A peer-reviewed version of this preprint was published in PeerJ on 21 April 2016.

View the peer-reviewed version (peerj.com/articles/1897), which is the preferred citable publication unless you specifically need to cite this preprint.

Plastome data reveal multiple geographic origins of *Quercus* Group Ilex

Marco Cosimo Simeone, Guido W Grimm, Alessio Papini, Federico Vessella, Simone Cardoni, Enrico Tordoni, Roberta Piredda, Alain Franc, Thomas Denk

Nucleotide sequences from the plastome are currently the main source for assessing taxonomic and phylogenetic relationships in flowering plants and their historical biogeography at all hierarchical levels. One exception is the large and economically important genus *Quercus* (oaks). Whereas differentiation patterns of the nuclear genome are in agreement with morphology and the fossil record, diversity patterns in the plastome are at odds with established taxonomic and phylogenetic relationships. However, the extent and evolutionary implications of this incongruence has yet to be fully uncovered.

The DNA sequence divergence of four Euro-Mediterranean Group Ilex oak species (*Quercus ilex* Qcoccifera QaucheriQalnifolia ilex L., *Q. coccifera* L., *Q. aucheri* Jaub. & Spach., *Q. alnifolia* Poech.) was explored at three chloroplast markers (rbcL, trnK-matK, trnH-psbA). Phylogenetic relationships were reconstructed including worldwide members of additional 55 species representing all *Quercus* subgeneric groups. Family and order sequence data were harvested from gene banks to better frame the observed divergence in larger taxonomic contexts. We found a strong geographic sorting in the focal group and the genus in general that is entirely decoupled from species boundaries. Main plastid haplotypes shared by distinct oak lineages from the same geographic region and high plastid diversity in members of Group Ilex are indicative for a polyphyletic origin of their plastomes. The results suggest that incomplete lineage sorting and repeated phases of unidirectional introgression among ancestral lineages of Group Ilex and two other main Groups of Eurasian oaks (Cyclobalanopsis and Cerris) caused this complex pattern. Comparison with the current phylogenetic synthesis also suggests an initial high- versus mid-latitude biogeographic split within *Quercus*. High plastome plasticity of Group Ilex reflects geographic area disruptions, possibly linked with high tectonic activity of past and modern distribution ranges, that did not leave imprints in the nuclear genome of modern species and infrageneric lineages.
Title: Plastome data reveal multiple geographic origins of Quercus Group Ilex

Authors
Marco Cosimo Simeone 1,*, Guido W. Grimm 2, Alessio Papini 3, Federico Vessella 1, Simone Cardoni 1, Enrico Tordoni 4, Roberta Piredda 5, Alain Franc 6,7, Thomas Denk 8

Affiliations
1 = Department of Agricultural and Forestry Science (DAFNE), Università degli Studi della Tuscia, via S. Camillo de Lellis, 01100 Viterbo, Italy
2 = Department of Palaeontology, University of Wien, Althanstrasse 14 (UZA II), 1090 Wien, Austria
3 = Dipartimento di Biologia Vegetale, Università degli studi di Firenze, via La Pira 4, 50121 Firenze, Italy
4 = Department of Life Science, Università degli studi di Trieste, via L. Giorgieri 10, 34127 Trieste, Italy
5 = Stazione Zoologica Anton Dohrn, Villa Comunale, 8012 Napoli, Italy
6 = INRA, UMR BIOGECO-1202, 69 route d’Arcachon, F-33612 Cestas, France
7 = Université Bordeaux, UMR BIOGECO-1202, Avenue des Facultés, F-33405 Talence, France
8 = Department of Palaeobiology, Swedish Museum of Natural History, Box 50007, 10405 Stockholm, Sweden

Corresponding Author
Dr. Marco Cosimo Simeone
Department of Agricultural and Forestry Science (DAFNE), Università degli Studi della Tuscia,
via S. Camillo de Lellis, 01100 Viterbo, Italy
E-Mail address: mcsimeone@unitus.it
Abstract

Nucleotide sequences from the plastome are currently the main source for assessing taxonomic and phylogenetic relationships in flowering plants and their historical biogeography at all hierarchical levels. One exception is the large and economically important genus *Quercus* (oaks). Whereas differentiation patterns of the nuclear genome are in agreement with morphology and the fossil record, diversity patterns in the plastome are at odds with established taxonomic and phylogenetic relationships. However, the extent and evolutionary implications of this incongruence has yet to be fully uncovered. The DNA sequence divergence of four Euro-Mediterranean Group Ilex oak species (*Quercus ilex* L., *Q. coccifera* L., *Q. aucheri* Jaub. & Spach., *Q. alnifolia* Poech.) was explored at three chloroplast markers (rbcL, trnK-matK, trnH-psbA). Phylogenetic relationships were reconstructed including worldwide members of additional 55 species representing all *Quercus* subgeneric groups. Family and order sequence data were harvested from gene banks to better frame the observed divergence in larger taxonomic contexts. We found a strong geographic sorting in the focal group and the genus in general that is entirely decoupled from species boundaries. Main plastid haplotypes shared by distinct oak lineages from the same geographic region and high plastid diversity in members of Group Ilex are indicative for a polyphyletic origin of their plastomes. The results suggest that incomplete lineage sorting and repeated phases of unidirectional introgression among ancestral lineages of Group Ilex and two other main Groups of Eurasian oaks (Cyclobalanopsis and Cerris) caused this complex pattern. Comparison with the current phylogenetic synthesis also suggests an initial high- versus mid-latitude biogeographic split within *Quercus*. High plastome plasticity of Group Ilex reflects geographic area disruptions, possibly linked with high tectonic activity of
past and modern distribution ranges, that did not leave imprints in the nuclear genome of modern species and infrageneric lineages.

**Keywords:** Fagaceae, Mediterranean, Plastome polyphyly, Ancient introgression, Incomplete lineage sorting, Decoupled phylogenies

**Introduction**

*Quercus* L. (oaks) is among the most ecologically diverse and economically important extratropical tree genera in the northern hemisphere (Govaerts and Frodin, 1998). *Quercus* is the largest genus in the order Fagales, comprising ca. 400–500 species. Oaks are concentrated in the Americas (Groups Quercus, Lobatae and Protobalanus; Flora of North America Editorial Committee, 1997) and Southeast Asia and southern China (Group Cyclobalanopsis; Flora of China Editorial Committee, 1999). In contrast, a relatively lower number of species can be found in western Eurasia and the Mediterranean (Groups Ilex and Cerris; Kubitzki, 1993; Menitsky, 2005). The six major infrageneric lineages of *Quercus* occur from the tropics to the high mountains of the temperate zone and to the boreal continental, cold temperate regions (Denk and Grimm, 2010). The northern limit of oaks in North America and Eurasia coincides with the border of $D_{fb}$ to $D_{fc}$ and $D_{wb}$ to $D_{wc}$ climates, snow climates with warm versus cool summers (Köppen, 1936; Kottek et al., 2006; Peel et al., 2007).

Recent molecular phylogenetic studies at and below the genus level focussed on the nucleome of oaks (Oh and Manos, 2008; Denk and Grimm, 2010; Hipp et al., 2014; Hubert et al., 2014). These studies consistently recovered two main lineages, the ‘New World Clade’ comprising the white oaks (Group Quercus), red oaks (Group Lobatae) and golden-cup oaks (Group Protobalanus), and the ‘Old World Clade’ consisting of the cycle-cup oaks (Group...
Cyclobalanopsis), the Ilex oaks (Group Ilex) and the Cerris oaks (Group Cerris). Evidence from nuclear markers and the fossil record suggests that the initial split in the ‘New World Clade’ was pre-Oligocene between the lineages leading to Group Lobatae and Group Protobalanus/Quercus (Bouchal et al., 2014; Hubert et al., 2014; Grimsson et al., 2015). This early radiation of the Quercus/Protobalanus lineage left its imprints in the molecular signatures of the few modern species of Group Protobalanus and two narrow endemic white oak species, *Quercus pontica* (north-eastern Turkey, south-western Georgia; Denk and Grimm, 2010) and *Q. sadleriana* (California; Hubert et al., 2014). Within the ‘Old World Clade’, the major split was established between the evergreen Groups Cyclobalanopsis and Ilex during the Eocene/Oligocene, whereas the chiefly temperate Group Cerris is suggested to have evolved (‘budded’) from a Group Ilex stock, possibly in Europe, not before the earliest Miocene (Denk and Grimm, 2009; Kmenta, 2011; Hubert et al., 2014; Velitzelos et al., 2014).

Nuclear amplicon data sets have also contributed to resolve the circumscription of these six groups and to delineate some intergroup and interspecies relationships (López de Heredia et al., 2007; Pearse and Hipp, 2009; Denk and Grimm, 2010; Hubert et al., 2014); well-resolved within-lineage relationships were recently obtained from phylogenomic data in the genetically least-diverged, but species-rich Group Quercus (Hipp et al., 2014). Nucleome-based studies, therefore, clearly indicate a strong correlation between morphology/speciation and nuclear differentiation in oaks. In contrast, oak plastid haplotypes are extensively shared between groups of species (Whittemore and Schaal, 1991; Belahbib et al., 2001; Manos and Stanford, 2001; Petit et al. 2002; Kanno et al., 2004; López de Heredia et al., 2007; Okaura et al., 2007; Neophytou et al., 2010; Gugger and Cavender-Bares 2013). Notably, this was also observed in other genera of Fagaceae such as *Fagus* (Fujii et al. 2002; Lei et al. 2012; Zhang et al. 2013b) and *Lithocarpus*.
(Cannon and Manos, 2003), and other Fagales such as the northern hemispheric *Carya* (Juglandaceae; Zhang et al., 2013a) and the South American *Nothofagus* (Nothofagaceae; Acosta and Premoli, 2010; Premoli et al., 2012). Plastomes of this large group of long-lived woody plants appear to retain molecular signatures of evolutionary events that cannot be investigated when considering the nuclear DNA alone (e.g., Cavender-Bares et al., 2011; Premoli et al., 2012). As such, they can provide additional information to complement hypotheses on diversification and speciation processes. However, the extent and evolutionary implications of nuclear-plastome incongruence in *Quercus* have yet to be fully uncovered.

