Wiley
Journal of Zoological Systematics and Evolutionary Research
Volume 2025, Article ID 7063146, 19 pages
https://doi.org/10.1155/jzs/7063146

WILEY

Research Article

Integrating Mitochondrial, Genomic and Species Distribution Model Approaches in the Reconstruction of the Sardinian Warbler (*Curruca melanocephala*) Phylogeography

Martina Nasuelli,[1,2] **Giovanni Boano**,[3] **Tamer Albayrak**,[4] **Marco Cucco**,[1,2] **Andrea Galimberti**,[2,5] **Luca Ilahiane**,[2,6] **Alexander N. G. Kirschel**,[7] **Flavio Mignone**,[8] **Michaella Moysi**,[7] **Marco Pavia**,[9] **Francesco Recco**,[8,10] **Gary Voelker**,[11] **and Irene Pellegrino**[1,2]

[1]*Department for Sustainable Development and Ecological Transition, University of Piemonte Orientale, Vercelli, Italy*
[2]*NBFC, National Biodiversity Future Center, Palermo, Italy*
[3]*Museo Civico di Storia Naturale di Carmagnola, Carmagnola, Italy*
[4]*Department of Biology, Dokuz Eylül University, İzmir, Türkiye*
[5]*Department of Biotechnology and Biosciences, University of Milano - Bicocca, Milano, Italy*
[6]*Department of Environmental Science and Policy, University of Milano - La Statale, Milano, Italy*
[7]*Department of Biological Sciences, University of Cyprus, Nicosia, Cyprus*
[8]*SmartSeq s.r.l, Devyser, Alessandria, Italy*
[9]*Department of Earth Science, University of Torino, Torino, Italy*
[10]*Department of Science and Technological Innovation, University of Piemonte Orientale, Alessandria, Italy*
[11]*Department of Ecology and Conservation Biology, Texas A&M University, College Station, Texas, USA*

Correspondence should be addressed to Martina Nasuelli; martina.nasuelli@uniupo.it

Received 5 September 2025; Revised 31 October 2025; Accepted 1 November 2025

Academic Editor: Savel Daniels

In the Western Palaearctic, ice ages played an important role in shaping the genetic diversity of birds, yet the phylogeography of species that persisted in the Mediterranean basin is understudied. Thus, we investigated the genetic diversity, phylogeography and demographic history of the Sardinian Warbler (*Curruca melanocephala*), a widespread Mediterranean songbird from the Canary Islands, North Africa, across southern Europe, to Turkey and the Middle East. Our study integrated mitochondrial (cytochrome C oxidase subunit I[(COI], cytochrome b [cytb]) and nuclear transforming growth factor-beta 2 (TGFB2) markers, single-nucleotide polymorphisms (SNPs) derived from a genotyping-by-sequencing (GBS) approach and species distribution models (SDMs). Both single markers and genomic SNPs identified four major clades. Several widespread haplotypes, extending from Iberia to Turkey and North Africa, indicate a panmictic pattern across the range of *C. m. melanocephala*. Among the remaining groups, one corresponds to *C. m. momus*, another includes individuals ascribable to subspecies*valverdei* and *leucogastra*, while the last consists of a highly divergent *leucogastra* haplotype. The ambiguity in subspecies range is further supported by nuclear markers and SNPs, which highlighted great levels of gene flow among populations. Phylogenetic reconstruction showed that the divergence of the clades within *C. melanocephala* occurred during the Pleistocene in the Middle East region. Demographic analyses suggest that the species maintained relatively stable effective population sizes (N_e) through time, findings supported by SDMs projections, which identify suitable habitats across the Mediterranean during the Last Glacial Maximum (LGM), implying broad persistence rather than contraction to isolated refugia. This study offers new insights into the historical processes that have shaped the apparent panmixia observed today in *C. melanocephala* and reveals that bird species originating in the Mediterranean may have partially escaped the constraints traditionally associated with glacial refugia along the region's coasts. The findings also emphasise the value of combining molecular and ecological approaches to disentangle the complex evolutionary dynamics of avian biodiversity hotspots.

Keywords: *Curruca melanocephala*; genetics; Last Glacial Maximum; Mediterranean; phylogeography; SNPs; species distribution models

1. Introduction

Phylogeographic studies allow researchers to infer the correlation between genetic data and the current geographic distribution of lineages within species [1]. Phylogeography has become an important tool to assess the evolutionary background of species, with implications for the understanding of the speciation process, taxonomic revisions, as well as assessment of biodiversity and conservation priorities [2, 3]. In the Western Palaearctic, Plio-Pleistocene glaciations and glacial refugia significantly shaped the genetic structure of animals and plants, with species moving southward and becoming isolated during ice ages and following recolonisation routes northward in interglacial periods [4–7]. In the Mediterranean basin, the different glacial refugia, such as the Iberian, Italian, Balkan peninsulas and Anatolia, along with their associated islands, and the geographical barrier posed by the sea itself, have played a key role in shaping the variability of bird populations at different levels (e.g., genetic, morphological and acoustic), hosting isolated populations with unique local adaptations that have persisted over time and are still often reflected in the current patterns of genetic diversity [8].

Despite numerous phylogeographical studies on Western Palaearctic birds [9], the knowledge regarding the phylogeography of species distributed in the Mediterranean basin is scant [10, 11]. In this biogeographical context, a few species have evolved in open Mediterranean maquis shrubland, but many of these are endemic (occurring in multiple habitats) to the basin, suggesting that the genetic isolation led to many speciation events in the avifauna of the region [10].

The genus *Curruca*, attributed to the family Sylviidae, represents a fine example of species inhabiting Mediterranean niches. This genus encompasses 25 species, once belonging to the genera *Sylvia* and *Parisoma*, associated with Mediterranean maquis vegetation and treed savannas in Africa [12, 13]. The ancestral origin of the Sylviidae family, including *Curruca*, is the Sino-Himalayan mountains, and the subsequent expansions towards the Mediterranean region might have followed the forested habitat distributed in Western and Central Asia [12, 14]. In support of this, Voelker and Light [15] found that the migratory behaviour was the inherited state for the *Sylvia* and *Curruca* clades, while populations developed the sedentary habit independently several times in more recent ages.

We focused this study on the Mediterranean coastal specialist Sardinian Warbler *Curruca melanocephala* to understand whether this species would show a phylogeographic pattern consistent with the glacial refugia model, as has been found in other species such as *Athene noctua* [16], *Acrocephalus scirpaceus* [17] and *Picus viridis* [18, 19].

The Sardinian Warbler is distributed from the Canary Islands, North Africa and the Iberian Peninsula in the west, across Mediterranean Europe, to the Middle East and Turkey [20]. Populations are predominantly resident, especially those on islands and coasts, occupying small territories [21], a trait that can favour geographical genetic speciation [22]. However, most of the migratory behaviour of this species is displayed by inland populations, especially individuals in the northern and easternmost part of the distribution range [23]. These populations perform long-distance migration southward in winter, reaching non-breeding areas extending from southern Europe to northern sub-Saharan Africa. The winter distribution is discontinuous, with key areas in western Mauritania, northern Senegal, Saharan oases in Niger and Sudan and across North Africa [20, 24]. The presence of vagrant individuals has been documented [25] and is further supported by the historical ringing data series for the species [26]. In addition, other short-range and elevational migrations, besides offspring dispersion, are commonly reported for this species [23]. The most recent assessment of the conservation status of the Sardinian Warbler has classified the species as least concern (LC) [27], consistent with previously documented expansions of its northernmost populations, particularly along the Atlantic coasts of Spain, southwestern France, and further into the Balkan region [28], as well as with the colonisation of new areas, such as Cyprus [29], where the species is now considered a common breeder [30]. Based on distribution and morphology, four extant and one extinct subspecies are currently accepted [31]: the nominate *C. m. melanocephala* distributed across most of the species' range, *C. m. leucogastra* in the Canary Islands, *C. m. momus* in Lebanon, Israel and Syria, and *C. m. valverdei* in Western Sahara and Morocco [20]. The subspecies *C. m. norrisae* is considered extinct, a probable consequence of the salinisation and vegetation degradation in Egypt [32].

The morphological variation among the subspecies shows a general tendency to become slightly smaller and paler in the southern and eastern parts of the distribution range. However, large morphological variation is found within each subspecies and a great overlap in characters: *C. m. leucogastra* is paler in the Eastern islands and darker in the Western islands, while *C. m. valverdei* is similar to the nominal subspecies but has paler upperparts. These colour variations are believed to result from specific ecological adaptations to habitat conditions and repeated recolonisation events from the mainland [23, 33, 34]. With regard to vocalisations, comparative studies between subspecies are lacking; only one structural difference in contact calls has been identified between the nominate subspecies and *leucogastra* [35].

From a molecular perspective, Blondel et al. [36] used DNA/DNA hybridisation techniques to investigate the historical biogeography of Mediterranean Warblers. They inferred that *C. melanocephala* belongs to a central-Mediterranean clade with *C. melanothorax*, C. *cantillans*, *C. rueppelli*, *C. mystacea* and *C. conspicillata*, closely related to the western-Mediterranean clade (*C. sarda*, *C. deserticola*, *C. undata*), with divergence from one another between 3.1 and 3.4 million years ago (mya). This hypothesis was not supported by the mitochondrial gene analysis of Voelker and Light [15], who developed an alternative phylogeny for *Sylvia* and *Curruca*, and estimated much older lineage divergences (i.e., 12.5–15 mya) as compared to Blondel et al. [36]. The Voelker and Light [15] results were confirmed by Cai et al. [12], who further assessed the Sylviidae family phylogeny with a reconstruction based on several mitochondrial and nuclear markers.

Focusing on Sardinian Warbler insular populations (Canary Islands), Dietzen et al. [34] found a low degree of genetic diversification between insular and mainland populations, contrasting with a significant diversity in plumage colouration; they also rejected the *C. m. leucogastra* subspecies. Similarly, Moysi et al. [37] discovered a low degree of differentiation between the Greece and Crete *C. melanocephala* populations, with the latter showing unexpectedly high genetic variability, suggesting a possible role of the Mediterranean islands as suitable refugia during glacial periods.

A tool that can help explain intricate phylogenies is species distribution models (SDMs), which have been used to assess the geographic distribution of species under current bioclimatic conditions. This approach allows present-day suitable environments (geographic areas) to be compared with past bioclimatic conditions, such as the Last Glacial Maximum (LGM), so that alternative distributional areas and recolonisation routes can be identified [11, 38]. In this study, we conducted the most extensive genetic sampling to date for the Sardinian Warbler, and combined data from multiple molecular approaches (i.e., single molecular markers and genome-wide single-nucleotide polymorphisms [SNPs]) with SDM to: (i) assess the role of glacial refugia and phylogeographical history in shaping the genetic variability of *C. melanocephala* across its breeding range, and to evaluate possible taxonomic discrepancy in relation to the currently recognised four subspecies; and (ii) determine the influence that the current and LGM climatic conditions have had on the present-day genetic population structure.

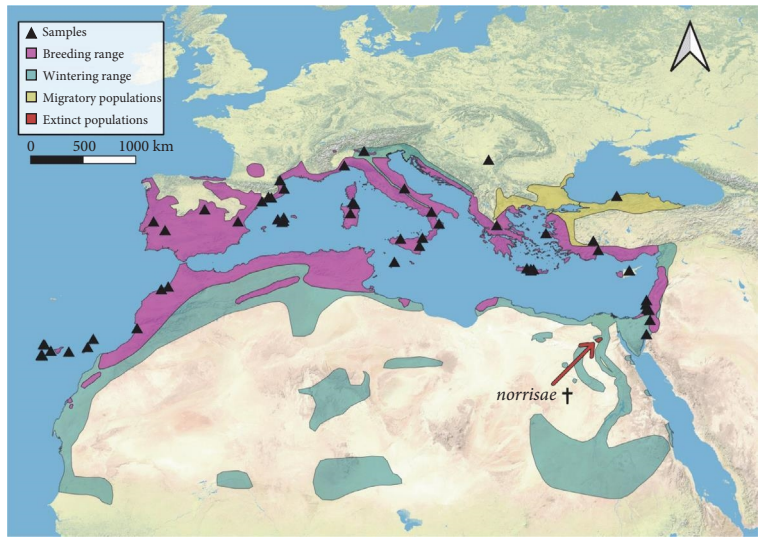
2. Materials and Methods

2.1. Samples Collection and DNA Extraction. We collected 95 samples of *C. melanocephala*, covering most of the breeding range (Figure 1a, a detailed list of the samples analysed in this study are provided in Supporting Information 1: Table S1) and including samples from regions not previously represented or that were underrepresented in other phylogeographic studies (Italy, Turkey, Serbia, Israel, Cyprus, mainland Spain, Balearic Islands, Greece). In particular, we obtained DNA extracts ($N = 5$) and fresh tissue samples ($N = 12$) from museum collections and other institutions, as well as blood samples ($N = 60$) and feathers ($N = 18$) collected by the authors or retrieved from museum specimens (Supporting Information 1: Table S1). All tissue and blood samples were preserved in 96%–100% ethanol and stored at −20°C shortly after collection. The sampling also included the distributional areas of all recognised subspecies currently attributed to the Sardinian Warbler [31]. Furthermore, we included in our bioinformatic analyses 35 cytochrome b (cytb) haplotypes available in GenBank (https://www.ncbi.nlm.nih.gov/genbank/) and attributed to *Curruca* (ex *Sylvia*) *melanocephala*, along with three cytochrome C oxidase subunit I (COI) haplotypes from the BOLD system [40] (Supporting Information 1: Table S1). Total genomic DNA was isolated with the NucleoSpin Tissue kit (Macherey-Nagel, Düren, Germany), adapting the manufacturer's protocol specifically to the starting material. For feathers and tissue samples, an overnight pre-lysis at 56°C was carried out, while blood samples preserved in ethanol were dried and later washed with saline solution and directly digested at 70°C. The subsequent steps were conducted as described by the employed DNA isolation kit. The eluted nucleic acids were stored at −20°C immediately after the extraction protocol until further wet lab protocols were performed. Moreover, considering the phylogenetic relatedness of the investigated species, we included one sample (i.e., CM011) of Cyprus Warbler *Curruca melanothorax* in our analyses, in addition to other outgroup species sequences from available online databases.

