UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

Centro de Investigaciones en Ecosistemas

Filogeografía e hibridación en cuatro especies del género *Quercus* (Fagaceae) en México

T E S I S

QUE PARA OBTENER EL GRADO ACADÉMICO DE

DOCTOR EN CIENCIAS

P R E S E N T A

JUAN MANUEL PEÑALOZA RAMÍREZ

TUTOR PRINCIPAL DE TESIS: Dr. Ken Oyama Nakagawa

COMITÉ TUTOR: Dr. Juan José Morrone Lupi
Dr. Juan Núñez Farfán

Morelia, Michoacán **Junio, 2011**

UNAM
POSGRADO
Ciencias Biológicas

COORDINACIÓN

Dr. Isidro Ávila Martínez
Director General de Administración Escolar, UNAM
P r e s e n t e

Por medio de la presente me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día 18 de enero del 2010, se acordó poner a su consideración el siguiente jurado para el examen de DOCTOR EN CIENCIAS del alumno **PEÑALOZA RAMÍREZ JUAN MANUEL** con número de cuenta **505013736**, con la tesis titulada: **"Filogeografía e hibridación en cuatro especies de *Quercus* (Fagaceae) en México"**, bajo la dirección del **Dr. Alberto Ken Oyama Nakagawa.**

Presidente:	Dr. Juan J. Morrone Lupi
Vocal:	Dr. Antonio González Rodríguez
Vocal:	Dr. Daniel Ignacio Piñero Dalmau
Vocal:	Dra. Susana Valencia Ávalos
Secretario:	Dr. Alberto Ken Oyama Nakagawa
Suplente:	Dr. Mauricio Quesada Avendaño
Suplente:	Dr. Juan Núñez Farfán

Sin otro particular, quedo de usted.

A t e n t a m e n t e
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, D.F., a 7 de junio del 2011.

Dra. María del Coro Arizmendi Arriaga
Coordinadora del Programa

c.c.p. Expediente del interesado

AGRADECIMIENTOS

A las instituciones de hicieron posible el desarrollo de esta tesis

A mi alma Mater, La Universidad Nacional Autónoma de México, por haberme dado la oportunidad de ser Puma y por darnos la oportunidad a los que tenemos menos recursos de llegar a soñar.

Al Posgrado en Ciencias Biológicas de la UNAM, por haberme aceptado como alumno, por todas las atenciones brindadas en todo momento, a la calidad humana que tienen todos.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por haberme otorgado una beca para realizar mis estudios de doctorado y estancias en el extranjero.

Al Centro de Investigaciones en Ecosistemas de la UNAM Campus Morelia, por haberme brindado un espacio como estudiante en el cual tengo la fortuna de pertenecer al laboratorio de Ecología genética y molecular.

Los recursos económicos para la realización de esta tesis fueron proporcionados por los proyectos: CONACYT 38550-V, CONACYT-SEMARNAT 2004-01-97 y 23728 del Dr. Ken Oyama.

A las personas que aportaron su trabajo conocimiento y experiencia

Agradezco principalmente al Dr. Ken Oyama, por el apoyo incondicional, por haber creído en mí como estudiante, por su pasión por la biología, por las inmensas cosas que he aprendido de tí, tu inagotable entusiasmo y la visión de tu parte para dirigir con atino esta investigación, por dedicarme horas de tu tiempo e inculcarme los más altos valores de excelencia, perseverancia, dedicación, honradez y lo buen amigo que te considero, ojala sea el principio de todo.

Agradezco al Dr. Antonio González Rodríguez por haberme guiado en este camino de la ciencia de tan excelente manera y por sus grandes aportaciones al trabajo y su gran honestidad y desempeño.

A los miembros de mi comité tutoral, mis amigos el Dr. Juan José Morrone Lupi y al Dr. Juan Núñez Farfán, gracias por todo el apoyo académico que me han dado, por haberme guiado durante los tutorales, por la calidad moral y la visión tan objetiva para corregir mi tesis.

Gracias a mis compañeros y amigos, al Dr. Pablo Cuevas Reyes, Dr. Omar Chassin Noria, Dr. Luis Felipe Mendoza Cuenca y al Dr. Mauricio Quesada Avendaño, gracias por ese tiempo invaluable que me otorgaron para poder aprender de ustedes que han sido piezas súper necesarias en mi formación académica.

A mis amigos franceses, el Dr. Antoine Kremer, Dr. Henri Caron, Dr. Remy Petit, Dr. Olivier Lepais y al Dr. Jeremy Derory, muchas gracias por haberme aceptado en su laboratorio en el INRA-Pierronton, Bordeaux, Francia, por su paciencia y por los grandes conocimientos que recibí de ustedes, por esa pasión y honestidad.

A los que revisaron la tesis como miembros de mi jurado, al Dr. Juan José Morrone Lupi, Dr. Antonio González Rodríguez, Dr. Daniel Piñero Dalmau, Dr. Mauricio Quesada Avendaño, Dra. Susana Valencia Ávalos, Dr. Ken Oyama y al Dr. Juan Núñez Farfán, por contribuir de manera precisa en la corrección de la tesis de forma considerable.

A la Dra. Susana Valencia Ávalos, gracias por su apoyo a entender las complejas variaciones morfológicas de los encinos, su gran paciencia y ayuda en la determinación taxonómica de mis ejemplares, y finalmente gracias por sus grandes aportes a entender el complejo mundo de los encinos.

A las personas que me ayudaron con el trabajo de campo: Ana Luisa Albarrán Lara, Enrique Pascual Alvarado, Antonio González Rodríguez, Pablo Cuevas Reyes, Yurixhi Maldonado, Luisa Herrera, Hermilo Sánchez, Rafael Aguilar, Erasto Hernández y a la gente de las comunidades a las que fuimos y que siempre nos trataron con mucho respeto, muchas gracias por acompañarme en tantos lugares por todo este hermoso México y sus magníficos paisajes.

Al Dr. Víctor Rocha Ramírez, por toda la ayuda que nos prestas en todo momento, por el apoyo logístico y por esas contribuciones inmensas en el campo de la biología molecular, por tu tenacidad, honestidad y amistad.

A Nidia Pérez, gracias por la amistad de todos estos años de apoyo como técnico, por esa labor de amiga en todo momento.

A la Sra. Dolores Lugo Aquino, Lolita, muchas gracias por mantener un laboratorio siempre impecable, por sus grandes conocimientos de todos los reactivos y buffers que usted ha preparado de la mejor manera, y sobre todo por su amistad invaluable que la hacen ser la master y guía de nuestro laboratorio.

A mis amigos y personas que tienen un lugar especial

A mis amigos del Laboratorio de Ecología Genética y Molecular del CIEco: Ana Luisa, Antonio, Chassin, Pablo, Lucho, Nancy, Luis Humberto (Machín), Rodo, Lalo, Mauricio, Lorena, Gabriela, Yuri, Yuni, Luisa, Enrique, Erasto, Hermilo, Nidia, Selene, Chucho, Nando, Wil, Coral, Paco-Beto, Cesar, Karla, Paulina, Esmeralda, Víctor, Pacheco, Rafa, Pavka, Rolo, Omar, Walehaa y Gume, por esos momentos en la convivencia diaria, el intercambio incesante de información “académica” y por supuesto de las fiestas interminables y las pláticas hasta altas horas de la noche.

A las personas como Lorena y Gabriela que más que enseñarles y guiarlas en este camino de la Biología, he aprendido mucho con ustedes mil gracias por toda su gran valía, honestidad, por esas ganas inmensas de aprender y por esa lección de vida que me han dado, espero que sigamos siendo amigos.

A mis amigos de aquellas noches entrañables de fiesta y por su amistad en todo momento: Ana, Antonio, Chassin, Pablo, Lucho, Gume, Yuri, Yuni, Nancy, Luis, Rodo, Mauricio, Lorena, Gabriela, Nidia, Hermilo, Alina y Selene, ha sido increíble el viaje que hemos tomado y espero que solo sea el comienzo de todo, gracias por considerarme su amigo.

A mis eternos amigos de la de la facultad de Biología de la UAEM: el Perro, Panda, Correa, Caco, Zobaco, Trucha, Tobata, la Tuza, Ofelia, etc. etc. que siempre nos reunimos a ponernos al corriente de todos los acontecimientos más importantes, en un ambiente bastante bueno.

Pero sobre todo a mis padres por su apoyo incondicional durante toda mi formación académica y personal. A mi padre Pedro Peñaloza Sánchez gracias por su compañía en esos años maravillosos, y en mi formación, siempre estarás en mi corazón donde quiera que te encuentres.

A mi madre Julia Victoria Guerrero Hernández, gracias por tener la valentía, la bravura y convicción necesarias para hacerse cargo de una familia como la nuestra, por esa lección de vida de todos los días, de superación y ganas de trabajar, eres mi mejor ejemplo en la vida, gracias dios por darme una madre así.

Y muy especialmente a mi esposa Ana Luisa Albarrán Lara (Aniuxilla) gracias por su compañía y amor en todos estos años extraordinarios de formación y de crecimiento personal, en las buenas y en las no tanto, realmente eres una parte imprescindible de mi vida, y espero que continuemos con esta increíble aventura, que es vivir la vida contigo.

A mis hermanitos Martha y su esposo Carlos Ocampo, Pedro (perrys), y Abraham (piloncillo) gracias por todos esos años en los que crecimos juntos y por no fragmentarnos en los malos momentos, que bueno es poder tenerlos. Agradezco a mi sobrina Carla Cheyenne Ocampo Peñaloza, que sea la luz que ilumina nuestra familia y nuestro ángel de la guarda, gracias por existir.

Igualmente agradezco a mi perrita la Chiqui que con tu compañía amor y fidelidad me reconfortas mi corazón, gracias por los paseos nocturnos de todos los días.

A mi familia por parte de mi madre gracias por todo el apoyo que he recibido de cada uno de ustedes de la misma manera agradezco a mi familia por parte de mi padre, por los consejos, el apoyo moral en las buenas y en las malas.

Agradezco también a la familia de mi esposa Ana Luisa: los Lara de Alvarado y a los Albarrán de Morelos, gracias por haberme aceptado en su familia y por el apoyo que me han brindado en todo momento

En el fondo, los científicos somos gente con suerte: podemos jugar a lo que queramos durante toda la vida. Lee Smolin

Hay una fuerza motriz más poderosa que el vapor, la electricidad y la energía atómica: la voluntad. Albert Einstein

Nuestra lealtad es para las especies y el planeta. Nuestra obligación de sobrevivir no es sólo para nosotros mismos sino también para ese cosmos, antiguo y vasto, del cual derivamos. Carl Sagan

Si no muriéramos no apreciaríamos la vida como lo hacemos.
Jacques Yves Cousteau

Las personas no son recordadas por el número de veces que fracasan, sino por el número de veces que tienen éxito. Thomas Alva Edison

¿La ilusión? Eso cuesta caro. A mí me costó vivir más de lo debido.
Juan Rulfo

No existe la libertad, sino la búsqueda de la libertad, y esa búsqueda es la que nos hace libres. Carlos Fuentes

Así es -suspiró el coronel-. La vida es la cosa mejor que se ha inventado.
Gabriel García Márquez

Un libro, como un viaje, se comienza con inquietud y se termina con melancolía.
José Vasconcelos

Prefiero morir de pie que vivir toda una vida arrodillado.
Los que no tengan miedo que pasen a firmar.
Emiliano Zapata

La autocrítica está muy bien, mientras no tenga que ver con uno mismo.
Carlos Monsiváis

Con flores escribes, Dador de la Vida, con cantos das color, con cantos sombreas a los que han de vivir en la tierra. Nezahualcóyotl

Has arribado a tu ciudad, aquí has venido a sentarte en tu solio, en tu trono. Oh, por tiempo breve te lo reservaron, te lo cuidaron los que ya se fueron, tus sustitutos. Los señores reyes Icoatzin, Motecuhmatzin el viejo, Axayácatl, Tízoc, Ahuízotl... Llega a la tierra y descansa. Toma posesión de tus casas reales, da refrigerio a tu cuerpo.
Encuentro de Moctezuma Xocoyotzin y Hernán Cortés.

CONTENIDO

I. RESUMEN

Las zonas de hibridación son consideradas como una fuente muy importante de recombinación genética y de diversidad en la evolución de las plantas. Estas zonas son a menudo el producto de contacto secundario entre poblaciones o especies que se han diferenciado previamente en alopatría, o también pueden surgir en respuesta a variables de selección espacialmente divergentes dentro de una misma especie. Casi todas las zonas híbridas que han sido estudiadas hasta el momento implican dos especies progenitoras y sus híbridos. No obstante, poco se sabe de los casos más complejos donde ocurre la hibridación simultánea entre tres o más especies progenitoras, por lo que la estructura genética y la dinámica del flujo génico en zonas híbridas con múltiples especies son poco conocidas. Las zonas de hibridación múltiple requieren primero de la generación de algunos genotipos híbridos F1 fértiles o retrocruzas entre al menos un par de especies. Posteriormente, estos individuos pueden dispersarse a lo largo de la zona híbrida donde la introgresión puede ocurrir entre los híbridos F1 y una tercera especie, o entre híbridos F1 generados a partir de diferentes combinaciones de especies. La evolución de las poblaciones híbridas en la zona de contacto dependerá de la capacidad relativa de los distintos genotipos híbridos con respecto a las especies progenitoras. Sin embargo, la naturaleza y la intensidad de los factores selectivos no son necesariamente homogéneas en el espacio. La selección exógena implica la adaptación local a su medioambiente, mientras que la selección endógena ocurre cuando los híbridos tienen menor competitividad, debido a las incompatibilidades entre los genomas progenitoras. Por lo tanto, las presiones selectivas que actúan sobre los híbridos puede ser consecuencia de los efectos combinados de factores ambientales como el clima, el suelo y las interacciones con otros organismos, y de factores independientes de la variación ambiental, tal como el rompimiento de los complejos de genes coadaptados y la acción de los genes asociados con la esterilidad. Sin embargo, aún no está muy claro cuál de estos modelos describa mejor las zonas de hibridación múltiples. En el primer capítulo se analizaron los patrones de variación en la forma de la hoja, además de la estructura genética en una zona de contacto de hibridación múltiple entre las especies de encinos rojos *Q. hypoleucoides*, *Q. sideroxyla* y *Q. scytophylla* en el noroeste de México. Analizamos un total de 247 árboles, diez poblaciones de referencia y 13 poblaciones presumiblemente intermedias y siete microsatélites nucleares como marcadores genéticos. Se realizó un análisis de la función discriminante para la variación de la

forma de la hoja, además de obtener estimaciones de la diversidad y la estructura genética y finalmente se practico un análisis de asignación genética Bayesiano por medio del programa STRUCTURE. Las poblaciones de referencia se dividieron en tres grupos completamente distintos según el análisis de función discriminante basado en los datos morfológicos, y mostraron una baja pero significativa diferenciación genética. Poblaciones en la zona de contacto contienen individuos morfológicamente intermedios entre pares de especies en diferentes combinaciones, o incluso entre las tres especies. El análisis de asignación genética Bayesiano reveló que tres grupos genéticos principales son los más idóneos en los datos, con una buena correspondencia de las poblaciones de referencia de cada especie a cada uno de sus grupos genéticos. No obstante, se encontraron varios grados de introgresión en las poblaciones de la zona de contacto. Las tres especies de encinos rojos forman una zona hibrida compleja que está geográficamente estructurada como mosaico, y que comprende una amplia gama de genotipos, incluidos los híbridos entre pares de especies diferentes, retrocruzas y probables híbridos triples.

Para el capítulo dos se evaluó la estructura filogeográfica, la dinámica del intercambio citoplasmático y el modelado de nicho ecológico de tres especies de encinos rojos *Q. hypoleucoides*, *Q. scytophylla* y *Q. sideroxyla* en toda su distribución geográfica en México. En total analizamos 434 individuos de 51 poblaciones de las especies de encino rojo con ocho loci de microsatélite de cloroplasto, de los cuales fueron estimados parámetros de diversidad y estructura genética dentro y entre poblaciones, incluso se probaron si existen señales históricas de expansión demográfica, finalmente se construyeron los modelos de nicho ecológico que fueron obtenidos para la distribución potencial presente y dos posible escenarios durante el ultimo glaciar máximo utilizando los modelos CCSM y MIROC que están implementados en el programa MAXENT. Los resultados obtenidos revelaron valores altos de diversidad dentro y entre poblaciones, así como para la riqueza haplotípica y para las distancias genéticas dentro de poblaciones (D^2sh). Por ejemplo, para *Q. hypoleucoides* altos valores de $H_S = 0.671$, $H_T = 0.932$, para *Q. scytophylla* con valores de $H_S = 0.708$ y $H_T = 0.972$ y para *Q. sideroxyla*, los valores de diversidad genética fueron $H_S = 0.695$, $H_T = 0.981$. Incluso, se sugiere la presencia de intercambio citoplasmático por la invasión de haplotipos de *Q. hypoleucoides* hacia *Q. scytophylla* y *Q. sideroxyla*, principalmente en poblaciones que se encuentran en las zona de contacto. También, se observo un aumento en la diversidad genética de algunas poblaciones en la zona de contacto

principalmente en el norte de México. Estos resultados indican que a pesar de la hibridación incluso se mantiene una moderada diferenciación genética en las poblaciones. Se encontró una fuerte estructura filogeográfica y un proceso de colonización de norte-sur en sus principales linajes a través de México. Observamos que las poblaciones que muestran señales de expansión han sido interpretadas como poblaciones que han servido como reservorios de diversidad genética durante las oscilaciones climáticas, para poblaciones relativamente grandes. Podemos sugerir que las especies han experimentado una moderada reducción en su distribución (10-20%) para los modelos del último glaciar máximo, es muy posible que las poblaciones hayan experimentado desplazamientos altitudinales en vez de migraciones o colonizaciones latitudinales como respuesta a los diferentes periodos glaciares. Adicionalmente y en consecuencia se han retenido altos niveles de variación genética, lo que implica tamaños efectivos de población grandes y una compleja dinámica en el intercambio citoplasmático, permitiendo el mantenimiento de ciertos niveles de conectividad en sus poblaciones.

Para el capítulo tres, los objetivos fueron inferir la estructura filogeográfica, de la especie de encino rojo *Quercus castanea* que tiene una amplia distribución geográfica en México además de probar si han existido señales históricas de expansión poblacional y finalmente reconstruir los modelos de distribución potencial presente y durante el último glaciar máximo (LGM) por medio del programa MAXENT. Colecté 341 individuos de 36 poblaciones de *Q. castanea* que se localizan en cuatro provincias morfotectónicas y biogeográficas de México comúnmente reconocidas: la Sierra Madre Occidental (SMO), la Meseta Central (PC), la Faja Volcánica Transmexicana (FVTM) y la Sierra Madre del Sur (SMS). Se utilizaron siete loci de microsatélite de cloroplasto (cpSSRs), los datos fueron analizados para obtener medidas de diversidad y estructura genética y para probar si existen señales históricas de expansión demográfica. Finalmente, los modelos de distribución potencial se obtuvieron para la distribución potencial actual y dos escenarios de las condiciones climáticas durante el último glaciar máximo usando los modelos CCSM y MIROC que son implementados en el programa MAXENT. Se encontraron un total de 90 haplotipos, de los cuales 28 fueron únicos o raros, la diversidad genética dentro de poblaciones fue alta h_S= 0.72, y la diferenciación genética entre-poblaciones indicaron una fuerte estructura filogeográfica (R_{ST} = 0.711> G_{ST} = 0.261, p <0.05). Además, las señales de expansión demográfica fueron identificadas en la mayoría de las poblaciones, con valores altos de τ (tiempo desde la

expansión poblacional), en las regiones de SMS, seguidos por la SMO, el TMVB y finalmente la CP. Los modelos de nicho ecológico sugieren una distribución similar pero moderadamente (30-50%) grande para los modelos del último glaciar máximo (CCSM y MIROC) en comparación con el modelo de la distribución potencial actual. En conclusión encontramos una diversidad genética alta, una fuerte estructura filogeográfica y los modelos de distribución potencial sugieren la permanencia *in situ* de poblaciones de *Q. castanea* con grandes tamaños efectivos durante los últimos miles años. Las señales de expansión poblacional, probablemente fueron aparentemente muy anteriores al último glaciar máximo, posiblemente, tan antiguas como 1.6×10^6, 1.9×10^5 años atrás, y sugieren que las expansiones ocurrieron siguiendo una dirección desde el sur hasta el norte de México. Finalmente, y como ha ocurrido con otras especies de árboles Mexicanos, la compleja historia geológica y climática de la FVTM contribuye a explicar el origen y mantenimiento de una gran parte de la diversidad genética que reside en esta especie de encino rojo Mexicano.

II. ABSTRACT

Hybrid zones are considered an important source of genetic recombination and diversity in plant evolution. These zones are most often the product of secondary contact between populations or species that have differentiated previously in allopatry, or may also arise in situ in response to spatially varying selection. Nearly all of the hybrid zones that have been studied involve two parental taxa and their hybrids. Nevertheless, more complex instances of simultaneous hybridization among three or more parental taxa also occur in nature, but these have been rarely analysed. Consequently, the genetic structure and the dynamics of gene flow in multispecies hybrid zones are poorly known. These situations require the production of some fertile hybrid genotypes (F1 and/or backcrosses) between at least one species pair. Subsequently, these individuals may disperse throughout the hybrid zone and crosses can occur between hybrids and a third species, or between hybrids from different species combinations. The evolution of hybrid populations in the zone of contact will depend on the relative fitness of the various hybrid and parental genotypes. Nevertheless, the nature and intensity of selective factors are not necessarily homogeneous across landscape. Exogenous selection implies adaptation to local environments, whereas endogenous selection occurs when hybrids have low fitness due to incompatibilities between the parental genomes. Therefore, the selective pressures acting on hybrids may result from the combined effects of environmental factors such as climate, soil and interactions with other organisms, and factors independent of environmental variation such as the disruption of coadapted gene complexes and the action of genes associated with sterility. However, it is not clear which of these models best describe multispecies hybrid zones.

In this study I examined the leaf shape variation patterns, as well as the genetic structure in a zone of contact of multiple hybridization between the red oaks species *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla* in northwestern Mexico. A total of 247 trees from ten reference and 13 presumed intermediate populations were characterized using leaf shape variation and geometric morphometrics, and seven nuclear microsatellites as genetic markers. Discriminant function analysis was performed for leaf shape variation, and estimates of genetic diversity and structure, and individual Bayesian genetic assignments were obtained with the program STRUCTURE. Reference populations formed three completely distinct groups according to discriminant function analysis based on the morphological data, and showed low, but significant, genetic differentiation. Populations from the zone of contact contained

individuals morphologically intermediate between pairs of species in different combinations, or even among the three species. The Bayesian admixture analysis found that three main genetic clusters best fitted the data, with good correspondence of reference populations of each species to one of the genetic clusters, but various degrees of admixture evidenced in populations from the contact area. The three oak species have formed a complex hybrid zone that is geographically structured as a mosaic, and comprising a wide range of genotypes, including hybrids between different species pairs, backcrosses and probable triple hybrids.

For chapter two I evaluated the phylogeographic structure, the cytoplasmic exchange dynamics and ecological niche modeling in the red oak species *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla* in all their geographic distribution in Mexico. In total we analyzed 434 individuals from 51 populations with eight chloroplast microsatellite loci, and then was estimated parameters within and among genetic diversity and structure, as well as the possible signals of historical demographic expansion. Also, the potential distribution models were obtained for the present-day distribution and two models for the last glacial maximum using CCSM and MIROC models which are implemented in the program MAXENT. The results obtained indicated higher levels for within and between total population genetic diversity, haplotype richness and the pairwise genetic distance within-population (D^2sh), in general were observed higher values for the three species *Q. hypoleucoides* with values from $h_S = 0.671$, $h_T = 0.932$, for *Q. scytophylla*, with values of $h_S = 0.708$ and $h_T = 0.972$ and for *Q. sideroxyla*, the values $h_S = 0.695$, $h_T = 0.981$ respectively. Also we found evidence for a cytoplasmic exchange invasion from *Q. hypoleucoides* haplotypes, toward *Q. sideroxyla* and *Q. scytophylla* mainly in populations that are in the zones contact. We observed an increase of genetic diversity of some populations in the zones of contact, although; despite the hybridization also maintain a moderate genetic differentiation in their populations. The three species had a strong phylogeographic structure and a north-south colonization routes in their main lineages throughout Mexico. We observed that populations with signals of expansions have been interpreted to populations that have served as a reservoir of genetic diversity during climate oscillations for a relatively large population. We can suggest that the species experienced a moderate reduction in their distribution for the LGM models, it is quite possible that populations experienced altitudinal displacements rather than latitudinal migrations or colonizations as responses to the different glacial periods and, as a

consequence high genetic variation levels and local haplotypes were retained, which implies large effective population sizes and a complex dynamic in the cytoplasmic exchange, thus allowing the maintenance o certain levels of connectivity in their populations.

For chapter three, the objectives were to infer the phylogeographic structure of the red oak species *Quercus castanea* with a wide geographical distribution in Mexico, as well as testing the possible signals of population expansions, and finally reconstruct the models for present and potential distribution for the last glacial maximum (LGM). We collected a total of 341 individuals from 36 populations of *Q. castanea* localized in four commonly recognized biogeographic and morphotectonic provinces of Mexico: the Sierra Madre Occidental (SMO), the Central Plateau (CP), the Trans-Mexican Volcanic Belt (TMVB) and the Sierra Madre del Sur (SMS). We utilized seven chloroplast microsatellite loci (cpSSRs), and data were analyzed to obtain standard statistics of genetic diversity and structure and to test for signals of historical demographic expansions. Finally, potential distribution models were obtained for present-day and two scenarios of LGM climatic conditions using the MAXENT algorithm. We found a total of 90 haplotypes were identified, with 28 of these restricted to single populations. Within-population genetic diversity was high (mean $h_S = 0.72$) and among-population genetic differentiation showed a strong phylogeographic structure ($R_{ST} = 0.711 > G_{ST} = 0.261$; $P < 0.05$). Signals of demographic expansion were identified in most populations, with larger τ (time to the population expansion) values in the SMS, followed by the SMO, the TMVB, and finally the CP. Ecological niche models suggested a similar but moderately larger (30-50%) potential distribution extent for the species under LGM conditions than at present. In conclusion, high genetic diversity, strong phylogeographic structure and potential distribution models suggest in situ permanence of *Q. castanea* populations with large effective sizes during the last thousands years. Signals of population expansions probably predating the LGM and possibly as old as 1.6×10^6 to 1.9×10^5 years before present were observed and suggested that the expansion occurred following a south-north direction. As with other Mexican tree species, the complex geologic and climatic history of the TMVB contributes to explain the origin and maintenance of a large proportion of the genetic diversity in this oak species.

III. INTRODUCCIÓN GENERAL

El género *Quercus* (Fagaceae) incluye árboles o arbustos, polinizados por viento, inflorescencias unisexuales y bellotas como frutos característicos (Nixon, 1993). Este género cuenta con aproximadamente 500 especies, que se encuentran en casi todos los bosques templados del hemisferio Norte, así como en algunas regiones tropicales y subtropicales del mismo (De Greef *et al.* 1998). Existen incluso algunas especies en hábitats más secos, como en el sureste de Asia y nororiente de África. En América se localizan desde Canadá, Estados Unidos y México hasta Colombia, incluyendo Cuba (Valencia, 2004). Este género se encuentra constituido por dos subgéneros *Ciclobalanopsis* y *Quercus* (Nixon, 1993). Existen dos centros de diversidad para el género *Quercus*, el primero se localiza en el sureste de Asia con 125 especies, y el segundo se localiza en México donde se encuentran aproximadamente 135-150 especies (Nixon, 1993). Recientemente, Valencia (2004) en una revisión taxonómica presentó una lista donde se consideran 161 especies de encinos en México, así, en nuestro país están representadas el 32-40% de las especies del género, lo que equivale a una tercera parte con respecto a la diversidad mundial. De los dos subgéneros (*Cyclobalanopsis* y *Quercus*) en México solo se encuentra representado *Quercus*, el cual está constituido por tres secciones: *Quercus* (encinos blancos) con 81 especies y 47 endémicas, *Lobatae* (encinos rojos) con 76 especies y 61 endémicas y *Protobalanus* (encinos intermedios) con 4 especies y 1 endémica (Nixon, 1993; Valencia, 2004). Estas cifras destacan una gran diversidad y endemismo, debido principalmente a las condiciones tan particulares de México como son una notable heterogeneidad de hábitats, variabilidad climática y una compleja historia geológica, que ha favorecido en gran manera el origen y diversificación de un número considerable de linajes de plantas en nuestro país (Rzedowski, 1993).

En México las especies de encinos se distribuyen desde el nivel del mar hasta los 3500 m (Rzedowski, 1993) y constituyen un elemento característico de los encinares, siendo muy importantes también en los bosques de pino, de pino-encino, mesófilos de montaña y en algunas variantes del matorral xerófilo, entre otros (González, 1993). En su distribución altitudinal y latitudinal los encinos se encuentran mejor representados en las zonas montañosas de México, principalmente en la Sierra Madre Occidental, Sierra Madre Oriental, Faja Volcánica Transmexicana, Sierra Madre del Sur, Sierras del Norte de Oaxaca, de Chiapas y Baja California (Rzedowski, 1978). En el aspecto ecológico los encinos son considerados comunidades vegetales muy características que, junto con los pinares constituyen la mayor parte de la cubierta vegetal en los bosques templados (Nixon, 1993). Asimismo, los bosques de encino son el hábitat de una gran cantidad de plantas epífitas y de animales vertebrados e invertebrados asociados estrechamente con los encinos (Valencia, 2004). De manera más particular se considera su importancia económica por sus productos maderables y ornamentales (Rzedowski, 1978).

Diversificación del género *Quercus*

Los procesos de diversificación en la familia Fagaceae iniciaron desde el Oligoceno inferior (30 millones de años), en donde se ubican los géneros fósiles *Castaneoides*, *Trigonobalanoides*, Quercoides y *Fagoides* (Crepet y Nixon, 1989), los cuales se distribuyeron por Asia, Europa, y América del Norte. En América del Norte el género *Quercus* logró una amplia distribución durante el Eoceno superior al Oligoceno inferior por lo que la diversificación de las especies de *Quercus* y su subsecuente dispersión ocurrió esencialmente en el Oligoceno superior (Crepet y Nixon, 1989). Axelrod (1983) propone que los encinos blancos de la sección *Quercus* evolucionaron durante el oligoceno dentro de América del Norte y que posteriormente migraron por el puente

terrestre del estrecho de Bering durante el Oligoceno superior a Asia y Europa, lo cual ocurrió hace aproximadamente 17 millones de años (Axelrod, 1983). Estos patrones de distribución pasada confirman la hipótesis general de la diversidad de especies actualmente presentes en Norte América y México con 125 especies de encinos blancos (i.e. sección *Quercus*) y donde además se encuentran los fósiles más antiguos, mientras que en Europa y Asia sólo existen 20 especies (Valencia, 2004).

Posteriormente Nixon (1993) propone que los encinos blancos entraron a México durante el Oligoceno-Mioceno migrando a través del este de Estados Unidos hasta la Sierra Madre Oriental en el este de México y posteriormente dispersándose hacia el sur de México vía la Sierra Madre del Sur. Se han encontrados fósiles de hojas de encinos blancos en los registros del Mioceno-Plioceno en el Paraje Solo Veracruz, los cuales muestran el paso de los encinos blancos por esta región (Suter, 1984). Esta información coincide con los registros actuales de diversidad de encinos blancos según los cuales existe una gran concentración de especies en las sierras de Nuevo León y Tamaulipas (Sierra Madre Oriental) y de la misma manera en la Sierra Madre del Sur (Nixon, 1993). Esta evidencia sugiere un escenario de dispersión de encinos blancos a través de estas dos sierras, desde el Oligoceno-Mioceno, época en que estas dos provincias se encontraban en pleno proceso de formación de sus rangos montañosos y pudieron permitir la colonización de estas especies de encinos (Graham, 1993). Los registros de polen y de hojas fósiles atribuibles a encinos blancos también nos permiten suponer una segunda dispersión de estos encinos a través de la Sierra Madre Occidental, como la demuestra la evidencia de fósiles de bellotas de *Quercus* en los registros del Eoceno de Oregon (Manchester, 1984). Esta evidencia coincide con el levantamiento de la Sierra Madre Occidental para el Oligoceno temprano (28 millones de años), y Mioceno temprano (20 millones de años), en los cuales podemos observar ya algunas de

las conformaciones montañosas a lo largo de toda la Sierra Madre Occidental (Rossotti *et al.* 2002).

En cuanto al origen y diversificación de los encinos rojos (*Lobatae*) según los estudios de Trelease (1924), Manos *et al.* (1999) y Manos y Stanford (2001) ellos sugieren que los encinos Americanos pertenecientes a la sección *Protobalanus* son los ancestros de las especies americanas de la sección *Lobatae* (encinos rojos) y sección *Quercus* (encinos blancos) con base en las características morfológicas y del análisis filogenético molecular. Incluso sugirieron que los eventos de divergencia entre encinos blancos y rojos se llevaron a cabo durante el Oligoceno-Mioceno (20 millones de años). De esta manera, el más reciente linaje de encinos rojo (*Lobatae*) pudo dispersarse por México a través de la Sierra Madre Occidental y Oriental que en esta época ya tenía formados sus rangos montañosos, por lo que representó un puente natural entre las biotas de América del Norte y del occidente de México vía las Montañas Rocallosas. Según Axelrod (1983) durante el Oligoceno los encinos rojos (sección *Lobatae*) y blancos (sección *Quercus*) tuvieron una rápida evolución y dispersión desde América del Norte, principalmente desde climas de condiciones más templadas y secas hacia climas con condiciones más cálidas y secas como las que ocurren en el suroeste de Estados Unidos, en México en la Sierra Madre Occidental, Oriental y Planicie Central de México, hasta Centroamérica. Posteriormente y con el surgimiento del Faja Volcánica Transmexicana la cual comenzó su desarrollo durante el Terciario temprano, pero experimentó su levantamiento principal y formación desde el Mioceno hasta el Cuaternario, hace 2.4 millones de años (Crepet y Nixon, 1989; Rossotti *et al.* 2002). Ésta probablemente sirvió de puente de enlace y de dispersión de encinos entre las Sierras Madre Occidental y Oriental, y sirvió para delimitar la abundancia de especies en el centro y sur de México, donde encontramos la mayor diversidad de especies de

encinos reportadas para el país con 60-75 especies (González, 1993; Nixon, 1993). Sin embargo, entre más al sur como por ejemplo hacia al sur en Chiapas y América Central la diversidad de especies disminuye dramáticamente hasta tan solo 26 especies reportadas para estas regiones (Breedlove, 1984).

IV. ANTECEDENTES

Filogeografía en encinos

La cuantificación de la variación genética dentro y entre especies es esencial para explicar la distribución de la diversidad biológica (Schaal *et al.* 1998). La distribución de esta variación genética está frecuentemente regulada por factores ecológicos, procesos de selección o por los patrones de intercambio genético entre las poblaciones (Avise, 1998; Bermingham y Moritz, 1998). Los estudios filogeográficos proveen análisis poderosos para esclarecer las relaciones genealógicas entre los haplotipos así como también su distribución geográfica (Avise, 2000). Además, pueden potencialmente determinar procesos como cuellos de botella o expansiones de área geográfica en las poblaciones (Schaal *et al.* 1998), así como también procesos de intercambio citoplasmático en las poblaciones de las especies (Dumolin-Lapegue *et al.* 1997; Le Corre y Kremer, 1998; Belahbib *et al.* 2001). Incluso, se han detectado los efectos de las fluctuaciones climáticas ocurridas en el Pleistoceno en la distribución de la diversidad y estructura genética de los bosques templados de Europa y Norteamérica (Petit *et al.* 2002; Magni *et al.* 2005).

Sin embargo, pocos estudios han incorporado datos medioambientales en la interpretación filogeográfica (Hugall *et al.* 2002; Cilibertii *et al.* 2009; Jakob *et al.* 2009). Recientemente, se ha dado un mayor énfasis al desarrollo de metodologías cuantitativas en la filogeografía en un intento de proveer una mayor confidencia

estadística para los patrones observados (Knowles y Madison, 2002; Manni *et al.* 2004). La implementación de los modelos de nicho ecológico como una herramienta para predecir las distribuciones presentes y pasadas, han provisto un punto de vista complementario para comparar con los datos filogeográficos y las hipótesis biogeográficas tradicionales (Hugall et al 2002; Waltari *et al.* 2007). Los modelos de nicho ecológico incorporan modelos que reconstruyen la distribución potencial de las especies basados en las coordenadas geográficas de las localidades, y que son extrapoladas a áreas que representen las condiciones ecológicas más adecuadas para las especies (Guisan y Thuiller, 2005; Richards *et al.* 2007). Estos modelos toman en cuenta variables como la temperatura, precipitación, elevación y otras variables que podrían potencialmente influenciar la distribución de las especies (Phillips *et al.* 2006). De esta manera la combinación de la filogeografía y los modelos de nicho ecológico juegan un papel preponderante ya que responden preguntas que no se pueden abordar en cada una de las áreas por separado. El modelado de nicho ecológico ha sido aplicado a los campos de la ecología, biogeografía, genética del paisaje, biología de la conservación entre otros (Guisan y Thuiller, 2005; Peterson *et al.* 2007).

Particularmente, en las especies del género *Quercus* se han realizado estudios extensivos en Europa y han revelado los efectos de las fluctuaciones climáticas del Cuaternario en la distribución y conectividad de las poblaciones, y han indicado la existencia de refugios Pleistocénicos situados en la Península Ibérica, Italia y los Balcanes, así como una fuerte estructura genética, como resultado de varias rutas de recolonización postglacial (Dumolin-Lapegue *et al.* 1997; Le Corre y Kremer, 1998; Belahbib *et al.* 2001; Petit *et al.* 2002). En contraste, las especies de encinos norteamericanos son caracterizados por una baja estructura genética (Magni *et al.* 2005) o por una alta diversidad genética (Grivet *et al.* 2006) a diferencia de los encinos

europeos (Marsico y Hellman, 2009). Estas diferencias han sido explicadas como el resultado de diferentes historias poblacionales en las especies de encinos de ambos continentes. Probablemente debido a que en las poblaciones de encinos de América del Norte, los cambios climáticos durante las glaciaciones del Pleistoceno fueron menos severas por lo tanto las especies de encinos fueron más persistentes localmente, con algunos eventos de expansión y contracción de rango, pero sin una extensiva colonización post-glacial (Magni *et al.* 2005; Grivet *et al.* 2006).

Sin embargo, poco sabemos de la historia de las especies de encinos en las regiones tropicales y subtropicales donde se encuentra las más alta riqueza de especies. Particularmente, México es considerado como un centro de diversificación secundario para el genero *Quercus* con aproximadamente 161 especies que representan cerca de 32-40% de la diversidad mundial (Valencia, 2004). Algunos estudios filogeográficos previos realizados específicamente en dos complejos de encinos rojos mexicanos como *Q. affinis-Q. laurina* y *Q. crassipes-Q. crassifolia*, han mostrado altos niveles de diversidad genética y una baja estructura genética en comparación con las especies de encinos europeos y norteamericanos (González-Rodríguez *et al.* 2004; Tovar-Sánchez *et al.* 2008). Estos resultados implican tamaños efectivos históricos de las poblaciones grandes y una dinámica compleja en el intercambio citoplasmático entre poblaciones. Es posible que en las latitudes tropicales los periodos glaciares durante el Pleistoceno hayan permitido la expansión de área y el incremento en la conectividad de las especies de los bosque templados, tal como ha sido sugeridos para algunas especies de pinos (Moreno-Letelier y Piñero, 2009; Rodríguez-Banderas *et al.* 2009). Sin embargo, los estudios palinológicos que han evaluado la cantidad y la intensidad de las fluctuaciones en el clima durante el Pleistoceno aún son incompletas (Lozano-García *et al.* 2005; Metcalfe, 2006). Los registros climáticos sugieren una reducción de la temperatura

promedio de 6° C durante el ultimo glaciar máximo (Metcalfe, 2006) con un incremento en los casquetes de los volcanes de aproximadamente 1200 m por debajo de la línea actual (Lachniet y Vázquez-Selem, 2005). Incluso, los registros palinológicos indican que las hierbas alpinas colonizaron extensivamente áreas con condiciones frías y áridas, mientras que bajo condiciones climáticas con una alta humedad los bosques templados se expandieron hacia altitudes bajas que en la actualidad (Lozano-García *et al.* 2005; Metcalfe, 2006). En resumen, los cambios climáticos durante el Pleistoceno fueron muy seguramente menos severos en los trópicos que en otras regiones de Europa y América del Norte (Grivet *et al.* 2005; Magri *et al.* 2006). Por lo que posiblemente el impacto en las poblaciones de árboles fue significativamente menor y posiblemente más heterogéneo, debido a que el clima depende tanto de la latitud como de la altitud (Lachniet y Vázquez-Selem, 2005; Metcalfe *et al.* 2000). En consecuencia, es necesario realizar estudios adicionales de especies de encinos mexicanos utilizando estas metodologías para tratar de entender el impacto de las glaciaciones del Pleistoceno en las comunidades de bosques templados de México.

