

# Genetic Diversity of Pedunculate Oak (*Quercus robur* L.) in Clonal Seed Orchards in Croatia, Assessed by Nuclear and Chloroplast Microsatellites

Ida Katičić Bogdan<sup>1\*</sup>, Davorin Kajba<sup>1</sup>, Zlatko Šatović<sup>2</sup>, Silvio Schüller<sup>3</sup>, Saša Bogdan<sup>1</sup>

(1) University of Zagreb, Faculty of Forestry, Department of Forest Genetics, Dendrology and Botany, Svetošimunska 25, HR-10000 Zagreb, Croatia; (2) University of Zagreb, Faculty of Agriculture, Department of Seed Science and Technology, Svetošimunska 25, HR-10000 Zagreb, Croatia; (3) Austrian Research Centre for Forests - BFW, Department of Forest Growth and Silviculture, Seckendorff-Gudent-Weg 8, A-1131 Vienna, Austria

\* Correspondence: e-mail: [ikaticic@sumfak.hr](mailto:ikaticic@sumfak.hr)

**Citation:** KATIČIĆ BOGDAN I, KAJBA D, ŠATOVIĆ Z, SCHÜLER S, BOGDAN S Genetic Diversity of Pedunculate Oak (*Quercus robur* L.) in Clonal Seed Orchards in Croatia, Assessed by Nuclear and Chloroplast Microsatellites. *South-east Eur for* 9 (1): 29-46. DOI: <https://doi.org/10.15177/see-for.18-09>

**Received:** 10 Apr 2018; **Revised:** 11 Jun 2018; **Accepted:** 15 Jun 2018; **Published online:** 30 Jun 2018

## ABSTRACT

**Background and Purpose:** Natural stands of pedunculate oak in Croatia have been delineated in seed areas, zones and regions. The current bylaw recommends that the transfer of reproductive material remains limited within zones, but that it is permitted within areas. Clonal seed orchards (CSOs) of pedunculate oak were established to increase genetic quality of seed and to acquire a more regular seed yield than in natural stands. In total 150 plus trees were selected within three seed regions. The selection included a number of favourable traits of tree size and stem quality. Three CSOs were planted with grafted plus-trees. We aimed to establish whether these orchards encompass enough genetic diversity to potentially produce genetically improved and sufficiently diverse reproductive material. We also wanted to assess neutral genetic differentiation between these orchards and compare it with the genetic diversity obtained from chloroplast DNA markers, depicting conserved lineages from recolonization routes. We wanted to investigate spatial genetic structure in the area of our research and provide additional information on the transfer of forest reproductive material.

**Materials and Methods:** Leaves were collected from all clones in the CSOs. Total genomic DNA was extracted and clones were analysed with eight nuclear and ten chloroplast microsatellite markers. Spatial autocorrelation analysis was performed with nuclear microsatellite data and original plus trees' coordinates, for each CSO separately, to determine whether shared favourable traits among the selected plus trees in smaller distances are the results of relatedness, which the sampling strategy tried to avoid.

**Results:** We found 28 chloroplast haplotypes belonging to four maternal lineages, and significant differentiation between CSOs, indicating origin from different refuges. Nuclear microsatellites' diversity in the CSOs is quite high and comparable to diversity found within a recent study of Croatian natural populations. Nuclear microsatellites did not show genetic differentiation between CSOs, i.e. between the seed regions and seed zones they represent. No genetic differentiation was found with nuclear microsatellites among haplotypic lineages. We found no genetic structure within the analysed regions.

**Conclusions:** Lack of differentiation between CSOs found with nuclear microsatellites confirms the permission for transfer of reproductive material between zones within the seed area 1 - Lowland Forests. If original differentiation between chloroplast haplotypic lineages was present after recolonization, it was erased by subsequent gene flow. Lack of genetic structure, with nuclear microsatellites within regions indicates successful sampling strategy.

**Keywords:** plus trees, clonal seed orchards, seed regions, genetic differentiation, haplotypic lineages, spatial autocorrelation analysis

## INTRODUCTION

Pedunculate oak (*Quercus robur* L.) is one of the ecologically and economically most important tree species in Croatia. Based on various ecological and management criteria, as well as research results, natural stands of pedunculate oak in Croatia were delineated in seed management units. According to the current bylaw on forest tree species provenances [1], they are divided into three seed areas, four seed zones and eight seed regions. To increase genetic quality of reproductive material and overcome irregular periodicity of masting in natural stands, a breeding programme for pedunculate oak improvement started in the 1980s [2]. Plus trees were selected in the seed area 1 - Lowland Forests, within three seed regions 1.1.2, 1.2.1, and 1.2.3, which belong to two seed zones, 1.1. and 1.2. (Table 1), based on multiple criteria of tree size and stem quality. Three clonal seed orchards (CSOs) were established with their vegetative copies (grafts).

The knowledge of the species' genetic diversity and the variability of different traits in various environmental conditions becomes especially important in the context of climate change. This particularly refers to the production of well-adapted forest reproductive material. Current research of quantitative traits' genetic variability of pedunculate oak in Croatia has shown that dominant component lies in intrapopulation variability [3-9].

Similar pattern of genetic variability (high intrapopulation and low interpopulation variability) and weak spatial structure of populations in greater distances can also be expected for nuclear microsatellites research on pedunculate oak, as well as sessile oak, when investigated populations belong to large uninterrupted complexes of these species [10, 11]. This was confirmed in recent research of natural pedunculate oak populations in Croatia with nuclear microsatellites, since only a few populations were significantly genetically differentiated and they did not belong to a more uniform complex with other populations [9]. In case of more fragmented populations, especially under strong anthropogenic influence, intrapopulation differentiation becomes higher, although still not as high as intrapopulation variability [12, 13].

Regarding plant species in general, biological factors such as limited gene flow by seed or pollen can lead to genetic differentiation even in continuous populations of species. In such cases, fertilization is not random, but rather conditioned by the distance between individuals [14]. Although the majority of other investigated plant species show pronounced genetic structures in short distances, this is rarely found in forest tree species. In these species the structure is weak or non-existent, sometimes described as almost random [15]. Interestingly, weak genetic structure is noticed in case of forest tree species with different reproductive systems and history, which can be attributed to their joint characteristics: long lifespan and stronger gene flow compared to other plant species. Selection can also contribute to creating genetic structure patterns and enhance or diminish genetic differentiation. Selection caused by diversity of microecologic conditions in the stand can trigger the creation of spatial patterns, while alternatively, in case of inbreeding depression, selection can enhance the dying off of inbred individuals and lead to weakening of genetic structure [16].

Regarding pedunculate oak, there is pronounced

asymmetry of gene flow by seed and pollen. Pollen is spread by wind, while seeds are large and heavy and can be carried to larger distances only by animals, humans or along river flows. Research shows that in pedunculate oak gene migration by pollen is approximately two hundred times stronger than migration by seed. This difference is manifested in weaker differentiation of populations when analysed with biparently inherited nuclear markers, while at the same time differentiation by uniparentally inherited cytoplasmic (in case of broadleaves chloroplast) markers can be substantial. Presumably, limited gene flow by seed primarily causes spatial patterns, but they are subsequently erased by constant and strong gene flow by pollen [15, 17, 18].

Research on correlation of chloroplast markers' diversity with nuclear molecular markers and quantitative traits showed complete incoherence of chloroplast lineages with quantitative traits, and weak concordance with allozyme frequencies [19]. It led to the following conclusions: (1) During the last glacial deciduous oaks were probably limited to three main refuges, which led to genetic differentiation of both nuclear and chloroplast genomes; (2) At the end of the glacial oaks migrated to the north, at which point spatial pattern of chloroplast diversity was established, which has remained highly conserved until today; (3) As oaks progressed in colonization towards central and northern Europe, gene flow by pollen reinstated communication between populations originating from western, or eastern refuges, which led to gradual loss of original differentiation in parts of nuclear genomes. During that process, differentiation in chloroplast genomes remained approximately the same; (4) Local selection pressures influencing the newly established populations gradually led to differentiation of parts of nuclear genomes whose new patterns were completely different from those immediately following the recolonization. During that whole period, chloroplast diversity remains conserved since it is not connected to quantitative traits. All that remains is a weak connection to nuclear markers that are not under strong influence of selection.

The aim of our research is the evaluation of two types of genetic diversities within and among three CSOs, established with grafts of plus trees selected within three legally delineated seed regions, belonging to two seed zones. The first type is genetic diversity of nuclear microsatellite markers, presumably neutral, highly polymorphic, biparently inherited markers that point to gene flow by both seed and pollen. The second type is diversity of maternally inherited chloroplast microsatellite markers, pointing to gene flow only by seed.

Genetic diversity analysis of clones in CSOs will demonstrate whether the selected parental populations, i.e. CSOs, consist of enough diverse and unrelated genotypes to potentially produce genetically diverse and improved seed for filling the natural stands and aiding their regeneration. Genotyping by nuclear microsatellites also lays ground for parentage analysis of crops produced in the CSOs, to determine the true efficacy of the CSOs in producing genetically diverse progeny. Although nuclear microsatellite markers in this study are selectively neutral, high genetic diversity of CSOs for these markers would also point to potentially wider adaptation potential, which is especially important in the context of climate and habitat change [20].

Considering the known positions of original plus trees in the stands, one of the aims is to analyse spatial distribution of nuclear microsatellite genetic diversity within each region. We want to investigate whether there is geographically influenced genetic structure at certain classes of distance, detected by a kinship degree of the selected trees within a region. This way we want to eliminate the possibility that multiple favourable traits shared among geographically close plus trees result from their kinship, which the sampling strategy of 50 m distance between the selected trees tried to avoid [2].

By using chloroplast microsatellites, which reflect the genetic diversity established by postglacial recolonization routes, we will analyse the diversity of haplotypes within and among regions. The results will demonstrate lesser or greater homogeneity of individual regions and possible anthropogenic influence in haplotype diversity of certain regions.

By comparing the analysis with nuclear and chloroplast markers we will try to establish whether the original differentiation by recolonization routes remained detectable in the analysed loci of the nuclear genome.

## MATERIALS AND METHODS

### Study Area

The leaves for DNA analysis were collected in three CSOs. Detailed information on CSOs, their position and the list of

forest districts where plus trees were selected is shown in Table 1, while the position of the CSOs is shown in Figure 1.

### DNA Extraction and Microsatellite Analysis

In spring we collected young leaves from one ramet of all clones in the three CSOs. The leaves were kept at -60°C until DNA extraction. We extracted total genomic DNA with GenElute Plant Genomic DNA Miniprep Kit (Sigma®).

We chose eight nuclear microsatellite markers for analysis: *ssrQrZAG96*, *ssrQrZAG108*, *ssrQrZAG87*, *ssrQrZAG112* and *ssrQrZAG31* from Kampfer *et al.* [21], and *ssrQpZAG110*, *ssrQpZAG9* and *ssrQpZAG36* from Steinkellner *et al.* [22]. PCRs were carried out following a slightly changed protocol from Kampfer *et al.* [21].

For chloroplast analysis, we chose 10 microsatellite markers: *μdt1*, *μdt3*, *μdt4*, *μcd4*, *μcd5*, *μkk3* and *μkk4* from Deguilloux *et al.* [23], and *ccmp2*, *ccmp10* and *ccmp6* from Weising *et al.* [24]. The PCRs were carried out following a slightly changed protocol from Deguilloux *et al.* [23].

Detailed description of the primers, multiplexes and PCR protocols can be found in Deguilloux *et al.* [23]. The 5' primers were labelled with different fluorescent dyes (WellRED Oligos, Proligo). All PCRs were carried out by PTC-100 thermal cycler (MJ Research®). Fragment separation was performed by capillary electrophoresis on CEQTM 8000, Genetic Analysis System (Beckman & Coulter), and alleles were scored with CEQTM 8000 Fragment Analysis Software.

