The latitudinal trend in genetic diversity and distinctiveness of *Quercus robur* rear edge forest remnants calls for new conservation priorities

Camilla Avanzi [1], Francesca Bagnoli [1], Edoardo Romiti [1,2], Ilaria Spanu [1], Yoshiaki Tsuda [3], Elia Vajana [1], Giovanni Giuseppe Vendramin [1], Andrea Piotti [1]

[1] Institute of Biosciences and Bioresources, National Research Council of Italy, Via Madonna del Piano 10, 50019 Sesto Fiorentino, Italy

[2] Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Viale delle Idee 30, I-50019 Sesto Fiorentino, Italy

[3] Sugadaira Montane Research Center, University of Tsukuba, 1278-294 Sugadaira-Kogen, Ueda, Nagano 386-2204, Japan

Abstract

Knowledge of the spatial distribution of intraspecific genetic variation is essential for conserving biodiversity, designing rational networks of protected areas, and informing translocation strategies. Although the Italian peninsula likely harbours unique genetic variation as a legacy of Quaternary migrations towards southern Europe, few genetic data are available for many Italian forest tree species and populations.

Here, we present the first, comprehensive characterisation of the nuclear genetic variation of Italian populations of pedunculate oak (*Quercus robur* L.), an iconic broadleaved species of lowland forests. A total of 745 individuals from 25 populations were sampled across the local range of the species and genotyped with 16 microsatellite markers. The genetic layout of *Q. robur* populations was assessed through various metrics of diversity and distinctiveness, as well as by Bayesian clustering and multivariate methods. Based on the inferred genetic structure, the demographic history of *Q. robur* gene pools was evaluated through Approximate Bayesian Computation (ABC) analysis.

We found a clear spatial trend in the estimates of genetic diversity and distinctiveness of Italian *Q. robur* populations. The genetic distinctiveness and private allelic richness values showed a linear decline with increasing latitude, while allelic richness reached its peak in central Italy. We also observed a south-to-north trend in the complexity of genetic structure, with peninsular Italy being characterised by a mosaic of gene pools, in contrast to the homogeneity exhibited by northern Italian populations. Demographic inference indicated that the southern gene pool has had separate dynamics for the last 135,000 years and, therefore, populations in central Italy did not originate from a northward migration route after the Last Glacial Maximum (LGM). Therefore, they were not only a 'melting pot' of multiple recolonization routes, but most likely a mosaic of small refugia where pedunculate oak persisted during the LGM.

Our results shed new light on conservation priorities for highly fragmented *Q. robur* populations, and call for consideration of whether treating the southern and central Italian populations as independent conservation units, or intervening to increase their genetic connectivity, is the best strategy for their long-term conservation.

Keywords

White oaks, conservation genetics, intraspecific genetic diversity, genetic structure, bottleneck, Approximate Bayesian Computation

Introduction

Conserving biodiversity means conserving its three essential components: genetic, specific and ecosystem diversity (McNeely et al. 1990). Intraspecific genetic diversity, i.e. the DNA-encoded heritable variation within and between populations of a species, is not only valuable in and of itself, as it underlies the vast variety of existing life forms, but it is also involved in maintaining the other two components. It indeed provides species with the ability to adapt and evolve (Hoban et al. 2020) and it influences ecological processes, with significant effects on community structure, primary production, nutrient cycling, and ecosystem resilience (Hughes et al. 2008). In the context of the current global crisis, with rapidly changing climatic and environmental conditions and the spread of new pests and diseases, adequate levels of genetic diversity are the essential prerequisite for species to keep pace with changing environments (Sgrò et al. 2011) and for ecosystems to continue to provide their ecological functions and services (Des Roches et al. 2021). However, genetic diversity is declining globally (IPBES 2019), with an estimated six per cent loss of genetic variation in wild species since the industrial revolution (Leigh et al. 2019). Due to its paramount importance for evolutionary, ecological and socio-economic processes, the conservation of genetic diversity has been the target of numerous global initiatives over the last 30 years (see Hoban et al. (2023) for a critical commentary on their development). The first draft of the post-2020 global biodiversity framework (CBD 2021) has for the first time set an ambitious (sensu Díaz et al. 2020) quantitative goal, that is preserving at least 90% of the genetic diversity within all species (relative to a 2020-baseline). Although 90% may sound like an ambitious threshold, this appears to be an appropriate target only for captive populations (Frankham et al. 2002). Simulations have shown that wild populations of both vertebrate and plant species will suffer a significant increase in inbreeding depression in 100 years even if 97% of genetic diversity is conserved between 2022 to 2050 (Frankham 2022). Thus, conserving at least 99% of intraspecific genetic diversity and maintaining all genetically distinct populations would likely be a more robust conservation goal (Hoban et al. 2023).

Despite broad scientific *consensus* on the urgency of conserving genetic diversity, and hence adaptive potential (DeWoody et al. 2021), conservation strategies provide few and vague guidance on how to translate this goal into practical action, particularly at national and subnational levels (Pierson et al. 2016; Cook and Sgrò 2017; Laikre et al. 2020). Hoban et al. (2020) proposed three pragmatic indicators, based on the number of populations with an effective population size (*Ne*) above the charismatic (but controversial) threshold of 500, the proportion of populations conserved for each species and, finally, the number of populations in which genetic diversity is assessed (and tracked) using molecular markers. Indeed, direct knowledge of the spatial distribution of intraspecific genetic variation and the degree of genetic connectivity among populations is indeed particularly useful for informing translocation strategies and/or designing optimal networks of protected areas (Aitken and Whitlock 2013; Fady et al. 2016; Andrello et al. 2022). In Europe, range-wide surveys of the genetic variation in forest tree species have long been conducted (e.g. Lagercrantz

and Ryman 1990; Vendramin et al. 2000, 2008; Petit et al. 2002b, 2003; Bucci et al. 2007; Tollefsrud et al. 2008; Liepelt et al. 2009; Wagner et al. 2015; Mayol et al. 2015; Pina-Martins et al. 2019; Postolache et al. 2021; Milesi et al. 2023). However, southern Europe has often been overlooked in range-wide genetic or genomic datasets, despite its disproportionate importance as a long-term repository of genetic diversity (Hampe and Petit 2005). Southern Europe has likely acted as a mosaic of small, favourable refugia for the persistence of tree species during the cold phases of the Quaternary (Bennett et al. 1991; Tzedakis et al. 1997), resulting in southern European populations often exhibiting complex patterns of genetic variation (Piotti et al. 2017; Scotti-Saintagne et al. 2021). Robust sampling within putative refugial areas is essential to quantitatively describe such complexity in the spatial structure of extant genetic variation (Piotti et al. 2017) and, thus, to plan more effective conservation strategies of forest genetic resources (Vries et al. 2015).

Pedunculate oak (*Quercus robur* L.) is an iconic broadleaved tree species of European lowland ecosystems. It has a wide distribution, ranging from the Iberian, Italian and Balkan peninsulas to Scandinavia and from Scotland to the Ural and Caucasus mountains (Eaton et al. 2016). Although *Q. robur* appears to be characterised by a continuous range according to the available distribution maps (Caudullo et al. 2017), the species suffers from a high degree of fragmentation, as lowland forests have been increasingly replaced by agroforestry and urban landscapes in large parts of the species distribution over the centuries (Petit et al. 2002c; Vakkari et al. 2006; Agnoletti et al. 2018; Degen et al. 2021). In recent decades, fragmentation has been exacerbated by episodes of decline and mortality, which have been fuelled by a complex interplay of changing biotic and abiotic conditions (Gibbs 1997; Ragazzi et al. 2000; Thomas et al. 2002; de Sampaio e Paiva Camilo-Alves et al. 2013; Stojanović et al. 2015; Gathercole et al. 2021). Such decline may be even more pronounced in rear edge populations, whose adaptive potential may be limited by their high degree of geographic isolation and small size. On the Italian peninsula, at one of the rear edges of the species distribution, a comprehensive and accurate study of the genetic layout of *Q. robur* populations is still lacking. To our knowledge, only Fineschi et al. (2002) assessed the chloroplast DNA variation of Italian *Q. robur* populations, as part of a larger investigation of cpDNA patterns of European white oaks (Petit et al. 2002a, b), but only partially covering the local distribution of the species and with very few samples collected from southern Italy, despite solid evidence that the southern Apennines were one of the main glacial refugia for white oak species (Brewer et al. 2002; Petit et al. 2002a). Few other studies have been carried out at the regional scale: Belletti et al. (2005) characterised the isozyme variation of five populations from northwestern Italy, while Ducci et al. (2007) focused only on a few populations from the Po Valley.

Our work is the first attempt to comprehensively characterise the forest genetic resources of pedunculate oak across the Italian peninsula. To this end, 25 populations from northern, central, and southern Italy were sampled and genotyped with 16 microsatellite markers. The specific aims of this study were to: i) assess the levels of genetic diversity and distinctiveness, ii) quantify the contribution of each population to the total allelic diversity and its components within and between populations, iii) identify which populations have an

effective population size <500 and which show signatures of past bottlenecks, iv) decipher the existing genetic structure of the species along the Italian peninsula, and finally v) date the divergence of oak gene pools, if any. As central and southern Italy are characterised by greater topographic heterogeneity than the Po Valley, and peninsular Italy was probably a multi-refugial area (Magri et al. 2015), genetic diversity and distinctiveness, as well as the complexity of genetic structure, were expected to decrease with latitude. Uncovering hotspots of genetic diversity and distinctiveness, as well as the existence of potential gene flow constraints among oak forest remnants, will be essential to inform management and conservation actions aimed at mitigating the negative effects of genetic drift in such an economically and ecologically important tree species.

Materials and methods

Sampling activities and genotyping

Twenty-five populations of *Quercus robur* L. were sampled along the Italian peninsula, sixteen in northern Italy, seven in central Italy and two in southern Italy (Figure 1). In each population, leaves were collected from approximately 30 individuals, spaced at least 20-30 meters to avoid the sampling of related trees. *Q. robur* stands were sometimes so small, especially in central Italy, that it was difficult to reach 30 individuals, but in any case, no fewer than 24 individuals per population were sampled (Table 1), reaching a total of 745 sampled trees. Each individual was tagged and georeferenced using a Garmin GPS tracker. Fresh leaves were dried in silica gel and then stored at -20 °C until DNA extraction. To be sure that we had not mistakenly sampled sessile oak (*Quercus petraea* (Matt.) Liebl.) individuals, we also collected leaves from some *Q. petraea* trees (Appendix 1 – Figure S1a) for successive species discrimination based on molecular markers.

DNA was extracted using the NucleoSpin 96 Plant II Kit (Macherey-Nagel, Germany), following the manufacturer's instructions. Approximately 30 mg of frozen leaves and a 3-mm tungsten bead were added to each well of 96-well plates for grinding leaf material. Then, plates were frozen at -80 °C before two cycles of 30-second disruption at 27 Hz using a Mixer Mill MM300 (Retsch, Germany). All individuals were genotyped with 16 nuclear microsatellite markers (SSRs) (PIE020, PIE223, PIE242, PIE102, PIE239, PIE227, PIE271, PIE267, PIE215: Durand et al. 2010; MsQ13: Dow et al. 1995; QrZAG7, QrZAG112, QrZAG20, QrZAG96: Kampfer et al. 1998; QpZAG15, QpZAG110: Steinkellner et al. 1997). All 16 SSRs were arranged into a 9- and 7-plex by slightly modifying the 12-plex and 8-plex developed by Guichoux et al. (2011). Multiplex reactions were carried out with the Type-it Microsatellite PCR Kit (Qiagen, Germany). The final volume of the PCR reactions was optimised to 6 µl. The PCR mix was composed of 3 µl Type-it Multiple PCR Master Mix, 0.6 µl of primer premix, 1.4 µl of H_2O and 1 µl of DNA. The cycling conditions indicated by Guichoux et al. (2011) were used for the 9-plex, while they were slightly changed for the 7-plex, with an initial step at 95°C for 5 min, followed by 32 cycles at 95°C for 30 s, 56°C for 90 s and 72°C for 45 s, and a final 30 min extension step

at 60°C. PCR products were run on an ABI 3500 automatic sequencer (Applied Biosystems, USA), with LIZ-500 (ThermoFisher Scientific) as internal size standard. The obtained results were sized using GeneMarker (Softgenetics).

