Molecular Ecology (2004) **13**, 3467–3476 doi: 10.1111/j.1365-294X.2004.02344.x

Chloroplast DNA variation in the *Quercus affinis–Q. laurina* complex in Mexico: geographical structure and associations with nuclear and morphological variation

A. GONZÁLEZ-RODRÍGUEZ,* J. F. BAIN,† J. L. GOLDEN† and K. OYAMA*
Centro de Investigaciones en Ecosistemas, Universidad Nacional Autónoma de México, Antigua Carretera a Pátzcuaro no. 8701, Col. Ex–Hacienda de San José de la Huerta, Morelia, 58190 Michoacán, México, †Department of Biological Sciences, University of Lethbridge, Lethbridge, Alberta, Canada T1K 3M4

Abstract

The geographical distribution of chloroplast DNA (cpDNA) variation in 39 populations of two hybridizing Mexican red oaks, *Quercus affinis* and *Q. laurina*, was investigated using polymerase chain reaction-restriction fragment length polymorphism (PCR–RFLP). Six haplotypes were identified. Of these, two (H1 and H4), separated by four mutations, had high frequencies (58 and 23% of the individuals, respectively) and were present across the whole geographical range of both species, often co occurring in the same populations. The other four haplotypes were rare, geographically restricted, and are probably derived from the two frequent haplotypes. Latitudinal or other clinal patterns in diversity levels or haplotype composition of populations were not apparent. The pattern of haplotype distribution was characterized by some mosaicism, with contrasting populations often situated in proximity. Average within-population diversity (h_S = 0.299) and population differentiation (G_{ST} = 0.499) were, respectively, higher and lower than values reported in previous studies of oak species. There was evidence for phylogeographical structure, as indicated by N_{ST} (0.566) being significantly higher than G_{ST}. Haplotypic variation was largely species-independent, although some very weak associations were detected between haplotypes H1 and H4 and morphological and nuclear molecular variation correspondingly characterizing *Q. affinis* and *Q. laurina*. These oaks probably did not experience a marked restriction to one or a few particular subregions of their present range during the last glacial cycle. It is more likely that substantial populations persisted throughout several episodes of climatic change, but experienced recurrent latitudinal and altitudinal migrations which may have caused the widespread distribution of haplotypes H1 and H4 and frequent intermixing of populations.

Keywords: chloroplast DNA, geographical structure, hybridization, Mexico, population history, *Quercus*

Received 21 April 2004; revision received 14 July 2004; accepted 9 August 2004

Introduction

Analysis of chloroplast DNA (cpDNA) variation within plant species is a useful tool to reconstruct historical events such as population expansions and contractions, migration and colonization (McCauley 1995; Ennos *et al.* 1999), as well as to detect ancient and contemporary hybridization (Rieseberg & Soltis 1991; Rieseberg *et al.* 1996). In a number of angiosperm species, particularly from northern latitudes, high differentiation among populations and clear-cut geographical distributions of cpDNA haplotypes have indicated restricted seed dispersal, and seemingly reflect drastic isolation of populations into refugia during the last glaciation, postglacial migration and range expansions (Sewell *et al.* 1996; Soltis *et al.* 1997; Lumaret *et al.* 2002; Palmé & Vendramin 2002; Cavers *et al.* 2003; Petit *et al.* 2003; Rendell & Ennos 2003). However, there are some species that appear to have been affected less severely by climatic changes. For example, *Salix caprea* was probably not confined to small and isolated refugia in southern

Correspondence: Antonio González-Rodríguez. Fax: + 52 55 56 23 27 19; E-mail: agrodrig@oikos.unam.mx

Europe during the last glacial maximum, but persisted in intermediate latitude refugia with large population sizes which, when combined with high dispersal ability, have resulted in the absence of geographical structure of cpDNA variation (Palmé *et al.* 2003). At lower latitudes it could have been more common for plant species to experience such comparatively moderate population size changes and range shifts, because recent climatic fluctuations were probably less extreme in these areas (Flenley 1998). Consequently, the spatial distribution of cpDNA variation in these cases is in general expected to reflect less markedly recent climatic fluctuations and, instead, patterns considerably more ancient than the last glacial cycle may be present in some species (Cannon & Manos 2003). However, few plant phylogeographical studies have been conducted so far in these areas.

The genus *Quercus* is represented in Mexico by 135–150 species, with approximately 86 of them endemic, which indicates that a major secondary radiation of this group occurred within this area (Nixon 1993; Manos *et al*. 1999). All the major mountain ranges in Mexico are rich in species of *Quercus* (Nixon 1993) and approximately 5.5% of the country is covered by oak forests and woodlands, and 13% by pine–oak forests. In all these communities the presence of several intermixed oak species is common (Rzedowski 1981). Some oaks also occur in xeric areas, cloud forests and tropical forests (Rzedowski 1981; Nixon 1993). However, the evolution of oaks in Mexico is poorly understood. Fossil pollen belonging to this genus has been identified in Mio–Pliocene formations situated in Guatemala and Panama (Graham 1999), which suggests a minimum late Miocene age for the arrival of oaks into what is now Mexico. Diverse topography, climate and habitat may have been very important in the processes of radiation and maintenance of *Quercus* diversity in this region. The impact of changes in temperature and precipitation regimes during the late Tertiary and the Quaternary on the populations of particular oak species is unknown, but it is probable that recurrent latitudinal and elevational displacements occurred, affecting patterns of intra- and interspecific gene flow dynamics and to an important extent shaping the geographical distribution and variation among species that we observe today. Human influence during the last few thousand years has almost certainly also caused some changes in the genetic diversity and structure of Mexican oak populations (Butzer & Butzer 1997; Conserva & Byrne 2002), but these have probably not been of great extent. In many regions oaks are used intensively at a local scale, mainly as a source of timber and firewood, and forests have been cleared and fragmented to create land for agriculture and grazing of cattle, but large-scale exploitation has been kept low (Rzedowski 1981). The establishment of populations through plantation and artificial seed transfer has been relatively infrequent.