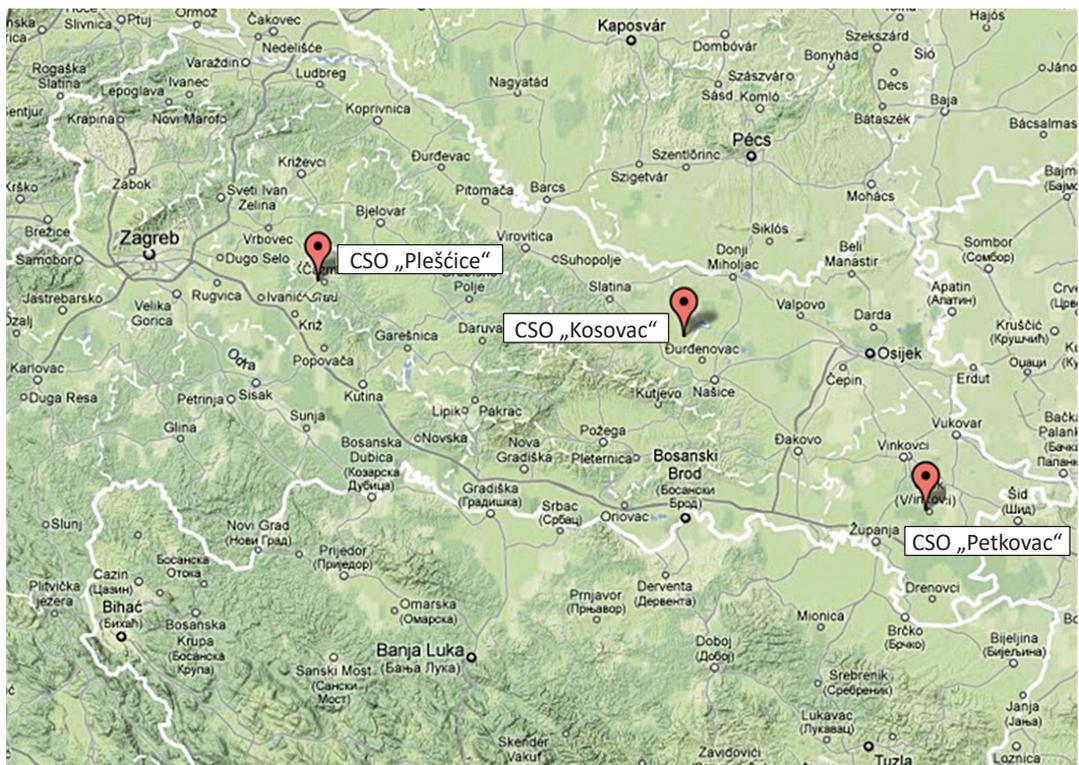


FIGURE 1. Position of *Quercus robur* clonal seed orchards in Croatia

**TABLE 1.** Information about *Quercus robur* clonal seed orchards (CSOs).

CSO	Year	Number of clones	Position	Seed area, zone and region	Selection of plus trees (No of trees in different FD)
Kosovac	1996	40	FA Našice“	SA 1. Lowland forests	FD „Slatina“ 26
			FD „Orahovica“	SZ 1.1 Podravina and Podunavlje	FD „Donji Miholjac“ 9
			MU „Obradovačke nizinske šume“	SR 1.1.2. Central Podravina	FD „Koška“ 4
					FD „Đurđenovac“ 1
Petkovac	2000	57	FA „Vinkovci“	SA 1. Lowland forests	FD „Gunja“ 19
			FD „Otok“	SZ 1.2 Posavina, central Croatia and Pokuplje	FD „Vrbanja“ 14
			MU „Slavir“	SR 1.2.1. Lower Posavina	FD „Otok“ 11
					FD „Županja“ 9
Plešćice	2001	53	FA „Bjelovar“	SA 1. Lowland forests	FD „Strošinci“ 4
			FD „Čazma“		FD „Čazma“ 15
			MU „Čazmanske nizinske šume“	SZ 1.2 Posavina, central Croatia and Pokuplje	FD „Bjelovar“ 10
				SR 1.2.3. Upper Posavina and Pokuplje	FD „Garešnica“ 10
					FD „Vrbovec“ 7
			FD „Grubišno Polje“ 5		
			FD „Veliki Grđevac“ 4		
			FD „Daruvar“ 2		

FA - Forest Administration, FD – Forest District, MU – Management Unit, SA – Seed Area, SZ – Seed Zone, SR – Seed Region

## Statistical Analysis

### A) Nuclear Microsatellites (nSSRs)

#### a1) Genetic Diversity and Differentiation

The total number of alleles per locus (Na), the observed HO and HE, expected heterozygosity and Polymorphism Information Content (PIC; [26]) for each microsatellite locus, as well as the average number of alleles (Navg), HO and HE in each region were calculated using PowerMarker v3.25 software [27]. The allelic richness (Nar) as the measure of the number of alleles per locus independent of sample size was calculated by FSTAT v.2.9.3.2 programme package [28], while the number of private alleles (Npr) per region was assessed by MICROSAT [29]. The estimates of Nar, HO and HE among regions were compared using the Kruskal-Wallis test in SAS Release 8.02 (SAS Institute, 2004).

To establish whether the analysed regions belong to Hardy-Weinberg equilibrium we used Wright's F statistics [30]. If there are deviations, the possible reasons are inbreeding within the regions or genetic differentiation between the regions. Entry data are allele frequencies of analysed markers in sampled regions. Potential heterozygote deficiencies, as a result of inbreeding, is described by fixation indices;  $F_{IS'}$  on individual level, and  $F_{ST}$  on population (in this case regional) level.

To calculate  $F_{IS}$  [31] for each locus in each region and to test congruence between the observed and expected genotypic frequencies under Hardy-Weinberg equilibrium we used GENEPOP v. 3.4 [32]. All probability tests are based on Markov chain method [33, 34], using 10000 steps of repeated memory, 100 data groups and 5000 repetitions per group. To correct it for multiple testing we used sequential Bonferroni adjustments [35, 36] implemented in the SAS Release 8.02 programme (SAS Institute, 2004).

Genetic differentiation among regions was estimated by FST using Weir and Cockerham method [31] as implemented in FSTAT. The significance of FST estimates was assessed by randomization. We calculated Nei's standard genetic distances [37] between pairs of regions with GENDIST programme implemented in PHYLIP ver 3.6b software [38].

#### a2) Analysis of Molecular Variance (AMOVA)

The proportion-of-shared-alleles distance [39] between pairs of individuals was calculated using MICROSAT and the distance matrix was subjected to the analysis of molecular variance (AMOVA; [40]) using ARLEQUIN version 2.000 [41]. AMOVA was used to partition the total microsatellite diversity into among-region and within-region components. The variance components were statistically tested by non-parametric randomisation tests using 10,000 permutations.

#### a3) Spatial Autocorrelation Analysis

Spatial autocorrelation analysis demonstrates relationships between specific measures of autocorrelation in different geographical distance classes, by means of correlograms, i.e. distograms. It is also possible to test significance of autocorrelation measure for each individual distance class. Since we analysed nuclear microsatellite markers, we used kinship coefficient for co-dominant markers by Loiselle *et al.* [42], described in detail in SPAGeDi 1.2g programme manual [43].

When plus trees were selected in the stands, their positions were mapped on sub-compartment management maps. Based on those sketches, we transferred their position on digitised management maps in ArcGis 9.2. (ESRI, 2007) programme and retrieved Gauss-Krüger coordinates for spatial analysis.

We performed spatial autocorrelation analysis with SPAGeDi 1.2g programme. Geographical distance between individuals for each region separately was calculated based on Gauss-Krüger coordinates of plus trees. Pairs of plus trees were classified into ten distance classes, containing approximately the same number of trees in each class. Based on nuclear marker analysis we calculated a matrix of kinship coefficient for co-dominant markers. The significance level of average kinship coefficients of individual distance classes was determined by 1000 permutations.

## B) Chloroplast Microsatellites (cpSSRs)

### b1) Lineages

#### b1.1) Cluster Analysis (UPGMA)

Genetic distances between haplotypes were calculated in MICROSAT [29] using absolute distance ( $D_{AD}$ ) metric based on the sum of the number of repeat differences between cpSSR haplotypes (described in detail in Deguilloux *et al.* [23]).

Cluster analysis based on distance matrix was performed using the unweighted pair-group method (UPGMA) as implemented in NEIGHBOR programme of the PHYLIP ver. 3.6b software [38]. The reliability of the UPGMA topology was assessed via bootstrapping [44] of over 1000 replicates generated by MICROSAT [29] and subsequently used in NEIGHBOR and CONSENSE programmes in PHYLIP.

#### b1.2) Median – Joining Network of Haplotypes

A Median-Joining (MJ) network [45] was constructed based on cpSSR haplotypes using the Network 4.5.1.6 programme [46]. This method is used in discovering very complex patterns of haplotypes. It uses parsimony criteria for finding median vectors, i.e. consensus sequences or mutually close marker sequences that present a biological equivalent of possible non-sampled or extinct ancestral haplotypes.

### b2) Genetic Diversity and Differentiation

Genetic diversity was estimated in each population and in the metapopulation as the number of haplotypes ( $n_h$ ), the number of haplotypes per the number of individuals ( $n_h/n$ ), the number of private haplotypes ( $n_{ph}$ ; i.e. haplotypes found in a single population), the effective number of haplotypes ( $n_e$ ), haplotype richness ( $n_r$ ), haplotype diversity ( $H_e$ ), and the frequency of the most common haplotype ( $f_1$ ). [47]

Haplotype richness ( $n_r$ ) as the measure of the number of haplotypes per population independent of sample size was calculated by rarefaction method [47] as implemented in Contrib [48].

An unbiased estimate of the haplotype diversity ( $H_e$ ) was calculated according to Nei [49].

Genetic differentiation was analysed by three different methods described in Pons and Petit [50, 51], using the Permut programme [52]:

- Unbiased estimation of  $G_{ST}$  [49] according to Pons and Petit [50], considering only haplotype frequencies (unordered haplotypes). Unbiased estimation of  $G_{ST}$  involved the calculation of the diversity of each region ( $h_r$ ), average within-region diversity ( $h_i$ ), and total diversity ( $h_t$ ).
- The estimation of  $N_{ST}$  [51] for which the genetic similarities between haplotypes (proportion of shared fragments) were taken into consideration (ordered haplotypes). The  $eof$   $N_{ST}$  involved the calculation of the diversity of each region ( $v_r$ ) by taking into account similarities between the haplotypes, average within-region diversity ( $v_i$ ), and total diversity ( $v_t$ ).
- The estimation of  $R_{ST}$  [53] for which the genetic distances at individual SSR loci were taken into consideration according to a microsatellite stepwise mutation model (SMM). The estimation of  $R_{ST}$  involved calculations similar to those used for the estimation of  $N_{ST}$  except that the distance between two haplotypes

was the sum (across all loci) of squared difference in the number of repeats [53]. This method was specifically designed for microsatellites by assuming the stepwise mutation model (SMM), while the  $N_{ST}$  method makes fewer assumptions about the nature and complexity of mutational patterns detected by fragment length analysis [51].

Detailed calculations of all methods can be found in Deguilloux *et al.* [23].

A comparison of differentiation and diversity measures for ordered ( $N_{ST}$  and  $R_{ST}$ ) vs. unordered ( $G_{ST}$ ) haplotypes was carried out according to Pons and Petit [51]. Significance was tested on the basis of 1,000 random permutations using the Permut programme [52].

### b3) Analysis of Molecular Variance (AMOVA)

The partition of cpSSR diversity within and among regions was studied using the Analysis of Molecular Variance (AMOVA, [40]) as implemented in ARLEQUIN [41], while the significance of AMOVA was tested based on 10,000 permutations of haplotypes among populations. Again, two genetic distance measures between pairs of haplotypes were used based on (1) the number of different alleles ( $F_{ST}$ ), as well as on (2) the sum of squared size differences ( $R_{ST}$ ).

### AMOVA of Haplotypic Lineages, with nSSRs

Based on the results of cluster analysis the haplotypes were grouped into lineages.

The partition of nSSR diversity within and among lineages (based on cpSSR) was tested using AMOVA in ARLEQUIN [41]. The variance components were tested statistically by non-parametric randomisation tests using 10,000 permutations.

## RESULTS

### A) Nuclear Microsatellites

#### a1) Genetic Diversity and Differentiation

Parameters of genetic diversity of analysed nuclear microsatellite loci are shown in Table 2. The values of PIC were in the range of 0.280 to 0.915, with an average of 0.739. Five out of eight markers had PIC values higher than the average, which makes them highly efficient in assessing genetic diversity of these regions. Markers *ssrQrZAG96*, *ssrQrZAG108* and *ssrQrZAG112* had lower levels of PIC, of which *ssrQrZAG96* was least polymorphic.

