Preliminary characterization of the *Quercus pubescens* complex in southern Italy using molecular markers

Romeo Di Pietro[1], Piera Di Marzio[2], Gaby Antonecchia[3], Antonio Luca Conte[2], Paola Fortini[2]*

[1] Department of Planning, Design, Architecture technology (PDTA), Herbarium Flaminio, Sapienza University of Rome, Rome, Italy

[2] Department of Bioscience and Territory, University of Molise, Pesche (IS), Italy

[3] Ministry of Education, University and Research (MIUR), Rome, Italy

Abstract – *Quercus pubescens* s.l. is a group of taxonomically intricate and highly debated deciduous white oaks widely distributed in southern Europe. The Apulia region occupies the south-easternmost part of the Italian Peninsula; the land-use pattern is based on extensive agricultural systems and only 10% is covered by forests that are mainly composed of oak woods. It is the region in Italy showing the highest number of oak species, among which four putative species of the *Quercus pubescens* group, have been reported in floras and checklists with uncertain taxonomic value because of the overlapping of diagnostic characters. In this paper, we carried out a molecular analysis on natural populations of *Q. pubescens* s.l. distributed throughout the Apulia region. Individuals from 24 pubescent oak populations were sampled and each tree was genotyped at 11 polymorphic microsatellite markers. Overall, the average expected heterozygosity (H_e) was 0.629, and the allelic richness (A_r) ranged between 2.130 and 7.187. No differentiation was observed among the populations investigated, and the genetic differentiation coefficient (F_{ST}) was 0.036. Gene flow among populations was found to be relatively high (Nm = 6.664). From a taxonomic point of view, the possibility of the coexistence of more than one species among the Apulian pubescent oaks reported in the taxonomic and syntaxonomic literature is not supported by the results of this molecular analysis.

Key words: genetic diversity; Italy; population structure; *Quercus pubescens*; SSR; taxonomy

Introduction

The genus *Quercus* L. subgenus *Quercus* has a wide distribution in the Northern Hemisphere, especially in central and southern Europe, where *Quercus* species form important forest communities in the Temperate and the Mediterranean bioclimatic regions (Nixon 1993, Govaerts and Frodin 1998).

Italy is the country showing the highest oak species diversity in Europe, although there is still incomplete agreement on the exact number of oak taxa occurring in the territory. Most of the taxonomic uncertainty regarding the Italian (and European) oaks concerns the so-called white oaks and in particular the pubescent oaks (*Quercus pubescens* s.l. subgen. *Quercus* Sect. *Quercus*). The most recent Italian flora and checklists (Pignatti et al. 2017, Bartolucci et al. 2018) report various pubescent oaks in southern Italy (*Q. amplifolia* Guss., *Q. apennina* Lam., *Q. congesta* C. Presl, *Q.*

dalechampii Ten., *Q. leptobalana* Guss., *Q. ichnusae* Mossa, Bacch. et Brullo, *Q. humilis* Mill., *Q. virgiliana* Ten.) which are all considered Steno-Mediterranean vicariant species of *Q. pubescens* Willd.

Apulia is the easternmost Italian administrative region and all Italian oak species, both evergreen and deciduous, occur in this region with the exception of *Q. petraea* (Matt.) Liebl. Despite this richness in oak species and a wide potential range for the oak woodlands, only 10% of the Apulian territory is covered by forests. The *Q. pubescens* s.l. forests are the most widespread and can be found in the flooded depressions of the plain, in the limestone plateaus and on the sub-montane rocky slopes.

Molecular marker studies have not yet been published for the genus *Quercus* in the Apulian region. However, the

evaluation of genetic variation can be of great importance for the conservation and management of forest ecosystems, especially in areas particularly vulnerable to climatic change (Peñuelas et al. 2017) and in forests whose original extent has been significantly reduced. This is the case of the oak forests of the Apulia region. In fact, the dry Mediterranean climate that characterizes this region places it at risk of desertification if the climate should experience a further increase in aridity. Moreover, the widespread practice of extensive agriculture (wheat, olive and vine) during the last century led to the destruction of most of the natural forest resources (Biondi et al. 2010).

Several reports on the population genetics of the white oaks *Q. robur* L., *Q. petraea*, *Q. pubescens* have been published for south-eastern Europe in the last two decades (Franjić et al. 2006, Curtu et al. 2007a, Jerše and Batič 2007, Slade et al. 2008, Curtu et al. 2009, Ballian et al. 2010, Enescu et al. 2013, Gailing et al. 2013). In Italy, only few data are available. Fineschi et al. (2002) and Fineschi and Vendramin (2004) presented a study on the genetic diversity of the Italian white oaks analyzing chloroplast DNA. Subsequently, several studies that compared leaf morphology and molecular data on some white oaks populations of Central Italy have been published (Fortini et al. 2009, 2013, 2015).

In the most recent papers published on the Apulia forests (Biondi et al. 2004, Di Pietro and Misano 2009) the following pubescent oaks were reported in the phytosociological tables: *Q. pubescens*, *Q. virgiliana*, *Q. amplifolia* and *Q. dalechampii*. A recent study (Di Pietro et al. 2016) statistically analyzed 25 morphological characters of leaves and fruits within 24 pubescent oak populations in the Apulia region, in order to identify possible diagnostic traits useful for taxonomic differentiation. The results provided no evidence for differentiation among tree individuals based on comparisons of morphological traits. In the present study plant material from the same pubescent oak populations was used to analyze their genetic diversity, structure and gene flow. The aims of this study are twofold. First, to assess whether any group of genetic structure could be identified among oak individuals and/or populations, notwithstanding the morphological and taxonomical uniformity highlighted in the previous morphological work (Di Pietro et al. 2016). Second, to investigate the pattern of genetic diversity in the Apulian pubescent oaks in order to provide useful information for further studies.

Material and methods

Study area

The Apulia region is located in the south-eastern part of the Italian Peninsula where it is largely open to the Adriatic and Ionian Seas. It is slightly sloping, with more than 60% of the territory occurring below 200 m a.s.l. Five physiographic units can be distinguished in the Apulia region and these are the Daunian sub-Apennine (1150 m), the Gargano promontory (1080 m), the Murgian Plateau (680 m), the Salento Peninsula and the large plain Tavoliere delle Puglie.

The bedrock of the mountainous systems is mainly composed of Cretaceous limestone and Paleocenic calcarenites. Marls occur in the lower parts of the Daunian sub-Apennine and in the Tavoliere plain together with conglomerates and Pleistocenic and Holocenic sand deposits. The bioclimate is Thermo-Mediterranean and Meso-Mediterranean with the Temperate region occurring only in the sub-and lower-montane belts. The rainiest areas are the Gargano, the Daunian sub-Apennine and the south-eastern part of the Salento Peninsula (slightly over 800 mm year^{-1}). Annual precipitation values on average of less than 500 mm year^{-1} are recorded in the western side to the Murgian Plateau and in the Tavoliere plateau whereas the remaining portion of the territory exhibits an average annual rainfall of between 500 and 700 mm year^{-1} (Blasi and Michetti 2007, Cotecchia et al. 2014). As far as the climax vegetation is concerned almost the entire Apulian territory would potentially be covered by oak forests except for some scattered coastal and subcoastal areas dynamically linked to the *Pinus halepensis* Mill. woods and maquis, and the inner part of the Gargano promontory where the wide "Foresta umbra" beech wood occurs (Biondi et al. 2004, 2010).

