

# Historical demography of the Eurasian green woodpecker: integrating phylogeography and ecological niche modelling to test glacial refugia hypothesis

Utku PERKTAŞ<sup>1,2\*</sup>, Hakan GÜR<sup>3</sup> and Eren ADA<sup>1</sup>

<sup>1</sup>Department of Biology (Zoology Section), Faculty of Science, Hacettepe University, 06800 Beytepe, Ankara, Turkey; e-mail: perktas@hacettepe.edu.tr

<sup>2</sup>Department of Ornithology, American Museum of Natural History, Central Park West at 79<sup>th</sup> Street, 10024 New York, NY, U.S.A.

<sup>3</sup>Department of Biology, Faculty of Arts and Sciences, Ahi Evran University, Bağbaşı, 40100 Kırşehir, Turkey

Received 23 June 2015; Accepted 15 October 2015

**Abstract.** In this study, we re-evaluated historical demography of the Eurasian green woodpecker (*Picus viridis*) on the basis of previously published multi-locus mitochondrial DNA (mtDNA) data and ecological niche modelling. We particularly aimed to test glacial refugia hypothesis during climatic oscillations of the late Quaternary for the Eurasian green woodpecker. Our results indicate that the Eurasian green woodpecker was sensitive to the effects of climate change. Prior to the Last Glacial Maximum (22000 years ago), the population contraction started and the Eurasian green woodpecker restricted its range to southern Europe (including France), Anatolia, and the Caucasus/Caspian region, and, afterwards, substantially expanded its range from this restricted area to its present range around 14500 years before present. Therefore, during the Last Glacial Maximum, we recognize a single large refugium for this species, located in southern Europe (including France), Anatolia, and the Caucasus/Caspian region.

**Key words:** *cyt-b* gene, demography, extended Bayesian skyline plot, *ND2* gene, *Picus viridis*, statistical phylogeography

## Introduction

Phylogeography plays a key role in biogeographic studies, since it helps in identifying processes that might lead to genetic differentiation between populations (Avise 2000). Historical events are known to play a vital role in shaping patterns of genetic differentiation and demography within species (Templeton et al. 1995). Along these lines, numerous phylogeographic studies indicate that the genetic differentiation of many bird species was affected by Pleistocene glacial cycles (Hewitt 1996, 1999). However, relatively few phylogeographic studies cover most of the distribution range of widespread bird species (Pavlova et al. 2003, Brito 2005, Perktas et al. 2011, Perktas & Quintero 2013). This is an essential point for phylogeographic studies, since sampling gaps are known to have significant effects on conclusions of especially genetic differentiation drawn from phylogeographic studies. Thus, one should be very careful in drawing phylogeographic

and historical demographic inferences from patterns of genetic differentiation between and within species, unless one is confident that they have sampled from the complete range of taxa involved.

At present, comprehensive phylogeographic studies mostly focus on genetic and demographic patterns of different organisms and comprehension of refugia in their distribution ranges. For example, demographic history of different species in Europe has been discussed with well-known refugia in southern Europe (Iberia, Italy, and Balkans). The principal theory for most temperate species in Europe has been that populations shifted their distribution range to the favorable areas of southern Europe (i.e. refugia) during the Last Glacial Maximum (22000 years ago). Yet, some areas in Europe have not been well discussed (e.g. Anatolia) and/or have been recently begun to discuss (e.g. extra-Mediterranean refugia, Schmitt & Varga 2012). For instance, phylogeographic studies from Anatolia have been increasing and this region

\* Corresponding Author

appeared to be an important region both as a refugium and as a substantial source of re-colonization for European biota (Korkmaz et al. 2014, Perktas et al. 2015 and references therein). In addition to the typical Mediterranean refugia, extra-Mediterranean refugia essentially describe cryptic northern refugia located at higher latitudes (e.g. south-central and southeastern Europe). This pattern of refugia has been observed in some vertebrate species, e.g. European pine marten (*Martes martes*), Ruiz-González et al. (2013).

In this study, we picked a widespread, genetically well-sampled, and well-studied woodpecker species in Europe, the Eurasian green woodpecker (*Picus viridis*). These features make this species an ideal organism to test the glacial refugia hypothesis using both a phylogeographic approach and ecological niche modelling. The Eurasian green woodpecker used to be considered as including four subspecies throughout its distribution range (Cramp 1985, Dickinson 2003). Perktas et al. (2011) proposed that three species should be recognized in Europe and the Near East. The Iberian taxon has been recognized as a different species, the Iberian green woodpecker (*Picus sharpei*); three subspecies have been recognized within the Eurasian green woodpecker (del Hoyo et al. 2014). Although this taxonomic arrangement has not been fully accepted in some recent publications (e.g. de Juana & Garcia 2015), in this study we follow Perktas et al. (2011) in recognizing the Eurasian green woodpecker and the Iberian green woodpecker as valid species; therefore, we do not include samples from the Iberian green woodpecker. Most of the global range of the Eurasian green woodpecker lies in Europe, where it is widely distributed, but absent from northern and eastern parts of Ireland, north of Scandinavia, and Iberia. Its range in Russia and Near East covers western Russia, mostly coastal areas in Turkey, the Caucasus, and northern and western Iran (Fig. 1). It is largely non-migratory, but it may perform altitudinal movements. Limited dispersal seems to be an important behavioral characteristic of this species, since it is not found on Mediterranean islands, which are not so far from mainland. Quantitative data on population sizes and trends are scarce in Europe and Russia, but it is known that over half of European population is thought to occur in France and Germany (Tucker & Heath 1994). An essential habitat requirement of this species consists of old deciduous trees (especially for nesting); this species also prefers semi-open landscapes, with small woodlands, hedgerows, and forest edges (Winkler et al. 1995). Although the Eurasian green woodpecker

