ELSEVIER

Forest Ecology and Management 156 (2002) 75–87

Forest Ecology and Management

www.elsevier.com/locate/foreco

Is there a correlation between chloroplastic and nuclear divergence, or what are the roles of history and selection on genetic diversity in European oaks?

Antoine Kremer[a,*], Jochen Kleinschmit[b], Joan Cottrell[c], Edward P. Cundall[c], John D. Deans[d], Alexis Ducousso[a], Armin O. König[e], Andrew J. Lowe[d], Robert C. Munro[d], Rémy J. Petit[a], B. Richard Stephan[e]

[a]*Institut National de la Recherche Agronomique (INRA), Station de Recherches Forestières, BP 45, F-33611 Gazinet Cedex, France*
[b]*Lower Saxony Forest Research Institute, Escherode, D-3513 Staufenberg, Germany*
[c]*Forestry Commission (FC), Forest Research, Northern Research Station, Roslin, Midlothian, EH25 9SY Scotland, UK*
[d]*Institute of Terrestrial Ecology (ITE), Bush Estate, Penicuik, Midlothian, EH26 0QB Scotland, UK*
[e]*Bundesforschungsanstalt für Forst- und Holzwirtschaft (BFH), Institut Für Forstgenetik und Forstpflanzenzüchtung, Sieker Landstrasse 2, D-22927 Grosshansdorf, Germany*

Abstract

The aim of this work was to investigate whether a correlation exists between maternal lineage, assessed by variation in maternally inherited chloroplast DNA (cpDNA) and nuclear controlled variation (phenotypic traits and gene markers). Variation in cpDNA and nuclear controlled traits (62 phenotypic traits, eight isozyme and 31 RAPD loci) was studied in deciduous oak trees (mostly *Q. petraea*) growing in 16 provenance tests. Results from two nuclear diversity studies were also included. The test for correlation was performed using two methods by: (1) comparing provenance mean values (or allele frequencies) among different lineages using ANOVA, (2) making pairwise comparisons of chloroplastic genetic distances (CGDs) with phenotypic differentiation index (DI) (or nuclear genetic distances (GD)) among all provenances using the Mantel test. Among the 62 phenotypic traits, only seven exhibited significant associations with maternal lineages when tested using ANOVA (six using Mantel test). This number decreased to two once correction for geographic distance was introduced in the calculation for Mantel test. There were stronger correlations between maternal lineage and nuclear gene markers. The existence of cytonuclear disequilibrium was shown by the significant differences in allozyme frequencies between the four maternal lineages (at least one allele for each locus). These associations were confirmed by significant correlations between CGD and GD. Finally associations were also found between levels of diversity in nuclear markers and maternal lineages. These results are discussed in the context of the glacial and postglacial history of oak populations in Europe. The analysis suggests that the processes which led to the current structure of chloroplastic diversity and variation for phenotypic traits can be subdivided into four major phases. (1) During the last long glacial period, deciduous oaks were probably confined to three major refugia which were genetically differentiated for chloroplastic and nuclear genes. (2) At the end of the glacial period, oaks migrated northwards and established a spatial pattern of distribution for chloroplastic genes which remains largely intact to this day. (3) As oaks progressively occupied the mid and northern parts of Europe, pollen flow established communication between stands originating from eastern and western refugia. This resulted in the gradual erosion of the original nuclear differentiation while conserving the chloroplastic variation installed during colonization. (4) Local selection pressures acting on the established populations eventually caused their genetic differentiation gradually to increase with time. New patterns of

* Corresponding author. Tel.: +33-5-57-97-90-74; fax: +33-5-57-97-90-88.
E-mail address: antoine.kremer@pierroton.inra.fr (A. Kremer).

0378-1127/02/$ – see front matter
PII: S0378-1127(01)00635-1

differentiation now exist which are totally different from those in place immediately following colonization. During this time, chloroplastic divergence remained unchanged. These processes led to the loss of any association between chloroplastic divergence and phenotypic traits, although some association with gene markers which are less affected by selection has been retained.

Keywords: *Quercus*; Provenance test; Chloroplast DNA; Genetic distance; Mantel test

1. Introduction

The previous contributions to this special issue have shown the distinct east–west distribution of chloroplast DNA (cpDNA) variation in European oaks. In this paper, we investigate whether genetic differentiation is also retained for complex phenotypic traits and other nuclear markers. In contrast to studies conducted in quantitative genetics with the analysis of quantitative trait loci, our hypothesis of the persistence of a genetic differentiation is not related to a causal relationship between cytoplasmic genes and phenotypic traits. Indeed, there is evidence to suggest that such a causal relationship probably does not exist, despite the complete linkage within the cpDNA molecule due to the absence of recombination. The chloroplast genome comprises only about 100 genes (Olmstead and Palmer, 1994), which is highly unlikely to be sufficient to account for variation observed in such complex characters as height growth, stem form and phenological traits. Instead our focus of interest lies in the study of disequilibrium between the cytoplasmic and nuclear genome which may result from several different evolutionary forces (Asmussen and Arnold, 1991). In contrast to Asmussen and Arnold (1991) who concentrated on the within population variation, the present study examines the correlation between chloroplast haplotype frequencies and nuclear frequencies at the between population level. We aim to test the hypothesis that common refugial origin may result in a correlation between cpDNA variation and nuclear variation.

Recent genetic data support earlier evidence from fossil pollen which indicates that oaks were isolated in separate refugia during the last glacial period (Petit et al., 2002a). The presence of four different chloroplast lineages originating from three major refugial zones indicates that genetic differentiation developed during the long period of genetic isolation in refugia. The last glacial period lasted about 100,000 years (from 115,000 to 15,000 BP) (Birks, 1986). Despite a few short warming periods during the last glacial era, oaks were restricted to the extremities of southern Europe for about 100,000 years in genetically isolated regions (Brewer et al., 2002). Even though oaks are able to maintain large population sizes, if complete separation of populations lasts over such a long period, genetic differentiation due to stochastic effects such as drift is likely to increase divergence between refugia of chloroplastic and nuclear gene pools. Based on these speculations, we may assume that when the warming period started (at 15,000 BP), Europe was recolonized from genetically differentiated gene pools.

