

# Fine-scale spatial genetic structure in a multi-oak-species (*Quercus* spp.) forest

Alexandru Lucian Curtu<sup>(1)</sup>, Iacob Craciunesc<sup>(1)</sup>, Cristian Mihai Enescu<sup>(1-2)</sup>, Amaryllis Vidalis<sup>(3)</sup>, Nicolae Sofletea<sup>(1)</sup>

Patterns of fine-scale spatial distribution of multilocus genotypes can provide valuable insights into the biology of forest tree species. Here we tested for the existence of spatial genetic structure (SGS) in a four-oak-species forest with contrasting species abundances and hybridization rates. A total of 483 adult trees were mapped over 8.6 ha and genotyped using 10 highly polymorphic genomic regions. A weak but significant SGS was observed in each of the four oak species, with *Quercus frainetto*, the species with the lowest density in the sampling plot, exhibiting the strongest SGS. The values of the *Sp* statistic were 0.0033, 0.0035, 0.0042, and 0.0098 for *Q. petraea*, *Q. robur*, *Q. pubescens*, and *Q. frainetto*, respectively. The spatial correlogram of the total population was significantly different when hybrids were removed from the analysis, which suggests that hybridization influenced the SGS. Interspecific SGSs were significantly correlated with the rates of hybridization. Implications of the obtained results for the conservation and management of forest genetic resources are discussed.

**Keywords:** Spatial Genetic Structure, *Quercus*, Oak Species, Population Density, Hybridization

## Introduction

Understanding the patterns of spatial genetic structure (SGS), *i.e.*, the distribution of genetic variants in space, can provide us with useful information for the conservation and management of genetic resources. Spatial clustering of closely related individuals within populations arises from the interaction of genetic and demographic factors, which are often difficult to disentangle (Epperson 1992, Kalisz et al. 2001, Vekemans & Hardy 2004, Hampe et al. 2010). The intensity of SGS is strongly influenced by species dispersal ability, *i.e.*, SGS is expected to increase in species with lower pollen and seed dispersal distances (Cavers et al. 2005). Demographic parameters, such as population density and the spatial arrangement of individuals, are also important determinants of SGS. The strength of SGS is expected to increase in low density, as op-

posed to high density populations, although larger gene dispersal distances in low density populations may compensate for the effect of genetic drift (Vekemans & Hardy 2004, Born et al. 2008). In mixed stands, high tree density can limit dispersal by serving as a physical barrier to the movements of propagules (Hardy 2009). A natural reduction in population size may lead to a decrease in SGS through successive life stages, from juvenile to adult generations (Chung et al. 2003). The degree of spatial clustering of related individuals may also reflect the impact of silvicultural practices (Cottrell et al. 2003, Rajendra 2011, Paffetti et al. 2012), the extent of clonal reproduction (Hoebee et al. 2006), and the influence of life history traits (Kalisz et al. 2001, Luna et al. 2005). The complex interaction of factors that influence spatial patterns can have unexpected results, such as an absence of SGS in a pristine

beech forest and the detection of SGS in an even-aged beech plantation (Piotti et al. 2013). The level of introgressive hybridization may also shape SGSs within populations due to an excess of mating events between nearby individuals of different species and the spatial aggregation of the resulting hybrids (Valbuena-Carabana et al. 2007).

Oak species (genus *Quercus*, family Fagaceae) are common elements of temperate and Mediterranean forest ecosystems. They are long-lived tree species, monoecious, highly outcrossing, and have wind pollination and gravity dispersed seeds, although a small proportion of seeds can be dispersed by animals (*e.g.*, birds). Most oak species do not produce seeds (acorns) every year; the periodicity of abundant acorn production can reach up to 10 years in some parts of their natural range (Stanescu et al. 1997). They are both ecologically and economically important species, and are well known for their propensity to hybridize (Burger 1975, Rushton 1993). Relatively few studies have addressed the issue of SGS within oak populations and its genetic and ecological determinants (Bacilieri et al. 1994, Berg & Hamrick 1995, Streiff et al. 1998, Valbuena-Carabana et al. 2007, Hampe et al. 2010). The SGS intensity, as quantified by the *Sp* statistic, is lower in oak populations than in other plant species (Vekemans & Hardy 2004, Valbuena-Carabana et al. 2007, Hampe et al. 2010). Differences in SGS between the two main oak species of temperate European forests, *Q. robur* and *Q. petraea*, were observed in mixed stands. The former species exhibited weaker spatial patterns, which may be explained by its greater acorn dispersal (Streiff et al. 1998, Cottrell et al. 2003, Jensen et al. 2003). Thus, only a single study has investigated the effects of oak hybridization on SGS in mixed stands of two species, *Q. petraea* and *Q. pyrenaica*, in Spain (Valbuena-Carabana et al. 2007). An interspecific component of SGS was found in one of the stands, which showed a high level of hybridization. However, no investigation of the possible effects of density and interspecific gene flow has been performed in more complex communities that consist of other oak species or more than two closely related species with varying pairwise hybridization rates.

In the present study, we combined information from a high number of multilocus genotypes and spatial data to study the fine-scale SGS in a four-oak-species plot, covering 8.6 ha, which was exhaustively sampled. More specifically, we addressed the following questions:

(i) Do genotypes have a non-random spatial distribution in a mixed forest of four closely related species? We tested the hypothesis of absence of SGS by conducting a

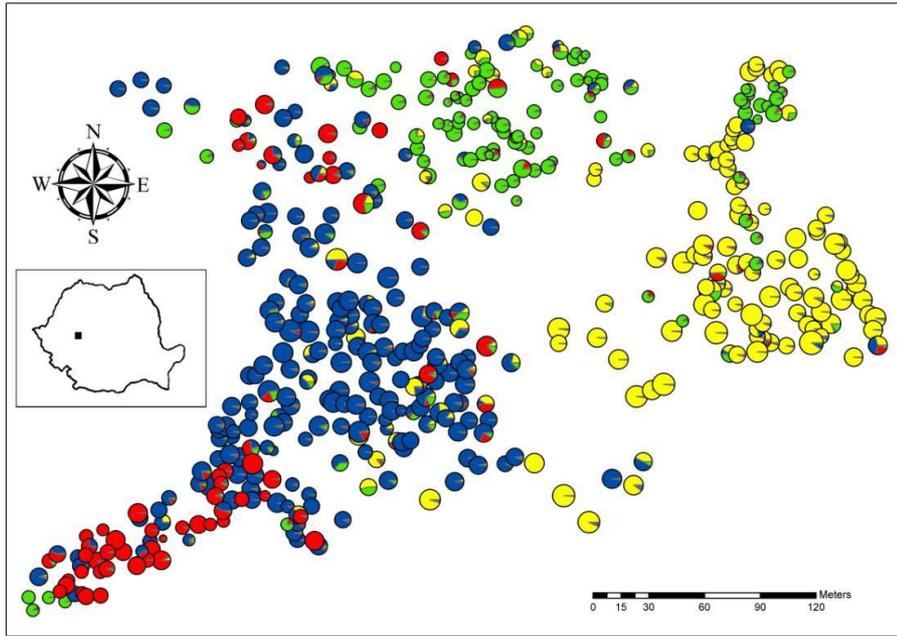
□ (1) Department of Forest Sciences, Transilvania University of Brasov, Sirul Beethoven-1, 500123 Brasov (Romania); (2) Department of Soil Sciences, University of Agronomic Sciences and Veterinary Medicine of Bucharest (Romania); (3) Umeå Plant Science Centre, Department of Ecology and Environmental Science, SE-901 87 Umeå (Sweden)

@ Alexandru Lucian Curtu (lucian.curtu@unitbv.ro)

Received: Oct 14, 2013 - Accepted: Aug 23, 2014

Citation: Curtu AL, Craciunesc I, Enescu CM, Vidalis A, Sofletea N, 2015. Fine-scale spatial genetic structure in a multi-oak-species (*Quercus* spp.) forest. iForest 8: 324-332 [online 2014-09-05] URL: <http://www.sisef.it/iforest/contents/?id=ifor1150-008>

Communicated by: Andrea Piotti



**Fig. 1** - Spatial distribution of oak trees at Bejan Oaks Reserve. Each individual is represented by a circle. Each color within the circle represents the membership in one of the four genetic clusters identified in the Bayesian analysis. The membership in *Q. robur*, *Q. pubescens*, *Q. petraea* and *Q. frainetto* cluster is represented by yellow, green, blue and red color, respectively. Circle size is proportional to DBH.

spatial autocorrelation analysis for each species, as well as for the total population. We expected the SGS to increase with reducing conspecific density, although the physical barriers to gene dispersal represented by the other tree species present in the stand may also affect the SGS.

