

Complex phylogeographic patterns indicate Central American origin of two widespread Mesoamerican *Quercus* (Fagaceae) species

Hernando Rodríguez-Correa¹ · Ken Oyama¹ · Mauricio Quesada¹ · Eric J. Fuchs² · Maura Quezada³ · Lilian Ferrufino⁴ · Susana Valencia-Ávalos⁵ · Alfredo Cascante-Marín² · Antonio González-Rodríguez⁶

Received: 6 October 2016 / Revised: 10 March 2017 / Accepted: 26 April 2017 / Published online: 15 May 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract The northern Neotropical region is characterized by a heterogeneous geological and climatic history. Recent studies have shown contrasting patterns regarding the role of geographic elements as barriers that could have determined phylogeographic structure in various species. Recently, the phylogeography and biogeography of *Quercus* species have been studied intensively, and the patterns observed so far suggest contrasting evolutionary histories for Neotropical species in comparison with their Holarctic relatives. The goal of this study was to describe the phylogeographic structure of two Neotropical oak species (*Quercus insignis* and *Quercus sapotifolia*) in the context of the geological and palaeoclimatic history of the northern Neotropics. Populations through the distribution range of both species were collected and characterized using nine chloroplast DNA microsatellite loci. Both oak species showed high levels of genetic diversity and strong

phylogeographic structure. The distribution of genetic variation in *Q. insignis* suggested an influence of two major barriers, the Isthmus of Tehuantepec and the Nicaraguan Depression, while *Q. sapotifolia* exhibited a genetic structure defined by the heterogeneity of the Chortis highlands. The haplotype networks of both species indicated complex histories, suggesting that colonization from the Sierra Madre de Chiapas to central Mexico and from the north of the Nicaraguan Depression to the Costa Rican mountains may have occurred during different stages, and apparently more than one time. In conclusion, the phylogeographic structure of Neotropical oak species seems to be defined by a combination of geological and climatic events.

Keywords Neotropical trees · Middle America · Palaeodistribution · Phylogeography · Historical demography

Communicated by A. Kremer

Electronic supplementary material The online version of this article (doi:10.1007/s11295-017-1147-7) contains supplementary material, which is available to authorized users.

✉ Antonio González-Rodríguez
agrodrig@cieco.unam.mx

¹ Escuela Nacional de Estudios Superiores Unidad Morelia, Universidad Nacional Autónoma de México, Antigua Carretera a Pátzcuaro No. 8701, Col. Ex-Hacienda de San José de la Huerta, 58190 Morelia, Michoacán, Mexico

² Escuela de Biología, Universidad de Costa Rica, San Pedro Montes de Oca, San José 11501-2060, Costa Rica

³ Herbario de San Carlos de Guatemala, Universidad San Carlos de Guatemala, Avenida la Reforma 0-63, Zona 10, 01010 Guatemala, Guatemala

⁴ Laboratorio de Histología Vegetal y Etnobotánica, Universidad Nacional Autónoma de Honduras, Edificio J1 (CB), Boulevard Suyapa, Tegucigalpa, Honduras

⁵ Herbario de la Facultad de Ciencias (FCME), Departamento de Biología Comparada, Universidad Nacional Autónoma de México, Avenida Universidad 3000, 04510 Coyoacán, Distrito Federal, Mexico

⁶ Laboratorio de Genética de la Conservación, Instituto de Investigaciones en Ecosistemas y Sustentabilidad, Universidad Nacional Autónoma de México, Antigua Carretera a Pátzcuaro No. 8701, Col. Ex-Hacienda de San José de la Huerta, 58190 Morelia, Michoacán, Mexico

Introduction

The northern Neotropical region comprises southern Mexico and Central America. This region is important for the study of evolutionary processes such as migration, extinction, and diversification of different organismal groups, in the context of the geological and climatic heterogeneity that characterizes it (Gutiérrez-García and Vázquez-Domínguez 2013). Even though consensual patterns have not emerged yet, several studies coincide on the importance of geological elements such as the Isthmus of Tehuantepec (IT), the Polochic-Motagua-Jocotán fault system (PMJ), the Nicaraguan Depression (ND), and the Isthmus of Panama (IP) as promoters of phylogeographic structure in different organisms (see Table 1 in Gutiérrez-García and Vázquez-Domínguez 2013), and also in the geographical organization of entire lineages such as *Quercus* (Rodríguez-Correa et al. 2015) and ferns (Ramírez-Barahona and Luna-Vega 2015). On the other hand, the lack of phylogeographic structure related to these geological elements also has been observed mainly (but not exclusively) in lowland species (see Table 1 in Gutiérrez-García and Vázquez-Domínguez 2013).

More interesting is the fact that even when the phylogeographic structure of the species appears to be determined by these barriers, it seems like processes of intra-specific divergence took place at different time periods, suggesting differential responses to the geological dynamics of the region, and therefore, lineage-specific evolutionary histories (Ornelas et al. 2013). Historical environmental fluctuations have also been considered important in structuring genetic diversity of the Neotropical species. However, available evidence (reviewed in Ramírez-Barahona and Eguiarte 2013) indicates differential responses to climate changes according to the distribution, ecology, and phylogenetic relationships of species. Therefore, the identification of common patterns and how these can be associated to the species' characteristics requires further studies dealing with codistributed species and the integration of palaeoclimatic, ecological, and molecular data (Ramírez-Barahona and Eguiarte 2013).

In northern latitudes, several *Quercus* species have been studied using phylogeographic approaches (Dumolin-Lapègue et al. 1997; Fineschi et al. 2002; Cottrell et al. 2002; Csaikl et al. 2002; Olalde et al. 2002; Petit et al. 2002a, b; Grivet et al. 2006; López de Heredia et al. 2007; Chen et al. 2012; Liu et al. 2013; Alexander and Woeste 2014) and the results obtained have allowed understanding the historical migration processes related to climatic fluctuations during the last glacial cycle, particularly in the European continent. Interestingly, oak species situated in the Neotropics exhibit contrasting patterns with respect to their Nearctic and Palearctic relatives. For example, Mexican oak species show

higher levels of within-population variation and lower among population differentiation than northern latitude species (Tovar-Sánchez et al. 2008; Ramos-Ortiz et al. 2016), probably because at lower latitudes, the effects of glaciations were less severe and oak populations did not isolate into small refugia but remained large and maintained relatively stable geographic ranges (González-Rodríguez et al. 2004; Cavender-Bares et al. 2011). In contrast, it seems like the most important effects of the Pleistocene climatic changes on oak species in the Neotropics were related to changes in their altitudinal distribution (Hooghiemstra and van der Hammen 2004; Rodríguez-Correa 2015).

Recently, it has been suggested that areas such as the IT, the PMJ, the ND, and the IP have acted as important barriers to the dispersal of oak species, influencing biogeographic patterns, species diversity, and niche divergence processes (Torres-Miranda et al. 2013; Rodríguez-Correa et al. 2015). However, a few oak species (*Quercus corrugata*, *Quercus insignis*, *Quercus oleoides*, and *Quercus sapotifolia*, and perhaps also *Quercus benthamii*, *Quercus cortesii*, and *Quercus salicifolia*) (Valencia-A 2004; Morales 2010) have geographical distributions that extend across almost all these barriers (except the IP), eliciting questions about the timing and process of dispersal and the influence of species' traits and ecological niche on dispersal success. Also, an evident hypothesis is that the mentioned barriers probably correspond to phylogeographic breaks within these species. However, the lowland *Q. oleoides* (distributed from northeastern Mexico to northern Costa Rica) is the only species so far analyzed (Cavender-Bares et al. 2011, 2015), using traditional genetic markers and next-generation sequence data. No genetic discontinuities across the IT and the PMJ were found, but there was strong differentiation across the ND, estimated at about 1.9 my, implicating the formation of the Nicaraguan Depression and associated volcanic activity as the cause of this pattern (Cavender-Bares et al. 2015).

In this study, we used chloroplast DNA microsatellite data to reconstruct phylogeographic patterns of two Neotropical oak species (*Q. insignis* and *Q. sapotifolia*) distributed from southern Mexico to northern Panama. Both species are cataloged as threatened in Mexico (*Q. insignis* as critically endangered and *Q. sapotifolia* as vulnerable; González-Espinosa et al. 2011). *Q. insignis* is a white oak (section *Quercus*) and *Q. sapotifolia* is a red oak (section *Lobatae*), but both have roughly similar geographical distributions and are mainly montane species. The specific questions for the study were as follows: (i) What are the phylogeographic patterns in the two oak species, *Q. insignis* and *Q. sapotifolia*? (ii) How did the major geological features of the northern Neotropics influence the evolutionary history of these two species? (iii) Did recent Pleistocene climatic changes affect the geographic distribution and connectivity patterns in these oaks across their range?

Methods

Studied species and sampled populations

The two oak species were sampled across their whole distribution range (Fig. 1). *Q. insignis* M. Martens & Galeotti is found in the Mexican states of Jalisco, Guerrero, Oaxaca, Chiapas, and Veracruz, and in Belize, Guatemala, Honduras, Nicaragua, Costa Rica, and Panama in Central America. The altitudinal range of *Q. insignis* varies from approximately 1500 to 2000 m (Valencia-A 2004). The species is emblematic for producing very large acorns (7–8 cm in diameter; Montes-Hernández and López-Herrera 2013), probably the largest of all oaks species. *Q. sapotifolia* Liebm. is distributed in the states of Hidalgo, Oaxaca, Chiapas, and Veracruz in Mexico, and in Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica, and Panama in Central America, with an altitudinal range between 250 and 2000 m (Valencia-A 2004). This species is characterized by having considerable morphological variation across its distribution. Despite their wide ranges, both species are rare and populations are difficult to locate because of their low individual density, which was also

reflected in the dissimilar population sample sizes for both species (see “Results”). In total, 13 populations of *Q. insignis* and 15 populations of *Q. sapotifolia* were sampled (Table 1).

