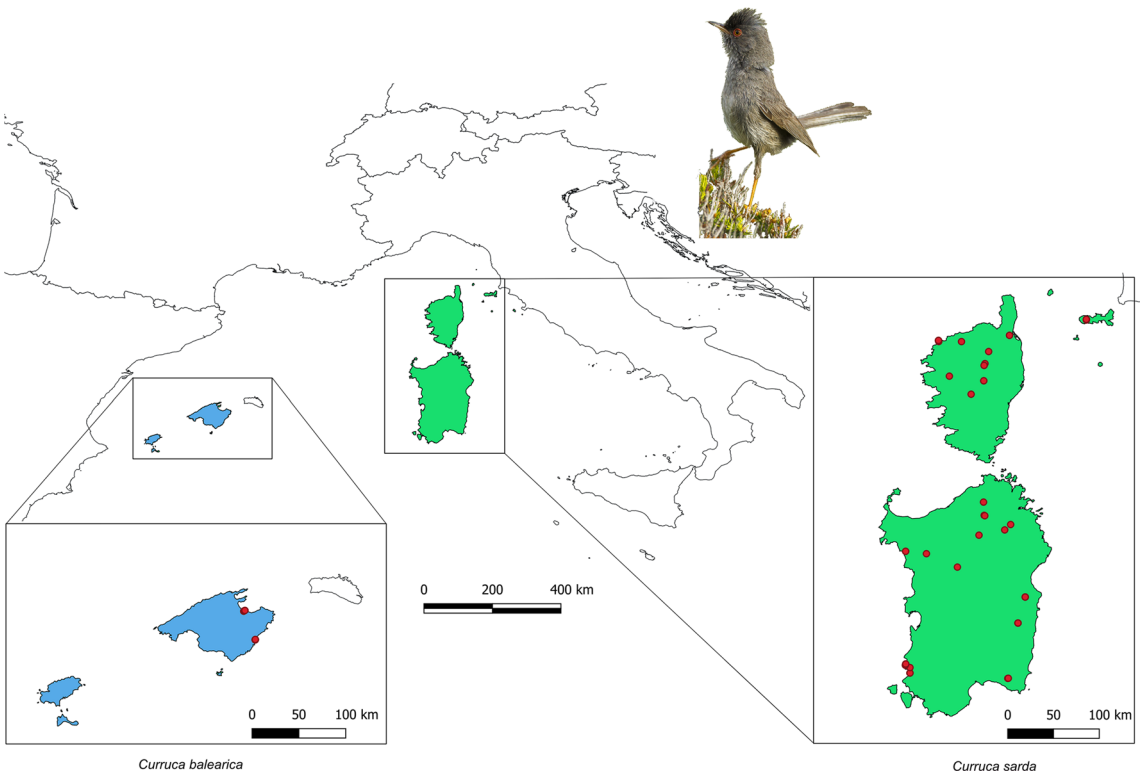

Fig. 1 Geographical distribution of sampling localities (in red): breeding range of *Curruca sarda* (in green) and *C. balearica* (in light blue)

2020; Ottenburghs et al. 2020). Specifically, we sequenced the second half of the mitochondrial cytochrome b (*cytb*) (537 bp, with primers 648L and 1137H, Brambilla et al. 2008) and the nuclear Glyceraldehyde-3-phosphate dehydrogenase (G3PDH) intron 11, spanning 358 bp, 274 bp of intron 11 and 84 bp of flanking exons, (primers G3P13b and G3P14b, sequenced with primers G3PIntl1 and G3P14b, Irestedt et al. 2002) for all the sampled individuals. Furthermore, in order to improve phylogenetic resolution between taxa, we also sequenced a fragment of 828 bp of the mitochondrial ND2 (with primers L5215, Hackett 1996 and H1064, Drovetski 2004) and a fragment of 661 bp of cytochrome c oxidase (COI; primers BirdF and BirdR1, Hebert et al. 2004) for a subsample of individuals representing the investigated populations (Supporting Information Appendix 1). Amplifications were carried out in 10-µl reactions using 2 µl of template DNA solution, 0.80 µl 10 × Taq buffer with 15 mM $MgCl_2$ (5Prime, Gaithersburg, MD, USA), 0.80 µl of 0.2% BSA, 0.15 µl of each 0.2 mM primer, 0.40 µl of 2.5 mM dNTP, 0.2 U of Taq polymerase (5Prime, Gaithersburg, MD, USA), with the following thermal profile: 94 °C for 2 min; 45 cycles at 94 °C for 30 s, Ta°C for 30 s, 72 °C for 30 s; 72 °C for 10 min; Ta was 55 °C for *cytb* and ND2 and 48 °C for COI. PCR temperature profile for the amplification of G3PDH consisted of the following steps: 94 °C for 2 min; 35 cycles at 94 °C for 35 s, 57 °C for 35 s, 72 °C for 50 s; 72 °C for 10 min. Amplicons were purified by enzymatic digestion of excess of primers with ExoSap-IT (Affymetrix, Santa Clara, CA) and the samples were bi-directionally sequenced using Big-Dye terminators (Thermo Fisher Scientific, Waltham, MA, USA) in an ABI Prism 3100 XL Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Phylogenetic analyses