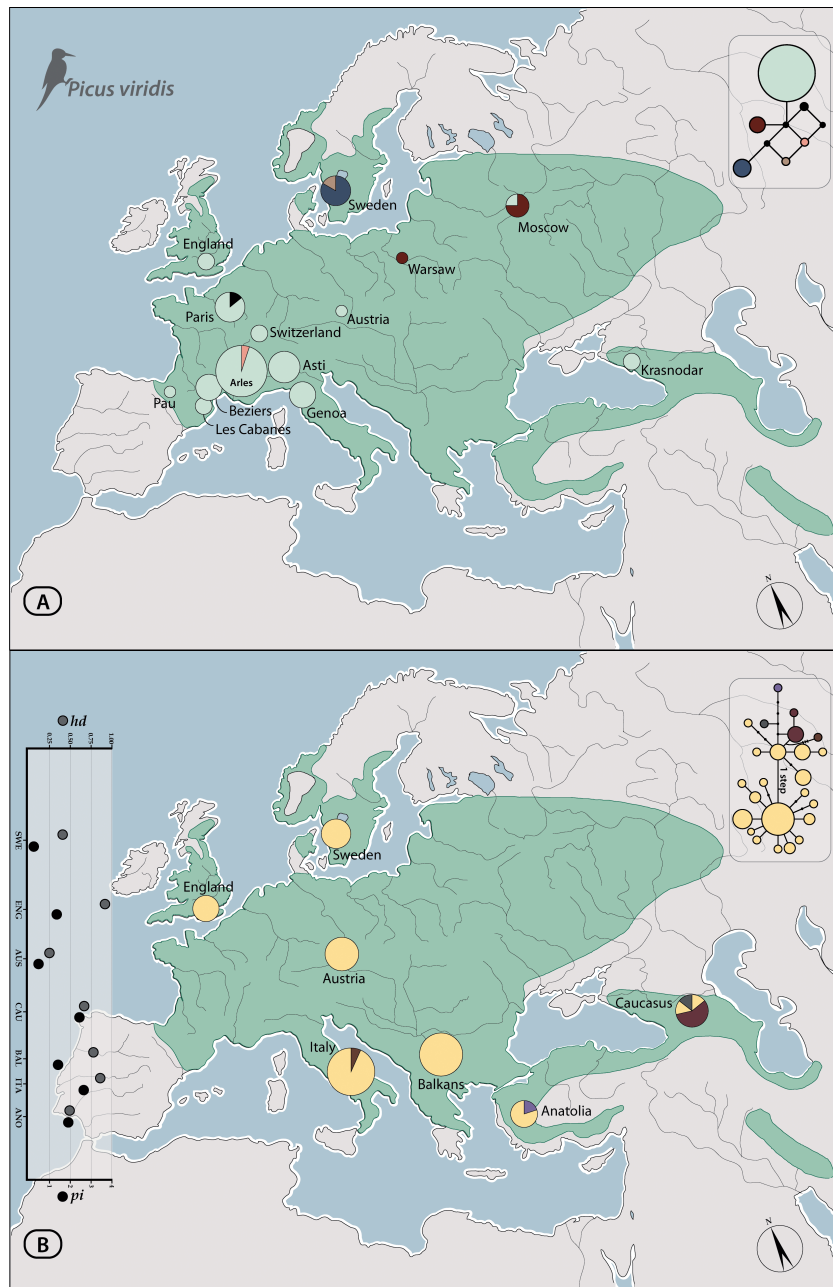
is recognized as a single species consisting of three subspecies, one of these subspecies, *P. v. innominatus*, which occurs in western Iran (Zagros Mountains), has been recognized as morphologically and genetically distinct from other subspecies (Perktas et al. 2011). It has been discussed that both morphologic and genetic differences are not sufficient to recognize this subspecies as a different species (del Hoyo et al. 2014), hence the Eurasian green woodpecker comprised of three subspecies in recently published checklists. However, we again follow the taxonomic discussion in Perktas et al. (2011). This subspecies has an allopatric distribution in Zagros Mountains in Iran, has high  $F_{st}$  values and is morphologically different from its European counterparts. Due to all of these features, this subspecies deserves a specific status (for more details, see discussion in Perktas et al. 2011). Hence, we also do not include any specimen or observation records from Zagros Mountains in this study.

Previous genetic studies investigated genetic variation and phylogeographic patterns in this species complex based on mitochondrial and nuclear genes (Pons et al. 2010, Perktas et al. 2011). Both studies showed post-glacial re-colonization events (expansion) associated with multiple refugia from east of Iberia, Italy, Balkans, Anatolia, and Caucasus region. Because of the published phylogeographic results and the taxonomic conclusions for this species, the Eurasian green woodpecker deserves a detailed demographic study to test glacial refugia hypothesis during climatic oscillations of the late Quaternary. Therefore, our objective in this study is to re-evaluate previously published multi-locus mitochondrial DNA (mtDNA) data (Pons et al. 2010, Perktas et al. 2011) through Bayesian-based demographic analysis and to reconstruct an ecological niche model in order to better understand species-specific demographic history of this species.

## Material and Methods

### Phylogeography

Samples and gene regions – this study aims to provide more detailed demographic analysis of the Eurasian green woodpecker based on previously published multi-locus data. It includes mitochondrial DNA (mtDNA) haplotypes of the cytochrome *b* (*cyt-b*) and the NADH dehydrogenase 2 (*ND2*) genes. Sample locations and haplotypes, together with their frequencies, were given in Pons et al. (2010) for the *cyt-b* gene, and in Perktas et al. (2011) for the *ND2* gene. The total sample size of the Eurasian green



**Fig. 1.** Distribution of the Eurasian green woodpecker (*Picus viridis*) in Europe, Anatolia, and Caucasus. Genetic structure of populations is indicated by pie charts (A, the *cyt-b* gene; B, the *ND2* gene); the area of each pie chart is proportional to the sampling size for each population. Both figures include haplotype networks. In figure B, scatter dot plot shows nucleotide (*pi*) and haplotype diversity (*hd*) based on the *ND2* gene for each sampling areas.

woodpecker for the phylogenetic analysis was 123 for mtDNA loci. Sample locations for mitochondrial loci were shown in Fig. 1.

Genetic variation and phylogenetics – standard measures of genetic variation, such as number of haplotypes, haplotype diversity, and nucleotide diversity, were estimated in Pons et al. (2010) and Perktas et al. (2011) for mtDNA loci. Thus, they were not estimated once again in this paper. Haplotype networks were constructed using

PopART (Population Analysis with Reticulate Trees) version 1.7.2 ([www.popart.otago.ac.nz](http://www.popart.otago.ac.nz)) with median joining network criteria implemented in NETWORK (Bandelt et al. 1999). The median joining network combines the minimum-spanning trees within a single haplotype network and yields the best gene genealogy. Therefore, it provides a good methodological approach to investigate phylogenetic relationships between mtDNA haplotypes and also to describe genetic population structure of the Eurasian

green woodpecker. On the basis of the results of haplotype network of mtDNA loci, we concluded that there was no phylogeographic structure among the Eurasian green woodpecker populations. In addition, Pons et al. (2010) used nuclear BRM intron in their phylogenetic analyses. They did not find significant variation within distribution range of the Eurasian green woodpecker. Thus, all sequences from all populations were combined to explore the historical demography of the Eurasian green woodpecker. This approach made the results of phylogeography more comparable with those of ecological niche modelling (e.g. Perktas et al. 2015).

**Historical demography** – to infer the demographic history of the Eurasian green woodpecker, we used the extended Bayesian skyline plot analysis (Heled & Drummond 2008) implemented in BEAST version 1.7.5 (Drummond et al. 2012). This analysis uses coalescent approach to estimate Bayesian inference of the effective population size changes through the time based on the information of multiple mtDNA loci. Before the extended Bayesian skyline plot runs, the best-fit substitution models were identified for each locus in MEGA version 6.06 (Tamura et al. 2013). These were the Tamura & Nei (1993) with a discrete Gamma distribution (TN93 + G,  $AIC_c = 3462.72$ ) for the *ND2* gene, Hasegawa, Kishino and Yano with a discrete Gamma distribution (HKY + G,  $AIC_c = 1466.53$ ) for the *cyt-b* gene. Because we focused on the different mtDNA data sets of the Eurasian green woodpecker, we first performed independent separate runs for each dataset (the *cyt-b* gene and the *ND2* gene) using extended Bayesian skyline plot (Heled & Drummond 2008) implemented in BEAST version 1.7.5. Because of the fact that results of separate extended Bayesian skyline plots were congruent with each other, we used both mtDNA loci (the *cyt-b* and the *ND2*) in one extended Bayesian skyline plot analysis. Multiple independent extended Bayesian skyline plot runs were performed using the following parameters: linear models, 100 million steps, parameters sampled every 10000 steps, and a burn in of 10 %. We used strict clock model with default mutation rate under uniform prior distribution, and allowed the analysis to estimate the rates, relative to the *cyt-b* gene, for the *ND2* gene. Then, we calculated the expansion time using a wide range of mutation rates of mtDNA, i.e. 1.65 % substitutions/site/million years as the mutation rate of the mtDNA *cyt-b* gene of Piciformes (Weir & Schluter 2008), and the range of 1-4 % (Brito 2005). Effective sample size values of the parameters were over 200 for each run.