The greatest number of alleles was found in markers *ssrQrZAG112*, *ssrQrZAG31* and *ssrQrZAG87*, and the lowest in *ssrQrZAG96*, with an average number of 20.25 across all loci.

Expected heterozygosity ( $H_e$ ) ranged from low 0.288 in *ssrQrZAG96* to 0.915 in *ssrQrZAG31*, with an average of 0.750 across all loci. Observed heterozygosity ( $H_o$ ) in most loci ranged between the expected values, with an exception of locus *ssrQrZAG31*, and minor deviations in loci *ssrQrZAG108* and *ssrQrZAG87*. In loci *ssrQpZAG9* and *ssrQpZAG36*  $H_o$  was greater than  $H_e$ .

$F_{ST}$  values showed significant differentiation between regions only for loci *ssrQrZAG96* and *ssrQrZAG108*. On average, there was no significant differentiation between regions.

The values of inbreeding coefficient  $F_{IS}$  for regions and loci are shown in Table 3.  $F_{IS}$  values were significant for the following markers within regions: *ssrQrZAG31* values point to significant deviations from Hardy–Weinberg equilibrium (HWE) in all three regions. Values for *ssrQrZAG87* and *ssrQrZAG112* are also significant for region represented by CSO Kosovac, while *ssrQrZAG108* is only significant for CSO Plešćice.

Parameters of genetic diversity of the analysed regions (CSOs) are shown in Table 4. The regions are quite uniform across all parameters, with most private alleles belonging to CSO Petkovac, and least to CSO Kosovac. Allelic richness ( $N_{ar}$ ), corrected for different sized samples with rarefaction index,

does not show significant differentiation between regions, nor do  $H_O$  and  $H_E$ .

Genetic distance parameters are shown in Table 5. Nei's standard genetic distances between regions are almost equal, and differentiation between regions is not significant.

## a2) Analysis of Molecular Variance (AMOVA)

The results of AMOVA between regions are shown in Table 6. There is no significant differentiation between regions, with 99.96% of diversity lying at the within-region level, meaning that the majority of diversity is caused by differences between individual trees of all three regions.

**TABLE 2.** Allelic diversity of the microsatellite loci scored in three *Quercus robur* CSOs.

Locus	$N_a$	$H_O$	$H_E$	PIC	$F_{ST}^{\#}$
<i>ssrQpZAG9</i>	20	0.880	0.866	0.852	-0.002
<i>ssrQrZAG31</i>	25	0.260	0.920	0.915	0.000
<i>ssrQpZAG36</i>	18	0.884	0.872	0.860	-0.002
<i>ssrQrZAG87</i>	25	0.757	0.871	0.860	-0.002
<i>ssrQrZAG96</i>	9	0.248	0.288	0.280	0.014**
<i>ssrQrZAG108</i>	17	0.475	0.695	0.678	0.015*
<i>ssrQpZAG110</i>	22	0.589	0.603	0.593	-0.001
<i>ssrQrZAG112</i>	26	0.866	0.881	0.872	-0.002
Mean	20.25	0.620	0.750	0.739	0.001

$N_a$ , the total number of alleles;  $H_O$ , observed heterozygosity;  $H_E$ , expected heterozygosity; PIC, Polymorphic Information Content  
<sup>#</sup>Significant at  $P < 0.001$ ; \* Significant at  $P < 0.01$ ; \* Significant at  $P < 0.05$ ; not indicated - non-significant value

**TABLE 3.** Inbreeding coefficients  $f$  across eight microsatellite loci in three CSOs of *Quercus robur*.

Locus	CSO		
	Plešćice	Kosovac	Petkovac
<i>ssrQpZAG9</i>	0.047 <sup>ns</sup>	-0.041 <sup>ns</sup>	-0.033 <sup>ns</sup>
<i>ssrQrZAG31</i>	0.726 <sup>***</sup>	0.753 <sup>***</sup>	0.685 <sup>***</sup>
<i>ssrQpZAG36</i>	-0.040 <sup>ns</sup>	-0.060 <sup>ns</sup>	0.049 <sup>ns</sup>
<i>ssrQrZAG87</i>	0.141 <sup>ns</sup>	0.251 <sup>***</sup>	0.047 <sup>ns</sup>
<i>ssrQrZAG96</i>	0.150 <sup>ns</sup>	0.162 <sup>ns</sup>	0.080 <sup>ns</sup>
<i>ssrQrZAG108</i>	0.400 <sup>***</sup>	0.250 <sup>ns</sup>	0.260 <sup>ns</sup>
<i>ssrQpZAG110</i>	0.062 <sup>ns</sup>	-0.037 <sup>ns</sup>	0.033 <sup>ns</sup>
<i>ssrQrZAG112</i>	0.011 <sup>ns</sup>	0.076 <sup>***</sup>	0.000 <sup>ns</sup>
All loci	0.197 <sup>***</sup>	0.175 <sup>***</sup>	0.124 <sup>***</sup>

Inbreeding coefficients ( $f$ ) were calculated according to Weir and Cockerham (1984)

Significant deviations from Hardy–Weinberg equilibrium: “\*\*\*” corresponds to significance at the 1% nominal level, “\*\*” significance at the 5% nominal level, and “ns” depicts non-significant values after sequential Bonferroni corrections

**TABLE 4.** Genetic variation within *Quercus robur* CSOs at eight microsatellite loci

CSO	$n$	$N_{avg}$	$N_{ar}$	$N_{pr}$	$H_O$	$H_E$
Plešćice	50	14.50	12.67	18	0.617	0.755
Kosovac	40	13.25	12.58	12	0.604	0.730
Petkovac	54	15.25	13.06	26	0.635	0.738
p(KW)			0.89		0.98	0.93

$n$ , sample size;  $N_{avg}$ , average number of alleles per locus;  $N_{ar}$ , allelic richness;  $N_{pr}$ , number of private alleles;  $H_O$ , observed heterozygosity;  $H_E$ , expected heterozygosity; p(KW), probabilities of the Kruskal–Wallis tests among CSOs

**TABLE 5.** Nei's standard genetic distance (upper diagonal) and pairwise  $F_{ST}$  values (lower diagonal) among three *Quercus robur* CSOs.

CSO (region)	1	2	3
1 Plešćice		0.0466	0.0450
2 Kosovac	0.0004 <sup>ns</sup>		0.0464
3 Petkovac	0.0016 <sup>ns</sup>	0.0013 <sup>ns</sup>	

P-values as obtained by randomizations: "\*\*\*" corresponds to significance at the 1% nominal level, "\*\*" to significance at the 5% nominal level, and "ns" depicts non-significant values

**TABLE 6.** AMOVA analysis for the partitioning of microsatellite diversity among and within CSOs.

Source of variation	df	Variance components	Percentage of variation	$\Phi_{ST}$	P( $\Phi$ )
Among	2	0.00014	0.04	0.0004	0.423 <sup>ns</sup>
Within	141	0.336	99.96		

P( $\Phi$ ) -  $\Phi$ -statistics probability level after 10,000 permutations; ns – not significant

### a3) Spatial Autocorrelation Analysis

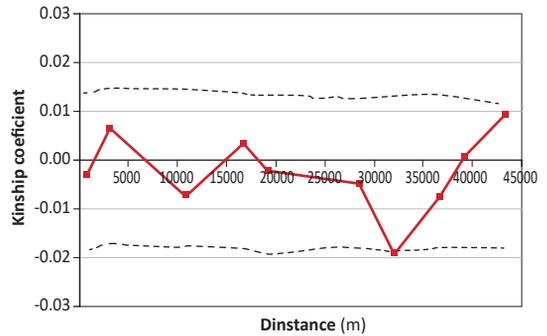
Spatial autocorrelation analysis was conducted separately for three regions, based on the analysis with nSSRs and Gauss–Krüger coordinates of original plus trees. The results are shown in Figure 2. Geographical distances were divided in ten classes in which pairs of trees were distributed, according to their mutual geographical distance. Average kinship coefficients for co-dominant markers between all pairs in each class [42] were calculated. Figure 2 shows upper and lower boundaries of 95% confidence intervals. Those intervals represent the boundaries above or below which, for a certain geographical distance, pairs of individuals are more or less related than pairs of randomly chosen individuals. The Figure shows that in the studied regions, based on our sample of selected plus trees, we found no geographical genetic structure. There was no class of geographic distance detected, containing pairs of individuals significantly more or less related than randomly chosen pairs of individuals.

## B) Chloroplast Microsatellites (cpSSRs)

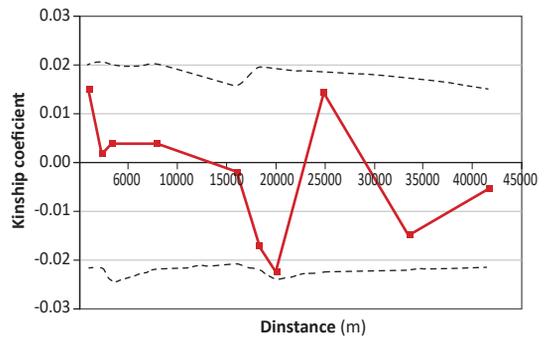
### b1) Distribution of Haplotypes

Table 7 shows the definition of cpSSR haplotypes found in three regions, based on nine cpSSR markers. Marker  $\mu\text{kk}3$  was monomorphic and was therefore excluded from further analysis. The numbers represent allele size in base pairs (bp). Table 7 also shows numbers of individuals within regions, belonging to different haplotypes. In haplotypes H07, H14, H15, H17 and H18 a specific long allele for  $\mu\text{dt}4$  was found. The authors in [23, 54, 55] did not mention encountering such an allele in their research, but it was definitely found in the inventory of Austrian oak populations [56]. When this allele is included in statistical analysis in its full size, it strongly influences calculations, especially those based on SMM models. In Provan *et al.* [57], the authors

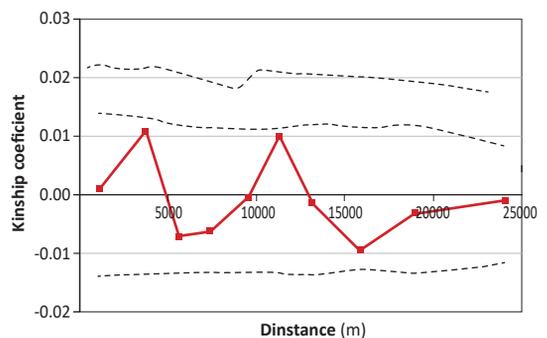
### CSO Plešćice



### CSO Kosovac



### CSO Petkovac

**FIGURE 2.** Charts of spatial autocorrelation analysis, separately for three CSOs, with 10 distance classes per CSO and eight nSSR loci. Dotted lines represent 95% confidence intervals.

also encountered an unusually long allele, sequenced it and determined that it was due to an indel of a sequence close to the microsatellite, so they treated this entire indel as a single mutation step. We suspect that something similar occurred in the sequence of  $\mu\text{dt}4$  marker, so following the example from Provan *et al.* [57] we also corrected the size of this allele, taking the difference of 18bp between allele 148 and 166 as a single mutation step in all further statistical analysis.

