Copyright © 1997 by the Genetics Society of America

Phylogeographic Structure of White Oaks Throughout the European Continent

S. Dumolin-Lapègue,* B. Demesure,*[,1] S. Fineschi,†[,2] V. Le Corre* and R. J. Petit*

*Institut National de la Recherche Agronomique, Laboratoire de Génétique et d'Amélioration des Arbres Forestiers, F-33611 Gazinet Cedex, France and †Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft, CH-8903 Birmensdorf, Switzerland

Manuscript received December 26, 1996
Accepted for publication May 6, 1997

ABSTRACT

Patterns of chloroplast DNA (cpDNA) variation were studied in eight white oak species by sampling 345 populations throughout Europe. The detection of polymorphisms by restriction analysis of PCR-amplified cpDNA fragments allowed the identification of 23 haplotypes that were phylogenetically ordered. A systematic hybridization and introgression between the eight species studied is evident. The levels of subdivision for unordered (G_{ST}) and ordered (N_{ST}) alleles are very high and close (0.83 and 0.85). A new statistical approach to the quantitative study of phylogeography is presented, which relies on the coefficients of differentiation G_{ST} and N_{ST} and the Mantel's test. Based on pairwise comparisons between populations, the significance of the difference between both coefficients is evaluated at a global and a local scale. The mapped distribution of the haplotypes indicates the probable routes of postglacial recolonization followed by oak populations that had persisted in southern refugia, especially in the Iberian peninsula, Italy and the Balkans. Most cpDNA polymorphisms appear to be anterior to the beginning of the last recolonization. A subset of the preexisting haplotypes have merely expanded north, while others were left behind in the south.

IN the last decade, intraspecific phylogeography, the study of the relationship between the phylogeny of variants and their geographic distribution (Avise *et al.* 1987), has become a popular field in evolutionary science (see reviews in Hewitt 1996; Soltis *et al.* 1997). Such studies are attractive because historical insights can be obtained by identifying the genealogy of variants, which are then overlaid upon a geographic frame, reflecting the spatiotemporal dynamics of the organism studied.

Among the historical factors likely to have played a role in determining the present pattern of genetic variation observed in natural populations of the temperate zone, a major one appears to be the last glacial period and the subsequent recolonization when animal and plant species expanded in the lands newly released by ice. Such migrations of taxa have occurred repeatedly during the Quaternary and are considered to be a response to orbitally forced climate changes at time scales of 10^5 years (*e.g.*, Huntley and Webb 1989). In Europe, numerous species were confined to southern refugia during the last glaciation, which started about 115,000 before present (BP) and reached its maximum 18,000 BP. This was followed by recolonization during the present interglacial that started ~13,000 BP.

Corresponding author: R. J. Petit, Institut National de la Recherche Agronomique, Laboratoire de Génétique et d'Amélioration des Arbres Forestiers, B.P. 45, F-33611 Gazinet Cedex, France.
E-mail: remy@pierroton.inra.fr

[1] *Present address:* Conservatoire des Ressources Génétiques Forestières, Office National des Forêts, Ardon F-45160 Olivet, France

[2] *Present address:* Istituto per l'agroselvicoltura, Via Marconi 2, I-05010, Italy.

Deciduous species of oaks recolonized most of Europe early, from refugia located in southern Iberia, Italy and the Balkans (Huntley and Birks 1983). Deciduous oaks are major constituents of the mixed temperate forests (Huntley 1990). Forest tree species have responded individualistically to climatic changes, so that the results obtained concerning the postglacial routes of one tree species may not be extrapolated to other tree species (*e.g.*, Davis 1976; Huntley and Birks 1983). But oak forests are housing a particularly rich assemblage of plant and animal species (*e.g.*, Speight and Wainhouse 1989), some of which are narrowly dependent on oaks. These species may have followed the same pattern of recolonization than the oaks. Hence, the precise knowledge of the phylogeography of oaks at a continental scale may be important beyond the study of the oaks themselves.

Because chloroplast DNA (cpDNA) is maternally inherited in oaks (Dumolin *et al.* 1995), it allows a direct study of seed-mediated dispersal and gene flow and is consequently particularly useful to infer recolonization routes. As a consequence of the reduced gene flow for this marker, cpDNA polymorphisms will be much more structured than nuclear polymorphisms (Petit *et al.* 1993a,b). Another consequence of uniparental inheritance is a clonal mode of evolution, well-suited for reconstructing phylogenies. Contrary to animal mitochondrial DNA (mtDNA), cpDNA evolves very slowly (four times slower than the plant nuclear genome; Wolfe *et al.* 1987), a trend that is even more pronounced in long-lived tree species such as Fagaceae (Frascaria *et al.* 1993). Nevertheless, variation in angio-

sperms has now been demonstrated in many species, including forest trees, as a consequence of the development of improved molecular techniques and of the use of larger sample sizes (SOLTIS *et al.* 1992; DEMESURE *et al.* 1995).

In oaks, rangewide studies of cpDNA variability have revealed a strong pattern of genetic structure in Europe, which was interpreted as reflecting largely the routes of postglacial recolonization (FERRIS *et al.* 1993; PETIT *et al.* 1993a). Furthermore, extensive cytoplasmic exchanges between sympatric species of white oaks were observed both in eastern United States (WHITTEMORE and SCHAAL 1991) and in Europe (FERRIS *et al.* 1993; PETIT *et al.* 1993a). However, these studies were based on a limited number of polymorphisms. Here, phylogenetic analyses based on numerous cpDNA mutations enable us to introduce a temporal dimension in the description of the genetic structure. In addition, an attempt is made to test the continental phylogeographic pattern, based on an extensive sampling of populations.

Despite the multiplication of phylogeographic studies, evidence that there is an association between the geographic location of the haplotypes and their position within a phylogenetic tree remains largely limited to the casual observation of variants on geographic maps. Even if the existence of a genetic structure is tested, this does not prove that related haplotypes have more similar distribution than what could have been expected by chance. TEMPLETON (1993) and TEMPLETON *et al.* (1995) have developed a nested cladistic analysis of geographic distances to test the null hypothesis of no geographic association among haplotypes. We will follow a different approach first described by PONS and PETIT (1996) and based on the comparison of the coefficient of genetic differentiation for ordered (N_{ST}) and unordered (G_{ST}) alleles. Indeed, the existence of a phylogeographic structure should yield higher N_{ST} than G_{ST} estimates for the same data set. However, in the method of PONS and PETIT (1996), no use is made of the information on the geographic distance between the populations. We therefore extend the test of PONS and PETIT (1996) to take into account this information by using Mantel's tests procedures (MANTEL 1967).

MATERIALS AND METHODS

Plant material: A total of 345 oak populations (1412 individual trees) were sampled throughout Europe. Oaks are members of the family Fagaceae, and the species studied belong to the subgenus Lepidobalanus (CAMUS 1936–1939). Populations belonging to the following eight species were included: *Quercus robur* L. (133 populations), *Q. petraea* (Matt.) Liebl. (143 populations), *Q. pubescens* Willd. (76 populations), *Q. pyrenaica* Willd. (seven populations), *Q. faginea* Lam. (three populations), *Q. frainetto* Ten. (14 populations), *Q. macranthera* Fisch. and Mey (two populations) and *Q. canariensis* Willd. (one population) (since several species can be represented in a population, the total exceeds 345). The oak material (acorns, buds or leaves) was collected in forests or in provenance tests. Particular care was taken to collect material from old trees of ancient forests to reduce the risk of sampling introduced populations. The cumulative range of the eight species is illustrated in Figure 1 together with the location of the populations analyzed. The list of the populations is available upon request. Since we were initially mostly interested in *Q. robur* and *Q. petraea*, the regions where these species are rare or absent were not sampled thoroughly (Greece, southern Spain, North Africa, Turkey). On the other hand, some parts of France were more extensively sampled, to check if rare representatives of the different cpDNA lineages could be detected in a given area by increasing the sample size.

PCR-RFLP procedure: Total DNA was extracted following the procedure described in DUMOLIN *et al.* (1995). This DNA was used as a template in PCR reactions involving a set of conserved primers homologous to the most conserved coding regions of cpDNA and that allow the amplification of the more variable noncoding regions. The pairs of primers are listed in Table 1 and their sequences can be found in TABERLET *et al.* (1991), DEMESURE *et al.* (1995) and DUMOLIN-LAPÈGUE *et al.* (1997). For this study, four pairs of primers (in bold characters in Table 1) were used to characterize all oak populations. These pairs of primers were chosen as they exhibited numerous polymorphisms in a first screening of 30 populations (DEMESURE 1996). The amplifications were performed as detailed in DEMESURE *et al.* (1995). The haplotypes identified were further characterized with the eight other pairs of primers in order to study their relationships. All 12 PCR products (5 μl) were digested with one or two four-base recognition restriction endonucleases (5 units): *Cfo*I, *Dpn*II, *Hin*fI, *Msp*I and *Taq*I, representing 14 PCR-fragment/enzyme combinations. In most cases, PCR fragments were analyzed with a single enzyme (usually *Hin*fI); this way, there was no risk to count the same mutation twice, especially the insertion/deletions (indels). The DNA restriction fragments were separated by electrophoresis on 8% polyacrylamide gels as described in DUMOLIN *et al.* (1995).

