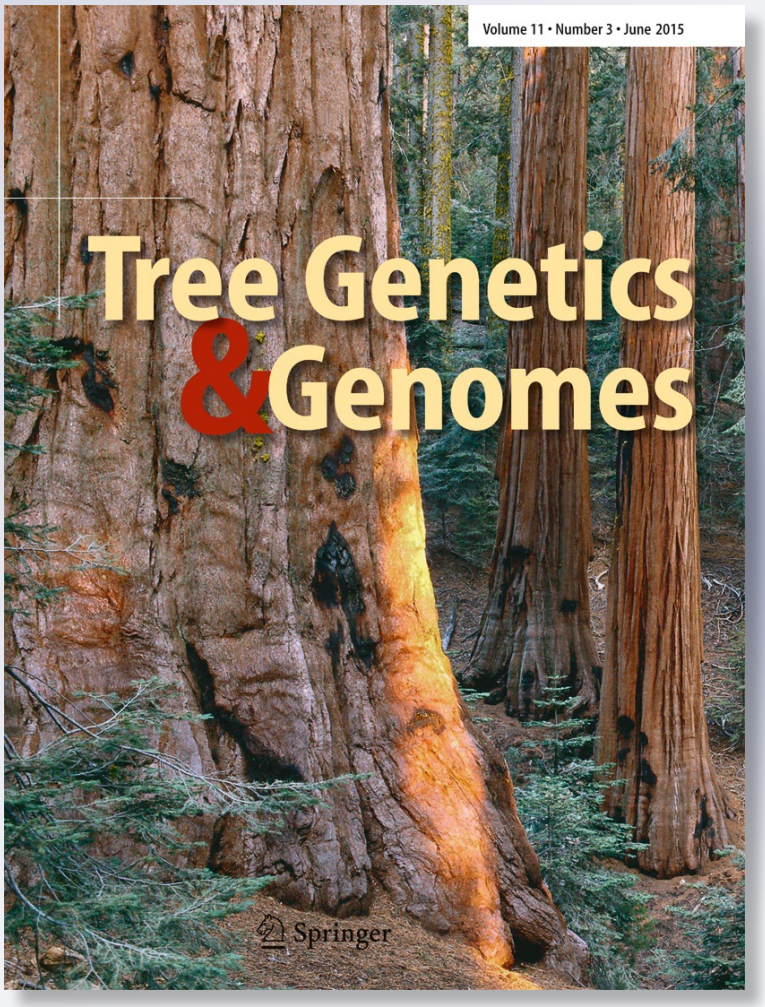
Genetic variation of Central European oaks: shaped by evolutionary factors and human intervention?

Charalambos Neophytou, Stefanie M Gärtner, Rodrigo Vargas-Gaete & Hans-Gerhard Michiels

Tree Genetics & Genomes

ISSN 1614-2942
Volume 11
Number 4

Tree Genetics & Genomes (2015)
11:1-15
DOI 10.1007/s11295-015-0905-7



Springer

Tree Genetics & Genomes (2015) 11:79
DOI 10.1007/s11295-015-0905-7

ORIGINAL ARTICLE

Genetic variation of Central European oaks: shaped by evolutionary factors and human intervention?

Charalambos Neophytou[1] · **Stefanie M Gärtner**[2] · **Rodrigo Vargas-Gaete**[1,3] · **Hans-Gerhard Michiels**[1]

Received: 9 February 2015 / Revised: 9 May 2015 / Accepted: 24 June 2015

Abstract Oak species (*Quercus* spp.) in Central Europe grow on a relatively wide range of sites. Due to the economic importance of oak for its wood and other products, oak forests have long been managed by humans. This raises the question whether adaptation and/or human activities—especially the moving of propagules—have left their footprints on the genetic variation of oak populations. To address this question, we focused on the Upper Rhine Valley, a densely populated area today that was settled by humans early on. Here, the three most common native Central European oak species can be found. We studied their genetic variation across a large number of oak stands, growing on different sites and having different silvicultural histories, using neutral and EST-derived microsatellite markers. At the interspecific level, we showed that *Quercus robur* is relatively well delimited, while *Quercus petraea* and *Quercus pubescens* are more closely related. Natural hybridization might explain the increased genetic introgression between these two species. Within species, we found a low differentiation among populations of *Q. robur* and *Q. petraea*. In spite of forest fragmentation, we detected no spatial genetic barriers. However, we found that populations of *Q. pubescens*, a species with a marginal distribution in the study area were spatially structured. Genetic drift but also unidirectional introgressive hybridization with *Q. petraea* may account for this. Regarding the question of adaptation, we considered soil flooding, texture, drainage, and calcium carbonate in the upper horizons as physiologically important site condition variables. But with multivariate statistics, we could not find any significant effects of these parameters on genetic differentiation. Although there was no evidence for natural selection due to adaptation in stands of *Q. robur*, we demonstrated that age had a significant effect on their genetic variation and that stands established after the end of the Second World War had higher genetic diversity. We interpret these findings as being the result of an increase in large-scale transfers of reproductive materials during this time period and discuss arguments supporting this hypothesis. Finally, we consider the implications of these results for forest management.

Keywords *Quercus* · Genetic differentiation · Genetic structure · Adaptation · Human impact

Communicated by A. Kremer

Electronic supplementary material The online version of this article (doi:10.1007/s11295-015-0905-7) contains supplementary material, which is available to authorized users.

✉ Charalambos Neophytou
chneophytou@gmail.com

[1] Forest Research Institute (FVA) Baden-Württemberg, Wonnhaldestrasse 4, 79100 Freiburg, Germany

[2] Chair of Site Classification and Vegetation Science, Institute of Forest Sciences Faculty of Environment and Natural Resources, University of Freiburg, Tennenbacherstr. 4, 79085 Freiburg, Germany

[3] Laboratorio de Análisis Cuantitativo de Recursos Naturales, Departamento de Ciencias Forestales, Universidad de la Frontera, Temuco, Chile

Introduction

The Upper Rhine Valley, a low-lying strip of land between southwestern Germany and eastern France, has a multitude of forest types that contain different oak species (*Quercus* spp.). They are influenced by diverse site conditions such as water availability and flooding (Oberdorfer 1992; Boeuf

Published online: 12 July 2015

2014). Moreover, the Upper Rhine is an area of early human settlement and because oaks have been a valuable source of wood and other products since antiquity, oak forests have been affected by human activities throughout history (Behre 1988; Küster 1996). For these reasons, the Upper Rhine Valley is an ideal area to study the genetic variation of Central European oaks resulting from natural evolutionary factors and human impacts.

Oak species migrated to the Upper Rhine Valley relatively early after the onset of the Holocene (about 9,500 to 9,000 years before present), as revealed by palynological evidence (Brewer et al. 2002; Goepp 2007). Currently, there is a good knowledge about post-glacial recolonization pathways based on evidence from chloroplast DNA (cpDNA). It has been shown that all three main ice age refugia (Iberian, Apennine, and Balkan) were involved in the recolonization of the area which is a meeting point of several large-scale recolonization pathways (Petit et al. 2002; Neophytou and Michiels 2013). The fact that current spatial structures of cpDNA still carry the imprints of this historical migration event strongly supports that the majority of extant populations are autochthonous. Moreover, it suggests that oaks could adapt to the local conditions over a large number of generations.

Indeed, oak populations in the Upper Rhine Valley can be found on a wide variety of sites because the three indigenous species have different environmental requirements. *Quercus robur*, the species with the widest distribution in the area, in contrast to the other two species, can tolerate waterlogging. However, *Quercus petraea* requires well-drained soils and can perform well on acidic soils. *Quercus pubescens* is less competitive than the other two species and only has a marginal distribution in the area, but it is the most drought-tolerant of the three and differs from *Q. petraea* in that it mostly occurs on alkaline soils. In spite of these differences, the environmental requirements and distribution of all three species overlap in the area. In addition, differences in site conditions can be found among populations of the same species. For instance, *Q. robur* can grow on periodically flooded and also on relatively dry sites (Oberdorfer 1992; Boeuf 2014). Such different adaptations may have led to within-species genetic differentiation through natural selection (Herzog 1996; Herzog and Krabel 1999).

Apart from natural selection and migration, genetic drift could also have shaped the genetic variation of the Upper Rhine oaks. In general, inbreeding levels in oak populations have been shown to be low, due to large-scale gene flow through wind pollination (Buschbom et al. 2011; Gerber et al. 2014) and mechanisms preventing self-fertilization (Abadie et al. 2012). However, an increase in isolation can lead to a decrease in gene flow towards the isolated population (Gerber et al. 2014). Under these circumstances, marginal populations may be more prone to genetic drift (Chybicki et al. 2012). In the Upper Rhine, *Q. pubescens* is a species with a marginal distribution limited to small, isolated stands and therefore prone to genetic drift. But also *Q. robur* and *Q. petraea* populations show some degree of fragmentation, mainly due to agriculture and settlements.