Population sample information

Individual mature oak trees were randomly selected in each population trying to maintain a distance of at least 20 meters from each other. Leaves were randomly collected from 379 individuals within 24 populations (subdivided in squares 50 × 50 m), which covered approximately uniformly the physiographic units of the Apulia region (Fig. 1, Tab. 1, On-line Suppl. Tab. 1). Both the individuals and the populations they belong to are the same on which the morphological analysis of leaves and acorns was performed in the paper by Di Pietro et al. (2016). The collected oak specimens were dried on silica gel, and stored at room temperature until analysis. Voucher specimens for each individual sampled were stored at the Herbarium of the University of Molise.

DNA extraction, EST-SSR amplification

Total genomic DNA was extracted from 0.5 g of dried leaves from all samples using spin columns of "Invisorb® Spin Plant Mini Kit" and following the protocol of the manufacturer.

Eleven microsatellite loci or simple sequence repeats (EST-SSRs), developed by Durand et al. (2010) (PIE020, PIE102, PIE152, PIE215, PIE223, PIE227, PIE239, PIE242, PIE243, PIE267, PIE271) and polymorphic in white oaks (Guichoux et al. 2011, Antonecchia et al. 2015) were selected. Polymerase chain reaction (PCR) amplification was performed in a single multiplex reaction in a DNA Engine Tetrad (MJ Research Bio-Rad) thermocycler and PCR products were run on an ABI 3730xl capillary sequencer (Applied Biosystems) using Genescan 600 LIZ internal size standard (Applied Biosystems). Alleles were scored using STR and software version 2.3.106 (Toonen and Hughes 2001) and al-
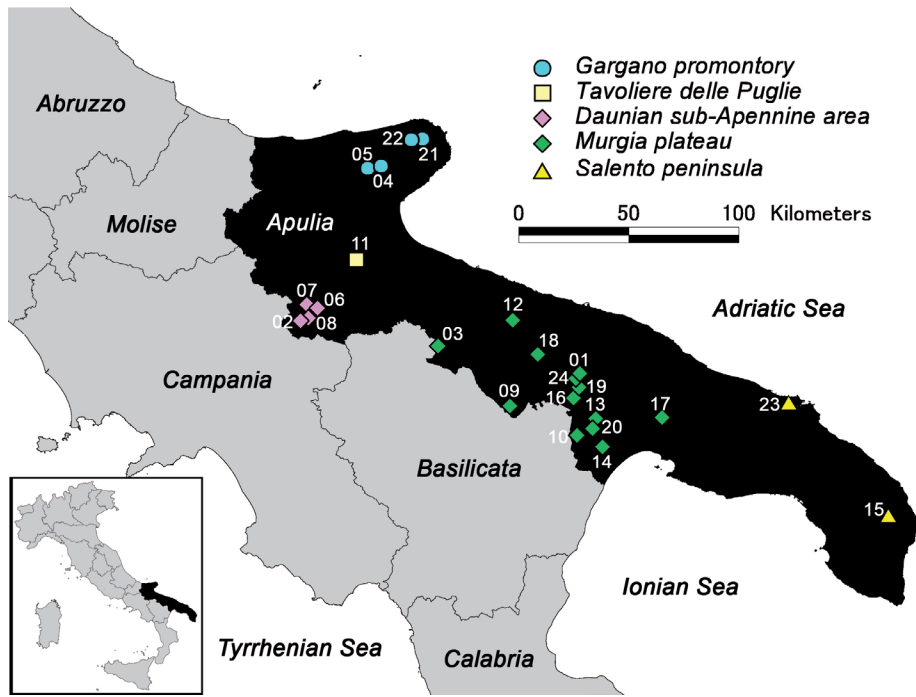
Fig. 1. Locations of the 24 oak collection stands.

Tab. 1. *Quercus* species composition of the 24 oak populations (according to Biondi et al. 2004, Biondi and Guerra 2008, Di Pietro and Misano 2009) with number of genotyped individuals and coordinates.

Population code	Analyzed individuals	Genotyped individuals	Latitude (N)	Longitude (E)	*Quercus* species
Pop01	16	16	4524805	650840	*Q. pubescens*
Pop02	16	16	4557220	526723	*Q. pubescens*
Pop03	16	16	4624669	567854	*Q. pubescens*
Pop04	16	16	4624043	561581	*Q. virgiliana, Q. dalechampii, Q. pubescens*
Pop05	16	15	4562416	534533	*Q. virgiliana, Q. dalechampii, Q. pubescens*
Pop06	16	15	4564231	529921	*Q. pubescens*
Pop07	16	16	4558496	530175	*Q. pubescens*
Pop08	16	16	4512469	618357	*Q. pubescens*
Pop09	18	16	4497035	647716	*Q. pubescens*
Pop10	16	16	4582747	553553	*Q. pubescens, Q. dalechampii, Q. virgiliana*
Pop11	16	16	4551101	622237	*Q. virgiliana, Q. dalechampii, Q. amplifolia*
Pop12	15	15	4504016	656599	*Q. pubescens*
Pop13	12	12	4491270	658881	*Q. pubescens, Q. dalechampii, Q. virgiliana*
Pop14	16	16	4446957	278106	*Q. virgiliana, Q. amplifolia, Q. dalechampii*
Pop15	15	15	4514201	647317	*Q. pubescens*
Pop16	16	11	4502657	686693	*Q. pubescens*
Pop17	16	13	4534971	632507	*Q. pubescens, Q. dalechampii, Q. virgiliana*
Pop18	16	14	4523011	649110	*Q. dalechampii, Q. virgiliana, Q. pubescens*
Pop19	16	14	4499728	654902	*Q. pubescens*
Pop20	17	14	4635503	582284	*Q. pubescens, Q. dalechampii, Q. virgiliana*
Pop21	16	16	4504890	744270	*Q. virgiliana, Q. dalechampii, Q. pubescens*
Pop22	14	11	4518821	650238	*Q. virgiliana, Q. dalechampii, Q. pubescens*
Pop23	16	6	4540867	597285	*Q. pubescens*
Pop24	16	7	4635677	587298	*Q. pubescens*

lele binning was performed as described in Guichoux et al. (2011). MICROCHECKER (Van Oosterhout et al. 2004) was used to check for the presence of null alleles and to check for potential problems related to allele dropout with 1000 randomizations and a 95% confidence interval.

Data analysis

Polymorphic EST-SSR markers were used to analyze the genetic diversity of 338 *Q. pubescens* s.l. individuals in 24 populations. EST-SSR amplification was unsuccessful for 41 of the 379 individuals sampled. Two populations (Pop23 and Pop24) were not included in the statistical analyses because of the low number of amplified individuals.

The basic statistics for each locus: number of alleles (NA), number of individuals (N), observed heterozygosity (H_o), expected heterozygosity (H_e), and number of private alleles (PA) were computed using the statistical program GenAlEx version 6.503 (Peakall and Smouse 2012).

Polymorphic information content (PIC), deviation from Hardy-Weinberg equilibrium (HW), and null allele frequency (F(Null)) were computed for each locus using Cervus software version 3.0.7 (Kalinowski et al. 2007).

The basic statistics for each population: number of different alleles, number of effective alleles, allelic richness, observed heterozygosity, expected heterozygosity, index of fixation (F_{IS}) calculated as F = 1 – (Ho / He) and number of private alleles, were computed using Arlequin software version 3.5.2.2 (Excoffier and Lischer 2010). Allelic richness per locus and population were computed using FSTAT version 2.9.4 (Goudet 2003).

Analysis of molecular variance (AMOVA) was performed using GenAlEx in order to partition the total microsatellite diversity, among and within populations, and within the individuals of each *Q. pubescens* s.l. population, based on 11 EST-SSR markers loci. The variance components were statistically tested by non-parametric randomization tests of significance based on 1000 bootstraps. Gene flow among populations was estimated using the indirect method based on the number of migrants per generation (Nm) using the formula, Nm = 0.25 (1 – F_{ST}) / F_{ST} (Slatkin and Barton 1989). To investigate population differentiations, pairwise F_{ST} between all pairs of populations was computed, F_{ST} was calculated according to Weir and Cockerham (1984) in program FSTAT 2.9.4.

