



Genetic and morphometric insights into the Sardinian deciduous oaks (subgen. *Quercus*, sect. *Quercus*): taxonomy and conservation

Elisa Proietti¹ · Romeo Di Pietro^{1,2} · Piera Di Marzio^{2,3} · Katharina Birgit Budde^{4,6} · Emmanuele Farris⁵ · Oliver Gailing⁶ · Leonardo Rosati⁷ · Paola Fortini^{2,3}

Received: 8 March 2026 / Accepted: 19 May 2026 / Published online: 25 June 2026
© The Author(s) 2026

Abstract

This study examines the genetic and morphological variability within the *Quercus pubescens* s.l. complex of Sardinian white oak populations, including *Q. congesta*, *Q. ichnusae*, and *Q. virgiliana*. These tree taxa act as umbrella species in several climax communities across the Sardinia Island. A comparison of actual and potential cover of the different types of Sardinian deciduous oak forests was carried out. Ten EST-SSR markers were analysed to examine genetic diversity and differentiation within and among populations. Genetic structure was assessed using Bayesian clustering methods while relationships between populations were assessed by means of multivariate analyses. Five additional *Q. pubescens* populations from Italy, Spain, and Croatia and three populations belonging to *Q. frainetto*, *Q. petraea*, and *Q. robur* from peninsular Italy were also included in the genetic analyses. The morphometric analysis was performed on twelve leaf morphological traits focussing exclusively on the Sardinian populations. The results showed that deciduous oak forests occupy only a fraction of their potential area. Xero-thermic pubescent oak forests showed the greatest loss of their potential coverage. Neither genetic nor morphometric data revealed sufficient levels of diversity between populations to justify the current taxonomic classification of Sardinian pubescent oaks into four different taxa. Minimal genetic differentiation was found within the entire group of *Q. pubescens* s.l., whereas a clear separation among *Q. frainetto*, *Q. petraea*, *Q. pubescens*, and *Q. robur* was highlighted. Nevertheless, two major gene pools were identified within the *Q. pubescens* s.l. group with geographic distance among populations having a predominant influence over taxonomic boundaries.

Keywords Genetic diversity · Leaf morphological traits · *Quercus congesta* · *Q. ichnusae* · *Q. pubescens* · *Q. virgiliana* · Sardinia island · Taxonomy · White oaks

Handling Editor: Karol Marhold.

✉ Romeo Di Pietro
romeo.dipietro@uniroma1.it

¹ Department of Planning, Design, and Technology of Architecture, University of Rome Sapienza, Via Flaminia 72, 00196 Rome, Italy

² NBFC, National Biodiversity Future Center, 90133 Palermo, Italy

³ Department of Biosciences and Territory, University of Molise, Fonte Lappone, 86090 Pesche, Italy

⁴ Department of Forest Genetic Resources, Northwest German Forest Research Institute, Professor-Oelkers-Strasse 6, 34346 Hann Muenden, Germany

⁵ Department of Chemical, Physical, Mathematical and Natural Sciences, University of Sassari, Via Vienna 2, 07100 Sassari, Italy

⁶ Department of Forest Genetics and Forest Tree Breeding, University of Göttingen, 37077 Göttingen, Germany

⁷ Department of Health Sciences, University of Basilicata, Via Ateneo Lucano 10, 85100 Potenza, Italy

Introduction

Islands and archipelagos tend to exhibit a high rate of plant endemism due to their isolation from surrounding landmasses. A high rate of tree endemism is quite a common phenomenon especially on large islands (e.g. Madagascar, New Zealand, New Caledonia, Cuba, etc.) and in the highly isolated oceanic archipelagos (e.g. Hawaii Islands, Canary Islands, etc.) where endemic genera and families of woody plants are found (see Takhtajan 1986). In contrast, islands and archipelagos situated within inland seas generally exhibit considerably lower level of tree endemism. Inland seas are significantly smaller in size and characterized by lower water depths than oceans. Accordingly, a partial or complete network of land connections between islands or between islands and adjacent landmasses may have repeatedly formed during stages of sea regression, facilitating the migration of tree species. In terms of biological conservation, the Mediterranean basin is considered one of the world's biodiversity hotspots (Myers 2003; Médail and Myers 2004). Its highly diverse flora and vegetation gave rise to high rates of rare and endemic species (Quezél 1999). Covered by forests to the tune of 82%, the Mediterranean basin certainly stands out, in the temperate climate context, for its diversity of forest tree species, boasting 245 native taxa, 46 of which are endemic (Médail et al. 2019). However, the occurrence of endemic species in individual islands is mostly linked to shrubby or herbaceous species and rarely involves tree species. This may be due to various factors, such as trees' long average generation time compared to herbs or shrubs, their high cross-fertilization linked to long-distance gene flow, or their natural tendency to give rise to climax forests that extend uninterruptedly over vast areas ensuring high gene flow continuity (Aitken and Bemmels 2015). The effectiveness of these factors is linked to the discontinuous geographical isolation in space and time that has characterized all the Mediterranean islands to a greater or lesser extent (Médail 2017; Thompson 2020). Paleoclimatic and paleogeographic events, such as the Messinian salinity crisis and sea regression during the Quaternary Ice Ages, have repeatedly resulted in land connections forming between the islands and continental areas (La Greca 1980). This has opened migration routes and allowed interbreeding between previously isolated populations of forest trees. Nevertheless, there are some known examples of tree species that are endemic to individual Mediterranean islands, such as *Cedrus brevifolia* (Hook.f.) A.Henry in Cyprus, *Abies nebrodensis* (Lojac.) Mattei, and *Betula aetnensis* Raf. in Sicily, or subendemic, such as *Phoenix theophrastii* Greuter in Crete. Regarding *Quercus*, the dominant tree genus in southern Europe's forested areas, there are reports

of endemic oak species among the Mediterranean islands, although there is no complete agreement on their taxonomic consistency. There should be no doubt that *Q. alniifolia* Poech (subgen. *Cerris*, sect. *Cerris*) deserves the rank of endemic species in Cyprus. Most floras and checklists grant taxonomic autonomy to *Q. trojana* subsp. *euboica* (Papaioannou) K.I.Chr., a subspecies that is endemic to the island of Crete.

Other examples of endemic oaks occur in Sicily and Sardinia, the two largest Mediterranean islands. However, the taxonomic situation is somewhat more controversial and debated.

According to the latest edition of the "Flora d'Italia" (Pignatti et al. 2017), Sardinia and Sicily are home to three endemic oaks. These are *Q. ichnusae* Mossa, Bacch. & Brullo for Sardinia, and *Q. leptobalanos* Guss. and *Q. gussonei* (Borzi) Brullo for Sicily. In addition to these three species, both islands are home to other oak species belonging to the pubescent white oak group [i.e. *Q. amplifolia* Guss., *Q. congesta* C.Presl, *Q. dalechampii* Ten., *Q. virgiliana* (Ten.) Ten. (Brullo et al. 1999, 2008; Bacchetta et al. 2004, 2009), which were originally described for Sicily (the first two) and southern Italy (the last two) and subsequently recognized also for Sardinia, the Italian Peninsula, and some areas of southern and eastern Europe (especially *Q. dalechampii* and *Q. virgiliana*)]. However, there is currently no taxonomic agreement regarding the taxonomic consistency of these species, either in Italy or across Europe. The recently published checklists of the Italian vascular flora (Bartolucci et al. 2018, 2024) have synonymized *Q. leptobalanos* and *Q. ichnusae* (along with the aforementioned four critical white oak species) with *Q. pubescens* subsp. *pubescens* and *Q. gussonei* with *Q. cerris*. The decision regarding *Q. ichnusae* and *Q. leptobalanos* and the other doubtful taxa (all deciduous species belonging to the subgen. *Quercus*, sect. *Quercus*) was based on taxonomic and genetic studies published in the last decade (Di Pietro et al. 2016, 2020a, 2020b, 2021; Fortini et al. 2022). These studies revealed that the extensive taxonomic and nomenclatural diversification proposed for the *Q. pubescens* complex in central and southern Italy is not supported by corresponding differentiation at either the morphological or molecular level. However, while studies have been published on the white oaks of Sicily in which genetic data of individuals and populations were coupled with leaf and acorn morphological analysis (Di Pietro et al. 2020a, 2021), only genetic data have been published on the Sardinian oaks so far (Di Pietro et al. 2021). For this reason, the present study analysed Sardinian populations of white oaks from a genetic and morphological point of view. These populations belong to the taxa reported by Pignatti et al. (2017) as *Quercus congesta*, *Q. virgiliana*, and *Q. ichnusae*.

The present study aimed to achieve the following objectives:

- (1) To confirm or disprove the previously obtained results on the genetic differences of Sardinian white oaks by considering a new set of populations,
- (2) to determine whether genetic differentiation aligns with morphometric diversity traits within the different taxa, which could support their taxonomic autonomy at the species level,
- (3) to assess which communities are most at risk in relation to land use and ongoing climate change based on actual and potential vegetation data from Sardinia's white oak forests.

To make the results representative on a larger scale, a more consistent analysis framework was created. This framework involved genetic comparisons of the Sardinian populations with populations of other white oak taxa from the Italian Peninsula and southern Europe.

Materials and methods

An outline of the forest vegetation in Sardinia

The island of Sardinia is home to several forest communities which can be dominated by both evergreen and deciduous trees. A comprehensive overview of the island's forest vegetation was provided by Bacchetta et al. (2009, 2010). According to these authors, the most widespread evergreen forest types in Sardinia are *Quercus ilex* and *Q. suber* woods, which act as potential vegetation (for the concept of "potential vegetation" as mature and stable vegetation see

Tüxen 1956 and Loidi et al. 2010) for most of the island territory (Fig. 1a), in the form of several phytosociological associations. Rare and scattered are the stands dominated by *Quercus coccifera* or by natural coniferous forests (Calvia et al. 2022). Regarding the deciduous broadleaf forests, the majority of the communities are oak woods dominated by *Quercus pubescens* s.l. These pubescent oak woods exhibit a different floristic composition primarily linked to the phytoclimatic belt (Canu et al. 2015) and the bedrock type occupied. The two associations *Loncomelo pyrenaici-Quercetum ichnusae* and *Glechomo sardoae-Quercetum congestae* occur on non-carbonate soils (e.g. basalts, andesite, and granite) and are widespread within a large bioclimatic belt ranging from meso-Mediterranean to the supra-temperate thermotypes (600–1700 m). In contrast, the *Lonicero implexae-Quercetum virgiliana* woods are found mainly on limestone substrates and are widespread at significantly lower altitudes (0–500 m) within the thermo-Mediterranean and meso-Mediterranean thermotypes. The potential and actual distribution of the deciduous oak forest communities are shown in Fig. 1b and c, respectively.

Study area and plant material

The study was focussed on six Sardinian populations of the collective group of *Quercus pubescens*, including two stands of *Q. congesta*, two of *Q. ichnusae*, and two of *Q. virgiliana*. *Quercus dalechampii* was not considered in this study because no Sardinian forest communities dominated or strongly characterized by this species are known in the phytosociological or forestry literature. The well-known

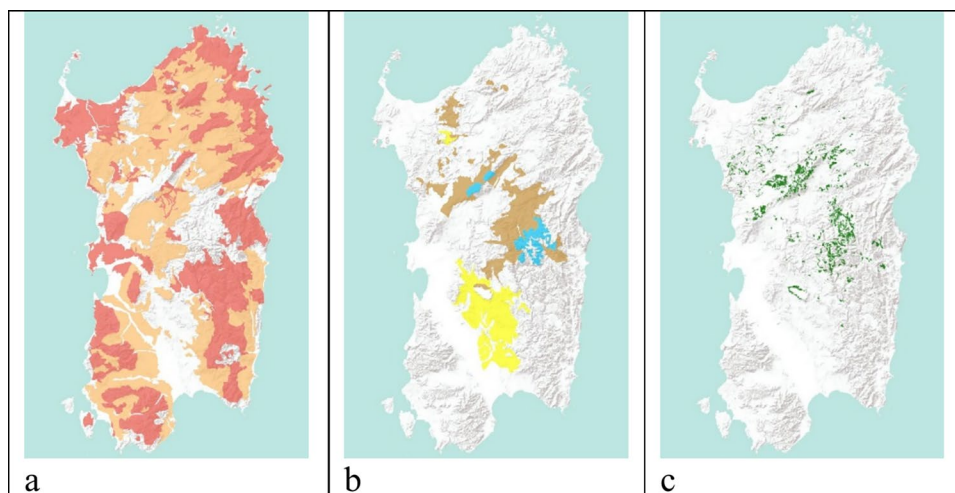


Fig. 1 **a** Potential distribution of evergreen sclerophyllous oak forests in Sardinia Island: *Q. ilex* (red), *Q. suber* (orange). **b** Potential distribution of deciduous oak forests in Sardinia based on their current taxonomical classification in Bacchetta et al. (2009) and Pignatti et al. (2017), namely *Q. virgiliana* (yellow), *Q. ichnusae* (brown), *Q. con-*

gesta (light blue). **c** Current distribution of deciduous oak forests in Sardinia. Data for arranging Fig. 1a and b refer to a 1:350.000 map published in Bacchetta et al. (2009), while those for Fig. 1c refer to a 1:50.000 scale map published in Camarda et al. (2011)

identification problems within the *Q. pubescens* complex and the taxonomical ambiguity of *Q. dalechampii* without clear diagnostic morphological traits (see Di Pietro et al. 2012) further justify this omission. An attempt to collect samples of the presumed *Q. dalechampii* in white oak communities where this species is reported as a low-frequency companion species mixed with other “critical” white oak species would not have ensured reliable identification. The same (even amplified) goes for *Q. amplifolia*, whose presence in Sardinian forests is reported as sporadic (see Fig. 2; Online Resource 1). In contrast, an extensive literature is available on forest communities dominated by *Q. congesta*, *Q. ichnusae*, and *Q. virgiliana* (Bacchetta et al. 2004).

