

RESEARCH ARTICLE

Open Access



Ancient events and climate adaptive capacity shaped distinct chloroplast genetic structure in the oak lineages

Mengxiao Yan^{1,2}, Ruibin Liu^{1,3}, Ying Li^{1,4}, Andrew L. Hipp^{5,6}, Min Deng^{1,2*}  and Yanshi Xiong¹

Abstract

Background: Understanding the origin of genetic variation is the key to predict how species will respond to future climate change. The genus *Quercus* is a species-rich and ecologically diverse woody genus that dominates a wide range of forests and woodland communities of the Northern Hemisphere. *Quercus* thus offers a unique opportunity to investigate how adaptation to environmental changes has shaped the spatial genetic structure of closely related lineages. Furthermore, *Quercus* provides a deep insight into how tree species will respond to future climate change. This study investigated whether closely related *Quercus* lineages have similar spatial genetic structures and moreover, what roles have their geographic distribution, ecological tolerance, and historical environmental changes played in the similar or distinct genetic structures.

Results: Despite their close relationships, the three main oak lineages (*Quercus* sections *Cyclobalanopsis*, *Ilex*, and *Quercus*) have different spatial genetic patterns and occupy different climatic niches. The lowest level and most homogeneous pattern of genetic diversity was found in section *Cyclobalanopsis*, which is restricted to warm and humid climates. The highest genetic diversity and strongest geographic genetic structure were found in section *Ilex*, which is due to their long-term isolation and strong local adaptation. The widespread section *Quercus* is distributed across the most heterogeneous range of environments; however, it exhibited moderate haplotype diversity. This is likely due to regional extinction during Quaternary climatic fluctuation in Europe and North America.

Conclusions: Genetic variations of sections *Ilex* and *Quercus* were significantly predicted by geographic and climate variations, while those of section *Cyclobalanopsis* were poorly predictable by geographic or climatic diversity. Apart from the different historical environmental changes experienced by different sections, variation of their ecological or climatic tolerances and physiological traits induced varying responses to similar environment changes, resulting in distinct spatial genetic patterns.

Keywords: *Quercus*, Spatial genetic structure, Climate, Geography, Local adaptation, Chloroplast genome

Background

Anthropogenic climate change has a marked impact on both species and ecosystems throughout the world [1, 2]. Contemporary spatial patterns of population genetic structure are products of (1) modern climate, landscapes, and range shift [3–8], (2) historical factors such as plate tectonics (e.g., orogeny) and climate change (e.g., the

glacial-interglacial cycle) [9–11], and (3) species-specific or lineage-specific attributes such as climatically-selected physiological traits, which shape species distribution and their ability to adapt [12–14]. Understanding the mechanisms that generate and structure population genetic variation will help to predict species responses to future climatic conditions and conserve diversity [15, 16]. Furthermore, this will help to unravel the details of how geography and environment influence spatial genetic diversity patterns, which is a central goal of ecological and conservation genetics [17, 18].

Climate change may create particular problems for tree species. While, trees tend to be capable of high rates

* Correspondence: dengmin@sibs.ac.cn

¹Shanghai Chenshan Plant Science Research Center, Chinese Academy of Sciences, Shanghai Chenshan Botanical Garden, Shanghai 201602, China

²Southeast Asia Biodiversity Research Institute, Chinese Academy of Sciences, Yezin, Nay Pyi Taw 05282, Myanmar

Full list of author information is available at the end of the article



of local adaptation [19], their long generation times and immobility render them vulnerable to rapid environmental change [1, 20, 21]. Trees, and, oaks (*Quercus*, Fagaceae) in particular are an important group to understand the impacts of environmental change on population genetic diversity for organisms with different traits and distributions [5, 20, 22]. Oaks are among the most widespread woody genera in the Northern Hemisphere, with important ecological functions and economical services for both ecosystem and humans [23–26]. With ca. 450 species, oaks are rich in life history strategies and ecological diversity [27–30]. They dominate a wide range of forest and woodland ecosystems, ranging from temperate deciduous forests and savannas of North America, Europe and Asia, to Mediterranean and desert scrub forests of the Americas and Europe, and to tropical montane forests of South America and Southeast Asia [31]. Moreover, oak seeds are typically recalcitrant (less drought-tolerant and thus limited in their longevity) and their long-range dispersal depend on mammals or birds [32]. As a consequence, the maternal contribution to the population genetic structure, which we can assess directly through plastome sequencing, tends to exhibit conservative, regionalized patterns of diversity [33, 34] that tend to be linked to climate [35]. Hence, oaks offer an ideal system for investigating how natural selection, environmental changes and spatial patterns of existing plants may shape genetic structure of closely related lineages. Due to their wide distribution and dominance, oaks can provide a deep insight into how geological events and climate changes imprint the evolutionary history of forests of the Northern Hemisphere.

Oak plastome haplotypes have been widely sampled, particularly in three major clades: sections *Cyclobalanopsis* and *Ilex*, which have an East and Southeast Asian and Eurasian (sub) tropical distribution, respectively; and section *Quercus*, which is widely distributed in the Northern Hemisphere [27]. Recent phylogeographic studies of the species of section *Cyclobalanopsis* [36–38] indicated that landscape, climate and local adaptation shape regional genetic diversity patterns. Similarly in section *Ilex*, geomorphological and climate changes from the Neogene onward have been demonstrated to have shaped the genetic structure [39–47]. The European white oaks of section *Quercus* have been particularly well studied [33, 48, 49], and the results consistently demonstrated that geographic patterns affect plastome diversity across species boundaries: consequently, it is where populations grow rather than which species they belong to that shapes their plastid haplotype diversity. The impacts of geography, geology and climate on North American members of the section have been explored by less studies [29, 50–53], and with less conclusive findings: while species play a very small role in structuring plastome haplotype diversity, and

geography is also not clearly associated in these studies. These studies point to the potential for multiple species to share plastomes, suggesting the potential power in cross-species studies of plastome diversity within oak clades. However, these studies are all restricted to species and clade. This limits the degree to which they can use cross-species, cross-clade comparative analyses to address the relative contributions of the environment and geography to the spatial-temporal pattern of genetic diversity in oaks.

This study addressed two interrelated sets of questions. First, do the closely related oak clades share a similar spatial genetic structure? If not, what induced the distinct genetic patterns of oak clades? Second, what roles have geographic distribution, ecological tolerance and historical environmental changes played in the genetic diversity of each clade? This study sampled plastome haplotypes of representative species from three oak clades (sections *Cyclobalanopsis*, *Ilex* and *Quercus*) that span the geographic range of each. We aim to figure out the underlying mechanisms that influence the maternal genetic pattern of oaks to predict how future climatic changes will affect acorn migration and to provide new insights into the response of trees to future climate change, and to offer suggestions for the sustainable management of forest tree populations.

Results

Genetic diversity and neutrality tests

Section *Cyclobalanopsis* exhibited the lowest mean p -distance (0–0.0020), nucleotide diversity (0.917) and F_{st} (0.3220) among the three sections (Table 1). In addition, the majority of its pairwise genetic distances were narrowly distributed at low values (Fig. 1). Section *Ilex* had the highest genetic diversity as estimated using p -distance (0–0.0090), genetic distances and nucleotide diversity (0.985), however, it had a median F_{st} of 0.4624 (Fig. 1, Table 1). The genetic diversity of section *Quercus* was in-between (Fig. 1). Its p -distance (0–0.0046), genetic distances and nucleotide diversity (0.934) were medium (Fig. 1, Table 1), however, section *Quercus* had the highest F_{st} of 0.7436.

Tajima's D , F_u and Li's F and Fay and Wu's H were significantly negative on section *Cyclobalanopsis* (Table 2), suggesting a recent selective sweep in section *Cyclobalanopsis*. Neutrality tests for section *Ilex* were nonsignificant. F_u and Li's D was moderately significantly ($p < 0.05$) negative and Fay and Wu's H was negative in section *Quercus*. This is likely the result of population shrinkage during the glacial period.

Phylogenetic and spatial structures

The earliest diverging branches of section *Cyclobalanopsis* (blue dots in Fig. 2a) were concentrated in southwest China (SW China), Vietnam and Nepal. These areas are near the western edge of the range of section *Cyclobalanopsis*, but

Table 1 Genetic diversity parameters of three sections of genus *Quercus*

Sections	N	Ns	p-distance	Vs		Pi	Hn	Hd	Indel	Fst
				number	%					
<i>Cyclobalanopsis</i>	147	28	0–0.019 (0.0020 ± 0.0031)	145	4.87	0.00252	54	0.917	56	0.3220
<i>Ilex</i>	121	28	0–0.018 (0.0090 ± 0.0041)	163	5.27	0.01289	66	0.985	184	0.4624
<i>Quercus</i>	89	27	0–0.020 (0.0046 ± 0.0014)	77	2.61	0.00415	30	0.934	90	0.7436

N number of samples, **Ns** number of species, **p-distance** (min.–max. (mean ± SD)), **Vs** variable sites, **Pi** nucleotide diversity, **Hn** number of haplotypes, **Hd** haplotype diversity

do not show any clear geographic structure. The majority of individuals in this section formed a large lineage encompassing the entire range of the section (green dots in Fig. 2a), which are widespread in subtropical East Asia, including Japan, mainland China and even stretching to Nepal. This main lineage likewise did not exhibit geographic structure. In total, seven haplotypes were observed in this section (H2– H3, H5–H7, H10, and H21; Fig. 2b), which were all shared among species. Haplotypes H5 and H6, were restricted to SW China, while all others were geographically widespread (Fig. 2b). Particularly haplotype H3 essentially spanned the entire geographic distribution of the section, from the East Asian subtropics to Northern Indochina and the southern slope of the Himalayas.

The chloroplast phylogeny of section *Ilex* showed a strong geographic structure with the Himalayas-Mediterranean lineage (purple dots in Fig. 3a) strongly separated from the East Asian lineage. The East Asian lineage further separated into the SW China lineage (blue dots in Fig. 3a) and the

Sino-Japan lineage (green dots in Fig. 3a), comprising individuals from the tropics and subtropics of both mainland China and Taiwan Island. The sharing of haplotypes among species (H68, H81, H89, H91, H96, H99 and H114) in this section was also identified and these interspecific haplotypes were locally distributed (Additional file 1: Table S1, Additional file 2a).