Sequences were manually proofread and corrected in Seqscape v2.5 (Applied Biosystems, Foster City, CA, USA) and aligned with ClustalW in MEGA7 (Kumar et al. 2016) with default parameters. In order to exclude the amplification of nuclear pseudogenes of mitochondrial origin (NUMTs) we verified the absence of stop codons by translation of all sequences with the EMBL-EBI Emboss Transeq tool (https://www.ebi.ac.uk/Tools/st/emboss_transeq/) with the subsequent options: frame = F (Forward three frames), codon table = "Vertebrate Mitochondrial". Since no G3PDH sequences contained insertions/deletions or double heterozygotes, genotype phasing was straightforward and carried out using Phase 2.1 implemented in DnaSP v5 (Librado & Rozas 2009) without ambiguity. All sequences were trimmed in Bioedit (Hall 1999) and collapsed to unique haplotypes using DnaSP v5 (Librado and Rozas 2009). The unrooted minimum spanning networks were obtained using the median-joining algorithm (Bandelt et al. 1999) implemented in PopART (http://popart.otago.ac.nz). All pairwise distances, within and among groups, were calculated using the p-distance method in MEGA 7; standard error estimates were obtained by a bootstrap procedure (1000 replicates). Estimators of molecular diversity, i. e. haplotype number, haplotype diversity (Hd), average number of nucleotide differences (k) and nucleotide diversity (π), Tajima's D (Tajima 1989) and Fu and Li's D* and F* (Fu and Li 1993) tests for neutrality were computed for the analysed genetic markers as implemented in the software DnaSP v5 for the entire dataset and for the single islands.

The identification of the best nucleotide substitution model for each dataset was carried out in jModelTest 2 (Darriba et al. 2012), among all the 203 possible models available (integrating the estimate of nucleotide frequencies, the proportion of invariable sites and a gamma model with 4 categories), using all the possible evaluation criteria (AICc, BIC, DT). The base tree was ML optimized and the base tree search parameter was set on "best". jModelTest identified as the best fit models for *cytb*, ND2, COI mitochondrial haplotypes, respectively, HKY + I, TPM1uf + I, GTR + I.

The Bayesian Inference trees were generated using MrBayes 3.2.6 (Ronquist et al. 2012) with the selected models for each gene and the concatenated dataset; 3 cold and 1 hot Metropolis-coupled Markov chain Monte Carlo chains were run for $2*10^7$ generations and sampled every 1000 generations, with default parameters. Two runs with different starting points were carried out. The first 10% of generations were discarded as burn-in, and the posterior probability (PP) values were calculated for the remaining $1.5*10^7$ generations or more until the average standard deviation of split frequency was below 0.01. Calculated trees were drawn using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree). We added, to the mtDNA datasets, one or more sequences to be used as an outgroup (*Curruca undata* sampled on Mallorca during the ringing campaign of this study and COI GenBank ID KP780415; *Sylvia atricapilla* GenBank ID AM889140, *Curruca communis* GenBank ID GU571643; *Curruca nisoria* GenBank ID GU571647). The concatenated dataset was obtained with only mitochondrial haplotypes due to the very different evolutionary history and relatively rapid mutation/substitution rates that the nuclear and mitochondrial genes have in animal taxa (Brown et al. 1979).

In the main text, we only show the tree originated from the concatenated mtDNA sequences (Fig. 3); individual gene trees are shown in the online Supporting Information Appendix 3.

Analysis of Molecular Variance (AMOVA) was conducted using Arlequin v3.5 (Excoffier and Lischer 2010) to examine the structure of molecular variation. AMOVA calculates Φ statistics, which are analogous to F-statistics, in order to assess how much of the molecular variation is

explained at different levels of structuring. Φ_{CT} values indicate the amount of variation explained among groups, Φ_{SC} values indicate the amount of variation explained among populations within groups and Φ_{ST} values indicate the amount of variation explained within populations. In the analyses we defined the two taxa as "groups" and the four islands as "populations" (Mallorca for *C. balearica*; Sardinia, Corsica and Elba for *C. sarda*).

Results

We sampled 62 individuals from 27 sites distributed in four different regions (13 individuals from Mallorca, 4 from Elba island, 31 from Sardinia and 14 from Corsica), from which we obtained a total of 191 sequences (Supporting Information Appendix 1). Genetic diversity indices calculated for the four markers are reported in Table 1. All mtDNA sequences were functionally translated without premature terminations, with no evidence of NUMTs.

Median Joining haplotype networks (Fig. 2) identified five unique haplotypes from *cytb* sequences, 12 from ND2 and seven from G3PDH, while only two haplotypes were found for COI (Table 1, Supporting Information Appendix 1). Haplotype sequences were deposited in GenBank (refer to Supporting Information Appendix 2 for accession numbers). No shared haplotypes were observed between the two species for both mitochondrial and nuclear markers. Haplotype *cytb*-3 (*cytb* marker) and *nd2-4* (ND2 marker) were the most frequent, as they were found in 29 and 31 individuals, respectively (Fig. 2). Haplotype *g3pdh-2* was the most frequent for G3PDH marker with 42 occurrences in *C. sarda*. All these haplotypes were shared by the three *C. sarda* populations. All *C. balearica* samples were homozygotes for a private allele, *g3pdh-1*, while the remaining haplotypes were shared by individuals sampled in at least two of the other islands (Sardinia, Corsica and Elba), except for *g3pdh-4*, which occurred only in Sardinian birds (Fig. 2). We identified eight haplotypes for the concatenated alignment, three found in Mallorca individuals and five in Elba, Corsica, and Sardinia individuals (Fig. 2). Concatenated mtDNA haplotypes showed 98 parsimony informative sites.

