Open access • Journal Article • DOI:10.1111/ZSC.12320

Significant Asia-Europe divergence in the middle spotted woodpecker (Aves, Picidae) — Source link

Laura Kamp, Laura Kamp, Gilberto Pasinelli, Pietro Milanesi ...+5 more authors

Institutions: University of Bern, Naturhistorisches Museum, National Museum of Natural History, Adam Mickiewicz University in Poznań ...+1 more institutions

Published on: 01 Jan 2019 - Zoologica Scripta (John Wiley & Sons, Ltd)

Topics: Middle spotted woodpecker and Divergence

Significant Asia-Europe divergence in the middle spotted woodpecker (Aves, Picidae)

Corresponding author:
Manuel Schweizer, Naturhistorisches Museum der Burgergemeinde Bern, Bernastrasse 15, CH-3005 Bern, Switzerland, 0041 350 72 83, manuel.schweizer@nmbe.ch

Significant Asia-Europe divergence in the middle spotted woodpecker (Aves: Picidae: *Dendrocoptes medius*)

LAURA KAMP, GILBERTO PASINELLI, PIETRO MILANESI, SERGEI V. DROVETSKI, ZIEMOWIT KOSINSKI, SERGUEI KOSSENKO, HUGO ROBLES, MANUEL SCHWEIZER

Running title: Phylogeography of the middle spotted woodpecker

Laura Kamp et al.

Abstract

Kamp, L. , Pasinelli, G., Pietro Milanesi P., Drovetski, S.V., Kosiński, Z., Kossenko, S., Robles, H., Schweizer, M. (2018) Significant Asia-Europe divergence in the middle spotted woodpecker (Aves: Picidae: *Dendrocoptes medius*). *ZoologicaScripta*, 00, 000-000.

Population divergence could be strongly affected by species' ecology and might not be a direct response to climate driven environmental change. We tested this in the middle spotted woodpecker (*Dendrocoptes medius*), a non-migratory, low-dispersal, habitat specialist associated with old deciduous forests of the Western Palearctic. We present the first phylogeographic study of this species integrating genetic data (three mitochondrial loci, one autosomal, and one Z-linked intron) with species distribution modelling. Based on this species' ecology we predicted that the middle spotted woodpecker could have colonized its current range from multiple Last Glacial Maximum (LGM) refugia and that strongly structured populations could be expected. Indeed, we discovered a strong genetic divergence between Asian and European populations, with a split estimated at around one million of years ago. This was surprising given only slight intraspecific variation in plumage and morphology. Although there was no significant phylogeographic structure within the Asian and European groups, we cannot exclude the possibility of multiple refugia within either group during the LGM. This has to be further investigated with more extensive geographic sampling and larger number of variable independently evolving markers. Future studies should also investigate potential differences in vocalizations and ecology between the two groups. Lineages showing similar level of genetic differentiation including woodpeckers are often treated as species-level taxa. Comparison of our results with the phylogeographic history of other woodpeckers, suggests that sympatric species with similar life-histories might have idiosyncratic phylogeographic patterns probably resulting from different ecological requirements or historic stochasticity.

Laura Kamp[1,2,3], Gilberto Pasinelli[3], Pietro Milanesi[3], Sergei V. Drovetski[4], Ziemowit Kosiński[5], Serguei Kossenko[6]
Hugo Robles[7], Manuel Schweizer[1]

[1] Naturhistorisches Museum der Burgergemeinde Bern, Bernastrasse 15, CH-3005 Bern, Switzerland

[2] Universität Bern, Institute of Ecology and Evolution, Baltzerstrasse 6,CH- 3012 Bern, Switzerland

[3] Schweizerische Vogelwarte, Seerose 1, CH-6204 Sempach, Switzerland

[4]Laboratories of Analytical Biology, National Museum of Natural History, Smithsonian Institution, 1000 Constitution Ave NW, Washington, DC, 20004, USA

[5]Department of Avian Biology and Ecology, Institute of Environmental Biology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland

[6]State Nature Biosphere Reserve "Bryansky Les", Nerussa Station, Suzemka District, Bryansk Region, 242180, Russia

[7]Evolutionary Ecology Group (EVECO), University of Antwerp, Campus Drie Eiken, Universiteitsplein 1, BE-2610, Wilrijk, Belgium

Introduction

Landscape alterations triggered by the Pleistocene climate change had a strong effect on biodiversity patterns observed today. These were particularly severe during the last one million of years when climate cycles shifted from a 41 kyr to a 100 kyr cycle and increased in amplitude leading to an accentuation of glacial and interglacial periods (e.g. G. Hewitt, 2000; G. M. Hewitt, 1999). As a consequence, repeated periods of range contractions fostering population fragmentation and divergence followed by range expansion and secondary contact were prevalent in many species. However, such population processes were generally not a direct response to environmental change, but could be strongly related to species' ecology (Hung *et al.*, 2017; Zamudio *et al.*, 2016). Many Holarctic temperate and boreal plants and animals showed range contraction and could survive in only one or a few refugia in the south of the continent during cold and arid glacials, but expanded their ranges and colonized areas to the north during warmer and wetter interglacials (e.g. G. Hewitt, 2000; Taberlet *et al.*, 1998; Weir & Schluter, 2004). In contrast, the ranges of arid-adapted species in ice free areas contracted during interglacials and expanded during glacials (Alaei Kakhki *et al.*; Garcia, Alda*, et al.*, 2011; Garcia, Mañosa*, et al.*, 2011). Some species also could have survived the glacial maxima at higher temperate latitudes in so-called 'cryptic refugia' (Provan & Bennett, 2008).

Population diversification might occur faster in ecological specialists as a consequence of smaller effective population sizes, mainly due to patchy distributions and repeated population bottlenecks caused by their high sensitivity to climate related environmental changes (Hung *et al.*, 2017; Li *et al.*, 2014). Population diversification might also be accelerated in resident species and those with low dispersal abilities due to reduced gene flow among populations and thus smaller effective populations sizes compared to migratory species or species with high dispersal (Arguedas & Parker, 2000; Hung *et al.*, 2017; Smith *et al.*, 2014). By comparing 30 forest-dwelling passerines in a Pleistocene forest refugium in the Caucasus with their European counterparts, Hung *et al.* (2017) indeed revealed habitat specialists to be generally more divergent compared to generalists as well as residents compared to migrants.

Hewitt (2000) emphasized the need for phylogeographic studies of ecologically diverse species in different geographical settings to understand the influence of past climate changes on biodiversity dynamics. Here, we investigate the phylogeography of the middle spotted woodpecker *Dendrocoptes medius* (Gill & Donsker, 2018; Fuchs & Pons, 2015), a non-migratory, low-dispersal, habitat specialist (Pasinelli, 2003) that lives in old deciduous forests across large parts of Europe, Anatolia, the Caucasus, and in an isolated area in the Zagros Mountains of Iran and Iraq (Winkler, Christie, Kirwan*, et al.*, 2018) (Fig. 1). Four recognized subspecies (Gill & Donsker, 2018) differ slightly in size and color patterns: *D. m. medius* in continental Europe*; D. m. caucasicus* in Northern Turkey, the Caucasus, Transcaucasia

and probably in North-West Iran*; D. m. anatoliae* in Western and Southern Anatolia; and *D. m. sanctijohannis* in the Zagros Mountains in South-Western Iran and North-East Iraq.

The middle spotted woodpecker primarily feeds on arthropods and is considered a habitat specialist depending on rough-barked deciduous trees with a high arthropod richness like old oaks *Quercus sp.* (Ciudad *et al.*, 2009; Jenni, 1983; Kosinski, 2006; Pasinelli & Hegelbach, 1997; Pettersson, 1983; H. Robles *et al.*, 2007). Trees suitable for cavity excavation are important resources for this woodpecker, since it typically excavates a new cavity each year and rarely reuses old cavities (Hebda *et al.*, 2016; Kosinski *et al.*, 2018; Kosinski & Winiecki, 2004; Pasinelli, 2000, 2007). Its breeding success is affected by weather conditions during the nestling period most likely through food abundance (Kossenko & Kaygorodova, 2007; Pasinelli, 2001). The middle-spotted woodpecker has a strongly sedentary life-style with relatively low dispersal capacity (Ciudad *et al.*, 2009; Kossenko, 2003; Pasinelli, 2001, 2003). Its patch occupancy dynamic is inversely related to habitat loss and fragmentation (Bühlmann & Pasinelli, 2012; Kosenko & Kaigorodova, 2001; Müller, 1982; Hugo Robles & Ciudad, 2012, 2017). Collectively, these ecological and behavioral traits likely make the middle spotted woodpecker highly vulnerable to environmental and climate change. Therefore, it could have been strongly affected by Pleistocene climate fluctuations and given its low dispersal capacity, we would expect that it has colonized its current relatively large range from more than one refugium after the last glacial maximum (LGM). Consequently, strongly genetically structured populations could be expected across its range. In an alternative scenario, its range could have been reduced to a single refugium and its current distribution area rapidly colonized after LGM. Only shallow genetic structure would be expected in the latter case and the described slight geographic phenotypic variation might reflect local adaptions to contemporarily different environments.

To test these hypotheses we conduct the first phylogeographic study of the middle spotted woodpecker integrating genetic data with species distribution modelling to gain independent information on the location of climatically suitable areas which could have acted as refugia for this species during LGM.

Materials and Methods

Samples and laboratory work

We used 87 samples of the middle spotted woodpecker represented primarily by blood and tissue samples, with a few toe pads from museum specimens and feathers also included (Table S1). The sampling covered the entire range of middle spotted woodpecker, except the Apennine Chain. The species range was divided into eight geographic areas represented by 5 - 20 samples each (Fig. 1, Table

S1). The great spotted woodpecker *Dendrocopos major* was chosen as outgroup because sequences for all markers used were available for this species in GenBank (Table S1). We sequenced two nuclear introns: z-linked Brahma protein 15 intron (BRM) and autosomal beta-Fibrinogen intron 7 (Fib7), and three mitochondrial loci: Control Region 1 D-Loop (ConReg), ATPase subunit 6 (ATP6), and Cytochrome b (CytB). Only two mitochondrial genes (ATP6 and CytB) were sequenced for the toe pad samples of the old museum specimens (Table S1) using two smaller overlapping fragments of about 260-320 base pairs long. We used either previously reported PCR primers or designed them using PRIMER3 0.4.0 (Rozen & Skaletsky, 2000) as implemented in GENEIOUS 8.1.6 (Kearse *et al.*, 2012) (Table S2). DNA extraction methods and PCR protocols are described in the "Laboratory Protocol" Appendix. Sequencing of the PCR products in both directions was performed by LGC Genomics (Berlin) using PCR primers.