Section *Quercus* was composed of two lineages, corresponding to the Eurasian lineage (purple dots in Fig. 3b) and the North American lineage (orange dots in Fig. 3b). Within regions, however, extensive sharing of haplotypes (H42–H43, H47–H50, H56–H58, H62 and H65) were observed among species of the section (Additional file 1: Table S1, Additional file 2b). This has also been reported in previous studies [33, 50, 53].

Climatic niches

To identify which climatic variables restrict the geographic distribution of different sections, the contribution of climatic variables was ranked (Additional file 1: Table S2).

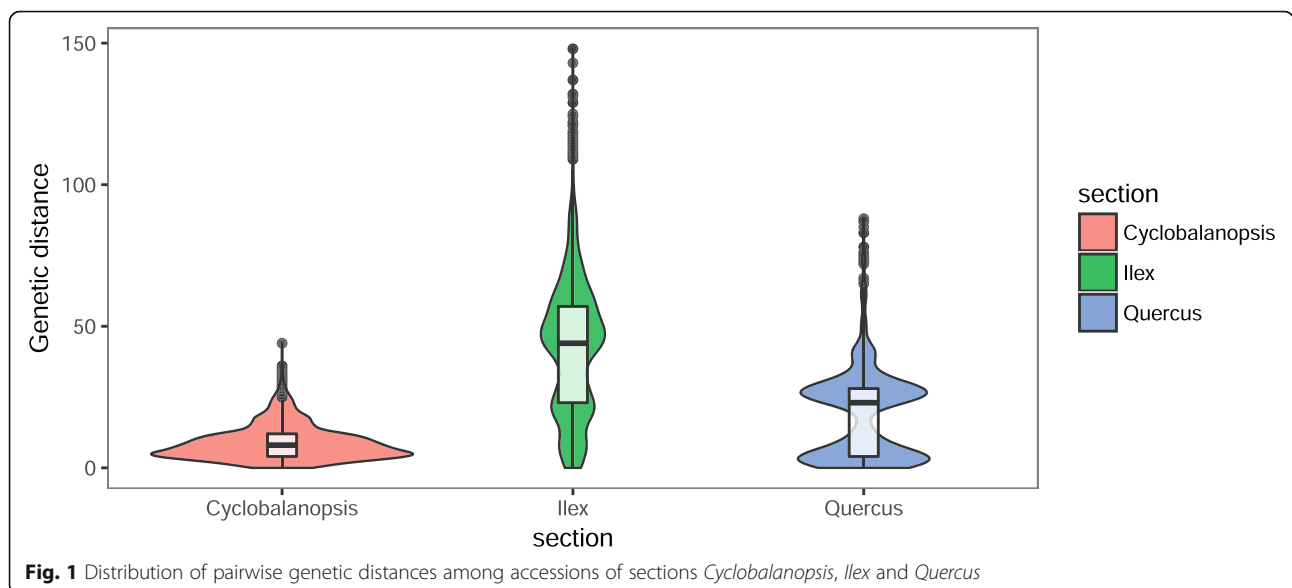


Table 2 Neutrality test of sections *Cyclobalanopsis*, *Ilex* and *Quercus*

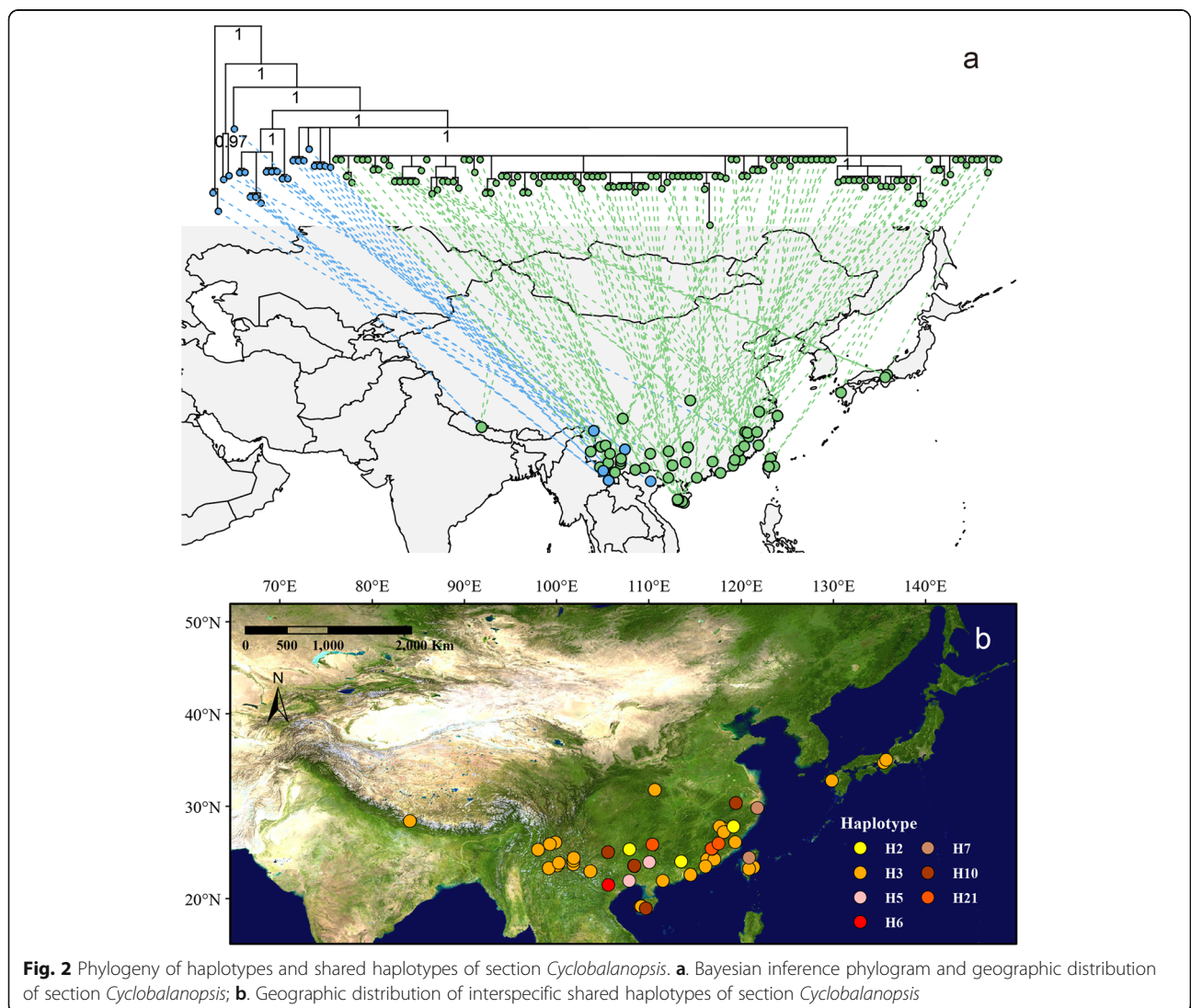
Section	Tajima's <i>D</i>	Fu and Li's <i>D</i>	Fu and Li's <i>F</i>	Fay and Wu's <i>H</i>
<i>Cyclobalanopsis</i>	-2.35202**	-1.95271	-2.55835*	-27.2651
<i>Ilex</i>	-1.01241	-0.44143	-0.83735	1.7057
<i>Quercus</i>	-0.92330	-2.34890*	-2.10962	-12.4408

* $p < 0.05$; ** $p < 0.01$

The top climatic factors that restricted species distribution in section *Cyclobalanopsis* were bio12 (annual precipitation), bio3 (isothermality), bio1 (annual mean temperature), and bio7 (temperature annual range), in the order of variable contribution. For section *Ilex*, the top climatic factors with the highest contributions were bio7, bio3, bio1 and bio12. Furthermore, for section *Quercus*, they were bio1, bio12, bio3 and bio19 (precipitation of coldest quarter).

Overall, annual mean temperature and annual precipitation were key climatic variables that restricted species distribution of the three oak sections.

Section *Cyclobalanopsis* had relatively narrow ranges of annual mean temperature and annual precipitation (Fig. 4), corresponding to low multidimensional climatic niche diversity (Fig. 5a). It had higher mean annual temperature than section *Quercus*, and higher mean annual precipitation



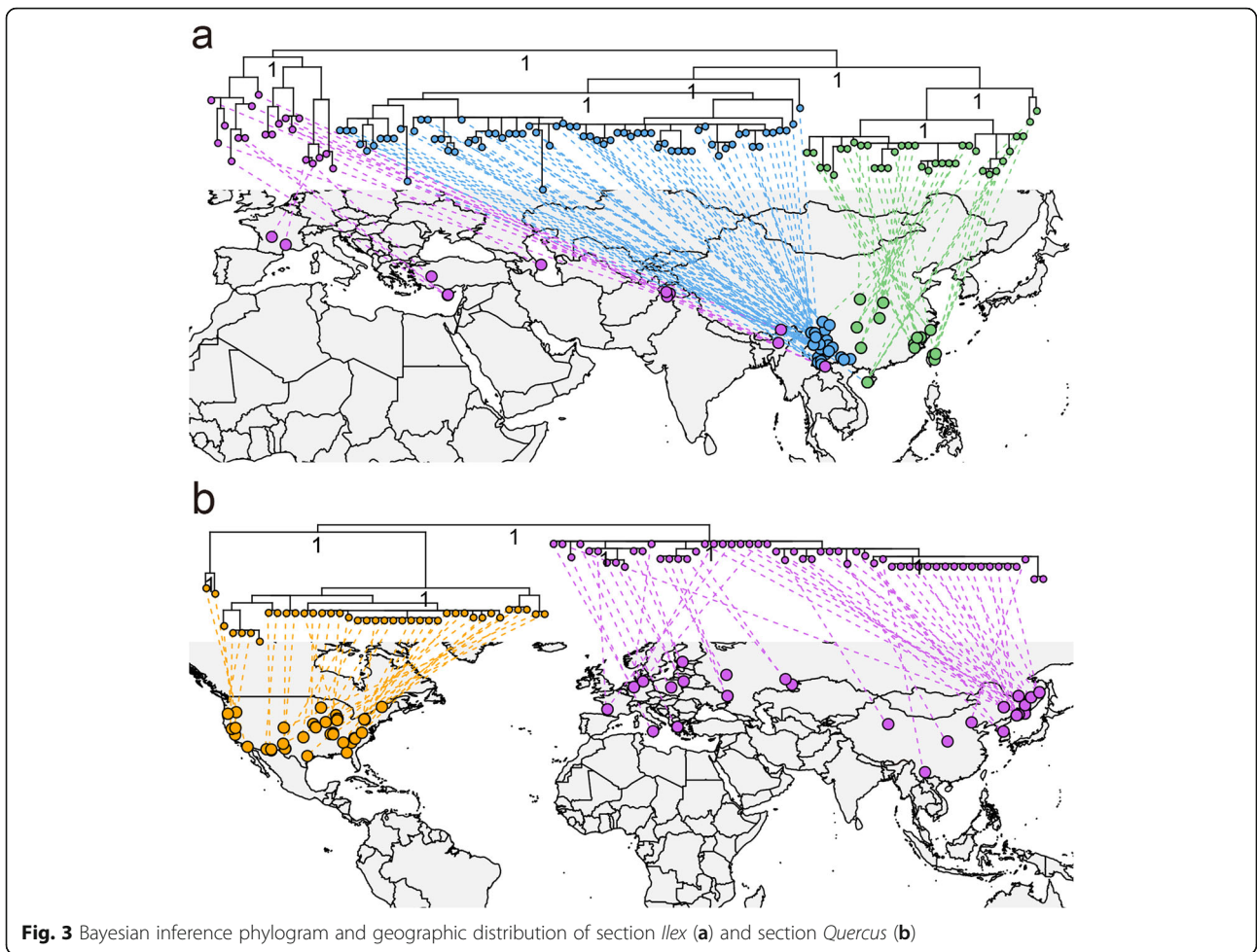


Fig. 3 Bayesian inference phylogram and geographic distribution of section *Ilex* (a) and section *Quercus* (b)

