ISSN 1022-7954, Russian Journal of Genetics, 2024, Vol. 60, No. 8, pp. 1022–1039. © Pleiades Publishing, Inc., 2024.

PLANT GENETICS

Differentiation and Taxonomic Identification of Roburoid Oaks in the Caucasian and Crimean Regions Using Nuclear Microsatellite Markers

S. A. Semerikova[a], *, Kh. U. Aliev[b], and V. L. Semerikov[a]

[a] *Institute of Plant and Animal Ecology, Ural Branch of Russian Academy of Science, Yekaterinburg, 620144 Russia*

[b] *Mountain Botanical Garden of Dagestan Federal Research Centre of the Russian Academy of Sciences, Makhachkala, 367000 Russia*

**e-mail: s.a.semerikova@ipae.uran.ru*

Received February 20, 2024; revised March 13, 2024; accepted March 25, 2024

Abstract—The inter- and intra-specific structure of genetic variability was studied using 18 microsatellite loci (nSSRs) in closely related roburoid white oaks in the Crimean-Caucasian region. The seven most widespread *Quercus* taxa in the region were studied in 29 morphologically pure populations from different parts of the North Caucasus, Transcaucasia, Crimea and northeastern Europe. Most taxa were studied using nSSR markers for the first time. Among the 492 trees studied, Bayesian clustering method implemented in STRUCTURE identified clusters corresponding to the pedunculate oak *Quercus robur*, the Hartwiss oak *Q. hartwissiana*, the Caucasian oak *Q. macranthera*, the downy oak *Q. pubescens* and three subspecies of sessile oak: *Q. petraea* ssp. *petraea*, *Q. petraea* ssp. *iberica*, *Q. petraea* ssp. *medwediewii*. Geographic structure was identified within *Q. robur*, *Q. pubescens* and *Q. p.* ssp. *petraea*. The 18 nSSR loci used are efficient in the taxonomic assignment of individuals, and identifying hybrids. The close relationship between the "long-peduncle" roburoid oaks (*Q. robur* and *Q. hartwissiana*) is shown, with a greater difference from other species. For one of the subspecies of sessile oak, widespread in the North Caucasus and Crimea *Q. petraea* ssp. *medwediewii* (syn. *Q. calcarea*), or limestone oak, significant differences from other taxa were found, reaching the interspecies level. The assumption of a possible hybrid origin of this taxon as a result of hybridization of *Q. petraea* and *Q. pubescens* is not confirmed by genetic analysis. The other two subspecies of *Q. petraea* (*Q. p.* ssp. *petraea* and Georgian durmast oak *Q. p.* ssp. *iberica*) are differentiated to a lesser extent and are related to each other, which confirms the legitimacy of distinguishing two geographically isolated taxa at the rank of subspecies. The highest variability was observed in *Q. pubescens* ($H_e = 0.777$). In *Q. p.* ssp. *medwediewii* variability was lower than in other widespread taxa ($H_e = 0.652$), and was approximately at the level of variability of *Q. hartwissiana* ($H_e = 0.633$) and *Q. macranthera* ($H_e = 0.659$). Clear differentiation of taxa by nuclear markers shows the limited introgression in closely related oak species in the Caucasus and Crimea. The identified genetic clusters can be used as reference groups for further population genetic studies of oaks in the Crimean-Caucasian region.

Keywords: *Quercus* spp., nuclear microsatellites, Caucasus, Crimea, nSSR, STRUCTURE, genetic differentiation, taxon assignment

DOI: 10.1134/S1022795424700492

INTRODUCTION

European roburoid oaks or European white oaks (*Quercus* L., section *Quercus*, Fagaceae) are important elements of temperate forests with centers of diversity located in the southeastern regions of Western Eurasia, including the Black Sea region, the Caucasus and Western Asia [1–3]. This is a young group [4, 5], containing a large number of morphologically poorly differentiated, capable for hybridization species with complex taxonomy. In the Caucasus and Crimea, the number of taxa in this group varies significantly depending on the classification [1–3, 6–10]. Currently, the Caucasian flora conspectus [6] indicates six species that belong to roburoid oaks, among which the most common European species are pedunculate oak (*Q. robur* L.) and sessile oak (*Q. petraea* (Matt.) Liebl.). Hartwiss oak (*Q. hartwissiana* Steven) is found in the Western Caucasus and Transcaucasia. Caucasian oak or Persian oak (*Q. macranthera* Fisch. et C.A. Mey. ex Hohen) grows in the mountainous regions of the Eastern Caucasus and Transcaucasia. Downy oak *Q. pubescens* Willd., a Mediterranean xerophytic species, is widespread in Crimea. In the Caucasus, its range is limited to the coastal regions of the Northwestern and Eastern Caucasus.

In addition to the generally recognized species, among the oaks growing in the Caucasus, the intraspecific taxa of *Q. petraea* deserve special attention, the taxonomy of which causes the greatest controversy ([7], review). The monographer of Caucasian oaks Yu.L. Menitsky [1–3, 6] considered several taxa within *Q. petraea* in the rank of subspecies, of which the most common in the region are the type subspecies of sessile oak (*Q. petraea* L. ex Liebl. ssp. *petraea*), Georgian durmast oak (*Q. petraea* ssp. *iberica* (Steven ex Bieb.) Krassiln.) and "limestone" oak (*Q. petraea* ssp. *medwediewii* (A. Camus) Menitsky). The latter was put into *Q. petraea* system, but with an indication of the intermediate morphological characteristics between *Q. petraea* and *Q. pubescens*, which suggests, according to Yu.L. Menitsky, its hybrid origin [1–3]. At the same time, N.D. Troitsky [9] previously described this taxon in Crimea in the species rank (*Q. calcarea* Troitsky).

Subspecies of sessile oak, with clear ecological differentiation, occupy different, although overlapping, areas in the Caucasus: *Q. p.* ssp. *medwediewii* predominates on the northern macroslope of the Greater Caucasus and in the Ciscaucasia, *Q. p.* ssp. *iberica*—in Transcaucasia and Dagestan, *Q. p.* ssp. *petraea*—in the Crimea and Western Ciscaucasia [3], while in some places vast areas of "intergradation" [1, 8], or introgressive hybridization between subspecies, as well as with other closely related species, are formed: these are primarily the regions of the Western Caucasus and Western Transcaucasia, as well as the Eastern Caucasus and Crimea.

Researchers in oak taxonomy have noted the difficulty of distinguishing taxa due to the significant overlap of diagnostic traits and the need to distinguish taxa using many characters, including pubescence structure, which is labor-intensive and not always available [1, 11–18]. Even so, for a confident diagnosis it is necessary to know the origin of the sample [1]. In addition to high morphological variability, hybridization poses a significant problem, which is facilitated by wide overlapping ranges and sympatric habitats. All this determines the need to use molecular markers to study the taxonomic structure and demographic processes in oak populations.

The relationships of the most common European Roburoid oaks have previously been extensively studied using nuclear microsatellite loci (nSSRs) [11–29]. Using different sets of nSSRs, the possibility of distinguishing some taxa (*Q. robur*, *Q. petraea*, *Q. pubescens*, *Q. frainetto* and *Q. pyrenaica*) was shown. In mixed oak stands, significant differences in the level of hybridization between populations have been found, depending on composition [11, 13–16, 24–26]. At the same time, it was impossible to distinguish several species close to *Q. pubescens* [18, 30–32]. For more accurate taxonomic identification, it is necessary to take into account all taxa present in the regions [24, 28], use "pure" reference groups when studying hybridization and composition of stands [18, 23, 24, 30] and a balanced number of samples and populations of different taxa included into analysis [33].

Our recent studies of five roburoid oak species in Crimea and the Caucasus using maternally inherited chloroplast DNA markers [34, 35] revealed a distinct geographic structure in the absence of taxonomic subdivision, which is explained by the presence of gene flow between related species. At the same time, 14 nSSR loci well separated *Q. robur* and *Q. petraea* s. l. in accordance with species affiliation in all studied areas of the eastern part of the range and in sympatric populations [36].

The variability of other *Quercus* species and their taxonomic structure in the Crimean-Caucasian region previously have not been studied using nuclear markers. In this work, we analyze for the first time the structure of variation of nuclear multilocus markers (nSSR) of all the most common taxa of roburoid oaks in the Caucasus and Crimea in order to identify the main genetic clusters and compare them with the taxonomic division previously proposed on the basis of morphological data [1, 2, 6]. The objectives of this study were: (1) testing the possibility of reliable species delimitation based on nSSR; (2) study of phylogenetic relationships between taxa and the nature of intraspecific differentiation; (3) clarification of the taxonomic status of controversial taxa. For this purpose, morphologically pure populations of seven studied species and subspecies, including three subspecies of *Q. petraea*, were selected in geographically remote regions of the North Caucasus, Transcaucasia, Crimea and North-Eastern Europe, and the variability of 18 nSSR loci was studied.

MATERIALS AND METHODS

The seven most common *Quercus* taxa in the Crimean-Caucasian region, taking into account three subspecies of sessile oak *Q. petraea* identified according to Yu.L. Menitsky [3, 6], were studied. In sessile oak *Q. petraea* s. l. 215 samples from 12 populations were studied, including 71 samples of *Q. petraea* ssp. *petraea* (two populations from the Kaliningrad region and two from Crimea), 64 samples from populations attributed to the durmast oak *Q. petraea* ssp. *iberica* (three locations from Transcaucasia, one from Dagestan), and 80 samples from four populations of limestone oak *Q. petraea* ssp. *medwediewii* from the Ciscaucasia, Central Caucasus and Crimea (Table 1; Fig. 1a). Some samples from three populations of *Q. petraea* s. l. (nos. 1, 5, 8) were used in our previous study of variation in *Q. robur* using 14 nSSR loci [36]. Downy oak *Q. pubescens* (99 samples) was studied in two geographically distant regions: three populations from the Crimea (59 trees) and two from the Eastern Caucasus – Dagestan (40 trees). Fifty trees of the Caucasian oak *Q. macranthera* from two populations of Foothill

Dagestan and Inner Dagestan and two localities of the Lesser Caucasus (Armenia) were studied. In the pedunculate oak *Q. robur*, the analysis included five populations from different botanical and geographical regions of the North Caucasus, previously studied for 14 loci and belonging to different genetic clusters of *Q. robur* [36]. In the present study, specimens with available data for 18 loci were analyzed, which slightly reduced the sample size compared to previous work [36]. In the Hartwiss oak *Q. hartwissiana*, 33 trees from three populations of the Western Caucasus were studied, two of which are located on the southern macroslope and one on the northern (Table 1; Fig. 1a). Thus, the number of samples per taxon ranged from 33 (*Q. hartwissiana*) to 99 (*Q. pubescens*), with an average of 70.3 specimens.

For our study, we followed an approach that involved collecting "pure" taxa unaffected by hybridization, primarily from areas where there are no closely related species or in populations where they do not grow sympatrically. The distribution areas of the studied taxa in the Caucasus and Crimea were considered according to Yu.L. Menitsky [3] and N.D. Troitsky [9]. Samples of *Q. pubescens* were taken from stands lacking *Q. petraea*. The material was collected in natural habitats, with a distance between trees of at least 50 m. The population samples averaged 17 individuals, in most cases more than 15 trees. In some cases, sample sizes were smaller in hard-to-reach areas (Georgia, Armenia) or for rare species (*Q. hartwissiana*). Preliminary taxonomic affiliation was established according to descriptions [3, 6], taking into account morphological characteristics: length of the peduncle, leaf shape and petiole length, features of the pubescence of the leaf and shoots. Voucher samples are stored in the Laboratory of Molecular Ecology of Plants (Institute of Plant and Animal Ecology, Russian Academy of Sciences). The names of taxa (in the rank of species or subspecies) are given according to the Caucasian flora conspectus, 2012 [6]. Genomic DNA was isolated using the CTAB method [37] from leaves dried in silica gel. DNA samples from our previous studies of chloroplast DNA were also used [34, 35].