Phylogeographic analysis

The sequences were visually examined in GENEIOUS, then trimmed and aligned with MAFFT 1.3.1 (Katoh & Standley, 2013) for each marker separately. For coding markers, alignments were translated into amino acids and checked for stop codons. Alleles in Fib7 were inferred in PHASE 2.1 (Stephens & Scheet, 2005; Stephens *et al.*, 2001) using default settings with input file generated with the online software SEQPHASE (http://seqphase.mpg.de/seqphase/) ((FLOT, 2010). To visualize genetic variation, a median joining haplotype networks were constructed for each marker separately with POPART (Leigh and Bryant, 2015) using the default settings (epsilon =0). As PopArt excludes positions with heterozygous individuals, we phased the seven individuals with heterozygous positions in BRM (see below) to show the overall geographic distribution of alleles. As no individual showed more than one heterozygous position, this was done manually. Now sex information was available for the remaining individuals and they were not phased consequently. A standard McDonald-Kreitman test performed online (http://bioinf3.uab.cat/mkt/MKT.asp) for ATP6 and Cytb for the two groups (see below) did not reveal any significant deviation from neutrality.

Standard genetic diversity indices: the number of polymorphic sites, gene and nucleotide diversities, were computed for each locus, except BRM due to its insufficient variability, for each geographic area: Spain, Central Europe, Balkans, Western Russia, Anatolia including the Lesvos Island, Greater Caucasus, Armenia, and Zagros (Figure 1) in ARLEQUIN 3.5.2 (Excoffier & Lischer, 2010). To estimate differentiation between each pair of geographic areas, we calculated FST (ΦST) values using 10000 permutations in ARLEQUIN to test for significance. To examine genetic structure within and between the eastern and western area groups (see below) we conducted an analysis of molecular variance (AMOVA; (Excoffier *et al.*, 1992) in ARLEQUIN using 10000 permutations for significance testing.

We reconstructed a dated gene tree for the three mtDNA markers in BEAST 2.4.3 (Bouckaert *et al.*, 2014) with unlinked site models and clock rates, but linked tree models. PARTITIONFINDER (LANFEAR *ET AL.*, 2012) was used to select best-fitting models of nucleotide evolution for each locus**.** A relaxed uncorrelated lognormal distribution was used for the clock model. A lognormal distribution with standard deviation of 0.05 was used as prior distribution for the clock rates with the following means (in real space) for the different markers based on published substitution rates taken from Lerner et al. (2011): ATP6 0.026 substitution/site/Million years (My), CytB 0.014 substitution/site/My, ConReg 0.026 substitution/site/My. Both, a coalescent exponential population and a coalescent constant population tree model were independently tested and their model-fit was compared with AICM (Akaike information criteria in Bayesian Monte Carlo context) comparison in TRACER 1.6 (Rambaut *et al.*, 2013) with 1000 bootstrap-replicates. For each tree model tested, three independent chains were run for 30 million generations and trees were sampled every 1000 generations. Tracer was used to assess parameter convergence among the three independent runs and to confirm appropriate burn-in and adequate effective sample sizes (ESS) of the posterior distribution. The three runs were then combined with 10% burn-in each using LogCombiner v. 2.4.3 (Bouckaert *et al.*, 2014). A maximum clade credibility tree was generated in TREEANNOTATOR 2.3.3 (Bouckaert *et al.*, 2014) with mean heights set for node heights and visualized in FigTree v. 1.4.2 (Rambaut, 2008).

We used *BEAST to estimate a dated species tree based on the three mtDNA markers and phased alleles of the autosomal Fib7, implementing unlinked sites models and clock rates, but a linked tree model for the mtDNA markers. BRM was not used in this analysis due to the lack of geographic structure and minimal variability (see below). The eight geographic sampling areas: Spain, Central Europe, Balkans, Western Russia, Anatolia including Lesvos Is., Greater Caucasus, Armenia, and Zagros (Figure 1) were treated as “taxa” in the species tree reconstruction. PARTITIONFINDER was used to estimate the best-fitting models of nucleotide evolution for Fib7. A relaxed uncorrelated lognormal distribution was implemented and the clock rates (ucld.mean parameters) were estimated using the same rate priors for the mtDNA markers and a substation rate of 0.0019 substitution/site/My for Fib7 (Lerner *et al.*, 2011) implemented again as a mean (in real space) of a lognormal distribution with a standard deviation of 0.05. We implemented a Yule processes as the species tree prior and different priors for the population function (constant, linear, linear with constant root) leading to three models. For each model, the MCMC analysis was run three times independently for 100 million generations with sampling every 1000 generations. Tracer was used to compare posterior distributions for all parameters to assess convergence among the independent runs and to confirm appropriate burn-in and the adequate effective sample sizes (ESS) for each parameter. The three independent runs were combined with 10% burn-in each using LogCombiner. Tracer was used to evaluate the best-fitting model using AICM comparisons based on 1000 bootstrap replicates. The resulting maximum clade

credibility species tree and the 95% highest posterior density (HPD) distributions of estimated node ages were calculated using TreeAnnotator based on the best-fitting model and visualized in FigTree. A cloudogram illustrating the uncertainty of the sampled topologies over the posterior distribution was generated in DensiTree 2.2.5 (Bouckaert & Heled, 2014).

To reconstruct the effective population sizes along a timescale for the eastern and western groups separately (see below), we constructed extended Bayesian Skyline plots (Drummond *et al.*, 2005) for the three mtDNA markers in BEAST using unlinked sites models and clock rates, but a linked tree model and the same priors as above for substitution models and clock rates. We initially included Fib7 (with a separate tree model), however, due to the lack of convergence in some parameters, we present our analyses of only the three mtDNA loci. For each group, MCMC analysis was run three times independently for 100 million generations with sampling every 1000 generations. Comparison of posterior distributions, assessment of convergence among the independent runs and the adequate effective sample sizes (ESS) for each parameter and combination of runs was done as above with 10% burn-in each. The parameter sum(indicators.alltrees) was used to check for the most likely numbers of demographic changes. Resulting data were analyzed and plotted using R scripts available at: https://github.com/CompEvol/beast2/tree/master/doc/tutorials/EBSP.

Species distribution modeling

Predictor variables

Initially, we considered 21 predictor variables (Table S3) representing current topographic and climatic conditions for which continuous spatial data were available for the entire study area. Notably, two topographic variables derived from a Digital Elevation Model (DEM) with a spatial resolution of 15 m (ASTER GDEM; gdem.ersdac.jspacesystems.or.jp) and 19 bioclimatic predictors derived from the WorldClim current dataset (www.worldclim.org) for the average values from 1950–2000 at a 2.5 min resolution (≈ 5 km; the highest resolution available for the last glacial maximum conditions, see below) were included. The two topographic variables were resampled at the same resolution as bioclimatic predictors.

We checked for multicollinearity among predictor variables through the Variance Inflation Factor (VIF)(Zuur *et al.*, 2010). For a given predictor, we calculated VIF by regressing it against the others and included the resulting R^2 values in the VIF formula given by Harrel (2001): $VIF = 1 / 1 - R_i^2$, where *i* is our target predictor. Following Zuur *et al.*, (2010) we considered predictors with VIF values > 3 as multicollinear with other predictors. Fourteen predictor variables were then removed and thus a total of seven predictors were considered in further analyses; namely altitude, slope, mean diurnal

range, temperature seasonality, mean temperature of wettest quarter, precipitation seasonality, and precipitation of the coldest quarter (Table S3).

The same set of variables (at the same spatial resolution) of current conditions was considered to predict distributions during the last glacial maximum (LGM, ≈ 22,000 years ago; Clark *et al.*, 2009) . Values for the variables altitude and slope were derived from NOOA ETOPO1 (www.ngdc.noaa.gov/mgg/global/relief/ETOPO1/data/ice_surface/grid_registered/georeferenced_tiff), while for paleoclimatic predictors, similarly to Ihlow et al. (2016), values were averaged over three different Global Climate Models (GCMs), namely CCSM4, MIROC-ESM and MPI-ESM-P (Hijmans *et al.*, 2005), obtained from the WorldClim past dataset.

Predictions of distributions under last glacial maximum conditions

We collected a total of 2,239 point locations from two online platforms: GBIF - the Global Biodiversity Information Facility (http://www.gbif.org/) and observation.org (www.observation.org; Figure S1). Based on the results of our genetic analysis showing two distinct intraspecific lineages (see below), we developed species distribution models (SDMs) combining locations in the range of the eastern and western groups separately.

The ensemble predictions (EPs) derived from 12 SDMs were estimated to assess distribution under current conditions, avoiding biased estimation due to single model uncertainty (Thuiller *et al.*, 2009). We averaged the predictions of: (i) artificial neural networks (ANN) (Ripley, 2007); (ii) boosted regression trees (BRT) (Friedman, 1991); (iii) classification tree analyses (CTA) (Breiman *et al.*, 1984); (iv) flexible discriminant analyses (FDA4) (Hastie *et al.*, 1994); (v) Gaussian Process SDMs using maximum *a posteriori* inference (GP-MAP) (Golding & Purse, 2016); (vi) Generalized Additive Models (GAM); (vii) Generalized Linear Models (GLM) (McCullagh & Nelder, 1989); (viii) factorial decomposition of Mahalanobis distances (MADIFA) (Calenge *et al.*, 2008); (ix) multivariate adaptive regression splines (MARS) (Friedman, 1991); (x) maximum entropy algorithm (MAXENT) (Phillips *et al.*, 2006) (xi) maximum-likelihood model (MAXLIKE) (Royle *et al.*, 2012) and (xii) random forests (RF) (Breiman, 2001). Point localities (see above) were used as presence data and 10,000 random points were generated to serve as pseudo-absence data in SDMs.

We initially found evidence of spatial autocorrelation among residuals of models and, therefore, similar to Pasinelli et al. (2016), X and Y coordinates of species locations and their interaction in SDMs were also included. Thereafter, model residuals where no longer spatially autocorrelated.

The predictive accuracy of SDMs was evaluated carrying out 10-fold cross-validations, using a random subsample of 90% of the locations to calibrate the models and the remaining 10% to validate them (Thuiller *et al.*, 2009). We estimated (i) the area under the receiver operating characteristic curve (AUC), (ii) the true skills statistic (TSS) and (iii) the Boyce index (BI).