DNA isolation and microsatellite amplification

Fresh leaf tissue was collected in the field and stored in silica gel until processing in the laboratory. Dried tissue was disrupted using liquid nitrogen, and DNA isolation was carried out using the DNeasy plant mini kit (Qiagen), following the manufacturer’s instructions. A set of nine chloroplast DNA (cpDNA) microsatellite loci (cmcs2, cmcs3, cmcs4, cmcs5, cmcs6, cmcs7, cmcs10, cmcs12, and cmcs14) previously characterized for Fagaceae species by Sebastiani et al. (2004) was tested for polymorphism in both *Q. insignis* and *Q. sapotifolia*. Two groups of primers (cmcs3, cmcs4, cmcs5, and cmcs6 in the first group and cmcs2, cmcs7, cmcs10, cmcs12, and cmcs14 in the second group), with different combinations of fluorescent dyes and expected size were employed in multiplexed polymerase chain reactions (PCRs). These were performed using the multiplex PCR kit (Qiagen) with a final volume of 5 μ L containing 1 \times multiplex

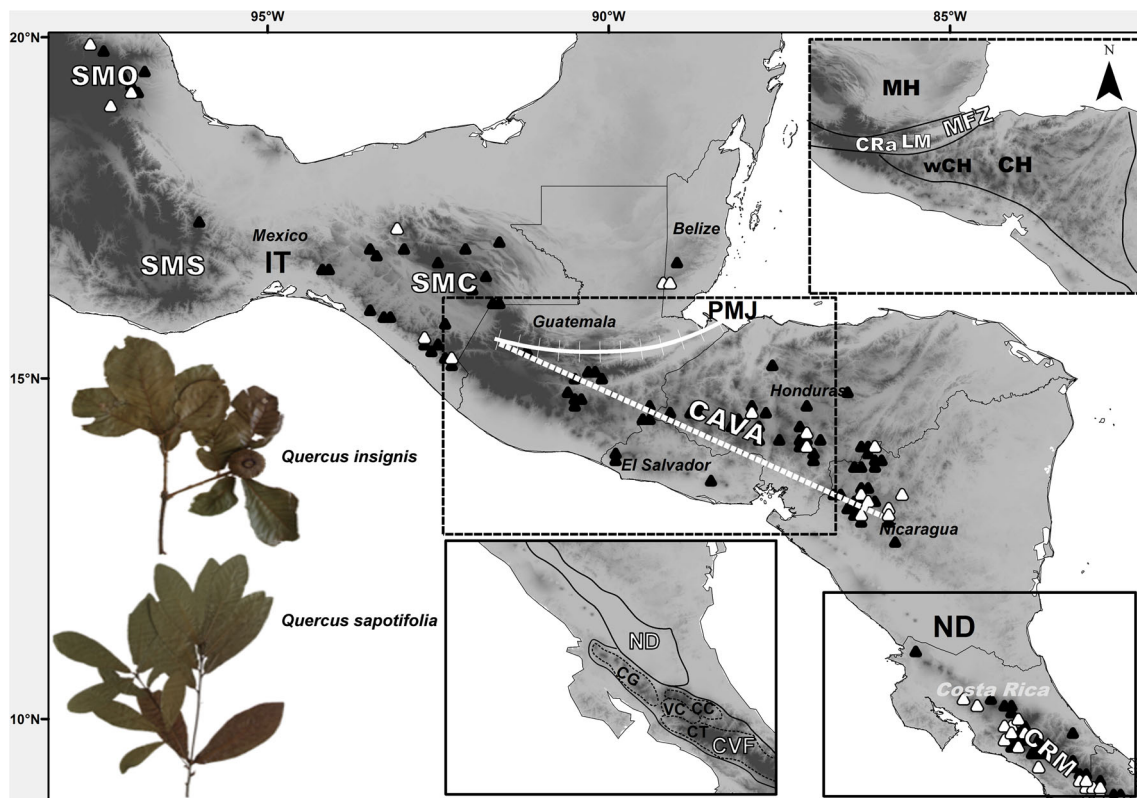


Fig. 1 *Quercus insignis* and *Q. sapotifolia* distribution represented by white and black triangles, respectively. Elevation is represented in a grayscale from lowlands (light gray) to highlands (black). SMO Sierra Madre Oriental, SMS Sierra Madre del Sur, IT Isthmus of Tehuantepec, SMC Sierra Madre de Chiapas, PMJ Polochic-Motagua-Jocotán fault system, CAVA (white dotted line) Central America volcanic arc, ND Nicaraguan Depression, CRM Costa Rica mountains. Details of Middle

Mesoamerica (MM) and Lower Mesoamerica (LM) are represented inside dotted and solid rectangles as follows: MH Maya highlands, CRA Chuacús range, LM Las Minas range, PMJ (white crossed line) Polochic-Motagua-Jocotán fault system, wCH western rifted Chortis highlands, CH central Chortis highlands, CG Cordillera de Guanacaste, VC Valle Central, CC Cordillera Central, CVF (solid line) Chorotega volcanic front

Table 1 Geographic information summary for *Quercus insignis* and *Q. sapotifolia* populations and genetic diversity and historical demography estimates

	ID	Longitude	Latitude	Country/ region	Elevation (m.a.s.l)	<i>n</i>	Total haplotypes	haplotype richness (rarefacted)	<i>h_S</i>	Fu's <i>F_S</i>	<i>p</i> value
<i>Quercus insignis</i>	i1	-82.93	9.4	CR	725	5	2	1.60	0.40	0.2	0.94
	i2	-83.98	9.88	CR	1715	7	4	2.46	0.80	0.5	0.15
	i3	-84.8	10.3	CR	1401	3	3	3.00	1.00	0.66	0.62
	i4	-84.01	9.95	CR	1372	5	2	1.60	0.40	0.68	0.94
	i5	-83.81	9.54	CR	2451	10	5	2.50	0.82	0.23	0.07
	LM					30	12	8.56	0.91	0.22	0.58
	i6	-87.95	14.53	HON	1638	7	3	1.86	0.52	0.18	0.76
	i7	-87.81	14.76	HON	1505	7	6	2.86	0.95	0.13	0.65
	i8	-87.95	14.76	HON	846	6	3	2.00	0.60	0.25	0.51
	i9	-87.12	14.24	HON	1754	5	2	1.90	0.60	0.37	0.4
	i10	-87.07	14.04	HON	1507	6	4	2.45	0.80	0.27	0.21
	MM					31	16	10.07	0.93	-1.89	0.25
	i11	-97.04	19.18	MEX	1408	7	2	1.71	0.47	0.72	0.93
	i12	-96.94	17.43	MEX	802	5	2	1.60	0.40	0.68	0.49
	i13	-96.96	19.37	MEX	1059	3	2	2.00	0.66	1	0.56
i14*	-92.9	16.8	MEX	1108	1	1	NA	NA	NA	NA	
UM					15	2	2	0.42	4.14	0.96	
<i>Quercus sapotifolia</i>	s1	-83.96	9.7	CR	1671	5	2	1.60	0.40	1.68	0.76
	s2	-84.25	10.14	CR	2276	9	3	2.05	0.63	1.85	0.48
	LM					14	5	5.00	0.79	1.35	0.77
	s3	-86.79	13.39	HON	1938	11	3	1.76	0.47	-0.65	0.1
	s4	-87.13	13.9	HON	1832	9	3	1.67	0.41	-1.08	0.07
	s5	-87.09	13.98	HON	1317	6	3	2.00	0.60	-0.85	0.08
	s6	-87.07	14.04	HON	1122	8	7	2.89	0.96	-2.62	0.05
	s7	-87.06	14.22	HON	2004	8	5	2.59	0.85	-1.74	0.06
	s8	-88.21	14.27	GUA	1298	10	3	1.60	0.37	0.286	1.45
	s9	-96.46	14.69	GUA	1468	9	4	1.99	0.58	2.21	0.86
	s10	-91.88	14.83	GUA	1524	14	2	1.54	0.36	0.65	0.45
	s11	-90.98	15.49	GUA	1507	18	2	1.44	0.29	0.46	0.38
	s12	-90.98	15.49	GUA	1914	3	1	1.00	0.00	NA	NA
	s13	-91.7	16.11	MEX	1501	6	2	1.50	0.33	1.6	0.72
	s14	-94.18	16.89	MEX	1908	3	2	2.00	0.66	0.38	0.2
MM					105	22	7.82	0.88	-0.65	0.47	
s15	-97.04	19.18	MEX	1178	20	7	5.64	0.64	3.02	0.91	
UM					20	7	5.64	0.64	3.02	0.91	

n sample size, *h_S* within-population genetic diversity, *UM* upper Mesoamerica, *MM* middle Mesoamerica, *LM* lower Mesoamerica, *CR* Costa Rica, *HON* Honduras, *GUA* Guatemala, *MEX* Mexico

PCR master mix, 0.25 mM of each primer, dH₂O and 20 ng of template DNA. Amplification was performed using an initial denaturation step for 15 min at 95 °C, followed by 35 cycles of 30 s at 95 °C, 1.5 min at 55 °C, 1 min at 72 °C, and a final extension step for 30 min at 60 °C. PCR products were run in a ABI-PRISM 3300 Avant sequencer (Applied Biosystems), and allele size was determined using a standard (GeneScan-600 LIZ) with the Peak Scanner program version 2.0 (Applied Biosystem).