**Isolation-by-distance** – a matrix of genetic distances between all pairs of populations was estimated from standard fixation index ( $F_{st}$ ) for two mtDNA loci (the *cyt-b* and the *ND2*) by DnaSP version 5.10 (Librado & Rozas 2009). Because each gene region was evaluated independently in different studies (the *cyt-b* gene in Pons et al. 2010, the *ND2* gene in Perktas et al. 2011), two different matrixes of geographic distances (km) between all pairs of populations from each study were estimated using Geographic Distance Matrix Generator (version 1.2.3). A Mantel test with 10000 random permutations was performed between matrixes of genetic ( $F_{st}$ ) and geographic (log) distances (Slatkin 1993, Rousset 1997).

**Coalescent simulations** – approximate Bayesian computation implemented in DIYABC version 2.0 beta (Cornuet et al. 2008) was used to infer most likely coalescent scenarios on the demographic history of the Eurasian green woodpecker. In approximate Bayesian computation methods, different population scenarios were modeled by simulating datasets fitting each defined scenario. Then, statistical tests were used to define which scenario best fits the observed dataset. For this study, two different scenarios were tested using the *ND2* gene, which was geographically the most complete data set in the species distribution range. First scenario simulated decreasing ancestral population effective size ( $N_{ef}$ ) before the Last Glacial Maximum (22000 years ago) to the refuge area, then after the Last Glacial Maximum  $N_{ef}$  started to increase. On the other hand, second scenario simulated stable effective population size through the time. So, no refuge area was taken into account in determining the second scenario. In both scenarios, the prior distributions for all historical parameters and mutation rate were set as uniform. Mutation rate was set between 5.00E-09 and 2.00E-08, which corresponded to 1 % and 4 % sequence divergence per Myr. The substitution model was Hasegawa-Kishino-Yano. Six different summary statistics were taken into account: number of haplotypes, number of segregating sites, mean of pairwise differences, variance of pairwise differences, Tajima's D, and private segregating sites. Two million simulated datasets were generated (1 million for each scenario). Then, posterior probability of each scenario was estimated using logistic regression on the simulated datasets.

#### *Ecological Niche Modelling*

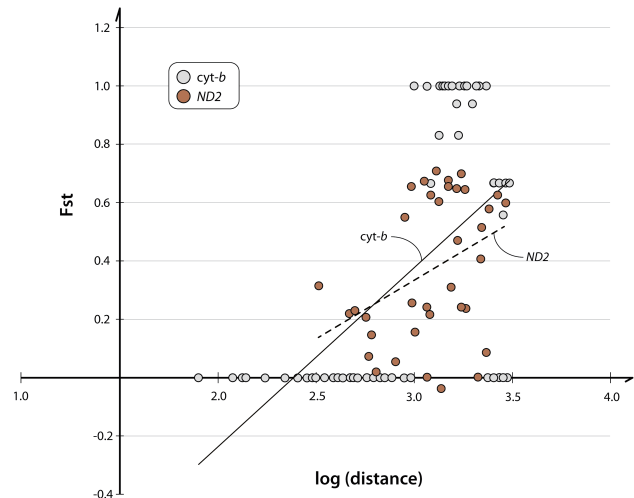
**Modelling** – species occurrence data were compiled from the following sources: GBIF (www.gbif.org), KuşBank (www.kusbank.org), and ORNIS (www.



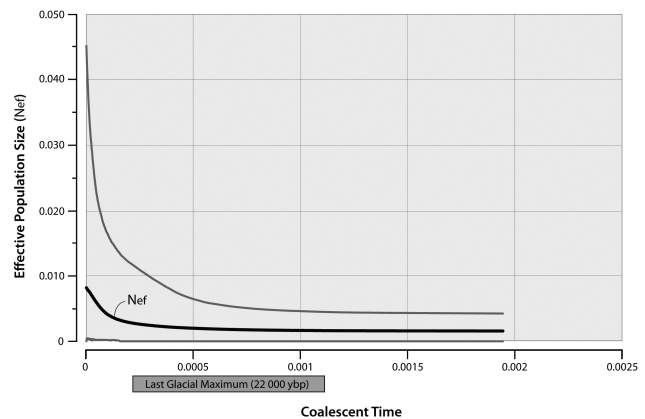
ornisnet.org). The occurrence records showed more intense sampling (e.g. France, Sweden) in the northwest of and much less (e.g. Poland, Romania) or no sampling (e.g. Russia, Ukraine) in the east of the Western Palearctic. To correct for sampling bias and to reduce spatial autocorrelation (Boria et al. 2014, Fourcade et al. 2014), the occurrence records were removed by reducing multiple records to a single record within 50 km distance, resulting 454 records used for ecological niche modelling. This reduced the spatial aggregation of the occurrence records, but could not correct the lack of occurrence records in the east of the Western Palearctic (see Fig. 5). Thus, a Gaussian kernel density of spatially filtered occurrence records was also created by a sampling bias distance of two decimal degrees, producing a bias grid used for ecological niche modelling (Fig. S1; see below; Elith et al. 2010, Fourcade et al. 2014).

Present (1950–2000) and Last Glacial Maximum (22000 years ago) bioclimatic data were downloaded from the WorldClim database (Hijmans et al. 2005, [www.worldclim.org](http://www.worldclim.org)) at a resolution of 2.5 arc-min (4.63 km at the equator). Last Glacial Maximum bioclimatic data are based on three general circulation model (GCM) simulations: CCSM4, MIROC-ESM, and MPI-ESM-P (for detailed information for downscaled paleoclimate data, see CMIP5 – Coupled Model Intercomparison Project Phase 5; [cmip-pcmdi.llnl.gov/cmip5](http://cmip-pcmdi.llnl.gov/cmip5)). Bioclimatic data include 19 bioclimatic variables derived from monthly temperature and precipitation values (for detailed descriptions of these bioclimatic variables, see [www.worldclim.org/bioclim](http://www.worldclim.org/bioclim)). Since the geographic distribution of the Eurasian green woodpecker covers a large range in the Western Palearctic, these bioclimatic variables were masked to include only  $-12^{\circ}$  to  $62^{\circ}$  E and  $29^{\circ}$  to  $72^{\circ}$  N (representing our hypothesized M area, as in Soberón & Peterson 2005). To reduce multicollinearity among bioclimatic variables (Dormann et al. 2013), some highly intercorrelated ( $r > 0.9$  or  $< -0.9$ ) variables were removed, leaving Bio 2-4, 8-11, and 15-19 (see [www.worldclim.org/bioclim](http://www.worldclim.org/bioclim)) as input variables used for ecological niche modelling. Correlation structure among these bioclimatic variables appeared to be stable and consistent across time periods (Mantel test:  $r \geq 0.857$ ,  $P < 0.0001$ ), suggesting that the model may be transferable to Last Glacial Maximum bioclimatic conditions (Jiménez-Valverde et al. 2009).