Phylogenetic analysis: The haplotypes were phylogenetically ordered using the Wagner parsimony method, which minimizes the total number of character state changes in the tree. This was performed with PAUP (phylogenetic analysis using parsimony, version 3.1, SWOFFORD 1993). For this purpose the data were scored as multistates characters: each polymorphic restriction fragment is a character and the states are the different sizes of this fragment. The length variants were noted from 1 to 6, 9 being reserved to restriction site mutations (Appendix A). The numbers increase from the highest to the lowest molecular weight fragments to facilitate the notation but this does not imply any mutational sequence. After a heuristic search we obtained the equally most parsimonious trees. The Consistency Index (*CI*), which measures the overall amount of homoplasy, was then calculated. Robustness of the resulting trees was tested by bootstrap analysis (FELSENSTEIN 1985) of 100 replicates with PAUP. In addition, the Neighbor-Joining tree was constructed with the same data matrix using PHYLIP 3.5 (FELSENSTEIN 1993). Bootstrap analysis of 2000 replicates was performed.

Genetic diversity analysis: The frequencies of the haplotypes were used to estimate the mean within-population gene diversity (h_S), the total gene diversity (h_T) and the coefficient of gene differentiation over all populations (G_{ST}) and their standard deviations following PONS and PETIT (1995). Then, the level of subdivision of chloroplast diversity for ordered alleles (N_{ST}) modified from LYNCH and CREASE (1990) was computed following PONS and PETIT (1996), the distance between two haplotypes being the number of different restriction fragments. N_{ST} was compared to G_{ST} using the *U*-statistics, which takes into account the covariance between N_{ST} and G_{ST},

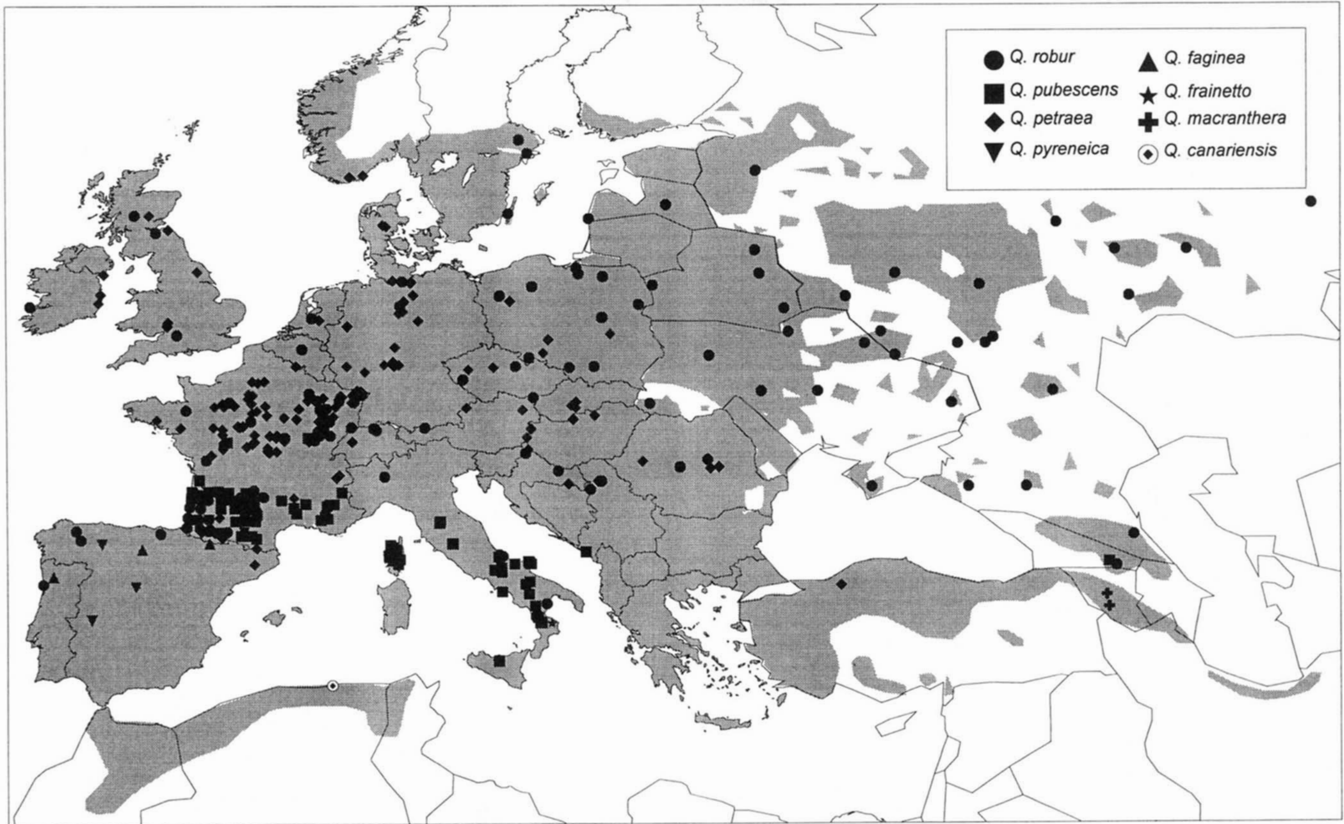

FIGURE 1.—Cumulative range of the eight species of Quercus studied according to SCHMUCKER (1942), JALAS and SUOMINEN (1976), and BROWICZ (1982, 1988) and location of the 345 populations analyzed.

and a one-sided test (see PONS and PETIT 1996). Only the populations having at least three individuals were taken into account. A subset of 283 populations and 1314 individuals (an average of 4.64 individuals per population) were therefore included in this analysis and the subsequent ones.

Mantel tests: The standardized Mantel coefficient, which is the product-moment correlation between elements of two dissimilarity matrices, derived from the Z-statistics of MANTEL (1967), was computed. This procedure is used to estimate the association between two independent dissimilarity matrices and to test whether the association is stronger than would be expected by chance. A matrix of geographic distances between the 283 populations was derived from the longitudinal and latitudinal coordinates. For this purpose, simple trigonometric rules were applied, by approximating the earth to a sphere of known diameter. The precision was sufficient as judged by a comparison with values obtained with a Geographic Information System (MapInfo Professional version 4.02) that was also used to draw the maps. In addition, two genetic distance matrices were derived, by computing the coefficient of gene differentiation between all the pairs of populations for ordered (N_{ST}) and unordered (G_{ST}) haplotypes. The estimators of the diversities are those described in PONS and PETIT (1996) for random populations. A third genetic matrix was built by computing the difference between the two parameters of differentiation for all pairs of populations ($\Delta N = N_{ST} - G_{ST}$). The goal was to test whether N_{ST} increased more quickly than G_{ST} with the geographic distance between the two populations. The standardized Mantel correlation coefficients between the matrix of geographic distance and each of the three genetic distance matrices were then calculated. The significance of the test was evaluated by the construction of a null distribution by a Monte-Carlo procedure: 1000 permutations of rows and columns of the geographic distances matrix were realized whereas the genetic distance matrices were kept constant. The parameter p is the number of permutations that gave a higher Mantel coefficient plus one, divided by the total number of permutations.

Description of the spatial structure: After the global test realized with the Mantel procedure, local tests were performed to describe more precisely the spatial structure. Therefore, a total of 83 classes of 50-km intervals were defined up to the interval 4100–4150 km. The same three parameters (G_{ST}, N_{ST} and ΔN) were computed for each distance class from the average of the within-population gene diversities and the total gene diversities. For each class, the observed value of ΔN was compared with its value obtained through a Monte-Carlo procedure as described above.