Hibridación en *Quercus*

Se ha planteado que los encinos tienen una tendencia a formar híbridos de manera más frecuente que en otros grupos de plantas (Palmer, 1948). La hibridación implica el apareamiento exitoso en la naturaleza entre individuos de poblaciones que pertenecen a especies que son ecológica, fisiológica y morfológicamente diferentes (Arnold, 1992). Este fenómeno es un importante mecanismo de migración de genes de una especie a otra, que incrementa la variación genética y morfológica que se presenta en los encinos y que en algunos casos dificulta el poder diferenciar de manera correcta algunas especies de encinos (Nixon, 1993). Los híbridos en *Quercus* ocurren comúnmente en la

naturaleza entre especies cercanamente emparentadas (Craft *et al.* 2002), aunque también se han documentado la existencia de híbridos en especies no tan cercanas filogenéticamente (Belahbib *et al.* 2001). La hibridación resulta, en algunos casos en árboles usualmente solitarios o en individuos intermedios agregados en pequeños grupos. (Silliman y Leisner, 1958; Hardin, 1975). Aunque, en algunos casos los individuos en las zonas de hibridación pueden estar ampliamente distribuidos y muestran una variabilidad bastante alta, por lo que se les ve frecuentemente colonizando nuevos hábitats que no ocupaban los progenitores (Vila y D´Antonio, 1998). En este sentido, la hibridación puede influenciar la evolución de los híbridos en las poblaciones de diferentes maneras: puede causar la fusión de las especies progenitoras (Seehausen, 2004), introgresión direccional de una de las especies hacia la otra (Turelli *et al.* 2001) o el reforzamiento de las barreras reproductivas entre especies incompletamente aisladas (Arnold, 1992). La transferencia de material genético entre las especies que hibridan puede facilitar su evolución adaptativa y finalmente el aumento de la variabilidad genética de estas (Grant y Grant, 1992). Sin embargo, es importante poder discernir con claridad la naturaleza de las zonas de hibridación, pudiendo definirlas como zonas de contacto primario y secundario (Hewitt, 1988). En las zonas de contacto primario, por ejemplo un gradiente en el medio ambiente favorece diferentes alelos en los extremos produciendo una clina morfológica, que podría ser gradualmente transformada por medio de selección natural modificando los genes hasta convertirlo en una zona de hibridación estrecha, entre dos genotipos internamente adaptados (Hewitt, 1988; Holman *et al.* 2003; Seehausen, 2004). Las zonas de contacto secundario son aquellas en la que dos especies que han permanecido geográficamente aisladas, posteriormente se encuentran simpátricamente, formando una zona de contacto donde hibridan (Hewitt, 1988; Ishida *et al.* 2003). En resumen, se ha observado que en las zonas de hibridación

son a menudo mantenidas para adaptarse a nuevos medioambientes, y por lo tanto les ha conferido mejores oportunidades en los procesos de adaptación y en los complejos mecanismos de especiación (Rieseberg, 1997). Sin embargo, aún no existen muchos estudios que involucren complejos que incluyan más de dos especies de encinos hibridando simultáneamente (Manos y Fairbrothers, 1987; Dodd y Afzal-Rafii, 2004; Curtu *et al.* 2007; Lepais *et al.* 2009). En estos casos, se generan genotipos F1 o retrocruzas (parcialmente fértiles) entre al menos un par de especies, los cuales pueden dispersarse a través de la zona de hibridación donde las cruzas ocurren con otra especie distinta a las progenitoras, o entre híbridos de diferentes combinaciones de especies (Kaplan y Fehrer, 2007). Estas complejas zonas de hibridación (de mosaico en su mayoría) dependen del balance entre el flujo génico de los progenitoras hacia los híbridos, de la extensión de la zona de contacto entre las especies y de la selección que exista en los genotipos intermedios y progenitoras en la permanencia de la zona de hibridación (Dodd y Afzal-Rafii, 2004).

Los encinos en México

En un contexto histórico, los encinos en México han sido caracterizados por una gama de cambios en su evolución: expansión y contracción en su distribución geográfica, como producto de sucesivos cambios climáticos a lo largo del tiempo, los cuales han fluctuado en el clima desde el Pleistoceno hasta la transición del Holoceno hace cerca de 11000 años (Martínez-Hernández, 1992; Metcalfe, 2006; Lozano-García & Xelhuantzi-López, 1997). Este proceso puede ser especialmente importante en lugares como México, donde la diversidad en topografía, clima y hábitat resultan en que las especies de encinos ecológica y morfológicamente diferenciadas, ocurran en simpatría y formen zonas de contacto (Bacon y Spellenberg, 1996). Tal parece que esta variación

climática puede haber incentivado la hibridación como un mecanismo importante en la evolución de los encinos (Arnold, 1992). Se ha reportado la aparición de zonas de contacto e hibridación, en algunas especies de encinos rojos que ocurren en la Sierra Madre Occidental, por ejemplo *Q. eduardii* y *Q. conzatti* (Bacon y Spellenberg, 1996), así como la especie híbrida *Q. basaseachicensis* y sus progenitores *Q. depressipes* y *Q. rugosa* (Spellenberg, 1995). De la misma forma existen algunas especies de encinos blancos que morfológicamente son muy similares y a menudo se les encuentra ocurriendo simpátricamente, por lo que se supone que estas especies hayan formado zonas de contacto donde forman enjambres de híbridos, este grupo de especies son: *Q. arizonica*, *Q. grisea*, *Q. laeta*, *Q. convallata*, *Q. obtusata*, *Q. chihuahuensis*, *Q. transmontana*, y *Q. striatula* (Bacon y Spellenberg, 1996). De la misma manera existen reportes de la especialista Susana Valencia, de otras especies simpátricas de encinos rojos como son *Q. castanea*, *Q. hypoleucoides*, *Q. scytophylla* y *Q. sideroxyla* que morfológicamente son muy similares en el norte de México por lo que presumiblemente han formado zonas de contacto donde hibridan entre si. Podemos pensar que los procesos de hibridación en el norte de México son más comunes por lo que podemos suponer que efectivamente ciertos procesos sucedidos en la Sierra Madre Occidental, hayan promovido patrones de evolución en encinos que resultan muy interesantes de dilucidar y por supuesto proponer escenarios que nos arrojen luz para poder comprender paso a paso la complicada historia de los encinos en México (Bacon, 1996).

Especies estudiadas

Se estudiaron cuatro especies de encinos rojos mexicanos (sección *Lobatae*): *Quercus castanea*, *Q. hypoleucoides*, *Q. scytophylla* y *Q. sideroxyla*, estas especies pertenecen a series diferentes de acuerdo con clasificaciones existentes por lo que muestran caracteres morfológicos diagnósticos, bien delimitados entre ellas (Trelease, 1924). Sin embargo, en el norte de la Sierra Madre Occidental se han encontrado inconsistencias morfológicas por lo que se les ha cambiado de nombre de especie en más de una ocasión, incluso se ha observado que presentan morfología intermedia cuando se sobrelapan espacialmente en algunas localidades en la Sierra Madre Occidental. Al parecer, las especies se sobrelapan no solo espacialmente, sino también lo hacen en su distribución altitudinal (1800 a 2500m) y en sus periodos de floración (Marzo a Junio). Por lo que posiblemente la interacción entre estas especies haya generado una zona de contacto secundario en la Sierra Madre Occidental. Haciendo una revisión rigurosa del material de herbario y de las colectas de las poblaciones de las cuatro especies que ya se habían realizado y con el apoyo de la especialista Susana Valencia, se pudieron identificar de manera eficiente que las poblaciones en el norte de México son el producto de la hibridación entre *Q. hypoleucoides*, *Q. scytophylla* y *Q. sideroxyla*. De esta manera en el capítulo uno y dos se trataron de evidenciar los procesos de hibridación en la Sierra Tarahumara haciendo uso de marcadores moleculares como microsatélites nucleares, de cloroplasto y metodologías para medir la variación en la forma de la hoja en las poblaciones de las tres especies.

En el primer capítulo se presenta el análisis morfológico y genético de la hibridación entre las especies de encinos rojos *Q. hypoleucoides*, *Q. scytophylla* y *Q. sideroxyla* en la Sierra Tarahumara, México. Las preguntas particulares de esta capítulo fueron (1) ¿Existe realmente un proceso de hibridación e introgresión entre estas

especies?, (2) ¿Cuál es la estructura geográfica de la variación morfológica y genética de las tres especies en la zona de contacto?, (3) ¿Existen asociaciones entre la variación fenotípica y molecular con el gradiente altitudinal que comprende la zona de hibridación?

En el segundo capítulo se analizó la estructura filogeográfica y se determinaron los patrones de intercambio citoplasmático entre las especies de encinos rojos *Q. hypoleucoides*, *Q. scytophylla* y *Q. sideroxyla* en su distribución en México. Este sistema es muy interesante ya que no existen muchos estudios que evalúen el impacto de las fluctuaciones climáticas durante el Pleistoceno en su diversidad y estructura genética, así como su historia demográfica de sus poblaciones. Los objetivos específicos fueron i) estimar la diversidad y estructura genética de poblaciones de las tres especies de encinos rojos a lo largo de su distribución geográfica en México, utilizando microsatélites de cloroplasto, ii) reconstruir los patrones filogeográficos de las tres especies, la historia demográfica de las poblaciones, e inferir los modelos de nicho ecológico para su distribución actual y durante el ultimo glaciar máximo, para iii) entender la dinámica de las poblaciones en sus áreas de contacto y conocer los procesos históricos que ha operado en la evolución de estas especies.

En cuanto a la especie *Q. castanea* la cual no hibrida en el norte de México y de la que ya se contaba con la mayoría de las poblaciones en su distribución geográfica se decidió reconstruir la historia biogeográfica y demográfica de la especie de encino rojo *Quercus castanea*, para lograr una mayor comprensión de la historia de la poblaciones de árboles templados de montaña, y sus respuestas a los últimos cambios climáticos en México. *Q. castanea* constituye un sistema de estudio interesante ya que tiene una de las más amplias distribuciones geográficas y altitudinales entre los encinos rojos Mexicanos. Los objetivos específicos de este estudio fueron (i) determinar los patrones

de diversidad y diferenciación genética de las poblaciones de *Q. castanea* en toda su distribución geográfica en México, usando microsatélites de cloroplasto como marcador molecular, (ii) probar si han existido señales de expansión demográfica histórica, y iii) contrastar las conclusiones derivadas de los datos genéticos con los modelos de la distribución potencial actual y durante el último glaciar máximo (LGM) de *Q. castanea*.

Quercus castanea Nee. Es una especie que presenta gran variación morfológica por lo que ha sido previamente descrita varias veces bajo diferentes nombres que ahora son consideradas como sinonimias de la especie (Figura 1), las cuales son: *Q. alamosencis, Q. axilaris, Q. castanea ssp. sublobata, Q. circummontana, Q. consociata, Q. crassivenosa, Q. impressa, Q. lanigera, Q. mucronata, Q. pulchella, Q. rossii, Q. seleri, Q. serrata, Q. serrulata, Q. spathulistipula, Q. subscrispata, Q. tepoxuchilensis, Q. tristis* y *Q. verrucosirama* (McVaugh, 1974). Son árboles de hasta de 18 m de alto, tronco de hasta 0.4 m de diámetro, ocasionalmente arbustos. Hojas maduras con pecíolos de 2-12.5 mm de largo y 0.4-0.7 mm de ancho, glabrescentes con algún indumento canosos; laminas coriáceas, obovadas a oblanceoladas o elípticas con 3-6 dientes cortos, cortamente aristados distribuidos en el último tercio o último medio apical de la hoja, haz glabro, lustroso, ligeramente rugoso, envés ampuloso, semiglabro o cubierto con mechones de pelos crispados o estipitados, dispuestos laxamente o llegando a cubrir casi totalmente la superficie. Flores masculinas sésiles o sobre un pedicelo de 1.5 mm de largo apretadamente dispuestas sobre el ráquis; flores femeninas sésiles o casi sésiles, dispuestas en las axilas de las hojas, en pares o rara vez solitarias. Frutos de maduración anual en pares o solitarios, a menudo asociado a las axilas de las hojas; cúpula hemisférica o ligeramente turbinada; bellota elíptica a anchamente elíptica, glabrescentes o glabras (Figura 1 y 2). Esta especies tiene un periodo de

floración de abril a mayo, frutos maduros de julio a noviembre. Los individuos de esta especie se localizan en bosques de *Quercus*, bosque tropical caducifolio; con menos frecuencia se presentan en el bosque mesófilo de montaña y en el bosque de coníferas y *Quercus* en altitudes de 1180-2600msnm. Tiene una distribución geográfica desde México hasta Centro América. En México se encuentra en los estados de Colima, Chiapas, Distrito Federal, Durango, Guanajuato, Guerrero, Hidalgo, Jalisco, Estado de México, Michoacán, Morelos, Nayarít, Oaxaca, Puebla, San Luis Potosí, Sinaloa, Sonora y Veracruz (McVaugh, 1974; Valencia, 1995).

Figura 1. Árbol representativo de la especie *Q. castanea,* Zapotlanejo, Jal (Foto, JM Peñaloza-Ramírez)

Figura 2. Acercamiento a las hojas y frutos de la especies *Q. castanea* (Foto, JM Peñaloza-Ramírez)

Quercus hypoleucoides A Camus. Es una especie que ha sido descrita como especie diferente bajo las sinonimias de *Q. hypoleuca,* y *Q. confertifolia* (Trelease, 1924). Son árboles medianos de 10-20m de altura, de corteza dura, hendida y negra. Hojas más o menos persistentes, variables, angostamente lanceoladas, elípticas, largamente ovadas o subobovadas, agudo-aristadas, de 5-8 y hasta 11cm de largo por 1.5-3 de ancho; haz liso o finamente estrellado-pubescente, de color glauco; envés con denso tomento blanquecino; nervaduras unas 10 a cada lado, con algunas intermedias qué se desvanecen, ramificadas y vagamente anastomosadas cerca del borde (Figura 3). Amentos masculinos de 3-6 cm de largo, blanco-pilosos; anteras elipsoides y salientes: amentos femeninos de 5 a 100 mm de largo, con 1 a 2 flores en la extremidad de un pedúnculo estrellado-tomentoso. Fruto de maduración anual solitario o en pares, sésil o casi sésil: cúpula algo turbinada, de unos 10 mm de diámetro; bellota oblonga a ovoide

de 12-14mm de largo por 7-9 mm de ancho (Figura 3 y 4). Esta especie tiene su periodo de floración de marzo-junio, frutos maduros de junio a octubre, es frecuente encontrarla en bosque de encino, bosques mixtos de pino-encino en altitudes de 2000-2500 m. Esta especie tiene una distribución geográfica en México en los estados de Sonora, Chihuahua, Coahuila y Durango, aunque también se le encuentra en el suroeste de Estados Unidos en los estados de Arizona, Nuevo México y Texas (McVaugh, 1974).

Figura 3 Árbol representativo de la especie *Q. hypoleucoides*, Batopilas, Chih (Foto, JM Peñaloza-Ramírez)

Figura 4. Acercamiento de hojas de la especie *Q. hypoleucoides* (Foto, JM Peñaloza-Ramírez).

Quercus scytophylla Liebm. Es una especie que ha a sido descrita al menos tres veces bajo las sinonimias de *Q. campanariensis*, *Q. epileuca* y *Q. incarnata* (Trelease, 1924). Son árboles de hasta de 20 m de altura y tronco de 0.7 m de diámetro (Figura 5). Hojas maduras con pecíolos de 9-35 mm de largo por 0.9-2.9 mm de ancho, ligeramente pardo-rojizos, glabrescentes, con mechones de pelos crispados sésiles; margen cartilaginoso con 1-6 dientes aristados que se distribuyen a cada lado de la hoja hacia la mitad distal; haz opaco, glabro, ocasionalmente con restos de mechones de pelos sésiles y púberulo pardo, envés cubierto con mechones de pelos sésiles crispados o rectos que ocultan totalmente a la epidermis ampulosa y papilosa. Flores masculinas regularmente distribuidas sobre el ráquis, gamotépalas, ocasionalmente imbricados. Flores femeninas generalmente en grupos de 2 de las axilas de las hojas, casi sésiles, ligeramente pilosas, estigmas cortos y ocasionalmente rodeados del perianto fusionado. Frutos de maduración bianuales solitarios o en grupos de 2-4; cúpulas obcónicas a depresamente ovoide; bellotas ovoides con residuos de tomento canoso (Figura 5 y 6). Esta especie

presenta sus periodos de floración entre febrero y marzo y tiene frutos maduros de julio a septiembre. Es frecuente encontrar a esta especies formando parte del bosque mesófilo de montaña, aunque también se encuentra en los bosques de encinos y de pino-encino, incluso se le ha encontrado en algunas transiciones con el bosque tropical seco, en altitudes de entre 1540-2600 msnm. Esta especie tiene una amplia distribución geográfica en los estados de Chihuahua, Durango, Sonora, Sinaloa, Zacatecas, Jalisco, Guerrero, Estado de México, Michoacán, Oaxaca y Chiapas (McVaugh, 1974; Valencia, 1995).

Figura 5. Individuo de la especie *Q. scytophylla*, Volcán Ceboruco, Nayarít (Foto, JM Peñaloza-Ramírez)

Figura 6. Acercamiento de hojas de *Q. scytophylla*, Volcán Ceboruco Nayarít (Foto, JM Peñaloza-Ramírez)

Quercus sideroxyla Humb. & Bonpl. Es una especie que ha sido descrita al menos dos veces con las sinonimias de *Q. omissa, Q. epileuca* (Trelease, 1924. Es por lo general un árbol pequeño de 7-8 m de alto, de diámetro delgado de 250-300 cm (Figura 7). hojas tardíamente decíduas muy engrosadas y rígidas, oblanceoladas, obovadas u oblongo elípticas, por lo general presentan de 1-5 aristas o dientes deltoides a cada uno de la hoja, generalmente confinados a la tercera parte distal de la hoja, haz duro, ruguloso sin brillo, de color verde seco amarillento, finamente glabro, envés con un denso tomento blanco-amarillento que cubre las nervaduras secundarias, sólo las primarias sobresalen entre los pelos; epidermis ampollada y papilosa; amentos masculinos tomentosos, el perianto en forma de copa, con el margen ciliado, las anteras de 1.3-1.6 mm de largo, oblongas y glabras, fruto de maduración bianual solitario o en pares, casi sésil; cúpula hemisférica o con la base prolongada, tomentosa, de color castaño, bellota ovoide con una tercera parte de su largo incluida en la cúpula (Figura 7

y 8). Esta especie tiene su periodo de floración entre junio a agosto y fructifica de noviembre a enero, y se le puede encontrar en bosques de encino, pino-encino. Esta especie es frecuente en altitudes de 2400-2800 msnm asociada a especies de pino de grandes altitudes. Presenta una moderada distribución geográfica en México en los estados de Chihuahua, Sonora, Durango, Zacatecas, Nayarit, Jalisco, Aguascalientes, Guanajuato, San Luis Potosí, Coahuila, Tamaulipas y Nuevo León (McVaugh, 1974).

Figura 7. Individuo representativo de la especie *Q. sideroxyla*, Pueblo Nuevo, Durango (Foto, JM Peñaloza-Ramírez)

Figura 8. Acercamiento a las hojas de *Q. sideroxyla*, El Tecuan, Durango (Foto, JM Peñaloza-Ramírez)

V.

Juan Manuel Peñaloza-Ramírez, Antonio González-Rodríguez, Luis Mendoza-Cuenca, Henri Caron, Antoine Kremer and Ken Oyama

Interspecific gene flow in a multispecies oak hybrid zone in the Sierra Tarahumara of Mexico

Annals of Botany, **105, 389-399. 2010**

Annals of Botany **105**: 389–399, 2010
doi:10.1093/aob/mcp301, available online at www.aob.oxfordjournals.org

ANNALS OF BOTANY
Founded 1887

Interspecific gene flow in a multispecies oak hybrid zone in the Sierra Tarahumara of Mexico

Juan Manuel Peñaloza-Ramírez[1,*], Antonio González-Rodríguez[1], Luis Mendoza-Cuenca[2], Henri Caron[3], Antoine Kremer[3] and Ken Oyama[1]

[1]*Centro de Investigaciones en Ecosistemas, Universidad Nacional Autónoma de México (UNAM), Antigua Carretera a Pátzcuaro No. 8701, Col. Ex-Hacienda la Huerta, 58190, Morelia, Mexico,* [2]*Facultad de Biología, Universidad Michoacana de San Nicolás de Hidalgo, Francisco J. Múgica, Colonia Felícitas del Río, C.P. 58030, Morelia, México and* [3]*UMR Biodiversity, Genes & Ecosystems, Institut National de la Recherche Agronomique, 69 route d'Arcachon, F-33612 Cestas cedex, France*
** For correspondence. E-mail jpenaloza@oikos.unam.mx*

Received: 16 July 2009 Returned for revision: 7 October 2009 Accepted: 17 November 2009 Published electronically: 6 January 2010

• *Background and Aims* Interspecific gene flow can occur in many combinations among species within the genus *Quercus*, but simultaneous hybridization among more than two species has been rarely analysed. The present study addresses the genetic structure and morphological variation in a triple hybrid zone formed by *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla* in north-western Mexico.
• *Methods* A total of 247 trees from ten reference and 13 presumed intermediate populations were characterized using leaf shape variation and geometric morphometrics, and seven nuclear microsatellites as genetic markers. Discriminant function analysis was performed for leaf shape variation, and estimates of genetic diversity and structure, and individual Bayesian genetic assignments were obtained.
• *Key Results* Reference populations formed three completely distinct groups according to discriminant function analysis based on the morphological data, and showed low, but significant, genetic differentiation. Populations from the zone of contact contained individuals morphologically intermediate between pairs of species in different combinations, or even among the three species. The Bayesian admixture analysis found that three main genetic clusters best fitted the data, with good correspondence of reference populations of each species to one of the genetic clusters, but various degrees of admixture evidenced in populations from the contact area.
• *Conclusions* The three oak species have formed a complex hybrid zone that is geographically structured as a mosaic, and comprising a wide range of genotypes, including hybrids between different species pairs, backcrosses and probable triple hybrids.

Key words: Altitudinal cline, hybridization, introgression, leaf shape variation, Mexico, nuclear microsatellites, *Quercus scytophylla*, *Quercus sideroxyla*, *Quercus hypoleucoides*, red oak.

INTRODUCTION

Hybrid zones are considered an important source of genetic recombination and diversity in plant evolution (Arnold, 1997; Rieseberg, 1997). These zones are most often the product of secondary contact between populations or species that have differentiated previously in allopatry, or may also arise *in situ* in response to spatially varying selection (Durrett *et al.*, 2000). Nearly all of the hybrid zones that have been so far studied involve two parental taxa and their hybrids. Nevertheless, more complex instances of simultaneous hybridization among three or more parental taxa also occur in nature, but these have been rarely analysed (Arnold, 1993; Kaplan and Fehrer, 2007). Consequently, the genetic structure and the dynamics of gene flow in multispecies hybrid zones are poorly known (Dodd and Afzal-Rafii, 2004; Curtu *et al.*, 2007; Kaplan and Fehrer, 2007; Lepais *et al.*, 2009). These situations require the production of some fertile hybrid genotypes (F_1 and/or backcrosses) between at least one species pair. Subsequently, these individuals may disperse throughout the hybrid zone and crosses can occur between hybrids and a third species, or between hybrids from different species combinations (Kaplan and Fehrer, 2007).

According to theoretical models, the evolution of hybrid populations will depend on the relative fitness of the various hybrid and parental genotypes (Barton and Gale, 1993; Arnold, 1997). Nevertheless, the nature and intensity of selective factors are not necessarily homogeneous across space. Both exogenous and endogenous selection have been invoked to explain the structure of hybrid zones (Barton and Hewitt, 1985; Moore and Price, 1993). Exogenous selection implies adaptation to local environments, whereas endogenous selection occurs when hybrids have low fitness due to incompatibilities between the parental genomes (Bronson *et al.*, 2003). Therefore, the selective pressures acting on hybrids may result from the combined effects of environmental factors such as climate, soil and interactions with other organisms (Fitzpatrick and Schaffer, 2004; James and Abbott, 2005; Raudnitschka *et al.*, 2007), and factors independent of environmental variation such as the disruption of co-adapted gene complexes and the action of genes associated with sterility (Barton

and Gale, 1993; Rieseberg and Wendell, 1993). If the main selection pressures are exogenous, the likely result is a clinal or mosaic hybrid zone, in which certain genotypes could be spatially segregated according to the distribution of habitats (Endler, 1977). In contrast, if the selection pressures are endogenous, the generation of tension zones is expected, wherein clines are maintained by the equilibrium between the movement of parental individuals and selection against the hybrid genotypes (Barton and Gale, 1993). However, it is not clear which of these models best describe multispecies hybrid zones (Dodd and Afzal-Rafii, 2004).

Although a high frequency of interespecific gene flow in many combinations has been inferred within the genus *Quercus* from morphological variation, relatively few studies have used genetic markers considering several species simultaneously (Whittemore and Schaal, 1991; Dumolin-Lapègue *et al.*, 1999; Dodd and Afzal-Rafii, 2004; Curtu *et al.*, 2007; Lepais *et al.*, 2009). Of particular interest has been the detection of widespread cytoplasmic introgression even among distantly related species at local scales (Whittemore and Schaal, 1991; Dumolin-Lapègue *et al.*, 1999). Nevertheless, few studies have been conducted using nuclear microsatellites, which allow the identification of different genealogical classes and detailed characterization of the genetic structure of hybrid zones (i.e. Gugerli *et al.*, 2008; Lepais *et al.*, 2009).

Mexico is considered a centre of species diversification for the genus *Quercus*, with 161 species (32–40 % of the worldwide diversity). In particular, 76 species of red oaks have been reported in Mexico, including 61 endemic species (Valencia, 2004). In recent years, several studies conducted on Mexican red oaks have focused on hybridization between two species (Valencia and Delgado, 2003; González-Rodríguez *et al.*, 2004, 2005; Tovar-Sánchez and Oyama, 2004; Tovar-Sánchez *et al.*, 2008). However, several oak species complexes comprising three or more species that probably hybridize simultaneously have been described in Mexico (Dodd and Kashani, 2003; McCauley *et al.*, 2007). In this context, understanding of the evolutionary dynamics and consequences of hybridization requires the consideration of such multispecies interactions as this is the way the phenomenon occurs in nature. The present study investigated a complex of three red oak species formed by *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla*. These species are well delimited by diagnostic morphological characters, although some individuals exhibit morphological intermediacy in different combinations in several localities across the Sierra Tarahumara in north-west Mexico where the species coexist. In this region, the species overlap partially in distribution along an altitudinal gradient (1800–2500 m) and also in the timing of their flowering periods (from March to August) (Valencia, 2004).

The purpose of this study was to analyse this multispecies oak hybrid zone by combining morphometric analysis of leaf shape and nuclear microsatellite molecular markers. The particular questions addressed here were: (1) Do hybridization and introgression occur between *Q. scytophylla*, *Q. sideroxyla* and *Q. hypoleucoides*?; (2) What is the geographical structure of morphological and genetic variation among the three species in the contact area; and (3) are there associations between molecular and phenotypic variation with the altitudinal gradient encompassed by the hybrid zone?

MATERIALS AND METHODS

Study species

The three species belong to the red oak section *Lobatae* (Nixon, 1993). *Quercus scytophylla* Liebm. is a tree about 20 m in height included in the series *Scytophyllae* (Trelease, 1924). The leaves of this species are elliptical or obovate, 5–17 cm in length and 2·5–8 cm in width, with 1–6 teeth with aristae. The petioles are 9–35 mm in length. Acorns are produced singly or in groups and have a peduncle 3–10 mm in length, with maturation periods of 2 years. It is the most widely distributed of the three species in the complex, present in the Sierra Madre Occidental, the Eje Neovolcánico Transversal and the Sierra Madre del Sur, at altitudes between 1400 and 2500 m (Fig. 1). *Quercus sideroxyla* Humb & Bonpl is a tree about 10 m in height that belongs to series *Sideroxylae* (Trelease, 1924), with populations occurring in the Sierra Madre Occidental and the Eje Neovolcánico Transversal, with an altitudinal distribution from 1800 to 2700 m (Fig. 1). It has obovate or oblanceolate leaves 3–6 cm in length and 2·5–3·5 cm in width, with 1–5 distal teeth with aristae, and flattened petioles 3–9 mm long. Fruits are biennial, almost sessile, and solitary or produced in pairs. Finally, *Quercus hypoleucoides* Camus is a tree 10–20 m in height included in series *Hypoleucae* (Trelease, 1924), and has a narrower geographical distribution than the other two species, with populations in the north of the Sierra Madre Occidental and in Arizona, New Mexico and Texas, at altitudes ranging from 2000 to 2500 m (Fig. 1). The species has narrow leaves that are lanceolate, elliptical or obovate, 5–11 cm long and 1·5–3 cm wide, with entire margins. Acorns are produced annually or biennially, are sessile or almost sessile and solitary or produced in pairs.

Sampling procedure

Leaf samples were collected in 13 populations located in the zone of contact (Fig. 2). As reference, two morphologically representative populations situated within the zone of contact were also sampled, and two representative isolated populations situated out of the zone of contact were sampled per species (Fig. 2), with the exception of *Q. hypoleucoides*, because its geographical distribution within Mexico is restricted to the study zone. Representative populations were chosen based on individuals with typical diagnostic characters of each species. In each locality, 8–12 individuals sampled haphazardly were chosen but at least 50 m apart from each other. Three branches from each individual were collected: one for taxonomic identification, one for morphometric analysis and the third to obtain fresh intact leaves for molecular analysis, which were placed on ice and then at −80 °C until genetic analyses were conducted.

Leaf shape morphometric analysis

Photographs were taken of the abaxial side of five intact mature light-exposed leaves from each of eight to ten individuals per population. Coordinates '*x*, *y*' of 21 unambiguous and repeatable marks (i.e. landmarks and semi-landmarks) were registered along the border of each leaf image using the

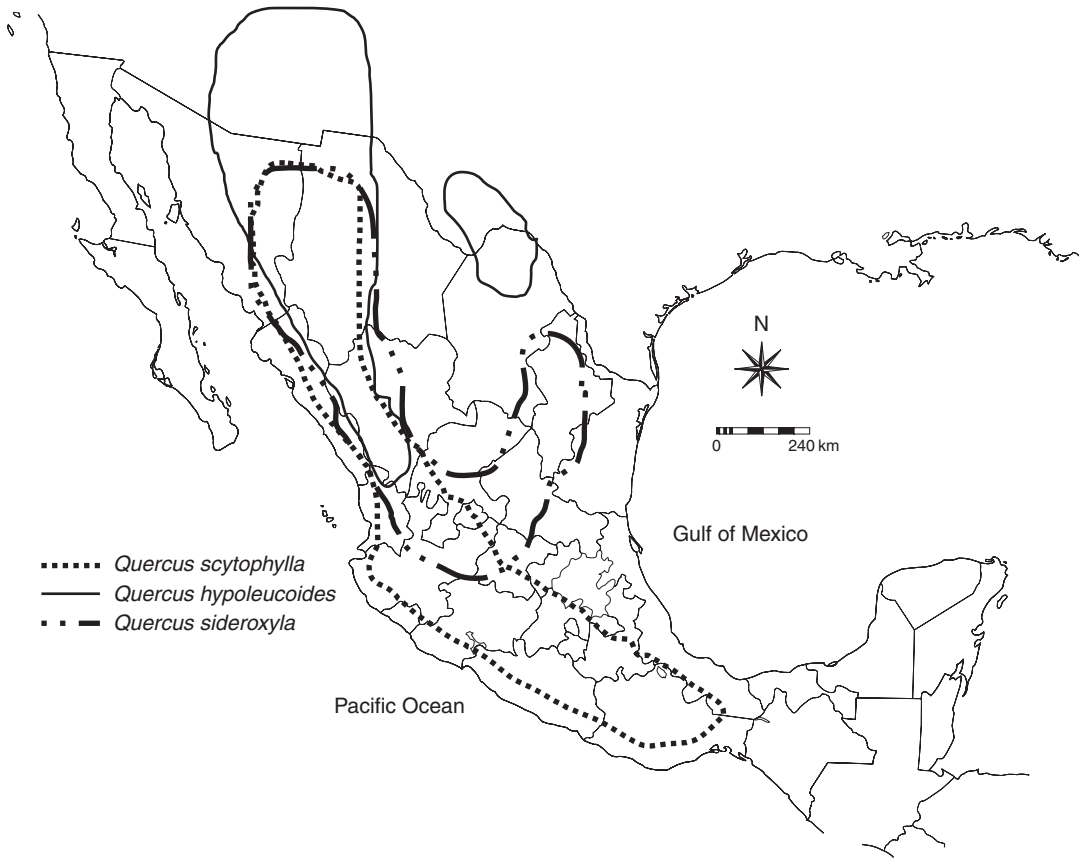


FIG. 1. Geographical distribution of the three red-oak species studied, *Quercus scytophylla*, *Q. hypoleucoides* and *Q. sideroxyla*.

program TpsDig (Rohlf, 2004). Three marks (apex, lamina base and petiole extreme) corresponded to homologous loci ('landmarks' *sensu* Bookstein, 1991), and the other 18 were semi-landmarks, i.e. morphological points that incorporate information about leaf contour in morphometric analysis, in areas that lack landmarks for all the individuals (Zelditch *et al.*, 2004). Landmarks '1' (lamina base) and '12' (apex) were used to construct a 'fan' (radial guidelines with equal angular spacing on images) with 80 radial guidelines covering the whole leaf contour, which was used to digitize the 18 semi-landmarks. The program MakeFan6 within the IMP software package (Integrated Morphometrics Package; http://www.canisius.edu/~sheets/morphsoft.html) was used for this procedure. A Procrustes superimposition analysis was developed with the CoordGen program in IMP. This analysis allows the calculation of leaf shape variation without the effect of the size. The resulting shape variables (Procrustes distances) for all individuals were subjected to a canonical discriminant analysis to determine the variation in leaf shape among individuals of each species and intermediate populations using SPSS 11·0 (Ferran, 1997). Individuals from reference populations of the three species were first analysed to obtain the canonical discriminant functions, and then the discriminant scores calculated with these functions were obtained for individuals from all populations. In this way, five morphological groups were identified, three of which corresponded to pure species individuals and two corresponded to intermediate individuals (see Results).

Microsatellite amplification

Genomic DNA was extracted from 100 mg of leaf material using the method proposed by Lefort and Douglas (1999). Seven nuclear microsatellite loci (OC11, OA01, 1C08, 2M04, OE09, 1H14, 1F07) previously designed for *Quercus rubra* (Aldrich *et al.*, 2002) were selected based on the quality of preliminary amplification trials. Polymerase chain reactions were carried out in a volume of 25 μL containing 20 ng of template DNA, 2 mM $MgCl_2$, 10 mM Tris-HCl (pH 9), 0·1 mM of each dNTP, 0·5 mg mL^{-1} bovine serum albumin, 2 μM of each primer and 0·3 units of TaqDNA polymerase (Gibco, Invitrogen, San Diego, CA, USA). The thermal cycling conditions consisted of 40 cycles, each at 94 °C for 1 min, annealing at 50 °C for 1 min and extension at 72 °C for 2 min. A final extension at 72 °C for 10 min was included. Amplified fragments were separated in polyacrylamide gels (6 %) in a semi-automatic sequencer LI-COR 4300 (LI-COR Biosciences, Lincoln, NB, USA). Fragment sizes were calculated using the program ImageJ by comparisons with internal and external standards. All individuals sampled from the 23 populations were genotyped.

Genetic analysis

The mean number of alleles per locus (N_a), mean effective number of alleles (N_e), mean observed heterozygosity (H_O), mean expected heterozygosity (H_E), mean fixation index (F)

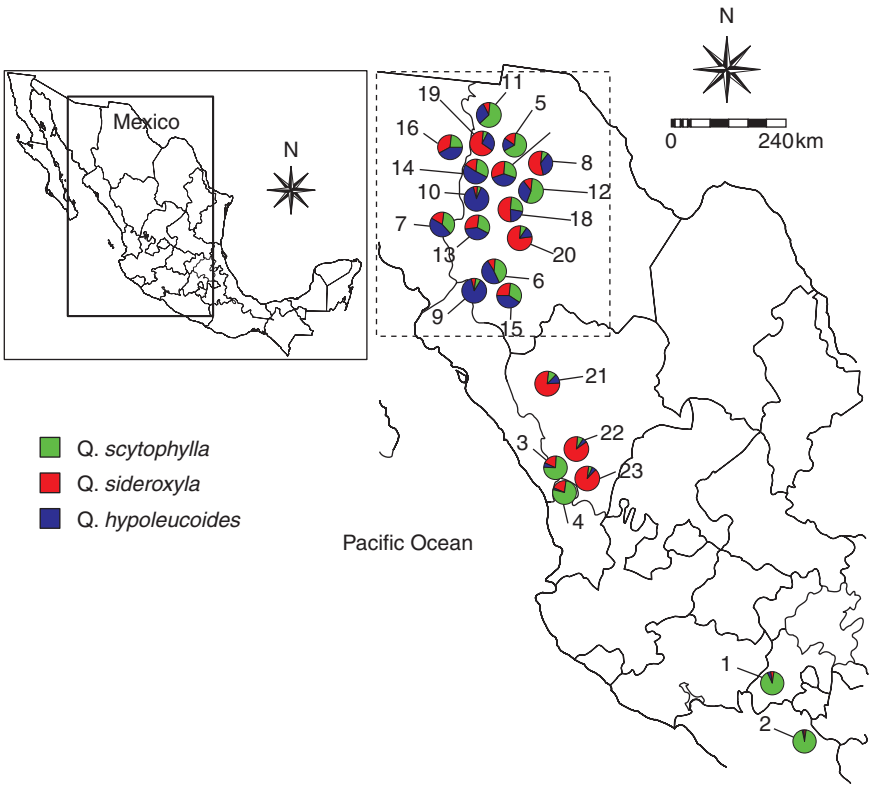


FIG. 2. Map showing sampling localities representing parental and intermediate populations. Each pie chart represents the proportions in each population of the three genetic groups as assigned by the program STRUCTURE. Green, red and blue represent the genetic groups corresponding to *Quercus scytophylla*, *Q. sideroxyla* and *Q. hypoleucoides*, respectively. Numbers next to each symbol correspond to the population numbers given in Table 1. The area of species contact is delineated with a dashed rectangle.

and their respective standard errors were calculated for each population using FSTAT (Goudet, 1995). Average values were also obtained for each of the five morphological groups defined on the basis of canonical discriminant function analysis. Pairwise genetic differentiation among the five morphological groups was calculated using the method of Weir and Cockerham (1984) with the GENETIX 4 program (Belkhir *et al.*, 2004) with 10 000 permutations for statistical significance. Within each of the morphological groups genetic differentiation among populations was estimated with analyses of molecular variance (AMOVA). The significances of the different variance components were estimated from distributions generated from 10 000 random permutations. These analyses were carried out using ARLEQUIN 3·0 (Excoffier *et al.*, 2005).

Bayesian admixture analysis

The genetic ancestry of each individual was inferred with the program STRUCTURE 2·3·1 (Pritchard *et al.*, 2000; Falush *et al.*, 2003; Hubisz *et al.*, 2009). This program is based on a Bayesian model clustering procedure to determine the proportions of ancestry of individuals derived from multiple populations. In this analysis, different probable values of K (number of genetic groups) were assayed, increasing the probable number from $K = 1$ to 13 (ten times for each value). The program was run using the admixture model with correlated allelic frequencies without prior population information. For all runs, a length of the burn-in period of 50^3 iterations was used, followed by 10^6 Markov chain Monte Carlo iterations. Additionally, the most probable value of K was determined using the maximum value of ΔK, following the rules of Evanno *et al.* (2005), and by the estimated ln probability of data, ln$P(D)$.

