doi: http://dx.doi.org/10.1098/rspb.2018.1606

A test of the European Pleistocene refugial paradigm, using a Western Palaearctic endemic bird species

Sergei V. Drovetski, Igor V. Fadeev, Marko Rakovic´, Ricardo J. Lopes, Giovanni Boano, Marco Pavia, Evgeniy A. Koblik, Yuriy V. Lohman, Yaroslav A. Red'kin, Sargis A. Aghayan, Sandra Reis, Sofya S. Drovetskaya, Gary Voelker

ABSTRACT

Hewitt's paradigm for effects of Pleistocene glaciations on European populations assumes their isolation in peninsular refugia during glacial maxima, followed by re-colonization of broader Europe during interstadials. This paradigm is well supported by studies of poorly dispersing taxa, but highly dispersive birds have not been included. To test this paradigm, we use the dunnock (Prunella modularis), a Western Palaearctic endemic whose range includes all major European refugia. MtDNA gene tree, multilocus species tree and species delimitation analyses indicate the presence of three distinct lineages: one in the Iberian refugium, one in the Caucasus refugium, and one comprising the Italian and Balkan refugia and broader Europe. Our gene flow analysis suggests isolation of both the Iberian and Caucasus lineages but extensive exchange between Italy, the Balkans and broader Europe. Demographic stability could not be rejected for any refugial population, except the very recent expansion in the Caucasus. By contrast, northern European populations may have experienced two expansion periods. Iberia and Caucasus had much smaller historical populations than other populations. Although our results support the paradigm, in general, they also suggest that in highly dispersive taxa, isolation of neighbouring refugia was incomplete, resulting in large super-refugial populations.

INTRODUCTION

TheWestern Palaearctic was greatly affected by Pleistocene glaciations. In a series of seminal papers, Hewitt [1–4] and Taberlet et al. [5] summarized the impact of Pleistocene glaciations in delimiting the current distribution of genetic lineages, for the majority of organisms inhabiting Europe. According to their widely accepted view, during glacial advances, a large part of Europe was covered by an ice sheet, which was preceded by a wide belt of tundra and permafrost. Collectively, the ice sheet, tundra and permafrost rendered most of the continent uninhabitable for temperate and boreal taxa. Those taxa retreated to the Mediterranean peninsular refugia of Iberia, the Apennine Peninsula (Italy hereafter), the Balkan Peninsula or the area south of the Greater Caucasus. Each of these refugia was also isolated from the rest of the European mainland by east–west-running mountain ranges (Pyrenees, Alps, Dinaric Alps and Balkan Mountains, and the Caucasus, respectively), each with their own ice caps.

Subsequently, dramatic demographic declines via range reduction coupled with geographical isolation during each glacial period facilitated genetic divergence of populations surviving in different refugia. During interstadials, refugial populations provided colonists for broader Europe. Although the patterns and speed of post-glacial expansion differed among species due to their dispersal abilities, ecology or local environmental conditions, the height of the mountain ranges separating refugia from

the rest of the continent appear to have influenced the success of post-glacial colonization from those refugia. The Balkan Mountains and Dinaric Alps are the lowest among the ranges isolating peninsular refugia, with the Pyrenees being higher than either of these ranges, but lower than the Alps. Accordingly, the Balkan Peninsula appears to have served as a colonization source most frequently, followed by Iberia, whereas Italy provided very few northern colonists [1]. The role of the tallest European mountain range, the Greater Caucasus, could not be assessed at the time, due to the lack of data, but recent work has suggested that Caucasus region houses unique lineages in a number of avian species and thus could also have served as a Pleistocene refugium [6].

Paradigm species which Hewitt [1–4] and Taberlet et al. [5] used in their reviews included mammals, trees, amphibians and insects characterized by poor dispersal abilities, except perhaps the brown bear (Ursus arctos), whose dispersal abilities are moderate. However, highly dispersive organisms like birds were conspicuously omitted from these reviews. Although birds still had to retreat into the peninsular refugia during glacial advances, the degree of the isolation among refugial populations is unclear since the distances separating refugial ranges is unlikely to exceed their flight endurance limits. Indeed, some European migrants regularly cross the Mediterranean Sea, and some species have colonized Atlantic islands such as the Azores and Canary Islands [7–9].

Several avian phylogeographic studies that sampled populations from both refugial and re-colonized areas of Europe have been published recently [10–14]. While some of these studies support the refugial paradigm in documenting genetic differentiation in refugia followed by northward recolonization, none included samples from all peninsular refugia and the Caucasus, and most used species that do not exhibit migratory behaviour.

Here, we test the European refugial paradigm by assessing the possible role of glacial refugia in driving divergence in the dunnock (Prunella modularis), one of a handful of avian species endemic to the Western Palaearctic, that is best known for its highly variable mating strategies [14]. Importantly, the dunnock has a distribution that includes sedentary populations in each peninsular refugium (Iberia, Italy, Balkan) and the Caucasus [15]. Northern European populations are short distance migrants to southern Europe, and each refugial area has an expanded presence of dunnocks in winter [15]. Using a multilocus dataset, we performed phylogeographic and species tree analyses to determine whether a signature indicative of isolation in Pleistocene refugia followed by post-glacial expansion is evident in the biogeographical history of the species. In addition, we used our data to assess (i) gene flow patterns among populations residing in potential refugia and breeding in northern Europe, (ii) effective population sizes of identified lineages through time, and (iii) the possibility that identified lineages are better recognized as species, using Bayesian species delimitation methodology.