Identification of misclassified individuals and quality control of the microsatellite dataset

Quercus robur and *Q. petraea* were present in sympatry in several populations of the present study. Although some "easy-to-use" discriminant morphometric traits were used during fieldwork to identify and sample *Q. robur* trees, molecular markers were subsequently used to further exclude the presence of *Q. petraea* mistakenly classified as *Q. robur*, which could bias estimates of genetic diversity and structure. Our set of SSRs was indeed proven to discriminate well between these two Quercus species (Guichoux et al. 2011). Such a check was done by applying a Principal Coordinate Analysis (PCoA) on the matrix of individual pairwise distances among all oak individuals (N_{tot} = 791, N_{robur} = 745 + $N_{petraea}$ = 46). As the two Quercus species turned to be well separated on the first axis (see Results, Appendix 1 – Figure S1), the highest first principal coordinate (PCo1) score among *Q. petraea* individuals was set as threshold, thus excluding all *Q. robur* trees with a lower PCo1 score from the final dataset.

The presence of null alleles and genotyping failures in the final dataset was assessed at the population level by using INEST 2.2 (Chybicki and Burczyk 2009). The model was run using a Gibbs sampler based on 200,000 iterations, with a burn-in of 20,000 and retaining one iteration every 200.

Genetic structure

The Bayesian clustering approach implemented in STRUCTURE (Pritchard et al. 2000) was used to identify the most likely number of gene pools (*K*). The analysis was based on the LOCPRIOR model (Hubisz et al. 2009), as well as on the correlated allele frequencies model (Falush et al. 2003). Thirty runs were performed for each value of *K*, ranging from 1 to 10. Each run consisted of 10,000 burn-in iterations and 50,000 data collection iterations. These settings for the total length of the MCMC were sufficient to obtain reliable genetic structures by comparing the results of initial runs with those involving longer chains. Different runs for the same *K* were then averaged using the software CLUMPAK (Kopelman et al. 2015). The most likely grouping scenarios were evaluated by using Evanno et al. (2005)'s Δ*K*, as well as by a visual assessment of the log-likelihood trend across the various tested scenarios (mean Ln P(D) and its variance among runs). Both Δ*K* and log-likelihood values were calculated by using the web-based software StructureSelector (Li and Liu 2018). A hierarchical analysis of molecular variance (AMOVA; Excoffier et al. 1992) according to the main genetic structure emerging from STRUCTURE results was performed by using GenAlEx. The statistical significance of the genetic divergence among gene pools, among populations within gene pools and within populations was evaluated by 999 permutations.

The genetic structure of Italian oak populations was also investigated with a multivariate approach, running a PCoA on the matrix of pairwise among-population genetic distance with GenAlEx.

Genetic diversity and distinctiveness

The average number of alleles per locus (*Na*), observed heterozygosity (H_O) and expected heterozygosity (H_E) were calculated for each population using GenAlEx (Peakall and Smouse 2012). The inbreeding rate (F_{IS}) of each population was assessed by using INEST. Allelic richness (*Ar*) and private allelic richness (*PAr*) were calculated by using HP-RARE (Kalinowski 2005). Rarefaction was applied based on a minimum sample size of 42 genes. The contribution of each population to the total allelic diversity (A_T), together with its within-subpopulation (A_S) and among-subpopulation (D_A) components, was assessed using METAPOP2 (López-Cortegano et al. 2019).

The genetic distinctiveness of each population was estimated using a population-specific index of differentiation (F_{ST}) between the target population and a common migrant gene pool from which it differs in varying degrees, using BayeScan v2.1 (Foll and Gaggiotti 2008). BayeScan is implemented within a Bayesian framework which offers the advantage of handling uncertainty on allele frequencies due to small sample sizes without the risk of introducing a bias. F_{ST} coefficients were estimated through 20 pilot runs of 10,000 iterations each, followed by a final MCMC made up of a burn-in of 50,000 and further 50,000 iterations with a thinning of 20.

Each of the above-mentioned estimates of genetic diversity and distinctiveness (i.e. Na, H_O, H_E, *Ar, PAr,* A_T, A_S, D_A, F_{ST}) was regressed against latitude, including a quadratic term to account for non-linear trends in the spatial distribution of genetic diversity and structure. Linear models were run within the R environment (R Core Team 2023).

Demographic parameters

Effective population size (*Ne*) was assessed using the linkage disequilibrium method implemented in NeEstimator v2.1 (Do et al. 2014). In forest tree species, due to their peculiar life-history traits, reliable estimates of *Ne* should be obtained only for isolated and small populations (typically less than 500 reproductive trees) (Santos-del-Blanco et al. 2022). In our dataset, the populations that met such requirements are shown in Table 1.

The occurrence of past bottlenecks within each population was evaluated by searching both for an excess of heterozygosity (*H*; Cornuet and Luikart 1996) and a deficiency in the ratio between the number of alleles and the range of allele size (MR; Garza and Williamson 2001) with INEST. The equilibrium H_{eq} and MR_{eq} values were computed with coalescent simulations, assuming a two-phase mutation model with a proportion of multi-step mutations of 0.22, and then tested against observed *H* and MR values using a one-tailed Wilcoxon signed-rank test with 1,000,000 permutations.

Approximate Bayesian Computation to trace back the divergence among oak gene pools

To infer the past demographic history of Italian *Quercus robur* populations, the Approximate Bayesian Computation (ABC) method implemented in the DIYABC v2.1 software (Cornuet et al. 2008, 2014) was applied. The information obtained at *K*=3 from the Bayesian clustering analysis was used to design demographic scenarios. Three 'demographic populations' were, thus, defined: Pop1 (northern Italy), Pop2 (central Italy), and Pop3 (strictly corresponding to the PAN population) (see Results). Pop1 and Pop2 were assembled by randomly selecting 60 individuals from the populations assigned to the relative gene pool with a membership ≥0.7, while Pop3 consisted of the entire PAN population (63 individuals). Three sets of preliminary ABC simulations were run to determine: A) which populations among Pop1, Pop2 and Pop3 should be traced back to an ancestor, B) which was the hierarchy of the split scenarios, and C) if, and when, asymmetric admixture events occurred. In principle, DIYABC does not consider the hypothesis of post-split gene flow within scenarios, but a recent study has modified DIYABC to include symmetric admixture between two populations (Chapuis et al. 2020), and Kitamura et al. (2022) have extended this approach to model directional admixture (e.g. from Pop1 to Pop2 and from Pop2 to Pop1), which could be considered analogous to gene flow. Thus, the parameter 'r#' relating to the admixture rate between two generic populations *a* and *b* can be considered as the relative amount of gene flow at time 't#', while '1-r#' represents the value of the relative amount of gene flow within the reference parent population. Details on the sets of preliminary ABC simulations, as well as posterior probabilities of the demographic scenarios tested at each step, are reported in Appendix 2.

At the end of the preliminary analyses, the resulting six best scenarios were compared in the final ABC analysis (Figure 2). In all tested scenarios, t1 to t7 represent the time scale measured in generation time, N1, N2 and N3 represent the effective population sizes of Pop1, Pop2 and Pop3, respectively, and ra to rd represent the admixture ratios. The six scenarios are as follows:

- *Scenario 1 – simple split model*: the three populations diverged at the same time (t7), with admixture from Pop2 to Pop1 (ra) and from Pop1 to Pop2 (rb) at times t1 and t2, respectively;
- *Scenario 2 – hierarchical split model I*: Pop1 was merged to Pop3 at t7 and Pop2 was merged to Pop1 at t5, with admixture from Pop2 to Pop1 (ra) and from Pop1 to Pop2 (rb) at times t1 and t2, respectively, and from Pop3 to Pop2 (rc) and from Pop2 to Pop3 (rd) at time t3 and t4, respectively;
- *Scenario 3 – hierarchical split model II*: this scenario differed from *Scenario 2* only in the fact that gene flow was allowed only from Pop2 to Pop1 (ra) at time t1 and from Pop3 to Pop2 (rc) at time t2;
- *Scenario 4 – hierarchical split model III*: Pop2 was merged to Pop3 at t7 and Pop1 was merged to Pop2 at time t5, with admixture from Pop2 to Pop1 (ra) at time t1 and from Pop3 to Pop2 (rc) at time t6;

- Scenario 5 – *hierarchical split model IV*: this scenario differed from *Scenario 4* only in the fact that admixture from Pop3 to Pop2 (rc) occurred at time t2, after the split of Pop1 from Pop2 at time t5;
- Scenario 6 – *hierarchical split model V*: this scenario differed from *Scenario 4* and *Scenario 5* for the reciprocal admixture between Pop1 and Pop2 (ra and rb) at time t1 and t2, and between Pop2 and Pop3 (rc and rd) at time t3 and t4.

The Generalized Stepwise Mutation model (GSM; Estoup et al. 2002) with Single Nucleotide Indels (SNI) was used for all ABC simulations. The default priors were modified to obtain better posterior distributions based on the results of the preliminary simulations (Appendix 2 – Table S4). The minimum and maximum priors for SSR mutation rate were set to 1×10^{-4}–1×10^{-3}. The mean number of alleles, mean genetic diversity, mean allele size variance and mean Garza-Williamson's M were used as summary statistics for individual populations. For each population pair, mean number of alleles, mean genetic diversity, mean allele size variance, F_{ST}, and mean individual assignment likelihood were calculated. One million simulations were run for each scenario. After a total of six million simulations, the most likely scenario was assessed by comparing posterior probabilities, using logistic regression. Goodness-of-fit was estimated for each scenario by model checking using the Principal Component Analysis (PCA) approach implemented in DIYABC, which measures the discrepancy between simulated and observed data.

Results

Identification of misclassified individuals and quality control of the microsatellite dataset

The Principal Coordinate Analysis showed that *Quercus petraea* and *Q. robur* individuals were separated on the first Principal Coordinate (PCo), which explained 5.47% of the total variance (Appendix 1 – Figure S1b). Seventeen trees that had been misclassified as *Q. robur* in the field were excluded from the final dataset as they grouped together with *Q. petraea* trees. The exclusion threshold was set on the highest first PCo score among all *Q. petraea* individuals (i.e. -0.234).

No marker showed consistent and/or significant problems of genotyping failure and null alleles (Appendix 1 – Figure S2). Estimates of genotyping failure rates ranged from 0.0004 to 0.1382, with an average value of 0.0037, while estimates of null allele frequency ranged from 0.0008 to 0.0890, with an average value of 0.0074.

Genetic structure

According to ΔK statistic (Evanno et al. 2005), the Bayesian clustering analysis revealed an optimal grouping at $K = 2$ (Appendix 1 – Figure S3a), which clearly distinguished a gene pool represented by the southernmost population of PAN from another gene pool represented by all the other Italian populations (Figure 3a). AMOVA analysis showed that the percentage of total genetic variation explained by the differences between

these two main gene pools was 7% (P = 0.001). Based on Δ*K*, the next strongest level of structuring was *K* = 3, which separated populations from central Italy from those from northern Italy (Figure 3b). Although Δ*K* was the highest when *K* = 2, the mean Ln P(*K*) increased with *K,* reaching its maximum at *K* = 7 (Appendix 1 – Figure S4b). However, as variance of Ln P(*D*) among runs was larger after *K* = 4, the point at which log-likelihood also started to plateauing, the level of structuring at *K* = 4 was also reported (Figure 3c). According to this scenario, subtle but distinct genetic features of the central Italian populations emerged, with one population (FAB) mainly assigned to a fourth gene pool ('pink' in Figure 3c).