Principal coordinate analysis (PCoA) with data standardization based on Nei's genetic distance was performed using GenAlEx.

Bayesian clustering analysis, using STRUCTURE software version 2.3.4 (Pritchard et al. 2000, 2010), with population IDs as Sampling Location Information, has been carried out. We opted for the admixture with correlated allele frequencies model between populations. We used a burn-in period of 25,000 and a number of MCMC of 100,000 cycles. Ten runs for each K (K = number of possible clusters)

value were done to test the consistency of the results. Each (K) value was tested from K = 1 to K = 10. The optimum K value was predicted using the web-based software STRUCTURE HARVESTER web version 0.6.94 which implements the Evanno method (Earl and vonHoldt 2012). Individuals with probabilities above 0.80 (Q ≥ 0.800) were assigned as putative purebred while all the other individuals were assigned as "putative hybrids" although the term "hybrid" could be a little misleading since no plant material of possible parental species was collected. For this reason we have opted for using the much more neutral term "off-type" in the rest of the text.

BOTTLENECK software version 1.2 (Cornuet and Luikart 1996) was used to detect the likelihood of a bottleneck effect. Heterozygosity excesses were displayed, identified based on the estimates of multilocus genotypes calculated by using the Wilcoxon signed rank tests and evaluating departures from the ratio 1:1 deficiency/excess (Cornuet and Luikart 1996, Luikart and Cornuet 1998, Piry et al. 1999). A two-phase mutation model (TPM) and a stepwise mutation model (SMM) for Wilcoxon signed-rank tests were used considering a 90% SMM proportion and TPM with a variance of 10, and 1000 repeats.

Results

SSR polymorphism and genetic diversity

All the loci were polymorphic (Pj = q ≤ 0.99). The total number of alleles identified was 100 out for the 325 individuals analyzed. The allele fragment sizes (Tab. 2) matched with the relative SSR reference size, as described in Guichoux et al. (2011). The number of alleles per locus (k) ranged between 4 and 17, with an average of 11 alleles. The PIC values varied from 0.233 (PIE227) to 0.879 (PIE152). According to (Hildebrand et al. 1992) eight EST-SSR loci were found to be informative (PIC ≥ 0.5), whereas two were found to be slightly informative (PIC < 0.4). The observed heterozygos-

Tab. 2. Allelic diversity of eleven microsatellite loci scored in 325 individuals of *Quercus pubescens* s.l.; S – size range fragments (bp); NA – Number of alleles at the locus, Ar – Allelic richness per locus, N – Number of individuals typed at the locus, H_o – Observed heterozygosity, H_e – Expected heterozygosity, PIC – polymorphic information content, HW – Significance of deviation from Hardy-Weinberg equilibrium (Hardy-Weinberg equilibrium test chi-square value, p value and significance with Bonferroni correction: ns = not significant, * = P < 0.05; ** = P < 0.01; *** = P < 0.001), F(Null) – Null allele frequency estimate, F_{IS} – Fixation index (* = P < 0.05).

Locus	S	NA	Ar	N	H_o	H_e	PIC	HW	F(Null)	F_{IS}
PIE020	97–109	7	3.386	312	0.583	0.585	0.537	35.434***	–0.026	–0.067
PIE102	139–167	13	5.367	321	0.523	0.750	0.726	87.243 ***	0.187	0.279*
PIE152	228–260	17	7.187	314	0.834	0.890	0.879	3.756 ns	0.032	0.043*
PIE215	185–218	12	6.295	281	0.384	0.707	0.673	193.668***	0.311	0.417*
PIE223	197–234	11	5.980	318	0.811	0.822	0.803	0.525 ns	0.006	–0.002
PIE227	156–165	4	3.387	323	0.223	0.251	0.233	3.776 ns	0.060	0.095*
PIE239	69–93	9	4.689	302	0.248	0.435	0.420	56.089***	0.274	0.399*
PIE242	102–132	14	5.924	318	0.748	0.841	0.821	16.677ns	0.058	0.093*
PIE243	204–224	10	4.392	282	0.688	0.683	0.649	37.002***	–0.029	–0.038
PIE267	84–106	12	4.029	313	0.425	0.667	0.622	82.591***	0.227	0.334*
PIE271	182–204	12	2.130	311	0.949	0.843	0.823	37.040***	–0.063	–0.153
Mean	–	11	–	308.64	0.583	0.679	0.653	–	0.094	–
Standard error	–	1.053	–	4.386	0.074	0.058	0.059	–	0.040	–

ity (H_o) for microsatellite loci ranged from 0.223 (PIE227) to 0.949 (PIE271), and the average was 0.583. A high level of H_o was detected in seven EST-SSR loci ranging from 0.523 to 0.949, whereas four loci detected low H_o with values ranging from 0.223 to 0.425. The expected heterozygosity (H_e) ranged from 0.251 to 0.890 with an average value of 0.679 and was higher than H_o for all the microsatellite loci, except for PIE243 and PIE271. The Hardy-Weinberg exact test for all populations (HW) revealed that seven loci (PIE020, PIE102, PIE 215, PIE239, PIE243, PIE267, PIE271) exhibited significant deviation from Hardy–Weinberg equilibrium (P<0.001) whereas four loci (PIE152, PIE223, PIE227, PIE242) did not show significant departures from Hardy-Weinberg equilibrium.

The indexes of genetic diversity of the 22 *Q. pubescens* s.l. populations are summarized in Tab. 3. The number of differ-

ent observed alleles (Na) and effective alleles (Ne) averaged across all loci ranged from 4.727 (Pop22) to 7.182 (Pop09) and 2.559 (Pop22) to 4.363 (Pop20), respectively. Nine private alleles (PA) was found and these are distributed within ten individuals in turn distributed in eight oak populations (two other private alleles were found in Pop24; On-line Suppl. Tab. 2). The allelic richness (Ar), ranged between 3.865 (Pop3) and 5.171 (Pop20). The average observed heterozygosity (aH_o) ranged from 0.487 (Pop03) to 0.665 (Pop11) with mean value of 0.583. The average expected heterozygosity (aH_e) ranged from 0.510 (Pop03) to 0.709 (Pop09) with mean value of 0.629. Overall, the observed heterozygosity showed mostly slightly lower values than the expected heterozygosity. The mean fixation index (F_{IS}) values are close to zero for all populations (ranging from 0.058 in Pop11 to 0.188 in Pop21) that are the values expected under a random mating.

Population genetic structure and gene flow

The overall population differentiation degree is reported in Tab. 4. AMOVA analysis showed that 3.6% (P < 0.001) of the genetic variations was among populations while 96.4% was within populations. The majority of molecular variance was partitioned within individuals (78.5%, P < 0.01) while 17.9% (P < 0.01) of variance was found to occur among individuals. The pairwise F_{ST} matrix between populations is shown in On-line Suppl. Tab. 3; eleven population pairs exhibited negative values but all of them showed no significant p values (all largely > 0.05). The differentiation coefficient between population pairs having significant P values showed that Pop11 (Incoronata-Tavoliere) and Pop15 (Palmariggi-Salento) was the smallest (F_{ST} = 0.0025; P ≤ 0.001); the differentiation coefficient between Pop02 (Deliceto-Daunian sub-Apennine) and Pop22 (Vico del Gargano-East Gargano) was the largest (F_{ST} = 0.1053; P < 0.01).