In addition to intraspecific gene flow, hybridization is also among the factors affecting the genetic variation of oaks. All three indigenous species of the Upper Rhine Basin are interfertile and can form fertile hybrids naturally (Lepais and Gerber 2011). Levels of hybridization between these species are variable, and the factors determining interbreeding are complex. Intrinsic reproductive barriers like pollen–pistil incompatibilities (Abadie et al. 2012) and species relative abundance in mixed stands (Lepais et al. 2009) are among the factors known to affect not only the intensity but also the direction of introgression. Furthermore, natural selection acting to promote or prevent hybrids may also affect patterns of introgressive hybridization (Petit et al. 2004). In the Upper Rhine region, the fact that all three species share the same spatial variation of cpDNA provides evidence that hybridization and introgression have happened in the past (Neophytou and Michiels 2013).

Finally, human activities may also have been a significant factor affecting oak genetic variation in the Upper Rhine Valley. Because oaks have many uses, they have been intensively used since antiquity, and their abundance has often been promoted by humans in Central Europe (Behre 1988). In addition to the possibility of their being indirectly impacted though forest fragmentation, other human activities could also have directly influenced genetic variation, for example through seed transfer and artificial regeneration. However, seed transfer during antiquity and the Medieval Ages would have been geographically restricted due to limitations on travel. Oak forests must have been predominantly regenerated with local reproductive material and continued to be until the first half of the twentieth century. The fact that current cpDNA spatial structure still reflects post-glacial recolonization supports this hypothesis (Neophytou and Michiels 2013; Petit et al. 2002).

On the other hand, since the development of the rail network and the establishment of new infrastructure during the nineteenth century, more intensive seed transfer over long distances has taken place (Krabel et al. 2010). For example, large amounts *Q. robur* seeds of Balkan origin were transferred to Central Europe in the second half of the nineteenth century (Gailing et al. 2007). A remarkable increase in the amount of forest tree seeds traded was observed during the twentieth century (Krabel et al. 2010). This may explain the erosion of the original cpDNA spatial pattern variation within younger oak stands established in the Upper Rhine Valley in the second half for the twentieth century compared to older stands, as presented recently in Neophytou and Michiels (2013). In particular, new (compared to the regional spatial pattern) and probably introduced cpDNA haplotypes were found in these

young stands in addition to the autochthonous ones (i.e., those that fit to cpDNA spatial pattern that arose through post-glacial recolonization). Compared to older populations, these were found to be significantly more diverse in terms of cpDNA. This result suggests a more intensive use of allochthonous reproductive material, accounting for the presence of new cpDNA haplotypes, while natural regeneration might have also occurred and contributed to the conservation of the autochthonous variants.

As a follow-up to the aforementioned study, we analyzed the genetic variation of oak populations in the Upper Rhine Valley based on nuclear markers. In particular, we investigated to what extent genetic variation can be explained by geographic location, population history, and site conditions. First, we analyzed interspecific differentiation and introgression among the indigenous oak species of the region. Second, we studied intraspecific differentiation by analyzing genetic structures with and without use of spatially explicit approaches and by testing for isolation by distance. Third, we looked at relationships between variations in genetics and site conditions. For our analysis, we took advantage of a large number of sampled populations covering most of the species' ecological amplitudes in the region. Fourth, we explored the influence of human management on the genetic variation of oak forests. We hypothesized that increased use of allochthonous reproductive material has led to an alteration of the genetic composition over time and to an increased genetic diversity in younger stands.

Materials and methods

Study area, sample collection, and genotyping

The study area was delimited by the Vosges Mountains to the west, the Jura Mountains to the south, and the Black Forest to the east (Fig. 1). A total of 2,048 *Q. robur*, *Q. petraea*, and *Q. pubescens* trees were systematically sampled from 76 forest stands in the Upper Rhine Valley in France and Germany. All of them were georeferenced. Depending on the season of sampling, one of three different types of plant tissue—cambium, leaves, or buds—were collected from each individual for the DNA analysis.

Four physiologically relevant variables were used to characterize the site conditions of each tree sampled based on digital site condition maps from the Forest Research Institute of Baden-Württemberg (FVA) in Germany and the Office National des Forêts (ONF) in France. Because of the different assessment protocols in France and Germany, we could only include them as a standardized binary scale. The included site conditions were soil texture (loam/clay or sand/gravel), periodic flooding (present or absent), temporary anaerobic conditions in the root area due to poor soil drainage (expressed by hydromorphic characteristics in the upper soil horizons, less than 40 cm below the soil surface; present or absent), and calcium carbonate in the upper soil horizons (upper 40 cm, detected through the release of carbon dioxide (CO_2) when soil is treated with hydrochloric acid (HCl); present or absent).

Also, each population was assigned a code for its location. We chose to include spatial information because (1) oak forests in the study area are fragmented to a certain extent (Fig. 1) and (2) several populations are located close to each other. Age information (time since stand establishment in decades) was taken from forest inventory data from both countries. Finally, information about the refugial origin of the study stands was obtained from Neophytou and Michiels (2013). In that study, the same stands had been partitioned into groups of related chloroplast DNA haplotypes indicating different refugial origins based on a spatial analysis of molecular variance (SAMOVA; Dupanloup et al. 2002). Details about the study stands are included in the supplementary material (Online Resource 1).

All sampled trees were analyzed at 21 nuclear microsatellite loci. The analyzed loci are presented in Table 1. Details on DNA extraction, PCR conditions, and genotyping were described in detail in Neophytou (2014).

Species identification and purebred characterization

For species assignments and purebred selection, we conducted a Bayesian clustering analysis using the software STRUCTURE v. 2.3.3 (Pritchard et al. 2000; Falush et al. 2003). Based on multilocus data, subdivision into species groups was detected and individuals were assigned a proportion of membership (q) to each of the species groups. For STRUCTURE analysis, we chose the admixture model and correlated allele frequencies. We performed 30 independent runs applying 100,000 burn-in replications followed by 100,000 MCMC iterations for three assumed subpopulations (K=3). Clustering problems with STRUCTURE using the same dataset had been previously detected (Neophytou 2014) and were taken into account. In particular, it had been shown that STRUCTURE tends to assign species represented by smaller samples to the same cluster. We identified runs with such spurious partitions through their reduced posterior probability of data and discarded them. For the remaining runs, which showed reasonable partition, we inferred the optimal cluster alignment and calculated the average membership proportions for each individual using the software CLUMPP v.1.1.2 (Jakobsson and Rosenberg 2007). Finally, we produced bar charts with individual admixture coefficients using the cluster visualization program DISTRUCT (Rosenberg 2004). Individuals with admixture coefficients (q) of at least 0.9 were characterized as purebreds, because this value maximizes the overall performance of assignment using the given marker set (Neophytou 2014).

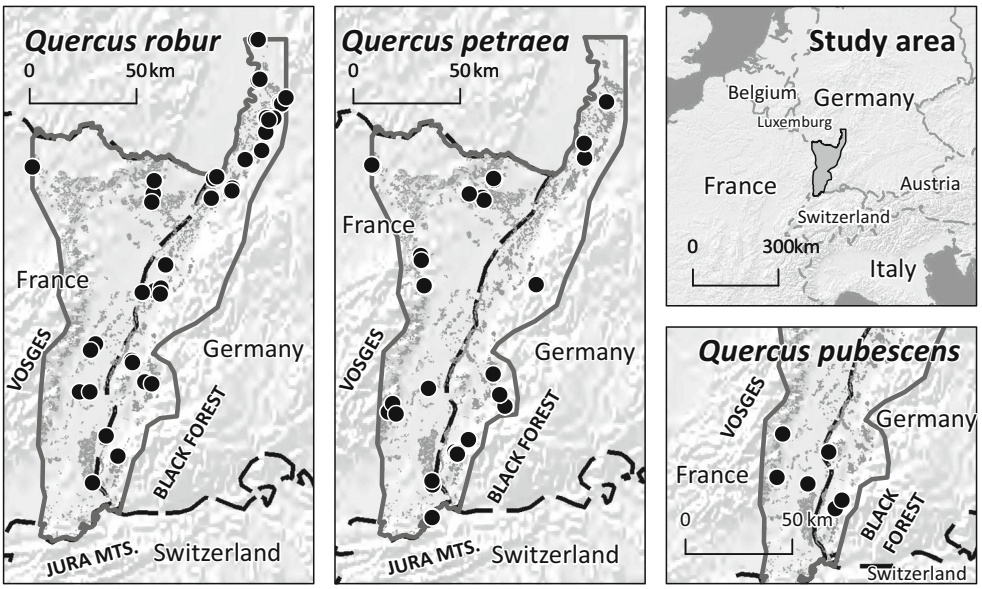


Fig. 1 Geographic location of the study area and study populations. The study area is outlined with a *continuous dark gray line*, national borders with a *dashed line*, and the study stands with *black circles*. The distribution of each species' population is shown on three separate maps. On these maps, forests with oak proportions making up more than 30 % of the canopy are the areas *shaded gray* within the study area

Population genetic structure within species

Next, we repeated STRUCTURE analyses within species, in order to detect possible population subdivisions. For these analyses, we used only genetically purebred individuals. The number of assumed subpopulations (K) varied, depending on species. For *Q. robur*, we set K to vary from 1 to 20, for *Q. petraea* from 1 to 10, and for *Q. pubescens* from 1 to 6. The number of assumed subpopulations was a compromise between the sample size of each species and computational resources. We performed ten independent runs for each K value. Analyses were performed once without a prior

Table 1 Name, allele length range found in our study, fluorescent labeling, and literature citation of the amplified loci and subdivision into multiplexes