The resulting continuous ensemble prediction maps (ranging between 0 and 100, i.e. low and high probability of occurrence, respectively) were converted into binary maps (indicating presence/absence areas) considering threshold values estimated by maximizing the True Skill Statistics (TSS) (Allouche *et al.*, 2006); values higher and lower than the thresholds represent suitable and unsuitable areas, respectively. The same threshold values were used to estimate the distribution in the LGM, projecting current SDMs into past conditions. We carried out all these analyses using the open-source software R 3.1.2 (http://www.R-project.org) and the packages `adehabitat' (Calenge, 2006), `biomod2' (Thuiller *et al.*, 2016), `ecospat' (Di Cola *et al.*, 2017), `GRaF' (Golding & Golding, 2014) and `maxlike' (Chandler & Royle, 2013).

Results

Phylogeographic analyses

The alignments consisted of 617 base pairs (bp) for ATP6 (number of samples n = 73), 452 bp for CytB (n = 83), 326 bp for ConReg (n = 68), 376 bp for BRM (n = 79) and 705 bp for Fib7 (n = 53; Table S1). When translated into amino acids, no coding sequences showed unexpected stop codons. There were three heterozygous positions in the sequences of Fib7 and two in BRM. Standard genetic diversity indices are given in Table S4.

The median joining networks for the mtDNA loci (ATP6, ConReg, CytB) revealed a deep divergence between eastern and western parts of the range roughly divided by a line running from the Sea of Marmara through the Sea of Azov. Sampling areas east and west of this line were separated by eight (ConReg) to 21 (ATP6) substitutions (Fig. 2 a-c). The total number of haplotypes ranged from six (CytB) to 11 (ConReg). The western groups had one common haplotype present in all populations and two (CytBb) to five (ATP6,ConReg) rare haplotypes divergent by one or two substitutions and mostly found in Central Europe and the Balkans. Spain had the lowest number of haplotypes across loci with only ConReg having more than one haplotype. The eastern group had five haplotypes in ConReg and three in each ATP6 and CytB. Only a haplotype was found in Anatolia in each mtDNA locus, while for the Greater Caucasus only a single haplotype was found in each ATP6 and CytB. Samples from Anatolia and the Greater Caucasus did not share any haplotypes. Accordingly, haplotype and nucleotide diversity varied among loci for each sampled geographic area. While Anatolia in the east and Spain in the west had the lowest mtDNA diversity, Armenia and the Zagros in the eastern group and Central Europe and the Balkans in the western group had the highest mtDNA diversity (Table S4). In Western Russia, diversity was high for ATP6 and ConReg but no variation was observed in Cytb.

Nuclear loci had little variation: Fib7 had four alleles (Fig. 2 d), however, eastern and western groups had divergent Fib7 alleles that differed by 3-4 substitutions. BRM had three alleles that differed by a single substitution each There was no geographic structure in BRM allele distribution (Fig. 2e). One allele was found in all geographic areas of the western group, one in all areas of both, eastern and western group except Western Russia, and one was restricted to one individual from Armenia.

High FST values significantly different from zero were found between the eastern and the western group in all markers (0.97 in ATP6, 0.97 in Cytb, 0.92 in ConReg, 0.95 in Fib7, $p<10^{-5}$). FST values between geographic areas within groups remained significantly different from zero after correction for multiple testing between the Greater Caucasus and all other areas of the eastern group for ATP6 ($p<0.006$), between Anatolia and Armenia or Greater Caucasus for ConReg ($p<4*10^{-4}$), and between Anatolia and all other eastern populations for Cytb ($p<4*10^{-4}$). The AMOVA showed that the vast majority of variance in the data was explained by differences between the eastern and western group (96% in ATP6, 97% in Cytb, 91% in ConReg, and 95% in Fib7). Differences among geographic areas within the groups (1.90% in ATP6, 1.85% in Cytb, 4.23% in ConReg, and 0.46% in Fib7) and among individual birds within areas (1.88% in ATP6, 1.52% in Cytb, 4.88% in ConReg, and 4.49% in Fib7) explained similarly small proportions of variance in the data.

In all BEAST analyses, ESS values for all parameters were >200 for all parameters and there was high convergence in parameter values among different runs. Substitution models evaluated for different loci are given in Table S5. A coalescence exponential tree prior was found as the best-fitting model. The maximum clade credibility mtDNA gene tree based on the three independent runs, revealed a deep and highly supported divergence between the populations east and west of a line running from the Sea of Marmara through the Black Sea and the Sea of Azov, that is the line effectively separating an Asian (including northern slope of the Greater Caucasus) and European clades (Fig. 3). The divergence time between the eastern and western clade in the dated trees was 1.42 Million of years ago (Ma) (95% highest posterior density (HPD): 0.93-1.99). The samples comprising the eastern clade coalesced at 0.13 Ma (95% HPD: 0.06-0.21), those comprising western clade at 0.11 Ma (95% HPD: 0.06-0.17). There was no robustly supported geographic structure within clades.

In the *BEAST analyses, a constant population function prior was revealed as the best-fitting model. However, there was no convergence in the species tree height parameter, and we thus considered a linear population function with a constant root as the best fitting model. After combining the three runs, ESS values were >200 for all parameters and there was high convergence in parameter values among the different runs. Our *BEAST maximum clade credibility species tree strongly supported divergence between the western and the eastern clade (Fig. 4). Their split was dated at 0.96 Ma (95% HPD: 0.1-1.66). None of the nodes within either clade was statistically supported. Because the

timeframe for these might be confounded by gene flow among the predefined geographic areas, they are not further discussed (Fig. 4).

A population expansion towards the present was indicated by the extended Bayesian Skyline Plots for both the eastern and the western groups (Fig. 5). While this expansion seems to have been rather recent in the last thousand years in the eastern group, there was a more steady increase for the western group in the last 50 kyr. However, the 95% HPD intervals were rather broad for both groups (Fig. 5). The median for the number of populations size change events was one for both groups with a 95% HPD intervals ranging from zero to three for the eastern group and from zero to four for the western group.

Present and past range modelling

The 10-fold cross-validations showed significant values for all evaluation methods of all SDMs for both the eastern and the western groups (Table S6). Area under the receiver operating characteristic curve (AUC) and the true skills statistic (TSS) values of ensemble predictions (EPs) for the eastern group (0.988 and 0.926, respectively) were higher than those for the western one (0.922 and 0.891, respectively) while the Boyce index (BI) value for the Western group (0.999) was higher than that for the Eastern group (0.952; Table S6). The prediction for the current distribution reflected the known breeding range overall; nevertheless areas predicted as suitable were found in regions outside of the actual range for both eastern and western groups (cf. Figures 1 and 6). Due to the expansion of ice sheets and permafrost, suitable areas during the LGM were predicted to be at the southern edge of the current range (Figure 7). For the eastern group, potential glacial refugia were primarily situated in coastal regions of southern Greece, Anatolia, along the eastern Mediterranean coast and at the Persian Gulf. For the western group, potential glacial refugia were located in southwestern France, the Apennine Peninsula and in some regions of the Balkan Peninsula (Figure 7).

Discussion

East-west divergence

In this study we conducted the first phylogeographic analysis of the middle spotted woodpecker - a Western Palearctic non-migratory, low-dispersal, habitat specialist that depends on rough barked deciduous trees like old oaks. The most striking result was the presence of a deep divergence between

European and Asian (including the North Caucasus) groups. While the European group is restricted to samples from the nominate subspecies, the Asian group comprises samples from the remaining three subspecies (*D. m. caucasicus*, *D. m. anatoliae,* and *D. m. sanctijohannis)*. In the mtDNA gene tree, this split was dated to around 1.5 Ma, whereas the species tree approach revealed a younger age at approximately one Ma. MtDNA gene coalescence dates are expected to overestimate divergence dates and have large confidence intervals (Drovetski *et al.*, 2015; Edwards & Beerli, 2000; Weir *et al.*, 2016) which might explain this discrepancy. The reliability of temporal aspects of phylogeographic reconstructions depends on the plausibility of the calibration of the applied substitution rate for the different markers. We utilized locus-specific rates calculated for Hawaiian Honeycreepers (members of Carduelinae) belonging to the order Passeriformes (Lerner *et al.*, 2011) that is not closely related to woodpeckers (Jarvis *et al.*, 2014; Prum *et al.*, 2015). This introduces uncertainty to our dating because of the lineage- and time-dependent variation in substitution rates (Ho *et al.*, 2011; Nabholz *et al.*, 2009). However, Lerner et al. (2011) found concordance of their rate estimates with published evolutionary rates from other bird groups indicating suitability of the rates across different avian orders, which was also suggested by Cibois et al. (2016).

The strong genetic differentiation between the European and Asian groups of the middle spotted woodpecker contrasts with the low level of morphological differences between the European and the three Asian subspecies (Winkler, Christie, Kirwan*, et al.*, 2018). The divergence occurred during the Calabrian stage of the Pleistocene, before or around the Early-Middle Pleistocene Transition 1.2 - 0.8 Ma when climatic cycles increased in duration and amplitude (Maslin *et al.*, 2014). This climate shift had a strong effect on distribution and diversification patterns of biota particularly in the Northern hemisphere (Head & Gibbard, 2005). Similar splits are often found between endemic bird lineages of the Caucasus and their European counterparts with the Caucasus being considered as an important Pleistocene forest refugium (Drovetski, 2003; Drovetski *et al.*, 2004; Hung *et al.*, 2012, 2016, 2017; Zink *et al.*, 2006). Divergence estimates between Caucasian and European populations of 30 forest-dwelling avian species ranged from 0-1.7 Ma with residents, habitat specialists, and species with low effective population sizes being more differentiated than migrants, habitat generalists, and abundant species (Hung *et al.*, 2017). The deep east-west divergence in the middle spotted woodpecker is thus consistent with its specialized ecology and resident life-history. While the populations north of the Black Sea are separated by several hundreds of km of unsuitable steppe habitat, the Sea of Marmara with the Dardanelles and Bosphorus Straits do not represent a strong physical barrier for a bird. The presence of the middle spotted woodpecker on the Lesvos Is. supports this assertion. During the LGM, there was a land bridge and many North Aegean islands were connected to each other and the mainland (Perissoratis & Conispoliatis, 2003). Moreover, indication for historical gene flow between Anatolia and Europe throughout the Pleistocene was found in various taxa such as bats (Bilgin *et al.*,

2009), grasshoppers (Korkmaz *et al.*, 2014) and plants (Ansell *et al.*, 2011; Ülker *et al.*, 2018). However, according to estimated SDMs, climates seem not to have been suitable for the middle spotted woodpecker in this region during the LGM and the area must have been colonized from the east and the west recently. Although there are several bird species with different subspecies described from either side of the Sea of Marmara (Roselaar, 1995) such a strong phylogeographic break has not been documented previously.