Nei's genetic distance among the populations belonging to the five units are graphically illustrated in Fig. 2, constructed on the basis of principal coordinate analysis (PCoA with the physiographic units superimposed). The first two axes of the PCoA explained 42.26% of the variation. The scattergram did not show an evident correlation between the physiographic units and the genetic similarities of the populations. In fact, populations belonging to four different physiographic units are closely associated along the first axis in the left part of the diagram. The Mantel test comparing the matrices of geographic and genetic distances among populations did not evidence any correlation between these two variables.

Tab. 3. Genetic diversity parameters for the *Quercus pubescens* s.l. populations (Pop) analyzed through eleven microsatellite loci. n – number of individuals; N_a – average number of different alleles; N_e – average number of effective alleles; Ar – allelic richness; aH_e – average expected heterozygosity; aH_o – average observed heterozygosity; F_{IS} – average of fixation index tested by 1023 permutation of gene copies between individuals within population (* = P < 0.05); PA – number of private alleles.

Pop	n	N_a	N_e	Ar	aH_o	aH_e	F_{IS}	PA
Pop01	16	6.182	3.641	4.642	0.581	0.604	0.035	0
Pop02	16	5.091	3.353	4.064	0.498	0.584	0.125*	0
Pop03	16	4.818	2.968	3.865	0.487	0.510	0.029	0
Pop04	16	6.364	3.684	4.605	0.572	0.638	0.094*	0
Pop05	15	5.182	3.097	4.192	0.604	0.610	-0.036	1
Pop06	15	5.182	3.055	3.940	0.529	0.590	0.037	0
Pop07	16	5.636	3.213	4.156	0.557	0.544	-0.059	0
Pop08	16	6.455	3.888	4.775	0.577	0.632	0.086*	0
Pop09	16	7.182	4.166	4.903	0.605	0.709	0.155*	0
Pop10	16	6.455	3.979	4.601	0.587	0.656	0.093*	0
Pop11	16	6.000	3.369	4.313	0.665	0.614	-0.112	0
Pop12	15	5.909	3.790	4.420	0.628	0.655	0.025	1
Pop13	12	5.545	3.883	4.760	0.595	0.705	0.092	2
Pop14	16	6.091	3.749	4.487	0.655	0.648	-0.018	1
Pop15	15	5.727	3.503	4.583	0.596	0.634	0.066	0
Pop16	11	5.273	3.400	4.606	0.627	0.647	0.006	1
Pop17	13	6.091	3.700	4.560	0.580	0.639	0.062	0
Pop18	14	5.909	3.688	4.651	0.605	0.642	0.005	1
Pop19	14	6.545	3.955	4.841	0.583	0.669	0.064	0
Pop20	14	7.091	4.363	5.171	0.617	0.692	0.120*	1
Pop21	16	6.000	3.753	4.510	0.557	0.684	0.116*	1
Pop22	11	4.727	2.559	3.941	0.527	0.540	-0.129	0
Mean	14.8	5.884	3.580	4.481	0.583	0.629	0.039	

Tab. 4. Analysis of molecular variance among populations, within populations, and within individuals of *Quercus pubescens* s.l. populations; Df – degree of freedom, SS – sum of squares, MS – mean squares, P – probability based on standard permutation across the full data set, Nm –average estimate of gene flow among populations.

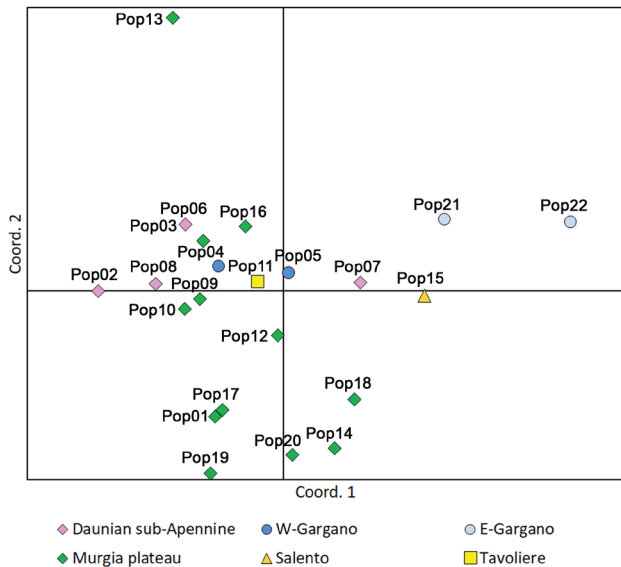| Source of variation	Df	SS	MS	Estimated variation	Percentage variation %	F-statistics	P-value	Gene flow
Among Populations	21	180.496	8.595	0.141	3.6	Fst = 0.036		Nm = 6.664
Among Individuals	303	1346.241	4.443	0.695	17.9	Fis = 0.185	0.001	
Within Individuals	325	992.500	3.054	3.054	78.5	Fit = 0.215		
Total	649	2519.237		3.889	100			

Fig. 2. Principal Coordinate Analysis (PCoA) via Distance matrix (Nei Genetic Distance) with data standardization. Daunian sub-Apennine populations: Pop02, 06, 07, 08; W-Gargano populations: 04, 05; E-Gargano populations: 21, 22; Murgia plateau populations: 01, 03, 09, 10, 12, 13, 14, 16, 17, 18, 19, 20; Salento population: 15; Tavoliere population: 11.

When STRUCTURE was run using population IDs as sampling location information, a maximum value of the rate of change in the log probability of data was revealed at K = 2, using Evanno's method (Evanno et al. 2005) (On-line Suppl. Fig. 1).

The individual membership proportion determined by Bayesian clustering analysis suggested a subdivision of the whole data-set of oak individuals (325) into two ge-

netic clusters: a first cluster including 305 pure individuals (94.15%) and a second one composed of 19 off-type individuals (5.85%) (Fig. 3a). The distribution of the 19 off-types involved 13 populations out of the 22 investigated (Fig. 3b, On-line Suppl. Tab. 4). Population 13, where the sampled individuals were selected from a mixed wood with *Q. pubescens* and *Q. robur* that developed in a humid area, is the population showing the highest number (4) of off-types. Populations 2, 9, and 10 count two off-types each and relate to *Q. pubescens* s.l. woods with the occurrence of *Q. frainetto*.

The Wilcoxon sign-rank test showed no significant results as regards bottleneck effect in all the populations analyzed, using both TPM and SMM models. The observed number of loci with heterozygosity excess (obs LH_{exc}) was found to be always lower than the expected number of loci with heterozygosity excess (exp LH_{exc}) except for population 13 where "obs LH_{exc}" is about 1.5 times greater than "exp LH_{exc}". However, the Wilcoxon sign-rank test was not significant (Tab. 5).