**TABLE 7.** Definition of cpSSR haplotypes and their distribution in the CSOs, based on nine cpSSR markers. The numbers represent allele lengths in base pairs.

cpSSR haplotypes	cpSSR markers									Distribution of haplotypes in the CSOs (number of individuals)			
	$\mu$ dt3	ccmp10	$\mu$ cd5	$\mu$ dt1	ccmp6	$\mu$ cd4	$\mu$ dt4	ccmp2	$\mu$ kk4	PI	K	P	$\Sigma$
H01	126	113	75	84	105	96	147	236	112	1	13	18	32
H02	125	113	75	84	105	96	147	236	112	0	0	23	23
H03	126	112	75	84	105	96	148	236	113	13	0	0	13
H04	127	113	75	84	105	96	147	236	112	2	9	0	11
H05	126	112	75	84	105	95	148	236	113	7	0	0	7
H06	128	113	75	84	105	96	147	236	112	0	6	0	6
H07	126	112	76	85	105	98	166*	237	112	0	3	1	4
H08	125	113	75	84	104	96	147	236	112	0	0	3	3
H09	126	113	75	84	105	97	147	236	112	1	1	1	3
H10	127	113	75	84	105	97	147	236	112	2	1	0	3
H11	126	112	75	84	104	96	148	236	113	2	0	0	2
H12	125	112	74	84	105	96	147	236	112	0	0	1	1
H13	125	112	75	84	105	95	148	236	113	1	0	0	1
H14	125	112	76	85	105	98	166*	237	112	1	0	0	1
H15	126	112	75	85	105	98	167*	237	112	1	0	0	1
H16	126	112	76	84	104	95	148	236	113	1	0	0	1
H17	126	112	76	85	104	98	166*	236	112	1	0	0	1
H18	126	112	76	85	104	98	166*	237	112	1	0	0	1
H19	126	113	75	84	103	96	147	236	112	0	0	1	1
H20	126	113	75	84	104	96	147	236	112	0	0	1	1
H21	126	113	75	84	104	97	147	236	112	1	0	0	1
H22	126	113	75	84	105	95	148	236	113	1	0	0	1
H23	126	114	75	84	105	96	147	236	112	0	0	1	1
H24	127	113	75	84	104	97	147	236	112	1	0	0	1
H25	127	113	75	84	105	96	147	235	112	0	1	0	1
H26	127	113	75	84	105	96	147	236	113	0	1	0	1
H27	127	114	75	84	105	96	147	236	112	0	1	0	1
H28	128	114	75	84	105	96	147	236	112	0	1	0	1
$\Sigma$										37	37	50	124

## b2) Lineages

### b2.1) Cluster Analysis (UPGMA)

Figure 3 shows UPGMA dendrogram of cpSSR haplotypes based on absolute distances ( $D_{AD}$ ). Based on bootstrap values it is visible that maternal lineages L3 and L4 create separate branches with strong statistical support, and that lineages L1 and L2 are more related.

The distribution of individuals across the lineages for each region is shown in Table 8. All but three individuals from CSO Petkovac belong to L1, while clones from CSO Kosovac are divided between lineages L1 and L2. Most clones from CSO Plešćice belong to the separate lineage L3, with some clones belonging to L1 and L2. For each CSO there is at least one individual belonging to L4, with the unusually long allele for  $\mu$ dt4.

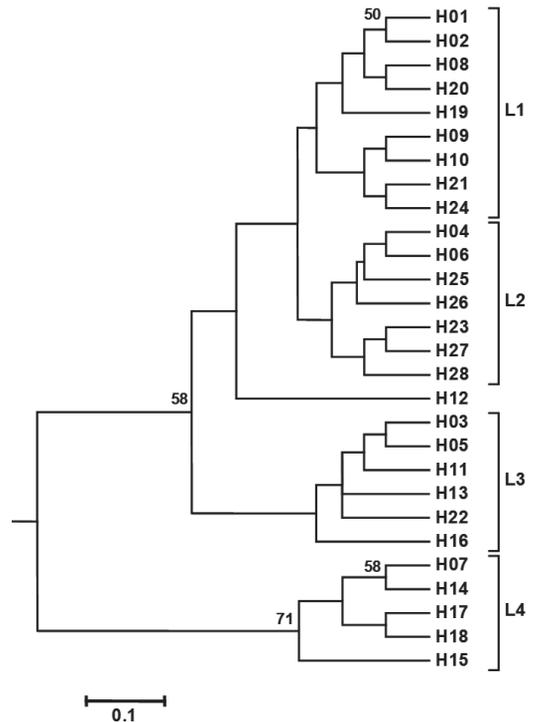
**TABLE 8.** The distribution of individuals from CSOs per haplotypic lineages.

Haplotypic lineages	Number / percentage of individuals in CSOs per lineages			$\Sigma$
	Plešćice	Kosovac	Petkovac	
1	6 (16.2 %)	15 (40.5 %)	47 (94.0 %)	64
2	2 (5.4 %)	19 (51.4 %)	1 (2.0 %)	26
H12	-	-	1 (2.0 %)	1
3	25 (67.6 %)	-	-	25
4	4 (10.8 %)	3 (8.1 %)	1 (2.0 %)	8
$\Sigma$	37	37	50	124

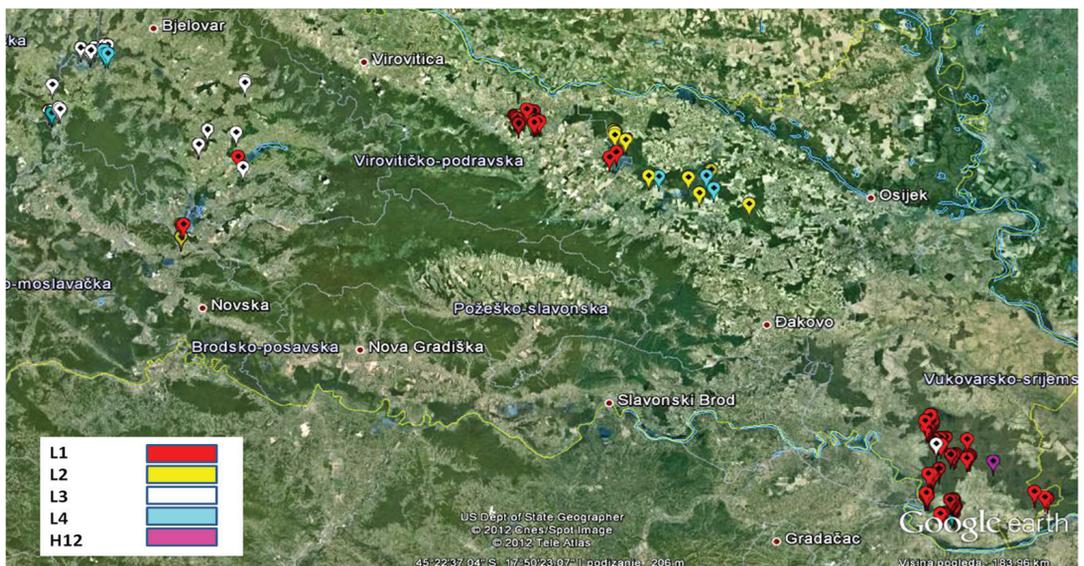
Spatial distribution of individual plus trees marked with the colour of lineage they belong to is shown in Figure 4. Trees belonging to different lineages are not randomly distributed in space, but rather grouped by lineages. For example, all plus trees from CSO Plešćice, belonging to lineages L1 and L2 were selected in Forest District "Garešnica", Management Unit "Međuvođe – Ilovski lug". L2 Trees included in CSO Kosovac are distributed evenly across the entire region's selection area; in the eastern part together with L4 trees, in the middle part with L1 trees, while L1 trees dominate over the western part. Almost all trees selected for CSO Petkovac belong to L1 lineage, with only individual L2 and L4 trees and a single tree with H12 haplotype, which does not belong to any lineages. In two CSOs where there is more than one L4 tree, they were also localized to a smaller area.

### b2.2) Median-Joining Network of Haplotypes

Figure 5 shows median-joining network of haplotypes. The sizes of the circles that mark haplotypes are proportional to the numbers of individuals belonging to that haplotype, and the colours in the circles mark the regions (CSOs) and are described in the legend. The size of the circle sector in a certain colour is proportional to the number of individuals of the corresponding region belonging to that haplotype. Lines connecting the haplotypes mark the number of mutation steps leading to the next haplotype. The colours of the lines mark cpSSR markers in which the change of size (bp) in that mutation step occurred, leading to the next haplotype. The colours are also described in the legend. In case when in a certain cpSSR the change in size bigger than 1bp occurred between two haplotypes, a number above the line marks the difference in bp. Haplotypes belonging to lineages defined by UPGMA algorithm are framed by blue dotted lines for comparison, and marked by numbers 1-4. Figure 5



**FIGURE 3.** UPGMA dendrogram of cpSSR haplotypes (cpSSR1S) based on absolute distance (DAD). Bootstrap support values of 1,000 replicates higher than 50% are given above the branches. Putative haplotype lineages (L1 to L4) are indicated on the right.



**FIGURE 4.** Spatial distribution of the selected plus trees in three seed regions, marked by their cpSSR haplotypic lineage. Left group – trees selected for CSO Plešćice (Seed Region Upper Posavina and Pokuplje), middle group - trees selected for CSO Kosovac (Seed Region Central Podravina), right group - trees selected for CSO Petkovac (Seed Region Lower Posavina).

also shows median vectors (mv 1-4), representing biological equivalents of possible non-sampled or extinct ancestral haplotypes, mostly found between groupings of sampled haplotypes.

**b3) Genetic Diversity and Differentiation**

Parameters of genetic diversity in the studied regions for cpSSRs are shown in Table 9. Greatest haplotype diversity presented as the number of haplotypes per the number of individuals ( $n_h/n$ ) is observed in CSO Plešćice (0.432), followed by CSO Kosovac (0.270), while CSO Petkovac was most uniform with 0.180 haplotypes per individual. Similar situation is found in other parameters: CSO Plešćice has

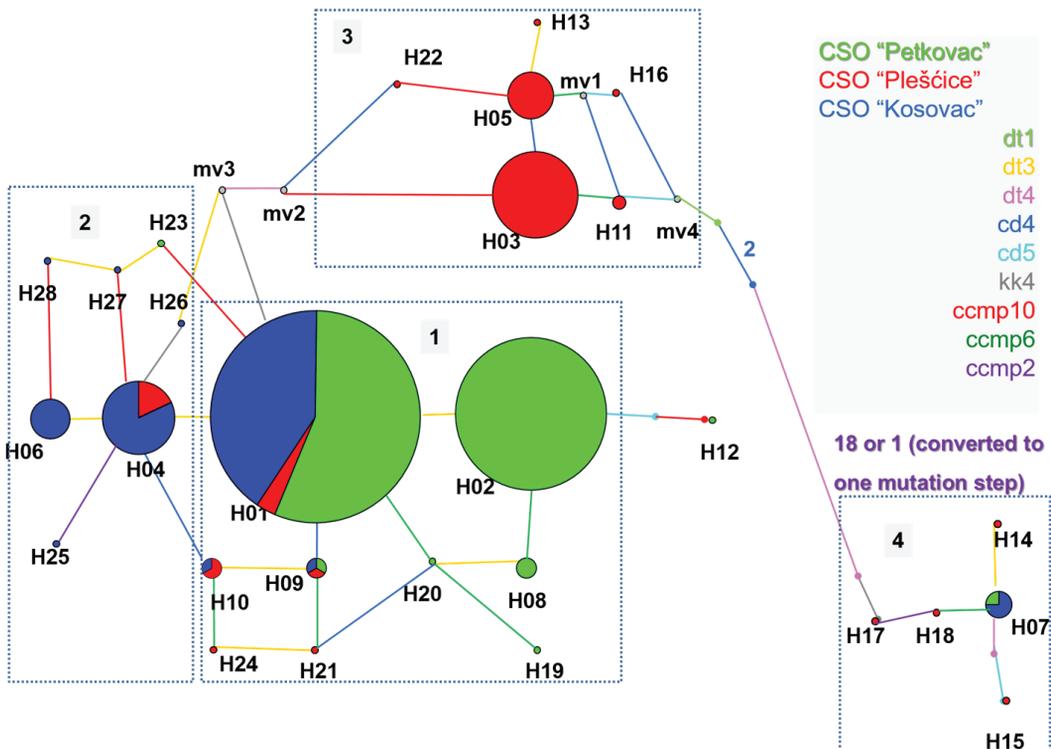
the greatest number of private haplotypes ( $n_{ph}$ ), double then other CSOs, as well as the highest effective number of haplotypes ( $n_e$ ), haplotype richness ( $n_{hr}$ ) and unbiased estimate of haplotype diversity ( $H_e$ ). Frequency of the most frequent haplotype was almost equal for CSOs Plešćice and Kosovac and somewhat higher for CSO Petkovac. In the metapopulation, the most frequent haplotype was H01, the overall  $n_h/n$  was 0.226 and  $n_e$  was 7.719.