Testing the potential of DNA barcoding in western Eurasian oaks, Simeone et al. (2013) recently found puzzling diversity in the plastid haplotypes of samples belonging to Group Ilex. In addition to interspecific haplotype sharing, paraphyly to Groups Cerris and Quercus and an underlying geographic partitioning was suggested. In the present study, we increased the geographic coverage and taxon sampling to explore the complex patterns of plastome evolution in *Quercus* Group Ilex. This species group is today confined to extra-tropical regions of Eurasia, spanning from arid Mediterranean maquis to high mountain and sub-alpine Himalayan forests and thickets, and to subtropical forests of SE Asia. Group Ilex includes some 35 evergreen, mostly sclerophyllous taxa, whose taxonomy is still controversial (see Table 1) and biogeographic history is not yet well understood (Menitsky, 2005; Denk and Grimm, 2010). In this work, we compiled plastid sequence data for 81 accessions of 20 oak taxa of Group Ilex. The main sampling effort was put into the four species currently occurring in the Mediterranean and adjacent regions in North Africa (Atlas Mountains) and northern Turkey (Black Sea region): the widespread *Quercus ilex* L. and *Q. coccifera* L., and the two East Mediterranean narrow endemics *Q. aucheri* Jaub. & Spach. and *Q. alnifolia* Poech. Data for additional 56 individuals of
ca. 40 species were also produced to integrate all subgeneric *Quercus* groups and their worldwide geographic distribution. Additionally, Fagales data sets were harvested from gene banks to allow interpretation of the observed divergence in the plastid markers within a larger taxonomic frame. Our objectives were: (1) to assess the extent of plastome diversity in the Euro-Mediterranean focal group; (2) to outline key phylogeographic patterns within *Quercus* Group Ilex; (3) to establish major evolutionary steps for the differentiation of the ‘Old World Clade’.

**Material and methods**

**Plant Material, DNA amplification and analyses**

Our analysis included 59 individuals of the four Mediterranean *Quercus* Group Ilex species (Table S1) covering their entire range in North Africa and western Eurasia. Additionally, 22 individuals of 16 Asian species of Group Ilex were analysed. The final dataset also included all species of the western North American Group Protobalanus (five species, 10 individuals), 16 species of Group Quercus (20 individuals, from North America and Eurasia), five species of the East Asian Group Cyclobalanopsis (11 individuals), seven species of the American Group Lobatae (eight individuals), and six species of Group Cerris (seven individuals). The outgroup set was represented by one sample each of the monotypic genera *Notholithocarpus* and *Chrysolepis* (western North America) and one species each of *Castanea* and *Castanopsis* [(NCBI GenBank accessions HQ336406 (complete plastid genome of *C. mollissima*), JN044213, JF941179, FJ185053). Based on their genetic (plastid) signatures these genera are the closest relatives of *Quercus* within the Fagaceae (Manos et al., 2008). For voucher information and accession numbers see Table S1. The molecular analyses included three plastid DNA regions: a
part of the *rbcL* gene, the *trnH-psbA* intergenic spacer and a portion of the *trnK/matK* region (3’ intron and partial gene). These markers were chosen based on the variability displayed in previous works (e.g. Manos et al., 2001; Okaura et al., 2007; Simeone et al., 2013) and on the high number of their sequences available on GenBank. DNA extractions, primers and PCR protocols were the same as in (Piredda et al., 2011) and (Simeone et al., 2013). Sequencing was performed at Macrogen (http://www.macrogen.com); electropherograms were edited with CHROMAS 2.3 (http://www.technelysium.com.au) and checked visually.

**Statistical tools**

The diversity of the investigated regions was evaluated with MEGA 5.2 (Tamura et al., 2011) and DNAsp 5.1 (Librado and Rozas, 2009). For comparisons of divergence patterns across all Fagales, available data in gene banks were processed using GBK2FAS (Göker et al., 2009); multiple sequence alignments were done with MAFFT v.7 (Katoh and Standley, 2013) using default setting and checked by eye to remove inconsistencies and erroneous sequences (taxa and sequence numbers in Supporting Information). To minimise the effect of alignment gaps, and since we were primarily interested in assessing intra- and intergeneric divergence, alignments included only subsets of the Fagales: 1) Nothofagaceae (data covering all four genera); 2) Fagaceae (10 genera including *Quercus*); 3) Betulaceae-*Ticodendron-Casuarinaceae* (11 genera); 4) Juglandaceae (9 genera); 5) Myricaceae (4 genera). Pairwise distance matrices (uncorrected *p*-distance, K2P, HKY, GTR+Γ) for each marker were calculated with PAUP* 4.0 (Swofford, 2002). Minimum intra-specific and minimum/maximum inter-specific distances (calculated with G2CEF; Göker and Grimm, 2008) within and between genera, subgenera in the case of *Fagus*, and infragenic groups in case of *Quercus*, are listed in Table S2.
Phylogenetic analyses

Multiple sequence alignments for the focal group were obtained with CLUSTALW 1.81 (Thompson et al., 1994) and checked by eye. The matrices were concatenated with the Python programme COMBINEX2_0.PY (PYTHON v. 2.6.4; BIOPYTHON 1.57).

Maximum likelihood trees were inferred with GARLI (Zwickl, 2006; run on the CIPRES portal, http://www.phylo.org/sub_sections/portal/) using four data partitions (rbcL and matK codons, trnK intron and trnH-psbA spacer). MRMODELTEST 2.0 (Nylander, 2004) and the Akaike Information Criterion (AIC; Akaike, 1974) were used to decide on the best-fitting substitution model for each partition.

MRMODELTEST2.0 results were also used for setting up Bayesian inference, performed with MrBayes 3.4b4 (Ronquist and Huelsenbeck 2003; Ronquist et al. 2012). RAxML v. 7.0.4 (Stamatakis et al., 2008) was used for calculating maximum likelihood bootstrap support (1000 replicates). Trees were edited with FigTree 1.3.1(Rambaut, 2014) and Mesquite v. 2.75 (Maddison and Maddison, 2011). Median-joining (MJ) haplotype networks were inferred with NETWORK 4.6.1.1 (http://www.fluxus-engineering.com/) for each gene region (rbcL, trnK/matK, trnH-psbA), treating gaps either as missing or 5th state. MJ algorithm was invoked with default parameters (equal weight of transversion/transition), in order to handle large datasets and multistate characters.

Primary data, analysis and supplementary files (S1–S3) are provided for anonymous download in an online supporting archive hosted at www.palaeogrimm.org/data/Smn15_OSA.zip

Results
Levels of intra- and interspecies plastome divergence in Quercus

The entire dataset included 423 plastid DNA sequences (141 samples, three markers each). Table 2 shows that \textit{trnH-psbA} was the most variable marker region (a 34-bp inversion occurring in approximately 50% of the samples was not considered). The least variable region was \textit{rbcL}, as expected. No indels were found in the \textit{rbcL} and \textit{matK} coding regions. The combined cpDNA dataset (\textit{trnH-psbA}, \textit{trnK/matK}, \textit{rbcL}) resulted in an alignment of 2082 characters (sites), of which 122 were variable (thereof 72 parsimony-informative; gaps not considered). The alignment had a nucleotide diversity of 0.0056 and included 74 different haplotypes of which 50 were unique (restricted to a single accession). As a result, the overall haplotype diversity was high ($H_d = 0.978 \pm 0.005$). With gaps considered, the number of haplotypes increased to 110, of which 89 were unique ($H_d = 0.9939$).

In general, the infrageneric divergence calculated in \textit{Quercus} is comparable to that found in other genera of the Fagaceae and Betulaceae, and higher than in Juglandaceae (Table 3). All three gene regions allow distinguishing the generic affinity of an oak individual; the same haplotype may be shared by several or many oak species (usually within the same infrageneric group; Table 3), but not with other genera of the Fagaceae.

At the infrageneric level in \textit{Quercus}, minimal inter-species distances can be zero for all three markers and within all infrageneric groups. Notably, maximal inter-species distances within infrageneric groups of \textit{Quercus} can reach or even exceed the level of inter-generic differentiation in Fagaceae (e.g. between \textit{Notholithocarpus, Lithocarpus, Castanopsis, Castanea, Chrysolepis}), Juglandaceae and Myricaceae. The maximum intra-specific distance found in Mediterranean
individuals of *Quercus* Group Ilex equals the maximum inter-specific divergence found within this group.

