

Pre-print of the final, accepted manuscript

Tóth EGy, Bede-Fazekas Á, Vendramin GG, Bagnoli F, Höhn M (in press): Mid-Pleistocene and Holocene demographic fluctuation of Scots pine (*Pinus sylvestris* L.) in the Carpathian Mountains and the Pannonian Basin: signs of historical expansions and contractions. *Quaternary International*. DOI: 10.1016/j.quaint.2017.11.024

Mid-Pleistocene and Holocene demographic fluctuation of Scots pine (*Pinus sylvestris* L.) in the Carpathian Mountains and the Pannonian Basin: signs of historical expansions and contractions

Endre Gy. Tóth¹, Ákos Bede-Fazekas^{2,3}, Giovanni G. Vendramin⁴, Francesca Bagnoli⁴, Mária Höhn¹ *

¹Szent István University, Dept. of Botany, 1118, Ménesi út 44., Budapest, Hungary, toth.endre@kertk.szie.hu, hohn.maria@kertk.szie.hu

²MTA Centre for Ecological Research, Inst. of Ecology and Botany, 2163, Alkotmány u. 2-4., Vácrátót, Hungary, bede-fazekas.akos@okologia.mta.hu

³MTA Centre for Ecological Research, GINOP Sustainable Ecosystems Group, 8237, Klebelsberg Kuno u. 3., Tihany, Hungary, bede-fazekas.akos@okologia.mta.hu

⁴CNR Institute of Biosciences and Bioresources, Via Madonna del Piano 10, 50019 Sesto Fiorentino (FI), Italy, giovanni.vendramin@ibbr.cnr.it, francesca.bagnoli@ibbr.cnr.it

* For correspondence. E-mail: hohn.maria@kertk.szie.hu,

Abstract

Climate fluctuations of the Quaternary caused radical changes in distribution of tree species and resulted in large-scale range shifts, population contractions and expansions. Scots pine (*Pinus sylvestris* L.) a widely distributed conifer of the boreal regions underwent spatio-temporal changes, which shaped the modern-day genetic structure and phylogeographic pattern of the species. By applying independent approaches, including molecular genetic data and historical climate models we aimed to describe demography and past distribution patterns of Scots pine populations from the highly fragmented southern periphery, the Carpathians and the Pannonian Basin. We used Approximate Bayesian Computation (ABC) approach based on nuclear microsatellite markers (nSSRs) and Maximum Entropy distribution modelling (MaxEnt) with temperature- and precipitation-related bioclimatic data. ABC results indicated that from an ancestral Scots pine population two genetic lineages have diverged that in the Mid-Pleistocene due to the favourable climatic conditions underwent population expansion leading to an admixture event. The outcome of the hindcasting confirmed the expansion that led to the admixture event revealed by the ABC analysis. This can be dated to the Late Glacial period (14,160–11,800 yrs BP), in which widespread distribution of Scots pine in accordance with palynological proxies was detected. Predictions for the Mid-Holocene period have shown large-scale reduction in distribution of Scots pine and low probability of its occurrence, leading to disjunction and population fragmentation.

Keywords: Carpathians, ABC, nuclear microsatellites, demography, distribution modelling, Maximum Entropy

1. Introduction

Molecular genetic surveys combined with fossil records–based historical reconstructions of the vegetation are nowadays the ultimate tools to reveal evolutionary history of species (Lascoux et al., 2004; Taberlet and Cheddadi, 2002; Slater et al., 2012). Coupling of available DNA information (e.g. diversity, differentiation, gene flow) with past distribution of species' populations has been increasingly used (Magri et al., 2006; Cheddadi et al., 2006), to reveal genetic implications of historical events and to circumscribe glacial refugia of species (Cheddadi et al., 2006; Zinck and Rajora, 2016).

By evaluating species demography and distribution with molecular and statistical approaches throughout an evolutionary time-scale, we are able to link the results with fossil evidences. Studying species' demography (ancestral dynamics of effective population size) is important in several contexts. It allows to identify population decline, expansion and divergence and by these, to infer various historical climatic oscillations in population demography at geological time scale (Lorenzen et al., 2011; Boitard et al., 2016). Alternatively, distribution modelling is also suitable to confirm historical events. Predictive species distribution modelling has an important role in ecology, biogeography, conservation biology and evolutionary history (Thuiller, 2007; Marmion et al., 2009).

During the Quaternary period European tree species were characterized with rapid large-scale range shifts, population contractions and extinctions (Petit et al., 2008). This was particularly experienced at the southern latitudes, where forests suffered remarkable altitudinal shifts and range fragmentation, unlike the mid-latitudinal continental sites, which were marked with larger forest development (with species re-shuffling and changes in abundances) after the end of the last glaciation (Colinvaux et al., 1996). Forest establishment on high-latitudinal regions (northern), previously being ice-covered, are estimated for the early Holocene period as a result of population colonization from the southern glacial refugia (Magri et al., 2006; Petit et al., 2008; Svenning et al., 2008).

Scots pine (*Pinus sylvestris* L.) is a widely distributed coniferous tree, growing primarily along the temperate boreal forest belt in Eurasia (Giertych and Mátyás, 1991; Matías and Jump, 2012). In Central-Eastern Europe, particularly in the Carpathian Mountains and the Pannonian Basin, species' natural distribution dramatically changed during the Quaternary period, resulting today in fragmented, isolated and relict populations, evidenced also by pollen records (Willis, 1994; Feurdean et al., 2007, 2011). Climatic oscillations and human activities largely influenced present distribution of Scots pine (Willis et al., 1998; Sümegei et al., 2002), the species was affected by interspecific competitive displacement (by *Picea abies* and *Quercus* spp.), that was indirectly favored by the climate change (Feurdean et al., 2007). As a result Mid-Holocene postglacial climate warming, forced Scots pine to immigrate into edaphically specialized habitat types and the populations survived only in sporadic small habitats with divergent (extreme) ecological conditions, including humid, cool peatbogs and sunny, dry, rocky outcrops (Pócs, 1960; Giertych and Mátyás, 1991).

Modern-day molecular genetic studies performed on Carpathian Scots pine populations revealed high level of genetic diversity and low level of genetic differentiation despite fragmentation and isolation by distance (Bernhardsson et al., 2016; Tóth et al., 2017). However, genetic barriers were also detected acting in restricting the gene flow between the spatially/geographically segregated genetic groups. Our former study, has detected structuring in the Carpathian populations, defining genetic groups identified with nuclear genetic markers. These were (1) Western Carpathians with the Pannonian Basin and the (2) Central-Island Mountains (Apuseni Mts.) with the Eastern Carpathians (Tóth et al., 2017).

Distribution modelling of forest tree species based on climate data has been successfully applied to investigate not only the future perspectives, but also the past historical events, i.e. hindcasting, (Svenning et al., 2008). Although, several conifer taxa have been studied by distribution modelling (van Zonneveld et al., 2009; Worth et al., 2013; Bede-Fazekas et al., 2014; Aguirre-Gutiérrez et al., 2015; Bede-Fazekas et al., 2015), fewer number of evaluations were focusing on past historical predictions (Alba-Sánchez et al. 2010). To our knowledge, Scots pine was only analyzed for large-scale predictions

to circumscribe historic habitat suitability and to estimate relative probability of presence (Cheddadi et al., 2006; Houston et al., 2016). Therefore, the historic distribution in small-scale details for the Central-European region based on long-term climatic variables in different evolutionary time periods, is yet unknown. Hindcasting is a tool increasingly used in phylogeographic studies for confirming or supplement the results achieved by traditional methods.

Relying on the former molecular results, our principal aim was to elucidate the spatio-temporal demographic and distributional changes in Scots pine from the Carpathians and the Pannonian Basin by independent approaches. Therefore, we (1) circumscribe historical divergence, expansion and contraction based on molecular genetic data and (2) reveal distributional changes by hindcasting past periods. Furthermore, we interpret the results of these independent approaches in the light of available fossil evidences.

2. Materials and methods

2.1. Regional setting and plant material

16 natural populations were collected from the Carpathian Mountains and the Pannonian Basin. The selected stands originate from the fragmented peripheral Scots pine populations from Central-Eastern Europe (Fig. 1, Table 1). Altogether, 331 individuals were analysed with nuclear simple-sequence repeat (SSR) markers. Five populations from Western Hungary (Pannonian Basin), four from the Slovakian Lower Tatra, and bottom of the westernmost ridges of the High Tatra (Western Carpathians), six populations from the Romanian Carpathians, where three originates from the Transylvanian Central Island Mountains (Apuseni), one from the Southern Carpathians and four from the Eastern Carpathians. Moreover one population from the Bulgarian Rila Mountains was involved. All of our sampling plots are from natural sites included in the frame of Nature Reserves, none of them managed by forestry, accordingly Scots pine regenerates naturally. Most of the populations persists in extreme edaphic conditions, like dry, sunny outcrops with low soil availability, acidic peatbogs or dry sandy substrate with low nutrient content.

One-year old needles were sampled from mature trees (8–30 from each population) with at least 30 m isolation distance between the individuals. However, in some of the populations samples are taken from shorter inter-individual distances due to relatively small population size (ranged from 0.02 to >5 km²) and limited sample availability. Plant material was stored on silica gel and frozen at -80 °C until DNA extraction.

2.2. Laboratory methods

Total genomic DNA was extracted from 20–25 mg of plant material by using the DNeasy Plant Mini Kit (QIAGEN Inc., Valencia, CA, USA) protocol. Eight nuclear microsatellite markers were used, which of two (SPAG 7.14, SPAC 11.4) developed by Soranzo et al. (1998) and six (psyl16, psyl17, psyl19, psyl36, psyl42, psyl57) by Sebastiani et al. (2012). All loci proved to be polymorphic in our populations, and were used within this study. Forward nuclear primers were fluorescently labelled with 6-FAM (SPAG 7.14, SPAC 11.4, psyl17, psyl19, psyl42, and psyl57) and NED (psyl16, psyl36).