In previous cpDNA phylogeographical studies conducted on *Quercus* species, most notably involving the European white oaks, extensive documentation of variation at a continental scale has revealed strong geographical structuring of cpDNA lineages, mainly along a longitudinal gradient (Petit *et al*. 2002a,b). This information, complemented with the examination of fossil pollen data, has been used to obtain a precise delineation of the location of the refugia where oaks survived the last glacial period and to assess the postglacial recolonization dynamics of these species (Brewer *et al*. 2002; Petit *et al*. 2002b). Together with the clear-cut geographical patterns revealed by these results, an overall picture of nearly species-independent sharing of cpDNA variation in *Quercus* has emerged, which implies systematic cpDNA exchanges between sympatric oak species through hybridization and introgression (Whittemore & Schaal 1991; Dumolin-Lapègue *et al*. 1999; Belahbib *et al*. 2001; Petit *et al*. 2002a,b). This makes it difficult to identify unambiguously the origin of shared cpDNA haplotypes, but nevertheless some initial associations between some white oak species and particular haplotype lineages have been suggested (Petit *et al*. 2002a).

Quercus affinis Scheidweiler and *Q. laurina* Humboldt et Bonpland are closely related Mexican red oaks. Phenotypic and genetic patterns of variation in this complex suggest that ancient events of secondary contact between the two species have resulted in frequent hybridization and introgression, leading to extensive morphological and genetic intergradation (Valencia 1994; González-Rodríguez *et al*. 2004). Hybridization probably also occurs between *Q. laurina* and at least four other red oak species (*Q. crassifolia*, *Q. crassipes*, *Q. mexicana* and *Q. rubramenta*), although with much lower frequency (Valencia 1994). Intergradation between *Q. affinis* and *Q. laurina* occurs mainly within a wide region of overlap in the distribution of the two oaks, whereas interspecific populations outside this area show clear genetic and phenotypic differentiation (González-Rodríguez *et al*. 2004). The area of overlap is situated in the eastern portion of the Trans-Mexican Volcanic Belt and northern Oaxaca, while representative populations of *Q. laurina* are distributed at altitudes between 2440 and 3065 m along the Sierra Madre del Sur and the western region of the volcanic belt, and representative populations of *Q. affinis* occur with an altitudinal range of 1600–2800 m in the Sierra Madre Oriental (Fig. 1). From these patterns of geographical variation, it has been suggested (Valencia 1994) that the divergence between the two species occurred in isolation (*Q. laurina* in the Sierra Madre del Sur and *Q. affinis* in the Sierra Madre Oriental). The time when this could have happened is not clear, but it was possibly during the mid- or late Pliocene (Valencia 1994). Subsequently, secondary contact probably took place in the volcanic belt and northern Oaxaca during episodes of range expansion favoured by climatic conditions. The glacial and interglacial