We estimated a haplotype diversity ranging between 0.154 and 0.758; the lowest value was detected in *C. balearica*

Table 1 Genetic diversity indices for mitochondrial and nuclear markers for *Curruca balearica* from Mallorca and three populations of *C. sarda*: number of individual (haploid) sequences, number of haplotypes, nucleotide diversity (π) ± standard deviation, haplotype diversity (Hd) ± standard deviation, and average number of nucleotide differences (k), Tajima's D (*D*), Fu & Li's F*(*F**) and D* (*D**)

Pop	Sequences	Haplotypes	π	Hd	k	*D*	*D**	*F**
cytb								
C. balearica Mallorca	13	2	0.00029 ± 0.00023	0.154 ± 0.126	0.15385	− 1.149	− 1.365	− 1.481
C. sarda Elba	4	2	0.00093 ± 0.00049	0.500 ± 0.265	0.50000	− 0.612	− 0.612	− 0.478
Sardinia	31	3	0.00108 ± 0.00012	0.546 ± 0.038	0.57634	0.307	− 0.748	− 0.519
Corsica	14	2	0.00049 ± 0.00025	0.264 ± 0.136	0.26374	− 0.341	0.715	0.508
Total	49	5	0.00100 ± 0.00010	0.509 ± 0.037	0.53401	0.341	− 0.887	− 0.609
ND2								
C. balearica Mallorca	13	3	0.00084 ± 0.00018	0.603 ± 0.088	0.692	0.220	− 0.476	− 0.337
C. sarda Elba	4	1	0	0	0	0	0	0
Sardinia	25	7	0.00077 ± 0.00027	0.403 ± 0.124	0.640	− 2.221**	− 3.550**	− 3.673**
Corsica	10	3	0.00048 ± 0.00025	0.378 ± 0.181	0.400	− 1.429	− 1.657	− 1.797
Total	39	9	0.00063 ± 0.00019	0.387 ± 0.097	0.52149	− 2.291**	− 3.897**	− 3.975**
G3PDH								
C. balearica Mallorca	24	1	0	0	0	0	0	0
C. sarda Elba	8	2	0.00120 ± 0.00047	0.429 ± 0.169	0.42857	0.333	0.888	0.825
Sardinia	52	6	0.00259 ± 0.00031	0.653 ± 0.061	0.92760	− 0.027	-0.113	− 0.101
Corsica	22	5	0.00306 ± 0.00038	0.758 ± 0.042	1.09524	0.179	0.114	0.153
Total	82	4	0.00259 ± 000,025	0,674 ± 0,043	0.92894	0.320	0.958	0.886

Population diversity indices on COI sequences are null because of the lack of variability within populations

Significance levels: *$P < 0.05$; **$P < 0.02$

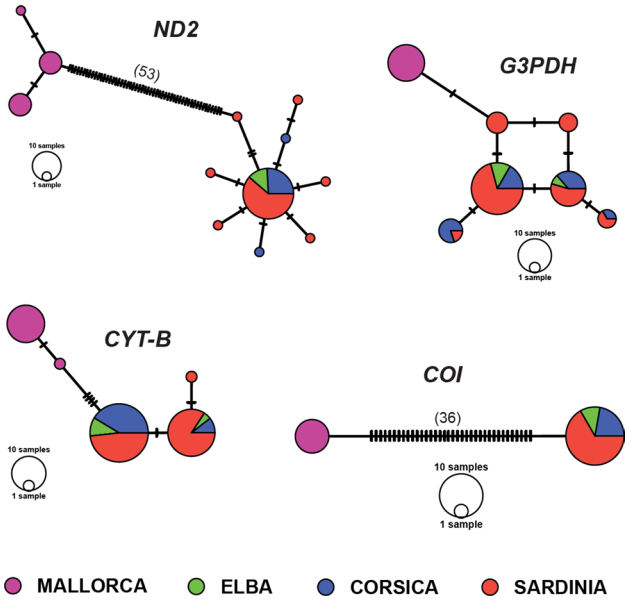

Fig. 2 Median-joining networks of ND2, G3PDH, *cytb* and COI haplotypes of *Curruca balearica* and *C. sarda* (see Supporting Information Appendix 1 and 2). Each circle represents a haplotype and circle size is proportional to haplotype frequency. Colours indicate different sampling islands. Dashes represent substitutions and when > 10, the number of substitutions is also reported within brackets

from Mallorca (*cytb* marker), while the largest value was detected in *C. sarda* from Corsica (nuclear marker). Except for these two extreme values, the other haplotypes' diversities ranged between 0.378 and 0.603. Nucleotide diversity ranged from 0.00029 (SD = 0.00023) in Mallorca (*cytb* marker) to 0.00306 (SD = 0.00038) in Corsica (G3PDH marker) (Table 1). The analysis of COI sequences showed a marked difference among the two taxa and lack of genetic diversity within populations. We found only two haplotypes: one for *C. balearica* (Mallorca population haplotype *coi-1*) and one for *C. sarda* (Elba, Sardinia and Corsica, haplotype *coi-2*) differing for 36 mutations. Tajima's D and Fu and Li's D* and F* tests showed positive but not significant values, except for the ND2 sequences of Sardinian populations, which showed negative and significant values; genetic diversity tests did not detect any bottleneck or directional selection events (Table 1).