PCR was performed in a 25- μ l reaction volume containing 20–80 ng DNA, 10X PCR reaction buffer, 2.5 mM MgCl₂, 0.02 mM dNTP mix, 2.5 μ mol of each primers, 1 unit of Taq DNA polymerase (Waltham, MA, USA) and Milli-Q ultrapure water. PCR reaction was performed in a thermocycler (Swift™ MaxPro Thermal Cycler, Esco Healthcare Pte, Singapore). Amplification protocol greatly varied between the primer pairs. Touch-down PCR protocol was carried out for SPAC 11.4 as follows: initial denaturation at 94 °C for 3 min, followed by 5 cycles of 30 s at 94 °C, 30 s at 65 °C and 30 s at 72 °C (decreasing the annealing temperature 1 °C per cycle), followed by 25 cycles of 30 s at 94 °C, 30 s at 60 °C and 30 s at 72 °C, and a final extension step of 5 min at 72 °C. Protocol for SPAG 7.14 consisted of an initial step of 94 °C for 3 min, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min, followed by 72 °C for 10 min. The amplification profile for psyl16, psyl17, psyl19, psyl36 and psyl42 consisted a first step at 94 °C for 4 min, and 35 cycles (94 °C for 30 s, 55 °C for 30

s, 72 °C for 40 s), followed by a final 72 °C for 8 min. Similar protocol was used for *psyl57*, but with an annealing temperature of 56 °C.

Electrophoresis to detect PCR products was carried on a 1 % (w/v) ethidium bromide-stained agarose gel in 1xTBE buffer. After initial amplification, strong yield of PCR product was diluted 25 to 30 times for fragment sizing, which were performed in an automated sequencer ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Band scoring was analysed by using Peak Scanner software 1.0 (Applied Biosystems 2006), and all size scores were visually checked.

2.3. Demographic history inferred by ABC

Prior to the demographic analysis, the Bayesian clustering approach implemented in STRUCTURE 2.3.4 (Pritchard et al., 2000) was used to infer groups or subpopulations in the dataset. We performed the analysis with an admixture model with correlated allele frequencies and a LOCPRIOR setup. This method uses sampling locations as prior information in the case of a relatively weak signal of structures (Hubisz et al., 2009). K value was set to 1-10 with a burn-in period of 10^5 steps followed by 10^6 repetitions of Markov Chain Monte Carlo (MCMC). Fifteen repetitions were set for each run. The web-based STRUCTURE HARVESTER (Earl, 2012) was used to apply the Evanno method (Evanno et al., 2005) to detect the value of K (the number of genetic groups) that best fit the data. The 15 simulations were averaged using CLUMPP v1.1.2 (Jakobsson and Rosenberg, 2007) and represented in the form of bar graphs using POPHELPER (Francis, 2016).

Expected heterozygosity (H_e) the standard estimator of genetic diversity was calculated for all genetic groups using GenAIEx v.6.5 software (Peakall and Smouse, 2006). Analysis of molecular variance (AMOVA) implemented in Arlequin v.3.5 software (Excoffier and Lischer, 2010) was used to determine the partition of the genetic variation among genetic groups, within and among populations. Significance tests were evaluated using a permutation approach with 999 replications.

Analysis of the historical demographic changes was studied with Approximate Bayesian Computation implemented in DIY ABC v2.0 software (Cornuet et al., 2014). We used the nuclear microsatellite data to estimate possible population divergence time, changes in the effective population sizes, and admixture events. ABC analysis allows to compare complex historical scenarios to infer past history of populations. To perform ABC analysis we defined three populations according to STRUCTURE results: Pop1 (Hungarian; HFE, HVE, HZA, HOR, HKO, Slovakian; SKV, SME, STU, SLI, and Romanian; RCO, RBI, RPA), Pop2 (Romanian; RFE, RPO, RMO), and the admixed Pop3 (Bulgarian–BYU). Population relationships design in the scenarios were mainly based on the result of STRUCTURE analysis and on calculated standard population genetic indices (e.g. H_e , F_{ST}). Several pilot runs were carried out to produce final competing scenarios. In all scenarios, Pop1 was specified as the population to be traced back to an ancestral population (as required for DIY ABC analysis), because this population had the highest heterozygosity (H_e). Six different scenarios were examined following a two-step DIY ABC procedure.

In the first step (Fig. 2A), the three populations (Pop1, Pop2 and Pop3) were analysed to infer historical divergences and times of admixture event, and to reveal history of recent population structure identified by STRUCTURE. The second analysis (Fig. 2B) focused on determining changes within population size by estimating effective sizes, and to refine potential ancient expansion-contraction in species history, based on the scenario selected by the first analysis. In all tested scenarios, $t\#$ refers to timescale expressed as generation time, and $N\#$ to effective population size of the corresponding population.

The scenarios are as follows (Fig. 2A):

Scenario 1 (hereafter: Sc1): (Hierarchical split model 1): Pop1 and Pop2 diverged at t_2 from an ancestral population N_a , and Pop3 diverged from Pop1 at t_1 .

Scenario 2 (Sc2): (Hierarchical split model 2): Pop1 and Pop2 diverged at t_2 from ancestral population N_a , and Pop3 diverged from Pop2 at t_1 .

Scenario 3 (Sc3): (Simple split model): Pop1 diverged from N_a at t_2 , and Pop2 and Pop3 diverged from Pop1 at t_1 .

Scenario 4 (Sc4): (Isolation with admixture model 1): Pop1 and Pop2 diverged at t_2 from ancestral population N_a , and Pop3 was generated by admixture of Pop1 and Pop2 at t_1 .

The most possible scenario (Sc4) was tested for demographic changes assuming an expansion at both Pop1 and Pop2.

The two final scenarios are as follows (Fig. 2B):

Scenario 5 (Sc5): (Isolation with admixture model 1): same as Scenario 4.

Scenario 6 (Sc6): (Isolation with admixture model 2): we assumed Pop1 and Pop2 expansion from the ancestral population. N1c and N2c represent a variation in the effective population sizes of Pop1 and Pop2, respectively, and are set bigger than Na and smaller than Pop1 and Pop2, respectively.

In all simulations, Generalized Stepwise Mutation model (GSM; Estoup et al., 2002) and a standard mutation rate of min. 1×10^{-3} and max. 1×10^{-4} were used (Table A.1, A.3). Mean number of alleles (N_A), genetic diversity (H_e) and size variance was calculated as summary statistics for single populations, and mean number of alleles, genetic diversity, size variance and differentiation (F_{ST}) for population pairs. 10^6 simulation was performed for each scenario, after the most likely scenario was evaluated by comparing posterior probabilities using logistic regression (Table A.4). Model checking option (Table A.2) was used to test the goodness of fit of the scenarios with Principal Coordinates Analysis (PCoA) (Fig. A.1) to measure deviation between simulated and real data.

2.4. Climate data

Observed and modelled climate surfaces were obtained from the WorldClim 1.4 database (Hijmans et al. 2005). Instead of the raw climate data, 19 bioclimatic variables (Nix, 1986) were used, that are preferred predictors of potential distribution of species in case of hindcasting (e.g. Mayol et al., 2015), prediction to the reference period (e.g. Kabaš et al., 2014; Yuan et al., 2015), and forecasting (eg. Ali et al., 2014; Somodi et al., in press) as well. For the reference period (1960–1990) observed and interpolated climate surfaces at 30 arc second horizontal resolution were used. Model prediction of ENCAR Community Climate System Model (CCSM) provided the climate data of the Mid-Holocene (c. 6000 yrs BP, Hijmans et al., 2005, www.cesm.ucar.edu/models/ccsm4.0), the Last Glacial Maximum (LGM, c. 22,000 yrs BP, Hijmans et al., 2005, www.cesm.ucar.edu/models/ccsm4.0), and the Last Inter-Glacial (LIG, 140–120,000 yrs BP, Otto-Bliesner et al., 2006; Kiehl and Gent, 2004) at resolutions of 30 arc second, 2,5 arc minute, and 30 arc second, respectively.

2.5. Distribution modelling

Final predictor set for model building was selected from the 19 bioclimatic variables based on pairwise correlation test (Pearson's r), multicollinearity test (Condition Number (CN), Variance Inflation Factor (VIF)), and ecologically-informed expert decision. The final predictor set contains four temperature-related and five precipitation-related variables: bio1 (Annual Mean Temperature), bio3 (Isothermality), bio7 (Temperature Annual Range), bio9 (Mean Temperature of Driest Quarter), bio15 (Precipitation Seasonality), bio16 (Precipitation of Wettest Quarter), bio17 (Precipitation of Driest Quarter), bio18 (Precipitation of Warmest Quarter), bio19 (Precipitation of Coldest Quarter). Acceptability of the multicollinearity and correlation measures of the final predictor set (CN = 17.59, max(VIF) = 47.61, mean(VIF) = 15.58, max(|rl|) = 0.73) was confirmed by the literature (Dormann et al., 2013; Elith et al., 2006; O'Brien, 2007; Chennamaneni et al., 2016).

Modelling was carried out in R Statistical Software (R Core Team, 2017). We build Maximum Entropy Model ver. 3.3.3k (MaxEnt, Phillips et al., 2006) implemented in the dismo package of R (Hijmans et al. 2016), which is one of the most widely used species distribution modelling algorithm in the field of phylogeography and phylogenetics (e.g. Carnaval and Moritz, 2008; Cordellier and Pfenninger, 2009; Mayol et al., 2015). MaxEnt is able to handle incomplete, presence-only training data (Phillips et al., 2006; Elith et al., 2011). For the two studied populations, two independent models were trained in the reference period, using presence-only records ($n_1 = 12$, $n_2 = 4$) with 50,000 pseudo-absence (background) data sampled from the study area. In contrast to presence-absence models, presence-pseudoabsence methods, such as MaxEnt, are not able to distinguish false negatives and false positives definitely from the true ones (Lobo et al., 2007; Gormley et al., 2011). Therefore, since MaxEnt is prone to overfitting (Phillips and Dudík, 2008; Baldwin, 2009), small training presence data (which is the case of Pop2) may result in over or under-prediction when, relative to the environmental hyperspace of training period, significant extrapolation is needed (which is the case of LGM). Although

models based on presence-only data face with several issues of their methodology and interpretability, nonetheless, the modelling without real absence data may be informative (Soberón et al., 2000; Graham et al., 2004; Elith et al., 2006; Zuckerman et al. 2011). The domain of the model building was 10°–30°E and 40°–55°N (WGS-84 system), while the prediction domain was 10°–30°E and 40°–51°N for all the four prediction period. We used the default convergence threshold (10^5) and logistic output that provides predicted probabilities between 0 and 1.