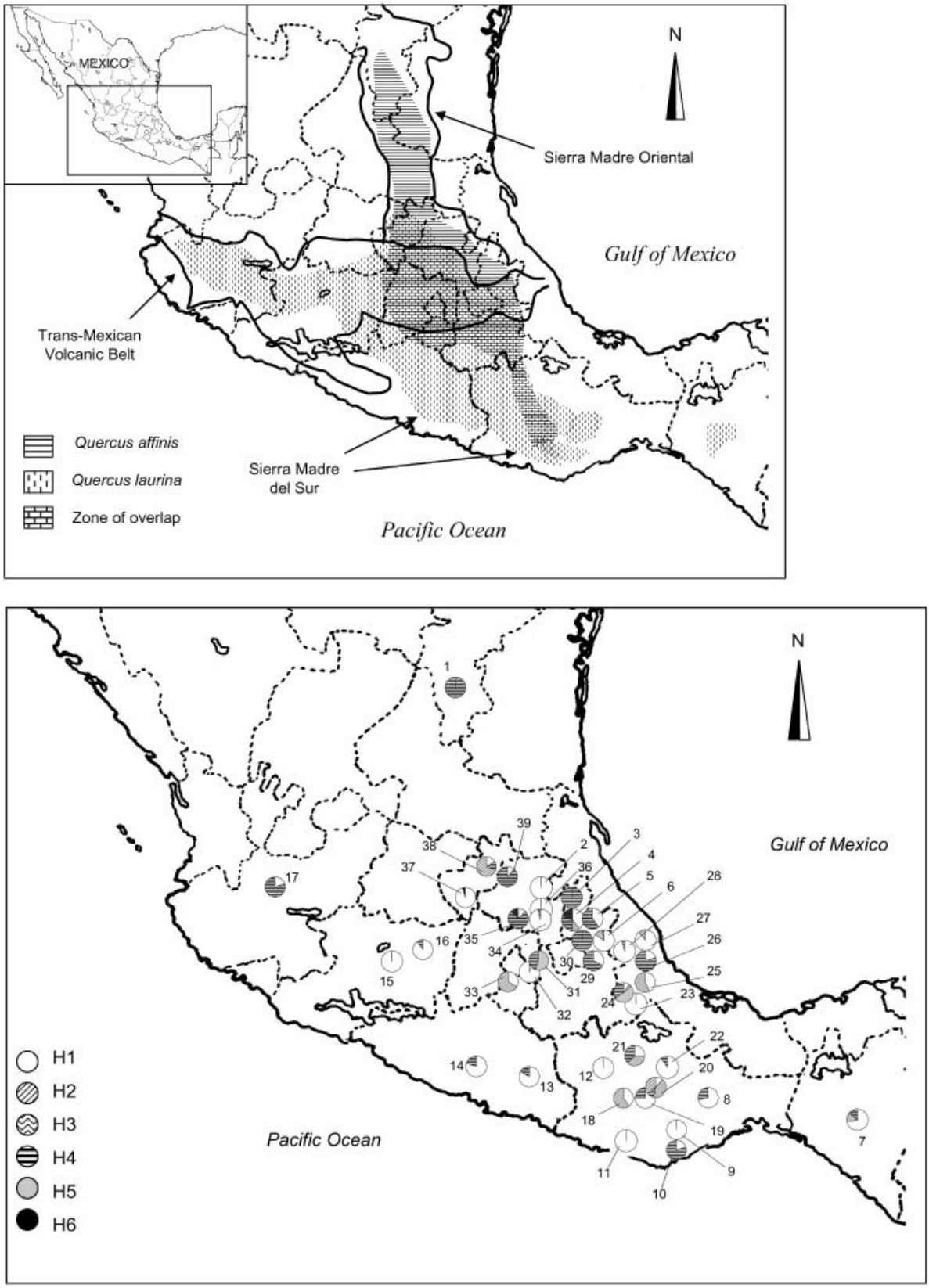

Fig. 1 Top panel, map of the geographical distribution of *Quercus affinis* and *Q. laurina*. The mountain ranges mentioned in text are sketched. Note the overlap area situated in the eastern portion of the volcanic belt and northern Oaxaca. Bottom panel, map of the populations and the geographical distribution of chloroplast PCR–RFLP haplotypes.

periods of the Pleistocene probably determined further episodic changes in the geographical distribution and, thus, in the intra- and interspecific gene flow dynamics of these oaks (Valencia 1994).

In this study we performed restriction enzyme analysis of polymerase chain reaction (PCR)-amplified regions [polymerase chain reaction-restriction fragment length polymorphism (PCR–RFLP)] to examine cpDNA variation in the *Q. affinis*–*Q. laurina* complex. First, we try to make some inferences about the spatio-temporal dynamics of these species, as suggested by the current distribution of cpDNA variants. We also wish to assess if cpDNA variation further supports that intergradation between *Q. affinis* and *Q. laurina* is the product of hybridization between previously allopatric species. If so, we expect the intermixing of rather divergent haplotypes in putative hybrid populations (Beckstrom-Sternberg *et al*. 1991; Holman *et al*. 2003). Finally, and if this occurs, we examine the degree of association between haplotypic variation and the morphological and nuclear molecular variation used previously to discriminate between both species (González-Rodríguez *et al*. 2004).

Materials and methods

Sampling

A total of 537 individuals of *Q. affinis* and *Q. laurina* from 39 localities across the whole distribution range of both species were included in the present study (Fig. 1 and Table 1). Of these, 355 individuals from 16 populations (Table 1) had been included previously in an analysis of hybridization between the two species using random amplified poymmorphic DNA (RAPD) markers and morphological variables (González-Rodríguez *et al*. 2004). At each site, young, undamaged leaves were collected from randomly chosen adult trees and frozen in liquid nitrogen, and then stored in a −80 °C freezer. In all cases, sampled individuals within a locality were separated by at least 100 m. A voucher specimen from each tree was also collected. Each individual was classified as *Q. affinis*, *Q. laurina* or intermediate (Table 1) using a previously obtained canonical discriminant function based on nine foliar traits (González-Rodríguez *et al*. 2004). Other sympatric oak species were observed in most localities, particularly *Q. rugosa* (a white oak) and *Q. candicans*, *Q. crassifolia* and *Q. crassipes* (red oaks). Individuals suspected to have resulted from hybridization of *Q. affinis* or *Q. laurina* with any other species were avoided in the present sampling.

PCR–RFLP

Genomic DNA was extracted from 100 mg of frozen leaf tissue using the protocol of Lefort & Douglas (1999) with minor modifications. PCR conditions were optimized for the amplification of six cpDNA regions (AS, CD, DT, HK, TF, ST), using universal primer pairs (Taberlet *et al*. 1991; Démesure *et al*. 1995). The thermal cycling programme was as follows for fragments CD, DT, HK, TF and ST: one cycle of 2 min at 94 °C; 30 cycles of 1 min at 94 °C, 1 min at 55 °C and 1 min 30 s at 72 °C. For fragment AS, the programme differed in that 35 cycles were used; the annealing temperature was 60 °C, and the extension time 2 min 30 s. A final step of 7 min at 72 °C was included in both programmes. All amplification reactions were carried out in 25 μL volumes, containing 10 ng template DNA; 1 × PCR buffer (Invitrogen); 2 mM $MgCl_2$; 0.1 mM each dNTP (Fermentas), 0.2 μM each primer; 5 μg bovine serum albumin (BSA); and 1.5 U *Taq* polymerase (Invitrogen).

The six amplified regions were screened for variation with nine restriction enzymes (Fermentas): *Alu*I, *Dra*I, *Mbo*I, *Hae*III, *Hha*I, *Hin*fI, *Msp*I, *Taq*I and *Vsp*I, using 20 individuals from 10 geographically separated populations. The restriction reactions were performed by adding 10 μL of PCR product to a mix containing 17.5 μL H_2O, 2 μL 10 × enzyme buffer (Fermentas), 5 U of the enzyme and 0.5 μL BSA if required. Reactions with all enzymes were incubated for 3 h at 37 °C, except *Taq*I, which was incubated at 65 °C. The restriction fragments were electrophoresed on 8% polyacrylamide gels using 1 × Tris-Borate-EDTA (TBE) buffer at 300 V for 3 h. The gels were stained with ethidium bromide and photographed under UV light.