(ii) Does hybridization influence SGS patterns? Purebreds and putative hybrids were identified in the oak forest based on genetic assignment. By quantifying the SGS for each species *sensu lato* (all trees) and *sensu stricto* (purebreds only), we tested whether hybrid individuals aggregated with their parents. If hybrids were randomly distributed, there would be no influence on the SGS once they are removed. To determine whether hybridization takes place between neighboring individuals of different species, we calculated correlograms based on restricted comparisons of pairs between two species (pairs within species were excluded).

**Material and methods**

*Study site and sampling*

This study was performed in the Bejan Oaks Reserve, a mixed deciduous forest, which is located in central Romania (45° 51' N, 22° 53' E), near the city of Deva (Fig. 1). This forest consists mainly of four closely related oak species (*Q. robur*, *Q. petraea*, *Q. pubescens* and *Q. frainetto*), as well as the more distant species *Q. cerris*. In this location, there are other scattered tree species such as *Tilia cordata*, *Acer campestre*, *A. platanoides*, *A. tataricum*, *Carpinus betulus*, *Fraxinus excelsior*, *Pyrus pyraeaster*, *Sorbus torminalis*, *S. domestica*, and *Malus sylvestris*. The sampling was exhaustive over an area of about 8.6 ha. A total of 483 oak trees were tagged and mapped with a Field-Map system (Fig. 1). *Q. cerris*, the second most common species after *Q. petraea*, was not included in the present study, although it

might have an effect on the pollen dispersal of the other oak species, thus influencing the SGS. Diameter at breast height (DBH) was also measured for all oak trees. Individuals were classified to the species level based on morphological traits (Stanescu et al. 1997). For a subset of 269 trees, a detailed leaf morphological examination was previously performed and only 16 (approximately 6%) individuals with intermediate morphologies were observed (Curtu et al. 2007a). Individuals with intermediate morphology were classified as the species they most closely resemble. Finally, we identified 190, 118, 113 and 62 individuals, which were *Q. petraea*, *Q. robur*, *Q. pubescens* and *Q. frainetto*-like, respectively (Tab. 1).

*Genotyping*

DNA was extracted from buds using the Qiagen DNeasy96 and Plant Mini Kit according to the manufacturer’s instructions, except for material disruption for which no liquid nitrogen was used. For a small subset of trees from the Bejan mixed forest, DNA was obtained using the cetyltrimethyl ammonium bromide (CTAB) method. Ten highly polymorphic microsatellite markers were used: QrZAG112, QrZAG96, QrZAG11, QrZAG20, QrZAG7, QrZAG87 (Kampfer et al. 1998), QpZAG110, QpZAG9 (Steinkellner et al. 1997) PIE215 and PIE223 (Durand et al. 2010). These microsatellite markers are among those selected to genotype European oak species in recent studies (Lepais et al. 2009, Guichoux et al. 2011, Neophytou et al. 2011). Several other markers were tested (e.g., QrZAG39 and QrZAG5b), but they were removed from the population screening due to the presence of null alleles or because the genotypes could not be reliably scored. Polymerase chain reactions (PCRs) were conducted in a *Corbett* thermal cycler. The forward primer was fluorescently labelled with Beckman dyes: D2, D3 or D4. The primer pairs were combined into three multiplex reactions. The first set comprised four markers, and dye type and primer concentrations (µM) are given in brackets: QrZAG112 (D4, 0.10), QrZAG96 (D3, 0.30), QrZAG11 (D3, 0.20) and QpZAG110 (D4, 0.40). The second set contained three markers, QrZAG20 (D3, 0.28), QrZAG7 (D4, 0.35) and QrZAG87 (D2, 0.40), and the last set comprised PIE215 (D4, 0.50), PIE223 (D3, 0.30) and QpZAG9 (D2, 0.35). The reactions were performed in a 10 µl volume containing approximately 10 ng of template DNA, 1x Promega colorless PCR buffer, 2 mM of MgCl<sub>2</sub>, 0.45 mM of each dNTP (Fermentas), and 0.2 U *Taq* DNA polymerase (Promega); for each primer concentration see the kit composition. The PCR profile was as follows: 3 min of denaturation at 94 °C followed by 31 cycles of 50 s denaturation at 94 °C, a 40 s annealing step at 52 °C, a 1

**Tab. 1** - Sample size per species (n), tree density and basic genetic statistics averaged across 10 microsatellite loci. (N<sub>a</sub>): mean number of alleles per locus; (H<sub>e</sub>): gene diversity; (F): fixation index.

Population (Long/Lat)	Species	n	Tree density ha <sup>-1</sup>	N <sub>a</sub>	H <sub>e</sub>	F
Bejan (45° 51' N / 22° 53' E)	<i>Q. robur</i>	118	13.72	19.5	0.787	0.047
	<i>Q. pubescens</i>	113	13.14	18.2	0.854	-0.005
	<i>Q. petraea</i>	190	22.09	19.8	0.786	0.037
	<i>Q. frainetto</i>	62	7.21	12.2	0.714	-0.029
Total		483	-	-	-	-

min 20 s elongation step at 70 °C and a final extension step at 70 °C for 12 min. Amplification products were run on a GenomeLab GeXP Genetic Analyser using the Frag-3 method and size standard 400. The products were then analyzed using Fragment Analysis Software using default parameters and PA ver1 dye correction.

First, we used the 10 microsatellite markers on a group of eight oak populations distributed across Romania (Tab. S1 in Appendix 1). Microsatellite markers were tested for genotyping errors due to large allele drop-outs, scoring of stutter peaks and non-amplified alleles using MICRO-CHECKER 2.2.0.3 (Van Oosterhout et al. 2004). No evidence of large allele drop-outs or scoring of stutter peaks was found in the data set. The software indicated the presence of null alleles at low frequencies (less than 8%) for three markers (QrZAG11, PIE223 and QrZAG96) in one or two external populations only. However, simulations showed that at such low frequencies, these null alleles do not affect the overall outcome of the genetic assignment (Carlsson 2008) or the robustness of SGS results (Piotti et al. 2013). Finally, we used the same set of genetic markers to genotype the oak individuals sampled at the Bejan oak mixed forest. Basic genetic statistics were calculated using GENALEX software v. 6.4 (Peakall & Smouse 2006).