Phylogeographic patterns and demographic history of Asian and European groups

Genetic differentiation was rather shallow within both, the Asian and European groups. There was no significant geographic structure, although populations from Anatolia and the Greater Caucasus did not share any mtDNA haplotypes. This lack of geographic structure in the rather large ranges of both groups was unexpected given the ecological specialization and presumed relatively low dispersal abilities of the middle spotted woodpecker (Pasinelli, 2003), and the presence of slight geographically structured phenotypic variation particularly in the Asian group (Vaurie, 1965; Winkler, Christie, Kirwan*, et al.*, 2018). Consequently, we cannot exclude this phenotypic variation reflects recent local adaptions to contemporarily different environments. However, the variability of the analyzed mtDNA markers might be insufficient to detect geographic segregation within different refugia during the last glacial period or divergence estimates might be confounded by gene flow between populations making them appear very recent. Nonetheless, the lack of genetic diversity more likely indicates that populations in both groups went through bottlenecks during the LGM loosing most of the within-group diversity.

The extended Bayesian Skyline Plots suggested very recent population expansions in the Asian group and a steady increase of the effective population size in the European group over the last 50 kyr. However, these analyses have to be interpreted with caution and we cannot reject constant population sizes with our data. Although the sample sizes likely were adequate to suggest population expansions, they could still have been inadequate to effectively capture their dimensions (Grant, 2015). Moreover, analyses were based on just a single gene tree and the limited genetic variation of the analyzed mtDNA loci could have additionally impeded tracing the recent population history.

In the European group, the relatively high haplotype and nucleotide diversity in Central Europe may indicate that this region was colonized from multiple LGM refugia (cf. Petit *et al.*, 2003). Indeed, SDMs for the LGM revealed different suitable areas that could have acted as refugia, particularly the southwestern France and the northeastern Iberia, the Apennine Peninsula and the Balkans. These refugia are concordant with those predicted for the common oak *Quercus robur* (Petit *et al.*, 2002; Ülker *et al.*, 2018) one of the key resource species for the middle spotted woodpecker (Pasinelli, 2003). The Iberian Peninsula probably was not a refugium given the low haplotype and nucleotide diversity

and the presence of only a single widespread haplotype there, with the exception of a second Control Region haplotype shared with European Russia. The lack of samples from the Apennine peninsula however, left the picture of the genetic diversity in the European group somewhat incomplete.

In the Asian group, haplotype and nucleotide diversity in the mtDNA markers was consistently higher for the Armenian and Zagros populations compared to those from the Greater Caucasus and Anatolia. Thus, the former areas have likely served as source areas for colonization of the latter areas. However, SDMs revealed suitable areas during the LGM not only near the Persian Gulf but also in Anatolia, southern Greece, and along the eastern Mediterranean coast. Anatolia retained much of its forest cover during the LGM and was a refugium for the common oak (Ülker *et al.*, 2018) and the Turkey oak *Quercus cerris* (Bagnoli *et al.*, 2016). The low genetic diversity of the middle spotted woodpecker in Anatolia, however, suggests that this area was not likely a source for postglacial population expansion. Biotic factors not included in our models such as the lack of appropriate vegetation for foraging, for example, may have rendered climatically suitable areas uninhabitable during the LGM.

The Green Woodpecker *Picus viridis,* another temperate Western Palearctic woodpecker shows a different phylogeographic pattern. It lacks mtDNA differentiation over large parts of its European range but its Iberian population is often treated as a distinct species *P. sharpei* (e.g. del Hoyo & Collar, 2014; Gill & Donsker, 2018). This Iberian taxon diverged from the European population during the Early-Middle Pleistocene Transition (Perktas *et al.*, 2011; Pons *et al.*, 2011) contemporaneously with the Asia-Europe split in the middle spotted woodpecker. In contrast to the middle spotted woodpecker, green woodpecker populations from the Caucasus and particularly from Anatolia were poorly differentiated from European populations, whereas the isolated subspecies from the Zagros Mountains was genetically distinct (Perktas *et al.*, 2011).

The great spotted woodpecker, which is an ecological generalist (Winkler, Christie, & Kirwan, 2018), shows a weak mtDNA structure over large parts of its Eurasian range (Perktas & Quintero, 2013; Zink *et al.*, 2002). In contrast to the middle spotted woodpecker, great spotted woodpecker populations from the Caucasus and Anatolia were not differentiated from those of Europe, while only populations from the south Caspian region and Transcaspia diverged from European populations about 0.5 Ma (Perktas & Quintero, 2013).

Conclusions

We discovered a strong genetic divergence between Asian and European populations of the middle spotted woodpecker. Although population structuring was expected for a resident habitat specialist with low dispersal such as the middle spotted woodpecker, its magnitude was surprising given only slight apparent geographic variation in plumage and morphology which has been described as partly

clinal by some authors (Cramp, 1985). The lack of shared mtDNA haplotypes and alleles in one nuclear marker might indicate a lack of recent gene flow between the two groups. However, our sampling and data do not allow us to reject the possibility of secondary contact with at least restricted gene flow between European and Asian groups and this calls for further investigation. Additional studies should also look for potential differences in vocalizations and ecology between the two groups. Lineages showing similar level of genetic differentiation including woodpeckers (green woodpecker s.l., great spotted woodpecker s.l.) are often treated as distinct species regardless of the species concepts considered (del Hoyo & Collar, 2014; Perktas *et al.*, 2015; Perktas & Quintero, 2013).

Even though there was no significant phylogeographic structure within the two clades, we cannot exclude the possibility of multiple LGM refugia in both Asian and European parts of the species range. Population structure has to be further investigated with more extensive geographic sampling and larger number of variable, independently evolving markers. Comparing our results with the phylogeographic history of other woodpeckers, we show that largely sympatric species with similar resident life-histories might have idiosyncratic phylogeographic patterns resulting, perhaps, from different species' ecology.

Acknowledgements

We would like to thank Hans-Martin Berg from the Natural History Museum Vienna, Sharon M. Birks from the University of Washington Burke Museum, Heather L. Prestridge from the Biodiversity Research and Teaching Collections at Texas A&M University, Thomas J. Trombone from the American Museum of Natural History and Kristof Zyskowski from the Yale University Peabody Museum Rolf Hennes, Bad Homburg, for providing samples from their institutions to us and Amelie Lindgren who kindly provided us ice sheets, glaciers, permafrost and sea shelf permafrost maps of the Last Glacial Maximum. Carlos Ciudad and Chano Robles kindly assisted in gathering the Spanish samples. This study was supported by the Swiss Ornithological Institute and the Natural History Museum of Bern. Samples from the Balkans, Greater Caucasus, and Armenia were collected with support by FEDER funds through the COMPETE programme, POPH/QREN/FSE funds, and by the Fundação para a Ciência e a Tecnologia/MEC (FCOMP-01-0124- FEDER-008941; FCT reference PTDC/BIA- BEC/103435/2008), Portugal and the National Geographic Society (S.V.D.)

References

Alaei Kakhki, N., Aliabadian, M., Förschler, M. I., Ghasempouri, S. M., Kiabi, B. H., Verde Arregoitia, L. D., et al. (2018). Phylogeography of the *Oenanthe hispanica–pleschanka–cypriaca* complex (aves, muscicapidae: Saxicolinae): Diversification history of open-habitat specialists based on climate niche models, genetic data, and morphometric data. *Journal of Zoological Systematics And Evolutionary Research*, 56, 408-427.

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (tss). *Journal of Applied Ecology*, 43, 1223-1232.

Ansell, S. W., Stenoien, H. K., Grundmann, M., Russell, S. J., Koch, M. A., Schneider, H., et al. (2011). The importance of anatolian mountains as the cradle of global diversity in arabis alpina, a key arctic-alpine species. *Annals of Botany*, 108, 241-252.

Arguedas, N., & Parker, P. G. (2000). Seasonal migration and genetic population structure in house wrens. *Condor*, 102, 517-528.

Bagnoli, F., Tsuda, Y., Fineschi, S., Bruschi, P., Magri, D., Zhelev, P., et al. (2016). Combining molecular and fossil data to infer demographic history of *Quercus cerris*: Insights on european eastern glacial refugia. *Journal Of Biogeography*, 43, 679-690.

Bilgin, R., Coraman, E., Karatas, A., & Morales, J. C. (2009). Phylogeography of the greater horseshoe bat, *Rhinolophus ferrumequinum* (Chiroptera: Rhinolophidae), in southeastern europe and anatolia, with a specific focus on whether the sea of marmara is a barrier to gene flow. *Acta Chiropterologica*, 11, 53-60.

BirdLifeInternationalandHandbookoftheBirdsoftheWorld 2016. Bird species distribution maps of the world. Version 6.0.

Bouckaert, R., & Heled, J. (2014). Densitree 2: Seeing trees through the forest. *bioRxiv, 2014*.

Bouckaert, R., Heled, J., Kuhnert, D., Vaughan, T., Wu, C. H., Xie, D., et al. (2014). Beast 2: A software platform for Bayesian evolutionary analysis. *Plos Computational Biology*, 10.

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. New York: Chapman and Hall.

Bühlmann, J., & Pasinelli, G. (2012). Analyse des bestandsrückgangs beim mittelspecht *Dendrocopos medius* von 1978–2002 im kanton zürich: Grundlagen für den nachhaltigen schutz einer gefährdeten waldvogelart. *Ornithol. Beob.*, 109, 73-94.

Calenge, C. (2006). The package "adehabitat" for the R software: A tool for the analysis of space and habitat use by animals. *Ecological Modelling*, 197, 516-519.

Calenge, C., Darmon, G., Basille, M., Loison, A., & Jullien, J. M. (2008). The factorial decomposition of the mahalanobis distances in habitat selection studies. *Ecology*, 89, 555-566.

Chandler, R., & Royle, A. (2013). The package “maxlike”. Model occurrence probability using presence-only data.

Cibois, A., Thibault, J. C., Rocamora, G., & Pasquet, E. (2016). Molecular phylogeny and systematics of blue and grey noddies (Procelsterna). *Ibis*, 158, 433-438.

Ciudad, C., Robles, H., & Matthysen, E. (2009). Postfledging habitat selection of juvenile middle spotted woodpeckers: A multi-scale approach. *Ecography*, 32, 676-682.

Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., et al. (2009). The last glacial maximum. *Science*, 325, 710-714.

Cramp, S. (1985). *The birds of the western palearctic. Vol 4. Terns to woodpeckers.* Oxford: Oxford University Press.

del Hoyo, J., & Collar, N. J. (2014). *Hbw and birdlife international illustrated checklist of the birds of the world. Volume 1: No-passerines*. Barcelona: Lynx Edicions.

Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F. T., D'Amen, M., Randin, C., et al. (2017). Ecospat: An R package to support spatial analyses and modeling of species niches and distributions. *Ecography*, 40, 774-787.

Drovetski, S. V. (2003). Plio-pleistocene climatic oscilations, holarctic biogeography and speciation in an avian subfamily. *Journal of Biogeography*, 30, 1173-1181.

Drovetski, S. V., Semenov, G., Red'kin, Y. A., Sotnikov, V. N., Fadeev, I. V., & Koblik, E. A. (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.

Drovetski, S. V., Zink, R. M., Rohwer, S., Fadeev, I. V., Nesterov, E. V., Karagodin, I., et al. (2004). Complex biogeographic history of a Holarctic passerine. *Proceedings of the Royal Society B-Biological Sciences*, 271, 545-551.

Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, 22, 1185-1192.

Edwards, S. V., & Beerli, P. (2000). Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, 54, 1839-1854.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564-567.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131, 479-491.

Flot, J. F. (2010). Segphase: A web tool for interconverting phase input/output files and fasta sequence alignments. *Molecular Ecology Resources*, 10, 162-166.

Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics*, 1-67.

Fuchs, J., & Pons, J. M. (2015). A new classification of the pied woodpeckers assemblage (dendropicini, picidae) based on a comprehensive multi-locus phylogeny. *Molecular Phylogenetics and Evolution*, 88, 28-37.

Garcia, J. T., Alda, F., Terraube, J., Mougeot, F., Sternalski, A., Bretagnolle, V., et al. (2011). Demographic history, genetic structure and gene flow in a steppe-associated raptor species. *Bmc Evolutionary Biology*, 11, 11.

Garcia, J. T., Mañosa, S., Morales, M. B., Ponjoan, A., García de la Morena, E. L., Bota, G., et al. (2011). Genetic consequences of interglacial isolation in a steppe bird. *Molecular Phylogenetics and Evolution*, 61, 671-676.

Gill, F. B., & Donsker, D. E. 2018. IOC world bird list (v 8.1.4): International Ornithologists' Union.

Golding, N., & Golding, M. N. (2014). Package "graf". Species distribution modelling using latent gaussian random fields.

Golding, N., & Purse, B. V. (2016). Fast and flexible bayesian species distribution modelling using Gaussian processes. *Methods in Ecology and Evolution*, 7, 598-608.

Grant, W. S. (2015). Problems and cautions with sequence mismatch analysis and Bayesian skyline plots to infer historical demography. *Journal of Heredity*, 106, 333-346.

Harrell, F.E. (2001). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Berlin: Springer.

Hastie, T., Tibshirani, R., & Buja, A. (1994). Flexible discriminant-analysis by optimal scoring. *Journal of the American Statistical Association*, 89, 1255-1270.

Head, M. J., & Gibbard, P. L. 2005. Early-middle pleistocene transitions: An overview and recommendation for the defining boundary. In M. J. Head & P. L. Gibbard (Eds) *Early-middle Pleistocene transitions: The land-ocean evidence* pp. 1-18). Bath: Geological Soc Publishing House.

Hebda, G., Wesolowski, T., & Rowinski, P. (2016). Nest sites of middle spotted woodpeckers *Leiopicus medius* in a primeval forest. *Ardea*, 104, 119-128.

Hewitt, G. (2000). The genetic legacy of the quaternary ice ages. *Nature*, 405, 907-913.

Hewitt, G. M. (1999). Post-glacial re-colonization of european biota. *Biological Journal of the Linnean Society*, 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. *International Journal of Climatology*, 25, 1965-1978.

Ho, S. Y. W., Lanfear, R., Bromham, L., Phillips, M. J., Soubrier, J., Rodrigo, A. G., et al. (2011). Time-dependent rates of molecular evolution. *Molecular Ecology*, 20, 3087-3101.

Hung, C. M., Drovetski, S. V., & Zink, R. M. (2012). Multilocus coalescence analyses support a mtDNA-based phylogeographic history for a widespread palearctic passerine bird, *Sitta europaea*. *Evolution*, 66, 2850-2864.

Hung, C. M., Drovetski, S. V., & Zink, R. M. (2016). Matching loci surveyed to questions asked in phylogeography. *Proceedings of the Royal Society B-Biological Sciences*, 283.

Hung, C. M., Drovetski, S. V., & Zink, R. M. (2017). The roles of ecology, behaviour and effective population size in the evolution of a community. *Molecular Ecology*, 26, 3775-3784.

Ihlow, F., Courant, J., Secondi, J., Herrel, A., Rebelo, R., Measey, G. J., et al. (2016). Impacts of climate change on the global invasion potential of the African clawed frog *Xenopus laevis*. *Plos One*, 11.

Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., et al. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, 346, 1320-1331.

Jenni, L. (1983). Habitatnutzung, nahrungserwerb und nahrung von mittel-und buntspecht (*D. endrocopos medius* und *D. major*) sowie bemerkungen zur verbreitungsgeschichte des mittelspechts. *Ornithol. Beob.*, 80, 29-57.

Katoh, K., & Standley, D. M. (2013). Mafft multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772-780.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28, 1647-1649.

Korkmaz, E. M., Lunt, D. H., Ciplak, B., Degerli, N., & Basibuyuk, H. H. (2014). The contribution of anatolia to european phylogeography: The centre of origin of the meadow grasshopper, *Chorthippus parallelus*. *Journal Of Biogeography*, 41, 1793-1805.

Kosenko, S. M., & Kaigorodova, E. Y. (2001). Effect of habitat fragmentation on distribution, density and breeding performance of the middle spotted woodpecker *Dendrocopos medius* (alves, picidae) in nerussa-desna polesye. *Zoologichesky Zhurnal*, 80, 71-78.

Kosinski, Z. (2006). Factors affecting the occurrence of middle spotted and great spotted woodpeckers in deciduous forests - a case study from poland. *Annales Zoologici Fennici*, 43, 198-210.

Kosinski, Z., Pluta, M., Ulanowska, A., Walczak, L., Winiecki, A., & Zarebski, M. (2018). Do increases in the availability of standing dead trees affect the abundance, nest-site use, and niche partitioning of great spotted and middle spotted woodpeckers in riverine forests? *Biodiversity and Conservation*, 27, 123-145.

Kosinski, Z., & Winiecki, A. (2004). Nest-site selection and niche partitioning among the great spotted woodpecker *Dendrocopos major* and middle spotted woodpecker *Dendrocopos medius* in riverine forest of central europe. *Ornis Fennica*, 81, 145-156.

Kossenko, S. M. 2003. A study of mechanisms underlying habitat fragmentation effects on the middle spotted woodpecker *Picoides medius*: A progress report. *International Woodpecker Symposium* pp. 91-103): Forschungsbericht, 48 Nationalparkverwaltung Berchtesgaden, Germany.

Kossenko, S. M., & Kaygorodova, E. Y. (2007). Reproduction of the middle spotted woodpecker *Dendrocopos medius* in the nerussa-desna woodland, sw russia, with particular reference to habitat fragmentation, weather conditions and food supply. *Ardea*, 95, 177-189.

Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). Partitionfinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29, 1695-1701.

Lerner, H. R. L., Meyer, M., James, H. F., Hofreiter, M., & Fleischer, R. C. (2011). Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Current Biology*, 21, 1838-1844.

Li, S. N., Jovelin, R., Yoshiga, T., Tanaka, R., & Cutter, A. D. (2014). Specialist versus generalist life histories and nucleotide diversity in caenorhabditis nematodes. *Proceedings of the Royal Society B-Biological Sciences*, 281.

Lindgren, A., Hugelius, G., Kuhry, P., Christensen, T. R., & Vandenberghe, J. (2016). Gis-based maps and area estimates of northern hemisphere permafrost extent during the last glacial maximum. *Permafrost and Periglacial Processes*, 27, 6-16.

Maslin, M. A., Brierley, C. M., Milner, A. M., Shultz, S., Trauth, M. H., & Wilson, K. E. (2014). East african climate pulses and early human evolution. *Quaternary Science Reviews*, 101, 1-17.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*. London: Chapman and Hall.

Müller, W. (1982). Die besiedlung der eichenwälder im kanton zürich durch den mittelspechts. *Der Ornithologisce Beobachter*, 79, 105-119.

Nabholz, B., Glemin, S., & Galtier, N. (2009). The erratic mitochondrial clock: Variations of mutation rate, not population size, affect mtdna diversity across birds and mammals. *Bmc Evolutionary Biology*, 9, 13.

Pasinelli, G. (2000). Oaks (quercus sp.) and only oaks? Relations between habitat structure and home range size of the middle spotted woodpecker (*Dendrocopos medius*). *Biological Conservation*, 93, 227-235.

Pasinelli, G. (2001). Breeding performance of the middle spotted woodpecker *Dendrocopos medius* in relation to weather and territory quality. *Ardea*, 89, 353-361.

Pasinelli, G. (2003). *Dendrocopos medius* middle spotted woodpecker. *Birds of the Western Palearctic Update*, 5, 49-99.

Pasinelli, G. (2007). Nest site selection in middle and great spotted woodpeckers D*endrocopos medius* & *D. major*: Implications for forest management and conservation. *Biodiversity and Conservation*, 16, 1283-1298.

Pasinelli, G., Grendelmeier, A., Gerber, M., & Arlettaz, R. (2016). Rodent-avoidance, topography and forest structure shape territory selection of a forest bird. *Bmc Ecology*, 16.

Pasinelli, G., & Hegelbach, J. (1997). Characteristics of trees preferred by foraging middle spotted woodpecker *Dendrocopos medius* in northern switzerland. *Ardea*, 85, 203-209.

Perissoratis, C., & Conispoliatis, N. (2003). The impacts of sea-level changes during latest pleistocene and holocene times on the morphology of the ionian and aegean seas (se alpine europe). *Marine Geology*, 196, 145-156.

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. *Biological Journal of the Linnean Society*, 104, 710-723.

Perktas, U., Gur, H., & Ada, E. (2015). Historical demography of the Eurasian green woodpecker: Integrating phylogeography and ecological niche modelling to test glacial refugia hypothesis. *Folia Zoologica*, 64, 284-295.

Perktas, U., & Quintero, E. (2013). A wide geographical survey of mitochondrial DNA variation in the great spotted woodpecker complex, *Dendrocopos major* (Aves: Picidae). *Biological Journal of the Linnean Society*, 108, 173-188.

Petit, R. J., Aguinagalde, I., de Beaulieu, J. L., Bittkau, C., Brewer, S., Cheddadi, R., et al. (2003). Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science*, 300, 1563-1565.