Discussion

The data-set exhibited the following levels of genetic diversity at the microsatellite loci examined: mean expected heterozygosity (H_e) = 0.679, mean observed heterozygosity (H_o) = 0.583, average number of alleles per locus (K) = 11.0 (from 4 to 17 alleles). Four (PIE102, PIE215, PIE239, PIE267), out of the eleven loci investigated, showed a positive F_{IS} (0.279-0417) which turned out to be statistically significant (Tab. 2). It is known that the fixation of alleles could be favored by directional selection (Andolfatto 2001), or be the results of genetic drift. Molecular studies regarding oth-
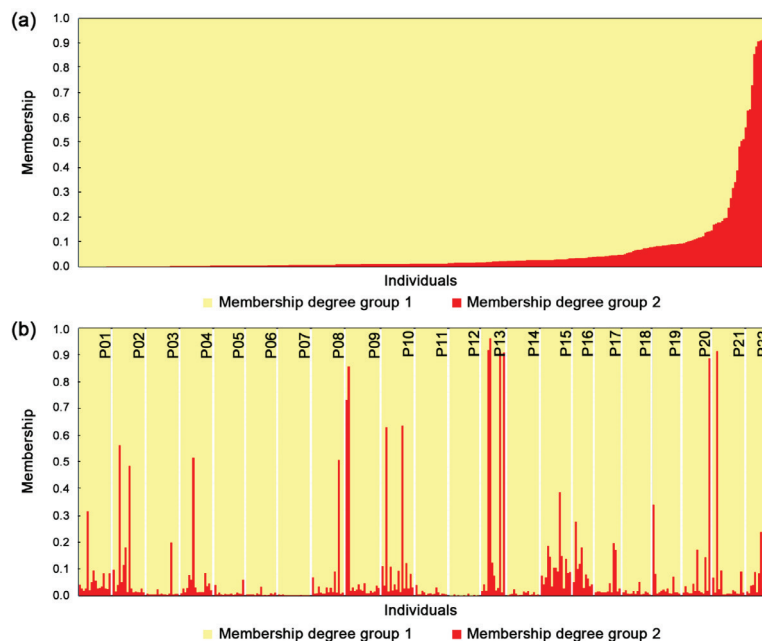

Fig. 3. Genetic assignment obtained with STRUCTURE clustering analysis. Oak individuals are distributed in progressive order along the horizontal axis while the vertical axis expresses the group membership percentage degree per single individual per K=2. Clustering of all the data-set oak individuals into two different genetic clusters with decreasing membership degrees for group 2 and increasing degrees for group 1 (a). Clustering of all the data-set oak individuals ordered per populations; the latter separated by white lines (b).

Tab. 5. Results of tests for genetic bottlenecks using the two-phase mutation model (TPM) and a stepwise mutation model (SMM), and the Wilcoxon sign-rank tests for heterozygosity excess; the number of polymorphic loci for all populations (Pop) is 11; obs LH_{exc} – observed number of loci with heterozygosity excess; exp LH_{exc} – expected number of loci with heterozygosity excess; P – probability of no significant heterozygosity excess.

Pop	TPM			SMM		
	obs LH_{exc}	exp LH_{exc}	P	obs LH_{exc}	exp LH_{exc}	P
Pop01	4	6.46	0.9492	2	6.31	0.9954
Pop02	6	6.44	0.7114	4	6.43	0.8398
Pop03	4	6.43	0.9385	4	6.43	0.9585
Pop04	4	6.37	0.9126	4	6.37	0.9585
Pop05	3	6.29	0.8608	3	6.29	0.9263
Pop06	5	6.39	0.7676	4	6.35	0.8799
Pop07	4	6.36	0.9585	4	6.38	0.9731
Pop08	5	6.49	0.9492	4	6.58	0.9663
Pop09	3	6.55	0.9492	2	6.50	0.9976
Pop10	5	6.27	0.7676	3	6.36	0.8970
Pop11	2	6.39	0.9895	2	6.49	0.9966
Pop12	5	6.49	0.5845	5	6.57	0.7114
Pop13	8	6.52	0.1602	8	6.61	0.2598
Pop14	4	6.59	0.8608	4	6.58	0.9585
Pop15	6	6.58	0.7676	5	6.51	0.8608
Pop16	4	6.59	0.9263	4	6.55	0.9663
Pop17	5	6.46	0.8799	2	6.51	0.9919
Pop18	3	6.47	0.9126	3	6.50	0.9126
Pop19	4	6.39	0.9126	4	6.50	0.9492
Pop20	4	6.48	0.8799	3	6.59	0.9919
Pop21	6	6.56	0.6812	4	6.57	0.8970
Pop22	3	6.56	0.9663	2	6.53	0.9954

er oak species stated that the genetic variation removed by genetic drift would affect the genome rather uniformly (Alberto et al. 2010). This would not seem to be the case of the Apulian pubescent oak populations, since the other EST-SSR loci examined showed only slightly positive values of F_{IS}, or even negative ones. Due to the relatively low number of sampled individuals per population, it cannot be completely excluded that these results could be addressed, at least in part, to the sampling effect. However, there is no uniformity of views at present on the minimum number of individuals per population that should be sampled to detect "reliable" results for genetic differentiation. Hale et al. (2012) stated that the number of individuals per population should be established using a "case by case" approach. On the other hand the genetic literature on this issue shows a wide range of opinions and experiences from which it emerges that the number of individuals sampled can be highly variable. In many cases it was considered as acceptable to take a minimum of 10 to 30 individuals. For some *Pinus sylvestris* L. populations in the Baltic area, the number of 20-25 individuals was found to be large enough to detect all the alleles (Danusevičius et al. 2016) whereas other studies have even considered significantly lower numbers (Kitamura et al. 2017, Rinaldi et al. 2019). In contrast, in Kalinowski (2005) it was shown particularly clearly that when F_{ST} is less than 0.01 it can be useful to sample up to 100 individuals. As far as we are concerned, we believe that these results, although deriving from

the analysis of populations composed of a seemingly limited number of samples, can still provide interesting information that should not be completely ignored.

The highest allelic variability was found in locus PIE152 with 17 alleles and PIC = 0.879. A similar variability was found in other pubescent oak populations investigated in southern Italy and central Italy (Antonecchia 2015) where locus PIE152 exhibited the highest polymorphic degree and in particular the highest values of allelic richness for *Q. pubescens*. The still insufficient knowledge about the *Q. pubescens* genome do not allow it to be established if the high variability in locus PIE152 might be associated with adaptive traits. More in-depth investigations would be needed to clarify this point.

The level of genetic variability within the pubescent oak populations of the Apulia region was here found to be lower than that reported for *Q. pubescens* in other countries. Curtu et al. (2007b) reported aH_e = 0.891, Ar = 17.8 on a study based on 6 microsatellite loci for 73 *Q. pubescens* individuals while Enescu et al. (2013) reported aH_e=0.847, k = 22 and Ar = 14.22 in a 7 microsatellite loci study in a geographically related area. The heterozygosity values reported in these studies, however, would not be directly comparable with those obtained by us due to the use of different type of markers (genomic SSRs vs. EST-SSRs). Some authors (Durand et al. 2010, Parthiban et al. 2018) showed that the use of EST-SSRs highlights the values of polymorphism slightly lower than those obtained from the use of genomic SSRs. However, there are some studies on the *Quercus* genus in which both these types of markers were tested. Curtu et al. (2011) investigated the genetic variability of 65 individuals of *Q. pubescens* using 7 microsatellite loci, of which 5 gSSR and 2 EST-SSRs. The overall average values for the genetic parameters in issue were the following: aH_e = 0.859, k = 17, Ar = 16.5. Calculating the values of these parameters solely for the two EST-SSRs markers we obtained values (aH_e = 0.819, k = 17, Ar = 16.8) that are lower than those obtained averaging both kind of markers but still significantly higher than those we obtained from our Apulian pubescent oak data-set.

The only available analyses of genetic differentiation within Italian white oak populations including *Q. pubescens* s.l. carried out using EST-SSRs markers are those reported for the mixed *Q. petraea*, *Q. frainetto*, *Q. pubescens* woods of the Mount Vairano range in southern Italy (Antonecchia et al. 2015). Comparison between these analyses and our data revealed that the genetic variability of the Apulian populations is significantly lower than that recorded in Mount Vairano (aH_e = 0.72, k = 9.46 Ar = 9.2) as regards aH_e and Ar, and slightly higher as regards k.

Also the level of genetic variability among the *Q. pubescens* s.l. populations of the Apulia region (F_{ST} = 0.036) was found to be lower than the F_{ST} values found for other oak species in other parts of the world, e.g. *Q. rubra* L. (F_{ST} = 0.080) (Sullivan et al. 2013), *Quercus mongolica* Fisch. ex Ladeb. (F_{ST} = 0.077) (Ueno and Tsumura 2008), *Q. variabilis* Bl. (F_{ST} = 0.063) (Shi et al. 2017).