#### Genetic assignment and hybrid identification

The Bayesian clustering method implemented in the STRUCTURE software v. 2.3.3 (Pritchard et al. 2000) was used to genetically assign individuals to clusters. Simulations were run for 100 000 steps following a burn-in period of 50 000 steps, considering values of  $K$  (number of clusters) from one to 10, with 10 replications for each value of  $K$ . The analysis was performed using admixture, correlated allele frequencies and no prior information regarding sampling location or morphological species. For each individual the program identifies the fraction of the genome that belongs to each one of the clusters.

First, we ran STRUCTURE for the set of eight

external pure populations (346 individuals), which consisted of two morphologically pure populations for each of the four species found in the mixed forest of Bejan. The rate of change in the log likelihood between successive  $K$  values (Evanno et al. 2005) was also estimated. The calculations were performed with STRUCTURE HARVESTER (Earl 2011). We then ran STRUCTURE for all samples (829 individuals) with the same parameters as before. For the most likely number of clusters ( $K = 4$ ), we averaged the values of the admixture coefficient ( $q$ ) over 10 runs. For the classification of individuals into purebred and hybrid classes, we followed the procedure described in a previous study on four French oak species with 10 microsatellite markers (Lepais et al. 2009). Each individual was considered to be a purebred when its  $q$  value (admixture coefficient) was larger than 0.90, and a hybrid when its  $q$  value was below this threshold. When an individual had a  $q$  value lower than 0.90 for one genetic cluster, and  $q$  value lower than 0.10 for each of the three remaining clusters, it was considered as purebred. Individuals with  $q$  values between 0.10 and 0.90 for three or four clusters were considered as hybrids between the two clusters (species) with the highest  $q$  value, no matter the  $q$  values for the third and fourth clusters.

#### Spatial genetic structure analysis

To study the spatial distribution of multilocus genotypes within both species and the total population, we computed the correlation coefficient  $r$  (Smouse & Peakall 1999) using the program GENALEX 6.5 (Peakall & Smouse 2006). The maximum pairwise distance between individuals sampled within the mixed forest varied between 315.8 m and 500.2 m for *Q. robur* and *Q. pubescens*, respectively. Because the sample size for each species varied considerably and the number of pairs within each distance class should be at least 100 (Hardy & Vekemans 2002), we set the distance intervals to 25 m. To compare the results between species, we also chose to conduct the analyses for 10 distance intervals between 0 and 250 m. Lower and upper 95% confidence intervals were com-

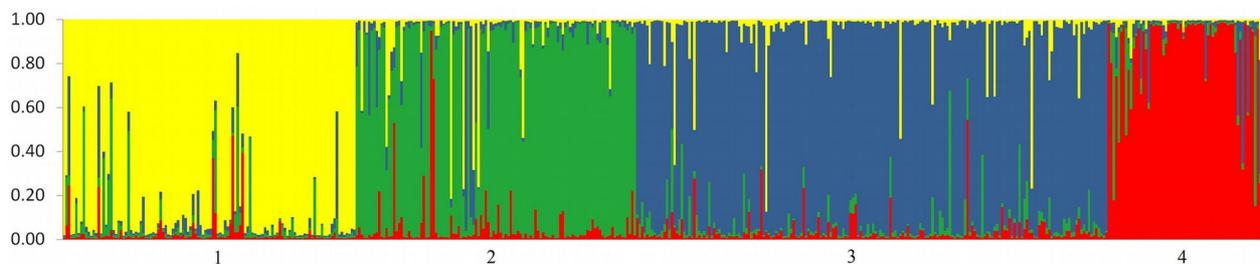
puted by random shuffling (1000 times) of individual geographic locations; 95% confidence intervals around the mean value of the correlation coefficient  $r$  were estimated by bootstrapping pairwise comparisons within each distance interval (1000 repeats). The nonparametric heterogeneity test of Smouse et al. (2008) was used to compare correlograms between species, with and without hybrids. The number of bootstrap resamplings was set to 999.

Additionally, we computed correlograms for interspecific pairs using Nason's kinship coefficient,  $F_{ij}$  (Loiselle et al. 1995), a widely used estimator of kinship for codominant markers (Hardy & Vekemans 2002). We also used the statistic  $Sp = -b_F/(1-F_1)$ , where  $F_1$  is the mean kinship coefficient between all pairs of individuals in the first distance class, and  $b_F$  ( $b$ -log) is the slope of the regression of kinship vs. the log of distance (Vekemans & Hardy 2004). The significance of the slope of the regression analysis was computed after 10 000 permutations.  $Sp$  confidence intervals were obtained as described in Valbuena-Carabana et al. 2007. All calculations were performed using the SPAGeDi 1.2 software (Hardy & Vekemans 2002). Correlations were calculated using STATISTICA v.8 software (StatSoft 2008).