2.2. Amplification and Sequencing of Mitochondrial and Nuclear Markers. To investigate the phylogeographic structure of the species, we selected different molecular markers: cytb, COI and the nuclear transforming growth factor-beta 2 (TGFB2) gene. To obtain the nucleotide sequences, we first performed a polymerase chain reaction (PCR) using specific published primers to target the region of interest. Thermocycling conditions were set as follow: initial denaturation at 95°C for 2 min, denaturation at 95°C for 30 s, annealing for 30 s and at different temperature based on the primer pair used: 48°C for the COI primer pair BirdF1 and BirdR2 [41], 53°C for the cytb primer pair L14850 and MtFs-H [42], and 58°C for the nuclear gene pair TGFB2.5F and TGFB2.6R [43]. PCR amplification was performed in a volume of 25 μL, including 10–20 ng of total DNA, 5 μL of 5x GoTaq Flexi buffer (Promega, Italy), 0.3 μL of each primer (10 μM), 2.5 μL of $MgCl_2$ (25 mM), 0.4 μL of dNTPs mix (10 mM), 0.14 μL of bovine serum albumine (BSA - 20 mg/mL, Thermo Fisher Scientific, Waltham, Massachusetts, USA) and finally 0.15 μL of GoTaq Flexi DNA Polymerase (5 u/μL). Reactions were checked through electrophoresis on a 1.8% agarose gel, including negative controls. PCR products were subsequently purified using ExoSAP-ITA PCR Product Cleanup Reagent (Thermo Fisher Scientific, Waltham, Massachusetts, USA), following the manufacturer's protocol. The purified amplicons were sequenced on an ABI 3730xl DNA Analyser by Macrogen Europe Inc. (Amsterdam, The Netherlands), using the same PCR primers.

2.3. Haplotype Inference. The obtained electropherograms were visualised with FinchTV 1.4 (https://finchtv.software.informer.com/1.4/). For nuclear markers, any heterozygous site was marked with an IUPAC code mixed base. Fasta-formatted files were then exported, divided by marker and merged into single datasets. All generated datasets were aligned with MAFFT 7.52 [44] implemented in Bioconda [45]. The nucleotide sequences were submitted to GenBank and to Zenodo, as reported in the Data Availability section. Haplotype inference was performed on each mitochondrial dataset using FaBox 1.6 [46], while for the nuclear sequences, we first resolved the TGFB2 alleles with PHASE 2.1 [47] available in DNASP 5.10 [48], using default settings and a probability threshold of 0.9.

The mtDNA sequences were combined to create concatenated sequences with the package *concatipede* [49] implemented in R [50] and RStudio (Posit [51]), excluding the nuclear sequences as they often return a conflicting or



(a)

	1	2	3	4
1	/	± 0.0016	± 0.0026	± 0.0039
2	**0.0054**	/	± 0.0025	± 0.0037
3	**0.0117**	**0.0123**	/	± 0.0036
4	**0.0217**	**0.0206**	**0.0213**	/

(c)

(b)

FIGURE 1: Localities, phylogenetic tree and genetic distances. (a) Sampling locations and distribution ranges of *C. melanocephala*. Each triangle indicates an individual sample. The species' known wintering (green), year-round (pink), and breeding (yellow) ranges follow BirdLife International [39] and Birds of the World [23], (b) Bayesian phylogenetic tree of the COI–cytb mtDNA dataset with estimated molecular dating for the major node splits. Red dots highlight nodes with a posterior probability >0.9. Tree branches are coloured to show the major haplogroups' division resulted according to the phylogenetic reconstruction, and (c) *p*-distances and standard deviation (above the diagonal) among the identified haplogroups.

weaker phylogenetic signal when analysed with mitochondrial genes [52].

2.4. Phylogenetic Reconstruction and Molecular Dating. We reconstructed the phylogenetic relationship between the populations by inferring the best substitution model for the single markers and for concatenated mitochondrial COI–cytb gene alignments. Available haplotypes (COI = 3, cytb = 35) from different online databases were included in the datasets. We used PartitionFinder 2.1 [53] to estimate the substitution model under the corrected Akaike information criterion (AICc), setting the partition divided by molecular marker and by codon position. The best-fitting models inferred were: GTR for COI_pos1, GTR + I for COI_pos2, TRN + I for COI_pos3 and cytb_pos3, GTR + G for cytb_pos1, and HKY + G for cytb_pos2 and TRN + X for TGFB2_pos1 and TGFB2_pos2 and K80 for TGFB2_pos3.

Using these partitions, we constructed a Bayesian phylogenetic tree using BEAST 1.10 [54], setting a strict molecular clock and a Yule speciation process as tree prior. The MCMC chains were run for 1×10^7 generations and were sampled every 1000 generations. Goodness of fit (ESS of the parameters > 200) was assessed in Tracer 1.7 [55]. We created a consensus tree using TreeAnnotator 1.10 [54], setting a Maximum Clade Credibility Tree as target and discarding 10% of the samples as burn-in. The consensus tree topology was visualised in FigTree 1.4 (http://tree.bio.ed.ac.uk/software/figtree) and plotted in RStudio using the *treeio* [56] and *ggtree* [57] packages.

Divergence times for the main haplogroup nodes were calculated using the locus-specific substitution rates for COI and cytb as estimated in Lerner et al. [58], 0.016 and 0.014, respectively, with a standard deviation of 0.001 for both. Following Nasuelli et al. [59], three different runs were performed, with set Bayesian parameters consisting of 1×10^8 generations, sampled every 10,000, under a strict clock model with a normal distribution and a coalescent tree prior. After the divergence estimation, we discarded 10% of the sampled trees and created a consensus tree with the remaining trees using TreeAnnotator 1.10.

The uncorrected pairwise genetic distances (*p*-distance) between and within haplogroups were estimated using MEGA 11 [60] on both the COI–cytb concatenated dataset and the cytb alignment, the latter integrated with cytb haplotypes recovered from previous studies [15, 34, 61–63].

Finally, median-joining haplotype networks were generated for each mitochondrial marker and for the concatenated dataset (COI–cytb) using PopART 1.7 [64], in order to highlight the geographic distribution of the haplotypes and their relative frequencies. In addition, we used the nuclear TGFB2 sequences to create a median-joining haploweb using HaplowebMaker [65], connecting heterozygous individuals with co-occurring alleles.

2.5. Genetic Indexes, Demographic History and Ancestral State Reconstruction. In order to shed light on the current genetic structure of breeding Sardinian Warblers, we estimated the number of haplotypes (h), number of polymorphic sites (S), percentage of $G + C$ content, nucleotide diversity (π) and haplotype diversity (H_d) for each mitochondrial and nuclear dataset using DNASP 5.10. We performed Tajima's D and Fu's F_s neutrality test with Arlequin 3.5 [66] to determine possible population expansions. We conducted 1000 simulated samples and set the threshold filter to 0.9.

To correlate the current geographical distribution of phylogenetic lineages with biogeographic events that might have had an important influence on past distributions, we inferred the possible paths of colonisation through the Western Palaearctic and verified if extant populations are the results of past events such as dispersal or vicariance. We used RASP v4.4 [67] to perform the ancestral area reconstruction. The phylogenetic analysis was conducted using BEAST, with the same overall parameters as those used for the concatenated dataset, but applying the best-fitting substitution models specifically selected for the cytb codon positions (i.e., TRN + G + X for cytb_pos1, TRN + I + G + X for cytb_pos2, and HKY + I + X for cytb_pos3). We subsequently implemented the Bayesian binary MCMC (BBM) method in RASP, with 1×10^7 generations, sampled every 1000 generations, under a fixed Jukes-Cantor model and analysed after a 10% burn-in. Four different geographical regions were considered, based on the distribution of individuals: (A) Canary Islands, (B) Morocco, (C) Mediterranean Europe and (D) Israel. Following the findings in Voelker and Light [15], we conducted the RASP analysis on both the complete cytb dataset and a dataset excluding the most divergent haplotypes derived from individuals on Gran Canaria Island.

To infer possible fluctuation in the effective population size (N_e) that might have occurred during the evolutionary time of the species, we calculated a Bayesian skyline plot (BSP) with BEAST 1.10 using the concatenated mtDNA dataset. We set a coalescent Bayesian skyline as tree prior, a strict clock model with a normal distribution prior, 5×10^7 as the chain length, and sampled every 50,000 generations. After the analysis, the output was imported into Tracer 1.7, setting the bin to 100, corresponding to a 10% burning.

2.6. Genotyping-by-Sequencing (GBS) Library Preparation and NGS Sequencing. We selected a subsample of individuals choosing representatives of different haplogroups and countries as identified in the mitochondrial phylogenetic analyses, from the available pool of *C. melanocephala* samples (Supporting Information 1: Table S1). Such sampling allowed us to assess genomic structure and potential gene flow among lineages more comprehensively. We applied a GBS analysis following the protocol of Elshire et al. [68] with some modifications developed in collaboration with SmartSeq s.r.l. and described in previous studies [69, 70]. Briefly, depending on its concentration, 100–200 ng of DNA from each sample was digested with PstI-HD restriction enzyme (New England Biolabs), and two different types of adapters were ligated with T4 ligase (New England Biolabs). Using the NucleoMag NGS kit (Macherey-Nagel, Düren, Germany), we selected the size of the fragment and pooled normalised samples in libraries, and each library was then denatured and amplified adding indexes and adapters for the Illumina sequencing. PCR products were pooled into a single library, which was

then purified, size-selected (200–800 bp) with the Nucleo-Mag NGS kit and the concentration was measured by fluorescence (Qubit 4, Thermo Fisher Scientific). Finally, the library was sequenced on an Illumina MiSeq System (Illumina, San Diego, CA, USA) with a MiSeq Reagent Kit V3 (600 cycles).

2.7. Reads Filtering and SNPs Data Analyses. Raw sequences were demultiplexed using an in-house pipeline based on sample-specific tags. Adapter removal was applied to the reads. Reads were aligned to the reference genome of the closest available relative species in NCBI (*Sylvia borin*, GCA_014839755.1_bSylBor1.pri, https://www.ncbi.nlm.nih.gov/) using the bwa aln function with the default parameters of the BWA software [71]. Alignment files were processed with SAMtools [72] for conversion and sorting. SNP calling was performed with bcftools [72], extracting variant sites only from regions covered by at least 10 reads. Variants were further filtered to retain only those covered in at least two-thirds (~66%) of samples. Variant allele frequencies (VAFs) thresholds were applied as follows: mutations with $\leq 30\%$ VAF were considered wild type, variants with VAF between 30% and 60% were considered as heterozygous (using IUPAC notation), variants with $\geq 60\%$ VAF were considered homozygous mutant.

The files containing the demultiplexed and aligned reads of the individual samples were imported into STACKS 2.64 [73], which included all SNPs called in comparison to the reference genome. We used the provided *ref_map.pl* script, with default settings and the *population* function to generate and export a single VCF file including all SNPs of the complete dataset.

We subsequently estimated the mean coverage per site and per individual with VCFtools [74] and filtered the data using PLINK 1.9 [75]. The parameters chosen for data filtering were mac = 3 and mind = 0.9. After the first data filtering, the retained samples and related SNPs were tested for linkage disequilibrium (LD) using the PLINK function *indep-pairwise*, by setting 50 as the test's window size, 5 as the shift frame and 0.5 as the variance inflation factor (VIF).