How differentiated were the refugial populations before recolonization? This question relates to the degree of genetic isolation between the three different refugia. That the Iberian refugium was genetically isolated from the other two is supported by geography (Brewer et al., 2002). In contrast it would appear that the Italian refugium was less well isolated from the Balkan refugium. First of all, these two regions now share the same chloroplast lineages (Petit et al., 2002b); second, the Adriatic Sea was partially filled (down to the 42.5° latitude); third southern Italy is only 75 km from southern Greece. Therefore, the possibility of pollen flow between the two eastern refugia cannot be ruled out. In this paper, we test whether the nuclear differentiation developed during the period of isolation in the refugia is still detectable in the current populations which are growing today, or if it has been entirely eroded as a result of either pollen flow among separate lineages and/or the action of local selection pressures. Earlier investigations performed to test the same hypothesis suggested only a vague persistence of the maternal lineage on isozyme variation (Le Corre et al., 1997; Kremer et al., 1998). This study was limited to 21 populations of *Quercus petraea*. Here the same methodology is applied to a large set of provenances that were raised in provenance tests in

Great Britain, Germany and France to test for the existence of a correlation between maternal lineage and isozymes as well as phenotypic traits. Important geographic trends of variation have already been reported for various phenotypic traits (Kleinschmit and Svolba, 1996 for growth and stem form; Stephan et al., 1996 for budburst and growth cessation; Ducousso et al., 1996 for bud burst; Deans and Harvey, 1996 for frost hardiness). Furthermore differentiation of phenotypic traits among populations was much higher than for isozymes (Kremer et al., 1997, 1998). We will compare whether these trends of variation are parallel to those observed for cpDNA variation and are related to the maternal origin of the provenances. The test is based on two different statistical methods. First of all ANOVA will be used to evaluate differences of mean provenances values for various phenotypic traits among the different maternal lineages. Second pairwise genetic distances between populations based on maternal lineages will be correlated to differences in phenotypic traits and tested by using multiple Mantel tests (Smouse et al., 1986).

2. Material and methods

2.1. Provenance evaluation for nuclear controlled traits

Two sets of data concerning traits controlled by nuclear genes were available. Firstly, phenotypic traits were assessed in 16 provenance tests established in western Europe since 1950s (Table 1) (Great Britain, Germany and France). All provenance tests

Table 1
List of provenance tests

Code of the test	Name	Species	Latitude	Longitude	N^a	Nb[b]	Nc[c]	Na[d]	Ne[e]	Nm[f]	Traits[g]
T-FR-001	Petite Charnie	*Q. petraea*	48.082	−0.165	94	55	7	21	3	8	BB6, H7, NS9, LR9
T-FR-002	Vierzon	*Q. petraea*	47.275	2.085	94	55	7	21	3	8	BB6
T-FR-003	Vincence	*Q. petraea*	46.950	3.633	94	55	7	21	3	8	BB6, H7
T-FR-004	Sillegny	*Q. petraea*	48.591	6.129	94	55	7	21	3	8	BB6, H7
T-DE-001	Sprakensehl	*Q. petraea*	52.768	10.484	14	4	3	2	1	4	H5, H8, S5, S8, F8
T-DE-002	Walkenried	*Q. petraea*	51.601	10.635	12	5	2	2	0	3	H5, H8, S5, S8, F8
T-DE-003	Bramwald	*Q. petraea, Q. robur*	51.468	9.601	50	10	27	9	0	4	H48, D48, F48, V48, DA50 for *Q. petraea* and *Q. robur* separately
T-DE-101	Wiesentheid	*Q. petraea*	49.817	10.267	32	7	6	8	0	11	H4
T-DE-102	Eitorf	*Q. petraea*	50.767	7.183	26	4	4	7	0	11	H4, H6, H8, S8, F8
T-DE-103	Eppenbrunn	*Q. petraea*	49.100	7.667	28	6	5	6	0	11	H4, H6, H8, S8, F8
T-DE-104	Müncheberg	*Q. petraea*	52.500	14.050	34	8	7	8	0	11	H4, H6, H8, S8, DA8
T-DE-105	Plön	*Q. petraea*	54.100	10.233	32	8	5	7	1	11	H4, H6, H7, H8, S8, F7, BB8, Bb8
T-GB-000	North York Moors 61	*Quercus petraea*	54.260	−0.540	22	9	4	3	2	4	H8, S8, F8
T-GB-001	Arden 4	*Quercus petraea*	52.328	−1.479	23	10	4	4	1	4	H8, S8
T-GB-002	Alice Holt 429	*Quercus petraea*	51.082	−0.555	20	9	4	3	0	4	H8, S8
T-GB-003	Dean 161	*Quercus petraea*	51.800	−2.543	20	8	4	4	0	4	H8, S8

[a] Total number of provenances used for the analysis.
[b] Number of provenances belonging to the B (yellow) lineage.
[c] Number of provenances belonging to the C (red) lineage.
[d] Number of provenances belonging to the A (blue) lineage.
[e] Number of provenances belonging to the E (green) lineage.
[f] Number of provenances comprising different lineages.
[g] BB: bud burst (measured as score); Bb: bud burst (measured in number of days); D: breast height diameter; DA: damages on the tree (due to frost or other causes); F: form of the trunk; H: height; LR: leaf retention in winter; S: survival; V: volume. Numbers after each character indicate the age of the tree when the phenotypic trait was assessed.

comprising enough provenances belonging to at least three among four major maternal lineages were included in the analysis. Data from other additional provenance tests were available but could not be used, because of the unbalanced distribution between the three lineages. The following four categories of traits were measured: survival, growth, stem form, and phenology (Table 1). In total 62 traits were assessed in the two species (*Q. petraea* and *Q. robur*) in all the provenance tests. There was only one test in which both species were present (T-DE-03); this test was subdivided and the calculations made separately for *Q. petraea* and *Q. robur*. Differences between provenance for phenotypic traits were tested separately within each provenance test by using standard analysis of variance. Provenance mean values for each test were then compiled in separate files by each partner of the project prior to further analysis.