2. Material and methods

(a) Sampling and molecular protocols

We obtained a total of 121 tissue or blood samples of the dunnock from the four potential glacial refugia, and various localities in northern Europe and the Urals that were likely to have been colonized after the Last Glacial Maximum (figure 1; electronic supplementary material, table S1). Two additional mitochondrial DNA (mtDNA) ND2 sequences were downloaded from GenBank (UK: accession number AF407038; Sweden: GU816839). We included samples of three black-throated (Prunella atrogularis) and six Radde's (Prunella ocularis) accentors as outgroups [16]. With the exception of a few individuals from the non-migratory Iberian population and a few potential migrants sampled in northeastern Europe, our samples represented breeding adults or local fledglings, with the latter collected at different localities or dates to avoid sampling family members (electronic supplementary material, table S1).

Total genomic DNA was extracted from ethanol-preserved tissue or blood samples using the JETQUICK Tissue DNA Spin Kit (Genomed, Loo¨hne, Germany) according to the manufacturer's instructions. We sequenced the complete mtDNA ND2 gene (1041 base pairs (bp)) for 122 individuals. For 114 individuals, we sequenced a z-specific intron 9 of the aconitase 1 gene (ACO1I9), and for a subset of 56 birds, we sequenced 9 autosomal introns linked to different chromosomes (electronic supplementary material, table S2).

We sequenced PCR fragments in both directions on an ABI 3730 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA) at the Macrogen Europe facility and aligned resulting sequences automatically in SEQUENCHER 5.4.1 (Gene Codes Corporation, Ann Arbor, MI). We manually verified alignments to ensure consistent placement of indels and to check for stop codons in mtDNA ND2 sequences. Alleles of heterogametic individuals that differed in length were resolved manually as described previously [16–22]. Equal-length alleles of heterogametic individuals with multiple substitutions were resolved in PHASE 2.1.1 [23,24] with two independent runs. The first 500 interactions were discarded as burn-in and the following 5000 iterations used a thinning interval of 10. We used hemi- and homozygous alleles and alleles that differed by a single substitution as known in our PHASE runs.

We conducted the McDonald–Kreitman test [25] implemented in: http://mkt.uab.es [26] and the HKA test [27] implemented in HKA software (https://bio.cst.temple.edu/ ~hey/software#hka-div) to test neutrality of the mtDNA ND2 gene and multilocus dataset, respectively. In both tests, we used three pairs of dunnock clades: Iberia versus Europe þ Italy þ Balkans, Iberia versus Caucasus and Europe þ Italy þ Balkans versus Caucasus. None of these tests was significant suggesting neutral evolution of included loci.

(b) mtDNA gene tree and multilocus species tree analyses

We used the standard template in BEAST 2.4.8 [28] to reconstruct an ND2 gene tree, and the StarBeast template to reconstruct multi-locus species trees (site models, clock models and tree models unlinked) and to estimate divergence times among taxa. In tree reconstructions, we used the mean rate of sequence evolution and associated 95% HPD interval reported for mtDNA ND2 (2.9 10^{-2} substitutions/site/Ma; 95% HPD interval 2.4– 3.3 10^{-2}) of Hawaiian honeycreepers [29]. In the species tree reconstructions, we allowed evolutionary rates for nuclear loci to be estimated relative to that of the mtDNA ND2 gene (electronic supplementary material, table S2).

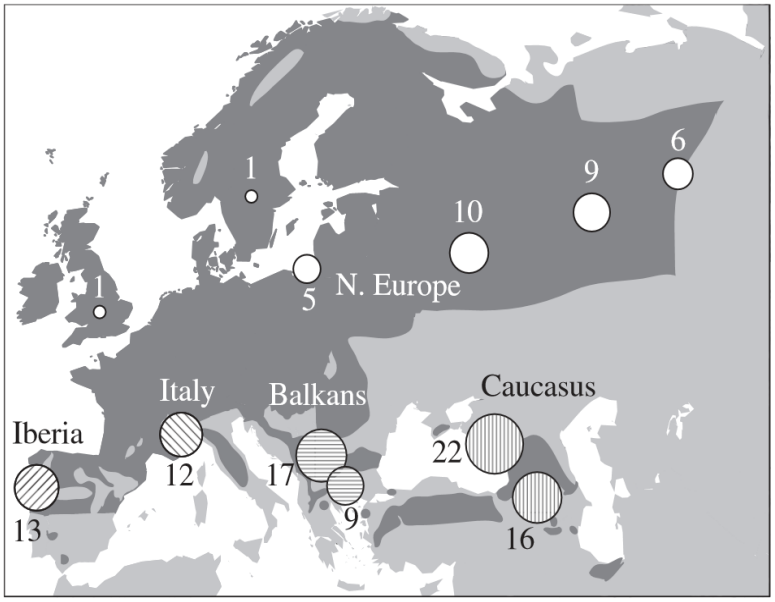

Figure 1. Species range, sampling localities and sample sizes. Area of the circles is proportional to the sample size. Different potential glacial refugia (Iberia, Italy, Balkans, Caucasus) are identified by different patterns of the circles. White circles represent northern European localities that were colonized after the Last Glacial Maximum.

We used the Bayesian information criterion (BIC) implemented in jModelTest 2.1.7 [30] to select substitution models (electronic supplementary material, table S2) for our BEAST analyses. Initially, we ran BEAST for each locus with a strict and relaxed lognormal clock priors and conducted a maximum-likelihood ratio test [31] to determine whether tree likelihoods differed significantly. We found no significant differences, and therefore we used the strict molecular clock prior in reported BEAST analyses.

Three separate MCMC analyses with Yule speciation prior were run for 10^8 generations with a 10^4 generations burn-in and parameters sampled every 10^4 steps for mtDNA ND2 gene and 10^9 generations with a 10^5 generations burn-in and parameters sampled every 10^5 steps for multilocus species tree. We combined independent runs in LogCombiner 2.4.8 [32]. Tree topologies were assessed using TreeAnnotator 2.4.8 [32] and visualized in FigTree 1.4.3 (http://tree.bio.ed.ac.uk/software/ figtree/).