Based on these data, we reconstructed the population genomic structure with a principal component analysis (PCA), calculating the eigenvalues and eigenvectors using the function *pca* in PLINK, extracting the 20 best principal components for the variation. The individuals considered in this inference were divided and analysed based on single-sample locality. The plot was generated in RStudio using the *ggplot2* package [76].

We also evaluated the percentage of shared ancestry between the individuals coming from the different geographical areas included in this study. We used ADMIXTURE [77], setting the K values from 2 to 4 and performing 10 runs with random seeds for each putative group. The variation in the cross-validation error rate was estimated to infer the best K value. The ADMIXTURE analysis was performed on groups of samples, each created based on what was found in the previous analysis and divided into higher geographical ranking (i.e., Canary Islands, Morocco, Mediterranean Europe, Israel). Finally, in order to assess the current degree of gene flow between the populations in this study, we estimated the migration patterns with divMigrate-online [78], using Alcala's estimation parameter [79] and setting 500 bootstrap replicates and a filter threshold of 0.15.

Pairwise population summary statistics (such as Fst and Nm) and within-population statistics (H_o, H_e, and F_{is}) were calculated for each group and population using Genetix [80].

2.8. SDMs. Occurrence data were obtained from the Global Biodiversity Information Facility (GBIF.org). The raw data consisted of 488,423 records (GBIF Occurrence download: https://doi.org/10.15468/dl.accw63). The data cleaning procedure included removing the records without coordinates or with coordinate uncertainty > 1000 m ([81]; N = 718), observations not made in the main breeding period (outside the period encompassing April–July: N = 306,286), and incidental data from outside the breeding range delimited by the EBBA2 atlas ([28]; N = 639). To avoid overfitting of SDMs, we removed duplicate records by keeping only a single occurrence in each occupied cell of a grid with the same resolution (30 arcsec, ~1 km) and extent of the Bioclim environmental predictors.

The cleaning procedure led to a final dataset of 14,713 occurrences. To predict the current species distribution, we used worldwide climate surfaces representing 19 temperature-precipitation related bioclimatic variables available from WorldClim (version 2, http://www.worldclim.org; variable descriptions in [82] and [83]). To examine past environmental conditions in the study area, we utilised the data provided by WorldClim for the refugium-driving LGM cold scenario (21,000 years ago). We used both climate surfaces representing past conditions according to the 'Community Climate System Model' (CCSM) and the 'Model for Interdisciplinary Research on Climate' (MIROC), version 3.2, downloaded from worldclim.org [84, 85]. Using the *raster* R package [86], current and past climate surfaces were cropped to the south European–North African geographical boundaries (spatial extent: longitude 19°W–45°E, latitude 0°N–50°N).

SDMs were developed using the R package *sdm* [87]. The process included (i) reducing the number of climatic variables to avoid collinearity problems, (ii) generating pseudo-absence data, (iii) estimating current times SDMs using different modelling techniques, (iv) calculating an ensemble model from the previous models and (v) utilising the current model conditions to estimate areas of possible occurrence during the past LGM. To reduce the number of external variables and select only uncorrelated ones (highly correlated variables are prone to inflation errors), we utilised a VIF approach [81]. To select variables from the 19 input variables that could have a collinearity problem, we excluded the VIF values higher than the threshold value (VIF > 10, Supporting Information 3: Table S5) suggested in Naimi and Araújo [87]; also [88].

The resulting set comprised nine variables (Supporting Information 3: Table S5) including annual mean diurnal temperature range (BIO2), isothermality (BIO3), mean temperature of the wettest quarter (BIO8), temperature in summer (BIO9),

precipitation of wettest month (BIO13), precipitation of driest month (BIO14), precipitation seasonality (BIO15), precipitation in the warmest quarter (BIO18) and precipitation of coldest quarter (BIO19). We utilised the built-in function in *sdm* to generate 14,713 random pseudo-absence points. The number of pseudo-absence data was equal to the number of true presence data, following the recommendation of Barbet-Massin et al. [89].

We utilised several statistical techniques to model Sardinian Warbler distribution, as prediction differences between different techniques can be large [90–92]. We first employed all modelling algorithms available in the package *sdm*: bioclim, boosted regression trees (BRTs), classification and regression trees (CARTs), flexible discriminant analysis (FDA), generalised additive models (GAMs), generalised linear models (GLMs), multivariate adaptive regression spline (MARS), maximum entropy models (MaxEnt), mixture discriminant analysis (MDA), MLP and RBF artificial network analysis (ANN), random forests (RFs), recursive partitioning (Rpart), support-vector machine (SVM), detailed in [87]. Each model was replicated 3 times using 70% of the data as the training dataset and the remaining 30% as the testing dataset [87]. Predictive performances of SDMs were assessed by measuring the true skill statistics (TSSs; [93]) and the area under the receiver operating characteristic curve (AUC). We then excluded models with TSS < 0.89 and AUC < 0.97, and the models retained (RF, SVM and MLP) were utilised to build an ensemble model by applying a TSS-weighted average [93]. The current distribution was estimated first, and the result was used to project the Sardinian Warbler model on past climate layers using two distinct LGM models, the CCSM and the MIROC. The resulting models were then utilised to build two ensemble models, both obtained by applying TSS-weighted averages.

3. Results

3.1. Genetic Diversity and Phylogenetic Reconstruction. We obtained a total of 25 haplotype sequences of 619 bp for COI from 94 individuals and 52 haplotype sequences of 898 bp for cytb from 82 individuals. Compared with the haplotypes already in the online databases, we identified 23 new haplotypes for cytb and 18 for COI. The TGFB2 nuDNA dataset after the phasing process returned 9 different alleles of 531 bp, differentiated at three variable sites. All information on samples and haplotypes is reported in Supporting Information 1: Table S1, including GenBank accession numbers of the generated nucleotide sequences.

The Bayesian tree reconstruction based on the concatenated COI–cytb dataset of 1517 bp highlighted the presence of four well-supported haplogroups (posterior probability ~1) (Figure 1b). The largest group in terms of number of haplotypes, hereafter referred to as haplogroup 1, includes individuals from Portugal, Spain and Balearic Islands, France, Italian peninsula including Sardinia and Sicily Islands, Greece, Crete and Cyprus Islands, Serbia and Turkey, along with several individuals from the Canary Islands and Morocco. Based on the sample's information and distribution, this haplogroup consists mostly of individuals belonging to the nominate subspecies *C. m. melanocephala*, together with samples ascribed, based on morphology, to *C. m. leucogastra*. The attribution of Moroccan individuals to *C. m. valverdei* should be excluded since no other information (e.g., morphology), besides sample locality, is known. Conversely, the reconstructed haplogroup 2 includes individuals from Morocco and from the Canary Island of El Hierro, the latter ascribed to *C. m. leucogastra*. The third clade, haplogroup 3, shows a clear separation of the individuals sampled in Israel, where the *C. m. momus* subspecies is described from. The last clade, haplogroup 4, includes one distant haplotype sampled in the Canary Islands (Gran Canaria), and described as *C. m. leucogastra*. In contrast, the tree calculated from TGFB2 haplotypes returned two main haplogroups, which lacked geographical structure, except for a subclade formed by two alleles from El Hierro Island and a second subclade comprising two alleles from Israel (Supporting Information 2: Figure S2).

The divergence between haplogroups 1 and 2 is estimated at 0.3115 mya (height 95% HPD 0.214–0.4476). The Israeli haplogroup 3 diverged from these European/Moroccan clades 0.5599 mya (height 95% HPD 0.3501–0.7821), while the most basal divergence separating haplogroup 4 from the other haplogroups is estimated at 0.9769 mya (height 95% HPD 0.691–1.3713). The divergence time estimated for the *C. melanocephala* clade is 5.7672 mya (height 95% HPD 4.2519–7.9948, Figure 1b).

The uncorrected pairwise genetic distances (*p*-distance) calculated between the identified haplogroups of the concatenated dataset ranged from 0.0054 between haplogroups 1 and 2 to 0.0217 between haplogroups 1 and 4 (Figure 1c).

The genetic distances within each haplogroup (excluding haplogroup 4, which includes only one sequence) showed the highest value for haplogroup 2 (Morocco and Canary Islands) (0.0023 ± 0.0008), followed by haplogroup 3 (Israel and Cyprus) (0.0022 ± 0.0007) and haplogroup 1 (widespread in Europe, Morocco and Canary Islands) (0.0014 ± 0.0004).

The intergroup *p*-distances calculated on the mtDNA cytb dataset only, integrated with GenBank sequences, supported the results obtained from the concatenated dataset: the distance between the haplogroups ranges from 0.00690 (± 0.0023, haplogroups 1 vs. 2) to 0.02217 (± 0.00402, haplogroups 2 vs. 4), while the highest *p*-distance values were reached between Gran Canaria and the others islands (from 0.01503 ± 0.0031 to 0.01762 ± 0.0032).

The median-joining haplotype network constructed on each mitochondrial marker and COI–cytb concatenated haplotypes showed a similar genetic pattern as phylogenetic trees identifying four haplogroups (Figure 2 and Supporting Information 2: Figure S1A,B). In particular, the network calculated on the concatenated dataset highlighted that haplogroup 3 (Israel and Cyprus) differs by nine mutations from the larger, widely distributed haplogroup 1, in which Cmel1, Cmel8 and Cmel19 were the most represented haplotypes across individuals. Haplogroup 2 (Morocco and El Hierro

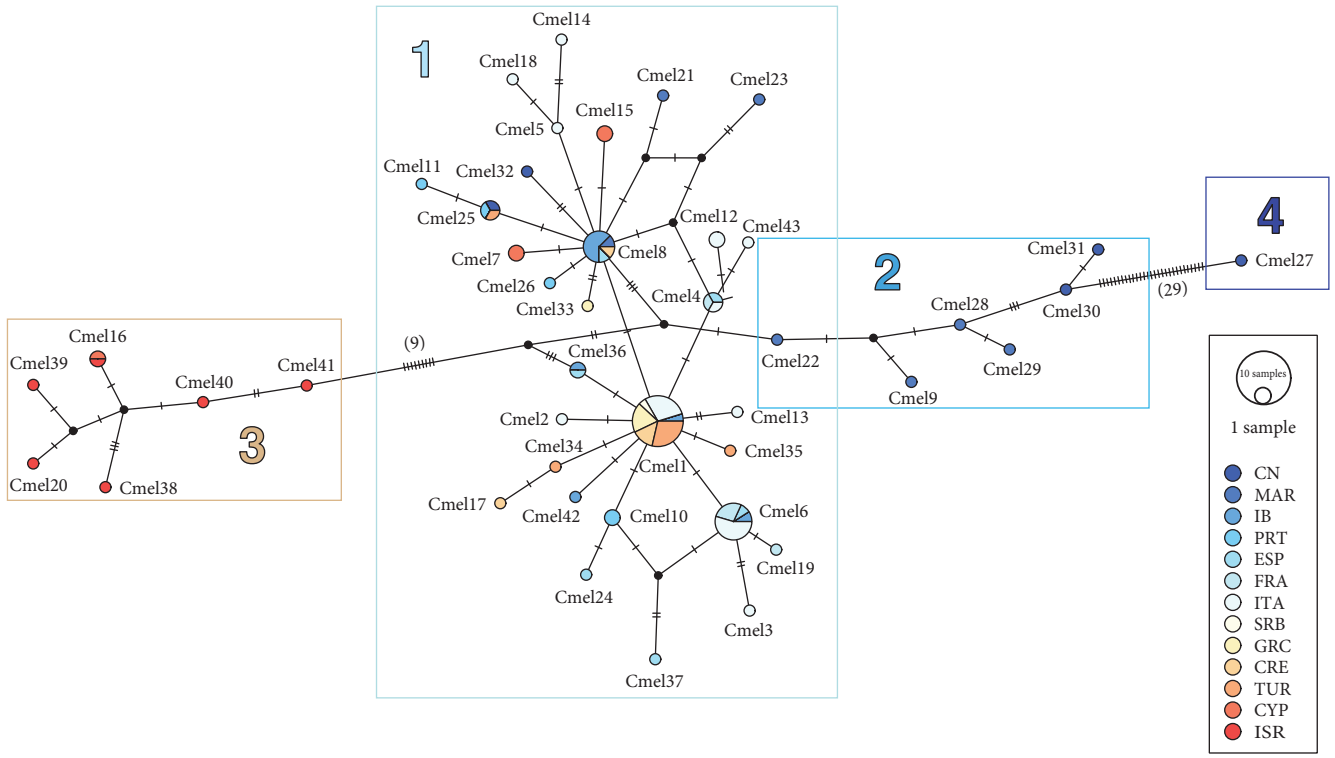


FIGURE 2: Median-Joining haplotypes network of the COI–cytb dataset. Haplotypes inference started from 93 samples, subsequently collapsed into 43 haplotypes and coloured based on sample localities. The number of mutations is reported in brackets, and the haplogroups are pointed up by the ellipses.