A Bayesian tree, computed with the concatenated mitochondrial sequence alignment (*cytb*, COI and ND2, 2026 bp long; 8 haplotypes; 23 individuals) identified two clusters corresponding to the two species (Fig. 3): one clade included all the individuals of *C. balearica* sampled in Mallorca, and the other one included all the individuals from Sardinia, Corsica and Elba. These structure and clade separations were also confirmed by the Bayesian trees obtained using each dataset independently (*cytb*, ND2 and COI; Supporting Information Appendix 3).

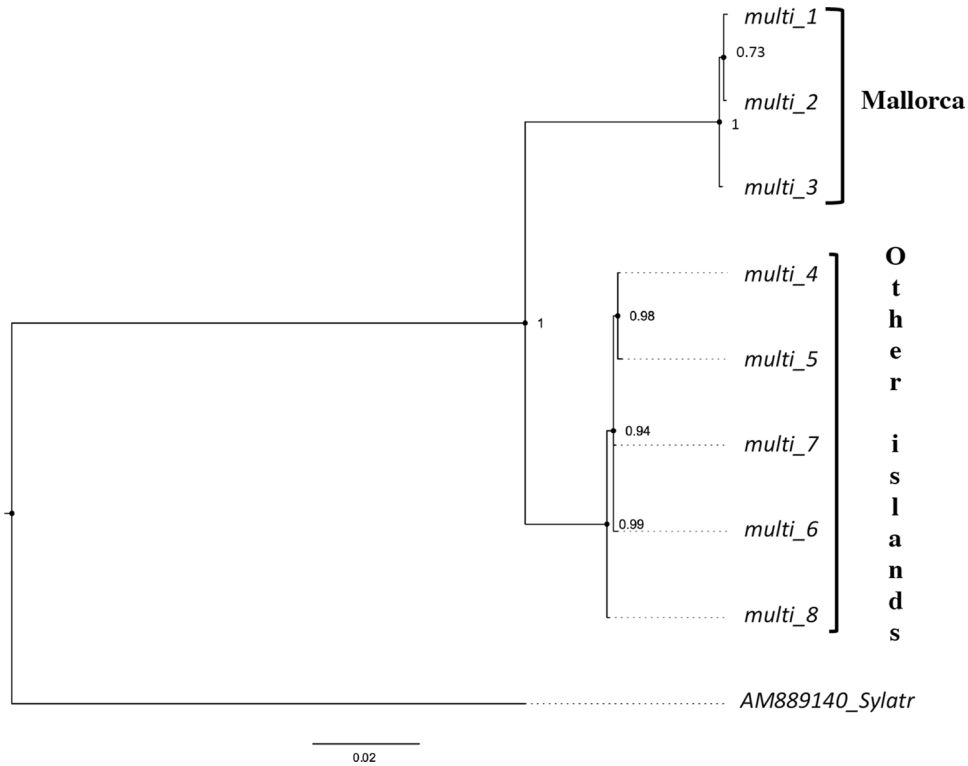

Fig. 3 50% majority consensus rule Bayesian Inference tree inferred from 2026 bp concatenated mitochondrial haplotypes. Bayesian Posterior Probabilities of each node are reported

Table 2 Genetic distances (uncorrected p-distances ± S.E.) within and among different island populations. Standard errors were calculated by means of 1000 bootstrap replicates

G3PHD	Within group	*C. sarda* All pop	Sardinia	Corsica	Elba
C. balearica Mallorca	0	0.0059 ± 0.0034	0.006 ± 0.003	0.007 ± 0.004	0.005 ± 0.003
C. sarda all pop	0.002 ± 0.001				
Sardinia	0.002 ± 0.002			0.003 ± 0.002	0.003 ± 0.001
Corsica	0.003 ± 0.002				0.003 ± 0.002
Elba	0.004 ± 0.001				
cytb		*C. sarda*	Sardinia	Corsica	Elba
Mallorca	0.0003 ± 0.0003	0.0100 ± 0.0040	0.0103 ± 0.0042	0.0094 ± 0.0041	0.0096 ± 0.0041
C. sarda	0.0010 ± 0.0009				
Sardinia	0.0011 ± 0.001			0.0011 ± 0.0011	0.0011 ± 0.0010
Corsica	0.0005 ± 0.0005				0.0006 ± 0.0006
Elba	0.0009 ± 0.001				
ND2		*C. sarda*	Sardinia	Corsica	Elba
C. balearica—Mallorca	0.0008 ± 0.0006	0.0692 ± 0.008	0.067 ± 0.0090	0.0674 ± 0.00903	0.0671 ± 0.0001
C.sarda					
Sardinia	0.0008 ± 0.002			0.0006 ± 0.0001	0.0004 ± 0.0001
Corsica	0.0004 ± 0.002				0.0002 ± 0.0002
Elba	0.0000 ± 0.000				
COI		*C. sarda*	*C. sarda* Sardinia	*C. sarda* Corsica	*C. sarda* Elba
Mallorca	0	0.0545 ± 0.0088	0.0545 ± 0.0088	0.0545 ± 0.0088	0.0545 ± 0.0088
C. sarda					
Sardinia	0			0	0
Corsica	0				0
Elba	0				
Concatenated haplotypes		*C. sarda*	Sardinia	Corsica	Elba
Mallorca	0.0004 ± 0.0003	0.0478 ± 0.0048	0.0501 ± 0.005	0.0478 ± 0.0047	0.0477 ± 0.0047
C. sarda	0.0005 ± 0.0003				
Sardinia	0.0006 ± 0.0003			0.0004 ± 0.0002	0.0004 ± 0.0003
Corsica	0.0003 ± 0.0003				0.0002 ± 0.0002
Elba	0.0005 ± 0.0005				