Parameters of average within-region diversity ( $h_s, v_s$ ), total diversity ( $h_t, v_t$ ) and genetic differentiation ( $G_{ST}, N_{ST}$  and  $R_{ST}$ ) are shown in Table 10, as well as their comparison, considering three different methods for calculating genetic differentiation:

**TABLE 9.** CpSSR variation in three *Quercus robur* populations.

Population	n	$n_h$	$n_h/n$	$n_{ph}$	$n_e$	$n_{hr}$	$H_e$	$f_h$
Plešćice	37	16	0.432	12	5.680	15.000	0.847	0.351
Kosovac	37	10	0.270	5	4.548	9.000	0.802	0.351
Petkovac	50	9	0.180	6	2.880	6.425	0.666	0.460
Metapopulation	124	28	0.226	-	7.719	-	0.878	0.258

n - number of individuals;  $n_h$  - number of haplotypes;  $n_h/n$  - number of haplotypes per number of individuals;  $n_{ph}$  - number of private haplotypes;  $n_e$  - effective number of haplotypes;  $n_{hr}$  - haplotype richness;  $H_e$  - unbiased haplotype diversity;  $f_h$  - frequency of the most common haplotype (H03 in Plešćice, H01 in Kosovac, H02 in Petkovac, H01 in the metapopulation)



**FIGURE 5.** Median-joining haplotype network. H01-H28; haplotypes, mv1-mv4; median vectors.

- (A) unordered alleles (taking into consideration only the frequencies of haplotypes) –  $G_{ST}$ ;
- (B) ordered alleles (taking into consideration genetic similarities between haplotypes, i.e. the proportion of shared alleles) –  $N_{ST}$ ;
- (C) differences based on Stepwise Mutation Model (SMM) (taking into consideration genetic distances in individual cpSSR loci) –  $R_{ST}$ .

There were no significant differences between methods (A) and (B) and (A) and (C) for the calculation of parameters of average within-region diversity ( $h_s$  and  $v_s$ ) and total diversity ( $h_T$  and  $v_T$ ). However, there are significant differences between  $N_{ST}$  and  $G_{ST}$ , as well as between  $R_{ST}$  and  $G_{ST}$ , which points to the existence of significant genetic differentiation between these regions, for cpSSR loci.

#### a2) Analysis of Molecular Variance (AMOVA)

The results of AMOVA based on cpSSRs are shown in Table 11. Two methods were used for variance components partition: (1) The number of different alleles ( $F_{ST}$ ) and (2) the sum of squared size differences ( $R_{ST}$ ).

In both versions, the between-region component of diversity is significant, confirming that the regions are differentiated regarding cpSSR loci. Values obtained by both methods are similar and highly significant.

#### AMOVA of Haplotypic Lineages, with nSSRs

The results of AMOVA for partition of nSSR diversity in components within and between UPGMA haplotypic lineages are shown in Table 12.

In the within-lineage component lies 99.29% of diversity. The between-lineage component is not statistically significant.

#### DISCUSSION

##### Analysis with Nuclear Microsatellites (nSSRs) Genetic Diversity and Differentiation

Table 13 provides a comparison of  $H_o$  and  $H_e$  values of nSSR loci from this study with values from various studies of *Quercus robur* natural populations.

The values of genetic diversity parameters for most commonly used markers *ssrQpZAG9* and *ssrQpZAG36* are similar to those from other studies in Croatia [9] and all around Europe [15, 58-63]. Marker *ssrQrZAG87* was also in concordance with other studies, with slightly bigger difference between  $H_o$  and  $H_e$  than in Lepais *et al.* [64], but lesser than in Neophytou *et al.* [58]. Marker *ssrQpZAG110* was lower than in other studies and *ssrQrZAG112* was higher, but in both markers in this study,  $H_o$  and  $H_e$  were

**TABLE 10.** Levels of diversity and differentiation based on (A) unordered alleles, (B) ordered alleles, and (C) stepwise SSR differences.

Parameter	(A) Unordered alleles	(B) Ordered alleles <sup>a</sup>	(C) stepwise SSR differences <sup>b</sup>
Average within-population diversity	$h_s = 0.772$	$v_s = 0.641^{ns}$	$v_s = 0.633^{ns}$
Total diversity	$h_T = 0.945$	$v_T = 1.009^{ns}$	$v_T = 1.013^{ns}$
Genetic differentiation	$G_{ST} = 0.184$	$N_{ST} = 0.365^*$	$R_{ST} = 0.375^*$

<sup>a</sup>Test of the hypothesis that  $v_s$ ,  $v_T$ , and  $N_{ST}$  differ from  $h_s$ ,  $h_T$ , and  $G_{ST}$ , respectively

<sup>b</sup>Test of the hypothesis that  $v_s$ ,  $v_T$ , and  $R_{ST}$  differ from  $h_s$ ,  $h_T$ , and  $G_{ST}$ , respectively

<sup>ns</sup> not significant ( $P > 0.05$ ), \* significant at  $0.01 < P < 0.05$ ), \*\* significant at  $P < 0.01$ , \*\*\* significant at  $P < 0.001$

**TABLE 11.** AMOVA analysis for the partitioning of cpSSR diversity (A) based on the number of different alleles ( $F_{ST}$ ), and (B) based on squared size differences ( $R_{ST}$ ).

Analysis	Source of variation	df	Variance components	%Total variation	$\phi_{ST}$
(A) Based on the number of different alleles ( $F_{ST}$ )	Among populations	2	0.575	37.11	0.371***
	Within populations	121	0.974	62.89	
(B) Based on squared size differences ( $R_{ST}$ )	Among populations	2	0.940	38.97	0.390***
	Within populations	121	1.472	61.03	

<sup>ns</sup> not significant ( $P > 0.05$ ), \* significant at  $0.01 < P < 0.05$ ), \*\* significant at  $P < 0.01$ , \*\*\* significant at  $P < 0.001$

**TABLE 12.** AMOVA analysis for the partitioning of nuclear SSR diversity within and among putative cpSSR haplotypic lineages.

Classification	Source of variation	df	Variance components	%Total variation	$\phi_{ST}$
UPGMA	Among lineages	3	0.002	0.71	0.007 <sup>ns</sup>
	Within lineages	117	0.337	00.29	

<sup>ns</sup> not significant ( $P > 0.05$ ), \* significant at  $0.01 < P < 0.05$ ), \*\* significant at  $P < 0.01$ , \*\*\* significant at  $P < 0.001$

**TABLE 13.** Comparison of genetic diversity on nSSR loci between three CSOs from this study and natural populations of *Quercus robur* from other studies.

Locus		N (number of trees) / P (number of populations)	H <sub>o</sub>	H <sub>E</sub>	
ssrQpZAG9	<b>3 CSOs Croatia</b>	<b>117 / 3 (regions)</b>	<b>0.88</b>	<b>0.87</b>	
	Morić 2016 [8]	330 / 17		0.87	
	Neophytou <i>et al.</i> 2010 [57]	48 / 1	0.92	0.85	
		48 / 1	0.90	0.86	
	Muir and Schlotterer 2005 [58]	48 / 1	0.78	0.82	
		103 / 12	0.88	0.88	
	Cottrell <i>et al.</i> 2003 [60]	387 / 1	0.86	0.87	
		58 / 1	0.78	0.89	
	Bakker <i>et al.</i> 2003 [61]	48 / 1	0.81	0.85	
		48 / 1	0.88	0.84	
Buitveld <i>et al.</i> 2001 [62]	57 / 1	-	0.86		
Degen <i>et al.</i> 1999 [63]	210 / 1	0.88	-		
	83 / 1	0.84	-		
Streiff <i>et al.</i> 1998 [14]	183 / 1	0.79	0.85		
ssrQrZAG31	<b>3 CSOs Croatia</b>	<b>104 / 3 (regions)</b>	<b>0.26</b>	<b>0.92</b>	
	Muir and Schlotterer 2005 [58]	89 / 12	0.43	0.95	
ssrQpZAG36	<b>3 CSOs Croatia</b>	<b>121 / 3 (regions)</b>	<b>0.88</b>	<b>0.87</b>	
	Muir and Schlotterer 2005 [58]	105 / 12	0.80	0.87	
	Cottrell <i>et al.</i> 2003 [60]	387 / 1	0.85	0.89	
		58 / 1	0.81	0.94	
	Bakker <i>et al.</i> 2003 [61]	48 / 1	0.67	0.82	
		48 / 1	0.92	0.86	
	Buitveld <i>et al.</i> 2001 [62]	57 / 1	-	0.88	
	Degen <i>et al.</i> 1999 [63]	213 / 1	0.88	-	
		85 / 1	0.88	-	
	Streiff <i>et al.</i> 1998 [14]	183 / 1	0.79	0.88	
ssrQrZAG87	<b>3 CSOs Croatia</b>	<b>144 / 3 (regions)</b>	<b>0.76</b>	<b>0.87</b>	
	Morić 2016 [8]	330 / 17	-	0.88	
	Neophytou <i>et al.</i> 2010 [57]	48 / 1	0.43	0.75	
		48 / 1	0.34	0.87	
		48 / 1	0.50	0.84	
	Lepais <i>et al.</i> 2006 [64]	30 / 1	0.81	0.85	
	Muir and Schlotterer 2005 [58]	101 / 12	0.82	0.89	
	ssrQrZAG96	<b>3 CSOs Croatia</b>	<b>144 / 3 (regions)</b>	<b>0.25</b>	<b>0.29</b>
		Morić 2016 [8]	330 / 17	-	0.42
		Neophytou <i>et al.</i> 2010 [57]	48 / 1	0.35	0.33
48 / 1			0.42	0.39	
48 / 1			0.27	0.27	
Lepais <i>et al.</i> 2006 [64]		30 / 1	0.37	0.37	
Muir and Schlotterer 2005 [58]		101 / 12	0.34	0.42	
ssrQrZAG108		<b>3 CSOs Croatia</b>	<b>144 / 3 (regions)</b>	<b>0.48</b>	<b>0.70</b>
		Muir and Schlotterer 2005 [58]	96 / 12	0.40	0.81
ssrQpZAG110		<b>3 CSOs Croatia</b>	<b>129 / 3</b>	<b>0.59</b>	<b>0.60</b>
	Neophytou <i>et al.</i> 2010 [57]	48 / 1	0.75	0.74	
		48 / 1	0.81	0.84	
		48 / 1	0.81	0.74	
	Lepais <i>et al.</i> 2006 [64]	30 / 1	0.71	0.66	
	Muir and Schlotterer 2005 [58]	105 / 12	0.71	0.76	
ssrQrZAG112	<b>3 CSOs Croatia</b>	<b>135 / 3 (regions)</b>	<b>0.87</b>	<b>0.88</b>	
	Morić 2016 [8]	330 / 17	-	0.88	
	Neophytou <i>et al.</i> 2010 [57]	48 / 1	0.77	0.83	
		48 / 1	0.85	0.81	
		48 / 1	0.94	0.84	
	Lepais <i>et al.</i> 2006 [64]	30 / 1	0.74	0.66	
	Muir and Schlotterer 2005 [58]	102 / 12	0.63	0.75	

almost equal. Marker *ssrQrZAG96* had low values for both heterozygosities and PIC (Table 2). This was especially pronounced in this study, although the values were quiet low in other studies as well. This locus also showed significant differentiation between regions in this study (Table 2). The specificity of this marker is that it differentiates between pedunculate and sessile oak based on  $H_o$  and  $H_e$  on the population level. These values are systematically low for pedunculate oak and high for sessile oak populations from the same area [59]. Marker *ssrQrZAG112* shows similar patterns, but vice versa: high for pedunculate and low for sessile oak. In this study *ssrQrZAG112* values were very high, but same as the average in other Croatian populations [9]. Since plus trees were all selected in lowland areas, in a typical pedunculate oak habitat, with not much possibility for hybridization with sessile oak, very low *ssrQrZAG96* and very high *ssrQrZAG112* values are not surprising.

Markers *ssrQrZAG31* and *ssrQrZAG108* show significantly lower  $H_o$  than  $H_e$  in this study. Similar result, although not so pronounced, was shown by *ssrQrZAG31* in the study of twelve populations from Portugal, Spain, Ireland, Austria, France, Italy, Slovenia and Serbia [59]. In the same study, *ssrQrZAG108* had even bigger differences between  $H_o$  and  $H_e$  than in this study.  $F_{ST}$  value for this locus also demonstrates significant differentiation between the regions, although in lesser extent than in a similar study of Bosnian populations [65]. In that research all the loci showed significant differentiation between populations ( $F_{ST}$ ) and inbreeding coefficients ( $F_{IS}$ ), which the authors explained by intensive anthropogenic influence and populations' fragmentation throughout history, but they also presumed that the pattern of diversity was partially influenced by different ecological conditions and remnants of historical migration processes. In our study, only two loci are significantly differentiated between regions, while  $F_{ST}$  values for other loci, as well as the overall  $F_{ST}$ , are not significant (Table 2). Lack of differentiation usually points to effects of balancing evolutionary factors, like gene flow and balancing selection (similar selection pressures acting in all studied populations). It can also result from common descent, in absence of divergent evolutionary factors like random drift or divergent selection (adaptive differentiation due to different selection pressures acting in the studied populations).