Genotype simulation and assignment

To evaluate the ability of the assignment procedure to recognize accurately the different genotypic classes potentially present in our samples (i.e. pure parental, hybrids between the different pairs of species and triple hybrids), the procedure developed by Vähä and Primmer (2006) was followed. Individuals that in the assignment analysis described above had an assignment coefficient (Q) higher than 0·90 (indicating a high probability of belonging to a single genetic cluster) were used to estimate allelic frequencies of the three species. Thirty-nine such individuals were identified for *Q. scytophylla*, 21 for *Q. hypoleucoides* and 44 for *Q. sideroxyla*. Pure species and hybrid genotypes were then simulated with HYBRIDLAB 1·0 (Nielsen *et al.*, 2006) using the calculated allelic frequencies. Four hundred genotypes were simulated for each species, 37 for the different possible F_1 hybrids, 113 for backcrosses and 14 for triple hybrids (triple hybrids were considered the product of crossing an F_1 hybrid from a pair of species with a pure individual of the third species). The number of simulated hybrid genotypes

is similar to the number observed in the real data. The simulated data were then analysed with the program STRUCTURE, with the same number of genetic groups previously inferred ($K = 3$), and the same settings. Thereafter, the performance of STRUCTURE in hybrid and pure-bred individual identification was evaluated via the parameters of efficiency (number of individuals correctly assigned), accuracy (proportion of an identified group that truly belongs to that category) and performance (efficiency multiplied by accuracy), following Vähä and Primmer (2006). Finally, the optimal threshold values of Q to assign individuals to the different genotypic categories were determined.

Geographical patterns of variation

To determine if variation patterns in genetic composition and morphology of populations could be explained by geographical variables, we conducted stepwise regression analyses of the population average values of canonical discriminant scores and proportions of genetic ancestry in each of the genetic groups inferred against the latitude, longitude and altitude of the populations.

RESULTS

Leaf shape morphometric analysis

The first two discriminant functions (F1 and F2), derived from the Procrustes analysis of leaf shape, explained all the observed variation (55·1 and 44·9 %, respectively), indicating that discrimination among reference populations of the three species was absolute and highly significant (for F1, Wilks' $\lambda = 0{\cdot}027$, d.f. = 82, $P < 0{\cdot}001$; for F2, Wilks' $\lambda = 0{\cdot}180$, d.f. = 40, $P < 0{\cdot}001$), supporting the taxonomic delimitation of the three taxa. This is clearly observed in Fig. 3, with the three species forming completely separated clusters. F1 contributed most to the discrimination between *Q. hypoleucoides* and the other two species, while F2 discriminated between *Q. scytophylla* and *Q. sideroxyla*. Individuals from the 13 populations considered to show indications of morphological intermediacy had score values identical to one of the reference clusters or were intermediate either along F1, F2 or both. These populations formed two recognizable groups (Fig. 3). The first group (hereafter called Mixed 1) included individuals from populations Río Chico, Querari and Yécora, most of which were similar to reference individuals of *Q. hypoleucoides*, and individuals from population Madera that were clearly intermediate between *Q. hypoleucoides* and *Q. sideroxyla*. The second group (hereafter Mixed 2) contained individuals similar to *Q. scytophylla* or intermediate between this species and the other two, and belonged to populations Huracán A, Huracán B, Guadalupe, Km 346, Guachochic, Amarilla B, Poleo, Amarilla A and Babícora.

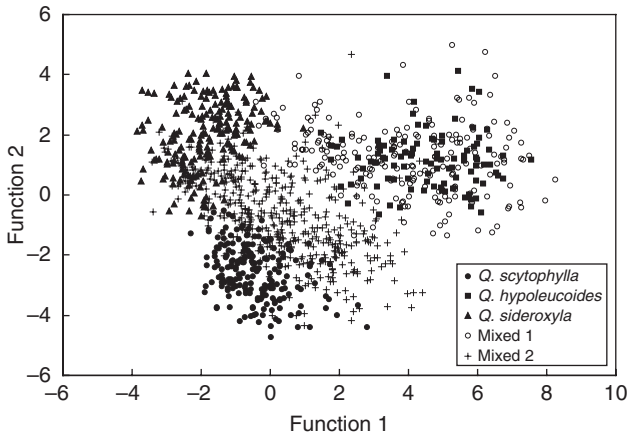


FIG. 3. Scatterplot of individual scores from discriminant function analysis of populations in the oak complex based on foliar geometric morphometric data. Closed circles, squares and triangles symbolize representative populations of *Quercus scytophylla*, *Q. hypoleucoides* and *Q. sideroxyla*, respectively, and open circles and crosses are the Mixed 1 and Mixed 2 population groups, respectively. See text for details.

Genetic structure

Seven microsatellite loci revealed high levels of genetic diversity and low but significant genetic differentiation between and within the three species (Tables 1 and 2). The mean number of alleles per locus (N_a) within populations varied from 3·71 to 10·42 and the mean of the effective number of alleles (N_e) was between 2·55 and 7·25 (Table 1). High values of H_O (ranging from 0·633 to 0·905) and H_E (from 0·555 to 0·841) were observed (Table 1). H_O, H_E, N_a and N_e did not differ among the five morphological groups (three reference and two intermediate groups) according to Wilcoxon tests (data not shown). Among 161 inbreeding coefficients (F) calculated (23 populations and seven loci), only four were significant after applying a Bonferroni correction: for locus OA01 in populations Espinazo ($F = 0{\cdot}744$, $P < 0{\cdot}00001$) and Tecuán ($F = 0{\cdot}226$, $P < 0{\cdot}0001$), and for locus 2M04 in populations Temosachic ($F = -0{\cdot}337$, $P < 0{\cdot}0001$) and Tecuán ($F = -0{\cdot}313$, $P < 0{\cdot}00001$). The program MICROCHECKER (Van Oosterhout *et al.*, 2004) was used to test whether these deviations could be due to the presence of null alleles, stuttering or large allele dropout at those loci. Only population Espinazo showed signs of the presence of a null allele at high frequency (in four of ten individuals) at locus OA01. For further analysis, this locus was coded as missing data in this population.

According to Weir and Cockerham's θ, genetic differentiation was significant in all pairwise comparisons among the five morphological groups, except between the Mixed 1 and Mixed 2 groups (Table 2). Differentiation among populations was also significant within the three species, but not within the Mixed 1 and Mixed 2 groups (Table 3).

Estimates of admixture

The highest posterior probability was obtained for three genetic clusters according to the values of ln$P(D)$ (Fig. 4A). This result was also confirmed by the ΔK values (Fig. 4B). Using the admixture model in the program STRUCTURE, the proportion of ancestry (Q) of each individual and population in each of the three genetic groups was also estimated. The allopatric reference populations had a high proportion of ancestry from a single genetic group (Figs 2 and 5). For *Q. scytophylla*, populations Tejupilco and Pozo Largo had genetic ancestries of 0·913 and 0·947 in one genetic cluster

TABLE 1. *Name, geographical coordinates, altitude, mean observed heterozygosity (H_O), mean expected heterozygosity (H_E), mean number of alleles (N_a) and mean number of effective alleles per locus (N_e) for 23 populations of* Quercus scytophylla, Q. hypoleucoides *and* Q. sideroxyla

Number and name of locality	State	Coordinates	Sample size	Genetic diversity			
				H_O	H_E	N_a	N_e
Q. scytophylla							
1. Tejupilco	México	19°02′/100°03′	9	0·889 (0·12)	0·772 (0·13)	6·85 (2·96)	5·46 (0·84)
2. Pozo Largo	Guerrero	17°36′/99°05′	10	0·742 (0·19)	0·810 (0·05)	7·14 (2·05)	5·61 (0·55)
3. Espinazo	Sinaloa	23°34′/105°05′	10	0·814 (0·29)	0·805 (0·08)	8·14 (2·26)	5·83 (0·77)
4. Loberas	Durango	23°29′/105°51′	10	0·857 (0·12)	0·832 (0·03)	8·42 (0·61)	6·25 (0·49)
Mixed 1							
5. Río Chico	Chihuahua	29°36/108°10′	12	0·657 (0·35)	0·729 (0·11)	6·71 (3·03)	4·52 (0·94)
6. Querari	Chihuahua	27°11′/107°32′	11	0·831 (0·24)	0·819 (0·08)	9·42 (2·50)	6·41 (0·85)
7. Yécora	Sonora	28°22′/109°02′	12	0·839 (0·13)	0·831 (0·08)	9·85 (3·48)	7·07 (1·03)
8. Madera	Chihuahua	29°32′/108°09′	11	0·889 (0·09)	0·801 (0·05)	7·86 (0·59)	5·35 (0·53)
Q. hypoleucoides							
9. Batopilas	Chihuahua	27°08′/107°34′	10	0·885 (0·16)	0·782 (0·08)	7·00 (2·58)	5·09 (0·58)
10. Temosachic	Chihuahua	28°59′/108°13′	11	0·766 (0·22)	0·728 (0·16)	7·00 (2·76)	4·74 (0·85)
Mixed 2							
11. Huracán A	Chihuahua	29°40′/108°15′	12	0·853 (0·09)	0·824 (0·05)	8·42 (1·51)	6·13 (0·62)
12. Guadalupe	Chihuahua	29°11′/107°57′	10	0·846 (0·17)	0·765 (0·14)	7·14 (2·03)	5·19 (0·68)
13. Km 346	Chihuahua	28°26′/108°32′	11	0·633 (0·47)	0·555 (0·38)	3·71 (2·21)	2·55 (0·67)
14. Amarilla B	Chihuahua	29°12′/108°14′	12	0·657 (0·26)	0·681 (0·14)	6·42 (3·86)	4·09 (0·97)
15. Guachochic	Chihuahua	26°56′/107°08′	8	0·905 (0·12)	0·799 (0·03)	7·57 (1·81)	5·11 (0·32)
16. Poleo	Sonora	29°37′/108°98′	12	0·809 (0·16)	0·829 (0·09)	10·42 (3·40)	7·25 (1·21)
17. Amarilla A	Chihuahua	29°11′/108°14′	10	0·778 (0·09)	0·796 (0·08)	8·14 (2·11)	5·53 (0·68)
18. Babícora	Chihuahua	29°13′/107°49′	10	0·724 (0·22)	0·778 (0·13)	7·57 (2·99)	5·53 (0·85)
19. Huracán B	Chihuahua	29°40′/108°15′	12	0·734 (0·22)	0·764 (0·08)	7·71 (2·21)	4·69 (0·56)
Q. sideroxyla							
20. Maycoba	Chihuahua	28°17′/108°06′	10	0·857 (0·11)	0·783 (0·13)	7·85 (3·02)	5·75 (0·91)
21. Cuevecillas	Durango	25°02′/106°16′	11	0·844 (0·15)	0·841 (0·07)	9·71 (2·75)	7·17 (0·88)
22. Buenos Aires	Durango	23°42′/105°43′	11	0·831 (0·11)	0·841 (0·04)	8·71 (2·13)	6·71 (0·62)
23. Tecuan	Durango	23°55′/105°01′	12	0·821 (0·04)	0·833 (0·06)	10·28 (1·89)	6·64 (0·79)

Standard errors are given in parentheses. The five morphological groups (three parental species and two mixed groups) were defined on the basis of geometric morphometrics.

TABLE 2. *Pairwise values of genetic differentiation (F_{ST}) among the five morphological groups estimated with the method of Weir and Cockerham (1984)*

	Q. scytophylla	*Q. hypoleucoides*	*Q. sideroxyla*	Mixed 1	Mixed 2
Q. scytophylla	–	**0·022**	**0·071**	**0·031**	**0·039**
Q. hypoleucoides		–	**0·036**	**0·022**	**0·035**
Q. sideroxyla			–	**0·029**	**0·054**
Mixed 1				–	0·014
Mixed 2					–

Numbers in bold indicate statistically significant values ($P < 0·05$) based on 10 000 random permutations.

and, for *Q. sideroxyla*, individuals from populations Buenos Aires and Tecuan also were assigned mostly to a single genetic group (Figs 2 and 5). In contrast, morphologically representative populations of the three species from within the contact zone showed some indications of introgression. Populations of *Q. hypoleucoides* had slight contributions from the other two genetic groups, while populations Espinazo and Loberas, belonging to *Q. scytophylla*, were influenced by the *Q. sideroxyla* genetic group. Finally, some introgression of the *Q. hypoleucoides* genetic cluster was detected in populations Maycoba and Cuevecillas, belonging to *Q. sideroxyla* (Figs 2 and 5). In the morphologically mixed populations, there was evidence of admixture among the three genetic clusters in different proportions. The Mixed 1 group was characterized by a transition from a higher proportion of the *Q. scytophylla* genetic cluster at lower altitudes (i.e. populations Río Chico and Huracán A) to more even contributions from the three species at higher altitudes (Km 346 and Yecora). The Mixed 2 group showed a change from even genetic proportions from the three genetic clusters towards a predominance of ancestry from *Q. sideroxyla* at high altitudes (Figs 2 and 5).

TABLE 3. *Genetic structure within the five morphological groups estimated with Φ_{ST} obtained from AMOVA*

Source of variation		d.f.	SS	Variance components	Percentage of variation	Fixation index
F_{ST}						
	Q. scytophylla					
	Among populations	3	18·1	0·157	5·01	$\Phi_{ST} = 0{\cdot}05$***
	Within populations	74	220·06	2·973	94·99	
	Q. hypoleucoides					
	Among populations	1	8·89	0·292	9·54	$\Phi_{ST} = 0{\cdot}09$***
	Within populations	40	110·89	2·772	90·46	
	Q. sideroxyla					
	Among populations	3	17·96	0·135	4·26	$\Phi_{ST} = 0{\cdot}04$***
	Within populations	84	254·39	3·028	95·74	
Mixed 1						
	Among populations	3	7·18	0·008	0·36	$\Phi_{ST} = 0{\cdot}004$ n.s.
	Within populations	88	194·5	2·21	99·64	
Mixed 2						
	Among populations	7	−3·43	−0·118	−6·44	$\Phi_{ST} = -0{\cdot}06$ n.s.
	Within populations	158	307·83	1·95	106·44	

Asterisks indicate statistically significant values ($P < 0{\cdot}05$) and n.s. indicates non-significant values. Tests were based on 10 000 random permutations.

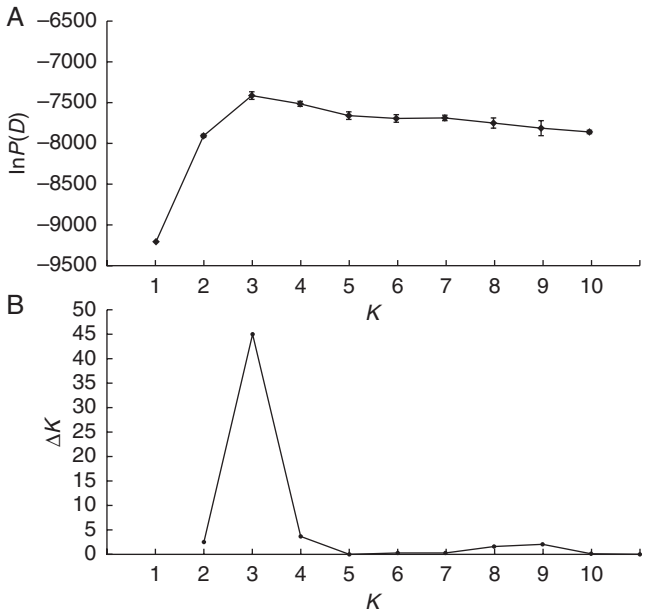


FIG. 4. (A) Mean and standard deviation of ln$P(D)$ for ten independent runs of STRUCTURE plotted against the number of genetic groups (K) used in the analysis. (B) Values of ΔK plotted against K. In both cases the peak indicates the most probable number of genetic groups given the data.

Performance of STRUCTURE with simulated genotypes

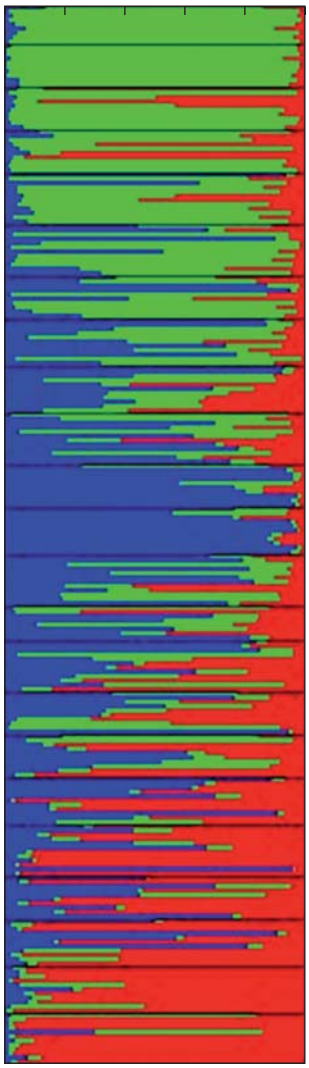
Analysis of the simulated genotypes indicated that a threshold value in the admixture coefficient of $Q \geq 0{\cdot}90$ allowed adequate separation between pure species and hybrid individuals (Table 4). The genetic assignment of pure individuals from the three species showed high levels of performance, with values of 93 % for *Q. scytophylla*, 96 % for *Q. sideroxyla* and 95 % for *Q. hypoleucoides* (Table 4). For hybrid classes, performance values in most cases were moderately lower in comparison with pure-species assignment. Even for triple hybrids there was a reasonable identification rate.

Given these results, individuals from all populations in the real samples were assigned to a genotypic class as follows: individuals with $Q \geq 0{\cdot}90$ were considered to be pure bred, trees with $Q < 0{\cdot}90$ from two genetic groups were considered hybrids between two species, and trees with ancestry from the three genetic groups in more or less equal proportions ($Q = 0{\cdot}3–0{\cdot}4$) were considered triple hybrids (Fig. 6). Populations from the contact zone were largely dominated by hybrid individuals (Fig. 6). As previously noted for the admixture coefficient of each population, the proportion of the different genotypes also changed with altitude. At lower altitudes, the predominant genotypes corresponded to crosses between *Q. scytophylla* and *Q. sideroxyla*, which were replaced by crosses between *Q. scytophylla* and *Q. hypoleucoides*, and finally by crosses between *Q. sideroxyla* and *Q. hypoleucoides* with increasing altitude. Triple hybrids were situated in populations at mid altitudes, particularly in populations Km 346 and Amarilla A, where also almost all the other possible genotypes were present.

Associations among leaf shape morphology, genetic ancestry and altitude

The results of the regression analysis indicated that the proportions of genetic ancestry of *Q. scytophylla* and *Q. sideroxyla* in the populations were strongly correlated with altitude ($R^2 = 0{\cdot}78$, $P < 0{\cdot}0001$ and $R^2 = 0{\cdot}72$, $P = 0{\cdot}0002$, respectively) but not with latitude or longitude. In turn, the proportion of the *Q. hypoleucoides* genetic group was weakly correlated with latitude only ($R^2 = 0{\cdot}4$, $P = 0{\cdot}03$). Finally, the canonical scores of morphological variation showed a marginally significant tendency to be correlated with latitude ($R^2 = 0{\cdot}3$, $P = 0{\cdot}06$).

Nevertheless, as reference populations of the three species were situated at the extremes of the altitudinal range, it might be that the significant correlations obtained are mostly due to the different altitudinal requirements of the pure species. To verify this explanation, correlations were calculated again without the reference populations and practically the same results were obtained (data not shown).



Locality	Morphological group	Genetic assignment			
		'scytophylla'	*'hypoleucoides'*	*'sideroxyla'*	Altitude (m)
Tejupilco[1]	*Q. scytophylla*	0·913	0·046	0·041	1725
Pozo Largo[2]	*Q. scytophylla*	0·947	0·023	0·030	1860
Espinazo[3]	*Q. scytophylla*	0·771	0·034	0·196	1955
Loberas[4]	*Q. scytophylla*	0·742	0·055	0·203	*2000*
Río Chico[5]	Mixed 1	0·652	0·185	0·163	1785
Huracán A[11]	Mixed 2	0·621	0·289	0·090	1800
Guadalupe[12]	Mixed 2	0·541	0·339	0·120	1852
Querari[6]	Mixed 1	0·416	0·478	0·106	1900
Km 346[13]	Mixed 2	0·307	0·400	0·283	1910
Yécora[7]	Mixed 1	0·365	0·449	0·186	1970
Batopilas[9]	*Q. hypoleucoides*	0·074	0·871	0·056	1900
Temosachic[10]	*Q. hypoleucoides*	0·049	0·899	0·052	1950
Amarilla B[14]	Mixed 2	0·310	0·518	0·172	2101
Guachochic[15]	Mixed 2	0·325	0·407	0·268	2177
Poleo[16]	Mixed 2	0·233	0·422	0·345	2240
Amarilla A[17]	Mixed 2	0·291	0·386	0·323	2350
Babícora[18]	Mixed 2	0·266	0·223	0·510	2240
Madera[8]	Mixed 1	0·083	0·368	0·549	2300
Huracán B[19]	Mixed 2	0·057	0·274	0·669	2450
Maycoba[20]	*Q. sideroxyla*	0·082	0·130	0·788	2375
Cuevecillas[21]	*Q. sideroxyla*	0·054	0·155	0·781	2500
Buenos Aires[22]	*Q. sideroxyla*	0·073	0·055	0·872	2622
Tecuán[23]	*Q. sideroxyla*	0·055	0·045	0·900	2798

FIG. 5. Genetic assignment of individuals and populations according to the Bayesian method implemented in the program STRUCTURE. One sample, considered to explain the data best, of ten iterated runs is shown. Each thin horizontal line represents an individual and the proportion of each colour is the proportion of ancestry derived from each of the three main genetic groups ($K = 3$) inferred. Populations are separated by black lines. Values of the proportions of ancestry for each population are given in the table.

DISCUSSION

The combined results of geometric morphometrics of leaf shape and genetic analysis with microsatellites indicate that a high rate of hybridization and introgression have occurred among the three oak species studied, which have resulted in a complex hybrid zone containing a wide array of genotypes. Furthermore, this hybrid zone is geographically structured as a mosaic, with a patchy spatial distribution of pure and mixed populations with different genetic compositions. Altitude was, apparently, the most important geographical variable explaining the distribution of genetic and morphological variation.

The utility of geometric morphometrics to quantify leaf shapes and to assess hybridization in plants is well established (e.g. Jensen *et al.*, 2002). In species of *Quercus*, foliar shape is particularly informative because other organs (i.e. flowers) show little variation (Kaul, 1985). The morphometric assessment performed in this study, based on Procrustes analysis of landmarks followed by canonical discriminant function analysis, substantiated the complete morphological differentiation of individuals from representative populations of the three oak taxa *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla*. The morphometric analysis also demonstrated that some individuals were intermediate in various combinations with respect to representative populations, as was initially presumed from field observations. Intermediate individuals were present in a series of populations scattered along a wide area of contact among the three oak species. These populations showed clear differences in their phenotypic composition and could be divided into two groups. The first

TABLE 4. *Number of simulated individuals (rows), which were assigned to one of the three species or into a hybrid category (columns)*

Simulated/ assigned	Qscytop	Qhypole	Qsidero	Hyb QscytoQsidero	Hyb QscytoQhypole	Hyb QsideroQhypole	Triple Hybrid	Total
Qscytop	**391**	–	–	9	–	–	–	400
Qhypole	–	**398**	–	–	2	–	–	400
Qsidero	–	–	**395**	–	–	5	–	400
F_1 Qscyto-Qsidero	2	–	–	**3**	–	2	–	7
BX Qscyto	11	–	–	**24**	–	2	–	37
F_1 Qscyto-Qhypole	–	–	–	2	**14**	–	–	16
BX Qhypole	–	12	–	–	**32**	–		44
F_1 Qsidero-Qhypole	–	–	–	–	1	**13**	–	14
BX Qsidero	–	–	9	–	1	**21**	1	32
Triple Hybrid	–	–	–	2	–	3	**9**	14
Total	404	410	404	40	50	46	10	1364
Efficiency	97·75	99·50	98·75	61·3636	76·666	73·913	64·28	
Accuracy	96·782	97·073	97·777	67·50	92·00	73·914	90·00	
Performance	94·60	96·58	96·55	41·42	70·53	54·63	57·85	

Parameters of efficiency, accuracy and overall performance of the assignment method are given in per cent. Note individuals correctly assigned are in bold type. Hyb, hybrids; F_1, first-generation hybrids; BX, backcrosses; Qscyto, *Quercus scytophylla*; Qhypole, *Q. hypoleucoides*; Qsidero, *Q. sideroxyla*.

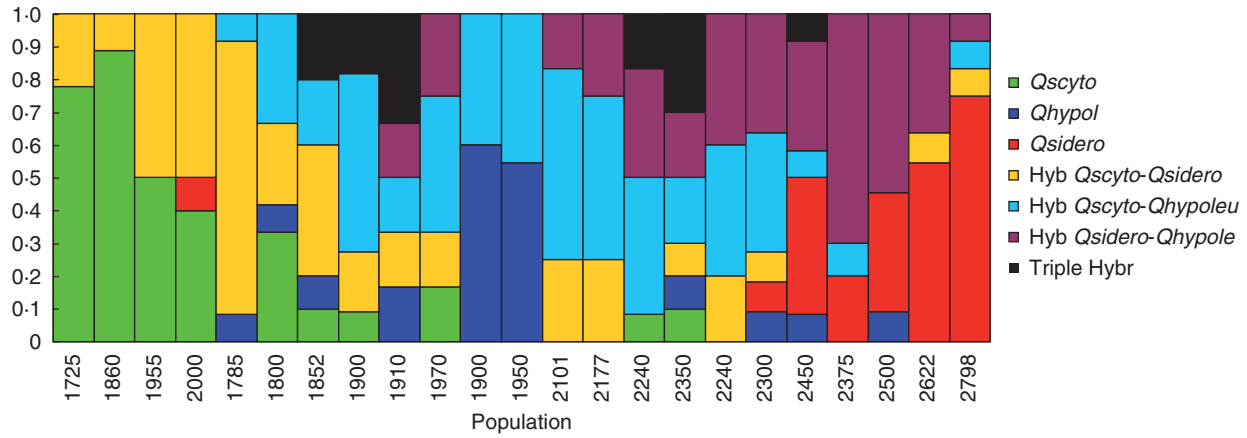


FIG. 6. Frequency of the different genotypic classes observed in each population. Individuals were assigned to each category (*Q. scytophylla*, *Q. hypoleucoides*, *Q. sideroxyla*, hybrids between *Q. scytophylla* and *Q. sideroxyla*, hybrids among *Q. scytophylla* × *Q. hypoleucoides*, hybrids between *Q. sideroxyla* and *Q. hypoleucoides* and triple hybrids, as indicated in the key), depending on their individual coefficient of admixture derived from STRUCTURE. Performance of the assignment procedure was previously assessed by analysing simulated genotypes (see text for details).

group included a series of sites (i.e. Río Chico, Querari, Yécora and Madera) in which the predominant morphology was similar to *Q. hypoleucoides*, *Q. sideroxyla* or intermediate between them. The second group (i.e. Huracán, Guadalupe, Km 346, Guachochic, Amarilla B, Poleo, Amarilla A and Babícora) consisted of populations showing evidence of morphological intergradation between *Q. scytophylla* and the other two species.

The degree of genetic differentiation among the three species was significant but very low (Table 2). However, similar comparisons of oak species belonging to several complexes have repeatedly found low interspecific genetic differentiation based on nuclear markers, ranging from 0·02 to 0·17 (Dodd and Kashani, 2003; González-Rodríguez *et al.*, 2005; Muir and Schlötterer, 2005; Craft and Ashely, 2006; Curtu *et al.*, 2007; Gugerli *et al.*, 2007). Although it has been argued that low genetic differentiation among oak species does not necessarily imply introgression, but can be accounted for by the sharing of ancestral polymorphisms (e.g. Muir and Schlötterer, 2005), in most instances interspecific gene flow is a more parsimonious explanation for this observation (Lexer *et al.*, 2006).

In the specific case studied here intermediate populations occurred within a defined area with the characteristics of a mosaic hybrid zone (see below). Furthermore, the Bayesian analysis suggested a clustering into three genetic groups, which agrees with the number of taxa involved in this complex, with isolated representative populations assigned to a single genetic cluster. However, the same analysis also suggested the occurrence of various degrees of admixture among the three genetic groups in populations from within the area of contact. The above evidence strongly suggests that introgression has occurred among the three oak taxa and contributed to shape the patterns of variation observed. As shown, both the morphological and the genetic data indicate that some of the populations in the zone of contact are the

result of introgression between a pair of species, but others clearly seem to be an admixture of all three taxa.

At the level of individuals, a high proportion of hybrids resulting from crosses between different species pairs were identified within the contact zone, and 14 trees were assigned as probable triple hybrids. Simulations of genotypes provided reasonable support to the assignment method, providing some confidence on the identification of even these complex hybrids (Vähä and Primmer, 2006). Studies on closely related European white oaks have suggested that five or six microsatellite loci are sufficient to distinguish between pure species and introgressed individuals (Curtu *et al.*, 2007; Gugerli *et al.*, 2007). A recent study involving four oak species suggested that patterns of interspecific crosses may be more complex than those modelled in simulations, possibly leading to the existence of third- or later-generation hybrids as well as to hybridization involving more than two species (Lepais *et al.*, 2009). Triple hybridization first requires the production of fertile hybrid genotypes between at least two species, and then the crossing between hybrids and a third species, or between hybrids from different species combinations (Kaplan and Fehrer, 2007).

Mosaic hybrid zones are described as areas where pure species populations and mixed populations are patchily distributed across a zone of overlap (Harrison and Rand, 1989). The morphological variation in the contact area in the Sierra Tarahumara clearly follows this pattern. Populations with typical parental species morphology and those belonging to the Mixed 1 and Mixed 2 morphological groups are geographically scattered. Genotypic variation was also distributed as a mosaic and showed little correlation with latitude and longitude. Strong associations among specific genotypes and certain habitats are also characteristic of mosaic hybrid zones, and exogenous selection is considered important in such structuring (Harrison and Rand, 1989). From the analyses presented, it was evident that the genetic and morphological composition of the oak populations studied is strongly associated with altitude. However, as can be judged from the lack of similarity among geographically proximate populations, dispersal does not seem to be the factor determining these correlations. Alternatively, the influence of environmental variables on the structure of this hybrid zone can be suggested as an explanation for these results. In their study on four Californian red oak species, Dodd and Afzal-Rafii (2004) obtained results implying that environmental gradients rather than pollen dispersal determine the extent of introgression. In other plant species, hybrid zones associated with altitudinal gradients have been reported and the patterns of variation are often coupled to environmental variables that change with elevation (e.g. James and Abbott, 2005; Kimball, 2008). For oaks, there is some direct evidence indicating that the frequency of hybridization, the direction of introgression and the performance of hybrids depend on habitat conditions (Williams and Ehleringer, 2000; Williams *et al.*, 2001; Himrane *et al.*, 2004), but also on species' relative abundance (Lepais *et al.*, 2009). These factors could be contributing to the structure of the hybrid zone in the Sierra Tarahumara. Nevertheless, ecological studies are required to gain more detailed insight into the dynamics of this hybrid zone. In conclusion, the evidence presented supports the existence of a complex hybrid zone that has formed from pairwise and triple hybridization among *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla* in the mountains of the Sierra Tarahumara in north-western Mexico, which is structured as a mosaic probably in response to environmental variables associated with altitude.

ACKNOWLEDGEMENTS

We thank E. de Luna, A. Guerrero, A. Albarrán-Lara, E. Pascual, L. Herrera, O. Chassin-Noria, P. Cuevas-Reyes, N. Pérez-Nasser, V. Rocha, J. J. Fuentes-Junco and F. Alvarado-Ramos for technical assistance and suggested improvements to the manuscript. We thank Alex Buerkle, Felix Gugerli and an anonymous reviewer for their suggestions that greatly improved the manuscript. J.M.P.-R. received a PhD scholarship from CONACyT. This project was supported by CONACyT [grant 38550-V to K.O.], CONACyT-SEMARNAT [23728 to K.O.] and CONACyT-ECOS NORD [M03-A01] to A. K. and K.O.

LITERATURE CITED

Aldrich PR, Michler CH, Sun W, Romero-Severson J. 2002. Microsatellites markers for northern red oak (Fagaceae: *Quercus rubra*). *Molecular Ecology Notes* **2**: 472–474.

Arnold ML. 1993. *Iris nelsonii* (Iridaceae): origin and genetic composition of a homoploid hybrid species. *American Journal of Botany* **80**: 577–583.

Arnold ML. 1997. *Natural hybridization and evolution*. Oxford: Oxford University Press.

Barton NH, Gale KS. 1993. Genetic analysis of hybrid zones. In: Harrison RG. ed. *Hybrid zones and the evolutionary process*. Oxford: Oxford University Press, 13–45.

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**: 113–148.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 1996–2004. *GENETIX 4·05, logiciel sous Windows TM pour la génétique des populations*. Montpellier: Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II.

Bookstein FL. 1991. *Morphometric tools for landmarks data: geometry and Biology*. Cambridge: Cambridge University Press.

Bronson CL, Grubb TC, Jr, Braun MJ. 2003. A test of the endogenous and exogenous selection hypotheses for the maintenance of a narrow avian hybrid zone. *Evolution* **57**: 630–637.

Curtu AL, Gailing O, Finkeldey R. 2007. Evidence for hybridization and introgression within a species-rich oak (*Quercus* spp.) community. *BMC Evolutionary Biology* **7**: 218–233.

Craft KJ, Ashley MV. 2006. Population differentiation among three species of white oak in northeastern Illinois. *Canadian Journal of Forest Research* **36**: 206–215.

Dodd RS, Afzal-Rafii Z. 2004. Selection and dispersal in a multispecies oak hybrid zone. *Evolution* **58**: 261–269.

Dodd RS, Kashani N. 2003. Molecular differentiation and diversity among the California red oaks (Fagaceae; *Quercus* section Lobatae). *Theoretical and Applied Genetics* **197**: 884–892.

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

Durrett R, Buttel L, Harrison R. 2000. Spatial models for hybrid zones. *Heredity* **84**: 9–19.

Endler JA. 1977. *Geographic variation, speciation, and clines*. Princeton, NJ: Princeton University Press.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**: 47–50.

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

Ferran M. 1997. *SPSS para WINDOWS. Programación y análisis estadístico*. Madrid: McGraw Hill.

Fitzpatrick BM, Schaffer HB. 2004. Environment-dependent admixture dynamics in a tiger salamander hybrid zone. *Evolution* **58**: 1282–1293.

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

González-Rodríguez A, Arias DM, Oyama K. 2005. Genetic variation and differentiation of populations within the *Quercus affinis* – *Quercus laurina* (*Fagaceae*) complex analyzed with RAPD markers. *Canadian Journal of Botany* **83**: 153–162.

Goudet J. 1995. Fstat version 1.2: a computer program to calculate F statistics. *Journal of Heredity* **86**: 485–486.

Gugerli F, Walser JC, Dounavi K, Holderegger R, Finkeldey R. 2007. Coincidence of small-scale spatial discontinuities in leaf morphology and nuclear microsatellite variation of *Quercus petraea* and *Q. robur* in a mixed forest. *Annals of Botany* **99**: 713–722.

Gugerli F, Brodeck S, Holderegger R. 2008. Utility of multilocus genotypes for taxon assignment in stands of closely related European white oaks from Switzerland. *Annals of Botany* **102**: 855–863.

Harrison RG, Rand DM. 1989. Mosaic hybrid zones and the nature of species boundaries. In: Otte D, Endler J. eds. *Speciation and its consequences*. Sunderland, MA: Sinauer Associates, 110–133.

Himrane H, Camarero JJ, Gil-Pelegrín E. 2004. Morphological and ecophysiological variation of the hybrid oak *Quercus subpyrenaica* (*Q. faginea* × *Q. pubescens*). *Trees* **18**: 566–575.

Hubisz M, Falush D, Stephens M, Pritchard J. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* **9**: 1322–1332.

James JK, Abbott RJ. 2005. Recent, allopatric, homoploid hybrid speciation: the origin of *Senecio squalidus* (Asteraceae) in the British Isles from a hybrid zone on Mount Etna, Sicily. *Evolution* **59**: 2533–2547.

Jensen RJ, Ciofani KM, Miramontes LC. 2002. Lines, outlines, and landmarks: morphometric analyses of leaves of *Acer rubrum*, *Acer saccharinum* (Aceraceae) and their hybrid. *Taxon* **51**: 475–492.

Kaul RB. 1985. Reproductive morphology of *Quercus* (Fagaceae). *American Journal of Botany* **72**: 1962–1977.

Kaplan Z, Fehrer J. 2007. Molecular evidence for a natural primary triple hybrid in plants revealed from direct sequencing. *Annals of Botany* **99**: 1213–1222.

Kimball S. 2008. Links between floral morphology and floral visitors along an elevational gradient in a *Penstemon* hybrid zone. *Oikos* **117**: 1064–1074.

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.

Lepais O, Petit RJ, Guichoux E, *et al*. 2009. Species abundance and direction of introgression in oaks. *Molecular Ecology* **18**: 2228–2242.

Lexer C, Kremer A, Petit RJ. 2006. Shared alleles in sympatric oaks: recurrent gene flow is a more parsimonious explanation than ancestral polymorphism. *Molecular Ecology* **15**: 2007–2012.

McCauley RA, Cortés-Palomec AC, Oyama K. 2007. *Phylogeography and historical gene flow patterns in disjunct Quercus across the Sierra Madre Occidental and southern Cordillera of Mexico*. Botanical Society of America/American Society of Plant Biologists, Chicago, IL. Available at http://www.2007.botanyconference.org/engine/search/index.php

Moore WS, Price JT. 1993. Nature of selection in the northern flicker hybrid zone and its implications for speciation theory. In: Harrison RG. ed. *Hybrid zones and the evolutionary process*. New York: Oxford University Press, 196–255.

Muir G, Schlötterer C. 2005. Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Molecular Ecology* **14**: 549–561.

Nielsen EE, Bach LA, Kotlick P. 2006. Hybrilab (version 1.0): a program for generating simulated hybrids from population samples. *Molecular Ecology Notes* **6**: 971–973.

Nixon KC. 1993. The genus *Quercus* in Mexico. In: Ramamoorthy TP, Bye R, Lot A, Fa J. eds. *Biological diversity of Mexico: origins and distribution*. New York: Oxford University Press, 447–458

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

Raudnitschka D, Hensen I, Oberprieler C. 2007. Introgressive hybridization of *Senecio hercynus* and *S. ovatus* (Compositeae, Senecioneae) along an altitudinal gradient in Harz national Park (Germany). *Systematics and Biodiversity* **5**: 333–344.

Rieseberg LH. 1997. Hybrid origins of plant species. *Annual Review of Ecology and Systematics* **28**: 359–389.

Rieseberg LH, Wendell JF. 1993. Introgression and its consequences in plants. In: Harrison RG. ed. *Hybrid zones and the evolutionary process*. New York: Oxford University Press, 70–109.

Rohlf FJ. 2004. *TPSDIG, digitize landmarks and outlines*, version 2.0. Stony Brook, NY: Department of Ecology and Evolution, State University of New York.

Tovar-Sánchez E, Oyama K. 2004. Natural hybridization and hybrid zones between *Quercus crassifolia* and *Quercus crassipes* (Fagaceae) in Mexico: morphological and molecular evidence. *American Journal of Botany* **91**: 1352–1363.

Tovar-Sánchez E, Mussali-Galante P, Esteban-Jiménez R, *et al*. 2008. Chloroplast DNA polymorphism reveals geographic structure and introgression in *Quercus crassifolia* × *Q. crassipes* complex in México. *Botany* **86**: 228–239.

Trelease W. 1924. The American Oaks. *Memories of the National. Academy of Sciences* **20**: 1–255.

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

Valencia S. 2004. Diversidad del género *Quercus* (Fagaceae) en México. *Boletín de la Sociedad Botánica de México* **75**: 33–53.

Valencia S, Delgado A. 2003. Los tricomas foliares en la caracterización de un grupo de especies del género *Quercus*, sección Lobatae (Fagaceae). *Anales del Instituto de Biología* **74**: 5–15.

Van Oosterhout C, Hutchinson WF, Wills DPM. 2004. Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**: 535–538.

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

Williams DG, Ehleringer JR. 2000. Carbon isotope discrimination and water relations of oak hybrid populations in southwestern Utah. *Western North American Naturalist* **60**: 121–129.

Williams JH, Boecklen W J, Howard DJ. 2001. Reproductive processes in two oak (*Quercus*) contact zones with different levels of hybridization. *Heredity* **87**: 680–690.

Whittemore AT, Schaal BA. 1991. Interspecific gene flow in sympatric oaks. *Proceedings of the National Academy of Sciences USA* **88**: 2540–2544.

Zelditch ML, Swiderski DL, Sheets HD, Fink WL. 2004. *Geometric morphometrics for biologists: a primer*. New York: Elsevier.

VI.

Juan Manuel Peñaloza-Ramirez, Natalia Trujillo-Arias, and Ken Oyama.

Phylogeographic structure, population demography and ecological niche modeling, in the multispecies red oak hybrid complex formed by *Quercus hypoleucoides*, *Q. scytophylla*, and *Q. sideroxyla* (Fagaceae) in Mexico.

Enviado a Molecular Ecology mayo 2011.