	Locus name	Measured allele length span (bp)	Fluorescence labeling	Citation
Multiplex kit 1	QpZAG110	189–256	Green (HEX)	Steinkellner et al. (1997)
	QrZAG7	111–163	Blue (6-FAM)	Kampfer et al. (1998)
	QrZAG11	241–310	Yellow (Atto-550)	Kampfer et al. (1998)
	QrZAG30	170–278	Blue (6-FAM)	Kampfer et al. (1998)
	QrZAG96	140–192	Yellow (Atto-550)	Kampfer et al. (1998)
	QrZAG112	69–137	Green (HEX)	Kampfer et al. (1998)
Multiplex kit 2	QpZAG9	185–285	Green (HEX)	Steinkellner et al. (1997)
	QpZAG15	104–155	Green (HEX)	Steinkellner et al. (1997)
	QpZAG104	179–255	Blue (6-FAM)	Steinkellner et al. (1997)
	QrZAG101	123–181	Blue (6-FAM)	Kampfer et al. (1998)
	MSQ13	187–231	Green (HEX)	Dow et al. (1995)
Multiplex kit 3	PIE020	96–120	Blue (6-FAM)	Durand et al. (2010)
	PIE102	139–172	Green (VIC)	Durand et al. (2010)
	PIE152	224–261	Blue (6-FAM)	Durand et al. (2010)
	PIE215	191–215	Yellow (Atto-550)	Durand et al. (2010)
	PIE223	197–227	Blue (6-FAM)	Durand et al. (2010)
	PIE227	153–178	Red (PET)	Durand et al. (2010)
	PIE242	107–130	Green (VIC)	Durand et al. (2010)
	PIE243	206–229	Green (VIC)	Durand et al. (2010)
	PIE267	89–114	Yellow (Atto-550)	Durand et al. (2010)
	PIE271	183–214	Red (PET)	Durand et al. (2010)

distribution and once using location (as defined above) information to define prior distribution for genetic assignments which is recommended for weakly structured data (Hubisz et al. 2009). In order to choose the most appropriate number of clusters we used the method of Evanno et al. (2005). According to this method, ΔK, an ad hoc statistic based on the rate of change of the maximum posterior probability of data is calculated for each value of K. The value of K for which ΔK is maximized indicates the uppermost hierarchical level of population subdivision. Clustering alignments and the visualization of the results were carried out as described above.

Given that the geographical coordinates of the sampled trees were available, we additionally applied SAMOVA (Dupanloup et al. 2002) in order to detect possible spatial structure. This method aims to identify groups of geographically adjacent populations with minimum genetic differentiation within and maximum genetic differentiation among groups. In this way, it is possible to locate spatial genetic barriers caused by fragmentation. We set the predefined number K of such groups from 2 to 20 in *Q. robur*, from 2 to 10 in *Q. petraea*, and from 2 to 5 in *Q. pubescens*. We ran 1,000 iterations for each K for each of 100 simulated annealing processes. To choose the optimal group number (K), we compared F_{CT} values among different K configurations. On one hand, the highest F_{CT} represents the optimal configuration and on the other higher K configurations containing single population groups indicate lack of genetic structure (Heuertz et al. 2004). Thus, among all K values without single population groups we chose the K value that resulted in the highest F_{CT} as the optimal clustering solution.

Interpopulation genetic differentiation and diversity

We calculated genetic differentiation and diversity within and between species by treating purebreds of different species from a specific forest stand as separate populations. First, for each population, we calculated the number of alleles per locus, expected heterozygosity, allelic richness with rarefaction (Petit et al. 1998), a measure which is largely independent of sample size, as well as F statistics based on Weir and Cockerham (1984). Rarefaction size was four, corresponding to the minimum number of genotyped loci in the smallest population. Second, by treating the species as continuous populations, we computed interspecific F_{ST}s. For all aforementioned analyses, we used the software FSTAT (Goudet 1995).

To illustrate patterns of genetic differentiation within and between species, we computed Nei's (1978) genetic distances (D_A) between all population pairs using the software GenAlEx (Peakall and Smouse 2006). Subsequently, we constructed a neighbor-joining tree using the Phylip software (Felsenstein 2002), which we plotted with the software HyperTree (Bingham and Sudarsanam 2000).

Genetic variation in relation to geographic location, environmental conditions, and human management

The Mantel test was used, first to analyze isolation by distance (IBD), second isolation by adaptation (IBA), and third as a partial regression on three distance matrices (Diniz-Filho et al. 2013): genetic dissimilarity (Nei's genetic distances), environmental dissimilarity (sum of differences over site condition variables: equal value for a specific variable yielded 0, a different value yielded 1), and geographic location (Euclidian distance). The distance matrices were calculated in GenAlEx (Peakall and Smouse 2006). In the case of environmental dissimilarity, for constructing input data sheets, each site condition variable was treated as a locus (obtaining values of 0 or 1) using haplotypic data format. Mantel tests were done for *Q. robur* and *Q. petraea* (*Q. pubescens* was represented only by six populations), using the functions mantel and partial.mantel of the package vegan 2.0-10 (Oksanen et al. 2013) of the R statistical sofware (R Development Core Team 2013) based on 999 permutations and Kendall's rank correlation tau.

To evaluate the influence of the site conditions (as described by the variables mentioned previously), human influence (by including age as described above assuming that long-distance seed transfer has increased in the last decades) and refugial origin (for the refugial origin, we assigned the populations to genetically differentiated groups established by SAMOVA (Dupanloup et al. 2002) carried out using cpDNA data from subsamples of the same populations—presented in Neophytou and Michiels (2013)) as well as their interaction on genetic differentiation, we conducted a permutational multivariate analysis of variance (PERMANOVA) (Anderson 2001). The analysis partitions the sums of squares of a multivariate data set analogous to the multivariate analysis of variance but uses semimetric and metric distance matrices (Nei's genetic distances and Euclidian distances for the other variables). Significance tests (a=0.05) were done using F tests based on sequential sums of squares from permutations of the raw data (function anosim in the package vegan, Oksanen et al. 2013). Input data for the PERMANOVA are presented in Online Resource 2.

Furthermore, we tested for differences in genetic diversity and differentiation between different age classes and species. For this purpose, we carried out two-sided permutation-based tests comparing allelic richness, expected heterozygosity, and F_{ST} and F_{IS} between "young" stands (with an age of less than 70 years) and "old" stands (with an age of more than 100 years) separately within *Q. robur* and *Q. petraea* using FSTAT (comparison among groups of samples, Goudet 1995). Finally, we carried out a series of tests within *Q. robur* and *Q. petraea* by splitting the populations into two categories (if they existed) for each one of the binary environmental variables. A Bonferroni correction to adjust α values was applied depending on the number of tests within each species. For

each test, we applied 1,000 permutations (i.e., stands) between groups (age classes or species).

Results

Species identification, genetic introgression, and purebred assignment

According to the results, out of 30 runs of Bayesian clustering analysis, eight partitions into species were identified. Individuals could be assigned to three clusters corresponding to the three species, based on their multilocus genotypes. About 88 % of the analyzed trees were assigned to one of the three species clusters with a membership proportion of at least 90 % ($q \geq 0.9$), while about 95 % showed a membership proportion of at least 75 % ($q \geq 0.75$; Table 2 and Fig. 2). For the remaining individuals, intermediate q values suggest a hybrid origin. Most of them were intermediates between *Q. robur* and *Q. petraea*. Intermediates between *Q. petraea* and *Q. pubescens* formed the second largest group of putative hybrids and were overrepresented given the relatively small sample sizes of their parental species.

Interpopulation genetic differentiation, within and among species

Separate Bayesian clustering analyses using purebreds of *Q. robur* or *Q. petraea* revealed that further population genetic subdivision within these species was weak. In particular, posterior probability of data was, in both cases, relatively high for $K=1$ and did not increase for higher order K values, either when data were analyzed without use of any prior or when location was used for defining prior distribution for genetic assignments (Fig. 3). When genetic subdivision is present, posterior probability of data increases once the "real" K is reached and then plateaus (Evanno et al. 2005), which was not the case within *Q. robur* and *Q. petraea*. On the other hand, in *Q. pubescens*, posterior probability of data increased between $K=1$ and $K=3$ and fluctuated for higher values, while ΔK was maximized for $K=2$ or $K=3$, depending on the analysis parameters, suggesting genetic structure. Inspection of individual membership proportions, down to the derived clusters, further supports this subdivision within *Q. pubescens* (Fig. 4). For example, for $K=4$, northwestern populations from the foothills of the Vosges clustered together, central populations formed another cluster, while one of the southeastern populations from the foothills of the Black Forest was highly differentiated from all other populations. For $K=5$, results suggest further subdivision within the aforementioned groups.

In contrast to *Q. pubescens*, patterns of population membership proportions in *Q. robur* or *Q. petraea* were not particularly distinctive. In *Q. petraea*, for all values of K, there was a slight latitudinal gradient with membership proportions to specific clusters changing from north to south (Fig. 4). In *Q. robur*, such gradients were even less obvious. These results may suggest some slight local differentiation, either within or between regions (Fig. 4). However, this species presented the lowest values of ΔK without any obvious maxima and, thus, evidence for genetic partition was weak (Fig. 3).