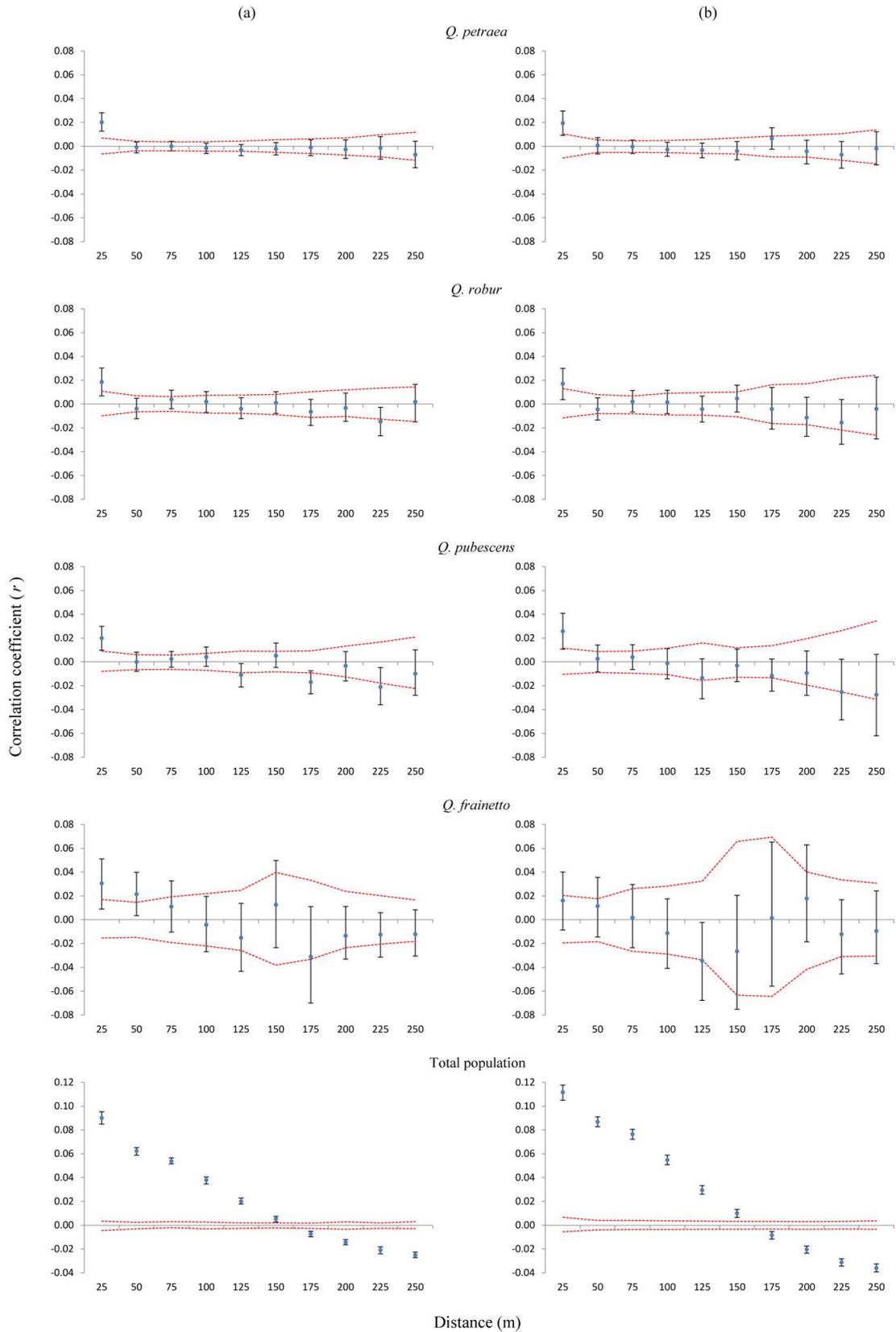
## Results

#### Classification of purebred and hybrid individuals in the mixed oak forest

By running the STRUCTURE software for the set of eight external oak populations, we found that the optimal number of genetic clusters, which corresponds to the highest posterior probability, was  $K = 4$  (additional data are given in Fig. S1 - Appendix 1). The highest value of  $\Delta K$  statistics was obtained for  $K = 2$ , although the statistics also gave support for  $K = 4$ . We then ran the software for the whole sample set, including the 483 trees sampled in the mixed forest. For the most likely number of clusters ( $K = 4$ ), a clear correspondence was observed between the genetic clusters identified in STRUCTURE and the morphological classification into four species (Fig. 2). The overall percentage



**Fig. 2** - Genetic structure revealed by 10 microsatellite markers. Each individual tree is represented by a thin vertical line which is divided into four color segments that are proportional to its membership in the four genetic clusters ( $K=4$ ) inferred in the Bayesian analysis. Populations are separated by a thin black line. Populations 1-4 consist of oaks sampled in Bejan mixed forest, which are grouped to the species they most closely resemble: *Q. robur* (1), *Q. pubescens* (2), *Q. petraea* (3) and *Q. frainetto* (4).



**Fig. 3** - Autocorrelograms for each oak species and the total population, considering (a) all trees (b) only trees genetically assigned as pure types (without hybrids). On x-axis are 25 m distance intervals. On y-axis is the correlation coefficient  $r$  of Smouse & Peakall (1999). Confidence intervals of 95% obtained through random shuffling (red line) are given. Black lines around the  $r$  value represent 95% confidence intervals generated by bootstrapping.

of putative hybrids in the mixed forest was high (23.6%) and varied among the four species. The highest proportions of hybrids, 30.1% and 25.6%, were observed among individuals phenotypically classified as *Q. pubescens* and *Q. frainetto*, respectively. The most genetically pure species was *Q. robur*, with only 16.1% of the sampled individuals showing a hybrid origin. The highest hybridization rate was observed between *Q. pubescens* and *Q. frainetto* (12.0%), whereas the lowest was between *Q. robur* and *Q. frainetto* (1.7%).

*Intraspecific spatial genetic structure*

We found values of the correlation coefficient *r* outside the 95% confidence intervals in each of the four species (Fig. 3a). Regardless of the species, oak individuals in the shortest distance interval appeared to be more genetically related than expected based on a random distribution. The correlation coefficient was significantly positive for the first two distance classes (up to 50 m) only for *Q. frainetto*, the least abundant species in the mixed forest. The SGS patterns were characterized by a peak *r* value at the first distance class, followed by a more or less steady decrease for the remaining distance intervals (Fig. 3a). Overall, the heterogeneity test of SGS between each pair of species was not significantly different, nor for each distance class (data not shown).

When only purebred individuals were considered, the correlograms revealed a similar pattern (Fig. 3b). However, the SGS was no longer significant for *Q. frainetto*, which is very likely due to the small sample size of this species once putative hybrids were removed. The statistical tests for heterogeneity between pairs of species were also not significant (data not shown).

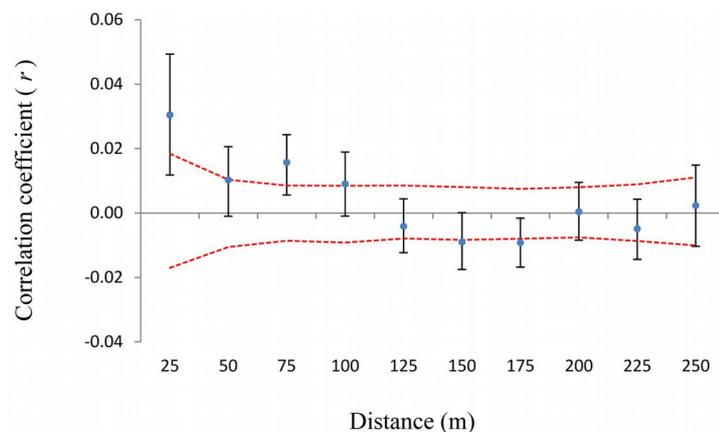
The value of the *Sp* statistic was highest for *Q. frainetto*, and this value was significantly different from those of *Q. robur* and *Q. petraea*, which had the lowest values (Tab. 2). The value of *b<sub>F</sub>* was significantly different from zero in all four species (Tab. 2). After hybrids removal, *Sp* statistic values were no longer significantly different between species. Interestingly, *Q. petraea* and *Q. robur* had nearly the same values for the *Sp* statistic, with and without the inclusion of putative hybrids. After the removal of hybrids, *Sp* statistic value only increased for *Q. pubescens*. This species showed the highest proportion of hybrids (30.01%) among the four species in Bejan. Species density was inversely proportional to the *Sp* parameter, although the correlations were not significant, either with (*r* = -0.79, *p* = 0.20) or without hybrids (*r* = -0.87, *p* = 0.13).

*Spatial genetic structure in the total population*

The intensity of the SGS was much higher

**Tab. 2** - Parameters describing spatial genetic structure for each species, the total population (overall) and hybrids. (*n*): number of individuals considered in the analysis; (*F*<sub>1</sub>): average kinship coefficient between individuals of the first distance class (0-25 m); [*b<sub>F</sub>* (*b-log*)]: slope of the regression of kinship coefficient *F*<sub>ij</sub> computed among all pairs of individuals against geographical distance; (*Sp*): intensity of spatial genetic structure; (SE): standard error; (95% CI): 95% confidence intervals; (\*): *P*<0.05; (\*\*): *P*<0.01; (\*\*\*): *P*<0.001.

SGS Analysis	Species	<i>F</i> <sub>1</sub>	<i>b<sub>F</sub></i> ( <i>b-log</i> ) (±SE)	<i>Sp</i> (95% CI)
<i>Sensu lato</i> (with the inclusion of putative hybrids)	<i>Q. petraea</i>	0.0114***	-0.0032 ± 0.0006***	0.0033 (0.0019-0.0045)
	<i>Q. robur</i>	0.0103**	-0.0035 ± 0.0014***	0.0035 (0.0007-0.0063)
	<i>Q. pubescens</i>	0.0099***	-0.0041 ± 0.0014***	0.0042 (0.0014-0.0069)
	<i>Q. frainetto</i>	0.0146**	-0.0096 ± 0.0017***	0.0098 (0.0064-0.0131)
	Overall ( <i>n</i> = 483)	0.0511***	-0.0286 ± 0.0083***	0.0301 (0.0129-0.0471)
<i>Sensu stricto</i> (pure-bred only)	<i>Q. petraea</i>	0.0110***	-0.0032 ± 0.0009***	0.0033 (0.0015-0.0052)
	<i>Q. robur</i>	0.0097**	-0.0034 ± 0.0014**	0.0034 (0.0006-0.0063)
	<i>Q. pubescens</i>	0.0123***	-0.0051 ± 0.0015***	0.0051 (0.0020-0.0083)
	<i>Q. frainetto</i>	0.0074	-0.0053 ± 0.0017*	0.0053 (0.0020-0.0087)
	Overall ( <i>n</i> = 369)	0.0637***	-0.0375 ± 0.0102***	0.0400 (0.0182-0.0613)
All hybrids	Hybrids ( <i>n</i> = 114)	0.0172***	-0.0069 ± 0.0029***	0.0069 (0.0011-0.0127)



**Fig. 4** - Spatial autocorrelogram for hybrids. On x-axis are 25 m distance intervals. On y-axis is the correlation coefficient *r* of Smouse & Peakall (1999). Confidence intervals of 95% obtained through random shuffling (red line) are given. Black error bars represent 95% confidence intervals generated by bootstrapping.

