ORIGINAL ARTICLE

Phylogeography of *Quercus aquifolioides* provides novel insights into the Neogene history of a major global hotspot of plant diversity in south-west China

Fang K. Du[1],*, Meng Hou[1,2], Wenting Wang[3], Kangshan Mao[4] and Arndt Hampe[5,6]

[1]College of Forestry, Key Laboratory for Forest Resources and Ecosystem Processes of Beijing, Beijing Forestry University, Beijing 100083, China, [2]School of Life Science, Lanzhou University, Lanzhou 730000, China, [3]School of Mathematics and Computer Science, Northwest University for Nationalities, Lanzhou 730030, China, [4]Key Laboratory for Bio-resources and Eco-environment, College of Life Science, Sichuan University, Sichuan 610065, China, [5]INRA, UMR1202 BIOGECO, F-33610 Cestas, France, [6]Univ. Bordeaux, BIOGECO, UMR 1202, F-33615 Pessac, France

ABSTRACT

Aim Hotspots of biodiversity are often associated with areas that have undergone orogenic activity during recent geological history. Mountain uplifts are known to catalyse species radiation but their impact on evolutionarily stable taxa such as many trees remains little understood. The oak *Quercus aquifolioides* is endemic to yet widely distributed across the Hengduanshan Biodiversity Hotspot in the Eastern Himalayas. Here, we investigate how the region's Neogene and Quaternary history has driven the species' past population dynamics and the resulting extant patterns of intraspecific diversity.

Location Hengduanshan Biodiversity Hotspot in SW China.

Methods We sampled 58 populations throughout the species range and genotyped a total of 959 individuals at four chloroplast DNA fragments and 11 nuclear microsatellite loci. Phylogenetic reconstructions, molecular dating techniques and ancestral area reconstructions were used in combination with population genetic statistics to infer the biogeographical history of *Q. aquifolioides*. The phylogeographical study was complemented by a survey of fossil records and a niche modelling exercise.

Results Combined molecular and fossil evidence indicates that *Q. aquifolioides* descended during the late Miocene from the central Qinghai-Tibet Plateau into Tibet and the western Sichuan Plateau, and from there, into the area of highest endemism in the Hengduan Mountains *sensu lato*. Great apparent population stability and a haplotype 'radiation' in this area contrasted with marked extinction–recolonization dynamics and reduced population diversity in Tibet. We found evidence for extremely limited seed gene flow but extensive pollen gene flow (global F_{ST}: cpDNA = 0.98, nSSR = 0.07) with signals of asymmetric pollen dispersal from the Hengduan Mountains into Tibet.

Main conclusion Our results provide insights of unprecedented detail into the ancient biogeographical history of the Hengduanshan Biodiversity Hotspot, suggesting that past environmental changes in the region may have catalysed radiative diversifications within species much in the same way as among species.

*Correspondence: Dr Fang K. Du, College of Forestry, Beijing Forestry University, Beijing 100083, China.
E-mail: dufang325@bjfu.edu.cn

INTRODUCTION

The evolutionary history of the world's major hotspots of biodiversity remains incompletely understood despite intense research (Qian & Ricklefs, 2000; Sechrest *et al.*, 2002; Hoorn *et al.*, 2010). Many of the world's top hotspots are located in tectonically active areas that have experienced significant geological and climatic changes since the late Neogene. Prominent examples are the Californian Floristic Province, the tropical Andes, New Zealand or the eastern Himalayas

(Myers *et al.*, 2000). These hotspots contain a large fraction of evolutionarily young taxa that result from rapid radiations driven by environmental changes (Linder, 2008; Hughes & Atchison, 2015). Radiative speciation processes have received much attention as putative keys to a better understanding of the emergence of diversity hotspots (e.g. Hoorn *et al.*, 2010; Favre *et al.*, 2015). On the contrary, we know very little about how past geological and climatic changes in these hotspot areas have affected the population dynamics of older and evolutionarily more stable taxa (Petit & Hampe, 2006). Yet, imprints of this history persisting in the extant genetic structure of long-lived taxa could help elucidate past environmental changes and biotic responses that cannot be tracked through the fossil or geological records alone (Donoghue, 2008; Pennington *et al.*, 2010).