In addition to the Sardinian populations of white oaks, five populations of *Q. pubescens* Willd. subsp. *pubescens* from northern, central, and southern Italy, as well as Spain and Croatia, were taken into consideration to carry out a genetic comparison with the three Sardinian putative pubescent oak taxa investigated. To further explore the genetic differences within the *Q. pubescens* collective group, the sample design also included one population of *Q. frainetto* Ten., one of *Q. petraea* (Matt.) Liebl. subsp. *petraea*, and one of *Q. robur* L. subsp. *robur*, all selected from pure,

unmixed sites to provide a clear reference for a comparison in terms of “macrospecies”.

The collection sites of the *Q. pubescens* collective group were chosen based on a biogeographic/taxonomical criterion (Table 1; Fig. 3). Specifically, one population of *Q. ichnusae* was collected in the *locus classicus* of this taxon [Bau Nou, near Senis, OR; (Mossa et al. 1999)], while the stands of *Q. congesta* and *Q. virgiliana* were selected based on data from phytosociological relevés available for Sardinia (Bacchetta et al. 2004). The five populations of *Q. pubescens* Willd. subsp. *pubescens* were chosen to enable comparisons not only across the three Italian biogeographic sectors (northern, central, and southern Italy), but also with the Iberian Peninsula, formerly contiguous with Sardinia prior to the opening of the western Mediterranean, and with the Balkans, which belong to biogeographical units other than Sardinia and could therefore represent distinct gene pools within the *Q. pubescens* collective group.

A total of 279 adult trees from 14 populations were sampled (Table 1). The oak specimens collected were identified provisionally in the field, at the time of sampling, based on the information gathered from taxonomic, floristic and phytosociological scientific papers published for those areas (Brullo et al. 1999; Mossa et al. 1999; Pignatti et al. 2017)

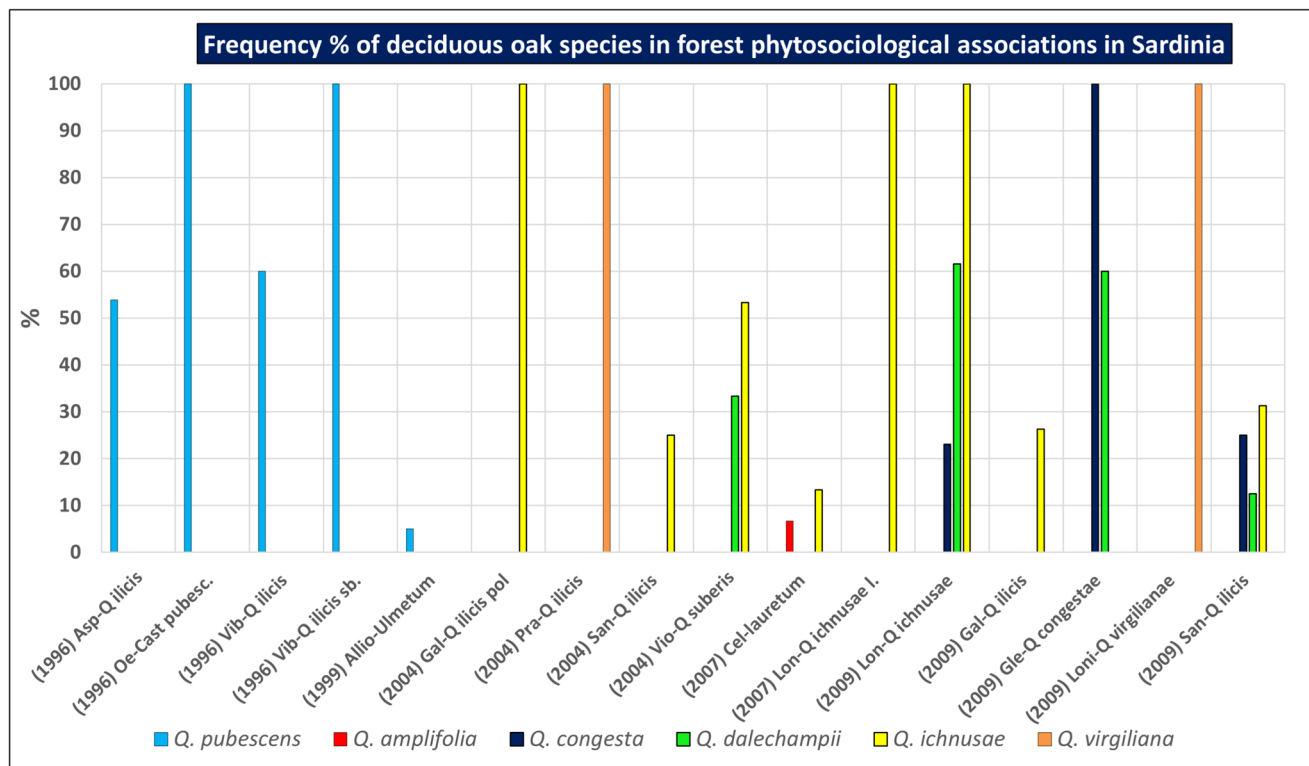


Fig. 2 Relative frequency of the single deciduous white oak species occurring in the phytosociological tables of all the forest communities (associations) comprising deciduous oak species in the tree layer, as identified and published for Sardinia, is shown. The forest commu-

nities are ordered from left to right according to their year of publication from 1996 to 2009. The complete names of the forest associations are reported in Online Resource 1

Table 1 Populations of *Quercus* taxa analysed in this study at nuclear microsatellite markers. Taxa with asterisk (*) belong to the *Q. pubescens* collective group. Lat: latitude; Lon: longitude; Alt: altitude (m a. s. l)

ID population	Taxon	Country	Administrative region	Locality	No. of samples	Lat	Lon	Alt
IT22	<i>Q. pubescens</i> subsp. <i>pubescens</i> *	Italy	Lombardy	Ome	20	45.4314	10.6282	76
IT64	<i>Q. pubescens</i> subsp. <i>pubescens</i> *	Italy	Lazio	Aurunci Mountains	20	41.3588	13.5272	535
IT89	<i>Q. pubescens</i> subsp. <i>pubescens</i> *	Italy	Basilicata	Catarozzolo	20	40.0555	16.0886	392
ES45	<i>Q. pubescens</i> subsp. <i>pubescens</i> *	Spain	Navarre	Zilbeti	20	42.9519	−1.4778	780
HR05	<i>Q. pubescens</i> subsp. <i>pubescens</i> *	Croatia	Split-Dalmatia County	Muč	20	43.7025	16.4253	470
IT52	<i>Q. congesta</i> *	Italy	Sardinia	Fonni (Muggiana)	20	40.1096	9.249	1060
IT53	<i>Q. congesta</i> *	Italy	Sardinia	Fonni (Govossai)	20	40.1292	9.3024	978
IT48	<i>Q. ichnusae</i> *	Italy	Sardinia	Senis	20	39.8159	8.9486	227
IT50	<i>Q. ichnusae</i> *	Italy	Sardinia	Monte Traessu	20	40.4686	8.6722	580
IT49	<i>Q. virgiliana</i> *	Italy	Sardinia	Pau	19	39.7871	8.7892	395
IT51	<i>Q. virgiliana</i> *	Italy	Sardinia	Ittiri	20	40.6138	8.5034	120
IT24	<i>Q. frainetto</i>	Italy	Lazio	Selvapiana Wood (Amaseno)	20	41.4862	13.3106	90
AT02	<i>Q. petraea</i> subsp. <i>petraea</i>	Austria	Kremstal	Pfaffenberg	20	48.2611	16.3056	374
IT13	<i>Q. robur</i> subsp. <i>robur</i>	Italy	Basilicata	Oasis Pantano Policoro	20	40.1741	16.6998	4

and the collectors' expert knowledge. The taxa considered in this study are named in accordance with IPNI (2025). Version 3.20 of the QGIS programme was used to prepare the maps (QGIS.org 2025).

Nuclear microsatellite marker analyses

The entire dataset of 14 populations was subjected to microsatellite (simple sequence repeat—SSR) marker analyses. Except for population IT49 (19 trees), 20 trees per population were analysed. Genomic DNA was extracted from dried leaf tissue (~20 mg per tree) using two commercial kits: the Qiagen DNeasy™ Plant Kit (Qiagen, Hilden, Germany) and the EchoLUTION Plant DNA Kit (BioEcho, Cologne, Germany), following the manufacturer's instructions. Samples were genotyped at 12 Expressed Sequence Tag-Simple Sequence Repeat (EST-SSR) loci, selected for their high polymorphism and previously used in studies of European white oaks (Durand et al. 2010; Guichoux et al. 2011; Antonecchia et al. 2015; Curtu et al. 2015). Primer pairs were combined into three multiplex PCR reactions: Mu1 (PIE239, PIE227, PIE223, PIE215), Mu2 (PIE020, PIE152, PIE243, PIE242), and Mu3 (PIE267, PIE102, PIE258, PIE271) (Online Resource 2). PCR amplifications were carried out under standardized conditions for all taxa and across the 12 markers (Di Pietro et al. 2021). A touchdown PCR programme was applied using a Biometra

Thermocycler TOptical Gradient 96 (Biometra, Göttingen, Germany). The program began with an initial denaturation at 95 °C for 15 min, followed by 10 cycles of denaturation at 94 °C for 1 min, annealing at 60 °C (−1 °C per cycle) for 1 min, and extension at 72 °C for 1 min. The second step consisted of 25 cycles of 1 min denaturation at 94 °C, 1 min annealing at 50 °C and 1 min extension at 72 °C, followed by a final extension step at 72 °C for 20 min. EST-SSR fragment sizes were determined via capillary electrophoresis using an ABI 3130XL Genetic Analyzer (Applied Biosystems, USA). The GeneScan 500 ROX internal size standard was used. Fragment analysis and genotyping were performed in GeneMapper v4.0 (Applied Biosystems). Automatically detected peaks were inspected for reliability and corrected manually when needed. To eliminate potential genotyping errors, uncertain results were reanalysed. Allele calling and binning were completed by visual inspection.

Data analysis

All the genetic analyses were performed exclusively on 10 microsatellite loci, since markers PIE239 and PIE258 did not amplify consistently, and the binning showed a significant overlap in bin boundaries.

SPAGeDI v.1.5.d (Hardy and Vekemans 2002) was used to assess, for each population, the mean number of alleles (N_a), the effective number of alleles (N_e), the allelic richness