*Phylogenetic placement of Mediterranean Quercus Group Ilex plastid haplotypes*

Individuals of the Mediterranean species of *Quercus* Group Ilex cluster in three well supported distinct clades (Fig. 1). The first clade (‘Euro-Med’) accommodates most accessions of *Q. ilex* and *Q. coccifera*. In the second clade (‘Cerris-Ilex’), accessions of *Q. ilex*, *Q. coccifera*, and one of the five samples of *Q. aucheri* group together with all representatives of *Quercus* Group Cerris and two Himalayan-East Asian species of Group Ilex. Sister to this clade are the three representatives of the single Japanese species of Group Ilex (*Q. phillyraeoides*). In the third clade (West Asia-Himalaya-East Asia; ‘WAHEA’) the remaining specimens of *Q. aucheri* form a subclade along with the Cypriote endemic *Q. alnifolia*, and several Eastern Mediterranean *Q. coccifera*. The second, more divergent and poorly supported subclade comprises two western Himalayan species (*Q. baloot*, *Q. floribunda*), two individuals of Himalayan-East Asian species of *Quercus* Group Ilex, and one Central China accession of a Cyclobalanopsis member (*Q. oxyodon*) sympatric with many group Ilex oaks, including *Q. semecarpifolia*, *Q. leucotrichophora*, *Q. floribunda* (Menitsky, 2005). In contrast to Group Ilex, all other infrageneric groups show relatively high chlorotypic coherence, usually forming clades or grouped within the same subtree. The actual root of the tree is obscured; representatives of *Castanea*, *Castanopsis*, and *Notholithocarpus*/*Chrysolepis* that could be used as putative outgroups are placed in different subtrees.
Evolutionary significance of plastid haplotypes in western Mediterranean oaks of Quercus

Group Ilex

The MJ network for the plastid region with the highest overall variability (trnH-psbA, only length-homogenous parts considered; Fig. 2) highlights the evolutionary significance of the three main haplotypes, ‘Euro-Med’, ‘Cerris-Ilex’, and ‘WAHEA’. Three main clusters differ by a minimum of two conserved mutations: 1) Group Quercus, Protobalanus and Lobatae (‘New World Oaks’); 2) individuals with ‘Euro-Med’ haplotypes; 3) individuals with ‘Cerris-Ilex’ and ‘WAHEA’ haplotypes, representatives of Group Cerris and East Asian species of Group Ilex and Group Cyclobalanopsis (‘Old World Oaks’). In general, haplotypes (File S3 includes MJ-networks for the other three regions, rbcL gene, matK gene, 3’ trnK intron,) found in the western Eurasian members of Group Ilex represent unique or ancestral variants. Unique haplotypes of Group Cerris are directly derived from the Group Ilex or shared ‘Cerris-Ilex’ haplotypes. Haplotypes of Group Cyclobalanopsis are identical to or can be derived from East Asian members of Group Ilex. The graphs further highlight a close relationship of haplotypes of Chrysolepis and Notholithocarpus with those of the ‘New World’ oaks; those of Castanea and Castanopsis can be derived from the ‘Old World’ oaks basic type.

Figures 1 and 2 clearly illustrate that differentiation in the plastid sequences of Quercus (and related Fagaceae) is independent from the formation or at least the genetic homogenization (lineage sorting) of the modern clades.

Phylogeographic structure in Quercus Group Ilex

Haplotypes forming the ‘Euro-Med’, ‘Cerris-Ilex’ and ‘WAHEA’ lineages are geographically sorted. The phylogenetically isolated ‘Euro-Med’ haplotypes are encountered in the western
Mediterranean populations of *Q. ilex* and *Q. coccifera* (North Africa, Iberia, Southern France, Italy), along the Adriatic coast and into Central Greece (Fig. 3). Also included here are isolated populations of *Q. ilex* from Crete and the southern Black Sea coast. ‘Cerris-Ilex’ and ‘WAHEA’ haplotypes are confined to the eastern Mediterranean region. ‘Cerris-Ilex’ haplotypes are found in the Aegean region (*Q. ilex*, *Q. coccifera* and *Q. aucheri* individuals) and replaced by ‘WAHEA’ haplotypes (*Q. coccifera*, *Q. aucheri*, *Q. alnifolia*) in south-western Turkey and extending to the east (Levant; Fig. 3). The ‘Cerris-Ilex’ type is also found in the *Q. coccifera* individual from northern Turkey, representing the north-easternmost population of this species.

**Discussion**

Despite resolution issues due to weak signals regarding intergeneric relationships, all data on Fagaceae show a deep incongruence between nuclear and plastid data. Nuclear phylogenies unambiguously point towards an inclusive common origin of all oaks, i.e. a monophyletic (s. str.) genus *Quercus* (Oh and Manos, 2008; Denk and Grimm, 2010; Hubert et al., 2014). At the same time plastid data repeatedly failed to resolve all oaks as one clade (Manos et al., 2008; this study). Instead, a split emerges (with varying support) between the North American *Notholithocarpus* and North American/northern temperate clade of oaks, the 'New World Oaks', and the Eurasian *Castanea*, *Castanopsis* and oak lineages, the 'Old World oak' clade; an observation that holds independent from the exact placement of the root in a plastid tree. If we accept the monophyly of the genus *Quercus*, which is backed also by morphology and evidence from the fossil record, haplotypes of *Castanea/Castanopsis* and *Notholithocarpus* that group with the 'New World' and 'Old World' oaks, respectively, can hence only be the result of
incomplete lineage sorting during the formation of the modern genera. In addition, the plastid
genepool of the earliest oaks must have shown a genetic gradient that was to some extent caused
by biogeographic patterns. Although it is impossible to pinpoint the place of origin of oaks, it is
clear that the ancient oaks must have been widespread, allowing them to pick up and propagate
genetic signatures inherited from their common ancestors with *Notholithocarpus, Castanea*
and Castanopsis. Geographic signatures in two Mediterranean species of Group Ilex, *Q. ilex* and
*Q. coccifera*, are discussed in the following.

Major trends of plastome differentiation

The overall low genetic intra- and intertaxonomic (intragenic lineages, genera) distances
suggest low evolutionary rates for the chloroplast genomes of Fagales, at least at the examined
loci. However, the data coverage is far from sufficient for most genera and families to precisely
evaluate the plastome potential variation within this plant group. In Fagaceae, a comparison with
the (genetically) more diverse Nothofagaceae and Betulaceae families reveals that haplotype
variation at the *trnH-psbA* locus can be sufficiently high to allow inferences at the
phylogeographic and systematic level (see Premoli et al., 2012; Grimm and Renner, 2013). In
analogy, haplotypes of intragenic lineages of *Quercus* differ in this marker. Furthermore, a
geographic pattern is evident for the most widely sampled groups. Groups Ilex, Lobatae and
Quercus appeared the most variable, whereas Group Cerris exhibited the lowest differentiation
rates. Interestingly, interclade differentiation among all *Quercus* groups equalled or exceeded the
values scored by the four outgroup genera (*Castanea, Castanopsis, Notholithocarpus* and
*Chrysolepis*). As a consequence, the outgroup taxa appear scattered across the tree, rather than
being culled in a distinct subtree, rendering the plastome of *Quercus* ‘non-monophyletic’.
Outgroup selection as a potential source of topological ambiguity was already pointed out by Hubert et al. (2014; 108 oak taxa, eight nuclear markers). Ambiguous relationships among this group of genera independently of the strength of the obtained phylogenetic signal were also suggested by a recent study on Fagales (based on molecular, fossil and reproductive syndromes analyses), which resolved the majority of inter-generic relationships in each family except in the Quercoideae group making Castanopsis and Quercus non-monophyletic (Xiang et al., 2014).

Plastid phylogeny does not conform to the current synopsis of oak evolution

Figure 4 highlights the incongruence of the plastid genealogy tree with the current understanding of the evolution of Fagaceae and oaks based on molecular sequence data from the non-coding nuclear gene regions (Manos et al., 2008; Denk and Grimm, 2010), a recent time-calibrated nuclear phylogeny of oaks (Hubert et al. 2014), and the fossil record of modern lineages as documented by pollen investigated under the scanning-electron microscope (Grímsson et al., 2015; see also Denk and Grimm, 2009). Two evolutionary mechanisms (incomplete lineage sorting, reticulation) may account for the observed, highly complex pattern.

Firstly, speciation processes in Quercus do not immediately leave imprints on the plastome (e.g. Neophytou et al., 2010; Cavender-Bares et al., 2011) as also well documented for Nothofagus (Acosta and Premoli, 2010; Premoli et al., 2012). Low mutation rate and long generation time can contribute to slow evolutionary rates and incomplete lineage sorting of organellar genomes (Cavender-Bares et al., 2015; Besnard et al., 2007). In addition, reiterated extinctions and re-colonisations involving bottlenecks, genetic drift, and founder effects may cause random fixation of haplotypes, increasing the probability for retaining ancestral traits. Oaks in general, and especially the Mediterranean taxa, are also characterised by a marked resprouting ability in...
response to disturbances of varying frequencies and intensities, including the action of man, herds, and wild fires (Barbero et al., 1990). This could also have contributed to clonally preserve and transmit ancestral plastid lineages (maternally inherited) during multiple and reiterated unfavourable conditions since the origin of the Mediterranean region (Blondel and Aronson, 1999). At the same time, different environmental selection and adaptation, large population sizes, and long distance pollen dispersal, might have homogenised the nuclear genomes in local populations of a species but not their organelle genomes.