As regards gene flow, the present study showed that the Apulian *Q. pubescens* s.l. populations are characterized by a relatively high values (Nm = 6.664). Although *Q. pubescens* is a typical wind-pollinated and outcrossing species such high values are quite surprising considering the progressive contraction suffered by the Apulian oak woods in the last millennium. The value of gene flow is determined by many factors, such as the biological means of spreading via pollen and seeds, the occurrence of physical barriers among populations, population dimension and so forth; in addition, it can be facilitated by physical proximity of the populations. It is probable that the short distance separating some of the Apulian pubescent oak populations and the lack of significant mountain barriers have contributed to keep the average gene flow value high. It is not possible, at present, to establish whether the above gene-flow values are stable over time, or are the result of a contraction due to over-exploitation of forests. In studies on other oak species (Shi et al. 2017) it was established that gene-flow values can decrease significantly in a relatively short time (ten years) following the transition from primary forest to natural secondary forest. In the Apulia region the millennial exploitation of the woodlands led to the disappearance of primeval forests as early as Roman times. However, the secondary forest communities that have since replaced the primary ones show a high degree of naturalness and a rather stable floristic composition. It is conceivable therefore that the gene flow rate will remain stable over time and that the only parameter capable of negatively acting on the gene-flow rate is the spatial reduction of the population size.

The apparent relatively low level of intraspecific diversity observed for the Apulian *Q. pubescens* populations could be addressed to their marginal geographical position not only in the Italian Peninsula but also in the central Mediterranean context. In terms of biological conservation of deciduous oak forests the Apulia region is particularly exposed to environmental risks due to climate change drying effects which overlap a bioclimatic pattern dominated by the Thermo-Mediterranean thermotype the latter having in the Mediterranean maquis the main type of potential vegetation (Rivas-Martínez et al. 2004, Ladisa et al. 2012). Threats to pubescent oak woods are further increasing in many localities that are experiencing greater aridity and water deficiency due to the overexploitation of water reserves resulting from the ever increasing consumption for agricultural purposes and by tourist settlements. In the long term, the combined effect of these two factors could lead to a change in the floristic composition of the upper structural layers of the thermophilous deciduous oak forests with a progressive replacement of the pubescent oak individuals by evergreen woody species, such as *Quercus ilex* L. *Phillyrea latifolia* L., *Rhamnus alaternus* L., *Arbutus unedo* L. and *Pistacia lentiscus* L. Conversely, the high gene flow values observed in the study area seem to move in the opposite direction and might positively contribute to preventing possible activations of genetic drift. Yet the lack of bottleneck effects in all investigated populations implies a certain degree of resilience despite the relatively small areas covered by the single oak populations.

As regards possible taxonomic implications arising from this study, the results of STRUCTURE Bayesian clustering showed that the populations of Apulian pubescent oaks could not be divided into groups. None of the populations investigated displayed significant differences in their genetic composition. However, the highest percentage of the possible "putative hybrids" that we simply named "off-types", was found in Pop13, which also showed the highest percentage of private alleles. Pop13 is located within a flat sub-humid area in the western side of the Murgian plateau in the municipality of Laterza, and lies in spatial contact with a small stand of *Q. robur* wood (the only known site of *Q. robur* currently recorded for the Apulia region). It is conceivable that in this area *Q. robur* was much more abundant in the past and that both climate change and the millennial work of deforestation carried out by the local populations, has led to its current extreme impoverishment. It is very probable that during the climatic oscillations that took place during the Quaternary, environmental situations favorable to the development of mixed forests of *Q. robur* and *Q. pubescens* s.l. were created several times with consequent hybridization or introgression between these two species. In fact, a possible key of interpretation of genetic structure of the *Q. pubescens* Apulian populations cannot fail to consider some physiographic and geographical features of this region such as the lack of sizeable mountain systems and its deep wedges jutting into the Adriatic Sea. These features led to the Apulian Peninsula experiencing the effects of quaternary glaciations only mildly and allowed some parts of it to work as glacial refuges for the thermophilous oak forests during the Quaternary cold periods, when most of the Italian Peninsula was covered by steppic grasslands (Follieri et al. 1988). The occurrence of restricted areas where the climate forced the various species of white oaks to coexist for long (cold) periods may have favored the generation of hybrids. The innumerable past hybridization or introgression events, which led to the current high morphological variability of *Q. pubescens* s.l., may also have played a role in enlarging the gene flow of the Apulian populations. White oaks commonly live in sympatry, making them particularly suitable for gene exchange and production of hybrid individuals (Rushton 1993, Williams et al. 2001, Lepais and Gerber 2011). According to Burger (1975), applying the biological concept of species to the genus *Quercus* would cause an upheaval of the nomenclature, since the binomial species would then not correspond to the biological one. This makes the systematics of this genus extremely susceptible to disagreement among botanists and the hope of arriving at the proposal of a largely shared and possibly unambiguous taxonomic framework (in particular for the south-European white oaks) is something not soon to be realized. The problematic classification of pubescent oaks in the Apulian Peninsula and in the entire Italian Peninsula is perfectly evidenced by the discrepancy that emerges when comparing the taxonomy of the genus *Quercus* as reported in latest versions of Flora d'Italia (Pignatti et al. 2017) and in the Checklist of the Italian vascular flora (Bartolucci et al. 2018). Although both works are worthy of the utmost respect and

consideration, they show enormous differences in both the numbers and the names of the taxa considered good species for the same taxonomic group (white oaks). Actually, based on what emerges from our work, the high taxonomic differentiation among the Italian pubescent oaks as reported in most of the Italian floras and checklists does not seem to have much foundation (see also Di Pietro et al. 2012) or at least does not seem to have it with regard to the pubescent oaks of Apulian Peninsula. In fact, our results established that it was not possible to identify genetic clusters among the pubescent oak populations of the Apulian region basing on 11 highly polymorphic markers as it was not possible to identify morphological clusters basing in a previous biometric study on the morphological traits of leaves and fruits on the same set of specimens (Di Pietro et al. 2016).

Summarizing, the presence of more than one taxon at species rank belonging to *Q. pubescens* s.l. (e.g., *Q. virgiliana*, *Q. dalechampii*, *Q. amplifolia*, etc.) is not confirmed (at least at present) for the Apulia region. At the same time, the phytosociological frameworks that are currently supposed to classify different types of pubescent oak forests (phytosociological associations), dominated by different pubescent oak species and including up to five different species of pubescent oaks in the dominant tree layer, need at the very least to be reconsidered. Obviously we do not consider this work to be a solution to all the problems concerning the taxonomy of white pubescent oaks, not even to some of them of topical relevance

in south-eastern Europe, such as the eternal debate on the dualism between *Q. pubescens* and *Q. virgiliana* (cf. Borazan and Babaç 2003, Škvorc et al. 2005, Franjić et al. 2006, Jerše and Batič 2007, Trinajstić 2007, Enescu et al. 2013). We simply believe that this work represents a small step forward, a substantial contribution, albeit on a regional scale, to being able to arrive in the near future at a possible solution of the problem on a global scale. Finally, it is worth emphasizing, that all the taxonomical and phytosociological considerations mentioned in this paragraph should not be viewed as useful only for updating floras, checklists or vegetation maps. In fact, they have direct implications for the application of European Policies for Nature conservation. It is only necessary to recall the 92/43/EEC European directive where the identification of some of the forest habitats listed in the Manual of Interpretation (e.g., 91H0* "Pannonian woods with *Quercus pubescens*", 91AA* "Eastern white oak woods", 91M0 "Pannonian-Balkanic turkey oak-sessile oak forests", 91Y0 "Dacian oak and hornbeam forests"), is also based on the occurrence of species such as *Q. virgiliana* and *Q. dalechampii*.