Data analysis

Haplotype frequencies were calculated and graphed for each population. Genealogical relationships among haplotypes were depicted in a minimum-spanning network generated using the ARLEQUIN version 2000 program (Schneider *et al*. 2000). This software was also used to obtain haplotype diversity values (h_S) within each population according to the method of Nei (1987), which is based only on haplotype frequencies (unordered alleles). Average within-population haplotype diversity, total diversity (h_T) and the coefficient of population differentiation (G_{ST}) were obtained with the program PERMUT (http://www.pierroton.inra.fr/genetics/labo/Software/). This same program was used to estimate *v*-type parameters (Pons & Petit 1996), which take into account the genealogical relationships among haplotypes (ordered alleles). The permutation test implemented in PERMUT (Burban *et al*. 1999) was employed to compare parameters of population differentiation with unordered and ordered alleles (G_{ST} and N_{ST}, respectively). A value of N_{ST} significantly higher than the G_{ST} value indicates that genealogically more closely related haplotypes tend to occur together within populations, signalling a phylogeographical structure in the distribution of haplotypes (Pons & Petit 1996). The influence of spatial separation on the degree of differentiation among populations was

Table 1 Details of the geographical location of collection sites and number of *Quercus affinis*, *Q. laurina* and intermediate individuals sampled, frequency of PCR–RFLP haplotypes in each population and estimated within population diversity (h_S)

				Number of individuals			Frequency of haplotypes						
No.	Population name	Lat N/Long W	Geographic region	*Q. affinis*	*Q. laurina*	Intermediate	H1	H2	H3	H4	H5	H6	h_S (SE)
1	Cerro Potosí	24°55′/100°15′	A	8	0	1	0	0	0	9	0	0	0.000 (0.000)
2	Zacualtipán*	20°39′/98°40′	A	13	0	0	13	0	0	0	0	0	0.000 (0.000)
3	Huauchinango	20°05′/98°07′	A	7	0	4	0	0	0	11	0	0	0.000 (0.000)
4	Zacatlán	20°12′/98°03′	A	9	0	1	4	1	0	3	0	2	0.779 (0.091)
5	Teteles	19°50′/97°28′	A	4	0	1	2	0	0	3	0	0	0.600 (0.175)
6	Zacapoaxtla*	19°50′/97°40′	A	34	0	6	34	5	0	1	0	0	0.268 (0.084)
7	San Cristobal	16°50′/92°45′	L	0	7	0	5	1	0	1	0	0	0.524 (0.209)
8	Ayutla	17°01′/96°03′	L	0	7	0	5	0	0	2	0	0	0.476 (0.171)
9	Suchiltepec I	16°12′/96°30′	L	0	8	0	8	0	0	0	0	0	0.000 (0.000)
10	Suchiltepec II	16°10′/96°31′	L	0	5	0	1	0	0	4	0	0	0.400 (0.237)
11	Lachao	16°15′/97°10′	L	0	6	0	6	0	0	0	0	0	0.000 (0.000)
12	Shini Yucu	17°15′/97°45′	L	0	9	0	9	0	0	0	0	0	0.000 (0.000)
13	Tlapa	17°30′/98°55′	L	0	6	0	5	0	0	1	0	0	0.333 (0.215)
14	Filo de Caballo	17°35′/99°50′	L	0	5	0	4	0	0	1	0	0	0.400 (0.237)
15	Cerro Burro	19°25′/101°30′	L	0	6	0	6	0	0	0	0	0	0.000 (0.000)
16	Mil Cumbres*	19°40′/100°55′	L	0	7	3	9	0	0	1	0	0	0.200 (0.154)
17	Tequila*	20°50′/103°48′	L	0	20	0	4	0	0	16	0	0	0.337 (0.110)
18	Santa Inés*	17°03′/96°55′	O	0	15	5	8	12	0	0	0	0	0.505 (0.056)
19	Corral de Piedra	17°05′/96°32′	O	0	6	2	6	0	0	2	0	0	0.476 (0.171)
20	Llano de Flores*	17°30′/96°30′	O	5	10	2	2	7	6	2	0	0	0.721 (0.068)
21	Pápalo	17°50′/96°49′	O	3	3	2	2	3	0	3	0	0	0.750 (0.096)
22	Comaltepec	17°43′/96°30′	O	0	5	5	9	0	0	1	0	0	0.200 (0.154)
23	Zoquitlán	18°17′/97°03′	O	0	10	2	12	0	0	0	0	0	0.000 (0.000)
24	Puerto Aire*	18°45′/97°30′	O	2	22	5	3	1	0	10	15	0	0.623 (0.059)
25	Orizaba	18°45′/97°05′	O	2	8	4	6	8	0	0	0	0	0.527 (0.064)
26	Perote	19°35′/97°10′	O	0	4	1	1	0	0	4	0	0	0.400 (0.237)
27	Tonayan	19°51′/96°55′	O	3	0	6	8	1	0	0	0	0	0.222 (0.166)
28	Jalacingo*	19°45′/97°15′	O	8	2	12	21	0	0	1	0	0	0.091 (0.081)
29	Tetela	19°45′/97°56′	O	1	2	3	2	0	0	4	0	0	0.533 (0.172)
30	Chignahuapan	20°30′/98°32′	O	1	1	3	0	0	0	5	0	0	0.000 (0.000)
31	Río Frío	19°20′/98°37′	O	0	6	3	0	0	0	4	5	0	0.556 (0.090)
32	Ozumba*	19°05′/98°42′	O	2	16	10	28	0	0	0	0	0	0.000 (0.000)
33	Cuernavaca*	19°05′/99°15′	O	0	14	4	6	12	0	0	0	0	0.471 (0.082)
34	Cerro Navajas*	20°12′/98°30′	O	4	16	9	28	0	0	1	0	0	0.069 (0.063)
35	El Chico*	20°05′/98°40′	O	0	24	7	4	0	0	23	0	4	0.453 (0.099)
36	El Zembo*	20°15′/98°32′	O	4	14	7	25	0	0	0	0	0	0.000 (0.000)
37	Amealco*	20°10′/100°20′	O	1	17	3	20	0	0	0	0	1	0.095 (0.084)
38	Pinal de Amoles*	21°01′/99°40′	O	6	11	8	4	0	17	4	0	0	0.507 (0.099)
39	Jacala*	20°50′/99°05′	O	2	1	6	1	0	0	8	0	0	0.222 (0.166)
Total				109	294	134	311	51	23	125	20	7	

*The 16 populations previously analysed for morphological and RAPD variation (González-Rodríguez *et al*. 2004).
Geographic region: A = distribution area of *Q. affinis*; L = distribution area of *Q. laurina*; O = overlap area between the two species.

investigated by calculating pairwise G_{ST} and N_{ST} values with the program DISTON (http://www.pierroton.inra.fr/genetics/labo/Software/), and plotting means of these parameters against geographical distance classes.