Genetic analysis

Genetic diversity and genetic structure

Each unique combination of size variants for the nine loci was defined as a different haplotype. Genetic diversity was described in terms of rarefacted haplotype richness (AR), haplotype diversity with unordered alleles (*h* sensu Pons and Petit

1996), nonstandardized haplotype diversity with ordered alleles (v sensu Pons and Petit 1996), and the pairwise genetic distance among individuals within a population under a stepwise mutation model (D^2_{SH} ; Goldstein et al. 1995) using SPAGeDi version 1.1 (Hardy and Vekemans 2002). These estimators were calculated both at the population and region levels. The regions were defined as follows: (1) populations located to the north/east of the IT belonged to the Upper Mesoamerica (UM) region, (2) populations found between the IT and the ND belonged to the Middle Mesoamerica (MM) region, and (3) populations to the south of the ND formed the lower Mesoamerica (LM) group. Population i14 of *Q. insignis* (from Chiapas, Mexico) only had one individual and was considered for graphical purposes only, but it was excluded from population structure and historical demography analysis.

To infer haplotype relationships, a minimum spanning network was computed using Network version 4.6 (available at www.fluxus-engineering.com) with the median-joining method (Bandelt et al. 1999) and a maximum parsimony search (Polzin and Daneschmand 2003). Genetic differentiation and phylogeographic structure were analyzed by calculating G_{ST} and N_{ST} (a G_{ST} analogue which takes into account the genetic distances among haplotypes) with SPAGeDi version 1.1 (Hardy and Vekemans 2002). A permutation test implemented in this program was used to test whether the value of N_{ST} was significantly greater than the value of G_{ST} , which indicates that there is phylogeographic structure in the populations (Pons and Petit 1996). The F_S statistic (Fu 1997) was estimated for each population and region in order to test for events of demographic expansion. Calculations were carried out using Arlequin version 3.5 (Excoffier et al. 2005) and codifying cpSSR data in binary form as suggested by Navascués et al. (2009).

The partitioning of the genetic variation among regions, among populations within regions and within populations was described using a hierarchical analysis of molecular variance (AMOVA) implemented in Arlequin version 3.5 (Excoffier et al. 2005). The grouping of populations in regions was the same as explained above. AMOVA was calculated using both F_{ST} (based on the infinite alleles mutation model, IAM) and R_{ST} (based on the stepwise mutation model, SMM) with 10,000 permutations to determine significance of the estimates (Excoffier et al. 2005). To further understand the geographic distribution of genetic variation, a spatial analysis of molecular variance was implemented using SAMOVA version 1.0 (Dupanloup et al. 2002). The number of groups was evaluated using values of K from 2 to 10 so that the genetic differentiation among the groups (Φ_{CT}) was maximized. Consistence between different runs was determined by repeating twice every K-value evaluated.

Also, the geographical location of the most important genetic discontinuities among populations was determined using

the Monmonier's maximum difference algorithm implemented in BARRIER ver. 2.2 (Manni et al. 2004). To provide bootstrap support to the observed genetic barriers, a set of 100 pairwise matrices of average square genetic distance (ASD; Goldstein et al. 1995; Slatkin 1995) among populations was generated by manual resampling of individuals in the original matrix. Finally, genetic diversity indexes (AR , h , v , D^2_{SH}) were correlated with the altitude and latitude of the sampled populations and also compared among regions using a Kruskal-Wallis rank sum test in order to determine the existence of geographical patterns in the distribution of genetic diversity. These tests were implemented in R ver. 3.0.2.

Ecological niche modelling

Ecological niche models (ENMs) were estimated using a maximum entropy approach implemented in MAXENT version 3.3.3a (Phillips et al. 2006) in order to determine the climatically suitable areas for both *Q. insignis* and *Q. sapotifolia* during the Last Interglacial (LIG; ~120 ka BP), the Last Glacial Maximum (LGM; ~21 ka BP), and the present-day (PD) periods using the most common and therefore comparable climatic layers reported in previous phylogeographic studies for Neotropical species (e.g., Ramírez-Barahona and Eguiarte 2014; see below). In order to build the models, occurrence data were downloaded from a public repository (Global Biodiversity Information Facility; <http://www.gbif.org>). The data were filtered using as reference the reported distribution and altitudinal range of the species (Valencia-A 2004).

Climatic information used to run the models was based on the 19 bioclimatic variables proposed by Hijmans et al. (2005; www.worldclim.com) at a spatial resolution of 30 arc-seconds (~1 km). However, in order to avoid correlation among variables, a correlation matrix was calculated, and from each pair of highly correlated variables ($r > 0.7$), the more specific variable was discarded. To decrease possible effects of spatial autocorrelation due to the aggregation of records, we only used points separated by more than 0.1 decimal degrees from the nearest neighbor. ENMs were estimated after 100 replicates using the bootstrap resampling method in MAXENT. Thirty percent of the presence records were used to calculate estimators of quality, and the remaining 70% was used to run the models. The ENMs were projected onto past climate scenarios (LGM and LIG), with two models for the LGM, the Community Climate System Model (CCSM; Collins et al. 2004) and the Model for Interdisciplinary Research on Climate (MIROC; Hasumi and Emori 2004). Both models simulate climatic conditions as they are calculated to have been for the LGM, with a stronger temperature decrease assumed in CCSM compared to MIROC (Otto-Bliesner et al. 2006). The original LGM data were downloaded from the

PMIP2 website (<http://www.pmip2.cnrs-gif.fr/>) (Braconnot et al. 2007) and LIG from www.worldclim.com.

Model quality was determined with a threshold-independent method, the area under the receiver operating characteristic (ROC) curve analysis (AUC; Fielding and Bell 1997), and the fixed cumulative value 1 logistic threshold (FCV) was applied in order to obtain binomial output. FCV was used as it has been observed as an appropriated threshold value in other Neotropical species analyzed with palaeoecological methods involving ENM (Rodríguez-Correa 2015). Finally, in the models, we determined the areas where the distribution of both species remained stable during the LIG, LGM, and PD periods. The location of the main phylogeographic breaks obtained with the molecular data was compared with the discontinuities in climatically suitable areas obtained in the MAXENT models to evaluate the possible role of present-day and historical climatic conditions in structuring the genetic diversity of the two species.

Results

Genetic diversity, genetic structure, and historical demography

A total of 76 individuals from 14 populations and 139 individuals from 15 populations were respectively characterized for *Q. insignis* and *Q. sapotifolia* (Table 1) (Fig. 1). For *Q. insignis*, the total number of haplotypes was 28, 14 of them corresponding to singletons. The number of haplotypes per population ranged from 1 to 6. At the region level, the number of haplotypes was as follows: 2 haplotypes were observed in UM, 16 haplotypes (9 singletons) in MM, and 12 haplotypes (5 singletons) in LM. Within-population haplotype diversity (h_S) ranged between 0.4 and 1. Mean (SE) h_S and total diversity (h_T) were 0.65 (0.05) and 0.95 (0.03), respectively. Genetic differentiation among populations was 0.315 (0.05) for unordered alleles (G_{ST}) and 0.716 (0.07) for ordered alleles (N_{ST}). G_{ST} and N_{ST} values were significantly different ($P < 0.001$), indicating the presence of phylogeographic structure. Fu's F_S values for all populations and geographic regions were nonsignificant ($P > 0.02$; Table 1).

Q. sapotifolia exhibited 34 haplotypes, with 17 singletons. The number of haplotypes per population ranged from 1 to 7. At the regional level, haplotype richness was 7 (4 singletons) in UM, 22 (11 singletons) in MM, and 5 (2 singletons) in LM. Within-population genetic diversity (h_S) ranged between 0.4 and 0.95. Mean (SE) h_S and total haplotype diversity (h_T) were 0.51 (0.06) and 0.96 (0.01), respectively. Genetic differentiation among populations was 0.471 (0.06) for unordered alleles (G_{ST}) and 0.752 (0.05) for ordered alleles (N_{ST}). A significant difference between G_{ST} and N_{ST} ($P < 0.001$) was observed, indicating the presence of phylogeographic structure. Fu's F_S

values for all populations and geographic regions were not significant ($P > 0.02$; Table 1).

The partitioning of genetic variation in *Q. insignis* according to the AMOVA (Table 2) showed that for both F_{ST} and R_{ST} , the differences among regions explained a considerable amount of variation (20 and 30%, respectively; $P < 0.001$). Meanwhile, most of the genetic variation was distributed among populations within regions under both mutation models, 42.22% ($P < 0.001$) for IAM and 45.18% ($P < 0.001$) for SMM. The haplotype network (Fig. 2a) indicated a complex scenario. Despite the presence of haplotypes with high frequencies such as H27 (found in 15 individuals), there is not a widely distributed haplotype through the whole *Q. insignis* distribution. The distribution of haplotypes in MM and SM suggested that several haplotypes are more closely related between regions than within regions (e.g., H10 and H5, H6 and H7). In *Q. sapotifolia*, partitioning of the genetic variation according to the AMOVA results (Table 2) showed that for F_{ST} , differences among regions explained 21% ($P < 0.001$) of the variation. However, for the case of R_{ST} , the differentiation value was very small and nonsignificant (2%; $P > 0.05$). The observed differences between F_{ST} and R_{ST} should be related to the effect of high mutation rates and the mutation models to estimate population structure (for a complete discussion see Rousset (1996)). On the other hand, most of the genetic variation was distributed among populations within regions under both mutation models (68%; $P < 0.001$, for IAM and 64%; $p < 0.001$ for SMM). The haplotype network (Fig. 3a) showed that the majority of haplotypes are found in the MM region. Interestingly, the haplotypes found within the UM region seem to belong to two lineages independently derived from haplotypes H34 and H3, and a similar pattern seems to be true for the haplotypes present in the LM region (Fig. 3a).