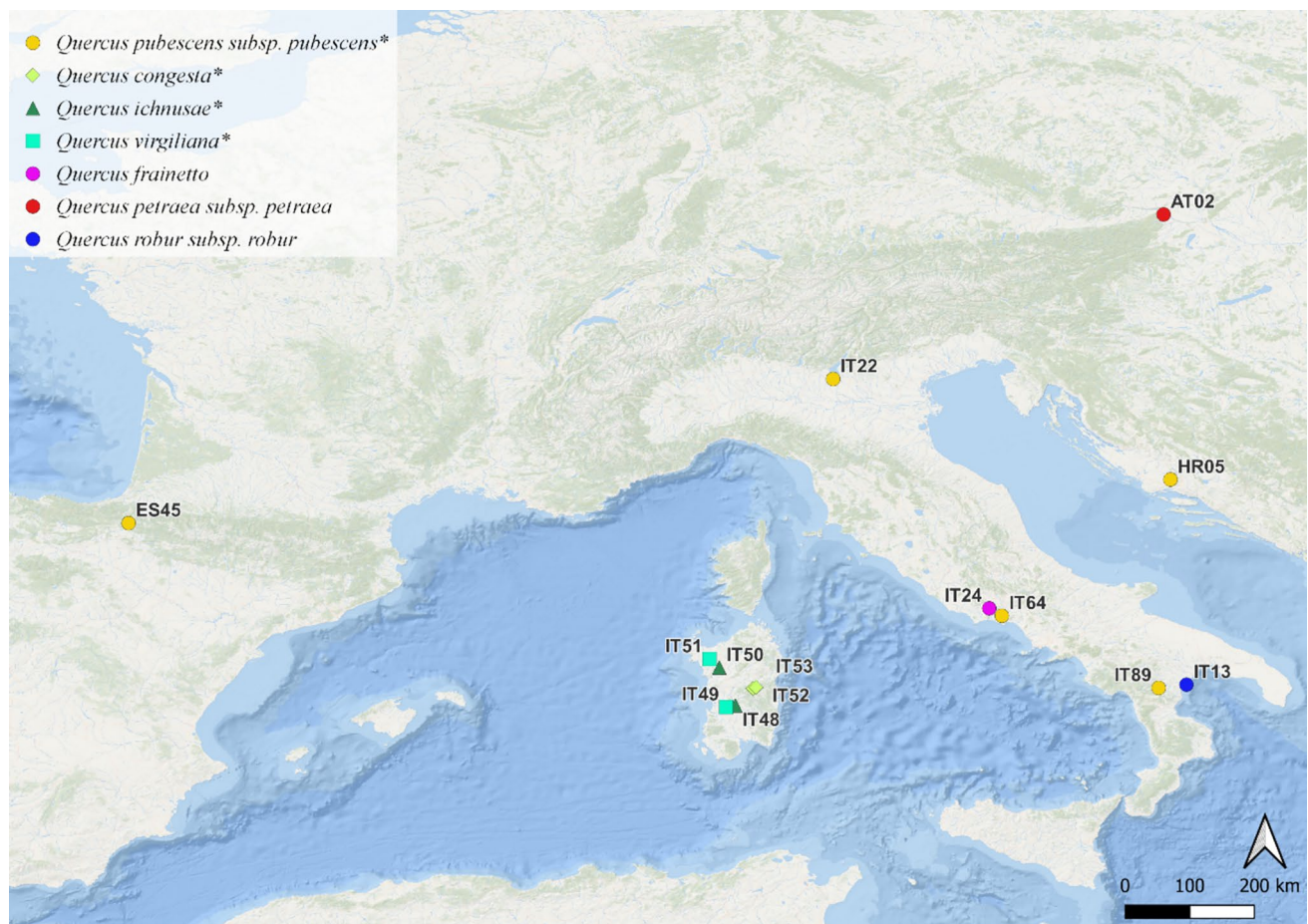


Fig. 3 Distribution map of the 14 populations of *Quercus* analysed at nuclear microsatellite markers. Taxa with an asterisk (*) in the legend belong to the *Q. pubescens* collective group

(A_r , after rarefaction to 30 allele copies), the observed (H_o) and expected (H_e) heterozygosity. Inbreeding coefficients (F_{IS}) were estimated for each population using INEST 2.3 (Chybicki and Burczyk 2009), which accounts for the presence of null alleles. Estimates were obtained using the Bayesian Individual Inbreeding Model (IIM). Markov Chain Monte Carlo (MCMC) simulations were run for 50,000 iterations with a burn-in of 10,000 iterations. For each population, two models were compared: a full model including null alleles, inbreeding, and genotyping failures (nfb), and a model including null alleles and genotyping failures (nb). Support for inbreeding was evaluated by comparing deviance information criterion (DIC) values between models. The pairwise F_{ST} values between populations, assessed following Weir and Cockerham (1984), and the private alleles (N_p) were estimated using the R package hierfstat (Goudet 2005). The significance of the F_{ST} indices was tested with 1000 permutations. To assess the partitioning of genetic variation, an analysis of molecular variance (AMOVA) was performed in GenAlEx v. 6.5 (Peakall and Smouse 2012). The analysis was conducted both for all 14 populations and

separately for the 11 populations of the *Q. pubescens* group, estimating variance components at two hierarchical levels: among populations and among individuals within populations (average of 20 individuals per population). The Bayesian clustering algorithm implemented in the software STRUCTURE 2.3.4 (Pritchard et al. 2000) was used to analyse the population genetic structure of the 14 populations. The admixture model with correlated allele frequencies without prior information on the geographical location of samples or the taxonomical classification was selected. For every K value between 1 and 8 (number of taxa included in the study plus one) 20 replications were run, with a burn-in length of 50,000 and a run length of 100,000 iterations.

The optimal number of clusters (K) was determined using StructureSelector (Li and Liu 2018), following the Puechmaillie (2016) method, which accounts for uneven sampling. A bar plot of STRUCTURE assignment results at $K=5$ was realized in R (version 4.3.3; R Core Team 2025), employing the tidyverse (Wickham et al. 2019) package and ggplot2 (Wickham 2016). The membership coefficient (Q) describes the proportion of assignment of the sample to each genetic

cluster, while the q_i defines proportion of membership of a population in a cluster. Samples with membership coefficients $Q \geq 0.90$ were considered as clearly belonging to a cluster. To determine the genetic relationships between the various populations of *Quercus* from a taxonomic and geographic perspective, a rooted tree was built using the UPGMA method (unweighted pair group method with arithmetic mean), based on Nei's minimum genetic distance (1973), implemented in the Populations software v. 1.2.32 (Langella 1999). Using the grouped population option, 1000 replicates over loci were used for bootstrapping. MEGA V11.0.13 (Tamura et al. 2021) was used to plot the resultant UPGMA tree.

A Principal Coordinate Analysis (PCoA) on the genetic distances among individuals was carried out using the `dist` function in the R package `poppr` (Kamvar et al. 2014). Furthermore, as a cross-check for the clustering pattern, a Discriminant Analysis of Principal Components (DAPC), using the `dapc` function in the R package `ade4` (Jombart and Ahmed 2011; Jombart and Collins 2015) was performed on individual tree-level data from a subset of 219 samples restricted to the *Q. pubescens* collective group.

Analyses of the leaf morphological traits

Leaf morphological traits were examined exclusively in the six Sardinian populations [(IT48, IT49, IT50, IT51, IT52, IT53); see Online Resource 3 for the variables considered]. For this analysis, nine leaves were sampled from each of the 119 trees, resulting in a total of 1071 leaves. The leaves were scanned using an Epson GT-15000 scanner at 300 dpi, with the abaxial side facing upward. Twelve morphological traits were measured on the scanned images following the protocol of Kremer et al. (2002). Dimensional variables, measured with ImageJ (Rasband 1997), included leaf area (Area), perimeter (Per), lamina length (LL), petiole length (PL), total lamina plus petiole length (LL+PL), lobe width (LW), sinus width (SW), and the height of the widest point from the leaf base (WP). Because leaf size traits often vary with environmental or seasonal conditions, several "derived traits" were also calculated, following Kremer et al. (2002) and Rellstab et al. (2016). These included lobe depth ratio (LDR), lobe width ratio (LWR), lamina shape or obversity (Ob), and petiole ratio (Pr). For each leaf trait, differences among populations were tested using the nonparametric Kruskal–Wallis test by means of the `kruskal.test` function. When the Kruskal–Wallis test indicated significant differences ($p < 0.05$), a post-hoc pairwise comparison using Dunn's test with Bonferroni correction was performed to identify significantly differing population pairs. Dunn's test was carried out using both the FSA (Ogle et al. 2025) and `rstatix` (Kassambara 2025) packages. For each trait, boxplots were generated using `ggplot2` (Wickham 2016) to display the

distribution of measurements across populations. Significant post-hoc results were annotated on the plots with `ggpubr`'s `stat_pvalue_manual` (Kassambara 2025). Additionally, Principal Component Analysis (PCA) was carried out on the correlation matrix of the "derived traits" (LDR, LWR, Ob, Pr), using the average value of the nine leaves for each tree (119 trees). PCA was conducted using the `prcomp` function in R. All graphical outputs were generated using `ggplot2` (Wickham 2016). These analyses were run in R (version 4.3.3; R Core Team 2025).

Results

Deciduous oak forests' distribution

From a taxonomic point of view, the graph in Fig. 2 shows that about until the end of the twentieth century, the populations of Sardinian pubescent oaks were all classified under the single binomial *Q. pubescens*. Since the 2000s, based on some previous influential studies (e.g. Brullo et al. 1999; Mossa et al. 1999), the binomial *Q. pubescens* s.l. was progressively excluded from the island and substituted by the other taxa of pubescent white oaks, the object of this study. Fig. 1a and 1b shows that, in terms of potential vegetation (climax) of the whole Sardinian territory, there is a clear dominance of sclerophyllous broadleaved evergreen forests (868,978 hectares) over deciduous forests (358,901 hectares), the latter being restricted to inland sectors nearby the main mountainous ranges. The comparison of the potential area for deciduous oak forests with their actual covered area (64,190.8 hectares), showed that the deciduous oak forests currently occupy only 17.9% of their potential area. The greatest disproportion between the potential area and the currently covered area concerns *Quercus virgiliana* (*Lonicero-Quercetum virgilianae*) forests. These are widespread within the meso-Mediterranean bioclimatic belt and have undergone a decrease of 97.9%. Conversely, the meso-temperate forest of *Quercus congesta* (*Glechomo-Quercetum congestae*) and *Q. ichnusae* (*Loncomelo-Quercetum ichnusae*) located in the submontane and lower-montane belts conserved a significantly greater part of their potential area (59 and 54%, respectively).

Results from the nuclear microsatellite marker analyses

Ten EST-SSR markers were used in the analysis. Private alleles (N_p) were identified in 12 of the 14 populations. Population-level genetic statistics indicated that the effective number of alleles (N_e) ranged from 3.62 in IT24 to 5.62 in IT50. Allelic richness (A_r) varied between 5.31 in IT13 and 7.47 in IT50. Observed heterozygosity

(H_o) ranged from 0.55 in IT24 to 0.80 in HR05, whereas expected heterozygosity (H_e) ranged from 0.63 in IT24 to 0.80 in ES45. The null allele corrected inbreeding coefficient (F_{IS}) ranged from 0.0093 to 0.0817 across populations (mean = 0.0420). On the basis of model comparison, inbreeding was supported only in population IT52 of *Q. congesta*, whereas no evidence for inbreeding was found in the other populations (Table 2).

Pairwise population differentiation (F_{ST}) revealed a strong genetic divergence among populations belonging to the four taxonomic macrogroups (*Q. frainetto*, *Q. petraea*, *Q. pubescens* collective group, and *Q. robur*), showing the highest differentiation (F_{ST} = 0.219) between the two populations of *Q. frainetto* (IT24) and *Q. robur* (IT13). Within the *Q. pubescens* collective group, the highest significant pairwise F_{ST} value was observed between the Spanish population ES45 and the Sardinian one IT49 (F_{ST} = 0.077), while the lowest (F_{ST} = 0.009) occurred between two Sardinian populations of *Q. congesta* IT52 and IT53 (Fig. 4).

The AMOVA showed that when all 14 populations are considered (including *Q. frainetto*, *Q. robur*, and *Q. petraea*) a moderate genetic differentiation among populations (15%) is observed, suggesting a limitation of gene flow probably related to effective boundaries between species. When only the populations belonging to the *Q. pubescens* group are considered, genetic differentiation among populations decreases to 7%, indicating an increased gene flow (Table 3).

According to the estimators implemented in StructureSelector (Li and Liu 2018) following Puechmaille (2016), the most likely number of genetic clusters is five ($K=5$) (Online Resources 4, 5).

In the bar plot of Fig. 5, the five clusters are depicted with distinct colours. The genetic clusters 1, 2, and 5 are mainly characterized by individuals from the populations of *Q. robur* (blue), *Q. petraea* (red), and *Q. frainetto* (magenta). In contrast, the gene pools associated with clusters 3 and 4 are distributed among the 11 populations belonging to the *Q. pubescens* group. Individual membership coefficient (Q) values are provided in Online Resource 6 and proportions of membership (q_i) for all 14 populations are detailed in Online Resource 7 and visually summarized in Fig. 6.

In the UPGMA tree constructed using Nei's minimum genetic distance (Fig. 7) the populations of *Q. robur* and *Q. frainetto* appeared clearly separated and distinct with a high bootstrap support (BS = 100%).

The *Q. petraea* population (AT02) was differentiated from the collective group of *Q. pubescens* with relatively moderate bootstrap support (BS = 62%) reflecting the known closer genetic relationship with *Q. pubescens* than with the other two white oak taxa mentioned above.

Within cluster D, which includes all the taxa belonging to the *Q. pubescens* collective group, some internal sub-clusters were highlighted, although these were based on low bootstrap values. Notably, two populations of *Q. pubescens* subsp. *pubescens* (IT22 and HR05) formed a distinct cluster separate from the other populations (BS = 20%).