Phylogeographic structure, population demography and ecological niche modeling in the multispecies red oak hybrid complex formed by *Quercus hypoleucoides, Q. scytophylla* and *Q. sideroxyla* (Fagaceae) in Mexico

J. M. Peñaloza-Ramírez*, N. Trujillo-Arias+ and K. Oyama*

**Centro de Investigaciones en Ecosistemas, Laboratorio de Ecología Genética, Universidad Nacional Autónoma de México (UNAM), Antigua Carretera a Pátzcuaro No. 8701, Col. Ex-Hacienda la Huerta, 58190, Morelia (Phone: + 443 322-27-04).*

+Universidad del Quindío, Avenida Bolívar Calle 12 norte Armenia, Colombia.

Keywords: cpSSR, cytoplasmic exchange, ecological niche modeling, Mexico, phylogeography, population history and demography, *Quercus hypoleucoides*, *Quercus scytophylla*, *Quercus sideroxyla*, red oak species.

Corresponding author: *Peñaloza-Ramírez Juan Manuel. *Centro de Investigaciones en Ecosistemas, Laboratorio de Ecología Genética, Universidad Nacional Autónoma de México (UNAM), Antigua Carretera a Pátzcuaro No. 8701, Col. Ex-Hacienda la Huerta, 58190, Morelia (Fax: + 443 322-27-19).* mail: jpenaloza@oikos.unam.mx

Running title: Phylogeography and niche modeling in the red oak hybrid complex in Mexico

Abstract

In this study, we evaluated the phylogeographic structure, the cytoplasmic exchange dynamics and ecological niche modeling in the multispecies red oak hybrid complex formed by *Q.hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla* in their geographic distribution in Mexico. In total we analyzed 51 populations with eight polymorphic chloroplast microsatellite loci to study the patterns of genetic diversity, structure and the signals of historical population expansion. The potential distribution models were also obtained for present-day and two models of the last glacial maximum (LGM, 21000 years) using MAXENT algorithm. We found higher levels of genetic diversity and low genetic structure, possibly due to the amount of cytoplasmic exchange from *Q. hypoleucoides* haplotypes toward *Q. sideroxyla* and *Q. scytophylla* for some populations in the contact zone. The three species had a strong phylogeographic structure and a north-south haplotypes distribution in their main lineages that suggest an in-situ origin, differentiation and then followed by colonization in different migrations routes. Also star-like shape forms in the haplotype network revealed population expansion events. Populations with expansions signals may have served as a reservoir of genetic diversity during climate oscillations. Also, the LGM models indicated that the species experienced a moderate reduction in their latitudinal distribution but we suggest that populations experienced altitudinal displacements rather than latitudinal migrations or colonizations as responses to the different glacial periods and, as a consequence high genetic variation levels and local haplotypes were preserved, which implies large effective population sizes and a complex dynamics in the cytoplasmic exchange.

Introduction

Studies about genetic variation between and within species are an important factor to elucidate the distribution of biological diversity (Schaal *et al.* 1998). The distribution of this variation is often modulated by factors such as ecological events, selective process or by the patterns of genetic exchange (Avise 2000). Phylogeographic approaches can potentially determine process such as the genealogical relationships between haplotypes, population bottlenecks or expansions (Schaal *et al.* 1998; Avise 2000). Phylogeographic studies also detect the coincidence or concordance of geographic variation in genotypes or in their genealogies to understand microevolution and speciation processes (Schaal *et al*. 1998; Templeton 1998). Nevertheless, few studies have incorporated environmental data in phylogeographic interpretation (Hugall *et al*. 2002; Cilibertii *et al*. 2009; Jakob *et al*. 2009). Ecological niche modeling has been incorporated to phylogeographic studies to predict the potential distribution of species based on present locality points extrapolated to areas with the most suitable ecological condition (Guisan & Thuiller 2005; Richards *et al*. 2007). Transferring the current climate conditions of a species into climate models of the past gives an estimation of the species potential paleodistribution assuming stable ecological niches of species (Wiens & Graham 2005). These models take into account temperature, precipitation, elevation and other variables that potentially influence the distribution of species (Phillips *et al*. 2006). Recently, ecological niche modeling has been applied to the fields of ecology, biogeography, evolution and conservation biology (Guisan & Thuiller 2005; Richards *et al*. 2007).

Temperate oak forests have been exposed to very different scenarios in their evolutionary history, and in several cases have resulted in common patterns. For instance, some studies have provided information about population history, genetic

structure, bottlenecks, and hybridization (Dumolin-Lapegue *et al*. 1997; Le Corre & Kremer 1998; Belahbib *et al*. 2001; Magni *et al*. 2005). In others, the effects of Pleistocene glaciations have often documented the existence of several refugia in Europe and North America where oak species occurred in common areas (Petit *et al*. 2002; Grivet *et al*. 2006), that resulted in extensive cytoplasmic exchanges between sympatric species of white oaks in United States (Whittemore & Schaal 1991; Marsico & Hellman 2009) and Europe (Dumolin-Lapegue *et al*. 1999; Belahbib *et al*. 2001; Magri *et al.* 2006).

Several studies have suggested that climatic fluctuation during glaciations was less severe in the tropics than in other regions (Grivet *et al*. 2005; Magri *et al*. 2006). In Mexico, fossil records of pollen indicate a decrease in 5-6 °C in temperature during Pleistocene glaciations (Bradbury 1997; Lozano-García & Xelhuantzi-López 1997) contrasting with the 10°C decrement occurred in North America and Europe (Martínez-Hernández 1992; Metcalfe 2006). In consequence, in the tropical and sub-tropical forests it could be expected that the strength of climatic fluctuations was possibly less severe and more variable (Flenley 1998; Metcalfe *et al*. 2000; Lachniet & Vázquez-Selem 2005). Some studies in Mexico revealed a higher level of genetic diversity, moderate genetic structure and high cytoplasmic exchange between species (González-Rodríguez *et al*. 2005; Tovar-Sánchez *et al*. 2008; Albarrán-Lara *et al*. 2010; Peñaloza-Ramírez *et al*. 2010). Nevertheless, less is known about the impact of climatic fluctuations in the population history, demography, genetic structure, and establishment of sympatric zones of contact and hybridization in the Mexican oak species. Particularly, the Sierra Madre Occidental have experienced climate changes over time, progressive warming alternated with cold periods, which fluctuated moderately in climate since the Pleistocene to the transition of the Holocene (Martínez-Hernández

1992; Metcalfe 2006). Changes in the distribution of oak forests in northern Mexico promoted the formation of a zone of contact and hybridization between the species under study *Q. scytophylla*, *Q. sideroxyla* and *Q. hypoleucoides* (Peñaloza-Ramirez *et al*. 2010). The aim of this study is to reconstruct the phylogeographic patterns and determine the patterns of cytoplasmic exchange between the red oak species: *Q. scytophylla*, *Q. sideroxyla* and *Q. hypoleucoides* in Mexico. The specific objectives were to i) estimate the genetic diversity and structure and determine the amount of genetic exchange of *Q. scytophylla*, *Q. sideroxyla* and *Q. hypoleucoides*, ii) reconstruct the phylogeographic structure and test the possible signals of population expansions, and iii) compare the findings with the ecological niche modeling inferences to understand the historical processes that have operated in the evolution of these species.

Materials and Methods

Study species

Quercus hypoleucoides A Camus. belongs to section *Lobatae*, it is a medium-sized tree 10-20 m in height. Leaves are narrowly lanceolate, lower surface with dense white tomentum. This species are frequent between 2000 to 2500 m in elevation, mainly in oak forests or in associations with pine-oak forests. The flowering period occurs during March-June and the acorns mature from August to December. This species is distributed in southwestern United States in Arizona, New Mexico and Texas, and in Mexico has a northern distribution in the states of Sonora, Chihuahua, Durango, Sinaloa and Coahuila (Valencia 2004).

Quercus scytophylla Liebm. is a red oak species which belongs to section *Lobatae*, it is a moderate tall tree that measure up to 20m in height, leaves are oblanceolate often with 1-6 teeth distributed toward the middle distal, underside

covered with white compressed sessile hair. Populations are found between 1540-2600 m of altitude, mostly in oak forest and pine-oak forest, but sometimes are frequent in association with tropical dry forest. The flowering period occurs between February to March, acorns mature from July to September. This species has a wide geographic distribution in the states of Chihuahua, Sonora, Sinaloa, Durango, Zacatecas, Jalisco, Michoacán, Estado de México, Guerrero, Oaxaca and Chiapas (Valencia 2004).

Quercus sideroxyla Humb. & Bonpl. is a red oak species that belongs to *Lobatae* section, it is a small tree about 7-8m in height, leaves are thickened and rigid, often with 1-5 deltoid teeth, confined to the distal third of the leaf, lower surface with a dense yellowish-white tomentum. This species is frequent at altitudes between 2400-2600 m in oak and pine-oak forests. The flowering period comprises the months from June to August and acorns mature from November to January. This species has a moderate geographic distribution in the states of Chihuahua, Coahuila, Tamaulipas, Nuevo León, Durango, Zacatecas, Jalisco, Aguascalientes, Guanajuato and San Luis Potosi (Valencia 2004).

Sampling procedure and molecular techniques

Samples were collected in 51populations of the three red oak species *Q. hypoleucoides, Q. scytophylla*, and *Q. sideroxyla* in their distribution in Mexico, as follows: 14 populations of *Q. hypoleucoides,* 21 of *Q. scytophylla* and 16 of *Q. sideroxyla*. Within populations, 8 to 12 individuals were randomly selected with at least 10 m of separation between consecutive samples (Table 1). Genomic DNA was obtained from 100 mg of leaf material using the technique designed by Lefort and Douglas (1999). Eight chloroplast DNA (cpDNA) microsatellite loci were chosen and amplified in multiplex polymerase reactions (PCR). We designed, two groups of

primers that were situated according to allele size and fluorescent labels. The first group was conformed by the primers pairs UDT1, UDT 3, UDT4 and UCD 5 (Deguilloux *et al*. 2003). The second group of primers included CMCS6, CMCS10, UKK3 and UKK4 (Sebastiani *et al*. 2004). PCR was performed using the QIAGEN Multiplex PCR kit in a volume of 5 µl containing 1X Multiplex PCR Master Mix, 2 µM each primer, dH20, and 40 ng template DNA. Two different thermal cycling conditions were performed for each primer pairs, and consisted of 40 cycles, each at 95°C for 1 min, but the annealing for the first primer-pairs was set at 48°C and for the second primers-pairs was set at 43°C. The rest of conditions are the same, 1 min and extension at 72°C for 2 min. a final extension at 72 °C for 10 min was included. Multiplex PCR products were combined with a GeneScan-500 LIZ size standard and then run in an ABI-PRISM 3100-Avant sequencer (Applied Biosystems). All the cpDNA fragments were then analyzed with the Peak Scanner program 1.0 (Applied Biosystems).

Genetic analysis

In order to compare the genetic diversity between populations from *Q. hypoleucoides*, *Q.scytophylla* and *Q. sideroxyla*, we utilized the RAREFAC program in populations with different sample size to estimate with rarefaction the mean within population genetic diversity (h_S), total gene diversity (h_T), haplotype richness (Petit *et al*. 1998) , as well we estimate the population differentiation for unordered alleles (G_{ST}) and for ordered alleles (R_{ST}) with the CPSSR program (Pons & Petit 1996), if R_{ST}, which takes into account the genetic differences between the haplotypes, is higher than G_{ST}, indicates the presence of phylogeographic structure in the populations, all the calculations were performed at 10^4 permutations (Pons & Petit 1996). We performed a hierarchical test of population structure by means of AMOVA in ARLEQUIN 3.5

(Excoffier *et al*. 2005). We used the two different assemblage criteria to estimate the variance components of genetic variation: i) considering all the three species simultaneously, and ii) considering each species separately, all significance of partitions was tested using 10^4 permutations in ARLEQUIN 3.5 (Excoffier *et al.* 2005). Finally, we construct a haplotype network considering each species separately utilizing a Median-Joining Network approach, this method combines the topology of a minimum spanning tree with a parsimony-based MP (Maximum Parsimony) search to identify and remove unnecessary links between haplotypes, all these methods were performed in NETWORK 4.6.0 (Bandelt *et al*. 1995).

Historical population demography

We performed the tests Fu´s F_S statistic (Fu 1997) and mismatch distribution (Rogers & Harpendig 1992) to detect signals of historical demographic fluctuations. Particularly, the negative and significant values for Fu test F_S signifies an excess of recent substitutions events (i.e. presence of rare haplotypes) produced by population growth. But also, the mismatch distribution population expansions events are expected to have an unimodal distribution, whereas a multimodal distribution is predicted for constant populations (Rogers & Harpendig 1992). In order to use ARLEQUIN for the analysis, the cpSSR data were binary coded following Navascués *et al*. (2006), the number of repeats was coded with '1' and shorter alleles were coded filling the differences in repeats with '0'. Significance of *Fs* values was evaluated with 10^4 data bootstraps (Excoffier *et al*. 2005), for the mismatch distribution analysis, the Harpending´s Raggedness index was used to evaluate the goodness of fit of the observed distributions to those expected under the model of population expansions. Additionally, the distributions were utilized to estimate the time to the population

expansion ($\tau=2l\mu t$) with the initial and present effective population sizes scaled by mutation rate ($\Theta_0=2\mu N_0$ and $\Theta_1=2\mu N_1$), where l is the number of cpSSR loci, μ represents the mutation rate, t the number of generations since the population expansion occurred and N_0 and N_1 are the effective population sizes previous and after expansion (Rogers & Harpendig 1992). For this analysis, we used the pseudo-likelihood method implemented in the program LMSE 1.1 (Navascués *et al*. 2009), this method takes into account the homoplasy in the calculations, which is common for the cpSSRs data. However, there is no estimate of mutation rate available for cpSSR for oak species. But recently, mutation rates in the range of 10^{-5} to 10^{-4} mutations per generation per locus have been usually utilized in the literature, as well as generation times for trees around 25 to 100 years (Navascués *et al*. 2009).

Pleistocene and present ecological predictive models

We modeled the ecological niche models for *Q. hypoleucoides, Q. scytophylla* and *Q. sideroxyla* using a total of 912 locality records obtained from herbarium specimens at the Herbario Nacional de México (MEXU), the TROPICOS database (Missouri Botanical Garden), and the Global Biodiversity Information Facility database (GBIF; http://www.data.gbif.org/). In order to reduce spatial autocorrelation a 0.1º filter was applied, resulting in a total of 566 occurrences records to run the model We used the Maximum Entropy algorithm implemented in MAXENT 3.3.1 (Phillips *et al*. 2004) to project the LGM (e.g. LGM, ~21ka) potential distribution of the species using data from 19 climatic variables. Present climate layers were based on the WorldClim database (Hijmans & Graham 2006; http://www.worldclim.org) and past climate layers on two past climate scenarios developed by the Paleoclimate Modelling Intercomparison Project Phase II (Braconnot *et al.* 2007): the Community Climate

System Model (CCSM; Collins *et al.* 2004) and the Model for Interdisciplinary Research on Climate (MIROC, Hasumi & Emori 2004). Layers grid sized used for both projections was 0.0083333338*0.0083333338 (~30 Arcseg). A subset of 75% of Q. castanea records were used to run the model and the remaining 25% were used to test the model. We utilized a threshold independent method for model validation, using the area under the receiver operating characteristic curve (e.g. AUC; Cilibertii *et al*. 2009). Ten independent runs using bootstrap replications were performed for each model, to assure the convergence on similar solutions. For the three distribution models (i.e. present, CCSM and MIROC) we performed a Jackknife test to measure the relative importance of each climatic variable on the occurrence prediction. As a validation test, two subgroups randomly generated from the original occurrences were used to run two different MAXENT models for both LGM projections and to compare if the resulting distribution extension converges with the one observed for the full points runs. Finally, to evaluate the proportion of changes in the distribution of the three species between the LGM and the present, an average value of distribution extent was obtained for each model from the ten replicate runs, and the averages were those compared between the models.

Results

Diversity and genetic structure

Genetic diversity analysis for *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla* in their distribution in Mexico, resulted in 108 total haplotypes distributed as follows: *Q. hypoleucoides* had 24 haplotypes in total, 1 ancestral, 5 shared within the species, and 9 unique haplotypes; *Q. scytophylla* had 57 haplotypes, 5 ancestral, 23 unique, and 22 shared within the species; and *Q. sideroxyla* included 41 haplotypes, 4

ancestral, 13 unique, and 15 shared within the species (Fig. 1). Only haplotype 14 was shared between the three species, five shared by *Q. hypoleucoides* and *Q. sideroxyla* (haplotypes 2, 3, 4, 8, 16), three shared by *Q. hypoleucoides* and *Q. scytophylla* (9, 10, 15) and three shared among *Q. scytophylla* and *Q. sideroxyla* (41, 42, 50) (Fig. 1). Parameters of within-population genetic diversity for *Q. hypoleucoides* resulted in moderate to high genetic diversity values (h_S = 0.671, h_T = 0.932), the differentiation for ordered alleles (R_{ST} = 0.643) was significantly higher than unordered alleles (G_{ST} = 0.280). In the case of *Q. scytophylla*, the genetic diversity values were h_S = 0.708 and h_T =0.972, and the differentiation for ordered and unordered alleles were R_{ST} =0.708 and G_{ST} =0.271, respectively. For *Q. sideroxyla*, the genetic diversity values were h_S =0.695 and h_T =0.981, and the differentiation for ordered and unordered alleles resulted in R_{ST} =0.809 and G_{ST} =0.291, respectively. For all the species, the differentiation for ordered alleles (R_{ST}) was significantly higher than for the unordered alleles (G_{ST}) indicating a strong phylogeographic structure in their populations (Pons & Petit 1996). Estimates of gene diversity corrected for unequal population sizes with RAREFAC, haplotype richness and genetic distance among individuals within population D^2sh (Table 1) resulted in higher values in populations where the cytoplasmic exchange was more frequent, such as Huracán A and B, Guadalupe, Pinos Altos and Querari that belong to *Q. hypoleucoides*, Poleo, Espinazo, Cuarenteño, and that correspond to *Q. scytophylla*, and Maycoba, Guachochic, P. Nuevo and Ceboruco that belong to *Q. sideroxyla* (Table 1). The analysis of molecular variance (AMOVA), considering all species together and taking into account each species separately (Table 2), indicated that most of the cpDNA diversity was located among groups with about 53%, 63%, 71% and 79%, respectively, while the differentiation among populations within groups was 34% for all species together, and the differentiation within populations accounted for 11%, 36%, 28% and

20% of the total variation. In general, we observed a high degree of structure between the three species analyzed (Table 2). The haplotype network for the 51 populations of the three species calculated independently (Fig. 2) showed for *Q. hypoleucoides* the presence of only one ancestral haplotype, and haplotypes 9, 10 and 15 shared with *Q. scytophylla* (in green) and the haplotypes 2, 3, 4, 8 and 16 shared with *Q. sideroxyla* (in blue). With respect to *Q. scytophylla*, we observed 5 ancestral haplotypes and in general the network showed a structured genealogy. The haplotypes 41, 42 and 50 were shared with *Q. sideroxyla* (in brown) and the haplotypes 9, 10 and 15 were shared with *Q. hypoleucoides* (in green) (Fig. 2). The network of *Q. sideroxyla* also showed a structured genealogy with 4 ancestral haplotypes. We identified the haplotypes 41, 42 and 50 shared with *Q. scytophylla* (in brown) and the haplotypes 2, 3, 4, 8 and 16 shared with *Q. hypoleucoides* (in green) (Fig. 2). In general, the networks for the three species showed a strong phylogeographic structure where haplotypes genealogically related occurs in geographical proximity. In the same way, ancestral haplotypes in the networks showed a widespread distribution, with several recently derived haplotypes in the peripheries.

Historical population demography

In most of the populations of *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla*, Fu's (F_S) values were significantly negative and non significant mismatch distribution in the Raggedness index, indicating that the observed distribution did not deviate significantly from an unimodal shape (Table 3). These results are frequently interpreted as populations that have undergone a recent range expansion. The average values of the time to expansion (parameter τ) for the populations in the three species were for *Q. hypoleucoides* $\tau = 1.1621$, for *Q. scytophylla* $\tau = 2.5930$, and for *Q.*

sideroxyla τ =1.1286. The estimation of the time since expansion for populations of *Q. hypoleucoides* occurred between 726,315 years ago (low mutation rate, high generation time) and 18,157 years ago (high mutation rate, short generation time). The estimates for *Q. scytophylla* were between 1,584,868 and 40, 516 years ago, and for *Q. sideroxyla* between 718,014 and 17,950 years ago.

Pleistocene and present potential distribution models

The AUC values (averaged over ten runs) for the training and test data in the current, MIROC and CCSM models were 0.990 and 0.962, respectively, indicating a good performance for the three models. The Jackknife analysis indicates the climate variables with the highest relative contribution to the present distribution and the LGM models were those related to the driest month and the minimum temperature of the coldest month in the case of *Q. hypoleucoides*. For *Q. scytophylla*, the most informative variables were those related to precipitation, the mean of the coldest quarter, and the temperature of the wettest quarter. For *Q. sideroxyla*, the main useful variables were those related to precipitation, seasonality and mean temperatures of the wettest quarter. Comparisons between the distribution extent for the present and the two LGM models indicated that between 30% and 40% of the areas currently occupied were considerably restricted during the last glacial maximum; there were some differences between both LGM models on the predicted distribution extent for the three species at specific geographic regions (Fig. 3). Although, the results showed that after the contraction occurred in the LGM, populations of the three species experienced moderate altitudinal migrations and colonization in some areas. Also the models showed considerably restrictions for *Q. hypoleucoides* in the most northern populations in the Sierra Madre Occidental (Fig. 3a-c). In contrast, for *Q. scytophylla* and *Q. sideroxyla* (Figs. 3 d-f and

g-i, respectively), the models resulted in similar areas of distribution for both species, with 10-20% of the areas currently occupied were moderately restricted for the LGM models, mainly for the northwestern populations in the Sierra Madre Occidental, but also for some populations along the Trans-Mexican Volcanic Belt and Sierra Madre del Sur.

Discussion

Diversity and genetic structure

In this study, we analyzed the variation of eight cpSSR loci in 51 populations from three red oak species in their distribution in Mexico. Higher values for the within and the total population genetic diversity, were observed for *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla*. Particularly, we found higher values in populations where the cytoplasmic exchange was more frequent. This could be indicative that the process of genetic exchange between these species could increase the genetic variation, as were observed in the red oak complex *Q. affinis* and *Q. laurina* and *Q. crassifolia*x*Q. crassipes* in which hybrid populations had higher values of genetic diversity and number of haplotypes compared with pure or allopatric populations of their putative parents (González-Rodríguez *et al*. 2004; Tovar-Sanchez *et al*. 2008). We also found evidences of chloroplast capture in northern populations of the Sierra Madre Occidental, although the direction in the cytoplasmic exchange is mediated by *Q. hypoleucoides* that showed the most invasive haplotypes toward *Q. scytophylla* and *Q. sideroxyla*. In this sense, in the Sierra Madre Occidental the occurrence of zones of contact and hybridization have been reported, such as the red oaks species *Q. eduardii* and *Q. conzatti* (Bacon & Spellenberg 1996), white oak species *Q. arizonica*, *Q. grisea*, *Q. laeta*, *Q. convallata*, *Q. obtusata*, *Q. chihuahuensis*, *Q. transmontana*, and *Q. striatula*,

and the hybrid species *Q. basaseachicensis* and their parents *Q. depressipes* and *Q. rugosa* (Spellenberg 1995). Bacon & Spellenberg (1996) suggested that the process of altitudinal expansion and contraction occurred during the Pleistocene in the Sierra Madre Occidental may have stimulated the cyclic hybridization as an important mechanism in the evolution of this plant species. And more importantly, the cytoplasm exchange among species promotes the acquisition of new characters, with the subsequent increase of genetic diversity and the colonization of these hybrid gene combinations to new environments (Rieseberg *et al.* 2007). It is quite possible that some characteristics in the environment are directly related to the success of the hybrid, as were indicated by Peñaloza-Ramirez *et al.* (2010) where the hybrid genotypes are also more delimited with the altitude of each population, as a possible effect of environmental dependent selection that operates in the distribution of hybrid genotypes.

We observed lower values of differentiation for *Q. hypoleucoides* (G_{ST} =0.280), *Q. scytophylla* (G_{ST} =0.271), and *Q. sideroxyla* (G_{ST} =0.291), that resulted similar to those reported in the red oak species in Mexico, such as the *Q. affinis* and *Q. laurina* complex (G_{ST} =0.49) (González-Rodríguez *et al.* 2005), for the widespread red oak species *Q. castanea* (G_{ST} = 0.261) (Peñaloza-Ramirez *et al.* unpub. data), and the differentiation for the northern red oak *Q. rubra* (G_{ST} = 0.47) (Magni *et al.* 2005). In contrast, white oaks from the western United States *Q. garryana* showed a higher genetic differentiation (G_{ST} = 0.88) (Marsico & Hellman 2009) as well as the California white oak *Q. lobata* (G_{ST} = 0.709) (Grivet *et al.* 2006). Also, white oaks in Europe also showed higher values of differentiation (*Q. robur*: G_{ST} =0.78; *Q. petreae*: G_{ST} = 0.86; *Q. pyrenaica*: G_{ST} =0.96; *Q. ilex*: G_{ST} = 0.92; and *Q. suber*: G_{ST} = 0.84) (Petit *et al.* 2002). The differences observed highlighted the low levels of differentiation in red oak species in contrast with white oaks in North America and Europe. One of the potential causes

could be associated to the dissimilar evolutionary history of oak species in both continents (Magni *et al.* 2005). These differences must have resulted in unique colonization histories, for example events such as long distance colonization might have been more frequent in Europe than in North America (Magni *et al.* 2005). White oaks were more restricted to the different refugia during the last glaciations producing a strong genetic structure in their populations (Petit *et al.* 2002). In contrast, paleoenvironmental records in Mexico suggest that the impact of climate changes during the Pleistocene (Metcalfe *et al.* 2000) in red oaks species was not as dramatic in their populations genetic structure, thus allowing the maintenance of higher level of diversity, low genetic structure and certain levels of exchange among populations (González-Rodríguez *et al.* 2005; Tovar-Sánchez *et al.* 2008). The analysis of molecular variance (AMOVA) suggest that despite the presence of chloroplast exchange between these three species, we found a higher significant differentiation among *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla* and among populations of each species separately. These results indicated that the populations also maintain certain levels of genetic structure (Dood & Kashani 2003; Magni *et al*. 2005), and that through the time have allowed the accumulation and maintenance of more genetic diversity (Le Corre & Kremer 1998; Grivet *et al*. 2006). Furthermore, we observed that the genetic differentiation decrease for some populations immersed in the zone of contact, whereas the differentiation augmented much more when increasing geographical distance with respect to contact zone, suggesting that possibly natural selection is operating to maintain the genetic identity of the species (Howard *et al.* 1997).

The network of haplotypes for the three species showed a strong phylogeographic structure according to G_{ST} and R_{ST} values, which indicated that populations genealogical related were also in geographical proximity (Pons & Petit,

1996). We found one shared haplotype between the three species and 11 shared haplotypes between different pairs of species; all shared haplotypes are distributed mainly in the northern zone of contact in the states of Chihuahua, Sonora and Durango. We observed a north-south haplotype distribution for the three species in their main lineages, suggesting and in-situ origin, differentiation and then followed by the colonization in different migrations routes. Also the star-like shape forms in the network for each species revealed population expansion events (Schaal *et al.* 1998). This finding indicates a great structure in the lineages and implies that the distribution of haplotypes in the three species possibly reflects the successive effect of altitudinal expansions-contractions, and migrations of oak populations during successive glacial and interglacial periods (Hewitt 1999; Widmer & Lexer 2001).

Historical population demography

The Fu's (F_S) neutrality tests and the mismatch distribution analysis suggest demographic expansions (Harpending *et al*. 1993), and are frequently associated with higher values of genetic diversity, population size and by the star-like shape form observed in the haplotype network (Schneider & Excoffier 1999). The shape of distribution also responds to time of expansion, where older expansion events are characteristic to right-shifted unimodal peak, and large negative values in the neutrality test F_S (Rogers & Harpending 1992). The oldest population expansion was for *Q. scytophylla* which dated between 1, 584,868 and 40, 516 years ago, for *Q. hypoleucoides* the expansion occurred between 726, 315 and 18, 157 years ago and for *Q. sideroxyla* between 718, 014 and 17, 950 years ago. Populations with signals of expansions have been interpreted to populations that have served as a reservoir of genetic diversity during climate oscillations (Hewitt 2000; Moreno-Letelier and Piñero

2009) for a relatively large population (Widmer & Lexer 2001). Also, the distribution models for the LGM for the three species we observed a moderate reduction in their geographic distribution, thus allowing the maintenance of certain levels of connectivity in their populations.

Pleistocene and present potential distribution models

The distribution models for the present and for the Last glacial maximum models (e.g. CCSM and MIROC) resulted in higher values for the AUC test (Cilibertii *et al*. 2009). The current distribution models differ substantially from the LGM models. For instance, we observed a much more significant reduction in the populations of the three species across the SMO for the LGM models, than the observed distribution for the current models. In this sense, palynological records in the mid Pleistocene suggest that alpine grasslands were much more extended in areas of Baja, Sonora and Chihuahua with cold and arid conditions (Lachniet & Vázquez-Selem 2005; Metcalfe 2006), associated with a retreat of forests into the middle altitudes, and the lowlands areas surrounded by deserts or grasslands (Martínez-Hernández 1992). In the late Pleistocene, temperatures were significantly cooler than present (5-6 °C) with more winter precipitation (Metcalfe *et al.* 2000; Lozano-García *et al.* 2005). These findings indicate a successive periods of altitudinal expansion-contractions in the populations of the three species that possibly promotes the hybridization and introgression between the populations in the SMO, that are in accordance with the cytoplasmic exchange invasion from *Q. hypoleucoides*, toward *Q. sideroxyla* and *Q. scytophylla* mainly in populations that are in the zones contact that also showed a significant increase of genetic diversity.

In the case of *Q. scytophylla* and *Q. sideroxyla* we observed a significant reduction of suitable habitat indicated in populations in the north and eastern of the

TMVB, but also in populations from SMS, than the distribution observed for the current distribution models. The paleo-limnological records from TMVB showed very complex patterns in the late Pleistocene; there may be a contrast between the western part of the region that was more humid than the present with the presence of species of *Alnus*, *Carpinus*, *Corylus* and *Salix*, and the eastern part which was drier than the present with *Cuppressaceae* and *Juniperus* species (Lozano-García & Xelhuantzi-Lopez 1997; Metcalfe 2006). The differences in the available habitat for oak populations presumably resulted in a vicariant effect between the populations across the Trans-Mexican Volcanic Belt as was indicated by the genetic differentiation between these populations. One of the potential reasons could be associated to the restrictions in available areas where the species could survive during glaciations, thus remaining moderately isolated to certain areas of the total distribution (Hewitt 2000; Widmer & Lexer 2001). Populations of *Q. scytophylla* in the SMS were moderately reduced according to the LGM models. In this sense, the pollen records during the last glacial maximum indicate a tendency of dry and cold conditions, causing the development of grasslands around the Basin (Lozano-García *et al.* 2005). The forests were reduced by a combination of two main factors, the reduction of humidity and volcanic activity in the area (Lozano-García & Ortega-Guerrero 1998; Metcalfe *et al.* 2000).

The oak species studied here evidenced a reduction in their distribution according to the LGM models. It is possible that populations experienced altitudinal displacements rather than latitudinal migrations or colonizations as responses to the different glacial periods and, as a consequence high genetic variation levels and local haplotypes were retained. Temperate forests at tropical latitudes during the glacial periods in the Pleistocene allowed the expansions and the increase of connectivity in the temperate trees as were suggested in some species of pines (Moreno-Letelier and Piñero

2009; Rodríguez-Banderas *et al.* 2009), as well as in oak species from Mexico have been largely defined by the geological and climatic history of the country, producing a relatively moderate contractions and dispersal, larger effective population sizes, and a complex dynamics in the cytoplasm exchange (González-Rodríguez *et al.* 2005; Tovar-Sánchez *et al.* 2008; Peñaloza-Ramírez *et al*. unpub. data). Finally, the process of colonization should be mediated by the particular ecological requirements from each species (Magri *et al*. 2006). For instance, species with adaptations to be in warm climatic conditions could recolonize better during the interglacial periods (e.g. Sosa *et al*. 2009), although species with clear adaptations to temperate climatic conditions should respond better to recolonize during the glacial periods (Hewitt 2000).

Conclusion

We found evidence of cytoplasmic exchange from *Q. hypoleucoides* toward *Q. sideroxyla* and *Q. scytophylla* that also increase the genetic diversity in populations from contact areas. Strong phylogeographic structure was found and a north-south colonization routes in their main lineages throughout Mexico. The distribution of the species according to the LGM models suggests a reduction of suitable habitats compared with the current models with wider distribution in their populations. We suggest that the effects of glacial and interglacial periods in the tropics were less severe (Flenley 1998; Metcalfe *et al.* 2000). The impact in their populations was minor but significant and more variable, possibly due because the climate is susceptible by the effect of latitude and more importantly by altitude (Metcalfe *et al*. 2000; Lacniet & Vázquez-Selem 2005). In consequence, populations of the three species had different responses during the glacial periods (Magni *et al*. 2005; Marsico & Hellman 2009). For instance, we observed in their populations of *Q. hypoleucoides*, *Q. sideroxyla* and *Q.*

scytophylla altitudinal expansions rather than latitudinal colonization, thus allowing the maintenance of higher levels of genetic diversity, certain levels of connectivity and larger effective population size.

Acknowledgements

The authors thank to J.J. Morrone, J. Núñez-Farfán, V. Rocha-Ramírez, D. Lugo-Aquino, A. Albarrán-Lara, H. Rodríguez-Correa, A. González-Rodríguez, O. Chassin-Noria, N. Pérez-Nasser, J.V. Guerrero-Hernández, P. Peñaloza for technical assistance and suggestions to improve the manuscript. This paper constitutes a partial fulfillment of the Graduate Program in Biological Sciences of the Universidad Nacional Autónoma de México, UNAM. J.M. Peñaloza-Ramírez received a Ph.D. scholarship provided by the National Council of Science and Technology (CONACyT) and UNAM. This project was supported by CONACyT grant 38550-V, CONACyT-SEMARNAT 23728 and 2004-C01-97 to K.O. Acknowledge to the grants of RLB-Andrew W. Mellon Foundation Grant2010-2011 (RLB2010-P13).

References

Albarrán-Lara AL, Mendoza-Cuenca L, Valencia-Avalos S, González-Rodríguez A, Oyama K (2010) Leaf fluctuating asymmetry increases with hybridization and introgression between *Quercus magnoliifolia* and *Quercus resinosa* (Fagaceae) through an altitudinal gradient in Mexico. *International Journal of Plant Sciences*, **171**, 310-322.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, London, 447 pp.

Bacon JR, Spellenberg R (1996) Hybridization in two distantly related Mexican black oaks *Q. conzatti* and *Q. eduardii* (Fagaceae: *Quercus*: Sección *Lobatae*). *Sida*, **17**, 17-41.

Bandelt HJ, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations. *Genetics*, **141**, 743-753.

Belahbib N, Pemonge MH, Quassou A, Sbay H, 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.

Bradbury JP (1997) Sources of glacial moisture in Mesoamerica. *Quaternary International*, **44**, 97-110.

Braconnot P, Otto-Bliesner B, Harrison S, Joussaume S, Peterchmitt JY, Abe-Ouchi A, Crucifix M, Driesschaert E, Fichefet T, Hewitt CD, Kageyama M, Kitoh A, Loutre MF, Marti O, Merkel U, Ramstein G, Valdes P, Weber L, Yu Y, Zhao Y (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. *Climate Past*, **3**, 279-296.

Cilibertii P, De Jongi H, Schouten MA, Verweji PA, Heitmansi WRB (2009) Ecological niche models for four native cockroach species (Blattaria: *Blattellidae*: *Ectobiinae*) in the Netherlands. *Contributions to Zoology*, **78**, 149-156.

Collins WD, Blackmon M, Bitz C, Bonan G, Bretherton CS (2004) The community climate system model: CCSM3. *Journal of Climate*, **19**, 2122-2143.

Deguilloux MF, Dumolin-Lapegue S, Gielly L, Grivet D, Petit RJ (2003) A set of primers for the amplification of chloroplast microsatellites in *Quercus*. *Molecular Ecology Notes*, **3**, 24-27.

Dodd RS, Kashani N (2003) Molecular differentiation and diversity among the California red oaks section *Lobatae*. *Theoretical and Applied Genetics*, **107**, 884-892.

Dumolin-Lapegue S, Demesure B, Fineschi S, Le Corre V, Petit RJ (1997) Phylogeographic structure of white oaks throughout the European continent. *Genetics*, **146**, 1475-1487.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver 3.01. An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47-60.

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

Fu YX (1997) Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics*, **147**, 915-925.

González-Rodríguez A, Bain JF, Golden JL, Oyama K (2005) Chloroplast DNA variation in the *Quercus affinis*-*Q. laurina* complex in Mexico: Geographical structure and associations with nuclear an morphological variation. *Molecular Ecology*, **13**, 3467-3476.

Grivet D, Deguilloux MF, Petit RJ Sork VL (2006) Contrasting patterns of historical colonization in white oaks (*Quercus* spp.) in California and Europe. *Molecular Ecology*, **15**, 4085-4093.

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993-1009.

Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. *Current Anthropology*, **34**, 483-496.

Hasumi H, Emori S (2004) K-1 coupled GCM (MIROC) description. Tokyo: Center for Climate System Research, University of Tokyo, Japan.

Hewitt G (1999) Postglacial recolonization of European Biota. *Biological Journal of the Linnean Society*, **68**, 87-112.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907-913.

Hijmans RJ, Graham CH (2006) The ability of climatic envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 2272-2281.

Howard DJ, Preszler RW, Williams J, Fenchel S, Boecklen WJ (1997) How discrete are oak species? Insights from a hybrid zone between *Quercus grisea* and *Quercus gambelii. Evolution*, **51**, 747-755.

Hugall A, Moritz C, Moussalli A, Stanisic J (2002) Reconciling paleodistribution models and comparative phylogeography in the wet tropics rainforest land snail *Gnarosophia bellendenkerensis*. *Proceedings of the National Academy of Sciences, USA*, **99**, 6112–6117.

Jakob SS, Martínez-Meyer E, Blattner FR (2009) Phylogeographic analyses and paleodistribution modeling indicate Pleistocene in situ survival of *Hordeum* species (*Poaceae*) in Southern Patagonia without genetic or spatial restriction. *Molecular Biology and Evolution*, **26**, 907-923.

Lachniet MS, Vázquez-Selem L (2005) Last Glacial Maximum equilibrium line altitudes in the circum-Caribbean (Mexico, Guatemala, Costa Rica, Colombia, and Venezuela). *Quaternary International*, **138**, 129-144.

Le Corre V, Kremer A (1998) Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping stone model. *Journal of Evolutionary Biology*, **11**, 495-512.

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 and Zacapu. *Quaternary International*, **43**, 117-123.

Lozano-Garcia MS, Ortega Guerrero B (1997) Late Quaternary environmental changes of the central part of the Basin of Mexico; correlation between Texcoco and Chalco Basins. *Review of Palaeobotany and Palynology*, **99**, 77-93.

Lozano-García S, Sosa-Nájera S, Sugiera Y (2005) 23, 000 yr of vegetation history of the Upper Lerma, a tropical high-altitude basin in Central Mexico. *Quaternary Research*, **64**, 70-82.

Magni CR, Ducousso A, Caron H, Petit RJ, Kremer A (2005) Chloroplast DNA variation of *Quercus rubra* L. in North America and comparison with other Fagaceae. *Molecular Ecology*, **14**, 513-524.

Magri D, Vendramin GG, Comps B (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, **171**, 199-221

Marsico TD, Hellmann JJ (2009) Patterns of seed dispersal and pollen flow in *Quercus garryana* (Fagaceae) following post-glacial climatic changes. *Journal of Biogeography*, **36**, 929-941.

Martínez-Hernández E (1992) La vegetación y el clima de México. Historia del Cretácico tardío al Terciario basada en evidencias palinológicas. *Ciencia*, **43**, 69-72.

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

Metcalfe SE (2006) Late Quaternary environments of the northern deserts and central transvolcanic belt of Mexico. *Annals of the Missouri Botanical Garden*, **93**, 258-273.

Moreno-Letelier A, Piñero D (2009) Phylogeographic structure of *Pinus strobiformis* Engelm. accross the Chihuahuan Desert filter-barrier. *Journal of Biogeography*, **36**, 121-131.

Navascués M, Emerson CC (2006) Chloroplast microsatellites: measures of genetic diversity and the effect of homoplasy *Molecular Ecology*, **14**, 1333-1341.

Navascués M, Hardy OJ, Burgarella C (2009) Characterization of demographic expansions from pairwise comparisons of linked microsatellite haplotypes. *Genetics*, **181**, 1013-1019.