Petit, R. J., Brewer, S., Bordacs, S., Burg, K., Cheddadi, R., Coart, E., et al. (2002). Identification of refugia and post-glacial colonisation routes of european white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, 156, 49-74.

Pettersson, B. (1983). Foraging behavior of the middle spotted woodpecker *Dendrocopos medius* in sweden. *Holarctic Ecology*, 6, 263-269.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231-259.

Pons, J. M., Olioso, G., Cruaud, C., & Fuchs, J. (2011). Phylogeography of the eurasian green woodpecker (*picus viridis*). *Journal of Biogeography*, 38, 311-325.

Provan, J., & Bennett, K. D. (2008). Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution*, 23, 564-571.

Prum, R. O., Berv, J. S., Dornburg, A., Field, D. J., Townsend, J. P., Lemmon, E. M., et al. (2015). A comprehensive phylogeny of birds (aves) using targeted next-generation DNA sequencing. *Nature*, 526, 569-U247.

Author 2008. Figtree 1.4, published by the author.

Rambaut, A., Suchard, M. A., & Drummond, A. J. (2013). Tracer v 1.6, available fromhttp://tree.Bio.Ed.Ac.Uk/software/tracer/.

Ripley, B. D. (2007). *Pattern recognition and neural networks*. Cambridge: Cambridge university press.

Robles, H., & Ciudad, C. (2012). Influence of habitat quality, population size, patch size, and connectivity on patch-occupancy dynamics of the middle spotted woodpecker. *Conservation Biology*, 26, 284-293.

Robles, H., & Ciudad, C. (2017). Floaters may buffer the extinction risk of small populations: An empirical assessment. *Proceedings of the Royal Society B-Biological Sciences*, 284.

Robles, H., Ciudad, C., Vera, R., Olea, P. P., Purroy, F. J., & Matthysen, E. (2007). Sylvopastoral management and conservation of the middle spotted woodpecker at the south-western edge of its distribution range. *Forest Ecology and Management*, 242, 343-352.

Roselaar, C. S. (1995). *Songbirds of turkey. An atlas of biodiversity of turkish passerine birds.* . Robertsbridge & GMB, Harlem: Pica Press.

Royle, J. A., Chandler, R. B., Yackulic, C., & Nichols, J. D. (2012). Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods in Ecology and Evolution*, 3, 545-554.

Rozen, S., & Skaletsky, H. J. 2000. Primer3 on the www for general users and for biologist programmers. In S. Krawetz & S. Misener (Eds) *Bioinformatics methods and protocols: Methods in molecular biology.* pp. 365-386). Totowa, NJ: Humana Press.

Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D., et al. (2014). The drivers of tropical speciation. *Nature*, advance online publication.

Stephens, M., & Scheet, P. (2005). Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *American Journal of Human Genetics*, 76, 449-462.

Stephens, M., Smith, N. J., & Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, 68, 978-989.

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, 7, 453-464.

Thuiller, W., Georges, D., Engler, R., Breiner, F., Georges, M. D., & Thuiller, C. W. (2016). The package "biomod2". Species distribution modeling within an ensemble forecasting framework.

Thuiller, W., Lafourcade, B., Engler, R., & Araujo, M. B. (2009). Biomod - a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369-373.

Ülker, E. D., Tavsanoglu, C., & Perktas, U. (2018). Ecological niche modelling of pedunculate oak (*Quercus robur*) supports the 'expansion-contraction' model of Pleistocene biogeography. *Biological Journal Of The Linnean Society*, 123, 338-347.

Vaurie, C. (1965). *The birds of the palearctic fauna. Non passeriformes.* London: Witherby Ltd.

Weir, J. T., Haddrath, O., Robertson, H. A., Colbourne, R. M., & Baker, A. J. (2016). Explosive ice age diversification of kiwi. *Proceedings of the National Academy of Sciences of the United States of America*, 113, E5580-E5587.

Weir, J. T., & Schluter, D. (2004). Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 271, 1881-1887.

Winkler, H., Christie, D. A., & Kirwan, G. M. 2018. Great spotted woodpecker (*Dendrocopos major*). In J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie & E. de Juana (Eds) *Handbook of the birds of the world alive.* . Barcelona: Lynx Edicions.

Winkler, H., Christie, D. A., Kirwan, G. M., & de Juana, E. 2018. Middle spotted woodpecker (*Leiopicus medius*). In J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie & E. de Juana (Eds) *Handbook of the birds of the world alive.* . Barcelona: Lynx Edicions.

Zamudio, K. R., Bell, R. C., & Mason, N. A. (2016). Phenotypes in phylogeography: Species' traits, environmental variation, and vertebrate diversification. *Proceedings of the National Academy of Sciences of the United States of America*, 113, 8041-8048.

Zink, R. M., Drovetski, S. V., & Rohwer, S. (2002). Phylogeographic patterns in the great spotted woodpecker *Dendrocopos major* across Eurasia. *Journal of Avian Biology*, 33, 175-178.

Zink, R. M., Drovetski, S. V., & Rohwer, S. (2006). Selective neutrality of mitochondrial ND2 sequences, phylogeography and species limits in *Sitta europaea*. *Molecular Phylogenetics and Evolution*, 40, 679-686.

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1, 3-14.

Figure legends

Fig. 1. Breeding range of the middle spotted woodpecker *Dendrocoptes medius* with ranges of different subspecies shown in different colors. Sampling localities were grouped into eight geographic areas identified by different colors of the circles. Samples size (n) is given for each population. Distribution areas are modified from BirdLife International and the Handbook of the Birds of the World (Bird Life International and Handbook of the Birds of the World, 2016).

Fig. 2. Haplotype networks of ATPase subunit 6 (ATP6) (a), Control Region 1 D-Loop (ConReg) (b), Cytochrome b (CytB) (c), autosomal beta-Fibrinogen intron 7 (Fib7) (d), z-linked Brahma protein 15 intron (BRM) (e). Regions of sample origin are marked by different colors. Dashed lines indicate substitution steps and the size of the circles is proportional to the number of sampled haplotypes or alleles.

Fig. 3. Time-calibrated maximum clade credibility mtDNA gene tree reconstructed in BEAST and based on the coalescence exponential tree prior. a) eastern (Asian) and western (European) clades collapsed, b) western clade expanded c) eastern clade expanded. Colored circles indicate the population of origin for the different individuals. Mean ages are shown next to nodes and the blue bars represent their 95% highest posterior density distributions. Posterior probability values > 0.5 are also indicated next to nodes. Bird figures were taken with permission from Winkler et al. (2018).

Fig. 4. Left: Time-calibrated, maximum-clade-credibility species tree reconstructed in *BEAST based on three mtDNA and one nuclear loci. Geographic areas were used as “taxa”. Posterior probability values and mean ages are shown next to nodes. The blue bars represent 95% highest posterior density distributions for node ages. Right: Cloudogram illustrating the uncertainty of the sampled topologies over the posterior distribution. Different topologies from the posterior distribution are shown in different colors. Blue: Most popular topologies. Red: 2nd most popular topologies. Pale green: 3rd most popular topologies. Dark green: 4th most popular topologies.

Fig. 5. Extended Bayesian Skyline Plots based on three mtDNA markes for the Asian and European groups. Time scale in millions of years before present. The dashed lines show median values. The 95% highest posterior density interval (HPD) is indicated in grey. Population size is given as the product of the effective population size and the generation time. Bars represent a histogram of tree events per time slice.

Fig. 6. Ensemble prediction maps derived by from 12 species distribution models for the Asian (a) and the European group (b) under current climate conditions. Red-green scale represents lower to higher suitability for occurrence. Binary map representing potentially suitable/unsuitable areas for the eastern (c) and western (d) group. Grey and red colours represent modelled absence and presence areas, respectively.

Fig. 7. Ensemble prediction maps derived from 12 species distribution models for the eastern (Asian) (a) and the western (European) group (b) under past climate conditions (Last Glacial Maximum LGM; ~22,000 years BP). Ice sheets and glaciers (in light blue), permafrost and sea shelf permafrost (in light brown and dark blue, respectively) were obtained from Lindgren et al. (2016). Black lines indicate current surface above sea level. Red-green scale represents lower to higher suitability for occurrence. Binary map representing modelled suitable/unsuitable areas for the eastern (c) and western (d) groups. Grey and green colours represents absence and presence areas, respectively).

Supporting Information

Table S1. List of all used samples with Museum number/samples, country, geographic area according to Figure 1 and GenBank sequences accession numbers.

Table S2. List of all primers used in this study with their name, sequence, reference and annealing temperature used for PCR reactions.

Table S3. Variables considered in the analysis. Variables with a variance inflation factor (VIF) > 3 were removed from further analysis (*) due to multi-collinearity with other variables.

Table S4. Standard genetic indices for each geographic area and marker.

Table S5. Best-fitting substitution models.

Table S6. Model evaluation with 10-fold cross-validation of the 12 species distribution methods and their ensemble prediction (EP) for the Eastern and Western clade. Area Under the Curve (AUC) ranges between 0 and 1 (worse than a random model and best discriminating model, respectively). True Skill Statistic (TSS) and Boyce's Index (BI) ranges between −1 and 1 (higher values indicate a good predictive accuracy, while 0 indicates random prediction). Average values are shown (*: $P < 0.001$).

Figure S1. Occurrence point localities of the middle spotted woodpecker used for the maximum entropy niche modelling taken from GBIF databank (http://www.gbif.org/) and from www.observation.org. Observations between February and July (corresponding with the breeding period and season of high territoriality; Pasinelli *et al.* 2001) after 1980 and with a coordinate certainty of less than 500 meters were considered.

Laboratory Protocol

Supporting Information

Table S1. List of all used samples with Museum number/samples, country, geographic area according to Figure 1 and GenBank sequences accession numbers. AMNH: American Museum of Natural History; NMW: Natural History Museum Vienna; YPM: Yale University Peabody Museum of Natural History; UWBM: University of Washington Burke Museum.