RESULTS

cpDNA polymorphisms: First, four fragments were amplified with the pairs of primers DT, AS, TF and CD. After digestion with, respectively, *Taq*I, *Hin*fI, *Cfo*I and *Hin*fI, and *Taq*I, 16 restriction fragments were polymorphic and allowed the detection of 25 length variants and four point mutations. All the 1412 individuals from the 345 populations were studied with the four pairs of primers. With these 29 polymorphisms, it was possible to distinguish 23 haplotypes. The distribution of the haplotypes is given in Figure 2 where all the haplotypes are mapped. A subset of 23 individuals representing the complete collection of haplotypes were further studied

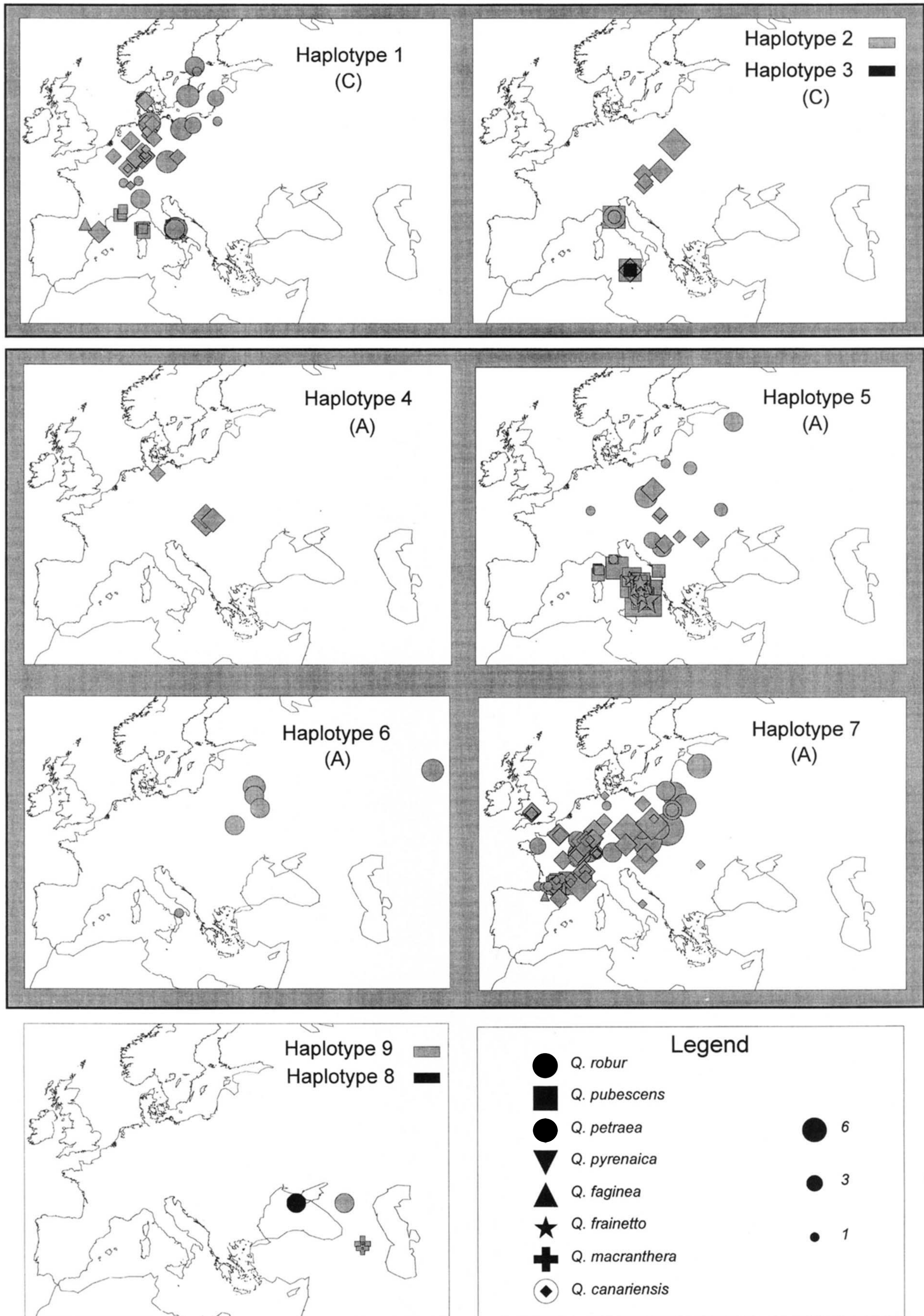

FIGURE 2.—Map of the 23 haplotypes found in the study. The symbols represent the populations and are proportional to the number of individuals analyzed. Maps are grouped by lineage (A, B and C) according to Figure 3.

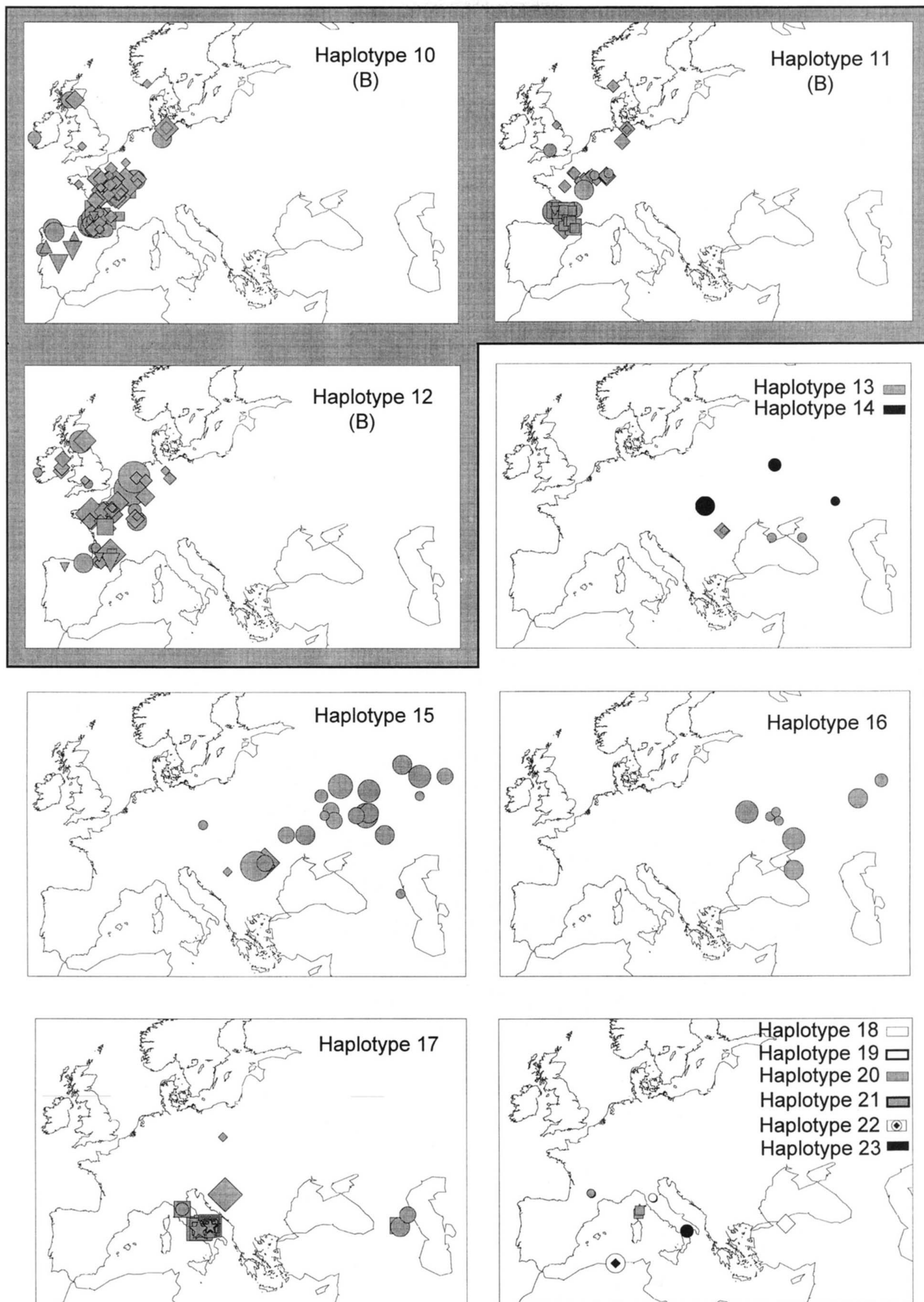

FIGURE 2.— *Continued*

TABLE 1

List of the pairs of primers used in this study

Primer 1	Primer 2	Abbreviation[a]	Reference[b]
trnH [tRNA-His (GUG)]	*trn*K [tRNA-Lys (UUU) exon 1]	HK	1
*trn*K [tRNA-Lys (UUU) exon 1]	*trn*K [tRNA-Lys (UUU) exon 2]	K1K2	1
*trn*Q [tRNA-Gln (UUG)]	*trn*R [tRNA-Arg (UCU)]	QR	2
*trn*C [tRNA-Cys (GCA)]	*trn*D [tRNA-Asp (GUC)]	**CD**	1
*trn*D [tRNA-Asp (GUC)]	*trn*T [tRNA-Thr (GGU)]	**DT**	1
*trn*T [tRNA-Thr (GGU)]	*psb*C [PSII 44 KD protein]	TC	2
*psb*C [PSII 44 KD protein]	*trn*S [tRNA-Ser (UGA)]	CS	1
*trn*S [tRNA-Ser (UGA)]	*trn*fM [tRNA-fMet (CAU)]	SfM	1
*psa*A [PS I (P 700 apoprotein A1)]	*trn*S [tRNA-Ser (GGA)]	**AS**	1
*trn*S [tRNA-Ser (GGA)]	*trnT* [tRNA-Thr (UGU)]	ST	1
*trn*T [tRNA-Thr (UGU)]	*trnF* [tRNA-Phe (GAA)]	**TF**	3
*trn*V [tRNA-Val (UAC) 3′ exon]	*rbc*L [RuBisCO large subunit]	VL	2