At *K* = 2, the F values – analogue to F_{ST} between the assumed common ancestral population and the gene pools retrieved by STRUCTURE – were 0.218 (PAN) and 0.044 (all remaining populations). At *K* = 3, the F values were 0.210 (PAN), 0.035 (northern Italy) and 0.020 (central Italy). At *K* = 4, the F values were 0.214 (PAN), 0.046 (northern Italy), 0.017 (central Italy) and 0.206 (FAB).

The results of the Principal Coordinate Analysis (PCoA) on the matrix of pairwise genetic distance among populations confirmed the genetic structure observed with the Bayesian clustering approach (Appendix 1 – Figure S4). The first three coordinates explained 60.3% of the total variance of the SSRs data. Oak populations were distributed across the axis of the first Principal Coordinate (PCo) according to a latitudinal gradient. The linear regression of first PCo scores against latitude was indeed significant (ANOVA test: F-statistic = 169.9, *P* <0.001). The southernmost population of PAN was well separated from all the other populations both on the first and second PCo axis. FAB population showed a clear separation from all the other populations on the third PCo axis.

Genetic diversity and distinctiveness

The 16 SSRs showed a total and mean number of alleles of 245 and 15.3, respectively. Estimates of population genetic diversity and distinctiveness are reported in Table 2. The average number of alleles per locus (*Na*) ranged from 6.250 to 8.188, while allelic richness (*Ar*) ranged from 5.801 to 7.877. *Ar* showed a quadratic dependence on latitude (*lat*: F-value = 5.19, *P* = 0.033; lat^2: F-value = 5.04, *P* = 0.035), with populations from central Italy showing the highest *Ar* values (Figure 4a). Observed (H_O) and expected (H_E) heterozygosity ranged, respectively, from 0.640 to 0.735 and from 0.624 to 0.726, indicating an overall small but not significant heterozygote excess (mean F_{IS} = 0.014). Private allelic richness (*PAr*) ranged from 0.005 to 0.231, and it had a linear negative relationship with latitude (F-value = 10.42, *P* = 0.004; Figure 4c). Genetic distinctiveness (population-specific F_{ST}) ranged from 0.008 to 0.128, and it also had a linear negative relationship with latitude (F-value = 37.79, P = 2.865×10^{-6}; Figure 4d). No population showed an inbreeding coefficient (F_{IS}) significantly greater than zero (Appendix 1 – Figure S5).

The population that contributed the most to the total allelic diversity was ORV (A_T = 0.619) while the one that contributed the least was FAB (A_T = -0.504) (Figure 5). When within-subpopulation (A_S) and among-subpopulation (D_A) components of A_T were disentangled, ORV still had the largest A_S value (0.454), while PAN

had the lowest value of A_S (-0.671) but the largest value of D_A (0.585) (Figure 5). D_A exhibited a linear negative relationship with latitude (F-value = 75.24, P = 1.041×10^{-8}), while A_S a quadratic dependence on latitude (*lat*: F-value = 5.18, P = 0.033; lat^2: F-value = 5.04, P = 0.035), with populations from central Italy showing the highest values.

Demographic parameters

Estimates of effective population size (*Ne*) did not overlap to infinite in ten cases out of 25 (Table 2). In these ten populations, mean *Ne* values were always <500 and, in three cases, they were even <50.

No population exhibited a significant heterozygosity excess, while a significant deficiency in the MR value with respect to the values expected at the equilibrium was found in 16 populations out of 25 (Table 3), providing evidence for the presence of bottlenecks alongside most population histories.

Divergence between oak gene pools

In the final DIYABC analysis, *Scenario 3* had the highest posterior probabilities (0.2698, 95% CI: 0.2445-0.2952), although its 95% CI overlapped with that of *Scenario 5* (0.2249, 95% CI: 0.1986-0.2511) (Table 4). Only two out of the 30 summary statistics calculated on the simulated data for both *Scenario 3* and *Scenario 5* were significantly different from the observed data (Appendix 2 – Table S6) which, together with the visual inspection by Principal Component Analysis (Appendix 2 – Figure S4), suggests a high goodness-of-fit of the observed data to the results of these two scenarios. We have chosen to describe only the parameter estimates for *Scenario 3* here, but Appendix 2 also reports parameter estimates for *Scenario 5*, which are indeed highly correlated with each other. Median values of the effective population sizes (N#) for *Scenario 3* were 6,400 (95% CI: 2,140-9,680), 22,900 (95% CI:11,400-29,400), and 4,930 (95% CI: 1,600-9,270) for Pop1, Pop2 and Pop3, respectively (Appendix 2 – Table S5). Posterior parameters showed that N2 was larger than N1 and N3, even considering the CI values. Median values of t1, t2, t5 and t7 were 140 (95% CI: 24.9-490), 737 (95% CI: 179-2,330), 2,360 (95% CI: 625-7,110), and 4,510 (95% CI: 1,190-9,600) generations ago, respectively (Appendix 2 – Table S5). Assuming 30 years as generation time, gene flows from Pop2 to Pop1 (t1) and from Pop3 to Pop2 (t2) were dated to 4,200 (95% CI: 747-14,700) and 22,110 (95% CI: 5,370-69,900) years ago, respectively, while divergence times of Pop1 and Pop2 (t5) and of Pop1 and Pop3 (t7) were scaled to 70,800 (95% CI: 18,750-213,300) and 135,300 (95% CI: 35,700-288,000) years ago, respectively. Median values of ra and rc were 0.862 (95% CI: 0.333-0.992) and 0.753 (95% CI: 0.161-0.986), respectively (Appendix 2 – Table S5). The median values of the mean mutation rate of SSR and SNI were 2.17×10^{-4} (95% CI: 1.14×10^{-4}-5.52×10^{-4}) and 1.52×10^{-7} (95% CI: 1.18×10^{-8}-2.48×10^{-6}), respectively (Appendix 2 – Table S5). The mean value of the mean P was 0.291 (95% CI: 0.217-0.300) (Appendix 2 – Table S5).

Discussion

Quercus robur is a keystone species in lowland ecosystems and provides a wide range of ecosystem services and wood products. Nonetheless, there is limited knowledge on the genetic diversity of its rear edge populations, despite they reasonably harbour unique genetic variation as a legacy of southward retreatment during the Quaternary glaciations. To our knowledge, the only assessment of the genetic diversity of *Q. robur* across the Italian peninsula was carried out by Fineschi et al. (2002), as part of a broader European-scale survey of white oak chloroplast DNA diversity (Dumolin-Lapègue et al. 1997; Petit et al. 2002a,b). Based on the spatial distribution of chloroplast haplotypic diversity, the Authors hypothesised the existence of three main post-glacial recolonization routes traversing the Italian peninsula: two starting from southern Italian and Balkan refugia, respectively, and moving northward, and a third one starting from the northern Balkans and moving westward. According to this hypothesis, central Italy, which was characterised by the highest haplotypic diversity, acted as a 'melting pot' between two migration routes coming from separate refugial areas (i.e. southern Italy and southern Balkans). Although these genetic studies shed the first light on post-glacial history of white oaks in Italy, their conclusions about this area were mainly based on *Quercus pubescens* specimens (~75% of the total analysed samples), with limited coverage of the local *Q. robur* range (64 samples from 15 populations, and only eight samples from southern Italy).

Our study represents the first, comprehensive characterisation of the nuclear genetic diversity of *Q. robur* genetic resources across the Italian peninsula. We exhaustively covered the entire distribution of the species, sampling 745 trees from 25 populations from both large and extremely small remnant forests. Likewise Fineschi et al. (2002), we observed the highest genetic diversity in central Italy and the lowest in southern Italy. Beyond Fineschi et al. (2002), we found that genetic distinctiveness and private allelic richness, as well as the contribution to genetic differentiation, showed a south-to-north linear decrease. We observed a similar trend in the complexity of the genetic structure, with peninsular Italy being characterised by a mosaic of different gene pools, in contrast to the relative genetic homogeneity of populations located in northern Italy. Genetic structure is expected to be more complex within multi-refugial areas (de Lafontaine et al. 2013), and growing evidence is showing that such pattern holds true for several forest tree species in the Italian peninsula (e.g. Piotti et al. 2017; De Dato et al. 2020; Fortini et al. 2023). Although our results must be taken with caution because they are limited to Italy, they showed interesting hints about the species dynamics in the last 300,000 years. First, we found that the southern gene pool, which is almost unique to the southernmost PAN population, was the most ancestral among the ones detected, from which the other gene pools diverged. We can rule out the populations nowadays present in central Italy originated from a migration route starting from southern Italy and moving northward after the Last Glacial Maximum (LGM). In fact, according to ABC median estimates, the southern gene pool has had separate genetic dynamics for the last 135,000 years. Apparently, since the onset of the Eemian interglacial the southern gene pool had no longer received any gene flow from gene pools currently present in northern and central Italy while asymmetrically acting as a source of northward gene flow, particularly at the end of the LGM. Similarly, we

can rule out that central Italy was simply a 'transit area' where multiple recolonization routes got in contact. On the contrary, we showed that central Italy acted as a mosaic of small refugia where pedunculate oak persisted during the LGM while accumulating private genetic diversity. Recent paleobotanical surveys indeed reported significant percentages of deciduous *Quercus* pollen in peninsular Italy at the onset of the Holocene, with especially high values in central Italy (Magri et al. 2015). Although a reconstruction at the species level is hampered by difficulties in distinguishing their pollen grains, palynological data indicate that deciduous *Quercus* spp. had limited migrations throughout the Holocene in Italy, and major demographic events were linked to increasing or decreasing their local density (Magri et al. 2015). Our data, not showing significant clines in genetic diversity throughout possible recolonization routes involving specific gene pools (e.g. latitudinal clines in northern Italy), supported this hypothesis.

At the highest hierarchical level both Bayesian clustering and multivariate analysis revealed the existence of two main gene pools among Italian populations: the first, at the very rear edge of the species distribution, consisting of only one population (Bosco Pantano, PAN, along the Ionian coast), and a second one made up of all remaining populations. This genetic structure suggests that PAN represents the last remnant of larger southern populations that likely survived (at least) the last two glacial periods in southern Italy, isolated from the other detected gene pools. PAN was indeed the most genetically distinct population in our dataset, with the highest contribution to genetic differentiation and the second largest private allelic richness value. Palynological and genetic evidence postulated the existence of a glacial refugium for pedunculate oak in the surroundings of Laghi di Monticchio, 200 km north of PAN (Watts 1996), or further south (Brewer et al. 2002; Fineschi et al. 2002; Petit et al. 2002a; Magri et al. 2015), an area for which, unfortunately, palynological records are scarce (see, for instance, Fig. 1 in Magri et al. 2015). Nowadays, PAN is an extremely small (i.e. 64 remaining trees, all surveyed and genotyped) and isolated remnant of lowland forest, mostly surrounded by cultivated fields and coastal pine forests (Travascia et al. 2022, Pericolo et al. 2023). The population, which has undergone a strong reduction in its census size over the last two centuries due to massive human overexploitation, showed reduced genetic variation, with an average ~20% decrease in allelic richness with respect to the other studied populations. It also exhibited a M-ratio value significantly lower than what expected at the equilibrium, which is usually interpreted as a signature of a severe and prolonged bottleneck (Williamson-Natesan 2005). The F values of the PAN gene pool (0.218, 0.210, 0.214 for K equal to 2, 3 and 4, respectively) also indicated that PAN experienced the effects of intense genetic drift (Falush et al. 2003). Nevertheless, no inbreeding was observed and the effective population size, albeit low, matched the census size of the population (*Ne* = 63, $CI_{95\%}$ = 51-81), probably indicating that the local gene pool was well represented in the reproductive dynamics leading to the inspected generation (Sandurska et al. 2019). To this regard, it is important to stress that very often the effects of habitat fragmentation are not detectable in forest trees because investigations are focused on pre-fragmentation cohorts (Lowe et al. 2015). The strongest reduction in census size due to extensive cutting in the PAN population happened in the 1950s of

the twentieth century (Travascia et al. 2022) and, therefore, deviations in parameters more sensible to recent demographic dynamics could only be detected in the regeneration established in the last decades, which is almost absent in PAN.