Secondly, Fagaceae lineages are susceptible to hybridisation and introgression (Arnold, 2006). This may lead to the formation of morphologically unambiguous individuals of a species with plastid signatures of another (Whittemore and Schaal, 1991; Petit et al., 2004). There is increasing evidence for local introgression in oak communities with morphologically distinct species in the case of European white oaks (Group Quercus; Q. robur, Q. petraea, Q. pyrenaica, Q. pubescens, Q. frainetto; Curtu et al., 2007; Valbuena-Carabaña et al., 2007; Lepais et al., 2009), as well as in members of Quercus subsection Virentes, a subgroup of Group Quercus, in North America (Cavender-Bares et al., 2015), and across a wide range of Group Lobatae (Dodd and Afzal-Rafii, 2004; Peñaloza-Ramírez et al., 2010; Moran et al., 2012; Valencia-Cuevas et al., 2015). In our focal group, hybrids and different levels of genetic introgression among morphologically pure individuals were molecularly documented in Q. ilex/Q. coccifera (Ortego and Bonal, 2010) and, to a lesser extent, in Q. coccifera/Q. alnifolia (Neophytou et al., 2011).

Also, the potential for inter-group hybridisation was experimentally demonstrated for Q. ilex and Q. robur (Group Quercus; Schnitzler et al., 2004), and natural introgression in Q. ilex/Q. suber was identified in Southern France (Mir et al., 2009) and Iberia (Burgarella et al., 2009). Therefore, it is possible that ancient hybridization and introgression, favoured by the well-known
sexual promiscuity between closely related taxa and their ability to disperse pollen over long distances, obscure the true evolutionary origin of an oak species or entire lineage. In the Mediterranean, the dramatic geological and ecological changes during the Neogene (Blondel and Aronson, 1999) likely caused extinction, re-colonisation, range fragmentation and hybridisation linked to secondary contact, especially when species were still young and reproductive barriers likely weaker than today. Taken together, incomplete sorting of ancestral traits and introgression of haplotypes thus appear highly likely mechanisms to decrease inter-species plastid differentiation while at the same time increasing intra-species variation. This provides an explanation why the widespread, morphologically and genetically (at the nuclear level) distinct species *Q. coccifera* and *Q. ilex* have accumulated three distinct main plastid haplotypes, which we discuss in the following.

**Polyphyletic clues in Quercus Group Ilex**

The most striking finding of this study is the plastid polyphyly and a clear geographic pattern displayed by a group of Mediterranean Ilex oaks (*Q. ilex*, *Q. coccifera* and *Q. aucheri*). Phylogeographic patterns reflecting distant vicariant events and a complex history of range expansions and contractions have been previously inferred for other Mediterranean woody species (Besnard et al., 2007; Desamore et al., 2011; Migliore et al., 2012; Chen et al., 2014). Mediterranean *Laurus* (Rodriguez-Sanchez et al., 2009), for instance, also comprises three plastid haplotype lineages roughly corresponding to biogeographic patterns as seen in the Mediterranean Ilex oaks: (1) an eastern lineage in Turkey and the Near East, (2) a second one in the Aegean region, and (3) a probably ancestral lineage of central and western Mediterranean populations. The importance of the Mediterranean basin in shaping the intraspecific divergence
of Tertiary plant species is therefore well-acknowledged (Nieto Feliner, 2014). Unfortunately, none of the above studies investigated multiple species, leaving the question whether the observed divergence is restricted to a single species or shared by several species as in the case of \textit{Quercus} Group Illex.

Extensive sampling is more likely to document hidden polyphyly (Wiens and Servedio, 2000); besides a weak phylogenetic signal resulting in poorly resolved and potentially inaccurate gene trees (countered here by using median-joining and bipartition networks), incomplete sorting of ancestral lineages and introgression represent very general sources of polyphyly. However, distinguishing the separate effects of these mechanisms can be very difficult in the absence of nuclear markers and (palaeo-)geography as complementing information (Funk and Omland, 2003). As a general rule, the phylogenetically basal position of polyphyly usually hints at retained ancestral polymorphism, while recently introgressed haplotypes may assume a highly derived position in a gene tree. At the same time, incomplete sorting is not predicted to promote the geographic proximity of interspecifically shared haplotypes that may be seen under local introgression (Hare and Avise, 1998, Masta et al., 2002).

In our reconstruction, Group Illex oaks appear to conform to both instances. Nuclear data covering the entire range of \textit{Q. ilex} and \textit{Q. coccifera} in the Mediterranean region unambiguously resolved the two species as close, but mutually monophyletic sister taxa (Denk and Grimm, 2010). In the absence of nucleome data for all here included individuals, it is impossible to infer to which degree introgression and incomplete lineage sorting contributed to the plastid gene pool of the Mediterranean species of Group Illex. Nevertheless, the most straightforward explanation for the observed scenario would be a combined effect of the two mechanisms: asymmetrical introgression of ancestral haplotypes resulting in local genetic clusters decoupled from
taxonomic boundaries, in which plastome accessions of species or species complexes may form
grades or multiple clades in phylogenetic trees, thus appearing polyphyletic (e.g. Rieseberg and

As modelled by Excoffier et al. (2009), interspecific interactions during historical range
fluctuations can profoundly affect the observed phylogeographic patterns, and manifest as
paraphyly or reticulation. In fact, most range expansions do not occur in completely uninhabited
areas, and interbreeding between local and an expanding (invasive) species with subsequent
asymmetrical introgression can develop also in absence of selection (Lepais et al., 2009). Plastid
haplotypes referring to the original (‘lost’) species are indeed likely to persist over long
evolutionary periods, and may still be found in the invading species. Noteworthy, environmental
changes and disturbance of local communities have been shown to increase hybridisation rates
(Lagache et al., 2013), hence, the potential for widespread, imbalanced introgression. In Group
Ilex oaks, the interspecific capture of plastids among sexually incompletely isolated species
likely occurred on the geological timescale, concealing the species relationships at various stages
in the history of the genus. In a comprehensive study of the genus *Ilex* (Manen et al., 2010), the
high incongruence between a taxonomically compatible nuclear gene tree and a geographically
structured plastid tree was explained with extensive extinctions between the Cretaceous and
Miocene and multiple hybridization and introgression events between distantly related lineages.
This has been documented also for *Platanus* (Grimm and Denk, 2010) and more recently
suggested for the evergreen white oaks of *Quercus* subsection Virentes (Eaton et al., 2015).

Similar ancient lateral transfers have been also inferred to explain the paraphyly of the
maternally inherited mtDNA of *Picea* (Bouillè et al., 2011) and *Pinus* (Tsutsui et al., 2009). As
noted above, introgression has been widely demonstrated across a wide range of oaks (Dodd and
Temporal and spatial framework of plastome evolution

The three distinct plastid haplotypes observed in modern Mediterranean members of *Quercus* Group Ilex may reflect three radiation phases (range extensions), followed by range disruptions and isolation of plastome lineages within the ‘Old World Clade’ of *Quercus*. Considering the high diversity of haplotypes in Group Ilex as compared to other major oak lineages (or other genera in the Fagales; see Table 3; Table S2) it can be assumed that the geographical disruptions in the plastome of the ancestors of Group Ilex and interacting lineages predate the manifestation of modern taxa (species and infrageneric groups; Fig. 1). Haplotypes shared between members of Group Ilex and its sister lineages Group Cerris and Group Cyclobalanopsis may indicate common (geographic) origin or may be the result of secondary contact and unidirectional introgression.

Independent from the position of the plastid root, the divergence of the ‘Euro-Med’ haplotype must have coincided with the initial differentiation in *Quercus* (Fig. 1). Oaks had achieved a wide northern hemispheric range by the Eocene. Unequivocal fossils are known from high latitudes (North America, Greenland, North Europe; Crepet and Nixon, 1989; Manchester, 1994; Grimsson et al., 2015) and mid latitudes (Central Europe, South East Asia; Kvaček and Walther, 1989; Hofmann, 2010). All major lineages of oaks were established by the end of the Eocene, ca.
35 Ma, as evidenced by the fossil record and molecular dating using eight nuclear gene regions 
(Bouchal et al., 2014: fig. 14; Hubert et al., 2014; Grimsson et al., 2015). During this time, one 
fraction of oaks must have been geographically and reproductively isolated which would have 
caused a major split in the plastid gene pool (Fig. 1). Today, the ‘Euro-Med’ haplotype is the 
only one exclusively shared by just two, but widespread Mediterranean species of Quercus 
Group Ilex, Q. ilex and Q. coccifera. Two evolutionary scenarios can explain the establishment 
of this haplotype in Q. ilex-Q. coccifera (Fig. 5): (i) The ‘Euro-Med’ haplotype is the remnant of 
an extinct oak lineage that was introgressed (invaded) and consumed by members of Group Ilex; 
under this scenario Group Ilex would have migrated into Europe at some point prior to the 
Miocene where it came into contact with this extinct oak lineage. (ii) The ‘Euro-Med’ haplotype 
represents the original plastome of Group Ilex; under this scenario, the first split within the 
modern ‘Old World clade’ would have been between a western Group Ilex and an eastern Group 
Cyclobalanopsis (Fig. 5). Fossil evidence and available phylogenies (discussed in the following) 
lend high credibility to scenario (i) as the most plausible explanation.