[a] The pairs of primers used for the whole set of populations are in bold characters.
[b] 1, Demesure *et al.* (1995); 2, Dumolin-Lapègue *et al.* (1997); 3, Taberlet *et al.* (1991).

with the eight other pairs of primers (Table 1). After digestion with *Hin*fI, *Taq*I, *Msp*I or *Dpn*II, 14 restriction fragments were polymorphic and allowed the detection of an additional 15 length variants and three point mutations.

Phylogenetic relationships between the haplotypes: Among the 23 haplotypes, 47 polymorphisms were detected in 30 restriction fragments, since many of these fragments had several length variants (see the matrix in the Appendix A). There were 19 phylogenetically informative polymorphisms (synapomorphies), shared by at least two individuals, including 12 length variants and the seven point mutations. A Wagner parsimony analysis was made with the complete data set and more than 100 most parsimonious trees were obtained after a heuristic search. Their *CI* value was 0.795 when excluding the autapomorphies. The Neighbor-Joining tree is represented in Figure 3. The length of the branches are correlated to the distances between the haplotypes. The values of bootstrap higher than 50% are added on the nodes and three main lineages A, B and C are defined according to these values. The Neighbor-Joining tree exhibited branching patterns very similar to the ones obtained with the Wagner parsimony analysis. The differences between the numerous equally parsimonious trees obtained with this latter analysis are located at the end of the branches, but the bootstrap analysis supports the definition of the main lineages in the same way. Haplotypes 8 and 9 were not grouped in a specific lineage as they were represented by too few populations. One haplotype (number 23), found only in two trees of one *Q. robur* population in southern Italy, was very divergent.

The mutations detected in this study are mainly short length variants, presumably indels. Although indels are sometimes site dependent and thus can contribute to homoplasy in evolutionary studies (Downie and Palmer 1992; Clegg *et al.* 1994), mutations longer than two bases that do not belong to tandem repetitions proved to be good phylogenetic markers (Gielly and Taberlet 1994). This is only true when the level of divergence is sufficiently low, as in intraspecific phylogenies. As a matter of fact, the two types of analyses performed here (phenetic and cladistic) were congruent in determining lineages.

Distribution of the haplotypes across species: In this study, the systematic local sharing of cytoplasms observed by Petit *et al.* (1993a) for *Q. robur, Q. petraea, Q. pubescens* and *Q. pyrenaica* is confirmed and can be extended to *Q. faginea, Q. frainetto* and *Q. macranthera.* Indeed, Table 2 shows that different species share the same haplotype and that in most species different haplotypes can be found. More samples are still necessary in the case of *Q. canariensis,* collected in only one population (in Algeria) and represented by a new, though not particularly divergent, haplotype (22).

Geographic distribution of the haplotypes: A first glance at the distribution maps of the haplotypes (Figure 2) makes it clear that the haplotypes are not geographically randomly distributed: each haplotype is circumscribed to a particular area. Moreover, haplotypes belonging to the same lineage are often located in the same part of the range. This is obvious for haplotypes 10, 11 and 12 (lineage B), which are restricted to the western part of Europe. This is also true for haplotypes of lineage C, which are distributed on a relatively narrow part of central Europe, as well as for the haplotypes of lineage A, which are found in a wider area but are also absent or rare at the extreme east and west of Europe. Finally, the other relatively frequent haplotypes are almost all located in the eastern part of Europe and do not belong to lineages A, B or C. However, haplotype 17, characterized by a rather central position in the phylogenetic tree (Figure 3), is found in Italy, the Balkans and the Caucasus. Note that nearly all haplotypes that were detected only once

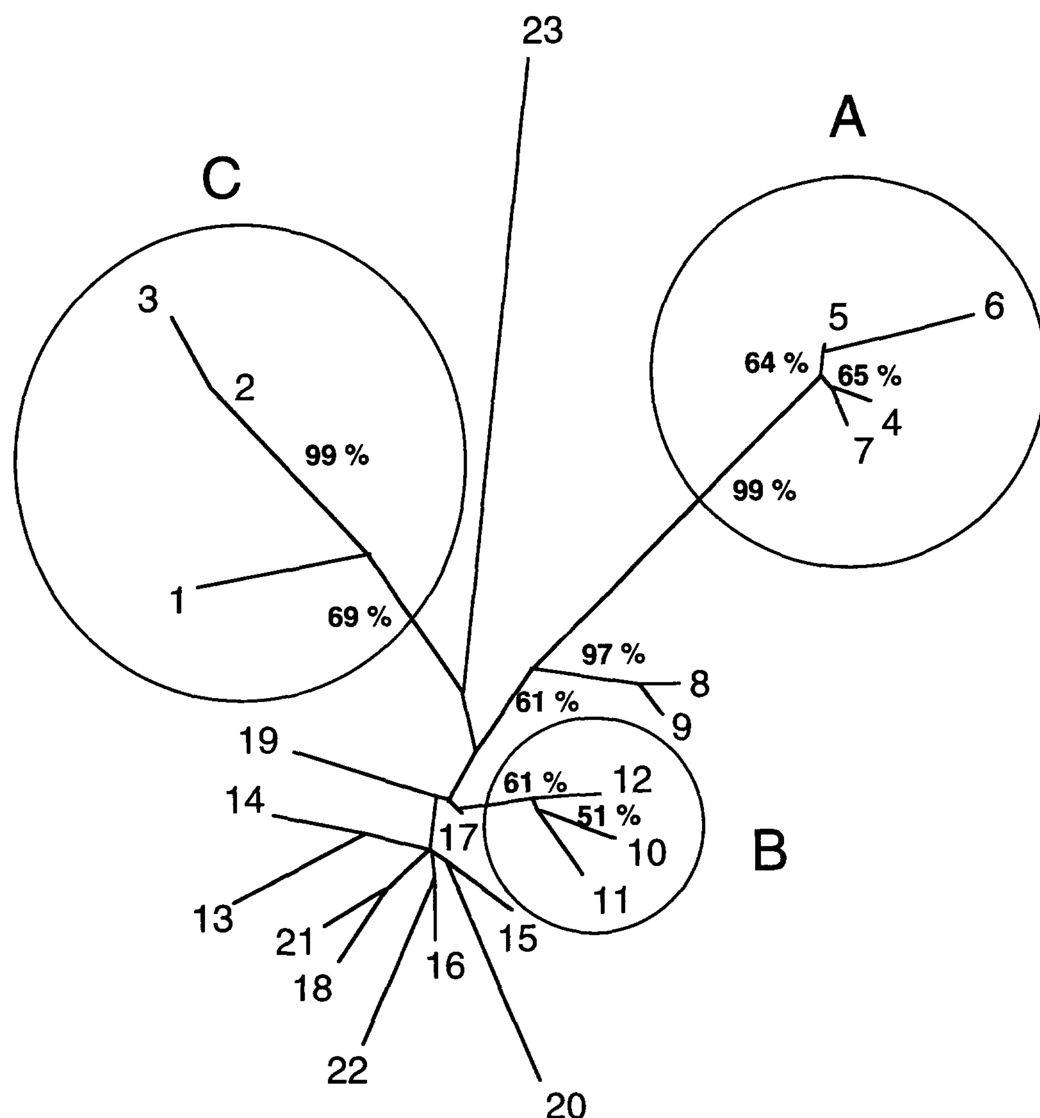

FIGURE 3.—Neighbor-joining tree of cpDNA haplotypes. The bootstrap values higher than 50% are represented. Three lineages A, B and C are identified according to this criteria. Haplotypes are numbered as in Appendix A.

(haplotypes 3, 8, 18 19, 21, 22, 23) are located in the southern part of Europe.