At the next strongest levels of structuring, the gene pool comprising all remaining *Q. robur* populations was further subdivided. At K = 3 and K = 4, one and two differentiated gene pools emerged in central Italy, respectively, while northern Italian populations kept belonging to a single homogeneous gene pool. Northern populations showed an overall low genetic distinctiveness (average population-specific F_{ST} = 0.015), clustered tightly on the first three PCo axes, and showed very similar coancestry coefficients for all relevant levels of structuring in the Bayesian clustering analysis. These results were in contrast with those from a regional-scale study which reported one order of magnitude higher differentiation values (F_{ST} = 0.107) for *Q. robur* populations in the Po Valley based on isozyme markers (Ducci et al. 2007). However, the low levels of genetic differentiation we found in this area of ~50,000 km^2 (450 km between MER and FAG, the most distant populations sampled in the Po Valley) were in line with several other studies reporting low to moderate levels of genetic differentiation at the regional scale (e.g. G_{ST} = 0.003 in Burger et al. 2021; F_{ST} = 0.031 in Belletti et al. 2005; F_{ST} = 0.032 in Kesić et al. 2021; F_{ST} = 0.051 in Ballian et al. 2010; F_{ST} = 0.054 in Degen et al. 2021). *Q. robur* is indeed a wind-pollinated tree, with a high potential for long-distance dispersal and extensive gene flow. Effective pollen dispersal distances up to 80 km in isolated populations are reported for this species (Buschbom et al. 2011). Pollen and seed flow rates were equal to 60% and 40%, respectively, across several populations from the entire species' range, even though with large variation among populations (21-88%, and 20-66%) (Gerber et al. 2014). Altogether, these pieces of evidence about *Q. robur* gene flow capabilities and the scarce differentiation found in such a large area outlined a scenario in which *Q. robur* populations in the Po Valley constituted a single, large meta-population, which was genetically connected despite the considerable degree of fragmentation of the species remnants, at least until the reproductive events that generated the inspected cohort. The peculiar topography of the Po Valley, with no significant geographic barriers, coupled with the current presence of typical elements of the agroforestry landscape, such as rows of *Q. robur* trees and isolated individuals, may have favoured pollen exchanges among pedunculate oak shrinking populations, leading to the extremely low differentiation that we observed.

The role of landscape features in shaping the genetic structure of *Q. robur* populations can be discussed also by shifting the focus on populations from central Italy. For instance, the distance between the two populations of ORV and FAB was ~15 km and their pairwise genetic differentiation (G_{ST}) was equal to 0.026. As a comparison, the easternmost and westernmost populations of the Po Valley (i.e. MER and FAG) were around 450 km apart but were much less differentiated (G_{ST} = 0.006). The differentiation between ORV and FAB was intercepted also by the Bayesian structuring at K = 4, where FAB emerged as a highly genetically peculiar population. Although the two populations were geographically close, they were separated by a small ridge that might have acted as an effective barrier to pollen and seed gene flow. As fine-scale topography

may determine contrasting micro-climates across very small distance (Scherrer and Körner 2011), individuals from the two populations could also experiment unsynchronised phenological cycles due to separate post-glacial dynamics, which would further hamper pollen gene flow. Despite the potential for extensive gene flow in this species, a similar case of large differentiation at a very fine scale was observed in Spain, at the southern margin of the species distribution (Moracho et al. 2016). The Authors studied ten *Q. robur* stands located across a 20-km long valley, identifying four distinct gene pools, and finding that most pollination events occurred within each stand. They suggest that the combination of rugged topography, dense canopy and the humid microclimate of riparian forests could have limited long-distance pollen dispersal.

According to the last National Forest Inventory, the so-called 'temperate oaks' (i.e. *Q. robur*, *Q. petraea* and *Q. pubescens*) cover about 1.1 million hectares in Italy (Gasparini et al. 2022), but those specifically occupied by *Q. robur* should be between 65,000 and 80,000 hectares, with the largest presence in northern Italy (Calvo 2007). Distribution maps of the species at the whole-distribution level (e.g. Caudullo et al. 2017) strongly overestimates, at least for Italy, the extent and continuity of such *Q. robur*-dominated formations, whose distribution is in fact largely fragmented (Belletti et al. 2007). Despite their high level of fragmentation, we argue that the genetic status of pedunculate oak populations is not of the utmost concern as previously suggested by some regional studies. Ducci et al. (2007), in their survey of *Q. robur* forest genetic resources in the Po Valley, reported a strong genetic structure and disproportionally high levels of inbreeding (F_{IS} values up to 0.5), suggesting that genetic connections through gene flow had long been interrupted. Our results depicted a rather different scenario, with low genetic differentiation among northern populations and a complete absence of inbreeding (no population showed a F_{IS} significantly larger than 0). This picture points out that *Q. robur* populations, at least in the Po plain where the density of the species is rather continuous although highly heterogeneous, were still genetically connected at least 50 years ago. In fact, it is important to further stress that, as our investigation was focused on the adult cohort and not on the regeneration, some concerns about the genetic status of such genetic resources may arise, as it seems, for instance, from the effective population size (*Ne*) estimates.

Ne defines the size of an ideal population that would result in the same level of inbreeding and genetic drift as that of the population under study (Jamieson and Allendorf 2012) and roughly corresponds to the number of reproductive individuals of a population that contribute to the next generation. Thus, as it quantifies the rate at which a population loses genetic variation, *Ne* is commonly used in conservation genetics to evaluate the state and viability of populations, both in the short and long run. As a rule of thumb, *Ne* should not be <50 and <500 in the short and long term, respectively (Franklin 1980). A *Ne* > 50 would be beneficial for avoiding the harmful effects of inbreeding depression, while a *Ne* > 500 would be necessary for maintaining sufficient adaptive potential for coping with changing environmental conditions. We found that ten populations had a *Ne* <500 (three of them <50). These populations represented the 25%, 57% and 100% of the northern, central and southern Italian populations, respectively. Low *Ne* estimates for *Q. robur*

populations were often reported (Dering and Chybicki 2012; Vranckx et al. 2014; Sandurska et al. 2019). For instance, four Belgian stands of the species exhibited *Ne* estimates ranging from 23 to 58, but, as in our case, no signals of inbreeding (Vranckx et al. 2014). Our findings may raise some concerns, especially for the future of central and southern Italian populations, but they should be interpreted with caution. There is in fact an intense debate not only on how to handle the 50/500 rule of thumb proposed by Franklin (1980) (see for example Traill et al. 2010 and Jamieson and Allendorf 2012 for different philosophical framings of the uses of the upper threshold of 500) but also on the reliability of *Ne* values for forest tree species (see the concerns raised by Fady and Bozzano (2021) and the reply by Hoban et al. (2021)). Santos-del-Blanco (2022) demonstrated that reliable estimates of *Ne* can be obtained only for isolated tree populations with less than 500 reproductive individuals. To complicate the picture even further, in forest trees low *Ne* and census size values were sometimes reported to be unrelated to reduced levels of genetic diversity (Sagnard et al. 2011) and to fatal evolutionary trajectories (Vendramin et al. 2008) or, some other times, correlated with a significant amount of 'genetic load' (Jaramillo-Correa et al. 2020). In this framework, we may consider our results on the *Ne* of *Q. robur* populations as a very early signal of genetic decline that may intensify in future generations. A negative relationship between *Ne* and the level of correlated paternity was observed in *Q. robur* (Vranckx et al. 2014), suggesting that inbreeding might increase in the future as the probability of mating among half- and full-sibs increases too. Monitoring natural regeneration is crucial to investigate how neutral and adaptive genetic variation is transmitted through generations (Lowe et al. 2015; Vajana et al. 2023). The low to moderate *Ne* values of our populations might predict a significant depauperization in the juvenile cohorts. However, juveniles sampled in five of the studied populations (i.e. FGN, GER, GRO, CDC and PAL) did not show any reduction in genetic diversity, although *Ne* values were markedly below 500 in two of them (unpublished data).

Conclusions

Although post-LGM recolonization and northward expansion of tree species from Italian refugia have been widely discussed in the last two decades, this is the first study to evaluate the genetic structure and demographic history of *Quercus robur*, one of the key elements of European floodplain forests, in the Italian peninsula by using biparental inherited DNA variation. The genetic information collected and analysed here clarified that at least three isolated refugia were present in the Italian peninsula and this shed a new light on the priorities to be set for the conservation of *Q. robur* highly fragmented populations. On one hand, since not only the southernmost populations but also stands in central Italy showed high genetic distinctiveness, as well as private and shared allelic richness, they should be treated as independent conservation units. On the other hand, as these populations are often severely isolated, we need to consider whether treating them as independent conservation units or intervening to increase their genetic connectivity (e.g. through assisted migration) is the best strategy for their long-term maintenance. The key to informing this choice is to assess

the adaptive potential and risk of non-adaptedness of *Q. robur* populations by population genomics approaches in future studies.

Acknowledgements

This research was carried out in the framework of the projects ResQ "Oak dieback in lowland forests: a multidisciplinary study for the selection of resistant genetic resources" co-funded by Lombardia Region through Bando 2018 d.d.s. n. 4403, and "L'Ultima Foresta Incantata" 2018- AMB-00138 funded by Fondazione con il Sud. We also acknowledge the contribution of LIFE14 IPE/IT/000018 project "LIFE IP GESTIRE 2020 – Nature Integrated Management to 2020", and the project "Climate Change" (CNR DTA.AD003.474) funded by the Italian Ministry of University and Research (FOE-2019). We wish to thank all the Parks and Natural Reserves where sampling sites are located for permissions and logistic support. Special thanks go to the staff of the Ticino and Boschi di Carrega Regional Parks, the Circeo National Park, the Natural Reserves of Bosco Siro Negri, Bosco della Mesola, Bosco Fontana, Montefalcone and Bosco Pantano, Oasi of WWF Policoro Herakleia and Oasi dei Quadris, Veneto Agricoltura, ERSAF Lombardia, as well as the Lame del Sesia and Groane Natural Park, for technical support during field campaigns. Marco Andrello, Catia Boggi, Fulvio Caronni, Maria Castellaneta, Maurizio Conticelli, Antonio Lapolla, Maurizio Marchi, Giorgio Matteucci, Libero Middei, Paola Nola, Danny Pietroni, Francesco Ripullone, Antonio Romano, Maurizio Rosito, Carlo Urbinati and Paolo Ventura have made important contribution at various stages of this research.

References

Agnoletti M (2018) Storia del bosco. Il paesaggio forestale italiano. Editori Laterza, Rome, Italy.

Andrello M, D'Aloia C, Dalongeville A, et al (2022) Evolving spatial conservation prioritization with intraspecific genetic data. Trends Ecol Evol 37:553–564. https://doi.org/10.1016/j.tree.2022.03.003

Aitken SN, Whitlock MC (2013) Assisted gene flow to facilitate local adaptation to climate change. Ann Rev Ecol Evol Syst 44:367–388. https://doi.org/10.1146/annurev-ecolsys-110512-135747

Ballian D, Belletti P, Ferrazzini D, et al (2010) Genetic variability of pedunculate oak (*Quercus robur* L.) in Bosnia and Herzegovina. Period Biol 112:353–362.