Taxon	Museum number/sample name	Locality/Region	Country	Geographic are	Type of sample	ATPase subunit 6	Control Region	Cytochrome b	beta-Fibrinogen intron 7	Brahma protein 15 intron
Dendrocoptes medius	21	Cantabrian Mountains	Spain	Spain	Blood					
Dendrocoptes medius	22	Cantabrian Mountains	Spain	Spain	Blood					
Dendrocoptes medius	23	Cantabrian Mountains	Spain	Spain	Blood					
Dendrocoptes medius	24	Cantabrian Mountains	Spain	Spain	Blood					
Dendrocoptes medius	25	Cantabrian Mountains	Spain	Spain	Blood					
Dendrocoptes medius	26	Cantabrian Mountains	Spain	Spain	Blood					
Dendrocoptes medius	27	Cantabrian Mountains	Spain	Spain	Blood					
Dendrocoptes medius	28	Cantabrian Mountains	Spain	Spain	Blood					
Dendrocoptes medius	AMNHDOT14706	Nordwestmecklenburg	Germany	Central Europe	Tissue					
Dendrocoptes medius	NMWID3063	Höflein an der Donau, Lower Austria	Austria	Central Europe	Tissue					
Dendrocoptes medius	NMWVS1101	Füllersdorf, Lower Austria	Austria	Central Europe	Tissue					
Dendrocoptes medius	7839075	Bad Homburg vor der Höhe	Germany	Central Europe	Feather					
Dendrocoptes medius	7874021	Bad Homburg vor der Höhe	Germany	Central Europe	Feather					
Dendrocoptes medius	7874027	Bad Homburg vor der Höhe	Germany	Central Europe	Feather					
Dendrocoptes medius	7874050	Bad Homburg vor der Höhe	Germany	Central Europe	Feather					
Dendrocoptes medius	7874065	Bad Homburg vor der Höhe	Germany	Central Europe	Feather					
Dendrocoptes medius	7874115	Bad Homburg vor der Höhe	Germany	Central Europe	Feather					

Dendrocoptes medius	F1	Września District	Poland	Central Europe	Feather
Dendrocoptes medius	F2	Września District	Poland	Central Europe	Feather
Dendrocoptes medius	F3	Września District	Poland	Central Europe	Feather
Dendrocoptes medius	T101301	Bad Homburg vor der Höhe	Germany	Central Europe	Feather
Dendrocoptes medius	T101302	Bad Homburg vor der Höhe	Germany	Central Europe	Feather
Dendrocoptes medius	SH101429	Schaffhausen	Switzerland	Central Europe	Feather
Dendrocoptes medius	TG152304	Thurgau	Switzerland	Central Europe	Feather
Dendrocoptes medius	TG152305	Thurgau	Switzerland	Central Europe	Feather
Dendrocoptes medius	ZH152306	Zurich	Switzerland	Central Europe	Feather
Dendrocoptes medius	ZH152307	Zurich	Switzerland	Central Europe	Feather
Dendrocoptes medius	ZH152308	Zurich	Switzerland	Central Europe	Feather
Dendrocoptes medius	YPM84372	Pcinja District	Serbia	Balkans	Tissue
Dendrocoptes medius	YPM84373	Pcinja District	Serbia	Balkans	Tissue
Dendrocoptes medius	YPM140529	East Macedonia, Paranesti	Greece	Balkans	Tissue
Dendrocoptes medius	YPM140530	East Macedonia, Paranesti	Greece	Balkans	Tissue
Dendrocoptes medius	YPM141470	East Macedonia, Paranesti	Greece	Balkans	Tissue
Dendrocoptes medius	YPM140579	East Macedonia, Paranesti	Greece	Balkans	Tissue
Dendrocoptes medius	YPM142773	East Macedonia, Rhodope Mountains	Greece	Balkans	Tissue
Dendrocoptes medius	YPM142709	East Macedonia, Rhodope Mountains	Greece	Balkans	Tissue
Dendrocoptes medius	YPM142654	East Macedonia, Rhodope Mountains	Greece	Balkans	Tissue
Dendrocoptes medius	02_03	Nerussa-Desna woodland, Bryansk Region	Russia	Western Russia	Blood
Dendrocoptes medius	03_03	Nerussa-Desna woodland, Bryansk Region	Russia	Western Russia	Blood
Dendrocoptes medius	07_03	Nerussa-Desna woodland, Bryansk Region	Russia	Western Russia	Blood
Dendrocoptes medius	09_03	Nerussa-Desna woodland, Bryansk Region	Russia	Western Russia	Blood
Dendrocoptes medius	10_03	Nerussa-Desna woodland, Bryansk Region	Russia	Western Russia	Blood
Dendrocoptes medius	12_03	Nerussa-Desna woodland, Bryansk Region	Russia	Western Russia	Blood
Dendrocoptes medius	YPM140405	Lesvos, Polichnitos	Greece	Anatolia	Tissue
Dendrocoptes medius	YPM140425	Lesvos, Polichnitos	Greece	Anatolia	Tissue
Dendrocoptes medius	YPM141384	Lesvos, Polichnitos	Greece	Anatolia	Tissue
Dendrocoptes medius	YPM140494	Lesvos, Polichnitos	Greece	Anatolia	Tissue

Dendrocoptes medius	YPM140467	Lesvos, Mount Olympus	Greece	Anatolia	Tissue
Dendrocoptes medius	YPM140468	Lesvos, Mount Olympus	Greece	Anatolia	Tissue
Dendrocoptes medius	YPM144969	Lesvos, Vrisa	Greece	Anatolia	Tissue
Dendrocoptes medius	YPM145110	Lesvos, Vrisa	Greece	Anatolia	Tissue
Dendrocoptes medius	YPM145145	Lesvos, Petra	Greece	Anatolia	Tissue
Dendrocoptes medius	YPM58966	Icel	Turkey	Anatolia	Toe pad
Dendrocoptes medius	YPM58967	Icel	Turkey	Anatolia	Toe pad
Dendrocoptes medius	YPM58968	Icel	Turkey	Anatolia	Toe pad
Dendrocoptes medius	YPM58970	Icel	Turkey	Anatolia	Toe pad
Dendrocoptes medius	YPM58971	Icel	Turkey	Anatolia	Toe pad
Dendrocoptes medius	YPM58972	Antalya	Turkey	Anatolia	Toe pad
Dendrocoptes medius	YPM58973	Mugla	Turkey	Anatolia	Toe pad
Dendrocoptes medius	YPM58974	Mugla	Turkey	Anatolia	Toe pad
Dendrocoptes medius	YPM58975	Izmir	Turkey	Anatolia	Toe pad
Dendrocoptes medius	YPM58976	Izmir	Turkey	Anatolia	Toe pad
Dendrocoptes medius	YPM101807	Krasnodar Krai	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	YPM140852	Krasnodar Krai	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	UWBM61173	Krasnodarskiy Kray	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	UWBM61406	Krasnodarskiy Kray	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	UWBM61424	Krasnodarskiy Kray	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	UWBM61425	Krasnodarskiy Kray	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	UWBM61426	Krasnodarskiy Kray	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	UWBM64628	Krasnodarskiy Kray	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	UWBM71227	Krasnodarskiy Kray	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	UWBM64673	Krasnodarskiy Kray	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	UWBM64715	Krasnodarskiy Kray	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	UWBM64847	Krasnodarskiy Kray	Russia	Greater Caucasus	Tissue
Dendrocoptes medius	SVD5612	Tavush Marz, Zikarta	Armenia	Armenia	Blood
Dendrocoptes medius	SVD5617	Tavush Marz, Zikarta	Armenia	Armenia	Blood
Dendrocoptes medius	KS82246	Nerkin Hand	Armenia	Armenia	Blood

Dendrocoptes medius	KS82270	Tavush Marz, Zikarta	Armenia	Armenia	Blood					
Dendrocoptes medius	KS82275	Tavush Marz, Zikarta	Armenia	Armenia	Blood					
Dendrocoptes medius	EAK444	Nerkin Hand	Armenia	Armenia	Blood					
Dendrocoptes medius	KS82243	Nerkin Hand	Armenia	Armenia	Blood					
Dendrocoptes medius	IVF1221	Nerkin Hand	Armenia	Armenia	Blood					
Dendrocoptes medius	AMNH462277	Bakhtiari	Iran	Zagros	Toe pad					
Dendrocoptes medius	AMNH777492	Bakhtiare	Iran	Zagros	Toe pad					
Dendrocoptes medius	AMNH462278	Bakhtiari	Iran	Zagros	Toe pad					
Dendrocoptes medius	AMNH777493	Bakhtiare	Iran	Zagros	Toe pad					
Dendrocoptes medius	AMNH777495	Bakhtiare	Iran	Zagros	Toe pad					
Dendrocopos major						KC813280	NC_028174	DQ923085	DQ479210	KF445363

Table S2. List of all primers used in this study with their name, sequence, reference and annealing temperature used for PCR reactions.

Gene	Primer Name	Sequence 5'-3'	Reference	Annealing Temperature
mtDNA Control Region 1 D-Loop	DenD1F	ANG CCA AAG ACT GAA GAN ACA CA	Rutkowski *et al.* (2008)	55°C
	DenD1R	GGC NGT TAC CAT GGA CTT C		
Brahma Protein Intron 15	BRM15F	AGC ACC TTT GAA CAG TGG TT	Goodwin (1997)	54°C
	BRM15R	TAC TTT ATG GAG ACG ACG GA		
ATPase subunit 6	L9245	CCT GAA CCT GAC CAT GAA C	Eberhard, Berningham (2004)	56°C
	H9947	CAT GGG CTG GGG TCR ACT ATG TG		
	ATP1F	CTA TTC CCA GCC CTC CTC CT	This study	58°C
	ATP1R	GGA GAA GGT GGC CTA GGG A	This study	
	ATP2F	AAC ATA GCC CTT GCC TTC CC	This study	58°C
	ATP2R	ACG TAG GCT TGA ATC ATG GCT	This study	
Cytochrome b	L15424	CGA TTC TTC GCT TTA CAC TTC CTC C	Fuchs *et al.* (2008)	58°C
	H15916	ATG AAG GGA TGT TCT ACT GGT TG		
	CytB1F	CTG CCG ATT CAG GTG AGG AC	This study	58°C
	CytB1R	GCT ATT CTC ACC MAA CCT CCT	This study	
	CytB2F	GGG GGT GTG ACT AGG GGG	This study	58°C
	CytB2R	CAC TTC CTC CTC CCA TTC CT	This study	
Beta Fibrinogen Intron 7	Fib7F	ATC TCT GTG AGT TGC CAG CC	This study	58°C
	Fib7R	AAA GTG AAA GAT GAA CTA TAA GCA AAC	This study	

Table S3. Variables considered in the analysis. Variables with a variance inflation factor (VIF) > 3 were removed from further analysis (*) due to multi-collinearity with other variables.