Coefficients of differentiation and phylogeographic structure: The levels of subdivision of the chloroplast diversity estimated from the 283 populations, taking into account or not the relationships between haplotypes, are very high: respectively, 85% (N_{ST}) or 83% (G_{ST}) of the diversity is distributed among populations. These two parameters are significantly different ($U = 1.90$, $P = 0.029$). A Mantel procedure was used to test

TABLE 2

Number of populations per species that contain each haplotype

	Haplotype																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
Q. robur	16	3			13	6	36	1	1	31	12	15	2	3	20	8	3	1			1		1	173
Q. pubescens	9	3	1		19		17			8	11	2					14		1	2				87
Q. petraea	21	9		4	6		38			40	18	28	2		5		2							173
Q. pyrenaica										3	1	3												7
Q. faginea	1						1			2														4
Q. frainetto	1				6												6							13
Q. macranthera									2															2
Q. canariensis																						1		1
Total	48	15	1	4	44	6	92	1	3	84	42	48	4	3	25	8	25	1	1	2	1	1	1	460

TABLE 3

Results of the analysis of diversity

	Value[a]	Mantel coefficient[b]
Number of populations ≥ 3 individuals	283	
Arithmetic mean	4.64	
Harmonic mean	4.08	
h_S	0.151 (0.016)	
h_T	0.874 (0.008)	
G_{ST}	0.828 (0.018)	0.244[c]
v_S	0.992 (0.132)	
v_T	6.610 (0.170)	
N_{ST}	0.850 (0.020)	0.270[c]
ΔN	0.022	0.110[c]

[a] Standard deviations are in parentheses.
[b] Comparison with the geographical distance.
[c] The P value is inferior to 0.001.

the overall existence of a geographic structure and to assess the significance of the phylogeographic component of this geographic structure. The results show that both G_{ST} and N_{ST} are highly correlated with the geographic distance, with correlations of 0.24 and 0.27, respectively, corresponding to $P < 0.001$ (Table 3). Moreover, their difference ΔN, which measures the contribution of the phylogenetic relationship among haplotypes to the overall coefficient of gene differentiation for ordered alleles (N_{ST}), is also highly significantly correlated with the geographic distance ($r = 0.11$, $P < 0.001$). After this global test, more detailed analyses were realized. The same computations were performed but within distance classes of 50 km. Figure 4a represents the two parameters G_{ST} and N_{ST} according to the distance separating the populations. As expected for unbiased estimates, the weighted mean over all distance classes are equal to the same parameters computed directly, if the mean of these coefficients of differentiation are obtained from the ratio of the means and not the mean of the ratios. We tested the hypothesis of equality of the parameters in each class with 1000 permutations and the results are represented in Figure 4b. In 34 of 83 cases, the test was significant. In particular, for the five lowest 50-km distance classes, ΔN was significantly lower than the mean permutation value.

DISCUSSION

Migration and genetic structure: As stressed in earlier studies (Petit *et al.* 1993a,b), the level of subdivision for cpDNA polymorphisms in oaks is much higher than for nuclear polymorphisms: the mean coefficient of genetic differentiation for 11 isozyme loci is 0.03 in *Q. petraea* (Zanetto and Kremer 1995), as compared to 0.83 for cpDNA in the present study. In plants, such a contrast between the level of structuration for nuclear and maternally inherited genes has now been shown to be the rule rather than the exception (*e.g.*, Ennos 1994; McCauley 1995; Demesure *et al.* 1996; El Mousadik and Petit 1996). Nevertheless, the difference in the case of oaks is particularly striking. Several factors contribute to increase genetic structure for organelle genes in comparison to nuclear genes. (1) Effective gene flow will be limited to seeds for maternally inherited genomes (Petit *et al.* 1993b). (2) Drift will be twice as strong for a haploid genome as compared to a diploid one. (3) In hermaphrodite species such as the oaks, the flowering and fruiting pattern result in an effective number of trees contributing to the next generation as females that is much reduced as compared to the effective number of trees acting as males (Demesure *et al.* 1996; Dow and Ashley 1996).

Another factor may explain the low level of intrapopulation diversity. Indeed, simulations have demonstrated that the patchy distribution of the haplotypes that is found at a *regional* scale (Ferris *et al.* 1995; Petit *et al.* 1996; Petit *et al.* 1997) can be generated by incorporating rare long-distance dispersal events and subsequent foundations in the simulations (Ibrahim *et al.* 1996; Le Corre *et al.* 1997). Actually, the rapidity of postglacial recolonization of oaks inferred from palynological studies is more concordant with such a model. Successive bottlenecks would therefore have occurred at the head of the migration front. This "leading edge" hypothesis has been advocated in the case of several plant species from the Pacific Northwest (Cwynar and MacDonald 1987; Soltis *et al.* 1997). In oaks, it could account for the difference of haplotype diversity between the north and the south of Europe: successive bottlenecks may have reduced the diversity during the northward expansion while some haplotypes were left behind in the refugia.

This genetic structure appears to be independent of the species, since in populations where several species are represented the same haplotype is usually shared. Moreover, the more frequent haplotypes are always shared by several species. These results confirm and generalize those of the previous studies of cpDNA diversity in this group of oaks. Systematic hybridization and introgression between these species is therefore the rule. It would be interesting to delimit exactly the extent of this complex of hybridizing oak species based on cpDNA evidence. Preliminary results indicate that *Q. afares* Pomel, *Q. cerris* L., *Q. ilex* L. and *Q. suber* L. do not belong to the complex studied here, as judged from their different cpDNA restriction patterns (results not shown).

Test of the phylogeographic structure: As can be seen in Figure 2, haplotypes belonging to the same lineage often occupy a similar range. However, it would be useful to quantify and test the phylogeographic structure. Here, the approach developed by Pons and Petit (1996) based on contrasting the coefficients of genetic differentiation for ordered (N_{ST}) and unordered (G_{ST})

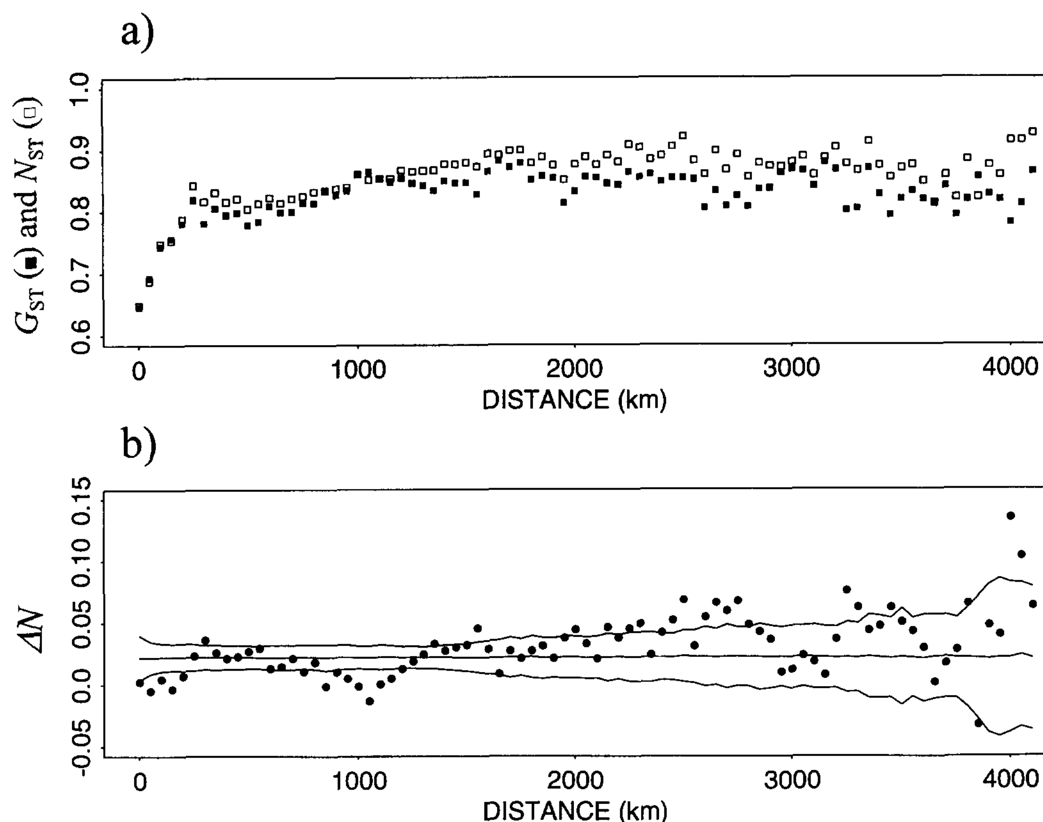

FIGURE 4.—(a) Differentiation parameters G_{ST} and N_{ST} computed from the average of the within-population gene diversities and the total gene diversities of all the pairs of populations for each 50-km distance class. (b) ΔN parameter and 95% confidence interval of its mean permutation value for each distance class after 1000 random permutations of the data.

alleles was used first. Despite their very close values, N_{ST} is significantly higher than G_{ST} in our study. Possible interpretations for such a difference are extensively discussed in PONS and PETIT (1996) (see in particular their Figure 2). Briefly, N_{ST} will be higher than G_{ST} if, on average, pairs of *different* haplotypes from the same population have more similar sequences than pairs of *different* haplotypes from separate populations.