Belletti P, Leonardi S, Monteleone I, Piovani P (2005) Allozyme variation in different species of deciduous oaks from northwestern italy. Silvae Genet 54:9–16. https://doi.org/10.1515/sg-2005-0002

Belletti P, Calvo E, Ducci F, et al (2007) La foresta della Carpaneta e il programma per le dieci grandi foreste lombarde. Studi e ricerche preliminari. In: In: Ducci F (Eds.), Le Risorse genetiche della farnia della Val Padana. Progetto "Foresta della Carpaneta", Regione Lombardia – ERSAF, Milano, pp. 10-15.

Bennett KD, Tzedakis PC, Willis KJ (1991) Quaternary refugia of north European trees. J Biogeogr 18:103–115. https://doi.org/10.2307/2845248

Brewer S, Cheddadi R, de Beaulieu JL, Reille M (2002) The spread of deciduous *Quercus* throughout Europe since the last glacial period. For Ecol Manag 156:27–48. https://doi.org/10.1016/S0378-1127(01)00646-6

Bucci G, González-Martínez SC, Le Provost G, et al (2007) Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. Mol Ecol 16:2137–2153. https://doi.org/10.1111/j.1365-294X.2007.03275.x

Burger K, Müller M, Rogge M, Gailing O (2021) Genetic differentiation of indigenous (*Quercus robur* L.) and late flushing oak stands (*Q. robur* L. subsp. *slavonica* (Gáyer) Mátyás) in western Germany (North Rhine-Westphalia). Eur J Forest Res 140:1179–1194. https://doi.org/10.1007/s10342-021-01395-8

Buschbom J, Yanbaev Y, Degen B (2011) Efficient long-distance gene flow into an isolated relict oak stand. J Hered 102:464–472. https://doi.org/10.1093/jhered/esr023

Calvo E (2007) Consistenza delle risorse in Italia settentrionale. In: Ducci F (Eds.), Le Risorse genetiche della farnia della Val Padana. Progetto "Foresta della Carpaneta", Regione Lombardia – ERSAF, Milano, pp. 28–41.

Caudullo G, Welk E, San-Miguel-Ayanz J (2017) Chorological maps for the main European woody species. Data Brief 12:662–666. https://doi.org/10.1016/j.dib.2017.05.007

CBD (2021) First draft of the post-2020 global biodiversity framework. CBD/WG2020/3/3. UNEP. URL https://www.cbd.int/doc/c/914a/ eca3/24ad42235033f031badf61b1/wg2020-03-03-en.pdf

Chapuis MP, Raynal L, Plantamp C, et al (2020) A young age of subspecific divergence in the desert locust inferred by ABC random forest. Mol Ecol 29:4542–4558. https://doi.org/10.1111/mec.15663

Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. J Hered 100:106–113. https://doi.org/10.1093/jhered/esn088

Cook CN, Sgrò CM (2017) Aligning science and policy to achieve evolutionarily enlightened conservation.

Conserv Biol 31:501–512. https://doi.org/10.1111/cobi.12863

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014. https://doi.org/10.1093/genetics/144.4.2001

Cornuet JM, Santos F, Beaumont MA, et al (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. Bioinformatics 24:2713–2719. doi: 10.1093/bioinformatics/btn514

Cornuet J-M, Pudlo P, Veyssier J, et al (2014) DIYABC 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. https://doi.org/10.1093/bioinformatics/btt763

de Dato GD, Teani A, Mattioni C, et al (2020) Genetic analysis by nuSSR markers of silver birch (*Betula* pendula Roth) populations in their southern European distribution range. Front Plant Sci 11: 310. https://doi.org/10.3389/fpls.2020.00310

de Lafontaine G, Ducousso A, Lefèvre S, et al (2013) Stronger spatial genetic structure in recolonized areas than in refugia in the European beech. Mol Ecol 22:4397–4412. https://doi.org/10.1111/mec.12403

de Sampaio e Paiva Camilo-Alves C, da Clara MIE, de Almeida Ribeiro NMC (2013) Decline of Mediterranean oak trees and its association with *Phytophthora cinnamomi*: a review. Eur J Forest Res 132:411–432. https://doi.org/10.1007/s10342-013-0688-z

Degen B, Yanbaev Y, Ianbaev R, et al (2021) When does habitat fragmentation lead to changes in populations gene pool of pedunculate oak (*Quercus robur* L.)? For Ecol Manag 499:119617. https://doi.org/10.1016/j.foreco.2021.119617

Dering M, Chybicki I (2012) Assessment of genetic diversity in two-species oak seed stands and their progeny populations. Scand J For Res 27:2–9. https://doi.org/10.1080/02827581.2011.631934

Des Roches S, Pendleton LH, Shapiro B, Palkovacs EP (2021) Conserving intraspecific variation for nature's contributions to people. Nat Ecol Evol 5:574–582. https://doi.org/10.1038/s41559-021-01403-5

DeWoody JA, Harder AM, Mathur S, Willoughby JR (2021) The long-standing significance of genetic diversity in conservation. Mol Ecol 30:4147–4154. https://doi.org/10.1111/mec.16051

Díaz S, Zafra-Calvo N, Purvis A, et al (2020) Set ambitious goals for biodiversity and sustainability. Science 370:411–413. https://doi.org/10.1126/science.abe1530

Do C, Waples RS, Peel D, et al (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol Ecol Resour 14:209–214. https://doi.org/10.1111/1755-0998.12157

Dow BD, Ashley MV, Howe HF (1995) Characterization of highly variable (GA/CT) microsatellites in the bur oak, *Quercus macrocarpa*. Theor Appl Genet 91:137–141. https://doi.org/10.1007/BF00220870

Ducci F, Belletti P, Monteleone I, et al (2007) Gestione di popolazioni di farnia e rovere in Val Padana. In: Ducci F (Eds.), Le Risorse genetiche della farnia della Val Padana. Progetto "Foresta della Carpaneta", Regione Lombardia – ERSAF, Milano, pp. 42–67.

Dumolin-Lapègue S, Demesure B, Fineschi S, et al (1997) Phylogeographic structure of white oaks throughout the European continent. Genetics 146:1475–1487.

https://doi.org/10.1093/genetics/146.4.1475

Durand J, Bodénès C, Chancerel E, et al (2010) A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. BMC Genomics 11:570. https://doi.org/10.1186/1471-2164-11-570

Eaton E, Caudullo G, Oliveira S, de Rigo D (2016) *Quercus robur* and *Quercus petraea* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz J, de Rigo D, Caudullo G, et al (Eds.), European Atlas of Forest Tree Species. Publ. Off. EU, Luxembourg, pp. E01c6df+

Estoup A, Jarne P, Cornuet J-M (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol Ecol 11:1591–1604. https://doi.org/10.1046/j.1365-294X.2002.01576.x

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491. https://doi.org/10.1093/genetics/131.2.479

Fady B, Cottrell J, Ackzell L, et al (2016) Forests and global change: what can genetics contribute to the major forest management and policy challenges of the twenty-first century? Reg Environ Change 16:927–939. https://doi.org/10.1007/s10113-015-0843-9

Fady B, Bozzano M (2021) Effective population size does not make a practical indicator of genetic diversity in forest trees. Biol Conserv 253:108904. https://doi.org/10.1016/j.biocon.2020.108904

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 164:1567–1587. https://doi.org/10.1093/genetics/164.4.1567

Fineschi S, Taurchini D, Grossoni P, et al (2002) Chloroplast DNA variation of white oaks in Italy. For Ecol Manag 156:103–114. https://doi.org/10.1016/S0378-1127(01)00637-5

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a bayesian perspective. Genetics 180:977–993. https://doi.org/10.1534/genetics.108.092221

Fortini P, Di Pietro R, Proietti E, et al (2023) Dissecting the continuum and unravelling the phylogeographic knot of plastid DNA in European white oaks (*Quercus* sect. *Quercus*): ancient signatures and multiple diversity reservoirs. Eur J For Res. https://doi.org/10.1007/s10342-023-01610-8

Frankham R (2022) Evaluation of proposed genetic goals and targets for the Convention on Biological Diversity. Conserv Genet 23:865–870. https://doi.org/10.1007/s10592-022-01459-1

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to Conservation Genetics. Cambridge University Press, Cambridge, UK.

Franklin I (1980) Evolutionary change in small populations. In: Soulé M, Wilcox BA (Eds.), Conservation biology: and evolutionary-ecological perspective, Sinauer Associates. Inc., Sunderland, US, pp. 135-150.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. Mol Ecol 10:305–318. https://doi.org/10.1046/j.1365-294x.2001.01190.x

Gasparini P, Di Cosimo L, Floris A, De Laurentis (2022) Italian National Forest Inventory—Methods and Results of the Third Surve. Springer, Berlin/Heidelberg, Germany.

Gathercole LAP, Nocchi G, Brown N, et al (2021) Evidence for the widespread occurrence of bacteria implicated in Acute Oak Decline from incidental genetic sampling. Forests 12:1683. https://doi.org/10.3390/f12121683

Gerber S, Chadœuf J, Gugerli F, et al (2014) High rates of gene flow by pollen and seed in oak populations across Europe. PLoS ONE 9:e85130. https://doi.org/10.1371/journal.pone.0085130

Gibbs J (1997) Biotic and abiotic factors affecting the dying back of pedunculate oak *Quercus robur* L. Forestry 70:399–406. https://doi.org/10.1093/forestry/70.4.399

Guichoux E, Lagache L, Wagner S, et al (2011) Two highly validated multiplexes (12-plex and 8-plex) for species delimitation and parentage analysis in oaks (*Quercus* spp). Mol Ecol Resour 11:578–585. https://doi.org/10.1111/j.1755-0998.2011.02983.x

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters: rear edges and climate change. Ecol Lett 8:461–467. https://doi.org/10.1111/j.1461-0248.2005.00739.x

Hoban S, Bruford M, D'Urban Jackson J, et al (2020) Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. Biol Conserv 248:108654. https://doi.org/10.1016/j.biocon.2020.108654

Hoban S, Paz-Vinas I, Aitken S, et al (2021) Effective population size remains a suitable, pragmatic indicator of genetic diversity for all species, including forest trees. Biol Conserv 253:108906. https://doi.org/10.1016/j.biocon.2020.108906

Hoban S, Bruford MW, da Silva JM, et al (2023) Genetic diversity goals and targets have improved, but remain insufficient for clear implementation of the post-2020 global biodiversity framework. Conserv Genet 24:181–191. https://doi.org/10.1007/s10592-022-01492-0

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x

Hughes AR, Inouye BD, Johnson MTJ, et al (2008) Ecological consequences of genetic diversity: ecological effects of genetic diversity. Ecol Lett 11:609–623. https://doi.org/10.1111/j.1461-0248.2008.01179.x

IPBES (2019) Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Díaz S, et al (Eds). IPBES, Bonn, Germany. 56 pages.