To test for associations of cpDNA variation with morphological variation and nuclear markers differentiating the two species, we employed the data sets of nine foliar variables and nine semidiagnostic RAPD markers obtained previously for 355 of the same individuals analysed here (González-Rodríguez *et al*. 2004). The morphological data set was complemented with new measurements obtained for the rest of the individuals. Because two haplotypes were largely predominant, the rare haplotypes were excluded, and *t*-tests were used to compare the nine morphological variables between the two groups of individuals defined by the possession of each haplotype. The two groups were also compared for the frequency of the nine RAPD markers with Fisher's exact test (Sokal & Rohlf 1995).

Results

All individuals were screened using the five polymorphic fragment/enzyme combinations that were identified, AS-*Taq*I, DT-*Hha*I, HK-*Hha*I, HK-*Taq*I and TF-*Dra*I. These combinations revealed three insertion/deletion mutations (indels) and three restriction site mutations, constituting six haplotypes (Fig. 2, Table 2). Haplotype H1 was the most frequent and geographically widespread (present in 35 populations and 58% of the individuals), followed by haplotype H4, which was found in 26 populations and 23% of the individuals. These two more frequent haplotypes were separated from each other by four mutations. Haplotype H2 was distinguished from haplotype H1 by only one change and appeared in 10 populations and 9.6% of the individuals. Haplotypes H3, H5 and H6 differed from haplotype H4 by one mutation each. These three haplotypes were rare (4.3%, 3.7% and 1.3% of the individuals, respectively) and each restricted to two or three populations, that were in relative geographical proximity in the case of haplotypes H5 and H6 and quite distant in the case of haplotype H3 (Fig. 1, Table 1).

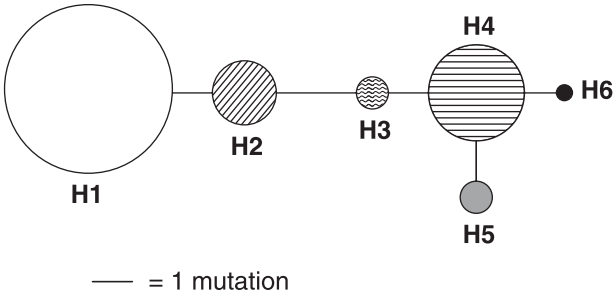

Fig. 2 Minimum-spanning network of the six haplotypes identified in populations of *Quercus affinis* and *Q. laurina*. One unit of branch length represents a single mutation. The sizes of the circles represent the frequency of each haplotype in the total sample set.

A large proportion of the populations (28; 76%) was polymorphic, with a weighted mean number of haplotypes in these populations of 2.55. Sample sizes within locations were not correlated with the levels of haplotype diversity detected ($r = -0.04$; $P = 0.787$). The spatial distribution of cpDNA variation did not suggest patterns of clinal change or clear differences in diversity levels between geographical areas (Fig. 1). The average within-population diversity, total diversity and differentiation estimates for the analysis with unordered alleles were $h_S = 0.299$, $h_T = 0.597$ and $G_{ST} = 0.499$; and for ordered alleles $v_S = 0.260$, $v_T = 0.598$ and $N_{ST} = 0.566$. The permutation test indicated a significantly higher value for the N_{ST} estimate than for the G_{ST} estimate ($P = 0.006$), indicating an association between the genealogical relationships among haplotypes and their geographical distribution. Specifically, haplotype H2 always occurred in populations in which haplotype H1 was also present, and haplotypes H3, H4 and H5 were found together with haplotype H4 in all but one occasion (Population Amealco, no. 34). Nonetheless, there were several populations in which mixing of haplotypes differing by four or five mutations occurred. In particular, the two more common haplotypes, H1 and H4 (differing by four mutations), were present together in 22 populations. These populations are situated along the distribution area of both species and not only in the region of overlap. Three

Table 2 Description of the cpDNA haplotypes identified in *Q. affinis*–*Q. laurina* populations by PCR–RFLP

	Polymorphic fragment/enzyme combinations					
Haplotype	AS-*Taq*I indel, band 1	AS-*Taq*I indel, band 2	DT-*Hha*I site	HK-*Hha*I site	HK-*Taq*I indel	TF-*Dra*I site
H1	1	1	0	1	1	1
H2	2	1	0	1	1	1
H3	2	1	1	0	1	1
H4	2	1	1	0	2	1
H5	2	1	1	0	2	0
H6	2	2	1	0	2	1

Table 3 Comparison of nine morphological variables between groups of individuals with haplotypes H1 and H4

Morphological variable	Mean values of morphological variables (standard error)				
	H1	H4	*t*†	***P***	R^2
Total leaf length	7.94 (0.102)	8.66 (0.172)	–3.59	0.0004	0.037
Lamina length	7.32 (0.092)	7.93 (0.155)	–3.39	0.0008	0.033
Petiole length	0.64 (0.019)	0.75 (0.032)	–2.67	0.0079	0.021
Maximal width	2.41 (0.037)	2.71 (0.063)	–4.11	< 0.0001	0.048
Distance from base to point of maximal width	3.73 (0.064)	3.87 (0.124)	–1.25	0.2139	0.005
Aristae number	3.02 (0.281)	2.49 (0.160)	–2.41	0.016	0.017
Petiole length/total length	0.08 (0.002)	0.08 (0.003)	–1.56	0.1193	0.007
Maximal width/lamina length	0.33 (0.004)	0.34 (0.006)	–1.74	0.0835	0.009
Distance from base to point of maximal width/lamina length	0.49 (0.004)	0.48 (0.007)	0.43	0.6707	0.001
Combined probability				< 0.0001	

†*t*-test statistic; d.f. = 536.
R^2, proportion of the total variance in each trait related to the possession of either haplotype.