The maximum entropy machine learning algorithm in the software MaxEnt version 3.3.3k (Phillips et al. 2006, Phillips & Dudík 2008, Elith et al. 2011; [www.cs.princeton.edu/~schapire/maxent](http://www.cs.princeton.edu/~schapire/maxent)) was used

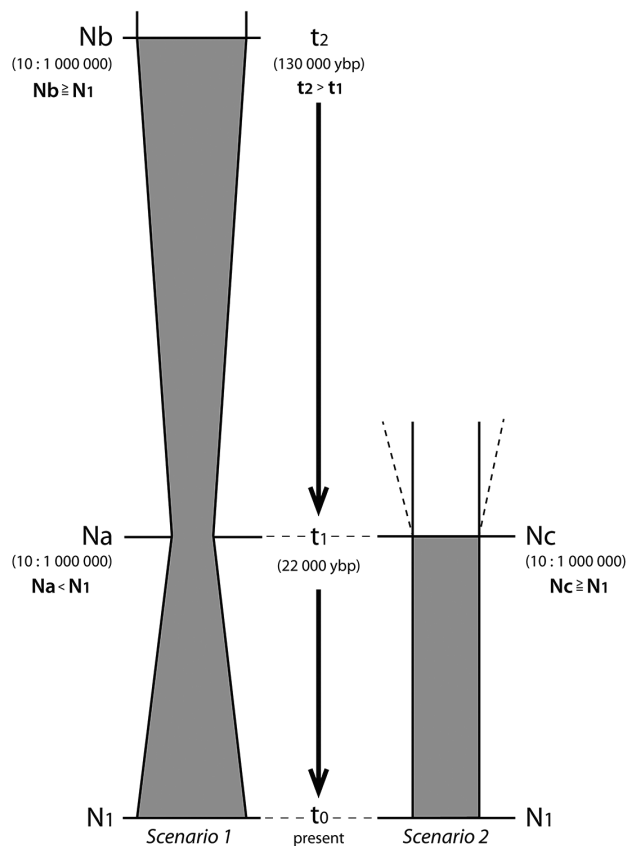


**Fig. 2.** Isolation-by-distance of populations of the Eurasian green woodpecker (*Picus viridis*) based on two loci (the *cyt-b* and the *ND2* genes).



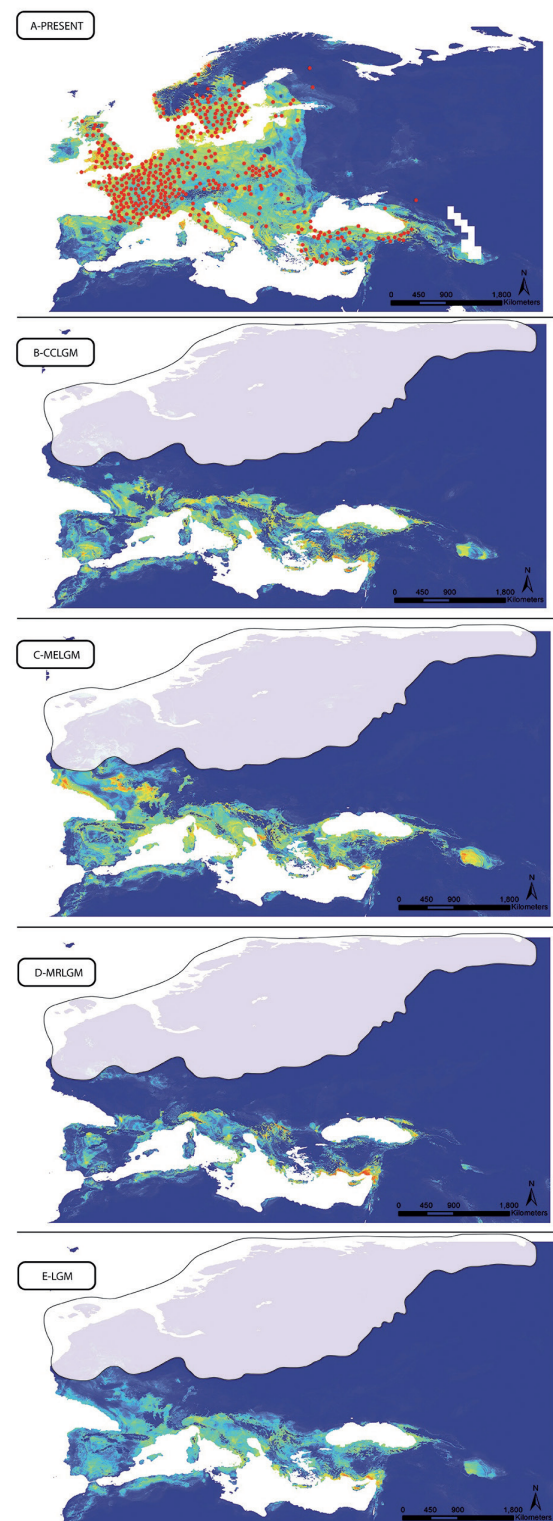
**Fig. 3.** The extended Bayesian skyline plot for the Eurasian green woodpecker (*Picus viridis*). The bar under the coalescent time axis shows the period of the Last Glacial Maximum based on different mutation rates used in this study.

to predict the geographic distribution of the Eurasian green woodpecker under present (1950–2000) and reconstructed past (Last Glacial Maximum, 22000 years ago) bioclimatic conditions. MaxEnt, which is among the most effective methods of ecological niche modelling (Elith et al. 2006), was run with the default settings, except for the cases noted below. To optimize the model performance (Shcheglovitova & Anderson 2013), 40 models were tested, using all combinations of five sets of feature types (1. linear, 2. linear and quadratic, 3. hinge, 4. linear, quadratic, and hinge, and 5. linear, quadratic, hinge, product, and threshold, as in Brown 2014) and eight values of regularization multiplier (0.05, 0.25, 0.50, 0.75, 1, 2, 5, and 10, as in Richmond et al. 2010, Gür 2013). The best model was chosen in the following order of preference: (1) model with the lowest omission rate, (2) model with

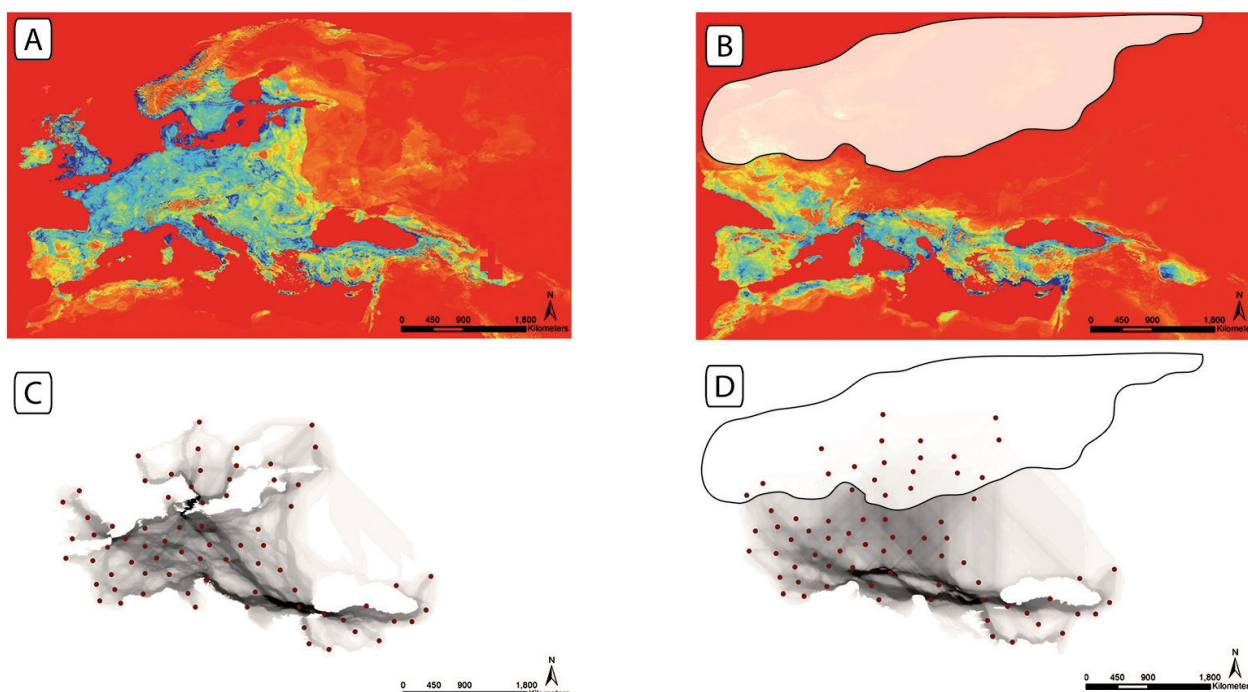


**Fig. 4.** Alternative biogeographic scenarios for the Eurasian green woodpecker (*Picus viridis*).  $t$  indicates the time scale and  $N$  indicates the effective population size. The prior distributions for the time scale and the effective population size are given in the brackets.