Analysis of the variability of 18 nSSR loci was carried out in accordance with that described previously [36] (Table S1, Supplementary materials), using two multiplexes [20], excluding the PIE223 and PIE258 loci, since nonspecific amplification products were observed at these loci in our study of *Q. robur* and *Q. petraea* [36]. Chromatograms were converted to genotypic data using GeneMapper 3.5 (Applied Biosystems). Individuals with missing data at more than four loci were excluded from further analysis. It should be noted that when changing the brand of the size standard, in some loci there was a significant shift in the determined allele sizes (from 1 to 4 bp). Bias was identified and corrected using reference samples with known allele sizes, including 2–4 reference samples in each analyzed 96-well PCR plate. Also, compared to the previous study, the composition of the multiplex kit-1 was optimized: for the locus PIE215, the TAMRA fluorescent dye was replaced by FAM (Table S1, Supplementary materials), which improved its amplification and eliminated the mixing of TAMRA and R6G signals. Replacing the fluorescent dye led to an underestimation of allele sizes by 5 bp, which was also corrected using reference samples and re-analysis of some previously analyzed samples with a new dye. The resulting data (allele sizes) were adjusted to the data set obtained for *Q. robur* and *Q. petraea* [36].

Bayesian clustering of genotypes was carried out using STRUCTURE 2.3.4 [38], applying the Admixture model and Correlated Allele Frequencies model. The K value was set from 1 to 12 or from 1 to 10, depending on the sample set (see Results). 10 runs of the STRUCTURE program were performed for each value of K. Each run included a burn-in of 50,000 and subsequent 100000 iterations without prior information on the population of origin. The optimal number of clusters was assessed using the ΔK method [39] and the $LnP(D)$ method [38] via the StructureSelector [40]. The averaging of multiple runs was carried out using the Clumpak [41].

STRUCTURE analysis was performed on several data sets. Set 1 included 492 samples from 29 populations of all taxa studied. The optimal number of K according to the ΔK method in this case was equal to two (see Results), while the clusterization corresponded to the division into two groups of species—"sessile-flower" (*Q. pubescens*, *Q. macranthera*, *Q. petraea* s. l.) and "long-peduncle" (*Q. robur*, *Q. hartwissiana*). Set 2 STRUCTURE included "sessile flowers" oaks (364 samples, 21 populations) and set 3 STRUCTURE—"long-peduncle" *Q. robur* and *Q. hartwissiana* (128 samples, 8 populations).

It is known that low representation of individual taxa or populations in the data set studied using STRUCTURE leads to erroneous assignment of the corresponding individuals to the wrong taxon or hybrid [42]. To investigate the importance of the presence in the STRUCTURE analysis of reference groups that correspond to all represented taxa, we conducted different variants of analysis ("simulated" options) that modeled the possible omission of individual taxa.

"Simulated" option 1—absence of reference groups of *Q. p.* ssp. *medwediewii*: all samples of *Q. p.* ssp. *medwediewii* were removed from the STRUCTURE analysis, but four samples with >90% of the *Q. p.* ssp. *medwediewii* cluster were left. "Simulated" option 2—absence of the *Q. hartwissiana* reference group: all samples of *Q. hartwissiana* were removed from the analysis apart from four samples with 90% of the *Q. hartwissiana* cluster.

To examine the ability to differentiate taxa with fewer loci, STRUCTURE analyzes were performed

Table 1. Characteristics of the studied populations of *Quercus* species

No.	Populations	Latitude	Longitude	Altitude, m	*n*	N_a	N_e	H_o	H_e	*F* (18 loci)	*p*-value (18 loci)	*F* (15 loci)	*p*-value (15 loci)
	Q. petraea ssp. *petraea*												
1	Svetlogorsk	54°55′56″	20°08′35″	50	15	7.278	4.499	0.769	0.725	–0.062	0.7741	–0.070	0.5927
2	Skrytnoe Lake	54°45′24″	20°14′24″	64	18	8.278	4.923	0.719	0.722	0.031*	0.2129	–0.023	0.7832
3	Kastel Mt.	44°38′52″	34°23′40″	120	18	8.444	4.749	0.719	0.725	0.019*	<0.001	–0.061	0.4008
4	Seraus Mt.	44°38′38″	34°22′09″	395	20	8.500	4.650	0.706	0.727	0.039*	0.0338	0.007	0.1757
	Mean (SE)				17.8	8.125	4.705	0.728	0.725	0.007	<0.001	–0.037	0.4911
						0.409	0.276	0.025	0.018	0.025		0.017	
	Q. petraea ssp. *iberica*												
5	Tskhinval	42°14′57″	43°59′47″	1200	24	8.889	4.799	0.679	0.697	0.019*	0.0830	0.008	0.9104
6	Orbeli	42°38′51″	42°48′52″	951	8	5.444	3.768	0.660	0.651	–0.015	0.5540	–0.036	0.6414
7	Borjomi	41°46′18″	43°28′22″	1473	8	6.111	4.110	0.688	0.683	–0.021	0.8621	–0.030	0.9158
8	Erpeli (Pt)	42°47′21″	46°58′40″	946	24[a]	9.944	5.366	0.771	0.757	–0.009	0.0022	–0.045	0.2895
	Mean (SE)				16	7.597	4.511	0.699	0.697	–0.007	0.0433	–0.026	0.9089
						0.405	0.268	0.025	0.024	0.016		0.016	
	Q. macranthera												
9	Termenlik	42°44′59″	46°59′70″	1400	23	8.389	4.487	0.669	0.670	–0.011	0.0558	–0.011	0.5213
10	Rosnob	42°02′44″	46°25′08″	1710	17	7.000	4.167	0.689	0.671	–0.034	0.5014	–0.038	0.6214
11	Dilijan	40°42′04″	44°50′49″	1587	6[a]	5.500	3.852	0.654	0.638	–0.030	0.6156	–0.034	0.5736
12	Vaik[b]	39°41′42″	45°34′25″	1391	4[b]	–	–	–	–	–	–	–	–
	Mean (SE)				12.5	6.963	4.169	0.671	0.659	–0.025	0.2273	–0.028	0.7384
						0.513	0.350	0.030	0.029	0.020		0.022	
	Q. petraea ssp. *medwediewii*												
13	Russkaya lesnaya dacha	45°02′13″	41°52′10″	671	17	6.944	3.905	0.654	0.640	–0.016	0.3964	–0.026	0.5586
14	Balta	42°55′12″	44°38′21″	1000	17[a]	6.722	3.862	0.623	0.649	0.036	<0.013	0.004	0.6581
15	Beshtau	44°07′16″	43°03′44″	529	23	7.500	3.875	0.635	0.654	0.022	0.0478	–0.032	0.6294
16	Agarmish Mt.	45°01′10″	35°01′53″	350	23	8.278	3.949	0.638	0.662	0.038	0.0910	0.014*	0.0925
	Mean (SE)				20	7.361	3.898	0.637	0.652	0.020	0.0025	–0.010	0.4492
						0.358	0.242	0.027	0.025	0.018		0.017	

Table 1. (Contd.)

No.	Populations	Latitude	Longitude	Altitude, m	*n*	N_a	N_e	H_o	H_e	*F* (18 loci)	*p*-value (18 loci)	*F* (15 loci)	*p*-value (15 loci)
	Q. pubescens												
17	Novyi Svet	44°50′17″	34°55′03″	200	18	10.278	6.324	0.768	0.793	0.028*	<0.001	0.013	0.3621
18	Arpat	44°51′27″	34°43′16″	260	22	10.889	6.686	0.813	0.797	–0.017	0.4689	–0.047	0.7460
19	Koreiz trail	44°27′23″	34°04′18″	275	19	9.944	6.015	0.728	0.783	0.071*	<0.001	0.042*	0.0967
20	Chipcharak Mt.	43°07′29″	46°59′49″	290	16	8.500	4.604	0.764	0.755	–0.014	0.1633	–0.034	0.3348
21	Zidyan-Kazmalyar	42°05′18″	48°09′30″	160	24	8.944	4.818	0.770	0.754	–0.021	0.6753	–0.055	0.9940
	Mean (SE)				19.8	9.711	5.689	0.769	0.777	0.009	<0.001	–0.016	0.6261
						0.357	0.266	0.016	0.014	0.013		0.011	
	Q. robur												
22	Samur	41°50′16″	48°31′19″	7	23	7.611	4.115	0.653	0.700	0.080*	<0.001	–0.010	0.0995
23	Erpeli (Rb)	42°47′21″	46°58′40″	946	10	6.500	4.045	0.677	0.695	0.029*	0.0942	–0.068	0.9495
24	Budennovsk	44°45′23″	44°09′55″	93	20	7.333	4.159	0.728	0.710	–0.030	0.2694	–0.063	0.6869
25	Semigorsky	44°53′53″	37°36′10″	100	22	8.389	4.835	0.733	0.732	–0.005	0.3263	–0.032	0.7763
26	Ubinskaya	44°42′20″	38°31′27″	133	20	8.222	4.471	0.673	0.713	0.051*	<0.001	–0.015	0.5671
	Mean (SE)				19	7.611	4.325	0.693	0.710	0.025	<0.001	–0.038	0.7577
						0.299	0.204	0.023	0.016	0.023		0.014	
	Q. hartwissiana												
27	Mzymta	43°29′00″	39°59′26″	80	8	5.000	3.393	0.639	0.627	–0.028	0.7018	–0.027	0.6856
28	Anastasievka	44°09′55″	39°16′02″	120	7	4.722	3.327	0.634	0.646	0.029*	0.9048	0.034*	0.8795
29	Perevalka	44°02′27″	40°45′30″	700	18	6.500	3.626	0.633	0.626	–0.011	<0.065	–0.021	<0.878
	Mean (SE)				11	5.407	3.448	0.635	0.633	–0.003	<0.541	–0.005	<0.538
						0.339	0.234	0.033	0.027	0.029		0.034	
	Mean, all species (SE)					7.716	4.478	0.696	0.700	0.006	<0.001	–0.023	0.8785
						0.152	0.103	0.010	0.008	0.008		0.007	

n—sample size; N_a—average number of alleles; N_e—effective number of alleles; H_o—observed heterozygosity; H_e—expected heterozygosity; *F*—inbreeding coefficient; *P*-value—results of the "Probability" test for deviation from the Hardy–Weinberg equilibrium.

* Deviation from the Hardy–Weinberg equilibrium according to the results of the "Global" test ($P < 0.05$).

[a] Variation measures were calculated after removing a specimen that did not match the population taxon assignment (see Results).