than both sections *Quercus* and *Ilex* (Fig. 4). This indicates their adaptation to a rather narrowly warm and humid environment compared with the other two sections. In contrast, section *Ilex* was distributed in areas with moderate annual mean temperature and both humid to arid

regions (Figs. 4, 5a). The transcontinental section *Quercus* had the widest annual mean temperature range (Fig. 4, 5a), adapted to the winter cold and aridity of Eurasia and northern North America, which corresponding to the highest multidimensional climatic niche diversity (Fig. 5a).

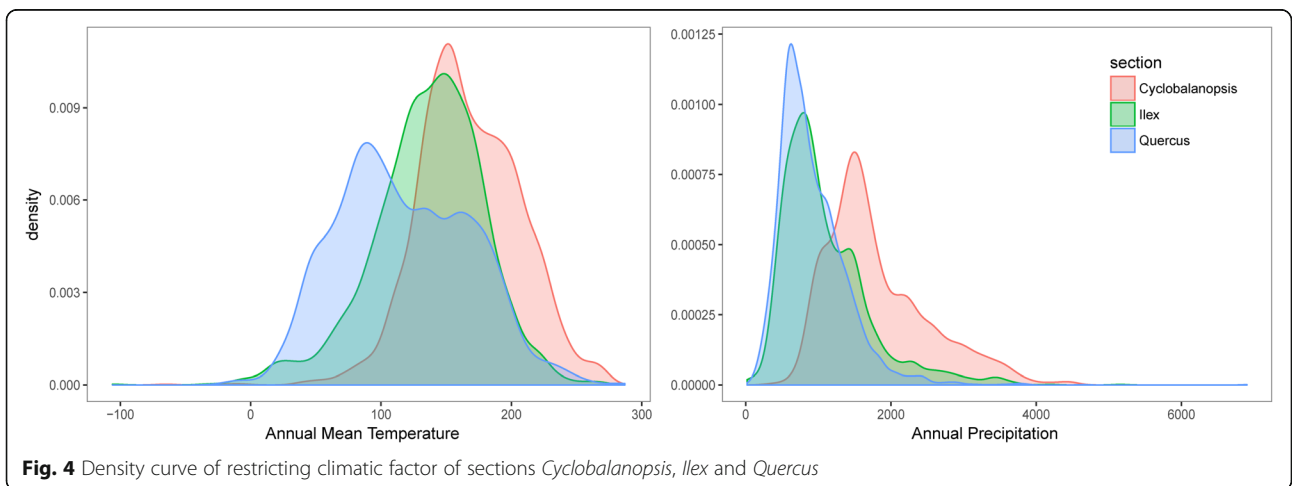
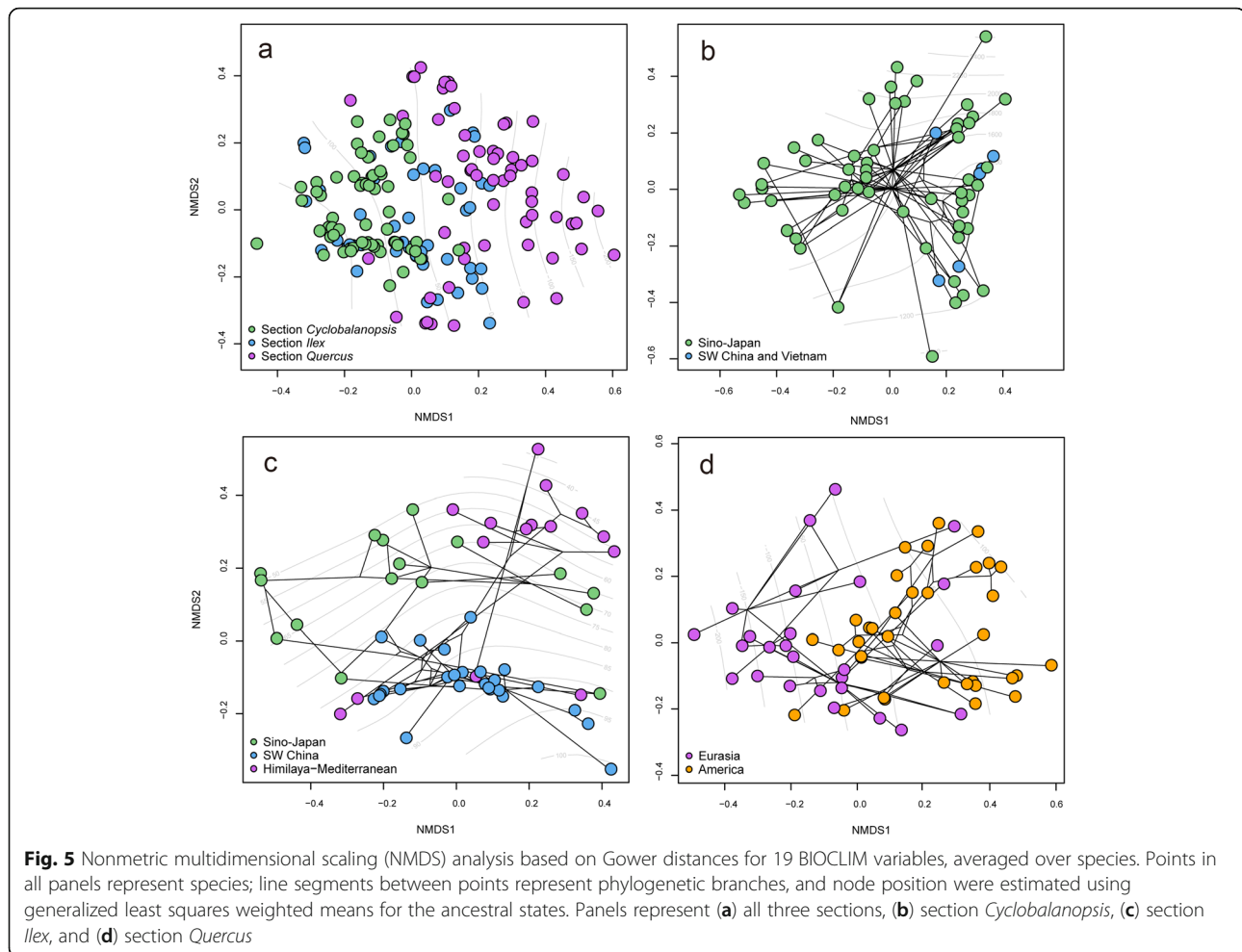


Fig. 4 Density curve of restricting climatic factor of sections *Cyclobalanopsis*, *Ilex* and *Quercus*



Effect of geography and climate on the spatial genetic diversity pattern

Nonmetric multidimensional scaling (NMDS) analysis of section *Cyclobalanopsis* showed limited climatic differentiation between the SW China lineage and the Sino-Japan lineage (Fig. 5b). In contrast, three main climatic clusters were identified within section *Ilex*, corresponding to Sino-Japan, SW China and Himalaya-Mediterranean, which was consistent with its phylogenetic and geographic structure (Fig. 5c). Similarly, section *Quercus* segregated into largely Eurasian and North American clusters in the climatic space with few intermediates (Fig. 5c).

Genetic distances within section *Cyclobalanopsis* and its two lineages were poorly predicted by either geographic or climatic diversity (r^2 in all models >0.05 ; Table 3). In contrast, genetic distances of section *Ilex* were significantly predicted by geographic distances ($r^2 = 0.3603$, $p = 0.001$) but not by climate (Table 3). Genetic distances within the Sino-Japan and the Himalayas-Mediterranean lineages were predicted well by geography and climate. However, the SW China lineage showed weak correlation between genetic distances and geography ($r^2 = 0.033$, $p = 0.022$) and no

correlation with climate (Table 3). Geographic distance ($r^2 = 0.3713$, $p = 0.001$) and climatic distances ($r^2 = 0.1458$, $p = 0.001$) were both strong predictors of genetic variation within section *Quercus*. Within geographic lineages, where plastomes were shared relatively freely, North American genetic distances were well predicted by geography ($r^2 = 0.1471$, $p = 0.001$) and climate ($r^2 = 0.1654$, $p = 0.001$), however, the Eurasian lineage showed no significant correlations (Table 3).