The six Sardinian populations did not cluster strictly according to their taxonomic identity but showed affinity with the Spanish and mainly with the central and southern Italian populations of *Q. pubescens* subsp. *pubescens*. The two populations of *Q. congesta* (IT52 and IT53) consistently segregate together in the same sub-cluster (bootstrap 97%),

Table 2 Genetic diversity indices to the 14 populations of *Quercus* analysed. Taxa with asterisk (*) are the taxa belonging to the *Q. pubescens* collective group. N_p : private alleles; N_a : mean number of alleles; N_e : effective number of alleles (Nielsen et al. 2003); A_r : allelic

richness (rarefied to 30 gene copies); H_o : observed heterozygosity; H_e : expected heterozygosity; F_{IS} : fixation index. F_{IS} values significantly different from 0 and indicating inbreeding are shown in bold

ID population	Taxon	No. of samples	N_p	N_a	N_e	A_r ($K=30$)	H_o	H_e	F_{IS}
IT22	<i>Q. pubescens</i> subsp. <i>pubescens</i> *	20	4	7.7	5.07	7.13	0.67	0.74	0.0286
IT64	<i>Q. pubescens</i> subsp. <i>pubescens</i> *	20	2	7.5	4.46	6.94	0.62	0.70	0.0635
IT89	<i>Q. pubescens</i> subsp. <i>pubescens</i> *	20	2	6.5	4.39	6.17	0.66	0.72	0.0322
ES45	<i>Q. pubescens</i> subsp. <i>pubescens</i> *	20	1	7.6	5.59	7.23	0.78	0.80	0.0391
HR05	<i>Q. pubescens</i> subsp. <i>pubescens</i> *	20	6	7.3	4.73	6.77	0.80	0.74	0.0093
IT52	<i>Q. congesta</i> *	20	0	7.9	5.25	7.36	0.67	0.74	0.0817
IT53	<i>Q. congesta</i> *	20	6	7.5	5.24	7.02	0.75	0.74	0.0116
IT48	<i>Q. ichnusae</i> *	20	1	7.8	5.11	7.24	0.72	0.73	0.0383
IT50	<i>Q. ichnusae</i> *	20	1	8.0	5.62	7.47	0.71	0.74	0.0217
IT49	<i>Q. virgiliana</i> *	19	1	6.5	4.54	6.20	0.59	0.66	0.0650
IT51	<i>Q. virgiliana</i> *	20	0	6.9	5.02	6.50	0.69	0.74	0.0429
IT24	<i>Q. frainetto</i>	20	2	6.0	3.62	5.53	0.55	0.63	0.0561
AT02	<i>Q. petraea</i> subsp. <i>petraea</i>	20	7	7.6	3.97	6.88	0.75	0.72	0.0174
IT13	<i>Q. robur</i> subsp. <i>robur</i>	20	12	5.5	4.09	5.31	0.62	0.69	0.0781

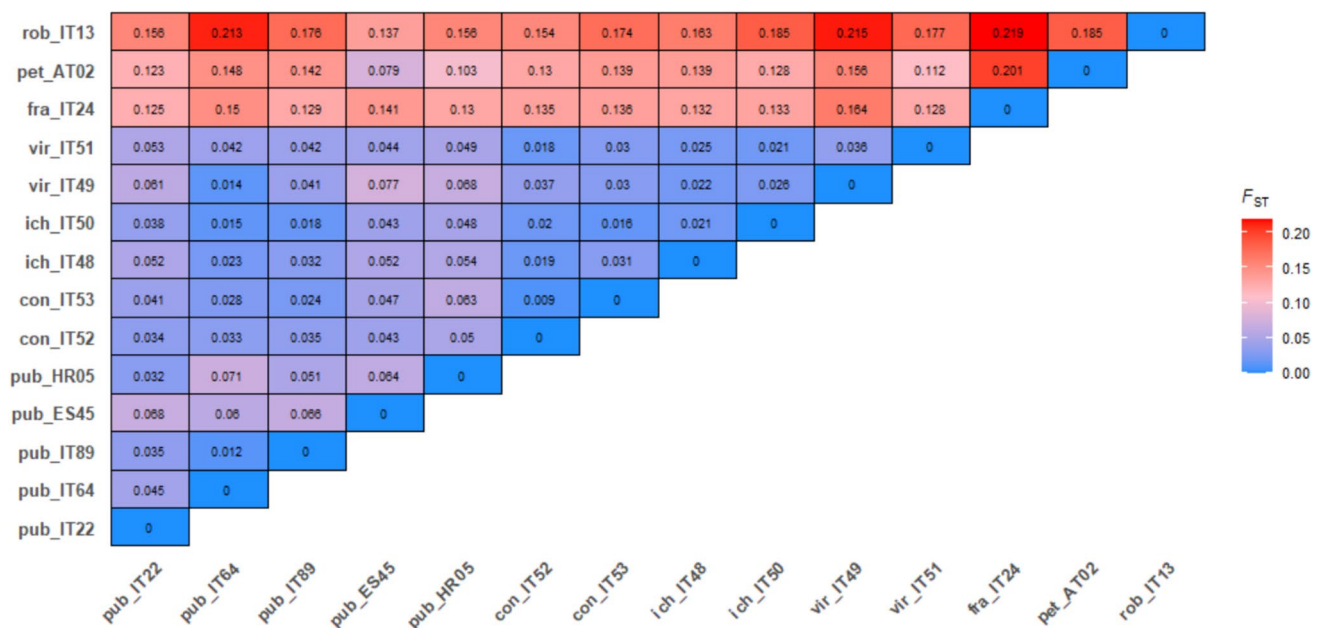


Fig. 4 The heatmap illustrates pairwise population differentiation (F_{ST}) estimates (Weir and Cockerham 1984) among 14 populations of the genus *Quercus*. The colours represent the degree of genetic divergence, with blue corresponding to low values and red to high values.

Population labels consist of a three-letter abbreviation of the taxon name (con: *Q. congesta*, fra: *Q. frainetto*, ich: *Q. ichnusae*, pet: *Q. petraea* subsp. *petraea*, pub: *Q. pubescens*, rob: *Q. robur*, vir: *Q. virgiliana*), followed by a four-character code identifying the population

Table 3 Analysis of molecular variance (AMOVA) showing the partitioning of genetic variation among and within populations for all 14 populations and for the 11 populations of the *Quercus pube-*

scens group. The degrees of freedom (df), sum of squares (SS), mean squares (MS), estimated variance components (Est. Var.), and percentage of variation (%) are presented for each source of variation

Dataset	Source of variation	df	SS	MS	Est. Var	%
14 populations (incl. <i>Q. robur</i> and <i>Q. frain.</i>)	Among populations	13	447.627	34.433	1.334	15
	Within populations	265	2077.900	7.841	7.841	85
	Total	278	2525.527		9.176	100
11 populations (<i>Q. pubescens</i> complex only)	Among populations	10	195.792	19.579	0.597	7
	Within populations	208	1599.863	7.692	7.692	93
	Total	218	1795.655		8.289	100

although the separation from the sub-cluster including most Italian and Sardinian pubescent oak populations is very low (bootstrap 1%). The individual populations of the other two pairs of Sardinian oaks, i.e. *Q. virgiliana* (IT49 and IT51) and *Q. ichnusae* (IT48 and IT50), segregate separately in different sub-clusters.

The PCoA1 vs. PCoA2 plot revealed a clear separation along both axes between samples from the three populations of *Q. frainetto*, *Q. petraea*, and *Q. robur* and those belonging to the *Q. pubescens* collective group. A similar pattern of separation was observed along the third axis in the PCoA1 vs. PCoA3 plot. PCoA plots also showed no clear separation among the six Sardinian populations, regardless of their taxonomic identity (Fig. 8). This pattern is particularly evident in the DAPC including only populations of the *Q. pubescens*

group, where the close genetic proximity between the two *Q. pubescens* subsp. *pubescens* populations, IT22 and HR05, was also evident (Fig. 9).

Results from the morphometric analyses

The summary results obtained for the complete set of morphological leaf trait measurements from the six Sardinian populations of *Q. congesta*, *Q. ichnusae*, and *Q. virgiliana* are shown in Online Resource 8.

The nonparametric Kruskal–Wallis test revealed significant differences among populations for several leaf morphological traits ($p < 0.05$), indicating variation in these characteristics across the studied populations. Post-hoc pairwise comparisons using Dunn's test with Bonferroni correction

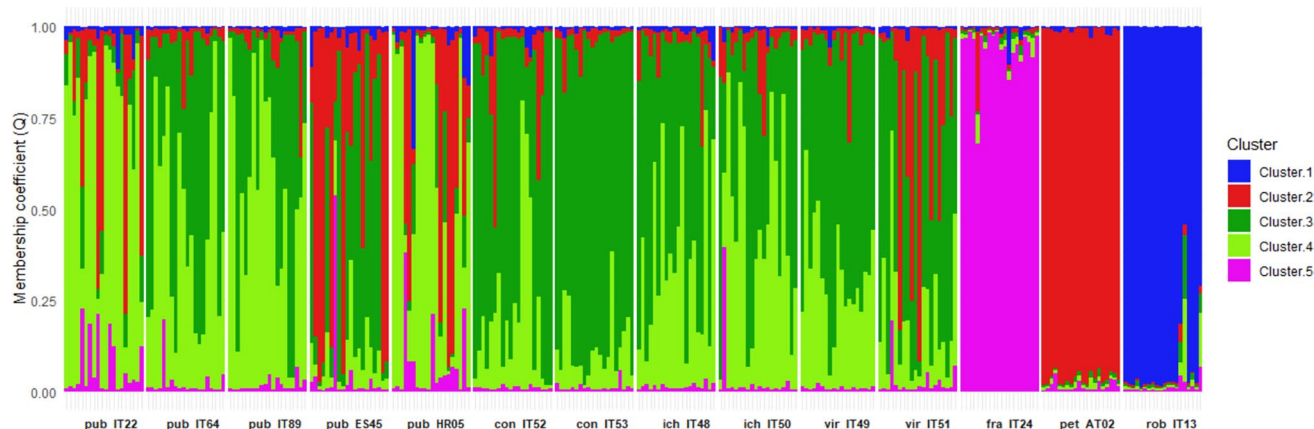


Fig. 5 Bar plot of STRUCTURE assignment results at $K=5$. Different colours represent each genetic cluster, while each vertical bar represents an individual's membership coefficients (Q) across clusters.

Populations are separated by white vertical lines. The names of the populations follow the same coding as in Fig. 4

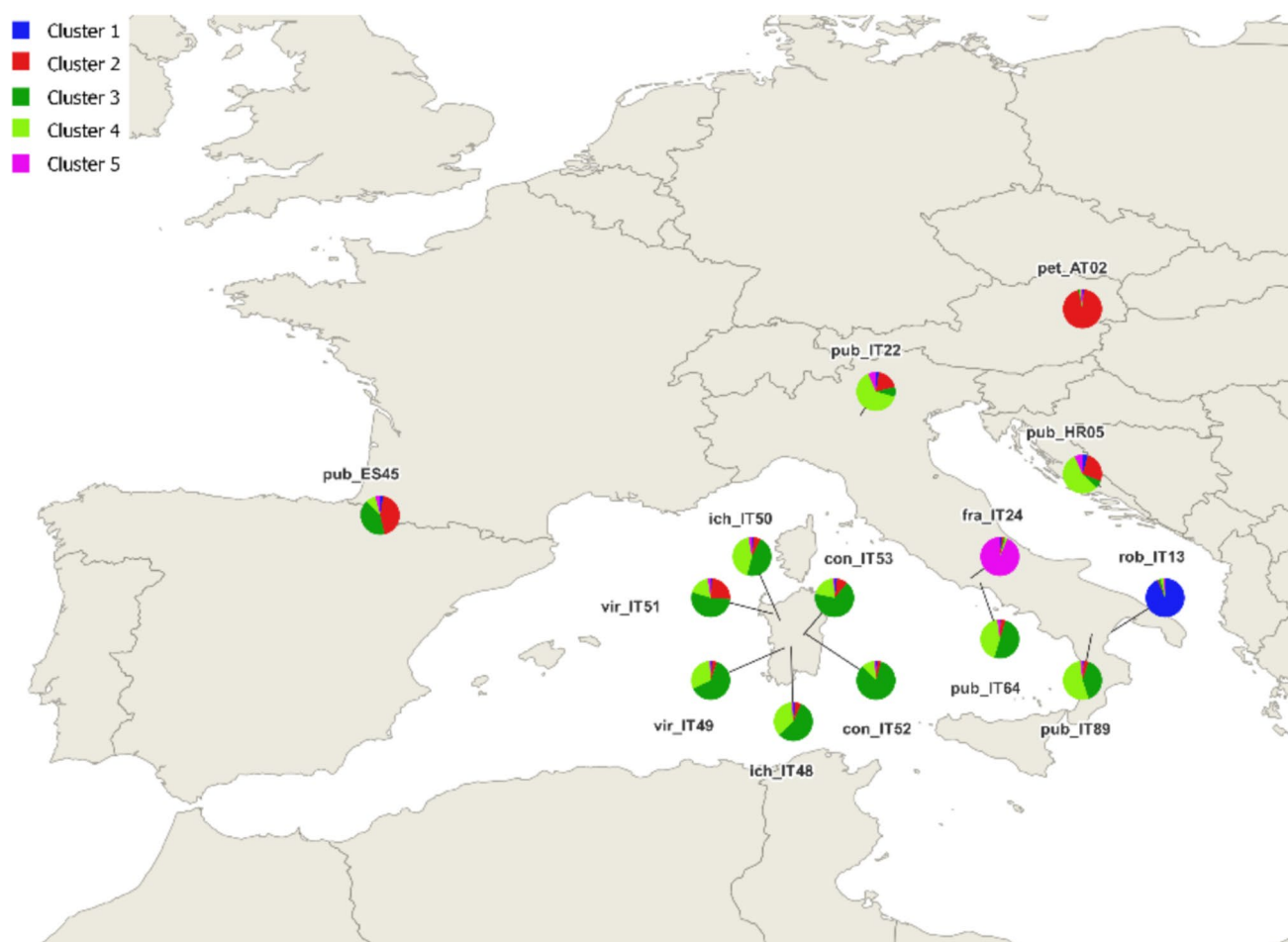


Fig. 6 Map showing the distribution of the 14 *Quercus* populations analysed based on 10 microsatellite loci, with results from the Bayesian cluster analysis at $K=5$. Each population is represented by a pie

chart, where the colours indicate the proportion of membership (q_i) values assigned to the five genetic clusters inferred by STRUCTURE. The names of the populations follow the same coding as in Fig. 4