[b] Due to the small sample size the Vaik population was combined with the Dilijan population to calculate the variability parameters

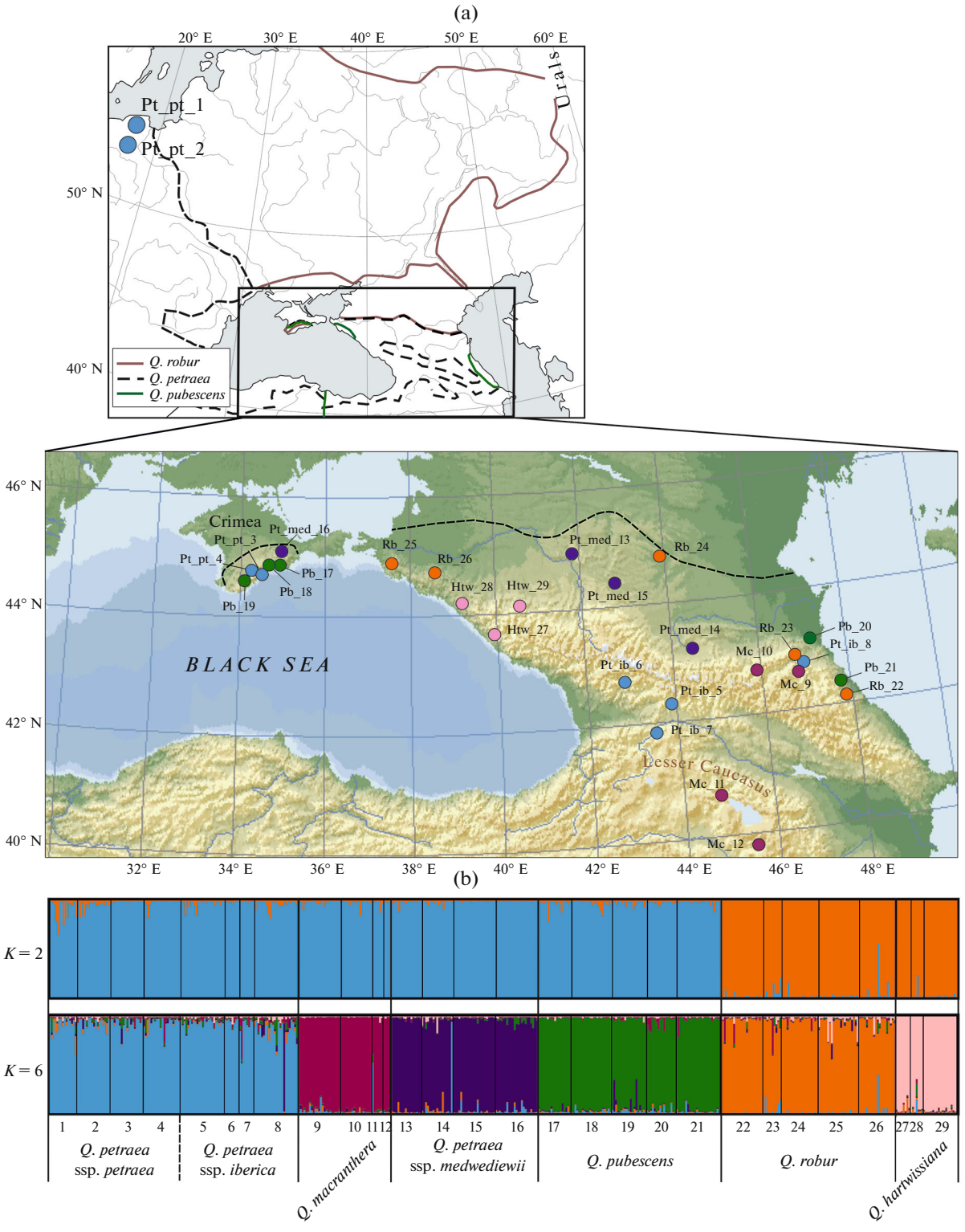

Fig. 1. (a) Geographical distribution of the studied populations of seven roburoid oaks taxa in the Crimean-Caucasian and North European parts of the range. The population numbers correspond to those indicated in the Table 1. Designation of taxa: Rb—*Q. robur*, Pb—*Q. pubescens*, Mc—*Q. macranthera*, Htw—*Q. hartwissiana*, Pt_pt—*Q. petraea* ssp. *petraea*, Pt_ib—*Q. petraea* ssp. *iberica*, Pt_med—*Q. petraea* ssp. *medwediewii*. The color coding of populations corresponds to the predominant STRUCTURE clusters at $K = 6$ (Fig. 1b). The dotted line marks the northern limit of the oak's range in the region. (b) Distribution of genetic clusters calculated by STRUCTURE based on 18 microsatellite loci for 29 populations (492 individuals) of five *Quercus* species at $K = 2, 6$. Each sample is represented by a vertical column divided into two or six colored segments proportional to the contribution of each genetic cluster.

Table 2. Characteristics and variability of 18 microsatellite loci

Locus	Size, bp	*A*	N_a	N_e	H_o	H_e	F_{IS}	*P*-value	F_{ST}
kit-1									
PIE020	97–127	13	5.214	3.193	0.708	0.658	–0.076*	0.842	0.137
PIE152	233–281	31	10.571	6.197	0.846	0.827	–0.023	0.958	0.088
PIE242	107–135	15	8.571	5.644	0.821	0.817	–0.005*	0.174	0.082
PIE102	137–181	23	7.679	3.755	0.670	0.671	0.003	0.776	0.135
PIE243	201–241	21	6.964	3.389	0.690	0.657	–0.051	0.770	0.204
PIE267	90–120	15	6.071	3.816	0.723	0.715	–0.012	0.594	0.147
PIE215	192–225	13	6.571	4.299	0.740	0.733	–0.009	0.480	0.119
PIE239	86–104	16	3.964	2.038	0.270	0.405	0.334*	<0.001	0.303
PIE227	154–181	10	4.143	2.272	0.532	0.508	–0.047	0.812	0.275
PIE271	186–212	17	7.107	3.903	0.732	0.711	–0.030	0.676	0.110
Mean		17.4	6.686	3.851	0.673	0.670	0.008		0.160
kit-2									
QrZAG7	116–162	24	11.786	7.906	0.893	0.853	–0.046	0.962	0.081
MsQ13	188–250	26	8.179	4.953	0.796	0.763	–0.043	0.612	0.135
QrZAG112	77–121	22	6.536	3.712	0.569	0.552	–0.030*	0.031	0.251
QrZAG20	157–197	23	9.250	5.791	0.808	0.803	–0.006*	0.159	0.093
QrZAG96	143–189	24	9.464	5.521	0.655	0.715	0.084*	<0.001	0.210
QrZAG11	231–285	50	11.071	6.102	0.689	0.813	0.153*	<0.001	0.139
QpZAG15	107–147	19	6.250	2.934	0.613	0.618	0.008	0.666	0.188
QpZAG110	202–256	27	9.500	5.173	0.770	0.781	0.014	0.257	0.084
Mean		26.9	9.005	5.262	0.724	0.737	0.017		0.148
Mean over all loci		21.6	7.716	4.478	0.696	0.700	0.012	<0.001	0.155
SE			0.152	0.103	0.010	0.008	0.023		0.016
Total		389							

A—total number of alleles; N_a—average number of alleles per population; N_e—average effective number of alleles per population; H_o—average observed heterozygosity per population; H_e—average expected heterozygosity per population; F_{IS}, F_{ST}—fixation indices; *P*-value – results of the "Probability" test for deviation from the Hardy–Weinberg equilibrium.
* Deviation from the Hardy–Weinberg equilibrium according to the results of the "Global" test ($P < 0.05$).

separately for two sets of loci (multiplexes): ten loci (kit-1) and eight loci (kit-2), for all 492 individuals.

Specimens (three individuals in total) that were not assigned by STRUCTURE to the taxon of the particular population (see Results) were excluded from the analysis of variability and taxon differentiation. Using the GenAlEx v 6.5 [43], the following statistics were calculated: total number of alleles (*A*), number of alleles per locus (N_a), effective number of alleles (N_e), observed H_o and expected H_e heterozygosity, inbreeding coefficient *F* for each population, fixation indices (F_{IS}, F_{ST}) for each locus (Tables 1, 2). Deviations from the Hardy–Weinberg equilibrium across populations and loci were assessed using GENEPOP [44], using the "Probability" test, as well as the "Global" test with the alternative hypothesis H1 = lack of heterozygotes/excess of heterozygotes. The presence of null alleles at loci was assessed using the MICRO-CHECKER [45].

Using GenAlEx, a hierarchical analysis of molecular variation (AMOVA) was performed for all taxa and populations, pairwise inter-population F_{ST} were calculated, and principal coordinate analysis (PCoA) was performed based on Nei genetic distances [46]. To visualize the degree of genetic differences between populations and taxa, a dendrogram was constructed based on a matrix of pairwise inter-population F_{ST} using the unweighted pair-group average method (UPGMA) in the NTSYS-pc software [47].

RESULTS

Variability of Microsatellite Loci

All *Quercus* taxa examined showed successful amplification at all loci. A total of 389 alleles were

found in 29 populations at 18 loci. Alleles number per locus varied from 10 (PIE227) to 50 (QrZAG11), with an average of 21.6 per locus (Table 2). The effective number of alleles (N_e) varied from 2.038 (locus PIE239) to 7.906 (QrZAG7), with an average per locus of 4.478. Most loci showed high variability in all species, with the exception of the PIE239 locus, which was almost monomorphic in *Q. macranthera* and *Q. hartwissiana*. The allele sizes in loci (Table 2; Table S2, Supplementary materials) generally correspond to those given in [20] and our previous study of *Q. robur* [36]. However, it should be noted that these sizes determined using capillary electrophoresis (without sequencing, only based on a size standard) can vary significantly between studies, including depending on the size standard used and the fluorescent dye of a given locus (see Material and Methods).

At some loci, alleles that did not correspond to the motif length were found in small numbers. Most of these were observed in QrZAG11, which led to a large number of alleles at this locus. Most loci lacked non-specific products and additional alleles. It should be noted that in 90% of the *Q. macranthera* samples within the range of the MsQ13 locus, a 192 bp product was initially considered to be an allele, but since the product was in most cases observed in combination with a heterozygote for the other two alleles, it was excluded from the analysis as non-specific, after which no deviations from the Hardy–Weinberg equilibrium were observed in *Q. macranthera* at this locus.

The presence of null alleles was detected in the QrZAG11, PIE239 and QrZAG96 loci in ten, nine and five populations, respectively. For the PIE242 locus, the presence of a null allele was determined in two populations, and for six more loci—in one population. The loci QrZAG11, PIE239 and QrZAG96 showed a significant deviation from the Hardy–Weinberg equilibrium towards a lack of heterozygotes (Table 2). When counted for individual taxa, these loci deviated from equilibrium in most cases (Table S2, Supplementary materials). In addition to the three loci with the maximum deviation, in the overall analysis of all populations, the QrZAG112 locus also deviated from equilibrium, but when calculated for individual taxa, the deviation of the QrZAG112 locus was significant only for *Q. p.* ssp. *iberica* and *Q. hartwissiana*. In some cases, the "Probability" test showed a significant deviation from the Hardy–Weinberg equilibrium: for *Q. p.* ssp. *medwediewii*—in loci PIE242 and QrZAG110, *Q. p.* ssp. *petraea*—at locus PIE227, *Q. macranthera*—QrZAG20, *Q. pubescens*—MsQ13, *Q. hartwissiana*—PIE242 (Table S2). For a total of 18 loci and for all populations, the deviation from equilibrium was significant both across all populations (Table 1) and for most taxa (Table S3, Supplementary Information). The "Probability" test, when counting 18 loci, showed a significant deviation for eight populations. When conducting the "Global" test, a significant lack of heterozygotes was found in 10 populations (Table 1).

The test was repeated without the three loci at which the greatest deviation from equilibrium was observed (PIE239, QrZAG96, QrZAG11). For 15 loci, the "Probability" test did not show a significant deviation in any of the populations (Table 1). The "Global" test for heterozygote deficiency was significant for three populations of different taxa (nos. 16, 19, 28). In total, for all loci and for all populations, no significant deviation from equilibrium was observed ($F = -0.023$, $P = 0.8785$) (Table 1).