However, no spatial information is taken into account in this test, which may restrict its power. Hence, we extended this approach to take into account the geographic distance between populations. A well-established method to assess the degree of concordance between distance matrices (genetic and geographic distances in our case) is the Mantel's procedure (MANTEL 1967), which generates statistical significance levels for correlation measures of similarity between distance matrices (DOW and CHEVERUD 1985). Here, the existence of a highly significant positive correlation between ΔN (the contribution of the genetic divergence between haplotypes to the overall coefficient of gene differentiation: $N_{ST} = G_{ST} + \Delta N$) and the geographic distance was demonstrated. This means that pairs of increasingly distant populations are also increasingly different genetically, not only in terms of haplotype frequencies (*cf.* the increase of both G_{ST} and N_{ST} with distance) but also in terms of the mean molecular divergence between their haplotypes (*cf.* the increase of ΔN with distance, *i.e.*, the faster increase of N_{ST} compared to G_{ST} with distance).

The next step was to analyze the global phylogeographic structure in more details, by computing the different genetic parameters within distance classes of 50 km. Both coefficients of differentiation increase up to distances of ~250 km. But N_{ST}, initially similar to G_{ST} (0.65), becomes then larger, with some fluctuations around its mean value of 0.85. Hence, below this threshold of 250 km, only the haplotype frequencies matter to analyze the genetic structure. Afterwards, it becomes useful to take into account the distance between haplotypes.

The advantage of this approach is to provide a statistical confidence to analyses based on classical genetic parameters. The global phylogeographic test based on Mantel's test, if significant, can then be applied to smaller distance classes. Of course, this requires that sufficient and well-distributed data are available. All these analyses are based on the estimation of coefficients of differentiation for pairs of populations that depend on the genetic divergence between alleles only when at least one population is polymorphic. Given the high frequency of genetically fixed populations in our sample, it may be worth investigating other genetic parameters for these tests (such as genetic distances or other measures of differentiation). The only requirement is that their value can be computed and compared for both ordered and unordered alleles. As discussed in PONS and PETIT (1996), the proposed tests are purely descriptive and make no inference on the underlying evolutive forces: only sampling effects are considered

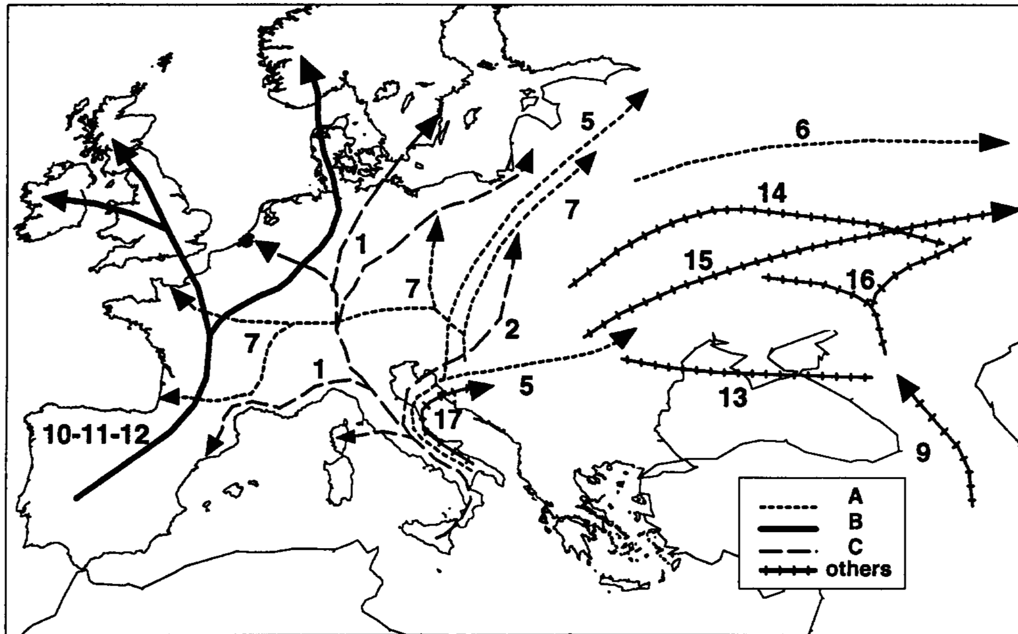

FIGURE 5.—Possible postglacial colonization routes of the haplotypes inferred from their distribution throughout the range. Each type of arrow corresponds to a lineage (A–C and others) and the number of the haplotype involved is given.

here. Such analyses may be followed by more complex ones aimed at testing specific evolutionary models, once a significant pattern has been demonstrated.

Inferences of colonization routes: According to palynological data, oaks were known to be present in refugia localized in southern Europe: mainly in the Iberian peninsula, in southern Italy and in the Balkans. By comparing these palaeoecological data, and particularly the isopollen maps of HUNTLEY and BIRKS (1983) with our results (Figure 2), we constructed a map (Figure 5) that summarized the possible postglacial colonization routes of the haplotypes. In Iberia and in the Balkans, fossil pollen evidence suggest the presence of refugia in regions where no thorough sampling had been carried out. Therefore, the colonization routes that are indicated may be incomplete for some haplotypes.

The statistical analyses developed above have demonstrated that phylogenetically related haplotypes are likely to occupy similar geographic regions. In particular, the three haplotypes of lineage B are likely to have expanded from a refugia in the Iberian peninsula northeastward, resulting in a similar distribution along the western part of the range of the oaks. But this is not always true. For instance, haplotypes 1 and 2 from lineage C seem to have expanded northward from an Italian refugia, but only haplotype 1 is found along the Mediterranean coast of France and Spain, whereas haplotype 2 has taken a more eastern route, like haplotype 5, which belongs to a different lineage (A).

For haplotypes that do not display a typically latitudinal distribution, the direction of colonization is not so obvious: a Balkanian origin can be supposed for haplotypes 7, 13, 14 and 15. However, the distribution of haplotypes 9, 16 and even 13 and 14 could also suggest a Caucasian refugia (Figure 5). Actually, authors do not agree about the existence of possible refugia in the Caucasus having sheltered temperate trees during the last glaciation (see for instance ZEIST and BOTTEMA 1988; FRENZEL 1992). Hence, even for wind-pollinated tree species that are well represented in Quaternary sediments, genetic studies could potentially answer questions that have remained unsolved in palaeoecological studies. However, in this particular case, additional information from the Caucasus and the Balkans is clearly needed.

Another Fagaceae, the European beech (*Fagus sylvatica*), which has presently a similar though more circumscribed geographic range than the oaks, was apparently absent from the Iberian peninsula during the last ice age (HUNTLEY and BIRKS 1983). The study of cpDNA variation in this species has indeed failed to detect a western lineage (DEMESURE *et al.* 1996). There was only one lineage restricted to Italy and another one present throughout most of the remaining of the range of the beech. Furthermore, for this tree, which recolonized Europe much later than the oaks, the Italian populations did not expand north of the Alps during the postglacial, a result that contrasts remarkably with those described here for the oaks. It appears therefore that not only the number and location of the refugia but also the dynamics at the outset of recolonization can vary from species to species and have major impacts on the present genetic structure of these tree species.

Although there is more chloroplast diversity in the oaks than in the beech, oak cpDNA haplotypes that are present in the north of Europe were also detected in southern Europe, *i.e.*, across a very broad latitudinal gradient. They have therefore remained indistinguishable with our technique during this extensive northward spread. Hence, most polymorphisms detected are

probably anterior to the beginning of the last recolonization. These results emphasize the low rate of evolution of the chloroplast genome in these species. A phylogenetic study of range expansion, based on characters that have evolved during recolonization (THORPE 1984), does not seem possible for these oak species, unless even more resolutive and systematic methods of detection of cpDNA or mtDNA mutations could be developed.