The second set of data corresponds to genetic diversity surveys of *Q. petraea* populations across Europe based on isozyme and DNA markers (Table 2). Results of the survey were published in earlier reports (Zanetto and Kremer, 1995; Le Corre et al., 1997, 1998). Frequencies of 56 alleles belonging to eight isozyme loci were available for the 89 populations. Only alleles with a frequency greater than 0.10 but less than 0.90 were included and thus the number of alleles included in the analysis was reduced to 18. Because allele frequencies are bounded to 1 in each population, rare alleles do not provide additional information to frequent alleles. In addition to the frequencies, diversity measures as allelic richness (*A*), observed heterozygosity (Ho), expected heterozygosity (He) and Wright's fixation index (*F*) for each population were also used for comparisons between maternal lineages. Estimation procedures for the diversity measures are given elsewhere (Zanetto and Kremer, 1995).

2.2. Molecular analysis of chloroplast DNA

Leaves or buds were collected from five trees of each provenance that was present in a provenance test. Total genomic DNA was extracted following standard procedures and amplified with four universal primer pairs: *trn*D-*trn*T, *psa*A-*trn*S, *trn*C-*trn*D (Demesure et al., 1995) and T-F (Taberlet et al., 1991). Amplified products were digested with *Taq*I (in combination with *trn*D-*trn*T and *trn*C-*trn*D), *Alu*I (*trn*D-*trn*T and T-F) or *Hin*fIII (*psa*A-*trn*S). The cpDNA haplotypes were scored after separating the different fragments by electrophoresis on 8% polyacrylamide gels and staining with ethidium bromide.

2.3. Data analysis

2.3.1. Difference between maternal lineages

Each provenance was assigned to a given maternal lineage according to its cpDNA haplotype. The maternal lineages (A, B, C, D and E) are those indicated on Figs. 1 and 2 of Petit et al. (2002a). These lineages have different geographic distributions: A (blue) has a widespread distribution, B (yellow) is preferentially spread in the west of Europe, C (red) in the central part and E (green) in the eastern part of Europe. The majority of populations tested contained a single cpDNA haplotype but there were a few cases of populations containing cpDNA haplotypes from more than one lineage. Provenances that comprised haplotypes from different lineages were discarded

Table 2
Description of the gene diversity survey (Quercus petraea)

Gene marker	No. of loci	No. of alleles	N^a	Nb[b]	Nc[c]	Na[d]	Ne[e]	Nm[f]	Traits
Isozymes	8	56	89	35	10	26	5	13	Frequencies of 18 alleles, *A*, He, Ho, *F*
RAPD	31	62	21	8	3	5	2	3	He

[a] Total number of provenances used for the analysis.
[b] Number of provenances belonging to the B (yellow) lineage.
[c] Number of provenances belonging to the C (red) lineage.
[d] Number of provenances belonging A (blue) lineage.
[e] Number of provenances belonging E (green) lineage.
[f] Number of provenances comprising different lineages.

from the analysis and considered as mixed populations (Tables 1 and 2). The following ANOVA model was then used to test for genetic differences among maternal lineages:

$$Y_{ij} = \mu + M_i + P_{ij} \quad (1)$$

where Y_{ij} is the mean value of a provenance for a given phenotypic trait, M_i the fixed effect due to maternal lineage i and P_{ij} the effect due to provenance j within lineage i. The maternal lineage is one of the four lineages (A, B, C and E) that were identified (Petit et al., 2002b).

2.3.2. Differences between gene pools

ANOVA was also used to test for differences between the two supposedly isolated refugial gene pools (Iberian and Balko-Italian) according to the following model:

$$Y_{ij} = \mu + \mathrm{GP}_i + P_{ij} \quad (2)$$

where Y_{ij} is the mean value of a provenance for a given phenotypic trait, GP_i the fixed effect due to refugial gene pool i and P_{ij} the effect due to provenance j within gene pool i. The gene pool is either Iberian (B lineage) or Balko-Italian (A, C and E lineage).

2.3.3. Correlation between chloroplastic divergence and nuclear divergence

A chloroplastic genetic distance (CGD) based on maternal lineages was calculated for all pairs of provenances within a provenance test. CGD was defined as the number of restriction fragment polymorphisms separating the two populations. Similarly for phenotypic traits, a differentiation index (DI) was computed between all pairs of populations as the absolute value of the difference between mean values of phenotypic traits of two provenances. For each provenance test, a square matrix of CGD (Table 3) is constructed and compared with a square matrix of phenotypic distances (DI). The comparison is made after computing the product moment correlation between CGD and DI. Significance of the correlation coefficient was tested with the help of the Mantel test (Mantel, 1967). The Mantel test consists of constructing a null hypothesis (H: the two distances CGD and DI are not correlated) with the help of a Monte Carlo procedure. Cells of CGD are randomly permuted, whereas cells of DI are maintained as such. For each permutation the product moment correlation coefficient between DI and CGD is computed. The procedure is repeated 1000 times, and the actual value is

Table 3
Matrix of CGD among the 13 different haplotypes (the codes of the haplotypes are as in Petit et al. (2002a))

Haplotypes (1)	1	2	4	5	7	10	11	12	13	15	17	18	21
1	0												
2	5	0											
4	13	14	0										
5	12	13	1	0									
7	14	15	2	2	0								
10	7	8	9	9	10	0							
11	8	9	10	10	11	1	0						
12	7	8	9	9	10	1	2	0					
13	8	7	9	8	10	5	6	5	0				
15	8	9	10	10	11	3	4	4	4	0			
17	6	7	8	8	9	1	2	2	4	2	0		
18	9	6	11	11	12	4	5	5	3	3	3	0	
21	7	4	9	9	10	2	3	3	1	1	1	2	0

compared to the distribution corresponding to the null hypothesis.