Thus, the populations of the taxa under study, as shown by 15 nSSR loci, are in equilibrium, which indicates their taxonomic homogeneity. The significant lack of heterozygotes at the PIE239, QrZAG96 and QrZAG11 loci was presumably caused by the presence of null alleles at these loci, which has also been noted for these loci in other studies [14, 23, 30, 32, 36]. However, it was shown [23] that the presence of null alleles with a low frequency, including for the QrZAG11 and QrZAG96 loci, does not affect the accuracy of taxonomic determination. In our case, only a small part of the populations had a significant content of null alleles; in addition, the F_{ST} subdivision value was higher than the mean value for two (PIE239, QrZAG96) of the three disequilibrium loci, i.e. they made a significant contribution to the differentiation of taxa. As a result, it was decided to leave the PIE239, QrZAG96, and QrZAG11 loci in further analysis.

Variation rates in a multiplex of ten EST-SSR loci (kit-1) were on average lower than those in eight genomic SSR loci (kit-2) (Table 2), but the F_{ST} subdivision value was on average consistent between multiplexes. In a combined analysis of all populations, F_{ST} was significant for all loci and varied from a minimum $F_{ST} = 0.081$ (locus QrZAG7) to a maximum $F_{ST} = 0.303$ (PIE239), with a mean across loci $F_{ST} = 0.155$ (Table 2). F_{ST} values above the average were found in each of the two multiplexes, at loci PIE243 (0.204), PIE239 (0.303), PIE227 (0.275), QrZAG112 (0.251), QrZAG96 (0.210) and QrZAG15 (0.188).

The average expected heterozygosity varied between taxa from $H_e = 0.633$ (*Q. hartwissiana*) to $H_e = 0.777$ (*Q. pubescens*) (Table 1). The highest diversity was observed in populations of downy oak, varying from 0.754 to 0.797 and reaching maximum values in the Crimean populations of *Q. pubescens*. In the type subspecies of sessile oak *Q. p.* ssp. *petraea* also showed high diversity, changing little between populations (from 0.722 to 0.727), which was higher than that of the Transcaucasian subspecies *Q. p.* ssp. *iberica*. In the latter, it varied significantly, from high in Dagestan (0.757) to reduced in Transcaucasia (0.697). In *Q. macranthera*, compared to sessile oak, diversity was reduced, $H_e = 0.659$ (0.638–0.671). In the studied Caucasian populations of *Q. robur*, the average expected heterozygosity was $H_e = 0.710$ (0.695–0.732). In *Q. hartwissiana* the diversity was lower than

in *Q. robur*, H_e = 0.633 (0.626–0.646). In *Q. p.* ssp. *medwediewii*, the diversity measures were lower than those of other widespread taxa of the Caucasus, H_e = 0.652 (0.640–0.662), and were approximately at the level of *Q. hartwissiana* and *Q. macranthera*.

Taxonomic and Geographical Structure of Variability

Alleles are largely shared across taxa (Table S2, Supplementary Information). The most common alleles at loci are commonly found in all taxa, but at different frequencies. We did not find diagnostic alleles or loci that could be used to unambiguously identify a taxon or separate one taxon from another. Taxon-specific alleles occurred at low frequencies. Several loci showed single-allele dominance and low diversity in some species, while being highly variable in other species. For example, at the PIE239 locus (with the maximum F_{ST} value), the "88" allele occurs in all taxa with a frequency varying from 0.136 (*Q. pubescens*), 0.351 (*Q. robur*) and 0.411 (*Q. p.* ssp. *medwediewii*) to 0.980 and 0.970 in *Q. macranthera* and *Q. hartwissiana*. In the QrZAG96 locus, the frequency of the "short" allele "143" varied from 0.601 (*Q. robur*) and 0.576 (*Q. hartwissiana*) to 0.127 (*Q. p.* ssp. *petraea*) and 0.123 (*Q. p.* ssp. *iberica*), reaching almost zero values for *Q. p.* ssp. *medwediewi* (0.025), *Q. pubescens* (0.015) and *Q. macranthera* (0.0). Interestingly, in *Q. macranthera* the QrZAG96 locus is dominated by the "145" allele (frequency 0.806), which is rare in other taxa.

Based on AMOVA results, the level of genetic differentiation F_{RT} between the seven taxa was 0.102 (P = 0.001), which was 3 times higher than the differentiation within taxa (F_{SR} = 0.032, P = 0.001). Within taxa (Table S2, Supplementary Information), the maximum differentiation between populations was found in *Q. robur*, F_{ST} = 0.077, significant for 17 loci. Also differentiated populations of *Q. pubescens* (F_{ST} = 0.042), *Q. p.* ssp. *petraea* (F_{ST} = 0.051), *Q. hartwissiana* (F_{ST} = 0.059), *Q. p.* ssp. *iberica* (F_{ST} = 0.041). The least subdivided taxa are *Q. macranthera* (F_{ST} = 0.031) and *Q. p.* ssp. *medwediewii* (F_{ST} = 0.032), in which for most loci the F_{ST} values are not significant between populations.

The UPGMA dendrogram of the studied populations (Fig. S1, Supplementary materials), constructed on the basis of pairwise F_{ST}, is in good agreement with the results of the STRUCTURE analysis, which are reflected in the form of frequency diagrams of the corresponding clusters for each population (at K = 6, set 1, see below). The deepest divergence is observed between clade I (*Q. robur* and *Q. hartwissiana*, cycle *Pedunculatae* Maleev [1], species with a long peduncle, hereinafter referred to as "long-peduncle" taxa) and clade II (*Q. p.* ssp. *petraea-iberica*, *Q. p.* ssp. *medwediewii*, *Q. pubescens* of the cycle *Sessiliflorae* Maleev, as well as *Q. macranthera*, species with a short peduncle, hereinafter referred to as "sessile-flower" taxa), with subsequent division into subclades of individual taxa. The average values of pairwise F_{ST} between taxa are given in Table S3, Supplementary materials. The pairwise F_{ST} values between taxa of clade II and *Q. robur* average 0.106, with *Q. hartwissiana*—0.130. Within clade II, pairwise F_{ST} values between populations of different taxa are lower: on average 0.081 (from 0.067 to 0.106). Within taxa pairwise inter-population F_{ST} values were less than 0.05.

The differences between taxa are clearly illustrated by the ordination of populations based on genetic distances (Fig. 2). According to the first component, populations of "sessile-flower" taxa are differentiated from "long-peduncle" taxa; according to the second, more closely related taxa are subdivided (Fig. 2a). When considering only the group of "sessile-flower" oaks on PCoA (Fig. 2b), *Q. p.* ssp. *medwediewii* turns out to be the most differentiated, *Q. pubescens* is also separated by the first component, then *Q. macranthera* is separated from *Q. petraea* and the differentiation of *Q. pubescens* into western and eastern groups can be traced. In PCoA analyzes limited to only populations of *Q. petraea* without *Q. p.* ssp. *medwediewii* (Fig. S2, Supplementary Information), the subspecies *Q. p.* ssp. *petraea* and *Q. p.* ssp. *iberica* differ in the first component and *Q. p.* ssp. *petraea* is divided into European and Crimean populations according to the second.

STRUCTURE analysis for 18 nSSR loci of all studied taxa of roburoid oaks was carried out for 492 samples (set 1). With a complex taxonomic composition with varying degrees of relatedness, STRUCTURE analysis reveals a hierarchical structure: with the number of clusters K = 2 (maximum value of ΔK, Fig. S3a), the taxa under study are divided into two groups: the combined cluster of *Q. robur* and *Q. hartwissiana* (orange) and the combined cluster of remaining taxa (blue) (Fig. 1b). At K = 3, 4, 5, the division into clusters between runs was unstable, with individual taxa (*Q. p.* ssp. *medwediewii*, *Q. pubescens*, *Q. macranthera*) alternately separating from the rest. The second peak of ΔK was observed at K = 6 (Fig. S3a). With this number of clusters, a high value of $LnP(K)$ was also achieved (Fig. S3b) and, accordingly, the number of clusters equal to six was chosen as optimal. Figure 1b shows the results of the STRUCTURE analysis for the number of clusters K = 2 and K = 6. At K = 6, genetic clusters are identified corresponding to the following species, subspecies and groups: (1) a cluster common to two subspecies of *Q. petraea* ssp. *petraea* and *Q. p.* ssp. *iberica* (blue cluster); (2) *Q. macranthera* cluster (purple); (3) cluster *Q. petraea* ssp. *medwediewii* syn. *Q. calcarea* (dark violet); (4) *Q. pubescens* cluster (green); (5) *Q. robur* cluster (orange); (6) *Q. hartwissiana* cluster (pink). At K = 6, a minor configuration is also revealed (minor, 5/10, data not shown), in which the blue cluster is divided into two in accordance with the subspecies—*Q. petraea* ssp. *petraea* and *Q. petraea*

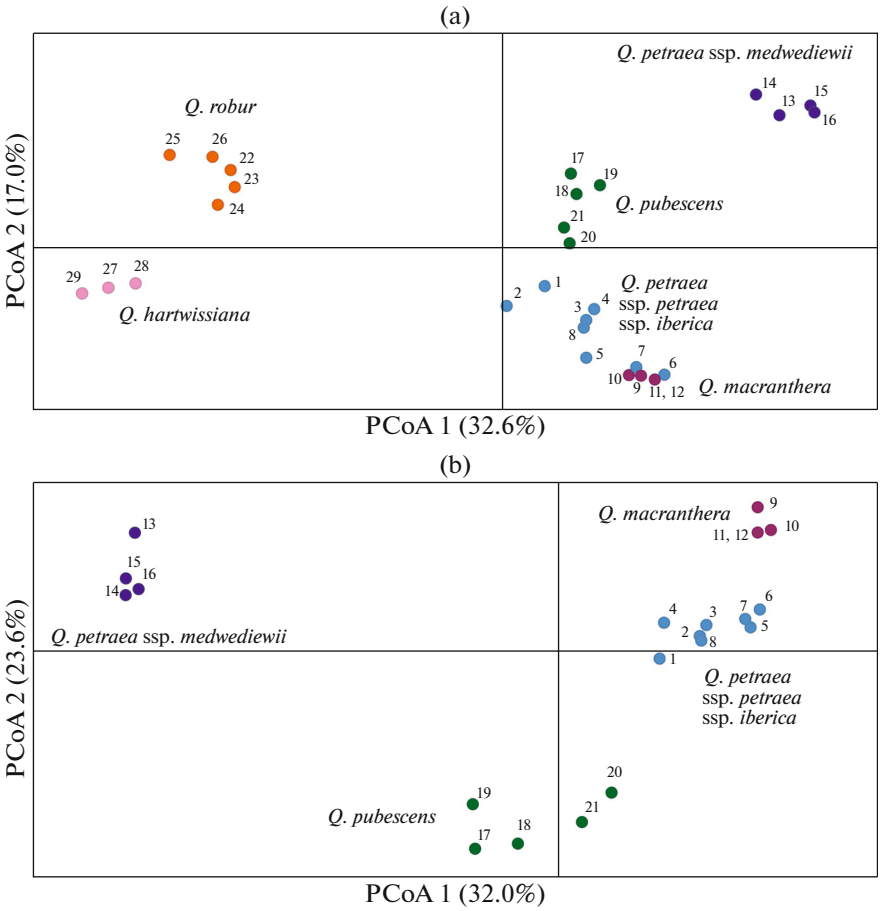

Fig. 2. PCoA-ordination based on the variability of 18 nSSR loci. (a) Populations of seven taxa of roburoid oaks; (b) populations of sessile-flower oaks *Q. petraea* s. l., *Q. macranthera*, *Q. pubescens*. Populations are color-coded according to the dominant cluster STRUCTURE with the number of clusters $K = 6$ in set 1 (see Fig. 1b).

ssp. *iberica*, while *Q. robur* and *Q. hartwissiana* remain in the same cluster.