the highest AUC (see below), and (3) model with the simplest set of feature types (Brown 2014) and included all feature types (linear, quadratic, hinge, product, and threshold) and a regularization multiplier of 0.05. To test whether the model performance is robust to variation in the selection of the occurrence records for training and testing, a ten-fold cross-validation was performed in which a different 90 % of the occurrence records were used to train the model and 10 % were used to test it for each of 10 runs (Gür 2013). To up-weight the occurrence records with fewer neighbors in the geographic landscape, the bias grid (Fig. S1; see above) was implemented in the “bias file” option (Elith et al. 2010). To remove the degree of clamping (negative effects caused by novel bioclimatic conditions) from the predictions under Last Glacial Maximum bioclimatic conditions, fade by clamping was performed (Habel et al. 2011), because the MESS (multivariate environmental similarity surface, Elith et al. 2010) maps showed that the model was required to extrapolate into novel Last Glacial Maximum bioclimatic conditions. However, novel bioclimatic conditions encountered during



**Fig. 5.** (A) Present and (B, C, D, E) Last Glacial Maximum predictions of the geographic distribution of the Eurasian green woodpecker (*Picus viridis*). The visible area in maps is  $-12^{\circ}$  to  $62^{\circ}$  E and  $29^{\circ}$  to  $72^{\circ}$  N, representing our hypothesized M area. The red circles indicate the occurrence records and warmer colors (more red) indicate areas of relatively high suitability. The white shading shows the ice extent during the LGM (Svendsen et al. 2004). Note that Last Glacial Maximum coastlines differ from present coastlines, because the sea level was lower in the Last Glacial Maximum than in the present.



**Fig. 6.** Creation of present and Last Glacial Maximum dispersal networks for the Eurasian green woodpecker (*Picus viridis*). (A) Present and (B) Last Glacial Maximum friction layers. (C) Present and (D) Last Glacial Maximum dispersal networks. The red circles indicate the occurrence records and warmer colors (more blue in A and B, more black in C and D) indicate areas of low dispersal cost. The white shading shows the ice extent during the LGM (Svendsen et al. 2004). Note that Last Glacial Maximum coastlines differ from present coastlines, because the sea level was lower in the Last Glacial Maximum than in the present.

projecting were observed mainly in areas covered by ice during the Last Glacial Maximum (Fig. S2). Thus, extrapolation should not be a critical issue.

The area under the receiver operating characteristic (ROC) curve (AUC) was used to evaluate the model performance (Fielding & Bell 1997). An  $AUC > 0.5$  indicates that the model performs better than a random prediction, i.e.  $AUC \geq 0.9$  = very good;  $0.9 > AUC \geq 0.8$  = good; and  $AUC < 0.8$  = poor (Gassó et al. 2012). Landscape connectivity – first, the occurrence records used for ecological niche modelling were further removed by reducing multiple records to a single record within 200 km distance, resulting 73 records (see Fig. 6), to reduce time necessary for analysis and three predictions (i.e. logistic values) under Last Glacial Maximum bioclimatic conditions (i.e. three GCM simulations) were averaged to generate a final summary prediction. Then, to create dispersal networks for once again spatially filtered occurrence records in the present and the Last Glacial Maximum, the predictions under present and Last Glacial Maximum bioclimatic conditions were inverted for use as friction layers, i.e. areas of high suitability were converted to areas of low dispersal cost. The least-cost corridors (LCC) were calculated in three classes (the lowest 1 % LCPs, lowest 2 % LCPs, and lowest 5 % LCPs) by percentage of “the least-cost path (LCP) values” method. The weighed

values were applied to each class of LCCs (values of 5, 2, and 1 to low, mid, and high classes, respectively), because all LCCs were summed to create dispersal networks (Chan et al. 2011, Brown 2014).

Except that the predictions under present and Last Glacial Maximum bioclimatic conditions were obtained by the software MaxEnt, all analyses were conducted using the software SDMtoolbox version 1.1b (Brown 2014, [www.sdmtoolbox.org](http://www.sdmtoolbox.org)) implemented in the software ArcGIS version 10.1 ([www.arcgis.com](http://www.arcgis.com)).

## Results

### Phylogeography

In this study, we focused on previously published multilocus mtDNA dataset of the Eurasian green woodpecker. In total, we detected six unique haplotypes for the *cyt-b* gene and 24 unique haplotypes for the *ND2* gene. Seventeen variable sites occurred in the *cyt-b* gene, six of which were parsimony informative. Eight of transitions were A/G, eight were C/T, and only one transversion was detected, which was A/T. The *ND2* gene contained more variable sites than the *cyt-b* gene. 35 variable sites were found, 11 of which were parsimony informative. 13 of transitions were A/G, 21 were C/T, and only one transversion was detected, which was C/A.



**Table 1.** Relative posterior probabilities with 95 % credibility intervals for each scenario.

	Scenario 1		Scenario 2	
2000	0.9991	(0.9978, 1.0000)	0.0009	(0.0000, 0.0022)
4000	0.9988	(0.9976, 1.0000)	0.0012	(0.0000, 0.0024)
6000	0.9985	(0.9974, 0.9997)	0.0015	(0.0003, 0.0026)
8000	0.9985	(0.9972, 0.9997)	0.0015	(0.0003, 0.0028)
10000	0.9985	(0.9973, 0.9997)	0.0015	(0.0003, 0.0027)
12000	0.9985	(0.9974, 0.9996)	0.0015	(0.0004, 0.0026)
14000	0.9985	(0.9974, 0.9995)	0.0015	(0.0005, 0.0026)
16000	0.9984	(0.9974, 0.9994)	0.0016	(0.0006, 0.0026)
18000	0.9984	(0.9974, 0.9994)	0.0016	(0.0006, 0.0026)
20000	0.9984	(0.9974, 0.9993)	0.0016	(0.0007, 0.0026)

Haplotypes were not geographically structured in neither the *cyt-b* gene nor the *ND2* gene (Fig. 1). The *ND2* gene had higher genetic diversity than the *cyt-b* gene. Estimated mutation rate of the *ND2* gene was almost 4.5 times faster than the *cyt-b* gene (for details see discussion). According to the variation within the *ND2* gene, genetic diversity of southern populations was higher than northern populations. Interestingly, however, a population from United Kingdom had relatively higher nucleotide diversity (see Table 1 in Perktas et al. 2011). Four populations in Europe (Austria, Balkans, Italy and United Kingdom) contained the most common haplotype of the *ND2* gene. However, northern Europe (Sweden), Caucasus, and Turkey did not contain this haplotype. Geographic origin of the most common haplotype of the *ND2* gene was south and almost mid-latitudes of Europe.

Haplotype network based on the *cyt-b* gene had low frequency haplotypes differed from the most common haplotype by one, two or three base pairs. A similar network pattern was found for the *ND2* gene. However, the only difference between both haplotype networks was the number of divergent haplotypes in the *ND2* gene. These haplotypes differed from the most common haplotype by up to six base pairs, and located in Italy, Turkey, Caucasus, and Russia; they made genetic diversity higher in southern part of distribution range of the Eurasian green woodpecker in Europe, Anatolia and Caucasus (Fig. 1).