Geographic distribution of the genetic variation

The *Q. insignis* results of the SAMOVA analysis (Fig. 2b) indicated that the maximum value of genetic differentiation among groups of populations corresponded to four groups ($\Phi_{CT} = 0.72$; $P < 0.001$). According to this result, the first group was formed by three populations from Veracruz, Mexico (i13, i12, i11) and three populations from Honduras (i10, i8, and i6), thus including populations from both UM and MM. A second group was constituted by two populations from Honduras (MM; i9 and i7) and two populations from Costa Rica (LM; i4 and i2). A third group included two populations from Costa Rica (i5 and i3), and the fourth group included only the southernmost population (i1). *Q. sapotifolia* SAMOVA results (Fig. 3b) showed that genetic differentiation among groups is maximum when $K = 5$ ($\Phi_{CT} = 0.72$; $P < 0.001$). According to this result, the grouping of populations was as follows: group 1 formed by UM and

Table 2 Hierarchical analysis of molecular variance (AMOVA) using F_{ST} and R_{ST} for *Quercus insignis* and *Q. sapotifolia*

		Source of variation	d.f.	s.s.	Variance components	Percentage of variation		
<i>Quercus insignis</i>	F_{ST}	Among groups	2	42.33	0.6	20.57	$\Phi_{CT} = 0.20^*$	
		Among populations within groups	10	83.24	1.24	42.22	$\Phi_{SC} = 0.53^*$	
		Within populations	63	69.19	1.09	37.2	$\Phi_{ST} = 0.62^*$	
		Total	75	194.77	2.95			
		R_{ST}	Among groups	2	550.86	9.01	30.41	$\Phi_{CT} = 0.30^*$
	Among populations within groups		10	848.43	13.38	45.18	$\Phi_{SC} = 0.64^*$	
	Within populations		63	455.6	7.23	24.41	$\Phi_{ST} = 0.75^*$	
	Total		75	1854.67	29.62			
	<i>Quercus sapotifolia</i>		F_{ST}	Among groups	2	65.96	0.52	21.21
		Among populations within groups		12	142.36	1.33	53.93	$\Phi_{SC} = 0.68^*$
Within populations		124		76.58	0.61	24.87	$\Phi_{ST} = 0.75^*$	
Total		138		284.91	2.48			
R_{ST}		Among groups		2	588.97	0.87	2.79	$\Phi_{CT} = 0.02$
		Among populations within groups	12	2099.73	19.54	62.39	$\Phi_{SC} = 0.64^*$	
		Within populations	124	1352.21	10.9	34.82	$\Phi_{ST} = 0.65^*$	
		Total	138	4040.92	31.32			

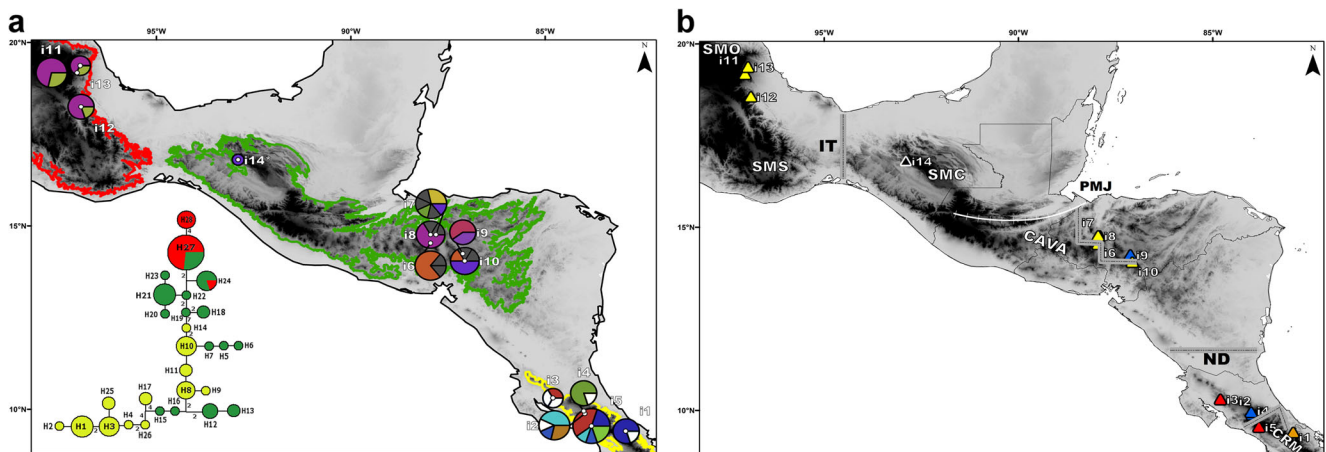


Fig. 2 a Haplotype distribution and haplotype network for *Quercus insignis*. In the network, circle diameter is proportional to haplotype frequency. On the map, circle diameter is proportional to sample size. The Northern Neotropics area is divided in upper Mesoamerica (UM; red line), middle Mesoamerica (MM; green line), and lower Mesoamerica (LM; yellow line). The same colors are used in the haplotype network to indicate the presence of haplotypes in each region. On the map, haplotype composition was represented by different colors reflecting richness and shared haplotypes. Gray and white haplotypes correspond to unique haplotypes in MM and LM,

respectively. **b** Geographical location of the most important genetic discontinuities (gray lines) for *Q. insignis* populations according to the BARRIER analysis. Also, the different colors of the triangles indicate the grouping of the populations according to SAMOVA. SMO Sierra Madre Oriental, SMS Sierra Madre del Sur, IT Isthmus of Tehuantepec, SMC Sierra Madre de Chiapas, PMJ Polochic-Motagua-Jocotán fault system, CAVA Central America volcanic arc, ND Nicaraguan Depression, CRM Costa Rica mountains. Elevation is represented in a grayscale from lowlands (light gray) to highlands (black)

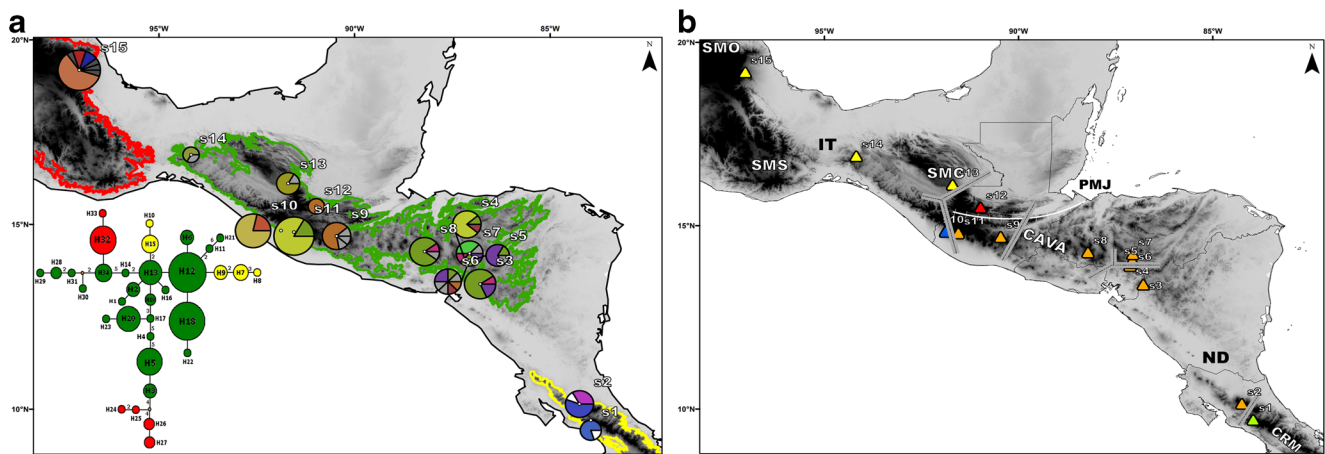


Fig. 3 **a** Haplotype distribution and haplotype network for *Quercus sapotifolia*. In the network, circle diameter is proportional to haplotype frequency. On the map, circle diameter is proportional to sample size. The Northern Neotropics area is divided in upper Mesoamerica (UM; red line), middle Mesoamerica (MM; green line), and lower Mesoamerica (LM; yellow line). The same colors are used in the haplotype network to indicate the presence of haplotypes in each region. On the map, haplotype composition was represented by different colors reflecting richness and shared-haplotypes. Dark gray, light gray, and white haplotypes correspond to unique haplotypes in UM, MM, and LM,

northern MM populations (s13, s14, and s15), groups 2 and 3 each formed by one population from MM (s10 and s12, respectively), group 4 was constituted by eight MM and one LM populations (s2, s3, s4, s5, s6, s7, s8, s9, and s11), and group five formed by population s1 (LM).

The analysis of genetic discontinuities along the *Q. insignis* distribution (Fig. 2b) suggests that the IT and the ND are important barriers for the species (94 and 95% bootstrap support). In Honduras, the two southernmost populations i7 and i10 were separated from the northern populations (95% bootstrap support). In Costa Rica, the heterogeneous composition of populations was indicated by the presence of barriers separating populations (95% bootstrap support) into three groups corresponding to northern, central, and southern Costa Rica (Fig. 2b). Finally, all the correlations of genetic diversity statistics with elevation and latitude were not significant ($P > 0.05$), as well as the comparisons among regions ($P > 0.05$). The analysis of *Q. sapotifolia* genetic discontinuities (Fig. 3b) showed several significant barriers that define the geographic configuration of the genetic differentiation. From north to south, the first barrier separates the Mexican populations from the remaining populations (100% bootstrap support), and it is located at the SMC close to the PMJ. Adjacent to this barrier and to population s10, a second barrier separates this population (s10) from the remaining groups (98% bootstrap support). At the middle portion of MM, a third barrier separates Guatemalan from Honduran populations (90% bootstrap support). In the Honduras territory, population s7 differentiates from

respectively. **b** Geographical location of the most important genetic discontinuities (gray lines) for *Q. sapotifolia* populations according to the BARRIER analysis. Also, the different colors of the triangles indicate the grouping of the populations according to SAMOVA. SMO Sierra Madre Oriental, SMS Sierra Madre del Sur, IT Isthmus of Tehuantepec, SMC Sierra Madre de Chiapas, PMJ Polochic-Motagua-Jocotán fault system, CAVA Central America volcanic arc, ND Nicaraguan Depression, CRM Costa Rica mountains. Elevation is represented in a grayscale from lowlands (light gray) to highlands (black)

the rest of the populations. LM populations were also separated in two groups (100% bootstrap support) similarly as for *Q. insignis*: northern and central Costa Rica. Correlations of genetic diversity measures with elevation and latitude as well as comparisons among geographic regions did not show any significant result ($P > 0.05$).