According to Greogorius *et al.* [66], generally speaking it is expected that the effect of gene flow, as a balancing factor, is manifested in a greater number of loci, as compared to balancing selection, which usually affects only a few loci. The effect of common descent is also mainly visible in a smaller number of loci, because it is not likely for large parts of the genome to stay intact through a great number of generations. In this study, regions were analysed with eight nSSR loci, most of which demonstrate lack of differentiation (Table 2). Regions were not significantly differentiated for allelic richness, nor for  $H_e$  and  $H_o$ . The AMOVA between-region component was very small and not significant (Table 6), Nei's distances between regions were almost equal, and  $F_{ST}$  values between pairs of regions were not significant (Table 5). Taking all that into consideration, we can conclude that between these regions gene flow is the

prevailing balancing factor, especially because in this whole area, the distribution of pedunculate oak is continuous enough to provide a corridor for gene migration. This is confirmed by recent research of 17 Croatian natural oak populations, with 10 nSSRs [9], where also no differentiation between natural populations from this area was found. Studies of quantitative traits in open pollinated progeny trials from these CSOs show a similar pattern: no significant differentiation between regions [4-6, 8].

Inbreeding coefficients ( $F_{IS}$ ) show significant deviations from HWE for two loci in CSO Plešćice, three loci for CSO Kosovac and one locus in CSO Petkovac (Table 3). It led to overall significant deviations from HWE in all three regions.  $F_{IS}$  values for AG31 [59] are very high for all three CSOs, corresponding to great differences between  $H_o$  and  $H_e$  in this, but also in other studies [59]. High  $F_{IS}$  values point to inbreeding only in case when they are significant in most loci, which was the case in Bosnian populations [65]. Otherwise, high  $F_{IS}$  at only certain loci can point to the selection for these loci, selection for genes linked to these loci or the occurrence of null alleles. We used the MICRO-CHECKER programme [67] to check for genotyping mistakes and it showed a possibility that high  $F_{IS}$  values for *ssrQrZAG31*, *ssrQrZAG87* and *ssrQrZAG108* were caused by null alleles, especially in *ssrQrZAG31*, where estimated values for null allele frequencies rise up to 35% in all three regions. Therefore, in this locus obviously null alleles were the main reason for significant  $F_{IS}$  values. Estimated frequencies for locus *ssrQrZAG87* and *ssrQrZAG108* are much smaller, but there is reason to suspect that null alleles influenced  $F_{IS}$  values for these loci as well.

Another possible reason for significant  $F_{IS}$  values only on individual loci is, as mentioned above, the selection on those loci or the genes to which the locus is linked. Locus *ssrQrZAG87* is situated in linkage group LG2 of *Quercus robur* map, in one of the genome regions coding for bud burst. Considering its position, it might be possible that high  $F_{IS}$  value for this locus in CSO Kosovac comes from the uniformity of phenological forms in that CSO (almost all clones are early leafing [68-70]), as opposed to other two CSOs with a much wider range of leafing forms. However, in this study it is much more likely that null alleles caused those significant  $F_{IS}$  values. In other studies, significant  $F_{IS}$  values were also often accompanied by significant frequencies of null alleles [15, 58]. Additionally in this study, most  $F_{IS}$  values are positive, i.e. there is a slight, although for most loci not significant, homozygote surplus. Considering that for each CSO plus trees were selected on the level of seed region, i.e. a wider range that covers more populations (stands within forest districts and management units), joint analysis of all the trees from a CSO as one population (called region in this study) can lead to the apparent surplus of homozygotes, caused by subpopulation structure, or the so-called Wahlund effect. Overall genetic diversity for nSSR loci encompassed in these three CSOs is similar to the diversity of natural pedunculate oak populations of this area.

### Spatial Structure of nSSR Genetic Diversity

Genetic diversity spatial patterns in plants are mostly determined by their reproduction system and life form.

Weakest geographical structure, assessed with biparently-inherited markers, is found in forest tree species, which are mostly outcrossing and have a long lifespan [71]. Considering that oaks belong to that group, studies with nSSRs in oaks have confirmed weak spatial structure of their diversity, usually confined to 30–50 m distances between trees [15, 61, 71–73]. Weak spatial structure in oaks is probably the result of two opposing forces (not considering selection, as an additional factor forming genetic diversity). First, limited natural seed-flow (by gravity and animals) results in groups of related individuals. On the other hand, strong pollen-flow has a reverse balancing effect and disrupts the creation of related groups, resulting from limited seed-flow. Computer simulations [73] confirmed this scenario. In another study [15], it is further confirmed that weak geographic structure in oaks also results from weak survival of inbred individuals, due to inbreeding depression [15, 60]. Considering all that, the results of spatial analysis are as expected, especially since one of the selection criteria was the distance of minimum 50 m between the selected trees. Nonetheless, since there is significant pollen-flow in pedunculate oak, and since the trees were not sampled randomly, but selected based on multiple shared favourable traits, we wanted to confirm that similar phenotypes of trees from closer distance classes are not the result of their relatedness. The results confirmed that pairs of individuals within any distance class are not significantly more related than pairs of randomly sampled individuals (Figure 2).

### Chloroplast Microsatellites (cpSSR)

In angiosperms, chloroplast genomes are inherited maternally and represent a useful tool for analysis and differentiation of populations considering postglacial recolonization routes.

A comprehensive research of chloroplast diversity of main European oak species from 2613 populations was conducted with PCR-RFLP analysis. Thirty-two cpRFLP variants were found grouped in six main maternal lineages. Their distribution, along with some results of palynological studies, were used to reconstruct recolonization routes from refuges after the last glacial [17, 18].

Some Croatian populations were also included in the study by Bordács *et al.* [74], and analysed in detail by Slade *et al.* [75]. In the area of Croatia, Bosnia and Herzegovina, Montenegro, Albania and Kosovo they found seven cpRFLP haplotypes, four belonging to Balkan refuge lineage, two to Apennine, and one to Pyrenees or Southern Alps lineage. In general, total diversity, as well as within-population chloroplast diversity of the Balkan area, is one of the highest in Europe [18].

Chloroplast markers can be of great use in tracking geographical origin of individuals and populations and therefore useful for certification of seed sources, control of wood and timber market, forensics etc., especially for species of great economic value like pedunculate oak [76, 77]. Since cpRFLP analysis is complicated, laborious, demands high quality DNA and sometimes has replicability issues, there was a need for a simpler method. The cpSSR markers were developed and comparatively tested with cpRFLP markers in the same populations [23, 55]. It has been found that the

results are abutting, if not more precise with cpSSRs, with resolution increasing with added cpSSR loci.

Total genetic diversity ( $h_T$ ) of these three regions, assessed with nine cpSSR markers in 124 individuals is very high ( $h_T=0.95$ ), the same as in the analysis of overall Croatian natural populations [9] (340 individuals from 17 populations). The  $h_T$  of French populations with 10 cpSSRs was 0.65, but based only on 30 individuals, targeted to include certain cpRFLP haplotypes. In a study of Grivet *et al.* [54] on 367 individuals from 92 populations from France and the Iberian Peninsula, with six cpSSR markers, eleven haplotypes were found. If we had used only these six markers, we would have gotten 22 haplotypes for our three regions and 124 individuals. This points to great cpSSR diversity of our studied area. Since Croatia is situated on the crossroad of different recolonization routes [75], relatively close to some main refuges (Balkan and Apennine peninsula), it is not surprising to find great and well-preserved haplotypic diversity. The reason for differences between AMOVA partition of diversity between and within regions in this study (within-region diversity greater than between-region) and the study of overall Croatian populations [9] is in the sampling. We had many plus trees sampled in three rather close regions, compared to a smaller number of trees in more separate populations in that study, so it is expected that between population the component in Morić [9] will be higher. Nevertheless, both in this study and in Morić [9], differentiation is highly significant.

Considering a great number of analysed individuals per region and nine analysed cpSSRs, in this study resolution is high, so 28 haplotypes were grouped by UPGMA in four maternal lineages. Comparing our results with populations from study of Slade *et al.* [75], belonging to the same seed regions (population Zdenački Gaj for CSO Plešćice, Slatina and Đurđenovac for CSO Kosovac and Vrbanja and Gunja for CSO Petkovac) we presume that more closely related cpSSR lineages L1 and L2 belong to cpRFLP lineage 5, which is not clearly delineated and originates either from Balkan or Apennine refuge. In both studies (this study and study of Slade *et al.* [75]) most diverse region is the one where trees were selected for CSO Plešćice, were lineages from all three refuges come in contact. The region for CSO Kosovac was somewhat poorer in haplotypes, while CSO Petkovac region was the most uniform. Lineages L3 and L4 are strongly differentiated both from L1 and L2 and between each other, so it is quite possible that they originate from other refuges and belong to cpRFLP lineages 2 (Apennine refuge) or 7 (Pyrenees or Southern Alps refuge). Lineage L3 is mostly found in northwestern part of Croatia (CSO Plešćice) where, according to Slade *et al.* [75] both cpRFLP 2 and cpRFLP 7 lineage is found, with 7 being a bit more frequent.

In spite of the correction made in this study for the long allele (see Material and Methods), haplotypes of L4 lineage were strongly grouped together, pointing to their different origin from L1 and L2 lineages, as well as L3.

The cpSSR haplotype distribution can be formed not only by recolonization routes, but also by reproductive material transfer. For example, in the CSO Plešćice region in management unit "Međuvođe – Ilovski lug", we found haplotypes of L1 and L2 lineages, common in a more eastern

forest complex. Also in forest district "Slatina" (CSO Kosovac) almost all selected trees belong to L1, although this forest district is situated further west from the dominant complex of this lineage. Moreover, in the same region to the east, which is closer to the L1 dominant complex, L2 individuals are prevailing. While it is possible that such distribution is caused by recolonization route admixture or single long-distance dispersal events [78], we cannot exclude the possibility of seed transfer from eastern to further western regions, especially because L1 individuals are mostly found within a complex known for pedunculate oak stem quality, for example in Spačva Basin. Reproductive material from that area was exported to other states for its quality. For example, Slavonian pedunculate oak was intensively planted in Germany at the end of the 19th century and entire stands of that origin are still present in that area, recognizable by stem quality and late flushing. Therefore, it is quite possible that reproductive material from that area was transferred also within Croatia. Based on cpRLFP analysis across Europe, anthropogenic influence on the distribution of genetic diversity of pedunculate oak is stronger than on any other oak species because of its economic value [17, 18].

#### AMOVA of Haplotypic Lineages, with nSSRs

In an extensive overview of studies of different genomes [79], the authors compared diversities of chloroplast, mitochondrial and nuclear markers in different plant species and found that in most cases differentiation found with cytoplasmic markers does not correspond to the one found with nuclear markers. Also, in oak species, gene flow by pollen is approximately 200 times greater than by seed. This difference is manifested in weak differentiation of populations by nuclear, biparentally inherited markers, as compared to maternally inherited cytoplasmic markers. It is presumed that primarily nuclear diversity corresponded to cytoplasmic markers, but was later erased by strong gene flow by pollen and rearranged by selection pressures.

According to Kremer *et al.* [19], parallel analysis with nuclear and cytoplasmic markers eventually shows weak concordance when diversity of nuclear markers is not strongly influenced by selection. With this in mind, we wanted to compare cpSSR diversity with presumably neutral nSSR loci. The assessment of genetic differentiation is hugely influenced by the choice of markers and by sampling [79]. In our study, considering the small number of sampled regions and many sampled trees per region, differentiation for nSSR was virtually non-existing. As mentioned before, there is a corridor for gene flow between these regions, which obviously led to balancing genetic diversity of nSSRs in the entire studied area. Considering weaker possibilities of seed transfer between regions compared to pollen, there was significant differentiation for cpSSRs. AMOVA for partitioning nSSR diversity between UPGMA chloroplast lineages showed that only about 1% is explained by between-lineage components and it was not significant. Obviously, strong gene flow by pollen erased recolonization patterns of nuclear diversity.