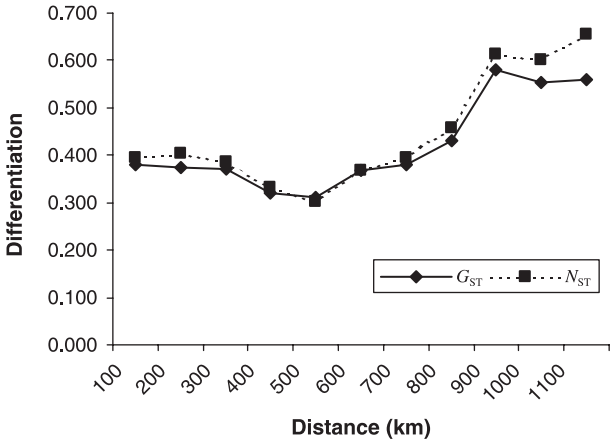

Fig. 3 Relationship of average pairwise G_{ST} and N_{ST} values with the geographical distances separating populations.

Table 4 Comparisons of the frequencies of nine semidiagnostic RAPD markers between groups of individuals with haplotypes H1 and H4

RAPD marker	Frequency of presence of the RAPD marker		Fisher's exact test *P*
	H1	H4	
A05–07	0.54	0.41	0.058
A07–09	0.39	0.52	0.068
A08–01	0.82	0.81	0.168
A10–03	0.94	0.88	0.132
B17–04	0.59	0.40	0.016
B17–06	0.56	0.52	0.386
C09–03	0.20	0.12	0.100
C09–05	0.23	0.29	0.250
I07–02	0.33	0.32	0.472
Combined probability			0.004

populations had particularly high levels of haplotype diversity (Table 1): Llano de Flores (no. 16, with haplotypes H1, H2, H3 and H4), Zacatlán (no. 27, with haplotypes H1, H2, H4 and H6) and Puerto Aire (no. 36, with haplotypes H1, H2, H4 and H5).

The plot of mean pairwise differentiation measures vs. geographical distances (Fig. 3) indicated that, up to a distance of about 700 km, the difference between G_{ST} and N_{ST} is small, and both measures do not show a clear increase from the initial values. Above this distance class there is an increase in differentiation, and the difference between G_{ST} and N_{ST} becomes larger.

The means of nine morphological variables and the frequency of nine RAPD markers that have been shown previously to be discriminatory between *Q. affinis* and *Q. laurina* (González-Rodríguez *et al*. 2004) were compared between individuals possessing the two more frequent haplotypes, H1 and H4. Results are given in Tables 3 and 4. The *t*-tests indicated that five of the morphological traits differed significantly between the two groups of individuals (Table 3), but these differences were weak, because only a very small proportion of the total variation in these five traits was related to the possession of one haplotype or the other (R^2 = 0.017–0.048). An overall test of these morphological differences between individuals with the two haplotypes was also performed with a meta-analysis in which the probabilities from the individual significance tests for the nine morphological variables were combined (Sokal & Rohlf 1995), and the result was highly significant (Table 3). In general, there was some tendency for leaves of trees with haplotype H4 to be longer, wider, with a longer petiole and fewer aristae, compared to those with haplotype H1 (Table 3). This is the same direction of the differentiation in these five traits between typical individuals of *Q. laurina*

and *Q. affinis* (González-Rodríguez *et al*. 2004). These results suggest a weak but still detectable association of haplotype H4 with the morphology of *Q. laurina*, and of haplotype H1 with the morphology of *Q. affinis*. In the case of RAPD markers, the frequency of only one (B17–4) differed significantly between individuals with haplotypes H1 and H4 according to Fisher's exact test (Table 4). However, from these tests probabilities close to significance were obtained for several of the other RAPD markers, and the overall combined probability obtained with the meta-analysis was significant (Table 4). For all markers, the frequency differences between individuals with the two haplotypes were in the direction of associating haplotype H1 with *Q. affinis* and H4 with *Q. laurina*. For example, the estimated frequencies of marker B17–4 were 0.844 and 0.000 in representative populations of *Q. affinis* and *Q. laurina*, respectively (González-Rodríguez *et al*. 2004).

Discussion

Oaks species have seeds with low dispersal potential, and are characterized in general by low population cpDNA diversity, relatively high levels of overall diversity and high population differentiation (h_S = 0.025–0.183; h_T = 0.635–0.847; G_{ST} = 0.781–0.961 for eight European white oaks, Petit *et al*. 2002b). In this study we found comparatively higher levels of average within population variation (h_S = 0.299) and lower, but still considerable, among-population variation (G_{ST} = 0.499). Furthermore, the geographical distribution of haplotypes was not as clearly defined as in previous studies involving species in genus *Quercus* (Belahbib *et al*. 2001; Lumaret *et al*. 2002; Petit *et al*. 2002a,b). The pattern of spatial cpDNA variation in the *Q. affinis*–*Q. laurina* complex is characterized by the presence of two distantly related haplotypes (H1 and H4) at high frequency throughout the species' range, distributed in a rather mosaic fashion and co-occurring frequently, and at least four less common, most probably derived haplotypes that are geographically more restricted. Some phylogeographical structuring of the haplotype distributions was indicated because N_{ST} (0.566) was significantly higher than G_{ST}. There is no reason to think that dispersal in *Q. affinis* and *Q. laurina* may be more efficient than in any other oak species, so these patterns imply that processes allowing the two probably ancestral haplotypes H1 and H4 to reach a widespread distribution and causing extensive mixing of populations have been important in the historical dynamics of these species. Available palaeoenvironmental reconstructions (Metcalfe *et al*. 2000) suggest that the climatic changes during the late Pleistocene were not as drastic in Mexico as to reduce oak species to small populations isolated into a few refugia. These reconstructions indicate significant oscillations between cold–moist, cold–dry, cool–moist and warm–dry conditions in Northern and Central Mexico, although these changes were smaller in magnitude than in other parts of the northern hemisphere tropics and subtropics (Metcalfe *et al*. 2000). More significantly, important amounts of *Quercus* pollen are almost constantly present, with some fluctuations, in palynological records from several locations in Mexico extending back to 44 000 years BP (Lozano-García & Xelhuantzi-López 1997; Metcalfe *et al*. 2000). As a result of the climatic changes, oak species presumably experienced geographical displacements according to their particular ecological requirements, but forests in general have been probably present and widespread for a long time. The continued persistence of *Q. affinis* and *Q. laurina* populations is plausible, although with several latitudinal and elevational displacements and range fragmentations and expansions, which combined could have produced the observed extensive distribution of the probably ancestral haplotypes H1 and H4, as well as their frequent intermixing. The other four haplotypes, but particularly haplotypes H5 and H6, are probably more recently derived variants that have had comparatively less time to disperse. However, the disjunct distribution of haplotype H3 in two localities separated by 500 km (Fig. 1) seems difficult to explain as a result of seed dispersal. It is possible that the mutation characterizing this haplotype occurred independently in the two populations. Another tentative explanation is that this haplotype was acquired through hybridization with another related species.