Despite mostly low-level genetic variation between populations of the Eurasian green woodpecker in Europe, Anatolia and Caucasus (maximum divergence was 0.007 for the *cyt-b* gene and 0.0077 for the *ND2* gene), genetic distances were positively correlated with geographic distances for both mtDNA genes (Fig. 2), which indicate an evidence of isolation-by-distance (the *cyt-b* gene,  $r = 0.568$ ,  $P < 0.0001$ ; the *ND2* gene,  $r = 0.374$ ,  $P = 0.001$ ). All of these results led us to justify combining all sequences of two different mtDNA genes from all populations while assessing historical demography of the Eurasian green woodpecker. Extended Bayesian skyline plot showed a good resolution of effective population size fluctuations over the history of the Eurasian green woodpecker. Therefore, the result of the extended Bayesian skyline plot detected a clear population expansion after the Last Glacial Maximum (Fig. 3). Finally, statistical phylogeography (i.e. coalescent simulations) supported the result of the extended Bayesian skyline plot and suggested that the best explanation of population expansion was scenario 1 (Fig. 4), which indicated a population expansion after the Last Glacial Maximum. The posterior probability was 0.9984 for this scenario (Table 1). Based on the approximate posterior distribution of parameters under scenario 1, population expansion began about 14500 years before present and the effective population size increased from 215000 to 375000 (Table 2).

**Table 2.** Properties of the approximate posterior distribution of parameters under the scenario 1.

Parameter	Mean	Median	Mode	q (0.025)	q (0.075)
N1	4.08E + 05	3.75E + 05	2.70E + 05	1.44E + 05	8.25E + 05
t1	1.41E + 04	1.45E + 04	1.70E + 04	5.54E + 03	2.06E + 04
Na	2.77E + 04	2.15E + 04	7.30E + 03	1.19E + 03	8.67E + 04
t2	8.08E + 04	8.28E + 04	1.02E + 05	2.53E + 04	1.28E + 05
Nb	6.96E + 05	7.20E + 05	9.03E + 05	2.76E + 05	9.88E + 05



### Ecological Niche Modelling

The model performed better than a random prediction (AUC for training data (mean  $\pm$  standard deviation (SD), range) =  $0.928 \pm 0.001$ , 0.926-0.930; AUC for test data (mean  $\pm$  SD, range) =  $0.842 \pm 0.013$ , 0.822-0.861, based on the 10-fold cross-validation runs). The small SDs for the mean AUCs suggested that the model performance was robust to variation in the selection of the occurrence records for training and testing.

Under present bioclimatic conditions, areas of relatively high suitability were predicted across most of the known distribution of the Eurasian green woodpecker. However, the model underpredicted the distribution in the east of the Western Palearctic (e.g. Russia, Ukraine) and overpredicted the distribution in Iberia (Fig. 5). The former is most likely to be an effect of sampling bias (e.g. no sampling in the east of the Western Palearctic), although correction for sampling bias was applied (Fig. 5, Fig. S1), and/or of that the distribution is poorly described in that region. The latter is most likely to be an effect of biotic interaction, because a sister species (Iberian green woodpecker) is distributed in Iberia, i.e. two species show parapatric distributions suggestive of competitive interaction.

Under the Last Glacial Maximum bioclimatic conditions, areas of relatively high suitability were predicted across southern Europe (including France), Anatolia, and the Caucasus/Caspian region. Three predictions under three GCM simulations were largely similar in this respect (Fig. 5).

During present bioclimatic conditions, a relatively high degree of population connectivity was observed across most of the Western Palearctic, with the higher degree of connectivity from the northern Anatolia through the Balkans and central Europe to England and Scandinavia. During the Last Glacial Maximum bioclimatic conditions, a relatively high degree of population connectivity was observed across southern Europe (including France), Anatolia, and the Caucasus/Caspian region, with the higher degree of connectivity from the northwestern Anatolia through the Balkans to Italy (Fig. 6).

All these results suggested that the Eurasian green woodpecker survived the Last Glacial Maximum in a single large glacial refugium, encompassing southern Europe (including France), Anatolia, and the Caucasus/Caspian region, and, afterwards, substantially expanded its range from this LGM refugium to its present range (Fig. 5, 6).

### Discussion

In this study, we aimed to evaluate historical demography of the Eurasian green woodpecker

using phylogeography (for review see Avise 2000) and ecological niche modelling (Guisan & Thuiller 2005), and tested glacial refugia hypothesis (e.g. Brito 2005). Unlike previous studies of the Eurasian green woodpecker (Pons et al. 2010, Perktas et al. 2011), this study integrated two different approaches to infer robust demographic history of this species. Thus, our study exemplified the utility of integrating phylogeography and ecological niche modelling.

The glacial refugia hypothesis has stated that temperate zone species in Europe, particularly climate-sensitive species, harbored in a low-latitude refugium or refugia during the Last Glacial Maximum. In Europe, southern peninsulas (i.e. Iberia, Italia, Balkans, and Anatolia) served as refugial areas for many bird species (Merilä et al. 1997, Bensch & Hasselquist 1999, Griswold & Baker 2002, Brito 2005, Perktas et al. 2011, Perktas & Quintero 2013). Our results from phylogeography, together with ecological niche modelling, revealed additional new insights into the historical demography of the Eurasian green woodpecker. However, these results are not far apart from a phylogeographic pattern congruent with the glacial refugia hypothesis.

The genetic diversity statistics (i.e. Tajima's D and Fu's F) from two previous studies (Pons et al. 2010, Perktas et al. 2011) suggested that the Eurasian green woodpecker experienced a population expansion. In addition, our extended Bayesian skyline plot analysis showed that there is a clear recent expansion event after the Last Glacial Maximum, which is concordant with previous results. Due to a higher mutation rate in the *ND2* gene, extended Bayesian skyline plot analysis for this gene showed a clear expansion pattern for the Eurasian green woodpecker. As stated for grackles and allies (Johnson & Lanyon 1999), the *ND2* gene was evolving significantly faster than the *cyt-b* gene in the Eurasian green woodpecker. Taken together, these results supported recent historical colonization of Europe dating back to just after the Last Glacial Maximum. Therefore, they are concordant with phylogeographic studies of the three-toed woodpecker (*Picoides tridactylus*, Zink et al. 2002) and the great spotted woodpecker (*Dendrocopos major*, Perktas & Quintero 2013). These two woodpecker species experienced a rapid population expansion, which precluded genetic differentiation within each species. The lack of geographic differentiation in large continental regions suggested that woodpeckers possess the potential for rapid population and range expansion given suitable conditions. However, both haplotype networks showed that a more or less

continuous structure with low frequency haplotypes with various mutational changes might indicate a shallow gene tree. This pattern was thus in accordance with Avise's phylogeographic category IV (Avise 2000), which refer to shallow gene trees and sympatric lineages refer to expansion events from small or modest number of founders (e.g. Merilä et al. 1997). Therefore, this result might not provide enough evidence of multiple glacial refugia. In addition, the most common haplotype of both mtDNA markers, i.e. the one at the center in both haplotype networks, was only found in mid- or southern latitudes in Europe. This kind of haplotype network is usually accepted as a "star phylogeny", which is another assumption of the category IV. In star phylogeny, central haplotype is usually accepted as an ancestral haplotype and recent derivatives are connected by short branches (Freeland et al. 2011). Hence, locations that had higher percentage of occurrence of the most common haplotypes were indicative of ancestral area or refugium for the Eurasian green woodpecker. Also, a positive and significant isolation-by-distance pattern for this species provided an evidence of colonization event from a single refugium (Zigouris et al. 2013, Manthey et al. 2014, Perktas et al. 2015). Taking all these results together, the most likely hypothesis is the presence of a single large glacial refugium in mid- or southern latitudes in Europe for this species.