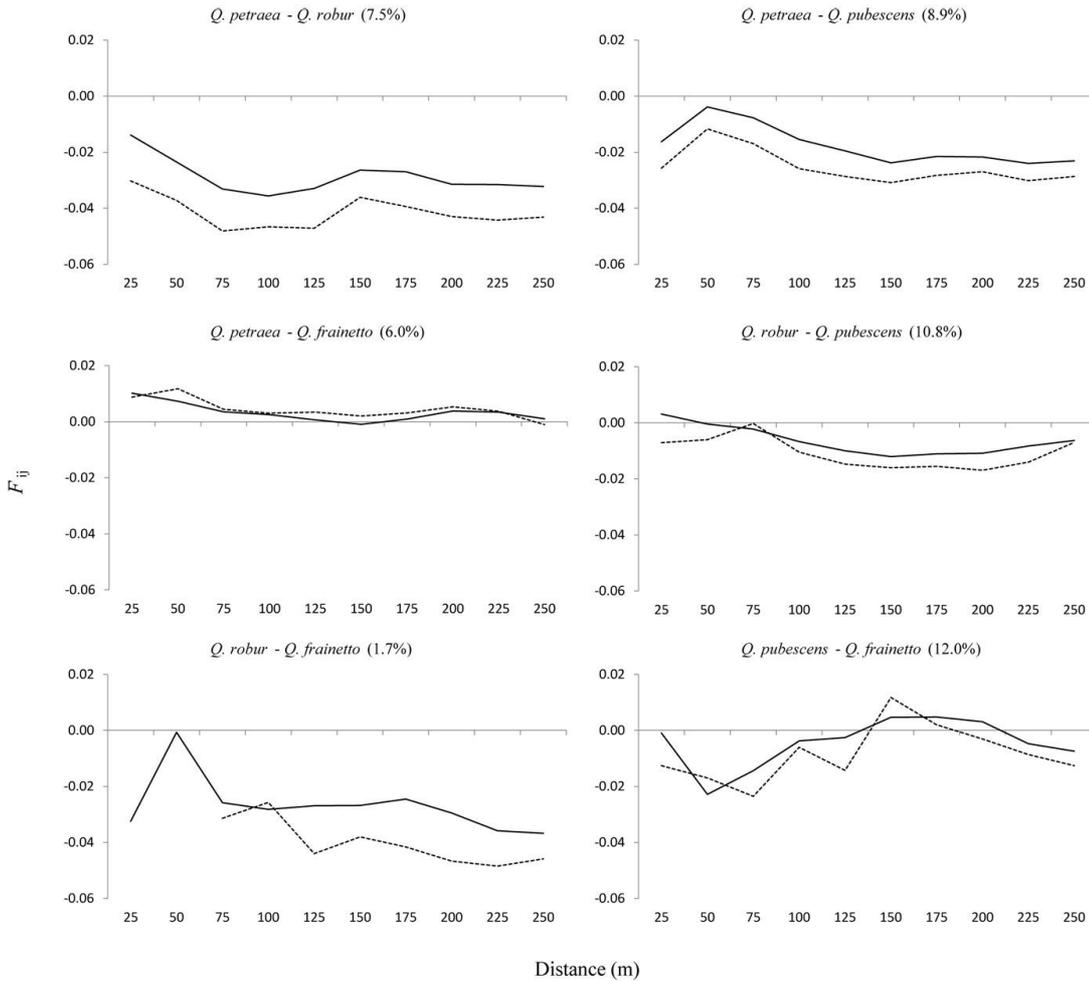
in the total population compared with each species (Fig. 3a), and it became more intense after the exclusion of hybrids from the analysis (Fig. 3b). The heterogeneity test of SGS between the total population, with and without hybrids, was significant (*p*<0.01).

The same pattern was observed for the *Sp* parameter (*Sp* = 0.0301 for 483 trees, including both pure types and hybrids, compared with *Sp* = 0.0400 without hybrids - Tab. 2). A spatial structure was also observed when

all 114 hybrids alone were analyzed (Fig. 4).

*Interspecific spatial genetic structure*

When the analysis of SGS was restricted to mixed pairs between species, values of the kinship coefficients (*F*<sub>ij</sub>) were negative in most comparisons (Fig. 5). However, positive values of the kinship coefficient (*F*<sub>ij</sub>) for shorter distance classes were observed for two pairs, *Q. petraea* - *Q. frainetto* and *Q. robur* - *Q. pubescens* (Fig. 5), while the va-



**Fig. 5** - Interspecific autocorrelograms, considering all trees (continuous line) and only trees genetically assigned as pure types (dotted line). On x-axis are 25 m distance intervals. On y-axis is the Nason kinship coefficient ( $F_{ij}$ ) of Loiselle et al. (1995). Estimation of the hybridization rate is given in brackets.

lue of  $b_F$ , which is independent of distance intervals, was not significantly different from zero ( $-0.0022^{NS}$  vs.  $-0.0029^{NS}$ , respectively). A very low value of  $b_F$  was obtained for the *Q. pubescens*-*Q. frainetto* pair ( $-0.0003^{NS}$ ), where distance pairs ranging from 150 to 200 m appeared to be more genetically related than those of the shorter distances (Fig. 5). The hybridization rate between pair of species (Fig. 5) was inversely correlated with the  $Sp$  value ( $r = -0.85$ ,  $p < 0.05$ ). Lower values of the kinship coefficient were obtained for the smallest distance intervals in three out of the five cases once putative hybrids were removed from the dataset (Fig. 5). The comparison between *Q. frainetto* and *Q. robur* was not taken into account due to the low number of pairs. Consequently, values of  $b_F$  were lower without hybrids (data not shown).

**Discussion**

*Intraspecific spatial genetic structure*

A weak though significant SGS up to 25-50 m was detected in each of the four oak species present in the mixed forest. The existence of a fine-scale SGS at the level of each oak species can primarily be explained by

the limited dispersal of propagules (seeds and pollen) from adult trees (Streiff et al. 1998, Sork et al. 2002, Vekemans & Hardy 2004). Mating usually takes place between nearby trees, and a reduced number of parental genotypes is overrepresented in the offspring generation (Hampe et al. 2010). The weak kinship structure fits well with our expectations because this mixed forest is at an old ontogenetic stage, in which family structures are weak due to mortality through life cycle stages, as was suggested for an unmanaged stand of *Q. petraea* (Jensen et al. 2003) and *Fagus sylvatica* (Dounavi et al. 2010). Similar values for the  $Sp$  statistics were obtained in adult oak stands composed of one or two species: 0.005 in a *Q. robur* stand (Hampe et al. 2010), 0.0029 and 0.0082 for *Q. robur* and *Q. petraea*, respectively (Streiff et al. 1998, calculated in Vekemans & Hardy 2004), and 0.0033 and 0.0132 for *Q. petraea*, depending on the stand (Valbuena-Carabana et al. 2007).