The eastern Himalayas and adjacent mountain ranges in south-west China harbour one of the world's major hotspots of plant diversity with around 12,000 species including 3,500 endemics (Myers *et al.*, 2000). A unique combination of geological and climatic circumstances has allowed this diversity to evolve and persist through an extended period of time. The region stands out for its topographic diversity with some of the world's most rugged mountain ranges and altitudinal gradients spanning over 5,000 m (from < 10,000 to > 6,000 m a.s.l.) (Favre *et al.*, 2015). This topography is the result of an exceptionally intense orogeny that has taken place since the late Neogene (or somewhat earlier: Lippert *et al.*, 2014; Renner, 2016). The eastern Himalayas and adjacent mountain ranges form an extension of the youngest part of the Qinghai-Tibet Plateau (QTP). Their main uplift has been suggested to have occurred in 10 Ma (Mulch & Chamberlain, 2006), reaching peak elevation already shortly before the Late Pliocene (Sun *et al.*, 2011). The orogeny went along with a distinction between a south-western area that maintained subtropical evergreen broad-leaved forest vegetation and a north-eastern area that experienced a decline in precipitation and the development of a vertical vegetation zonation (Sun *et al.*, 2011). The varied topography and the macrogeographical situation of the region mitigated the impact of subsequent global-scale climate changes, converting the area in a major climatic refugium for temperate plants (Qiu *et al.*, 2011). Several fossil records indicate indeed that the vegetation of the region has changed relatively little since the late Pliocene (Sun *et al.*, 2011). The intense orogeny and the resulting landscape heterogeneity render the so-called Hengduanshan Biodiversity Hotspot a highly suited natural laboratory for studying both the evolution (as 'cradle') and the maintenance (as 'museum') of plant diversity (López-Pujol *et al.*, 2011).

The area harbours an outstanding number of (mostly herbaceous) endemic plant taxa that have evolved since the Pliocene through numerous radiations and vicariance events (Wen *et al.*, 2014; Hughes & Atchison, 2015). On the contrary, palaeoendemic taxa are less abundant than in other parts of southern China (López-Pujol *et al.*, 2011). The relative poverty in palaeoendemics suggests that the geological

and climatic changes during the late Miocene and early Pliocene could have caused the disappearance of numerous taxa previously present in the area. Unfortunately, the fossil record provides no clear evidence for such a process. Numerous plant phylogeographical studies have targeted the QTP and its surroundings in recent years (reviewed in Qiu *et al.*, 2011; Wen *et al.*, 2014; Favre *et al.*, 2015; Hughes & Atchison, 2015). However, most have focused either on processes of speciation and evolutionary diversification or on Quaternary population dynamics and the question whether taxa persisted *in situ* at high altitude on the QTP through the Last Glacial Maximum (LGM) or recolonized the area from peripheral refugia. Some studies of woody species have looked further back in time to elucidate pre-Quaternary population dynamics in the region. These studies found clear imprints of the region's Neogene history on patterns of within-species diversity, which they related with the formation of major genetic lineages as a consequence of the QTP uplift (e.g. Wang *et al.*, 2010; Xu *et al.*, 2010; Li *et al.*, 2013; Liu *et al.*, 2013; Sun *et al.*, 2014). However, more detailed insights into the complex Neogene biogeography of the Hengduanshan Biodiversity Hotspot have to date been precluded by constraints such as incomplete geographical coverage of sampling, low genetic polymorphism of samples or the lack of fossil records to complement molecular data.

Here, we report on the Neogene and Quaternary population dynamics of *Quercus aquifolioides* Rehder & E.H. Wilson, an evergreen oak species that is endemic to but widely distributed across the Hengduanshan Biodiversity Hotspot. We combine evidence from Neogene and Quaternary fossil records, molecular analyses based on a dense range-wide sampling of populations and niche modelling to infer how geological and climatic changes in the region have driven past distribution shifts, population dynamics, gene flow and resulting patterns of genetic variation across the present-day range of the species. Our specific objectives were to: (1) examine range-wide patterns of intraspecific genetic diversity of *Q. aquifolioides* at chloroplast DNA (cpDNA) and nuclear markers; (2) determine divergence times of the major intraspecific lineages; (3) infer past population and range dynamics that have led to the species' current distribution and understand their underlying environmental causes; and (4) confront our reconstruction with the known late Neogene and Quaternary geological history of the Hengduanshan Biodiversity Hotspot. Our study provides insights of unprecedented detail into the complex history of the area, shedding light on some of the biogeographical processes that have allowed this hotspot to acquire and maintain its exceptional plant diversity.

MATERIALS AND METHODS

Study organism

Quercus aquifolioides is the most widely distributed member of the evergreen sclerophyllous *Quercus* section *Heterobalanus*

(Oerst.) Menits (Zhou, 1992). Its range spans western Sichuan, northern Yunnan and eastern Tibet, extending into the adjacent areas of Guizhou, Myanmar and Bhutan (Fig. 1a). This area largely coincides with the floristic region that has been defined as the Sino-Himalayan Forest subkingdom (Qiu *et al.*, 2011), rendering the species an emblematic representative of this area. *Quercus aquifolioides* occurs across a broad elevational range (1,900–4,600 m a.s.l.; Huang *et al.*, 1999) and environments. Its geographical distribution and ecological amplitude render *Q. aquifolioides* a highly suited model to investigate the effects of past geological and climatic changes in the region on the population dynamics of long-lived, woody species. Like other oaks, *Q. aquifolioides* is wind-pollinated and its acorns are putatively dispersed by animals.