3. Results

3.1. Spatial structure

Bayesian clustering for genetic assignment evaluated with STRUCTURE indicated that the most likely number of groups of populations was $K=2$, supported by the ΔK criterion (Evanno et al., 2005). At $K=2$ the two identified genepools were: (Pop1) Western Hungarian with Slovakian populations accompanied by populations from the Apuseni Mts. and the Southern Carpathians, (Pop2) Eastern Carpathian populations. Although, the Bulgarian population shared the group with the Eastern Carpathian populations, this population presented a high scale of admixture, hence it has been considered as a distinct group within the ABC analysis (Pop3).

Pop1 presented the highest expected heterozygosity ($H_e = 0.615$, SE: 0.027) while Pop2 and Pop3 showed slightly lower values ($H_e = 0.511$, SE: 0.054 and $H_e = 0.584$, SE: 0.099). AMOVA analysis between Pop1 and Pop2 resulted in $F_{ST}=0.084$ ($p<0.001$), which was the highest compared to the other population pairs, meaning that over 8% of the genetic variance resides among and 92% within populations. Pop2 vs. Pop3 and Pop1 vs. Pop3 resulted much lower differentiation ($F_{ST}=0.031$ ($p<0.001$) and $F_{ST}=0.058$ ($p<0.001$)), meaning that 3% and 6% of genetic variance resides among and 97% and 94% within populations, respectively (Table 2).

3.2. Demographic history

In the first ABC analysis (Fig. 2A) the most likely scenario resulted Sc4, the admixture model, where posterior probability (0.7914; 95% CI: 0.7671–0.8156) was significantly higher than that for Sc1 (0.1909; 95% CI: 0.1670–0.2148), Sc2 (0.0145; 95% CI: 0.0083–0.0206) and Sc3 (0.0033; 95% CI: 0.0000–0.0088). In Sc4, the effective size of Na ancestral population was determined to be 479, and population sizes for N1 (Pop1), N2 (Pop2) and N3 (Pop3) were 16,000, 5540 and 5370, respectively. Estimated time for ancient divergence t_2 (divergence of Pop1 and Pop2) was 3090 (95% CI: 867–8170) generations ago, and for recent admixture event and then divergence (Pop1-2-3) at t_1 was 180 (95% CI: 23.6–823) generations ago. Assuming 50 years as generation time for extreme peripheral populations of Scots pine, t_2 value scaled to 154,500 yrs, and t_1 to 9000 yrs BP. Considering a longer 60 years generation time, these increases to 185,400 yrs for t_2 , and 10,800 yrs for t_1 . Out of the 21 summary statistics, only 1 showed a significant difference between the observed and simulated data based on the posterior distributions, furthermore PCA analysis also showed that both simulated prior and posterior plots were clustered around the observed dataset, suggesting that scenario 4 generally fitted to the observed dataset. In case of Sc4, estimated Type I error rate was 0.444 and the average Type II error rate was 0.0766.

The results of the second analysis (Fig. 2B, 3A-B) indicated that Sc6 had the highest posterior probability (0.5670; 95%CI: 0.5307–0.6033), whereas Sc5 was less supported (0.4330; 95%CI: 0.3967–0.4693). The median values of the effective population sizes for the best scenario indicated that the ancestral population (Na, estimated to 430 (95%CI: 16.3–1790) was 38.37, 14 and 13.58 times lower than Pop1, Pop2 and Pop3 actual populations. Accordingly, growth in effective population sizes showed that both Pop1 and Pop2 increased from Na to 5790 (95%CI: 1180–9770) and 2520 (95%CI: 423–7350) by t_2 -db time (915; 95%CI: 59.8–1920). Assessment of most recent population sizes resulted a further increase to 16,500 for N1, 6020 and 5840 for N2 and N3, respectively. Determined divergence times slightly differed from Sc4, and resulted t_2 to 3560 (95%CI: 1040–8810) and t_1 to 236 (95%CI: 29.4–876) in the final scenario. Assuming 50 years as generation time t_2 value scaled to 178,000 yrs and t_1 to 11,800 yrs BP. With a longer, 60 years assumed generation time t_2 value scaled to 213,600 yrs and t_1 to 14,160 yrs BP.

3.3. Distribution modelling

Based on the known occurrences of Scots pine, we generated geographic distribution maps predicting areas where species might occurred in divergent time periods within the Quaternary, in the Pleistocene and the Holocene periods. The niche modelling accurately predicted the observed distribution for both defined genetic groups in the reference period (Fig. 4). In case of Pop1, the model selected linear/quadratic features and converged after 180 iterations, while in case of Pop2, only linear feature was used and the algorithm converged after 100 iterations. Pop1 model was more balanced (relying on four predictors, bio3, bio9, bio15 and bio18, which contributed to the model more than 20%) than Pop2 (Table 3).

Since (1) bio1 (annual temperature) has outstanding contribution (>50%) to the Pop2 model (Fig. A.2, A.3), and (2) no lower limit could be found by the model within the training area, the inverse relationship of the bioclimatic variable and the predicted probability can cause overestimation of the predicted distribution in cold regions/periods. This can be observed in case of the LGM period in the Northern part of the studied area, where the prediction might be unreliable.

The predicted distribution showed continuous presence of *P. sylvestris* in all studied geographic regions within the Carpathian Mountains and the western Pannonian Basin. Maps also indicated that Scots pine had potentially suitable habitats in all time periods, including the Last Inter-Glacial, the Last Glacial Maximum, the Mid-Holocene and also the Reference time period.

Although, we have no available climate data estimations for the early to Mid-Pleistocene time periods, LIG (c. 140,000–120,000 yrs BP) map showed presence of Scots pine in the Western and Eastern Carpathian regions, likewise in the western Pannonian Basin. The predicted probability of presence (0.3–0.6) was reduced, compared to the map dating from the LGM, but the Carpathian region showed to be separated forming distinct units.

The map predicting the LGM distribution (c. 22,000 yrs BP) was only suitable for Pop1. The prediction revealed that the potential distribution is larger than identified in the Mid-Pleistocene and formerly separated regions are merged. Also, result indicate higher probability values (0.4–0.9). At this time, species' presence is continuous along the Carpathians and the Eastern Alps.

The Mid-Holocene (c. 6000 yrs BP) prediction indicated reduction in Scots pine presence (0.4–0.8), and segregated the formerly joined distributions. The highest probability (0.5–1.0) was estimated for the Eastern Alps and to the pre-Alpine regions of the western Pannonian Basin.

The reference time period (1960–1990) indicated large areas potentially available for Scots pine and showed elevated probability of its presence (0.5–0.9) for both genetic groups (Pop1 and Pop2), hence higher potential availability outside the species' current natural distribution range.

4. Discussion

Demographic history of Scots pine using molecular markers and involving macrofossil and pollen remains has been studied formerly across the European range (Cheddadi et al., 2006; Pyhäjärvi et al. 2007). However, the species' distribution pattern in the Quaternary on a regional scale including the Carpathian Mountains and the Pannonian Basin, has not been investigated yet by coalescent-based genetic analysis using Approximate Bayesian Computation approach and distribution modelling with climate data. Fossil records from the east-central European region are available showing putative history of the species, like in the case of the Southern-Carpathians (Farcas et al., 1999; Magyari et al., 2012) or Northern and Western Carpathians (Magyari et al., 2014a, b). Therefore, the results of our ABC analysis and distribution modelling can provide valuable and synergic information for linking fossil evidence to present day genetic pattern of Scots pine.

4.1. Historical demographic estimation

The ABC analysis using eight nSSR loci supported an admixture scenario, the Sc6, in which the two main detected gene pools (Pop1 the Hungarian; HFE, HVE, HZA, HOR, HKO, the Slovakian; SKV, SME, STU, SLI, and the Romanian; RCO, RBI, RPA) and Pop2 the Romanian; RFE, RPO, RMO) have been separated at the same time, rather than the hierarchically split gradual divergence or simple split scenarios (Fig. 3A-B). Pop3 (Bulgarian population), as a highly mixed population detected in our STRUCTURE analysis genetically infer with the two main populations Pop2 and Pop1 in distinct times

presumably due to an admixture event. Admixture event has been confirmed along the run of demographic estimations, where gradual expansion of the populations were detected. This might have caused admixture of Pop1, Pop2 and Pop3, respectively. The estimated divergence times are strongly affected by the generation time of Scots pine, and generally can greatly vary in conifers. Grotkopp et al. (2002) has estimated 5 years minimum generation time (MGT) while Provan et al. (1999) up to 100 years as generation time for *Pinus sylvestris*. Therefore based on our experiences we assumed a generation time to be approx. 50–60 years under extreme environmental conditions where the studied populations grow. If we assume this generation time, the first divergence time (t_2), from the ancestral population, falls within 213,600 and 178,000 yrs BP and the admixture event (t_1) to 14,160 and 11,800 yrs BP. Furthermore, we detected a population expansion taking place from t_2 time, when diverged populations expanded and their effective population size increased from 430 individual up to 16,500 (Pop1), 6020 (Pop2) and 5840 (Pop3) by the time of the admixture event.

Although, it is hard to make conclusions due to the lack of long time pollen records (going back to Pleistocene) from the Carpathians, there are strong evidences that *Pinus* (diploxylon) species were dominating from the Mid-Pleistocene's transition towards the glacial until the early Holocene interglacial period. Deep pollen cores from the Tenaghi Philippon peatland in Greece, showed an overall increase of *Pinus* pollen for the first time by 10% from 129,000 yrs BP, which later steadily increased to 45% by 113,000 BP (Milner et al., 2013). Moreover, sedimentological proxies from a recent study by Sadori et al. (2015) from lake Ohrid (western Balkan region, Albania) highlighted the high abundance of *Pinus* pollen, 10–87% between 245–189,000, 14–83% between 161–121,000 and 9–77% between 70–12,000 yrs BP. Accordingly, the rate of the *Pinus* pollen remained high during the Mid-Pleistocene until the LGM (Bertini et al., 2016). These findings fit well to our detected population expansion by molecular markers causing extensive distribution of the species. Expansion might have caused an admixture event among the gene-pools that provided later the nowadays geographically distant populations. Since there is close relation between ice volumes, climate and forest expansion/contraction (Tzedakis et al., 2006), it is certain that due to the favourable climate of the long lasting glacials coniferous species have maintained their population sizes in the region, despite the upper-Pleistocene's warmer interglacial and the short dry/wet climate oscillations.