Fig. 7 UPGMA tree constructed using Nei's minimum genetic distance (1973) among the 14 *Quercus* populations. The horizontal branches indicate evolutionary changes measured in genetic divergence (numbers below). Bootstrap support values (from 1000 replicates) are displayed beside the branches. Letters mark the nodes that distinguish the four principal taxonomic groups. The names of the populations follow the same coding as in Fig. 4

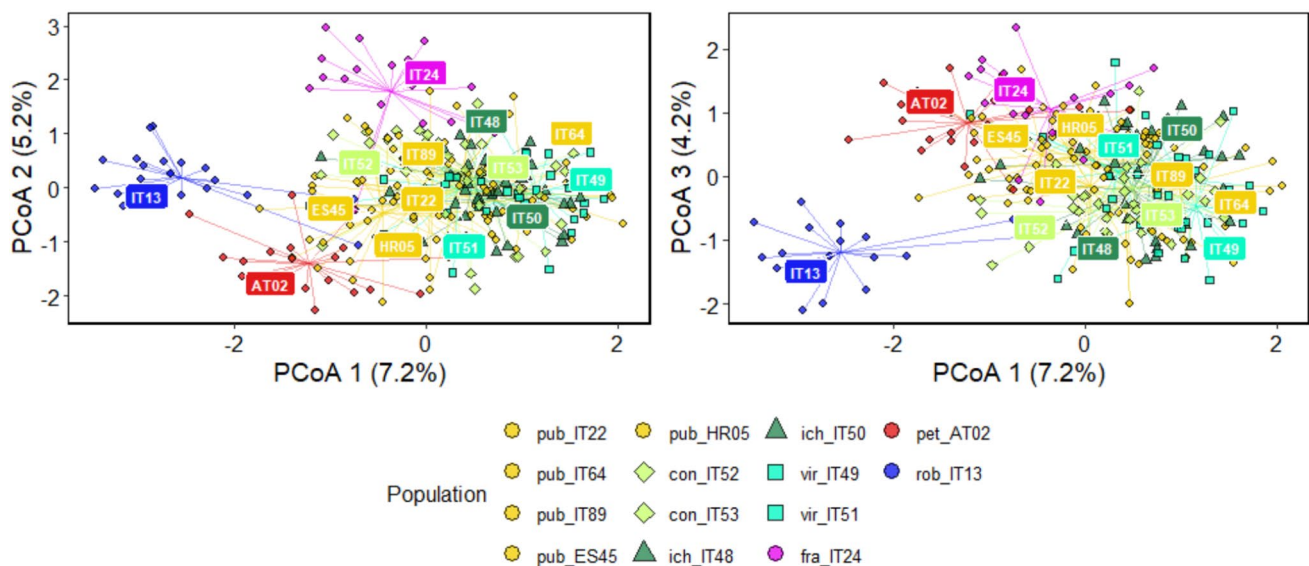
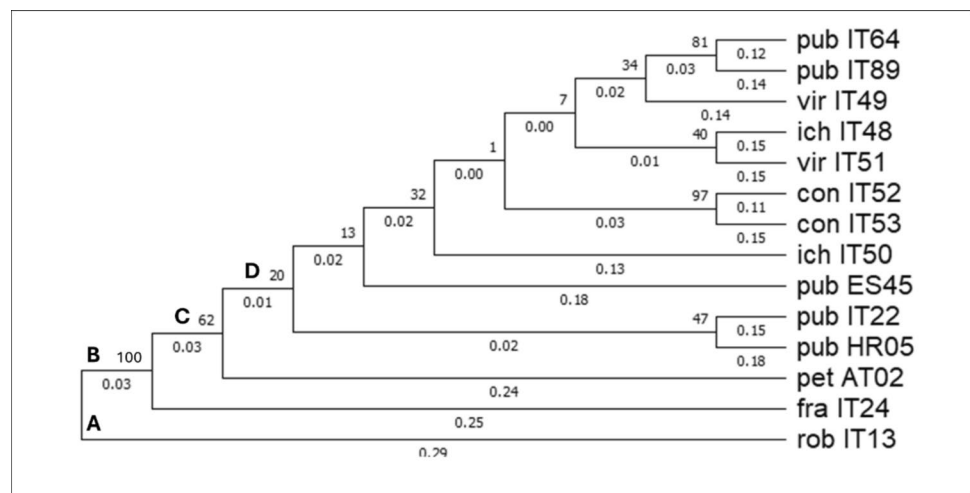


Fig. 8 Three-dimensional Principal Coordinate Analysis (PCoA) scatterplots of the 14 *Quercus* populations genotyped at 10 microsatellite loci. PCoA1 vs. PCoA2 (on the left) and PCoA1 vs. PCoA3 (on

the right). Axes 1, 2 and 3 explain 7.2%, 5.2% and 4.2% of the total variance, respectively

identified specific population pairs with statistically significant differences. These findings are displayed in the boxplots in Fig. 10.

The PC1 vs. PC2 and PC1 vs. PC3 scatterplots from the principal component analysis (PCA) on four “derived traits” of the leaf showed no clear clustering, indicating limited morphological differentiation among the six Sardinian populations and the three taxa analysed (Fig. 11 and Online Resource 9). In the PC1 vs. PC2 plot, the population IT50 of *Q. ichnusae* showed a slight separation along the first principal component, primarily due to the variation in lobe width ratio (LWR) and petiole length (Pr). While lobe depth ratio (LDR) and obversity (Ob) were more influential in distinguishing population IT49. The other populations

showed minimal separation influenced by the remaining variables. Similar separation patterns across populations were observed in the PC1 vs. PC3 scatterplot.

Discussion

This study aimed to explore genetic and morphological differences that could support the current taxonomic classification of white oak species in Sardinia. The Sardinian–Corsican block has been isolated in the central Mediterranean since ancient times. Such long-term geographic isolation promotes speciation and the conservation of relict species (Contandriopoulos 1962; Médail and Quézel

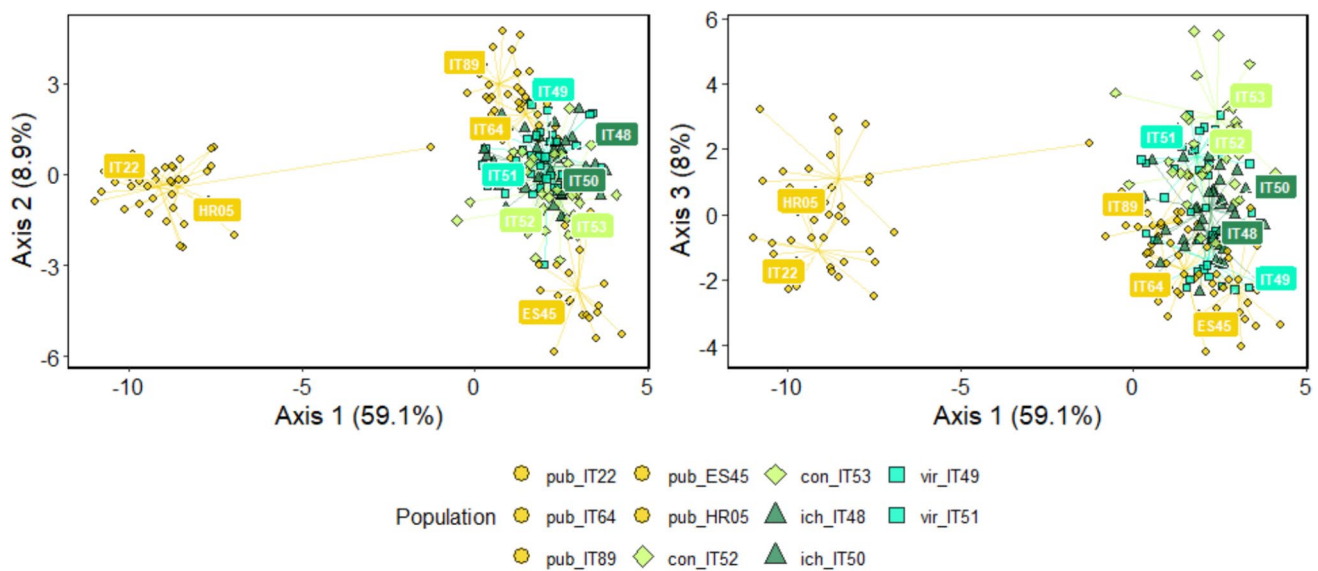


Fig. 9 Scatterplot of discriminant analysis of principal components (DAPC) displaying the distribution of 219 samples from 11 populations of the *Quercus pubescens* collective group. The analysis, based

on 10 microsatellite loci, is shown along the first three linear discriminants, which account for 59.1%, 8.9%, and 8.0% of the total variance, respectively

1997; Whittaker 1998; Quezél 1999, Tóth et al. 2019). Sardinia is an important biodiversity hotspot hosting 195 exclusive plant taxa (Fois et al. 2022) including several “paleoendemics” (i.e. relict endemic species with a long evolutionary history of isolation), such as the monophyletic genus *Morisia* (Viv.) Asch., *Rhamnus persicifolia* Moris, *Genista morisii* Colla, *Ribes sardous* Martelli, etc. (Arrigoni 2006).

Phylogeographic studies have revealed that Sardinian white oaks possess unique plastid haplotypes that are not found in other Mediterranean regions, and that Sardinia has a significantly higher number of endemic haplotypes than the Italian Peninsula and Sicily (Fineschi et al. 2002; Fortini et al. 2024; Di Pietro et al. 2024).

Phytosociological studies of forest vegetation on this island report communities dominated by various species of pubescent white oak (Bacchetta et al. 2009).

However, despite its tendency to conserve ancient flora and produce high levels of endemism, the combined genetic and morphometric analysis of Sardinian white oaks found no evidence of endemic *Quercus* species on the island.

The genetic parameters considered (private alleles, mean and effective number of alleles, allelic richness, observed and expected heterozygosity), reveal population-level genetic statistics that are comparable to those reported in previous studies on other Sardinian white oak populations (Di Pietro et al. 2021). Of the three pairs of Sardinian populations belonging to *Q. congesta*, *Q. ichnusae*, and *Q. virgiliana*, respectively, the latter shows slightly lower average values for all parameters.

Analysis of allelic variability across all the populations studied revealed that the *Q. robur* subsp. *robur* (IT13) population had the highest number of private alleles (12). Compared to the other oak populations considered (all belonging to the *Q. pubescens* collective group) this genetic distinctiveness reflects taxonomic distinctiveness and suggests limited interspecific gene flow. In contrast, five of the six Sardinian populations (IT48–IT52) belonging to the three presumed Sardinian pubescent oak taxa exhibited few (one or two) or no private alleles, indicating high gene flow between these populations and genetic similarity. This interpretation is further supported by the highest values of allelic richness (7.47) and effective number of alleles (5.62), which were both recorded in one Sardinian population (IT50–*Q. ichnusae*). Conversely, a relatively high number of private alleles (six) was detected in the *Q. congesta* population (IT53). This may be due to the topographic isolation of this population, which is located at a relatively high altitude (978 m a.s.l.) compared to the other Sardinian populations, potentially limiting gene exchange.