Sampling, sequencing and microsatellite genotyping

We sampled leaf material from 959 individuals in 58 wild populations (separated by at least 30 km from each other) from throughout the species' distribution range (see Fig. 1 and Appendix S1 in Supporting Information). Leaves were rapidly dried in silica gel and stored until DNA isolation. Voucher specimens of each population were stored in the herbarium of Beijing Forestry University. Total genomic DNA was isolated using a Plant Genomic DNA Extraction Kit (Tiangen, Beijing, China). Five hundred and ninety individuals were sequenced at four cpDNA loci [*trn*H-*psb*A, *rp*S16, *trn*S (GCU) – *trn*T (GGU) and *trn*Q (UUG) – *trn*S (GGU); Kanno *et al.*, 2004; Shaw *et al.*, 2005] after initial tests with 16 loci. All 959 samples were genotyped at 11 nuclear microsatellite (nSSR) loci (MSQ13, ssrQpZAG9, ssrQpZAG16, ssrQpZAG110, ssrQrZAG7, ssrQrZAG11, ssrQrZAG30, ssrQrZAG87, ssrQrZAG96, ssrQrZAG112 and PIE_271; Dow *et al.*, 1995; Kampfer *et al.*, 1998; Durand *et al.*, 2010) after initial tests with 34 loci. Primer selection, sequencing and genotyping procedures are described in Appendices S2 and S3.

Data analysis

Phylogenetic analyses

We aligned cpDNA sequences with ClustalW in mega 5.1 (Tamura *et al.*, 2011) and checked them manually. Chloroplast haplotypes were identified with DnaSP 5 (Librado & Rozas, 2009) and a haplotype network was constructed using network 4.6.11 with the median-joining model (Bandelt *et al.*, 1999). We derived phylogenetic relationships among cpDNA haplotypes using beast 1.7.4 (Drummond & Rambaut, 2007). *Trigonobalanus doichangensis* (A. Camus) Farman was used as outgroup for rooting the phylogenetic tree after testing without success a more closely related species (*Quercus ilex* L.). We refrained from a fossil calibration because some of the reported oak species are quite similar and the determination of their fossils is a matter of debate

(Zhou *et al.*, 2007). Instead, we used the cpDNA substitution rate of 0.96×10^{-9} s/s/y that was recently reported for *Quercus* by Xu *et al.* (2015) based on four Asian *Quercus* fossils. For details on the analysis settings, see Appendix S2.

We performed a statistical-dispersal vicariance analysis (S-DIVA) as implemented in rasp 3.2 (Yu *et al.*, 2015) after comparing this approach with two other commonly used approaches for ancestral range construction [Bayesian Binary MCMC (BBM) and Statistical Dispersal-Extinction-Cladogenesis (S-DEC)] with the software R package 'Bio-GeoBears' (Matzke, 2012) and confirming the superiority of S-DIVA (weighted AIC: S-DIVA = 0.20, BBM = 0.098, DEC = 0.001; see Appendix S2). Adopting the delimitation of the Hengduan Mountains by Li (1987), we assigned our populations to one of three areas separated by the Boshualing Mountains (A/C) and the Yalong River (B/C) respectively (see Fig. 2a).

Genetic diversity and differentiation

We estimated haplotype (H_d) and nucleotide (π) diversity at cpDNA markers for each population using DnaSP 5. Average gene diversity within populations (H_S), total gene diversity (H_T) and population differentiation indices (G_{ST} and N_{ST}) were calculated using the program permut and the differences between G_{ST} and N_{ST} were assessed using a permutation test with 1,000 random permutations (Pons & Petit, 1996).

Detailed information on the initial screening of the microsatellite loci (including tests for identical genotypes, null alleles, departure from Hardy–Weinberg equilibrium and traces of selection) is shown in Appendix S3. Three nSSR loci departed significantly from the simulated F_{ST} distribution based on neutral loci, suggesting that they could be under disruptive selection or linked to a locus under selection. These loci were therefore excluded from all subsequent analyses. Descriptive statistics for nSSRs including the number of alleles, observed and expected heterozygosity were estimated using GenAlEx 6.5 (Peakall & Smouse, 2012). Allelic richness and private allele richness for a sample size of six were estimated with ADZE (Szpiech *et al.*, 2008).

We performed a hierarchical analysis of genetic differentiation for both cpDNA and SSR markers using an analysis of molecular variance (AMOVA) with 1,000 permutations as implemented in Arlequin 3.5 (Excoffier & Lischer, 2010). We conducted a global analysis with all samples pooled as well as three specific analyses considering separately each of the three major lineages identified.