Inferences from phylogeography and historical demography analyses were consistent with those of ecological niche modelling. Ecological niche modelling results indicated a range expansion of the Eurasian green woodpecker in Europe, reflecting climatic oscillations in the Late Pleistocene, particularly noticeable after the Last Glacial Maximum. The importance of the Balkans and Italy as glacial refugia for northern temperate bird species during the Last Glacial Maximum has been well documented (e.g. Brito 2005, Pellegrino et al. 2014, 2015), and is supported here by ecological niche modelling of the Eurasian green woodpecker. The unique insight that is specific for this study is a single large glacial refugium in southern Europe (including southern France and

perhaps southern United Kingdom) for the Eurasian green woodpecker. Our landscape connectivity analysis suggested a higher than moderate connectivity among southern populations of the Eurasian green woodpecker. This result was yet another indication of a single large glacial refugium in this region. A similar result was shown in the Leisler's bat, *Nyctalus leisleri* (Boston et al. 2015), and the mole, *Talpa europaea* (Feuda et al. 2015). In the case of the Leisler's bat, the star phylogeny indicates recent expansion experience of this species. All European haplotypes of this species were closely related; and the most common mtDNA haplotype, which was probably the ancestral one, occurred in continental Europe, England, and Ireland within its distribution range. Moreover, our ecological niche modelling analysis was concordant with those of phylogeography of the Leisler's bat, which indicated the expansion from cryptic refugia located in along the coastal line of Europe together with the refugia located in southern Europe during the Last Glacial Maximum. In the case of the mole, on the other hand, a large single refugium located all along southern European belt from southern France to Balkans and northern coastal line of the Black Sea. Consequently, the demographic history of these two species has a distinctive signal of cryptic and/or extra-Mediterranean refugia (Schmitt & Varga 2012) in southern and western coast of Europe, which is almost complete match with the demographic history of the Eurasian green woodpecker. These results deserve to be researched more deeply yet, because it is plausible to say that European biogeography is more complex than previously thought.

## Acknowledgements

We are grateful to Prof. Jan Zima and Lenka Glosová to their help during the review process of this manuscript. We also extend our sincerest thanks to three referees who significantly improved this manuscript. Author contribution: U. Perktas and H. Gür conceived and designed this study, U. Perktas conducted phylogeographic analyses, H. Gür conducted ecological niche modelling analyses, E. Ada helped both phylogeographic and ecological niche modelling analyses, and U. Perktas and H. Gür prepared and edited the manuscript.

## Literature

- Avise J.C. 2000: Phylogeography: the history and formation of species. *Harvard University Press, Cambridge*.
- Bandelt H.J., Forster P. & Rohlf A. 1999: Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16: 37–48.
- Bensch S. & Hasselquist D. 1999: Phylogeographic population structure of great reed warblers: an analysis of mtDNA control region sequences. *Biol. J. Linn. Soc.* 66: 171–185.
- Boria R.A., Olson L.E., Goodman S.M. & Anderson R.P. 2014: Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecol. Model.* 275: 73–77.
- Boston E.S.M., Montgomery W.I., Hynes R. & Prodöhl P.A. 2015: New insights on postglacial colonisation in Western Europe: the phylogeography of the Leisler's bat (*Nyctalus leisleri*). *Proc. R. Soc. Lond. B* 282: 20142605.

- Brito P. 2005: The influence of Pleistocene glacial refugia on tawny owl genetic diversity and phylogeography in western Europe. *Mol. Ecol.* 14: 3077–3094.
- Brown J.L. 2014: SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods Ecol. Evol.* 5: 694–700.
- Chan L.M., Brown J.L. & Yoder A.D. 2011: Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Mol. Phylogenet. Evol.* 59: 523e537.
- Cornuet J.M., Santos F., Beaumont M.A., Robert C.P., Marin J.M., Balding D.J., Guillemaud T. & Estoup A. 2008: Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24: 2713–2719.
- Cramp S. (ed.) 1985: The birds of Western Palearctic, vol. IV: terns to woodpeckers. *Oxford University Press, Oxford*.
- de Juana E. & Garcia E. 2015: The birds of the Iberian Peninsula. *Bloombury, U.K.*
- del Hoyo J., Collar N.J., Christie D.A., Elliott A. & Fishpool L.D.C. 2014: HBW and BirdLife International illustrated checklist of the birds of the world. *Lynx Edicions and BirdLife International Barcelona, Spain and Cambridge, U.K.*
- Dickinson E.C. 2003: The Howard and Moore complete checklist of the birds of the world, 3<sup>rd</sup> edition. *Christopher Helm, London*.
- Dormann C.F., Elith J., Bacher S., Buchmann C., Carl G., Carré G., Marquéz J.R.G., Gruber B., Lafourcade B., Leitão P.J., Münkemüller T., McClean C., Osborne P.E., Reineking B., Schröder B., Skidmore A.K., Zurell D. & Lautenbach S. 2013: Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36: 27–46.
- Drummond A.J., Suchard M.A., Xie D. & Rambaut A. 2012: Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29: 1969–1973.
- Elith J., Graham C.H., Anderson R.P., Dudík M., Ferrier S., Guisan A., Hijmans R.J., Huettmann F., Leathwick J.R., Lehmann A., Li J., Lohmann L.G., Loiselle B.A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J.M., Peterson A.T., Phillips S.J., Richardson K., Scachetti-Pereira R., Schapire R.E., Soberón J., Williams S., Wisz M.S. & Zimmermann N.E. 2006: Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129–151.
- Elith J., Kearney M. & Phillips S.J. 2010: The art of modelling range-shifting species. *Methods Ecol. Evol.* 1: 330–342.
- Elith J., Phillips S.J., Hastie T., Dudík M., Chee Y.E. & Yates C.J. 2011: A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* 17: 43–57.
- Feuda R., Bannikova A.A., Zemlemerova E.D., Febbraro F.D., Loy A., Hutterer R., Aloise G., Zykov A.E., Annesi F. & Colangelo P. 2015: Tracing the evolutionary history of the mole, *Talpa europaea*, through mitochondrial DNA phylogeography and species distribution modelling. *Biol. J. Linn. Soc.* 114: 495–512.
- Fielding A.H. & Bell J.F. 1997: A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24: 38–49.
- Fourcade Y., Engler J.O., Rodder D. & Secondi J. 2014: Mapping species distributions with MaxEnt using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PLoS ONE* 9 (5): e97122.
- Freeland J.R., Kirk H. & Petersen S. 2011: Molecular ecology. *John Wiley and Sons, Chichester*.
- Gassó N., Thuiller W., Pino J. & Vilà M. 2012: Potential distribution range of invasive plant species in Spain. *NeoBiota* 12: 25–40.
- Griswold C. & Baker A. 2002: Time to the most recent common ancestor and divergence times of populations of common chaffinches (*Fringilla coelebs*) in Europe and North Africa: insights into Pleistocene refugia and current levels of migration. *Evolution* 56: 143–153.
- Guisan A. & Thuiller W. 2005: Predicting species distribution: offering more than simple habitat models? *Ecol. Lett.* 8: 993–1009.
- Gür H. 2013: The effects of the late Quaternary glacial-interglacial cycles on Anatolian ground squirrels: range expansion during the glacial periods? *Biol. J. Linn. Soc.* 109: 19–32.
- Habel J.C., Rödder D., Schmitt T. & Nève G. 2011: Global warming will affect the genetic diversity and uniqueness of *Lycaena helle* populations. *Glob. Change Biol.* 17: 194–205.
- Heled J. & Drummond A.J. 2008: Bayesian inference of population size history from multiple loci. *BMC Evol. Biol.* 8: 289.
- Hewitt G.M. 1996: Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* 58: 247–276.
- Hewitt G.M. 1999: Post-glacial re-colonization of European biota. *Biol. J. Linn. Soc.* 68: 87–112.
- Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G. & Jarvis A. 2005: Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25: 1965–1978.
- Jiménez-Valverde A., Nakazawa Y., Lira-Noriega A. & Peterson A.T. 2009: Environmental correlation structure and ecological niche model projections. *Biodiversity Informatics* 6: 28–35.
- Johnson K.P. & Lanyon S.M. 1999: Molecular systematics of the grackles and allies, and the effect of additional sequence (cyt b and ND2). *Auk* 116: 759–768.
- Korkmaz E.M., Lunt D.H., Çıplak B., Değerli N. & Başıbuyuk H.H. 2014: The contribution of Anatolia to European phylogeography: the centre of origin of the meadow grasshopper, *Chorthippus parallelus*. *J. Biogeogr.* 41: 1793–1805.
- Librado P. & Rozas J. 2009: DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
- Manthey J.D., Klicka J. & Spellman G.M. 2014: Effects of climate change on the evolution of brown creeper (*Certhia americana*) lineages. *Auk* 131: 559–570.
- Merilä J., Björklund M. & Baker A.J. 1997: Historical demography and present day population structure of the greenfinch (*Carduelis chloris*) – an analysis of mtDNA control-region sequences. *Evolution* 51: 946–956.
- Pavlova A., Zink R.M., Drovetski S.V., Red'kin Y. & Rohwer S. 2003: Phylogeographic patterns in *Motacilla flava* and *Motacilla citreola*: species limits and population history. *Auk* 120: 744–758.
- Pellegrino I., Negri A., Boano G., Cucco M., Kristensen T.N., Pertoldi C., Randi E., Šálek M. & Mucci N. 2015: Evidence for strong genetic structure in European populations of the little owl *Athene noctua*. *J. Avian Biol.* 46: 462–475.