On the other hand, the much longer glacial periods must have led to the divergence of cpDNA haplotypes in the different southern refugia. During the postglacial expansion, the initial genetic structure may have been merely extended northward. However, the contact of expanding lineages could have generated an overlap in the distribution of haplotypes originating from the various refugia. Our results indicate that the contact between expanding oak populations originating from different refugia resulted in a relatively sharp transition: this is particularly clear between the western lineage B and haplotypes from the lineages A and C. Nevertheless, haplotype 7 (lineage A) did extend far to the west, resulting into some interpenetration with lineage B. The presence of haplotype 7 in Wales could be due either to an early migration from western France or to an artificial introduction during historical times. Detailed studies of the distribution of these haplotypes in regions of contact could improve our understanding of the dynamic of colonization in these oaks.

Clearly, the distribution of refugia during previous ice ages along with the random sampling of haplotypes during each interglacial and the timing of the outset of recolonization in each refugia probably played the greatest role in determining the present geographic pattern of cytoplasmic diversity in these oaks. But other factors such as human influences cannot be ruled out. The study of long-distance seed transfers should now be greatly facilitated by the availability of detailed distribution maps of the chloroplast molecular variants in these oaks.

The following people sent material used in our analyses: G. AAS, N. BAHRMAN, A. BAUER, M. P. BONVICINI, S. BORDACZ, J. BRACH, K. BURG, G. BURZINSKI, C. E. DUREL, Y. EFIMOV, V. ENESCU, G. ERIKSSON, A. FLETCHER, L. GIL, D. GÖMÖRY, G. IEVLEV, D. JACQUES, M. KIZMAZ, J. KLEINSCHMIT, A. KÖNIG, J. KOPLIKA, V. KOSTRIKIN, H. LE BOULER, K. LAGERSTROM, K. LIEPE, S. MAC AN TSAOIR, S. MADSEN, O. MATYAS, F. MÜLLER, K. PAN, L. PAULE, S. RADU, P. RADOSTA, J. SABOR, V. SHIZNIN, Y. SHUTYAEV, T. SKROPPA, S. SOKOL, V. STANESCU, B. STIMM, A. TUCOVIC, B. VAN DER AA, M. C. VARELA, A. ZANETTO, the French National Forest Service (ONF), and all the participants to the technical meeting of the EC project (FAIR1 PL95-0297). The technical assistance of M. BERTOCCHI, J.-M. LOUVET, M.-H. PEMONGE and G. ROUSSEL is greatly acknowledged. We are particularly indebted to A. DUCOUSSO who was much involved in the collection of oak material, particularly in the east of Europe, and to J. TIMBAL for his help in the study of the chorology and taxonomy of the oak species. We thank P. TABERLET for his help with the phylogenetic analyses. During the whole study, the encouragement and help of A. KREMER, who additionally participated in the collection of the material, were decisive. Comments on an earlier version of the manuscript from J.-L. DE BEAULIEU and P. TABERLET were greatly appreciated. The research was supported by the EC Research programs in Biotechnology (BIO-CT9-0373) and Forest (MA2B-CT91-0022).

LITERATURE CITED

AVISE, J. C., J. ARNOLD, R. M. BALL, E. BERMINGHAM, T. LAMB *et al.*, 1987 Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. **18:** 489–522.

BROWICZ, K., 1982 *Chorology of Trees and Shrubs in South-West Asia and Adjacent Regions.* Vol. 1. Polish Scientific Publishers, Poland.

BROWICZ, K., 1988 *Chorology of Trees and Shrubs in South-West Asia and Adjacent Regions.* Vol. 6. Polish Scientific Publishers, Poland.

CAMUS, A., 1936–1939 *Les Chênes. Monographie du Genre Quercus.* Le Chevallier, Paris.

CLEGG, M. T., B. S. GAUT, G. H. LEARN, JR., and B. R. MORTON, 1994 Rates and patterns of chloroplast DNA evolution. Proc. Natl. Acad. Sci. USA **91:** 6795–6801.

CWYNAR, L. C., and G. M. MACDONALD, 1987 Geographical variation of lodgpole pine in relation to population history. Am. Nat. **129:** 463–469.

DAVIS, M. B., 1976 Pleistocene biogeography of temperate deciduous forests. Geoscience and Man **13:** 13–26.

DEMESURE, B., 1996 Analyse de la diversité chloroplastique en utilisant des fragments-PCR chez des Fagacées: *Fagus sylvatica* L. et *Quercus* ssp. Ph.D. Thesis, University of Bordeaux I, France.

DEMESURE, B., N. SODZI and R. J. PETIT, 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.

DEMESURE, B., B. COMPS and R. J. PETIT, 1996 Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. Evolution **50:** 2515–2520.

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

DOW, M. M., and J. M. CHEVERUD, 1985 Comparison of distance matrices in studies of population structure and genetic microdifferentiation: quadratic assignment. Am. J. Phys. Anthropol. **68:** 367–373.

DOWNIE, S. R., and J. D. PALMER, 1992 Use of chloroplast DNA rearrangements in reconstructing plant phylogeny, pp. 14–35 in *Molecular Systematics of Plants,* edited by P. S. SOLTIS, D. E. SOLTIS and J. J. DOYLE. Chapman and Hall, New York.

DUMOLIN, S., B. DEMESURE and R. J. PETIT, 1995 Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. Theor. Appl. Genet. **91:** 1253–1256.

DUMOLIN-LAPÈGUE, S., M.-H. PEMONGE and R. J. PETIT, 1997 An enlarged set of consensus primers for the study of organelle DNA in plants. Mol. Ecol. **6:** 393–397.

EL MOUSADIK, A., and R. J. PETIT, 1996 Chloroplast DNA phylogeography of the argan tree of Morocco. Mol. Ecol. **5:** 547–555.

ENNOS, R. A., 1994 Estimating the relative rates of pollen and seed migration among plant populations. Heredity **72:** 250–259.

FELSENSTEIN, J., 1985 Phylogenies and the comparative method. Am. Nat. **125:** 1–15.

FELSENSTEIN, J., 1993 PHYLIP (Phylogeny Inference Package) version 3.5c. Department of Genetics, University of Washington, Seattle.

FERRIS, C., R. P. OLIVER, A. J. DAVY and G. M. HEWITT, 1993 Native oak chloroplasts reveal an ancient divide across Europe. Mol. Ecol. **2:** 337–344.

FERRIS, C., R. P. OLIVER, A. J. DAVY and G. M. HEWITT, 1995 Using chloroplast DNA to trace postglacial migration routes of oaks into Britain. Mol. Ecol. **4:** 731–738.

FRASCARIA, N., L. MAGGIA, M. MICHAUD and J. BOUSQUET, 1993 The *rbc*L gene sequence from chestnut indicates a slow rate of evolution in the Fagaceae. Genome **36:** 668–671.

FRENZEL, B., 1992 Vegetation during the maximum cooling of the last glaciation, in *Atlas of Palaeoclimates and Palaeoenvironments of the Northern Hemisphere,* edited by B. FRENZEL, B. PECSI and A. A. VELICHKO. INQUA/Hungarian Academy of Sciences, Budapest.

GIELLY, L., and P. TABERLET, 1994 Chloroplast DNA polymorphism

at the intrageneric level and plant phylogenies. C. R. Acad. Sci. Paris **317:** 685–692.

HEWITT, G. M., 1996 Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. **58:** 247–276.

HUNTLEY, B., 1990 European vegetation history: palaeovegetation maps from pollen data—13 000 yr BP to present. J. Quater. Sci. **5:** 103–122.

HUNTLEY, B., and H. J. B. BIRKS, 1983 *An Atlas of Past and Present Pollen Maps for Europe, 0–13,000 Years Ago.* Cambridge University Press, Cambridge, U.K.

HUNTLEY, B., and T. WEBB III, 1989 Migration: species' response to climatic variations caused by changes in the earth's orbit. J. Biogeogr. **16:** 5–19.

IBRAHIM, K. M., R. A. NICHOLS and G. M. HEWITT, 1996 Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. Heredity **77:** 282–291.

JALAS, J., and J. SUOMINEN, 1976 *Atlas Florae Europaeae. 3. Salicaceae to Balanophoraceae.* Helsinki.

LE CORRE, V., N. MACHON, R. J. PETIT and A. KREMER, 1997 Colonisation with long-distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. *Genet. Res. Cambr.* (in press).

LYNCH, M., and T. J. CREASE, 1990 The analysis of population survey data on DNA sequence variation. Mol. Biol. Evol. **7:** 377–394.

MANTEL, N., 1967 The detection of disease clustering and a generalized regression approach. Cancer Res. **27:** 209–220.

MCCAULEY, D. E., 1995 The use of chloroplast DNA polymorphism in studies of gene flow in plants. TREE **10:** 198–202.