Discussion

Evolution of section *Cyclobalanopsis*

Although section *Cyclobalanopsis* originated from comparable ages as sections *Ilex* and *Quercus*, which is dated to Oligocene [29, 54], the individuals of section *Cyclobalanopsis* show very high genetic similarity and low genetic diversity. Haplotypes are shared widely among species (Additional file 1: Table S1), and most interspecifically shared haplotypes are geographically widespread: haplotype H3 shows particularly wide geographic range and interspecies sharing (Fig. 2b). Haplotype sharing among species across large geographic distances could

Table 3 Mantel test of the relationship between genetic distances and geographic and climatic distances in *Quercus* sections *Cyclobalanopsis*, *Ilex*, and *Quercus* (Him-Med: the Himalayas-Mediterranean)

Sections	Lineage	Genetic~geo	Genetic~ clim	Genetic~ geo + clim
<i>Cyclobalanopsis</i>	all	$r^2 = 0.0005$	$r^2 = 0.0007$	$r^2 = 0.0007$
	Sino-Japan	$r^2 = 0.0005$	$r^2 = 0.0008$	$r^2 = 0.0009$
	SW China	$r^2 = 0.0144$	$r^2 = 0.0122$	$r^2 = 0.0155$
<i>Ilex</i>	all	$r^2 = 0.3603^{**}$	$r^2 = 0.0038$	$r^2 = 0.3608^{**}$
	Sino-Japan	$r^2 = 0.1563^{**}$	$r^2 = 0.2363^{**}$	$r^2 = 0.2390^{**}$
	SW China	$r^2 = 0.0337^*$	$r^2 = 0.0178$	$r^2 = 0.0337$
	Him-Med	$r^2 = 0.1333^*$	$r^2 = 0.1546^*$	$r^2 = 0.1661$
<i>Quercus</i>	all	$r^2 = 0.3713^{**}$	$r^2 = 0.1458^{**}$	$r^2 = 0.3880^{**}$
	Eurasia	$r^2 = 0.0293$	$r^2 = 0.0088$	$r^2 = 0.0223$
	North America	$r^2 = 0.1471^{**}$	$r^2 = 0.1654^{**}$	$r^2 = 0.1799^{**}$

* $p < 0.05$; ** $p < 0.01$

be a result of incomplete lineage sorting [55–57]. In turn, incomplete lineage sorting could result from rapid radiation during early diversification, random sharing or extinction of genotypes, and/or extremely large effective population sizes [58]. In section *Cyclobalanopsis*, interspecific shared haplotypes appear to postdate the early diversification of the lineage, thus making this explanation relatively unlikely. Moreover, the chloroplast DNA is haploid; therefore, its effective size is very small [59]. Furthermore, the haplotypes of section *Cyclobalanopsis* are not randomly shared. Haplotype H3 is shared by 50% of individuals, and interspecific shared haplotypes (H2–H3, H5–H7, H10, and H21) are shared by 69.6% of individuals. Therefore, incomplete lineage sorting is unlikely explain the large-scale sharing of haplotypes.

Alternatively, the high genetic similarity and large-scale haplotype sharing that was observed within section *Cyclobalanopsis* could be a result of selective sweeps on the chloroplast. This is supported by the results of the conducted neutrality tests (Table 2). Significantly negative Tajima's D and Fu and Li's F (based on coalescent) indicate a recent selective sweep, as well as population expansion after a recent bottleneck [60]. However, Fay and Wu's H is suited to distinguish variation due to recent population expansion or direction sweep [60]. A negative Fay and Wu's H indicates an excessive high-frequency derived single nucleotide polymorphisms [61], suggesting a recent selective sweep in section *Cyclobalanopsis*. The majority of individuals in this study (viz, the Sino-Japan lineage) are genetically similar and dominated by few haplotypes, which is exactly as expected under purifying selection. The chloroplast genome can undergo trans-specific selective sweeps, as previously reported, e.g., in *Salix* [58], *Begonia* [62] and *Silene* [63]. Weak reproductive isolation in section *Cyclobalanopsis* [64–66], likely facilitates plastid capture [67, 68], which is a first step in spread of a selected plastome.

Both the timing and geography in section *Cyclobalanopsis* could then have intensified the selection for attenuated plastome diversity. Diversification and range expansion of section *Cyclobalanopsis* from SW China-Indochina to East Asia were driven by a combination of climatic and orogenic factors, initiated by the extrusion of Indochina since the Oligocene [69–71] and an intensification of East Asian monsoon [72] since the Miocene [36, 73]. The current distribution of section *Cyclobalanopsis* is relatively uniform in climate. Section *Cyclobalanopsis* is a dominant woody lineage of warm and humid Asian (sub) tropical evergreen broadleaved forests [73] (Fig. 4, 5b). Thus, the history of diversification combined with the present distribution may have selected for narrowed plastid diversity, since many functionally important genes are encoded by the chloroplast [68]. Moreover, a cyto-nuclear linkage disequilibrium has been reported in plants [74, 75]. Therefore, selection for nuclear genes may also have resulted in selective sweeps on the plastid. While the identified phylogeny shows no clear correlation with climate (Fig. 5b, Table 3), it remains to be seen whether any functional plastome variation is associated with climatic gradients in this section.

Evolution of section *Ilex*

While the genetic homogeneity of section *Cyclobalanopsis* implicates selection as a driving force in its plastid evolution, section *Ilex* exhibits the hallmarks of geographic isolation. Its three lineages correspond to Sino-Japan, SW China and the Himalayas-Mediterranean. Moreover, geography explains a significant proportion of genetic diversity both among ($r^2 = 0.3603$) and within ($r^2 = 0.0337$ – 0.1563) these lineages (Table 3), supporting previous findings in favor of plastid geographic differentiation in this section [40, 43, 47]. Geographic isolation in section *Ilex* is strongly congruent with geological and paleobotanical data. Fossils of section *Ilex* in Oligocene-Miocene sediments from East Asia [76, 77], Himalayas

[78, 79], Asian minor [80, 81], Middle East [80, 81] and Mediterranean [82–86] suggest that the ancestral lineage of section *Ilex* was once widespread throughout East Asia and Tethys-Paratethys. This widespread lineage was then segmented gradually. Initially, orogeny of the Himalayas beginning in the Middle Eocene [87–90] could explain the split between the Asian lineage and the Himalayas-Mediterranean lineage [39]. The uplift of the Hengduan Mountains and the associated cooling from the middle Miocene onward, appears to have triggered the divergence of the Sino-Japan lineage and the SW China lineage [40, 47]. These barriers are particularly likely to have affected plastid genetic diversity, since the chloroplast is maternally inherited in oaks [91], and Eurasian dispersal of oak acorns is mainly by gravity, stream, rodents and Corvids [92–94], which is not very efficient for a long distance dispersal for a species with highly fragmented distribution. Moreover, the seeds of oaks are typical recalcitrant, and thus have a very short viability if they become desiccated [95], which further limits their potential dispersal range. As a result, physical barriers encountered by section *Ilex* were particularly effective at differentiation of the main lineages.

Local adaptation may also have shaped the distribution of plastid diversity as seen in this section. Section *Ilex* has high ecological diversity, ranging from semi-savanna maquis in pan-Mediterranean, semi-arid subalpine scrubs and deserts of the Himalayas, and evergreen broadleaved forests of East Asia and Northern Indochina [42]. This work also demonstrates that section *Ilex* has a wide climatic niche (Figs. 4, 5a). Meanwhile, individuals of section *Ilex* form corresponding clusters in climatic ordinations (Fig. 5c), suggesting that geographic differentiation may have worked in concert with the climatic selection to shape the genetic divergences of today. In addition, the genetic diversities of the Sino-Japan and Himalayas-Mediterranean lineages can be significantly predicted by both geography and climate (Table 3). Within the Sino-Japan lineage of section *Ilex*, three main sublineages (South China, East China and Central China) were inferred (Fig. 3a). Similarly, three to four distinct sublineages were identified in the individuals of section *Ilex* from the Mediterranean [42, 43]. The main interruptions of haplotype migration coincide with the mountain ranges of Anatolia, Greece and the Balkans, the Sea of Sardinia and the Libyan Sea [43]. It may be that geological barriers resulted in allopatric diversification, and that local adaptation contributed to genetic differentiation within lineages.

Evolution of section *Quercus*

Section *Quercus* is widely distributed in European and East Asian forests, and in wetlands of the upper Midwestern USA to droughty mountains of the arid southwest. In North America alone, they range from southern Canada

to Central America [26, 96, 97]. This study demonstrates that this section exhibits a particularly wide climatic range (Fig. 4, 5b), as has been previously suggested in a sampling of the American taxa [29]. The genetic diversity of section *Quercus*, however, is lower than that of section *Ilex*. This is perhaps a result of the geographic range of white oak was strongly impacted by Quaternary climatic fluctuations. During Quaternary glaciation periods, the continental ice shield covered much of Europe and North America [98], which led to both widespread extinction and biotic shifts. In Europe, white oaks retreated to southern refugia, e.g., Iberian Peninsula, southern Italian Peninsula and the southern Balkan Peninsula [99]. Extensive continental ice dominated North America for the 2.4 Myr of the Pleistocene [100], and the eastern North American oak diversity was limited to the unglaciated southeastern portion of the continent [101–103]. Cold temperatures drove the local extinction of northern populations of these white oak species, presumably reducing genetic diversity. In addition, during inter-glacial and post-glacial recolonization, their genetic diversity decreased due to continuous bottleneck effects [91, 100]. In contrast, the European species of section *Ilex* are mainly distributed around the pan-Mediterranean, where they were less impacted by Pleistocene climatic fluctuations and could thus maintain their population sizes and higher genetic diversity.

The plastomes of the white oak derive from ancient and rapid diversification into four major plastome lineages in sections *Protobalanus* (not sampled here) and *Quercus* in only the plastome data, while nuclear data show clear monophyly of the sections [50, 104]. Moreover, the plastome resolves into reciprocally monophyletic lineages, which corresponding to the Eurasian and American white oaks distribution in section *Quercus*. This result seems most plausibly raised from long-term geographic isolation of the two groups of white oaks due to submersion of the North Atlantic land bridge and Bering land bridge since the late Neogene [105]. Climate cooling [106] and Central Asian aridity [107–109] since the middle Miocene are likely causes the vicariance between the European and East Asian white oak groups.

However, the genetic variation of the Eurasian lineage does not correlate with geography or climate (Table 3). The European white oaks typically exhibit strong geographic structure [33, 49] due to founder effects during post-glacial recolonization [33, 48, 49]. In contrast, the genetic variation of the North American lineage shows a significant correlation to geography and climate (Table 3), which is in large part, due to the geographic and reproductive isolation of the Californian species from the eastern North American species [50]. However, white oaks in Eastern North America do not demonstrate chloroplast genetic structure, presumably due to their capacity

for long-distance gene flow, their large population sizes and their relative continuous Pleistocene refugia [50].