Peñaloza-Ramírez JM, González-Rodríguez A, Mendoza-Cuenca L, Caron H, Kremer A, Oyama K (2010) Interspecific gene flow in a multispecies oak hybrid zone in the Sierra Tarahumara of Mexico. *Annals of Botany*, **105**, 389-399.

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844-855.

Petit RJ, Csaikl UM, Bordacs S, Burg K, Coart E, Cottrell J, Van Dam B, Deans J.D, Dumolin-Lapegue S, Fineschi S, Finkeldey R, Gillies A, Glaz I, Goicoechea PG, Jensen JS, Konig AO, Lowe AL, Madsen SF, Matyas G, Munro RC, Olalde M, Pemonge MH, Popescu F, Slade D, Tabbener H, Taurchini D, de Vries SGM, Ziegenhagen B, Kremer A (2002) 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.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.

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

Richards CL, Carstens BC, Knowles LL (2007) Distribution modeling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographic hypotheses. *Journal of Biogeography*, **34**, 1833-1845.

Rieseberg LH, Kim SC, Randell RA, Whitney KD, Gross BL, Lexer C, Clay C (2007) Hybridization and the colonization of novel habitats by annual sunflowers *Genetica*, **129**, 149-165.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552-569.

Rodríguez-Banderas A, Vargas-Mendoza CF, Buonamici A, Vendramin GG (2009) Genetic diversity and phylogeographic analysis of *Pinus leiophylla*: a post-glacial range expansion. *Journal of Biogeography*, **36**, 1807-1820.

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, **7**, 465-474.

Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079-1089.

Sebastiani F, Carnevale S, Vendramin GG (2004) A new set of mono- and dinucleotide chloroplast microsatellites in *Fagaceae Molecular Ecology Notes*, **4**, 259-261.

Sosa V, Ruiz-Sánchez E, Rodríguez-Gómez FC (2009) Hidden phylogeographic complexity in the Sierra Madre Occidental: the case of the Mexican tulip poppy *Hunnemannia fumariifolia* (Papaveraceae). *Journal of Biogeography*, **36**, 18-27.

Spellenberg R (1995) On the hybrid nature of *Quercus basaseachicensis* (Fagaceae, sect. *Quercus*). *Sida*, **16**, 439-437.

Templeton AR (1998) Species and speciation. Geography, population structure, ecology, and gene trees. In: *Species and Speciation* (ed. Howard DJ, Berlocher SH), 32-43 pp. Oxford Univ. Press, New York.

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 *Quercus crassifolia* x *Q. crassipes* complex in México. *Botany*, **86**, 228-239.

Valencia AS (2004) Diversidad del género *Quercus* (*Fagaceae*) en México. *Boletín de la Sociedad Botánica de México*, **75**, 33-53.

Widmer A, Lexer C (2001) Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology and Evolution*, **16**, 267-269.

Wiens JJ, Graham CH (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, **36**, 519-539.

Whittemore AT, Schaal BA (1991) Interspecific gene flow in sympatric oaks. *Proceedings of the National Academy of Sciences, USA*, **88**, 2540-2544.

Figure legends

Figure 1. Map showed haplotype frequencies in populations of *Q. hypoleucoides* with a pointed line (populations 1-14), *Q. scytophylla* no line (populations 15-35) and *Q. sideroxyla* two points line (populations 36-51) in Mexico. All 108 haplotypes are represented by different combinations of colors and patterns.

Figure 2. Median-joining haplotype networks for the 108 haplotypes identified in 51 populations of *Q. hypoleucoides* (i.e. white circles), *Q. scytophylla* (i.e. yellow circles) and *Q. sideroxyla* (i.e. gray circles) in Mexico. Each circle represents an individual haplotype and circle size is proportional to the frequency of the haplotype. Shared haplotypes in the three species are represented in blue for *Q. hypoleucoides* and *Q. sideroxyla*, in green shared haplotypes between *Q. hypoleucoides* and *Q. scytophylla*, in brown haplotypes shared between *Q. scytophylla* and *Q. sideroxyla*, and in red are the haplotypes shared among the three species.

Figure 3. Maps showing the potential distribution as probability of occurrence (green = 0-0.2, red = 0.8 - 1.0) for *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla* in their distribution in Mexico under present climatic conditions (figures a, d and g), at the Last Glacial Maximum (LGM; 21 ka) based on the CCSM model (figures b, e and h), and at the LGM based on the MIROC 3.2 model (figures c, f and i).

Table 1 Geographical location and estimates of genetic diversity for 51 populations of *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla* in Mexico. N, sample size; *hS*, within-population genetic diversity; D^2sh, mean pairwise genetic distance among individuals within a population under a stepwise mutation model. Allelic richness is reported after a rarefaction analysis to standardize for unequal sample sizes.

	Populations	*N*	*Coordinates*	*Altitude (m)*	*Allelic richness*	h_S	D^2sh
1	Huracan A	8	29°40'/108°15'	2208	3.0	0.750 (0.139)	0.053
2	Huracan B	8	29°37'/108°22'	2300	5.0	0.929 (0.084)	0.041
3	Rio Chico	8	29° 36'/108°10'	2006	3.0	0.750 (0.139)	0.053
4	Agua A	8	29°11'/108°14'	2250	2.0	0.464 (0.200)	0.017
5	Agua B	8	29°12'/108°14'	2321	1.0	0.429 (0.169)	0.026
6	Guadalupe	10	29°11'/107°57'	1852	3.7	0.861 (0.087)	0.043
7	Temosachic	8	29°8'/107°56'	2182	1.0	0.429 (0.169)	0.026
8	Babicora	8	29°13'/107°49'	2240	2.0	0.607 (0.164)	0.035
9	Pinos Altos	12	28°15'/108°19'	2190	3.4	0.833 (0.069)	0.038
10	Pto Cruz	8	28°22'/109°01'	1944	3.0	0.821 (0.101)	0.106
11	Yecora	8	28°22'/109°02'	1660	2.0	0.464 (0.200)	0.418
12	Km 346	9	28°26'/108°32'	1510	2.0	0.464(0.200)	0.038
13	Querari	10	27°11'/107°32'	2188	4.1	0.844 (0.103)	0.032
14	Batopilas	9	27°08'/107°34'	1900	2.	0.750 (0.112)	0.027
	Q. hypoleucoides (pooled)						
1	Poleo	8	29°45'/108°32'	2240	5.0	0.929 (0.084)	0.119
2	Madera	8	29°5'/108°3'	2360	3.0	0.750 (0.139)	0.043
3	Maguarichi	8	27°53'/107°57'	1852	2.0	0.607 (0.164)	0.097
4	Canelas	8	25°07'/106°30'	2601	1.0	0.250 (0.180)	0.562
5	Loberas	8	23°29'/105°51'	1955	1.0	0.429 (0.169)	0.026
6	Espinazo Diablo	8	23°34'/105°50'	1955	5.0	0.929 (0.084)	0.192
7	Puertecito	8	21°39'/103°09'	2367	3.0	0.750 (0.139)	0.045
8	Surutato	10	25°46'/107°33'	1557	2.7	0.778 (0.091)	0.039
9	Cuarenteño	8	21°27'/105°0'	1279	4.0	0.857 (0.108)	0.267
10	Villa Madero	8	19°19'/101°22'	2326	2.9	0.800 (0.089)	0.227
11	Tzararacua	8	19°21'/102°4'	1512	5.0	0.600 (0.129)	3.107
12	Valle de Bravo	8	19°10'/100°07'	1908	5.0	0.929 (0.084)	0.847
13	Temascal EM	8	19°02'/99°59'	1907	3.0	0.821 (0.101)	0.080
14	Tejupilco	9	18°51'/100°27'	1724	2.7	0.694 (0.147)	0.030
15	Atlixtac	9	17°34'/99°02'	2028	2.0	0.722 (0.097)	0.239
16	Pozo Largo	8	17°36'/99°05'	1858	2.0	0.679 (0.122)	0.222
17	Juquila	9	16°13'/97°09'	1957	5.0	0.893 (0.111)	0.143
18	Santos Reyes	8	16°29'/96°59'	1735	1.0	0.429 (0.169)	0.026
19	SJ Lachao	8	16°13'/97°8'	1867	1.0	0.536 (0.123)	0.033
20	Ayutla	8	17°02'/96°04'	1977	3.0	0.821 (0.101)	0.035
21	Duraznillo	9	17°0'/96°06'	1927	1.8	0.667 (0.105)	0.031
	Q. scytophylla (pooled)						
1	Maycoba	8	28°17'/108°06'	2375	3.6	0.857 (0.108)	0.049
2	Lajas	8	27°55'/107°54'	2218	1.0	0.536 (0.123)	0.033
3	Samachique	8	27°12'/107°31'	2210	1.0	0.571 (0.094)	0.035
4	Guachochic	8	26°56'/107°08'	2577	4.0	0.905 (0.103)	0.040
5	Cuevecillas	8	25°02'/106°16'	2600	2.6	0.643 (0.184)	0.084
6	Buenos Aires	9	23°42'/105°43'	2725	1.7	0.556 (0.165)	0.846
7	Tecuan	9	23°55/105°00'	2598	1.0	0.500 (0.128)	0.031
8	P. Nuevo	10	23°29'/105°21'	2752	2.3	0.711 (0.117)	0.048
9	Peñas	8	23°36'105°24'	2820	3.7	0.893 (0.086)	0.131
10	Temascal I	8	23°21'/104°17'	2583	1.0	0.429 (0.169)	0.107
11	Temascal II	8	23°23'/104°17'	2585	3.6	0.857 (0.108)	0.434
12	Temascal III	8	23°24'/104°16'	2585	2.7	0.750 (0.139)	0.030
13	Ceboruco	9	21°06'/104°30'	1065	4.0	0.889 (0.091)	0.039
14	Noria	10	21°29'/104°59'	1557	1.6	0.511 (0.164)	0.097
15	Rosa de Lima	8	21°04'/101°11'	2548	2.7	0.786 (0.113)	0.446
16	Calvillo	11	21°05'/101°09'	2460	2.4	0.733 (0.101)	0.028
51	***Q. sideroxyla (pooled)***						

TABLE 2 Hierarchical analysis of molecular variance (AMOVA) using R_{ST} for the three species together and for populations of *Q. hypoleucoides, Q. scytophylla* and *Q. sideroxyla* separately. Asterisks indicate statistically significant values ($P < 0.01$). Tests were based on 10^4 random permutations.

Source of variation	d.f.	SS	Variance components	Percentage of variation	Fixation index
Three species					
Among groups	2	4092.8	13.747 Va	53.74	Φ_{CT} =0.537***
Among populations within groups	48	3776.6	8.9107 Vb	34.83	
Within populations	383	1119.6	2.9344 Vc	11.43	
Total	433	8989.1	25.5818		
Q. hypoleucoides					
Among groups	13	132.437	1.09738 Va	63.25	Φ_{ST} =0.632***
Within populations	108	68.858	0.63758 Vb	36.75	
Total	121	201.295	1.73496		
Q. scytophylla					
Among groups	20	1537.43	8.8484 Va	71.24	Φ_{ST} =0.712***
Within populations	153	546.464	3.5716 Vb	28.76	
Total	173	2036.89	12.42014		
Q. sideroxyla					
Among groups	15	2106.80	15.8173 Va	79.28	Φ_{ST} =0.792***
Within populations	122	504.355	4.13406 Vb	20.72	
Total	137	2611.15	19.9514		

Table 3 Test for historical population expansions and estimation of demographic parameters for 51 populations of *Q. hypoleucoides*, *Q. scytophylla* and *Q. sideroxyla* in Mexico. Θ_0 and Θ_1 are, respectively, ancestral and current population sizes scaled by mutation rate; τ is the number of generations since the expansion occurred, scaled by mutation rate; -log [CL] is the likelihood of the model. See text for details.

No	Populations	*Fu´s Fs*	*P-value*	*Raggedness index*	*P-value*	*Population growth parameters (with homoplasy)*			
						Θ_0	Θ_1	τ	-log[CLh]
	Q. hypoleucoides								
1	Huracan A	-10.243	0.000	0.146	0.550	$3.21x10^{-5}$	$1.72x10^{13}$	1.148	22.32
2	Huracan B	-7.334	0.000	0.160	0.410	$2.29x10^{-5}$	$1.39x10^{13}$	2.238	9.699
3	Rio Chico	-10.240	0.000	0.146	0.640	$3.21x10^{-5}$	$1.54x10^{13}$	1.148	22.32
4	Agua A	-14.522	0.000	0.167	0.760	$1.95x10^{-5}$	$2.23x10^{12}$	0.516	13.14
5	Agua B	-15.453	0.000	0.204	0.600	$6.84x10^{-6}$	$1.29x10^{12}$	0.440	23.48
6	Guadalupe	-10.376	0.000	0.137	0.300	$2.17x10^{-5}$	$2.94x10^{13}$	1.942	17.55
7	Temosachic	-15.453	0.000	0.204	0.590	$6.84x10^{-6}$	$1.57x10^{12}$	0.440	23.48
8	Babicora	-11.911	0.000	0.956	0.840	$1.89x10^{-5}$	$1.57x10^{10}$	0.826	18.12
9	Pinos Altos	-17.688	0.000	0.168	0.130	$1.19x10^{-5}$	$3.16x10^{13}$	1.471	30.46
10	Pto Cruz	-7.334	0.000	0.149	0.490	$0.6x10^{-4}$	$3.31x10^{5}$	1.563	18.40
12	Yecora	-6.381	0.000	0.381	0.940	7.10326	0.5687	0.38	31.05
11	Km 346	-15.403	0.000	0.179	1.00	$6.06x10^{-6}$	0.9326	1.336	20.56
13	Querari	-12.033	0.000	0.097	0.490	$2.50x10^{-5}$	$2.06x10^{13}$	1.749	14.06
14	Batopilas	-7.760	0.000	0.136	0.530	$3.26x10^{-5}$	$1.46x10^{13}$	1.066	15.67
	Q. scytophylla								
1	Poleo	-5.531	0.000	0.056	0.810	2.847	365.71	1.417	14.59
2	Madera	-10.797	0.000	0.230	0.240	$2.10x10^{-5}$	$1.33x10^{13}$	1.025	20.82
3	Maguarichi	-9.180	0.000	0.248	0.320	$1.36x10^{-5}$	2.766	2.115	25.23
4	Canelas	-8.566	0.000	0.687	0.670	14.427	0.1488	0.201	58.77
5	Loberas	-15.453	0.000	0.204	0.520	$6.84x10^{-6}$	$1.57x10^{12}$	0.440	23.48
6	Espinazo Diablo	-4.610	0.002	0.155	0.240	$5.08x10^{-5}$	22.805	6.223	14.88
7	Puertecito	-9.180	0.000	0.299	0.170	$2.28x10^{-5}$	$7.41x10^{12}$	1.439	14.58
8	Surutato	-10.570	0.000	0.080	0.600	$5.14x10^{-5}$	$1.04x10^{13}$	1.296	22.06
9	Cuarenteño	-4.890	0.002	0.461	0.05	$5.6x10^{-3}$	8.743	7.838	21.84
10	Villa Madero	-6.959	0.000	0.204	0.440	$1.30x10^{-6}$	4.418	3.931	32.77
11	Tzararacua	-2.825	0.040	0.207	0.280	$5.24x10^{-5}$	$3.11x10^{13}$	3.600	20.843
12	Valle de Bravo	-2.620	0.040	0.096	0.590	$2.90x10^{-5}$	$2.20x10^{12}$	2.185	42.414
13	Temascal EM	-7.668	0.000	0.144	0.520	$3.24x10^{-5}$	$2.11x10^{13}$	2.061	21.04
14	Tejupilco	-9.567	0.000	0.087	0.600	$2.50x10^{-5}$	$1.52x10^{13}$	1.066	15.676
15	Atlixtac	-8.093	0.000	0.564	0.01	$4.30x10^{-5}$	6.608	3.721	44.41
16	Pozo Largo	-6.889	0.000	0.565	0.08	$2.18x10^{-6}$	5.193	3.700	33.119
17	Juquila	-5.241	0.001	0.128	0.210	$7.01x10^{-5}$	22.576	7.859	13.697
18	Santos Reyes	-15.453	0.000	0.204	0.640	$6.48x10^{-6}$	$1.29x10^{12}$	0.440	23.480
19	SJ Lachao	-14.111	0.000	0.292	0.320	$7.71x10^{-6}$	$3.24x10^{12}$	0.554	27.651
20	Ayutla	-9.602	0.000	0.218	0.220	$8.77x10^{-6}$	$2.30x10^{13}$	1.313	13.830
21	Duraznillo	-6.003	0.000	0.155	0.500	$1.84x10^{-5}$	$2.36x10^{10}$	0.817	23.145
	Q. sideroxyla								
1	Maycoba	-7.849	0.000	0.063	0.760	$1.04x10^{-14}$	$1.22x10^{13}$	1.964	11.48
2	Lajas	-14.111	0.000	0.292	0.300	$7.71x10^{-6}$	$3.24x10^{12}$	0.554	27.65
3	Samachique	-13.731	0.000	0.346	0.160	$1.18x10^{-5}$	$3.89x10^{12}$	0.552	28.96
4	Guachochic	-5.663	0.000	0.132	0.550	$2.88x10^{-5}$	$2.85x10^{13}$	1.829	10.58
5	Cuevecillas	-3.956	0.01	0.170	0.680	$1.58x10^{-5}$	$2.49x10^{12}$	1.592	38.96
6	Buenos Aires	-4.729	0.005	0.320	0.990	$3.12x10^{-5}$	$2.91x10^{12}$	1.421	29.342
7	Tecuan	-9.830	0.000	0.201	0.310	$2.39x10^{-6}$	$2.78x10^{12}$	0.516	33.81
8	P. Nuevo	-8.047	0.000	0.617	0.860	$2.62x10^{-5}$	$8.71x10^{12}$	1.298	29.89
9	Peñas	-7.334	0.000	0.082	0.660	$2.6x10^{-2}$	$1.04x10^{13}$	2.943	20.83
10	Temascal I	-11.432	0.000	0.693	0.940	$4.26x10^{-6}$	1.2730	1.6759	37.77
11	Temascal II	-3.978	0.009	0.117	0.540	9.559	$7.31x10^{4}$	$1.7x10^{-6}$	27.79
12	Temascal III	-10.243	0.000	0.146	0.640	$3.21x10^{-5}$	$1.11x10^{13}$	1.1480	12.75
13	Ceboruco	-9.011	0.000	0.060	0.810	$4.19x10^{-5}$	$4.29x10^{13}$	2.210	10.30
14	Noria	-10.986	0.000	0.137	1.00	1.7045	0.006595	0.0012	42.32
15	Rosa de Lima	-6.689	0.000	0.193	0.280	8.2849	0.71320	0.6478	32.62
16	Calvillo	-16.945	0.000	0.188	0.250	$1.06x10^{-5}$	$1.97x10^{10}$	0.7618	33.75

Figure 1

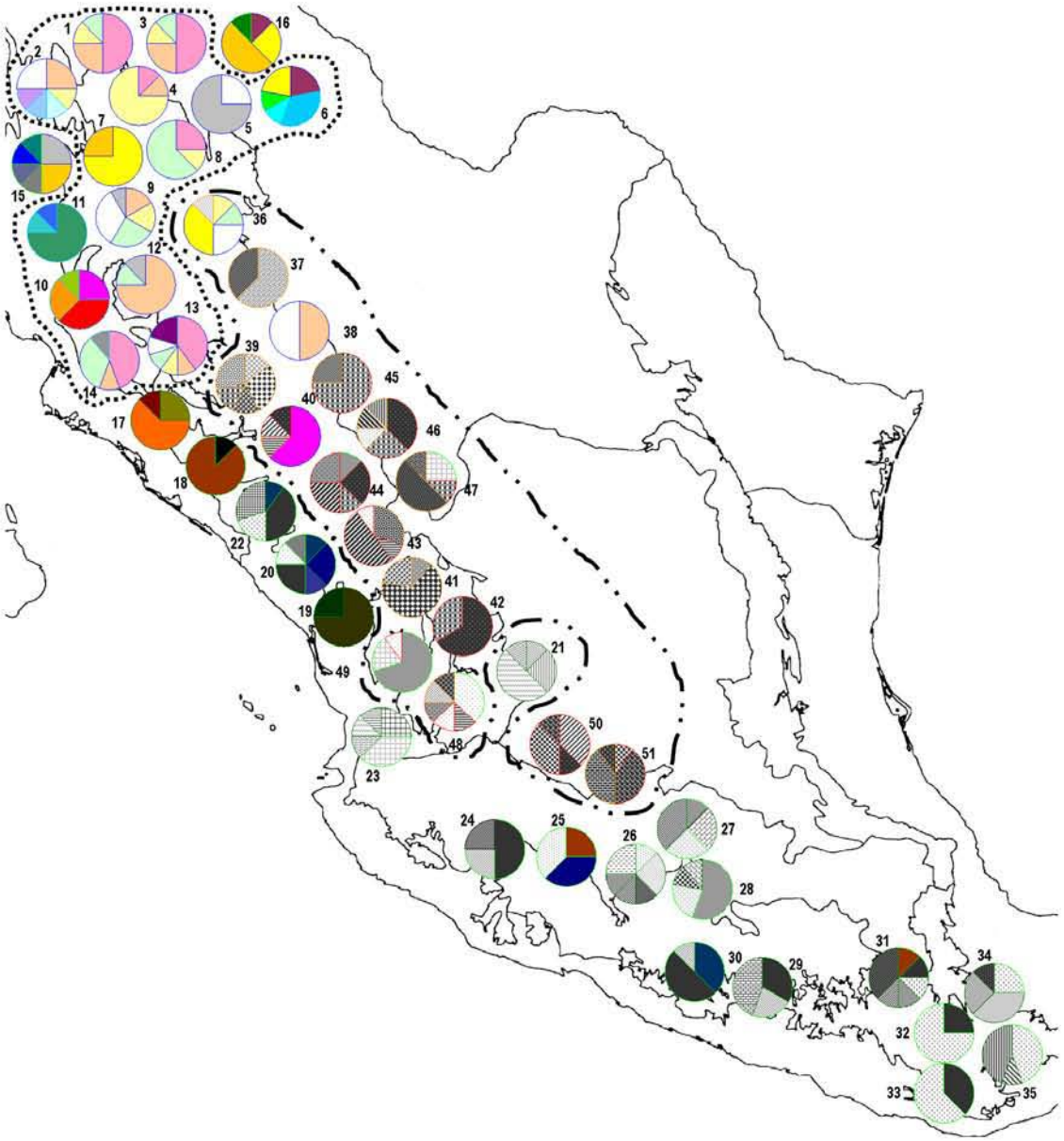

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

Figure 2

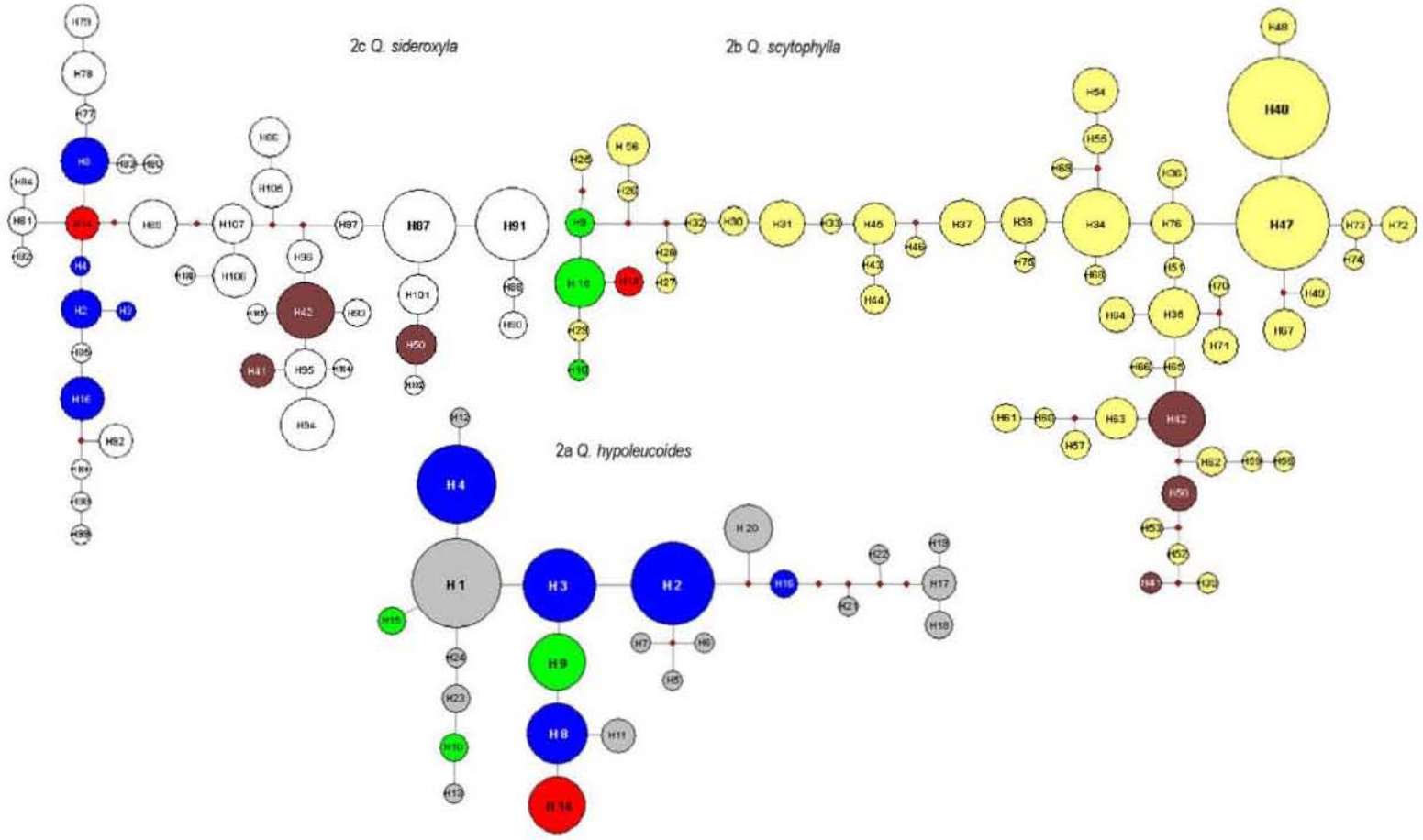
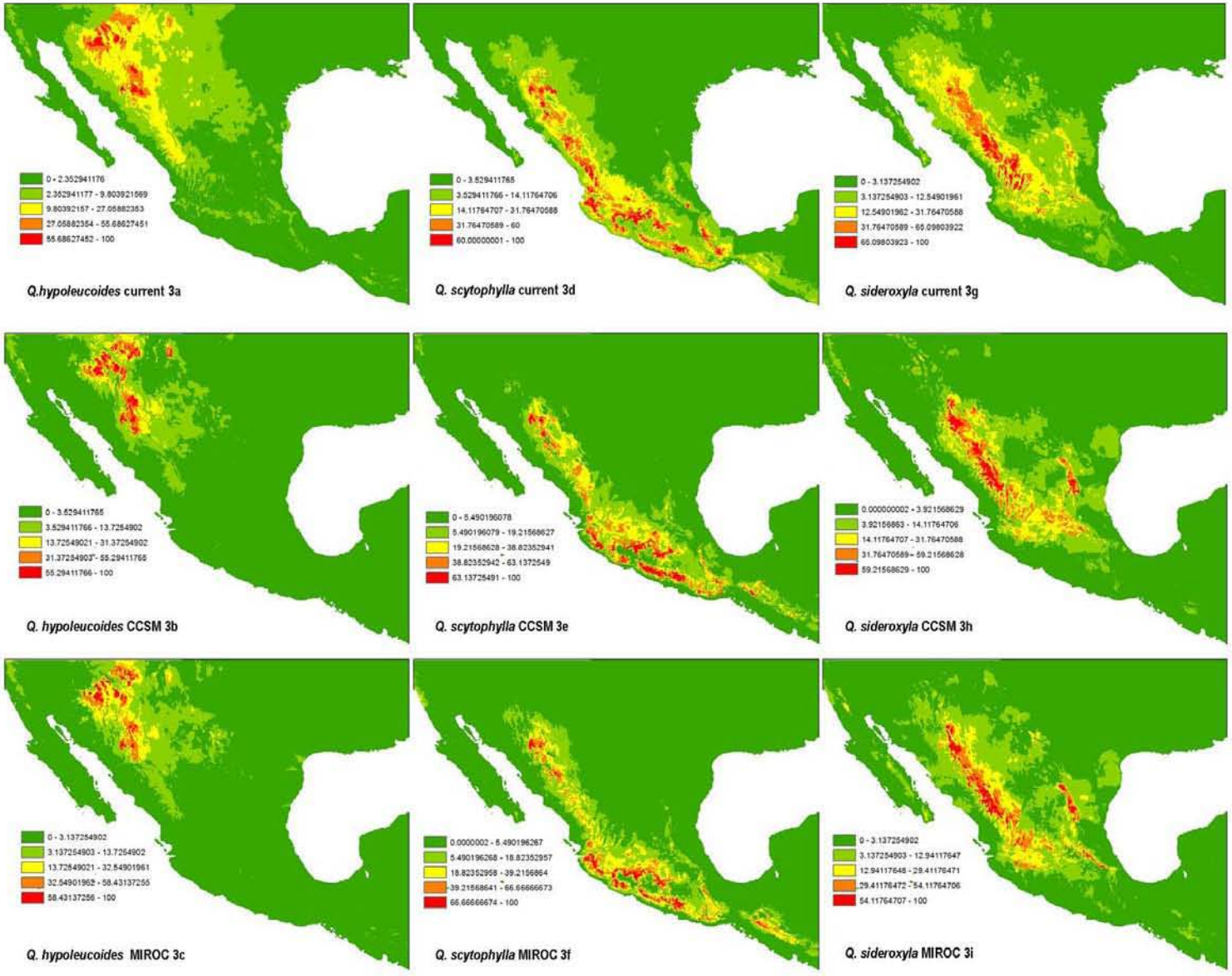


Figure 3


0 - 2.352941176
2.352941177 - 9.803921569
9.80392157 - 27.05882353
27.05882354 - 55.68627451
55.68627452 - 100
Q.hypoleucoides current 3a
0 - 3.529411765
3.529411766 - 14.11764706
14.11764707 - 31.76470588
31.76470589 - 60
60.00000001 - 100
Q. scytophylla current 3d
0 - 3.137254902
3.137254903 - 12.54901961
12.54901962 - 31.76470588
31.76470589 - 65.09803922
65.09803923 - 100
Q. sideroxyla current 3g
0 - 3.529411765
3.529411766 - 13.7254902
13.72549021 - 31.37254902
31.37254903 - 55.29411765
55.29411766 - 100
Q. hypoleucoides CCSM 3b
0 - 5.490196078
5.490196079 - 19.21568627
19.21568628 - 38.82352941
38.82352942 - 63.1372549
63.13725491 - 100
Q. scytophylla CCSM 3e
0.000000002 - 3.921568629
3.92156863 - 14.11764706
14.11764707 - 31.76470588
31.76470589 - 59.21568628
59.21568629 - 100
Q. sideroxyla CCSM 3h
0 - 3.137254902
3.137254903 - 13.7254902
13.72549021 - 32.54901961
32.54901962 - 58.43137255
58.43137256 - 100
Q. hypoleucoides MIROC 3c
0.0000002 - 5.490196267
5.490196268 - 18.82352957
18.82352958 - 39.2156864
39.21568641 - 66.66666673
66.66666674 - 100
Q. scytophylla MIROC 3f
0 - 3.137254902
3.137254903 - 12.94117647
12.94117648 - 29.41176471
29.41176472 - 54.11764706
54.11764707 - 100
Q. sideroxyla MIROC 3i

VII.

Juan Manuel Peñaloza-Ramírez, Antonio González-Rodríguez, Victor Rocha-Ramírez, Hernando Rodríguez-Correa, Enrique Martínez-Meyer and Ken Oyama.

Phylogeography and ecological niche modeling reveal Quaternary population history in the widespread Mexican red oak *Quercus castanea* Née.

Aceptado con cambios mayores en Journal of Biogeography, 14 septiembre 2010, correcciones enviadas a finales de mayo 2011.

Original article

Phylogeography and ecological niche modelling reveal Quaternary population history in the widespread Mexican red oak *Quercus castanea* Née.

Juan Manuel Peñaloza-Ramírez*, Antonio González-Rodríguez, Víctor Rocha-Ramírez, Hernando Rodríguez-Correa, Enrique Martínez Meyer[1] and Ken Oyama

Centro de Investigaciones en Ecosistemas, Universidad Nacional Autónoma de México (UNAM), Antigua Carretera a Pátzcuaro No. 8701, Col. Ex-Hacienda de San José de la Huerta, 58190, (Phone: +52 (443) 322-27-04) Morelia, México.

[1]Instituto de Biología, UNAM. Ciudad Universitaria. México, D.F.

*Correspondence: Juan Manuel Peñaloza-Ramírez

E-mail: jpenaloza@oikos.unam.mx

Running head: Phylogeography and ecological niche modelling of *Quercus castanea*

ABSTRACT

Aim To infer the Quaternary population history of the widespread Mexican red oak species *Quercus castanea*, through the analysis of phylogeographic structure, historical demographic processes, and present and Last Glacial Maxima (LGM) potential distribution models.

Location Four commonly recognized morphotectonic and biogeographic provinces of Mexico: the Sierra Madre Occidental (SMO), the Central Plateau (CP), the Trans-Mexican Volcanic Belt (TMVB), and the Sierra Madre del Sur (SMS).

Methods A total of 341 individuals from 36 populations of *Q. castanea* were genotyped with seven chloroplast microsatellite loci (cpSSRs). Data were analyzed to obtain standard statistics of genetic diversity and structure and to test for signals of historical demographic expansions. Potential distribution models were obtained for present-day and two scenarios of the Last Glacial Maximum (LGM) climatic conditions using the MAXENT algorithm.

Results A total of 90 haplotypes were identified, with 28 of these restricted to single populations. Within-population genetic diversity was high (mean $h_S = 0.72$), and among-population genetic differentiation showed a strong phylogeographic structure ($R_{ST} = 0.711 > G_{ST} = 0.261$; $P < 0.05$). Signals of demographic expansion were identified in most populations, with larger τ (time to the population expansion) values in the SMS, followed by the SMO, the TMVB, and finally the CP. Ecological niche models indicated a larger (~ 30-50 %) potential distribution extent for the species under LGM conditions than at present but in the same geographic regions.

Main conclusions High genetic diversity, strong phylogeographic structure and potential distribution models suggest in situ permanence of *Q. castanea* populations with large effective sizes during the Pleistocene and Holocene. Signals of population expansions probably predating the LGM and possibly as old as 1.6×10^6 - 1.9×10^5 years before

present were observed and suggested that the expansion occurred following a south-north direction. As with other Mexican tree species, the complex geologic and climatic history of the TMVB contributes to explain the origin and maintenance of a large proportion of the genetic diversity in this oak species.

Keywords

Chloroplast microsatellites, ecological niche modelling, historical demographic expansions, Mexico, phylogeography, *Quercus castanea*, Quaternary population history, red oaks.

Editor: Jack Williams

INTRODUCTION

Plant phylogeographic studies based on maternally inherited chloroplast DNA (cpDNA) are useful to understand the historical processes of gene exchange, population bottlenecks or expansions, and isolation between population lineages, which explain the current population genetic structure of a species (Ennos *et al.,* 1999; Avise, 2000). The increasing number of studies of this kind has also allowed comparisons across different taxa distributed within the same geographic regions and the detection of some common patterns indicating similar historical responses to geological or climatic events (e. g. Schneider *et al.,* 1998; Taberlet *et al.,* 1998; Riddle *et al.,* 2000). A recent alternative approach to evaluate the effects of this sort of events on the population history of plant species is to compare the same or closely related taxa between areas that experienced contrasting conditions and processes (Grivet *et al.,* 2006).

Extensive phylogeographic studies conducted on European oaks have revealed the effects of climate fluctuations in the Quaternary on the distribution and connectivity of their populations, indicating the existence of Pleistocene glacial refugia situated in the Iberian Peninsula, Italy and the Balkans, as well as strong patterns of genetic structure resulting from post-glacial recolonization routes (Dumolin-Lapegue *et al.,* 1997; Le Corre & Kremer, 1998; Belahbib *et al.,* 2001; Petit *et al.,* 2002). In contrast, North American oak species are characterized by a lower level of genetic structure (Magni *et al.,* 2005) or by a higher genetic diversity (Grivet *et al.,* 2006) than the European counterparts (but see Marsico & Hellman, 2009). These differences have been explained as a result of different population histories for oak species between the two continents, probably because in North America oak populations were less severely affected by climatic changes during the last glaciation, thus being more stable locally, with more restricted range of contractions and expansions, and less post-glacial colonization events (Grivet *et al.,* 2006).

Less is known about the history of oaks in the subtropical and tropical regions where the highest species richness occurs. Mexico has more than 160 species and is considered a major center of diversification for the genus *Quercus* in the world (Valencia, 2004). Previous phylogeographic studies conducted on two Mexican red oak species complexes (i.e. *Q. affinis*-*Q. laurina* and *Q. crassipes*-*Q. crassifolia*) have shown higher levels of genetic diversity and lower genetic structure in comparison to both European and North American species (González-Rodríguez *et al.,* 2004; Tovar-Sánchez *et al.,* 2008), implying large historical effective population sizes and dynamic gene flow among populations. It is possible that at tropical latitudes the glacial periods of the Pleistocene allowed the expansion and increased connectivity of montane temperate tree populations, as has been inferred for some pine species (Moreno-Letelier & Piñero, 2009; Rodríguez-Banderas *et al.,* 2009). However, the paleoenvironmental reconstructions of the extent and intensity of Pleistocene climatic changes in Mexico and the responses of plant communities are still incomplete (Lozano-García *et al.,* 2005; Metcalfe, 2006). Available data indicate an average temperature reduction of 6 °C during the Last Glacial Maximum (Metcalfe, 2006) with a 1200 m decrease in the altitude of glacier equilibrium lines (Lachniet & Vázquez-Selem, 2005). The palynological record suggests that alpine grasslands extended in areas with cold and arid conditions, while temperate forests were present at lower altitudes with higher humidity than at present (Lozano-García *et al.,* 2005; Metcalfe, 2006).

Phylogeographic studies complemented with pollen and macrofossil data have obtained a clearer picture of the history of the taxon under study (Petit *et al.,* 2002; Magri *et al.,* 2006; Afzal-Rafii & Dodd, 2007). However, in Mexico, the scarcity of data, the low resolution of the pollen record below the genus level, and the high species richness, make palynological information of little use for reconstructing the history of individual *Quercus* species. As an alternative, it has been argued that the modelling of species' current and past

distribution is a powerful tool to generate independent information for proposing and testing phylogeographic hypotheses (Richards *et al.,* 2007). The usefulness of this approach has been demonstrated in the identification of Pleistocene glacial refugia for 14 North American terrestrial vertebrates, since a significant correlation was found between the predictions of past distribution models and phylogeographic data (Waltari *et al.,* 2007).

In this study, we aimed at reconstructing the biogeographic and demographic history of the Mexican red oak, *Quercus castanea* Née on the basis of genetic information, and to contrast these results with models of both present and historical (Last Glacial Maximum, LGM) distribution of the species, with the far reaching goal of gaining further insight into the population history of montane temperate tree populations and their responses to past climatic changes in Mexico. *Quercus castanea* constitutes an interesting study system since it has one of the broadest geographical and altitudinal distributions among the Mexican red oaks, with populations in the Sierra Madre Occidental (SMO), the Central Plateau (CP), the Trans-Mexican Volcanic Belt (TMVB), and the Sierra Madre del Sur (SMS) (Valencia 2004; Fig. 1). During episodes of climatic change conditions were heterogeneous across regions (Lozano-García *et al.,* 2005; Metcalfe, 2006), and consequently it could be possible that populations situated in the different areas experienced contrasting historical processes. Particularly interesting for the present study is the TMBV. This morphotectonic province crosses the country in its central part from the Pacific coast to the coast of the Gulf of Mexico (Ferrusquía-Villafranca, 1993) and constitutes the transition between the Nearctic and Neotropical biogeographic regions in America (Ramamoorthy *et al*., 1993). Its complex topography, altitude and climate variability, coupled with its geographical position and geological history provide a mosaic of environments, habitats and microhabitats for a large number of species (Rzedowski & Rzedowski, 1989). The importance of this province as an area of diversification, species richness and endemism of several animal and plant

taxa, including oaks, has been highlighted (Rzedowski & Rzedowski, 1989; Ferrusquía-Villafranca, 1993; Nixon, 1993; Valencia 2004; Kapelle, 2006). However, few studies have tested its role in the origin and maintenance of genetic diversity within particular species (Albarran-Lara *et al.,* 2010).