PETIT, R. J., A. KREMER and D. B. WAGNER, 1993a Geographic structure of chloroplast DNA polymorphisms in European oaks. Theor. Appl. Genet. **87:** 122–128.

PETIT, R. J., A. KREMER and D. B. WAGNER, 1993b Finite island model of organelle and nuclear genes in plants. Heredity **71:** 630–641.

PETIT, R. J., B. DEMESURE and E. PINEAU, 1996 Chloroplast DNA differentiation at a local and continental scale in European oak species: the importance of historical factors, pp. 145–164 in *Inter- and Intra-Specific Variation in European Oaks: Evolutionary Implications and Practical Consequences,* edited by A. KREMER and H.-J. MUHS. Office for Official Publications of the European Communities, Brussels.

PETIT, R. J., E. PINEAU, B. DEMESURE, R. BACILIERI, A. DUCOUSSO *et al.,* 1997 Chloroplast DNA footprints of postglacial recolonization by oaks. Proc. Natl. Acad. Sci. USA (in press).

PONS, O., and R. J. PETIT, 1995 Estimation, variance and optimal sampling of gene diversity. I. Haploid locus. Theor. Appl. Genet. **90:** 462–470.

PONS, O., and R. J. PETIT, 1996 Measuring and testing genetic differentiation with ordered *vs.* unordered alleles. Genetics **144:** 1237–1245.

SCHMUCKER, T., 1942 *The Tree Species of the Northern Temperate Zone and Their Distribution.* Centre International de Sylviculture, Berlin-Wannsee.

SOLTIS, D. E., P. S. SOLTIS and B. G. MILLIGAN, 1992 Intraspecific chloroplast DNA variation: systematic and phylogenetic implications, pp. 117–150 in *Molecular Systematics of Plants,* edited by P. S. SOLTIS, D. E. SOLTIS and J. J. DOYLE. Chapman and Hall, New York.

SOLTIS, D. E., M. A. GITZENDANNER, D. D. STRENGE and P. S. SOLTIS, 1997 Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Plant Syst. Evol. (in press).

SPEIGHT, M. R., and D. WAINHOUSE, 1989 *Ecology and Management of Forest Insects.* Clarendon Press, Oxford.

SWOFFORD, D. L., 1993 Phylogenetic analysis using parsimony. Vers. 3.1. Ill. Nat. Hist. Surv., Champaign.

TABERLET, P., L. GIELLY and J. BOUVET, 1991 Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. **17:** 1105–1109.

TEMPLETON, A. R., 1993 The "Eve" hypotheses: a genetic critique and reanalysis. Am. Anthropol. **95:** 51–72.

TEMPLETON, A. R., E. ROUTMAN and C. A. PHILLIPS, 1995 Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum.* Genetics **140:** 767–782.

THORPE, R. S., 1984 Primary and secondary transition zones in speciation and population differentiation: a phylogenetic analysis of range expansion. Evolution **38:** 233–243.

WHITTEMORE, A. T., and B. A. SCHAAL, 1991 Interspecific gene flow in oaks. Proc. Natl. Acad. Sci. USA **88:** 2540–2544.

WOLFE, K. H., W.-H. LI and P. M. SHARP, 1987 Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Natl. Acad. Sci. USA **84:** 9054–9058.

ZANETTO, A., and A. KREMER, 1995 Geographic structure of gene diversity in *Quercus petraea* (Matt.) Liebl. 1. Monolocus patterns of variation. Heredity **75:** 506–517.

ZEIST, VAN W. and BOTTEMA, S. 1988 Late Quaternary vegetational and climatic history of Southwest Asia, pp. 129–148 in *Palaeoclimatic and Palaeoenvironmental Change in Asia During the Last 4 Million Years,* Indian National Science Academy, New Delhi.

Communicating editor: W-H. LI

APPENDIX A

Description of the 23 individuals used in the survey and complete data set for the chloroplast DNA

Type	Population name	Country	Species	Polymorphic fragments 1	2	3	4	5	6	7	8	9	10
1	Grogna	Central Italy	*Q. robur*	9	1	2	1	2	2	2	2	2	2
2	Klostermarienberg	Austria	*Q. petraea*	9	1	2	1	1	2	9	2	2	2
3	Petrusa	South Italy	*Q. pubescens*	9	1	2	2	1	2	9	2	2	2
4	Nagybatony	Hungria	*Q. petraea*	1	1	1	1	1	3	2	2	2	2
5	Casale	Central Italy	*Q. pubescens*	1	1	2	1	1	3	2	2	2	2
6	Policoro	South Italy	*Q. robur*	2	1	2	1	1	3	2	2	2	2
7	Lucatelli	South Italy	*Q. frainetto*	1	1	5	1	1	3	2	2	2	2
8	Crimea	Ukrainia	*Q. robur*	1	1	2	1	1	5	2	2	2	1
9	Ereven	Armenia	*Q. macranthera*	1	1	2	1	1	4	2	2	2	1
10	Mimizan	South-West France	*Q. robur*	1	2	3	1	1	2	2	2	2	2
11	Arès	South-West France	*Q. robur*	1	2	3	1	1	2	2	2	2	2
12	Hostens	South-West France	*Q. robur*	1	2	4	1	1	2	2	2	2	2
13	Brasov	Rumania	*Q. petraea*	1	1	2	1	1	2	2	2	2	2
14	Briansk	Russia	*Q. robur*	1	1	2	1	1	2	2	2	1	2
15	Dumbrava	Rumania	*Q. robur*	1	1	3	1	1	2	2	2	2	2
16	Tchernigov	Ukrainia	*Q. robur*	1	1	3	1	1	2	2	2	2	2
17	Chieuti	South Italy	*Q. pubescens*	1	1	3	1	1	2	2	2	2	2
18	Bolu	Turkey	*Q. petraea*	1	1	3	1	1	2	9	2	2	2
19	Renacci	North Italy	*Q. robur*	1	1	3	1	1	2	2	1	2	2
20	Abazzia	Corsica	*Q. pubescens*	1	1	3	1	1	1	2	2	2	2
21	Carmaux	South-West France	*Q. robur*	1	1	3	1	1	2	9	1	2	2
22	Djijeli	Algeria	*Q. canariensis*	1	1	3	1	1	2	2	1	2	2
23	Policoro	South Italy	*Q. robur*	1	1	6	1	2	3	1	3	2	2

Type	Polymorphic fragments 11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	2	1	1	1	3	2	3	1	2	2	9	2	9	2	2	1	1	2	2	1
2	2	1	1	9	2	2	3	1	2	2	9	2	9	1	2	1	1	2	2	1
3	2	1	1	9	2	2	3	1	2	2	9	2	9	1	2	1	1	2	2	1
4	2	9	1	1	2	2	2	1	1	2	1	2	1	2	2	1	2	1	2	2
5	2	9	1	1	2	2	2	1	1	2	1	2	1	2	2	1	2	1	2	2
6	2	9	1	1	1	2	2	1	1	2	1	2	1	2	2	1	2	1	2	2
7	2	9	1	1	2	2	2	3	1	2	1	2	1	2	2	1	2	1	2	2
8	2	1	1	1	2	2	3	1	2	2	1	2	1	2	2	1	2	2	2	1
9	2	1	1	1	2	2	3	1	2	2	1	2	1	2	2	1	2	2	2	1
10	2	1	1	1	2	2	3	1	2	2	1	2	1	2	2	1	1	2	2	1
11	1	1	1	1	2	2	3	1	2	2	1	2	1	2	2	1	1	2	2	1
12	2	1	1	1	2	2	3	1	2	2	1	2	1	2	2	1	1	2	2	1
13	2	1	1	1	2	2	2	1	2	2	1	2	1	1	2	9	1	2	2	1
14	2	1	1	1	2	2	3	1	2	2	1	2	1	2	2	9	1	2	2	1
15	2	1	1	1	2	1	3	1	2	2	1	2	1	2	2	9	1	2	2	1
16	2	1	2	1	2	2	3	1	2	2	1	2	1	2	2	9	1	2	2	1
17	2	1	1	1	2	2	3	1	2	2	1	2	1	2	2	1	1	2	2	1
18	2	1	1	1	2	2	3	1	2	2	1	2	1	1	2	9	1	2	2	1
19	2	1	1	1	2	2	3	1	2	2	1	2	1	2	2	1	1	2	1	1
20	2	1	1	1	2	2	3	1	2	1	1	2	1	2	1	9	1	2	2	1
21	2	1	1	1	2	2	3	1	2	2	1	2	1	2	2	9	1	2	2	1
22	2	1	3	1	2	2	1	1	2	2	1	2	1	2	2	9	1	2	2	1
23	2	1	1	1	4	2	3	2	2	3	1	1	2	2	2	1	1	2	2	1