Factors resulting in different spatial genetic structure of sections *Cyclobalanopsis*, *Ilex* and *Quercus*

The disparate spatial genetic patterns of these oak clades appear to result from variance in their ecological / climatic tolerances, migration histories, the strength of climatic selection and physiological traits. Climatic tolerance is likely the easiest of these to characterize. Section *Cyclobalanopsis* is restricted to warm and humid subtropical Asia and has an almost continuous distribution within its range. While section *Ilex* has a wider climatic range, it is still unsuited to the dry and cold climate of central Asia, which shapes its disjunct distribution in East Asian-Himalayas and Mediterranean. White oaks, in contrast, have adapted to a wide spectrum of environments and have the widest geographic distribution, comprising the most diverse and productive (in terms of number of species and total biomass) woody plant genus of North America [25]. From the Cretaceous to the Late Neogene, both the Bering Land Bridge and the North Atlantic Land Bridge served as filters that allowed specific plant taxa to migrate between Eurasia and the Americas [110, 111]. White oaks were well adapted to cool climates and thus distributed at high latitudes. Consequently, they were capable to cross the land bridge(s) during the latest Miocene, as inferred by pollen found on Iceland and Greenland [111]. Thus, the climatic niche breadth of white oaks contributed to their wide distribution in throughout the Northern Hemisphere, in contrast to other sections.

In addition, adaptive traits can induce varying responses to similar environment changes. Several species of section *Ilex* from East Asia share similar geographic distributions with section *Cyclobalanopsis*. Both sections originated during the late Eocene and experienced similar topological and climate changes in the Himalayas-East Asia [54, 73]. However, section *Ilex* exhibits clear geographic differentiation, while section *Cyclobalanopsis* does not. As discussed above, this may be partially explained by selective sweeps, which affect the chloroplast genome of section *Cyclobalanopsis* and results rapid migration of selected haplotypes. However, varying adaptive traits, e.g. reproductive capacity and seed dispersal distance may account for their different spatial genetic structure. The geographic distribution of section *Cyclobalanopsis* is mostly concentrated in low to middle elevation montane areas with rugged topography [112]. Therefore, gene flow is not strongly structured by topography. In contrast, most species of section *Ilex* inhabit highly fragmented landscapes, e.g. sky-island such as subalpine areas or river gorges in mountains [112], which disfavors long distance seed dispersal. Furthermore, sections *Cyclobalanopsis* and *Ilex* have different life-forms and tradeoffs between vegetative and reproductive growth. For example, most species of section

Ilex are smaller trees and shrubs, and produce fewer seeds than species of section *Cyclobalanopsis* [112], and have prominent mast years. Therefore, maternal gene flow is more limited in section *Ilex* than in section *Cyclobalanopsis*, and more influenced by both local adaptation and genetic drift, thus increasing their genetic spatial structure.

Comparison of plastome and nuclear genetic diversity patterns

The plastome genetic diversity pattern of *Quercus* does not always match the pattern of nuclear genetic diversity. Section *Cyclobalanopsis* has high species richness [112, 113], which is influenced by environmental variables, mostly by water availability [113]. However, the plastome genetic diversity of section *Cyclobalanopsis* is very low and could not be predicted by climatic factors. While, the spatial phylogenetic structures of sections *Ilex* and *Quercus* were confirmed by nuclear DNA. Based on RADseq nuclear markers, the species of sections *Ilex* from Mediterranean-Himalaya formed a monophyletic clade, though species from Sino-Himalaya and Sino-Japan mixed together [54]. Eurasian white oaks and North American white oaks formed two monophyletic clades, which were supported by RADseq nuclear markers [29, 114]. Furthermore, the nuclear genetic variation of white oaks could partly predict the plastome genetic variation ($r^2 = 0.311$, $p < 0.002$) [50].

Although nuclear genetic variation may show some similarity to plastome markers, they had very different evolution histories. The chloroplast is maternally inherited in oaks, and can only be transferred through seeds [91], while nuclear DNA is biparentally inherited. Therefore, differences between plastome and nuclear genetic structure reflects the efficiency of seed-induced gene flow and pollen-induced gene flow, respectively, which has also been reported for other oak species as well [39, 115–117]. Moreover, the whole chloroplast genome could be transferred among different species during hybridization processes, since hybridization frequently occurred among oaks of the same section [66, 114, 118]. However, only half of the nuclear homologues can be passed to the offspring. Furthermore, few nuclear genes can penetrate among species due to the natural selection. As a result, nuclear DNA is less impacted by hybridization. All these factors can lead to different genetic structure between plastome DNA and nuclear DNA in oaks.

Conclusions

This study clarifies the individual diversification history of the sections *Cyclobalanopsis*, *Ilex*, and *Quercus*, as manifest by their chloroplast diversity. The results highlight the importance of geological events and ecological adaptive capacity for the spatial genetic pattern of oak clades and provides detailed insights into the formation mechanism of their contemporary diversity. Further insights into the divergence

history of this groups will originate from a combination of whole-chloroplast sequencing and nuclear genetic data of deeper population sampling. Finally, association mapping can be used to investigate the relationship between genetic polymorphisms and environment, which will help to identify the relative effects of the climatic, edaphic variation, and migration history on genetic variation in multiple clades.

Methods

Sampling and molecular biology experiments

This study included 147 individuals from 29 species of *Quercus* section *Cyclobalanopsis* from China, Japan, Vietnam and Nepal as well as 121 individuals of 28 species of section *Ilex* from China, Himalayas, Azerbaijan, and Mediterranean. One individual of *Lithocarpus henryi* was used as outgroup to root the tree of genus *Quercus*. These data were obtained from our previous study [119], including four chloroplast DNA regions: *atpI-atpH*, *matK*, *psbA-trnH* and *ycf1*. Eighty-nine individuals from 27 species of section *Quercus* from Eurasia and North America were included in this study. Among these, seven individuals of the six species were obtained from our previous study [119]. The remaining samples (82 individuals and 21 species) were obtained from a previous study [50]. The four *cpDNA* regions were aligned using Muscle [120] implemented in MEGA 7.0.21 [121] (available at <http://www.megasoftware.net/>) and the extra end of long sequences were trimmed. An inversion of 60 bp found in the *psbA-trnH* region was replaced with its reverse complement (to maximize homology in the phylogenetic analyses). Relevant data on samples and sequences used in this study are included in the additional file (Additional file 1: Table S1).

Genetic diversity and phylogenetic analysis

Haplotypes were extracted by using DnaSP 5.10 [122]. Genetic diversity pattern were profiled using variable sites, average pairwise differences per base pair between sequences (nucleotide diversity) [123], haplotype diversity (Hd), Fst in DnaSP 5.10 [122] and *p*-distances in MEGA 7.0.21 [121] (available at <http://www.megasoftware.net/>). Genetic distances were calculated in GenAlEx 6.5 [124]. Neutrality tests (Tajima's *D*, Fu and Li's *D*, Fu and Li's *F* and Fay and Wu's *H*) were conducted using DnaSP 5.10 [122].

Bayesian trees were constructed using MrBayes 3.2.6 [125]. The nucleotide substitution model was selected by Modeltest 3.7 [126] based on the Akaike information criterion (AIC). Two parallel Markov chain Monte Carlo (MCMC) runs were performed for 20 million generations. The trees were sampled every 1000 generations and inspected via Tracer 1.6 (available at <http://tree.bio.ed.ac.uk/software/tracer/>) to ensure effective sample size (ESS) exceeding 200. The first 15% of trees were discarded as burn-in. Phylogenetic trees were plotted on the world map using R package phytools [127].

Geographic and climatic data

Georeferenced occurrence records were collected for each section from public specimen databases and the publications listed below and were augmented by our own collections. Data were initially cleaned by removing all records that were outside of published range records. Coordinates of these specimens were then used to extract 19 BIOCLIM variables from WorldClim 1.4: Current conditions (~ 1960–1990) [128] using the raster [129] and dismo packages [129] in R. Geographic records of sections *Cyclobalanopsis*, *Ilex* and *Quercus* were obtained from our field collection database, Chinese Virtual Herbarium (<http://www.cvh.ac.cn>), Global Biodiversity Information Facility (<https://www.gbif.org>) and National Herbarium of the Netherlands (<http://herbarium.naturalis.nl/nhn/explore>). Additional geographic occurrence data were added from previous studies for section *Cyclobalanopsis* [36–38, 130], section *Ilex* [39, 43, 47, 131], and section *Quercus* [29, 50, 130, 132–134]. The sampling records of sections *Ilex* and *Quercus* were uneven, which would have affect climate analysis. Therefore, the dense distribution records of sections *Ilex* and *Quercus* were filtered to ensure that the geographic records were even between main regions using the R packages raster [129] and dismo [129]. The filtering of section *Ilex* in west Europe retained only one occurrence record in each raster of 0.45 km × 0.45 km, but one record for each raster of 0.35 km × 0.35 km in SW China. The filtering resolution of section *Quercus* in Europe and North America retained one record in each raster of 0.04 km × 0.04 km. Records of each section were plotted on a map and the number of records were counted to ensure that the point densities were even between different regions.

Multicollinearity among 19 BIOCLIM variables was examined using a Pearson correlation matrix estimated in the R package psych [135]. The subsets of variables with high correlations ($r > 0.8$) were reduced to single variables. The restricted climatic variable for each section was calculated in MAXENT 3.4.1 [136] based on the reduced subsets of variables. Density curves of restricted climatic variables were plotted using the R package ggplot2 [137].

Modeling the effect of geography and climate on chloroplast phylogeny

Nonmetric multidimensional scaling (NMDS) was used on all 19 BIOCLIM variables to characterize the climatic niche for the three oak sections (Additional file 1: Table S2). Ordination was conducted on a Gower distance matrix which could handle variables with different physical units and of mixed precision levels [138]. NMDS was conducted from $K = 1$ to $K = 10$, and stress was plotted to select the best-fit number of dimensions for ordination. For visualization, tree topologies were projected into the two-dimensional

($K = 2$) ordination space. Analyses were conducted using the R packages *vegan* [139] and *phytools* [127].

Mantel and multiple mantel tests were performed in the R package *phytools* [127] using the following regression models: genetic distances predicted by geographic distances ($\Delta_{\text{genetic}} \sim \Delta_{\text{geo}}$); genetic distances predicted by climatic distances ($\Delta_{\text{genetic}} \sim \Delta_{\text{clim}}$); genetic variation predicted by geographic and climatic distances ($\Delta_{\text{genetic}} \sim \Delta_{\text{geo}} + \Delta_{\text{clim}}$). GenAlEx 6.5 [124] was used to calculate the genetic distance and the geographic distance of each section. Gower distances of 19 BIOCLIM variables were calculated in R using *vegan* [139].

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12862-019-1523-z>.

Additional file 1. Detailed information and data.