The ‘Cerris-Ilex’ haplotype is shared between all species of Quercus Group Cerris (western 
Eurasian and East Asian), East Mediterranean (Aegean) individuals and two East Asian species 
of Group Ilex. This is in agreement with Denk and Grimm (2009) who suggested that Quercus 
Group Cerris evolved from Group Ilex by budding (a hypothesis further confirmed by the 8-
nuclear gene data set used by Hubert et al., 2014), and the low support for a Group Ilex clade in 
an all-Fagaceae (excluding Fagus) tree based on over 1000 nuclear ITS sequences (Denk and 
Grimm, 2010). Hubert et al. (2014) inferred a Miocene age for this budding event, which 
corresponds to the earliest unequivocal fossil of Quercus Group Cerris (Kmenta, 2011) and is 
younger than the earliest definite fossil record of Quercus Group Ilex in Europe (early
Oligocene, Cospuden; Denk et al., 2012). Also, dispersed pollen from the Paleogene Changchang Formation, Hainan (Hofmann, 2010), resembles both Quercus Group Ilex and Group Cyclobalanopsis; the age of this formation is considered late early to early late Eocene (Lei et al., 1992). The most closely related haplotype to the ‘Cerris-Ilex’ haplotype is encountered in the widespread East Asian Q. phillyraeoides, the only species of Group Ilex extending to Japan (the East Asian members of Group Cerris have a much wider range in north-eastern Asia; Menitsky, 2005). Regarding its phylogenetic position, the emergence of the ‘Cerris-Ilex’ haplotype appears linked with a major taxonomic sorting event in Eurasian Fagaceae, resulting in distinct haplotypes restricted to genera and intrageneric groups of Quercus (Fig. 1). Based on the palaeobotanical record, these lineages (Castanopsis, Castanea, Quercus Group Ilex, Quercus Group Cyclobalanopsis) were well established at least by the Eocene (Table 4, Fig. 5; Grimsson et al., 2015); a deep divergence is reflected by their distinctly different nuclear genomes (Oh and Manos, 2008; Denk and Grimm, 2010; Hubert et al., 2014). Two evolutionary scenarios can explain the occurrence of the ‘Cerris-Ilex’ haplotype in Aegean individuals of Q. ilex and Q. coccifera and the westernmost Q. aucheri: (i) Group Cerris evolved in western Eurasia/Himalaya from an (extinct) subtropical to temperate sublineage of Group Ilex, which left its imprint in the Aegean members of Group Ilex, and Q. spinosa, Q. engleriana and Q. phillyraeoides; (ii) Group Cerris shares a common ancestry with the north-east Asian Q. phillyraeoides. Under this scenario, the budding event of the group took place in north-eastern Asia, from where it migrated into western Eurasia and the Aegean region; in relatively recent times, Group Cerris came into contact with the Mediterranean members of Group Ilex and were locally introgressed.
The high similarity of ‘Cerris-Ilex’ haplotypes lends some credibility to the second scenario. Furthermore, there is evidence for current introgression and occasional hybridization of *Q. suber* (Group Cerris) and *Q. ilex* in the western Mediterranean (Burgarella et al., 2009; Mir et al., 2009). However, it is difficult to explain why *Q. ilex-coccifera* should only intrograde into populations of Cerris oaks at a large scale in the Aegean region. Today, Group Cerris is more diverse than Group Ilex in the East Mediterranean (*Q. brantii, Q. cerris, Q. ithaburensis, Q. macrolepis, Q. libani, Q. trojana*) with some species adapted to distinctly continental climates (Browicz and Zieliński, 1982; Menitsky, 2005), outside the range of Group Ilex. This diversity and the vast distribution of only two species of Group Cerris in East Asia may point towards a young radiation of the group. During the early and middle Miocene, Cerris did not play an important role in western Eurasia. In contrast, two distinct species complexes of *Quercus* Group Ilex were prominently represented in eastern Mediterranean and Paratethyan plant assemblages, *Q. drymeja* Unger and *Q. mediterranea* Unger (e.g. Velitzelos et al., 2014). Intriguingly, the *Q. drymeja* complex includes morphotypes found today in *Q. ilex* as part of its intraspecific variation, and of a range of East Asian species including *Q. engleriana. Quercus mediterranea* is the morphological equivalent of *Q. ilex, Q. coccifera* and a range of East Asian species including *Q. spinosa*. Hence, the fossil record clearly favours a western Eurasian-Himalayan origin of Group Cerris (scenario i).

The West Asian-Himalayan-East Asian (WAHEA) haplotype represents Eastern Mediterranean members of *Quercus* Group Ilex and is sister to a clade comprising several Asian species of Group Ilex (Himalayas to the mountains of Southeast Asia). It reflects the second radiation within the Old World Clade and allies after the isolation of the ‘Euro-Med’ original lineage and prior to the radiation and subsequent sorting within the clade comprising the Cerris-Ilex
The modern distribution of species with the WAHEA haplotype follows the Himalayan corridor (Kitamura, 1955; Güner and Denk, 2012). The Himalayan corridor is a narrow band along the southern flanks of the Himalaya with a subtropical to temperate climate (Cwa, Cwb; Peel et al. 2007) providing a refuge for plants that were more widespread before the Himalayan uplift. In addition to Quercus Group Ilex (Zhou, 1992; Velitzelos et al., 2014), prominent relic taxa include species of Acer, Aesculus, Cedrus, Cotinus, Juglans, Platanus, and Rhododendron among others. The ‘WAHEA’ haplotype represents the western counterpart to the haplotype lineage found in East Asian species of Group Ilex and Cyclobalanopsis. The relic Q. alnifolia, today preserved only in the mid-montane region of Cyprus (Mt. Troodos), would be a witness of this expansion (Menitsky, 2005).

Towards an integrated biogeography of oaks

Plant biogeographic studies at the genus level have commonly relied on few to many chloroplast markers and a single or very few accessions per taxon. In the case of woody angiosperms with a subtropical to temperate distribution such as for example Nothofagaceae (Svenson et al., 2001; Knapp et al., 2005), Rhus (Yi et al., 2004), Cornus (Xiang et al., 2005), Carpinus (Yoo and Wen, 2007), Castanea (Lang et al., 2007), Juglans (Aradhya et al. 2007), and Carya (Zhang et al. 2013a), such an approach runs the risk of capturing but a limited aspect of the evolutionary history of the focal group. Mere combination with e.g. nuclear ITS data can be problematic, too (compare data shown here with data provided by Denk and Grimm, 2010, on western Eurasian members of Group Ilex). The decoupled evolutionary signals in plastomes and the nucleome/morphology as documented for Nothofagus (Acosta and Premoli, 2010; Premoli et al., 2012) and Quercus Group Ilex (this study) suggests that the traditional placeholder sampling
strategy is not ideal. Signals from few-marker/many-samples data sets are likely to be complex or even puzzling (Figs 1–4), but at the same time provide entirely new perspectives on plant evolution worth exploring. For Quercus Group Ilex, our pilot study focussing on Mediterranean species reveals a crucial aspect of oak evolution not seen in the combined nuclear, morphological, and fossil data: large-scale introgression and incomplete lineage sorting among ancestral lineages of modern major groups and species. The new data corroborate hypotheses that Group Cerris evolved (‘budded’) relatively recent from Group Ilex (over 600 ITS and over 900 5S-IGS accessions covering all western Eurasian oak species, Denk and Grimm, 2010; signal from six single-copy nuclear regions, Hubert et al., 2014). Quercus Group Cerris probably evolved in western Eurasia and the Himalayas when the then chiefly subtropical low latitude Group Ilex radiated into temperate niches. Within modern members of Group Cerris, a wide spectrum of leaf traits is found from pseudo-evergreen in Q. suber, to semi-evergreen in Q. brantii, Q. ithaburensis, Q. trojana (partly) and fully deciduous in Q. acutissima, Q. castaneifolia, Q. cerris, Q. libani and Q. variabilis. The conspicuous plastid diversity in the Mediterranean species of Group Ilex and the lineage in general (Figs 1, 2; Table 2) reflects the highly complex geographical history of this group. The ‘Euro-Med’ haplotype evidences an initial phase of west-east differentiation in low-latitude Eurasian oaks, the ‘Old World Clade’, probably triggered by the complex topography within its potential range essentially since the Eocene (Fig. 5). During its evolution, Group Ilex was continuously affected by range disruptions caused by tectonic activity south of the Paratethys linked to the collision of Africa and the Indian subcontinent with Eurasia (Fig. 6); progressive rarefaction of the original haplotypes and the occurrence of (repeated) invasion and introgression events that left imprints in the plastome even within the same species is highly likely.
Although decoupled from taxonomy, the plastid phylogeny provides important, independent information on the geographic differentiation of *Quercus* prior to the formation of modern species/species groups. The major split within oaks has traditionally been considered between ‘New World’ and ‘Old World’ oaks (Manos and Stanford, 2001) because of the current distribution of the major lineages of oaks. The plastid data presented here strongly suggest that the early evolution of oaks instead was geographically bound to high latitude Arctic regions and to low latitude subtropical regions (Fig. 5). The high latitude lineages remained genetically homogeneous in the nucleome, but also in the plastome to some degree. Continuous circum-polar distribution prevented pronounced genetic drift in the high latitude lineage, which became the ‘New World Clade’, and explains low genetic differentiation in deciduous high and mid latitude white oaks until today (Denk and Grimm, 2010). At the same time, the Atlantic, the proto-Mediterranean, and the Paratethys isolated the Eurasian low latitude lineage. Our data should only be viewed as a first step towards a more complete understanding of the biogeography and evolution of oaks. The next step would be to map the plastid variation of *Quercus* Group Ilex across its entire range by sampling multiple stands of the Himalayan and East Asian species to characterise the geographic and taxonomic ranges of the various plastid lineages.