Furthermore, SAMOVA did not support any spatial genetic subdivision in any of the species. For any of the predefined K values tested, there were groups consisting of a single population (results not shown).

Analysis of genetic variation based on F statistics revealed different patterns within each of the study species. Genetic differentiation among populations of *Q. pubescens* was higher than among populations within the other two species as shown by a higher F_{ST} value (Table 3). On the other hand, *Q. petraea* demonstrated the highest within-population heterozygote deficit (F_{IS}). This species showed the highest overall heterozygote deficit (F_{IT}), while the lowest F_{ST} and F_{IT} values were observed in *Q. robur*. With respect to genetic diversity, the average number of alleles per locus was highest in *Q. robur* and lowest in *Q. pubescens*. Furthermore, allelic richness per population was highest in *Q. pubescens*, with *Q. petraea* showing the lowest value. The difference in allelic richness

Table 2 Genetic assignment of individuals to species groups or hybrids according to their multilocus genotypes

Assignment	Membership proportion	Number of individuals and percentage (in parentheses)	
Q. robur	$q \geq 0.9$	1,182 (0.58)	1,254 (0.61)
	$0.9 > q \geq 0.75$	72 (0.04)	
Q. petraea	$q \geq 0.9$	511 (0.25)	569 (0.28)
	$0.9 > q \geq 0.75$	58 (0.03)	
Q.pubescens	$q \geq 0.9$	105 (0.05)	124 (0.06)
	$0.9 > q \geq 0.75$	19 (0.01)	
Hybrids	*Q. robur* × *Q. petraea*	59 (0.03)	101 (0.05)
	Q. robur × *Q. pubescens*	10 (0.00)	
	Q. pubescens × *Q. petraea*	32 (0.02)	

In order to assign intermediates, the two clusters with the highest proportion of membership (q) were selected

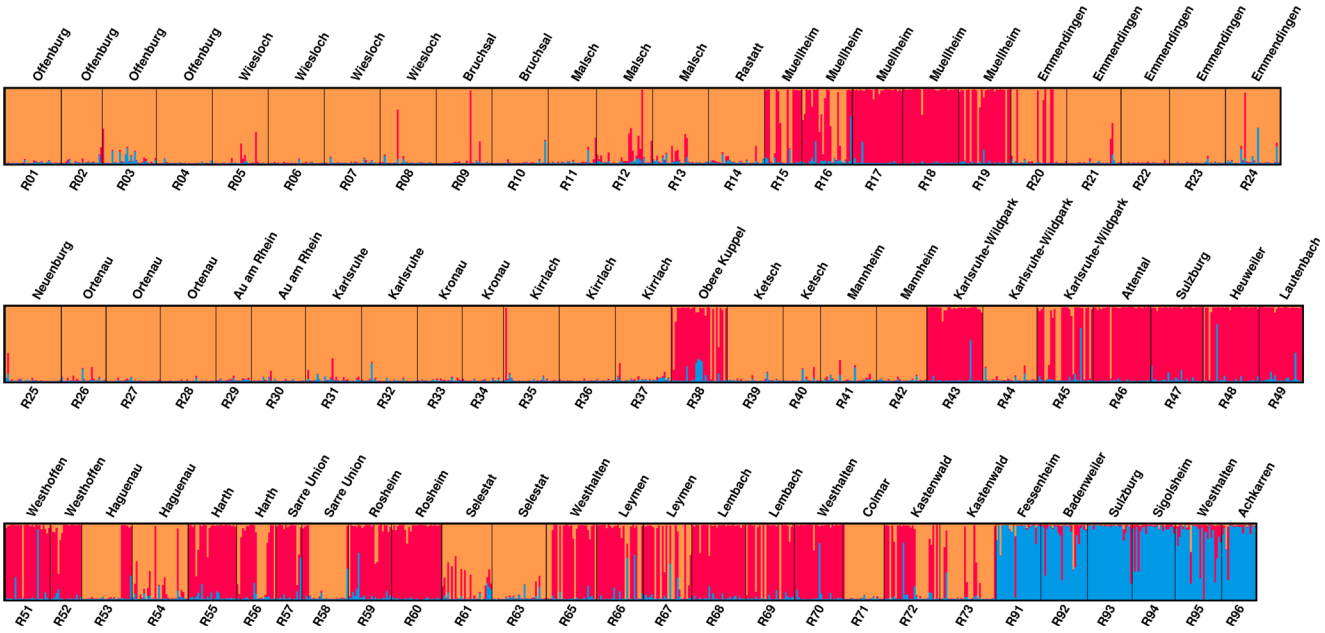


Fig. 2 Individual membership proportion determined by a STRUCTURE analysis across all populations and species. Each individual is represented with a *vertical bar*, and each inferred cluster is marked with a *different color*. *Q. robur* cluster: *orange* (e.g., R01), *Q. petraea* cluster: *red* (e.g., R60), *Q. pubescens* cluster: *blue* (e.g., R93)

between *Q. robur* and *Q. petraea* was highly significant (two-sided permutation-based test, $p=0.001$, see Online Resource 3).

The different patterns of genetic variation within each species became visible in the distance-based neighbor-joining tree (Fig. 5). Cluster length was maximal in *Q. pubescens* and shortest in *Q. robur*, which had demonstrated the highest and lowest within-species genetic differentiation among populations, respectively. We did not observe any biologically explainable clustering patterns within any of the species. However, some individual populations showed relatively high genetic differentiation (e.g., four to six populations within *Q. robur*). Pairwise F_{ST} values involving these populations were mostly significant (Online Resource 4). In addition to intraspecific genetic relationships among populations, the neighbor-joining tree illustrated phylogenetic relationships among the three species. *Q. petraea*

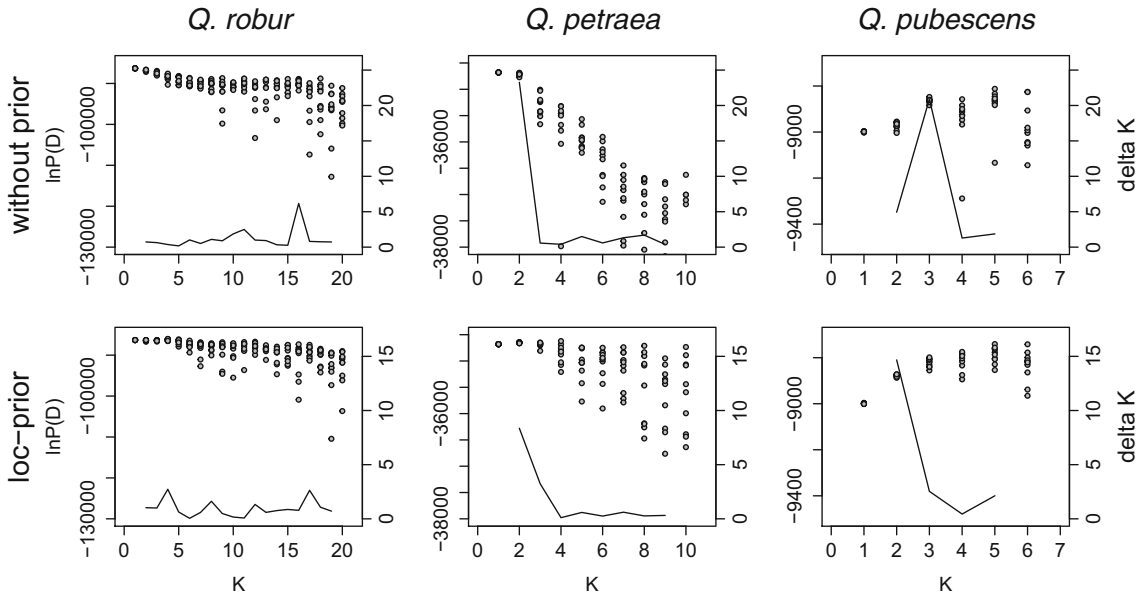


Fig. 3 Posterior probabilities of data—lnP(*D*)— demonstrated by *grey-filled circles* and ΔK (delta *K*) values demonstrated by a *continuous line* for different values of assumed subpopulations (*K*) for intraspecific Bayesian clustering analyses with STRUCTURE. Ten runs for each *K* value were carried out. Results from analyses without prior are presented in the *first row*. Results from analyses using location data as a prior (*loc-prior*) are presented in the *second row*