When analyzing the two main groups separately (set 2—sessile-flower oaks and set 3—long-peduncle oaks), a division into taxa can be traced similar to set 1, with further division into geographical groups of populations within taxa with an increase in the number of clusters. Figure 3 shows the results of the STRUCTURE analysis for set 2 (sessile-flower taxa) with the number of clusters $K = 4, 5, 6, 7$. The highest ΔK value was obtained for $K = 4$, the highest $LnP(K)$ value for $K = 6$ (Fig. S4). At the most supported number of clusters $K = 4$, the samples are divided into four groups, coinciding with the division at $K = 6$ for set 1. Next, at $K = 5$, the blue cluster is divided into two clusters in accordance with the subspecies *Q. p.* ssp. *petraea* and *Q. p.* ssp. *iberica*. At $K = 6$ *Q. p.* ssp. *petraea*, in turn, is divided into two groups: European populations of *Q. petraea* (nos. 1, 2) and Crimean populations of *Q. petraea* (nos. 3, 4). At $K = 7$, *Q. pubescens* is divided into two clusters: western (Crimea, nos. 17–19) and eastern (Dagestan, nos. 20, 21).

When analyzing only *Q. robur* and *Q. hartwissiana* samples (set 3), the maximum ΔK value was obtained for $K = 2$ (Fig. S5a), which separates individuals according to species into *Q. robur* and *Q. hartwissiana*. The second peak was observed at $K = 5$. The $LnP(K)$ value consistently increased, reaching a maximum at $K = 5$ (Fig. S5b). Figure S6 shows the results of STRUCTURE for $K = 2, 3, 4, 5$. As the number K increases, *Q. robur* is divided into geographical groups, which is not observed for *Q. hartwissiana*. At $K = 3$, *Q. robur* is divided into clusters of East Caucasian (nos. 22–24) and West Caucasian (nos. 25, 26) populations. Further, at $K = 4$, differentiation is observed in the Western Caucasian cluster, and at $K = 5$, in the Eastern Caucasian cluster.

Pie charts of frequencies of STRUCTURE clusters by population are presented in Fig. S1. Cluster content was calculated for populations at set 1 ($K = 6$), excluding three individuals that did not match taxonomic assignment (see below). The proportion of the cluster corresponding to the appropriate taxon was more than 85% (on average 93%) in all populations: in the populations of *Q. robur*—91% (from 85 to 95%), *Q. hartwis-*

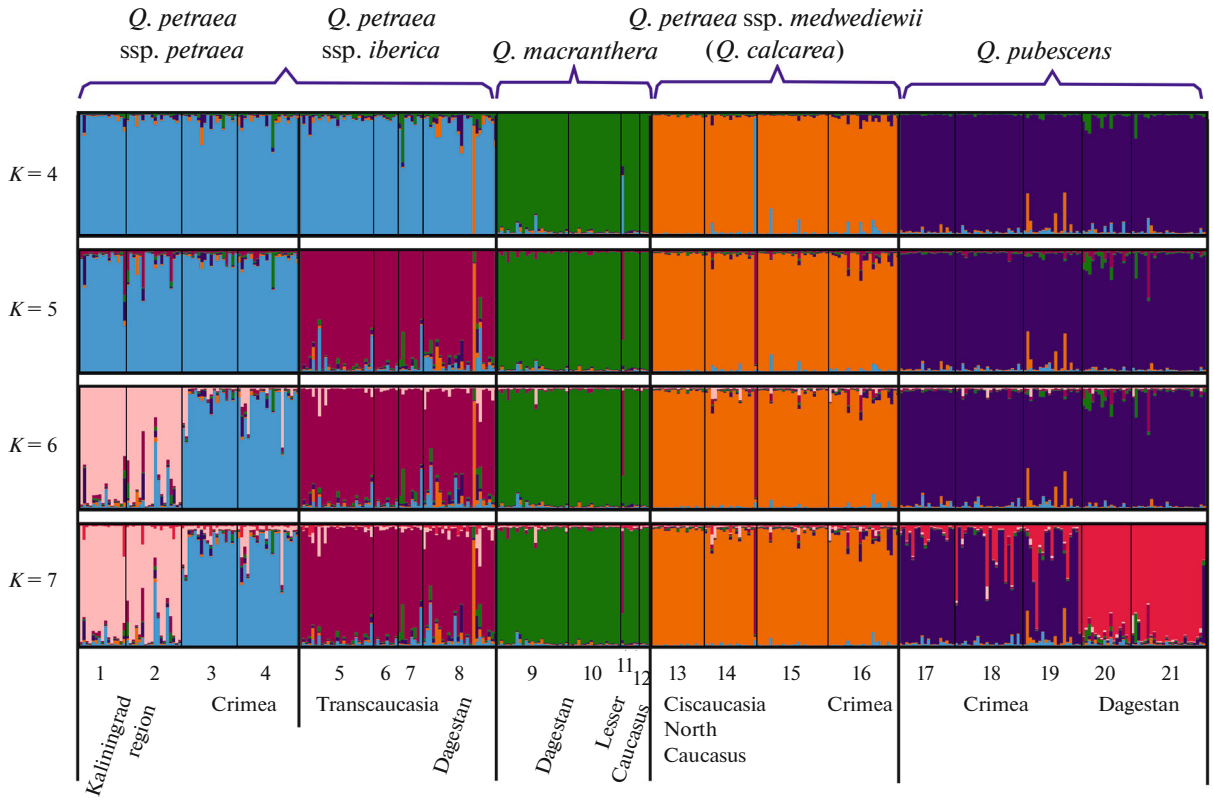

Fig. 3. Distribution of genetic clusters calculated using STRUCTURE based on 18 microsatellite loci for 364 individuals of *Q. petraea* s. l., *Q. macranthera*, *Q. pubescens* (set 2) at $K = 4, 5, 6, 7$. Each sample is represented by a vertical column, divided into segments proportional to the participation of each genetic cluster.

siana—94% (from 92 to 97%), *Q. p.* ssp. *petraea*—92% (from 90 to 93%), *Q. p.* ssp. *iberica*—91.5% (from 85 to 95%), *Q. pubescens*—93% (from 85 to 95%), *Q. p.* ssp. *medwediewii*—93% (from 91 to 96%), *Q. macranthera*—95% (from 94 to 97%).

Specificity of Genotyping in Microsatellite Analysis

It was found that the content of a cluster of extraneous taxon in a population sample can be about 2%, even if current hybridization is impossible. For example, STRUCTURE shows the content of a 2% cluster of the Western Caucasian species *Q. hartwissiana* in population *Q. robur* no. 23 and 3% in population *Q. p.* ssp. *medwediewii* no. 14 from the Eastern and Central Caucasus, 2.2% of the cluster of eastern and Transcaucasian *Q. macranthera* in population *Q. petraea* no. 4 in Crimea. That is, possible hybridization can be assumed only if the percentage of the extraneous cluster is increased to 4–5% in the population and with the potential for gene flow due to overlapping areas. An increased content of the *Q. robur* cluster is observed in populations no. 1 (4.2%) and no. 27 (3.7%), the *Q. p.* ssp. *medwediewii* cluster—in population no. 19 (5.3%), *Q. pubescens* cluster—no. 8 (4.3%), *Q. hartwissiana* cluster—no. 25 (5.9%), *Q. macranthera* cluster—5% in populations nos. 20 and 23. All these cases coincide with the possibility of hybridization due to sympatry.

Since the species identity of the individual trees was determined visually in the field, erroneously identified individuals were found in the collected material, which is expected, given the significant overlap of diagnostic characters in the taxa under consideration. In two populations each (no. 8 and no. 14), one individual was identified, which the STRUCTURE analysis assigned to a cluster of another subspecies within *Q. petraea* (Figs. 1b; 3), which corresponded to the possible geographical distribution of subspecies with overlapping ranges (see Discussion). In one of the populations of *Q. macranthera* (no. 11), an individual was identified in which the rate between clusters corresponding to *Q. p.* ssp. *iberica* and *Q. macranthera* according to STRUCTURE, was approximately 65 : 35%, which probably indicates a hybrid with a predominance of *Q. p.* ssp. *iberica*. Based on the herbarium material, it turned out that this specimen lacked the pubescence characteristic of *Q. macranthera*. The results of microsatellite analysis made it possible to correctly establish the taxonomic affiliation of some

samples, but, in addition, a certain number of apparently true hybrids were encountered: one individual in the population of *Q. p.* ssp. *iberica* no. 7 had 35% of the *Q. macranthera* cluster. In *Q. pubescens* population no. 19, two samples each contained 25–30% of the *Q. p.* ssp. *medwediewii* cluster. In *Q. robur* population no. 25, three samples contained 25–30% of the *Q. hartwissiana*, in *Q. robur* population no. 23, one sample shared 35% of the *Q. macranthera*, in *Q. robur* population no. 26, one sample had 45% of the *Q. petraea*.

To investigate the possibility to identify a taxon using STRUCTURE when it is poorly represented in the data, a series of experiments were conducted (see Material and Methods). In "simulated" option 1, *Q. p.* ssp. *medwediewii* was removed from the analysis except four specimens. With the most supported number of clusters $K = 4$, these samples were assigned by the algorithm to the *Q. pubescens* cluster, then with an increase in the number of clusters—to hybrid samples of several taxa, and only with the number $K = 9$, a separate cluster of four *Q. p.* ssp. *medwediewii* samples was identified (data not shown). In "simulation" option 2, which excluded reference populations of *Q. hartwissiana* from the analysis, the four retained Hartwiss oak samples were assigned to *Q. robur* and were not allocated to a separate cluster until $K = 12$ was reached. Thus, the reduced abundance of a particular taxon in the STRUCTURE analysis, compared to the representation of other groups, leads to incorrect identification of the sample with the most plausible number of clusters and reveals a taxon-specific cluster only with a significant number K (which usually does not meet the optimality criteria and is not considered in studies).

When performing the STRUCTURE analysis with different species sets (sets 1, 2, 3), it was shown that in the absence of individual taxa in the analysis, there may be misinterpretation of hybrids. For example, with set 3 (including only *Q. hartwissiana* and *Q. robur*) in the *Q. hartwissiana* population no. 28, one sample contained 50% of the *Q. robur* cluster (Fig. S6). With set 1 (Fig. 1b), this individual turns out to be a mixed hybrid with a contribution of *Q. petraea*/*Q. pubescens* without the participation of *Q. robur*. Some hybrids are not detected at all when taxa are omitted, for example, in *Q. robur* population no. 26, set 1 revealed a hybrid with 45% of the *Q. petraea* cluster, which was not shown with set 3.

To investigate the ability to differentiate taxa with fewer loci, STRUCTURE analyzes were performed separately for two multiplexes: ten loci (kit-1) and eight loci (kit-2). The composition of the clusters into which the populations were divided did not change (data not shown), which shows the possibility of identifying taxa for relatively pure populations with a smaller set of loci. However, the assignment of individual specimens to certain taxa may be incorrect: with a small number of loci, in individuals the frequencies of clusters of other species increase, sometimes to a significant value, making these individuals potentially "hybrid." For example, in populations of *Q. pubescens* from coastal locations of the southern macroslope of the Main Ridge of the Crimean Mountains for three individuals the participation of *Q. robur* cluster was shown to be 60–80% (kit-2 set, eight loci), which is unlikely in the complete absence of *Q. robur* on the southern macroslope. This is not observed when using both sets of loci. Thus, for preliminary differentiation of taxa, a smaller number of loci can be used (for example, only kit-2), which reduces costs, but for a more accurate identification of each individual, especially when studying hybridization processes in mixed populations, analysis of a large number of loci is necessary (in our case—the use of two multiplexes).