The lack of a latitudinal or other obvious clinal pattern in diversity levels or haplotype composition of populations indicates further that *Q. affinis* and *Q. laurina* did not experience a recent restriction to one or a few particular subregions of their present range that later acted as sources of population expansions. South–north partitioning of haplotypic composition, as well as a northwards decrease in diversity levels within populations, has been observed in several temperate plant species (Démesure *et al*. 1996; Sewell *et al*. 1996; Soltis *et al*. 1997; King & Ferris 1998; Petit *et al*. 2002a,b). These patterns imply the existence of two different areas of glacial refugia situated at different latitudes in the case of some species (Sewell *et al*. 1996; Soltis *et al*. 1997), or more commonly a leading-edge effect (Hewitt 2001) during recolonization from southern refugia. This model predicts a successive decrease in the genetic diversity of populations along the direction of recolonization due to consecutive founder effects (Hewitt 2001). In general the pattern of geographical distribution of cpDNA variation in *Q. affinis* and *Q. laurina* is probably the accumulated result of several episodes of change possibly covering the whole Pleistocene, rather than predominantly the product of the last glacial cycle.

The geographical distribution of cpDNA variation in *Q. affinis* and *Q. laurina* also suggests that, in addition to recurrent migration, limited gene flow between populations, drift and founder effects have played a significant role. The

mosaic structure of haplotype distribution is evident, and contiguous populations often have very different genetic compositions and levels of variation (Fig. 1). This structure is reflected in the high average G_{ST} (0.381) between population pairs separated by a maximum of 100 km (Fig. 3), which are expected to be the genetically more similar. Furthermore, there is no clear increasing trend for average G_{ST} values until the 700–800 km distance class. The mosaic distribution of haplotypes at these comparatively shorter geographical distances seems to largely reflect the interplay of stochastic events such as drift and founder effects, produced by a general low average cpDNA exchange among populations combined with rare events of long distance dispersal. Isolation by distance and phylogeographical structuring become more evident between populations pairs situated at larger distance classes (Fig. 3).

Although haplotype variation in the studied populations appears to be almost completely independent of the morphological and nuclear genetic variation that demarcates the two oak species, there are some weak indications of an association of haplotype H1 with *Q. affinis*, and of haplotype H4 with *Q. laurina*. This result is in agreement with previous studies showing that haplotype sharing in many combinations is the rule rather than the exception among oak species, even when they are not closely related (Whittemore & Schaal 1991; Dumolin-Lapègue *et al*. 1999; Belahbib *et al*. 2001; Petit *et al*. 2002a,b). A scenario of hybridization and cytoplasmic introgression would explain haplotype sharing in the *Q. affinis*–*Q. laurina* complex although, unfortunately, we cannot evaluate the degree to which incomplete lineage sorting could be also responsible for haplotype sharing in these species, because with the present data it is impossible to infer the original extent of cpDNA differentiation between the two oaks. The divergence between H1 and H4 could have occurred coupled to the separation of *Q. affinis* and *Q. laurina*, or they could represent more ancient polymorphisms, each haplotype being initially only associated in part with a species. These issues will probably be clarified to some extent by analysing a larger sample of species and locations, so we can understand better the taxonomic and geographical distribution of cpDNA variation in Mexican red oaks and its past evolutionary dynamics.

Acknowledgements

Financial support for this study was provided by DGAPA, UNAM (grant IN212999 to K. Oyama) and CONACyT (grant 38550-V to K. Oyama). The authors thank J. L. Ibarra and L. Valencia for help with field collections and Nidia Pérez-Nasser for laboratory assistance. We appreciate the useful comments by L. E. Eguiarte and M. Lara throughout the development of this work, and the constructive contributions of three anonymous reviewers to the first version of the manuscript. A. González-Rodríguez was supported by a doctoral scholarship from CONACyT.

References

Beckstrom-Sternberg SM, Rieseberg LH, Doan K (1991) Gene lineage analysis in populations of *Helianthus niveus* and *H. petiolaris* (Asteraceae). *Plant Systematics and Evolution*, **175**, 125–138.

Belahbib N, Pemonge MH, Ouassou A, Sbay H, Kremer A, Petit RJ (2001) Frequent cytoplasmic exchanges between oak species that are not closely related: *Quercus suber* and *Q. ilex* in Morocco. *Molecular Ecology*, **10**, 2003–2012.

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

Burban C, Petit RJ, Carcreff E, Jactel H (1999) Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudii* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Molecular Ecology*, **8**, 1593–1602.

Butzer KW, Butzer EK (1997) The 'natural' vegetation of the Mexican Bajío: archival documentation of a 16th-century savanna environment. *Quaternary International*, **43**, 161–172.