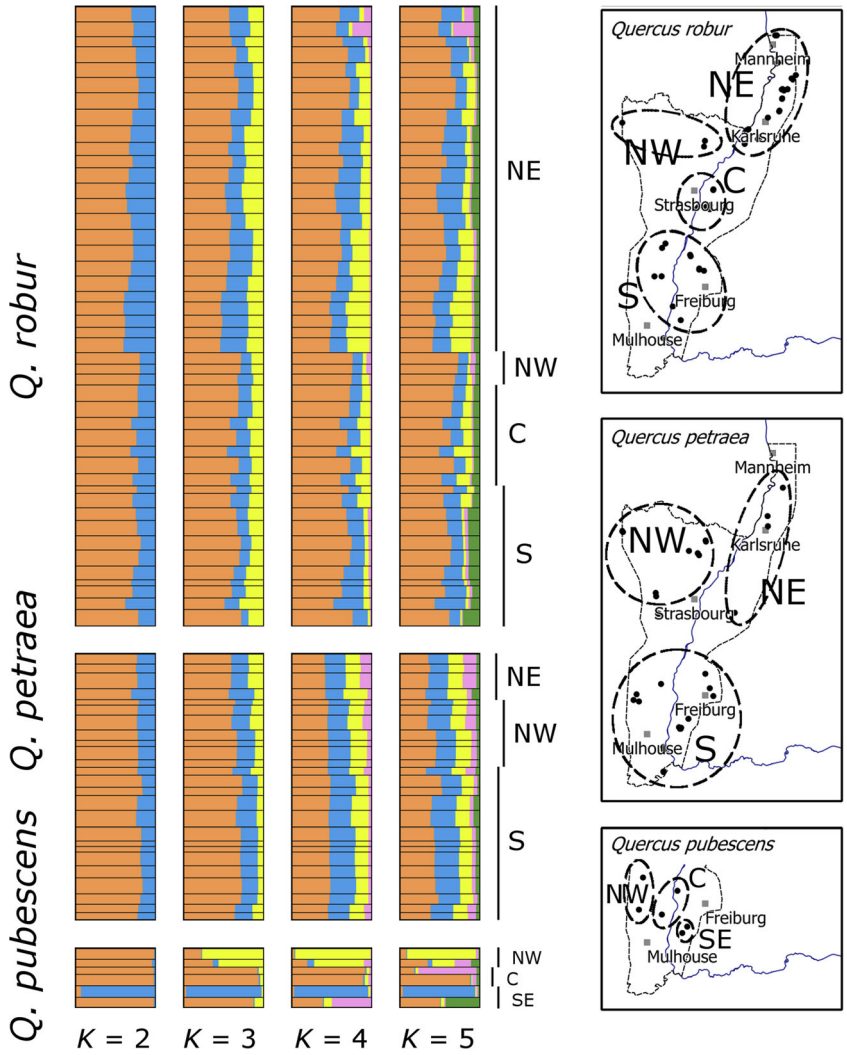


Fig. 4 Results from separate Bayesian clustering analysis within each species using STRUCTURE for K=2, 3, 4, and 5. *Each box* corresponds to a population, and its width is proportional to the population size. The derived clusters are illustrated with *different colors*. Based on their geographic location (not on the results from clustering analysis), populations (designated with *black dots*) are divided into regions shown on the maps. On the graph and within each region, they are sorted by decreasing latitude

and *Q. pubescens* appeared to be more closely related to each other, while *Q. robur* was more differentiated from these two species. When the species were treated as single populations, pairwise F_{ST} values were 0.048 between *Q. petraea* and *Q. pubescens*, 0.097 between *Q. robur* and *Q. pubescens*, and 0.116 between *Q. petraea* and *Q. robur*.

Table 3 Genetic variation within species

	Q. robur (N_{pop}=45)	*Q. petraea* (N_{pop}=25)	*Q. pubescens* (N_{pop}=6)
n_a/L	22.667±3.721	18.286±2.189	15.143±1.540
A	4.378±0.017	4.267±0.019	4.674±0.054
F_{ST}	0.008±0.001	0.010±0.002	0.033±0.004
F_{IS}	0.022±0.007	0.049±0.016	0.008±0.013
F_{IT}	0.030±0.008	0.058±0.016	0.040±0.015

n_a/L average number of alleles per locus, A average allelic richness per population based on a rarefaction size of four individuals, F_{IS}, F_{ST}, and F_{IT} F statistics according to Weir and Cockerham (1984)

Intraspecific genetic variation in relation to geographic location, environmental conditions and human management

The simple Mantel test revealed that there was no correlation between the genetic differentiation of *Q. robur* and the site conditions (rm=0.10; p=0.063). However, there was a very weak but significant effect of isolation by distance (rm=0.08; p=0.025). Thus, the main issue was to test if there was a correlation between genetic differentiation and site conditions after taking geographic distances into account. No significant correlation was found (partial rm=0.096, p=0.063).

In the case of *Q. petraea*, there was neither an isolation by distance effect (rm=0.03; p=0.348) nor a correlation to the site conditions (rm=0.01; p=0.408). Thus, no further partial analyses were needed.

The PERMANOVA showed a highly significant effect of stand age on the genetic differentiation of *Q. robur* (Table 4). Significant impacts were also shown by cpDNA and the interaction of flood and cpDNA.

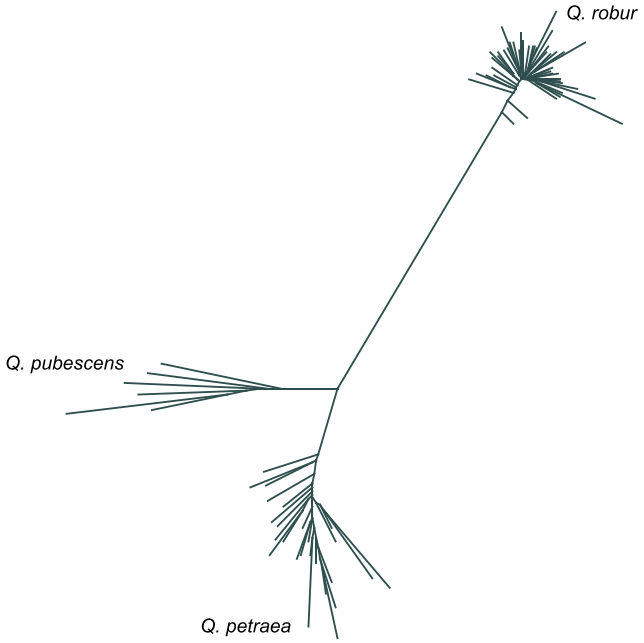


Fig. 5 Neighbor-joining tree constructed based on interpopulation D_A (Nei 1978) distances

In the case of *Q. petraea*, there were no significant effects found in the results of the PERMANOVA (Table 5). In the case of a larger sample, the cpDNA and the interaction between cpDNA and soil might significantly influence the genetic variability. In the present case, these were just tendencies on a significance level of 10 %.

Finally, a comparison of variation and differentiation between different age classes by means of a permutation-based two-sided test (according to Goudet 1995) revealed a significantly higher allelic richness and heterozygosity within younger stands of *Q. robur* (p=0.004 and p=0.021, respectively; significant at the α=0.05 level after Bonferroni correction). In particular, average allelic richness was 4.422 in younger and 4.328 in older stands of *Q. robur* (Fig. 6). In contrast, differences of allelic richness between older and younger stands of *Q. petraea* were minimal (4.267 and 4.266, respectively) and non-significant (p=0.993). No significant differences of F_{ST} or F_{IS} values were found between the two age classes in either species. Moreover, variation measures (allelic richness and heterozygosity) did not differ significantly between population groups growing under different site conditions (Online Resource 3). With respect to F statistics, the F_{IS} value was higher in flooded compared to non-flooded *Q. robur* populations, and the F_{ST} value was higher among populations with calcium carbonate present in the upper soil horizons (populations with Ca presence displayed a higher F_{ST}; Online Resource 3). However, these differences were not significant after Bonferroni correction (significance level declined from 0.05 to 0.01 due to five tests in *Q. robur*).

Table 4 Results of PERMANOVA testing for the effects of site conditions stand age, cpDNA, and their interaction on the genetic differentiation of *Quercus robur*

	Df	*F*	R^2	Pr(>*F*)
Soil	1	149.628	0.02671	0.132
Water	1	113.909	0.02033	0.362
Flood	1	190.102	0.03393	0.063
$CaCO_3$	1	0.96942	0.01730	0.498
cpDNA	4	185.490	0.13244	0.025*
Age	1	256.561	0.04580	0.008**
Soil/water	1	118.060	0.02107	0.315
Soil/flood	1	161.638	0.02885	0.107
Water/flood	1	194.243	0.03467	0.056
Soil/$CaCO_3$	1	121.335	0.02166	0.295
Water/$CaCO_3$	1	132.214	0.02360	0.231
Soil/cpDNA	1	161.313	0.02879	0.101
Water/cpDNA	2	139.351	0.04975	0.161
Flood/cpDNA	2	201.678	0.07200	0.025*
$CaCO_3$/cpDNA	1	112.538	0.02009	0.315
Soil/age	1	0.61900	0.01105	0.837
Water/age	1	105.459	0.01882	0.383
Flood/age	1	0.65366	0.01167	0.784
$CaCO_3$/age	1	112.440	0.02007	0.377
cpDNA/age	1	149.238	0.02664	0.147
Soil/flood/cpDNA	1	0.70888	0.01265	0.767
Soil/$CaCO_3$/cpDNA	1	0.57973	0.01035	0.836
Soil/$CaCO_3$/age	1	146.508	0.02615	0.179
Residuals	16		0.28560	
Total	44		1.00000	

Soil soil texture (loam/clay or sand/gravel), *flood* periodic flooding (present or absent), *water* temporary anaerobic conditions in the root area due to poor soil drainage (present or absent), *$CaCO_3$* calcium carbonate in the upper soil horizons (present or absent), cpDNA membership to spatial genetic clusters based on chloroplast DNA differentiation, *age* age of study stand in decades

***p<0.001, **p<0.01, *p<0.05

Discussion

Good species delimitation of *Q. robur* vs. a closer genetic affinity and higher levels of genetic introgression between *Q. petraea* and *Q. pubescens*

One of the primary aims of the present study was the investigation of genetic differentiation of the Upper Rhine oaks within and between species. Bayesian clustering analysis based on 21 nuclear microsatellites could efficiently subdivide them into species. The great majority of individual trees could be assigned to species clusters with high membership proportions. A relatively small fraction of the sampled trees displayed affinity to more than one cluster, suggesting a hybrid origin.