Acknowledgments

References

Alberto, F., Niort, J, Derory, J., Lepais, O., Vitalis, R., Galop, D., Kremer, A., 2010: Population differentiation of sessile oak at the altitudinal front of migration in the French Pyrenees. Molecular Ecology 19, 2626–2639.

Andolfatto, P., 2001: Adaptive hitchhiking effects on genome variability. Current Opinion in Genetics and Development 11, 635–641.

Antonecchia, G., 2015: Analisi della variabilità genetica del sottogenere *Quercus* Oerst. in Italia centro-meridionale. PhD thesis, Department of Bioscience and Territory, University of Molise, Pesche (IS). Retrieved April 10, 2016 from http://hdl.handle.net/2192/296.

Antonecchia, G., Fortini, P., Lepais, O., Gerber, S., Léger, P., Scippa, G. S., 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). Annals of Forest Research 58, 205–216.

Ballian, D., Ivanković, M., Gračan, J., Perić, S., Marjanović, H., Bobinac, M., Slade, D., 2010: Analysis of pubescent oak (*Quercus pubescens* Willd.) in the western part of the Balkan peninsula. Acta Societatis Botanicorum Poloniae 79, 189–195.

Bartolucci, F., Peruzzi, L., Galasso, G., Albano, A., Alessandrini, A., Ardenghi, N.M.G., 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., Iamonico, D., Iberite, M., Jiménez-Mejías, P., Lattanzi, E., Marchetti, D., Martinetto, E., Masin, R.R., Medagli, P., Passalacqua, N.G., Peccenini, S., Pennesi, R., Pierini, B., Poldini, L., Prosser, F., Raimondo, F.M., Roma-Marzio, F., Rosati, L., Santangelo, A., Scoppola, A., Scortegagna, S., Selvaggi, A., Selvi, F., Soldano, A., Stinca, A., Wagensommer, R.P., Wilhalm, T., Conti, F., 2018: An updated checklist of the vascular flora native to Italy. Plant Biosystems 152, 179–303.

Biondi, E., Casavecchia, S., Beccarisi, L., Marchiori, S., Medagli, P., Zuccarello, V., 2010: Le serie di vegetazione della Regione Puglia. In: Blasi, C. (ed.), La vegetazione d'Italia, 391–409. Palombi Editori, Roma.

Biondi, E., Casavecchia, S., Guerra, V., Medagli, P., Beccarisi, L., Zuccarello, V., 2004: A Contribution towards the knowledge of semideciduous and evergreen woods of Apulia (southeastern Italy). Fitosociologia 41, 3–28.

Biondi, E., Guerra, V., 2008: Vegetazione e paesaggio vegetale delle gravine dell'arco jonico. Fitosociologia 45, 57–125.

Blasi, C., Michetti, L., 2007: Biodiversity and climate. In: Blasi, C., Boitani, L., La Posta, S., Manes, F., Marchetti, M. (eds.), Biodiversity in Italy. Contribution to the National Biodiversity Strategy, 57–66. Palombi Editori, Roma.

Borazan, A., Babaç, M.T., 2003: Morphometric leaf variation in oaks (*Quercus*) of Bulu, Turkey. Annales Botanici Fennici 40, 233–242.

Burger, W.C., 1975: The Species Concept in *Quercus*. Taxon 24, 45–50.

Cornuet, J.M., Luikart, G., 1996: Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144, 2001–2014.

Cotecchia, V., Simeone, V., Gabriele S., 2014: Caratteri climatici della Regione Puglia. In: Ispra Ambiente (ed.), Memorie descrittive della Carta geologica d'Italia 92, 338–369. ISPRA serv. Geologico d'Italia, Roma.

Curtu, A.L., Gailing, O., Finkeldey, R., 2007a: Evidence for hybridization and introgression within a species-rich oak (*Quercus* spp.) community. BMC Evolutionary Biology 7, 218.

Curtu, A.L., Gailing, O., Finkeldey, R., 2009: Patterns of contemporary hybridization inferred from paternity analysis in a four-species forest. BMC Evolutionary Biology 9: 284.

Curtu, A.L., Gailing, O., Leinemann, L., Finkeldey, R., 2007b: Genetic variation and differentiation within a natural community of five oak species (*Quercus* spp.). Plant Biology 9, 116–126.

Curtu, A.L., Moldovan, I.C., Enescu, C.M., Craciunesc, J., Sofletea, N., 2011: Genetic Differentiation between *Quercus frainetto* Ten. and *Q. pubescens* Willd. in Romania. Notulae Botanicae Horti Agrobotanici Cluj-Napoca 39, 275–282.

Danusevičius, D., Kavaliauskas, D., Fussi, B., 2016: Optimum sample size for SSR-based estimation of representative allele frequencies and genetic diversity in Scots pine populations. Baltic Forestry 22, 194–202.

Di Pietro, R., Di Marzio, P., Medagli, P., Misano, G., Silletti, G.N., Wagensommer, R.P., Fortini, P., 2016: Evidence from multivariate morphometric study of the *Quercus pubescens* complex in southeast Italy. Botanica Serbica 40, 83–100.

Di Pietro, R., Misano, G., 2009: Analisi fitosociologica e considerazioni sintassonomiche sulla vegetazione forestale delle Gravine occidentali dell'Arco Ionico (Murge pugliesi e lucane, Italia meridionale). Informatore Botanico Italiano 41, 215–246.

Di Pietro, R., Viscosi, V., Peruzzi, L., Fortini, P., 2012: A review of the application of the name *Quercus dalechampii*. Taxon 61, 1311–1316.

Durand, J., Bodénès, C., Chancerel, E., Frigerio, Vendramin, G., Sebastiani, F., Buonamici, A., Gailing, O., Koelewijn, H.P., Villani, F., Mattioni, C., Cherubini, M., Goicoechea, P.B., Herrán, A., Ikaran, 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 ESTSSR markers, oak as a case study. BMC Genomics 11, 570.

Earl, D.A., von Holdt, B.M., 2012: STRUCTURE HARVESTER, a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4, 359–361.

Enescu, C.M., Curtu, A. L., Sofletea, N., 2013: Is *Quercus virgiliana* a distinct morphological and genetic entity among European white oaks? Turkish Journal of Agriculture and Forestry 37, 632–641.

Evanno, G., Regnaut, S., Goudet, J., 2005: Detecting the number of clusters of individuals using the software STRUCTURE, a simulation study. Molecular Ecology 14, 2611–2620.

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

Fineschi, S., Taurchini, D., Grossoni, P., Petit, R.J., Vendramin, G.G., 2002: Chloroplast DNA variation of white oaks in Italy. Forest Ecology and Management 156, 103–114.

Fineschi, S., Vendramin, G.G., 2004: La diversità cloroplastica delle querce italiane: evidenze di una maggiore ricchezza genetica nelle popolazioni meridionali e insulari. Forest@ 1, 82–87.

Follieri, M., Magri, D., Sadori, L., 1988: 250,000-year pollen record from Valle di Castiglione (Roma). Pollen et Spores 30, 329–256.

Fortini, P., Antonecchia, G., Di Marzio, P., Maiuro, L., Viscosi, V., 2013: Role of micromorphological leaf traits and molecular data in taxonomy of three sympatric white oak species and their hybrids (*Quercus* L.). Plant Biosystems 149, 546–558.

Fortini, P., Di Marzio, P., Di Pietro, R., 2015: Differentiation and hybridization of *Quercus frainetto*, *Q. petraea*, and *Q. pube-scens* (Fagaceae), insights from macro-morphological leaf traits and molecular data. Plant Systematics and Evolution 301, 375–385.