TABLE 1: Summary of genetic diversity indices inferred from the concatenated COI–cytb mtDNA dataset and from the mitochondrial and nuclear loci separated, considering all the retrieved sequences.

Marker	*N*	*h*	*S*	*G* + *C*	*Π* (± s.d.)	H_d (± s.d.)	Tajima's *D* (*p*-value)	Fu's F_s (*p*-value)
Concatenated COI–cytb	93	43	73	0.495	0.00375 (± 0.0006)	0.917 (± 0.021)	−1.9771*	−25.0646**
COI	97	25	26	0.503	0.0035 (± 0.0006)	0.615 (± 0.059)	−1.7172	−16.7795**
cytb	121	52	64	0.488	0.00548 (± 0.0008)	0.920 (± 0.016)	−1.8603*	−25.4590**
TGFB2	108	9	7	0.442	0.00281 (± 0.0002)	0.703 (± 0.036)	0.2811	−1.1870

*$p < 0.05$.
**$p < 0.01$.

Island) differs by four mutations from the largest haplogroup 1. The most genetically distant haplogroup 4 (including Cmel27) differs by 29 mutations from haplogroup 2 and consists of only one haplotype from Gran Canaria.

A haploweb network calculated for the nuclear marker TGFB2 indicated that many phased alleles are shared across haplogroups (Supporting Information 2: Figure S1C). The most common allele is CmelTGFB2_1, shared by samples from all sampling localities, followed by CmelTGFB2_2 and CmelTGFB2_3. The least common alleles were two alleles from the Canary Islands (El Hierro, CmelTGFB2_7 and CmelTGFB2_8) belonging to the same heterozygous individual, one allele from Israel (CmelTGFB2_9) and one allele from Italy (CmelTGFB2_5). The haplotypes from El Hierro and Israel clustered separately in the phylogenetic tree calculated in BEAST, as did the exclusively Italian allele, found in an individual from Linosa, an island located about equidistant between Sicily and Tunisia (Supporting Information 2: Figure S2).

3.2. Genetic Indices, Demographic History and Ancestral State Reconstruction. The mtDNA diversity estimated on the total number of samples showed a haplotype diversity of 0.920 (± 0.016) for cytb and 0.615 (± 0.059) for COI (Table 1). The highest number of haplotypes (*h*) was found for cytb (52), while COI returned 25 haplotypes. Significant negative values of Tajima's *D* and Fu's F_s were estimated for the cytb dataset, while only a significant negative value for the Fu's F_s index was estimated from the COI dataset. The TGFB2 dataset of nine different alleles had an estimated haplotype diversity of 0.703 (± 0.036). No significant values of Tajima's *D* or Fu's F_s were calculated for this gene (Supporting Information 3: Table S2).

A BSP based on the concatenated dataset exhibited some fluctuations with an initial decline followed by a slight increase of N_e (Figure 3b).

The BBM analysis calculated on the complete cytb dataset identified the Canary Islands (Gran Canaria) as the possible centre of origin of *C. melanocephala* (posterior probability =

99%). Gran Canaria (consistent with haplogroup 4) is also identified as the most likely ancestral area at all divergences between haplotype groups, with the Saharan regions identified as a much less probable ancestral area (posterior probability = 12%) (Figure 3a). Although the ancestral state reconstruction based on the complete cytb dataset returned limited support for this latter scenario, the analysis, excluding the divergent haplotypes from Gran Canaria, indicated that the most likely origin of the ancestral colonisation of the Mediterranean basin was from the Saharo-Arabian regions (posterior probability = 71%). This supports the lower probability previously found in the same node that resulted in the division of the *C. momus* clade from the other haplogroups, as a consequence of both vicariance and dispersal phenomena (Figure 3a). The analysis also highlighted that the ancestral state for haplogroups 1 and 2 resulted in C (i.e., Mediterranean Europe) with a high posterior probability (85%). The divergence and vicariance events estimated with the BBM analysis on both the cytb dataset (Supporting Information 3: Table S3) revealed a consistent biogeographic pattern. The Mediterranean Europe area (C) was identified as the main centre of diversification for the species during its evolutional history, as high values of within-area diversification and lineage dispersal movements were inferred. Morocco (B) acted more as a sink area, receiving multiple historical lineages from area A (Canary Islands) and C, thus exporting fewer. The Canary Islands showed, particularly in the complete cytb dataset analysis, a more balanced biogeographic role, with similar values estimated for both incoming and outgoing lineage dispersal, alongside a relatively high level of within-area diversification. The Israeli region, coherently with the ancestral state reconstruction, acted as a source, providing historical lineages, while also undergoing local differentiation.

3.3. Genome-Wide Population Structure and Differentiation. The total number of reads obtained was 13,641,102, ranging from a minimum of 65,874 to a maximum of 1,032,676 reads. After the filtering step, as indicated in the method section, 66,970 SNPs and 30 samples, including individuals from different haplogroups and their respective countries, were retained.

The structure and the main groups identified by mtDNA markers were confirmed by the PCA calculated on the SNPs dataset: all individuals from Israel clustered together, as well as some individuals from the Canary Islands. All other individuals were recovered in one large cluster (Figure 4b).

The ADMIXTURE analysis of the ancestral populations showed that the optimal number of clusters (K) was 4 (Figure 4b). At $K = 2$, a first generalised differentiation between a cluster including European Mediterranean and Moroccan individuals and a cluster with the samples from the Canary Islands and Israel, at $K = 3$, a specific genomic pattern for the Moroccan individuals is highlighted, with detectable traces found in two individuals from the Canary Islands and several individuals from the European Mediterranean group. The division of the Israeli and Canarian cluster is confirmed. At K = 4, four main clusters are described, the same ones recognised by the PCA and found in the analyses on single genes, and in line with the recognised *C. melanocephala* subspecies.

Finally, the divMigrate analysis on the SNPs dataset showed that gene flow is largely from Israel, Canary Islands and Morocco towards the European populations, while very low values of gene flow were estimated amongst Morocco, Israel and Canary Islands (Figure 4c).

We found significantly high genetic differentiation between Israel and the other three populations, alongside a moderate differentiation between Canary Islands and Mediterranean Europe (Table 2). The correlation between $F_{st}/(1 - F_{st})$ and log (N_m) indicated that the increase of genetic distance is strongly associated with a decrease of gene flow, consistent with the isolation by distance model (Mantel test: r = 0.836, p-value < 0.0001; Supporting Information 2: Figure S3). In contrast, we found low levels of observed heterozygosity (H_o) across all four populations analysed, consistently lower than the expected heterozygosity (H_e), with $F_{is} > 0$ for all populations (Supporting Information 3: Table S4). These values indicated a heterozygote deficiency, which may suggest limited gene flow and potential isolation and inbreeding.

3.4. SDMs. The RF, SVM and Rpart models showed high levels of predictive performance as assessed from AUC and TSS values (Supporting Information 3: Table S5). Our analysis indicated that climatic suitability is directly proportional to the temperature isothermality (BIO3; importance of 12%), temperature of the wettest quarter (BIO8; 10.7%) and precipitation of the wettest month (BIO13; 9.4%). Thus, our model showed that the species' climatic requirements are related to warm and dry conditions outside the wettest quarter.

The present potential range based on climatic variables (Figure 5b) is basically consistent with the observed current range (Figure 5a) where the most suitable conditions are concentrated in southern Europe, the Middle East and northwestern Africa.

The palaeodistribution model showed high levels of predictive performance as assessed from AUC and TSS (Supporting Information 3: Table S5). The map of predicted suitability under the LGM scenario is shown in Figure 5c. Our reconstruction showed that under the LGM scenario, the Sardinian Warbler would largely occur in northern Africa, on the Atlantic coasts of Portugal, Spain and France, and along countries bordering the northern boundary of the Mediterranean Sea. The estimated climatic suitability during the LGM did not show a noticeable contraction of the suitable range, with the main areas of suitable habitat identified largely in the same areas of the current distribution (Figure 5c). The outcome was not dependent on the LGM model (CCSM or MIROC) used for the inference. In general, the LGM distribution inferred with CCSM predicts a lower probability of occurrence than MIROC, with a highly similar range.

4. Discussion

This work represents the first comprehensive phylogeographic investigation of *C. melanocephala*, a species of Mediterranean origin that has undergone significant range expansion in recent decades [28]. Given its strictly Mediterranean distribution and origin [36], this species offers a

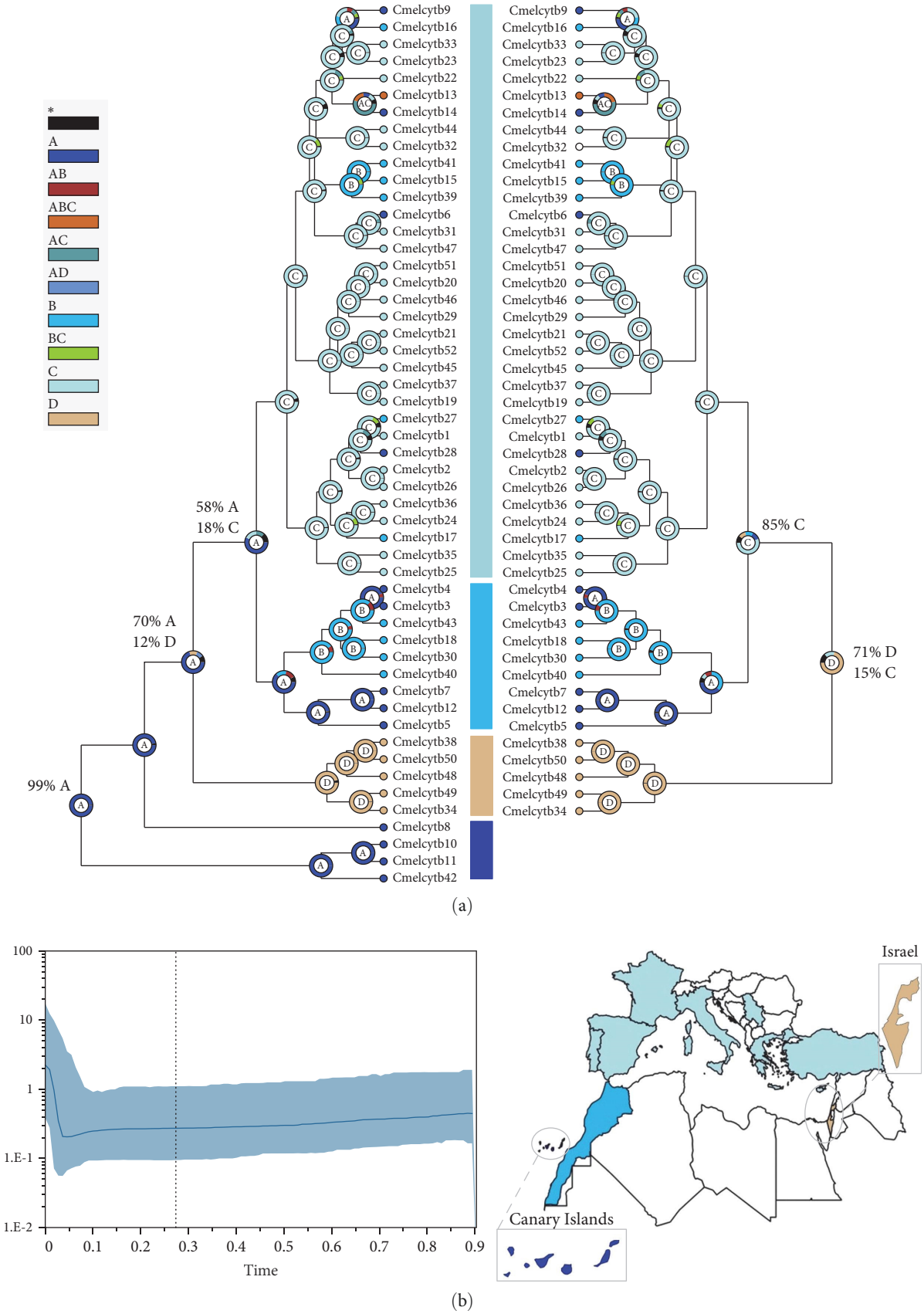
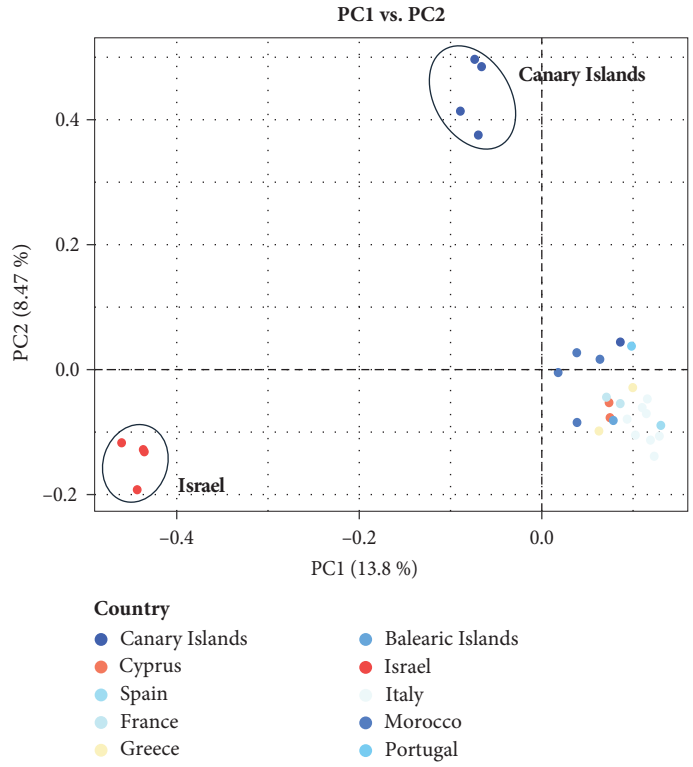