At Bejan, the highest value of the  $Sp$  statistics was obtained for the least abundant species, *Q. frainetto*. This value was significantly different from those obtained for the most common species, *Q. petraea* and *Q. robur*, being the former characterized by a

restricted seed dispersal as compared with the two latter species. Moreover, results of a previous paternity analysis carried out in the same multi-oak stand suggest a restricted pollen flow in *Q. frainetto*: only 18-24% of the pollen grains originated from outside of the study plot (Curtu et al. 2009). In contrast, four *Q. petraea* mother trees received on average at least 59% of their pollen from outside the sampling plot. The intensity of SGS was similar in *Q. robur* and *Q. petraea*, although *Q. robur* acorns are heavier than those of *Q. petraea*, and it is widely accepted that the former can be transported larger distances, as they are preferred by birds (Petit et al. 2004). Indeed, the fine-scale SGS was stronger for *Q. petraea* than for *Q. robur* in a mixed forest with similar densities of both species (Streiff et al. 1998). Hence, other factors should explain the nearly identical  $Sp$  values obtained for *Q. robur* and *Q. petraea* in Bejan, although the kinship coefficient at the first distance interval ( $F_1$ ) was slightly higher for *Q. petraea* than for *Q. robur* (Tab. 2). The strength of the SGS, as indicated by the  $Sp$  parameter, is significantly related to population density, that is, the  $Sp$  value is higher in low density, as compared with high density populations (Vekemans & Hardy

2004). However, no correlation between the SGS intensity and the conspecific density was found at Bejan. Although the density of *Q. robur* trees in Bejan was lower than that of *Q. petraea*, the higher seed and pollen dispersal ability of *Q. robur* (Petit et al. 2004, Lagache et al. 2013) may have contributed to the decrease relative to *Q. petraea*, in SGS, thereby leading to similar SGS intensities in the sampling plot.

#### Within-population spatial genetic structure

A very strong and statistically significant SGS was detected over the 8.6 ha of the total population, which consisted of all 483 sampled oak individuals. The values of the *Sp* statistic were 3-9 times higher in the total population than in each of the four species. The spatial distribution of species in the study plot had a considerable influence on the SGS (Fig. 1). Each of the four oak species is predominant in one part of forest, where habitat conditions better correspond to its ecological requirements. Thus, *Q. robur* trees are mostly found in deep soil along a small valley in the eastern part of the study site (Fig. 1). *Q. pubescens* grows in shallow soil, on a gradual south-facing slope in the upper part of the sampling plot. *Q. petraea* is found in better aerated and drained soil, whereas *Q. frainetto* grows in clay-rich soils (Stanciu 1995).

#### Influence of hybridization on spatial genetic structure

Hybridization appears to be another factor influencing fine-scale SGS in plant populations (Valbuena-Carabana et al. 2007). At an intraspecific level, almost no changes were detected for *Q. robur* and *Q. petraea* when the putative hybrids were removed, likely because these two species exhibited the lowest number of hybrids. Nearly identical values of the *Sp* parameter, with and without the inclusion of hybrids, suggest that hybrids contributed little to the formation of the SGS in these two species, and that their hybrids were not spatially aggregated with their pure species in the sampling plot. In contrast, the SGS for *Q. pubescens* was stronger in the absence of hybrids, which suggests that hybridization has weakened the SGS. It is worth mentioning that the proportion of hybrids was the highest amongst trees phenotypically classified to *Q. pubescens*. At an interspecific level, our results suggest that the hybridization rate between two species is related to the intensity of the SGS, as quantified by the *Sp* parameter. That is, the higher the gene flow between two species, the weaker the interspecific SGS. However, the rate of hybridization is dependent on species characteristics and the environmental context (Seehausen et al. 2008, Lagache et al. 2013). Positive values of the kinship coefficient in

the first distance class were observed in those pairs of species showing a high degree of intermixing at Bejan: *Q. petraea* - *Q. frainetto* and *Q. robur* - *Q. pubescens*. A study performed in Spain revealed the presence of both SGS components, intra- and inter-specific, in one stand of *Q. petraea* and *Q. pyrenaica* that had a high hybridization rate, although the interspecific component was missing in a second stand where hybridization was rare (Valbuena-Carabana et al. 2007). At Bejan, the exclusion of hybrid individuals led to a decrease in the pairwise kinship coefficient in the first distance class (0-25 m) in three out of five species comparisons, therefore suggesting that hybridization could predominantly take place between nearby trees belonging to different species. This result is consistent with that reported for a Spanish mixed oak stand (Valbuena-Carabana et al. 2007).

#### Implications for conservation and management

The existence of SGS in each of the four oak species supports the hypothesis of the natural origin of this multi-oak-species stand, and confirms the small amount of information available on stand history from forest records (Stanciu 1995). The presence of any SGS was hypothesized to be a hint that the study plot is part of an ancient forest (Jensen et al. 2003). Moreover, the strongest SGS occurred in the least disturbed, old-growth Eastern white pine forest (Marquardt et al. 2007). However, Piotti et al. (2013) reported no significant SGS in a 250 to 300-years-old beech stand located in a pristine Austrian forest. In the same study, a weak, but significant, SGS was also detected in a single beech plantation, suggesting that the presence of SGS alone does not necessarily imply that the stand is of natural origin. The autochthony of the Bejan forest is also supported by the identification of region-specific chloroplast DNA haplotypes that belong to lineage B (Curtu et al. 2007b). These chloroplast haplotypes were observed in the same geographic region during a large-scale survey (Petit et al. 2002).

Certainly, there has been great anthropogenic pressure on the Bejan forest until its establishment as an Oaks Reserve in 1936. Since then, only sanitation cuttings of low-intensity have been performed in different parts of the reserve, including the study plot (Stanciu 1995). This low-intensity forest management might have contributed to the preservation of the SGS found by spatial autocorrelation analysis in Bejan. More intensive silvicultural interventions, e.g., the shelterwood system, had also a weak to negligible effect on the SGS when comparing managed and unmanaged beech stands having the same biogeographic history (Piotti et al. 2013). However, second growth stands did

not exhibit a SGS, although a SGS was found in primary growth stands of Eastern white pine in North America (Marquardt & Epperson 2004). Hence, the decrease of pairwise relatedness with spatial distance within forest stands provide useful information on the stand history, which is essential for conservation strategies and sustainable forest management practices.

#### Acknowledgements

This work was financially supported by UEFISCDI-CNCS, project number PNII-RU-TE-73/2010. Iacob Craciunesc was supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the project number POSDRU/159/1.5/S/134378. We thank Andras Tothpal for his assistance during the field work. We are grateful to three anonymous reviewers for their constructive comments on the manuscript.