Cannon CH, Manos PS (2003) Phylogeography of the Southeast Asian stone oaks (*Lithocarpus*). *Journal of Biogeography*, **30**, 211–226.

Cavers S, Navarro C, Lowe AJ (2003) Chloroplast DNA phylogeography reveals colonization history of a Neotropical tree, *Cedrela odorata* L., in Mesoamerica. *Molecular Ecology*, **12**, 1451–1560.

Conserva ME, Byrne R (2002) Late Holocene vegetation change in the Sierra Madre Oriental of Central Mexico. *Quaternary Research*, **58**, 122–129.

Démesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution*, **50**, 2515–2520.

Démesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*, **6**, 393–397.

Dumolin-Lapègue S, Kremer A, Petit RJ (1999) Are chloroplast and mitochondrial DNA variation species independent in oaks? *Evolution*, **53**, 1406–1413.

Ennos RA, Sinclair WT, Hu X-S, Langdon A (1999) Using organelle markers to elucidate the history, ecology and evolution of plant populations. In: *Molecular Systematics and Plant Evolution* (eds Hollingsworth PM, Bateman RM, Gornall RJ), pp. 1–19. Taylor & Francis Ltd, London.

Flenley JR (1998) Tropical forests under the climates of the last 30 000 years. *Climatic Change*, **39**, 177–197.

González-Rodríguez A, Arias DM, Valencia S, Oyama K (2004) Morphological and RAPD analysis of hybridization between *Quercus affinis* and *Q. laurina* (Fagaceae), two Mexican red oaks. *American Journal of Botany*, **91**, 401–409.

Graham A (1999) The Tertiary history of the northern temperate element in the northern Latin America biota. *American Journal of Botany*, **86**, 32–38.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography — or seeing genes in space and time. *Molecular Ecology*, **10**, 537–549.

Holman JE, Hughes JM, Fensham RJ (2003) A morphological cline in *Eucalyptus*: a genetic perspective. *Molecular Ecology*, **12**, 3013–3025.

King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology*, **7**, 1151–1161.

Lefort F, Douglas GC (1999) An efficient micro-method of DNA isolation from mature leaves of four hardwood tree species *Acer*, *Fraxinus*, *Prunus* and *Quercus*. *Annals of Forest Science*, **56**, 259–263.

Lozano-García MS, Xelhuantzi-López MS (1997) Some problems in the late Quaternary pollen records of Central Mexico: basins of Mexico and Zacapu. *Quaternary International*, **43**, 117–123.

Lumaret R, Mir C, Michaud H, Raynal V (2002) Phylogeographical variation of chloroplast DNA in holm oak (*Quercus ilex* L.). *Molecular Ecology*, **11**, 2327–2336.

Manos PS, Doyle JJ, Nixon KC (1999) Phylogeny, biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). *Molecular Phylogenetics and Evolution*, **12**, 333–349.

McCauley DE (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology and Evolution*, **10**, 198–202.

Metcalfe SE, O'Hara SL, Caballero M, Davis SJ (2000) Records of Late Pleistocene–Holocene climatic change in Mexico — a review. *Quaternary Science Reviews*, **19**, 699–721.

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

Nixon KC (1993) The genus *Quercus* in Mexico. In: *Biological Diversity of Mexico: Origins and Distribution* (eds Ramamoorthy TP, Bye R, Lot A, Fa J), pp. 397–420. Oxford University Press, New York.

Palmé AE, Semerikov V, Lascoux M (2003) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity*, **91**, 465–474.

Palmé AE, Vendramin GG (2002) Chloroplast DNA variation, postglacial recolonization and hybridization in hazel, *Corylus avellana*. *Molecular Ecology*, **11**, 1769–1779.

Petit RJ, Aguinagalde I, de Beaulieu J-L *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.

Petit RJ, Brewer S, Bordács S *et al.* (2002a) Identification of refugia and post-glacial colonization routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156**, 49–74.

Petit RJ, Csaikl UM, Bordács S *et al.* (2002b) Chloroplast DNA variation in European white oaks: phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management*, **156**, 5–26.

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144**, 1237–1245.

Rendell S, Ennos RA (2003) Chloroplast DNA diversity of the dioecious European tree *Ilex aquifolium* L. (English holly). *Molecular Ecology*, **12**, 2681–2688.

Rieseberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants*, **5**, 65–84.

Rieseberg LH, Whitton J, Randal Linder C (1996) Molecular marker incongruence in plant hybrid zones and phylogenetic trees. *Acta Botanica Neerlandica*, **45**, 243–262.

Rzedowski J (1981) *Vegetación de México*. Limusa, México DF, México.

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN* Version 2000*: a Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, Department of Ecology and Anthropology, University of Geneva, Switzerland.

Sewell MM, Parks CR, Chase MW (1996) Intraspecific DNA variation and biogeography of North American *Liriodendron* L. (Magnoliaceae). *Evolution*, **50**, 1147–1154.

Sokal RR, Rohlf FJ (1995) *Biometry, the Principles and Practice of Statistics in Biology Research*. Freeman, New York.

Soltis DE, Gitzendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, **206**, 353–373.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.

Valencia S (1994) *Contribución a la delimitación taxonómica de tres especies del género* Quercus *subgénero* Erythrobalanus. Tesis de Maestría, Facultad de Ciencias, Universidad Nacional Autónoma de México.

Whittemore AT, Schaal BA (1991) Interespecific gene flow in sympatric oaks. *Proceedings of the National Academy of Sciences, USA*, **88**, 2240–2544.

This study is part of the Ph.D. thesis of A. González-Rodríguez on the population genetics, hybridization processes and phylogeography of *Quecus affinis* and *Q. laurina*. J. F. Bain and J. L. Golden are interested in plant taxonomy, molecular systematics and phylogeography. The research of K. Oyama has focused on plant population ecology and genetics and plant–insect interactions.