DISCUSSION

Differentiation of Quercus Taxa in the Crimean-Caucasian Region Based on Microsatellite Data

Using 18 nSSR loci, the genetic diversity of *Q. robur*, *Q. hartwissiana*, *Q. petraea* ssp. *petraea*, *Q. p.* ssp. *iberica*, *Q. p.* ssp. *medwediewii*, *Q. pubescens* and *Q. macranthera* was studied in the Crimea, the Caucasus and northeastern Europe (Fig. 1a). All taxa, including those studied for the first time using nSSRs (*Q. hartwissiana*, *Q. macranthera*, *Q. p.* ssp. *iberica*, *Q. p.* ssp. *medwediewii*), showed stable amplification and fairly high variability at all loci. There is a great deal of overlap in the allele sets of species, and species differ mainly in allele frequencies. The loci used are good at the taxonomic assignment of trees and identifying hybrids. Based on the results of STRUCTURE (Figs. 1b; 3), genetic clusters corresponding to all seven taxa were recognized.

At $K = 2$ (Fig. 1b), all samples are divided according to the ratio of the lengths of the fruiting stalk (peduncle) and leaf petiole into a group of long-peduncle oaks (*Q. robur*, *Q. hartwissiana*) and a group of sessile-flower oaks (all others, including *Q. macranthera*), which is consistent with the results of a phylogenetic study based on RADseq markers [5], where the maximum likelihood tree contained the clade of *Q. robur*, *Q. hartwissiana* and *Q. canariensis*, sister to the clade containing *Q. petraea*, *Q. pubescens* and *Q. macranthera*. It is interesting that the relatedness between *Q. robur*, *Q. hartwissiana* and *Q. canariensis* was previously assumed by Yu.L. Menitsky [2] based on the morphological characters. In addition, in previous studies based on nSSR markers, the differences between *Q. robur* and sessile-flowers oaks (*Q. pubescens*, *Q. petraea*, *Q. frainetto*, etc.) were higher than between the latter [12, 15, 16, 24, 29]. At the same time, the division of the studied taxa into two groups does not correspond to the classification of Menitsky, who distributed these taxa into three subsections [6–9] (Subsect. *Quercus*, Subsect. *Macrantherae*, Sub-

sect. *Galliferae*), as well as other recent classifications of European oaks section *Quercus* [48].

All studied taxa of sessile-flowers oaks (*Q. pubescens*, *Q. macranthera* and three subspecies of *Q. petraea*) are differentiated from each other, which corresponds to STRUCTURE (Fig. 3) and PCoA results (Fig. 2; Fig. S2). According to the AMOVA, the level of genetic differentiation of F_{RT} between the five sessile-flowers taxa (including subspecies) was quite high and amounted to 0.085 ($P = 0.001$), which is 4 times higher than the differentiation within taxa ($F_{SR} = 0.022$, $P = 0.001$). The species level is confirmed for *Q. pubescens* and *Q. macranthera* and is suggested for one of the subspecies of sessile oak (*Q. petraea* ssp. *medwediewii*), which showed high differentiation from the other two subspecies of *Q. petraea*.

In the studied populations, when several taxa live together, hybrid individuals were found, but a significant predominance of "pure" individuals confirms that hybridization between species is difficult, probably due to ecological and phenological isolation [8].

Hartwiss Oak Q. hartwissiana and Pedunculate oak Q. robur

If the pedunculate oak was previously studied using nSSR markers in the Caucasian region [36], then the Hartwiss oak *Q. hartwissiana* was studied for the first time. The analysis of *Q. hartwissiana* included 33 individuals from three localities of several regions of the Caucasus (Fig. 1). Reliable differences in microsatellite markers were found between the Hartwiss oak and other oak species, including the closest species, *Q. robur* (Figs. 1, 2; Fig. S6), and the species status of this morphologically well-differentiated species was confirmed [1–3].

Within *Q. robur* in the Caucasus at $K \geq 3$, a clear geographic pattern is observed, with a consistent division of populations according to geographic groups that coincide with those previously identified based on the variability of 14 nSSR loci [36]. In contrast to *Q. robur*, populations of *Q. hartwissiana* from the northern (no. 29) and southern (nos. 27, 28) macroslopes of the Greater Caucasus remain homogeneous according to the STRUCTURE results. However, the subdivision of the three Hartwiss oak populations (pairwise $F_{ST} = 0.040–0.049$) was quite high, probably due to significant fragmentation of *Q. hartwissiana* populations, whose habitats are usually limited to river valleys [3]. There is obviously a gene flow between the pedunculate and Hartwiss oaks. For example, in *Q. robur* populations, according to STRUCTURE results, the proportion of the *Q. hartwissiana* cluster is higher in the Western Caucasus (population no. 25, Fig. S6), where the latter is widespread.

Intra-population variability in *Q. hartwissiana,* compared to *Q. robur*, is reduced in all populations. It is likely that this thermophilic, mesophilic species experienced particularly serious population declines during unfavorable climatic periods. Most researchers consider *Q. hartwissiana* to be close to the most ancient form of Tertiary Roburoids [1, 8]. At the same time, the ecological properties of the Hartwiss oak—being confined to warm and humid biotopes and belonging to the core of the tertiary flora—indicate the relict nature of this species [8]. Yu.L. Menitsky notes that the Hartwiss oak is close in leaf morphology to the Georgian durmast oak *Q. petraea* ssp. *iberica* [1], but such features as flowering simultaneously with the blossoming of leaves and long peduncle reflect the beginning of adaptation to a cold climate and bring Hartwiss oak closer to the pedunculate oak, which indicates their common origin [8].

Differentiation of Limestone Oak Q. petraea ssp. medwediewii from Sessile Oak Q. petraea

In the studied areas of the Caucasus and Crimea, according to the taxonomic system of Yu.L. Menitsky [1, 3, 6], three widespread subspecies of *Q. petraea* grow. If the type subspecies *Q. p.* ssp. *petraea* has been studied in numerous studies in Western, Central and Eastern Europe [15, 16, 19, 23, 24, 26, 28, 33], two other, more eastern sessile oak subspecies were studied for the first time using nuclear markers. The most unexpected and important result is that the *Q. petraea* subspecies do not group together on the PCoA ordination and on the UPGMA dendrogram (Fig. 2; Fig. S1), but form two clusters, as in the STRUCTURE analysis (Fig. 3). One of the clusters corresponds to the subspecies *Q. petraea* ssp. *medwediewii* (A. Camus) Menitsky (syn. *Q. calcarea* Troitsky), judging by the geographic distribution and morphological description. The degree of differences of this taxon in nSSRs both from the other two sessile oak subspecies and from *Q. pubescens* and *Q. macranthera* confirms the interpretation of this taxon in the rank of species—*Q. calcarea* Troitsky [9].

Yu.L. Menitsky's assumption about the possible hybrid origin of *Q. p.* ssp. *medwediewii* [1–3] during the historical hybridization of *Q. petraea* and *Q. pubescens*, based on the presence of transitional morphological characters, is not confirmed by genetic analysis. Populations of *Q. p.* ssp. *medwediewii* are not intermediate on the PCoA ordination (Fig. 2), and individuals do not contain a mixture of genetic pools of other species according to STRUCTURE results (Fig. 3) and retain their genetic identity not only in the areas where it grows separately from other subspecies of *Q. petraea* and the rest sessile-flowers oaks (Central Caucasus and Ciscaucasia, Fig. 1a), but also in areas where limestone oak grows together with other closely related taxa (for example, no. 16 in Crimea).

Existence and interpretation of the taxon "limestone oak" *Q. p.* ssp. *medwediewii* (=*Q. calcarea*), which differs morphologically and ecologically from other sessile-flowers oaks and occupies a large range

on the northern macroslope of the Greater Caucasus, in the Ciscaucasia and in the Crimea, has been the subject of debate for many years [7]. In various taxonomic classifications, this oak [1, 7, 10] was considered as a separate species, a subspecies of downy oak, a subspecies of sessile oak, or was not recognized as a separate taxon. Yu.L. Menitsky considered *Q. p.* ssp. *medwediewii* as a synonym of *Q. dalechampii* Ten.—a European species [49], the systematic position and interpretation of which, in turn, remain very uncertain and have not yet received proper resolution (review [7, 49–52]). After a recent revision in the scope and composition of the taxon *Q. dalechampii* and its division into two parts belonging to the systems *Q. petraea* and *Q. pubescens* [50], a new name was proposed for Central Europe—*Quercus banatus* P. Kucera [51].

The discrepancy between taxonomic classifications is explained not only by morphological similarity and largely overlapping characters between closely related taxa, but also by different degrees of study of taxa in the regions. Although the morphological variability of the Caucasian and Crimean populations of limestone oak was studied on a large material in the works of Yu.L. Menitsky [1] and N.D. Troitsky [9], but an extended comparison with European populations of *Q. dalechampii* was not carried out, which probably determines a different understanding of the volume of species (systems of species) of sessile oak and downy oak. Our study shows that the limestone oak (*Q. p.* ssp. *medwediewii* = *Q. calcarea*) is clearly differentiated from all taxa we studied and, in terms of the level of differences, can be recognized as a separate species, since the differences are at the level of differences between downy oaks and sessile oaks. In general, the confusion of classifications only increases the need for joint research of likely related taxa and expansion of geography with the involvement of material from Europe, including *Q. dalechampii* (or *Q. banatus*).

Variability of *Q. p.* ssp. *medwediewii* is less than in other common species, presumably as a result of reduced effective population size. Weak differentiation of *Q. p.* ssp. *medwediewii* populations within the Crimean-Caucasian region is likely due to recent dispersal from a single source. It is possible that low effective population sizes are also responsible for the increased genetic differences of *Q. p.* ssp. *medwediewii* with other taxa. Approximately equal differentiation with closely related species—downy oak (average paired F_{ST} = 0.078), sessile oak (0.086) and durmast oak (0.089) makes the phylogenetic relationships of *Q. p.* ssp. *medwediewii* uncertain.

The other two studied subspecies of *Q. petraea* (type subspecies *Q. p.* ssp. *petraea* and Georgian durmast oak *Q. p.* ssp. *iberica*) are differentiated from each other to a much lesser extent than from limestone oak *Q. p.* ssp. *medwediewii* and other species (Fig. 2b; Fig. S1), however, the geographical distribution of populations with a noticeable participation of the corresponding STRUCTURE clusters confirms the validity of the subspecies *Q. p.* ssp. *iberica* within *Q. petraea*. (Fig. 3; Fig. S2). A phylogenetic study based on the variability of ITS sequences [53] also confirmed the differences between the two taxa, which do not exclude their species rank with a clear relatedness between the subspecies.

Interestingly, the type subspecies *Q. p.* ssp. *petraea* possess a geographic structure. Populations of Crimea differ from *Q. petraea* from the Kaliningrad region (Fig. 3; Fig. S1), while not being transitional between durmast oak and sessile oak (Fig. S2). The genetic characteristics of *Q. petraea* from Crimea may explain the assignment, based on morphological characteristics, of the Crimean populations to both *Q. iberica* [9] and *Q. p.* ssp. *petraea* [1]. To more fully identify the intraspecific structure of *Q. petraea*, it is necessary to study the populations of the Black Sea coast of the Caucasus [1, 6], where both subspecies and hybrids between them are noted, as well as the inclusion in the analysis of *Q. petraea* from Asia Minor and the Balkan Peninsula, where both taxa may possibly grow [54].

Detection in population no. 14 (Balta) among the samples of *Q. p.* ssp. *medwediewii* of one individual belonging to *Q. p.* ssp. *iberica*, is an important confirmation of the observations of Yu.L. Menitsky, who described the range of the durmast oak as almost never extending north of the Main Caucasus Range, but noted its presence in the Verkhnetersky floristic region (Central Caucasus) [6], in particular in the Saurgom Gorge, near the village of Chmi [1], i.e. in close proximity to population no. 14.