Genetic comparisons among all populations reveal substantial diversification into four main groups corresponding to the four “macrospecies”, *Q. frainetto*, *Q. petraea*, *Q. pubescens*, and *Q. robur*, as indicated by high pairwise F_{ST} values (Fig. 4). In contrast, variation within the *Q. pubescens* group is relatively low, and the analysed taxa do not reveal differentiation associated with their current taxonomic classification. Populations of the *Q. pubescens* group in particular exhibit a higher degree of genetic similarity based on their geographic distribution than on their taxonomic identity. This pattern is evident in the two Sardinian populations

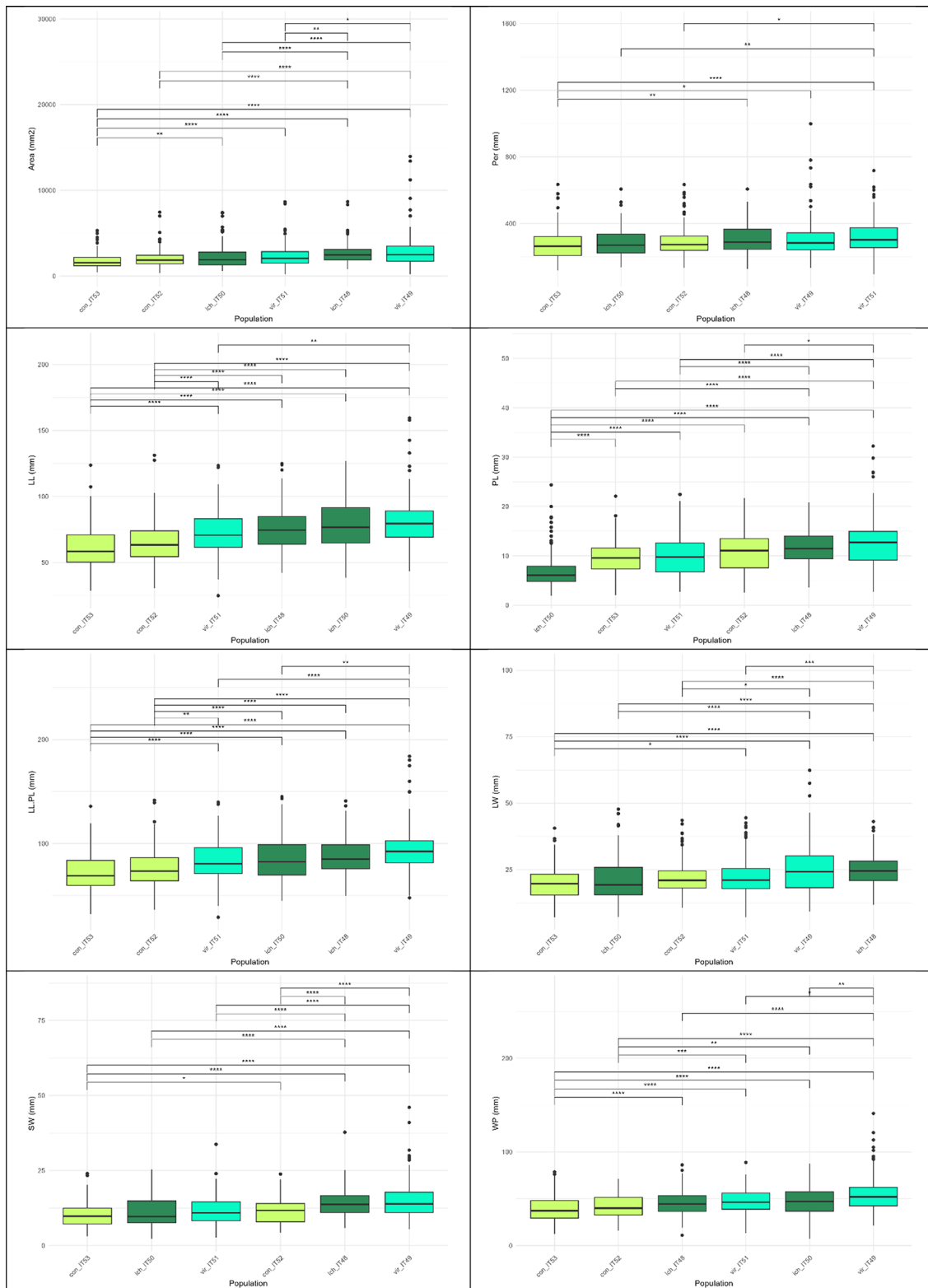


Fig. 10 Boxplots displaying the distribution of leaf morphological traits across six Sardinian populations of *Quercus pubescens* complex. Significant differences between pairs of populations, assessed with Bonferroni-corrected pairwise Dunn's tests after Kruskal-Wallis tests, are shown on the plots. Boxes represent the interquartile range (IQR), and the median is shown as the central line. Whiskers extend to 1.5 times the IQR, with outliers displayed as individual points. Colours denote different taxa (light green: *Q. congesta*; aquamarine: *Q.*

virgiliana; dark green: *Q. ichnusae*). The variables represented include: leaf area (Area), perimeter (Per), lamina length (LL), petiole length (PL), total lamina plus petiole length (LL+PL), lobe width (LW), sinus width (SW), the height of the widest point from the leaf base (WP), lobe depth ratio (LDR), lobe width ratio (LWR), lamina shape or obversity (Ob), and petiole ratio (Pr). All variables are described in Online Resource 3

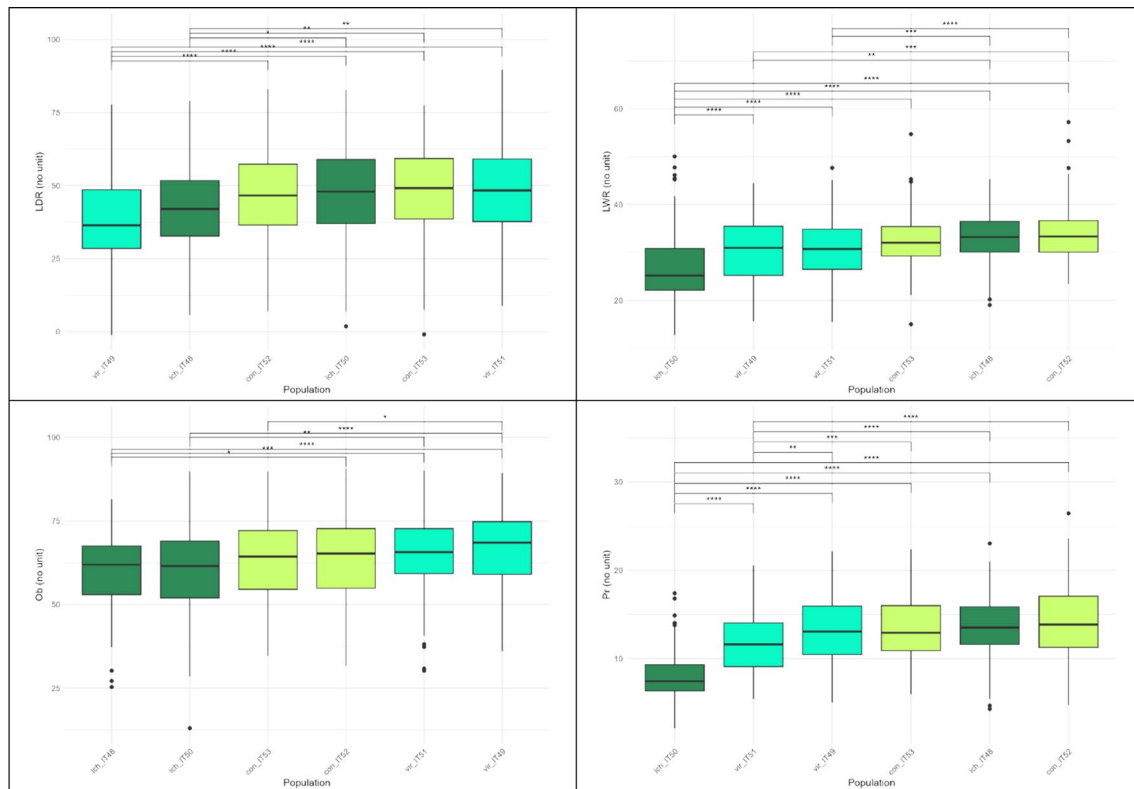


Fig. 10 (continued)

classified in the literature as *Q. congesta* (IT52 and IT53), which are only 9 km apart and exhibit minimal genetic differentiation ($F_{ST}=0.009$).

The two *Q. pubescens* populations from central and southern Italy (IT64 and IT89) showed a higher degree of similarity with the Sardinian populations than to the other *Q. pubescens* population collected in the Italian Peninsula (IT22). The latter, from northern Italy, shows its greatest genetic similarity with the Croatian population HR05 ($F_{ST}=0.032$). These genetic relationships are consistent with the biogeographic connections between Sardinia and the Tyrrhenian coastal sector of the Italian Peninsula (Rivas-Martínez et al. 2004).

The clear separation between the white oak “macro-species” *Q. frainetto*, *Q. petraea*, *Q. pubescens*, and *Q. robur* is evident in the STRUCTURE assignment bar plot (Fig. 5). However, according to the Puechmaile (2016) method, the optimal number of clusters is five rather than four. In addition to the genetic clusters associated with *Q. frainetto*, *Q. petraea*, and *Q. robur*, the pubescent oak group is divided into two main genetic clusters (clusters 3 and 4), which are depicted in the bar plot by two different shades of green. While this suggests a genetic separation within the *Q. pubescens* group, it is not useful for discriminating between different taxa of Sardinian

white oaks. Nevertheless, this result is interesting from a large-scale biogeographical perspective. It shows a separation between populations in Sardinia, central and southern Italy, and Spain that express a gene pool linked to cluster 3, and populations in northern Italy and the Dinaric Alps that express a gene pool linked to cluster 4. The frequency of cluster 4 decreases progressively towards central and southern Italy and Spain, where a southwestern gene pool, which is dominant within the Sardinian populations, prevails.

The overrepresentation of *Q. pubescens* s.l. populations in the dataset could explain the observed internal genetic differentiation (two distinct gene pools). This pattern was confirmed by a genetic structure analysis conducted exclusively on *Q. pubescens* populations (see Online Resource 10), as well as by examining the values of Nei’s genetic distances. The latter confirmed the presence of two main gene pools within the *Q. pubescens* complex. However, this study could not determine the exact position of the boundary line between these two gene pools. More extensive sampling and broader geographic coverage are required to fully characterize this genetic transition.

As previous studies (Leroy et al. 2017; Fortini et al. 2022; Šijačić-Nikolić et al. 2023) have highlighted, this study also showed that the *Q. pubescens* s.l. group is more likely to

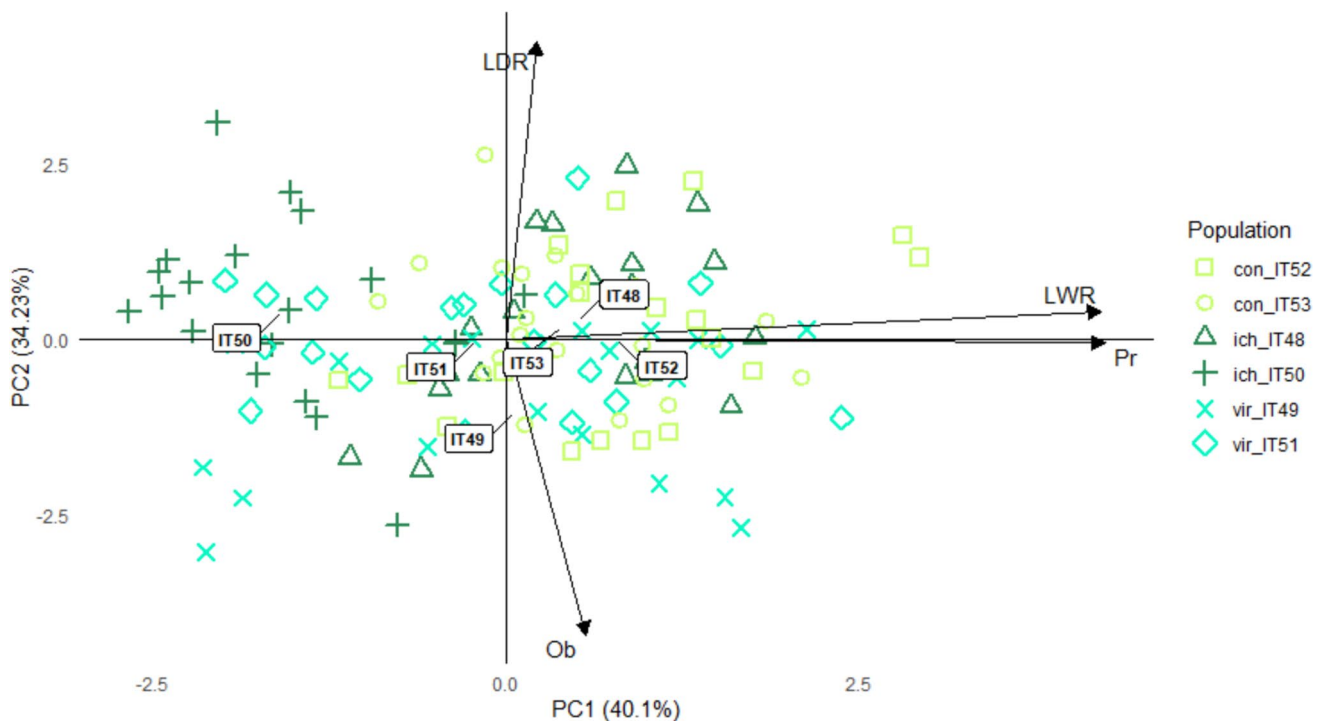


Fig. 11 PCA biplot of 119 *Quercus* samples from six Sardinian populations based on four leaf morphological traits (LDR: lobe depth ratio; LWR: lobe width ratio; Ob: obversity; Pr: petiole ratio). Shown are principal component axes 1 and 2 (PC1 vs. PC2), explain-

ing 74.33% of total morphological variation. Colours denote different taxa. The names of the populations follow the same coding as in Fig. 4

hybridize with *Q. petraea* than with *Q. robur* or *Q. frainetto*. Consequently, populations of *Q. pubescens* found in areas where *Q. petraea* is prevalent (cluster 2) exhibit a higher proportion of *Q. petraea* than other oak species. An apparent exception to this rule is the contribution of the *Q. petraea* genome to one Sardinian *Q. virgiliana* population. This contribution from *Q. petraea* is surprising for a number of reasons. Firstly, *Q. petraea* is absent from Sardinia. Secondly, the pubescent oak identified as *Q. virgiliana* is considered the most thermophilic of the Italian white oaks (Brullo et al. 1999; Bacchetta et al. 2004; Pignatti et al. 2017) making it the least likely to form mixed oak communities with a mesophilic species such as *Q. petraea*.