- Pellegrino I., Negri A., Cucco M., Mucci N., Pavia M., Šálek 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.
- Perktas U., Barrowclough G.F. & Groth J.G. 2011: Phylogeography and species limits in the green woodpecker complex (Aves: Picidae): multiple Pleistocene refugia and range expansion across Europe and the Near East. *Biol. J. Linn. Soc.* 104: 710–723.
- Perktas U., Gür H., Sağlam İ.K. & Quintero E. 2015: Climate-driven range shifts and demographic events over the history of Kruper's nuthatch *Sitta krueperi*. *Bird Study* 62: 14–28.
- Perktas U. & Quintero E. 2013: 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.
- Phillips S.J., Anderson R.P. & Schapire R.E. 2006: Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190: 231–259.
- Phillips S.J. & Dudík M. 2008: Modeling of species distributions with MaxEnt: new extensions and a comprehensive evaluation. *Ecography* 31: 161–175.
- Pons J.M., Olmos G., Cruaud C. & Fuchs J. 2010: Phylogeography of the Eurasian green woodpecker (*Picus viridis*). *J. Biogeogr.* 38: 311–325.
- Richmond O.M.W., McEntee J.P., Hijmans R.J. & Brashares J.S. 2010: Is the climate right for Pleistocene rewilding? Using species distribution models to extrapolate climatic suitability for mammals across continents. *PLoS ONE* 5: e12899.
- Rousset F. 1997: Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145: 1219–1228.
- Ruiz-González A., Madeira M.J., Randi E., Abramov A.V., Davoli F. & Gómez-Moliner B.J. 2013: Phylogeography of the forest-dwelling European pine marten (*Martes martes*): new insights into cryptic northern glacial refugia. *Biol. J. Linn. Soc.* 109: 1–18.
- Schmitt T. & Varga Z. 2012: Extra-Mediterranean refugia: the rule and not the exception? *Front. Zool.* 9: 22.
- Shcheglovitova M. & Anderson R.P. 2013: Estimating optimal complexity for ecological niche models: a jackknife approach for species with small sample sizes. *Ecol. Model.* 269: 9–17.
- Slatkin M. 1993: Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47: 264–279.
- Soberón J. & Peterson A. 2005: Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics* 2: 1–10.
- Svendsen J.I., Alexanderson H., Astakhov V.I., Demidov I., Dowdeswell J.A., Funder S., Gataullin V., Henriksen M., Hjort C., Houmark-Nielsen M., Hubberten H.W., Ingolfsson O., Jakobsson M., Kjær K.H., Larsen E., Lokrantz H., Lunkka J.P., Lysa A., Mangerud J., Mantioukhov A., Murray A.S., Möller P., Niessen F., Nikolskaya O., Polyak L., Saarnisto M., Siegert C., Siegert M.J., Spielhagen R.F. & Stein R. 2004: Late Quaternary ice sheet history of northern Eurasia. *Quat. Sci. Rev.* 23: 1229–1271.
- Tamura K. & Nei M. 1993: Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10: 512–526.
- Tamura K., Stecher G., Peterson D., Filipinski A. & Kumar S. 2013: MEGA6: Molecular Evolutionary Genetics Analysis, version 6.0. *Mol. Biol. Evol.* 30: 2725–2729.
- Templeton A.R., Routman E. & Phillips C. 1995: Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140: 767–782.
- Tucker G.M. & Heath M.F. 1994: Birds in Europe: their conservation status. *BirdLife International (Conservation Series No. 3)*, Cambridge, U.K.
- Weir J.T. & Schluter D. 2008: Calibrating the avian molecular clock. *Mol. Ecol.* 17: 2321–2328.
- Winkler H., Christie D.A. & Nurney D. 1995: Woodpeckers: a guide to the woodpeckers, piculets and wrynecks of the world. *Pica Press, Robertsbridge, U.K.*
- Zigouris J., Schaefer J.A., Fortin C. & Kyle C.J. 2013: Phylogeography and post-glacial recolonization in wolverines (*Gulo gulo*) from across their circumpolar distribution. *PLoS ONE* 8: e83837.
- Zink R.M., Rohwer S., Drovetski S., Blackwell-Rago R.C. & Farrell S.L. 2002: Holarctic phylogeography and species limits of three-toed woodpeckers. *Condor* 104: 167–170.

#### Supplementary online materials

**Fig. S1.** The bias grid used for ecological niche modelling to up-weight the occurrence records with fewer neighbors in the geographic landscape, scaled from 0 (blue) to 10 (red) (for detailed information, see Elith et al. 2010). The red circles indicate the occurrence records (URL: [http://www.ivb.cz/folia/download/perktas\\_et\\_al.\\_fig.\\_s1.jpg](http://www.ivb.cz/folia/download/perktas_et_al._fig._s1.jpg)).

**Fig. S2.** Comparison of climates across time periods, using bioclimatic variables used for ecological niche modelling and the MESS method (for detailed information, see Elith et al. 2010). Brown indicates areas having one or more bioclimatic variables outside the range present in the training data (URL: [http://www.ivb.cz/folia/download/perktas\\_et\\_al.\\_fig.\\_s2.tif](http://www.ivb.cz/folia/download/perktas_et_al._fig._s2.tif)).