Additional file 2. Distribution of shared haplotypes of section *Ilex* (a) and section *Quercus* (b).

Abbreviations

SW: Southwest

Acknowledgements

We thank anonymous reviewers for their insightful and valuable suggestions/comments for further improving the manuscript.

Authors' contributions

MD conceived and designed the research. MXY, RBL and YL analyzed data. MXY, AH and MD wrote the paper. MXY, YL, RBL and YSX collected georeferenced occurrence records. All authors read and approved the final manuscript.

Funding

This work was supported by grants from the National Natural Science Foundation of China (31972858), the Shanghai Municipal Administration of Forestation and City Appearances (G172406, G162404, G182417, G182427), Southeast Asia Biodiversity Research Institute, Chinese Academy of Sciences (Y4ZK111B01). These funding bodies did not have a role in the design of the study, the collection, analysis, and interpretation of data, or in writing the manuscript.

Availability of data and materials

DNA sequences are from references [50, 119]. Sampling distribution and climate data are uploaded as additional files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that there are no competing interests.

Author details

¹Shanghai Chenshan Plant Science Research Center, Chinese Academy of Sciences, Shanghai Chenshan Botanical Garden, Shanghai 201602, China. ²Southeast Asia Biodiversity Research Institute, Chinese Academy of Sciences, Yezin, Nay Pyi Taw 05282, Myanmar. ³College of Life Sciences, Shanghai Normal University, Shanghai 200234, China. ⁴The Ecological Technique and Engineering College, Shanghai Institute of Technology, Shanghai 201418, China. ⁵The Morton Arboretum, 4100 Illinois Route 53, Lisle, IL 60532, USA. ⁶The Field Museum, 1400 S Lake Shore Drive, Chicago, IL 60605, USA.

Received: 30 January 2019 Accepted: 1 October 2019

Published online: 04 November 2019

References

1. Parmesan C. Ecological and evolutionary responses to recent climate change. *Annu Rev Ecol Evol S.* 2006;37(1):637–69.
2. Scheffers BR, De ML, Bridge TC, Hoffmann AA, Pandolfi JM, Corlett RT, Butchart SH, Pearce-Kelly P, Kovacs KM, Dudgeon D. The broad footprint of climate change from genes to biomes to people. *Science.* 2016;354(6313):aaf7671.
3. He K, Jiang XL. Sky islands of Southwest China. I: an overview of phylogeographic patterns. *Chin Sci Bull.* 2014;59(7):585–97.
4. Mosca E, Eckert A, Di Piero E, Rocchini D, La Porta N, Belletti P, Neale D. The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. *Mol Ecol.* 2012;21(22):5530–45.
5. Gugger PF, Ikegami M, Sork VL. Influence of late Quaternary climate change on present patterns of genetic variation in valley oak, *Quercus lobata* nee. *Mol Ecol.* 2013;22(13):3598–612.
6. Arenas M, Ray N, Currat M, Excoffier L. Consequences of range contractions and range shifts on molecular diversity. *Mol Biol Evol.* 2011;29(1):207–18.
7. Mona S, Ray N, Arenas M, Excoffier L. Genetic consequences of habitat fragmentation during a range expansion. *Heredity.* 2014;112(3):291.
8. Arenas M, Mona S, Trochet A, Sramkova Hanulova A, Currat M, Ray N, Chikhi L, Rasteiro R, Schmeller D, Excoffier L: The scaling of genetic diversity in a changing and fragmented world. In: *Scaling in Ecology and Biodiversity Conservation*. Edited by Henle K, Potts S, Kunin W, Matsinos Y, Similä J, Pantis J, Grobelenik V, Penev L, Settele J. Sofia: Pensoft Publishers; 2014: 55–60.
9. Hewitt G. The genetic legacy of the quaternary ice ages. *Nature.* 2000; 405(6789):907–13.
10. Comes HP, Kadereit JW. The effect of quaternary climatic changes on plant distribution and evolution. *Trends Plant Sci.* 1998;3(11):432–8.
11. Favre A, Päckert M, Pauls SU, Jähnig SC, Dieter U, Ingo M, Muellner-Riehl AN. The role of the uplift of the Qinghai-Tibetan plateau for the evolution of Tibetan biotas. *Biol Rev.* 2015;90(1):236–53.
12. Aitken SN, Yeaman S, Holliday JA, Wang T, Curtismcclane S. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evol Appl.* 2008;1(1):95–111.
13. Barrett R, Schluter D. Adaptation from standing genetic variation. *Trends Ecol Evol.* 2008;23(1):38.
14. Cavender-Bares J, Ramírez-Valiente JA. Physiological evidence from common garden experiments for local adaptation and adaptive plasticity to climate in American live oaks (*Quercus* Section *Virentes*): implications for conservation under global change. In: Gil-Pelegrín E, Peguero-Pina JJ, Sancho-Knapik D, editors. *Oaks Physiological Ecology. Exploring the Functional Diversity of Genus Quercus*: Springer; 2017. p. 107–35.
15. Bay RA, Rose N, Barrett R, Bernatchez L, Ghalambor CK, Lasky JR, Brem RB, Palumbi SR, Ralph P. Predicting responses to contemporary environmental change using evolutionary response architectures. *Am Nat.* 2017;189(5):463–73.
16. Bay RA, Harrigan RJ, Underwood VL, Gibbs HL, Smith TB, Ruegg K. Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science.* 2018;359(6371):83–6.
17. Marcott SA, Mix AC. A reconstruction of regional and global temperature for the past 11,300 years. *Science.* 2013;339(6124):1198–201.
18. Li L, Abbott RJ, Liu BB, Sun YS, Li LL, Zou JB, Wang X, Miehle G, Liu JQ. Pliocene intraspecific divergence and Plio-Pleistocene range expansions within *Picea likiangensis* (Lijiang spruce), a dominant forest tree of the Qinghai-Tibet plateau. *Mol Ecol.* 2013;22(20):5237.
19. Petit RJ, Hampe A. Some evolutionary consequences of being a tree. *Annu Rev Ecol Evol S.* 2006;37:187–214.
20. Sork VL, Davis FW, Westfall R, Flint A, Ikegami M, Wang H, Grivet D. Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata* Née) in the face of climate change. *Mol Ecol.* 2010;19(17):3806–23.
21. Somero GN, Barnes B, Gordon M, Sato K, Hoppeler H. The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine 'winners' and 'losers'. *J Exp Biol.* 2010;213(6):912–20.
22. Ortego J, Riordan EC, Gugger PF, Sork VL. Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. *Mol Ecol.* 2012;21(13):3210–23.