## CONCLUSIONS

Molecular analysis with nuclear and chloroplast microsatellite markers of plus trees selected in three regions (upper Posavina, lower Posavina and Pokuplje, and central Podravina) within two seed zones (first two regions – seed zone Posavina, central Croatia and Pokuplje, third region – seed zone Podravina and Podunavlje) and included in three clonal seed orchards of pedunculate oak led to the following findings:

1. Analysis with eight nSSRs shows homogeneity of presumably neutral genetic diversity in the entire studied area, with no genetic differentiation between regions. The regions are not differentiated by genetic parameters, and AMOVA between-region component was very small and not significant. The results point to the existence of strong gene flow between regions, as prevailing balancing factor in this area.
2. Some nSSR loci (AG31, *ssrQrZAG108* and AG87) displayed deviations from Hardy-Weinberg equilibrium in some or all CSOs. The most probable cause of increased homozygosity in these loci are null-alleles. Additional reason for positive  $F_{IS}$  values for these and other loci is probably the sampling method. The analysis was conducted on the level of seed regions, which are themselves composed of more populations, which probably resulted in substructure of regions, i.e. Wahlund effect.
3. Spatial autocorrelation analysis, with nSSRs and Gauss-Krüger coordinates of original plus trees showed that for no class of mutual distance between the selected plus trees within regions pairs of trees were more related than randomly selected pairs of trees. This points to good sampling strategy and successful avoidance of selecting related genotypes.
4. Molecular analysis with nine polymorphic cpSSRs resulted in 28 haplotypes, belonging to four maternal lineages, by UPGMA algorithm. Lineages L1 and L2 are more related and lineages L3 and L4 are each grouped separately, with greater support. With these markers, significant differentiation was found between regions. Greatest haplotypic diversity was found in CSO Plešćice, smaller in CSO Kosovac and the smallest in CSO Petkovac. Probable cause is the distribution and admixture of recolonization lines from different refuges, which is more pronounced in western parts of Croatia, and possible reproductive material transfer. Analysis with these markers provides good basis for tracking the origins of reproductive material.

These findings lead to the following conclusions and recommendations.

The results of this study with nSSRs show no differentiation between regions in this area and are in concordance with previous studies on quantitative traits [3-9] and a recent study of overall Croatian populations with nSSRs [9]. CpSSR

analysis showed geographic structure of haplotypes, but it is connected to recolonization routes and according to another research [19] has no connection with adaptively important traits. Moreover, in this research, smaller differentiation for cpSSRs was observed between the regions belonging to other seed zones (CSO Kosovac and CSO Petkovac) than between those belonging to the same zone (CSO Plešćice and CSO Petkovac). Therefore, the change implemented in current Croatian bylaw on forest tree species provenances (NN 147/11, 96/12, 115/14 and 114/15), by which the transfer of forest reproductive material of pedunculate oak is permitted within the entire seed area, and not just within the seed zone, is justified. This is especially true in the light of climate change, when needs for changes in policy of forest reproductive material transfer are discussed on much larger scales, on the level of whole Europe [80].

We conclude that genetic diversity levels encompassed in these CSOs are comparable to those in natural populations [9] and provide a good basis for the production of quality and genetically diverse progeny, if other conditions for optimal reproduction are achieved. For now, the clones in these CSOs still have not reached their full size and the clones are not equally contributing to the crops [25]. Genotyping with nSSRs also provides the foundation for molecular analysis of the progeny, which would provide a very precise estimation

of genetic diversity produced in the CSOs in the given year, and level of contamination with pollen from outside sources. The results from such analysis could then be compared to simpler and much cheaper estimations, based on flowering and fruiting, to find the most appropriate method for assessing genetic diversity of crops. Along with flowering and fruiting observations, it is necessary to regularly observe climatic and other environmental conditions in the CSOs, so that the observed trends could be correlated to various factors and possibly explained. The clones in some of these orchards are grouped by phenology, which creates substructures and prevents random reproduction [69, 70]. According to our results, phenologically compatible clones from all three orchards could be combined in a separate orchard, in order to achieve maximum synchrony and enhance genetic diversity of the crops.

### Acknowledgments

We thank Univ. Prof. Dr. Thomas Geburek and Thomas Thalmayr from the BFW Institute in Vienna for all the help in the laboratory part of the study and Msc. Marina Antulov, Msc. Bernarda Bukovac, and Msc. Mladen Šimunić from Croatian Forests Ltd. for their help in collecting samples.

### REFERENCES

1. Bylaw on forest tree species provenances (NN 147/11, 96/12, 115/14 and 114/15) *Official gazette* 147/11 (document 2987); 96/12 (document 2157); 115/14 (document 2186); 114/15 (document 2183)
2. VIDAKOVIĆ M 1996 Establishment of clonal Seed Orchard of Pedunculate Oak (in Croatian and English). In: KLEPAC D (ed) Pedunculate oak (*Quercus robur* L.) in Croatia. Croatian Academy of Sciences and Arts, Science center Vinkovci and "Croatian Forests" Ltd., Vinkovci, Zagreb, Croatia, pp 127-138
3. PERIĆ S, GRAČAN J, DALBELO-BAŠIĆ J 2000 Flushing variability of pedunculate oak (*Quercus robur* L.) in provenance experiment in Croatia. *Glas Šum Pokuse* 37: 395-412
4. BOGDAN S, KAJBA D, KATIČIĆ I 2004 Genetic Variation in Growth Traits in a *Quercus robur* L. Open-Pollinated Progeny Test of the Slavonian Provenance. *Silvae Genet* 53 (5-6): 198-201
5. BOGDAN S, BEDENIKOVIĆ D, IVANKOVIĆ M 2008 Results Of The Field Trial With Open-Pollinated Families From Pedunculate Oak (*Quercus robur* L.) Plus Trees Originating From The Seed Region "Central Podravina" (in Croatian with English summary). *Rad – Šumar Inst Jastrebar* 43 (2): 93-114
6. KAJBA D, BOGDAN S, KATIČIĆ I 2006 Estimation of genetic gain for vigorous growth by clonal seed orchards of pedunculate oak (*Quercus robur* L.). *Glas Šum Pokuse* 5: 251-260
7. BALLIAN D, MEMIŠEVIĆ M, BOGUNIĆ F, BAŠIĆ N, MARKOVIĆ M, KAJBA D 2010 Morphological Variability Of Pedunculate Oak (*Quercus robur* L.) In The Region Of Croatia And Western Balkans (in Croatian with English summary). *Sumar list* 134 (7-8): 371-386
8. KAJBA D, KATIČIĆ I, BOGDAN S 2011 Estimation of Genetic Parameters in Open Pollinated Progeny Trials from Plus Trees of Pedunculate Oak (*Quercus robur* L.) Selected in Posavina and Podravina and Podunavlje Seed Zones. *Croat J For Eng* 32 (1): 190-192
9. MORIĆ M 2016 Genetic diversity of pedunculate oak (*Quercus robur* L.) in field trials with progeny from selected seeds stands (in Croatian with English summary). PhD thesis, University of Zagreb, Faculty of Forestry, Zagreb, Croatia, 243 p
10. 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. *Mol Ecol* 19(13): 2626-2639. DOI: <https://doi.org/10.1111/j.1365-294X.2010.04631.x>
11. MARIETTE S 2001 Measurement of Intra and Inter-Population Genetic Diversity: Intra-Genome Sampling, Choice of Markers and Experimental Applications on *Quercus petraea*, *Quercus robur* and *Pinus pinaster*. PhD thesis, ENGREF, Paris, France, 90 p
12. ORTEGO J, BONAL R, MUÑOZ A 2010 Genetic Consequences of Habitat Fragmentation in Long-Lived Tree Species: The Case of the Mediterranean Holm Oak (*Quercus ilex* L.). *J Hered* 101 (6): 717-726, DOI: <https://doi.org/10.1093/jhered/esq081>
13. BRUSCHI P, VENDRAMIN GG, BUSSOTTI F, GROSSONI P 2003 Morphological and Molecular Diversity among Italian Populations of *Quercus petraea* (Fagaceae). *Ann Bot-London* 91 (6): 707-716. DOI: <https://doi.org/10.1093/aob/mcg075>
14. WRIGHT S 1969 Evolution and the Genetics of Natural Populations. The University of Chicago Press, Chicago, USA
15. STREIFF R, LABBE T, BACILIERI R, STEINKELLNER H, GLÖSSL J, KREMER A 1998 Within-Population Genetic Structure in

- Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. Assessed with Isozymes and Microsatellites. *Mol Ecol* 7 (3): 317-328. DOI: <https://doi.org/10.1046/j.1365-294X.1998.00360.x>
16. HAMRICK JL, GODT MJW 1996 Effects Of Life History Traits On Genetic Diversity In Plant Species. *Philosophical Transactions of the Royal Society in London Series B*, 351: 1291-1298
  17. PETIT RJ, BREWER S, BORDÁCS S, BURG K, CHEDDADI R, COART E, COTTRELL J, CSAIKL UM *et al.* 2002 Identification of refugia and post-glacial colonization routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecol Manag* 156 (1-3): 49-74. DOI: [https://doi.org/10.1016/S0378-1127\(01\)00634-X](https://doi.org/10.1016/S0378-1127(01)00634-X)
  18. PETIT RJ, CSAIKL UM, BORDÁCS S, BURG K, COART E, COTTRELL J, VAN DAM B, DEANS JD *et al.* 2002 Chloroplast DNA variation in European white oaks: phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecol Manag* 156 (1-3): 5-26. DOI: [https://doi.org/10.1016/S0378-1127\(01\)00645-4](https://doi.org/10.1016/S0378-1127(01)00645-4)
  19. KREMER A, KLEINSHMIT J, COTRELL J, CUNDALL EP, DEANS JD *et al.* 2002 Is There a Correlation between Chloroplastic and Nuclear Divergence, or What are the Roles of History and Selection on Genetic Diversity in European Oaks? *Forest Ecol Manag* 156 (1-3): 75-87. DOI: [https://doi.org/10.1016/S0378-1127\(01\)00635-1](https://doi.org/10.1016/S0378-1127(01)00635-1)
  20. KAHILAINEN A, PUURTINEN M, KOTIAHO JS 2014 Conservation Implications of Species–Genetic Diversity Correlations. *Global Ecology and Conservation* 2: 315-323. DOI: <https://doi.org/10.1016/j.gecco.2014.10.013>
  21. KAMPFER S, LEXER C, GLOSSL J, STEINKELLNER H 1998 Characterization Of (GA), Microsatellite Loci From *Quercus robur*. *Hereditas* 129 (2): 183-186. DOI: <https://doi.org/10.1111/j.1601-5223.1998.00183.x>
  22. STEINKELLNER H, FLUCH S, TURETSCHKE E, LEXER C, STREIFF R, KREMER A, BURG K, GLÖSSL J 1997 Identification and characterization of (GA/CT)<sub>n</sub>- microsatellite loci from *Quercus petraea*. *Plant Mol Biol* 33 (6): 1093-1096. DOI: <https://doi.org/10.1023/A:1005736722794>
  23. DEGUILLLOUX MF, DUMOLIN-LAPEGUE S, GIELLY L, GRIVET D, PETIT RJ 2003 A set of primers for the amplification of chloroplast microsatellites in *Quercus*. *Mol Ecol Notes* 3 (1): 24-27. DOI: <https://doi.org/10.1046/j.1471-8286.2003.00339.x>
  24. WEISING K, GARDNER RA 1999 A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* 42 (1): 9-19. DOI: <https://doi.org/10.1139/g98-104>
  25. KATIČIĆ BOGDAN I 2012 Genetic diversity of pedunculate oak (*Quercus robur* L.) in clonal seed orchards in Croatia (*in Croatian with English summary*). PhD thesis, University of Zagreb, Faculty of Forestry, Zagreb, Croatia, 166 p
  26. BOTSTEIN D, WHITE RL, SHOLNICK M, DAVID RV 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32 (3): 314-331
  27. LIU K, MUSE SV 2005 PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21 (9): 2128-2129. DOI: <https://doi.org/10.1093/bioinformatics/bti282>
  28. GOUDET J 2002 FSTAT, a program for Windows (95 and above) to estimate and test gene diversities and fixation indices (version 2.9.3). Institut d'Ecologie, Bâtiment de Biologie, Université de Lausanne, Dorigny, Switzerland
  29. MINCH E, RUIZ-LINARES A, GOLDSTEIN D, FELDMAN M, CAVALLI-SFORZA LL 1997 MICROSAT: a computer program for calculating various statistics on microsatellite allele data, ver. 1.5d. Stanford University, Stanford, CA, USA
  30. WRIGHT S 1931 Evolution in Mendelian populations. *Genetics* 16 (2): 97-159
  31. WEIR BS, COCKERHAM CC 1984 Estimating F-statistics for the analysis of population structure. *Evolution* 38 (6): 1358-1370. DOI: <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x>
  32. RAYMOND M, ROUSSET F 1995 GENEPOP (version 1.2): population Genetics software for exact tests and ecumenicism. *J Hered* 86 (3): 248-249. DOI: <https://doi.org/10.1093/oxfordjournals.jhered.a111573>
  33. GUO SW, THOMPSON EA 1992 Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics* 48 (2): 361-372. DOI: <https://doi.org/10.2307/2532296>
  34. RAYMOND M, ROUSSET F 1995 An exact test for population differentiation. *Evolution* 49 (6): 1283-1286. DOI: <https://doi.org/10.1111/j.1558-5646.1995.tb04456.x>
  35. HOLM S 1979 A simple sequentially rejective multiple test procedure. *Scand J Stat* 6 (2): 65-70.
  36. RICE WR 1989 Analyzing tables of statistical tests. *Evolution* 43 (1): 223-225. DOI: <https://doi.org/10.1111/j.1558-5646.1989.tb04220.x>
  37. NEI M 1972 Genetic distance between populations. *Am Nat* 106 (949): 283-292. DOI: <https://doi.org/10.1086/282771>
  38. FELSENSTEIN J 2004 PHYLIP (Phylogeny Inference Package), version 3.6b. Department of Genetics, University of Washington, Washington, D.C., USA
  39. BOWCOCK AM, RUIZ-LINARES A, TOMFOHRDE J, MINCH E, KIDD JR, CAVALLI-SFORZA LL 1994 High resolution human evolutionary trees with polymorphic microsatellites. *Nature* 368: 455-457. DOI: <https://doi.org/10.1038/368455a0>
  40. EXCOFFIER L, SMOUSE PE, QUATTRO JM 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction sites. *Genetics* 131 (2): 479-491
  41. SCHNEIDER S, ROESSLI D, EXCOFFIER L 2000 ARLEQUIN Version 2.000: A Software for Population Genetic Data Analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland
  42. LOISELLE BA, SORK VL, NASON J, GRAHAM C 1995 Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot* 82 (11): 1420-1425. DOI: <https://doi.org/10.1002/j.1537-2197.1995.tb12679.x>
  43. HARDY OJ, VEKEMANS X 2007 SPAGEdi 1.2: a program for Spatial Pattern Analysis of Genetic Diversity - User's manual. Laboratoire Eco-éthologie Evolutive, Université Libre de Bruxelles, Belgium
  44. FELSENSTEIN J 1985 Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39 (4): 783-791. DOI: <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x>
  45. BANDEL T, FORSTER P, ROHL A 1999 Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16 (1): 37-48. DOI: <https://doi.org/10.1093/oxfordjournals.molbev.a206036>
  46. Network 4.5.1.6. URL: [www.fluxus-engineering.com/sharenet.htm](http://www.fluxus-engineering.com/sharenet.htm) (12 March 2018)
  47. PETIT RJ, EL MOUSADIK A, PONS O 1998 Identifying populations for conservation on the basis of genetic markers. *Conserv Biol* 12: 844-855. DOI: <https://doi.org/10.1046/j.1523-1739.1998.96489.x>
  48. CONTRIB. URL: <https://www6.bordeaux-aquitaine.inra.fr/biogeco/Production-scientifique/Logiciels/Contrib-Permut/Contrib> (12 March 2018)
  49. NEI M 1987 Molecular Evolutionary Genetics. Columbia University Press, New York, USA
  50. PONS O, PETIT RJ 1995 Estimation, variance and optimal sampling of gene diversity. I. Haploid locus. *Theor Appl Genet* 90 (3-4): 462-470. DOI: <https://doi.org/10.1007/BF00221991>