Table 5 Results of PERMANOVA testing for the effects of site conditions: stand age, cpDNA, and their interaction on the genetic differentiation of *Quercus petraea*

	Df	F. model	R^2	Pr(>*F*)
Soil	1	0.78990	0.03494	0.705
$CaCO_3$	1	0.63875	0.02825	0.809
cpDNA	2	145.009	0.12828	0.075
Age	1	0.48511	0.02146	0.938
Soil/cpDNA	2	152.396	0.13481	0.078
$CaCO_3$/cpDNA	1	0.24384	0.01079	0.990
Soil/age	1	0.94780	0.04192	0.502
cpDNA/age	2	0.57850	0.05117	0.931
Soil/cpDNA/age	1	0.39829	0.01762	0.947
Residuals	12		0.53077	
Total	24		1.00000	

Soil soil texture (loam/clay or sand/gravel), *$CaCO_3$* calcium carbonate in the upper soil horizons (present or absent), *cpDNA* membership in spatial genetic clusters based on chloroplast DNA differentiation, *age* age of study stand in decades

***$p<0.001$, **$p<0.01$, *$p<0.05$

Most of them were assigned as hybrids between *Q. robur* and *Q. petraea* (59 trees with $q \leq 0.75$), which was expected, given that these two species were represented by the largest samples in our study (1,182 *Q. robur* and 511 *Q. petraea*). Evidence for genetic introgression between *Q. robur* and *Q. pubescens* was even weaker (ten individuals with $q \leq 0.75$ compared to 1,182 *Q. robur* and 105 *Q. pubescens*). A large number of studies have shown that interspecific mating between *Q. robur* and the other two species is relatively limited and does not lead to widespread genetic introgression (Streiff et al. 1999; Curtu et al. 2009; Gerber et al. 2014). Pre-mating and post-mating reproductive barriers have a restrictive effect on gene flow (Abadie et al. 2012; Lepais et al. 2013).

Fig. 6 Allelic richness distribution within two different age classes in *Q. robur*. The class of old stands includes populations at least 100 years old and the class of young stands includes populations up to 70 years old

However, hybridization between *Q. petraea* and *Q. pubescens* seems to be more prevalent. Compared to other species combinations in our study, intermediates between *Q. petraea* and *Q. pubescens* were overrepresented, given their relatively small sample size. In addition to hybridization, close genetic affinity exacerbating genetic assignment might account for this result. Indeed, low genetic differentiation reduces the accuracy of Bayesian clustering analyses (Vähä and Primmer 2006). On the one hand, the use of 21 marker loci has proved adequate to efficiently recognize purebreds of *Q. petraea* and *Q. pubescens* in our case. Contrariwise, it has been shown that STRUCTURE analysis, based on this marker set, is only moderately efficient in identifying first generation hybrids between these two species, and it only poorly identifies backcrosses; many of which are often classified as purebreds (Neophytou 2014). Therefore, genetic introgression between *Q. petraea* and *Q. pubescens* in the Upper Rhine could be even higher than indicated by the results of our analysis. Previous studies based on morphology also suggest a high degree of introgression between the two species (Aas 1989; Dupouey and Badeau 1993; Müller 1999; Bruschi et al. 2000). Based on leaf trichomes and morphometric traits, Aas (1989) suggested that Swiss populations of *Q. pubescens* located in the northern Alps (in an area adjacent to our study area) should be regarded as hybrid swarms mixed with *Q. petraea*.

Direct evidence about hybridization between *Q. petraea* and *Q. pubescens* has been provided in other studies. By means of controlled pollination, it has been shown that interspecific crosses between *Q. petraea* and *Q. pubescens* are more successful than crosses of the two species with *Q. robur* (Müller 1999). Furthermore, under natural conditions, reproductive success of *Q. pubescens* as a pollen donor during interspecific crosses with *Q. petraea* is much higher than in reciprocal crosses (Müller 1999; Lepais and Gerber 2011; Lepais et al. 2013). Such directional introgression is a possible explanation for the higher degree of genetic diversity among populations of *Q. pubescens* in the Upper Rhine, compared to the other two species. If *Q. pubescens*, but also hybrids between the two species, are more likely to pollinate *Q. petraea* than the reverse (Lepais and Gerber 2011), then it is likely that allelic variants from *Q. petraea* have been acquired by *Q. pubescens* through pollen-swamping events. Moreover, the higher interpopulation differentiation in *Q. pubescens*, compared to the other two species in this study, could also be due to this unidirectional introgression. Depending on species relative abundance and the competition between the two species, populations assigned to *Q. pubescens* may show different degrees of introgression at each site, leading to high genetic differentiation among the populations of this species.

Lack of spatial genetic structure within *Q. robur* and *Q. petraea* indicates sufficient large-scale gene flow

In contrast to *Q. pubescens*, the other two species appeared to be much more genetically homogenous. First, global F_{ST} values, either in *Q. robur* or *Q. petraea*, were very low (around or below 0.01). Second, we could not find convincing evidence for intraspecific genetic structure. Both spatially explicit methods used (STRUCTURE with location prior and SAMOVA) did not identify any subdivision in any of the two species. Even if some STRUCTURE results may suggest some slight differentiation among regions, especially in *Q. petraea*, this signal of genetic structure is very weak (as shown for instance by the patterns of lnP(D) and ΔK). In light of these results, we cannot confirm the presence of any spatial genetic discontinuities, despite forest fragmentation in the area. Third, Mantel tests showed a very weak correlation between genetic and geographic distances which also supports a high homogeneity among populations of *Q. robur*, as well as *Q. petraea*.

The low genetic differentiation within these two species which are widespread in our study area is not surprising. Even if the forests of each species are not continuous, large distance gene flow might account for a high degree of genetic exchange among populations resulting in a highly homogenous gene pool. This has been confirmed by paternity analyses carried out in progenies of isolated populations in other studies with the same study species. For instance, significant external gene flow (up to 35 %) was found in a marginal population of *Q. robur* which was isolated by 80 km from any other individual of the species (Buschbom et al. 2011). In another example, Gerber et al. (2014) could detect mating events between oak trees growing in different scattered populations lying approximately 10 km from each other. Our *Q. robur* and *Q. petraea* study forests were often broken up by inhabited areas or agricultural lands, but the distances between disjoint populations were limited to only a few kilometers, whereas populations of *Q. pubescens* were further apart from each other. In this species, genetic drift due to isolation and the small size of the forests may have additionally contributed to the higher interpopulation genetic differentiation.

No relationship between environmental and genetic variation: no evidence of natural selection at the analyzed marker loci

Given the wide distribution and wide ecological amplitude of *Q. robur* and *Q. petraea* populations in the study area, we considered natural selection as an additional factor potentially affecting the genetic constitution of these species. However, we could not provide any strong evidence for an effect of site conditions on genetic variation as a result of adaptation. For none of the four site condition variables could we find a significant relationship between site condition variation and genetic differentiation at the given locus set in the two species. Furthermore, the Mantel test for IBA was not significant for both *Q. petraea* and *Q. robur*. However, the probability value was low in the latter species. Additionally, in *Q. robur*, the interaction of periodic flooding with cpDNA found in PERM ANOVA was significant at the 0.05 level. Given the relatively high probability values, even if these results indicate a trend, we cannot consider them as strong evidence that the genetic and site condition differences are related to each other. Use of more detailed site condition data and genotypic data from a higher number of loci, especially from DNA regions coding for physiologically relevant traits, could lead to a more powerful test of the aforementioned hypothesis.

In spite of recent developments in genomic research, other studies based on molecular markers have not yet provided strong evidence about genetic differences between oak populations due to natural selection. In *Q. robur* and *Q. petraea*, genetic differentiation at expressed loci was often proven to be of the same magnitude as at neutral loci, while adaptive loci did not necessarily show higher genetic differentiation among contrasting habitats, compared to neutral ones (Derory et al. 2010; Homolka et al. 2013; Vidalis et al. 2013). Herzog and Krabel (1999) reported allele frequency and variation differences at certain isozyme loci between two *Q. robur* populations from the same region of the Upper Rhine plain with different flooding regimes. However, it could not be proved whether this was a result of adaptation. On the other hand, trade-offs between genetic variation at neutral loci and site condition differences could be found at the multivariate level in other oak species (Grivet et al. 2008; Sork et al. 2010). These findings are probably due to the polygenic character of the great majority of adaptive traits resulting in minor to moderate effects of natural selection across the genome ("soft selective sweeps"; Pritchard et al. 2010). Thus, the use of large molecular marker sets (both adaptive and neutral) with high genome coverage (aiming at "outlier" detection) rather than candidate gene approaches, along with detailed site condition data, might be the most appropriate strategy for studying molecular signatures of natural selection.

Furthermore, physiological and phenotypic plasticity should not be overlooked as factors leading to the adaptation of the Upper Rhine oaks—especially *Q. robur*—on a variety of sites. They have been shown to be essential for the establishment of plants with pioneer traits like our study species (Zenni et al. 2014), and they may also act to limit the effects of natural selection on the genome.