Jamieson IG, Allendorf FW (2012) How does the 50/500 rule apply to MVPs? Trends Ecol Evol 27:578–584. https://doi.org/10.1016/j.tree.2012.07.001

Jaramillo-Correa JP, Bagnoli F, Grivet D, et al (2020) Evolutionary rate and genetic load in an emblematic Mediterranean tree following an ancient and prolonged population collapse. Mol Ecol 29:4797–4811. https://doi.org/10.1111/mec.15684

Kalinowski ST (2005) hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Mol Ecol Notes 5:187–189. https://doi.org/10.1111/j.1471-8286.2004.00845.x

Kampfer S, Lexer C, Glössl J, Steinkellner H (1998) Characterization of $(GA)_n$ microsatellite loci from *Quercus*

robur. Hereditas 129:183–186. https://doi.org/10.1111/j.1601-5223.1998.00183.x

Kesić L, Cseke K, Orlović S, et al (2021) Genetic diversity and differentiation of pedunculate oak (*Quercus robur* L.) populations at the southern margin of its distribution range—Implications for Conservation. Diversity 13:371. https://doi.org/10.3390/d13080371

Kitamura K, Namikawa K, Tsuda Y, et al (2022) Possible northern persistence of Siebold's beech, *Fagus crenata*, at its northernmost distribution limit on an island in Japan Sea: Okushiri Island, Hokkaido. Front. Plant Sci 13:990927. doi: 10.3389/fpls.2022.990927

Kopelman NM, Mayzel J, Jakobsson M, et al (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour 15:1179–1191. https://doi.org/10.1111/1755-0998.12387

Lagercrantz U, Ryman N (1990) Genetic structure of Norway spruce (*Picea abies*): concordance of morphological and allozymic variation. Evolution 44:38–53. https://doi.org/10.1111/j.1558-5646.1990.tb04278.x

Laikre L, Hoban S, Bruford MW, et al (2020) Post-2020 goals overlook genetic diversity. Science 367:1083–1085. https://doi.org/10.1126/science.abb2748

Leigh DM, Hendry AP, Vázquez-Domínguez E, Friesen VL (2019) Estimated six per cent loss of genetic variation in wild populations since the industrial revolution. Evol Appl 12:1505–1512. https://doi.org/10.1111/eva.12810

Li Y-L, Liu J-X (2018) StructureSelector: a web-based software to select and visualize the optimal number of clusters using multiple methods. Mol Ecol Resour 18:176–177. https://doi.org/10.1111/1755-0998.12719

Liepelt S, Cheddadi R, de Beaulieu J-L, et al (2009) Postglacial range expansion and its genetic imprints in *Abies alba* (Mill.) — A synthesis from palaeobotanic and genetic data. Rev Palaeobot Palynol 153:139–149. https://doi.org/10.1016/j.revpalbo.2008.07.007

López-Cortegano E, Pérez-Figueroa A, Caballero A (2019) metapop 2: Re-implementation of software for the analysis and management of subdivided populations using gene and allelic diversity. Mol Ecol Resour 19:1095–1100. https://doi.org/10.1111/1755-0998.13015

Lowe AJ, Cavers S, Boshier D, et al (2015) The resilience of forest fragmentation genetics—no longer a paradox—we were just looking in the wrong place. Heredity 115:97–99. https://doi.org/10.1038/hdy.2015.40

Magri D, Agrillo E, Di Rita F, et al (2015) Holocene dynamics of tree taxa populations in Italy. Rev Palaeobot Palynol 218:267–284. https://doi.org/10.1016/j.revpalbo.2014.08.012

Mayol M, Riba M, González-Martínez SC, et al (2015) Adapting through glacial cycles: insights from a long-lived tree (*Taxus baccata*). New Phytol 208:973–986. https://doi.org/10.1111/nph.13496

McNeely JA, Miller KR, Reid WV, et al (1990) Conserving the World's Biological Diversity. IUCN, World Resources Institute, Conservation International, WWF-US and the World Bank, Washington, DC.

Milesi P, Kastally C, Dauphin B, et al (2023) Synchronous effective population size changes and genetic stability of forest trees through glacial cycles. BioRXiv. https://doi.org/10.1101/2023.01.05.522822

Moracho E, Moreno G, Jordano P, Hampe A (2016) Unusually limited pollen dispersal and connectivity of pedunculate oak (*Quercus robur*) refugial populations at the species' southern range margin. Mol Ecol 25:3319–3331. https://doi.org/10.1111/mec.13692

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics 28:2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pericolo O, Camarero JJ, Colangelo M, et al (2023) Species specific vulnerability to increased drought in temperate and Mediterranean floodplain forests. Agric For Meteorol 328:109238. https://doi.org/10.1016/j.agrformet.2022.109238

Petit RJ, Aguinagalde I, de Beaulieu J-L, et al (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300:1563–1565. https://doi.org/10.1126/science.1083264

Petit RJ, Brewer S, Bordács S, et al (2002a) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. For Ecol Manag 156:49–74. https://doi.org/10.1016/S0378-1127(01)00634-X

Petit RJ, Csaikl UM, Bordacs S, et al (2002b) Chloroplast DNA variation in European white oaks. Phylogeography and patterns of diversity based on data from over 2600 populations. For Ecol Manag 156:5–26. https://doi.org/10.1016/S0378-1127(01)00645-4

Petit RJ, Latouche-Hallé C, Pemonge M-H, Kremer A (2002c) Chloroplast DNA variation of oaks in France and the influence of forest fragmentation on genetic diversity. For Ecol Manag 156:115–129. https://doi.org/10.1016/S0378-1127(01)00638-7

Pierson JC, Coates DJ, Oostermeijer JGB, et al (2016) Genetic factors in threatened species recovery plans on three continents. Front Ecol Environ 14:433–440. https://doi.org/10.1002/fee.1323

Pina-Martins F, Baptista J, Pappas G, Paulo OS (2019) New insights into adaptation and population structure of cork oak using genotyping by sequencing. Glob Change Biol 25:337–350. https://doi.org/10.1111/gcb.14497

Piotti A, Leonarduzzi C, Postolache D, et al (2017) Unexpected scenarios from Mediterranean refugial areas: disentangling complex demographic dynamics along the Apennine distribution of silver fir. J Biogeogr 44:1547–1558. https://doi.org/10.1111/jbi.13011

Postolache D, Oddou-Muratorio S, Vajana E, et al (2021) Genetic signatures of divergent selection in European beech (*Fagus sylvatica* L.) are associated with the variation in temperature and precipitation across its distribution range. Mol Ecol 30:5029–5047. https://doi.org/10.1111/mec.16115

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959. https://doi.org/10.1093/genetics/155.2.945

Ragazzi A, Dellavalle I, Moricca S, et al (2000) Decline of oak species in Italy, problems and perspectives. Accademia Italiana di Scienze Forestali, Florence, Italy.

Sagnard F, Oddou-Muratorio S, Pichot C, et al (2011) Effects of seed dispersal, adult tree and seedling density on the spatial genetic structure of regeneration at fine temporal and spatial scales. Tree Genet Genomes 7:37–48. https://doi.org/10.1007/s11295-010-0313-y

Sandurska E, Ulaszewski B, Burczyk J (2019) Genetic diversity and differentiation of coexisting populations of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. Acta Biol Crac Ser Bot 61:17–28. https://doi.org/10.24425/ABCSB.2019.127739

Santos-del-Blanco L, Olsson S, Budde KB, et al (2022) On the feasibility of estimating contemporary effective population size (Ne) for genetic conservation and monitoring of forest trees. Biol Conserv 273:109704. https://doi.org/10.1016/j.biocon.2022.109704

Scherrer D, Körner C (2011) Topographically controlled thermal-habitat differentiation buffers alpine plant diversity against climate warming. J Biogeog 38:406-416. https://doi.org/10.1111/j.1365-2699.2010.02407.x

Scotti-Saintagne C, Boivin T, Suez M, et al (2021) Signature of mid-Pleistocene lineages in the European silver fir (*Abies alba* Mill.) at its geographic distribution margin. Ecol Evol 11:10984–10999. https://doi.org/10.1002/ece3.7886

Sgrò CM, Lowe AJ, Hoffmann AA (2011) Building evolutionary resilience for conserving biodiversity under climate change: Conserving biodiversity under climate change. Evol Appl 4:326–337. https://doi.org/10.1111/j.1752-4571.2010.00157.x

Steinkellner H, Lexer C, Turetschek E, Glössl J (1997) Conservation of $(GA)_n$ microsatellite loci between *Quercus* species. Mol Ecol 6:1189–1194. https://doi.org/10.1046/j.1365-294X.1997.00288.x

Stojanović DB, Levanič T, Matović B, Orlović S (2015) Growth decrease and mortality of oak floodplain forests as a response to change of water regime and climate. Eur J Forest Res 134:555–567. https://doi.org/10.1007/s10342-015-0871-5

Thomas FM, Blank R, Hartmann G (2002) Abiotic and biotic factors and their interactions as causes of oak decline in Central Europe. Forest Pathol 32:277–307. https://doi.org/10.1046/j.1439-0329.2002.00291.x

Tollefsrud MM, Kissling R, Gugerli F, et al (2008) Genetic consequences of glacial survival and postglacial colonization in Norway spruce: combined analysis of mitochondrial DNA and fossil pollen. Mol Ecol 17:4134–4150. https://doi.org/10.1111/j.1365-294X.2008.03893.x

Traill LW, Brook BW, Frankham RR, Bradshaw CJA (2010) Pragmatic population viability targets in a rapidly changing world. Biol Conserv 143:28–34. https://doi.org/10.1016/j.biocon.2009.09.001

Travascia D, Castellaneta M, Colucci A, et al (2022) Valutazione dei fattori di vulnerabilità della riserva naturale Bosco Pantano di Policoro: l'approccio dinamico SWOT - AHP. Forest@ 19:67-73. https://doi.org/10.3832/efor4133-019

Tzedakis PC, Andrieu V, de Beaulieu J-L, et al (1997) Comparison of terrestrial and marine records of changing climate of the last 500,000 years. Earth Planet Sci Lett 150:171–176. https://doi.org/10.1016/S0012-821X(97)00078-2

Vajana E, Bozzano M, Marchi M, Piotti A (2023) On the inclusion of adaptive potential in species distribution models: towards a genomic-informed approach to forest management and conservation. Environments 10:3. https://doi.org/10.3390/environments10010003

Vakkari P, Blom A, Rusanen M, et al (2006) Genetic variability of fragmented stands of pedunculate oak (*Quercus robur*) in Finland. Genetica 127:231–241. https://doi.org/10.1007/s10709-005-4014-7

Vendramin GG, Anzidei M, Madaghiele A, et al (2000) Chloroplast microsatellite analysis reveals the presence of population subdivision in Norway spruce (*Picea abies* K.). Genome 43:68–78. https://doi.org/10.1139/g99-093

Vendramin GG, Fady B, González-Martínez SC, et al (2008) Genetically depauperate but widespread: the case of an emblematic mediterranean pine. Evolution 62:680–688. https://doi.org/10.1111/j.1558-5646.2007.00294.x

Vranckx G, Jacquemyn H, Mergeay J, et al (2014) Transmission of genetic variation from the adult generation to naturally established seedling cohorts in small forest stands of pedunculate oak (*Quercus*

robur L.). For Ecol Manag 312:19–27. https://doi.org/10.1016/j.foreco.2013.10.027

Vries SMG de, Alan M, Bozzano M, et al (2015) Pan-European strategy for genetic conservation of forest trees and establishment of a core network of dynamic conservation units, European Forest Genetic Resources Programme (EUFORGEN). Bioversity International, Rome, Italy.

Wagner S, Liepelt S, Gerber S, Petit RJ (2015) Within-range translocations and their consequences in European larch. PLoS ONE 10:e0127516. https://doi.org/10.1371/journal.pone.0127516

Watts W (1996) Vegetation history and climate of the last 15,000 years at Laghi Di Monticchio, southern Italy. Quat Sci Rev 15:113–132. https://doi.org/10.1016/0277-3791(95)00038-0

Williamson-Natesan EG (2005) Comparison of methods for detecting bottlenecks from microsatellite loci. Conserv Genet 6:551–562. https://doi.org/10.1007/s10592-005-9009-5

Table 1. Geographic coordinates and sampling size of *Quercus robur* populations sampled for the present study.