Genetic distances within and between different populations are reported in Table 2. Distances within populations were mainly low, ranging from 0 (no divergence in COI within any population) to 0.004 (Elba population for G3PHD), whereas pairwise comparisons between the two species showed higher values, ranging from 0.001 for *cytb* to 0.069 for ND2. The degrees of divergence between *C. sarda* populations (Sardinia, Corsica and Elba) were comparable with their respective within-population variation levels.

AMOVA results for each mtDNA marker and the concatenated dataset are reported in Supplementary Information Appendix 4. Analyses at both loci revealed that most part of the molecular variation (ranging from 90.48% to 100% for mtDNA markers and about 70% for the nuclear G3PDH) was attributable to differences between the two species. The lowest variation occurred among different populations within each taxon.

Hierarchical structure levels were significant at $\alpha = 0.05$. In particular, group-level structuration always obtained a *p*-value close to the theoretical minimum *p*-value.

Discussion

Our results indicate a clearly defined genetic differentiation between the two taxa previously suggested to be distinct species, *C. balearica* and *C. sarda*. All *C. balearica*

individuals possessed private haplotypes for each analysed marker, whereas *C. sarda* individuals from distinct populations largely shared the remaining haplotypes. The existence of private alleles was indicative of a prolonged isolation and a degree of differentiation which is coherent with the recently proposed two-species taxonomic suggestion. AMOVA again highlighted that most of the molecular divergence between individuals of the *C.* [*sarda*] complex is mainly due to the substructure in two different blocks, coinciding with the two currently recognized species. The divergence between individuals inhabiting different islands of the Mediterranean (Sardinia, Corsica, and Elba) was very low and far lower than the average differentiation level between individuals inhabiting any single island.

The amount of divergence between *C. sarda* and *C. balearica* is quite high for populations separated by only a few hundreds of kilometres and greater than the divergence between other sister species within the *Curruca* genus (e.g. *C. crassirostris* and *C. hortensis*; Voelker and Light 2011). The genetic distance calculated between *C. sarda* and *C. undata* on COI marker (4.75%) was similar to that detected between *C. sarda* and *C. balearica* (5.8%), while the genetic distance between Dartford Warbler *C. undata* and *C. balearica* (3.51%) showed smaller values; this evidence supports previous conclusions that *C. undata* and *C. balearica* would be regarded as sister species (Voelker and Light 2011). Similar results were found by Cai et al. (2019) in a multi-locus phylogeny of babblers, confirming the paraphyly of *C. sarda* as a traditionally defined polytypic species including *C. balearica* as a subspecies and reinforcing the need for treating *C. sarda* and *C. balearica* as distinct species. Overall, the huge genetic differences found between *C. sarda* and *C. balearica* suggest a long-term interruption of gene flow between their populations; a similar pattern was found by Pons et al. (2016) between continental and insular populations of *Muscicapa* (*M. striata vs M. tyhrrenica*) and, in a striking parallel with our case study, between *Muscicapa tyrrhenica thyrrenica* (inhabiting Corsica and Sardinia) and *M. t. balearica* (Balearic Islands); in both these cases the Mediterranean islands seem to have played an important role in taxa isolation and differentiation.

Our results clearly indicate the existence of two different and independent genetic characters (nuclear and mitochondrial markers) which allow to distinguish between *C. sarda* and *C. balearica* individuals; the two taxa may, therefore, be considered "full" species when strictly applying the original Phylogenetic Species Concept (Cracraft 1983). Moreover, these genetic differences are supported by divergences in vocalizations, biometry, plumage and bare-parts colouration (detailed in Shirihai et al. 2001). The existence of multiple independent traits allowing the recognition of two or more taxa (or Operational Taxonomic Units) is considered a solid criterion of species delimitation according to the paradigm of "integrative taxonomy" (Padial et al. 2010; Schlick-Steiner et al. 2010; Galimberti et al. 2012). On the whole, the present study improves our knowledge on two Mediterranean birds, allowing to recognize two taxa as two distinct species. It would be worth exploring, in the near future, other interesting biological attributes of the two species, in particular ecological needs and possible competition with other warbler species within the Mediterranean region. Moreover, detailed monitoring and estimates of population size of these species endemic to Mediterranean islands are needed to assess their conservation status and evaluate potential threats for their long-term survival. Indeed, small and isolated island populations may be at immediate risk of local extinction arising from stochastic environmental events and the adverse genetic effects deriving from inbreeding (as observed in different avian species, e.g. Pertoldi et al. 2012; Andersen et al. 2017; Kangas et al. 2018).