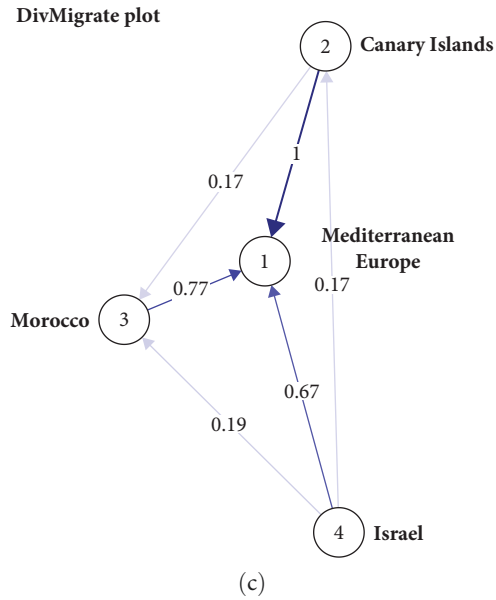
FIGURE 3: Historical biogeography reconstruction and effective populations size past trends of *C. melanocephala*. (a) Bayesian binary MCMC tree of the cytb mtDNA dataset. The probabilities of ancestral area reconstruction analysis are shown for the nodes with the pie chart colour for the different region as shown in the legend: A = Canary Islands, B = Morocco, C = Mediterranean Europe, and D = Israel, and all the combination in the case of vicariance or dispersal events, colour black with an asterisk indicates other ancestral ranges. (b) Bayesian skyline plot reporting the trend in the effective population (N_e) size during the last 0.9 million years, corresponding to the speciation event that had led to the origin of the Sardinian Warbler.



(a)

(b)

FIGURE 4: Continued.



(c)

FIGURE 4: Population genomic structure, shared ancestry and gene flow population. (a) Principal components analysis (PCA) on the identified SNPs attributed to each sample. The legend refers to the sample locality; (b) ADMIXTURE analysis plot to highlight the percentage of shared ancestry of the single-nucleotide polymorphism belonging to the considered groups (i.e., Canary Islands, Morocco, Mediterranean Europe, Israel) and (c) DivMigrate network highlighting the percentage of migration between the sample groups analysed.

TABLE 2: Population pairwise F_{st} values inferred from SNPs dataset.

Area	Morocco	Canary Islands	Israel
Mediterranean Europe	0.049*	0.097*	0.181*
Morocco	—	0.945	0.236
Canary Islands	—	—	0.228

*$p<0.05$.

valuable case for testing whether patterns of genetic structure were shaped by isolation in glacial refugia during the Pleistocene, as has been found in other European species [94–97]. Previous knowledge on the species was mostly derived from molecular studies using either a single mitochondrial marker or an SNP dataset [34, 37]. In contrast, this study provides the first comprehensive assessment of the species' phylogeography combining multilocus and ecological approaches, allowing us to investigate the genetic structure across its entire range. Our approach allowed us to assess the validity of the four extant subspecies of Sardinian Warbler: *C. m. momus* in Israel, *leucogastra* in the Canary Islands, *valverdei* from Western Sahara and Morocco and *melanocephala* from the rest of the distribution range (Figure 1a) [20], although few genetic data are available to support this distinction. Despite some morphological and colour differences, a previous molecular study did not support the *C. m. leucogastra* in the Canary Islands [34], while no information was thus far available for other subspecies.

4.1. Phylogeographic Structure. The phylogenetic reconstruction on mtDNA COI and cytb gene fragments supported the identification of a clade pertaining to *C. m. momus* in Israel, high genetic diversity in Canary Island populations, a distinct clade including individuals from Morocco and Canary Islands, and a clade widespread throughout the range, except Israel. This latter group could be identified as the *C. m. melanocephala* clade distributed from Iberia to Turkey. Data from the TGFB2 nuclear marker did not reveal a well-resolved phylogeny, an unsurprising result, given that while this gene (and other nuclear markers) is often integrated with mitochondrial genes in phylogeographic studies, the mtDNA vs nuclear phylogenies often differ due to different sorting times [98, 99]. While all individuals from Israel are grouped in the same well-supported haplogroup, the situation for the individuals from the Canary Islands and Morocco is less clear. In fact, some individuals clustered with the 'European' haplogroup, while others clustered in a unique clade that could be consistent with the subspecies *C. m. valverdei* from Morocco and Western Sahara. Finally, in the Canary Islands, where the subspecies *C. m. leucogastra* is assigned, we also found a strongly differentiated clade from Gran Canaria, highlighted by mitochondrial markers.

Our findings on populations from the Canary Islands and Morocco agree with what was found by Dietzen et al. [34]: overall low genetic divergence between the Canary Islands and the mainland, with notably higher within-population variation on Gran Canaria due to the presence of distinct relict haplotypes absent from continental populations and the other islands. Our study indicates the existence of three different clades in this archipelago: Mediterranean-European, North African and the one confined to Gran Canaria. As Dietzen et al. [34], we found that our haplotype Cmelcytb13 (corresponding to Hm01 in that study) is widespread from Crete to Morocco. Overall, the complexity of the phylogenetic relationships of *C. melanocephala* in Morocco and the Canary Islands

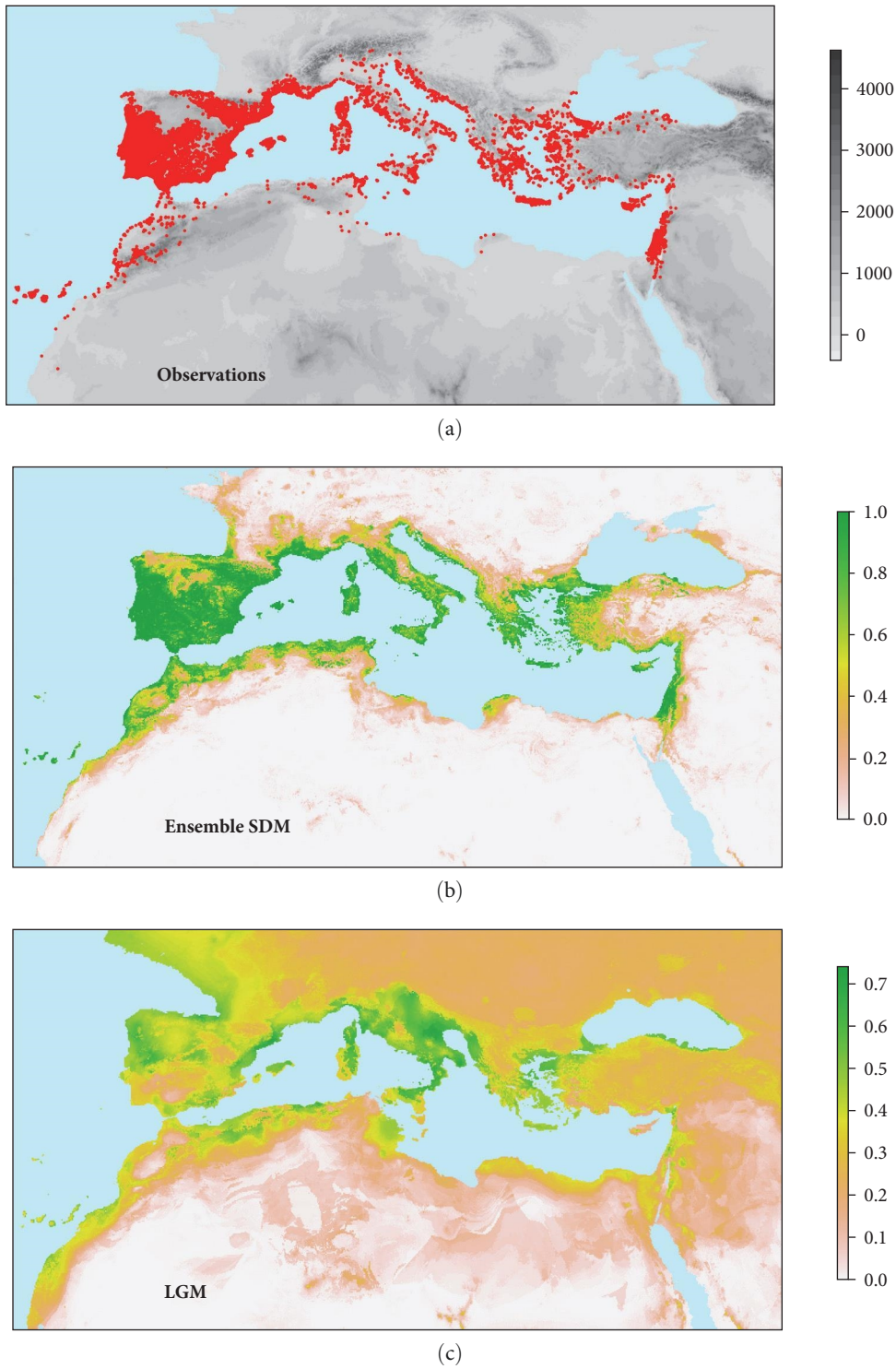


FIGURE 5: Species distribution models. (a) Occurrence locations of Sardinian Warbler *Curruca melanocephala* (red dots) in the entire range of Eurasia and Africa ($N = 14{,}713$), (b) probability of occurrence of Sardinian Warbler under the current climatic scenario (ensemble SDM) and (c) occurrence prediction for the Sardinian Warbler in Last Glacial Maximum (LGM) climatic scenario. Probabilities of occurrence are shown with a colour scale.

suggest that a comprehensive, integrative approach, including broader geographic and taxonomic sampling, additional molecular markers, and biometric and acoustic data, will be needed to resolve *C. melanocephala* phylogeography.

In our study, we integrated an extensive mitochondrial and nuclear DNA analysis with a GBS approach applied to a selected subset of individuals. This combined strategy allowed us to explore different levels of genetic variation and to detect

complementary signals of demographic history, gene flow and potential barriers to dispersal [100]. While mitochondrial and nuclear markers provided insights into historical lineage divergence and broad-scale population structure, the GBS data allowed the detection of gene flow and fine-scale spatial patterns. The recovered SNPs supported four major clades: a well-differentiated cluster corresponding to the Israeli population (*C. m. momus*), a main Mediterranean cluster, a third group that included most of the individuals from Morocco and a fourth one with several individuals from the Canary archipelago (Figure 4). Notably, the results also confirm the presence of Mediterranean European and Moroccan genotypes within populations from the Canaries, suggesting historical or ongoing gene flow between these regions.

Regarding *C. m. momus*, both mtDNA and SNP data confirmed its genetic distinctiveness. The results showed a marked genetic differentiation between this population and those from other regions. Interestingly, the Israeli clade also includes a single individual from Cyprus sampled early in the breeding period (middle April). This finding revealed the presence of two lineages on Cyprus due either to dispersal or, less likely, to a late migrant. This result confirms the recent colonisation of Cyprus by the Sardinian Warbler [37] and supports the existence of two possible colonisation sources: one from Israel and the other from other Mediterranean areas.