Specifically, the objectives of the present study were (i) to determine the patterns of genetic diversity and differentiation of populations of *Q. castanea* across its entire geographical range using chloroplast DNA microsatellite loci (cpSSR) as molecular markers, (ii) to test for signals of historical demographic expansions, and iii) to contrast the conclusions derived from genetic data with models of the current and LGM potential distribution of *Q. castanea*. Finally, we contrasted the phylogeographic patterns revealed in this study with the regionalization based on morphotectonic or biogeographic provinces, and identified the centres of biodiversity in terms of genetic diversity and the process of genetic diversification.

MATERIALS AND METHODS

Study species

Quercus castanea Née belongs to section *Lobatae* (Nixon, 1993). It is a moderately large forest tree 10-18 m in height. Populations are found between 1100 and 2600 m of elevation, mostly in oak forests and pine-oak forests, but sometimes in association with tropical dry forest elements at lower altitudes. It grows in mountainous areas with warm to temperate humid climates, with average annual temperatures ranging from 10 to 26 °C, but more often from 12 to 20 °C (Kapelle, 2006), and where the rainy season occurs in the warm season of the year (Rzedowski, 1978). The flowering period occurs in April and May, and acorns mature from July to November (Valencia, 1995).

Sampling procedure and molecular techniques

Samples were collected from approximately 10 individuals in each of 36 populations of *Q. castanea* over all its distribution in Mexico (Table 1, Fig. 1). Within populations, individuals were randomly selected with at least 10 m of separation between consecutive samples. From each tree, five young intact leaves were collected and stored at -80°C for molecular analysis and a branch was pressed to prepare voucher specimens.

Genomic DNA was extracted from 100 mg of leaf material using the protocol designed by Lefort & Douglas (1999). Seven chloroplast DNA (cpDNA) microsatellite loci were selected and amplified in multiplex polymerase chain reactions (PCR). Two groups of primers were arranged according to allele size and fluorescent labels. The first group was formed by the primer pairs UDT1, UDT3, UCD5, UKK3 and UKK4 (Deguilloux *et al.,* 2003). The second group of primers included CMCS6 and CMCS10 (Sebastiani *et al.,* 2004). PCR was performed using the QIAGEN Multiplex PCR kit (QIAGEN) in a volume of 5 µl containing 1X Multiplex PCR Master Mix, 2 µM each primer, dH_20, and 20 ng template DNA. The thermal cycling conditions consisted of 40 cycles, each at 95°C for 1 min, annealing at 50°C for 1 min, and extension at 72°C for 2 min. A final extension at 72 °C for 10 min was included. Multiplex PCR products were combined with a GeneScan-500 LIZ size standard and then ran in an ABI-PRISM 3100-Avant sequencer (Applied Biosystems). Fragments were analyzed and sized with the Peak Scanner program 1.0 (Applied Biosystems).

Genetic diversity and structure

Genetic diversity parameters were estimated for each population and for four regions commonly recognized as morphotectonic and biogeographic units for Mexico (Ferrusquía-Villafranca, 1993; Morrone, 2005): Sierra Madre Occidental (SMO), Trans-Mexican

Volcanic Belt (TMVB), Central Plateau (CP), and Sierra Madre del Sur (SMS). The RAREFAC program (Petit *et al.,* 1998) was used to calculate for each population and region haplotype richness, the number of exclusive haplotypes and gene diversity (h_S) corrected for unequal sample sizes with rarefaction. D^2_{SH}, which is the mean pairwise genetic distance among individuals within a population under a stepwise mutation model (Goldstein *et al.,* 1995) was calculated using the LMSE program (Navascués *et al.*, 2009).

Overall genetic differentiation among populations was calculated by obtaining the mean within population genetic diversity (h_S), the total gene diversity (h_T) and the differentiation coefficients for unordered and ordered alleles (G_{ST} and R_{ST}, respectively) with the PERMUT-CpSSR software (Pons & Petit, 1996). If R_{ST}, which takes into account the genetic differences among the haplotypes, is significantly higher than G_{ST}, then there is phylogeographic structure in the populations (Pons & Petit, 1996).

Hierarchical tests of population structure were performed using analyses of molecular variance (AMOVA) with ARLEQUIN 3.5 (Excoffier *et al.,* 2005) to obtain estimates of both F_{ST} (based on the infinite alleles mutation model, IAM), and R_{ST} (based on the stepwise mutation model, SMM). The four above mentioned biogeographic provinces were used as grouping criteria to obtain the variance components of genetic variation among groups, among populations within groups, and within populations. Significance of partitions was tested using 10^4 permutations. Additionally, pairwise genetic differentiation among the four biogeographic provinces was estimated with both F_{ST} and R_{ST}. This analysis was also performed in ARLEQUIN 3.5 (Excoffier *et al.,* 2005) with 10^4 permutations.

A Median-Joining Network was constructed for *Q. castanea* haplotypes using NETWORK 4.5 (Bandelt *et al.,* 1995). This method combines the topology of a minimum

spanning tree with a maximum parsimony search to identify and remove unnecessary links between haplotypes.

To identify the geographic location of the most important genetic discontinuities among *Q. castanea* populations, the BARRIER 2.2 software (Manni *et al.,* 2004) was used. This program uses the Monmomier's maximum difference algorithm (Monmomier, 1973) to find "barriers" that correspond to the largest genetic distances among the populations involved in the barrier. For the barriers analysis, we used a pairwise matrix of average square genetic distance (ASD) (Goldstein *et al.,* 1995; Slatkin, 1995). Resampling random subsets of individuals within populations provided 100 bootstrap replicate distances to determine statistical significance for the predicted barriers.

Historical population demography

Tests for population expansions were performed using the F_S statistic (Fu, 1997) and mismatch distributions (Rogers & Harpending, 1992). F_S (Fu, 1997) evaluates the probability of observing a random neutral sample with a number of alleles similar or smaller than the observed value, given the observed number of pairwise differences, taken as an estimator of θ (Fu, 1997). Statistically significant F_S negative values suggest population expansion or genetic hitchhiking and result from an excess of recent mutations (i.e. rare haplotypes). In turn, positive F_S values indicate a reduction in the number of alleles, as would be expected from a recent population bottleneck or from overdominant selection. Simulations have shown that F_S is a more precise indicator of population expansion and genetic hitchhiking than Tajima's *D* (Ramos-Onsins & Rozas, 2002).

Expanding populations are also expected to have unimodal distributions for differences in repeat number in pairwise comparisons among individuals in a sample (i.e. the mismatch distribution), while in stationary populations the distribution is predicted to

be ragged and multimodal (Rogers & Harpending, 1992). Despite its dependency on the infinite-alleles model, analysis of mismatch distributions is robust to fairly large deviations from this model (Harpending *et al.,* 1993). In order to use ARLEQUIN 3.5 for these analyses, the cpSSR data were binary coded following Navascués *et al.* (2006). The number of repeats was coded with '1' and shorter alleles were coded filling the differences in repeats with '0'. Significance of F_S values was evaluated with 10^4 data bootstraps (Excoffier *et al.,* 2005). For the mismatch distributions, the raggedness index of Harpending (1994) was used to evaluate the goodness of fit of the observed distributions to those distributions expected under the model of population expansion. Also, the distributions were used to estimate the time to the population expansion ($\tau = 2\mu t$) and the initial and present effective population sizes scaled by mutation rate ($\Theta_0 = 2\mu N_0$ and $\Theta_1 = 2\mu N_1$), where μ represents the mutation rate, t the number of generations since the population expansion occurred and N_0 and N_1 are effective population sizes previous and after the expansion (Rogers & Harpending, 1992). The pseudo-likelihood method of Navascués *et al.* (2009) implemented in the LMSE program was used to estimate these parameters. This method takes homoplasy, which is common for cpSSRs, into account in the calculations.

Distribution modelling

A total of 523 *Q. castanea* occurrence records were compiled from herbarium specimens deposited at the Herbario Nacional de México (MEXU), the TROPICOS database (Missouri Botanical Garden) and the Global Biodiversity Information Facility database (GBIF; http://www.data.gbif.org/). In order to reduce spatial autocorrelation in climatic data, a 0.1º filter was applied to the initial set of occurrences, resulting in a total of 253 *Q. castanea* occurrence records to run the model. The Maximum Entropy algorithm

implemented in MAXENT ver. 3.3.1 (Philips et al., 2006) was used to project the LGM (Last Glacial Maximum ~21 ka) potential distribution of the species using data from 19 climatic variables. Present climate layers were based on the WorldClim database (Hijmans & Graham, 2006; http://www.worldclim.org) and past climate layers on two past climate scenarios developed by the Paleoclimate Modelling Intercomparison Project Phase II (Braconnot *et al*., 2007): the Community Climate System Model (CCSM; Collins *et al*., 2004), and the Model for Interdisciplinary Research on Climate (MIROC; Hasumi & Emori 2004). Layers grid size used for both projections was 0.0083333338 degrees (~30 arcseg). A subset of 75% of *Q. castanea* records were used to run the model and the remaining 25% were used to test the model. As a threshold independent method for model validation, we used the area under the receiver operating characteristic curve (AUC) (Fielding & Bell, 1997). Ten independent runs using bootstrap replication were performed for each model, to assure convergence on similar solutions. For the three distribution models (i.e. present, CCSM and MIROC) a jackknife test was used to measure the relative importance of each climatic variable on the occurrence prediction. As a model validation test, two subgroups randomly generated from the original occurrence data set were used to run two different MAXENT models for both the LGM and current climatic conditions, to compare if the resulting distribution extension converges with the one observed for the full set of data points. In order to evaluate if the models indicate changes in the distribution of *Q. castanea* between the LGM and the present, an average value of distribution extent was obtained for each model from the ten replicate runs, and the averages were compared among models.

RESULTS

Genetic diversity and structure

A total of 90 haplotypes were found in the 36 sampled populations of *Q. castanea* (Table 1; Fig. 1) with 28 haplotypes being unique to single populations. The number of haplotypes per population ranged from 2 to 7. At the level of morphotectonic provinces, the SMO (populations 1-2) had a total of seven haplotypes (all exclusive to this province), the TMVB (populations 3-22) had 58 haplotypes (43 exclusive), the CP (populations 23-26) had 14 haplotypes (six exclusive) and the SMS (populations 27-36) had 28 haplotypes (19 exclusive) (Fig. 1). Haplotype richness (corrected for unequal population size with RAREFAC) and genetic diversity (h_S) were higher for the TMVB (14.97 and 0.97, respectively) and the SMS (12.53, 0.95) than in the CP (9.45, 0.91) and the SMO (7, 0.87) (Table 1). Genetic diversity parameters were not correlated with latitude, longitude or altitude of the populations, except for D^2_{SH}, which increased with altitude ($r = 0.43$; $P = 0.008$).

According to the CpSSR program the mean (standard error) of within-population genetic diversity (h_S) was 0.730 (0.022), and the total gene diversity (h_T) was 0.989 (0.003). Population differentiation for ordered alleles (R_{ST}) was 0.711 (0.089) and for unordered alleles (G_{ST}) was 0.261 (0.022). The permutation test for the comparison of the two differentiation values was significant ($P < 0.05$), indicating phylogeographic structure in the populations of *Q. castanea*. The hierarchical analysis of molecular variance (AMOVA) showed that when distances among haplotypes are not taken into account (IAM), differentiation among the four regions is significant (10.16%, $P = 0.002$), but most of the genetic variation resides among populations within regions (49.92%, $P < 0.0001$), and 39.92% of the variation is within populations (Table 2). In contrast, if distances among haplotypes are taken into account (SMM) most of the genetic differentiation is among the four regions (53.63%, $P < 0.0001$), while differentiation among populations within regions is 25.91% ($P < 0.0001$) and 21.46% of the variation resides among populations (Table 2).

The pairwise matrix of differentiation among regions indicated that the lowest differentiation was between the TMVB and the SMS (F_{ST} = 0.105; P < 0.0001 and R_{ST} = 0.092; P < 0.0001) while the highest values of differentiation were among SMO and the other four regions (Table 3).

The network of the 90 haplotypes identified in *Q. castanea* is shown in Figure 2. In most of the cases, haplotypes were separated by only one mutational step. Some groups of closely related haplotypes were restricted to specific geographical regions. This is particularly clear for haplotypes H1-H7, which were the only haplotypes present in the SMO and were exclusive to this region. In the case of the CP, most of the haplotypes present were genealogically related, except for three haplotypes (H35, H36 and H44), which belong to a very different lineage and were present in populations 23 (H35 and H36), 24, 25 and 26 (H44) located in northwest TMVB. Within the SMS, populations 27 and 36 contained haplotypes H90, H88, H18 and H79, which are genealogically distant from the other haplotypes present in this region. Finally, the majority of haplotypes were present in the TMVB, with practically all lineages represented in this area.

Five genetic barriers with more than 50% bootstrap support were identified among the 36 populations of *Q. castanea* (Fig. 1). The first and most significant barrier (B1), with a bootstrap support of 95% was placed between the SMO and the TMVB. The second barrier (B2), with 89% bootstrap support value was within the TMVB region, suggesting that populations in the southern part of this region (populations 20, 21, 22) are differentiated from the rest of the TMVB populations and are more similar to populations in the SMS. The third barrier (B3) with 79% bootstrap support, separated populations 24, 25 and 26 in the CP from population 23 (also belonging to the CP), and from populations in the TMVB. The fourth barrier (B4) with 68% bootstrap support, indicated differentiation of the three westernmost populations in the TMVB (7, 8 and 9). Finally, the fifth barrier (B5)

with 59% bootstrap support, separated population 34 in the SMS from the other populations (Fig. 1).

Historical population demography

We obtained large negative and significant F_S values for all except two populations (Table 4), suggesting population expansions (Fu, 1997). Mismatch distributions also provided evidence for population expansions since in all cases the Harpending's raggedness index was not significant, indicating that the observed distributions did not deviate significantly from a unimodal shape (Table 4). Average values of the time to expansion parameter τ for populations in the four different regions were 2.238 for the SMS, 0.889 for the SMO, 0.826 for the TMVB and 0.267 for the CP. Since no reliable estimates of mutation rate for cpSSRs have been reported, the translation of these values into years is subjected to great uncertainty. However, mutation rates in the range of 10^{-5} to 10^{-4} mutations per generation per locus have been usually assumed in the literature, as well as generation times for trees between 25 and 100 years (Navascués *et al.,* 2009; Heuertz *et al.,* 2010). This means that the oldest expansion could have occurred in the SMS between 1.6×10^6 (low mutation rate, long generation time) and 39,964 years ago (high mutation rate, short generation time). In the same way, the most recent expansion in the CP would be dated between 190,714 and 4768 years ago. In general, a significant negative correlation between τ and latitude was found (r = -0.34; P = 0.039), indicating a tendency of more northern populations to have more recent population expansion times.

LGM and present potential distribution models

The AUC values (averaged over the ten runs) for the training and test data in the current, MIROC, and CCSM models were 0.969/0.941, 0.969/0.935 and 0.968/0.935, respectively.

These values indicate a good performance for the three models. The jackknife analysis indicated that the climatic variables with the highest relative contributions to the present distribution model were temperature seasonality, maximum temperature of the warmest month, isothermality and mean temperature of the warmest quarter. On the other hand, temperature seasonality, maximum temperature of the warmest month and mean temperature of the warmest quarter were the more important variables under the MIROC and CCSM models.

The comparison of the distribution extent of *Q. castanea* among the present and the two LGM models indicate that the current distribution represents, on average, 67.3% or 49.9% of the area occupied by the species during the LGM according to the MIROC and CCSM models, respectively (Fig. 3). There were some differences between both LGM models on the predicted distribution extent of *Q. castanea* at specific geographic regions. It is probable that *Quercus castanea* experienced a range contraction after the end of the glaciations, but virtually within the same regions without experiencing important shifts in distribution. Interestingly, the models also suggest long-term fragmentation and isolation of some areas. For example, all three models indicate that the suitable areas for *Q. castanea* situated in the SMO and CP are not continuous with areas in the TMVB. The present and CCSM models also coincided at pointing out other important discontinuities in distribution, such as between the westernmost part and the rest of the TMVB, the SMS and the TMVB, and the north and the south of the SMS (Fig. 3). The models using two random independent dataset partitions showed similar results, with some changes in the probability distribution but with similar distribution total extent and similar barriers between presence areas (Figure 1S).

DISCUSSION

In the present study, we analyzed variation at seven cpSSR loci in 36 populations from across the whole distribution area of *Q. castanea,* one of the most widespread red oak species in Mexico. A total of 90 haplotypes were identified, 28 of which were exclusive to single populations. Average within population diversity and total gene diversity were high (h_S = 0.730; h_T = 0.989) and population differentiation was moderate when distances among haplotypes are not taken into account (G_{ST} = 0.261), but high when distances are considered (R_{ST} = 0.711). Using sample sizes comparable to those in this study, Grivet *et al.* (2006) analyzed populations of the white oak *Q. lobata* in California, and five species of the European white oak complex, with six cpSSR loci (three in common with this study) and reported considerably lower haplotype richness (11 and 39 haplotypes in total for Europe and California, respectively) and h_S values (0.285 and 0.114, respectively), although similar h_T (0.755 and 0.979). The genetic diversity was much more structured (with unordered alleles) in the European and Californian oaks (G_{ST} = 0.849 and 0.709) than in *Q. castanea*, but with ordered alleles differentiation estimates were comparable (R_{ST} = 0.838 and 0.824). The higher haplotype richness found in this study could be explained in part by the inclusion of an additional cpSSR locus. However, deleting one locus at a time from our database resulted in haplotype numbers between 56 and 78. Also, there could be some intrinsic differences in variability among the loci employed in the two studies. Nevertheless, the three loci in common (*μdt1*, *μdt3* and *μcd5*) had 4-5 alleles in *Q. castanea* compared to 2-4 in California and Europe. Furthermore, the studies of Magri *et al.* (2006) and Marsico & Hellman (2009), which employed eight and five cpSSR loci to analyze the phylogeography of *Q. suber* and *Q. garryana*, respectively, reported only five and six haplotypes.

The results thus indicate that *Q. castanea* actually has considerably higher haplotype richness and intra-population genetic diversity than other oak species so far studied at more northern latitudes. For this diversity to arise and be maintained, large effective population sizes are required for a long period of time. It does seem like the most recent glacial and interglacial cycle of the Pleistocene had a moderate effect on the demography and distribution of *Q. castanea* populations, which probably persisted in situ without dramatic size fluctuations. The cpSSR data showed signatures of demographic expansions in almost all populations, according to Fu's F_S and Harpending's raggedness index of mismatch distributions (Table 4). Estimations of the time to population expansions using τ suggested that these could be as old as 1.6×10^6 years before present in the SMS. After this expansion, the species could have colonized and expanded in the TMVB and the SMO at maximum times of about 6×10^5 years ago and finally in the CP about 1.9×10^5 years ago. Nevertheless, some alternative processes other than population expansion could explain negative and significant F_S values. For example, genetic hitchhiking can produce the same patterns than population expansion on gene genealogies. In genetic hitchhiking, the adaptive fixation of strongly advantageous mutations decreases the level of variation (Innan & Stephan, 2000), because such fixations sweep out neutral polymorphism in the surrounding region whereas some variants hitchhike with the favored mutations (Maynard-Smith & Haight, 1974). Under population expansion, a similar pattern of polymorphism results from the association of one common haplotype with other low-frequency or private haplotypes (Rogers & Harpending 1992; Harpending *et al*. 1993). The use of other tests such as the raggedness index in combination with F_S can further support the demographic expansion hypothesis over the selection hypothesis (Ramos-Onsins & Rozas, 2002). In our case, another reason to favor the demographic expansion hypothesis is that very few studies have found evidence of selection acting on cpDNA (e. g. Wright & Gaut, 2005).

The MAXENT distribution models suggest that *Q. castanea* populations had a larger distribution extent during the LGM with a range contraction of ~30 to 50% between the LGM and in its current distribution. However, the areas occupied by the species during the LGM were practically the same as in the present distribution but with different changes depending on the biogeographic region. It is plausible that in response to climate variation, populations of this species mostly experienced altitudinal displacements and a series of expansions and contractions, rather than large latitudinal migrations or colonizations and, as a consequence, high genetic variation levels and local haplotypes were retained.

Interestingly, the few previous studies conducted on wide-ranging Mexican montane temperate trees have also revealed high levels of intra-population and total genetic variation, presence of phylogeographic structure and complex patterns of geographic distribution of haplotypes, indicating that these species have long-term population histories in the areas currently occupied by them (González-Rodríguez *et al.,* 2004; Tovar-Sánchez *et al.,* 2008; Moreno-Letelier & Piñero, 2009; Rodríguez-Banderas *et al.,* 2009). Downslope movements and expansion of populations during glacial periods and fragmentation into isolated higher altitude stands during warmer interglacials seems to have been the predominant response of these species (Jaramillo-Correa *et al.*, 2009).

Populations of *Q. castanea* situated in the four biogeographic provinces (SMO, TMVB, CP and SMS) were significantly differentiated according to the AMOVA analyses, particularly under the SMM (R_{ST}), indicating the accumulation of distinct mutations within the regions (Table 3). Pairwise differentiation was the highest between the SMO and the other three regions, while the lowest was between the TMVB and the SMS. However, the genetic barriers identified through the Monmonier's analysis were not completely concordant with the limits between biogeographic provinces (Fig. 1), except for the strongest barrier (B1), which isolated populations in the SMO from the rest, supporting the

distinctness and long-term isolation of the populations in this region. Interestingly, at the SMO the LGM distribution models suggested a similar suitable area for *Q. castanea* in comparison to the model of current distribution. Moreover, all models (e. g. present and LGM) showed a disjunction in the areas suitable for *Q. castanea* between the southern SMO and the northwestern TMVB. The long-term persistence of this separation might explain this phylogeographic break.

The placement of the other four barriers suggested some gene flow among regions and genetic discontinuities within biogeographic regions. Gene flow among regions is also apparent from the haplotype network. Recent, as well as more ancient gene exchanges are clear between the TMVB, the SMS and the CP, since the three regions share some haplotypes or closely related haplotypes from several lineages from across the genealogical tree. In contrast, populations of the southeastern part of TMVB separated by the second barrier (B2) have more haplotypes in common with populations in the northern SMS than with other populations in the TMVB. Similarly, population 23, which is situated in the CP, is more similar to populations in the TMVB, as recognized by the third barrier (B3). The fourth (B4) and fifth (B5) barriers indicated some limitations to gene flow between westernmost and central portions of the TMVB and between the northern and southern portions of the SMS. These patterns could have arisen due to geologic and climatic events causing recurrent expansions and fragmentation of areas suitable for *Q. castanea*. Particularly, the TMVB has a complex history that has probably played an important role in shaping population genetic structure of many species. Geological data indicate that the western part of this mountain chain started to form about 23 million years ago (Ma), while the eastern part formed 2.5 Ma (Ferrusquía-Villafranca, 1993), implying very different characteristics and historical processes between the two parts. In fact, biogeographic analyses have suggested that the TMVB is not a homogeneous unit (e. g. Corona *et al.,*

2007). Also, palynological records indicate drier environmental conditions in the eastern part of the TMBV compared with the western part at the end of the Pleistocene (Bradbury, 1997; Lozano-García & Ortega-Guerrero, 1997). Congruently, our LGM distribution models showed a greater suitability for *Q. castanea* in the western part of the TMBV. Thus, the presence of some phylogeographic breaks among *Q. castanea* populations within this region is not completely surprising.

Genetic exchange between regions can be explained by similar factors. For example, the SMS probably had historical connections with the eastern side of the TMBV (Corona *et al.,* 2007), which could have facilitated migration of ancient population lineages between the two regions. More recently, in the late Pleistocene, palynological records from Ixtacyola and Ixtapa cores in the SMS indicated cold-humid conditions (McNeish *et al*., 1972; Piperno *et al*., 2007) associated to expansions of temperate forests to lowlands forming large continuous stands (Martínez-Hernández, 1992; Lozano-García & Ortega-Guerrero, 1997), which is also suggested by our distribution models of *Q. castanea*. At the end of the Pleistocene the tropical forest corridor expanded (e.g. Cuenca del Balsas) and temperate forests contracted to higher altitudes in the mountains (Metcalfe, 2006). These repetitive periods of altitudinal expansions and contractions presumably resulted in periods of differentiation followed by contact between the pine-oak forests from the south of the TMVB and the north of the SMS. This scenario is congruent with the clear genetic affinities of *Q. castanea* populations between these two areas. In the CP, a similar history of altitudinal migration, expansion and contraction of temperate forests might have occurred and is suggested by our distribution models and by palynological records that indicate abundance of temperate forests in valleys and lowlands during the LGM (Martínez-Hernández, 1992; Metcalfe, 2006). These cycles could also have caused some

genetic exchanges between the CP and the TMBV as well as some differentiation within the CP as suggested by barrier B3.

A weak coincidence between the genetic discontinuities among populations of temperate tree taxa and the traditionally delimited Mexican biogeographic zones has been previously noted (Jaramillo-Correa *et al.,* 2009). However, it is interesting that some of the phylogeographic breaks occurring in other tree species seem to coincide with those found in *Q. castanea*. For example, in *Pinus strobiformis*, a phylogeographic break was found between the populations of the SMO and the populations in the TMVB and the Sierra Madre Oriental (Moreno-Letelier & Piñero, 2009) and in the *P. montezumae-P. hartweggii* complex, a subdivision occurs between populations in the eastern and western parts of the TMVB (Matos & Schaal, 2000).

In conclusion, the results presented in this study suggest that some of the genetic patterns found in *Q. castanea* could be related to complex geological and climatic events, encompassing at least the whole Pleistocene, which caused cyclic expansions and fragmentation of suitable areas for this species. Moreover, similar patterns already reported for other tree species might indicate similar population histories in several temperate elements with similar ecological requirements as *Q. castanea*.

The origin and maintenance of a large proportion of the genetic diversity in *Q. castanea* has occurred within the TMVB. This region is also a hotspot for oak species diversity (Nixon, 1993) and has acted as a natural bridge between other biogeographic provinces. Moreover, some studies have shown that active hybridization and introgression processes between red oak species have taken place in this area (González-Rodríguez *et al.,* 2004; Tovar-Sánchez *et al.,* 2008), which has constituted a centre for active evolution of the Mexican biota.

ACKNOWLEDGEMENTS

The authors thank to J.J. Morrone, J. Núñez-Farfán, D. Lugo-Aquino, A. Albarrán-Lara, E. Pascual, L. Herrera, O. Chassin-Noria, N. Pérez-Nasser, J.J. Fuentes-Junco and S. Valencia for technical assistance and suggestions to improve the manuscript. This paper constitutes a partial fulfilment of the Graduate Program in Biological Sciences of the National Autonomous University of México (UNAM). J.M. Peñaloza-Ramírez received a Ph.D. scholarship provided by the National Council of Science and Technology (CONACyT), and UNAM. This project was supported by CONACyT grant 38550-V to K.O., CONACyT-SEMARNAT 2004-01-97 and 23728 to K.O.

REFERENCES

Albarrán-Lara, A.L., Mendoza-Cuenca, L., Valencia-Avalos, S., González-Rodríguez, A., & Oyama, K. (2010) Leaf fluctuating asymmetry increases with hybridization and introgression between *Quercus magnoliifolia* and *Quercus resinosa* (Fagaceae) through an altitudinal gradient in Mexico. *International Journal of Plant Sciences*, **171**, 310-322.

Afzal-Rafii, Z., & Dodd, R.S. (2007) Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (*Pinus nigra*) in Western Europe. Molecular Ecology, 16, 723-736

Avise, J.C. (2000) *Phylogeography: The history and formation of species*. Harvard University Press, London.

Bandelt, H.J., Forster, P., Sykes, B.C. & Richards, M.B. (1995) Mitochondrial portraits of human populations. *Genetics*, **141**, 743-753.

Belahbib, N., Pemonge, M.H., Quassou, A., Sbay, H. & Petit, R.J. (2001) Frequent cytoplasmic exchanges between oak species that are not closely related: *Quercus suber* and *Q. ilex* in Morocco. *Molecular Ecology*, **10**, 2003-2012.

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J.Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, T., Hewitt, C.D., Kageyama, M., Kitoh, A., Loutre, M.F., Marti, O., Merkel, U., Ramstein, G., Valdes, P., Weber, L., Yu, Y. & Zhao, Y. (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. *Climate Past*, **3**, 279-296.

Bradbury, J.P. (1997) Sources of glacial moisture in Mesoamerica. *Quaternary International*, **44**, 97–110.

Collins, W.D., Blackmon, M., Bitz, C., Bonan, G. & Bretherton, C.S. (2004) The community climate system model: CCSM3. *Journal of Climate*, **19**, 2122-2143.

Corona, A.M., Toledo, V.H. & Morrone, J. J. (2007) Does the Trans-Mexican Volcanic Belt represent a natural biogeographical unit? An analysis of the distributional patterns of Coleoptera. *Journal of Biogeography*, **34**, 1008-1015.

Deguilloux, M.F., Dumolin-Lapegue, S., Gielly, L., Grivet, D. & Petit, R.J. (2003) A set of primers for the amplification of chloroplast microsatellites in *Quercus*. *Molecular Ecology Notes*, **3**, 24-27.

Dumolin-Lapegue, S., Demesure, B., Fineschi, S., Le Corre, V. & Petit, R.J. (1997) Phylogeographic structure of white oaks throughout the European continent. *Genetics*, **146**, 1475-1487.

Ennos, R.A., Sinclair, W.T., Hu, X.S. & Langdon, A. (1999) Using organelle markers to elucidate the history, ecology and evolution of plant populations. *Molecular Systematics*

and plant evolution (ed. by P.M. Hollingsworth, R.M. Bateman and R.J. Gornall), pp. 1-19. Taylor & Francis Ltd, London.

Excoffier, L., Laval, G. & Schneider, S. (2005) Arlequin ver 3.5. An Integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47-60.

Ferrusquia-Villafranca, I. (1993) Geology of Mexico: a synopsis. *Biological diversity of Mexico: origins and distribution.* (ed. by T.P. Ramamoorthy, R. Bye, A. Lot and J. Fa), pp. 3-107. Oxford University Press, New York.

Fielding, A.H., & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Fu, Y.X. (1997) Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics*, **147**, 915-925.

Goldstein, D.B., Ruiz-Linares, A., Cavalli-Sforza, L.L. & Feldman, M.W. (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics*, **139**, 463-471.

González-Rodríguez, A., Bain, J.F., Golden, J.L. & Oyama, K. (2004) Chloroplast DNA variation in the *Quercus affinis-Q. laurina* complex in Mexico: Geographical structure and associations with nuclear an morphological variation. *Molecular Ecology*, **13**, 3467-3476.

Grivet, D., Deguilloux, M.F., Petit, R.J. & Sork, V.L. (2006) Contrasting patterns of historical colonization in white oaks (*Quercus* spp.) in California and Europe. *Molecular Ecology*, **15**, 4085-4093.

Harpending, H.C., Sherry, S.T., Rogers, A.R., & Stoneking, M. (1993) The genetic structure of ancient human populations. *Current Anthropology*, **34**, 483-496.

Harpending, H.C. (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591-600.

Hasumi, H. & Emori, S. (2004) *K-1 coupled GCM (MIROC) description.* Tokyo: Center for Climate System Research, University of Tokyo, Japan.

Heuertz, M., Teufel, J., González-Martínez, S.C., Soto, A., Fady, B., Alía, R. & Vendramin, G.G. (2010) Geography determines genetic relationships between species of mountain pine (*Pinus mugo* complex) in Western Europe. *Journal of Biogeography*, **37**, 541-556.

Hijmans, R.J. & Graham, C.H. (2006) The ability of climatic envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 2272-2281.

Innan, H., & Stephan, W. (2000) The coalescent in an exponentially growing metapopulation and its application to Arabidopsis thaliana. *Genetics*, **155**, 2015–2019.

Jaramillo-Correa, J.P., Beaulieu, J., Khasa, D.P. & Bousquet, J. (2009) Inferring the past from 12 the present phylogeographic structure of North American forest trees: seeing the forest for the 13 genes. *Canadian Journal of Forest Research*, **39**, 286-307.

Kappelle, M., (2006) *Ecology and Conservation of Neotropical Montane Oak Forests*. Ecological Studies Series, Vol. 185. Springer Verlag. Berlin – Heidelberg – New York. 483 pp.

Lachniet, M.S. & Vázquez-Selem, L. (2005) Last Glacial Maximum equilibrium line altitudes in the circum-Caribbean (Mexico, Guatemala, Costa Rica, Colombia, and Venezuela). *Quaternary International*, **138**, 129-144.

Le Corre, V. & Kremer, A. (1998) Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and steppingstone model. *Journal of Evolutionary Biology*, **11**, 495-512.

Lefort, F. & Douglas, G.C. (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-Garcia, M.S., & Ortega-Guerrero, B. (1997) Late Quaternary environmental changes of the central part of the Basin of Mexico; correlation between Texcoco and Chalco Basins. *Review of Palaeobotany and Palynology*, **99**, 77-93.

Lozano-García, S., Sosa-Nájera, S. & Sugiera, Y. (2005) 23, 000 yr of vegetation history of the Upper Lerma, a tropical high-altitude basin in Central Mexico. *Quaternary Research*, **64**, 70-82.

McNeish, R.S., Peterson, F.A. & Neely, J.A. (1972) The archaeological reconnaissance. *The prehistory of the Tehuacan Valley. Vol. 5. Excavations and reconnaissance* (ed. by R.S. McNeish), pp. 496–504. University of Texas Press, Austin, TX.

Magni, C.R., Ducousso, A., Caron, H., Petit, R.J. & Kremer, A. (2005) Chloroplast DNA variation of *Quercus rubra* L. in North America and comparison with other Fagaceae. *Molecular Ecology*, **14**, 513-524.

Magri, D., Vendramin, G.G. & Comps B. (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, **171**, 199-221.

Manni, F., Guerard, E. & Heyer, E. (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by ‘‘Monmonier’s algorithm’’. *Human Biology*, **76**, 173-190.

Marsico, T.D. & Hellmann, J.J. (2009) Patterns of seed dispersal and pollen flow in *Quercus garryana* (Fagaceae) following post-glacial climatic changes. *Journal of Biogeography*, **36**, 929-941.

Martínez-Hernández, E. (1992) La vegetación y el clima de México. Historia del Cretácico tardío al Terciario basada en evidencias palinológicas. *Ciencia*, **43**, 69-72.

Matos, J.A. & Schaal, B.A. (2000) Chloroplast evolution in the *Pinus montezumae* complex: a coalescent approach to hybridization. *Evolution*, **54**, 1218–1233.

Maynard-Smith, J., & Haigh, J. (1974) The hitch-hiking effect of a favourable gene. *Genetic Resources*, **23**, 23–35.

Metcalfe, S.E. (2006) Late Quaternary environments of the northern deserts and central transvolcanic belt of Mexico. *Annals of the Missouri Botanical Garden*, **93**, 258-273.

Monmonier, M. (1973) Maximum–difference barriers: an alternative numerical regionalization method. *Geographical Analysis*, **3**, 245-261.

Moreno-Letelier, A. & Piñero, D. (2009) Phylogeographic structure of *Pinus strobiformis* Engelm. across the Chihuahuan Desert filter-barrier. *Journal of Biogeography*, **36**:121-131.

Morrone, J.J. (2005) Hacia una síntesis biogeográfica de México. *Revista Mexicana de Biodiversidad*, **76**, 207-252.

Navascués, M., & Emerson, C.C. (2006) Chloroplast microsatellites: measures of genetic diversity and the effect of homoplasy. *Molecular Ecology*, **14**, 1333-1341.

Navascués, M., Hardy, O.J. & Burgarella, C. (2009) Characterization of demographic expansions from pairwise comparisons of linked microsatellite haplotypes. *Genetics*, **181**, 1013-1019.

Nixon, K.C. (1993) The genus *Quercus* in Mexico. *Biological diversity of Mexico: origins and distribution* (ed. by T.P, Ramamoorthy R. Bye, A. Lot and J. Fa), pp. 447-458 Oxford University Press. New York, USA.

Petit, R.J., El Mousadik, A. & Pons, O. (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844-855.

Petit, R.J., Csaikl, U.M., Bordacs, S., Burg, K., Coart, E., Cottrell, J., Van Dam, B., Deans, J.D., Dumolin-Lapegue, S., Fineschi, S., Finkeldey, R., Gillies, A., Glaz, I., Goicoechea, P.G., Jensen, J.S., Konig, A.O., Lowe, A.L., Madsen, S.F., Matyas, G., Munro, R.C., Olalde, M., Pemonge, M.H., Popescu, F., Slade, D., Tabbener, H., Taurchini, D., de Vries, S.G.M., Ziegenhagen, B. & Kremer, A. (2002) 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.

Phillips**,** S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.

Piperno, D.R., Moreno, J.E., Iriarte, J., Holst, I., Lachniet, M., Jones, J.G., Ranere A.J. & Castanzo, R. (2007) Late Pleistocene and Holocene environmental history of the Iguala Valley, Central Balsas Watershed of Mexico. *Proceedings of the National Academy of Sciences*, **104**, 11874–11881.

Pons, O. & Petit, R.J. (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, **144**, 1237-1245.

Ramamoorthy, T.P., Bye, R., Lot, A., & Fa, J. (1993) *Biological diversity of Mexico: origins and distribution*. Oxford University Press.

Ramos-Onsins, S.E., & Rozas, J. (2002) Statistical properties of neutrality test against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

Richards, C.L., Carstens, B.C. & Knowles, L.L. (2007) Distribution modeling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographic hypotheses. *Journal of Biogeography*, **34**, 1833-1845.

Riddle, B.R., Hafner, D.J., Alexander, L.F. & Jaeger, J.R (2000) Cryptic vicariance in the historical assembly of a Baja California Peninsular Desert biota. *Proceedings of the National Academy of Sciences USA*, **97**, 14438-14443.

Rodríguez-Banderas, A., Vargas-Mendoza, C.F., Buonamici, A. & Vendramin, G.G. (2009) Genetic diversity and phylogeographic analysis of *Pinus leiophylla*: a post-glacial range expansion. *Journal of Biogeography*, **36**, 1807-1820.

Rogers, A.R. & Harpending, H. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552-569.

Rzedowski, J. (1978) *Vegetación de México*. CONABIO, Editorial Limusa. México.

Rzedowski, J., & Rzedowski, G.C., (1989) Sinopsis numérica de la flora Fanerogámica del Valle de México. *Acta Botánica Mexicana*, **8**, 15-30.

Schneider, C.J., Cunningham, M. & Moritz, C. (1998) Comparative phylogeography and the history of endemic vertebrates in the Wet Tropics rainforest of Australia. *Molecular Ecology*, **7**, 487-498.

Sebastiani, F., Carnevale, S. & Vendramin, G.G. (2004) A new set of mono- and dinucleotide chloroplast microsatellites in Fagaceae. *Molecular Ecology Notes*, **4**, 259-261.

Slatkin, M. (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457-462.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.G. & Cosson, J.F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453-464.

Tovar-Sánchez, E., Mussali-Galante, P., Esteban-Jiménez, R., Piñero, D., Arias, D.M., Dorado, O. & Oyama, K. (2008) Chloroplast DNA polymorphism reveals geographic

structure and introgression in *Quercus crassifolia* x *Q. crassipes* complex in México. *Botany,* **86**, 228-239.

Valencia, A.S. (1995) *Contribución al conocimiento del género Quercus (Fagaceae) en el estado de Guerrero, México*. Coordinación de Servicios Editoriales, Facultad de Ciencias, UNAM.

Valencia, A.S. (2004) Diversidad del género *Quercus* (Fagaceae) en México. *Boletín de la Sociedad Botánica de México*, **75**, 33-53.

Waltari, E., Hijmans, R.J., Peterson, A.T., Nyari, Á.S., Perkins, S.L. & Guralnick, R.P. (2007) Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *Plos ONE,* **7**, e563-11.

Wright, S.I., & Gaut, B.S. (2005) Molecular population genetics and the search for adaptive evolution in plants. *Molecular Biology and Evolution*, **22**:506–519.

BIOSKETCHES

Juan Manuel Peñaloza-Ramírez is interested in population genetics, molecular evolution and phylogeography with particular focus on historical demography and ecological niche modelling.

Table 1 Geographical location and estimates of genetic diversity for 36 populations of *Q. castanea* in four biogeographic provinces of Mexico. *n*, sample size; h_S, within population genetic diversity; D^2sh, mean pairwise genetic distance among individuals within a population under a stepwise mutation model. Allelic richness is reported after a rarefaction analysis to standardize for unequal sample sizes.