Significant effect of age on genetic variation and high diversity in young stands of *Q. robur*: a result of intensified human impact through large-scale seed transfer?

Out of all of the factors, tested by means of a PERMANOVA, only age was shown to have a highly significant ($p<0.01$)

effect on a population's genetic variation, and this could be observed only in *Q. robur*. Additionally, younger populations of *Q. robur* displayed a higher genetic diversity. We hypothesize that this result was likely due to more intensive long-range seed transfer in the last seven decades rather than due to the effects of natural selection or other evolutionary factors. A similar pattern of genetic diversity contrasting young and old stands was also observed in our study populations based on independently inherited (from nuclear ones) and selectively neutral chloroplast DNA markers. In particular, younger stands (less than 70 years old) of *Q. robur* showed similar spatial structure but higher genetic diversity and more haplotypes (also not regionally specific ones) per population when compared to older ones (more than 100 years old; Neophytou and Michiels 2013). This supports the theory of increased long-distance seed transfer in the last 70 years which might have led to the introduction of new genetic variants. Moreover, the fact that cpDNA as well as the interaction of cpDNA and flooding showed a significant effect on nuclear genetic differentiation in the present study (with a higher probability than age) could also be the result of human activities. Also, seed transfer directly affects the distribution of cpDNA and may have resulted in local patterns which are different than the large-scale pattern shaped by post-glacial recolonization (also discussed in Neophytou and Michiels 2013). Furthermore, the intensity of seed transfer may have been different between flooded and non-flooded areas. Therefore, we conclude that human impact, accounts for the differences at nuclear DNA loci observed in the present study.

Alternatively, natural or artificial (i.e., from silvicultural practices) selection may also have accounted for genetic differentiation over time and may be reflected in the differences between old and young stands. For instance, in early isozyme-based population genetic studies in *Q. robur* and *Q. petraea*, higher genetic differentiation was found among older rather than among juvenile stands, and this has been attributed to natural selection (Kremer et al. 1991; Herzog 1996). Although we used a fairly large number of populations, we could not find support for this hypothesis. First, even if the F_{ST} values among older populations in our study were slightly higher than among younger ones in both *Q. robur* and *Q. petraea*, the differences were not significant (Online Resource 3). Second, if the differences were due to natural selection, then this should not affect the variation of chloroplast DNA haplotypes.

Changes in silvicultural treatments during the last several decades also support our hypothesis. Until the beginning of twentieth century, coppice-with-standards had been the predominant silvicultural practice in oak forests. Coppicing was applied at short intervals, while standards, i.e., single stemmed trees, were retained for longer periods and used as seed sources for natural or artificial regeneration (Sebald et al. 1998). This silvicultural treatment promoted the conservation of autochthonous genetic types. A change from the coppice-with-standards system to high forests started in the nineteenth century but was especially prevalent in the twentieth. In high forests, oaks (especially *Q. robur*) are characterized by a reduced ability to compete with more shade-tolerant species like beech (*Fagus sylvatica*) thus making natural regeneration even more difficult (Reif and Gärtner 2007; Uhl et al. 2008). Currently, planting and seeding are the predominant methods used to regenerate oak stands in Germany (Sebald et al. 1998). At the same time, the development of transportation infrastructure facilitated the trade of acorns. The amount of seed traded strongly increased during the first half of the twentieth century, especially after a period of limited activity between the two World Wars (König et al. 2002; Krabel et al. 2010). This is in agreement with the differences in genetic diversity found in our study between stands of *Q. robur* less than 70 years old on one hand and those more than 100 years on the other.

Finally, differences in ecological traits and silvicultural treatments between *Q. robur* and *Q. petraea* might explain why the previously mentioned differences between older and younger stands were not observed in both species. Due to a higher requirement for light, natural regeneration of *Q. robur* is less likely compared to *Q. petraea* which has later successional characteristics and is more shade-tolerant. Thus, *Q. robur* has been disadvantaged as a result of the move away from the coppice-with-standards system to a system of high forests. On the contrary, *Q. petraea* is a natural component of mixed high forests with beech (Sebald et al. 1998) and could probably adapt better to this change in forestry system. Furthermore, the draining of former alluvial sites in the second half of the nineteenth century resulted in a reduction in competitiveness and the failure of natural regeneration in *Q. robur* (Uhl et al. 2008). On the French side, natural regeneration has been successfully used in *Q. petraea* stands but not in *Q. robur* stands. On the German side, there is a higher potential for natural regeneration in *Q. petraea* stands and is recommended as a means to regenerate them. In addition, the amount of imported *Q. robur* acorns to Germany in recent years (2003–2008) was at least 40 times higher than for *Q. petraea* (Krabel et al. 2010). Thus, artificial regeneration with allochthonous reproductive material might have been much more widespread in *Q. robur* forests resulting in (1) an alteration of the genetic basis of autochthonous forests and (2) an increase of genetic diversity.

Conclusions and implications for conservation and forest management

By and large, both *Q. robur* and *Q. petraea* display a high degree of genetic homogeneity in the Upper Rhine Valley. Neither landscape fragmentation nor adaptation to variable site conditions has left any significant footprints on the genetic

variation of either species, at least not at the analyzed loci. In France and Germany, the Upper Rhine Valley is defined as a single provenance region for *Q. robur* and for *Q. petraea* (BLE 2013a, b; MAAF 2014). Our results support the treatment of the whole biogeographic sub-region as a single provenance since there is no evidence of a genetic subdivision. More detailed genetic and site condition data might be required to firmly conclude that natural selection has left imprints on the genetic variation of the Upper Rhine oaks. Thus, it should not be ruled out that natural selection has taken place locally. Given the autochthony of most of the oak trees in the region, natural regeneration should be strongly promoted because if local adaptations exist, it ensures that they will be transferred to the next generation. Otherwise, local reproductive material should be selected. The results from this paper, along with the findings from Neophytou and Michiels (2013), indicate that large-scale seed transfer has led to an alteration of the genetic constitution of oak forests in the Upper Rhine. Further guidelines are necessary to control and restrict the transfer of forest reproductive materials between different provenance regions.

Different conclusions can be drawn from our results concerning *Q. pubescens*, a species with a marginal distribution and currently without economic importance in the Upper Rhine Valley, but potentially of interest due to future climate change. Unlike *Q. robur* and *Q. petraea*, it has a high degree of genetic differentiation as each of the study stands have a well-delimited gene pool. Stands of *Q. pubescens* in the region have a high degree of biodiversity, and many of them are within protected areas. Given the high genetic differentiation among populations, special focus should be given to the conservation of its genetic resources.

Acknowledgments The current study has been conducted within the framework of the Interreg-IV project "The regeneration of the oaks in the Upper Rhine lowlands," funded by the European Regional Development Fund (ERDF), the regional government authority of Baden-Württemberg in Freiburg (Regierungspräsidium Freiburg; RPF), the National Office of Forests (Office National des Forêts; ONF) in France, and the Regional Directory of Food, Agriculture and Forestry of Alsace (Direction Régionale de l'Alimentation, de l'Agriculture et de la Forêt d'Alsace; DRAAF). We express our gratitude to all of the people from the ONF, RPF, and the FVA who worked for the project in the field, in the lab, or in the office. We thank three anonymous reviewers for providing valuable comments on this article. We thank Bernhard Thiel for improving the English of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

Data archiving statement Genotypic data used for this study are available at Dryad: 10.5061/dryad.b64b4.

References

Aas G (1989) Untersuchungen zur Trennung und Kreuzbarkeit von Stiel- und Traubeneiche (*Quercus robur* L. und *Quercus petraea* (Matt.) Liebl.). Doctoral Dissertation, Technische Universität München, Munich

Abadie P, Roussel G, Dencausse B, Bonnet C, Bertocchi E, Louvet J-M, Kremer A, Garnier-Géré P (2012) Strength, diversity and plasticity of postmating reproductive barriers between two hybridizing oak species (*Quercus robur* L. and *Quercus petraea* (Matt) Liebl.). J Evol Biol 25:157–173

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26:32–46

Behre K-E (1988) The role of man in European vegetation history. In: Huntley B, Webb T III (eds) Vegetation history. Springer, Netherlands, pp 633–672

Bingham J, Sudarsanam S (2000) Visualizing large hierarchical clusters in hyperbolic space. Bioinformatics 16:660–661

BLE (2013a) Forstliche Herkunftsgebiete (HKG) der Baumarten, die unter das FoVG fallen: *Quercus robur* L. Stieleiche. Available from: <http://fgrdeu.genres.de/index.php?tpl=fv_hkg&id_art=49> (accessed 3.11.14)

BLE (2013b) Forstliche Herkunftsgebiete (HKG) der Baumarten, die unter das FoVG fallen: *Quercus petraea* (Matt.) Liebl. Traubeneiche. Available from: <http://fgrdeu.genres.de/index.php?tpl=fv_hkg&id_art=47> (accessed 3.11.14)

Boeuf R (2014) Les végétations forestières d' Alsace, vol 1. Imprimerie Scheuer, Drulingen

Brewer S, Cheddadi R, de Beaulieu JL, Reille M (2002) The spread of deciduous *Quercus* throughout Europe since the last glacial period. Forest Ecol Manag 156:27–48