4.2. Historical Biogeography. The split between *C. melanothorax* and *C. melanocephala* was dated at 5.7 mya, earlier than that estimated by Blondel et al. [36], but in agreement with Cai et al. [12] and Voelker and Light [15]. Within the species, the Gran Canaria clade split at 0.97 mya, a glacial period of early Pleistocene. The Israeli clade originated ~0.56 mya, during a cold and arid phase of the Middle Pleistocene. The divergence between the Canarian/Moroccan and Mediterranean Europe clades occurred around 0.31 mya, during a warm and humid interglacial period. These results support the hypothesis that the differentiation among the clades was driven by colonisation of new areas through dispersal movements, which were not limited by the environmental and climatic conditions in the LGM. The dispersals occurred from mainland to the islands, and the presence of several lineages on the Canary archipelago can be a consequence of repeated cycles of dispersal and colonisation, alternated with isolation and genetic drift, as described for other species, for example *Erithacus rubecula superbus* [101] and *Lanius meridionalis koenigi* [102], or lead to endemic species such as *Cyanistes teneriffae* [103].

Ancestral area reconstruction, conducted on the complete cytb dataset, identified the *C. melanocephala* ancestral area on Gran Canaria. This result is likely due to highly divergent haplotypes from this island, as found by Dietzen et al. [34], while other Canary Islands host haplotypes shared with Morocco and Mediterranean Europe. The reconstruction, estimated without the divergent clade from Gran Canaria, identified Israel as the most likely centre of origin. This discrepancy was initially found by Voelker and Light [15] and could be attributed to the presence of extinct or missing haplotypes, or alternatively, to the influence of different relict lineages on the reconstruction process. In the light of the putative origin of the *Sylviidae* and of the genus *Curruca*, given in Asia [12], and considering the presence of possible relict mtDNA lineages in Gran Canaria Island, the ancestral state reconstruction of *C. melanocephala* in the easternmost of its breeding range is most likely, as our comprehensive survey encompassed individuals from Israel, Morocco, the Canary Islands and North Africa. Further analyses should include samples from areas not included in this study, such as eastern and northern Africa, to disentangle this aspect. In support of the ancestral state reconstruction, the BBM analysis highlighted the important role the Mediterranean refugia had in the intraspecific diversification of the species, while supporting the genetic differences found within individuals located in the *C. m. momus* breeding range, in some populations found within the Canary Islands (especially Gran Canaria) and in the Mauritania region in North Africa.

The current low genetic differentiation observed within most populations revealed limited geographic structuring. This pattern suggests a general panmictic structure, possibly resulting from repeated dispersal events by a small number of individuals, which may facilitate gene flow across the species' range.

Although the Sardinian Warbler is generally considered a strongly sedentary species [21, 104], with marked territorial behaviour [105, 106], occasional cases of long-distance dispersal have been reported. These include trans-Mediterranean movement from Spain to Algeria [25], as well as regular winter observations in African and Arabian regions. Genomic data confirm dispersal movements showing signatures of gene flow from the Canary Islands, Morocco and Israel to the rest of the Mediterranean areas.

Dispersive movements, as mentioned above, were almost certainly the basis for the recent colonisation of Cyprus, where, after the arrival of *C. melanocephala*, a noticeable decline of the endemic Cyprus Warbler *C. melanothorax* was observed [30, 107–110]. A genomic study on Cyprus, Greece and Crete populations showed a low differentiation between continental and island individuals, confirming the recent colonisation, within the past 30 years, from other neighbouring breeding populations [37].

Genetic analyses revealed evidence of a possible demographic expansion following an initial population contraction during LGM, as suggested by the BSP and supported by significantly negative values of Tajima's *D* and Fu's F_s. Currently, the range of the species is expanding towards the North and East (EBBA2, [28]). *C. melanocephala* is not the only member of the Sylviidae family exhibiting this pattern; a similar distribution has also been observed in *Curruca cantillans* (EBBA2, [28]). Both species show a preference for open maquis habitats and display a more generalist behaviour compared to other *Sylvia/Curruca* species. They tolerate degraded or anthropogenic environments better than, for example, *C. melanothorax*, with a discrete ability to persist after fires, signalling some territorial tenacity even after disturbance [107]. In light of these traits, the habitat changes driven by global warming may facilitate the expansion of these species [111].

While mitochondrial markers indicated signals of expansion, inbreeding coefficient (F_{is}) and observed heterozygosity

from genomic data suggest a heterozygote deficit, these values may reflect founder effects or genetic drift; however, caution is warranted given the low sample coverage in the GBS data, and further research efforts should seek to increase the number of samples per location, including islands and breeding populations not yet investigated.

4.3. Species Distribution Predictions for the LGM. The low differentiation between populations found all along the northern and southern parts of the Mediterranean basin could be explained by gene flow between suitable areas. These movements were not impeded under the LGM conditions. In fact, in our models, during the LGM, there was not a large contraction of suitable areas in the northern part of the current distribution boundaries and there were always areas with suitable climatic conditions in Mediterranean Iberia, France, Italy and Greece. In this scenario, populations of the Sardinian Warbler were not separated and could likely continue to exchange genetic material during the LGM period. In the southern part of its distribution, climatic conditions would be appropriate for Sardinian Warbler presence during the LGM from Morocco to Egypt. It is possible that the expansion of the Sahara during the LGM [112] would have acted as a barrier between the sub-Saharan Africa and Eurasia regions [113–115], thereby hampering a large expansion of the species southward. Climatic suitability was predicted to have undergone only minor spatial shifts, either slight expansions or reductions, since the LGM. However, the overall spatial connectivity between the main regions with favourable conditions remained largely stable. These findings therefore support the idea that the species' broad-scale genetic structure was shaped by the persistence of climatically suitable habitats throughout the LGM. This model differs from that of other birds, including the congeneric Mediterranean species *C. cantillans* [11, 116] and non-passeriformes species, such as Piciformes [18, 117, 118] and Strigiformes [95, 119], for which the LGM deeply influenced the current genetic structure.

5. Conclusions

This study highlights the importance of integrating genetic, genomic and ecological data to better understand population genetic differentiation in an iconic Mediterranean species such as *C. melanocephala*. By adopting a multilocus approach, we were able to provide new insights into the origin and patterns of genetic diversity among populations across the species' range. The results suggest that lineage divergence was not driven by glacial refugia during the Late Pleistocene, but instead originated in the Middle Pleistocene, likely triggered by dispersal events and the colonisation of new areas with suitable habitat and climatic conditions. Our data confirm the existence of a well-differentiated Israeli clade corresponding to the *momus* subspecies, as well as three additional well-supported clades: one primarily restricted to the Canary Islands, one in Morocco, and one broadly distributed across the remainder of the species' range. Both genetic and genomic evidence reveal gene flow between geographical regions. Further research should focus on previously unstudied populations and conduct more comprehensive genomic analyses of the Canarian populations to clarify the yet unresolved population structure.

This phylogeographic evidence fits into a particularly interesting ecological context, that of a species which is currently expanding and is capable of colonising semi-natural and heavily anthropised environments, including degraded habitats such as abandoned olive groves, secondary scrubland and peri-urban hedgerows. This ecological plasticity, combined with reduced genetic structuring across much of its range, could favour its competitive role against more specialised congeneric species, potentially leading to local replacement dynamics. In this sense, the species is an interesting model for studying the effects of Mediterranean landscape transformations on bird diversity, as well as the mechanisms that favour the expansion of generalist taxa in a context of large-scale environmental change.

Data Availability Statement

All mitochondrial and nuclear sequences from this study are available in GenBank (https://www.ncbi.nlm.nih.gov/genbank/, accession numbers in Supporting Information 1: Table S1) and in Zenodo at the following link: https://doi.org/10.5281/zenodo.17512012.

Disclosure

The present original paper is number 25 of the ZooDiversity Lab (University of Piemonte Orientale), publication number 1722 involving the Biodiversity Research and Teaching Collections at Texas A&M University, and number 14 of the Widespread Integration of Neoaves Evolution and Systematics (W.I.N.E.S.) collaborative research group.

Conflicts of Interest

The authors declare no conflicts of interest.

Author Contributions

Martina Nasuelli and Irene Pellegrino performed DNA extraction, PCR amplification, library preparation, statistical and population genetic analyses. Marco Pavia, Giovanni Boano, Gary Voelker, Alexander N. G. Kirschel, Tamer Albayrak and Michaella Moysi provided samples. Marco Cucco performed SDM analyses. Andrea Galimberti, Irene Pellegrino, Marco Pavia, Giovanni Boano, Gary Voelker, Marco Cucco, Luca Ilahiane and Martina Nasuelli jointly interpreted the results in an ecological and evolutionary context and wrote the manuscript. Flavio Mignone and Francesco Recco conducted quality control, alignment and variant calling. Irene Pellegrino supervised the research workflow.

Funding

This work was partially supported by the National Biodiversity Future Center, NBFC (Spoke 5) under the cascading project 'BIOURBAN-IMON' funded by the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4 – Call for Tender Number 3138 of December

16, 2021, rectified by Decree n.3175 of December 18, 2021 of Italian Ministry of University and Research funded by the European Union, NextGenerationEU. Project code: CN_00000033, Concession Decree Number 1034 of June 17, 2022, adopted by the Italian Ministry of University and Research (Grant CUP H43C22000530001).

Acknowledgments

The authors would like to thank the curators and the institutions that provided the biological samples used in this study: Martin Päckert of the Senckenberg Society for Nature Research, Petros Lymberakis of the Natural History Museum of Crete, Lars Erik Johannessen of the Natural History Museum of Oslo, Javier Quesada of the Museum of Natural Science of Barcelona, Amos Belmaker of the Jerusalem Bird Observatory and Manolo Suárez of the Grup Balear d'Ornitologia i Defensa de la Naturalesa. In addition, the authors would like to thank ISPRA, represented by Lorenzo Serra, as well as the Regions of Sardinia, Calabria and Abruzzo, for providing the permits required for the collection of biological samples in the field. The authors' gratitude is also extended to Gloria Ramello, Gianluca Congi and Elisa Mancuso for their help in the field activities.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting Information 1. Table S1: List of all the samples involved in the study, consisting of sampling dates, information on countries, type of starting biological material, sex and age, haplotype attribution, GenBank accession numbers and the collection of origin. In addition, the records of all GenBank and BOLD accession numbers of the sequences used in the study are provided

Supporting Information 2. Figures S1–S3: File reporting all the Supporting Figures, including the mtDNA Median-Joining haplotype networks, the TGFB2 allele haploweb with the related phylogenetic tree, and finally the plot representing isolation-by-distance estimated using the identified genomic SNPs.

Supporting Information 3. Table S2–S5: List of supporting information tables reporting the genetic indices estimated on the mtDNA and SNPs dataset, BBM output summaries and the detailed information on the SDM reconstruction.

References

[1] J. C. Avise, J. Arnold, R. M. Ball, et al., "Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics," *Annual Review of Ecology and Systematics* 18, no. 1 (1987): 489–522.

[2] S. V. Drovetski, I. V. Fadeev, M. Raković, et al., "A Test of the European Pleistocene Refugial Paradigm, Using a Western Palaearctic Endemic Bird Species," *Proceedings of the Royal Society B: Biological Sciences* 285, no. 1889 (2018): 20181606.

[3] D. Nespoli, I. Pellegrino, M. Galaverni, et al., "Disentangling the Taxonomic Status and Phylogeographic Structure of Marmora's (*Curruca sarda*) and Balearic Warbler (*Curruca balearica*): A Genetic Multi-Marker Approach," *Journal of Ornithology* 162, no. 3 (2021): 909–918.

[4] J. C. Avise and D. E. Walker, "Pleistocene Phylogeographic Effects on Avian Populations and the Speciation Process," *Proceedings of the Royal Society of London. Series B: Biological Sciences* 265, no. 1395 (1998): 457–463.

[5] P. Taberlet, L. Fumagalli, A. Wust-Saucy, and J. Cosson, "Comparative Phylogeography and Postglacial Colonization Routes in Europe," *Molecular Ecology* 7, no. 4 (1998): 453–464.

[6] G. M. Hewitt, "Post-Glacial Re-Colonization of European Biota," *Biological Journal of the Linnean Society* 68, no. 1-2 (1999): 87–112.

[7] G. M. Hewitt, "Genetic Consequences of Climatic Oscillations in the Quaternary," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359, no. 1442 (2004): 183–195.

[8] J. Blondel, *The Mediterranean Region: Biological Diversity in Space and Time* (Oxford University Press, 2010).

[9] L. G. Pârâu and M. Wink, "Common Patterns in the Molecular Phylogeography of Western Palearctic Birds: A Comprehensive Review," *Journal of Ornithology* 162, no. 4 (2021): 937–959.

[10] R. Covas and J. Blondel, "Biogeography and History of the Mediterranean Bird Fauna," *Ibis* 140, no. 3 (1998): 395–407.

[11] U. Perktaş, T. N. De Silva, E. Quintero, and Ç. Tavşanoğlu, "Adding Ecology Into Phylogeography: Ecological Niche Models and Phylogeography in Tandem Reveals the Demographic History of the Subalpine Warbler Complex," *Bird Study* 66, no. 2 (2019): 234–242.