The correlation between chloroplastic divergence and nuclear divergence may have been blurred by the clear-cut geographic separation between the four lineages (Petit et al., 2002a). Because provenances belonging to different lineages occupy different ecological regions, their differentiation may have been increased due to recent (since the last glaciation) local selection pressures and/or lack of gene flow. Therefore the differentiation that was generated during the glacial period by the isolation between refugia may have been inflated by selection pressures occurring during colonization. Hence geographic distances (GeoD) have to be used for correcting the calculation of the correlation between CGD and DI. This is the reason why the computation of the partial correlation between DI and CGD at constant GeoD r((DI, CGD), GeoD) is proposed as an alternative to test whether the maternal origin of a provenance has still an impact on nuclear controlled traits (Smouse et al., 1986). Interestingly this partial correlation coefficient can be compared to r((DI, GeoD), CGD) (correlation between phenotypic distance and GeoD at constant chloroplastic distance). The comparison between these two coefficients enables a distinction to be drawn between the two possible causes which formed the basis of the differentiation between populations:

- Either r((DI, CGD), GeoD) is significant, suggesting that population differences are due to the maternal origin of the provenances (historical cause).
- Or r((DI, GeoD), CGD) is significant suggesting that population differences are due to recent selection pressures and/or gene flow (geographic cause).

To calculate the two partial correlation coefficients, GeoDs were computed between all pairs of provenances as the Euclidean distance between the Mercator-projected coordinates.

Similar calculations were made for gene markers to test whether there is a cytonuclear disequilibrium between the population level (Asmussen et al., 1987). First of all to assess if differentiation of allele frequencies at nuclear loci was related to differentiation for chloroplast markers, the nuclear genetic distance (GD) was calculated according to Nei (1987) for isozymes (eight loci) and RAPDs (31 loci) and compared to CGD with the help of the Mantel test (Mantel, 1967). In a second step we also compared the DI for diversity statistics (allelic richness, observed and expected heterozygosity, fixation index) and compared it to CGD following the Mantel test procedure.

3. Results

3.1. Differences among maternal lineages

In the ANOVA (Model 1) only seven out of the 62 F-tests that were computed showed a significant effect ($p > 5\%$) of cpDNA lineage on the phenotypic traits assessed in the provenance tests. This number is slightly greater than the null hypothesis expectation considering the high number of tests that were conducted (5% of 62). Among the seven significant associations between maternal lineages and phenotypic traits, five are related to height (H7 in T-FR-01, T-FR-02, T-FR-04, H8 in T-GB-00 and T-GB-01), one to survival (S8 in T-GB-01), and one to leaf retention (LR9 in FR-01) (Table 4). The height growth data were quite consistent across the provenance tests where significant differences among maternal lineages were detected. In general provenances belonging to the blue (A) or yellow (B) lineages grew better than the others.

In contrast to the absence of an association between maternal lineages and the majority of phenotypic traits, there were significant effects of lineages on several nuclear markers. Among the 18 allozyme frequencies that were available, eight showed significant differences among maternal lineages corresponding to seven loci (among the eight that were used for the allozyme diversity study (Table 2)). These differences followed either an increasing or decreasing order in the following sequence of maternal lineage: B (yellow), C (red), A (blue) and E (green). Furthermore, there were significant differences among lineages for gene diversity statistics. In general, the observed and expected heterozygosity values differed markedly among the four lineages (Table 5); heterozygosity values were higher for populations belonging to the B (yellow) lineage, than to the A (blue) or C (red) lineage. The lowest values were observed in the E (green) lineage. Although the differences were not significant, the same trend was observed for RAPD markers. A striking feature of the calculations is the

Table 4
Differences among maternal lineages for phenotypic traits

	Test (trait)						
	T-FR-01 (H7)[a]	T-FR-03 (H7)[a]	T-FR-04 (H7)[a]	T-GB-00 (H8)[a]	T-GB-01 (H8)[a]	T-GB-01 (S8)[b]	T-FR-01 (LR9)[c]
Lineage							
B yellow	122.4	112.3	132.9	161.1	154.1	94.6	4.06
C red	122.2	103.7	124.5	143.5	93.5	76.3	5.38
A blue	127.8	109.2	129.8	180.2	165.5	93.5	3.83
E green	113.3	101.4	119.7	126.0	134.8	94	1.17
F-test (probability)	3.01 (0.04)	4.71 (0.005)	4.8 (0.004)	6.65 (0.025)	18 (0.0008)	10.6 (0.004)	3.61 (0.02)

[a] Height at age 7 and 8 in cm.
[b] Survival at age 8 in percentage.
[c] Leaf retention at age 9 (measured as a score).

negative correlation between heterozygosity and allelic richness. The lineage with highest heterozygosity exhibited the lowest allelic richness (Table 5).

3.2. Differences among gene pools of the refugia

There were only a few significant differences among the two gene pools (Iberian and Balko-Italian) for any trait assessed in the provenance tests according to the *F*-test values of the ANOVAs (Model 2). Among the few characters (seven among 62) that exhibited differences between the maternal lineages, only two (H7 in T-FR-03 and T-FR-04) showed significant differences between the Iberian and the Balko-Italian gene pools. Variation in allozyme frequencies and gene diversity measures were maintained and significant between the two major gene pools. The Iberian gene pool exhibited significantly higher heterozygosities (observed and expected) but was lower in allelic richness than the Balko-Italian gene pool.