Population	Label	Longitude	Latitude	N sampling	N final[§]
Bosco Merlino	MER	7.71317	44.78999	30	30
Lame del Sesia	LDS	8.39053	45.44156	30	30
Turbigaccio	TUR	8.71116	45.56480	26	26
Fagiana	FGN	8.82743	45.43451	30	30
Geraci	GER	8.98614	45.27786	30	30
Bosco Siro Negri	SIR	9.05683	45.21020	26[a]	26
Groane	GRO	9.09376	45.66090	30	30
Paladina	PAL	9.62455	45.72850	30	30
Capriano del Colle	CDC	10.14840	45.45162	30[a]	30
Bosco Fontana	FON	10.74915	45.19796	30	29
Moriago della Battaglia	MOR	12.13213	45.86176	30	30
Cavalier	CAV	12.55146	45.76378	30[a]	30
Fagagna	FAG	13.08088	46.13015	30	30
Bosco della Mesola	MES	12.25539	44.85905	30	29
Bosco della Panfilia	PFL	11.37699	44.78402	30[a]	30
Boschi di Carrega	BDC	10.20724	44.72643	24[a]	24
Riserva di Montefalcone	MON	10.70773	43.74992	38	27
Cascine di Tavola	TAV	11.06098	43.84190	24[a]	24
Fabriano	FBN	12.89251	43.35388	25[a]	22
Fabro	FAB	11.98985	42.90642	24[a]	24
Orvieto	ORV	12.09845	42.71113	24[a]	24
Foglino	FOG	12.71482	41.47387	26[a]	26
Circeo	CRC	13.04465	41.34019	24	24
Pollutri	POL	14.64420	42.14798	30	30
Bosco Pantano	PAN	16.68015	40.15543	64[a]	63

[a] small and isolated *Q. robur* populations, with less than 500 reproductive trees

[§] final number of individuals retained for each population after excluding trees that were misclassified as *Q. robur* in the field but that turned out to be *Q. petraea* individuals based on the results of a Principal Coordinate Analysis on the matrix of pairwise individual genetic distance

Table 2. Genetic characteristics of sampled populations. *Na*: mean number of alleles, *Ar*: allelic richness, *PAr*: private allelic richness, H_O: observed heterozygosity, H_E: expected heterozygosity, F_{IS}: inbreeding coefficient, F_{ST}: plot-specific differentiation index according to Foll and Gaggiotti (2008), *Ne* and *Ne* $CI_{95\%}$: effective population size and its parametric 95% confidence interval.

Population	***Na***	***Ar***	***PAr***	H_O	H_E	F_{IS}	F_{ST}	***Ne***	***Ne*** $CI_{95\%}$
MER	7.375	6.771	0.106	0.719	0.679	0.007	0.020	115	66-339
LDS	7.875	7.194	0.051	0.690	0.698	0.033	0.012	655	143-∞
TUR	7.688	7.253	0.010	0.676	0.691	0.026	0.013	∞	244-∞
FGN	7.875	7.146	0.059	0.679	0.690	0.014	0.010	∞	196-∞
GER	7.938	7.192	0.044	0.667	0.675	0.009	0.012	97	61-214
SIR	7.750	7.353	0.067	0.725	0.726	0.015	0.016	1330	139-∞
GRO	7.625	6.955	0.013	0.698	0.684	0.010	0.013	∞	382-∞
PAL	8.188	7.409	0.005	0.735	0.697	0.006	0.012	522	134-∞
CDC	7.375	6.762	0.121	0.671	0.684	0.021	0.023	106	61-315
FON	7.563	6.939	0.020	0.640	0.673	0.027	0.012	272	101-∞
MOR	7.625	6.972	0.091	0.721	0.683	0.004	0.025	46	35-65
CAV	7.938	7.224	0.019	0.688	0.696	0.016	0.008	∞	240-∞
FAG	7.375	6.718	0.028	0.654	0.654	0.011	0.025	1061	144-∞
MES	7.625	7.028	0.091	0.695	0.699	0.015	0.017	3248	158-∞
PFL	7.750	7.103	0.045	0.685	0.694	0.019	0.014	285	107-∞
BDC	7.813	7.525	0.033	0.720	0.696	0.013	0.008	∞	256-∞
MON	7.563	7.055	0.078	0.679	0.682	0.016	0.026	163	76-∞
TAV	7.625	7.340	0.231	0.734	0.701	0.006	0.023	1348	144-∞
FBN	7.688	7.612	0.139	0.721	0.696	0.007	0.027	55	38-92
FAB	6.250	6.037	0.101	0.679	0.624	0.006	0.080	22	17-29
ORV	8.188	7.877	0.032	0.732	0.725	0.008	0.021	∞	184-∞
FOG	7.313	6.878	0.156	0.734	0.692	0.009	0.053	36	27-54
CRC	7.625	7.325	0.075	0.732	0.706	0.007	0.038	141	73-965
POL	7.000	6.447	0.116	0.679	0.666	0.011	0.055	60	41-102
PAN	7.313	5.801	0.151	0.657	0.663	0.021	0.128	63	51-81

Table 3. Heterozygosity excess and M-ratio deficiency tests for assessing the occurrence of past genetic bottlenecks in *Quercus robur* populations. *H* and MR are, respectively, the observed values of expected heterozygosity and the M-ratio, while Δ*H* and ΔMR are the differences between such observed values and the values expected at the equilibrium. *P* is the *P*-value of the Wilcoxon signed-rank test used to evaluate the statistical significance of such differences (* *P*<0.1, ** *P*<0.05, *** *P*<0.01).

Population	***H***	**Δ*H***	***P***		**MR**	**ΔMR**	***P***	
MER	0.6895	0.0517	0.6833		0.6960	0.0913	0.0223	**
LDS	0.7104	0.0601	0.9529		0.7133	0.0623	0.0889	*
TUR	0.7051	0.0662	0.9627		0.6597	0.1102	0.0526	*
FGN	0.7016	0.0642	0.8840		0.7357	0.0411	0.3920	
GER	0.6866	0.0796	0.9711		0.6639	0.1128	0.0031	***
SIR	0.7407	0.0268	0.6278		0.6884	0.0821	0.0373	**
GRO	0.6958	0.0642	0.9416		0.6679	0.1128	0.0077	***
PAL	0.7088	0.0625	0.9414		0.7299	0.0449	0.2176	
CDC	0.6957	0.0476	0.7195		0.6997	0.0861	0.0370	**
FON	0.6854	0.0686	0.9128		0.7021	0.0787	0.1494	
MOR	0.6940	0.0681	0.9631		0.6656	0.1138	0.0469	**
CAV	0.7085	0.0555	0.9280		0.6930	0.0856	0.0531	*
FAG	0.6649	0.0778	0.9955		0.6829	0.1045	0.0368	**
MES	0.7117	0.0511	0.8629		0.6802	0.0972	0.0415	**
PFL	0.7059	0.0485	0.8739		0.6469	0.1335	0.0091	***
BDC	0.7107	0.0668	0.9035		0.7023	0.0603	0.1383	
MON	0.6954	0.0730	0.9934		0.7249	0.0474	0.1623	
TAV	0.7152	0.0575	0.9631		0.7061	0.0604	0.2182	
FBN	0.7119	0.0522	0.8627		0.7288	0.033	0.2811	
FAB	0.6369	0.0643	0.9478		0.6801	0.1178	0.0531	*
ORV	0.7410	0.0534	0.9962		0.7643	-0.0091	0.4931	
FOG	0.7050	0.0384	0.9283		0.6830	0.0977	0.0195	**
CRC	0.7212	0.0416	0.9537		0.7577	0.011	0.5511	
POL	0.6777	0.0482	0.8391		0.7023	0.0902	0.0528	*
PAN	0.6689	0.0443	0.8841		0.6683	0.1458	0.0017	***

Table 4. Posterior probabilities of the six demographic scenarios tested and their 95% confidence intervals based on DIYABC logistic estimation.

Scenario	Posterior probability	95% CI (lower - upper)
1	0.1296	0.1105 - 0.1487
2	0.1394	0.1211 - 0.1577
3	0.2698	0.2445 - 0.2952
4	0.1365	0.1188 - 0.1541
5	0.2249	0.1986 - 0.2511
6	0.0999	0.0836 - 0.1161

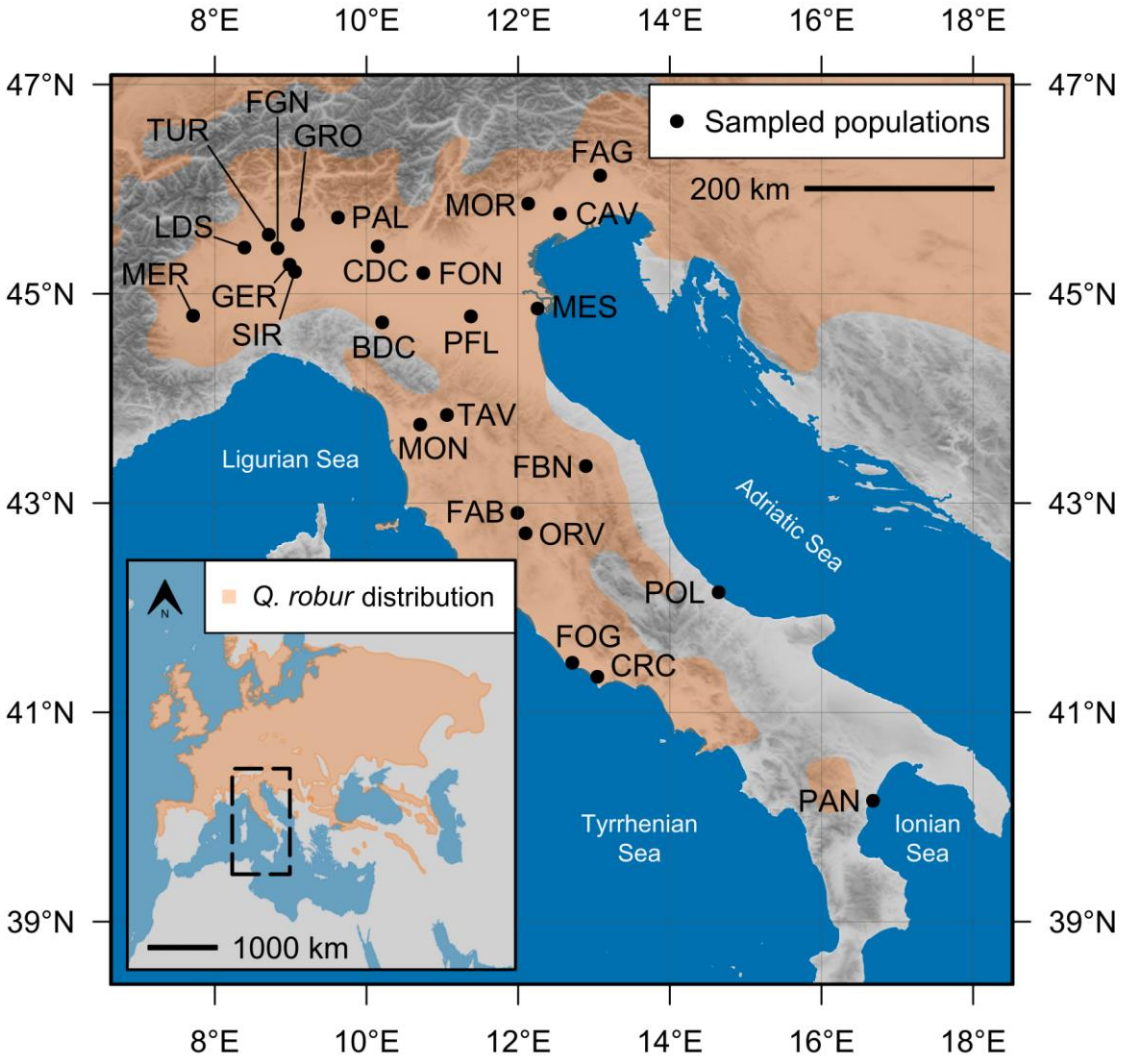

Figure 1. Map of *Quercus robur* populations sampled for the present study. The inset represents the distribution of the species in Europe according to Caudullo et al. (2017).