Bruschi P, Vendramin GG, Bussotti F, Grossoni P (2000) Morphological and molecular differentiation between *Quercus petraea* (Matt.) Liebl. and *Quercus pubescens* Willd. (Fagaceae) in Northern and Central Italy. Ann Bot 85:325–333

Buschbom J, Yanbaev Y, Degen B (2011) Efficient long-distance gene flow into an isolated relict oak stand. J Hered 102:464–472

Chybicki IJ, Oleksa A, Kowalkowska K, Burczyk J (2012) Genetic evidence of reproductive isolation in a remote enclave of *Quercus pubescens* in the presence of cross-fertile species. Plant Syst Evol 298:1045–1056

Curtu AL, Gailing O, Finkeldey R (2009) Patterns of contemporary hybridization inferred from paternity analysis in a four-oak-species forest. BMC Evol Biol 9:284

Derory J, Scotti-Saintagne C, Bertocchi E, Le Dantec L, Graignic N, Jauffres A, Casasoli M, Chancerel E, Bodenes C, Alberto F, Kremer A (2010) Contrasting relations between diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. Heredity 105:401–411

Diniz-Filho JAF, Soares TN, Lima JS, Dobrovolski R, Landeiro VL, de Campos Telles MP, Rangel TF, Bini LM (2013) Mantel test in population genetics. Genet Mol Biol 36:475–485

Dow B, Ashley M, Howe H (1995) Characterization of highly variable (GA/CT) n microsatellites in the bur oak, *Quercus macrocarpa*. Theor Appl Genet 91:137–141

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11:2571–2581

Dupouey JL, Badeau V (1993) Morphological variability of oaks (*Quercus robur* L, *Quercus petraea* (Matt) Liebl, *Quercus pubescens* Willd) in northeastern France: preliminary results. Ann Sci For 50(Suppl1):35–40

Durand J, Bodénès C, Chancerel E et al (2010) A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. BMC Genomics 11:570

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

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587

Felsenstein J (2002) PHYLIP: phylogeny inference package. Version 3(6):a3

Gailing O, Wachter H, Schmitt H, Curtu A, Finkeldey R (2007) Characterization of different provenances of Slavonian pedunculate oaks (*Quercus robur* L.) in Munsterland (Germany) with chloroplast DNA markers: PCR-RFLPs and chloroplast microsatellites. Allg Forst Jagdztg 178:85–90

Gerber S, Chadœuf J, Gugerli F, Lascoux M, Buiteveld J, Cottrell J, Dounavi A, Fineschi S, Forrest LL, Fogelqvist J, Goicoechea P, Jensen JS, Salvini D, Vendramin GG, Kremer A (2014) High rates of gene flow by pollen and seed in oak populations across Europe. PLoS ONE 9, e85130

Goepp S (2007) Origine, histoire et dynamique des Hautes-Chaumes du massif vosgien. Déterminismes environnementaux et actions de l'Homme. Doctoral Dissertation. Université Luis-Pasteur, Strasbourg

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Grivet D, Sork VL, Westfall RD, Davis FW (2008) Conserving the evolutionary potential of California valley oak (*Quercus lobata* Née): a multivariate genetic approach to conservation planning. Mol Ecol 17:139–156

Herzog S (1996) Genetic inventory of European oak populations: consequences for breeding and gene conservation. Ann Sci For 53:783–793

Herzog S, Krabel D (1999) Genetic structures of a flooded and a non-flooded oak (*Quercus robur*) population from the floodplains of the Rhein river. Ekológia (Bratislava) 18:160–163

Heuertz M, Fineschi S, Anzidei M, Pastorelli R, Salvini D, Paule L, Frascaria-Lacoste N, Hardy OJ, Vekemans X, Vendramin GG (2004) Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. Mol Ecol 13: 3437–3452

Homolka A, Schueler S, Burg K, Fluch S, Kremer A (2013) Insights into drought adaptation of two European oak species revealed by nucleotide diversity of candidate genes. Tree Genet Genom 9:1179–1192

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Kampfer S, Lexer C, Glössl J, Steinkellner H (1998) Characterization of (GA) n microsatellite loci from *Quercus robur*. Hereditas 129:183–186

König AO, Ziegenhagen B, van Dam BC, Csaikl UM, Coart E, Degen B, Burg K, de Vries SMG, Petit RJ (2002) Chloroplast DNA variation of oaks in western Central Europe and genetic consequences of human influences. For Ecol Manag 156:147–166

Krabel D, Liesebach M, Schneck V, Wolf H (2010) Transfer von saat-und pflanzgut innerhalb Europas. Forst und Holz 65:2–7

Kremer A, Petit RJ, Zanetto A, Fougère V, Ducousso A, Wagner D, Chauvin C (1991) Nuclear and organelle gene diversity in *Quercus robur* and *Q. petraea*. In: Müller-Starck G and Ziehe M (eds) Genetic variation in European populations of forest trees. Sauerländer's Verlag, Frankfurt am Main, pp. 125–140

Küster H (1996) Auswirkungen von klimaschwankungen und menschlicher landschaftsnutzung auf die arealverschiebung von pflanzen und die ausbildung mitteleuropäischer wälder. Forstwiss Cent Ver Mit Tharandter Forstl Jahrb 115:301–320

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

Lepais O, Petit RJ, Guichoux E, Lavabre JE, Alberto F, Kremer A, Gerber S (2009) Species relative abundance and direction of introgression in oaks. Mol Ecol 18:2228–2242

Lepais O, Roussel G, Hubert F, Kremer A, Gerber S (2013) Strength and variability of postmating reproductive isolating barriers between four European white oak species. Tree Genet Genomes 9:841–853

MAAF (2014) Liste de régions de provenance des espèces forestières - Liste actualisée en mai 2014. Available from: <http://agriculture.gouv.fr/Fournisseurs-especes-et-provenances-forestieres> (accessed 3.11.14)

Müller B (1999) Variation und Hybridisierung von *Quercus pubescens*. Eidgenössische Technische Hochschule Zürich

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Neophytou C (2014) Bayesian clustering analyses for genetic assignment and study of hybridization in oaks: effects of asymmetric phylogenies and asymmetric sampling schemes. Tree Genet Genomes 10: 273–285

Neophytou C, Michiels HG (2013) Upper Rhine Valley: a migration crossroads of middle European oaks. Forest Ecol Manage 304:89–98

Oberdorfer E (1992) Süddeutsche Pflanzengesellschaften Band IV (Wälder und Gebüsche) Fischer Verlag, Stuttgart

Oksanen J, Guillaume Blanchet F, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Henry H, Stevens H, Wagner H (2013) VEGAN: community ecology package. R package version 2.0-10. http://CRAN.R-project.org/package=vegan

Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol 12: 844–855

Petit RJ, Csaikl UM, Bordács S, Burg K, Coart E, Cottrell J, van Dam B, Deans JD, Dumolin-Lapègue S, Fineschi S (2002) Chloroplast DNA variation in European white oaks: phylogeography and patterns of diversity based on data from over 2600 populations. For Ecol Manag 156:5–26

Petit RJ, Bodénès C, Ducousso A, Roussel G, Kremer A (2004) Hybridization as a mechanism of invasion in oaks. New Phytol 161:151–164

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

Pritchard JK, Pickrell JK, Coop G (2010) The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. Curr Biol 20:R208–R215

R Development Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, <http://www.R-project.org/>

Reif A, Gärtner S (2007) Die natürliche Verjüngung der laubabwerfenden Eichenarten Stieleiche (*Quercus robur* L.) und Traubeneiche (*Quercus petraea* Liebl.)–eine Literaturstudie mit besonderer Berücksichtigung der Waldweide. Waldökologie online 5:79–116

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138

Sebald O, Seybold S, Philippi G, Wörz A, Kleinsteuber A, Gottschlich G, Böhling N, Baumann H (1998) Die Farn-und Blütenpflanzen Baden-Württembergs. Ulmer, Stuttgart

Sork VL, Davis FW, Westfall R, Flint A, Ikegami M, Wang H, Grivet D (2010) Gene movement and genetic association with regional

climate gradients in California valley oak (*Quercus lobata* Née) in the face of climate change. Mol Ecol 19:3806–3823

Steinkellner H, Lexer C, Turetschek E, Glössl J (1997) Conservation of (GA)n microsatellite loci between *Quercus* species. Mol Ecol 6: 1189–1194

Streiff R, Ducousso A, Lexer C, Steinkellner H, Gloessl J, Kremer A (1999) Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. Mol Ecol 8:831–841

Uhl A, Reif A, Gärtner S (2008) Naturverjüngung der Stieleiche (*Quercus robur* L.) im Gebiet der „Trockenaue "am südlichen Oberrhein (Südwestdeutschland). Carolinea 66:15–34

Vähä J-P, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Mol Ecol 15:63–72

Vidalis A, Curtu AL, Finkeldey R (2013) Novel SNP development and analysis at a NADP+-specific IDH enzyme gene in a four species mixed oak forest. Plant Biol 15:126–137

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

Zenni RD, Lamy J-B, Lamarque LJ, Porté AJ (2014) Adaptive evolution and phenotypic plasticity during naturalization and spread of invasive species: implications for tree invasion biology. Biol Invasions 16:635–644