Ecological niche modelling

The ENMs of *Q. insignis* for the present-day period (Fig. 4a) showed a good performance according to the AUC value (0.991, SD 0.005). This model showed two main distribution areas, the first one located from the Sierra Madre de Chiapas (SMC) through the Central America volcanic arc (CAVA) in Guatemala, Honduras, and northern Nicaragua. The second distribution area is located in the Costa Rica mountains (CRM) (Fig. 4a). The northernmost part of the species' range (i.e., southern Mexico) is characterized by a fragmented distribution. ENMs also suggested an increase in the area available for *Q. insignis* during the LGM (Fig. 4b, c) and an important reduction in the available climatic niche during the LIG (Fig. 4d). The increase in available area during the LGM could have favored the contact or the proximity of populations across barriers such as the IT and ND, suggesting the possibility of intermittent gene flow across these barriers at some time periods. Finally, it can be observed that distribution areas that have remained stable are located in SMS, SMC, CAVA (including an important portion of the Nicaraguan southern region), and CRM (Fig. 4e).

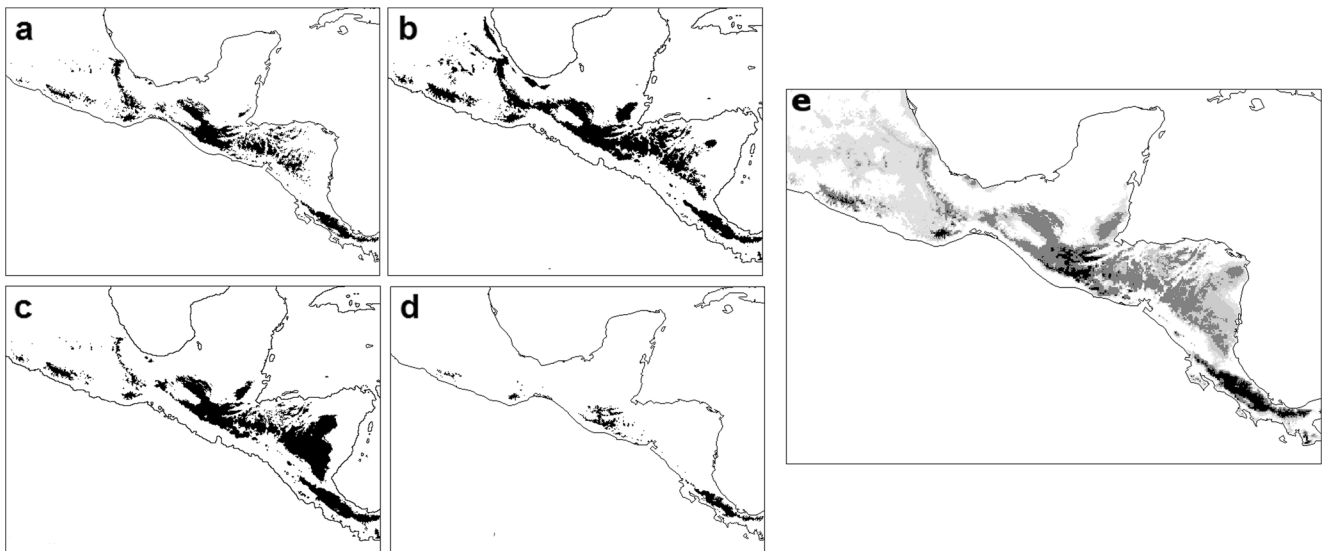


Fig. 4 **a** Present-day (1950–2000). **b, c** Last Glacial Maximum (~21 ka BP; CCSM and MIROC models, respectively). **d** Last Interglacial (~120 ka BP) ecological niche models for *Quercus insignis*. **e** *Q. insignis* stable areas of distribution since the Last Interglacial (~120 ka BP). Areas in black for **a–d** correspond to suitable climatic conditions for *Q. insignis*. In **e**, solid black corresponds to the areas

where the distribution coincides in the four models used (PD, CCSM, MIROC and LIG; see text for details). Dark gray corresponds to the areas where models for the PD and LGM coincide, and light gray corresponds to the areas where only one of the models predicted the presence of the species

Q. sapotifolia ENM for the present-day period (Fig. 5a) exhibited a good performance according to the AUC value (0.90, SD 0.001). This model predicted two main distribution areas as was also observed for *Q. insignis*. These areas are formed by (i) the MM region including the Sierra Madre de Chiapas (SMC) and the trans-isthmian mountains (CAVA) in Guatemala, Honduras, and northern Nicaragua, and (ii) the

CRM in LM (Fig. 5a). The first area showed a more continuous distribution than the observed in *Q. insignis*, particularly at Honduras and northern Nicaragua (Fig. 5a). The *Q. sapotifolia* distribution at UM is almost restricted to the southern SMS and SMO. For the LGM, both models (CCSM and MIROC) suggested an increase in the connectivity among areas across the IT region, but a gap between the

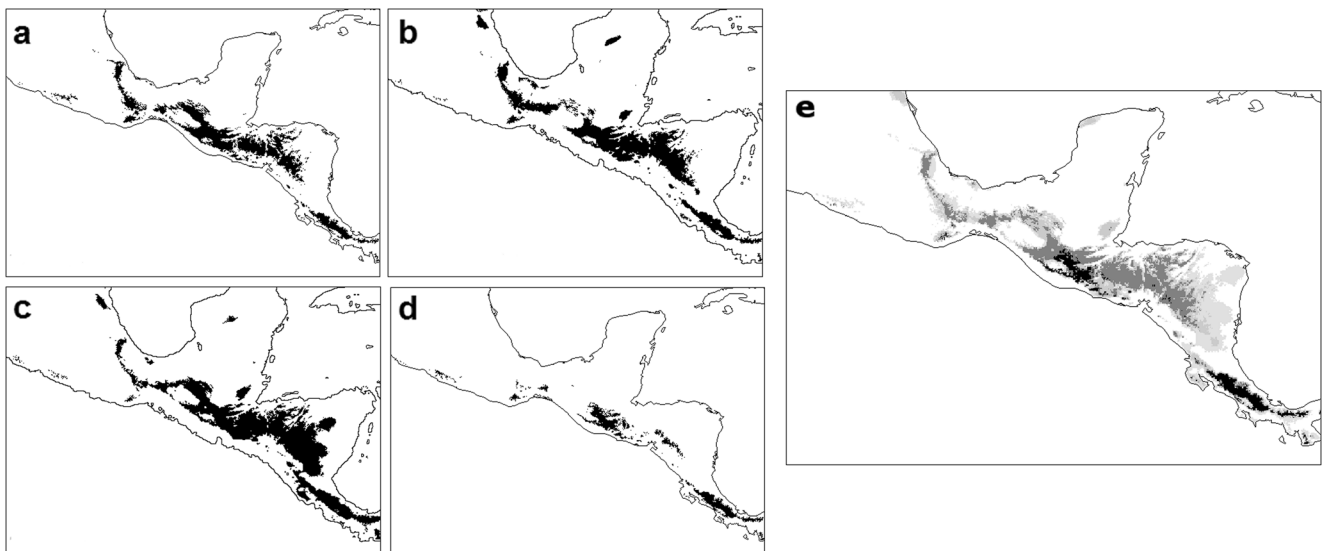


Fig. 5 **a** Present-day (1950–2000). **b, c** Last Glacial Maximum (~21 ka BP; CCSM and MIROC models, respectively). **d** Last Interglacial (~120 ka BP) ecological niche models for *Quercus sapotifolia*. **e** *Q. sapotifolia* stable areas of distribution since the Last Interglacial (~120 ka BP). Areas in black for **a–d** correspond to suitable climatic conditions for *Q. sapotifolia*. For **e**, solid black corresponds to the areas

where the distribution coincides in the four models used (PD, CCSM, MIROC, and LIG; see text for details). Dark gray corresponds to the areas where models for the PD and LGM coincide, and light gray corresponds to the areas where only one of the models predicted the presence of the species

SMS and CAVA at least in the CCSM model (Fig. 5b, c). At the ND, both models showed an important gap with a shallow connection between Nicaragua and northwestern Costa Rica according to the MIROC model (Fig. 5c). During the LIG, the climatic niche distribution is restricted to the SMS and northern CAVA, southern Honduras and northern Nicaragua portions of the CAVA and the CRM (Fig. 5d). When all models were considered together, it was observed that, as in *Q. insignis*, most stable areas were located between the SMS and the CAVA, as well as at the CRM (Fig. 5e). In contrast to the climatic niche extension in *Q. insignis*, *Q. sapotifolia* has more restricted suitable and stable areas at southern Nicaragua.

Discussion

Phylogeography in Neotropical oak species

Recent studies regarding the Central American biota have focused on the description of evolutionary processes such as speciation, extinction, and diversification of flora and fauna considering the intricate geologic history of this region, its habitat diversity, palaeoclimatic dynamics, and tectonic history (Gutiérrez-García and Vázquez-Domínguez 2013). Gutiérrez-García and Vázquez-Domínguez (2013) proposed the presence of at least three evolutionary groups in the northern Neotropics based on the coincidence of genetic differentiation patterns: the Mayan group (located between the IT and the PMJ), the Mid-Central group (between the PMJ and the Hess escarpment), and the Panamanian group (from the HE to the Andes). Within this area, other oak species studied have shown a strong differentiation between Costa Rica and Honduras associated to the formation of the ND (*Q. oleoides*; Cavender-Bares et al. 2011, 2015). However, little is known so far about the effects of past geological and climatic events on the population history of other oak species in the same area, but that differ from *Q. oleoides* in climatic niche and habitat.