Fortini, P., Viscosi, V., Maiuro, L., Fineschi, S., Vendramin, G.G., 2009: Comparative leaf surface morphology and molecular data of five oaks of the subgenus *Quercus* Oerst (Fagacee). Plant Biosystems 143, 543–554.

Franjić, J., Liber, B., Škvorc, Z., Idžojtić, M., Šostarić, R., Stančić, Z., 2006: Morphological and molecular differentiation of the croatian populations of *Quercus pubescens* Willd. (Fagaceae). Acta Societatis Botanicorum Poloniae 75, 123–130.

Gailing, O., Bodénès, C., Finkeldey, R., Kremer, A., Plomion, C., 2013: Genetic mapping of EST-derived Simple Sequence Repeats (EST-SSRs) to identify QTL for leaf morphological characters in a *Quercus robur* full-sib family. Tree Genetics and Genomes 9, 1361–1367.

Goudet, J., 2003: FSTAT (version 2.9.4), a program (for Windows 95 and above) to estimate and test population genetics parameters. Retrieved November 26, 2015 from http://www2.unil.ch/izea/softwares/fstat.html

Govaerts, R., Frodin, D.G., 1998: World checklist and bibliography of Fagales (Betulaceae, Corylaceae, Fagaceae and Ticodendraceae). Royal Botanic Gardens: Kew, Richmond.

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoittevein, C., Malausa, T., Revardel, E., Salin, F., Petit, R.J., 2011: Current trends in microsatellite genotyping. Molecular Ecology Resources 11, 591–611.

Hale, M.L., Burg, T.M., Steeves, T.E., 2012: Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. PLOS ONE 7, e 45170.

Hildebrand, C.E., Torney, D.C., Wagner, R.P., 1992: Informativeness of polymorphic DNA markers. Los Alamos Science 20, 100–102.

Jerše, M., Batič, F., 2007: Morphological analysis of pubescent oak (*Quercus pubescens* Willd.) in Slovenia. Zbornik gozdarstva in lesarstva 83, 35–45.

Kalinowski, S.T., 2005: Do polymorphic loci require large sample sizes to estimate genetic distances? Heredity 94, 33–36.

Kalinowski, S.T., Taper, M.L., Marshall, T.C., 2007: Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16, 1099–1106.

Kitamura, K., Namikawa, K., Kawahara, T., Matsumoto, A., San Jose-Maldia, L., 2017: Genetic structure of remnant *Quercus serrata* populations at the northernmost limit of their distribution in Japan. Acta Phytotaxonomica et Geobotanica 68, 1–15.

Ladisa, G., Todorovic, M., Trisorio, L.G., 2012: A GIS-based approach for desertification risk assessment in Apulia region, SE Italy. Physics and Chemistry of the Earth 49, 103–113.

Lepais, O., Gerber, S., 2011: Reproductive patterns shape introgression dynamics and species succession within the European white oak species complex. Evolution 65, 156–170.

Luikart, G., Cornuet, J.M., 1998: Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology 12, 228–237.

Nixon, K.C., 1993: Infrageneric classification of *Quercus* (Fagaceae) and typification of sectional names. Annals of Forest Science 50 (Suppl. 1993), 25s–34s.

Parthiban, S., Govindaraj, P., Senthilkumar, S., 2018: Comparison of relative efficiency of genomic SSR and EST-SSR markers in estimating genetic diversity in sugarcane. 3 Biotech 8, 144.

Peakall, R., Smouse, P.E., 2012: GenAlEx 6.5, genetic analysis in Excel. Population genetic software for teaching and research an update. Bioinformatics 28, 2537–2539.

Peñuelas, J., Sardans, J., Filella, I., Estiarte, M., Llusià, J., Ogaya, R., Carnicer, J., Bartrons, M., Rivas-Ubach, A., Grau, O., Peguero, G., Margalef, O., Pla-Rabés, S., Stefanescu, C., Asensio, D., Preece, C., Liu, L., Verger, A., Barbeta, A., Achotegui-Castells, A., Gargallo-Garriga, A., Sperlich, D., Farré-Armengol, G., Fernández-Martínez, M., Liu, D., Zhang, C., Urbina, I., Camino-Serrano, M., Vives-Ingla, M., Stocker, B.D., Balzarolo, M., Guerrieri, R., Peaucelle, M., Marañón-Jiménez, S., Bórnez-Mejías, K., Mu, Z., Descals, A., Castellanos, A., Terradas, J., 2017: Impacts of global change on Mediterranean forests and their services. Forests 8, 463.

Pignatti, S., Guarino, R., La Rosa, M., 2017: Flora d'Italia. New Business Media, Bologna.

Piry, S., Luikart, G., Cornuet, J.-M., 1999: Computer note. BOT-TLENECK, a computer program for detecting recent reductions in the effective size using allele frequency data. The Journal of Heredity 90, 502–503.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000: Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Pritchard, J.K., Wen, X., Falush, D., 2010: Documentation for STRUCTURE software. Retrieved March 1, 2019 from http://pritch.bsd.uchicago.edu/structure.html

Rinaldi, R., Cafasso, D., Strumia, S., Cristaudo, A., Sebastiani, F., Fineschi, S., 2019: The influence of a relict distribution on genetic structure and variation in the Mediterranean tree, *Platanus orientalis*. AoB PLANTS 11: plz002.

Rivas-Martínez, S., Penas, A., Díaz, T.E., 2004: Bioclimatic and biogeographic maps of Europe. University of León. Retrieved March 15, 2018 from www.globalbioclimatics.org/form/maps.htm

Rushton, B.S., 1993: Natural hybridization within the genus *Quercus* L. Annals of Forest Science 50 (Suppl. 1993), 73s–90s.

Shi, X., Wen, Q., Cao, M., Guo, X., Xu, L., 2017: Genetic diversity and structure of natural *Quercus variabilis* population in China as revealed by microsatellites markers. Forests 8, 495.

Škvorc, Ž., Franjić, J., Idžojtić, M., 2005: Population structure of *Quercus pubescens* Willd. (Fagaceae) in Croatia according to morphology of leaves. Acta Botanica Hungarica 47, 193–206.

Slade, D., Škvorc, Ž., Ballian, D., Gračan, J., Papeš, D., 2008: The chloroplast DNA polymorphisms of White Oaks of section *Quercus* in the Central Balkans. Silvae Genetica 57, 227–234.

Slatkin, M., Barton, N.H., 1989: A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43, 1349–1368.

Sullivan, A.R., Lind, J.F., McCleary, T.S., Romero-Severson, J., Gailing, O., 2013: Development and characterization of genomic and gene-based microsatellite markers in north American red oak species. Plant Molecular Biology Reporter 31, 231–239.

Toonen, R.J., Hughes, S., 2001: Increased throughput for fragment analysis on an ABI PRISM 377 automated sequencer using a membrane comb and STRand software. Biotechniques 31, 1320–1324.

Trinajstić, I., 2007: About the problem of differentiation between the oaks *Quercus pubescens* Willd and *Quercus virgiliana* (Ten.). Šumarski list 131, 57–60.

Ueno, S., Tsumura, Y., 2008: Development of ten microsatellite markers for *Quercus mongolica* var. *crispula* by database mining. Conservation Genetics 9, 1083–1085.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P., 2004: MICRO-CHECKER, software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4, 535–538.

Weir, B.S., Cockerham, C.C., 1984: Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370.

Williams, J.H., Boecklen, W.J., Howard, D.J., 2001: Reproductive processes in two oak (*Quercus*) contact zones with different levels of hybridization. Heredity 87, 680–690.