(c) Molecular species delimitation

We used BPP 3.1 [33,34] to estimate the joint probability of the species tree and speciation probability across lineages. A speciation probability of 1.0 on a node indicates that all species delimitation models visited by the algorithm support the hypothesis that the lineages descending from that node represent distinct taxa. We included only nuclear genes to avoid the possibility that mtDNA would overwhelm signal from the nuclear genes and assumed no variation in gene rates. A gamma prior is used on the population size parameters (u) and the age of the root in the species tree (t_0), with other divergence time parameters being parameterized with a Dirichlet prior. We assumed a small population size (e.g. refugial) and shallow divergences. We included five P. modularis geographical groups (Europe, Balkans, Italy, Iberia and Caucasus; see results) and P. atrogularis as an outgroup. We ran the analysis for 5 10^5 generations, with a burn-in period of 1 10^5, with sampling frequency set to 5.

(d) Gene flow and effective population size analyses

We estimated gene flow among geographical areas in Migrate 3.6.11 [35]. We conducted three independent runs with gene flow allowed only between adjacent areas. In the initial run, we employed default settings for the population (u) and migration (M) parameters. In the following two runs, we used estimates of u and M from the previous run as starting values. We used Bayesian inference, random seeds, varying mutation rates estimated from data, four replicates of a single long Markov chain with 10^4 recorded

steps and 10^3 steps sampling increment. The number of sampled parameter values was 4 10^7 and 10^4 trees were discarded per chain (burn-in). We used the static heating scheme of 5 chains with temperatures 1 000 000.00, 3.00, 1.50, 1.00 and swapping interval of 2.

We constructed extended Bayesian skyline plots (EBSP) [36] in BEAST 2.4.8 for each geographical area to determine their population dynamics through time. We used a strict molecular clock with the evolutionary models and rates described above, and the coalescent extended Bayesian skyline prior. We ran 10^9 generations with 10^5 burn-in and recorded parameters every 10^5 steps with 10 initialization attempts. Population scale factors were set to 2 for autosomal loci, to 1.5 for ACO1I9 and to 0.5 for mtDNA ND2. We set the upper limit for the clock rate priors to 1, and set the hyper-prior on the mean of the population size distribution (populationMean.alltrees) to normally distributed with the mean of 1 and standard deviation of 0.1. We changed the following operator weights: Bit Flip: indicators.alltrees EBSP bitflip operator to 330; Sample Off Values: popSizes.alltrees indicators.alltrees, EBSP indicator sampler to 165; Scale: indicators.alltrees popSizes.alltrees EBSP population sizes to 165 as recommended by the EBSP tutorial available at http://evomicsorg.wpengine.netdna-cdn. com/wp-content/uploads/2015/11/ebsp2-tut1.pdf. We examined Effective Sample Sizes to confirm convergence for all parameters in TRACER 1.6. The plots were made in R 3.3.3 using plotEBSP.r script available at https://github.com/laurenchan/ scripts/blob/master/figures/PlotEBSP.r. Limit values for xand y-axes were set to best visualize posterior values and 95% central posterior density (CPD) intervals. We assumed a 1-year generation time.

3. Results

(a) mtDNA gene tree and multilocus species trees

Our mtDNA ND2 gene tree strongly supported monophyly of the dunnock (PP ¼ 1) and revealed three divergent clades (figure 2). The first clade included all Caucasus birds (PP ¼ 1), the second included all Iberian birds (PP ¼ 1), and the third clade included birds from Italy, Balkans and the remainder of Europe (PP ¼ 1). The Iberian and widely distributed clade appeared as sisters on the tree, but the relationship was poorly supported (PP ¼ 0.72). The divergence date between the Iberian and widespread mtDNA clades was 0.408 Ma (0.275–0.495, 95% HPD interval), while the divergence between these two clades and the Caucasus clade was slightly older at 0.471 Ma (0.347–0.610). MtDNA divergence dates within the dunnock were comparable to the divergence between the outgroup non-sister taxa at 0.459 Ma (0.317–0.616).

Multilocus species trees based on 10 nuclear introns with or without mtDNA ND2 (figure 3) had a very similar topology to that of the mtDNA ND2 gene tree (figure 2). Both analyses recovered the same three major clades, with the Caucasus sister to the other two (PP ¼ 1). However, in contrast to the mtDNA tree, the sister relationship of Iberia and the widespread clade was strongly supported in both species trees (PP ¼ 1). Within the widespread clade, Italy and the Balkans were recovered as sisters relative to northern Europe, however, this relationship was not statistically significant in either species tree (10 introns PP ¼ 0.50; 10 introns þ ND2 PP ¼ 0.91), suggesting poor geographical structuring among these three areas.

Both multilocus divergence date estimates were more recent than those of the mtDNA ND2 gene tree. The divergence between Iberia and the widespread clade was estimated at 0.067 Ma (0.041–0.098 Ma) with 10 nuclear introns but increased to 0.090 Ma (0.046–0.148) when mtDNA ND2 was added to the dataset. The divergence of the Caucasus from the common ancestor of the other two clades was estimated at 0.188 Ma (0.108–0.268) without ND2 and at 0.286 Ma (0.168–0.419) with mtDNA ND2 included. Both multilocus species trees indicated that the age of divergence of the Caucasus from the common ancestor of the other two clades was similar to the black-throated and Radde's accentor divergence at 0.197 Ma (0.093–0.301) without ND2 and 0.312 Ma (0.152–0.492) with ND2 (figure 3).