Downy Oak Q. pubescens

According to the results of our study, it was shown that *Q. pubescens* from the Crimean-Caucasian part of its range belongs to the same taxon. Samples and individuals of *Q. pubescens* from Dagestan are close to the *Q. pubescens* populations of Crimea, forming a common cluster in the STRUCTURE analysis (Figs. 1b; 3) and a common group in the PCoA ordination (Fig. 2). Systematically and geographically (in terms of distribution in Crimea and the Caucasus), the selected group most closely coincides with Menitsky's interpretation [1, 6].

Within a group of five populations of *Q. pubescens* F_{ST} was 0.042 (P = 0.001). Among them, the greatest differences are observed between the eastern (Dagestan) and western (Crimea) populations. Within these regions, differentiation was low. According to the results of STRUCTURE (Fig. 3) with K = 7 (set 2), downy oak is divided into two clusters—western (Crimea) and eastern (Dagestan). The same division is observed in the UPGMA dendrogram (Fig. S1) and in PCoA (Fig. 2). In addition to isolation by distance, the reason for the differentiation of *Q. pubescens* may be associated with long-term processes of hybridiza-

tion when growing sympatrically with different taxa. In the Dagestan population samples, *Q. pubescens* contains more admixtures of *Q. macranthera* and *Q. p.* ssp. *iberica*, and in the Crimean populations—*Q. p.* ssp. *medwediewii* and *Q. p.* ssp. *petraea*.

The increased genetic diversity of downy oak compared to all studied taxa was previously noted (for example, [14, 22, 23]) and is explained by the high ability for hybridization, apparently historical, since in our case *Q. pubescens* forms a "pure" cluster, without a noticeable proportion of other taxa. We also propose another explanation for the increased allelic diversity—the high effective population size of the species, even despite its limited distribution in the Caucasus. This may have historical reasons and is associated with a wider distribution of this species in the past [1].

Caucasian or Persian Oak, Q. macranthera

This taxon was studied for the first time using microsatellite markers and the species status of *Q. macranthera* was confirmed. The populations and almost all individuals, preliminary assigned to *Q. macranthera*, on the base of nSSR markers differentiated from samples of other species into a separate STRUCTURE cluster (Figs. 1b; 3) and formed a separate clade at the UPGMA and a group at the PCoA ordination (Fig. 2b). Variability ($H_e = 0.659$) were slightly reduced; differentiation of populations within the taxon was not revealed, despite the inclusion in the analysis of populations from geographically distant regions (Dagestan and Armenian Highlands). A possible hybrid individual was discovered in the Dilijan location (population no. 11), with a predominance of the *Q. p.* ssp. *iberica* cluster (Fig. 3). Upon repeated examination of the herbarium of this specimen, an almost complete absence of pubescence characteristic of *Q. macranthera* was noted.

An increased content of the *Q. macranthera* cluster was found in populations of other species growing together with it in Dagestan and Transcaucasia (*Q. robur*—populations no. 22 and 23, *Q. p.* ssp. *iberica*—nos. 7 and 8, *Q. pubescens*—nos. 20 and 21). It is interesting that these species practically do not form mixed stands with *Q. macranthera*, growing at different altitudes. *Q. macranthera* is genetically close to other sessile-flower taxa and is included in clade II (UPGMA). According to phylogenetic studies based on RADseq data [5] and ITS sequences [53], this taxon is related to *Q. frainetto* and belongs to the clade containing *Q. pubescens* [5]. Further study of Caucasian taxa, involving closely related European oak species in the analysis, and demographic modeling based on the ABC method can more accurately determine the evolution of this group of species.

Efficiency of Taxa Recognition with Different Composition of Loci and Reference Groups

All loci used in our study have an F_{ST} between taxa of at least 0.081 (Table 2). Within taxa, subdivision between populations was significant, but pronounced geographic structure was revealed only within *Q. p.* ssp. *petraea*, *Q. pubescens* and *Q. robur*. Using only one of the multiplexes to discriminate taxa (which reduced time and costs) showed a division of populations into taxa similar to the two multiplexes, but the accuracy of assigning samples to a taxon decreased (see Results). It was previously shown that some loci discriminate better between species or groups of species, and their use to establish species identity gives a result no less effective than using a large set of loci [19]. Some studies have subsequently used a limited set of loci to discriminate between a pair of species (for example, [55] selected four loci to discriminate between *Q. robur* and *Q. petraea*). However, many examples have shown that for other species combinations other loci may be discriminatory (e.g., [13, 21, 22]). Analysis of our data showed that, although some loci differentiate individual taxa more, it is most advisable to use all 18 loci, since for the Caucasus and Crimea it is quite common for several taxa to be present in mixed stands or to grow in close proximity—up to four taxa in Crimea and Transcaucasia, up to five taxa—in the Western and Eastern Caucasus.

For one of the most famous so-called "outlier loci" QrZAG96 [19], described in a number of studies as the most discriminating against *Q. petraea* and *Q. robur* and potentially under the influence of selection, in European populations of *Q. robur* there was a tendency towards fixation of the shortest allele, which has previously been associated with the morphological trait of relative leaf petiole length. However, our work shows for the first time that an increased frequency of this allele was observed not only for *Q. robur* (average 0.601), but also for *Q. hartwissiana* (0.576), which is much higher than in sessile-flower oaks (from 0 to 0.127). This fact probably reflects the common origin of *Q. robur* and *Q. hartwissiana* and is not related to the length of the petiole, which in *Q. hartwissiana* is closer to *Q. petraea*. Accordingly, the association of the QrZAG96 locus with the hypothetical QTL controlling petiole length was not confirmed. In addition, the almost complete fixation of the short allele in *Q. robur* in more western European populations is replaced by a much lower frequency in the Caucasus (varying from 0.342 to 0.909, this study) and Turkey (0.210) [17].

Studying the possibility of identifying genetic structure using the Bayesian clustering method with a low representation of some taxa in the analyzed data ("simulated" options 1 and 2) showed the possibility of incorrect classification of pure individuals and hybrids of taxa that are rare in the analyzed data. These results confirm the conclusion of earlier studies [18,

23, 24] that it is necessary to include in the analysis a significant number of reference individuals of each of the taxa potentially present in the study area. With a sufficient amount of material for all taxa growing in the studied part of the range, the 18 nSSR loci we used discriminated them much better than a significant number of nuclear SNPs analyzed without taking into account the entire diversity of the Crimean-Caucasian *Quercus* [56].

The genetic clusters identified in our work can be used as reference ones for studying hybridization processes in areas where similar taxa grow together, to establish taxonomic affiliation in the absence of the possibility of classification based on morphology, to clarify the distribution of taxa and create programs for the protection of genetic resources in the Crimean-Caucasian region.

Thus, our study shows the effectiveness of using existing sets of microsatellite loci for identifying roburoid oaks in the Crimean-Caucasian region. The identified genetic clustering of samples was consistent with their taxonomic subdivision based on morphological characters and the distribution of seven *Quercus* taxa in the study areas [1, 3, 6]. Species status is confirmed for all species and, in addition, is assumed for the limestone oak *Q. petraea* ssp. *medwediewii*. The results obtained confirm the opinion that the introgression of closely related oak species in the region is highly limited [8], since it does not lead to the erasure of differences or disruption of the integrity and evolutionary independence of taxa.

SUPPLEMENTARY INFORMATION

The online version contains supplementary material available at https://doi.org/10.1134/S1022795424700492.

ACKNOWLEDGMENTS

The authors are grateful to E.V. Zinoviev, I.Yu. Isakov, N.V. Semerikov, Z.A. Guseinova, Z.M. Asadulaev for help in collecting oak samples and to an anonymous reviewer for constructive comments and suggestions.

FUNDING

The study was carried out within the framework of state contract of the Institute of Plant and Animal Ecology of the Ural Branch of the Russian Academy of Sciences, no. 122021000090-5. Collection of some samples in 2022–2023 was supported by a grant from the Russian Science Foundation no. 22-24-00667, https://rscf.ru/project/22-24-00667/.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This work does not contain any studies involving human and animal subjects.

CONFLICT OF INTEREST

The authors of this work declare that they have no conflicts of interest.

REFERENCES

1. Menitskii, Yu.L., *Duby Kavkaza: obzor kavkazskikh predstavitelei sektsii Quercus* (Oaks of the Caucasus: an Overview of the Caucasian Representatives of the Section *Quercus*), Leningrad: Nauka, 1971.
2. Menitskii, Yu.L., Oaks (*Quercus* L.) of Southwest Asia, *Nov. Sist. Vyssh. Rast.*, 1972, vol. 9, pp. 105—140.
3. Menitsky, Y.L., *Oaks of Asia*, Enfield, NH: Science Publ., 2005.
4. Kremer, A. and Hipp, A.L., Oaks: an evolutionary success story, *New Phytol.*, 2020, vol. 226, no. 4, pp. 987—1011. https://doi.org/10.1111/nph.16274
5. Hipp, A.L., Manos, P.S., Hahn, M., et al., Genomic landscape of the global oak phylogeny, *New Phytol.*, 2020, vol. 226, no. 4, pp. 1198—1212. https://doi.org/10.1111/nph.16162
6. *Konspekt flory Kavkaza v 3 tomakh* (Caucasian Flora Conspectus in 3 volumes), Kudryashova, G.L. and Tatanov, I.V., Eds., Moscow: KMK, 2012, vol. 3, part 2.
7. Schmidt, P.A., Oaks and oak forests in Caucasia, *Proc. Fourth Intern. Oak Conf.*, 2004, no. 15, pp. 9—29.
8. Semerikov, L.F., *Populyatsionnaya struktura drevesnykh rastenii (na primere vidov duba evropeiskoi chasti SSSR i Kavkaza)* (Population Structure of Arboraceous Plants (Exemplified by Oak Species of the European Part of the USSR and the Caucasus)), Moscow: Nauka, 1986.
9. Troitskii, N.D., Preliminary results of the oak trees study in the Crimean State Reserve and the adjacent region of the southern coast of Crimea (taxonomy in connection with growing conditions), *Zh. Russ. Bot. O-va.*, 1931, vol. 16, no. 4, pp. 313—354.
10. Cherepanov, S.K., *Sosudistye rasteniya Rossii i sopredel'nykh gosudarstv* (Vascular Plants of Russia and Neighboring Countries), St. Petersburg: Mir i Sem'ya, 1995.
11. Valbuena-Carabaña, M., González-Martínez, S.C., Hardy, O.J., and Gil, L., Fine-scale spatial genetic structure in mixed oak stands with different levels of hybridization, *Mol. Ecol.*, 2007, vol. 16, no. 6, pp. 1207—1219. https://doi.org/10.1111/j.1365-294X.2007.03231.x
12. Fortini, P., Viscosi, V., Maiuro, L., et al., Comparative leaf surface morphology and molecular data of five oaks of subgenus *Quercus* Oerst. (Fagaceae), *Plant Biosyst.*, 2009, vol. 143, no. 3, pp. 543—554. https://doi.org/10.1080/11263500902722980
13. Salvini, D., Bruschi, P., Fineschi, S., et al., Natural hybridization between *Quercus petraea* (Matt.) Liebl. and *Quercus pubescens* Willd. within an Italian stand as revealed by microsatellite fingerprinting, *Plant Biol.*, 2009, vol. 11, pp. 758—765. https://doi.org/10.1111/j.1438-8677.2008.00158.x
14. Antonecchia, G., Fortini, P., Lepais, O., et al., Genetic structure of a natural oak community in central Italy: evidence of gene flow between three sympatric white

oak species (*Quercus*, Fagaceae), *Ann. For. Res.*, 2015, vol. 58, no. 2, pp. 205—216. https://doi.org/10.15287/afr.2015.415