Our estimated admixture event for Scots pine, based on the ABC analysis, might have happened between 14,160 to 11,800 yrs BP (Sc6, Table A.1), when Scots pine displayed a vast expansion, at the Late Glacial period, when *Pinus* pollen percentages were at their maximum (Feurdean et al., 2011). Although, coniferous species like *Picea*, *Larix*, *Juniperus* including *Pinus sylvestris* survived the LGM in sparse populations restricted to areas with favourable environmental conditions in the Carpathians and in the Pannonian Basin (Rudner et al., 1995; Rudner and Sümegi, 2001; Magyari, 2011), a strong reduction of conifers and expansion of deciduous species has started only from the early and middle Holocene period (Feurdean et al., 2012). This decline in *Pinus* pollen abundance has been detected at several sites along the Carpathian mountain range (Tantau et al., 2003, 2006, 2014; Feurdean and Bennike, 2004; Feurdean et al., 2007), suggesting, that after the expansion and admixture, populations have suffered contraction causing fragmentation and reduction in population sizes.

Pollen sequences from the Bulgarian mountain massifs (including Rila and Pirin) indicate that diploxylon-type *Pinus* pollen varied greatly over time and space. Although, there is evidence for our detected admixture period from seven Rila mountain lakes (geographically close to our sampling site) presenting high abundance of *P. diploxylon*-type pollen (60–68%) between 15,100–12,900 yrs BP (Tonkov et al., 2011; 2013), other pollen sequences identify high proportion of pollen (over 50–55%) and enlargement of *Pinus* stands earlier about 16,500 yrs BP (Stefanova et al., 2006; Tonkov et al., 2008). A re-increase of pollen concentration has happened also in the Mid-Holocene period, while in the meantime multiple sites indicate only reduced level of arboreal pollen abundance attributed mainly to *P. sylvestris* and *P. mugo* and *P. peuce*. This second major re-increase characterising high elevation sites (above 2000 m a.s.l.) started about 7500–6700 yrs BP, when the favourable climatic and edaphic conditions facilitated the formation of the coniferous belt dominated in the upper part by *P. sylvestris* and *P. peuce* and in the lower part by *Abies alba* (Bozilova and Tonkov, 2000; Tonkov et al., 2002; Tonkov, 2003).

4.2. Potential historical distribution

Despite that the Last Inter-Glacial (LIG) period was not evidenced by our ABC analysis, climate based distributing modelling has indicated reduced probability and distribution of Scots pine in this period. This is in agreement with the former studies that have reported short and rapidly changing warming events in the LIG period (Boettger et al., 2009; Helmens, 2014). These warmer phases are characterized by reduction in cold adapted conifer trees, including *Pinus* taxa (Wolfe et al., 1999; Boettger et al., 2009). Warm tolerant tree species only dwelled in the early part of the interglacial interval, in which the mixed oak forest phase was dominating with *Quercus* and *Corylus* trees (Helmens, 2014). Additionally, the LIG Eemian period, to which our hindcasting period fits, it is characterized by an increase of herb pollen concentration and percentage (with xerothermic type grasses) in the northern outer Carpathians and Poland, accompanied also by a reduction in arboreal species content suggesting a common degradation of forest communities (Zagwijn, 1996; Komar et al., 2009; Majecka, 2014). The rapid and dynamic warming phases allowed only the development of open forests and scarce trees in forest-steppe coenoses (Novenko et al. 2008; Komar et al., 2009).

Our climate estimation for the LGM (c. 22,000 yrs BP) period provided strong evidences and revealed high probability distribution for Scots pine. This is congruent with the ABC analysis, in which an admixture of genetic lineages has been detected, as a result of an earlier expansion (Fig. 3, 4). Fossil evidences indicate high abundance of *Pinus* diploxylon pollen (including *P. sylvestris*) from the end of the LGM to the early Holocene, which could be aligned with our detected population expansion (Feurdean et al., 2011; Feurdean et al., 2014, 2015). Probably, Scots pine started to increase by the LGM and reached the largest extension in the Late Glacial and early Holocene period. Likewise, high pollen percentage was detected in the Southern Carpathians between c. 17,000–12,000 yrs BP in the Late Glacial interstadial, where glacial recolonization began with the development of *Pinus* forests (Tantau et al., 2006). This was also detected in the Eastern Carpathians between 15,000–10,000 yrs BP (Feurdean et al., 2007) and in the Western Carpathian region following the LGM (Petr et al., 2013). Studies indicate that the *Pinus* pollen concentration reached its maximum about 14,700–14,000 yrs BP shortly after the LGM (Feurdean et al., 2004, 2007; Tantau et al., 2006, 2014). Within the Bulgarian Rila Mountains higher altitudinal sites were colonized most probably earlier about 15,100–12,900 yrs BP (Tonkov et al., 2011; 2013). According to these, formerly dispersed and fragmented populations could have merged.

Reduction in distribution and lower occurrence probability of Scots pine was detected for the Mid-Holocene period (c. 6000 years BP) in the Carpathians, which is congruent with our ABC analysis and with our former SSR study on diversity and differentiation showing Holocene fragmentation and isolation of populations (Tóth et al., 2017). In the high mountains of Bulgaria, such as the Rila, distribution modelling showed a decreased but stable distribution of *P. sylvestris* in the Mid-Holocene, potentially because populations were able to sustain in the high montane sites as tree-line was higher than in present (Tonkov, 2003; Tonkov et al., 2008; Feurdean et al., 2014; Lazarova et al., 2015).

According to palynological evidences in the Mid-Holocene Scots pine abundance has become reduced not only in the Carpathian Mountains, but also on the Danubian lowland (Pannonian Basin), as a result of the spreading of broad-leaved woody species, drying and warming of the climate especially at low mountain elevations and foothill regions (Magyari et al., 2010; Willis et al., 1995, 1997). Our distribution analysis showed larger distribution than would be expected based on palynological records at sites in the north-east Hungarian region (Willis et al., 1995; Magyari et al., 2010) or the central Hungarian middle Tisza plain (Ronai, 1970). However patches of needle-leaf woodlands with open canopy were reported from the Transylvanian basin at the beginning of the Holocene (Feurdean et al. 2015) that most probably have declined following the droughts of hot summers and fire regimes, and since the Mid-Holocene or later the Bronze age, because of the increasing human activity.

An increased probability of presence for Scots pine shown for the recent time period (1960–1990) is due to range expansions of *Pinus* species (mainly *P. sylvestris*) as consequence of forest management evidenced also by palynological records in the Romanian Carpathians by Feurdean et al., (2011). For this period Tantau et al. (2003, 2014) have also reported modest increase in *Pinus* pollen percentages in the Eastern Carpathians.

5. Conclusion

Our combined modelling approaches provided evidences for historical range expansions, population admixture and later range contractions of *Pinus sylvestris* populations in East-Central Europe including the Carpathian sites and the Pannonian Basin. Based on two models we could predict species distribution beginning from the Mid-Pleistocene, Late-Interglacial, and the Last Glacial Maximum to Mid-Holocene. The divergence and expansion event from an ancient population leading to two genetic lineages revealed by the ABC in the period between 213,600–178,000 yrs BP cannot be detected by hindcasting, however latter predicted smaller distribution of the species between 140,000–120,000 yrs BP. Population expansion in the LGM indicated by the ABC was supported by hindcasting and predicted overall maximum expansion for Scots pine in the studied region. As the results of the expansion in the Late Glacial period, between 14,160–11,800 yrs BP, ABC showed admixture of the two lineages and indicated that Scots pine displayed a vast distribution in that period. Up to the Mid-Holocene, c. 6000 yrs BP, both models predict reduction in the species distribution that resulted in population fragmentation and disjunction. Accordingly, Scots pine in the Mid-Holocene period underwent a range contraction, being outcompeted by broad-leaf species that have expanded with the climate warming. Recent population structure is the result of the species' range contraction and fragmentation when Scots pine could only occupy edaphically specific habitat types. Our two models have been reconciled with palynological records that support the population demography described. Palynological studies have reported also spatio-temporal changes in the species' distribution within the last centuries influenced by the human activity.

Acknowledgements

We wish to thank the Editor for helpful comments and the Anonymous Reviewers who helped us improve the manuscript. The authors are also grateful to Tamás Pócs (Hungary), Ivan Iliev (Bulgaria), Tibor Baranec (Slovakia) for collection of plant material and for valuable discussions. GGv and FB were supported by the European Union's ProCoGen grant, under grant agreement n° 289841. EGYt and MH was supported by the National Research, Development and Innovation Office, Hungary by the grant of the Hungarian Scientific Research Fund [OTKA K101600]. ÁBF was supported by the GINOP-2.3.2-15-2016-00019 grant. Country borders were provided by EuroGeographics (original product (with terms of the licence) is available at www.eurogeographics.org).

Appendix A. Supplementary data

(Uploaded in a separate file.)

References

- Aguirre-Gutiérrez, J., Serna-Chavez, H.M., Villalobos-Arambula, A.R., Pérez de la Rosa, J.A., Raes, N., 2015. Similar but not equivalent: ecological niche comparison across closely-related Mexican white pines. *Diversity and Distributions* 21(3), 245–257.
- Alba-Sánchez, F., López-Sáez, J.A., Pando, B.B.D., Linares, J.C., Nieto-Lugilde, D., López-Merino, L., 2010. Past and present potential distribution of the Iberian *Abies* species: a phytogeographic approach using fossil pollen data and species distribution models. *Diversity and Distributions* 16(2), 214–228.
- Ali, K., Ahmad, H., Khan, N., Jury, S., 2014. Future of *Abies pindrow* in Swat district, northern Pakistan. *Journal of Forestry Research* 25(1), 211–214.
- Baldwin, R.A., 2009. Use of Maximum Entropy Modeling in Wildlife Research. *Entropy* 11, 854–866.
- Bede-Fazekas, Á., Horváth, L., Kocsis, M., 2014. Impact of climate change on the potential distribution of four Mediterranean pines. *Időjárás* 118(1), 41–52.