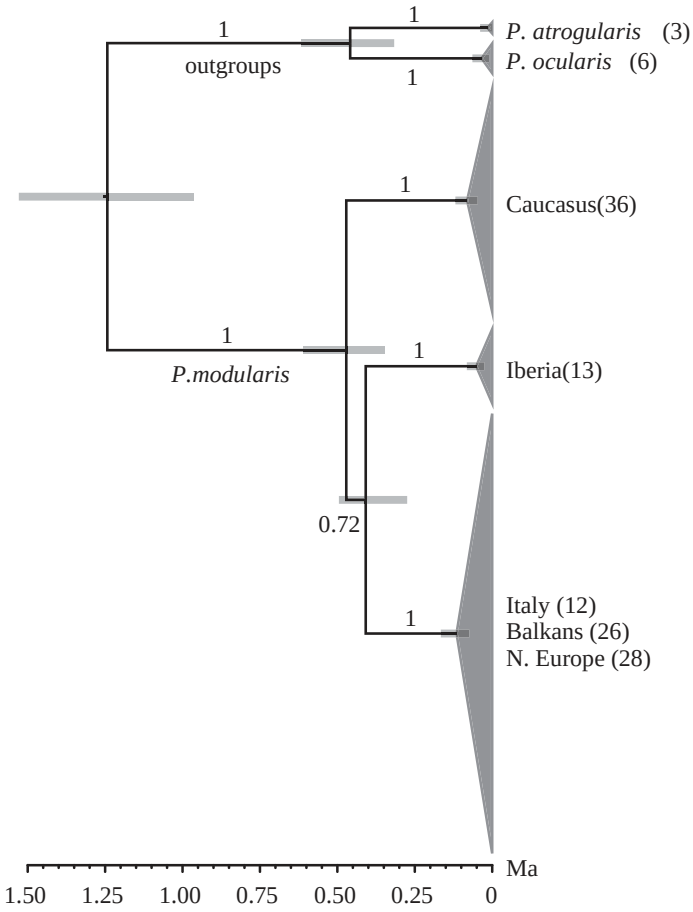

Figure 2. Phylogenetic tree based on MtDNA ND2 haplotypes. Numbers next to branches indicate their posterior probability. The grey bars next to nodes identify 95% HPD intervals for their age estimates. The time scale below the tree is in million years before present (Ma).

(b) Multispecies coalescent structure delimitation

The molecular species delimitation method BPP indicated that the dunnock is not a single lineage. Instead, BPP indicated the presence of three dunnock lineages, each strongly supported (PP ¼ 1.0). The first lineage comprises individuals from Iberia, the second comprises individuals from the Caucasus and a third comprises all individuals from northern Europe, Italy and the Balkans. Although species delimitation methods can return false-positive results for several reasons to include incorrect assignment of taxa to lineage and incorrect guide trees, our BPP results are completely consistent with our gene and species tree analyses (figures 2 and 3).

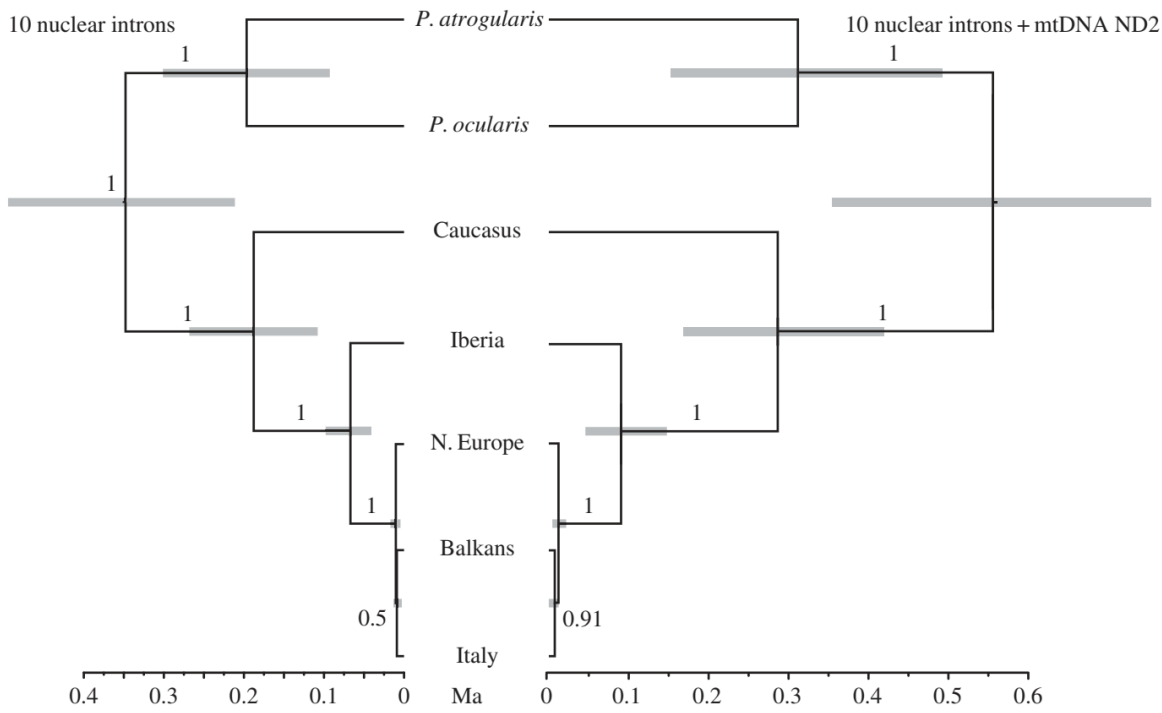

Figure 3. Multilocus species tree based on 10 nuclear introns (left) and 10 nuclear introns and MtDNA ND2 (right). Numbers next to branches indicate their posterior probability. The grey bars next to nodes identify 95% HPD intervals for their age estimates. The time scales below the trees are in million years before present (Ma).