23. Petit RJ, Carlson J, Curtu AL, Loustau M-L, Plomion C, González-Rodríguez A, Sork V, Ducouso A. Fagaceae trees as models to integrate ecology, evolution and genomics. *New Phytol.* 2013;197(2):369–71.
24. Cavender-Bares J. Diversity, distribution, and ecosystem services of the north American oaks. *Int Oaks.* 2016;27:37–48.
25. Cavender-Bares J. Diversification, adaptation, and community assembly of the American oaks (*Quercus*), a model clade for integrating ecology and evolution. *New Phytol.* 2019;221:669–92.
26. Nixon KC. Global and neotropical distribution and diversity of oak (genus *Quercus*) and oak forests. In: Ecology and conservation of neotropical montane oak forests. Edited by Kappelle M. Berlin/Heidelberg, Germany: Springer Verlag; 2006: 3–13.
27. Denk T, Grimm GW, Manos PS, Deng M, Hipp AL. An updated infrageneric classification of the oaks: review of previous taxonomic schemes and synthesis of evolutionary patterns. In: Gil-Pelegrín E, Peguero-Pina JJ, Sancho-Knapik D, editors. *Oaks Physiological Ecology. Exploring the Functional Diversity of Genus Quercus*. L: Springer; 2017. p. 13–38.
28. Denk T, Grimm GW. The oaks of western Eurasia: traditional classifications and evidence from two nuclear markers. *Taxon.* 2010;59(2):351–66.
29. Hipp AL, Manos PS, González-Rodríguez A, Hahn M, Kaproth M, McVay JD, Avalos SV, Cavender-Bares J. Sympatric parallel diversification of major oak clades in the Americas and the origins of Mexican species diversity. *New Phytol.* 2018;217(1):439–52.
30. Cavender-Bares J, Ackerly D, Baum D, Bazzaz F. Phylogenetic overdispersion in Floridian oak communities. *Am Nat.* 2004;163(6):823–43.
31. Hubert F, Grimm GW, Jousseil E, Bery V, Franc A, Kremer A. Multiple nuclear genes stabilize the phylogenetic backbone of the genus *Quercus*. *Syst Biodivers.* 2014;12(4):405–23.
32. Petit RJ, Bodénès C, Ducouso A, Roussel G, Kremer A. Hybridization as a mechanism of invasion in oaks. *New Phytol.* 2003;161(1):151–64.
33. Dumolin-Lapègue S, Demesure B, Fineschi S, Le Come V, Petit RJ. Phylogeographic structure of white oaks throughout the European continent. *Genetics.* 1997;146(4):1475–87.
34. Petit RJ, Excoffier L. Gene flow and species delimitation. *Trends Ecol Evol.* 2009;24(7):386–93.
35. Gerber S, Chadoeuf J, Gugerli F, Lascoux M, Buiteveld J, Cottrell J, Dounavi A, Fineschi S, Forrest LL, Fogelqvist J, et al. High rates of gene flow by pollen and seed in oak populations across Europe. *PLoS One.* 2014;9(1):e85130.
36. Xu J, Deng M, Jiang XL, Westwood M, Song YG, Turkington R. Phylogeography of *Quercus glauca* (Fagaceae), a dominant tree of east Asian subtropical evergreen forests, based on three chloroplast DNA interspace sequences. *Tree Genet Genomes.* 2015;11(1):805.
37. Xu J, Jiang XL, Deng M, Westwood M, Song YG, Zheng SS. Conservation genetics of rare trees restricted to subtropical montane cloud forests in southern China: a case study from *Quercus arbutifolia* (Fagaceae). *Tree Genet Genomes.* 2016;12(5):90.
38. Jiang XL, An M, Zheng SS, Deng M, Su ZH. Geographical isolation and environmental heterogeneity contribute to the spatial genetic patterns of *Quercus kerrii* (Fagaceae). *Heredity.* 2017;120(3):219–33.
39. Du FK, Hou M, Wang WT, Mao KS, Hampe A. Phylogeography of *Quercus aquifolioides* provides novel insights into the Neogene history of a major global hotspot of plant diversity in south-West China. *J Biogeogr.* 2016;44(2): 294–307.
40. Feng L, Zheng QJ, Qian ZQ, Yang J, Zhang YP, Li ZH, Zhao GF. Genetic structure and evolutionary history of three alpine sclerophyllous oaks in east Himalaya-Hengduan Mountains and adjacent regions. *Front Plant Sci.* 2016;7:1688.
41. Feng L, Zhang YP, Chen XD, Yang J, Zhou T, Bai GQ, Yang J, Li ZH, Peng Cl, Zhao GF. Allopatric divergence, local adaptation, and multiple Quaternary refugia in a long-lived tree (*Quercus spinosa*) from subtropical China. *bioRxiv.* 2017:112375. <https://doi.org/10.1101/112375>
42. Simeone MC, Grimm GW, Papini A, Vessella F, Cardoni S, Tordoni E, Piredda R, Franc A, Denk T. Plastome data reveal multiple geographic origins of *Quercus* group Ilex. *PeerJ.* 2016;4:e1897.
43. Vitelli M, Vessella F, Cardoni S, Pollegioni P, Denk T, Grimm GW, Simeone MC. Phylogeographic structuring of plastome diversity in Mediterranean oaks (*Quercus* group Ilex, Fagaceae). *Tree Genet Genomes.* 2017;13(1):3.
44. de Heredia UL, Jiménez P, Collada C, Simeone MC, Bellarosa R, Schirone B, Cervera MT, Gil L. Multi-marker phylogeny of three evergreen oaks reveals vicariant patterns in the Western Mediterranean. *Taxon.* 2007;56(4):1209.
45. Lumaret R, Mir C, Michaud H, Raynal V. Phylogeographical variation of chloroplast DNA in holm oak (*Quercus ilex* L.). *Mol Ecol.* 2002;11(11):2327–36.
46. Toumi L, Lumaret R. Genetic variation and evolutionary history of holly oak: a circum-Mediterranean species-complex [*Quercus coccifera* L./*Q. calliprinos* (Webb) Holmboe, Fagaceae]. *Plant Syst Evol.* 2010;290(1–4):159–71.
47. Meng HH, Su T, Gao XY, Li J, Jiang XL, Sun H, Zhou ZK. Warm-cold colonization: response of oaks to uplift of the Himalaya-Hengduan Mountains. *Mol Ecol.* 2017;26(12):3276–94.
48. Petit RJ, Pineau E, Demesure B, Bacilieri R, Ducouso A, Kremer A. Chloroplast DNA footprints of postglacial recolonization by oaks. *Proc Natl Acad Sci.* 1997;94(18):9996–10001.
49. Duminil J, Pemonge MH, Petit RJ. A set of 35 consensus primer pairs amplifying genes and introns of plant mitochondrial DNA. *Mol Ecol Notes.* 2002;2(4):428–30.
50. Pham KK, Hipp AL, Manos PS, Cronn RC. A time and a place for everything: phylogenetic history and geography as joint predictors of oak plastome phylogeny. *Genome.* 2017;60(9):720–32.
51. Grivet D, Deguilloux M, Petit RJ, Sork VL. Contrasting patterns of historical colonization in white oaks (*Quercus* spp.) in California and Europe. *Mol Ecol.* 2006;15(13):4085–93.
52. Riordan EC, Gugger PF, Ortego J, Smith C, Gaddis K, Thompson P, Sork VL. Association of genetic and phenotypic variability with geography and climate in three southern California oaks. *Am J Bot.* 2016;103(1):73–85.
53. Whittemore AT, Schaal BA. Interspecific gene flow in sympatric oaks. *Proc Natl Acad Sci.* 1991;88(6):2540–4.
54. Jiang XL, Hipp AL, Deng M, Su T, Zhou ZK, Yan MX. East Asian origins of European holly oaks (*Quercus* section *Ilex* Loudon) via the Tibet-Himalaya. *J Biogeogr.* 2019;00:1–12.
55. Avise JC. *Phylogeography: the history and formation of species.* Cambridge, MA: Harvard university press; 2000.
56. Funk DJ, Omland KE. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu Rev Ecol Evol S.* 2003;34(1):397–423.
57. Muir G, Schloetterer C. Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Mol Ecol.* 2005;14(2):549–61.
58. Percy DM, Argus GW, Cronk QC, Fazekas AJ, Kesanakurti PR, Burgess KS, Husband BC, Newmaster SG, Barrett SC, Graham SW. Understanding the spectacular failure of DNA barcoding in willows (*Salix*): does this result from a trans-specific selective sweep? *Mol Ecol.* 2014;23(19):4737–56.
59. Birky CW. Evolution and variation in plant chloroplast and mitochondrial genomes. In: Gottlieb LD, Jain SK, editors. *Plant Evol. Biol.* Netherlands: Springer; 1988. p. 23–53.
60. Hedrick WP. *Genetics of populations.* Sudbury, Mass.: Jones and Bartlett; 2005.
61. Fay JC, Wu CI. Hitchhiking under positive Darwinian selection. *Genetics.* 2000;155(3):1405–13.
62. Twyford AD, Kidner CA, Harrison N, Ennos RA. Population history and seed dispersal in widespread central American *Begonia* species (Begoniaceae) inferred from plastome-derived microsatellite markers. *Bot J Linn Soc.* 2012; 171(1):260–76.
63. Muir G, Filatov D. A selective sweep in the chloroplast DNA of dioecious *Silene* (section *Elisanthe*). *Genetics.* 2007;177(2):1239–47.
64. Tamaki I, Okada M. Genetic admixing of two evergreen oaks, *Quercus acuta* and *Q. sessilifolia* (subgenus *Cyclobalanopsis*), is the result of interspecific introgressive hybridization. *Tree Genet Genomes.* 2014;10(4):989–99.
65. Song YG, Deng M, Hipp AL, Li QS. Leaf morphological evidence of natural hybridization between two oak species (*Quercus austrocochinchinensis* and *Q. kerrii*) and its implications for conservation management. *Eur J For Res.* 2015;134(1):139–51.
66. An M, Deng M, Zheng SS, Jiang XL, Song YG. Introgression threatens the genetic diversity of *Quercus austrocochinchinensis* (Fagaceae), an endangered oak: a case inferred by molecular markers. *Front Plant Sci.* 2017;8:229.
67. Tsitrone A, Kirkpatrick M, Levin DA. A model for chloroplast capture. *Evolution.* 2003;57(8):1776–82.
68. Kapralov MV, Filatov DA. Widespread positive selection in the photosynthetic Rubisco enzyme. *BMC Evol Biol.* 2007;7(1):73.
69. Tapponnier P, Lacassin R, Leloup PH, Schärer U, Zhong D, Wu HW, Liu XH, Ji SC, Zhang LS, Zhong JY. The Ailao Shan/Red River metamorphic belt: tertiary left-lateral shear between Indochina and South China. *Nature.* 1990; 343(6257):431–7.
70. Lacassin R, Maluski H, Leloup PH, Tapponnier P, Hinthong C, Siribhakdi K, Chuaviroj S, Charoenravat A. Tertiary diachronic extrusion and deformation