- Bede-Fazekas, Á., Horváth, L., Trájer, A.J., Gregorics, T., 2015. An ArcGIS tool for modeling the climate envelope with feed-forward ANN. *Applied Artificial Intelligence* 29(3), 233–242.
- Bernhardsson, C., Floran, V., Ganea, S.L., García-Gil, M.R., 2016. Present genetic structure is congruent with the common origin of distant Scots pine populations in its Romanian distribution. *Forest Ecology and Management* 361, 131–143.
- Bertini, A., Sadori, L., Combourieu-Nebout, N., Donders, T.H., Kouli, K., Koutsodendris, A., Joannin, S., Masi, A., Mercuri, A.M., Panagiotopoulos, K., Peyron, O., Sinopoli, G., Torri, P., Zanchetta, G., Francke, A., Wagner, B., 2016. All together now: An international palynological team documents vegetation and climate changes during the last 500 kyr at lake Ohrid (SE Europe). *Alpine and Mediterranean Quaternary* 29(2), 201–210.
- Boettger, T., Novenko, E.Y., Velichko, A.A., Borisova, O.K., Kremenetski, K.V., Knetsch, S., Junge, F.W., 2009. Instability of climate and vegetation dynamics in Central and Eastern Europe during the final stage of the Last Interglacial (Eemian, Mikulino) and Early Glaciation. *Quaternary International* 207(1), 137–144.
- Boitard, S., Rodriguez, W., Jay, F., Mona, S., Austerlitz, F., 2016. Inferring population size history from large samples of genome-wide molecular data-an approximate Bayesian computation approach. *PLOS Genetics* 12(3), e1005877.
- Bozilova, E.D., Tonkov, S.B., 2000. Pollen from Lake Sedmo Rilsko reveals southeast European postglacial vegetation in the highest mountain area of the Balkans. *The New Phytologist* 148(2), 315–325.
- Carnaval, A.C., Moritz, C., 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography* 35(7), 1187–1201.
- Cheddadi, R., Vendramin, G.G., Litt T., Francois L., Kageyama, M., Lorentz, S., Laurent, J-M., de Beaulieu, J-L., Sadori, L., Jost, A., Lunt, D., 2006. Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecology and Biogeography* 15, 271–282.
- Chennamaneni, P.R., Echambadi, R., Hess, J.D., Syam, N., 2016. Diagnosing harmful collinearity in moderated regressions: A roadmap. *International Journal of Research in Marketing* 33(1), 172–182.
- Colinvaux, P.A., de Oliveira, P.E., Moreno, J.E., Miller, M.C., Bush, M.B. 1996. A long pollen record from lowland Amazonia: forest and cooling in glacial times. *Science* 274(5284), 85.
- Cordellier, M., Pfenninger, M., 2009. Inferring the past to predict the future: climate modelling predictions and phylogeography for the freshwater gastropod *Radix balthica* (Pulmonata, Basommatophora). *Molecular Ecology* 18(3), 534–544.
- Cornuet, J-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J-M., Estoup, A., 2014. DIY ABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism. DNA sequence and microsatellite data. *Bioinformatics* 30, 1187–1189.
- Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., García Marquéz, J.R., 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(1), 27–46.
- Earl, D.A., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4(2), 359–361.
- 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.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(2), 129–151.

- Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17(1), 43–57.
- Estoup, A., Jarne, P., Cornuet, J.M., 2002. Homoplasy and mutation model at microsatellite loci and their consequence for population genetics analysis. *Molecular Ecology* 11, 1591–1604.
- EUFORGEN, 2009. Distribution map of Scots pine (*Pinus sylvestris*), Rome. Italy. URL: www.euforgen.org.
- Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14(8), 2611–2620.
- Excoffier, L., Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10(3), 564–567.
- Farcas, S., de Beaulieu, J-L., Reille, M., Coldea, G., Diaconeasa, B., Goeury, C., Goslar, T., Jull T., 1999. First 14 C datings of Late Glacial and Holocene pollen sequences from Romanian Carpathes. *Comptes Rendues de l'Académie des Sciences de Paris* 322, 799–807.
- Feurdean, A., Bennike, O., 2004. Late Quaternary palaeoecological and palaeoclimatological reconstruction in the Gutaiului Mountains, northwest Romania. *Journal of Quaternary Science* 19, 809–827.
- Feurdean, A., Björkman, L., Wohlfarth, B., 2012. A paleoecological reconstruction of the Late Glacial and Holocene based on multidisciplinary studies at Steregoiu site (Gutai Mts., Romania). *Studia UBB Geologia* 46, 125–140.
- Feurdean, A., Marinova, E., Nielsen, A.B., Liakka, J., Veres, D., Hutchinson, S.M., Braun, M., Timar-Gabor, A., Astalos, C., Mosburgger, V., Hickler, T., 2015. Origin of the forest steppe and exceptional grassland diversity in Transylvania (central-eastern Europe). *Journal of Biogeography* 42(5), 951–963.
- Feurdean, A., Perşoiu, A., Tanţău, I., Stevens, T., Magyari, E.K., Onac, B.P., Marković, S., Andrić, M., Connor, S., Fărcaş, S., Gałka, M., Gaudeny, T., Hoek, W., Kalaczek, P., Kuneš, P., Lamentowicz, M., Marinova, E., Michczyńska, D.J., Perşoiu, I., Płóciennik, M., Słowiński, M., Stancikaite, M., Sumegi, P., Svensson, A., Tămaş, T., Timar, A., Tonkov, S., Toth, M., Veski, S., Willis, K.J., Zernitskaya, V., 2014. Climate variability and associated vegetation response throughout Central and Eastern Europe (CEE) between 60 and 8 ka. *Quaternary Science Reviews* 106, 206–224.
- Feurdean, A., Tanţău, I., Fărcaş, S., 2011. Holocene variability in the range distribution and abundance of *Pinus*, *Picea abies*, and *Quercus* in Romania; implications for their current status. *Quaternary Science Reviews* 30, 3060–3075.
- Feurdean, A., Wohlfarth, B., Björkman, L., Tantau, I., Bennike, O., Willis, K.J., Farcas, S., Robertsson, A.M., 2007. The influence of refugial population on Lateglacial and early Holocene vegetational changes in Romania. *Review of Palaeobotany and Palynology* 145, 305–320.
- Francis, R.M., 2016. Pophelper: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*. DOI:10.1111/1755-0998.12509.
- Giertych, M., Mátyás, C., 1991. Genetics of Scots pine. Amsterdam, Elsevier 3, 19–76.
- Gormley, A.M., Forsyth, D.M, Griffioen, P., Lindeman, M., Ramsey, D.S.L., Scroggie, M.P., Woodford, L., 2011. Using presence-only and presence–absence data to estimate the current and potential distributions of established invasive species. *Journal of Applied Ecology* 48, 25–34.
- Graham, C.H., Ferrier, S., Huettman, F., Mortiz, C., Peterson, A.T., 2004. New developments in museum based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution* 19(9), 497–503.
- Grotkopp, E., Rejmánek, M., Rost, T.L., 2002. Toward a causal explanation of plant invasiveness: seedling growth and life-history strategies of 29 pine (*Pinus*) species. *American Naturalist* 159, 396–419.

- Helmens, K.F., 2014. The Last Interglacial–Glacial cycle (MIS 5–2) re-examined based on long proxy records from central and northern Europe. *Quaternary Science Reviews* 86, 115–143.
- 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(15), 1965–1978.
- Hijmans, R.J., Phillips, S., Leathwick, J., Elith, J., 2016. *dismo*: Species Distribution Modeling. R package version 1.0-15. URL: CRAN.R-project.org/package=dismo.
- Houston D.T., de Rigo, D., Caudullo, G., 2016. *Pinus sylvestris* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A. (Eds.), *European Atlas of Forest Tree Species*. Publ. Off. EU, Luxembourg, pp. e016b94+
- Hubisz, M.J., Falush, D., Stephens, M., Pritchard, J.K., 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9(5), 1322–1332.
- Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23(14), 1801–1806.
- Kabaš, E., Batanjski, V., Glasnović, P., Vicić, D., Tanasković, A., Kuzmanović, N., Lakušić, D., Šinžar-Sekulić, J., 2014. Towards detecting bioclimatic niche–species distribution modelling in four maple species (*Acer* spp.). *Acta Botanica Croatica* 73(2), 401–417.
- Kiehl, J.T., Gent, P.R., 2004. The Community Climate System Model, Version 2. *Journal of Climate* 17(19), 3666–3682.
- Komar, M., Łanczont, M., Madeyska, T., 2009. Spatial vegetation patterns based on palynological records in the loess area between the Dnieper and Odra Rivers during the last interglacial–glacial cycle. *Quaternary International* 198(1), 152–172.
- Lascoux, M., Palmé, A.E., Cheddadi, R., Latta, R.G., 2004. Impact of Ice Ages on the genetic structure of trees and shrubs. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 359(1442), 197–207.
- Lazarova, M., Marinova, E., Tonkov, S., Snowball, I., (2015). A 5000-year pollen and plant macrofossil record from the Osogovo Mountain, Southwestern Bulgaria: Vegetation history and human impact. *Review of Palaeobotany and Palynology* 223, 1–9.
- Lobo, J.M., Jiménez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* 17(2), 145–151.
- Lorenzen, E.D., Nogués-Bravo, D., Orlando, L., Weinstock, J., Binladen, J., Marske, K.A., Ugan, A., Borregaard, M.K., Gilbert, M.T.P., Nielsen, R., Ho, S.Y.W., Goegel, T., Graf, K.E., Byers, D., Stenderup, J.T., Rasmussen, M., Campos, P.F., Leonard, J.A., Koepfli, K.-P., Froese, D., Zazula, G., Stafford, T.W.Jr., Aaris-Sørensen, K., Batra, P., Haywood, A.M., Singarayer, J.S., Valdes, P.J., Boeskorov, G., Burns, J.A., Davydov, S.P., Haile, J., Jenkins, D.L., Kosintsev, P., Kuznetsova, T., Lai, X., Martin, L.D., McDonald, H.G., Mol, D., Meldgaard, M., Munch, K., Stephan, E., Sablin, M., Sommer, R.S., Sipko, T., Scott, E., Suchard, M.A., Tikhonov, A., Willerslev, R., Wayne, R.K., Cooper, A., Hofreiter, M., Sher, A., Shapiro, B., Rahbek, C., Willerslev, E., 2011. Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* 479(7373), 359.
- Magri, D., Vendramin, G. G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., Latalowa, M., Litt, T., Paule, L., Roure, J.M., Tantau, I., van der Knaap, W.O., Petit, R.J., de Beaulieu, J.-L., 2006. A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, 171(1), 199–221.
- Magyari, E.K., 2011. Late Quaternary vegetation history in the Hortobágy steppe and Middle Tisza floodplain, NE Hungary. *Studia Bot Hung* 42, 185–203.
- Magyari, E.K., Chapman, J.C., Passmore, D.G., Allen, J.R.M., Huntley, J.P., Huntley, B., 2010. Holocene persistence of wooded steppe in the Great Hungarian Plain. *Journal of Biogeography* 37(5), 915–935.