15. Rellstab, C., Buhler, A., Graf, R., et al., Using joint multivariate analyses of leaf morphology and molecular-genetic markers for taxon identification in three hybridizing European white oak species (*Quercus* spp.), *Ann. For. Sci.*, 2016, vol. 73, no. 3, pp. 669—679. https://doi.org/10.1007/s13595-016-0552-7
16. Mačejovský, V., Schmidtová, J., Hrivnák, M., et al., Interspecific differentiation and gene exchange among the Slovak *Quercus* sect. *Quercus* populations, *Dendrobiology*, 2020, vol. 83, pp. 20—29. https://doi.org/10.12657/denbio.083.002
17. Yücedağ, C. and Gailing, O., Morphological and genetic variation within and among four *Quercus petraea* and *Q. robur* natural populations, *Turk. J. Bot.*, 2013, vol. 37, no. 4, pp. 619—629. https://doi.org/10.3906/bot-1205-18
18. Fortini, P., Marzio, P.D., Conte, A.L., et al., Morphological and molecular results from a geographical transect focusing on *Quercus pubescens/Q. virgiliana* ecological-altitudinal vicariance in peninsular Italy, *Plant Biosyst.*, 2022, vol. 156, no. 6, pp. 1498—1511. https://doi.org/10.1080/11263504.2022.2131923
19. Neophytou, C., Aravanopoulos, F.A., Fink, S., and Dounavi, A., Detecting interspecific and geographic differentiation patterns in two interfertile oak species (*Quercus petraea* (Matt.) Liebl. and *Q. robur* L.) using small sets of microsatellite markers, *For. Ecol. Manage.*, 2010, vol. 259, pp. 2026—2035. https://doi.org/10.1016/j.foreco.2010.02.013
20. Guichoux, E., Lagache, L., Wagner, S., et al., Two highly validated multiplexes (12-plex and 8-plex) for species delimitation and parentage analysis in oaks (*Quercus* spp.), *Mol. Ecol. Resour.*, 2011, vol. 11, pp. 578—585. https://doi.org/10.1111/j.1755-0998.2011.02983.x
21. Curtu, A.L., Gailing, O., Leinemann, L., and Finkeldey, R., Genetic variation and differentiation within a natural community of five oak species (*Quercus* spp.), *Plant Biol.*, 2007, vol. 9, pp. 116—126. https://doi.org/10.1055/s-2006-924542
22. Curtu, A.L., Moldovan, I.C., Enescu, C.M., et al., Genetic differentiation between *Quercus frainetto* Ten. and *Q. pubescens* Willd. in Romania, *Not. Bot. Horti Agrobot. Cluj-Napoca*, 2011, vol. 39, no. 1, pp. 275—282. https://doi.org/10.15835/nbha3915633
23. Curtu, A.L., Craciunesc, I., Enescu, C.M., et al., Fine-scale spatial genetic structure in a multi-oak-species (*Quercus* spp.) forest, *iForest*, 2015, vol. 8, no. 3, pp. 324—332. https://doi.org/10.3832/ifor1150-008
24. Lepais, O., Petit, R.J., Guichoux, E., et al., Species relative abundance and direction of introgression in oaks, *Mol. Ecol.*, 2009, vol. 18, pp. 2228—2242. https://doi.org/10.1111/j.1365-294X.2009.04137.x
25. Höltken, A.M., Buschbom, J., and Kätzel, R., Species integrity of *Quercus robur* L., *Q. petraea* (Matt.) Liebl. and *Q. pubescens* Willd. from the genetic point of view, *Allg. Forst-Jagdztg.*, 2012, vol. 183, nos. 5—6, pp. 100—110.
26. Gerber, S., Chadoeuf, J., Gugerli, F., et al., High rates of gene flow by pollen and seed in oak populations across Europe, *PLoS One*, 2014, vol. 9, p. e85130. https://doi.org/10.1371/journal.pone.0091301
27. Gugerli, F., Brodbeck, S., and Holderegger, R., Utility of multilocus genotypes for taxon assignment in stands of closely related European white oaks from Switzerland, *Ann. Bot.*, 2008, vol. 102, pp. 855—863. https://doi.org/10.1093/aob/mcn164
28. Sandurska, E., Ulaszewski, B., and Burczyk, J., Genetic diversity and differentiation of coexisting populations of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl., *Acta Biol. Cracov., Ser. Bot.*, 2019, vol. 61, no. 1, pp. 17—28. https://doi.org/10.24425/abcsb.2019.127739
29. Neophytou, C., Gärtner, S.M., Vargas-Gaete, R., et al., Genetic variation of Central European oaks: shaped by evolutionary factors and human intervention?, *Tree Genet. Genom.*, 2015, vol. 11, no. 79. https://doi.org/10.1007/s11295-015-0905-7
30. Enescu, C.M., Curtu, A.L., Şofletea, N., Is *Quercus virgiliana* a distinct morphological and genetic entity among European white oaks?, *Turk. J. Agric. For.*, 2013, vol. 37, no. 5, p. 14. https://doi.org/10.3906/tar-1210-28
31. Di Pietro, R., Di Marzio, P., Antonecchia, G., et al., Preliminary characterization of the *Quercus pubescens* complex in southern Italy using molecular markers, *Acta Bot. Croat.*, 2020, vol. 78, no. 2, pp. 107—115. https://doi.org/10.37427/botcro-2020-002
32. Di Pietro, R., Conte, A.L., Di Marzio, P., et al., Does the genetic diversity among pubescent white oaks in southern Italy, Sicily and Sardinia islands support the current taxonomic classification?, *Eur. J. For. Res.*, 2021, vol. 140, no. 2, pp. 1—17. https://doi.org/10.1007/s10342-020-01334-z
33. Neophytou, C., Bayesian clustering analyses for genetic assignment and study of hybridization in oaks: effects of asymmetric phylogenies and asymmetric sampling schemes, *Tree Genet. Genom.*, 2014, vol. 10, pp. 273—285. https://doi.org/10.1007/s11295-013-0680-2
34. Semerikova, S.A., Podergina, S.M., Tashev, A.N., and Semerikov, V.L., Phylogeography of oaks in the Crimea reveals Pleistocene refugia and migration routes, *Russ. J. Ecol.*, 2023, vol. 54, no. 3, pp. 197—212. https://doi.org/10.1134/S1067413623030049
35. Semerikova, S.A., Aliev, K.U., Semerikov, N.V., and Semerikov, V.L., Phylogeography of oak species in the Caucasus based on results of chloroplast DNA analysis, *Russ. J. Genet.*, 2023, vol. 59, no. 7, pp. 669—684. https://doi.org/10.1134/S1022795423070104
36. Semerikova, S.A., Tashev, A.N., and Semerikov, V.L., Genetic diversity and history of pedunculate oak *Quercus robur* L. in the east of the range, *Russ. J. Ecol.*, 2023, vol. 54, no. 5, pp. 423—438. https://doi.org/10.1134/S1067413623050089
37. Devey, M.E., Bell, J.C., Smith, D.N., et al., A genetic linkage map for *Pinus radiata* based on RFLP, RAPD and microsatellite markers, *Theor. Appl. Genet.*, 1996, vol. 92, no. 6, pp. 673—679. https://doi.org/10.1007/BF00226088

38. Pritchard, J.K., Stephens, M., and Donnelly, P., Inference of population structure using multilocus genotype data, *Genetics*, 2000, vol. 155, pp. 945—959.
39. Evanno, G., Regnaut, S., and Goudet, J., Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study, *Mol. Ecol.*, 2005, vol. 14, no. 8, pp. 2611—2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x
40. Li, Y.L. and Liu, J.X., StructureSelector: a web-based software to select and visualize the optimal number of clusters using multiple methods, *Mol. Ecol. Resour.*, 2018, vol. 18, no. 1, pp. 176—177. https://doi.org/10.1111/1755-0998.12719
41. Kopelman, N.M., Mayzel, J., Jakobsson, M., et al., CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K, *Mol. Ecol. Resour.*, 2015, vol. 15, pp. 1179—1191. https://doi.org/10.1111/1755-0998.12387
42. Puechmaille, S.J., The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem, *Mol. Ecol. Resour.*, 2016, vol. 16, pp. 608—627. https://doi.org/10.1111/1755-0998.12512
43. Peakall, R. and Smouse, P.E., GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update, *Bioinformatics*, 2012, vol. 28, no. 19, pp. 2537—2539. https://doi.org/10.1093/bioinformatics/bts460
44. Rousset, F., GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux, *Mol. Ecol. Notes*, 2008, vol. 8, pp. 103—106. https://doi.org/10.1111/j.1471-8286.2007.01931.x
45. Van Oosterhout, C., Hutchinson, W.F., Wills, D.P., and Shipley, P., MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data, *Mol. Ecol. Notes*, 2004, vol. 4, pp. 535—538. https://doi.org/10.1111/ j.1471-8286.2004.00684.x
46. Nei, M., Estimation of average heterozygosity and genetic distance from a small number of individuals, *Genetics*, 1978, vol. 83, pp. 583—590.
47. Rohlf, E.J., *NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System: Version 1.50*, New York: Exeter, 1988.
48. Vázquez, F.M., García, D., Márquez, F., and Vilaviçosa, C.M., Annotations to infrageneric nomenclature of *Quercus* L. (Fagaceae), *Fol. Bot. Extremadurensis*, 2023, vol. 17, pp. 7—64.
49. Kučera, P., *Quercus banatus* grows in Slovenia, *Thaiszia— J. Bot.*, 2019, vol. 29, no. 1, pp. 61—69. https://doi.org/10.33542/TJB2019-1-04
50. Di Pietro, R., Viscosi, V., Peruzzi, L., and Fortini, P., A review of the application of the name *Quercus dalechampii, Taxon*, 2012, vol. 61, no. 6, pp. 1311—1316. https://doi.org/10.1002/tax.616012
51. Kučera, P., New name for Central European oak formerly labelled as *Quercus dalechampii, Biologia*, 2018, vol. 73, no. 4, pp. 313—317. https://doi.org/10.2478/s11756-018-0048-z
52. Proietti, E., Filesi, L., Di Marzio, P., et al., Morphology, geometric morphometrics, and taxonomy in relict deciduous oaks woods in northern Italy, *Rend. Fis. Acc. Lincei*, 2021, vol. 32, pp. 549—564. https://doi.org/10.1007/s12210-021-01001-4
53. Papini, A., Simeone, M.C., Bellarosa, R., et al., *Quercus macranthera* Fisch. and Mey. ex Hohen. and *Quercus iberica* M. Bieb.: taxonomic definition and systematic relationships with European oaks inferred from nuclear internal transcribed spacer (ITS) data, *Plant Biosyst.*, 2011, vol. 145, no. 1, pp. 37—49. https://doi.org/10.1080/11263504.2010.502684
54. Tóth, E.G., Köbölkuti, Z.A., Cseke, K., et al., A genomic dataset of single-nucleotide polymorphisms generated by ddRAD tag sequencing in *Q. petraea* (Matt.) Liebl. populations from Central-Eastern Europe and Balkan Peninsula, *Ann. For. Sci.*, 2021, vol. 78, no. 43. https://doi.org/10.1007/s13595-021-01051-6
55. Jurkšienė, G., Baranov, O.Y., Kagan, D.I., et al., Genetic diversity and differentiation of pedunculate (*Quercus robur*) and sessile (*Q. petraea*) oaks., *J. For. Res.*, 2020, vol. 31, pp. 2445—2452. https://doi.org/10.1007/s11676-019-01043-3
56. Degen, B., Blanc-Jolivet, C., Mader, M., et al., Introgression as an important driver of geographic genetic differentiation within European white oaks, *Forests*, 2023, vol. 14, no. 12, p. 2279. https://doi.org/10.3390/f14122279