51. PONS O, PETIT RJ 1996 Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* 144 (3): 1237-1245
52. PERMUT CpSSR. URL: <https://www6.bordeaux-aquitaine.inra.fr/biogeco/Production-scientifique/Logiciels/Contrib-Permut/Permut> (15 March 2018)
53. SLATKIN M 1995 A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139 (1): 457-462
54. GRIVET D, DEGUILLLOUX MF, PETIT RJ, SORK VL 2006 Contrasting patterns of historical colonization in white oaks (*Quercus* spp.) in California and Europe. *Mol Ecol* 15 (13): 4085-4093. DOI: <https://doi.org/10.1111/j.1365-294X.2006.03083.x>
55. DEGUILLLOUX MF, PEMONGE MH, PETIT RJ 2004 Use of chloroplast microsatellites to differentiate oak populations. *Ann For Sci* 61 (8): 825-830. DOI: <https://doi.org/10.1051/forest:2004078>
56. SCHÜLLER S 2012 Ergebnisse des österreichischen Eichenherkunftsversuches. URL: [https://bfw.ac.at/rz/document\\_api/download?content=Wels2012\\_eichen.pdf](https://bfw.ac.at/rz/document_api/download?content=Wels2012_eichen.pdf) (15 March 2018)
57. PROVAN J, POWELL W, DEWAR H, BRYAN G, MACHRAY GC, WAUGH R 1999 An extreme cytoplasmic bottleneck in the modern European cultivated potato (*Solanum tuberosum*) is not reflected in decreased levels of nuclear diversity. *Proc R Soc Lond B Biol Sci* 266 (1419): 633-639. DOI: <https://doi.org/10.1098/rspb.1999.0683>
58. NEOPHYTOU C, ARAVANOPOULOS FA, FINK S, DOUNAVI A 2010 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. *Forest Ecol Manag* 259 (10): 2026-2035. DOI: <https://doi.org/10.1016/j.foreco.2010.02.013>
59. MUIR G, SCHLÖTTERER C 2005 Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Mol Ecol* 14 (2): 549-561. DOI: <https://doi.org/10.1111/j.1365-294X.2004.02418.x>
60. COTTRELL JE, MUNRO RC, TABBENER HE, MILNER AD, FORREST GI, LOWE AJ 2003 Comparison of fine-scale genetic structure using nuclear microsatellites within two British oakwoods differing in population history. *Forest Ecol Manag* 176 (1-3): 287-303. DOI: [https://doi.org/10.1016/S0378-1127\(02\)00289-X](https://doi.org/10.1016/S0378-1127(02)00289-X)
61. BAKKER EG, VAN DAM BC, VAN EEUWIJK FA, JACOBSEN E 2003 Population Genetics of indigenous *Quercus robur* L. populations and of derived half-sib families has implications for the reproductive management of the species. *Plant Biol* 5 (4): 393-399. DOI: <https://doi.org/10.1055/s-2003-42706>
62. BUIITVELD J, BAKKER EG, BOVENSCHEN J, DE VRIES SMG 2001 Paternity analysis in a seed orchard of *Quercus robur* L. and estimation of the amount of background pollination using microsatellite markers. *For Genet* 8 (4): 331-337
63. DEGEN B, STREIFF R, ZIEGENHAGEN B 1999 Comparative study of genetic variation and differentiation of two pedunculate oak (*Quercus robur*) stands using microsatellite and allozyme loci. *Heredity* 83: 597-603. DOI: <https://doi.org/10.1038/sj.hdy.6886220>
64. LEPAIS O, LÉGER V, GERBER S 2006 Short note: high throughput microsatellite genotyping in oak species. *Silvae Genet* 55 (1-6): 238-240. DOI: <https://doi.org/10.1515/sg-2006-0031>
65. BALLIAN D, BELLETTI P, FERRAZZINI D, BOGUNIĆ F, KAJBA D 2010 Genetic variability of Pedunculate Oak (*Quercus robur* L.) in Bosnia and Herzegovina. *Period Biol* 112 (3): 353-362
66. GREGORIUS HR, DEGEN B, KÖNIG A 2007 Problems in the analysis of genetic differentiation among populations – a case study in *Quercus robur*. *Silvae Genet* 56 (1-6): 190-199. DOI: <https://doi.org/10.1515/sg-2007-0029>
67. VAN OOSTERHOUT CW, HUTCHINSON F, WILLS DPM, SHIPLEY P 2004 Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4 (3): 535-538. DOI: <https://doi.org/10.1111/j.1471-8286.2004.00684.x>
68. KATIČIĆ I, BOGDAN S, SEVER K, ŠATOVIĆ Z, KAJBA D 2010 Genetic structure and variability of phenological forms of pedunculate oak (*Quercus robur* L.) from clonal seed orchards in Croatia. In: Vinceti B, NEATE P (eds) Book of Abstracts of Conference "Forest ecosystem genomics and adaptation". San Lorenzo de El Escorial, Madrid, Spain, 9-11 June 2010. Bioersity International, Rome, Italy and INIA, Madrid, Spain, pp 181
69. FRANJIĆ J, BOGDAN S, ŠKVORC Ž, SEVER K, KRSTONOŠIĆ D 2009 Phenological synchronisation of pedunculate oak clones from clonal seed orchards in Croatia. In: MATIĆ S, ANIĆ I (eds) Proceedings of the scientific symposium "Forests of pedunculate oak in changed site and management conditions", Croatian Academy of Sciences and Arts, Scientific Council for Agriculture and Forestry, Zagreb, Croatia, pp 153-168
70. FRANJIĆ J, SEVER K, BOGDAN S, ŠKVORC Ž, KRSTONOŠIĆ D, ALEŠKOVIĆ I 2011 Phenological Asynchronization as a Restrictive Factor of Efficient Pollination in Clonal Seed Orchards of Pedunculate Oak (*Quercus robur* L.). *Croat J For Eng* 32 (1): 141-156
71. VEKEMANS X, HARDY OJ 2004 New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* 13 (4): 921-935. DOI: <https://doi.org/10.1046/j.1365-294X.2004.02076.x>
72. HAMPE A, EL MASRI L, PETIT RJ 2010 Origin of spatial genetic structure in an expanding oak population. *Mol Ecol* 19 (3): 459-471. DOI: <https://doi.org/10.1111/j.1365-294X.2009.04492.x>
73. BERG EE, HAMRICK JL 1995 Fine-scale genetic structure of a turkey oak forest. *Evolution* 49 (1): 110-120. DOI: <https://doi.org/10.1111/j.1558-5646.1995.tb05963.x>
74. BORDÁCS S, PÉSCU F, SLADE D, CSAIKL UM, LESUR I, BOROVIĆ A, KÉZDY P, KÖNIG AO *et al.* 2002 Chloroplast DNA variation of white oaks in northern Balkans and in the Carpathian Basin. *Forest Ecol Manag* 156 (1-3): 197-209. DOI: [https://doi.org/10.1016/S0378-1127\(01\)00643-0](https://doi.org/10.1016/S0378-1127(01)00643-0)
75. SLADE D, ŠKVORC Ž, BALLIAN D, GRAČAN J, PAPEŠ D 2007 The chloroplast DNA polymorphism of white oaks of Section *Quercus* in the Central Balkans. *Silvae Genet* 57 (4-5): 227-234. DOI: <https://doi.org/10.1515/sg-2008-0035>
76. DEGUILLLOUX MF, PEMONGE MH, PETIT RJ 2002 Novel perspectives in wood certification and forensics: dry wood as a source of chloroplast, mitochondrial and nuclear DNA. *P Roy Soc Lond B Bio* 269 (1495): 1039-1046. DOI: <https://doi.org/10.1098/rspb.2002.1982>
77. DEGUILLLOUX MF, PEMONGE MH, BERTEL L, KREMER A, PETIT RJ 2003 Checking the geographical origin of oak wood: molecular and statistical tools. *Mol Ecol* 12 (6): 1629-1636. DOI: <https://doi.org/10.1046/j.1365-294X.2003.01836.x>
78. PETIT RRR, PINEAU E, DEMESURE B, BACILIERI R, DUCOUSSO A, KREMER A 1997 Chloroplast DNA footprints of postglacial recolonization by oaks. *PNAS* 94 (18): 9996-10001. DOI: <https://doi.org/10.1073/pnas.94.18.9996>
79. PETIT RJ, DUMINIL J, FINESCHI S, HAMPE A, SALVINI D, VENDRAMIN GG 2005 Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol Ecol* 14 (3): 689-701. DOI: <https://doi.org/10.1111/j.1365-294X.2004.02410.x>
80. KREMER A 2007 How well can existing forests withstand climate change? In: Koskela J, Buck A, Teissier Du Cros E (eds) Climate change and forest genetic diversity: implications for sustainable forest management in Europe. Bioersity International, Rome, Italy, pp 3-15