Code	Variable (units)	VIF	
		1950-2000	22,000 years ago
Altitude	Altitude (m a.s.l.)	2.081	1.861
Slope	Slope (°)	1.807	1.737
Bio 1 *	Annual mean temperature (°C)	> 3	> 3
Bio 2	Mean diurnal range (°C) (Mean of monthly (max temp - min temp))	2.649	2.331
Bio 3 *	Isothermality [(BIO2/BIO7) *100] (°C *100)	> 3	> 3
Bio 4	Temperature seasonality (standard deviation *100)	1.281	1.582
Bio 5 *	Max temperature of warmest month (°C)	> 3	> 3
Bio 6 *	Min temperature of coldest month (°C)	> 3	> 3
Bio 7 *	Temperature annual range (BIO5-BIO6) (°C)	> 3	> 3
Bio 8	Mean temperature of wettest quarter (°C)	1.577	1.537
Bio 9 *	Mean temperature of driest quarter (°C)	> 3	> 3
Bio 10 *	Mean temperature of the warmest quarter (°C)	> 3	> 3
Bio 11 *	Mean temperature of the coldest quarter (°C)	> 3	> 3
Bio 12 *	Annual precipitation (mm)	> 3	> 3
Bio 13 *	Precipitation of wettest month (mm)	> 3	> 3
Bio 14 *	Precipitation of driest month (mm)	> 3	> 3
Bio 15	Precipitation seasonality (coefficient of variation)	1.811	1.523
Bio 16 *	Precipitation of wettest quarter (mm)	> 3	> 3
Bio 17 *	Precipitation of the driest quarter (mm)	> 3	> 3
Bio 18 *	Precipitation of the warmest quarter (mm)	> 3	> 3
Bio 19	Precipitation of the coldest quarter (mm)	2.264	1.788

Table S4. Standard genetic indices for each geographic area and marker.

Marker	Geographic area	n	Haplotypes (private)		Haplotype diversity (S.D.)	Nucleotide diversity (S.D.)
ATP6	Spain	6	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Central Europe	16	5 (3)		0.7000 (0.0795)	0.001680 (0.001326)
	Balkans	8	2 (0)		0.5714 (0.0945)	0.000948 (0.000970)
	Western Russia	3	3 (1)		1.0000 (0.2722)	0.002186 (0.002246)
	Anatolia	15	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Greater Caucasus	12	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Armenia	4	3 (0)		0.8333 (0.2224)	0.001842 (0.001824)
	Zagros	4	2 (0)		0.6667 (0.2041)	0.001228 (0.001378)
	West	33		6	0.6534 (0.0718)	0.001465 (0.001166)
	East	38		3	0.5562 (0.0534)	0.001126 (0.001013)
CytB	Spain	8	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Central Europe	16	3 (1)		0.6083 (0.0898)	0.003337 (0.002372)
	Balkans	9	2 (0)		0.5000 (0.1283)	0.002212 (0.001856)
	Western Russia	6	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Anatolia	19	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Greater Caucasus	11	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Armenia	8	2 (0)		0.4286 (0.1687)	0.000978 (0.001115)
	Zagros	5	3 (0)		0.8000 (0.1640)	0.002283 (0.002125)
	West	39		3	0.4251 (0.0868)	0.002239 (0.001707)
	East	43		3	0.5936 (0.0358)	0.001557 (0.001344)
ConReg	Spain	7	2 (0)		0.2857 (0.1964)	0.001803 (0.001882)
	Central Europe	20	2 (0)		0.4789 (0.0720)	0.003022 (0.002402)
	Balkans	9	3 (1)		0.6389 (0.1258)	0.005608 (0.004050)
	Western Russia	5	4 (2)		0.9000 (0.1610)	0.006940 (0.005349)
	Anatolia	9	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Greater Caucasus	12	3 (2)		0.3182 (0.1637)	0.002172 (0.002012)
	Armenia	5	2 (1)		0.4000 (0.2373)	0.003704 (0.003278)
	West	41		6	0.5646 (0.0783)	0.004086 (0.002900)
	East	26		5	0.6092(0.0679)	0.005953 (0.003940)
Fib7	Spain	12	2 (0)		0.1667 (0.1343)	0.000473(0.000570)
	Central Europe	26	3 (0)		0.1508 (0.0927)	0.000529 (0.000581)
	Balkans	12	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Western Russia	10	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Anatolia	18	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	Greater Caucasus	12	2 (1)		0.4848 (0.1059)	0.000690 (0.000717)
	Armenia	14	1 (0)		0.0000 (0.0000)	0.000000 (0.000000)
	West	60		3	0.0977 (0.0519)	0.000322 (0.000426)
	East	44		2	0.1691 (0.0711)	0.000241 (0.000365)

Table S5. Best-fitting substitution models.

Marker	Model
ATPase subunit 6 (ATP6)	TN93
Control Region 1 D-Loop (ConReg)	TN93
Cytochrome b (CytB)	TN93
beta-Fibrinogen intron 7 (Fib7)	HKY

Table S6. Model evaluation with 10-fold cross-validation of the 12 species distribution methods and their ensemble prediction (EP) for the Eastern and Western clade. Area Under the Curve (AUC) ranges between 0 and 1 (worse than a random model and best discriminating model, respectively). True Skill Statistic (TSS) and Boyce's Index (BI) ranges between −1 and 1 (higher values indicate a good predictive accuracy, while 0 indicates random prediction). Average values are shown (*: $P < 0.001$).

Model	Eastern clade			Western clade		
	AUC	**TSS**	**BI**	**AUC**	**TSS**	**BI**
ANN	0.906*	0.875*	0.989*	0.905*	0.815*	0.961*
BRT	0.991*	0.908*	0.991*	0.904*	0.841*	0.976*
CTA	0.988*	0.959*	0.991*	0.921*	0.841*	0.999*
FDA	0.923*	0.872*	0.952*	0.919*	0.869*	0.988*
GP-MAP	0.979*	0.871*	0.939*	0.906*	0.872*	0.999*
GAM	0.949*	0.876*	0.918*	0.921*	0.815*	0.999*
GLM	0.976*	0.843*	0.915*	0.978*	0.821*	0.964*
MADIFA	0.948*	0.816*	0.955*	0.928*	0.892*	0.917*
MARS	0.971*	0.819*	0.976*	0.944*	0.815*	0.999*
MAXENT	0.963*	0.879*	0.979*	0.983*	0.872*	0.999*
MAXLIKE	0.912*	0.874*	0.903*	0.983*	0.811*	0.985*
RF	0.991*	0.999*	0.983*	0.999*	0.989*	0.999*
EP	0.988*	0.926*	0.952*	0.922*	0.891*	0.999*

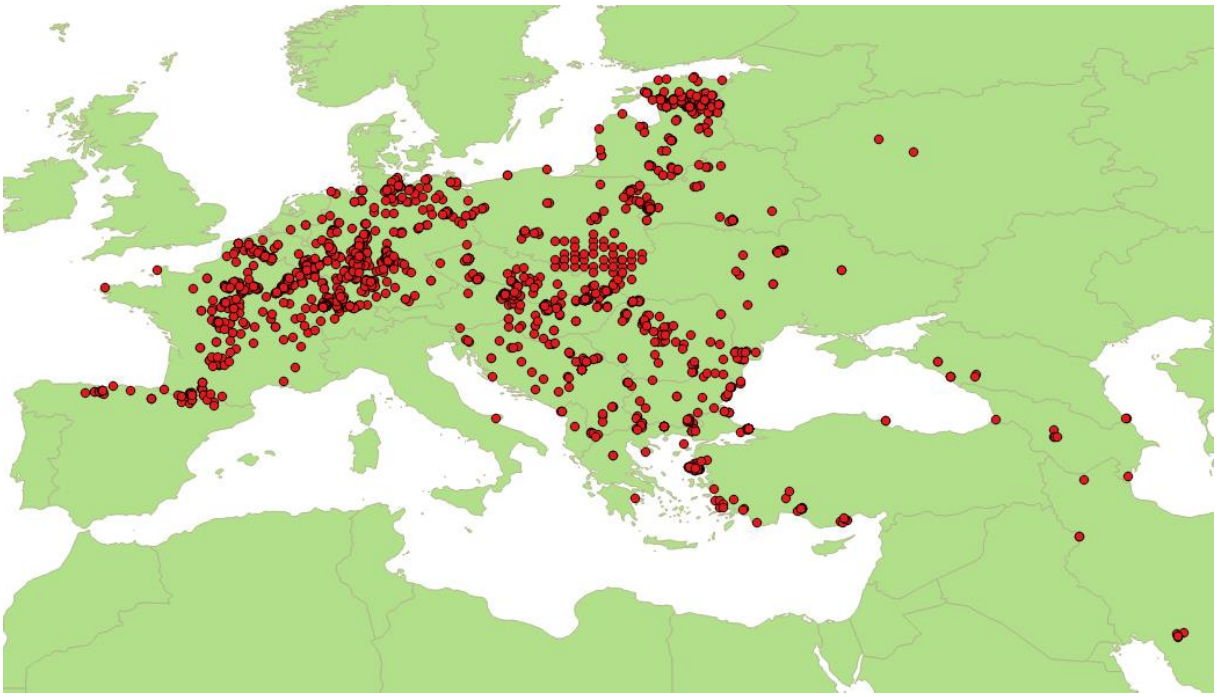

Figure S1. Occurrence point localities of the middle spotted woodpecker used for the maximum entropy niche modelling taken from GBIF databank (http://www.gbif.org/) and from www.observation.org. Observations between February and July (corresponding with the breeding period and season of high territoriality; Pasinelli *et al.* 2001) after 1980 and with a coordinate certainty of less than 500 meters were considered.

Laboratory Protocol

DNA extraction and PCR preparation were done in separate laboratories than post-PCR procession to avoid contamination from PCR products. All extractions from tissue, blood and feather samples were performed with the DNeasy Blood and Tissue kit (Qiagen) following the manufacturer's protocol. To gain as much DNA as possible from old toe pad samples, we used the sbeadex forensic kit (LGC Genomics). DNA extraction and PCR preparation for these toe pad samples were performed in a separate laboratory using different lab equipment than for fresh samples to minimize the risk of contamination from other samples. PCR amplification was done in 25 µl reaction volumes containing 12.5 µl GoTaq Hot Start Green Master Mix (Promega), 2 µl of each primer (10µM), 2 µl DNA and 6.5 µl ddH2O. PCR cycle started with 3 minutes at 94°C following a cycle of 0.5 minutes at 95°C, 0.5 minutes annealing and 1 minute at 72°C repeated for 40 times, followed by 7 minutes at 72°C and a final cooling at 10°C. The respective annealing temperatures for the different markers were adjusted depending on the melting temperature of the primer pairs (Table 1). We used agarose gel electrophoresis to confirm amplification of the target fragments. When PCR failed for single samples the whole reaction was redone and in the case of no success the extraction was redone.