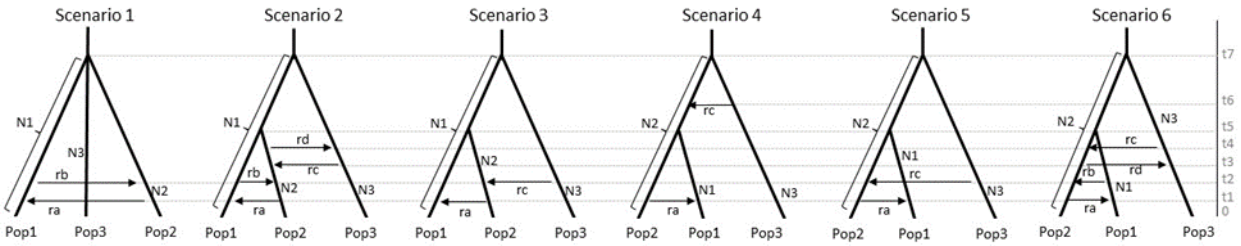

Figure 2. Six demographic models tested by using the Approximate Bayesian Computation approach implemented in DIYABC (Cornuet et al. 2008, 2014). Pop1, Pop2 and Pop3 are the three gene pools resulting from the Bayesian clustering analysis and represent, respectively, individuals from northern Italian populations, central Italian populations and the southernmost stand of PAN. t# = time-scaled by generation time; N# effective population size; r = admixture ratio.

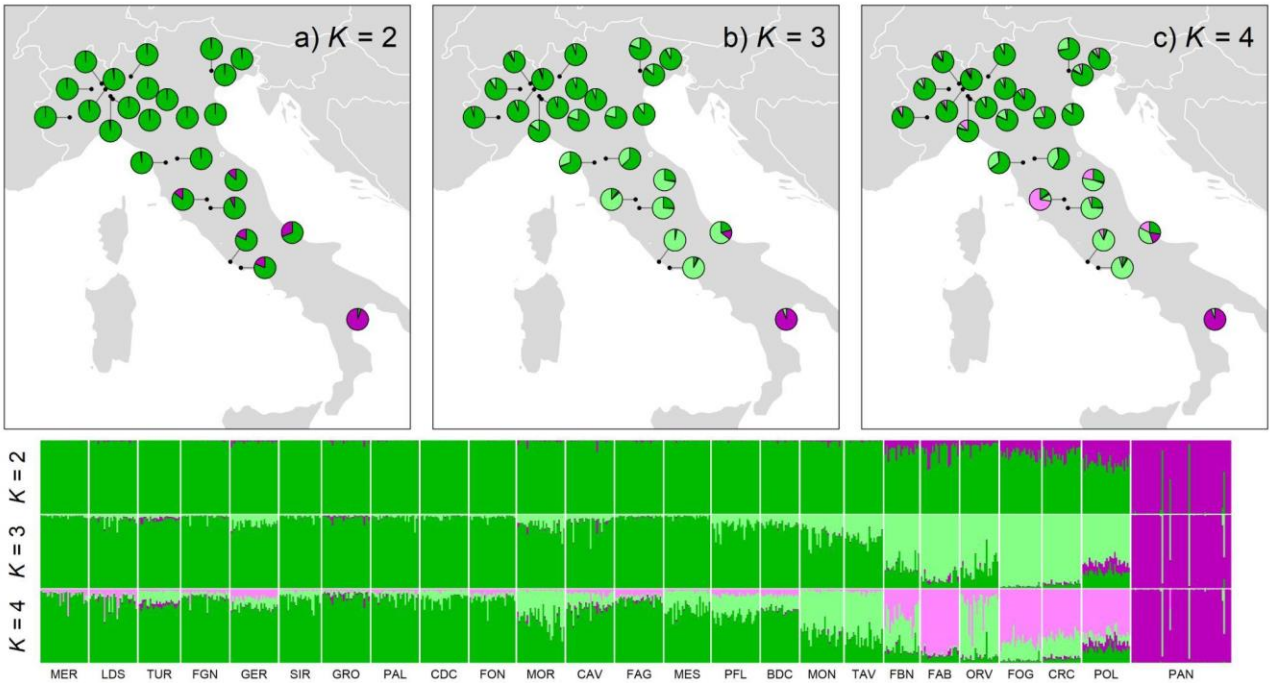

Figure 3. Results of the Bayesian clustering analysis carried out with STRUCTURE. For each population, average ancestry coefficients at (*a*) *K* = 2, (*b*) *K* = 3 and (*c*) *K* = 4 are shown. In the lower panel, individual ancestry coefficients are presented.

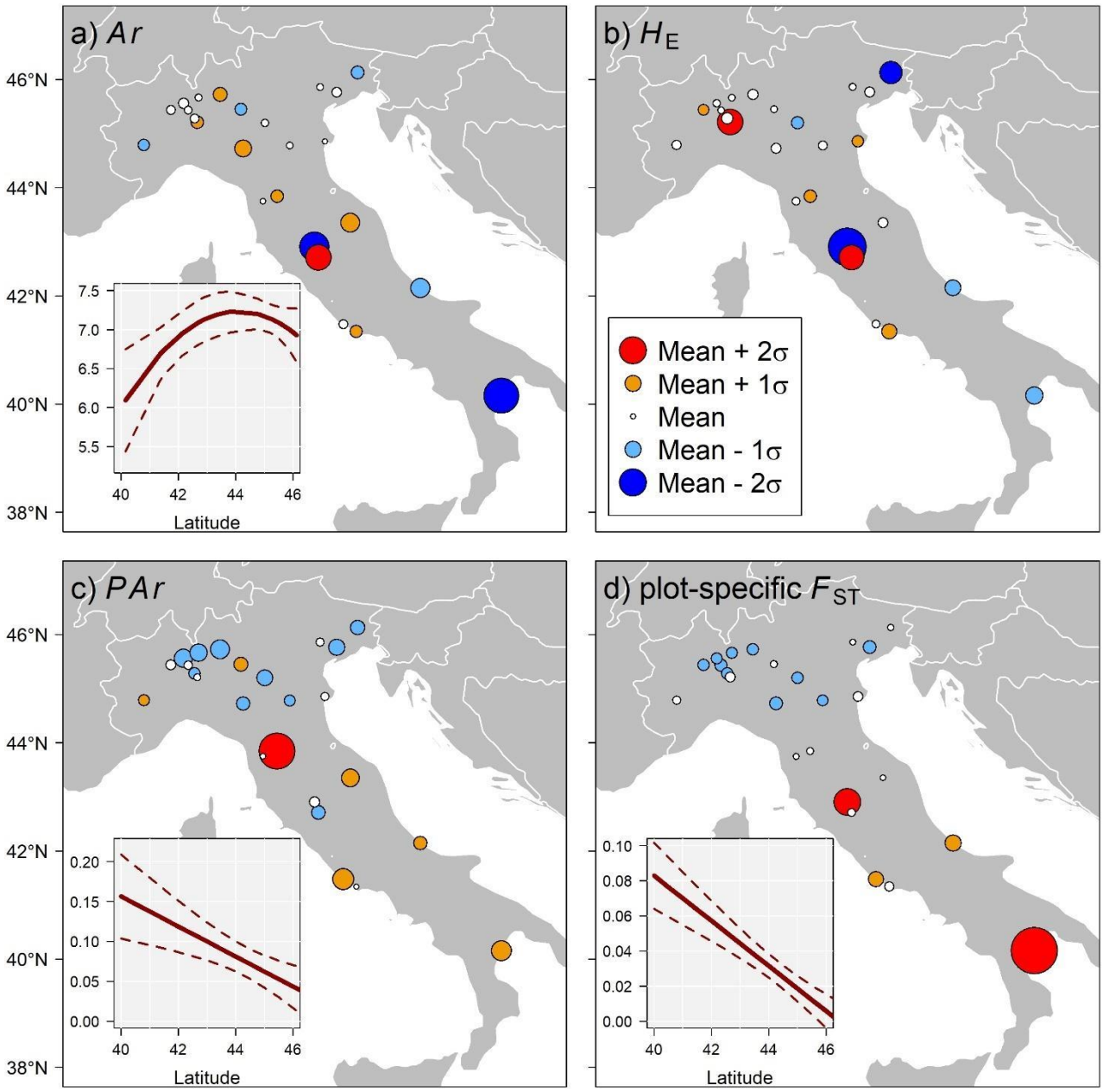

Figure 4. Estimates of allelic richness (*Ar*), expected heterozygosity (H_E), private allelic richness (*PAr*) and genetic distinctiveness (plot-specific F_{ST}) in the 25 *Quercus robur* populations of the present study. Circle sizes are proportional to the deviation from the mean. Colours indicate the direction of such deviations: white circles represent populations showing around-average values, orange and red circles populations showing above-average values, light blue and blue circles populations showing below-average values.

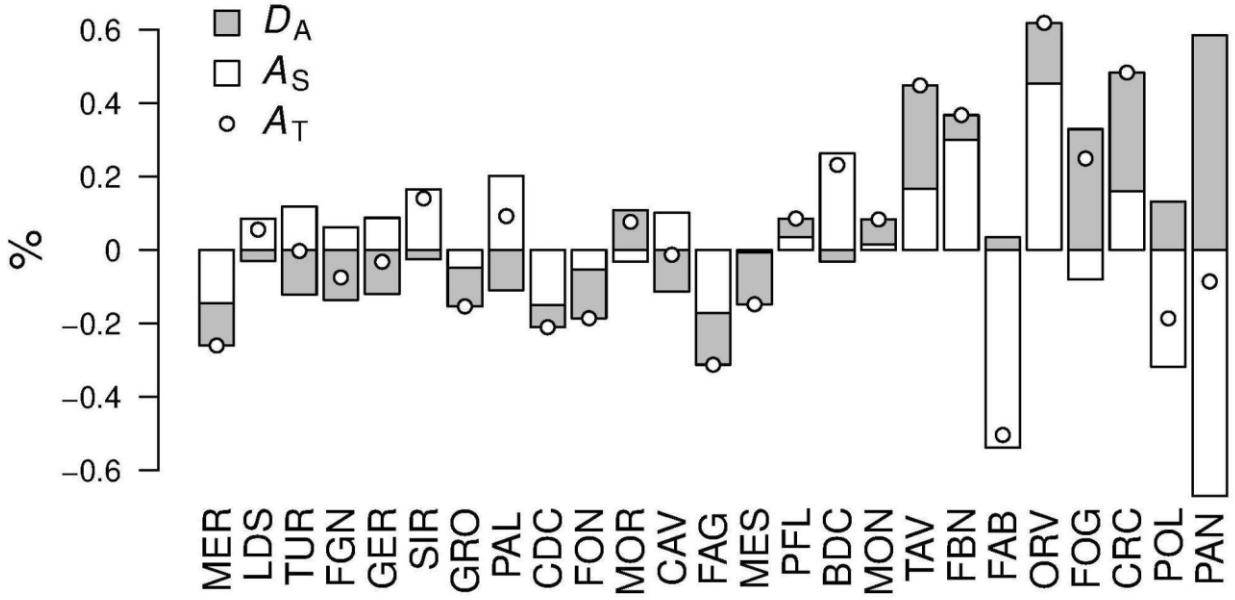

Figure 5. Contributions to within- (A_S; white), between- (D_A; grey) and total (A_T; white dots) allelic diversity of *Quercus robur* populations.