- Magyari, E.K., Jakab, G., Bálint, M., Kern, Z., Buczkó, K., Braun, M., 2012. Rapid vegetation response to Lateglacial and early Holocene climatic fluctuation in the South Carpathian Mountains (Romania). *Quaternary Science Reviews* 35, 116–130.
- Magyari, E.K., Kuneš, P., Jakab, G., Sümegi, P., Pelánková, B., Schäbitz, F., Braun, M., Chytrý, M., 2014b. Late Pleniglacial vegetation in eastern-central Europe: are there modern analogues in Siberia? *Quaternary Science Reviews* 95, 60–79.
- Magyari, E.K., Veres, D., Wennrich, V., Wagner, B., Braun, M., Jakab, G., Karátson, D., Pál, Z., Ferenczy, Gy., St-Onge, G., Rethemeyer, J., Francois, J.-P., von Reumont, F., Schäbitz, F., 2014a. Vegetation and environmental responses to climate forcing during the Last Glacial Maximum and deglaciation in the East Carpathians: attenuated response to maximum cooling and increased biomass burning. *Quaternary Science Reviews* 106, 278–298.
- Majecka, A., 2014. The palynological record of the Eemian interglacial and Early Vistulian glaciation in deposits of the Żabieniec Południowy fossil basin (Łódź Plateau, central Poland), and its palaeogeographic significance. *Acta Palaeobotanica* 54(2), 279–304.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R. K., Thuiller, W., 2009. Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions* 15(1), 59–69.
- Matías, L., Jump, A.S., 2012. Interactions between growth, demography and biotic interactions in determining species range limits in a warming world: the case of *Pinus sylvestris*. *Forest Ecology and Management* 282, 10–22.
- Mayol, M., Riba, M., González-Martínez, S.C., Bagnoli, F., de Beaulieu, J.-L., Berganzo, E., Burgarella, C., Dubreuil, M., Krajmerová, D., Paule, L., Romšáková, I., Vettori, C., Vincenot, L., Vendramin, G.G., 2015. Adapting through glacial cycles: insights from a long-lived tree (*Taxus baccata*). *New Phytologist* 208(3), 973–986.
- Milner, A.M., Müller, U.C., Roucoux, K.H., Collier, R.E.L., Pross, J., Kalaitzidis, S., Christanis, K., Tzedakis, P.C., 2013. Environmental variability during the Last Interglacial: a new high-resolution pollen record from Tenaghi Philippon, Greece. *Journal of Quaternary Science* 28, 113–117.
- Nix, H.A., 1986. A biogeographic analysis of Australian elapid snakes. In: Longmore R (ed.): Atlas of elapid snakes of Australia. Australian Flora and Fauna Series 7. Australian Government Publishing Service, Canberra. p.p. 4–15.
- Novenko, E.Y., Seifert-Eulen, M., Boettger, T., Junge, F.W., 2008. Eemian and Early Weichselian vegetation and climate history in Central Europe: A case study from the Klinge section (Lusatia, eastern Germany). *Review of Palaeobotany and Palynology* 151(1), 72–78.
- O'Brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity* 41(5), 673–690.
- Otto-Bliesner, B., Marshall, S.J., Overpeck, J.T., Miller, G.H., Hu, A., C.A.P.E. Last Interglacial Project members, 2006. Simulating arctic climate warmth and icefield retreat in the last interglaciation. *Science* 311(5768): 1751–1753.
- Peakall, R.O.D., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6(1), 288–295.
- Petit, R.J., Hu, F.S., Dick, C.W., 2008. Forests of the past: a window to future changes. *Science* 320(5882), 1450–1452.
- Petr, L., Žáčková, P., Grygar, T. M., Píšková, A., Křížek, M., Treml, V., 2013. Šúr, a former late-glacial and Holocene lake at the westernmost margin of the Carpathians. *Preslia* 85, 239–263.
- Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190(3–4), 231–259.
- Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with MaxEnt: new extensions and a comprehensive evaluation. *Ecography* 31, 161–175.

- Pócs, T., 1960. Die zonalen Waldgesellschaften Südwestungarns. (The zonal forest communities of southwestern Hungary) *Acta Botanica Hungarica* 6, 75–105.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Provan, J., Soranzo, N., Wilson, N.J., Goldstein, D.B., Powell, W., 1999. A low mutation rate for chloroplast microsatellites. *Genetics* 153, 943–947.
- Pyhäjärvi, T., Garcia-Gil, M.R., Knürr, T., Mikkonen, M., Wachowiak, W., Savolainen, O., 2007. Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. *Genetics* 177, 1713–1724.
- R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: www.R-project.org.
- Ronai, A., 1970. Lower and middle Pleistocene flora in the Carpathian basin. *Palaeogeography, Palaeoclimatology, Palaeoecology* 8(2-3), 265–285.
- Rudner, E., Sümegei, P., 2001. Recurring taiga forest-steppe habitats in the Carpathian Basin in the Upper Weichselian. *Quaternary International* 76, 177–189.
- Rudner, E., Sümegei, P., Tóth, I., Beszedá, I., Hertelendi, E., 1995. The vegetation of the upper-Weichselian in the central and southern part of the Great Hungarian Plain. Book of abstracts, 7th European Ecological Congress, pp 108.
- Sadori, L., Koutsodendris, A., Masi, A., Bertini, A., Combourieu-Nebout, N., Francke, A., Kouli, K., Joannin, S., Mercuri, A.M., Panagiotopoulos, K., Peyron, O., Torri, P., Wagner, B., Zanchetta, G., Donders, T.H., 2015. Pollen-based paleoenvironmental and paleoclimatic change at Lake Ohrid (SE Europe) during the past 500 ka. *Biogeosciences Discussions* 12, 15461–15493.
- Sebastiani, F., Pinzauti, F., Kujala, S.T., González-Martínez, S.C., Vendramin, G.G., 2012. Novel polymorphic nuclear microsatellite markers for *Pinus sylvestris* L. *Conservation Genetics Resources*, 4(2), 231–234.
- Slater, G.J., Harmon, L.J., Alfaro, M.E., 2012. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution* 66(12), 3931–3944.
- Soberón, J.M., Llorente, J.B., Onate, L., 2000. The use of specimen-label databases for conservation purposes: an example using Mexican Papilionid and Pierid butterflies. *Biodiversity & Conservation* 9(10), 1441–1466.
- Somodi, I., Molnár, Zs., Czúcz, B., Bede-Fazekas, Á., Bölöni, J., Pásztor, L., Laborczi A., Zimmermann, N.E., (in press): Implementation and application of Multiple Potential Natural Vegetation models – a case study of Hungary. *Journal of Vegetation Science*.
- Soranzo, N., Provan, J., Powell, W., 1998. Characterization of microsatellite loci in *Pinus sylvestris* L. *Molecular Ecology* 7(9), 1260–1261.
- Stefanova, I., Atanassova, J., Delcheva, M., Wright, H.E., 2006. Chronological framework for the Lateglacial pollen and macrofossil sequence in the Pirin Mountains, Bulgaria: Lake Besbog and Lake Kremensko-5. *The Holocene* 16(6), 877–892.
- Sümegei, P., Kertész, R., Hertelendi, E., 2002. Environmental Change and Human Adaptation in the Carpathian Basin at the lateglacial/postglacial transition. *British Archaeological Report* 1043, 171.
- Svenning, J.C., Normand, S., Kageyama, M., 2008. Glacial refugia of temperate trees in Europe: insights from species distribution modelling. *Journal of Ecology* 96(6), 1117–1127.
- Taberlet, P., Cheddadi, R., 2002. Quaternary refugia and persistence of biodiversity. *Science* 297(5589), 2009–2010.