(c) Gene flow among geographical areas and their population history

All migration parameter estimates for Iberia and Caucasus included 0 as the lower 95% confidence interval (figure 4) suggesting that the absence of gene flow between them and adjacent areas could not be rejected. The modes of migration parameters for these two areas regardless of the direction varied from 50 to 63.3. By contrast, modes of migration parameters among Italy, the Balkans and northern Europe varied from 163 to 9597, and in three cases (Italy to the Balkans, Italy to northern Europe and Balkans to northern Europe) the lower 95% confidence interval was substantially greater than 0 (5747, 593 and 1013, respectively; figure 4), suggesting extensive gene flow among these areas. However, gene flow appeared strongly asymmetric as in all three cases the gene flow parameters in opposing directions were much lower (1397 versus 9597, 450 versus 2443 and 163 versus 2637) and the lower 95% CI ¼ 0 (figure 4). In other words, gene flow from Italy to both the Balkans and northern Europe, and from the Balkans to northern Europe, is higher than in the opposite directions.

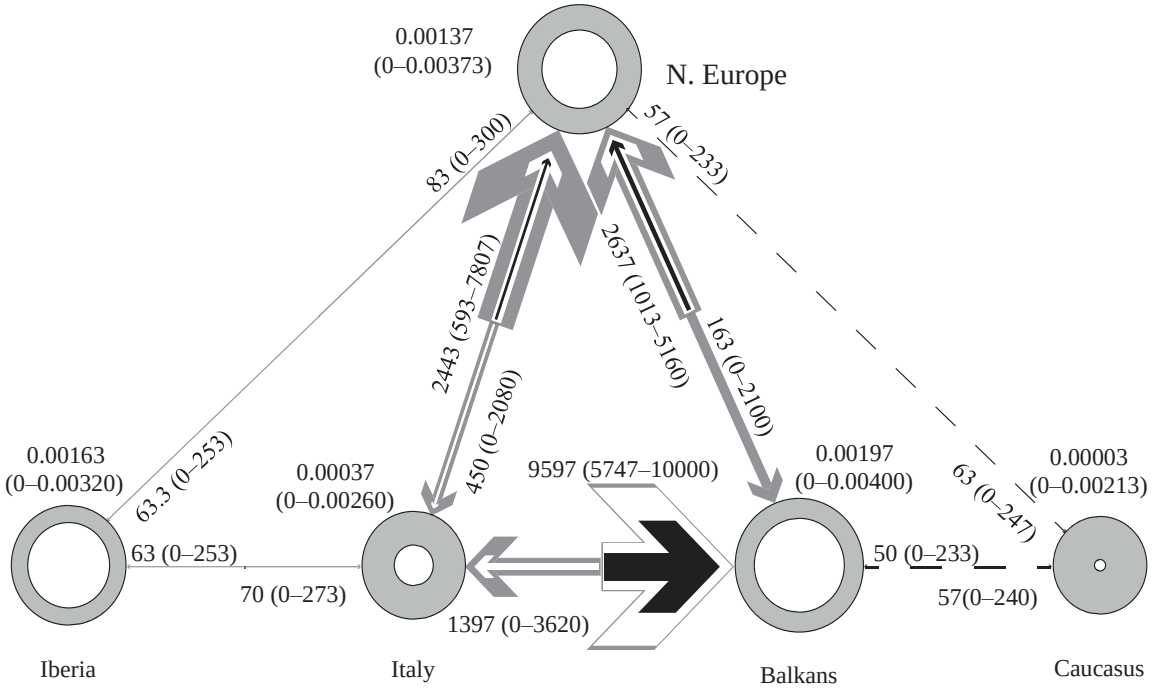

Figure 4. Results of the MIGRATE analysis. The area of the circles is proportional to the estimates of the population size parameter (u) and arrow thickness is proportional to migration parameter (M). Dashed lines represent M-values too small to be shown proportionally. Shades represent mode (white), lower 2.5% of the confidence interval (black) and upper 2.5% (grey). Size of the circles is proportional of the confidence interval (grey). Numbers next to circles and arrows represent parameter estimates—the mode and 95% confidence interval in the parentheses

Extended Bayesian skyline plots for individual geographical areas provided little temporal resolution. A stable low population size is indicated for the Iberian dunnocks over the last 10 ka, with no indication of population change through time (figure 4). A very recent (less than 1 kyr ago) rapid population growth is indicated for Caucasus dunnocks that had the lowest population of all sampled geographical areas prior to the recent growth (figure 5). Notably, the median number of population changes for Caucasian dunnocks was 1 and a stable population through time was rejected. The closely related and connected by gene flow populations of Italy, northern Europe and especially the Balkans had much larger historic population sizes. The rate and duration of the population increase appear to be inversely related with the historic population size. The Balkans had the largest population size and experienced a very slight, steady increase over the last 10 kyr. Northern Europe had an intermediate population size and its growth was faster than that of the Balkans overall, with rapid acceleration from approximately 3 kyr ago. Italy had a lower historic population size than Northern Europe and a stable population size until approximately 1.5 kyr ago when it started a rapid growth (figure 5). Although the median number of population changes for Italy and the Balkans was 1, a stable population could not be rejected because 0 was included into 95% central posterior density (CPD) interval. By contrast, the median number of population changes for northern Europe was 2 and a stable population was rejected.