[12] T. Cai, A. Cibois, P. Alström, et al., "Near-Complete Phylogeny and Taxonomic Revision of the World's Babblers (Aves: Passeriformes)," *Molecular Phylogenetics and Evolution* 130 (2019): 346–356.

[13] J. Fjeldså, L. Christidis, and P. G. P. Ericson, *The Largest Avian Radiation: The Evolution of Perching Birds, or the Order Passeriformes*, (2020).

[14] T. Cai, S. Shao, J. D. Kennedy, et al., "The Role of Evolutionary Time, Diversification Rates and Dispersal in Determining the Global Diversity of a Large Radiation of Passerine Birds," *Journal of Biogeography* 47, no. 7 (2020): 1612–1625.

[15] G. Voelker and J. E. Light, "Palaeoclimatic Events, Dispersal and Migratory Losses Along the Afro-European Axis as Drivers of Biogeographic Distribution in, *Sylvia*, Warblers," *BMC Evolutionary Biology* 11, no. 1 (2011): 163.

[16] I. Pellegrino, A. Negri, G. Boano, et al., "Evidence for Strong Genetic Structure in European Populations of the Little Owl *Athene noctua*," *Journal of Avian Biology* 46, no. 5 (2015): 462–475.

[17] L. Ilahiane, G. Boano, M. Pavia, et al., "Completing the Genetic Puzzle of the Reed Warbler Complex: Insights From Italy," *Bird Study* 67, no. 4 (2020): 440–447.

[18] U. Perktaş, G. F. Barrowclough, and J. G. Groth, "Phylogeography and Species Limits in the Green Woodpecker Complex (Aves: Picidae): Multiple Pleistocene Refugia and Range Expansion Across Europe and the Near East," *Biological Journal of the Linnean Society* 104, no. 3 (2011): 710–723.

[19] U. Perktaş, H. Gür, and E. Ada, "Historical Demography of the Eurasian Green Woodpecker: Integrating Phylogeography and Ecological Niche Modelling to Test Glacial Refugia Hypothesis," *Folia Zoologica* 64, no. 3 (2015): 284–295.

[20] H. Shirihai and L. Svensson, *Handbook of Western Palearctic Birds* (Larks to Warblers. Helm, 2018).
[21] J. M. Bas, P. Pons, and C. Gómez, "Home Range and Territory of the Sardinian Warbler *Sylvia melanocephala* in Mediterranean Shrubland," *Bird Study* 52, no. 2 (2016): 137–144.
[22] S. Belliure and M. Clobert, "Dispersal Distances Predict Subspecies Richness in Birds," *Journal of Evolutionary Biology* 13, no. 3 (2000): 480–487.
[23] R. Aymí and G. Gargallo, in *In Birds of the World*, ed. J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, E. de Juana, and S. M. Billerman, (Cornell Lab of Ornithology, 2024).
[24] S. Cramp, K. E. L. Simmons, and C. M. Perrins, "Handbook of the Birds of Europe," in *The Middle East and North Africa: The Birds of the Western Palearctic*, (Oxford University Press, 1988).
[25] G. Gargallo and R. Aymí, "Recovery in Algeria of a Sardinian Warbler (*Sylvia melanocephala*) Ringed in Catalonia (NE Spain)," *Butlletí del Grup Català d'Anellament* 12 (1995).
[26] F. Spina, S. R. Baillie, F. Bairlein, W. Fiedler, and K. Thorup, *The Eurasian African Bird Migration Atlas* (EURING, and the Max Planck Institute., 2022).
[27] Bird Life International, *European Red List of Birds* (Publications Office of the European Union, 2021).
[28] V. Keller, S. Herrando, P. Voříšek, et al., *European Breeding Bird Atlas 2: Distribution, Abundance and Change* (European Bird Census Council & Lynx Edicions, 2020).
[29] D. Pomeroy, F. Walsh, P. Flint, M. Hellicar, and P. Shaw, "A Sustained Decline in Cyprus Warbler *Sylvia melanothorax* Numbers in Western Cyprus, Coinciding With the Colonisation of Its Breeding Range by the Sardinian Warbler *S. melanocephala*," *Bird Conservation International* 26, no. 4 (2016): 436–450.
[30] N. Papanikolas, T. G. Hadjikyriakou, M. Sebastianelli, and A. N. G. Kirschel, "Habitat Selection and Interspecific Competition Between *Sylvia* Warblers in Cyprus Following the Rapid Expansion of a Recent Colonizer," *Avian Conservation and Ecology* 16, no. 2 (2021): art11.
[31] F. Gill, D. Donsker, and P. Rasmussen, *IOC World Bird List* (IOC World Bird List (v15.1), 2025).
[32] S. M. Goodman and P. L. Meininger, *The Birds of Egypt* (Oxford University Press, 1989).
[33] J. Cabot and C. Urdiales, "The Subspecific Status of Sardinian Warblers *Sylvia melanocephala* in the Canary Islands With the Description of a New Subspecies From Western Sahara," *Zenodo* 11 (2005).
[34] C. Dietzen, E. Garcia-Del-Rey, G. D. Castro, and M. Wink, "Phylogenetic Differentiation of *Sylvia* Species (Aves: Passeriformes) of the Atlantic Islands (Macaronesia) Based on Mitochondrial DNA Sequence Data and Morphometrics," *Biological Journal of the Linnean Society* 95, no. 1 (2008): 157–174.
[35] H. H. Bergmann, "Inseldialekte in den Alarmrufen von Weissbart- und Samtkopfgrasmücke (*Sylvia cantillans* und *S. melanocephala*)," *Die Vogelwarte* 28 (1976): 245–257.
[36] J. Blondel, F. Catzeflis, and P. Perret, "Molecular Phylogeny and the Historical Biogeography of the Warblers of the Genus *Sylvia* (Aves)," *Journal of Evolutionary Biology* 9, no. 6 (1996): 871–891.
[37] M. Moysi, B. O. Ogolowa, C. Nikiforou, et al., "Genomic Data Reveal Contrasting Patterns of Divergence Among Island and Mainland Birds of the Eastern Mediterranean," *Ibis* 165, no. 3 (2023): 829–843.
[38] L. Carrera, M. Pavia, and S. Varela, "Birds Adapted to Cold Conditions Show Greater Changes in Range Size Related to Past Climatic Oscillations Than Temperate Birds," *Scientific Reports* 12, no. 1 (2022): 10813.
[39] BirdLife International & Handbook of the Birds of the World, *Bird Species Distribution Maps of the World* (Data Zone, 2021).
[40] S. Ratnasingham, C. Wei, D. Chan, et al., "BOLD v4: A Centralized Bioinformatics Platform for DNA-Based Biodiversity Data," in *DNA Barcoding: Methods and Protocols*, (Springer, 2024): 403–441.
[41] P. D. N. Hebert, M. Y. Stoeckle, T. S. Zemlak, and C. M. Francis, "Identification of Birds Through DNA Barcodes," *PLoS Biology* 2, no. 10 (2004): e312.
[42] C. Dietzen, H.-H. Witt, and M. Wink, "The Phylogeographic Differentiation of the European Robin *Erithacus rubecula* on the Canary Islands Revealed by Mitochondrial DNA Sequence Data and Morphometrics: Evidence for a New Robin Taxon on Gran Canaria?" *Avian Science* 2, no. 3 (2003): 115–132.
[43] C. R. Primmer, T. Borge, J. Lindell, and G.-P. Sætre, "Single-Nucleotide Polymorphism Characterization in Species With Limited Available Sequence Information: High Nucleotide Diversity Revealed in the Avian Genome," *Molecular Ecology* 11, no. 3 (2002): 603–612.
[44] K. Katoh, K. Misawa, K. Kuma, and T. Miyata, "MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform," *Nucleic Acids Research* 30, no. 14 (2002): 3059–3066.
[45] B. Grüning, R. Dale, A. Sjödin, et al., "Bioconda: Sustainable and Comprehensive Software Distribution for the Life Sciences," *Nature Methods* 15, no. 7 (2018): 475-476.
[46] P. Villesen, "FaBox: An Online Toolbox for Fasta Sequences," *Molecular Ecology Notes* 7, no. 6 (2007): 965–968.
[47] M. Stephens, N. J. Smith, and P. Donnelly, "A New Statistical Method for Haplotype Reconstruction From Population Data," *The American Journal of Human Genetics* 68, no. 4 (2001): 978–989.
[48] P. Librado and J. Rozas, "DnaSP v5: A Software for Comprehensive Analysis of DNA Polymorphism Data," *Bioinformatics* 25, no. 11 (2009): 1451-1452.
[49] M. Bruneaux and M. Vecchi, "Tardipede/Concatipede:First Release of Concatipede (Versione v1.0.0)," *Zenodo* 27 (2021).
[50] R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2025).
[51] Posit Team, *RStudio: Integrated Development Environment for R. Posit Software* (Posit Team, 2025).
[52] M. C. Fisher-Reid and J. J. Wiens, "What Are the Consequences of Combining Nuclear and Mitochondrial Data for Phylogenetic Analysis? Lessons From Plethodonsalamanders and 13 Other Vertebrate Clades," *BMC Evolutionary Biology* 11, no. 1 (2011): 300.
[53] R. Lanfear, P. B. Frandsen, A. M. Wright, T. Senfeld, and B. Calcott, "PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses," *Molecular Biology and Evolution* 13 (2016): msw260.
[54] M. A. Suchard, P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond, and A. Rambaut, "Bayesian Phylogenetic and Phylodynamic Data Integration Using BEAST 1.10," *Virus Evolution* 4, no. 1 (2018).
[55] A. Rambaut, A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard, "Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7," *Systematic Biology* 67, no. 5 (2018): 901–904.