#### References

- Bacilieri R, Labbe T, Kremer A (1994). Intraspecific genetic structure in a mixed population of *Quercus petraea* (Matt.) Liebl. and *Q. robur* L. *Heredity* 73 (2): 130-141. - doi: [10.1038/hdy.1994.112](https://doi.org/10.1038/hdy.1994.112)
- Berg EE, Hamrick JL (1995). Fine-scale genetic-structure of a Turkey oak forest. *Evolution* 49 (1): 110-120. - doi: [10.2307/2410297](https://doi.org/10.2307/2410297)
- Born C, Hardy OJ, Chevallier M-H, Ossari S, Atteke C, Wickings EJ, Hossaert-Mckey M (2008). Small-scale spatial genetic structure in the Central African rainforest tree species *Aucoumea klaineana*: a stepwise approach to infer the impact of limited gene dispersal, population history and habitat fragmentation. *Molecular Ecology* 17 (8): 2041-2050. - doi: [10.1111/j.1365-294X.2007.03685.x](https://doi.org/10.1111/j.1365-294X.2007.03685.x)
- Burger WC (1975). The species concept in *Quercus*. *Taxon* 24: 45-50. - doi: [10.2307/1218998](https://doi.org/10.2307/1218998)
- Carlsson J (2008). Effects of microsatellite null alleles on assignment testing. *Journal of Heredity* 99 (6): 616-623. - doi: [10.1093/jhered/esn048](https://doi.org/10.1093/jhered/esn048)
- Cavers S, Degen B, Caron H, Lemes MR, Margis R, Salgueiro F, Lowe AJ (2005). Optimal sampling strategy for estimation of spatial genetic structure in tree populations. *Heredity* 95(4): 281-289. - doi: [10.1038/sj.hdy.6800709](https://doi.org/10.1038/sj.hdy.6800709)
- Chung MY, Epperson BK, Gi Chung M (2003). Genetic structure of age classes in *Camellia japonica* (Theaceae). *Evolution* 57(1): 62-73. - doi: [10.1111/j.0014-3820.2003.tb00216.x](https://doi.org/10.1111/j.0014-3820.2003.tb00216.x)
- 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 Ecology and Management* 176 (1-3): 287-303. - doi: [10.1016/S0378-1127\(02\)00289-X](https://doi.org/10.1016/S0378-1127(02)00289-X)
- Curtu AL, Gailing O, Finkeldey R (2007a). Evidence for hybridization and introgression within

- a species-rich oak (*Quercus* spp.) community. BMC Evolutionary Biology 7 (1): 218. - doi: [10.1186/1471-2148-7-218](https://doi.org/10.1186/1471-2148-7-218)
- Curtu AL, Gailing O, Leinemann L, Finkeldey R (2007b). Genetic variation and differentiation within a natural community of five oak species (*Quercus* spp.). Plant Biology 9(01): 116-126. - doi: [10.1055/s-2006-924542](https://doi.org/10.1055/s-2006-924542)
- Curtu AL, Gailing O, Finkeldey R (2009). Patterns of contemporary hybridization inferred from paternity analysis in a four-oak-species forest. BMC Evolutionary Biology 9 (1): 284. - doi: [10.1186/1471-2148-9-284](https://doi.org/10.1186/1471-2148-9-284)
- Dounavi A, Koutsias N, Ziehe M, Hattermer H (2010). Spatial patterns and genetic structures within beech populations (*Fagus sylvatica* L.) of forked and non-forked individuals. European Journal of Forest Research 129 (6): 1191-1202. - doi: [10.1007/s10342-010-0409-9](https://doi.org/10.1007/s10342-010-0409-9)
- Durand J, Bodenes C, Chancerel E, Frigerio J-M, Vendramin G, Sebastiani F, Buonamici A, Gailing O, Koelewijn H-P, Villani F, Mattioni C, Cherubini M, Goicoechea P, Herran A, Ikarán Z, Cabane C, Ueno S, Alberto F, Dumoulin P-Y, Guichoux E, de Daruvar A, Kremer A, Plomion C (2010). A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. BMC Genomics 11 (1): 570. - doi: [10.1186/1471-2164-11-570](https://doi.org/10.1186/1471-2164-11-570)
- Earl DA (2011). Structure harvester v0.6. Web site. [online] URL: [http://taylor0.biology.ucla.edu/struct\\_harvest/](http://taylor0.biology.ucla.edu/struct_harvest/)
- Epperson BK (1992). Spatial structure of genetic variation within populations of forest trees. New Forests 6: 257-278. - doi: [10.1007/BF00120648](https://doi.org/10.1007/BF00120648)
- Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology 14 (8): 2611-2620. - doi: [10.1111/j.1365-294X.2005.02553.x](https://doi.org/10.1111/j.1365-294X.2005.02553.x)
- Guichoux E, Lagache L, Wagner S, Leger P, Petit RJ (2011). Two highly validated multiplexes (12-plex and 8-plex) for species delimitation and parentage analysis in oaks (*Quercus* spp.). Molecular Ecology Resources 11(3): 578-585. - doi: [10.1111/j.1755-0998.2011.02983.x](https://doi.org/10.1111/j.1755-0998.2011.02983.x)
- Hampe A, El Masri L, Petit RJ (2010). Origin of spatial genetic structure in an expanding oak population. Molecular Ecology 19(3): 459-471. - doi: [10.1111/j.1365-294X.2009.04492.x](https://doi.org/10.1111/j.1365-294X.2009.04492.x)
- Hardy OJ, Vekemans X (2002). SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2 (4): 618-620. - doi: [10.1046/j.1471-8286.2002.00305.x](https://doi.org/10.1046/j.1471-8286.2002.00305.x)
- Hardy OJ (2009). How fat is the tail? Heredity 103 (6): 437-438. - doi: [10.1038/hdy.2009.120](https://doi.org/10.1038/hdy.2009.120)
- Hoebee SE, Menn C, Rotach P, Finkeldey R, Holderegger R (2006). Spatial genetic structure of *Sorbus torminalis*: the extent of clonal reproduction in natural stands of a rare tree species with a scattered distribution. Forest Ecology and Management 226 (1-3): 1-8. - doi: [10.1016/j.foreco.2005.12.024](https://doi.org/10.1016/j.foreco.2005.12.024)
- Jensen JS, Olrik DC, Siegismund HR, Lowe AJ (2003). Population genetics and spatial autocorrelation in an unmanaged stand of *Quercus petraea* in Denmark. Scandinavian Journal of Forest Research 18 (4): 295-304. - doi: [10.1080/02827580310005072](https://doi.org/10.1080/02827580310005072)
- Kalisz S, Nason JD, Hanzawa FM, Tonsor SJ (2001). Spatial population genetic structure in *Trillium grandiflorum*: the roles of dispersal, mating, history, and selection. Evolution 55 (8): 1560-1568. - doi: [10.1111/j.0014-3820.2001.tb00675.x](https://doi.org/10.1111/j.0014-3820.2001.tb00675.x)
- Kampfer S, Lexer C, Glössl J, Steinkellner H (1998). Characterization of (GA)<sub>n</sub> microsatellite loci from *Quercus robur*. Hereditas 129: 183-186. - doi: [10.1111/j.1601-5223.1998.00183.x](https://doi.org/10.1111/j.1601-5223.1998.00183.x)
- Lagache L, Klein EK, Guichoux E, Petit RJ (2013). Fine-scale environmental control of hybridization in oaks. Molecular Ecology 22 (2): 423-436. - doi: [10.1111/mec.12121](https://doi.org/10.1111/mec.12121)
- Lepais O, Petit RJ, Guichoux E, Lavabre JE, Alberto F, Kremer A, Gerber S (2009). Species relative abundance and direction of introgression in oaks. Molecular Ecology 18 (10): 2228-2242. - doi: [10.1111/j.1365-294X.2009.04137.x](https://doi.org/10.1111/j.1365-294X.2009.04137.x)
- Loiseau BA, Sork VL, Nason J, Graham C (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). American Journal of Botany 82 (11): 1420-1425. - doi: [10.2307/2445869](https://doi.org/10.2307/2445869)
- Luna R, Epperson BK, Oyama K (2005). Spatial genetic structure of two sympatric neotropical palms with contrasting life histories. Heredity 95 (4): 298-305. - doi: [10.1038/sj.hdy.6800655](https://doi.org/10.1038/sj.hdy.6800655)
- Marquardt PE, Epperson BK (2004). Spatial and population genetic structure of microsatellites in white pine. Molecular Ecology 13 (11): 3305-3315. - doi: [10.1111/j.1365-294X.2004.02341.x](https://doi.org/10.1111/j.1365-294X.2004.02341.x)
- Marquardt PE, Echt CS, Epperson BK, Pubanz DM (2007). Genetic structure, diversity, and inbreeding of eastern white pine under different management conditions. Canadian Journal of Forest Research 37 (12): 2652-2662. - doi: [10.1139/X07-114](https://doi.org/10.1139/X07-114)
- Neophytou C, Aravanopoulos F, Fink S, Dounavi A (2011). Interfertile oaks in an island environment. II. Limited hybridization between *Quercus alnifolia* Poech and *Q. coccifera* L. in a mixed stand. European Journal of Forest Research 130 (4): 623-635. - doi: [10.1007/s10342-010-0454-4](https://doi.org/10.1007/s10342-010-0454-4)
- Paffetti D, Travaglini D, Buonamici A, Nocentini S, Vendramin GG, Giannini R, Vettori C (2012). The influence of forest management on beech (*Fagus sylvatica* L.) stand structure and genetic diversity. Forest Ecology and Management 284(0): 34-44. - doi: [10.1016/j.foreco.2012.07.026](https://doi.org/10.1016/j.foreco.2012.07.026)
- Peakall R, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6 (1): 288-295. - doi: [10.1111/j.1471-8286.2005.01155.x](https://doi.org/10.1111/j.1471-8286.2005.01155.x)
- Petit RJ, Brewer S, Bordács S, Burg K, Cheddadi R, Coart E, Cottrell J, Csaikl U, van Dam B, Deans D, Espinel S, Fineschi S, Finkeldey R, Glaz I, Goicoechea PG, Jensen JS, König AO, Lowe AJ, Madsen SF, Mátyás G, Munro RC, Popescu F, Slade D, Tabbener H, de Vries SGM, Ziegenhagen B, de Beaulieu J-L, Kremer A (2002). Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. Forest Ecology and Management 156: 49-74. - doi: [10.1016/S0378-1127\(01\)00634-X](https://doi.org/10.1016/S0378-1127(01)00634-X)
- Petit RJ, Bodenes C, Ducousso A, Roussel G, Kremer A (2004). Hybridization as a mechanism of invasion in oaks. New Phytologist 161 (1): 151-164. - doi: [10.1046/j.1469-8137.2003.00944.x](https://doi.org/10.1046/j.1469-8137.2003.00944.x)
- Piotti A, Leonardi S, Heuert M, Buiteveld J, Geburek T, Gerber S, Kramer K, Vettori C, Vendramin GG (2013). Within-population genetic structure in beech (*Fagus sylvatica* L.) stands characterized by different disturbance histories: does forest management simplify population substructure? PLoS ONE 8 (9): e73391. - doi: [10.1371/journal.pone.0073391](https://doi.org/10.1371/journal.pone.0073391)
- Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genetics 155 (2): 945-959. [online] URL: <http://www.genetics.org/content/155/2/945.full.pdf&embedded=true>
- Rajendra K (2011). Spatial dynamics of intraspecific genetic variation in European beech (*Fagus sylvatica* L.). Georg-August University of Göttingen, Göttingen, Germany, pp. 125.
- Rushton BS (1993). Natural hybridization within the genus *Quercus*. Annales des Sciences Forestières 50 (Suppl. 1): 73-90. - doi: [10.1051/forest:19930707](https://doi.org/10.1051/forest:19930707)
- Seehausen O, Takimoto G, Roy D, Jokela J (2008). Speciation reversal and biodiversity dynamics with hybridization in changing environments. Molecular Ecology 17 (1): 30-44. - doi: [10.1111/j.1365-294X.2007.03529.x](https://doi.org/10.1111/j.1365-294X.2007.03529.x)
- Smouse PE, Peakall R (1999). Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity 82 (5): 561-573. - doi: [10.1038/sj.hdy.6885180](https://doi.org/10.1038/sj.hdy.6885180)
- Smouse PE, Peakall ROD, Gonzales EVA (2008). A heterogeneity test for fine-scale genetic structure. Molecular Ecology 17 (14): 3389-3400. - doi: [10.1111/j.1365-294X.2008.03839.x](https://doi.org/10.1111/j.1365-294X.2008.03839.x)
- Sork VL, Davis FW, Smouse PE, Apsit VJ, Dyer RJ, Fernandez-M JF, Kuhn B (2002). Pollen movement in declining populations of California Valley oak, *Quercus lobata*: where have all the fathers gone? Molecular Ecology 11(9): 1657-1668. - doi: [10.1046/j.1365-294X.2002.01574.x](https://doi.org/10.1046/j.1365-294X.2002.01574.x)
- Stanciu A (1995). Cercetari taxonomice, morfologice si ecologice privind hibridii genului *Quercus* din Rezervatia Stiintifica Bejan-Deva, județul Hunedoara [Research on the taxonomy, morphology and ecology of hybrids from genus *Quercus* in the Scientific Reserve Bejan-Deva, Hunedoara]. Facultatea de Silvicultura, Universitatea Transilvania Brasov, Brasov, Romania, pp. 140. [in Romanian]
- Stanescu V, Sofletea N, Popescu O (1997). Flora forestiera lemnoasa a Romaniei [Woody forest flora of Romania]. Editura Ceres, Bucuresti, Romania, pp. 451. [in Romanian]
- StatSoft (2008). STATISTICA for Windows. Software-System For Data Analysis, Tulsa, OK,