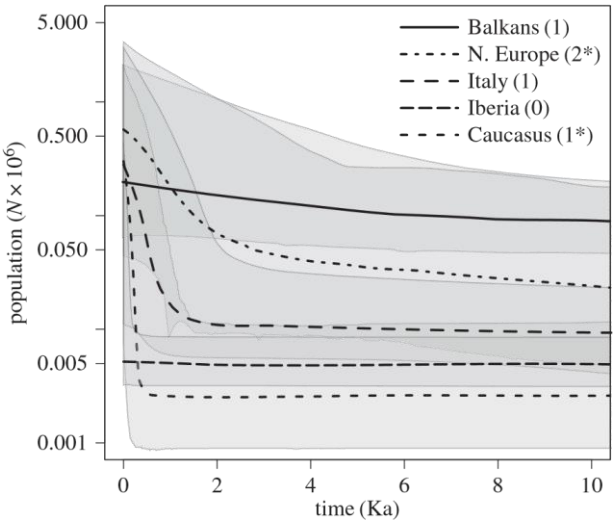

Figure 5. Extended Bayesian skyline plots for effective populations sizes (assuming generation time of 1 year) in sampled geographical areas. Lines represent median values and grey areas their 95% central posterior density (CPD) intervals. Numbers in the parenthesis behind geographical area names in the legend represent the median number of population changes and an asterisk indicates populations where 0 changes is outside the 95% highest posterior density (HPD) interval (i.e. populations in which stable size through time is rejected).

4. Discussion

This study is the first of which we are aware that directly assesses Hewitt's refugial paradigm by explicitly sampling nuclear introns

all four of the major European Pleistocene refugia (Iberia, Italy, Balkans and Caucasus). In using the Western Palaearctic breeding dunnock as our focal species, we were able to eliminate potentially confounding factors from populations outside the region of interest, which may have played a refugial or colonization role. All of our analyses—mtDNA gene tree, multilocus species tree, BPP species delimitation and MIGRATE gene flow estimates—agree with each other and clearly show three distinct, evolutionary independent dunnock groups. Further, all these analyses suggest that temperate and boreal Europe was colonized from the Balkans and Italy, which could have represented a single refugium. All trees recovered Iberia as sister to the widespread clade, although only the species trees produced statistical support for this.

Divergence estimates from the mtDNA ND2 gene tree are older as compared to the estimates derived from the species trees (figures 2 and 3). This result is not unexpected, but we feel that mtDNA may be more appropriate for lineage dating due to faster lineage sorting and lower introgression [17,18]. Furthermore, it has been argued that mtDNA dates better correspond to the palaeontological record [18]. While the palaeontological record for the dunnock is largely Late Pleistocene, there are two Middle Pleistocene records (one each from France and Italy) [37]. Therefore, the divergence estimates based solely on the mtDNA ND2 gene tree (figure 2) suggesting that the Caucasus lineage diverged from other dunnock lineages at 0.471 Ma, and that the Iberian versus widespread European lineage diverged at 0.408 Ma, are more consistent with Middle Pleistocene fossils than are the nuclear gene driven divergences dated to the Latest Pleistocene. Notably, the divergences among the three dunnock lineages are temporally similar to the divergence between the two outgroup taxa, which are not sisters (figure 2) [16]. One of these outgroup taxa (Prunella ocularis) is a near-endemic to the Caucasus [15], suggesting a shared geographical and temporal origin with the Caucasus dunnock lineage, just four ice ages ago. The dunnock is therefore a young species, and is, in fact, younger than suggested by our previous work on Prunella systematics, which lacked the detailed sampling we present here [16]. However, by comparison with other closely related lineages like the yellow and citrine wagtails (Motacilla species) complexes wherein lineages diverged extremely recently [17], the dunnock is not exceptionally young.

In assessing the role of refugia in the recolonization of post-glacial central and northern Europe, Hewitt [1,2] concluded that the Balkans were the most important refugial centre for that recolonization, given that the mountains there (Balkan Mountains and Dinaric Alps) are the lowest of the ranges separating refugia from broader Europe. The Pyrenees acted as barrier to northern recolonization from Iberia in some cases, but not in others. The Alps however seemed a more impermeable barrier for most lineages assessed, such that the Italian refugium provided few northern colonists [2]. Only in the case of the brown bear (Ursus arctos) were there data to suggest that the Caucasus refugium played a role in re-colonizing broader Europe [2].

Our results suggest that the Balkans were indeed important for re-colonizing broader Europe. Rapidly changing environmental conditions during glacial retreat would have allowed for rapid northward expansion by refugial populations. Therefore, it is especially interesting that northern Europe experienced two episodes of population growth (figure 5), which may indicate either colonization from two refugia (Italy and Balkans) or that northern Europe experienced separate waves of colonization (i.e. first from the Balkans to eastern Europe and later from Italy to western Europe). In this regard, because the Balkans and Italy are very similar genetically and have high levels of gene flow (figure 4), it may be that dunnock populations in those areas were part of a single refugial population; note that a
connection between these areas around the low-lying areas of Adriatic Sea would circumvent the need for crossing the Alps.

In a study of tawny owls (Strix aluco), Brito [10] recovered a similar pattern to the dunnock, where the Balkans were most important in re-colonization, followed by Italy. While just a few tawny owl individuals carrying Iberian haplotypes were recovered outside the region (i.e. southern France; see also green woodpecker (Picus viridis) [11]), individuals carrying either Balkan or Italian haplotypes were able to cross the Pyrenees into Iberia [10]. Our results for the dunnock indicate that Iberia, as well as the Caucasus, played no role in recolonizing broader Europe for the dunnock, and that the populations in both of these refugia are effectively isolated from the Balkans, Italy and broader Europe with respect to gene flow.

Furthermore, Iberia and the Caucasus had the lowest historical population sizes. This suggests that not only are these refugia are isolated, but also their small refugial population size would give colonizations from the Balkans and Italy an 'advantage'. Additionally, stable populations could not be rejected for all refugial populations, except the Caucasus, whereas for northern Europe it was rejected. The age of the demographic expansion in the Caucasus, however, is only a few hundred years old. This may indicate that the refugium was further south (e.g. in Iraq or Iran) with the Caucasus having been colonized very recently.