Q. insignis and *Q. sapotifolia* exhibited higher haplotype richness (28 and 34 haplotypes, respectively) in comparison to higher latitude oak species that have been characterized with comparable chloroplast microsatellite molecular markers, such as the European white oak complex (from France and the Iberian Peninsula; *Quercus robur*, *Q. petraea*, *Q. canariensis*, *Q. faginea*, and *Q. pyrenaica*) which showed 11 haplotypes (Grivet et al. 2006), *Q. suber* in the western Mediterranean with five haplotypes (Magri et al. 2007), and *Q. garryana* in the Pacific northwestern region of North America with six haplotypes (Marsico et al. 2009). However, haplotype richness was lower than in Mexican oak species such as *Q. castanea* (Peñaloza-Ramírez 2011) which had 90 haplotypes and similar to haplotype richness in the *Q. crassifolia* × *Q. crassipes* hybrid complex that

showed 26 haplotypes (Tovar-Sánchez et al. 2008) and the *Q. affinis* × *Q. laurina* hybrid complex that had 35 haplotypes (Ramos-Ortiz et al. 2016).

Also, the values were similar to haplotype richness in *Q. lobata* in California, with 39 haplotypes (Grivet et al. 2006). Within-population diversity (h_S) and total diversity (h_T) observed in *Q. insignis* (0.65 and 0.95, respectively) and *Q. sapotifolia* (0.51 and 0.96) were also higher than in most of the abovementioned mid- and high-latitude oak species (0.28 and 0.97 in *Q. lobata*; Grivet et al. 2006; 0.11 and 0.75 in the European white oak complex; Grivet et al. 2006; and 0.08 and 0.67 in *Q. garryana*; Marsico et al. 2009) and comparable to the Mexican species (0.73 and 0.98 for *Q. castanea*; 0.9 and 0.99 for the *Q. affinis* × *Q. laurina* hybrid complex; Ramos-Ortiz et al. 2016). Despite that diversity values seem to be clearly associated with the different geographic regions (higher values within the Neotropics and lower in the temperate zone), these comparisons should be treated with caution considering that not only the set but also the number of cpSSR loci used only partially coincide among studies.

In terms of the total number of haplotypes, the genetic diversity was higher at the regional level in the MM region than in the UM and LM regions for both *Q. insignis* and *Q. sapotifolia* (Table 1). However, the mean values of within-population diversity were not significantly different according to the Kruskal-Wallis test ($p > 0.05$). Several authors have attributed high levels of chloroplast DNA diversity to processes such as introgression and the persistence of large populations through different periods (Lumaret et al. 2002; Jiménez et al. 2004). It is probable for both Neotropical oak species (particularly *Q. sapotifolia*) that not only processes such as hybridization may have been important drivers of the high observed genetic diversity, but also that the persistence of large areas with suitable climatic niches at least since the LGM promoted large effective population sizes through time. This affirmation is also supported by the absence of the signal of a recent historical population demographic expansion (and therefore population equilibrium), for both species, not only at the population level but also regionally.

Geography and geology effects on the species phylogeographic structure

Both oak species exhibited considerable values of genetic structure. Moreover, significant differences between G_{ST} and N_{ST} ($p < 0.001$) indicated a well-defined phylogeographical structure. The AMOVA showed that 20.57–30.41% of the genetic variation in *Q. insignis* (depending on the mutation model, IAM and SMM, respectively), and 21.21–2.79% in *Q. sapotifolia* is explained by differences among the three predefined geographical regions (UM, MM, and LM). However, a SAMOVA maximizing the genetic variance

among population groups indicated that, in the case of *Q. insignis*, the IT and ND have not been effective barriers. Meanwhile, the BARRIER analysis suggested that the same areas (IT and ND) represent zones where genetic structure differ more than it would be expected by chance. The discrepancy between the SAMOVA and BARRIER results are probably due to methodological differences. SAMOVA searches for groups of populations that are geographically contiguous and maximally differentiated from each other, and BARRIER identifies boundaries between points or study sites based on differences of an attribute distance matrix. The lack of a concordant pattern of geographic structure between the results of BARRIER and SAMOVA may also be reflecting a complex evolutionary history, characterized by events of multiple colonization and recolonization. In particular, the haplotype network suggests an event of colonization of LM from Honduras, at least two events of migration on the opposite direction (LM to MM) and finally another colonization of MM to LM.

The BARRIER analysis also showed important differentiation of populations within MM and LM. Particularly, in the LM region, there was a substructure dividing the populations into northern, central, and southern groups. Also, both MM and LM regions are characterized by high levels of haplotype diversity and a considerable number of singletons. Complementarily, ENMs suggested that precisely, these areas (MM and LM) have exhibited stable climatic niche suitability since the LIG and, despite that geographic barriers (particularly IT and ND) seem to interrupt the suitable climatic areas for the species, through time there have been episodes when connectivity may have been possible.

Other important fact that may have influenced processes of colonization and recolonization between MM and LM is the unstable geological dynamics of the region, characterized by the persistence of intense volcanic activity (as most of southern Central America consists of a Neogene-Quaternary volcanic belt) and tectonic activity related to the Nicaraguan volcanic front, Los Gatusos and San Carlos lowlands and the Chorotega volcanic front described by Marshall (2007), which together configure a highly dynamic and heterogeneous geographic area. Despite it has been shown that human activity can have a major impact on forest species, and particularly in historical demographic expansions, extinctions, and bottlenecks of oak species (de Nascimento et al. 2009), recent studies have indicated that the most likely explanations for the genetic structure in the lowland *Q. oleoides* in Costa Rica are the volcanic activity and geologic events such as the formation of the Guanacaste mountains (Cavender-Bares et al. 2011).

In the case of *Q. insignis*, a similar effect of the complex geological history that characterizes the CRM can be observed. Populations in Costa Rica are distributed in the Cordillera de Guanacaste (CG; i3; Figs. 1 and 2), Cordillera Central (CC; i2 and i4; Figs. 1 and 2), and Cordillera de

Talamanca (CT; i1; Figs. 1 and 2). This distribution includes a Quarternary chain of shield-like stratovolcanos at CG, composite shield volcanoes and strong climate gradients at CC; the Central Valley (between CG, CC, and CT) consisting of a low-relief upland surface with deeply incised river canyons, active faulting and a thick accumulation of andesitic to dacitic lavas, pyroclastic rocks, and lacustrine sediments throughout the Quarternary. Meanwhile, the CT is characterized by a suite of Neogene-Quaternary intrusive and extrusive rocks and a rapid Quaternary uplift and glaciated peaks during the Pleistocene (Marshall 2007). Under such a geological diverse scenario mainly characterized by high volcanic activity and heterogeneous mountain systems, the differentiation of oak populations would be expected, even more considering the limited capability of the oak species to disperse their seeds. *Q. insignis* probably has the largest acorns of any oak in the world (between 7 and 8 cm in diameter; Montes-Hernández and López-Herrera 2013), which could have implied some restrictions in seed movement.

For *Q. sapotifolia*, both AMOVA and SAMOVA results suggested a pattern of genetic differentiation poorly congruent with our a priori regionalization based on the presence of the IT and ND. SAMOVA suggested that Mexican populations (s13, s14, and s15) formed a unique group, and results from BARRIER also showed a lack of differentiation associated to the IT. Three populations (s1, s10, and s12) did not group with other populations, and all the remaining sites formed a unique group including MM and LM samples. Again, the BARRIER analysis did not support significant differences across the ND. Most of the differentiation observed in the BARRIER analysis concentrated at the MM region, particularly near the PMJ. The first main barrier was located at the eastern portion of the IT in the SMC followed by a second barrier separating Guatemala from Honduras populations. The MM is an area of convergence among the Maya highlands, the Motagua fault zone, and the Chortis highlands. The potential distribution area of *Q. sapotifolia* in Guatemala and Honduras is interrupted by the Chuacus range and Las Minas range (in the Motagua fault zone) and the western rifted Chortis highlands. Differentiation among populations indicated by the BARRIER analysis could be attributed to prominent river valleys (such as the Motagua and Polochic valleys) and major structural depressions associated to the valleys as is it described by Marshall (2007).

The haplotype network of *Q. sapotifolia* indicates processes of multiple colonization from MM to UM and LM. Interestingly, the data suggested that the major barriers (IT and ND) were not an important factor defining the genetic structure of the species; meanwhile, barriers that were not expected as strong drivers of differentiation as the CRM showed an important effect on the distribution of genetic variation. This unexpected pattern may reflect for the CRM case the importance of the recent geological, volcanic, and climatic heterogeneity of the Chortis block. It is necessary

to consider that *Q. sapotifolia* has a wide altitudinal distribution (from 250 up to 2000 m) which suggests that this species could have been more successful during processes of dispersion through the lowlands than other oak species such as *Q. insignis* that distributes above the 1500 m. Second, seed dispersal in *Q. sapotifolia* should be easier due to the smaller size of their acorns (~1.5 cm in diameter) (Morales 2010). Therefore, even under a scenario of few climatic suitable areas connecting MM with UM and LM, it is possible to suggest long-distance dispersal (e.g., via birds such as passenger pigeons as it was hypothesized by Nixon (1985)) as a mechanism of dispersal through the main barriers. Finally, similarly to what was observed for *Q. insignis* in LM, *Q. sapotifolia* populations located in the CC are genetically differentiated from the population of the CT, suggesting that Central Valley and the characteristic geology of the CRM described above may have played an important role in the genetic differentiation of several *Quercus* species including *Q. sapotifolia* and *Q. insignis*.