- Tantau, I., Feurdean, A., de Beaulieu, J. L., Reille, M., Farcas, S., 2014. Vegetation sensitivity to climate changes and human impact in the Harghita Mountains (Eastern Romanian Carpathians) over the past 15 000 years. *Journal of Quaternary Science* 29, 141–152.
- Tantau, I., Reille, M., de Beaulieu, J.L., Farcas, S., 2006. Late Glacial and Holocene vegetation history in the southern part of Transylvania (Romania): pollen analysis of two sequences from Avrig. *Journal of Quaternary Science* 21, 49–62.
- Tantau, I., Reille, M., de Beaulieu, J.L., Farcas, S., Goslar, T., Paterne, M., 2003. Vegetation history in the Eastern Romanian Carpathians: pollen analysis of two sequences from the Mohoş crater. *Vegetation History and Archaeobotany* 12, 113–125.
- Thuiller, W., 2007. Climate change and the ecologist. *Nature* 448(2), 550–552.
- Tonkov, S., 2003. Holocene palaeovegetation of the Northwestern Pirin Mountains (Bulgaria) as reconstructed from pollen analysis. *Review of Palaeobotany and Palynology* 124(1), 51–61.
- Tonkov, S., Bozilova, E., Possnert, G., 2013. Postglacial vegetation history as recorded from the subalpine Lake Ribno (NW Rila Mts), Bulgaria. *Open Life Sciences* 8(1), 64–77.
- Tonkov, S., Bozilova, E., Possnert, G., Velčev, A., 2008. A contribution to the postglacial vegetation history of the Rila Mountains, Bulgaria: the pollen record of Lake Trilistnika. *Quaternary International* 190(1), 58–70.
- Tonkov, S., Panovska, H., Possnert, G., Bozilova, E., 2002. The Holocene vegetation history of Northern Pirin Mountain, southwestern Bulgaria: pollen analysis and radiocarbon dating of a core from Lake Ribno Ban derishko. *The Holocene* 12(2), 201–210.
- Tonkov, S., Possnert, G., Bozilova, E., 2011. The Lateglacial in the Rila Mountains (Bulgaria) revisited: The pollen record of Lake Ribno (2184m). *Review of Palaeobotany and Palynology* 166(1), 1–11.
- Tóth, E.Gy., Vendramin, G.G., Bagnoli, F., Cseke, K., Höhn, M., 2017. High genetic diversity and distinct origin of recently fragmented Scots pine (*Pinus sylvestris* L.) populations along the Carpathians and the Pannonian Basin. *Tree Genetics & Genomes* 13(2), 47.
- Tzedakis, P.C., Hooghiemstra, H., Pälike, H., 2006. The last 1.35 million years at Tenaghi Philippon: revised chronostratigraphy and long-term vegetation trends. *Quaternary Science Reviews* 25, 3416–3430.
- van Zonneveld, M., Jarvis, A., Dvorak, W., Lema, G., Leibing, C., 2009. Climate change impact predictions on *Pinus patula* and *Pinus tecunumanii* populations in Mexico and Central America. *Forest Ecology and Management* 257(7), 1566–1576.
- Willis, K.J., 1994. The vegetational history of the Balkans. *Quaternary Science Reviews* 13(8), 769–788.
- Willis, K.J., Bennett, K.D., Birks, H.J.B., 1998. The late Quaternary dynamics of pines in Europe. (In: Richardson D.M. (Ed), *Ecology and Biogeography of Pinus*). Cambridge, Cambridge University Press. pp. 107–121.
- Willis, K.J., Braun, M., Sümege, P., Tóth, A., 1997. Does soil change cause vegetation change or vice versa? A temporal perspective from Hungary. *Ecology* 78(3), 740–750.
- Willis, K.J., Sümege, P., Braun, M., Tóth, A., 1995. The late Quaternary environmental history of Bátorliget, NE Hungary. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 118(1-2), 25–47.
- Wolfe, B.B., Edwards, T.W., Aravena, R., 1999. Changes in carbon and nitrogen cycling during tree-line retreat recorded in the isotopic content of lacustrine organic matter, western Taimyr Peninsula, Russia. *The Holocene*, 9(2), 215–222.
- Worth, J.R., Sakaguchi, S., Tanaka, N., Yamasaki, M., Isagi, Y., 2013. Northern richness and southern poverty: contrasting genetic footprints of glacial refugia in the relictual tree *Sciadopitys verticillata* (Coniferales: Sciadopityaceae). *Biological Journal of the Linnean Society* 108(2), 263–277.

Yuan, H.-S., Wei, Y.-L., Wang, X.-G., (2015): Maxent modeling for predicting the potential distribution of Sanghuang, an important group of medicinal fungi in China. *Fungal Ecology* 17, 140–145.

Zagwijn, W.H., 1996. An analysis of Eemian climate in western and central Europe. *Quaternary Science Reviews* 15(5-6), 451–469.

Zinck, J.W., Rajora, O.P., 2016. Post-glacial phylogeography and evolution of a wide-ranging highly-exploited keystone forest tree, eastern white pine (*Pinus strobus*) in North America: single refugium, multiple routes. *BMC Evolutionary Biology* 16(1), 56.

Zuckerberg, B., Huettmann, F., Frair, J., 2011. Proper Data Management as a Scientific Foundation for Reliable Species Distribution Modeling. In: Drew, C.A., Wiersma, Y.F., Huettmann, F., (eds.) 2011. *Predictive Species and Habitat Modeling in Landscape Ecology. Concepts and Applications*. New York, NY, USA: Springer. pp. 45–70.

Tables and figures

Table 1: List of the 16 island-like populations of *Pinus sylvestris* used in this study from the Central and Eastern European peripheral distribution of the species, and their spatial description (lat. = latitude, long. = longitude) Country and region codes are as follow: RL: Rila Mountains, PB: Pannonian Basin, WC: Western Carpathians, EC: Eastern Carpathians, SC: Southern Carpathians and CIM: Central-Island Mountains (Apuseni).

Code	Country	Region	Site	Lat. (°N)	Long. (°E)	Altitude (m a.s.l.)	Approx. area (km ²)	Sample size
SKV	SK	WC	Kvačianska valley	49.182	19.539	799	0.48	20
SME	SK	WC	Medzi Bormi	49.268	19.630	815	0.07	20
STU	SK	WC	Turková	49.017	19.916	1107	0.70	20
SLI	SK	WC	Liptovský Hrádok	49.044	19.734	729	0.02	10
RFE	RO	EC	Fântâna Brazilor	46.507	25.255	953	0.32	26
RPO	RO	EC	Poiana Stampei	47.300	25.119	878	1.43	20
RMO	RO	EC	Mohos	46.135	25.903	1052	0.58	29
RCO	RO	EC	Suhardul Mic	46.799	25.796	981	0.04	19
RPA	RO	SC	Lătoriței	45.387	23.906	753	3.42	23
RBI	RO	CIM	Roșia	46.841	22.371	393	0.13	8
HFE	HU	PB	Fenyőfő	47.353	17.768	252	4.49	24
HZA	HU	PB	Pethőhenye	46.869	16.922	306	0.04	19
HVE	HU	PB	Szalafő	46.868	16.304	231	0.08	20
HOR	HU	PB	Csörötnek	46.930	16.350	296	0.10	23
HKO	HU	PB	Kőszeg	47.340	16.451	468	0.24	20
BYU	BG	RL	Yundola	42.069	23.831	1561	>5.00	30

Table 2: Analysis of molecular variance (AMOVA) for nSSR dataset of *Pinus sylvestris* at three spatial scales.

Spatial scale	Source of variation	D.f.	Sum of square	Mean square	Estimated variance	Variance (%)	Fixation index (F_{ST})
Pop1 vs. Pop2	among pop.	1	54.858	54.858	0.230	8%	0.084
	within pop.	602	1515.977	2.518	2.518	92%	
	total	603	1570.835		2.748	100%	
Pop2 vs. Pop3	among pop.	1	8.739	8.739	0.076	3%	0.031
	within pop.	208	489.514	2.353	2.353	97%	
	total	209	498.253		2.429	100%	
Pop1 vs. Pop3	among pop.	1	18.806	18.806	0.158	6%	0.058
	within pop.	508	1298.556	2.556	2.556	94%	
	total	509	1317.362		2.714	100%	

Table 3: Percentage of contribution of the nine bioclimatic predictors to the model of Pop1 and Pop2 in the distribution analysis of *Pinus sylvestris*.

Bioclimatic variable	Contribution to model of Pop1 (%)	Contribution to model of Pop2 (%)
bio1	0	53.1
bio3	24.8	24.5
bio7	0	1.6
bio9	21.3	0
bio15	20.1	0
bio16	0	0
bio17	0.6	16.8
bio18	27.8	0
bio19	5.4	4

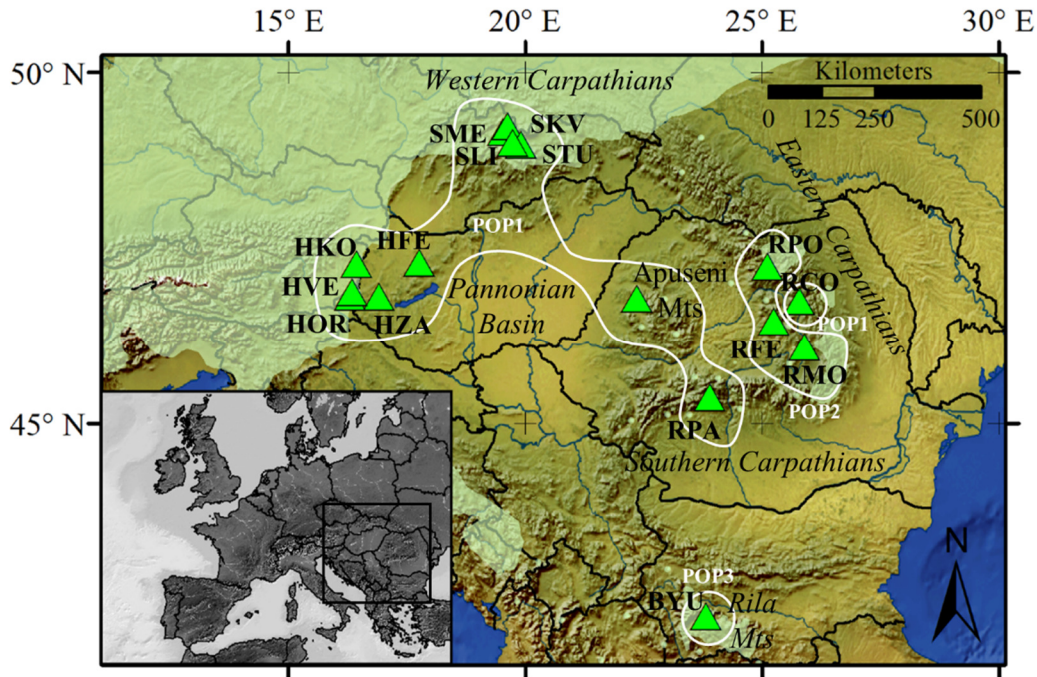
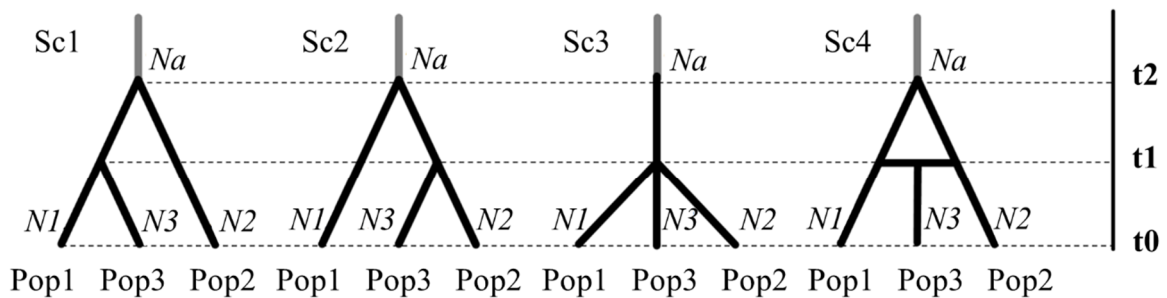


Fig. 1: Location of the studied populations of *Pinus sylvestris* from the Central and Eastern European Carpathians and Pannonian Basin. The acronyms stand for the population code in Table 1. The natural distribution of Scots pine is marked in green according to the EUFORGEN (2009) database.