5. Conclusion

The results of our study are consistent with part of Hewitt's [1,2] paradigm, where the Balkans serve as a major source area for the recolonization of northern Europe by dunnocks, but are inconsistent with the paradigm in that Italy played a prominent role in that recolonization as well; Italian and Balkan dunnocks seem not to have been isolated from one another. The latter suggests that dispersal abilities may make isolation of neighbouring refugia less likely than for species with low dispersal abilities. Under the paradigm, Iberia plays a role in some lineages, but not others; for the dunnock there appears to be no role in recolonization. The Caucasus, for which few data existed at the time, also played no role in the recolonization of broader Europe by the dunnock.

All of our analyses suggest that the dunnock comprises three evolutionary and demographically independent lineages. These findings pose an interesting question about the evolution of the unusual mating system in this species, where breeding associations include monogamy, polygyny, polyandry and polygynandry [14]. To our knowledge, this system has been studied in only one of these three lineages, and then only in a single geographical area (England) that has been colonized since the Last Glacial Maximum. Determining whether this is an ancestral system or a recently derived one in a newly colonized area will greatly improve our understanding of its evolution. Our results provide a framework for a comparative study of mating systems within the dunnock complex by identifying source populations of British colonists and two closely related refugial populations.

Ethics. All samples were obtained from existing museum collections, except those of Armenia, which were obtained via permits issued to the Scientific Center of Zoology, National Academy of Sciences, Republic of Armenia.

Data accessibility. All sequence data are available from GenBank; see

Authors' contributions. S.V.D. and G.V. conceived, designed and coordinated the study, carried out the statistical analyses and drafted the manuscript. R.J.L. participated in drafting the manuscript. All authors participated in field collection of samples. S.R., S.S.D., S.A.A. and R.J.L. carried out the molecular laboratory work.

Competing interests. We declare we have no competing interests.

Funding. FEDER funds through the COMPETE programme, POPH/ QREN/FSE funds to S.V.D. and NORTE2020/PORTUGAL funds (NORTE-01-0145-FEDER-AGRIGEN) to R.J.L., by the Fundac¸a˜o para a Cieˆncia e a Tecnologia/MEC to S.V.D. (FCOMP-01-0124FEDER-008941; PTDC/BIA-BEC/103435/2008) and R.J.L (SFRH/ BPD/84141/2012) and by the National Geographic Society to S.V.D., NHM Belgrade program 'Ptice zapadnog palearktika' to M.R. and ANSEF grant no. NS-zoo-2983 to S.A. for Armenian fieldwork.

Acknowledgements. We thank the Laboratories of Analytical Biology (LAB) facilities of the National Museum of Natural History, Smithsonian Institution for laboratory and computer support. This is publication number 1579 of the Biodiversity Research and Teaching Collections at Texas A&M University.

References

Hewitt GM. 2004 Genetic consequences of climatic oscillations in the Quaternary. Phil. Trans. R. Soc. B 359, 183–195. (doi:10.1098/rstb.2003.1388)

Hewitt GM. 2000 The genetic legacy of the Quaternary ice ages. Nature 405, 907–913. (doi:10. 1038/35016000)

Hewitt GM. 1999 Post-glacial re-colonization of European biota. Biol. J. Linn. Soc. 68, 87–112. (doi:10.1111/j.1095-8312.1999.tb01160.x)

Hewitt GM. 1996 Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. 58, 247–276. (doi:10.1111/j.1095-8312.1996.tb01434.x)

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF. 1998 Comparative phylogeography and postglacial colonization routes in Europe. Mol. Ecol. 7, 453–464. (doi:10.1046/j.1365-294x.1998.00289.x)

Hung C-M, Drovetski SV, Zink RM. 2017 The roles of ecology, behaviour and effective population size in the evolution of a community. Mol. Ecol. 26, 3775–3784. (doi:10.1111/mec.14152)

Rodrigues P, Lopes R, Drovetski SV, Reis S, Ramos J, da Cunha R. 2013 Phylogeography and genetic diversity of the Robin (Erithacus rubecula) in the Azores Islands: evidence of a recent colonisation. J. Ornithol. 154, 889–900. (doi:10.1007/s10336-013-0953-4)

Rodrigues P, Lopes RJ, Resendes R, Ramos JA, Cunha RT. 2016 Genetic diversity of the Azores Blackbirds Turdus merula reveals multiple founder events. Acta Ornithol. 51, 221–234. (doi:10.3161/00016454AO2016.51.2.008)

Valente L, Illera JC, Havenstein K, Pallien T, Etienne RS, Tiedemann R. 2017 Equilibrium bird species diversity in Atlantic Islands. Curr. Biol. 27, 1660–1666.e5. (doi:10.1016/j.cub.2017.04.053)

Brito P. 2005 The influence of Pleistocene glacial refugia on tawny owl genetic diversity and phylogeography in Western Europe. Mol. Ecol. 14, 3077–3094. (doi:10.1111/j.1365-294X.2005. 02663.x)

Pons JM, Olioso G, Cruaud C, Fuchs J. 2010 Phylogeography of the Eurasian green woodpecker (Picus viridis). J. Biogeogr. 38, 311–325. (doi:10.1111/j.1365-2699.2010.02401.x)

Pentzold S, Tritsch C, Martens J, Tietze DT, Giacalone G, Valvo ML, Nazarenko AA, Kvist L, Packert M. 2013 Where is the line? Phylogeography and secondary contact of western Palearctic coal tits (Periparus ater: Aves, Passeriformes, Paridae). Zoologischer Anzeiger 252, 367–382. (doi:10.1016/j.jcz.2012.10.003)

Perktas¸ U, Quintero E. 2012 A wide geographical survey of mitochondrial DNA variation in the great spotted woodpecker complex, Dendrocopos major (Aves: Picidae). Biol. J. Linn. Soc. 108, 173–188. (doi:10.1111/j.1095-8312.2012.02003.x)

Pellegrino I, Negri A, Cucco M, Mucci N, Pavia M, S˘a´lek M, Boano G, Randi E. 2014 Phylogeography and Pleistocene refugia of the Little Owl Athene noctua inferred from mtDNA sequence data. Ibis 156, 639–657. (doi:10.1111/ibi.12162)

Davies NB, Lundberg A. 1984 Food distribution and a variable mating system in the dunnock, Prunella modularis. J. Anim. Ecol. 53, 895–912. (doi:10.2307/4666)

Hatchwell BJ. 2005 Family Prunellidae (Accentors). In Handbook of birds of the world (eds J del Hoyo, A Elliott, DA Christie), p. 10. Barcelona: Lynx Edicions.