**Conclusion**

Taken all evidence together, the first major split of oaks, consistently found in all molecular phylogenies, would have been into two clades. A northern, high-latitude clade leading to white, red and golden-cup oaks, which evolved and diversified in the tectonically quiet parts of the
Northern Hemisphere. Its counterpart was a southern, mid-latitude clade made up by Group Ilex and Group Cyclobalanopsis (and later by Group Cerris), in the southern part of Eurasia, and perhaps western North America. Both *Quercus* Group Cyclobalanopsis and Group Ilex were present in southern Eurasia, close to the shores of the Tethys, and western North America by the middle Eocene. The outlined history of further steps in the southern clades appear to consist of two radiations: one lead to the Group Cerris clade (early Miocene) with migration westwards along the southern slopes of the Himalayas, and eastwards over China and Japan. The highly coherent ‘Cerris-Ilex’ haplotypes are key witnesses of this event indicating that westwards and eastwards migration and radiation of the monophyletic Group Cerris may have been relatively recent. A putative radiation centre is the East-Mediterranean Paratethys region, in which members of Group Ilex and Cerris coexist and share highly similar to identical haplotypes. The second radiation is likely to have occurred in the Miocene, too. From a Himalayan Group Ilex stock, the clade with ‘WAHEA’ haplotype expanded towards the eastern Mediterranean basin. Both the ancient western Eurasian clade, now extinct but evidenced by the ‘Euro-Med’ haplotype, and the originally Himalayan clade had been invaded by the late Neogene by the direct ancestors of today’s *Q. coccifera* and *Q. ilex*. Modern forms of these two Mediterranean oaks (and their two closest relatives, *Q. aucheri* and *Q. alnifolia*) would have preserved the ancestral haplotypes, keeping strong geographic indications of those events along with the formation of the Mediterranean region. Moreover, the Himalayan uplift coincided with the development of the modern monsoon climate; hence adaptations or exaptation to phases of drought may have played an important role at least since 15 million years ago (Wang and Wu 2015). Our reconstruction is still speculative, but consistent with (i) all fossil observations, (ii) known molecular phylogenies of plastid haplotypes, and (iii) known processes shaping the
diversity of plastid haplotypes by speciation and introgression, linked with invasions. Further
detailed studies, especially in the Himalayan and Chinese regions, are needed to validate or
disprove this scenario.

Because of similar strong correlation between plastome differentiation and geographic
distribution at the population level and the species/genus level, similar or identical plastid
haplotypes typically shared between co-occurring and often distantly related taxa, polyphyletic
signals and reproductive biology, the same processes could have likely played a key role in the
evolutionary history of other Fagaceae (e.g., *Fagus, Castanea, Castanopsis*). Broadening the
sampling efforts of phylogenetic analyses of the plastome could help decipher the speciation
history of these genera. At the same time, extended nucleome investigations will be obviously
necessary to definitely assess a clear molecular phylogeny of Fagaceae.
Acknowledgments

We thank all the Directors and curators of the Botanic Gardens and Arboreta who provided the investigated material: Anthony S. Aiello, Wolfgang Bopp, Anne Boscawen, Peter Brownless, Béatrice Chassé, Laszlo Csiba, Dirk De Meyere, Holly Forbes, Anett Krämer, Isabel Larridon, Gitte Petersen, David Scherberich, Patrick Thompson, Jef Van Meulder, Michael Wall, David Zuckerman, and all the colleagues and friends who contributed to the sampling design: Jeannine Cavender-Bares, Hanno Schaefer, Charalambos Neophytou, Martina Temunovic, Gianni Bedini, Laura Genco, Enara Otaegi Veslin.
References


Bouillé M, Senneville S, Bousquet J. 2011. Discordant mtDNA and cpDNA phylogenies indicate geographic speciation and reticulation as driving factors for the diversification of the genus *Picea*. *Tree Genetics and Genomes* 7: 469–484.


Huzioka K, Takahasi E. 1970. *The Eocene flora of the Ube coal-field, southwest Honshu, Japan*. Akita University, Mining College, Japan


Nylander JAA. 2004. *MrModeltest* 2.1. Program distributed by the author. Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden


Pearse IS, Hipp AL. 2009. Phylogenetic and trait similarity to a native species predict herbivory on non-native oaks. *Proceedings of the National Academy of Sciences of the USA* **106**: 18097-18102


Figure captions

Figure 1 ML tree of plastid accessions; tentatively rooted with the *Notholithocarpus-Chrysolepis* subtree. Stars indicate subtrees comprising accessions of Mediterranean members of *Quercus* Group Ilex. Number at branches indicate non-parametric bootstrap support under maximum likelihood using two different implementations and posterior probabilities calculated using Bayesian inference.

Figure 2 Haplotype network based on length-conserved portions of the *trnH-psbA* spacer. Colouration refers to the taxonomic affiliation of specimens.

Figure 3 Geographic pattern of plastid haplotype variation in Mediterranean members of *Quercus* Group Ilex. (A) Map showing the taxonomic identity of sampled specimens. (B) Map showing the plastid haplotypes of sampled specimens.

Figure 4 Mapping of chloroplast evolution in oaks (using the same rooting scenario as in Fig. 1) on current evolutionary synopsis (based on nuclear sequence data, morphology, and the fossil record; modified after Grímsson et al., 2015, fig. 16). Colouring of the plastid lineages refers to branches/subclades in Fig. 1: bluish, common (ancestral) and ‘New World’ oak/castanoids plastid haplotype lineages; green, lineages of the unique ‘Euro-Med’ plastid haplotype found only in Mediterranean members of Group Ilex; reddish, lineages of ‘Old World’ oaks and Eurasian castanoids. Note that members of Group Ilex keep plastid haplotypes of five different
evolutionary sources/systematic affinities. Abbreviations: C = Cretaceous, Pa = Paleocene, E = Eocene, O = Oligocene, M = Miocene, Pl = Plio-/Pleistocene

Figure 5 Eocene set-up and the origin of the ‘Euro-Med’ haplotype (palaeogeographic map © Ron Blakey, Colorado Plateau Geosystems). A Unequivocal fossil record of oaks in the Eocene mapped on a palaeotopographic map (© Ron Blakey, Colorado Plateau Geosystems) highlighting a primary split into a high-latitude and mid-latitude lineage that likely correspond to the deep phylogenetic split seen in nuclear and plastid sequence data of modern oaks between the ‘New World Clade’ (Groups Protobalanus, Quercus and Lobatae) and the ‘Old World Clade’ (Groups Cyclobalanopsis, Ilex, Cerris). B-C Scenarios that can explain the occurrence of the unique ‘Euro-Med’ haplotype in westernmost members of Quercus Group Ilex. B The ‘Euro-Med’ haplotype belonged to an extinct oak lineage geographically/biologically separated from both the ancestors of the New World and Old World Clade. Westward expansion of Himalayan members of Group Ilex and subsequent large-scale introgression/hybridisation homogenised the western members of Group Ilex and the extinct oak lineage, retaining and evolving the original haplotype in the Mediterranean region. C The ‘Euro-Med’ haplotype reflects geographic fragmentation within the Paleogene range of the Old World Clade that was overprinted to some degree after later radiation phases of Group Ilex

Figure 6 Tectonic activity during the Eocene (palaeogeographic map © Ron Blakey, Colorado Plateau Geosystems) and past and modern distribution of the New World (white) and Old World (yellow) groups within Quercus. Black lines indicate major subduction zones, red lines major orogenies. Note that the high latitude lineage of oaks (Quercus Group Lobatae, Group Quercus,
Group Protobalanus) evolved in tectonically stable regions, whereas the low latitude lineage (Quercus Group Ilex, Group Cyclobalanopsis, Group Cerris) evolved in tectonically unstable regions. Uppercase and lowercase letters refer to extant and extinct distribution areas of major oak lineages: P,p = Group Protobalanus; Q,q = Group Quercus; L,l = Group Lobatae; I,i = Group Ilex; C,c = Group Cerris; Y,y = Group Cyclobalanopsis
Supporting Information

Table S1 List of specimens and accessions included in this study

Table S2 Intra- and intertaxonomic minimum and maximum pairwise genetic distances

File S3 MJ networks based on different plastid gene regions

Data Archival Location

All sequence data generated as part of this study are available on GenBank (http://www.ncbi.nlm.nih.gov/genbank/) under accession numbers LM652969-LM653098; other relevant data are within the paper and its Supplementary files. Primary data and analyses are provided for anonymous download at www.palaeogrimm.org/data/Smn15_OSA.zip
Figure 1 (on next page)

ML tree of the investigated oak accessions

ML tree of plastid accessions; tentatively rooted with the Notholithocarpus-Chrysolepis subtree. Stars indicate subtrees comprising accessions of Mediterranean members of Quercus Group Ilex. Number at branches indicate non-parametric bootstrap support under maximum likelihood using two different implementations and posterior probabilities calculated using Bayesian inference.
Figure 2 (on next page)

Haplotype network based on the trnH-psbA spacer

Haplotype network based on length-conserved portions of the trnH-psbA spacer. Colouration refers to the taxonomic affiliation of specimens.
Figure 3 (on next page)

Plastid haplotype variation in Mediterranean members of Quercus Group Ilex

Geographic pattern of plastid haplotype variation in Mediterranean members of *Quercus* Group Ilex. (A) Map showing the taxonomic identity of sampled specimens. (B) Map showing the plastid haplotypes of sampled specimens
**Figure 4 (on next page)**

Map of chloroplast evolution in oaks

Mapping of chloroplast evolution in oaks (using the same rooting scenario as in Fig. 1) on current evolutionary synopsis (based on nuclear sequence data, morphology, and the fossil record; modified after (Grímsson et al., 2015, fig. 16). Colouring of the plastid lineages refers to branches/subclades in Fig. 1: bluish, common (ancestral) and ‘New World’ oak/castanoids plastid haplotype lineages; green, lineages of the unique ‘Euro-Med’ plastid haplotype found only in Mediterranean members of Group Ilex; reddish, lineages of ‘Old World’ oaks and Eurasian castanoids. Note that members of Group Ilex keep plastid haplotypes of five different evolutionary sources/systematic affinities. Abbreviations: C = Cretaceous, Pa = Paleocene, E = Eocene, O = Oligocene, M = Miocene, Pl = Plio-/Pleistocene
Figure 5 (on next page)