However, the genetic contribution of *Q. petraea* to a pubescent white oak population in northern Sardinia would be more understandable if it were considered evidence of the ancient presence of *Q. petraea* on the island. This presence would have occurred during periods when the island was connected to the continent via Corsica or the Tuscan Archipelago, i.e. during stages of sea regression coinciding with Quaternary glaciations. Indeed, *Q. petraea* is currently present on Corsica (Jeanmonod and Gamisans 2013) and is also common in the thermophilic oak forests of western Tuscany and adjacent north-western Lazio (Arrigoni et al. 1990; Scoppola and Filesi 1995; Di Pietro et al. 2010).

The clusters identified in the genetic structure analysis were closely matched by the UPGMA tree. In this tree the *Q. frainetto* and *Q. robur* populations were clearly separated from the others with complete bootstrap support. The *Q. petraea* cluster was also clearly separated from the pubescent oak populations but with a lower bootstrap support than *Q. robur* and *Q. frainetto*. This reflects a closer genetic relationship with *Q. pubescens*.

The PCoA scatterplot corroborates the UPGMA classification results, showing that the pubescent oak populations form a single macro-cluster. However, when the analysis is limited to pubescent oak populations only, and a DAPC is applied to maximize genetic variance among populations, those from northern Italy and the Western Dinaric Alps appear to be separated from the others. This is consistent with the STRUCTURE analysis results and the UPGMA tree. In contrast, the Sardinian populations do not form a distinct cluster, instead overlapping with all the other pubescent oak populations.

This wide genetic overlap among the Sardinian pubescent oaks, regardless of their taxonomic identity, is also evident in the morphometric analysis. Boxplots showing the range of variation for the most discriminating leaf morphological traits (Fig. 10) reveal no differences between *Q. congesta*, *Q. ichnusae*, and *Q. virgiliana* populations.

However, the IT50 population (*Q. ichnusae* from Mount Traessu) differs slightly from the other Sardinian populations in terms of two morphological traits related to the petiole (PL and Pr). This is also evident in Fig. 11 in the scatterplot of the principal component analysis (PCA) performed on the correlation matrix of the “derived leaf traits” (LDR, LWR, Ob, Pr).

As this variation pattern was not observed in the other *Q. ichnusae* population (IT53), which was the one collected in the *locus classicus*, it would be incorrect and misleading to consider petiole morphological traits as taxonomically diagnostic for *Q. ichnusae*. Nevertheless, this information could be useful to identify potential local morphotypes. The morphological diversity observed on the island, which has led some authors to propose the existence of different pubescent white oak taxa, is likely due to a combination of environmental factors and ecological conditions that nevertheless fall within the variability pattern of the same species.

The provisional finding that no endemic oaks occur in Sardinia or that the taxonomic diversity of white oaks is limited to a single species (i.e. *Q. pubescens*) rather than four should not result in these precious ecosystems receiving less attention or safeguarding. Despite an increase in total forest cover in Sardinia in recent decades (Puddu et al. 2012), the available data suggest that the current extent of deciduous oak forests in Sardinia is significantly lower than their potential. The most thermophilic communities, which are dominated by *Q. virgiliana*, currently cover only 2% of their potential distribution area. Regardless of whether *Q. virgiliana* is considered an established taxon, it plays an important role as an umbrella species for a diverse and high-fidelity flora in the underlying tree, scrub and undergrowth layers of the oak forest. Consequently, protecting it safeguards the entire community (*Lonicero implexae-Quercetum virgilianae*), which acts as a core area in the ecological network for numerous nemoral species that are closely linked to the microclimatic conditions of deciduous oak forests.

Even though detailed data on the temporal trends of deciduous forests are unavailable for Sardinia, it cannot be ruled out that thermophilic oak forests have been subject to differential deforestation and conversion into agricultural land in recent decades. Ongoing climate change, characterized by increased temperatures and summer droughts, will certainly not favour deciduous forests, especially given the overwhelming presence of surrounding sclerophyllous evergreen forests and Mediterranean maquis shrublands. The latter always benefit from short-term coppicing and subsequent soil erosion. As Blasi and Di Pietro (1998) reported for the coastal mountains of central Italy, one consequence of “unsustainable” *Q. pubescens* thermophilic forest management is the fragmentation of the original forest stands, followed by “sclerophyllization” and a progressive increase

in *Quercus ilex* and/or *Q. suber* as well as shrub species of Mediterranean maquis.

Conclusion

In this study, we analysed six Sardinian populations of pubescent white oak (belonging to the species *Q. congesta*, *Q. ichnusae*, and *Q. virgiliana*) from a genetic and morphological perspective. We compared these populations to other *Q. pubescens* populations and well-established white oak species from southern Europe. Our analysis of genetic and morphological variability in the six Sardinian populations did not reveal any taxon-specific traits. Instead, it revealed a significant degree of overlap between the three analysed taxa. Furthermore, the genetic findings did not allow the three Sardinian taxa (including the endemic *Q. ichnusae*) to be distinguished from other populations currently assigned to *Q. pubescens* from the Italian Peninsula, Spain, and Croatia. However, from a phylogeographic perspective, the genetic variability among the eleven populations of *Q. pubescens* s.l. considered in this study suggests a biogeographic structure characterized by two partially distinct gene pools. These were found to be related to their geographic distribution rather than to their taxonomic identity. Despite this internal diversity, however, all *Q. pubescens* s.l. populations tend to cluster together and are clearly separated from the *Q. frainetto*, *Q. petraea*, and *Q. robur* populations included for comparison in the study. This supports the hypothesis of a single pubescent white oak taxon for Sardinia capable of giving rise to forest types very different from each other in bioclimatic and ecological terms. The most thermophilic type, currently identified as *Lonicero-Quercetum virgilianae* is relatively rare and requires conservation strategies to be developed, although all Sardinian deciduous forests are to be considered at risk in this context of global warming. It would therefore be desirable that management plans prioritize the protection of all representative deciduous oak forests across the island. This would allow the conservation of the prevailing Sardinian gene pool within the *Quercus* genus, and to safeguard the floristic, ecological, and biogeographical uniqueness of white oak forests in their individual ecosystems. Accordingly, habitat fragmentation and land-use change should be minimized to preserve connectivity among forest stands. Long-term monitoring would help detect potential floristic and genetic diversity erosion under climate change scenarios.

Information on Electronic Supplementary Material

Online Resource 1. Phytosociological table showing the oak forest associations in Sardinia Island.

Online Resource 2. Observed size range, motif types, primer sequences, and dye type of the microsatellite loci.

Online Resource 3. Table of the morphological leaf traits.

Online Resource 4. Summary of StructureSelector results following the method of Puechmaille (2016). Median and mean estimates (MedMed, MedMean, MaxMed, MaxMean) across 20 replicates for each K are shown.

Online Resource 5. Summary of StructureSelector (Li and Liu 2018) results following the method of Puechmaille (2016).

Online Resource 6. Individual membership coefficient (Q) values of the 279 samples of *Quercus*.

Online Resource 7. Proportion of membership (q_i) to the six clusters of the 14 populations of *Quercus*.

Online Resource 8. Summary results obtained for the complete set of leaf variable measurements of the 6 Sardinian populations of *Quercus*.

Online Resource 9. PCA biplot of 119 *Quercus* samples from six Sardinian populations based on four leaf morphological traits.

Online Resource 10. Bar plot of STRUCTURE assignment results for $K=2$ and $K=3$.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00606-026-02009-9>.

Acknowledgements The authors wish to acknowledge anonymous referees, the handling Editor and the Editor-in-Chief for their valuable feedback on the earlier version of the manuscript. We express our gratitude to Alexandra Dolynska for her contribution to the genetic analyses at the laboratory of the “Department of Forest Genetics and Forest Tree Breeding” of the University of Göttingen (Göttingen, Germany). Thanks also to J. Loidi, W. Z. Skvorc, and W. Willner for their contributions as plant material collectors. A special thanks in memory of Professors Pier Virgilio Arrigoni and Edoardo Biondi, who passed away during the last year and who intensely studied the Sardinian flora and vegetation.

Author contribution R.D.P. and P.F. conceived the study; R.D.P., P.F., E.P. wrote the first draft of the paper; R.D.P., P.F., E.P., E.F., L.R. (together with the colleagues mentioned in the acknowledgements) collected the plant material; E.P., together with K.B.B. and O.G., carried out the genetic analysis while E.P. made morphometric measurements. E.P., P.D.M., P.F., and R.D.P. carried out the statistical analysis; all authors discussed the results and finalized the paper; all authors have read and agreed to the submitted version of the manuscript.

Funding Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement. Ministero dell’Università e della Ricerca. This study was supported by (a) MUR PRIN 2022, project code: 2022B3HA72; (b) the National Recovery and Resilience Plan (NRRP).

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are

included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aitken SN, Bemmels JB (2015) Time to get moving: assisted gene flow of forest trees. *Evol Appl* 9:271–290. <https://doi.org/10.1111/eva.12293>
- Antonecchia G, Fortini P, Lepais O, Gerber S, Legér P, Scippa GS, Viscosi V (2015) Genetic structure of a natural oak community in central Italy: evidence of gene flow between three sympatric white oak species (*Quercus*, Fagaceae). *Ann Forest Res* 58:205–216. <https://doi.org/10.15287/afr.2015.415>
- Arrigoni PV, Mazzanti A, Ricceri C (1990) Contributo alla conoscenza dei boschi della Maremma grossetana. *Webbia* 44:121–150
- Arrigoni PV (2006) Flora dell’Isola di Sardegna. Vol. I, II, III. Società Botanica Italiana. Carlo Delfino editore, Sassari
- Bacchetta G, Biondi E, Farris E, Filigheddu R, Mossa L (2004) A phytosociological study of the deciduous oak woods of Sardinia (Italy). *Fitosociologia* 41:53–64
- Bacchetta G, Bagella S, Biondi E, Farris E, Filigheddu RS, Mossa L (2009) Vegetazione forestale e serie di vegetazione della Sardegna. *Soc Ital Fitosociol Fitosociologia* 46(Suppl):1
- Bacchetta G, Bagella S, Biondi E, Farris E, Filigheddu R, Mossa L (2010) Le Serie di Vegetazione della regione Sardegna. In: Blasi C (ed) *La Vegetazione d’Italia*. Palombi & Partner S.r.l., Roma, pp 471–495
- Bartolucci F, Peruzzi L, Galasso G, Albano A, Alessandrini A, Ardenghi NMG, Astuti G, Bacchetta G, Ballelli S, Banfi E, Barberis G, Bernardo L, Bouvet D, Bovio M, Cecchi L, Di Pietro R, Domina G, Fascetti S, Fenu G, Festi F, Foggi B, Gallo L, Gottschlich G, Gubellini L, Iamónico D, Iberite M, Jiménez-Mejías P, Lattanzi E, Marchetti D, Martinetto E, Masin RR, Medagli P, Passalacqua NG, Peccenini S, Pennesi R, Pierini B, Poldini L, Prosser F, Raimondo FM, Roma-Marzio F, Rosati L, Santangelo A, Scoppola A, Scortegagna S, Selvaggi A, Selvi F, Soldano A, Stinca A, Wagensommer RP, Wilhalm T, Conti F (2018) An updated checklist of the vascular flora native to Italy. *Pl Biosyst* 152:179–303. <https://doi.org/10.1080/11263504.2017.1419996>
- Bartolucci F, Peruzzi L, Galasso G, Alessandrini A, Ardenghi NMG, Bacchetta G, Banfi E, Barberis G, Bernardo L, Bouvet D, Bovio M, Calvia G, Castello M, Cecchi L, Del Guacchio E, Domina G, Fascetti S, Gallo L, Gottschlich G, Guarino R, Gubellini L, Hofmann N, Iberite M, Jiménez-Mejías P, Longo D, Marchetti D, Martini F, Masin RR, Medagli P, Peccenini S, Prosser F, Roma-Marzio F, Rosati L, Santangelo A, Scoppola A, Selvaggi A, Selvi F, Soldano A, Stinca A, Wagensommer RP, Wilhalm T, Conti F (2024) A second update to the checklist of the vascular flora native to Italy. *Pl Biosyst* 158:219–296. <https://doi.org/10.1080/11263504.2024.2320126>
- Blasi C, Di Pietro R (1998) Two new phytosociological types of *Quercus pubescens* s.l. woodland community in southern Latium. *Pl Biosyst* 132:207–223. <https://doi.org/10.1080/11263504.1998.10654205>
- Brullo S, Guarino R, Siracusa G (1999) Revisione tassonomica delle querce caducifoglie della Sicilia. *Webbia* 54:1–72. <https://doi.org/10.1080/00837792.1999.10670670>
- Brullo S, Gianguzzi L, La Mantia A, Siracusa G (2008) La classe *Quercetea ilicis* in Sicilia. *Bull Accad Gioenia Sci Nat* 41:1–80