- of western Indochina: structural and 40Ar/39Ar evidence from NW Thailand. *J Geophys Res Solid Earth*. 1997;102(B5):10013–37.
71. Leloup PH, Arnaud N, Lacassin R, Kienast JR, Harrison TM, Trong TTP, Replumaz A, Tapponnier P. New constraints on the structure, thermochronology, and timing of the Ailao Shan-Red River shear zone, SE Asia. *J Geophys Res Solid Earth*. 2001;106(B4):6683–732.
 72. Wan SM, Li AC, Clift PD, Stuuw JBW. Development of the east Asian monsoon: mineralogical and sedimentologic records in the northern South China Sea since 20Ma. *Palaeogeogr Palaeoclimatol Palaeoecol*. 2007;254(3–4):561–82.
 73. Deng M, Jiang XL, Hipp AL, Manos PS, Hahn M. Phylogeny and biogeography of east Asian evergreen oaks (*Quercus* section *Cyclobalanopsis*; Fagaceae): insights into the Cenozoic history of evergreen broad-leaved forests in subtropical Asia. *Mol Phylogenet Evol*. 2017;119:170–81.
 74. Latta RG, Linhart YB, Mitton JB. Cytonuclear disequilibrium and genetic drift in a natural population of ponderosa pine. *Genetics*. 2001;158(2):843–50.
 75. Fields P, McCauley D, McAssey E, Taylor D. Patterns of cyto-nuclear linkage disequilibrium in *Silene latifolia*: genomic heterogeneity and temporal stability. *Heredity*. 2014;112(2):99.
 76. Xia K, Su T, Liu YS, Xing YW, Jacques FMB, Zhou ZK. Quantitative climate reconstructions of the late Miocene Xiaolongtan megaflora from Yunnan, Southwest China. *Palaeogeogr Palaeoclimatol Palaeoecol*. 2009;276(1–4):80–6.
 77. Li HM, Guo SX. The Miocene flora from Namling of Xizang. *Acta Palaeontol Sin*. 1976;15(1):7–18.
 78. Group of Cenozoic Plants of China. *Cenozoic Plants from China (Fossil Plants of China, Volume 3)*. Beijing: Science Press; 1978.
 79. Su T, Spicer RA, Li SH, Xu H, Huang J, Sherlock S, Huang YJ, Li SF, Wang L, Jia LB, et al. Uplift, climate and biotic changes at the Eocene-Oligocene transition in south-eastern Tibet. *Natl Sci Rev*. 2018;nwy062.
 80. Akhmetiev MA, Zaporozhets NI. Paleogene events in Central Eurasia: their role in the flora and vegetation cover evolution, migration of phytochore boundaries, and climate changes. *Stratigr Geol Correl*. 2014;22(3):312–35.
 81. Akkiraz MS. Vegetation and climate in the Miocene deposits of southern side of the Buyuk Menderes Graben, Şahinali-2 core SW Turkey. *Bull Geosci*. 2011;86(4):859–78.
 82. Ivanov D. Climate and vegetation change during the late Miocene in Southwest Bulgaria based on pollen data from the Sandanski Basin. *Rev Palaeobot Palynol*. 2015;221:128–37.
 83. Denk T, Velitzelos D, Güner TH, Bouchal JM, Grímsson F, Grimm GW. Taxonomy and palaeoecology of two widespread western Eurasian Neogene sclerophyllous oak species: *Quercus drymeja* Unger and *Q. mediterranea* Unger. *Rev Palaeobot Palynol*. 2017;241:98–128.
 84. Velitzelos D, Bouchal JM, Denk T. Review of the Cenozoic floras and vegetation of Greece. *Rev Palaeobot Palynol*. 2014;204:56–117.
 85. Zidianakis G, Mohr BAR, Fassoulas C. A late Miocene leaf assemblage from Vrysses, western Crete, Greece, and its paleoenvironmental and paleoclimatic interpretation. *Geodiversitas*. 2007;29(3):351–77.
 86. Palamarev E, Tsenov B. Genus *Quercus* in the Late Miocene flora of Baldevo formation (Southwest Bulgaria): taxonomical composition and palaeoecology. *Phytol Balcan*. 2004;10(2–3):147–56.
 87. Chen DM, Zhang XX, Kang HH, Sun X, Yin S, Du He YN, Gapare W, Wu HX, Liu CJ. Phylogeography of *Quercus variabilis* based on chloroplast DNA sequence in East Asia: multiple glacial refugia and mainland-migrated island populations. *PLoS One*. 2012;7(10):e47268.
 88. Clift PD, Shimizu N, Layne GD, Blusztajn J. Tracing patterns of erosion and drainage in the Paleogene Himalaya through ion probe Pb isotope analysis of detrital K-feldspars in the Indus Molasse India. *Earth Planet Sci Lett*. 2001;188(3–4):475–91.
 89. Chatterjee S, Goswami A, Scotese CR. The longest voyage: tectonic, magmatic, and paleoclimatic evolution of the Indian plate during its northward flight from Gondwana to Asia. *Gondwana Res*. 2013;23(1):238–67.
 90. Dupont-Nivet G, Lippert PC, Van Hinsbergen DJJ, Meijers MJM, Kapp P. Palaeolatitude and age of the Indo-Asia collision: palaeomagnetic constraints. *Geophys J Int*. 2010;182(3):1189–98.
 91. Dumolin S, Demesure B, Petit RJ. Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *Theor Appl Genet*. 1995;91(8):1253–6.
 92. Xiao ZS, Zhang ZB, Wang YS. Effects of seed size on dispersal distance in five rodent-dispersed fagaceous species. *Acta Oecol*. 2005;28(3):221–9.
 93. Johnson WC, Webb T III. The role of blue jays (*Cyanocitta cristata* L.) in the postglacial dispersal of fagaceous trees in eastern North America. *J Biogeogr*. 1989;561–71.
 94. Knight R. A model of episodic, abiotic dispersal for oaks (*Quercus robur*). *S Afr J Bot*. 1985;51(4):265–9.
 95. Xia K, Daws MI, Hay FR, Chen WY, Zhou ZK, Pritchard HW. A comparative study of desiccation responses of seeds of Asian Evergreen oaks, *Quercus* subgenus *Cyclobalanopsis* and *Quercus* subgenus *Quercus*. *S Afr J Bot*. 2012;78:47–54.
 96. Nixon KC. The genus *Quercus* in Mexico. In: Ramamoorthi R, Bye R, Lot A, Fa J, editors. *Biological diversity of Mexico: origins and distribution*. Oxford: University Press; 1993. p. 447–58.
 97. Huntley B. European vegetation history: Palaeovegetation maps from pollen data-13 000 yr BP to present. *J Quat Sci*. 1990;5(2):103–22.
 98. Zweck C, Huybrechts P. Modeling of the northern hemisphere ice sheets during the last glacial cycle and glaciological sensitivity. *J Geophys Res Atmos*. 2005;110:D07103.
 99. Petit RJ, Brewer S, Bordács S, Burg K, Cheddadi R, Coart E, Cottrell J, Csaiikl UM, van Dam B, Deans JD. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *For Ecol Manag*. 2002;156(1–3):49–74.
 100. Roberts DR, Hamann A. Glacial refugia and modern genetic diversity of 22 western North American tree species. *Proc R Soc London, Ser B*. 2015;282(1804):20142903.
 101. Delcourt PA, Delcourt HR. Vegetation maps for eastern North America, 40,000 years B.P. to the present. In: Ohio RCR, editor. *Plenum Geobotany: Proceedings of Geobotany Conference*. New York: Bowling Green State University; 1981. p. 123–65.
 102. Magni C, Ducousso A, Caron H, Petit RJ, Kremer A. Chloroplast DNA variation of *Quercus rubra* L. in North America and comparison with other Fagaceae. *Mol Ecol*. 2005;14(2):513–24.
 103. Delcourt HR, Delcourt PA. Ice age haven for hardwoods. *Nat Hist*. 1984;93:22–8.
 104. Manos PS, Doyle JJ, Nixon KC. Phylogeny, biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). *Mol Phylogenet Evol*. 1999;12(3):333–49.
 105. Milne IR. Northern hemisphere plant disjunctions: a window on tertiary land bridges and climate change? *Ann Bot*. 2006;98(3):465–72.
 106. Miao YF, Herrmann M, Wu FL, Yan XL, Yang SL. What controlled mid-Late Miocene long-term aridification in Central Asia? -global cooling or Tibetan plateau uplift: a review. *Earth-Sci Rev*. 2012;112(3–4):155–72.
 107. Guo ZT, Peng SZ, Hao QZ, Biscaye PE, An ZS, Liu TS. Late Miocene-Pliocene development of Asian aridification as recorded in the red-earth formation in northern China. *Glob Planet Chang*. 2008;41(3):135–45.
 108. Qiang XK, An ZS, Song YG, Chang H, Sun YB, Liu WG, Ao H, Dong JB, Fu CF, Wu F. New eolian red clay sequence on the western Chinese loess plateau linked to onset of Asian desertification about 25 ma ago. *Sci China Earth Sci*. 2011;54(1):136–44.
 109. Tang ZH, Ding ZL, White PD, Dong XX, Ji JL, Jiang HC, Luo P, Wang X. Late Cenozoic central Asian drying inferred from a palynological record from the northern Tian Shan. *Earth Planet Sci Lett*. 2012;302(3):439–47.
 110. Wen J, Nie ZL, Ickert-Bond Stefanie M. Intercontinental disjunctions between eastern Asia and western North America in vascular plants highlight the biogeographic importance of the Bering land bridge from late cretaceous to Neogene. *J Syst Evol*. 2016;54(5):469–90.
 111. Denk T, Grímsson F, Zetter R. Episodic migration of oaks to Iceland: evidence for a North Atlantic "land bridge" in the latest Miocene. *Am J Bot*. 2010;97(2):276–87.
 112. Huang C, Zhang Y, Bartholomew B: Fagaceae. In: *Flora of China 4 Cycadaceae Through Fagaceae*. Edited by Wu Z, Raven P. Beijing&St.Louis: Science Press and Missouri Botanical Garden Press; 1999. 314–400.
 113. Xu XT, Dimitrov D, Shrestha N, Rahbek C, Wang Z. A consistent species richness-climate relationship for oaks across the northern hemisphere. *Glob Ecol Biogeogr*. 2019;1051–66.
 114. McVay JD, Hipp AL, Manos PS. A genetic legacy of introgression confounds phylogeny and biogeography in oaks. *Proc R Soc London, Ser B*. 2017;284(1854):20170300.
 115. Kremer A, Petit RJ. Gene diversity in natural populations of oak species. *Ann Sci For*. 1993;50(Suppl1):186s–202s.
 116. Kremer A, Petit R, Zanetto A, Fougère V, Ducousso A, Wagner D, Chauvin C. Nuclear and organelle gene diversity in *Quercus robur* and *Q. petraea*. In: *Genetic variation in European populations of forest trees*. Edited by MullerStarck G, Ziehe M. Frankfurt: Sauerlander's Verlag; 1991. 141–166.

117. Toumi L, Lumaret R. Allozyme variation in cork oak (*Quercus suber* L.): the role of phylogeography and genetic introgression by other Mediterranean oak species and human activities. *Theor Appl Genet.* 1998;97(4):647–56.
118. Neophytou C, Dounavi A, Fink S, Aravanopoulos FA. Interfertile oaks in an island environment: I. high nuclear genetic differentiation and high degree of chloroplast DNA sharing between *Q. alnifolia* and *Q. coccifera* in Cyprus. A multipopulation study. *Eur J For Res.* 2011;130(4):543–55.
119. Yan MX, Xiong YS, Liu RB, Deng M, Song JJ. The application and limitation of universal chloroplast markers in discriminating east Asian evergreen oaks. *Front Plant Sci.* 2018;9:569.
120. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32(5):1792–7.
121. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 2016;33(7):1870–4.
122. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics.* 2009;25(11):1451–2.
123. Nei M, Li WH. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci.* 1979;76(10):5269–73.
124. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research—an update. *Mol Ecol Resour.* 2012;28:2537–9.
125. Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 2012;61(3):539–42.
126. Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. *Bioinformatics.* 1998;14(9):817–8.
127. Revell LJ. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol.* 2012;3(2):217–23.
128. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol.* 2005; 25(15):1965–78.
129. Hijmans RJ. Raster: Geographic Data Analysis and Modeling. R package version 2.6–7; 2017.
130. Menitsky LL. Oaks of Asia. St.Petersburg: Leningosed Sciences; 1984.
131. Yang SQ. The evolution and differentiation of distribution of *Quercus* section *Heterobalanus* (Fagaceae). Kuming: Kunming Institute of Botany, Chinese Academy of Sciences; 2008.
132. Komarov VL. Flora of the USSR, vol. 5. Moskva Leningrad: Izdatel'stvo Akademii Nauk SSSR; 1936.
133. Zeng YF, Liao WJ, Petit RJ, Zhang DY. Exploring species limits in two closely related Chinese oaks. *PLoS One.* 2010;5(11):e15529.
134. Zeng YF, Wang WT, Liao WJ, Wang HF, Zhang DY. Multiple glacial refugia for cool-temperate deciduous trees in northern East Asia: the Mongolian oak as a case study. *Mol Ecol.* 2015;24(22):5676–91.
135. Revelle W. Psych: procedures for personality and psychological research (version = 1.8.3.). Evanston: Northwestern University; 2018.
136. Phillips SJ, Dudík M, Schapire RE: Maxent software for modeling species niches and distributions (Version 3.4.1). 2018.
137. Wickham H. Ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.
138. Legendre P, Legendre L. Numerical ecology. 2nd. Amsterdam, NL: Elsevier Science; 1998.
139. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlenn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P et al: Vegan: Community Ecology Package. R package version 2.4–6. 2018.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