[56] L.-G. Wang, T. T.-Y. Lam, S. Xu, et al., "Treeio: An R Package for Phylogenetic Tree Input and Output With Richly Annotated and Associated Data," *Molecular Biology and Evolution* 37, no. 2 (2020): 599–603.
[57] G. Yu, D. K. Smith, H. Zhu, Y. Guan, and T. T.-Y. Lam, "ggtree: An R Package for Visualization and Annotation of Phylogenetic Trees With Their Covariates and Other Associated Data," *Methods in Ecology and Evolution* 8, no. 1 (2017): 28–36.
[58] H. R. L. Lerner, M. Meyer, H. F. James, M. Hofreiter, and R. C. Fleischer, "Multilocus Resolution of Phylogeny and Timescale in the Extant Adaptive Radiation of Hawaiian Honeycreepers," *Current Biology* 21, no. 21 (2011): 1838–1844.
[59] M. Nasuelli, L. Ilahiane, G. Boano, et al., "Phylogeography of *Lanius senator* in Its Breeding Range: Conflicts Between Alpha Taxonomy, Subspecies Distribution and Genetics," *The European Zoological Journal* 89, no. 1 (2022): 941–956.
[60] K. Tamura, G. Stecher, S. Kumar, and F. U. Battistuzzi, "MEGA11: Molecular Evolutionary Genetics Analysis Version 11," *Molecular Biology and Evolution* 38, no. 7 (2021): 3022–3027.
[61] P. Alström, P. G. P. Ericson, U. Olsson, and P. Sundberg, "Phylogeny and Classification of the Avian Superfamily Sylvioidea," *Molecular Phylogenetics and Evolution* 38, no. 2 (2006): 381–397.
[62] K. Böhning-Gaese, M. D. Schuda, and A. J. Helbig, "Weak Phylogenetic Effects on Ecological Niches of, *Sylvia*, Warblers," *Journal of Evolutionary Biology* 16, no. 5 (2003): 956–965.
[63] L. Valente, J. C. Illera, K. Havenstein, T. Pallien, R. S. Etienne, and R. Tiedemann, "Equilibrium Bird Species Diversity in Atlantic Islands," *Current Biology* 27, no. 11 (2017): 1660–1666.e5.
[64] J. W. Leigh and D. Bryant, "Popart: Full-Feature Software for Haplotype Network Construction," *Methods in Ecology and Evolution* 6, no. 9 (2015): 1110–1116.
[65] Y. Spöri and J. Flot, "HaplowebMaker and CoMa: Two Web Tools to Delimit Species Using Haplowebs and Conspecificity Matrices," *Methods in Ecology and Evolution* 11, no. 11 (2020): 1434–1438.
[66] L. Excoffier and H. E. L. Lischer, "Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses Under Linux and Windows," *Molecular Ecology Resources* 10, no. 3 (2010): 564–567.
[67] Y. Yu, C. Blair, and X. He, "RASP 4: Ancestral State Reconstruction Tool for Multiple Genes and Characters," *Molecular Biology and Evolution* 37, no. 2 (2020): 604–606.
[68] R. J. Elshire, J. C. Glaubitz, Q. Sun, et al., "A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species," *PLoS ONE* 6, no. 5 (2011): e19379.
[69] I. Pellegrino, L. Boatti, M. Cucco, et al., "Development of SNP Markers for Population Structure and Phylogeography Characterization in Little Owl (*Athene noctua*) Using a Genotyping- By-Sequencing Approach," *Conservation Genetics Resources* 8, no. 1 (2016): 13–16.
[70] I. Piccini, I. Pellegrino, D. Bellone, et al., "Combined Demographic, Ecological and Genetic Tools Reveal Connectivity Within a Fragmented Butterfly Population," *Global Ecology and Conservation* 54 (2024): e03095.
[71] H. Li and R. Durbin, "Fast and Accurate Short Read Alignment With Burrows–Wheeler Transform," *Bioinformatics* 25, no. 14 (2009): 1754–1760.
[72] P. Danecek, J. K. Bonfield, J. Liddle, et al., "Twelve Years of SAMtools and BCFtools," *GigaScience* 10, no. 2 (2021): giab008.
[73] N. C. Rochette, A. G. Rivera-Colón, and J. M. Catchen, "Stacks 2: Analytical Methods for Paired-End Sequencing Improve RADseq-Based Population Genomics," *Molecular Ecology* 28, no. 21 (2019): 4737–4754.
[74] P. Danecek, A. Auton, G. Abecasis, et al., "The Variant Call Format and VCF Tools," *Bioinformatics* 27, no. 15 (2011): 2156–2158.
[75] S. Purcell, B. Neale, K. Todd-Brown, et al., "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses," *The American Journal of Human Genetics* 81, no. 3 (2007): 559–575.
[76] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
[77] D. H. Alexander, J. Novembre, and K. Lange, "Fast Model-Based Estimation of Ancestry in Unrelated Individuals," *Genome Research* 19, no. 9 (2009): 1655–1664.
[78] L. Sundqvist, K. Keenan, M. Zackrisson, P. Prodöhl, and D. Kleinhans, "Directional Genetic Differentiation and Relative Migration," *Ecology and Evolution* 6, no. 11 (2016): 3461–3475.
[79] N. Alcala, J. Goudet, and S. Vuilleumier, "On the Transition of Genetic Differentiation From Isolation to Panmixia: What We Can Learn From Gst and D," *Theoretical Population Biology* 93 (2014): 75–84.
[80] K. Belkhir, P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme, *GENETIX 4.05 Population Genetics Software for Windows TM*University of Montpellier, 2004).
[81] B. Naimi, N. A. S. Hamm, T. A. Groen, A. K. Skidmore, and A. G. Toxopeus, "Where Is Positional Uncertainty a Problem for Species Distribution Modelling?" *Ecography* 37, no. 2 (2014): 191–203.
[82] R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis, "Very High Resolution Interpolated Climate Surfaces for Global Land Areas," *International Journal of Climatology* 25, no. 15 (2005): 1965–1978.
[83] S. E. Fick and R. J. Hijmans, "WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas," *International Journal of Climatology* 37, no. 12 (2017): 4302–4315.
[84] P. Braconnot, B. Otto-Bliesner, S. Harrison, et al., "Results of PMIP2 Coupled Simulations of the Mid-Holocene and Last Glacial Maximum – Part 1: Experiments and Large-Scale Features," *Climate of the Past* 3, no. 2 (2007): 261–277.
[85] B. L. Otto-Bliesner, S. J. Marshall, and J. T. Overpeck, "Simulating Arctic Climate Warmth and Icefield Retreat in the Last Interglaciation," *Science* 311, no. 5768 (2006): 1751–1753.
[86] R. J. Hijmans, "Raster: Geographic Data Analysis and Modeling," *R Package Version 3.6* 8 (2024).
[87] B. Naimi and M. B. Araújo, "Sdm: A Reproducible and Extensible R Platform for Species Distribution Modelling," *Ecography* 39, no. 4 (2016): 368–375.
[88] W. Werkowska, A. L. Márquez, R. Real, and P. Acevedo, "A Practical Overview of Transferability in Species Distribution Modelling," *Environmental Reviews* 25, no. 1 (2017): 127–133.
[89] M. Barbet-Massin, F. Jiguet, C. H. Albert, and W. Thuiller, "Selecting Pseudo-Absences for Species Distribution Models: How, Where and How Many?" *Methods in Ecology and Evolution* 3, no. 2 (2012): 327–338.
[90] M. B. Araújo, R. G. Pearson, W. Thuiller, and M. Erhard, "Validation of Species–Climate Impact Models Under Climate Change," *Global Change Biology* 11, no. 9 (2005): 1504–1513.
[91] F. Jiguet, V. Devictor, R. Ottvall, C. Van Turnhout, H. Van Der Jeugd, and Å. Lindström, "Bird Population Trends Are

Linearly Affected by Climate Change Along Species Thermal Ranges," *Proceedings of the Royal Society B: Biological Sciences* 277, no. 1700 (2010): 3601–3608.

[92] W. Thuiller, B. Lafourcade, R. Engler, and M. B. Araújo, "BIOMOD – a Platform for Ensemble Forecasting of Species Distributions," *Ecography* 32, no. 3 (2009): 369–373.

[93] O. Allouche, A. Tsoar, and R. Kadmon, "Assessing the Accuracy of Species Distribution Models: Prevalence, Kappa and the True Skill Statistic (TSS)," *Journal of Applied Ecology* 43, no. 6 (2006): 1223–1232.

[94] T. Albayrak, J. Gonzalez, S. V. Drovetski, and M. Wink, "Phylogeography and Population Structure of Krüper's Nuthatch *Sitta krueperi* From Turkey Based on Microsatellites and Mitochondrial DNA," *Journal of Ornithology* 153, no. 2 (2012): 405–411.

[95] P. H. Brito, "The Influence of Pleistocene Glacial Refugia on Tawny Owl Genetic Diversity and Phylogeography in Western Europe," *Molecular Ecology* 14, no. 10 (2005): 3077–3094.

[96] F. Hourlay, R. Libois, F. D'Amico, M. Sarà, J. O'Halloran, and J. R. Michaux, "Evidence of a Highly Complex Phylogeographic Structure on a Specialist River Bird Species, the Dipper (*Cinclus cinclus*)," *Molecular Phylogenetics and Evolution* 49, no. 2 (2008): 435–444.

[97] P. Salinas, F. Morinha, I. Literak, J. García, B. Milá, and G. Blanco, "Genetic Diversity, Differentiation and Historical Origin of the Isolated Population of Rooks *Corvus frugilegus* in Iberia," *Journal of Avian Biology* 52, no. 3 (2021): jav.02689.

[98] G. Song, R. Zhang, F. Machado-Stredel, et al., "Great Journey of Great Tits (*Parus Major* Group): Origin, Diversification and Historical Demographics of a Broadly Distributed Bird Lineage," *Journal of Biogeography* 47, no. 7 (2020): 1585–1598.

[99] R. M. Zink and G. F. Barrowclough, "Mitochondrial DNA Under Siege in Avian Phylogeography," *Molecular Ecology* 17, no. 9 (2008): 2107–2121.

[100] F. Vaux, L. Dutoit, C. I. Fraser, and J. M. Waters, "Genotyping-by-Sequencing for Biogeography," *Journal of Biogeography* 50, no. 2 (2023): 262–281.

[101] G. Sangster, J. A. Luksenburg, M. Päckert, C. S. Roselaar, M. Irestedt, and P. G. P. Ericson, "Integrative Taxonomy Documents Two Additional Cryptic *Erithacus* Species on the Canary Islands (Aves)," *Zoologica Scripta* 51, no. 6 (2022): 629–642.

[102] D. P. Padilla, L. G. Spurgin, E. A. Fairfield, J. C. Illera, and D. S. Richardson, "Population History, Gene Flow, and Bottlenecks in Island Populations of a Secondary Seed Disperser, the Southern Grey Shrike (*Lanius meridionalis koenigi*)," *Ecology and Evolution* 5, no. 1 (2015): 36–45.

[103] M. Päckert, J. Martens, J. Hering, L. Kvist, and J. C. Illera, "Return Flight to the Canary Islands – the Key Role of Peripheral Populations of Afrocanarian Blue Tits (Aves: *Cyanistes teneriffae*) in Multi-Gene Reconstructions of Colonization Pathways," *Molecular Phylogenetics and Evolution* 67, no. 2 (2013): 458–467.

[104] C. Coleiro, *Post-Breeding Dispersal and Adaptation of the Sardinian Warbler Sylvia Melanocephala in the Maltese Islands*L-Università ta' Malta, 2002).

[105] P. Beltrão, C. Godinho, R. Lourenço, and P. F. Pereira, "Behavioural Repeatability in Sardinian Warblers (*Sylvia melanocephala*): Larger Individuals Are More Aggressive," *Acta Ethologica* 24, no. 1 (2021): 31–40.

[106] K. J. Bensusan, B. Shorrocks, and K. C. Hamer, "Impacts of Passage Migrant Songbirds on Behaviour and Habitat Use of Resident Sardinian Warblers *Sylvia melanocephala* in Gibraltar: Impacts of Passage Migrant Songbirds," *Ibis* 153, no. 3 (2011): 616–621.

[107] P. Flint and A. McArthur, "Is the Sardinian Warbler *Sylvia melanocephala* Displacing the Endemic Cyprus Warbler S. Melanothorax on Cyprus?" *Sandgrouse* 36 (2014): 63–109.

[108] M. A. Hellicar and A. N. G. Kirschel, "Grazing Pressure and the Interaction Dynamics of the Endemic Cyprus Warbler *Sylvia melanothorax* and Its Recently Colonising Congener the Sardinian Warbler *S. melanocephala*," *Bird Conservation International* 31, no. 2 (2021): 239–254.

[109] C. Ieronymidou, N. J. Collar, and P. M. Dolman, "Endemic Cyprus Warbler *Sylvia melanothorax* and Colonizing Sardinian Warbler *Sylvia melanocephala* Show Different Habitat Associations," *Ibis* 154, no. 2 (2012): 248–259.

[110] D. Pomeroy and F. Walsh, "A European Endemic Warbler Under Threat? Population Changes in *Sylvia* Warblers on the Island of Cyprus," *Oryx* 36, no. 4 (2002): 342–348.

[111] N. Doswald, S. G. Willis, Y. C. Collingham, D. J. Pain, R. E. Green, and B. Huntley, "Potential Impacts of Climatic Change on the Breeding and Non-Breeding Ranges and Migration Distance of European, *Sylvia*, Warblers," *Journal of Biogeography* 36, no. 6 (2009): 1194–1208.

[112] J. A. Coetzee, "African Flora Since the Terminal Jurassic," in *Biological Relationships Between the Africa and South America*, ed. P. Goldbatt, (Yale University Press, 1993): 37–59.

[113] M. M. Fonseca, J. C. Brito, O. S. Paulo, M. A. Carretero, and D. J. Harris, "Systematic and Phylogeographical Assessment of the *Acanthodactylus erythrurus* Group (Reptilia: Lacertidae) Based on Phylogenetic Analyses of Mitochondrial and Nuclear DNA," *Molecular Phylogenetics and Evolution* 51, no. 2 (2009): 131–142.

[114] J. A. Godoy, J. J. Negro, F. Hiraldo, and J. A. Donázar, "Phylogeography, Genetic Structure and Diversity in the Endangered Bearded Vulture (*Gypaetus barbatus*, L.) as Revealed by Mitochondrial DNA," *Molecular Ecology* 13, no. 2 (2004): 371–390.

[115] T. R. Subedi, J. M. Peréz-García, S. Gurung, et al., "Global Range Dynamics of the Bearded Vulture (*Gypaetus barbatus*) From the Last Glacial Maximum to Climate Change Scenarios," *Ibis* 165, no. 2 (2023): 403–419.

[116] U. Perktaş and A. G. Gosler, "Species Status and Genetic Isolation in the Subalpine Warbler Complex: Insights From Mitochondrial DNA Analysis," *Avian Research* 16, no. 4 (2025): 100294.

[117] U. Perktaş and E. Quintero, "A Wide Geographical Survey of Mitochondrial DNA Variation in the Great Spotted Woodpecker Complex, *Dendrocopos major* (Aves: Picidae)," *Biological Journal of the Linnean Society* 108, no. 1 (2013): 173–188.

[118] R. M. Zink, S. V. Drovetski, and S. Rohwer, "Phylogeographic Patterns in the Great Spotted Woodpecker *Dendrocopos Major* Across Eurasia," *Journal of Avian Biology* 33, no. 2 (2002): 175–178.

[119] I. Pellegrino, A. Negri, M. Cucco, et al., "Phylogeography and Pleistocene Refugia of the Little Owl (*Athene noctua*) as Inferred From mtDNA Sequence Data," *Ibis* 156, no. 3 (2014): 639–657.