USA.

Steinkellner H, Fluch S, Turetschek 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 Molecular Biology* 33: 1093-1096. - doi: [10.1023/A:1005736722794](https://doi.org/10.1023/A:1005736722794)

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. *Molecular Ecology* 7: 317-328. - doi: [10.1046/j.1365-294X.1998.00360.x](https://doi.org/10.1046/j.1365-294X.1998.00360.x)

Valbuena-Carabana M, Gonzalez-Martinez SC, Hardy OJ, Gil L (2007). Fine-scale spatial genetic structure in mixed oak stands with different levels of hybridization. *Molecular Ecology* 16 (6): 1207-1219. - doi: [10.1111/j.1365-294X.2007.03231.x](https://doi.org/10.1111/j.1365-294X.2007.03231.x)

[07.03231.x](#)

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004). Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4 (3): 535-538. - doi: [10.1111/j.1471-8286.2004.00684.x](https://doi.org/10.1111/j.1471-8286.2004.00684.x)

Vekemans X, Hardy OJ (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* 13 (4): 921-935. - doi: [10.1046/j.1365-294X.2004.02076.x](https://doi.org/10.1046/j.1365-294X.2004.02076.x)

## Supplementary Material

### Appendix 1

**Tab. S1** - List of external populations per oak species, with sample size (*n*) and basic

genetic statistics averaged across 10 microsatellite loci.

**Fig. S1** - Number of clusters (*K*) derived from the Bayesian analysis implemented in STRUCTURE software. Mean and standard deviation of log probability of data over 10 runs are plotted. Calculations were performed with STRUCTURE HARVESTER.

**Fig. S2** - Estimation of number of clusters (*K*) from the rate of change in the log probability of data *L(K)* between successive *K* values, as measured by  $\Delta K$ . Calculations were performed with STRUCTURE HARVESTER.

**Link:** [Curtu\\_1150@suppl001.pdf](mailto:Curtu_1150@suppl001.pdf)