No	Population	n	Latitude/longitude	Altitude (m)	Total haplotypes/ exclusive haplotypes	Allelic richness	h_S	D^2sh
1	M. Escob	10	22°20'/103°36'	1732	4/3	2.43	0.78 (0.091)	0.11
2	Teul de Gzlez.	10	21°24'/103°30'	1916	4/3	2.33	0.73 (0.120)	0.07
	SMO (pooled)	**20**			**7/ 7**	**7.00**	**0.87 (0.038)**	**0.24**
3	Ceboruco	10	21°6'/104°30'	1065	5/3	2.67	0.76 (0.130)	0.05
4	Compostela	10	21°18'/104°54'	1192	5/3	3.00	0.84 (0.080)	0.04
5	Cuarenteño	10	21°29'/104°54'	1279	4/2	1.80	0.53 (0.180)	0.05
6	Tepatitlan	10	20°51'/102°37'	1889	4/2	2.17	0.71 (0.117)	0.14
7	Nevado Colima	10	19°37'/103°33'	2487	4/0	2.07	0.64 (0.152)	0.22
8	Tapalpa	10	19°56'/103°40'	2200	3/0	1.20	0.38 (0.181)	0.37
9	U. Guadalupe	10	19°51'/103°27'	1910	5/3	2.67	0.76 (0.130)	0.20
10	Zapotlanejo	10	20°29'/103°2'	2273	5/1	3.00	0.84 (0.080)	0.04
11	Mezcala	10	20°51'/102°47'	2115	4/2	2.33	0.73 (0.120)	0.12
12	Tequila	10	20°47'/103°50'	1912	4/1	2.07	0.64 (0.152)	0.12
13	Zitacuaro	9	19°26'/100°19'	2530	3/0	1.50	0.46 (0.200)	0.13
14	CD Hidalgo	10	19°42'/100°34'	2080	5/2	2.60	0.80 (0.089)	0.10
15	Zinapecuaro	10	19°48'/100°48'	1970	5/2	3.17	0.87 (0.071)	0.06
16	Temascaltepec	10	19°3'/100°3'	1907	6/4	3.50	0.89 (0.075)	0.25
17	Coatepec	10	18°54'/99°46'	2150	5/3	2.76	0.80 (0.100)	0.19
18	Santiago Laguna	10	19°15'/100°7'	1908	6/2	3.17	0.87 (0.071)	0.44
19	Amanalco	10	19°13'/100°7'	1901	4/1	2.07	0.64 (0.152)	0.12
20	Tepoztlan	8	19°0'/99°8'	2010	2/0	0.96	0.43 (0.169)	0.30
21	Sta Maria	10	18°59'/99°15'	2189	5/2	2.67	0.76 (0.130)	0.13
22	Cuernavaca	9	18°59'/99°12'	1980	3/0	1.91	0.72 (0.097)	0.40
	TMBV (pooled)	**196**			**58/43**	**14.97**	**0.97 (0.003)**	**0.006**
23	Cuquio	10	21°0'/103°0'	1907	5/2	2.76	0.80 (0.100)	0.10
24	Lima	10	21°4'/101°12'	2548	4/0	2.33	0.73 (0.120)	0.68
25	C. Medio	10	21°4'/101°11'	2150	6/2	3.26	0.84 (0.103)	0.17
26	Calvillo	10	21°5'/101°9'	2452	4/0	2.19	0.73 (0.101)	0.61
	CP (pooled)	**40**			**14/6**	**9.45**	**0.91 (0.020)**	**0.06**
27	Ixcateopan	10	18°32'/99°42'	2190	3/1	1.73	0.62 (0.138)	0.68
28	Chichila	10	18°33'/99°41'	2001	4/0	2.33	0.73 (0.120)	0.13
29	Atlixtac	10	17°32'/98°54'	2028	7/3	4.00	0.93 (0.062)	0.08
30	Tlaquiltepec	8	17°32'/98°51'	1890	5/3	3.21	0.86 (0.108)	0.12
31	Chilapa	6	17°34'/99°3'	1871	3/0	2.00	0.73 (0.155)	0.44
32	Juxtlahuaca	8	17°16'/98°0'	1960	4/0	2.71	0.81 (0.130)	0.07
33	Sola De Vega	6	16°34'/96°56'	1845	3/2	3.00	0.70 (0.218)	0.16
34	El Punto	9	17°8'/96°37	2430	3/0	2.32	0.68 (0.122)	0.18
35	Laguna	8	17°10'/97°53'	2485	4/2	2.71	0.81 (0.130)	0.42
36	B. Mixtepec	10	16°50'/96°54'	1748	4/4	1.80	0.53 (0.180)	0.38
	SMS (pooled)	**85**			**28/19**	**12.53**	**0.95 (0.009)**	**0.01**

Table 2 Hierarchical analysis of molecular variance (AMOVA) using F_{ST} and R_{ST} for 36 populations of *Q. castanea* grouped according to biogeographic provinces. Asterisks indicate statistically significant values ($P < 0.001$). Tests were based on 10 000 random permutations.

	Source of variation	d.f.	SS	Variance components	Percentage of variation	Fixation index
$\boldsymbol{F_{ST}}$						
	Among groups	3	79.391	0.213 Va	10.16	Φ_{CT} =0.101***
	Among populations within groups	32	346.69	1.05 Vb	49.92	Φ_{SC} =0.555***
	Within populations	307	258	0.840 Vc	39.92	Φ_{ST} =0.600***
	Total	342	681.08	2.105		
$\boldsymbol{R_{ST}}$						
	Among groups	3	1092.85	5.013 Va	53.63	Φ_{CT} = 0.526***
	Among populations within groups	32	816.17	2.467 Vb	25.91	Φ_{SC} =0.546***
	Within populations	307	627.60	2.044 Vc	21.46	Φ_{ST} =0.785***
	Total	342	2536.6	9.524		

Table 3 Pairwise values of genetic differentiation of *Q. castanea* among the four biogeographic provinces using F_{ST} (below diagonal) and R_{ST} (above diagonal). All values were highly significant ($P < 0.0001$) according to 10 000 random permutations.

$F_{ST}\backslash R_{ST}$	SMO	TMBV	CP	SMS
SMO	-	0.865	0.772	0.862
TMBV	0.268	-	0.214	0.092
CP	0.369	0.160	-	0.328
SMS	0.221	0.105	0.212	-

Table 4 Tests for historical population expansions and estimation of demographic parameters for 36 populations of *Q. castanea* in four biogeographic provinces of Mexico. Θ_0 and Θ_1 are, respectively, ancestral and current population sizes scaled by mutation rate; τ is the number of generations since the expansion occurred, scaled by mutation rate; -log [CL] is the likelihood of the model. See text for details.

No	Region	Population	Fu´s *Fs*	*P*-value	Raggedness index	*P*-value	Population growth parameters (with homoplasy)			
							Θ_0	Θ_1	τ	-log[CL]
1	SMO	M. Escob	-11.60	0.000	0.139	0.380	1.747	4.04×10^{12}	0.384	31.628
2		Teul de Gzlez	-13.4	0.000	0.400	0.060	1.710	1.12×10^{13}	1.393	33.14
3	TMBV	Ceboruco	-12.12	0.000	0.093	0.710	0.925	386245	0.868	15.54
4		Compostela	-12.50	0.000	0.102	0.500	5.8×10^{-5}	3.63×10^{13}	1.636	15.06
5		Cuarenteño	-15.42	0.000	0.125	1.000	0.992	5.37×10^{11}	0.088	17.83
6		Tepatitlan	-10.06	0.000	0.143	0.630	2.268	1.46×10^{11}	0.072	26.87
7		Nevado Colima	-9.41	0.000	0.170	0.580	3.408	0.001066	0.00018	30.78
8		Tapalpa	-10.70	0.000	0.530	0.390	6.740	0.00150	0.0012	30.74
9		U. Guadalupe	-8.62	0.000	0.115	0.980	3.904	37350	1.43×10^{-6}	37.96
10		Zapotlanejo	-13.20	0.000	0.186	0.180	1.94×10^{-5}	3.04×10^{13}	1.445	21.82
11		Mezcala	-11.20	0.000	0.034	0.990	3.53×10^{-5}	3.119	3.777	30.10
12		Tequila	-10.31	0.000	0.170	0.500	2.190	11290	5.9×10^{-7}	24.64
13		Zitacuaro	-4.03	0.000	0.121	0.760	2.070	0.0068	0.001	41.41
14		CD Hidalgo	-11.28	0.000	0.059	0.820	4.78×10^{-6}	5.48×10^{9}	1.778	1.0×10^{8}
15		Zinapecuaro	-11.44	0.000	0.097	0.560	6.54×10^{-5}	4.37×10^{13}	1.980	20.14
16		Temascaltepec	-7.45	0.000	0.093	0.680	5.230	1.44×10^{12}	0.215	19.48
17		Coatepec	-8.79	0.001	0.155	0.440	3.910	1.51×10^{11}	0.034	20.68
18		S Laguna	-7.42	0.000	0.123	0.480	5.982	2.59×10^{11}	0.058	35.20
19		Amanalco	-12.60	0.000	0.155	0.570	1.57×10^{-5}	2.85	4.381	28.45
20		Tepoztlan	-9.31	0.000	0.693	0.260	2.880	0.0026	0.0011	74.48
21		Sta Maria	-9.95	0.000	0.069	0.810	2.434	2.75×10^{11}	0.168	23.95
22		Cuernavaca	-4.64	0.000	0.138	1.000	3.210	11338	5.05×10^{-7}	52.63
23	CP	Cuquio	-10.77	0.000	0.152	0.910	1.929	2.27×10^{12}	0.474	21.21
24		Lima	-6.67	0.000	0.164	0.430	16.69	0.0427	0.011	50.32
25		C. Medio	-8.23	0.000	0.108	0.580	5.161	1.55×10^{11}	0.046	17.92
26		Calvillo	-6.55	0.000	0.238	0.120	5.090	3.65	0.536	22.67
27	SMS	Ixcateopan	-7.97	0.001	0.194	0.540	11.160	0.986	0.821	53.84
28		Chichila	-10.57	0.000	0.186	0.230	8.39×10^{-5}	5.001	3.312	28.38
29		Atlixtac	-9.46	0.000	0.064	0.670	7.25×10^{-5}	3.11×10^{13}	2.924	16.37
30		Tlaquiltepec	-5.09	0.000	0.089	0.630	2.60×10^{-5}	12.35	5.274	17.18
31		Chilapa	-3.07	0.011	0.293	0.420	8.430	0.680	0.568	50.15
32		Juxtlahuaca	-2.56	0.062	0.064	0.750	6.13×10^{-5}	8.06	2.848	21.41
33		Sola De Vega	-0.80	0.266	0.244	0.240	2.77×10^{-5}	2.629	3.063	42.34
34		El Punto	-3.85	0.016	0.150	0.420	2.90×10^{-16}	4.959	2.912	48.44
35		Laguna	-4.23	0.004	0.473	0.960	5.254	36827	7.28×10^{-7}	38.46
36		B. Mixtepec	-8.62	0.000	0.219	0.980	10.13	0.798	0.654	28.38

Figure legends

Figure 1 Haplotype frequencies in 36 sampled populations of *Q. castanea* in four morphotectonic and biogeographic provinces of Mexico: Sierra Madre Occidental (SMO: populations 1-2), Central Plateau (CP: 3-22), Trans-Mexican Volcanic Belt (TMBV: 23-26) and Sierra Madre del Sur (SMS: 27-36). All 90 haplotypes are represented by different combinations of colours and patterns. The location of the five main genetic barriers among populations is indicated by black lines labeled B1-B5 in order of importance. Numbers in each circle indicate the populations (see Table 1 for details).

Figure 2 Median-joining haplotype network for the 90 haplotypes identified in 36 populations of *Q. castanea* in Mexico. Each circle represents an individual haplotype and circle size is proportional to the frequency of the haplotype. Different colors indicate the presence of the haplotypes in the four biogeographic regions: in yellow haplotypes present in the SMO, in blue haplotypes present in the TMBV, in brown haplotypes present in the CP, and in orange haplotypes present in the SMS.

Figure 3 Maps showing potential distribution as probability of occurrence (green = 0 - 0.2, red = 0.8 - 1.0) for *Quercus castanea* in Mexico (a) under present climatic conditions, (b) at the Last Glacial Maximum (LGM; 21 ka) based on the CCSM model, and (c) at the LGM based on the MIROC 3.2 model.

Figure S1 Ecological niche models validation for *Quercus castanea* in Mexico, projecting current climate data (a, b) into two past (~21 *ka* BP) climate scenarios: CCSM (c, d) and MIROC 3.2 (e, f) models, using two independent dataset partitions.

Figure 1

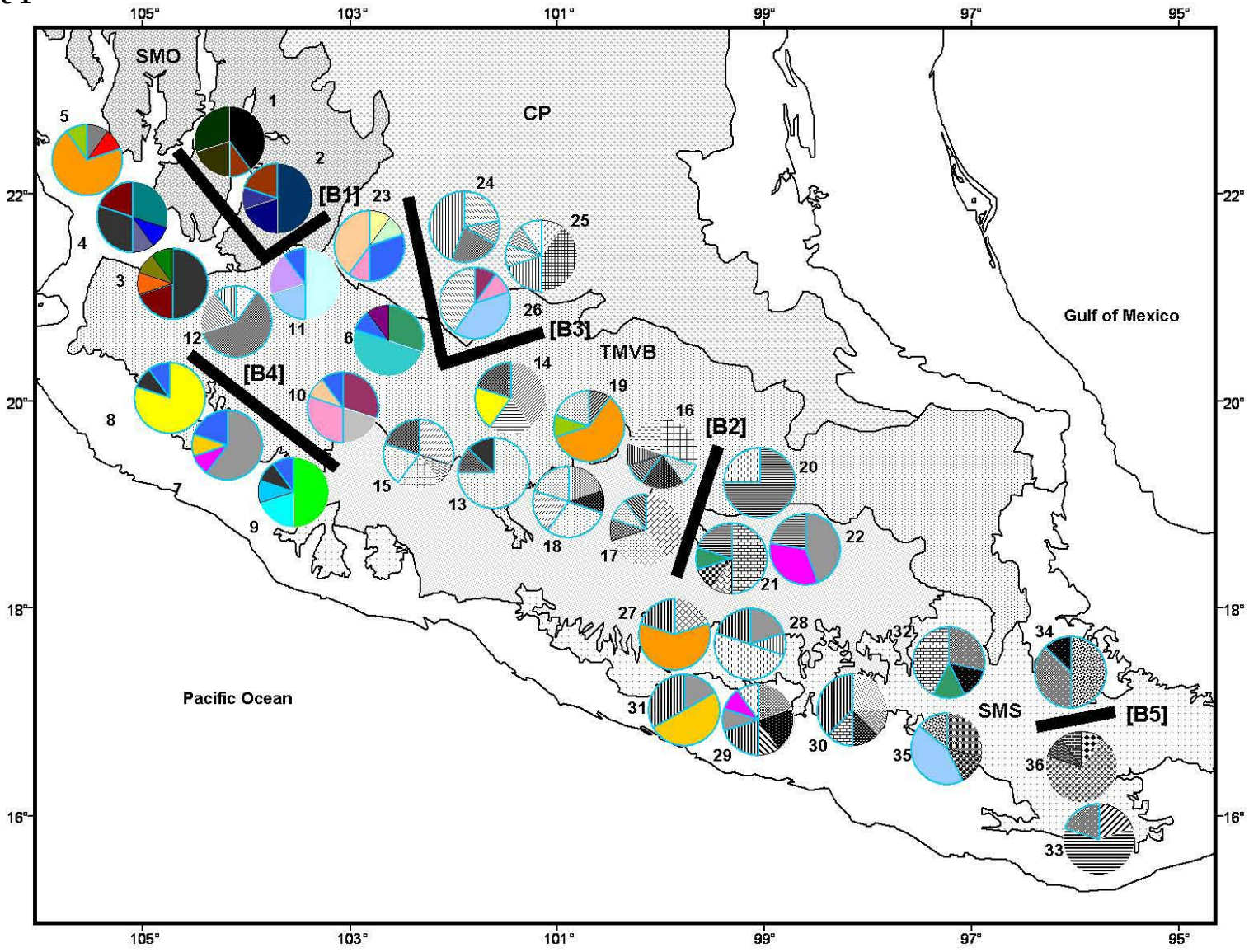

105°
103°
101°
99°
97°
95°
22°
20°
18°
16°
SMO
CP
TMVB
Gulf of Mexico
Pacific Ocean
SMS
[B1]
[B2]
[B3]
[B4]
[B5]
1
2
3
4
5
6
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

Figure 2

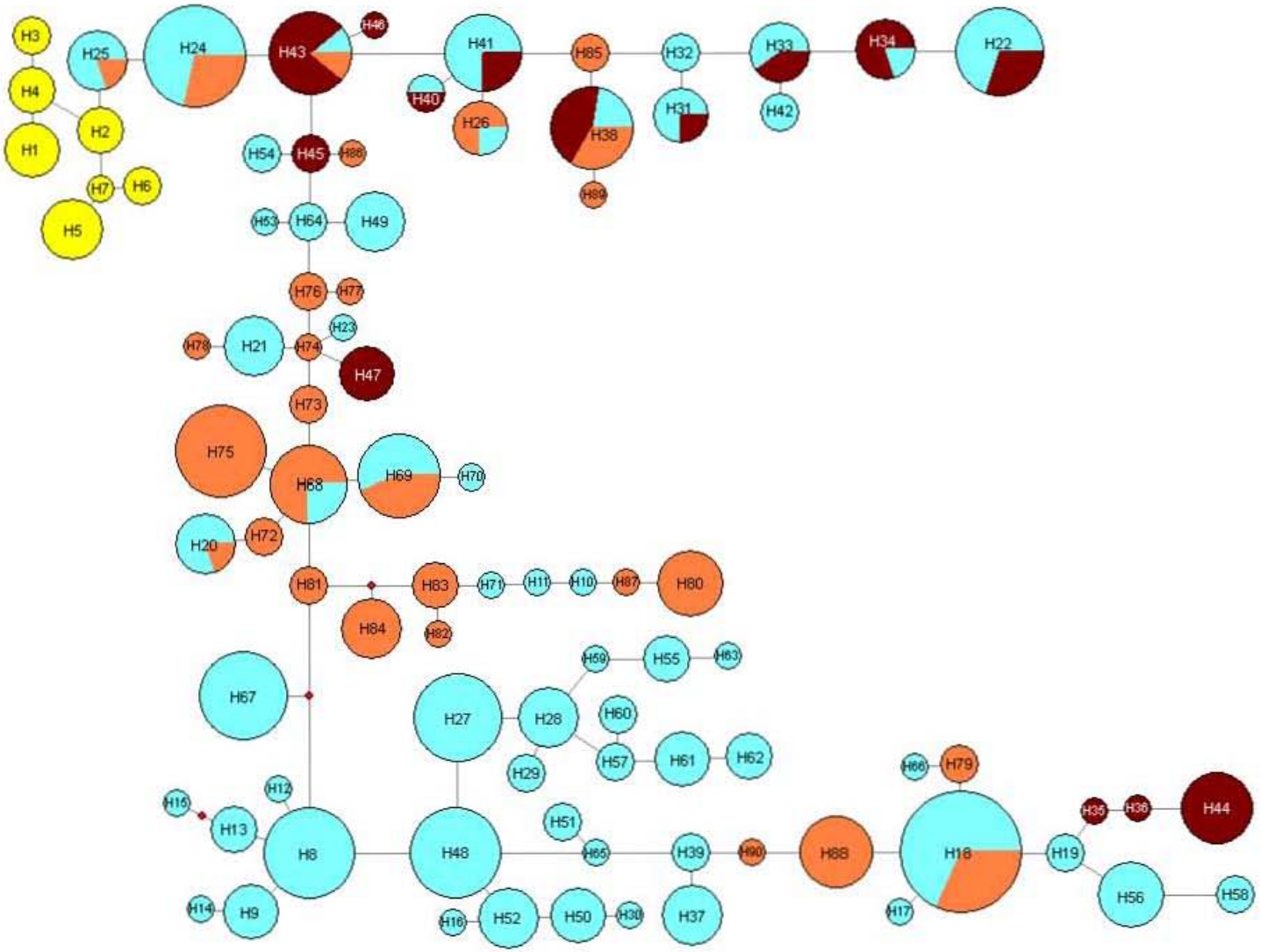

H3
H25
H24
H43
H46
H41
H85
H32
H33
H34
H22
H4
H40
H26
H38
H31
H42
H2
H1
H54
H45
H88
H89
H7
H6
H5
H63
H64
H49
H76
H77
H23
H78
H21
H74
H47
H73
H75
H68
H69
H70
H20
H72
H81
H83
H71
H11
H10
H87
H80
H84
H82
H59
H55
H63
H67
H27
H28
H60
H57
H61
H62
H29
H66
H79
H12
H16
H13
H8
H48
H51
H65
H39
H90
H88
H18
H35
H36
H44
H19
H56
H58
H14
H9
H16
H52
H50
H30
H37
H17

Figure 3

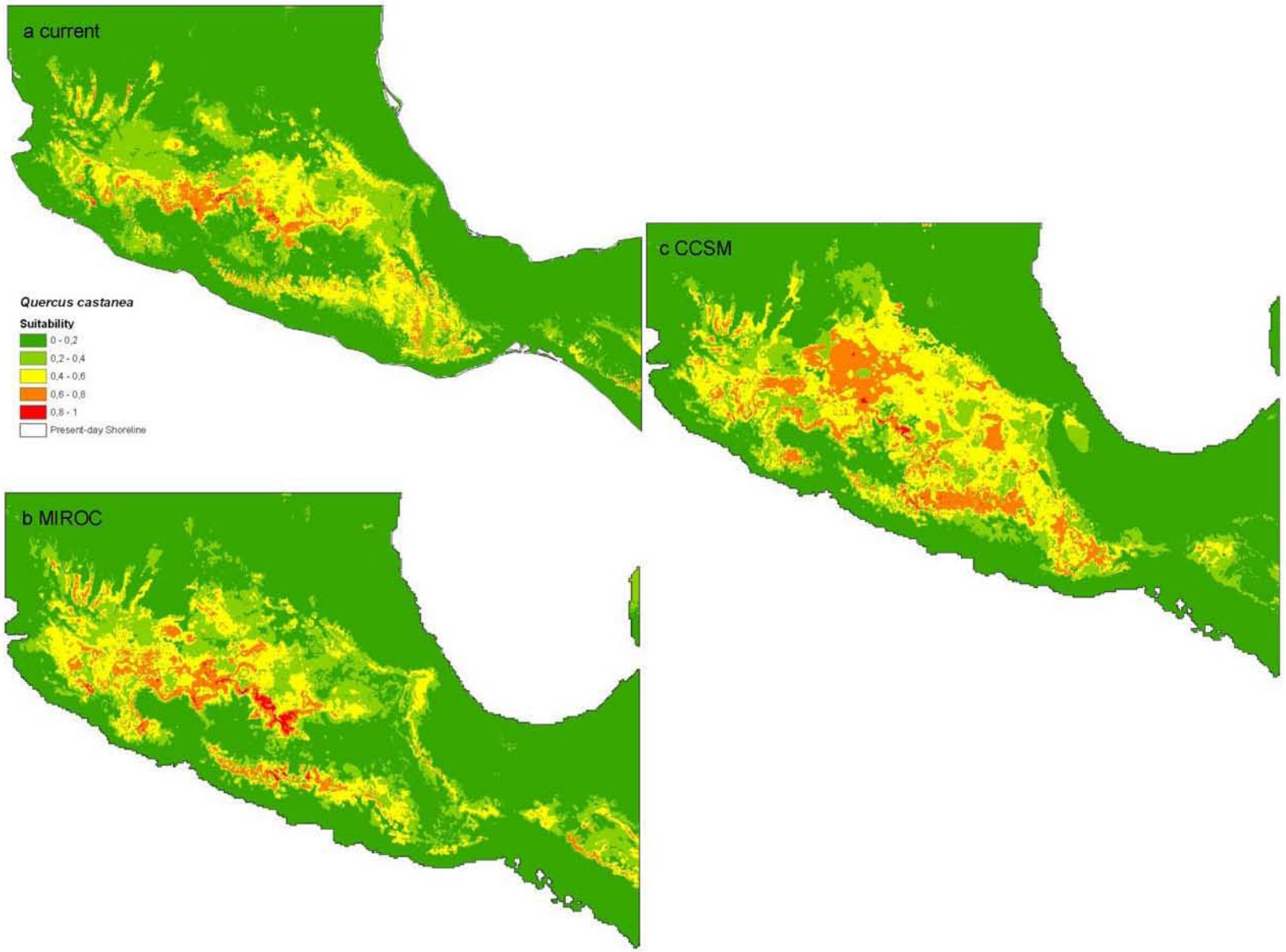
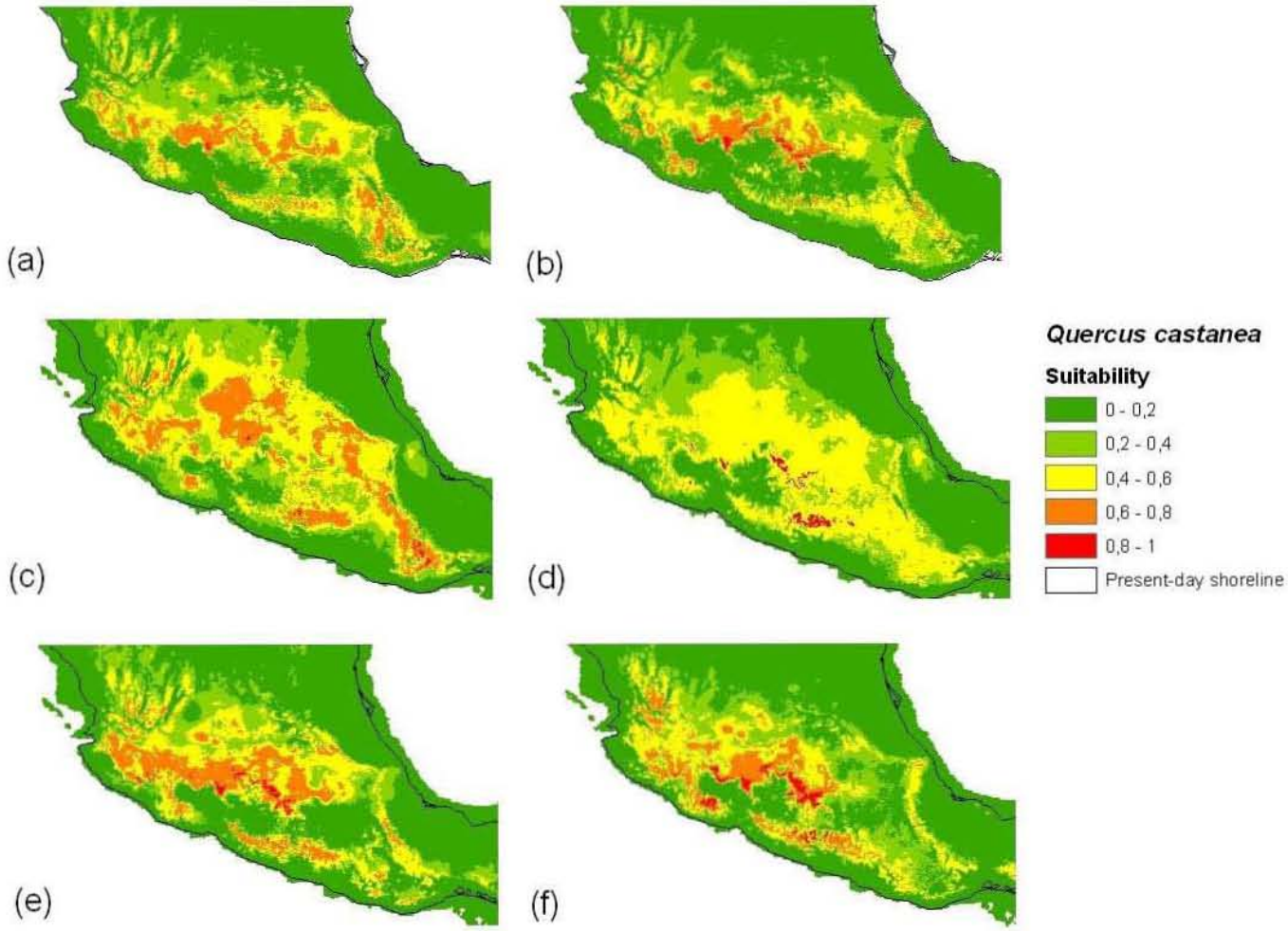

a current
Quercus castanea
Suitability
0 - 0,2
0,2 - 0,4
0,4 - 0,6
0,6 - 0,8
0,8 - 1
Present-day Shoreline
c CCSM
b MIROC

Figure S1


(a)
(b)
(c)
(d)
(e)
(f)
Quercus castanea
Suitability
0 - 0,2
0,2 - 0,4
0,4 - 0,6
0,6 - 0,8
0,8 - 1
Present-day shoreline

VIII. DISCUSION GENERAL

Dinámica de la zona de hibridación

La hibridación natural es muy frecuente e importante en la evolución y especiación de las plantas (Rieseberg, 1997). Este proceso ocurre debido a que las barreras reproductivas son débiles lo que permite el flujo genético con otras especies (Craft *et al.* 2002). El proceso de hibridación sucede cuando las especies se sobrelapan espacial y temporalmente en su época de floración y se reproducen, dando origen a una progenie viable o parcialmente fértil (Hewitt, 1989; Arnold, 1997). Las zonas de hibridación son vistas como un recurso potencial de generación de nuevos genotipos y por lo tanto de la colonización a nuevos medioambientes donde los progenitores no puedan competir con ellos (Arnold, 1997). El estudio de las zonas de hibridación es una de las áreas más dinámicas de la biología evolutiva y la ecología (Howard *et al.* 1997). Uno de los debates actuales más importantes, es el origen de las zonas de contacto en un contexto histórico y geográfico en el cual la introgresión ocurre (Arnold, 1997). Sin embargo, aún es necesario determinar que factores de selección están modelando la permanencia o desaparición de las zonas híbridas en la naturaleza (Barton y Hewitt, 1985). La naturaleza y la intensidad de los factores selectivos no son necesariamente homogéneas en el ambiente. Los factores de selección exógena y endógena, han sido requeridos para explicar la dinámica y estructura de las zonas híbridas (Barton y Hewitt, 1985; Moore y Price, 1993). La selección exógena implica la adaptación a las condiciones ambientales locales, es decir que la capacidad relativa de los individuos híbridos varían con su medio ambiente; mientras que la selección endógena ocurre cuando los híbridos tienen una baja adecuación, debido a incompatibilidades entre los genomas progenitoras (Bronson *et al*. 2003). Por lo tanto, las presiones selectivas que actúan en los híbridos pueden resultar de los efectos combinados de factores medioambientales tales como el clima, suelo y sus interacciones con otros organismos (Fitzpatrick y Schaffer,

2004; James y Abbot, 2005; Raudnitschka *et al.* 2007) así como de factores independientes del medioambiente tales como la disrupción de complejos de genes coadaptados y la acción de genes asociados con la esterilidad (Barton y Gale, 1993; Rieseberg y Wendell, 1993). De esta manera, si las principales presiones de selección son exógenas, se pudiera esperar que las zonas de hibridación resulten en clinas o de mosaico, en las cuales ciertos genotipos pudieran estar espacialmente segregados de acuerdo con la distribución de los hábitats (Endler, 1977). Sin embargo, si encontramos que la fuerza principal de selección es de naturaleza endógena, entonces muy probablemente se generara una zona de tensión, donde las clinas pueden ser mantenidas por el balance entre el movimiento de los individuos progenitoras y la selección en contra de los genotipos híbridos (Barton y Gale, 1993).

De esta manera en el primer capítulo de esta tesis, se trató de determinar los niveles de variación genética y en la forma de las hojas para tratar de evidenciar que tipo fuerzas selectivas son las que están determinando la distribución de los genotipos híbridos en cada una de las poblaciones de *Q. scytophylla*, *Q. hypoleucoides*, y *Q. sideroxyla* y en una serie de poblaciones con individuos intermedios, por medio de microsatélites nucleares y técnicas morfometría geométrica En este estudio se encontraron valores altos de diversidad genética para las especies examinadas *Q. scytophylla* N_a=7.81, H_E=0.883; *Q. hypoleucoides* N_a= 7.0, H_E=0.834; y *Q. sideroxyla,* N_a=9.54, H_E=0.894, que son comparables a los obtenidos en otras especies de encinos donde se analizan especies que hibridan (Cotrell, *et al.* 2002; Craft *et al.* 2002). Encontramos niveles de diferenciación genética significativa pero baja para las poblaciones de las tres especies, que nos permitieron poder discriminar entre ellas y que según Wright (1978) valores que van de 0 a 0.05 indican una baja diferenciación genética. Este patrón de baja diferenciación ha sido comúnmente observado en las especies de encinos y pudieran ser debidos a la existencia de polimorfismos ancestrales compartidos, este fenómeno ha sido documentado en especies cercanamente

emparentadas o en especies recientemente separadas con grandes tamaños efectivos poblacionales (e.g. Muir y Schloterer, 2005) o con problemas taxonómicos para reconocerlas por su morfología muy similar (Holman *et al.* 2003). En estas especies se encuentran polimorfismos ancestrales que aún se comparten y que resultan en filogenias parafiléticas o en procesos de introgresión entre las especies (Beckstrom-Strenberg *et al.* 1991). Sin embargo los casos de polimorfismos ancestrales necesitan condiciones muy particulares, por lo que en otros estudios el flujo genético interespecífico ha sido considerado como una explicación más adecuada para poder entender de manera adecuada los eventos de introgresión entre especies (Dodd y Kashani, 2003; González-Rodríguez *et al.* 2005; Craft y Ashley, 2006; Lexer *et al.* 2006; Curtu *et al.* 2007; Gugerli *et al.* 2007; Albarrán-Lara *et al.* 2010; Peñaloza-Ramírez *et al*. 2010a).

Usando STRUCTURE, se apoya la existencia de tres grandes grupos genéticos asociados a cada especie progenitora (*Q. scytophylla*, *Q. hypoleucoides* y *Q. sideroxyla*), cuando sus poblaciones se encuentran en alopatría, aunque cuando las poblaciones progenitoras ocurren en simpatría, los individuos mostraron una ascendencia intermedia. Estos resultados sugieren que cuando la hibridación ocurre entre las especies, la permanencia de los enjambres de híbridos puede provocar cruzas preferentemente entre híbridos y progenitores o retrocruzas causando gran discrepancia entre los fenotipos híbridos resultantes (Valbuena-Carabaña *et al.,* 2007). Los individuos intermedios (e.g. morfológica y genéticamente) distribuidos a lo largo del gradiente altitudinal en el norte de la Sierra Madre Occidental, mostraron claramente ser el resultado de la introgresión entre un par de especies aunque en diferente combinación. Sin embargo, en otras poblaciones se observa claramente la mezcla de las tres especies en la zona de contacto. Las zonas de hibridación generalmente se enfocan en la intergradación morfológica y genética de dos especies progenitoras en una zona de contacto (Arnold, 1997). Sin embargo, la

intergradación que ocurre entre más de dos especies necesariamente significa que el intercambio de genes ha tenido más que un efecto transitorio en al menos una especie, lo que podría considerarse evidencia de múltiples eventos de introgresión (Riesberg y Wendell, 1993; Kaplan y Fehrer, 2007). Estos resultados sugieren que la distribución de los genotipos híbridos está conformada en una zona de hibridación en mosaico y que de acuerdo a los modelos de zonas de hibridación son áreas donde las especies progenitoras y las poblaciones intermedias están distribuidas en parches a través de la zona de sobrelapamiento (Harrison y Rand, 1989), donde incluso se encontró una fuerte correlación con la altitud, latitud y longitud. Algunos estudios han evidenciado que la influencia de las variables medioambientales en la estructura de las zonas de hibridación, tal como se reporta en el estudio de Dodd y Afzal-Rafii (2004) donde se sugiere que la existencia de gradientes medioambientales pueden determinar el grado de introgresión. En otras especies de plantas los patrones de hibridación e introgresión han sido a menudo asociadas con variables medioambientales que cambian con la elevación (e.g. James y Abbott, 2005; Kimball, 2008). En este sentido, nuestros resultados sugieren que el limitado efecto de la selección endógena y la influencia del medioambiente en la distribución de los genotipos híbridos, puedan estar delimitando la dinámica de esta zona de hibridación multi especies como la que ocurre en la Sierra Tarahumara.

Filogeografía de encinos

Particularmente, en las especies de árboles de los bosques templados como los encinos del género *Quercus* se han realizado estudios filogeográficos exhaustivos en Europa y que han revelado en general un impacto mucho mayor de los cambios en el clima del Cuaternario en la distribución y conectividad de sus poblaciones (Dumolin-Lapegue *et al.* 1997; Le Corre y Kremer, 1998; Belahbib *et al.* 2001; Petit *et al.* 2002). En contraste,

las especies de encinos Norteamericanos a pesar del avance de la capa de hielo durante las glaciaciones, estas pudieron mantenerse en refugios Pleistocénicos donde las especies respondieron de manera diferente y dependiendo de su capacidad para afrontar los procesos de expansión y contracción de rango que dieron como resultado que algunas especies tuvieran una baja estructura genética (Magni *et al.* 2005) o una alta diversidad genética (Grivet *et al.* 2006) a diferencia de los encinos Europeos (Marsico y Hellman, 2009). Sin embargo, sabemos poco de la historia evolutiva de las especies de encinos en las regiones tropicales y subtropicales donde se encuentra las más alta riqueza de especies. Particularmente, México es considerado como un centro de diversificación secundario para el genero *Quercus* con aproximadamente 161 especies que representan cerca de 32-40% de la diversidad mundial (Valencia, 2004). Es posible que en las latitudes tropicales los periodos glaciares durante el Pleistoceno fueron probablemente menos severos que en otras regiones de Europa y América del Norte (Grivet *et al.* 2005; Magri *et al.* 2006). En consecuencia, es necesario realizar estudios adicionales en especies de encinos Mexicanos que complementen la filogeografía y el modelado de nicho ecológico para tratar de entender cual ha sido el impacto de las glaciaciones del Pleistoceno en las comunidades de bosques templados de México.

Para el segundo capítulo se trató de discernir los patrones filogeográficos, evidenciar el intercambio citoplasmático entre las especies de encinos rojos *Q. hypoleucoides, Q. scytophylla*, y *Q. sideroxyla* en México y comparar con los modelos de nicho ecológico. Los resultados obtenidos en 51 poblaciones de las tres especies de encinos rojos en toda su distribución en México revelaron altos valores de diversidad dentro y entre poblaciones de *Q. hypoleucoides* con valores de $H_S = 0.671$, $H_T = 0.932$, para *Q. scytophylla* con valores de $H_S = 0.708$ y $H_T = 0.972$ y para *Q. sideroxyla*, los valores fueron $H_S = 0.695$, $H_T = 0.981$. En particular se observaron valores altos de diversidad en

poblaciones donde el intercambio citoplasmático fue más frecuente, lo cual refleja que los procesos de hibridación entre las especies incrementan la variación genética tal como se había reportado para los complejos de encinos rojos *Q. affinis*-*Q. laurina* y *Q. crassifolia* x *Q. crassipes* (González-Rodríguez *et al.* 2004; Tovar-Sánchez *et al.* 2008). Al igual que en el capítulo 1, donde también se observaron evidencias de hibridación e introgresión en las poblaciones del norte de la Sierra Madre Occidental, aunque la dirección en el intercambio esta mediada por *Q. hypoleucoides* que presenta los haplotipos más invasivos hacia *Q. scytophylla* y *Q. sideroxyla*. Incluso en la red de haplotipos muestra tan solo un haplotipo compartido entre las tres especies, y algunos haplotipos compartidos entre diferentes pares de especies, por ejemplo cinco haplotipos compartidos entre *Q. hypoleucoides* y *Q. sideroxyla*, tres entre *Q. hypoleucoides* y *Q. scytophylla* y tres entre *Q. scytophylla* y *Q. sideroxyla*. Estos resultados coinciden con los resultados encontrados en el capitulo 1 donde observamos que las tres especies tienen las mismas oportunidades en la producción de híbridos F1 y retrocruzas, por lo que muy posiblemente las tres especies jueguen un papel esencial en la transmisión de los genotipos hacia las otras dos especies. Estos resultados al parecer señalan el hecho de que en estas especies los mecanismos de aislamiento reproductivo que impiden el desarrollo de los individuos híbridos, son más débiles y aún no se ha encontrado evidencia que sugiera que exista una dirección en el flujo de genes tal como se ha evidenciado en otros estudios (Bacilieri *et al.* 1996; Steinhoff, 1998). Incluso la abundancia relativa de cada especie en la zona de contacto (Lepais *et al.* 2009) puede jugar un papel preponderante en la frecuencia de la hibridación, la dirección de la introgresión dado que los posibles eventos de hibridación pueden estar mediados por las características del hábitat de las especies (Williams *et al.* 2001; Himrane *et al.* 2004). Sin embargo aún faltan estudios donde se realicen polinizaciones entre las especies para estimar la tasa de germinación y sobrevivencia de los individuos híbridos, para poder saber

que combinaciones de híbridos son los que tienen una mayor ventaja sobre los demás (Petit *et al.* 2003; Bacilieri *et al.* 1996)

Por otra parte, encontramos valores bajos de diferenciación para *Q. hypoleucoides* de G_{ST} = 0.280, *Q. scytophylla* G_{ST} = 0.271 y para *Q. sideroxyla* G_{ST} = 0.291 respectivamente. Estos valores contrastan con la alta diferenciación genética encontrada para los encinos blancos de América del Norte como *Q. garryana* con valores de G_{ST} = 0.88 (Marsico y Hellman, 2009), y *Q. lobata* con valores de G_{ST} = 0.709 (Grivet *et al.* 2006). De la misma manera, las especies de encinos blancos de Europa muestran valores altos para la diferenciación, por ejemplo *Q. robur* tiene un G_{ST} = 0.78, para *Q. petreae* G_{ST} = 0,86, para *Q. pyrenaica* G_{ST} = 0.96, para *Q. ilex* de G_{ST} = 0.92 *Q. suber* los valores de G_{ST} = 0,84 (Petit *et al.* 2002). Una de las posibles explicaciones es que a diferencia de las especies de Europa y de América del Norte, en las especies de encinos Mexicanos por el contrario, los registros palinológicos sugieren que el impacto de los cambios durante el Pleistoceno en las especies de encinos rojos no fue tan dramático (Metcalfe *et al.* 2000) en la estructura genética de sus poblaciones, lo que les permitió mantener cierto niveles de intercambio genético (Hewitt, 1999). Estos resultados indican que a pesar de la existencia de intercambio genético entre las especies, las poblaciones incluso mantienen ciertos niveles de estructura genética (Dodd y Kashani, 2003; Magni *et al.* 2005) y que a través del tiempo se han permitido la acumulación y mantenimiento de una mayor diversidad genética (Le Corre y Kremer, 1998; Grivet *et al.* 2006).