3.3. Correlation between chloroplastic divergence and phenotypic and nuclear divergence

The computation of product moment correlation between the DI for phenotypic traits and CGD revealed only six cases of significant relationships based on the Mantel test among the 62 tests that were made (Table 6). These traits which were correlated with CGD were bud burst (BB6 in T-FR-01, T-FR-02, T-FR-03 and T-FR-04) and height growth (H8 in T-GB-00, and T-GB-01). However only two associations remained significant when the partial correlation $r((DI, CGD), GeoD)$ was calculated (BB6 in T-FR-01 and H8

Table 5
Differences among maternal lineages for nuclear diversity statistics

Maternal lineage	Isozymes				RAPD (He)[a]
	A[b]	Ho[c]	He[a]	Fis[d]	
Lineage					
B yellow	3.16	0.356	0.392	0.09	0.238
C red	3.29	0.324	0.377	0.11	0.238
A blue	3.25	0.319	0.364	0.14	0.216
E green	3.33	0.267	0.363	0.21	0.232
F-test (probability)	1.31 (0.27)	10.10 (0.00002)	6.10 (0.001)	4.78 (0.004)	2.59 (0.09)

[a] Expected heterozygosity per population.
[b] Mean number of alleles per population.
[c] Observed heterozygosity per population.
[d] Fixation index of the population.

Table 6
Product moment correlation and partial correlation coefficients between DI of phenotypic traits and CGD or GeoDs

Test/trait[a]	r(DI, CGD)		r(DI, GeoD)		r((DI, CGD), GeoD)		r((DI, GeoD), CGD)	
	r^b	p^c	r^b	p^c	r^b	p^c	r^b	p^c
T-FR-01/BB6	0.20	0.008	0.47	0.002	0.12	0.024	0.45	0.002
T-FR-02/BB6	0.16	0.038	0.49	0.002	0.09	0.129	0.48	0.002
T-FR-03/BB6	0.15	0.018	0.27	0.016	0.11	0.064	0.25	0.016
T-FR-04/BB6	0.19	0.002	0.41	0.002	0.07	0.098	0.37	0.002
T-GB-00/H8	0.44	0.011	0.40	0.120	0.39	0.052	0.34	0.160
T-GB-00/H8	0.50	0.004	0.18	0.110	0.48	0.004	0.09	0.240

[a] BB6: bud burst at age 6, H8: height at age 8.
[b] Value of the correlation coefficient.
[c] Probability for a larger value than r according to the Mantel test.

Table 7
Product moment correlation and partial correlation coefficients between nuclear GDs and CGDs or GeoDs

Gene markers	r(GD, CGD)		r(GD, GeoD)		r((GD, CGD), GeoD)		r((GD, GeoD), CGD)	
	r^a	p^b	r^a	p^b	r^a	p^b	r^a	p^b
Isozymes	0.08	0.020	0.41	0.001	0.02	0.250	0.41	0.010
RAPDs	0.28	0.001	0.32	0.001	0.11	0.036	0.21	0.001

[a] Value of the correlation coefficient.
[b] Probability for a larger value than r according to the Mantel test.

in T-GB-01). There were 13 cases (among the 62) where the DI of phenotypic traits was significantly correlated to the GeoD and all of them remained significant when the partial correlation r((DI, GeoD), CGD) was calculated. Differences between populations were therefore more closely correlated to their GeoD than to their CGD.

In contrast to phenotypic traits, nuclear GDs computed for isozymes were significantly correlated to CGD, and the correlation remained significant for the RAPDs when it was computed as partial correlation (Table 7). Furthermore DI values of diversity parameters (allelic richness, observed and expected heterozygosities) were strongly correlated to CGD, but an important component of the relationship was the correlation between CGD and GeoD (Table 8). When the impact of GeoD is removed by the calculation of partial correlation, only the correlation

Table 8
Product moment correlation and partial correlation coefficients between DI of diversity parameters and CGDs or GeoDs

Gene markers	r(DI, CGD)		r(DI, GeoD)		r((DI, CGD), GeoD)		r((DI, GeoD), CGD)	
	r^a	p^b	r^a	p^b	r^a	p^b	r^a	p^b
Isozymes								
A	0.08	0.026	0.12	0.032	0.07	0.056	0.11	0.043
Ho	0.05	0.046	0.33	0.002	0.01	0.373	0.33	0.002
He	0.10	0.004	0.22	0.002	0.07	0.024	0.21	0.002
F	0.04	0.130	0.20	0.004	0.02	0.320	0.19	0.006
RAPD								
He	0.18	0.101	0.12	0.165	0.15	0.128	0.07	0.279

[a] Value of the correlation coefficient.
[b] Probability for a larger value than r according to the Mantel test.

for DI of He remained significant. Again correlations between DI and GeoD were higher and remained significant when calculated as partial correlation coefficients (Table 8).

4. Discussion

The two methods (ANOVAs and Mantel) that were used to test for associations between cpDNA lineage and nuclear controlled traits led to congruent conclusions. Both revealed that only a few phenotypic traits (seven for the ANOVA and six for the Mantel test among the 62 traits analyzed) exhibited significant associations with cpDNA variations. However, there were some discrepancies between the two methods related to the different types of association that are detected by each method. The Mantel test detects linear relationships between cpDNA variation and nuclear controlled traits, whereas ANOVA does not make any assumptions about the kind of associations that exists. On the other hand, the partial correlation enables a correction to be introduced for the effect of GeoD on the relationship between cpDNA and nuclear controlled traits. After these corrections were made only two phenotypic traits remained associated with cpDNA variation. The results were quite different with nuclear genetic markers where stronger associations were found due to either cytonuclear disequilibria or differences of nuclear genes among different maternal lineages. There were a few methodological constraints imposed by the existing available provenance tests. These included the unbalanced distribution of provenances among different maternal lineages in several tests (Table 1) and the reduced number of species (mostly *Q. petraea*) that were present. These limitations may have hampered the detection of association between chloroplastic and nuclear divergence, by decreasing the power of the statistical tests. However, even if the provenance tests were not originally planned for our purpose, their number and geographic distribution provided a large range of provenance sampling and expression of phenotypic traits. The results therefore provide evidence on which to base firm conclusions particularly in view of the fact that both the range wide provenance tests and the more regional provenance tests produced similar results. Therefore, the data can be reliably used to form genetic and evolutionary based interpretations of the results.