A Historical divergence estimation



B Demographic estimation

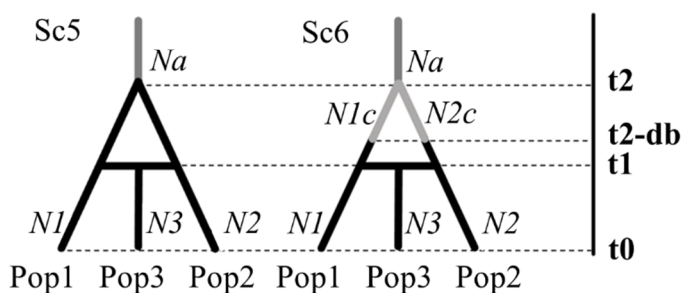


Fig. 2: The six demographic scenarios (A and B) tested on the central-eastern European Scots pine populations in DIY ABC. In the tested scenarios $t\#$ represent the estimated time expressed in generation time, $N\#$ is the effective size of population Pop1, Pop2 and Pop3 respectively. $N2c$ and $N1c$ indicate expansion of ancient population expressed in effective population size, as like $Na < N\#c < N\#$.

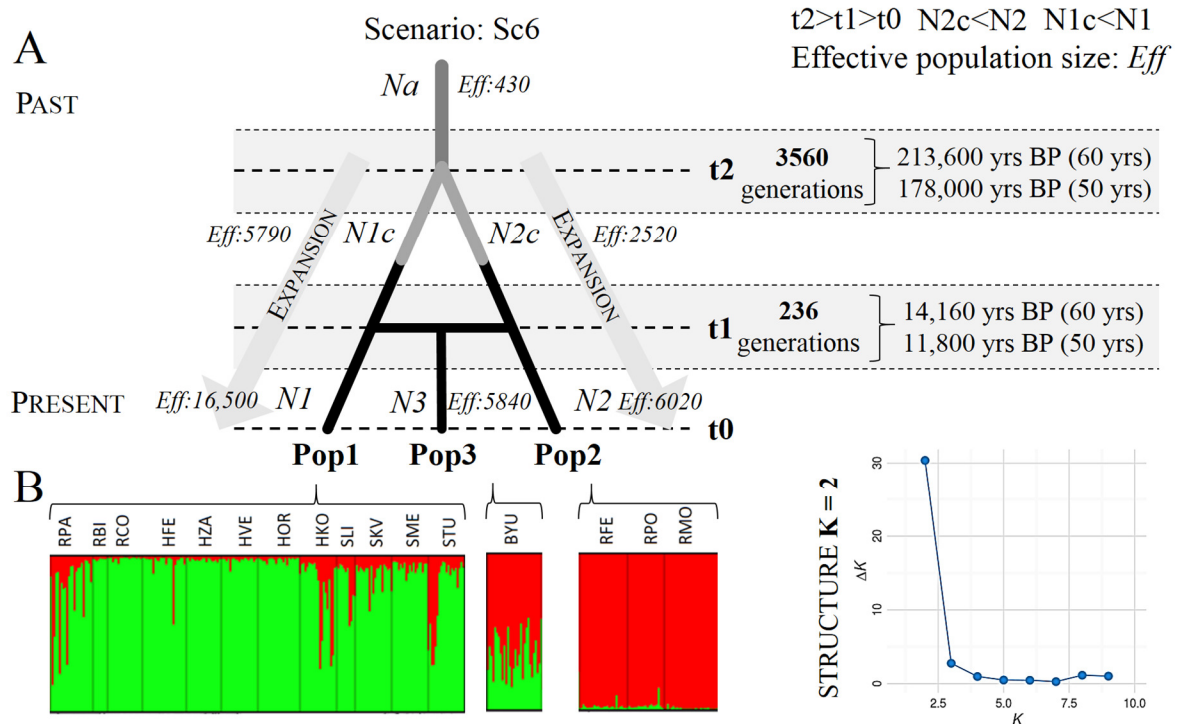


Fig. 3: The best supported demographic scenario Sc6 (A) with the highest posterior probability (0.5670, 95%CI: 0.5307–0.6033) resulted on the central-eastern European Scots pine populations in DIY ABC. Result of STRUCTURE analysis ($K = 2$) prior to ABC analysis (B). In the detected scenario $t\#$ represent the estimated time expressed in generation time, $N\#$ is the effective size of population Pop1, Pop2 and Pop3 respectively. N_{2c} and N_{1c} indicate expansion of ancient population expressed in effective population size, as like $Na < N_{\#c} < N\#$.

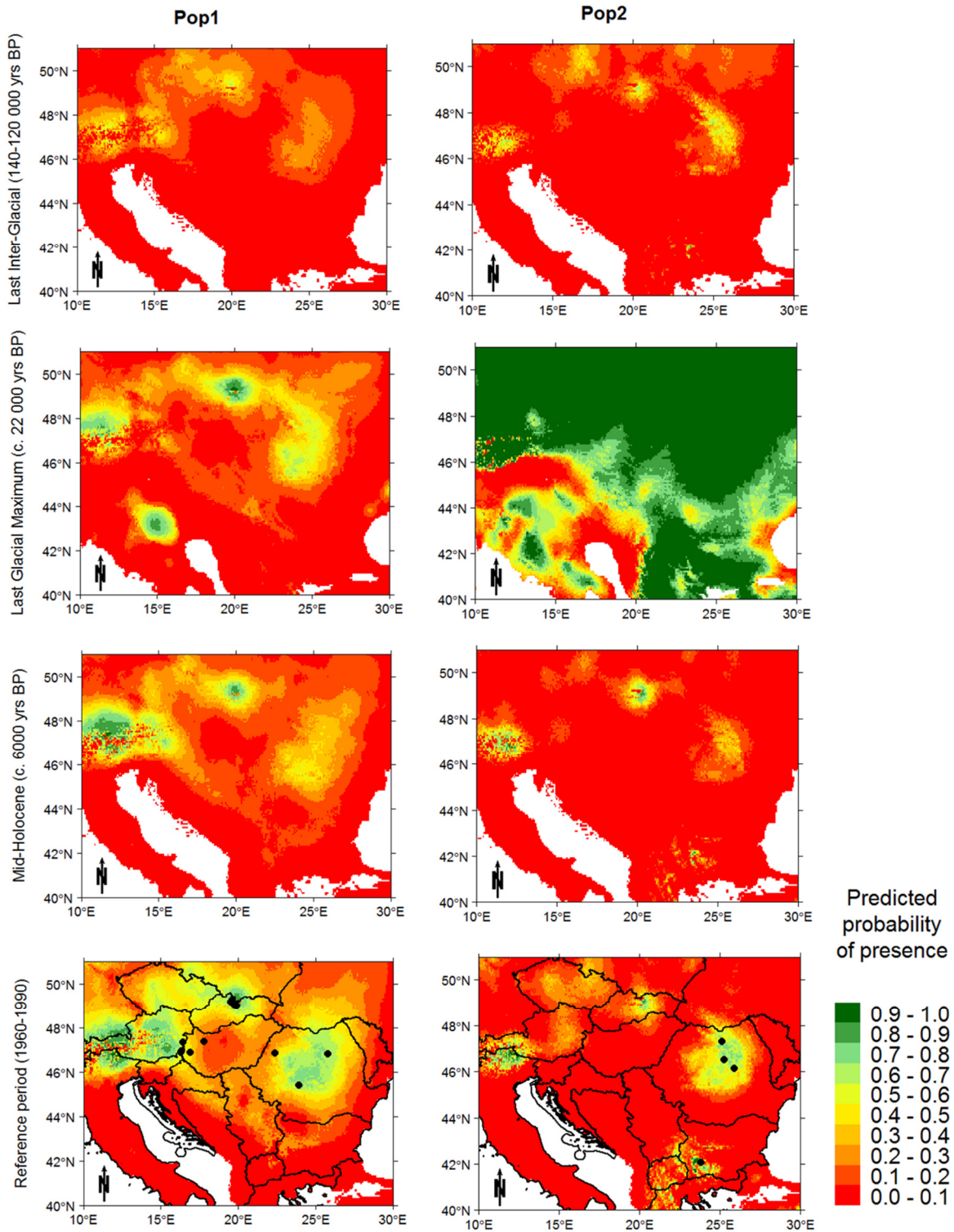


Fig. 4: Predicted distribution of *Pinus sylvestris* based on bioclimatic variables at four time periods: LIG, Last Inter-Glacial (c. 140,000–120,000 yrs before present, BP); LGM, Last Glacial Maximum (c. 22,000 yrs BP); Mid-Holocene (c. 6000 yrs BP); Reference period (1960–1990). The models were produced separately for the Pop1 and Pop2 gene pools. Darker green colors indicate higher probabilities of potential presence, while red color indicate lower predicted probability.