Since several Mediterranean ecosystems are at risk because of land-use change, urbanization, fragmentation and recurrent fires, it is essential to consider shared and island-specific threats to these endemic species characterized by such a very restricted distribution. In particular, *C. balearica* is sedentary and endemic to the Balearic Islands, excluding Menorca; recently, the impinging touristic development has been identified as a cause of disturbance and decline for birds and other vertebrate populations on these islands (Krawczyk et al. 2019). Recognizing *C. balearica* as a distinct species should call for special protection measures for this separate conservation unit. *C. sarda* has also a rather restricted range limited to Sardinia, Corsica and some little islands off Western Italy (plus small wintering areas in north Africa and Sicily; Shirihai et al. 2001) and is potentially vulnerable to the same threats affecting *C. balearica*. Moreover, both species could be subject to the negative effects of climate changes because of their restricted range and mostly insular distribution (Doswald et al. 2009).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10336-021-01885-x.

Acknowledgements We thank two anonymous referees for helpful comments on the previous version of our manuscript. We are indebted to Olivier Dehorter (CESCO, CRBPO) for his help during the planning of Corse field activities. All field activities and laboratory analyses comply with the current laws of the countries in which they were performed. In particular, ringing, sampling permits and access to genetic collected material (in compliance to the Nagoya Protocol) were obtained from all countries and regions involved in the field campaign: France (Permit n° 15722 released on 22/05/2015 by the Centre de Recherches sur la Biologie des Populations d'Oiseaux CRBPO), Spain (Permit n° CEP 18/2014 released on 15/03/2014 by the Govern de les Illes Balears, Conselleria d'Agricultura, Medi Ambient I Territori).

Authors' contributions Davide Nespoli and Joseph Sunyer ringed and sampled birds; Davide Nespoli, Irene Pellegrino, Marco Galaverni,

Romolo Caniglia, Andrea Galimberti and Chiara Mengoni performed genetic and data analyses; Ettore Randi, Diego Rubolini, Fernando Spina, Gabriel Gargallo and Mattia Brambilla were involved in planning, organization, logistics and theoretical validation of the study. All authors collaborated to the manuscript drafting.

Funding Open access funding provided by Università degli Studi del Piemonte Orientale Amedeo Avogrado within the CRUI-CARE Agreement. This work was part of the Ph. D. of Davide Nespoli, funded by a bursary of University of Milan. Genetic analyses were funded by BIO-CGE ISPRA and "Ricerca locale" to Irene Pellegrino.

Declarations

Conflicts of interest Authors do not have conflicts of interest.

Ethics approval Ringing and sampling permits were obtained from all respective national and local authorities: France (Permit n° 15,722 released on 22/05/2015 by the Centre de Recherches sur la Biologie des Populations d'Oiseaux CRBPO), Spain (Permit n° CEP 18/2014 released on 15/03/2014 by the Govern de les Illes Balears, Conselleria d'Agricultura, Medi Ambient I Territori).

Availability of data and material All genetic sequences are available on Genbank NCBI. Supplementary data are available through the publisher's website.

References

Ancillotto L, Bosso L, Smeraldo S, Mori E, Mazza G, Herkt M, Galimberti A, Ramazzotti F, Russo D (2020) An African bat in Europe, Plecotus gaisleri Biogeographic and ecological insights from molecular taxonomy and Species Distribution Models. Ecol Evol. https://doi.org/10.1002/ece3.6317

Andersen LH, Sunde P, Pellegrino I, Loeschcke V, Pertoldi C (2017) Using population viability analysis, genomics, and habitat suitability to forecast future population patterns of Little Owl *Athene noctua* across Europe. Ecol Evol. https://doi.org/10.1002/ece3.3629

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Bernini G, Bellati A, Pellegrino I, Negri A, Ghia D, Fea G, Sacchi R, Nardi P, Fasola M, Galeotti P (2015) Complexity of biogeographic pattern in the endangered crayfish *Austropotamobius italicus* in northern Italy: molecular insights of conservation concern. Conserv Genet. https://doi.org/10.1007/s10592-015-0767-4

Brambilla M, Vitulano S, Spina F, Baccetti N, Gargallo G, Fabbri E, Guidali F, Randi E (2008) A molecular phylogeny of the *Sylvia cantillans* complex: cryptic species within the Mediterranean basin. Mol Phylogenet Evol. https://doi.org/10.1016/j.ympev.2008.05.013

Brichetti P, Fracasso G (2010) Ornitologia Italiana. In: Perdisa OA ed. Sylviidae-Paradoxornithidae, vol 6, p 493. Bologna

Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA. https://doi.org/10.1073/pnas.76.4.1967

Cai T, Cibois A, Alström P, Moyle RG, Kennedy JD, Shao S, Zhang R, Irestedt M, Ericson PGP, Gelang M, Qu Y, Lei F, Fjeldså J (2019) Near-complete phylogeny and taxonomic revision of the world's babblers (Aves: Passeriformes). Mol Phylogenet Evol. https://doi.org/10.1016/j.ympev.2018.10.010

Covas R, Blondel J (1998) Biogeography and history of the Mediterranean bird fauna. Ibis. https://doi.org/10.1111/j.1474-919X.1998.tb04600.x

Cracraft J (1983) Species concept and speciation analysis. In: Johnston RF (ed) Current ornithology, vol 1. Plenum Press, New York, pp 159–187