4.1. Lack of a correlation between chloroplastic divergence and population differentiation for phenotypic traits

The extensive computation of data from various provenance tests installed in different countries has indicated that the imprint of the maternal origin on extant populations has in general no significant impact on the phenotypic traits which are of interest in forestry. Neither the ANOVA nor the correlation between chloroplastic and nuclear divergence has revealed any consistent trend of variation in growth, flushing date or form that could be attributed to the origin of the provenances. Various hypotheses can be invoked to explain why there is no significant association between chloroplastic divergence and phenotypic traits.

First, as there is no definite proof, the belief in the existence of an important nuclear genetic differentiation among refugial gene pools at the beginning of recolonization is based on speculations. Are there rationale grounds for such speculations? If the refugial gene pools were not differentiated, then the lack of association that is observed today may merely be due to the maintenance of the homogenization since recolonisation. This seems unlikely as the glacial period lasted over more than 100,000 years. Genetic drift over thousands of generations in the absence of any migration would lead to significant divergence. This trend may have even been reinforced for adaptive traits by selection if important ecological differences existed among refugial zones. As mentioned by Brewer et al. (2002) the refugial stands were located in mountainous areas at mid-altitude sites, which further increased their isolation. Isolation between the Italian and Balkan refugia may have been less complete due to the short distance separating potential refugial areas, and the restricted size of the Adriatic sea. Thus there is both geographic and historic evidence to support the theory of genetic isolation between refugial areas, at least between Iberian and Italian refuges. As a result, at the beginning of the recolonization process, 15,000 years ago, deciduous oaks were probably subdivided into genetically differentiated gene pools corresponding to the refugial zones.

Studies of isozyme gene diversity provide some indirect evidence for the existence of an original nuclear divergence among refugial populations. Allozymes, which are believed to be selectively neutral, have been found in a range wide study to exhibit a clear longitudinal trend of variation across Europe (Zanetto and Kremer, 1995; Kremer and Zanetto, 1997). This suggests that differences in allele frequencies existed between eastern and western populations. So, if genetic differentiation among refugial gene pools existed at the beginning of recolonization, why and how did it subsequently disappear?

There are at least two evolutionary forces that may have erased the original differentiation between refugial gene pools: these are pollen flow and local selection pressures. Pollen can be dispersed over large distances in oaks. Recent experiments using pollen traps have shown that oak pollen can migrate at several kilometers (Lahtinen et al., 1996). More recently these observations have been confirmed by paternity analysis conducted in European (Streiff et al., 1999) and American oaks (Dow and Ashley, 1996). Oaks have occupied their current range for more than 6000 years (Brewer et al., 2002), and have therefore had the opportunity to disseminate their genes by pollen from the eastern part to the western part (and vice versa). This long-term reciprocal movement of pollen will have contributed to the homogenization of the nuclear genetic composition of the oak stands throughout their natural range. As shown by the distribution of allozyme frequencies in a range wide collection of populations of *Q. petraea* (Zanetto and Kremer, 1995), population subdivision contributed 2.5% to the total allozyme diversity indicating that population differentiation—although significant—was extremely low. A geostatistical analysis of the same data set has shown that pollen dispersion with a standard deviation of less than 200 m is sufficient to achieve the observed low level of differentiation (Le Corre et al., 1998). This theoretical value for pollen dispersion is much lower than that inferred from paternity analysis (Streiff et al., 1999). These observations and results suggest that pollen flow is an extremely efficient way to erase population differentiation in oaks.

The second evolutionary force, natural selection, is likely to increase differentiation in a different direction from the original differentiation resulting from isolation during the last glacial period. Its net effect is to lead to the erosion of the initial genetic structure. The current distribution of *Q. robur* and *Q. petraea* is extremely wide both geographically and ecologically. These oak species grow in contrasting soil and climatic conditions, giving many opportunities for selection for diversifying populations. If selection pressures are strong, then differentiation can be achieved in a short number of generations. Lessons from exotic oak species introduced into Europe have shown that only a few turnover of generations are necessary for introduced populations to become differentiated from the original populations (Daubree and Kremer, 1993). Since oaks have occupied their actual range for more than 6000 years sufficient generations have elapsed to shape the distribution of genetic diversity among populations. As shown by results from common garden experiments and provenance tests, extensive variation between populations exists in oak species (Kleinschmit, 1993 for a review). The pattern of distribution of variation is dependent on the phenotypic trait measured. For example, bud burst and bud set exhibit a clinal altitudinal and latitudinal pattern of variation (Stephan et al., 1996; Ducousso et al., 1996), whereas growth traits and form follow a discontinuous pattern of variation (Kleinschmit, 1993). Phenological traits that are under photoperiodic or heat sum control exhibit geographic patterns at large spatial scales, whereas growth and form which are influenced by local soil and competition conditions show patterns at a microscale level. These results are confirmed by our own observations in this study; where there was a significant correlation between GeoD among populations and the DI for phenological traits, but no correlation between GeoD and DI for growth traits and form. Different patterns of geographic variation that are observed in existing provenance tests are indirect indications of recent (postglacial) selection pressures.