Palaeoclimate effects on the species phylogeographic structure

Despite important genetic barriers in both species were located at continuous areas of climatically suitable conditions during the present-day, LGM and LIG periods (such as the northern portions of the CAVA), there are coincidences between the presence of genetic discontinuities at the ND and the lack of climatic suitable areas in the same region according to the entire set of modeled climatic scenarios (Fig. 4a–d) for *Q. insignis*. Similarly, there was a lack of genetic discontinuities at the ND associated to the presence of continuous climatic suitable areas according to the MIROC and combined models (Fig. 5c, e) for *Q. sapotifolia*. Such differences between the observed genetic differentiation patterns and the results of the ecological niche modelling may suggest that the climatic oscillations during the LGM apparently were not major drivers of the intraspecific differentiation of the species, contrary to the patterns observed in cloud forest species (Ramírez-Barahona and Eguiarte 2014). Contrasting patterns of genetic structure between *Q. insignis* and *Q. sapotifolia* probably resulted from differences in the way the two species responded to environmental changes, as it has been observed in other *Quercus* species (Koehler et al. 2012). Therefore, patterns of genetic diversity should be interpreted on species-specific contexts, as it has been suggested by Ramírez-Barahona and Eguiarte (2013), for the Neotropical cloud forests. Despite this apparent limited effect of paleoclimate changes, it is important to mention that our ecological niche modelling did reflect higher and stable environmental suitability in the MM region, which we consider as a possible area of origin of the species analyzed.

Conclusions

Understanding the effect of historical processes such as tectonics, volcanism, and climate fluctuations on the evolutionary history of the biotas distributed in the northern Neotropics still requires considerable effort. In the present study, two *Quercus* species exemplify the diversity of possible evolutionary histories that can be observed and reconstructed in a common region. In the case of *Q. insignis*, we observed coincidences with previously proposed barriers that may have determined the species phylogeographic structure, the IT and the ND. The influence of both barriers has been observed in different biological groups such as mammals (Arellano et al. 2005), reptiles (Castoe et al. 2003; Hasbún et al. 2005; Daza et al. 2010), birds (DaCosta and Klicka 2008; Vázquez-Miranda et al. 2009), amphibians (Mulcahy et al. 2006), and plants (Cavender-Bares et al. 2011). All the mentioned studies attributed a role as barriers to the IT and ND based on different geological attributes (such as volcanism, tectonics, and physiography) and/or palaeoclimatic events.

On the other hand, the lack of influence of the IT and ND, but significant levels of population differentiation due to the geologic heterogeneity of areas such as the PMJ and the Central American volcanic arc have also been reported for different organismal lineages such as plants (Novick et al. 2003; Jardón-Barbolla et al. 2011) and mammals (Eizirik et al. 1998, 2001; Ordóñez-Garza et al. 2010). Once again, the mentioned studies not only considered different geological features but also palaeoclimatic information in order to explain the patterns observed in their studies. The findings of the present study suggest a complex evolutionary history of the oak populations in the northern Neotropics. The most striking indirect finding is the possibility of a MM origin of the species analyzed; this is particularly clear for *Q. sapotifolia*. This conclusion is based on the higher genetic diversity in this region and the evidence of multiple colonizations from MM to both UM and LM and even recolonizations from LM to MM. The evaluation of a possible Central American origin of this species should be considered in future studies, particularly considering that the northern Neotropics are the encounter point between Nearctic and Neotropical biotas under an extremely complex geologic and climatic scenario.

Acknowledgements The authors thank Jorge Lobo for laboratory access at the Universidad de Costa Rica (UCR) and Jesus Llanderal, German Sandoval, Wilson Zúñiga, Carlos Funes, Carlos O'Reilly, Iliam Rivera, and Katya Romero for support during the field stage. Paul C. Standley (EAP) and Cyril Hardy Nelson Sutherland (TEFH) herbariums provided logistic assistance in Honduras. The comments by Eduardo Ruiz-Sánchez to a previous version of the manuscript are gratefully acknowledged. H. Rodríguez-Correa specially thanks the Programa de Becas Posdoctorales (DGAPA, UNAM) for providing funding to develop postdoctoral studies at UNAM. HRC also thanks the financial support received by the Red Latinoamericana de Botánica-Andrew W. Mellon

Foundation Grant 2010–2011, and all authors acknowledge support from PAPIIT IV201015 and CONACYT 240136 grants.

Data archiving statement Genotypic data for the studied species will be submitted to the Dryad Digital Repository (datadryad.org/) following its data management policy, the accession numbers will be supplied once available. Meanwhile, data is provided as a [supplementary file](#).

References

- Alexander L, Woeste K (2014) Pyrosequencing of the northern red oak (*Quercus rubra* L.) chloroplast genome reveals high quality polymorphisms for population management. *Tree Genet Genomes* 10: 803–812
- Arellano E, González-Cozátl FX, Rogers D (2005) Molecular systematics of Middle American harvest mice *Reithrodontomys* (Muridae), estimated from mitochondrial cytochrome b gene sequences. *Mol Phylogenet Evol* 37:529–540
- Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48
- Braconnot P, Otto-Bliessner B, Harrison S, Joussaume S, Peterschmitt JY et al (2007) Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum—part 2: feedbacks with emphasis on the location of the ITCZ and mid- and high latitudes heat budget. *Clim Past* 3:279–296
- Castoe TA, Chippindale PT, Campbell JA, Ammerman LK, Parkinson CL (2003) Molecular systematics of the middle American jumping pitvipers (genus *Atropoides*) and phylogeography of the *Atropoides nummifer* Complex. *Herpetologica* 59:421–432
- Cavender-Bares J, González-Rodríguez A, Pahlich A, Koehler K, Deacon N (2011) Phylogeography and climatic niche evolution in the live oaks (*Quercus* series *Virentes*) from the tropics to the temperate zone. *J Biogeogr* 38:962–981
- Cavender-Bares J, González-Rodríguez A, Eaton D, Hipp A, Beulke A, Manos P (2015) Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): a genomic and population genetics approach. *Mol Ecol* 24:3668–3687
- Chen D, Zhang X, Kang H, Sun X, Yin S, Du H, Yamanaka N, Gapare W, Wu HX, Liu C (2012) Phylogeography of *Quercus variabilis* based on chloroplast DNA sequence in east Asia: multiple glacial refugia and mainland-migrated island population. *PLoS One* 7(10):e47268
- Collins WD, Blackmon M, Bitz C, Bonan G, Bretherton CS (2004) The community climate system model: CCSM3. *J Clim* 19:2122–2143
- Cottrell JE, Munro RC, Tabbener HE, Gillies ACM, Forrest GI, Deans JD, Lowe AJ (2002) Distribution of chloroplast DNA variation in British oaks (*Quercus robur* and *Quercus petraea*): the influence of postglacial colonization and human management. *Forest Ecol Manag* 156:191–195
- Csaikl UM, Burg K, Fineschi S, König AO, Mátyás G, Petit RJ (2002) Chloroplast DNA variation in the alpine region. *Forest Ecol Manag* 156:131–145
- DaCosta JM, Klicka J (2008) The Great American Interchange in birds: a phylogenetic perspective with the genus Trogon. *Mol Ecol* 17: 1328–1343
- Daza JM, Castoe TA, Parkinson CL (2010) Using regional comparative phylogeographic data from snake lineages to infer historical processes in Middle America. *Ecography* 33:343–354
- de Nascimento L, Willis KJ, Fernández-Palacios JM, Criado C, Whittaker RJ (2009) The long-term ecology of the lost forests of La Laguna, Tenerife (Canary Islands). *J Biogeogr* 36:499–514
- Dumolin-Lapègue S, Demesure B, Fineschi S, Le Corre V, Petit RJ (1997) Phylogeographic structure of white oaks throughout the European continent. *Genetics* 146:1475–1487
- Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Mol Ecol* 11:2571–2581
- Eizirik E, Bonatto SL, Johnson WE, Crawshaw PG Jr, Vié JC, Brousset DM, O'Brien SJ, Salzano FM (1998) Phylogeographic patterns and evolution of the mitochondrial DNA control region in two Neotropical cats (Mammalia, Felidae). *J Mol Evol* 47:613–624
- Eizirik E, Kim J-H, Menotti-Raymond M, Crawshaw PG Jr, O'Brien SJ, Johnson WE (2001) Phylogeography, population history and conservation genetics of jaguars (*Panthera onca*, Mammalia, Felidae). *Mol Ecol* 10:65–79
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver 3.5. An integrated software package for population genetics data analysis. *Evol Bioinforma* 1:47–60
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24:38–49
- Fineschi S, Turchini D, Grossoni P, Petit RJ, Vendramin GG (2002) Chloroplast DNA variation in white oaks in Italy. *Forest Ecol Manag* 156:103–114
- Fu YX (1997) Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics* 147:915–925
- Goldstein DB, Ruiz-Linares A, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139:463–471
- González-Espinosa M, Meave JA, Lorea-Hernández FG, Ibarra-Manríquez G, Newton AC (2011) The red list of Mexican cloud forest trees. *Fauna and Flora International (FFI)*, Cambridge
- González-Rodríguez A, Bain JF, Golden JL, Oyama K (2004) Chloroplast DNA variation in the *Quercus affinis-Quercus laurina* complex in Mexico: geographical structure and associations with nuclear and morphological variation. *Mol Ecol* 13:3467–3476
- González-Rodríguez A, Arias DM, Oyama K (2005) Genetic variation of populations within the *Quercus affinis-Quercus laurina* (Fagaceae) complex analyzed with RAPD markers. *Can J Bot* 83:155–162
- Grivet D, Deguilloux M, Petit RJ, Sork V (2006) Contrasting patterns of historical colonization in white oaks (*Quercus* spp.) in California and Europe. *Mol Ecol* 15:4085–4093
- Gutiérrez-García T, Vázquez-Domínguez E (2013) Consensus between genes and stones in the biogeographic and evolutionary history of Central America. *Quat Res* 79:311–324
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population level. *Mol Ecol Notes* 2:618–620
- Hasbún CR, Gómez A, Khöler GG, Lunt DH (2005) Mitochondrial DNA phylogeography of the Mesoamerican spiny-tailed lizards (*Ctenosaura quinquecarinata* complex): historical biogeography, species status and conservation. *Mol Ecol* 14:3095–3107
- Hasumi H, Emori S (2004) K-1 coupled GCM (MIROC) description. Center for Climate System Research (CCSR), University of Tokyo; National Institute for Environmental Studies (NIES); Frontier Research Center for Global Change (FRCGC). http://ccsr.aori.u-tokyo.ac.jp/~hasumi/miroc_description.pdf. Accessed 20 July 2014
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965–1978
- Hooghiemstra H, van der Hammen T (2004) Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. *Phil Trans R Soc Lond B* 359:173–181
- Jardón-Barbolla L, Delgado-Valerio P, Geada-López G, Vázquez-Lobo A, Piñero D (2011) Phylogeography of *Pinus* subsection *Australes* in the Caribbean basin. *Ann Bot-London* 107:229–241
- Jimenez P, López de Heredia U, Collada C, Lorenzo Z, Gil L (2004) High variability of chloroplast DNA in three Mediterranean evergreen oaks indicates complex evolutionary history. *Heredity* 93:510–515