- Calvia G, Bonari G, Angiolini C, Farris E, Fenu G, Bacchetta G (2022) Classification of the Sardinian pine woodlands. *Mediterr Bot* 43:e72699. <https://doi.org/10.5209/mbot.72699>
- Camarda I, Carta L, Laureti L, Angelini P, Brunu A, Brundu G (2011) Carta della Natura della Regione Sardegna: Carta degli habitat alla scala 1:50.000. ISPRA
- Canu S, Rosati L, Fiori M, Motroni A, Filigheddu R, Farris E (2015) Bioclimate map of Sardinia (Italy). *J Maps* 11:1–8. <https://doi.org/10.1080/17445647.2014.988187>
- Chybicki JJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. *J Hered* 100:106–113. <https://doi.org/10.1093/jhered/esn088>
- Contandriopoulos J (1962) Recherches sur la flore endémique de la Corse et sur ses origines. Impr. Louis-Jean, Gap
- Curtu AL, Craciunesc I, Enescu CM, Vidalis A, Sofletea N (2015) Fine-scale spatial genetic structure in a multi-oak-species (*Quercus* spp.) forest. *iForest* 8:324–332. <https://doi.org/10.3832/for1150-007>
- Di Pietro R, Azzella M, Facioni L (2010) The forest vegetation of the Tolfa-Ceriti mountains (northern Latium central Italy). *Hacquetia* 9:91–150. <https://doi.org/10.2478/v10028-010-0002-2>
- Di Pietro R, Viscosi V, Peruzzi L, Fortini P (2012) A review of the application of the name *Quercus dalechampii*. *Taxon* 61:1311–1316. <https://doi.org/10.1002/tax.616012>
- Di Pietro R, Di Marzio P, Medagli P, Misano G, Silletti GN, Wagensommer RP, Fortini P (2016) Evidence from multivariate morphometric study of the *Quercus pubescens* complex in southeast Italy. *Bot Serbica* 40:83–90. <https://doi.org/10.5281/zenodo.48865>
- Di Pietro R, Conte AL, Di Marzio P, Gianguzzi L, Spampinato G, Caldarella O, Fortini P (2020a) A multivariate morphometric analysis of diagnostic traits in southern Italy and Sicily pubescent oaks. *Folia Geobot* 55:163–183. <https://doi.org/10.1007/s12224-020-09378-0>
- Di Pietro R, Conte AL, Di Marzio P, Fortini P, Farris E, Gianguzzi L, Müller M, Rosati L, Spampinato G, Gailing O (2021) Does the genetic diversity among pubescent white oaks in southern Italy, Sicily and Sardinia islands support the current taxonomic classification? *Eur J Forest Res* 140:355–371. <https://doi.org/10.1007/s10342-020-01334-z>
- Di Pietro R, Quaranta L, Mattioni C, Simeone MC, Di Marzio P, Proietti E, Fortini P (2024) Chloroplast haplotype diversity in the white oak populations of the Italian Peninsula, Sicily, and Sardinia. *Forests* 15:864. <https://doi.org/10.3390/f15050864>
- Di Pietro R, Di Marzio P, Antonecchia G, Conte AL, Fortini P (2020b) Preliminary characterization of the *Quercus pubescens* complex in southern Italy using molecular markers. *Acta Bot Croat* 79:15–25
- Durand J, Bodénès C, Chancerel E, Frigerio J-M, Vendramin G, Sebastiani F, Buonamici A, Gailing O, Koelewijn H-P, Villani F, Mattioni C, Cherubini M, Goicoechea PG, Herrán A, Ikarán Z, Cabané C, Ueno S, Alberto F, Dumoulin P-Y, Guichoux E, de Daruvar A, Kremer A, Plomion C (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>
- Fineschi S, Taurichini D, Grossoni P, Petit RJ, Vendramin GG (2002) Chloroplast DNA variation of white oaks in Italy. *Forest Ecol Managem* 156:103–114. [https://doi.org/10.1016/S0378-1127\(01\)00637-5](https://doi.org/10.1016/S0378-1127(01)00637-5)
- Fois M, Farris E, Calvia G, Campus G, Fenu G, Porceddu M, Bacchetta G (2022) The endemic vascular flora of Sardinia: a dynamic checklist with an overview of biogeography and conservation status. *Plants* 11:601. <https://doi.org/10.3390/plants11050601>
- Fortini P, Di Marzio P, Conte AL, Antonecchia G, Proietti E, Di Pietro R (2022) Morphological and molecular results from a geographical transect focusing on *Quercus pubescens*/Q. *virgiliana* ecological-altitudinal vicariance in peninsular Italy. *PL Biosyst* 156:1498–1511. <https://doi.org/10.1080/11263504.2022.2131923>
- Fortini P, Di Pietro R, Proietti E, Cardoni S, Quaranta L, Simeone MC (2024) 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 Forest Res* 143:107–127. <https://doi.org/10.1007/s10342-023-01610-8>
- Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molec Ecol Notes* 5:184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Guichoux E, Lagache L, Wagner S, Chaumeil P, Lèger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ (2011) Current trends in microsatellite genotyping. *Molec Ecol Resources* 11:591–611. <https://doi.org/10.1111/j.1755-0998.2011.03014.x>
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molec Ecol Notes* 2:618–620. <https://doi.org/10.1046/j.1471-8286.2002.00305.x>
- IPNI (2025). International Plant Names Index. The Royal Botanic Gardens, Kew, Harvard University Herbaria & Libraries and Australian National Herbarium. Available at: <http://www.ipni.org>. Accessed 15 Sep 2025
- Jeanmonod D, Gamisans J (2013) Flora Corsica, 2e édition. Bull Soc Bot Centre-Ouest 39:1–1074
- Jombart T, Ahmed I (2011) Adegenet 1.3–1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27:3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>
- Jombart T, Collins C (2015) A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0. Available at: <https://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf>
- Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281. <https://doi.org/10.7717/peerj.281>
- Kassambara A (2025) ggpubr: ‘ggplot2’ based publication ready plots. R package version 0.6.1. Available at: <https://rpkgs.datanovia.com/ggpubr/>
- Kremer A, Dupouey JL, Deans JD, Cottrell J, Csaikl U, Finkeldey R, Espinel S, Jensen J, Kleinschmit J, Van Dam B (2002) Leaf morphological differentiation between *Quercus robur* and *Quercus petraea* is stable across Western European mixed oak stands. *Ann Forest Sci* 59:777–787. <https://doi.org/10.1051/forest:2002065>
- La Greca M (1980) Il contributo degli Ortoteri e dei Mantodei alla conoscenza della biogeografia di Sardegna. *Biogeographia* 8:557–575. <https://doi.org/10.21426/B68110206>
- Langella O (1999) Populations 1.2.32: a population genetic software. CNRS UPR9034. Available at: <https://mybiosoftware.com/populations-population-genetic-software.html>. Accessed 10 Sep 2025
- Leroy T, Roux C, Villate L, Bodénès C, Romiguier J, Paiva JAP, Dossat C, Aury J-M, Plomion C, Kremer A (2017) Extensive recent secondary contacts between four European white oak species. *New Phytol* 214:865–878. <https://doi.org/10.1111/nph.14413>
- Li YL, Liu JX (2018) StructureSelector: a web based software to select and visualize the optimal number of clusters using multiple methods. *Molec Ecol Resources* 18:176–177. <https://doi.org/10.1111/1755-0998.12719>
- Loidi J, Del Arco M, Pérez De Paz PL, Asensi A, Díez Garretas B, Costa M, Díaz González T, Fernández-González F, Izco J, Penas Á, Rivas-Martínez S, Sánchez-Mata D (2010) Understanding properly the ‘potential natural vegetation’ concept. *J Biogeogr* 37:2209–2211. <https://doi.org/10.1111/j.1365-2699.2010.02302.x>
- Médail F (2017) The specific vulnerability of plant biodiversity and vegetation on Mediterranean islands in the face of global change. *Regional Environm Change* 17:1775–1790. <https://doi.org/10.1007/s10113-017-1123-7>

- Médail F, Quézel P (1997) Hot-spots analysis for conservation of plant biodiversity in the Mediterranean basin. *Ann Missouri Bot Gard* 84:112–127. <https://doi.org/10.2307/2399957>
- Médail F, Monnet A, Pavon D, Nikolic T, Dimopoulos P, Bacchetta G, Arroyo J, Barina Z, Albassatneh MC, Domina G, Fady B, Mat-evski V, Mifsud S, Leriche S (2019) What is a tree in the Mediterranean Basin hotspot? A critical analysis. *Forest Ecosyst* 6:17. <https://doi.org/10.1186/s40663-019-0170-6>
- Médail F, Myers N (2004) Mediterranean Basin. In: Mittermeier RA, Robles Gil P (eds) Hotspots revisited: earth's biologically richest and most endangered terrestrial ecoregions, CEMEX (Monterrey), Conservation International, Washington and Agrupación Sierra Madre, pp 144–147
- Mossa L, Bacchetta G, Brullo S (1999) *Quercus ichnusae* (Fagaceae), a new species from Sardinia. *Israel J Pl Sci* 47:199–207. <https://doi.org/10.1080/07929978.1999.10676774>
- Myers N (2003) Biodiversity hotspots revisited. *Bioscience* 53:916–917
- Nei M (1973) The theory and estimation of the genetic distance. In: Morton NE (ed) Genetic structure of 1099 populations. University of Hawaii Press, Honolulu, pp 45–54
- Nielsen R, Tarpy DR, Reeve HK (2003) Estimating effective paternity number in social insects and the effective number of alleles in a population. *Molec Ecol* 12:3157–3164. <https://doi.org/10.1046/j.1365-294X.2003.01994.x>
- Ogle DH, Doll JC, Wheeler AP, Dinno A (2025) FSA: simple fisheries stock assessment methods. Available at: <https://cran.r-project.org/web/packages/FSA/index.html>, Accessed 3 Aug 2025. <https://doi.org/10.32614/CRAN.package.FSA>
- 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>
- Pignatti S, Guarino R, La Rosa M (2017) Flora d'Italia, 2nd edn. Edagricole, Bologna
- 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>
- Puddu G, Falcucci A, Maiorano L (2012) Forest changes over a century in Sardinia: implications for conservation in a Mediterranean hotspot. *Agroforest Syst* 85:319–330. <https://doi.org/10.1007/s10457-011-9443-y>
- Puechmaille SJ (2016) The program structure does not reliably recover the correct population structure when sampling is uneven: sub-sampling and new estimators alleviate the problem. *Molec Ecol Resources* 16:608–627. <https://doi.org/10.1111/1755-0998.12512>
- QGIS.org (2025) QGIS geographic information system. Open Source geospatial foundation project. Available at: <https://qgis.org/>
- Quézel P (1999) Les grandes structures de végétation en région méditerranéenne: Facteurs déterminants dans leur mise en place post-glaciaire. *Geobios* 32:19–32. [https://doi.org/10.1016/S0016-6995\(99\)80081-3](https://doi.org/10.1016/S0016-6995(99)80081-3)
- R Core Team (2025) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Rasband WS (1997) ImageJ. US National Institutes of Health, Bethesda
- Reilstab C, Bühler A, Graf R, Folly C, Gugerli F (2016) Using joint multivariate analyses of leaf morphology and molecular-genetic markers for taxon identification in three hybridizing European white oak species (*Quercus* spp.). *Ann Forest Sci* 73:669–679. <https://doi.org/10.1007/s13595-016-0552-7>
- Rivas-Martínez S, Penas A, Díaz TE (2004) Bioclimatic and biogeographic maps of Europe. University of Leon, Leon
- Scoppola A, Filesi L (1995) I boschi della Riserva Naturale Regionale di Monte Rufeno (VT). *Ann Bot (Roma)* 51(1993, Suppl 10): 241–277
- Šijačić-Nikolić M, Kerkez Janković I, Jovanović M, Milovanović J, Aleksić JM (2023) Genetic diversity and genetic structure of three sympatric oak species in Serbian landscape of outstanding features “Kosmaj” assessed by nuclear microsatellites. *South-East Eur Forest* 14:117–127. <https://doi.org/10.15177/see-for.23-17>
- Takhtajan AL (1986) Floristic region of the world. University of California Press, London
- Tamura K, Stecher G, Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. *Molec Biol Evol* 38:3022–3027. <https://doi.org/10.1093/molbev/msab120>
- Thompson JD (2020) Plant evolution in the Mediterranean: insights for conservation, 2nd edn. Oxford University Press, Oxford
- Tóth EG, Tremblay F, Housset JM, Bergeron Y, Carcaillet C (2019) Geographic isolation and climatic variability contribute to genetic differentiation in fragmented populations of the long-lived subalpine conifer *Pinus cembra* L. in the western Alps. *BMC Evol Biol* 19:190. <https://doi.org/10.1186/s12862-019-1510-4>
- Tüxen R (1956) Die heutige potentielle natürliche Vegetation als Gegenstand der Vegetationskartierung. *Angew Pflanzensoziol* 13:5–55
- Weir BS, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370. <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x>
- Whittaker RJ (1998) Island biogeography. Ecology, evolution, and conservation. Oxford University Press, Oxford
- Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Golemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019) Welcome to the tidyverse. *J Open Source Softw* 4:1686. <https://doi.org/10.21105/joss.01686>
- Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer-Verlag, New York. Available at: <https://ggplot2.tidyverse.org>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.