To sum up our conclusions for phenotypic traits, the scenario that led to the current structure of chloroplastic diversity and variation for phenotypic traits can be subdivided into four major phases. (1) The glacial period ended with the subdivision of the deciduous oaks in three major refugia that were genetically differentiated for chloroplastic and nuclear genes. At the end of the glacial period, there was probably an association between chloroplastic and nuclear divergence based on refugial origin. (2) As

recolonization begun, oaks migrated northwards and during migration installed their spatial distribution for chloroplastic genes. Because later seed migration between installed oak stands was limited, the chloroplastic spatial structure that was established during recolonization has never been changed (Petit et al., 1997). (3) As oaks occupied progressively the mid and northern part of Europe, pollen flow established communication between stands originating from eastern and western refugia. Pollen migration increased as stands from different origins merged, and reduced the initial nuclear differentiation between the gene pools of the different refugia. Pollen flow did not entirely obliterate genetic differentiation between refugia at the nuclear level as some vestige could still be detected in the selectively neutral isozyme markers. In the meantime, the chloroplastic divergence installed during colonization was conserved due to limited seed flow among established stands. (4) Finally local selection pressures acting on the installed populations contributed to their genetic differentiation as measured by phenotypic traits and constantly increased with time. The current patterns of differentiation for phenotypic traits are thus very different from those which were present immediately after colonization. At the same time, chloroplastic divergence remained still unaffected. As a result, there is no longer any association between chloroplastic divergence and phenotypic traits.

4.2. Maintenance of a chloroplastic–nuclear disequilibrium at the population level

In contrast to the results observed for phenotypic traits, a striking conclusion of our calculation is the existence of a disequilibrium between chloroplastic and nuclear genes, as shown by the correlation between CGD between populations and their GD (Table 7); these correlations were significant for isozymes and RAPDs. However, significance was maintained only for the RAPDs when partial correction taking into account GeoDs were computed. Our interpretation is that the cytonuclear disequilibria that were created during the ice age (see Section 4.1) were only partly erased after recolonization. Because isozymes are less susceptible to selective pressures, only gene flow contributed to the destruction of the cytonuclear association. As already mentioned, even if allozyme frequencies exhibit a low population differentiation across Europe in *Q. petraea* (Zanetto and Kremer, 1995), there is as clear a trend of variation from east to west as there is for cpDNA variation (Petit et al., 2002a). This geographic covariation is responsible for the significant disequilibrium which we observed.

4.3. Correlation between chloroplastic divergence and levels of nuclear diversity

Besides the cytonuclear disequilibrium, there were also significant differences of levels of nuclear diversity among populations belonging to different maternal lineages. Heterozygosity (observed and expected) values were much higher in the B (yellow) lineage, than in the E (green) or A (blue) lineage (Table 4). The higher heterozygosity values were surprisingly associated to lower allelic richness and lower fixation indexes. As shown by the correlation analysis (Table 8), these differences were related to GeoDs rather than to CGDs. These results mean that populations originating from the four main lineages and which are situated close to each other today (for example in Germany or France where several lineages meet) exhibit similar levels of diversity. Only populations belonging to the four lineages that are still separated by long distances maintain these differences. Again gene flow could be proposed to account for these patterns. Because gene flow among populations is more important when they are adjacent to each other, levels of diversity tend to be homogenized even between populations that were originally quite different for their nuclear diversity. However our results suggest that the original refugial gene pools were highly differentiated for their levels of nuclear diversity at the beginning of recolonization. They indicate that the eastern lineages had more alleles, but lower heterozygosities in comparison to the western refugia. These differences may be related to spatial distribution of populations during the glacial period. If refugial populations were split in scattered small stands within, for example the Balkan area, this structure would tend to decrease heterozygosity within populations, but lead to a higher number of alleles over the whole refugial area. On the other hand if refugial populations, for example in the Iberian area, were adjacent or in close contact by pollen, then higher

levels of heterozygosity would be maintained. An alternative hypothesis, would be the occurrence of bottlenecks in eastern populations, that would have contributed to restore more rapidly allelic richness than heterozygosity (Maruyama and Fuerst, 1985). Unfortunately, we lack accurate information concerning the distribution of deciduous oak forest during the glacial period. Hence we can only speculate on the reasons for the differences in levels of nuclear diversity among the four lineages.

Acknowledgements

This study has been carried out with financial support from the Commission of the European Communities, Agriculture and Fisheries (FAIR) specific RTD programme, PL95-0297 ''Synthetic maps of gene diversity and provenance performance for utilization and conservation of oak genetic resources in Europe''. It does not necessarily reflect its views and in no way anticipates the Commission's future policy in this area. We are very grateful to Katrin Groppe, Céline Latouche, Marie Hélène Pémonge, Birgit Ziegenhagen for helpful assistance in the laboratory, and to members of the Forest Research Technical Support Unit (Forestry Commission) in the United Kingdom, to members of Office National des Forêts in France, Edda Burchard, Katrin Groppe, Jean Marc Louvet, Sabine Steinhoff and Josef Svolba for collecting the data in the provenance tests.

References

Asmussen, M.A., Arnold, J., 1991. The effects of admixture and population subdivision on cytonuclear disequilibria. Theoret. Population Biol. 39, 273–300.

Asmussen, M.A., Arnold, J., Avise, J.C., 1987. Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic genotypes. Genetics 115, 755–768.

Birks, H.J.B., 1986. Late-quaternary biotic changes in terrestrial and lacustrine environments, with particular reference to north-west Europe. In: Berglund, E. (Ed.), Handbook of Holocene Palaecology and Palaehydrology. Wiley, New York, pp. 3–65.

Brewer, S., Cheddadi, R., Beaulieu, J.L., Reille, M., Data contributors, 2002. The spread of deciduous *Quercus* throughout Europe since the last glacial period. For. Ecol. Manage. 156, 27–48.