Drovetski SV, Semenov G, Drovetskaya SS, Fadeev IV, Red'kin YA, Voelker G. 2013 Geographic mode of speciation in a mountain specialist avian family endemic to the Palearctic. Ecol. Evol. 3, 1518–1528. (doi:10.1002/ece3.539)

Drovetski SV, Reeves AB, Red'kin YA, Fadeev IV, Koblik EA, Sotnikov VN, Voelker G. 2018 Multi-locus reassessment of a striking discord between mtDNA gene trees and taxonomy across two congeneric species complexes. Mol. Phylogenet. Evol. 120, 43–52. (doi:10.1016/j.ympev.2017.11.023)

Drovetski SV, Semenov G, Red'kin YA, Sotnikov VN, Fadeev IV, Koblik EA. 2015 Effects of asymmetric nuclear introgression, introgressive mitochondrial sweep, and purifying selection on phylogenetic reconstruction and divergence estimates in the Pacific clade of Locustella warblers. PLoS ONE 10, e0122590. (doi:10.1371/journal.pone.0122590)

Drovetski SV, Rakovic´ M, Semenov G, Fadeev IV, Red'kin YA. 2014 Limited phylogeographic signal in sex-linked and autosomal loci despite geographically, ecologically, and phenotypically concordant structure of mtDNA variation in the Holarctic avian genus Eremophila. PLoS ONE 9, e87570. (doi:10.1371/journal.pone.0087570)

Drovetski SV, Zink RM, Ericson PGP, Fadeev IV. 2010 A multi-locus study of pine grosbeak phylogeography supports the pattern of greater intercontinental divergence in Holarctic boreal forest birds compared to birds inhabiting other highlatitude habitats. J. Biogeogr. 37, 696–706. (doi:10.1111/j.1365-2699.2009.02234.x)

Hung C-M, Drovetski SV, Zink RM. 2012 Multilocus coalescence analyses support a mtDNA-based phylogeographic history for a widespread Palearctic passerine bird, Sitta europaea. Evolution 66, 2850–2864. (doi:10.1111/j.1558-5646.2012.01657.x)

Voelker G, Semenov G, Fadeev IV, Blick A, Drovetski SV. 2015 The biogeographic history of Phoenicurus redstarts reveals an allopatric mode of speciation and an out-of-Himalayas colonization pattern. Syst. Biodivers. 13, 296–305. (doi:10.1080/14772000.2014.992380)

Stephens M, Donnelly P. 2003 A comparison of Bayesian methods for haplotype reconstruction. Amer. J. Human Genet. 73, 1162–1169. (doi:10.1086/379378)

Stephens M, Smith NJ, Donnelly P. 2001 A new statistical method for haplotype reconstruction from population data. Amer. J. Human Genetics 68, 978–989. (doi:10.1086/319501)

McDonald JH, Kreitman M. 1991 Adaptive protein evolution at the Adh locus in Drosophila. Nature 351, 652–654. (doi:10.1038/351652a0)

Egea R, Casillas S, Barbadilla A. 2008 Standard and generalized McDonald–Kreitman test: a website to detect selection by comparing different classes of DNA sites. Nucleic Acids Res. 36, W157–W162. (doi:10.1093/nar/gkn337)

Hudson RR, Kreitman M, Aguade M. 1987 A test of neutral molecular evolution based on nucleotide data. Genetics 116, 153–159.

Bouckaert R, Heled J, Ku¨hnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014 BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537. (doi:10.1371/journal.pcbi.1003537)

Lerner HRL, Meyer M, James HF, Hofreiter M, Fleischer RC. 2011 Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. Curr. Biol. 21, 1838–1844. (doi:10.1016/j.cub.2011.09.039)

Darriba D, Taboada GL, Doallo R, Posada D. 2012 jModelTest 2: more models, new heuristics and parallel computing. Nature Meth. 9, 772. (doi:10.1038/nmeth.2109)

Huelsenbeck JP, Crandall KA. 1997 Phylogeny estimation and hypothesis testing using maximum likelihood. Ann. Rev. Ecol. Syst. 28, 437–466. (doi:10.1146/annurev.ecolsys.28.1.437)

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012 Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973. (doi:10.1093/molbev/mss075)

Yang Z. 2015 The BPP program for species tree estimation and species delimitation. Curr. Zool. 61, 854–865. (doi:10.1093/czoolo/ 61.5.854)

Yang Z, Rannala B. 2010 Bayesian species delimitation using multilocus sequence data. Proc. Natl Acad. Sci. USA 107, 9264–9269. (doi:10.1073/pnas.0913022107)

Beerli P, Palczewski M. 2010 Unified framework to evaluate panmixia and migration direction among multiple sampling locations. Genetics 185, 313–326. (doi:10.1534/genetics.109.112532)

Heled J, Drummond AJ. 2008 Bayesian inference of population size history from multiple loci. BMC Evol. Biol. 8, 289. (doi:10.1186/1471-2148-8-289)

Tyrberg T. 1998 Pleistocene birds of the Palearctic: a catalogue. Cambridge, MA: Nuttall Ornithological Club.