Origin of the ‘Euro-Med’ haplotype

Eocene set-up and the origin of the ‘Euro-Med’ haplotype (palaeogeographic map © Ron Blakey, Colorado Plateau Geosystems). A Unequivocal fossil record of oaks in the Eocene mapped on a palaeotopographic map (© Ron Blakey, Colorado Plateau Geosystems) highlighting a primary split into a high-latitude and mid-latitude lineage that likely correspond to the deep phylogenetic split seen in nuclear and plastid sequence data of modern oaks between the ‘New World Clade’ (Groups Protobalanus, Quercus and Lobatae) and the ‘Old World Clade’ (Groups Cyclobalanopsis, Ilex, Cerris). B-C Scenarios that can explain the occurrence of the unique ‘Euro-Med’ haplotype in westernmost members of Quercus Group Ilex. B The ‘Euro-Med’ haplotype belonged to an extinct oak lineage geographically/biologically separated from both the ancestors of the New World and Old World Clade. Westward expansion of Himalayan members of Group Ilex and subsequent large-scale introgression/hybridisation homogenised the western members of Group Ilex and the extinct oak lineage, retaining and evolving the original haplotype in the Mediterranean region. C The ‘Euro-Med’ haplotype reflects geographic fragmentation within the Paleogene range of the Old World Clade that was overprinted to some degree after later radiation phases of Group Ilex.
Tectonic activity during the Eocene

Tectonic activity during the Eocene (palaeogeographic map © Ron Blakey, Colorado Plateau Geosystems) and past and modern distribution of the New World (white) and Old World (yellow) groups within Quercus. Black lines indicate major subduction zones, red lines major orogenies. Note that the high latitude lineage of oaks (Quercus Group Lobatae, Group Quercus, Group Protobalanus) evolved in tectonically stable regions, whereas the low latitude lineage (Quercus Group Ilex, Group Cyclobalanopsis, Group Cerris) evolved in tectonically unstable regions. Uppercase and lowercase letters refer to extant and extinct distribution areas of major oak lineages: P,p = Group Protobalanus; Q,q = Group Quercus; L,l = Group Lobatae; I,i = Group Ilex; C,c = Group Cerris; Y,y = Group Cyclobalanopsis
Table 1 (on next page)

Species list

Species included in Quercus Group Ilex according to Denk and Grimm (2010); nomenclature followed Govaerts and Frodin (1998); species investigated in the present study are bolded. Taxonomic remarks and species distributions according to Govaerts and Frodin (1998), *Menitsky (2005) and **Flora of China Editorial Committee (1999).
Table 1 – Species included in *Quercus* Group Ilex according to Denk and Grimm (Denk and Grimm 2010); nomenclature followed Govaerts and Frodin (1998); species investigated in the present study are bolded. Taxonomic remarks and species distributions according to Govaerts and Frodin (1998), *Menitsky (2005) and **Flora of China Editorial Committee (1999).

<table>
<thead>
<tr>
<th>Species</th>
<th>Taxonomic remarks</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Q. acrodonta</em> Seemen</td>
<td>Includes <em>Q. handeliana</em> A. Camus*/<em>/</em></td>
<td>C, E and S China</td>
</tr>
<tr>
<td><em>Q. alnifolia</em> Poech</td>
<td></td>
<td>Cyprus</td>
</tr>
<tr>
<td><em>Q. aquifolioides</em> Rehder &amp; E.H.Wilson</td>
<td>Includes <em>Q. semecarpifolia</em> subsp. glabra*</td>
<td>Tibet, C and SW China to Myanmar</td>
</tr>
<tr>
<td><em>Q. aucheri</em> Jaub. &amp; Spach</td>
<td></td>
<td>SW Anatolia</td>
</tr>
<tr>
<td><em>Q. baloot</em> Griff.</td>
<td></td>
<td>Pakistan, Afghanistan</td>
</tr>
<tr>
<td><em>Q. baronii</em> Skan</td>
<td>Numerous morpho-ecological traits in common with members of Group Cerris*</td>
<td>NC and SW China</td>
</tr>
<tr>
<td><em>Q. bawanglingensis</em> C.C. Huang, Ze X. Li &amp; F.W. Xing</td>
<td>Poorly known; uncertain status, related to <em>Q. phillyreoides</em>**</td>
<td>SE China</td>
</tr>
<tr>
<td><em>Q. coccifera</em> L.</td>
<td>Includes <em>Q. calliprinos</em> Webb*</td>
<td>Mediterranean</td>
</tr>
<tr>
<td><em>Q. cocciferoides</em> Hand.-Mazz.</td>
<td>Includes <em>Q. taliensis</em> A. Camus**</td>
<td>CS China</td>
</tr>
<tr>
<td><em>Q. floribunda</em> Lindl. ex A. Camus</td>
<td>Basionym: <em>Q. dilatata</em> Lindl. ex A.DC. nom. illegit.</td>
<td>Pakistan, Afghanistan, Nepal</td>
</tr>
<tr>
<td><em>Q. dolicholepis</em> A. Camus</td>
<td>Includes <em>Q. fimbriata</em>**</td>
<td>CW to SW China</td>
</tr>
<tr>
<td><em>Q. engleriana</em> Seemen</td>
<td></td>
<td>Tibet to E China, Myanmar</td>
</tr>
<tr>
<td><em>Q. fimbriata</em> Y.C. Hsu &amp; H. Wei Jen</td>
<td>Included in <em>Q. semecarpifolia</em> or <em>Q. dolicholepis</em>**</td>
<td>C China</td>
</tr>
<tr>
<td><em>Q. franchetii</em> Skan</td>
<td></td>
<td>C China to N Vietnam</td>
</tr>
<tr>
<td><em>Q. gilliana</em> Rehder &amp; E.H. Wilson</td>
<td>Included in <em>Q. spinosa</em>***</td>
<td>Tibet, C and S China</td>
</tr>
<tr>
<td><em>Q. guyavifolia</em> H. Lév.</td>
<td>Included in <em>Q. semecarpifolia</em> subsp. glabra*; *Q.</td>
<td>C and S China</td>
</tr>
<tr>
<td>Q. ilex L.</td>
<td>pannosa or Q. aquifolioides var. rufescens**</td>
<td>Mediterranean</td>
</tr>
<tr>
<td>Q. kingiana Craib</td>
<td></td>
<td>C China to N Thailand</td>
</tr>
<tr>
<td>Q. lanata Sm.</td>
<td>Included in Q. leucotrichophora**</td>
<td>Buthan to Vietnam</td>
</tr>
<tr>
<td>Q. leucotrichophora A. Camus</td>
<td>Basionym: Q. incana Roxb. nom. illegit.</td>
<td>N Pakistan, N India, Nepal to N Vietnam</td>
</tr>
<tr>
<td>Q. lodicosa O.E. Warb. &amp; E.F. Warb.</td>
<td></td>
<td>SE Tibet to Myanmar</td>
</tr>
<tr>
<td>Q. longispica (Hand.-Mazz.) A. Camus</td>
<td>Includes Q. semecarpifolia subsp. glabra*; Q. rehderiana**</td>
<td>C and S China</td>
</tr>
<tr>
<td>Q. marlipoensis Hu &amp; Cheng</td>
<td>Poorly known*; very close to Q. engleriana**</td>
<td>C China</td>
</tr>
<tr>
<td>Q. monimotricha Hand.-Mazz.</td>
<td></td>
<td>C and S China</td>
</tr>
<tr>
<td>Q. oxyphylla Hand.-Mazz.</td>
<td>Includes Q. spathulata Seemen; included in Q. dolicholepis**</td>
<td>C and S China</td>
</tr>
<tr>
<td>Q. pannosa Hand.-Mazz.</td>
<td>Possibly conspecific with Q. semecarpifolia*</td>
<td>C and S China</td>
</tr>
<tr>
<td>Q. phillyreoides A. Gray</td>
<td>Includes Q. utilis*</td>
<td>C China to Japan</td>
</tr>
<tr>
<td>Q. pseudosemecarpifolia A. Camus</td>
<td>Includes Q. semecarpifolia subsp. glabra*, Q. rehderiana**</td>
<td>Tibet to CS China</td>
</tr>
<tr>
<td>Q. rehderiana Hand.-Mazz.</td>
<td>Included in Q. semecarpifolia subsp. glabra*; includes Q. longispica and Q. pseudosemecarpifolia**</td>
<td>Tibet to C and S China</td>
</tr>
<tr>
<td>Q. semecarpifolia Sm.</td>
<td>Includes: Q. fimbriata, Q. gujavifolia, Q. aquifolioides, Q. rehderiana, Q. longispica, Q. pseudosemecarpifolia*</td>
<td>Afghanistan to Myanmar</td>
</tr>
<tr>
<td>Q. senescens Hand.-Mazz.</td>
<td></td>
<td>E Himalaya, Tibet, C and S China</td>
</tr>
<tr>
<td>Q. setulosA Hickel &amp; A. Camus</td>
<td></td>
<td>C China to Vietnam</td>
</tr>
<tr>
<td>Q. spinosa David</td>
<td></td>
<td>NC and SW China to Taiwan</td>
</tr>
<tr>
<td>Q. tarokoensis Hayata</td>
<td></td>
<td>E Taiwan</td>
</tr>
<tr>
<td>Q. utilis Hu &amp; Cheng</td>
<td>Included in Q. phillyreoides subsp. fokiensis*</td>
<td>C China</td>
</tr>
</tbody>
</table>
Table 2 (on next page)

Diversity values of the DNA fragments used for the analyses
**Table 2.** Diversity values and models of DNA evolution of the fragments used for the analyses in 59 *Quercus* species (137 individuals) and 4 outgroup taxa. AL: Aligned length (bp); P: uncorrected p-distance (min. – max.); N hap: Number of identified haplotypes, brackets: with gaps considered; Hd: Haplotype diversity; S: Number of polymorphic sites; θ: Nucleotide polymorphism; π: Nucleotide diversity; PIC: Number of Parsimony Informative Characters; ME: Model of evolution.