Daubree, J.B., Kremer, A., 1993. Genetic and phenological differentiation between introduced and natural populations of *Quercus rubra* L. Ann. Sci. For. 50 (Suppl. 1), 271–280.

Deans, J.D., Harvey, F.J., 1996. Frost hardiness of provenances of *Quercus petraea* (Matt.) Liebl. In: Kremer, A., Muhs, H. (Eds.), Inter- and Intraspecific Variation in European Oaks: Evolutionary Implications and Practical Consequences. European Commission, EUR16717N, Brussels, pp. 217–238. ISBN 92-827-5993-8.

Demesure, B., Sodzi, N., Petit, R.J., 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol. Ecol. 4, 129–131.

Dow, B.D., Ashley, M.V., 1996. Microsatellite analysis for seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. Mol. Ecol. 5, 615–627.

Ducousso, A., Guyon, J.P., Kremer, A., 1996. Latitudinal and altitudinal variation of bud burst in western populations of sessile oak (*Quercus petraea* (Matt.) Liebl.). Ann. Sci. For. 53, 775–782.

Kleinschmit, J., 1993. Intraspecific variation of growth and adaptive traits in European oak species. Ann. Sci. For. 50 (Suppl. 1), 166s–185s.

Kleinschmit, J., Svolba, J., 1996. Intraspecific variation of growth and stem form in *Quercus robur* and *Quercus petraea*. In: Kremer, A., Muhs, H. (Eds.), Inter- and Intraspecific Variation in European Oaks: Evolutionary Implications and Practical Consequences. European Commission, EUR16717N, Brussels, pp. 217–238. ISBN 92-827-5993-8.

Kremer, A., Zanetto, A., 1997. Geographic structure of gene diversity in *Quercus petraea* (Matt.) Liebl. II. Multilocus patterns of variation. Heredity 78, 476–489.

Kremer, A., Zanetto, A., Ducousso, A., 1997. Multilocus and multitrait measures of differentiation for gene markers and phenotypic traits. Genetics 145, 1229–1241.

Kremer, A., Petit, R.J., Ducousso, A., 1998. Structure of gene diversity, gene flow and gene conservation in *Quercus petraea*. In: Turok, J., Kremer, A., de Vries, S. (Eds.), First European Meeting on Social Hardwoods. IPGRI, Roma, pp. 133–144. ISBN 92-9043-377-9.

Lahtinen, M.J., Pulkinnen, P., Helander, M.L., 1996. Potential gene flow by pollen between English oak (*Quercus robur* L.) stands in Finland. For. Stud. 28, 47–50.

Le Corre, V., Dumolin-Lapègue, S., Kremer, A., 1997. Genetic variation at allozyme and RAPD loci in sessile oak *Quercus petraea* (Matt.) Liebl.: the role of history and geography. Mol. Ecol. 6, 519–529.

Le Corre, V., Roussel, G., Zanetto, A., Kremer, A., 1998. Geographic structure of gene diversity in *Quercus petraea* (Matt.) Liebl. III. Patterns of variation identified by geostatistical analysis. Heredity 80, 464–473.

Mantel, N.A., 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27, 91–99.

Maruyama, T., Fuerst, P.A., 1985. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. Genetics 111, 675–689.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Olmstead, R.G., Palmer, J.F., 1994. Chloroplast DNA systematics: a review of methods and data analysis. Am. J. Bot. 81, 1205–1224.

Petit, R.J., Pineau, E., Demesure, B., Bacilieri, R., Ducousso, A., Kremer, A., 1997. Chloroplast DNA footprints of postglacial recolonisation by oaks. Proc. Natl. Acad. Sci. USA 94, 9996–10001.

Petit, R.J., Brewer, S., Bordács, S., Burg, K., Cheddadi, R., Coart, E., Cottrell, J., Csaikl, U.M., van Dam, B.C., Deans, J.D., Fineschi, S., Finkeldey, R., Glaz, I., Goicoechea, P.G., Jensen, J.S., König, A.O., Lowe, A.J., Madsen, S.F., Mátyás, G., Munro, R.C., Popescu, F., Slade, D., Tabbener, H., de Vries, S.M.G., Ziegenhagen, B., de Beaulieu, J.-L., Kremer, A., 2002a. Identification of refugia and postglacial colonization routes of European white oaks based on chloroplast DNA and fossil pollen evidence. For. Ecol. Manage. 156, 49–74.

Petit, R.J., Csaikl, U.M., Bordács, S., Burg, K., Coart, E., Cottrell, J., van Dam, B.C., Deans, J.D., Dumolin-Lapègue, S., Fineschi, S., Finkelday, R., Gillies, A., Glaz, I., Goicoechea, P.G., Jensen, J.S., König, A.O., Lowe, A.J., Madsen, S.F., Mátyás, G., Munro, R.C., Pemonge, M.-H., Popescu, F., Slade, D., Tabbener, H., Taurchini, D., de Vries, S.M.G., Ziegenhagen, B., Kremer, A., 2002b. Chloroplast DNA variation in European white oaks: phylogeography and patterns of diversity based on data from over 2,600 populations. For. Ecol. Manage. 156, 5–26.

Smouse, P.E., Long, J.C., Sokal, R., 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Syst. Zool. 35, 627–632.

Stephan, B.R., Venne, H., Liepe, K., 1996. Intraspecific variation of *Quercus petraea* in relation to bud burst and growth cessation. In: Kremer, A., Muhs, H.J. (Eds.), Inter- and Intraspecific Variation in European Oaks: Evolutionary Implications and Practical Consequences. European Commission, EUR16717N, Brussels, pp. 165–184. ISBN 92-827-5993-8.

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.

Taberlet, P., Gielly, L., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109.

Zanetto, A., Kremer, A., 1995. Geographic structure of gene diversity in *Quercus petraea* (Matt.) Liebl. I. Monolocus patterns of variation. Heredity 75, 506–507.