Los resultados de los modelos de distribución MAXENT para la distribución durante el último glaciar máximo en las tres especies muestran una significativa reducción en las poblaciones de la SMO, a diferencia de los modelos actuales. Incluso los registros palinológicos durante el Pleistoceno medio sugieren que los pastos alpinos estaban mucho más extendidos en áreas como Baja California, Sonora y Chihuahua que presentaban

condiciones frías y áridas (Lachniet y Vazquez-Selem, 2005; Metcalfe, 2006), que son asociadas con la retracción de los bosques templados en islas de bosque rodeados por desiertos o pastizales o vegetación xerofítica en las partes más bajas (Martínez-Hernández y Lozano-García, 1996). Aunque en el Pleistoceno tardío las condiciones climáticas parecen ser más frías que el presente (5-6° C menos) con un incremento en las lluvias de invierno (Metcalfe *et al.* 2000; Lozano-García *et al.* 2005). Esta evidencia sugiere que periodos sucesivos de expansión y contracción actitudinal ocurridos durante el Pleistoceno pudieran haber estimulado la hibridación cíclica como un mecanismo importante en la evolución de estas especies de encinos en la Sierra Madre Occidental (Bacon y Spellenberg, 1996). Además, este mecanismo puede ser particularmente importante en México, donde la gran heterogeneidad en su topografía, el clima y la extensa variedad de hábitats, probablemente ha promovido la formación de zonas híbridas entre diferentes especies de encinos rojos y blancos en nuestro país.

En el caso de *Q. scytophylla* y *Q. sideroxyla* podemos observar una reducción significativa en el hábitat más adecuado en poblaciones del norte y este de la FVTM, pero incluso en algunas poblaciones de la SMS, en contraste con la amplia distribución observada actualmente. En este sentido los registros de la Faja Volcánica Transmexicana muestra patrones muy complejos, por ejemplo durante el Pleistoceno tardío se observan condiciones contrastantes entre la región oeste de la FVTM que presenta condiciones en general más húmedas que en el presente (la presencia de *Alnus*, *Carpinus*, *Corylus* y *Salix*), y a diferencia de la región este que muestra en general condiciones más secas que en el presente (la presencia de *Cuppressaceae* y *Juniperus*) (Bradbury, 1997; Metcalfe, 2006). De esta manera pudiéramos sugerir que la posible reducción del hábitat potencial resulte en un proceso divergente entre las poblaciones de la FVTM para las dos especies tal como fue sugerido por la diferenciación genética entre esas poblaciones. Una de las posibles razones

pudiera estar asociada a las restricciones en las áreas disponibles donde las poblaciones de las especies puedan sobrevivir durante las glaciaciones, quedando aisladas a ciertas áreas del total de la distribución (Hewitt, 2000; Widmer y Lexer, 2001). Sin embargo, durante la transición del Pleistoceno-Holoceno las poblaciones se expandieron en su distribución según los modelos actuales. Esta evidencia es concordante con las pruebas de neutralidad que sugieren expansión en las poblaciones de *Q. scytophylla* y *Q. sideroxyla*. Incluso la red de haplotipos muestra una forma de estrella, que frecuentemente son asociadas a posibles eventos de expansión histórica (Schaal *et al.* 1998; Avise, 2000), esta conformación de haplotipos pudiera sugerir el origen in-situ de los linajes, diferenciación y colonización siguiendo cada uno diferentes rutas de migración.

Las poblaciones de *Q. scytophylla* en las SMS se encontraron reducidas tal como fue sugerido por los modelos para el último glaciar máximo. En este sentido los registros de polen fósil durante el último glaciar máximo indican una tendencia a condiciones climáticas frías y secas, lo que probablemente promovió la recolonización de los pastizales alrededor de la Cuenca del Balsas entre otras (Martínez-Hernández y Lozano-García, 1996). Con esta información, pudiéramos sugerir que la reducción de los bosques templados fue influenciada por la combinación de algunos factores principales, tal como la reducción de la humedad y la actividad volcánica en la zona (Lozano-García y Ortega-Guerrero, 1998; Lozano-García *et al.* 2005). Sin embargo, durante la transición del Pleistoceno-Holoceno la distribución geográfica de las poblaciones aumento según lo que sugieren los modelos de distribución actual. Incluso las pruebas de neutralidad muestran expansión para las poblaciones de *Q. scytphylla* en la SMS. Estos resultados indican que la distribución de los haplotipos posiblemente refleje periodos recurrentes de expansión-contracción y migración de las especies durante los periodos glaciares e inter-glaciares, tal como fue encontrado para los encinos de California (Dodd y Kashani, 2003).

En el capítulo dos se trató de reconstruir los patrones filogeográficos, su demografía histórica y comparar los resultados con modelos de nicho ecológico para la distribución actual y durante el último glaciar máximo de la especie de encino rojo *Q. castanea* que presenta una extensa distribución geográfica en México. Los resultados obtenidos en 36 poblaciones de *Q. castanea*, indican 90 haplotipos totales, 28 de los cuales son únicos, la diversidad dentro y entre poblaciones fue alta (h_S = 0,730; h_T = 0.989), así como también la diferenciación de las poblaciones que resulto moderada G_{ST}= 0.261. Para poder comparar estos valores de diversidad y diferenciación con otros estudios filogeográficos en encinos, se utilizaron tamaños de muestra y de loci genéticos comparables con aquellos estudios en especies de encinos rojos y blancos como los de Grivet *et al.* (2006), Magri *et al.* (2006) y Marsico y Hellman (2009). Los resultados sugieren que *Q. castanea* mantiene una considerable alta riqueza haplotípica y diversidad genética dentro de las poblaciones, que cualquier otra especie de encino estudiada en las latitudes del norte. Se han propuesto algunas hipótesis que puedan explicar el mantenimiento de esta alta diversidad en *Q. castanea*, por ejemplo que haya sido necesario grandes tamaños efectivos mantenidos durante largos periodos de tiempo. Incluso, parece que los más recientes ciclos glaciares e interglaciares del Pleistoceno han tenido un efecto moderado en la demografía y distribución de las poblaciones y haplotipos de *Q. castanea*, que probablemente han permanecido *in situ* sin grandes cambios en el tamaño.

Las pruebas de neutralidad mismatch distribution y las pruebas de Fu *FS* mostraron en general que las poblaciones de *Q. castanea* presentan distribuciones unimodales y un índice de Raggedness no significativo, además de valores negativos y significativos para las pruebas de neutralidad de Fu *FS*. Estos resultados sugieren la existencia de expansión poblacional (Rogers y Harpending, 1992) que con frecuencia son asociados con valores altos de diversidad genética, grandes tamaños poblacionales y por una forma de estrella en

la red de haplotipos (Schneider y Excoffier, 1999). Las estimaciones en años del tiempo desde la expansión poblacional, sugieren que las poblaciones de la SMS son las más antiguas de alrededor de 1.6×10^6 y 39,964 años antes del presente. Después de esta expansión, la especie probablemente colonizo y se expandió en las poblaciones de la FVTM y el SMO en tiempos de alrededor de 6×10^5 y años y finalmente las poblaciones de la PC se expandieron hace unos 197,714 y 4768 años atrás. Debido a que no existe un consenso en las tasas de mutación para encinos, decidimos utilizar dos tasas de mutación y dos tiempos generacionales distintos, por lo que encontramos dos estimaciones en años para los posibles eventos de expansión en las poblaciones. Estos resultados sugieren que dependiendo de la tasa de mutación utilizada se pueden encontrar varios posibles eventos de expansión en las poblaciones, por lo que es importante tener un intervalo como el que usamos para poder definir cuando posiblemente hayan existido eventos de expansión y no solamente utilizar una tasa de mutación y un tiempo generacional que se ajuste al tiempo que nosotros queramos. Por ejemplo, encontramos dos posibles eventos de expansión, uno más antiguo incluso que los modelos de distribución para el último glaciar máximo y el segundo más reciente que incluso son congruentes con los modelos de distribución que muestran una mayor distribución de las poblaciones de *Q. castanea* durante el último glaciar máximo.

Los modelos de distribución de MAXENT incluso sugieren cierta estabilidad de las poblaciones durante los últimos miles de años. La especie tenía probablemente una distribución mucho más grande durante el ultimo glaciar máximo (UGM), aunque en el Holoceno experimento una cierta contracción de rango moderada que fluctúa entre 30 y 50%, y que al parecer no causo una importante reducción de la diversidad genética. Es posible que las poblaciones de *Q. castanea* hayan experimentado desplazamientos altitudinales en vez de grandes migraciones o colonizaciones latitudinales, y que en

consecuencia haya permitido el mantenimiento de altos niveles de diversidad genética y que algunos haplotipos locales hayan sido retenidos. Estudios en especies de amplia distribución han mostrado en general altos niveles de variación genética total y dentro de poblaciones, además se ha encontrado la presencia de estructura filogeográfica y una compleja distribución geográfica de los haplotipos, lo cual indica que esas especies han tenido una compleja historia poblacional en las áreas que actualmente son ocupadas por ellas (González-Rodríguez *et al.* 2004; Tovar-Sánchez *et al.* 2008; Moreno-Letelier y Piñero, 2009; Rodríguez -Banderas *et al.* 2009).

Las poblaciones que se encuentran en las cuatro regiones biogeográficas (SMO, FVTM, PC y SMS) se diferenciaron significativamente según el análisis de AMOVA, lo cual indica que ha existido el tiempo suficiente para la acumulación de mutaciones distintas entre las regiones. Sin embargo, las barreras genéticas identificadas por medio del análisis de Monmonier's no fueron completamente concordantes con los limites entre las regiones biogeográficas. Algunos estudios ya habían detectado una coincidencia débil entre las discontinuidades genéticas entre las poblaciones de especies de árboles templados y las tradicionalmente delimitadas zonas biogeográficas de México, tal como lo había notado previamente Jaramillo-Correa *et al.* (2009), en *Pinus strobiformis* (Moreno-Letelier y Piñero, 2009) y en el complejo de especies de pino *P. montezumae-P. harttweggii* (Matos y Schaal, 2000).

Estos resultados sugieren que las regiones biogeográficas no son unidades homogéneas, sino que más bien existe una heterogeneidad geológica y climática que va a tener un efecto en las áreas de hábitat adecuadas para *Q. castanea*. Particularmente la región de la FVTM ha tenido una historia muy compleja que probablemente ha tenido un papel muy importante en el modelado de la estructura genética de muchas especies. Los registros geológicos indican que la porción oeste de esta cadena montañosa comenzó su

formación y levantamiento hace 23 millones de años, mientras que la porción este se formo hace 2.5 millones de años (Ferrusquia-Villafranca, 1993), lo que implica procesos históricos muy diferentes entre las dos porciones de la FVTM. Algunos análisis biogeográficos han sugerido que la FVTM no es una unidad homogénea (e.g. Corona *et al.* 2007). Incluso los registros palinológicos indican condiciones medioambientales húmedas y calidas en la porción este de la FVTM en contraste con la porción oeste que presenta condiciones más templadas y secas al final del Pleistoceno (Bradbury, 1997; Lozano-García y Ortega-Guerrero, 1997). En este sentido nuestros modelos de distribución para el último glaciar máximo muestran una gran cantidad de hábitat disponible para *Q. castanea* en la porción oeste de la FVTM. Esta información coincide con el número de linajes existentes en esta área y que se distribuyeron hacia otras poblaciones en la PC, la porción sureste de la FVTM y la SMS.

El intercambio genético entre las poblaciones en las regiones biogeográficas que fueron evidenciados por la red de haplotipos pueden ser explicados por factores similares. Por ejemplo en la región de la SMS probablemente han existido conexiones históricas con la porción sureste de la FVTM (Corona *et al.* 2007), lo que probablemente pudo haber facilitado la migración de antiguos linajes poblacionales entre las dos regiones. Más recientemente durante el Pleistoceno tardío, los registros palinológicos de Ixtacyola e Ixtapa en la región de la SMS indican condiciones húmedas y frías (McNeish *et al.* 1972; Piperno et al., 2007) asociados con expansiones de los bosques templados hacia las partes más bajas formado grandes poblaciones continuas (Martínez-Hernández, 1992; Lozano-García y Ortega-Guerrero, 1997), lo cual es incluso sugerido por nuestros modelos de distribución para *Q. castanea*. Sin embargo hacia el final del Pleistoceno el corredor de bosque tropical seco se expandió en muchas áreas (e.g. Cuenca del Balsas), por lo que los bosque templados se contrajeron en altitudes mayores en las montañas (Metcalfe, 2006). Es

posible que esos recurrentes periodos de expansión y contracción altitudinal posiblemente resulten en periodos de diferenciación seguidos por el contacto entre los bosques de pino-encino de la porción sur de la FVTM y el norte de la SMS. Este escenario es congruente con las claras afinidades genéticas de las poblaciones de *Q. castanea* entre esas dos áreas. En la PC encontramos una historia similar donde ocurrieron migraciones altitudinales, expansiones y contracciones de los bosques templados, y que es congruente por los modelos de distribución y por los registros palinológicos que muestran grandes abundancias de vegetación de bosque templado en los valles y las partes bajas durante el último glaciar máximo (Martínez-Hernández, 1992; Metcalfe, 2006).

Con estos resultados podemos sugerir que el efecto de estos ciclos recurrentes probablemente promovieron los patrones genéticos encontrados en *Q. castanea* y que muy posiblemente podrían estar relacionadas con los complejos eventos climáticos y geológicos que ocurrieron durante el Pleistoceno, lo cual seguramente causaron sucesivas expansiones y contracciones de los hábitats idóneos o característicos para esta especie. Finalmente y tal como se ha observado con otras especies, el origen y mantenimiento de una gran proporción de la diversidad genética en *Q. castanea* ha ocurrido dentro de la FVTM. Esta región es incluso considerada como un importante centro de diversificación para las especies de encinos (Nixon, 1993) y ha actuado como un puente natural entre otras provincias biogeográficas, por lo que la FVTM se ha constituido como un activo centro de evolución de la biota Mexicana.

En conclusión las especies de encinos rojos en este estudio han experimentado respuestas muy diferentes en la distribución de sus poblaciones para los modelos de nicho ecológico. En este sentido, al parecer las condiciones ecológicas óptimas donde las especies pueden desarrollarse como el tipo de suelo, clima, temperatura, precipitación y muy posiblemente por el intervalo actitudinal en el cual se desarrollan juegan un papel muy importante ya que especies que tienen grandes distribuciones geográficas como *Q. castanea* (Peñaloza-Ramírez et al en prep), *Q. candicans* y *Q. crassifolia* (Torres-Miranda *et al.* en prep) [i.e. de acuerdo a los modelos de nicho] pueden responder mejor durante los periodos glaciares. Mientras que las especies con distribuciones mucho más moderadas como *Q. hypoleucoides*, *Q. sideroxyla* y *Q. scytophylla* (Peñaloza-Ramírez *et al.* en prep), *Q. eduardii*, *Q. coccolobifolia*, *Q. emory*, y *Q. albocinta* (Torres-Miranda *et al.* en prep) pueden responder mejor durante la transición del Pleistoceno-Holoceno hacia condiciones menos frías. Tal como se ha observado en las especies de encinos Mexicanos al parecer han sido grandemente influenciadas por la historia geológica y climática, lo cual ha dado como resultado eventos recurrentes de contracción y dispersión, grandes tamaños efectivos poblacionales y una dinámica muy compleja en los procesos de intercambio citoplasmático entre las especies (González-Rodríguez *et al*. 2004, Tovar-Sánchez *et al.* 2008; Peñaloza-Ramírez *et al.* en prep).

IX. REFERENCIAS GENERALES

Albarrán-Lara, A.L., Mendoza-Cuenca, L., Valencia-Ávalos, S., González-Rodríguez, A., & Oyama, K. (2010). Leaf fluctuating asymmetry increases with hybridization and introgression between *Quercus magnoliifolia* and *Quercus resinosa* (Fagaceae) through an altitudinal gradient in Mexico. *International Journal of Plant Sciences*, **171**, 310–322.

Arnold, M.L. (1992). Natural hybridization as an evolutionary process. *Annual Review of Ecology, Evolution, and Systematics*, **23**, 237-261.

Arnold, M.L. (1997). *Natural hybridization and evolution*. Oxford, Oxford University Press.

Avise, J.C. (1998). The history and purview of phylogeography: A personal reflection. *Molecular Ecology*, **7**,371-381.

Avise, J.C. (2000). *Phylogeography. The history and formation of species*. Harvard University Press. Cambridge, England.

Axelrod, D.I. (1983). Biogeography of Oaks in the Arcto-Tertiary Province. *Annual Review of Ecology, Evolution, and Systematics*, **70**, 629-657.

Bacilieri, R., Ducousso, A., & Kremer, A. (1996) Comparison of morphological characters and molecular markers for the analysis of hybridization in sessile and pedunculate oak. *Annales des Sciences Forestières*, **53**, 79–91.

Bacon, J.R., & Spellenberg, R. (1996). Hybridization in two distantly related mexican black oaks *Q. conzatti* y *Q. eduardii* (Fagaceae: Quercus: Sección Lobatae). *Sida*, **17**, 17-41.

Bacon, J.R. (1996). Oaks of the Chihuahuan Desert Region. *Chihuahuan Desert Discovery*, **1**, 6-8.

Barton, N.H., & Gale, K.S. (1993). Genetic analysis of hybrid zones. **En**: *Hybrid zones and the evolutionary process* (eds. Harrison, R.G.), Oxford University Press, Oxford, pp 13-45.

Barton, N.H. & Hewitt, G.M. (1985). Analysis of hybrid zones. *Annual Review of Ecology, Evolution, and Systematics*, **16**, 113-148.

Belahbib, N., Pemonge, M.H., Quassou, A., Sbay, H., & Petit, R.J. (2001). Frequent cytoplasmic exchanges between oak species that are not closely related: *Quercus suber* and *Q. ilex* in Morocco. *Molecular Ecology*, **10**, 2003-2012.

Beckstrom-Sternberg, G.M., Rieseberg, L.H., & Doan, D.K. (1991). Gene lineage analysis in populations of *Helianthus niveus* and *H. petiolaris* (Asteraceae). *Pl. Syst. Evol*. **175**, 125-138.

Bermingham, E., & Moritz, C. (1998). Comparative phylogeography: concepts and applications. *Molecular Ecology*, **7**, 367-369.

Bradbury, J.P. (1997). Sources of glacial moisture in Mesoamerica. *Quaternary International*, **44**, 97–110.

Breedlove, D. (1986). Listados florísticos de México IV. Flora de Chiapas. Instituto de Biologia, UNAM. México.

Bronson, C.L., Grubb, Jr. T.C., & Braun, M.J. (2003). A test of the endogenous and exogenous selection hypotheses for the maintenance of a narrow avian hybrid zone. *Evolution*, **57**, 630-637.

Cilibertii, P., De Jongi, H., Schouten, M.A., Verweji, P.A. & Heitmansi, W.R.B. (2009). Ecological niche models for four native cockroach species (*Blattaria*: Blattellidae: Ectobiinae) in the Netherlands. *Contributions to Zoology*, **78**, 149-156.

Corona, A.M., Toledo, V.H. & Morrone, J.J. (2007). Does the Trans-Mexican Volcanic Belt represent a natural biogeographical unit? An analysis of the distributional patterns of Coleoptera. *Journal of Biogeography*, **34**, 1008–1015.

Cottrell, J.E., Munro, R.C., Tabbener, H.E., Gillies, A.C.M., Forrest, G.I., Deans, J.D., & Lowe, A.J. (2002). Distribution of chloroplast DNA variation in British oaks (*Quercus robur* and *Q. petrea*): the influence of postglacial colonization and human management. *Forest Ecology and Management*, **156**, 181-195.

Craft, K.J., Ashley, M.V., & Koenig, W.D. (2002). Limited hybridization between *Quercus lobata* and *Quercus douglasii* (Fagaceae) in a mixed stand in central coastal California. *American Journal of Botany*, **89**, 1792-1798.

Craft, K.J., & Ashley, M.V. (2006). Population differentiation among three species of white oak in northeastern Illinois. *Canadian Journal of Forest Research*, **36**, 206–215.

Crepet, W.L., & Nixon, K.C. (1989). Earliest Megafossil evidence of *Fagaceae*: Phylogenetic and Biogeographic implications. *American Journal of Botany*, **76**, 842-855.

Curtu, A.L., Gailing, O., & Finkeldey, R. (2007). Evidence for hybridization and introgression within a species-rich oak (*Quercus* spp.) community. *BMC Evolutionary Biology*, **7**, 218-233.

De Greef, B., Triest, L., De Cuyper, B., & Slycken, J.V. (1998). Assesment of intraespecific variation in half-sibs of *Quercus petraea* (Matt.) Liebl. “plus” trees. *Heredity*, **81**, 284-290.

Dodd, R.S., & Kashani, N. (2003). Molecular differentiation and diversity among the California red oaks section Lobatae. *Theoretical and Applied Genetics*, **107**, 884-892.

Dodd, R.S., & Afzal-Rafii. Z. (2004). Selection and dispersal in a multispecies oak hybrid zone. *Evolution*, **58**, 261-269.

Dumolin-Lapegue, S., Demesure, B., Fineschi, S., Le Corre, V., & Petit, R.J. (1997). Phylogeographic structure of white oaks throughout the European continent. *Genetics*, **146**, 1475-1487.

Endler, J.A. (1977). Geographic variation, speciation, and clines. Princeton University Press, Princeton, New Jersey, USA.

Ferrusquia-Villafranca, I. (1993). Geology of Mexico: a synopsis. **En**: *Biological diversity of Mexico: origins and distribution* (eds. Ramamoorthy, T.P., Bye, R., Lot, A., & Fa, J.), The Oxford University Press, New York, pp 3-107.

Fitzpatrick, B.M., & Schaffer, H.B., (2004). Environment-dependent admixture dynamics in a tiger salamander hybrid zone. *Evolution*, **58**, 1282-1293.

González, R. (1993). La diversidad de los encinos Mexicanos. *Revista de la Sociedad Mexicana de Historia Natural*, **XLIV**, 125-142.

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

González-Rodríguez, A., Bain, J.F., Golden, J.L., & Oyama, K. (2005). Chloroplast DNA variation in the *Quercus affinis-Q. laurina* complex in Mexico: Geographical structure and associations with nuclear an morphological variation. *Molecular Ecology*, **13**, 3467-3476.

Graham, A. (993). Historical Factors and Biological Diversity in Mexico. Cap. 2. **En**: *Biological diversity of México: Origins and distribution* (Eds. Ramamoorthy, T.P.R., Bye, R., Lot, A., & Fa, J.) Oxford University Press. New York, pp 3-107.

Grant, P.R., & Grant, B.R. (1992). Hybridization in bird species. *Science*, **256**, 193-197.

Grivet, D., Deguilloux, M.F., Petit, R.J. & Sork, V.L. (2006). Contrasting patterns of historical colonization in white oaks (*Quercus* spp.) in California and Europe. *Molecular Ecology*, **15,** 4085-4093.

Gugerli, F., Walser, J.C., Dounavi, K., Holderegger, R., & Finkeldey, R. (2007). Coincidence of small-scale spatial discontinuities in leaf morphology and nuclear microsatellite variation of *Quercus petraea* and *Q. robur* in a mixed forest. *Annals of Botany*, **99**, 713-722.

Guisan, A. & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Hardin, J.W. (1975). Hybridization and introgression in *Quercus alba*. *Journal of the Arnold Arboretum*, **56**, 336.363.

Harrison, R.G., & Rand, D.M. (1989). Mosaic hybrid zones and the nature of species boundaries. **En**: *Speciation and its consequences* (Eds. Otte, D., & Endler, J.), Sinauer, Sunderland, MA, pp 110-133.

Hewitt, G.M. (1988). Hybrid zones-natural laboratories for evolutionary studies. *TREE*, **3**, 158-167.

Hewitt, G.M. (1989). The subdivision of species by hybrid zones. **En**: *Speciation and itsconsequences* (Eds. Otte, D., & Endler, J.), Sinauer, Sunderland, MA, pp 85–110.

Hewitt, G. (1999). Postglacial recolonization of European Biota. *Biological Journal of the Linnean Society*, **68**, 87-112.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Himrane, H., Camarero, J.J., & Gil-Pelegrín, E. (2004). Morphological and ecophysiological variation of the hybrid oak *Quercus subpyrenaica* (*Q. faginea* × *Q. pubescens*). *Trees*, **18**, 566–575.

Holman, J.E., Hughes, J.M., & Fensham, R.J. (2003). A morphological cline in *Eucalyptus*: a genetic perspective. *Molecular Ecology*, **12**, 3013-3025.

Howard, D.J., Preszler, R.W., Williams, J., Fenchel, S., & Boecklen, W.J. (1997). How discrete are oak species? Insights from a hybrid zone between *Quercus grisea* and *Quercus gambelii*. *Evolution*, **51**, 747-755.

Hugall, A., Moritz, C., Moussalli, A. & Stanisic, J. (2002). Reconciling paleodistribution models and comparative phylogeography in the wet tropics rainforest land snail *Gnarosophia bellendenkerensis*. *Proceedings of the National Academy of Sciences*, **99**, 6112–6117.

Ishida, T., Hattori, K., Sato, H., & Kimura, M.T. (2003). Differentiation and hybridization between *Quercus crispula* and *Q. dentate* (Fagaceae): insights from morphological traits, amplified fragment length polymorphism markers and leafminer composition. *American Journal of Botany*, **90**, 769-776.

Jakob, S.S., Martínez-Meyer, E. & Blattner, F.R. (2009). Phylogeographic analyses and paleodistribution modeling indicate Pleistocene in situ survival of *Hordeum* species (Poaceae) in Southern Patagonia without genetic or spatial restriction. *Molecular Biology and Evolution*, **26**, 907-923.

James, J.K., & Abbott, R.J. (2005). Recent, allopatric, homoploid hybrid speciation: the origin of *Senecio squalidus* (Asteraceae) in the British Isles from a hybrid zone on Mount Etna, Sicily. *Evolution*, **59**, 2533-2547.

Jaramillo-Correa, J.P., Beaulieu, J., Khasa, D.P. & Bousqet, J. (2009). Inferring the past from 12 the present phylogeographic structure of North American forest trees: seeing the forest for the 13 genes. *Canadian Journal of Forest Research*, **39**, 286-307.

Kaplan, Z., & Fehrer, J. (2007). Molecular evidence for a natural primary triple hybrid in plants revealed from direct sequencing. *Annals of Botany*, **99**, 1213-1222.

Kimball, S. (2008). Links between floral morphology and floral visitors along an elevational gradient in a *Penstemon* hybrid zone. *Oikos*, **117**, 1064-1074.

Knowles, L.L, & Maddison, W.P. (2002). Statistical phylogeography. *Molecular Ecology*, **11**, 2623–35

Lachniet, M.S., & Vázquez-Selem, L. (2005). Last Glacial Maximum equilibrium line altitudes in the circum-Caribbean (Mexico, Guatemala, Costa Rica, Colombia, and Venezuela). *Quaternary International*, **138**, 129-144.

Le Corre, V., & Kremer, A. (1998). Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping stone model. *Journal of Evolutionary Biology*, **11**, 495-512.

Lepais, O, Petit, R.J., Guichoux, E, Lavabre, J.E., Alberto, F., Kremer, A., & Gerber, S. (2009). Species abundance and direction of introgression in oaks. *Molecular Ecology*, **18**, 2228-2242.

Lexer, C., Kremer. A., & Petit, R.J. (2006). Shared alleles in sympatric oaks: recurrent gene flow is a more parsimonious explanation than ancestral polymorphism. *Molecular Ecology*, **15**, 2007-2012.

Lozano-García, M.S., & Xelhuantzi-López, M.S. (1997). Some problems in the late Quaternary pollen records of central Mexico and Zacapu. *Quaternary International*, **43**, 117-123.

Lozano-García, M.S., & Ortega-Guerrero, B. (1997). Late Quaternary environmental changes of the central part of the Basin of Mexico; correlation between Texcoco and Chalco Basins. *Review of Palaeobotany and Palynology*, **99**, 77–93.

Lozano-García, S., Sosa-Nájera, S., & Sugiera, Y. (2005). 23, 000 yr of vegetation history of the Upper Lerma, a tropical high-altitude basin in Central Mexico. Quaternary *Research*, **64**, 70-82.

Magni, C.R., Ducousso, A., Caron, H., Petit, R. J., & Kremer, A. (2005). Chloroplast DNA variation of *Quercus rubra* L. in north America and comparison with other Fagaceae. *Molecular Ecology*, **14**, 513-524.

Magri, D., Vendramin, G.G., & Comps, B. (2006). A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, **171**, 199–221

Manos, P.S. & Fairbrothers, D.E. (1987). Allozyme variation in populations of six Northeastern American Red Oaks (Fagaceae: *Quercus* subg. Erythrobalanus). *Systematic Botany*, **12**, 365–373.

Manos, P.S., Doyle, J.J., & Nixon, K.C. (1999). Phylogeny, Biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). *Molecular Phylogenetics and Evolution*, **3**, 333.349.

Manos, P.S. & Stanford, A.M. (2001). The Historical Biogeography of Fagaceae: Tracking the Tertiary History of Temperate and Subtropical Forests of the Northern Hemisphere. *Journal of Plant Sciences*, **162**, S77-S93.

Manni, F., Guerard, E. & Heyer, E. (2004). Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by ‘‘Monmonier’s algorithm’’. *Human Biology*, **76**, 173-190.

Manchester, S.R. (1983). Eocene fruits, woods and leaves of the Fagaceae from the Clarno formation of Oregon. *American Journal of Botany*, **70**, 74-78.

Marsico, T.D., & Hellmann, J.J. (2009). Patterns of seed dispersal and pollen flow in *Quercus garryana* (Fagaceae) following post-glacial climatic changes. *Journal of Biogeography*, **36**, 929-941.

Martínez-Hernández, E. (1992). La vegetación y el clima de México. Historia del Cretácico tardío al Terciario basada en evidencias palinológicas. *Ciencia*, **43**, 69-72.

Martínez-Hernández, E., & Lozano-García, S. (1996). Volcanismo, Cambios climáticos y sucesión de floras, durante el Neogeno de la Cuenca de México. IX Coloquio de Paleobotánica y Palinología, p. 48-49.

Matos, J.A. & Schaal, B.A. (2000). Chloroplast evolution in the *Pinus montezumae* complex: a coalescent approach to hybridization. *Evolution*, **54**, 1218–1233.

McVaugh, R. (1974). Flora novo-galiciana (*Quercus*). Contributions from the University of Michigan Herbarium, *Ann Arbor, Michigan*, **12**, 1-93.

Metcalfe, S.E., O'Hara, S.L., Caballero, M., & Davies, S.J. (2000). Records of Late Pleistocene Holocene climatic change in Mexico: a review. *Quaternary Science Reviews*, **19**, 699-721.

McNeish, R.S., Peterson, F.A. & Neely, J.A. (1972) The archaeological reconnaissance. **En**: *The prehistory of the Tehuacan Valley. Vol. 5. Excavations and reconnaissance* (ed. by R.S. McNeish), pp. 496–504. University of Texas Press, Austin, TX.

Metcalfe, S.E. (2006). Late Quaternary environments of the northern deserts and central transvolcanic belt of Mexico. *Annals of the Missouri Botanical Garden*, **93**, 258-273.

Moreno-Letelier, A., & Piñero, D. (2009). Phylogeographic structure of *Pinus strobiformis* Engelm. across the Chihuahuan Desert filter-barrier. *Journal of Biogeography*, **36**, 121-131.

Moore, W.S., & Price, J.T. (1993). Nature of selection in the northern flicker hybrid zone and its implications for speciation theory. **En**: *Hybrid zones and the evolutionary process* (eds. Harrison, R.G.), Oxford University Press, Oxford, pp 196-255.

Muir, G., & Schlotterer, C. (2005). Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Molecular Ecology*, **14**, 549-561.

Nixon, K.C. (1993). The genus Quercus in Mexico. **En**: *Biological diversity of Mexico: origins and distribution* (eds. Ramamoorthy, T.P., Bye, R., Lot, A., & Fa, J.), Oxford University Press. New York USA, pp 447-458.

Palmer, E.J. (1948). Hybrid oaks of North America. *Journal of the Arnold Arboretum*, **29**, 1-48.

Peñaloza-Ramírez, J.M., González-Rodríguez, A., Mendoza-Cuenca, L., Caron, H., Kremer, A., & Oyama, K. (2010)a. Interspecific gene flow in a multispecies oak hybrid zone in the Sierra Tarahumara of Mexico. *Annals of Botany*, **105**, 389-399.

Peñaloza-Ramírez, J.M., González-Rodríguez, A., Rocha-Ramírez, V., Rodríguez-Correa, H., Martínez Meyer, E., & Ken Oyama (2010)b Phylogeography and ecological niche modelling reveal Quaternary population history in the widespread Mexican red oak *Quercus castanea* Née. Under review in *Journal of Biogoegraphy*.

Peterson, A.T., Papes, M., & Eaton, M. (2007). Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography*, 30, 550-560.

Petit, R.J., Csaikl, U.M., Bordacs, S., Burg, K., Coart, E., Cottrell, J., Van Dam, B., Deans J.D., Dumolin-Lapegue, S., Fineschi, S., Finkeldey, R., Gillies, A., Glaz, I., Goicoechea, P.G., Jensen, J.S., Konig, A.O., Lowe, A.L., Madsen, S.F., Matyas, G., Munro, R.C., Olalde, M., Pemonge, M.H., Popescu, F., Slade, D., Tabbener, H., Taurchini, D., de Vries, S.G.M., Ziegenhagen, B., & Kremer, A. (2002). 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.

Petit, R.J., Bodénès, C., Ducousso, A., Roussel, G., & Kremer, A. (2003). Hybridization as a mechanism of invasion in Oaks. *New Phytologist*, **161**, 151–164

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006). Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.

Piperno, D.R., Moreno, J.E., Iriarte, J., Holst, I., Lachniet, M., Jones, J.G., Ranere A.J. & Castanzo, R. (2007) Late Pleistocene and Holocene environmental history of the Iguala Valley, Central Balsas Watershed of Mexico. *Proceedings of the National Academy of Sciences*, **104**, 11874–11881.

Raudnitschka, D., Hensen, I., & Oberprieler, C. (2007). Introgressive hybridization of *Senecio hercynus* and *S. ovatus* (Compositeae, Senecioneae) along an altitudinal gradient in Harz national Park (Germany). *Systematics and Biodiversity*, **5**, 333-344.

Richards, C.L., Carstens, B.C. & Knowles, L.L. (2007). Distribution modeling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographic hypotheses. *Journal of Biogeography*, **34**, 1833-1845.

Rogers, A.R. & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552-569.

Rodríguez-Banderas, A., Vargas-Mendoza, C.F., Buonamici, A. & Vendramin, G.G. (2009). Genetic diversity and phylogeographic analysis of *Pinus leiophylla*: a post-glacial range expansion. *Journal of Biogeography*, **36**, 1807-1820.

Rossotti, A., Ferrari L., López-Martínez, M., & Rosas-Elguera, J. (2002). Geology of the boundary between the Sierra Madre Occidental and the Trans-Mexican Volcanic Belt in the Guadalajara region, western Mexico. *Revista Mexicana de Ciencias Geológicas*, **19**,1-15.

Rieseberg, L.H., & Wendell, J.F. (1993). Introgression and its consequences in plants. **En**: *Hybrid zones and the evolutionary process* (Eds. Harrison, R.G.), New York: Oxford University Press, pp 70–109.

Rieseberg, L.H. (1997). Hybrid origins of plant species. *Annual Review of Ecology, Evolution, and Systematics*, **28**, 359-389

Rzedowski, J. (1978). Vegetación de México. CONABIO, Ed. Limusa. México.

Rzedowski, J. (1993). Diversity and origins of the phanerogamic flora of Mèxico cap. 3. **En**: *Biological diversity of Mexico: origins and distribution* (Eds. Ramamoorthy, T.P., Bye, R., Lot, A., & Fa J.), Oxford University Press. New York USA, pp 447-458.

Schaal, B.A., Hayworth, D.A., Olsen, K.M., Rauscher, J.T., & Smith, W.A. (1998). Phylogeographic studies in plants: Problems and prospects. *Molecular Ecology*, **7**, 465-474.

Schneider, S. & Excoffier, L. (1999). Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.

Seehausen, O. (2004). Hybridization and adaptative radiation. *TREE*, **19**, 198-207.

Silliman, F.E., & Leisner, R.S. (1958). An analysis of a colony of hybrids Oaks. *American Journal of Botany*, **45**, 730-736.

Spellenberg, R. (1995). On the hybrid nature of *Quercus basaseachicensis* (Fagaceae, sect. *Quercus*). *Sida*, **16**, 439-437.

Suter, M. (1984). Cordilleran deformation along the eastern edge of Valles San Luis Potosí carbonated platform Sierra Madre Oriental fold-Thrust belt, eastern central Mexico. *Bulletin of Geological Society of America*, **95**, 1387-1397.

Steinhoff, S. (1998). Controlled crosses between pedunculate and sessile oak: results and conclusion. *Allgemeine Forst- und Jagdzeitung*, **169**, 163–168.

Tovar-Sanchéz, E., & Oyama, K. (2004). Natural hybridization and hybrid zones between *Quercus crassifolia* and *Quercus crassipes* (Fagaceae) in Mexico: morphological and molecular evidence. *American Journal of Botany*, **91**, 1352-1363.

Trelease, W. (1924). The American Oaks. *Memories of the National Acaddemy of Sciences*, **20**, 1-255.

Turelli, M., Barton, N.H., & Coyne, J.A. (2001). Theory and speciation. *TREE*, **16**, 330-343.

Valbuena-Carabaña, M., Gonzalez-Martinez, S.C., Hardy, O.J., & Gil, L. (2007). Fine-scale genetic structure in mixed oak stands with different levels of hybridization. *Molecular Ecology*, **16**, 1207-1219.

Valencia, A.S. (1995). Contribución al conocimiento del género *Quercus* (Fagaceae) en el estado de Guerrero, México. Coordinación de Servicios Editoriales, Facultad de Ciencias, UNAM.

Valencia, S. (2004). Diversidad del género *Quercus* (Fagaceae) en México. *Boletin de la Sociedad Botanica de Mexico*, **75**, 33-53.

Vila, M., & D'Antonio, C.M. (1998). Hybrid vigor for clonal growth in *Carpobrotus* (Aizoaceae) in coastal California. *Ecological Applications*, **8**, 1196–1205.

Waltari, E., Hijmans, R.J., Peterson, A.T., Nyari, A.S., Perkins, S.L., Guralnick, R.P., (2007). Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS One*, **2**:e563

Widmer, A., & Lexer, C. (2001). Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology & Evolution*, **16**, 267–269.

Williams, D.G., & Ehleringer, J.R. (2000). Carbon isotope discrimination and water relations of oak hybrid populations in southwestern Utah. *Western North American Naturalist,* **60**,121-129.

Williams, J.H., Boecklen, W.J., & Howard, D.J. (2001). Reproductive processes in two oak (*Quercus*) contact zones with different levels of hybridization. *Heredity*, **87**, 680–690.