<table>
<thead>
<tr>
<th>Markers</th>
<th>AL</th>
<th>P</th>
<th>Nhap</th>
<th>Hd</th>
<th>S</th>
<th>PICs</th>
<th>θw</th>
<th>π</th>
<th>ME</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>rbcL</em></td>
<td>743</td>
<td>0.00 – 0.008</td>
<td>28</td>
<td>0.846±0.027</td>
<td>26</td>
<td>18</td>
<td>0.0063</td>
<td>0.0027</td>
<td>HKY+I</td>
</tr>
<tr>
<td><em>trnH-psbA</em></td>
<td>634</td>
<td>0.00 – 0.035</td>
<td>37 (84)</td>
<td>0.944±0.008</td>
<td>38</td>
<td>23</td>
<td>0.0159</td>
<td>0.009</td>
<td>GTR+G</td>
</tr>
<tr>
<td><em>trnK/matK</em></td>
<td>705</td>
<td>0.00 – 0.022</td>
<td>49 (51)</td>
<td>0.952±0.008</td>
<td>59</td>
<td>31</td>
<td>0.0156</td>
<td>0.0064</td>
<td>n.d.</td>
</tr>
<tr>
<td><em>trnK</em> (intron)</td>
<td>401</td>
<td>n.d.</td>
<td>32 (34)</td>
<td>0.821±0.028</td>
<td>36</td>
<td>16</td>
<td>0.0169</td>
<td>0.0048</td>
<td>GTR+G</td>
</tr>
<tr>
<td><em>matK</em> (codons)</td>
<td>304</td>
<td>n.d.</td>
<td>25</td>
<td>0.925±0.008</td>
<td>23</td>
<td>15</td>
<td>0.0146</td>
<td>0.0086</td>
<td>n.d.</td>
</tr>
<tr>
<td><em>rbcL+trnK/matK+trnH-psbA</em></td>
<td>2082</td>
<td>0.00 – 0.014</td>
<td>74 (110)</td>
<td>0.978±0.005</td>
<td>122</td>
<td>72</td>
<td>0.0119</td>
<td>0.0056</td>
<td>Combined</td>
</tr>
<tr>
<td><em>rbcL + matK</em></td>
<td>1047</td>
<td>n.d.</td>
<td>49</td>
<td>0.965±0.006</td>
<td>49</td>
<td>33</td>
<td>0.0085</td>
<td>0.0044</td>
<td>n.d.</td>
</tr>
<tr>
<td><em>trnH-psbA + trnK</em></td>
<td>1035</td>
<td>n.d.</td>
<td>57 (103)</td>
<td>0.954±0.008</td>
<td>69</td>
<td>34</td>
<td>0.0161</td>
<td>0.0067</td>
<td>n.d.</td>
</tr>
<tr>
<td><em>trnH-psbA + trnK-matK</em></td>
<td>1339</td>
<td>n.d.</td>
<td>65 (110)</td>
<td>0.970±0.006</td>
<td>92</td>
<td>49</td>
<td>0.0155</td>
<td>0.0072</td>
<td>n.d.</td>
</tr>
</tbody>
</table>
Divergence patterns in other Fagaceae and Fagales

Divergence patterns in *Quercus* compared to other Fagaceae and Fagales based inter-species pair wise uncorrected p-distances of sequences retrieved from GenBank and produced in this study; \( N_t \) = number of taxa; \( N_s \) = number of sequences.
Table 3. Divergence patterns in *Quercus* compared to other Fagaceae and Fagales based inter-species pair wise uncorrected p-distances of sequences retrieved from GenBank and produced in this study; Nt = number of taxa; Ns = number of sequences.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Quercus</em></td>
<td>87/87/86</td>
<td>219/255/382</td>
<td>0.000</td>
<td>0.010</td>
<td>0.000</td>
<td>0.023</td>
<td>0.000</td>
<td>0.042</td>
<td>0.003 0.011 0.006 0.015 0.006 0.042</td>
</tr>
<tr>
<td><em>Fagus</em></td>
<td>8/9/6</td>
<td>33/30/19</td>
<td>0.000</td>
<td>0.014</td>
<td>0.000</td>
<td>0.007</td>
<td>0.000</td>
<td>0.042</td>
<td>0.024 0.036 0.091 0.098 0.120 0.147</td>
</tr>
<tr>
<td>Other Fagaceae</td>
<td>37/21/28</td>
<td>102/32/86</td>
<td>0.000</td>
<td>0.020</td>
<td>0.000</td>
<td>0.021</td>
<td>0.000</td>
<td>0.020</td>
<td>0.003 0.013 0.000 0.019 0.011 0.032</td>
</tr>
<tr>
<td>Nothofagaceae*</td>
<td>23/†/14</td>
<td>35/†/53</td>
<td>0.000</td>
<td>0.027</td>
<td>–</td>
<td>–</td>
<td>0.000</td>
<td>0.017</td>
<td>0.012 0.023 –</td>
</tr>
<tr>
<td>Betulaceae</td>
<td>55/19/77</td>
<td>131/34/247</td>
<td>0.000</td>
<td>0.011</td>
<td>0.000</td>
<td>0.006</td>
<td>0.000</td>
<td>0.069</td>
<td>0.006 0.024 0.014 0.033 0.011 0.079</td>
</tr>
<tr>
<td>Juglandaceae</td>
<td>18/†/21</td>
<td>23/†/28</td>
<td>0.000</td>
<td>0.005</td>
<td>–</td>
<td>–</td>
<td>0.000</td>
<td>0.007</td>
<td>0.000 0.021 –</td>
</tr>
</tbody>
</table>

* Values for *rbcL* may be over-estimated (data usually older than 15 years; sequences show features characteristic for sequencing and editing artifacts)

‡ Insufficient data.

‡ Values for *Quercus* and other Fagaceae not including *Fagus* (see *Fagus* for max. inter-generic divergence in Fagaceae)
Table 4 (on next page)

Eocene fossils of Quercus
<table>
<thead>
<tr>
<th>Locality, site, age</th>
<th>Reference</th>
<th>Taxon, organ</th>
<th>Affinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarno Fm., Oregon, western U.S.A.; ~48 Ma</td>
<td>Manchester 1994</td>
<td>“Quercus” paleocarpa Manchester; cupules and acorns</td>
<td>Group Cyclobalanopsis (? Lithocarpus)</td>
</tr>
<tr>
<td>Axel-Heiberg Island, Canadian Arctic; ~45 Ma</td>
<td>McIntyre 1991</td>
<td>Pollen and leaves</td>
<td>New World Clade; \textit{Quercus} Group Quercus/Lobatae Extinct/ancestral type</td>
</tr>
<tr>
<td>Hareøn, western Greenland; ~42 Ma</td>
<td>McIver and Basinger 1999 Grimmson et al. 2015</td>
<td>\textit{Quercus} sp. 4–5; pollen \textit{Quercus} sp. 6–7; pollen \textit{Quercus} sp. 1–3; pollen</td>
<td>New World Clade (aff. Group Protobalanus) \textit{Quercus} Group Quercus and/or Lobatae \textit{Quercus} Group Quercus</td>
</tr>
<tr>
<td>Baltic amber, northern Europe; ~45 Ma</td>
<td>Crepet and Nixon 1989; Weitschat and Wichard 2003</td>
<td>Flower and \textit{in situ} pollen</td>
<td>\textit{Quercus} Group Quercus</td>
</tr>
<tr>
<td>Königsau, near Aschersleben, Germany; middle Eocene (48–38 Ma)</td>
<td>Kvaček and Walther 1989</td>
<td>\textit{Quercus subhercynica} H. Walther &amp; Kvaček; leaf</td>
<td>\textit{Quercus} Group Lobatae</td>
</tr>
<tr>
<td>Ube coal-field, southwestern Honshu, Japan; middle Eocene (48–38 Ma)</td>
<td>Huzioka and Takahasi 1970</td>
<td>\textit{Cyclobalanopsis nathoi} Huzioka &amp; Takahashi; leaf, acorn (?), cupule (?)</td>
<td>\textit{Quercus}, affinity unclear</td>
</tr>
<tr>
<td>Changchang, Hainan; middle (?) Eocene (50–35 Ma)</td>
<td>Hofmann 2010</td>
<td>\textit{Quercus} pollen types 2-8, 10; pollen \textit{Quercus} pollen type 1; pollen \textit{Quercus} pollen type 9; pollen</td>
<td>\textit{Quercus}, affinity unclear (extinct, Group Quercus/Lobatae?, Group Protobalanus?) \textit{Quercus} Group Ilex \textit{Quercus} Group Cyclobalanopsis</td>
</tr>
</tbody>
</table>