Cuttelod A, Garcìa N, Abdul Malak D, Temple H, Katariya V (2008) The Mediterranean: a biodiversity hotspot under threat. In: Vié J-C, Hilton-Taylor C, Stuart SN (eds) The 2008 review of the IUCN Red List of Threatened Species. IUCN Gland, Cambridge

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. https://doi.org/10.1038/nmeth.2109

Díaz-Rodríguez J, Gehara M, Márquez R, Vences M, Gonçalves H, Sequeira F, Martínez-Solano I, Tejedo M (2017) Integration of molecular, bioacoustical and morphological data reveals two new cryptic species of *Pelodytes* (Anura, Pelodytidae) from the Iberian Peninsula. Zootaxa. https://doi.org/10.11646/zootaxa.4243.1.1

Dickinson EC, Christidis L (eds) (2014) The Howard and Moore Complete Checklist of the Birds of the World, vol 2, 4th edn. Passerines. Eastbourne: Aves Press.

Doswald N, Willis SG, Collingham YC, Pain DJ, Green RE, Huntley B (2009) Potential impacts of climatic change on the breeding and non-breeding ranges and migration distance of European *Sylvia* warblers. J Biogeog. https://doi.org/10.1111/j.1365-2699.2009.02086.x

Drovetski SV, Zink RM, Fadeev IV, Nesterov EV, Kobli EA, Red'kin YA, Rohwer S, (2004) Mitochondrial phylogeny of *Locustella* and related genera. J Avian Biol. https://doi.org/10.1111/j.0908-8857.2004.03217.x

Dufresnes C, Mazepa G, Rodrigues N, Brelsford A, Litvinchuk SN, Sermier R, Lavanchy G, Betto-Colliard C, Blaser O, Borzée A, Cavoto E, Fabre G, Ghali K, Grossen C, Horn A, Leuenberger J, Phillips BC, Saunders PA, Savary R, Maddalena T, Stöck M, Dubey S, Canestrelli D, Jeffries DL (2018) Genomic evidence for cryptic speciation in tree frogs from the apennine peninsula, with description of *Hyla perrini* sp. nov. Front Ecol Evol. https://doi.org/10.3389/fevo.2018.00144

Escoriza D (2020) Organization of Squamata (Reptilia) assemblages in Mediterranean archipelagos. Ecol Evol. https://doi.org/10.1002/ece3.6013

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genetics 133(3):693–709

Galimberti A, Assandri G, Maggioni D, Ramazzotti F, Baroni D, Bazzi G, Chiandetti I, Corso A, Ferri V, Galuppi M, Ilahiane L, La Porta G, Laddaga L, Landi F, Mastropasqua F, Ramellini S, Santinelli R, Soldato G, Surdo S, Casiraghi M (2021) Italian odonates in the Pandora's box: a comprehensive DNA barcoding inventory shows