- Koehler K, Center A, Cavender-Bares J (2012) Evidence of a freezing tolerance-growth rate trade-off in the live oaks (*Quercus* series *Virentes*) across tropical-temperate divide. *New Phytol* 193:730–744
- Liu H, Takeichi Y, Kamiya K, Harada K (2013) Phylogeography of *Quercus phillyraeoides* (Fagaceae) in Japan as revealed by chloroplast DNA variation. *J Forest Res-JPN* 18:361–370
- López de Heredia U, Carrión JS, Jiménez P, Collada C, Gill L (2007) Molecular and palaeoecological evidence for multiple glacial refugia for evergreen oaks on the Iberian Peninsula. *J Biogeogr* 34:1505–1517
- Lumaret R, Mir C, Michaud H, Raynal V (2002) Phylogeographic variation of chloroplast DNA in holm oak (*Quercus ilex* L.). *Mol Ecol* 11:2327–2336
- Magri D, Fineschi S, Bellarosa R, Buonamici A, Sebastiani F, Schirone B, Simeone MC, Vendramin GG (2007) The distribution of *Quercus suber* chloroplast haplotypes matches the paleogeographical history of the western Mediterranean. *Mol Ecol* 16:5259–5266
- Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by “Monmonier’s algorithm”. *Hum Biol* 76:173–190
- Marshall JS (2007) Geomorphology and physiographic provinces. In: Bunschuc J, Alvarado GE (eds) *Central America: geology, resources, hazards*, vol 1. Taylor & Francis, London, pp 75–122
- Marsico TD, Hellman JJ, Romero-Severson J (2009) Patterns of seed dispersal and pollen flow in *Quercus garryana* (Fagaceae) following post-glacial climatic changes. *J Biogeogr* 36:929–941
- Montes-Hernández B, López-Herrera (2013) Seedling establishment of *Quercus insignis*: a critically endangered oak tree species in southern Mexico. *Forest Ecol Manag* 310:927–934
- Morales JF (2010) Fagaceae. In: Hammel BE, Grayum MH, Herrera C, Zamora N (eds) *Manual de plantas de Costa Rica, Dicotiledóneas (Clusiaceae-Gunneraceae)*, vol 5. Missouri Botanical Garden, St. Louis, pp 777–781
- Mulcahy DG, Morrill BH, Mendelson JR III (2006) Historical biogeography of low-land species of toads (*Bufo*) across the Trans-Mexican Neovolcanic Belt and the Isthmus of Tehuantepec. *J Biogeogr* 33:1889–1904
- Navascués M, Hardy OJ, Burgarella C (2009) Characterization of demographic expansions from pairwise comparisons of linked microsatellite haplotypes. *Genetics* 181:1013–1019
- Nixon KC (1985) A biosystematics study of *Quercus* series *Virentes* (the live oaks) with phylogenetic analyses of Fagales, Fagaceae and *Quercus*. PhD Thesis, University of Texas, Austin, Tx
- Novick RR, Dick CW, Lemes MR, Navarro C, Caccone S, Bermingham E (2003) Genetic structure of a Mesoamerican population of big-leaf mahogany (*Swietenia macrophylla*) inferred from microsatellite analysis. *Mol Ecol* 12:2885–2893
- Olalde M, Herrán A, Espinel S, Goicochea PG (2002) White oaks phylogeography in the Iberian Peninsula. *Forest Ecol Manag* 156:89–102
- Ordóñez-Garza N, Matson JO, Strauss RE, Bradley RD, Salazar-Bravo J (2010) Patterns of phenotypic and genetic variation in three species of endemic Mesoamerican *Peromyscus* (Rodentia: Cricetidae). *J Mammal* 91:848–859
- Ornelas JF, Sosa V, Soltis DE, Daza JM, Gonzalez C, Soltis PS, Gutierrez-rodríguez C, Espinosa de los Monteros A, Castoe TA, Bell C, Ruiz-Sánchez E (2013) Complex evolutionary history of threatened cloud forest of northern Mesoamerica. *PLoS One* 8:e56283
- Otto-Bliesner BL, Marshall SJ, Overpeck JT, Miller GH, Hu A (2006) Simulating Arctic climate warmth and icefield retreat in the Last Interglaciation. *Science* 311:1751–1753
- Peñalosa-Ramírez JM (2011) Filogeografía e hibridación en cuatro especies del género *Quercus* (Fagaceae) en México. PhD Thesis, Universidad Nacional Autónoma de México, Mexico
- Petit RJ, Brewer S, Bordács S, Burg K, Cheddadi R, Coart E, Cottrel J, Csaikl M, van Dam B, Deans JD, Espinel S, Fineschi S, Finkeldey R, Glaz I, Goicochea PG, Jensen JS, König AO, Lowe AJ, Madsen SF, Mátyás G, Munro RC, Popescu F, Slade D, Tabbener H, de Vries SGM, Ziegenhagen B, de Beaulieu J, Kremer A (2002a) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecol Manag* 156:49–74
- Petit RJ, Csaikl UM, Bordács S, Burg K, Coart E, Cottrel J, van Dam B, Deans JD, Dumolin-Lapègue S, Fineschi S, Finkeldey R, Gillies A, Glaz I, Goicochea PG, Jensen JS, König AO, Lowe AJ, Madsen SF, Mátyás G, Munro RC, Olalde M, Pemonge M, Popescu F, Slade D, Tabbener H, Turchini D, de Vries SGM, Ziegenhagen B, Kremer A (2002b) Chloroplast DNA variation in European oaks phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecol and Manag* 156:5–26
- Phillips SJ, Anderson RP, Schapire PE (2006) Maximum entropy modelling of species geographic distributions. *Ecol Model* 190:231–259
- Polzin T, Daneschmand SV (2003) On Steiner trees and minimum spanning trees in hypergraphs. *Oper Res Lett* 31:12–20
- Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* 144:1237–1245
- Ramírez-Barahona S, Eguiarte L (2013) The role of glacial cycles in promoting genetic diversity in the Neotropics: the case of cloud forests during the Last Glacial Maximum. *Ecol Evol* 3:725–738
- Ramírez-Barahona S, Eguiarte L (2014) Changes in the distribution of cloud forest during the last glacial predict patterns of genetic diversity and demographic history of the tree fern *Alsophilla firma* (Cyatheaceae). *J Biogeogr* 41:2396–2407
- Ramírez-Barahona S, Luna-Vega I (2015) Geographic differentiation of tree ferns (Cyatheales) in tropical America. *Am Fern J* 105:73–85
- Ramos-Ortiz S, Oyama K, Rodríguez-Correa H, González-Rodríguez A (2016) Geographic structure of genetic and phenotypic variation in the hybrid zone between *Quercus affinis* and *Q. laurina* in Mexico. *Plant Spec Biol* 31:219–232
- Rodríguez-Correa H (2015) Patrones de distribución y filogeografía de los encinos (quercus: fagaceae) en Mesoamérica y Los Andes. PhD Thesis. Universidad Nacional Autónoma de México, Mexico
- Rodríguez-Correa H, Oyama K, Mac Gregor-Fors I, González-Rodríguez A (2015) How are oaks distributed in the Neotropics? A perspective from species turnover, areas of endemism, and climatic niches. *Int J Plant Sci* 176:222–231
- Rousset F (1996) Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics* 142:1357–1362
- Sebastiani F, Carnevale S, Vendramin GG (2004) A new set of mono- and dinucleotide chloroplast microsatellites in Fagaceae. *Mol Ecol Notes* 4:259–261
- Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139:457–462
- Torres-Miranda A, Luna-Vega I, Oyama K (2013) New approaches to the biogeography and areas of endemism of red oaks (*Quercus* L., Section *Lobatae*). *Syst Biol* 62:555–573
- Tovar-Sánchez E, Mussali-Galante P, Esteban-Jiménez R, Piñero D, Arias DM, Dorado O, Oyama K (2008) Chloroplast DNA polymorphism reveals geographic structure and introgression in the *Quercus crassifolia* x *Quercus crassipes* hybrid complex in Mexico. *Botany* 86:228–239
- Valencia-A S (2004) Diversidad del género *Quercus* (Fagaceae) en México. *B Soc Bot Mex* 75:33–53
- Vázquez-Miranda H, Navarro-Sigüenza AG, Omland KE (2009) Phylogeography of the rufous-naped red (*Camphylorhynchus rufinucha*): speciation and hybridization in Mesoamerica. *Auk* 126:765–778