taxonomic warnings at the Holarctic scale. Mol Ecol Res. https://doi.org/10.1111/1755-0998.13235
Galimberti A, Spada M, Russo D, Mucedda M, Agnelli P, Crottini A, Ferri E, Martinoli A, Casiraghi M (2012) Integrated Operational Taxonomic Units (IOTUs) in echolocating bats: a bridge between molecular and traditional taxonomy. PLoS ONE. https://doi.org/10.1371/journal.pone.0040122
Ghorbani F, Aliabadian M, Zhang R, Irestedt M, Hao Y, Sundev G, Lei F, Ma M, Olsson U, Alström P (2020) Densely sampled phylogenetic analyses of the Lesser Short-toed Lark (*Alaudala rufescens*)—Sand Lark (*A. raytal*) species complex (Aves, Passeriformes) reveal cryptic diversity. Zool Scr. https://doi.org/10.1111/zsc.12422
Hackett SJ (1996) Molecular phylogenetics and biogeography of tanagers in the genus Ramphocelus (Aves). Mol Phylogenet Evol. https://doi.org/10.1006/mpev.1996.0032
Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acid S 41:95–98
Hebert PD, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. PLoS Biol. https://doi.org/10.1371/journal.pbio.0020312
Del Hoyo J, Collar NJ, Christie DA, Elliott A, Fishpool LDC, Boesman P, Kirwan GM (2016) HBW and Birdlife International Illustrated Checklist of the Birds of the World, vol 2: Passerines. Lynx Edicions and BirdLife International, Barcelona, Spain and Cambridge
Irestedt M, Fjeldså J, Johansson US, Ericson PGP (2002) Systematic relationships and biogeography of the tracheophone suboscines (Aves: Passeriformes). Mol Phylogenet Evol. https://doi.org/10.1016/S1055-7903(02)00034-9
Jordans A von (1914) Die vogelfauna Mallorcas mit berücksichtigung Menorcas und der Pityusen. Bonn.
Juste J, Ruedi M, Puechmaille SJ, Salicini I, Ibáñez C (2018) Two new cryptic bat species within the *Myotis nattereri* species complex (Vespertilionidae Chiroptera) from the Western Palaearctic. Acta Chiropterol. https://doi.org/10.3161/15081109ACC2018.20.2.001
Kangas V-M, Carrillo J, Debray P, Kvist L (2018) Bottlenecks, remoteness and admixture shape genetic variation in island populations of Atlantic and Mediterranean common kestrels *Falco tinnunculus*. J Avian Biol 49:e01768. https://doi.org/10.1111/jav.01768
Krawczyk E, Hedman H, Pafilis P, Bergen K, Foufopoulos J (2019) Effects of touristic development on Mediterranean island wildlife. Landscape Ecol. https://doi.org/10.1007/s10980-019-00917-5
Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0. Mol Biol Evol. https://doi.org/10.1093/molbev/msw054
Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. https://doi.org/10.1093/bioinformatics/btp187
Mason NA, Burns KJ, Tobias JA, Claramunt S, Seddon N, Derryberry EP (2017) Song evolution, speciation, and vocal learning in passerine birds. Evolution. https://doi.org/10.1111/evo.13159
Myers N, Mittermeier RA, Mittermeier CG, daFonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature. https://doi.org/10.1038/35002501
Ottenburghs J, Honka J, Müskens G, Ellegren H (2020) Recent introgression between Taiga Bean Goose and Tundra Bean Goose results in a largely homogeneous landscape of genetic differentiation. Heredity. https://doi.org/10.1038/s41437-020-0322-z
Padial JM, Miralles A, De la Riva I, Vences M (2010) The integrative future of taxonomy. Front Zool. https://doi.org/10.1186/1742-9994-7-16
Pellegrino I, Cucco M, Harvey JA, Liberatore F, Pavia M, Voelker G, Boano G (2017) So similar and yet so different: taxonomic status of Pallid Swift *Apus pallidus* and Common Swift *Apus apus*. Bird Study. https://doi.org/10.1080/00063657.2017.1359235
Pellegrino I, Negri A, Boano G, Cucco M, Kistensen TN, Pertoldi C, Randi E, Šálek M, Mucci N (2015) Evidence for strong genetic structure in European populations of the little owl *Athene noctua*. J Avian Biol. https://doi.org/10.1111/jav.00679
Perktas U, Barrowclough GF, Groth J (2011) Phylogeography and species limits in the Green Woodpecker complex (Aves: Picidae): multiple Pleistocene refugia and range expansion across Europe and the Near East. Biol J Linn Soc. https://doi.org/10.1111/j.1095-8312.2011.01750.x
Pertoldi C, Pellegrino I, Cucco M, Mucci N, Randi E, Terp Laursen J, Sunde P, Loeschcke V, Kristensen TN (2012) Genetic consequences of population decline in the Danish population of the little owl (*Athene noctua*). Evol Ecol Res 14:921–932
Pons JM, Thibault JC, Aymí R, Grussu M, Muntaner J, Olioso G, Sunyer JR, Touihri M, Fuchs J (2016) The role of western Mediterranean islands in the evolutionary diversification of the spotted flycatcher *Muscicapa striata*, a long-distance migratory passerine species. J Avian Biol. https://doi.org/10.1111/jav.00859
Rato C, Harris DJ, Carranza S, Machado L, Perera A (2016) The taxonomy of the *Tarentola mauritanica* species complex (Gekkota: Phyllodactylidae): Bayesian species delimitation supports six candidate species. Mol Phylogenet Evol. https://doi.org/10.1016/j.ympev.2015.09.008
Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 32: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. https://doi.org/10.1093/sysbio/sys029
Sangster G, Collinson JM, Crochet PA, Kirwan GM, Knox AG, Parkin DT, Votier SC (2016) Taxonomic recommendations for Western Palearctic birds: 11th Report. Ibis. https://doi.org/10.1111/ibi.12322
Schlick-Steiner BC, Steiner FM, Seifert B, Stauffer C, Christian E, Crozier RH (2010) Integrative taxonomy: a multisource approach to exploring biodiversity. Ann Rev Entomol. https://doi.org/10.1146/annurev-ento-112408-085432
Shirihai H, Gargallo G, Helbig AJ (2001) Sylvia Warblers. Helm, London
Shirihai H, Svensson L (2018) Handbook of Western Palearctic birds: Passerines. Helm, London
Steinbauer MJ, Severin DHI, Beierkuhnlein C (2013) Elevation-driven ecological isolation promotes diversification on Mediterranean islands. Acta Oecol. https://doi.org/10.1016/j.actao.2012.11.004
Taberlet P, Fumagalli L, Wüst-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. Mol Ecol. https://doi.org/10.1046/j.1365-294x.1998.00289.x
Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595
Voelker G, Light JE (2011) Palaeoclimatic events, dispersal and migratory losses along the Afro-European axis as drivers of biogeographic distribution in *Sylvia* warblers. BMC Evol Biol. https://doi.org/10.1186/1471-2148-11-163
Wauters L, Amori G, Aloise G, Gippoliti S, Agnelli P, Galimberti A, Casiraghi M, Preatoni D, Martinoli A (2017) New endemic mammal species for Europe: *Sciurus meridionalis* (Rodentia Sciuridae). Hystrix. https://doi.org/10.4404/hystrix-28.1-12015
Zuccon D, Pons JM, Boano G, Chiozzi G, Gamauf A, Mengoni C, Nespoli D, Olioso G, Pavia M, Pellegrino I, Raković M, Randi E, Idrissi HR, Touihri M, Unsöld M, Vitulano S, Brambilla M (2020) Type specimens matter: new insights on the systematics, taxonomy and nomenclature of the subalpine warbler (*Sylvia cantillans*) complex. Zool J Linn Soc. https://doi.org/10.1093/zoolinnean/zlz169