The software structure 2.3.4 (Pritchard *et al.*, 2000) was used to examine the range-wide genetic structure at SSR markers. The program was run for $K = 1$–15 under the admixture model with independent allele frequencies. For each run, we set 100,000 MCMC cycles followed by 100,000 burn-in cycles. Twenty replications were performed for each K, and the optimal K was estimated according to Evanno *et al.* (2005).
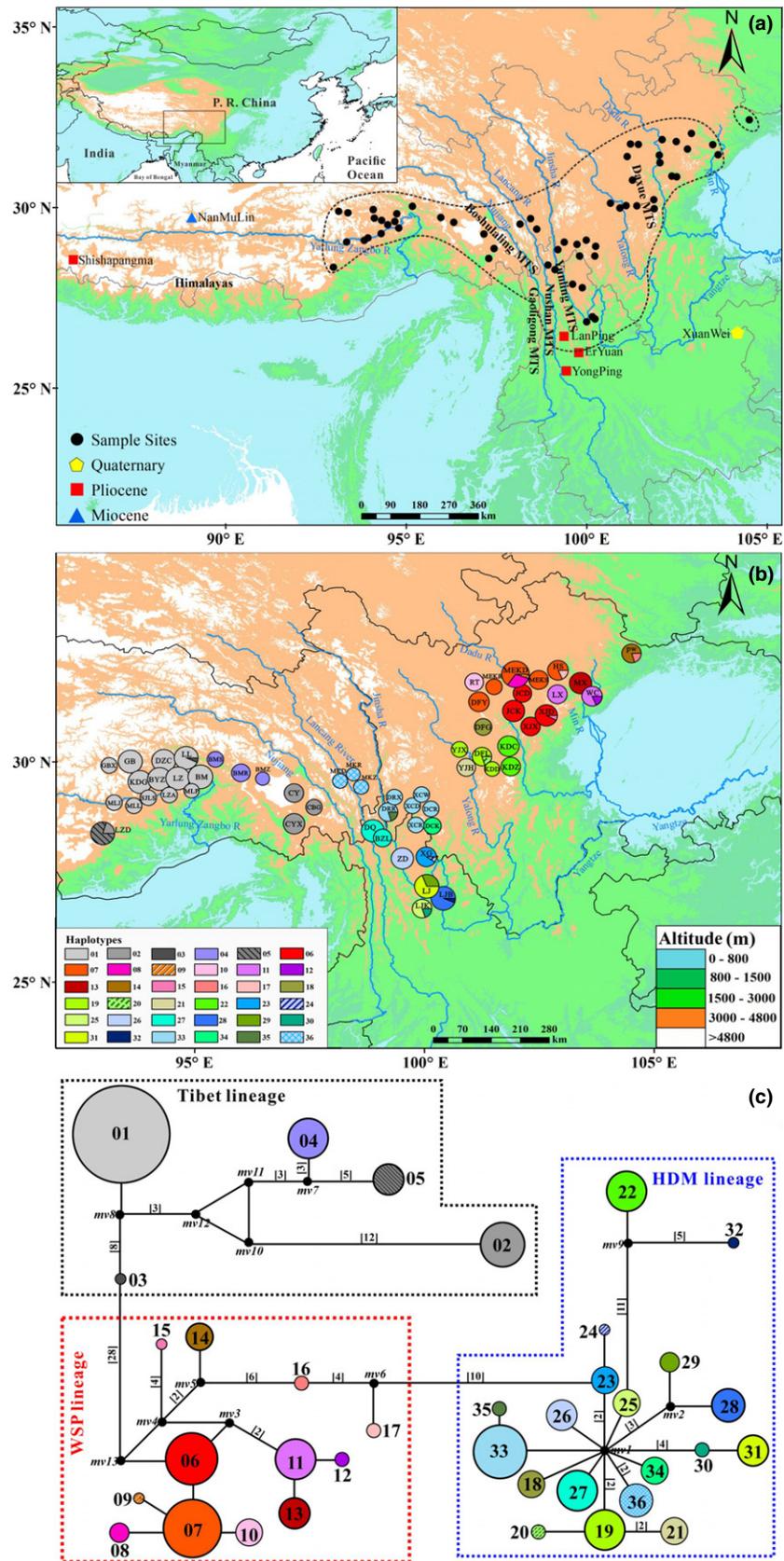
Figure 1 Geographical distribution, fossil records and cpDNA structure of *Quercus aquifolioides*. (a) Species range (dashed line), sampling locations and fossil records of *Q. aquifolioides* and its likely ancestor *Q. preguyavifolia* within the greater geographical context. The Gaoligong, Nushan and Yunling Mountains form the Hengduan Mountain massif that, together with the Daxue Mountains, harbours the core of the Hengduanshan Biodiversity Hotspot. (b) Geographical distribution of chloroplast haplotypes in the 58 populations. Circle sizes reflect the number of individuals genotyped (*n* = 5–20). (c) Inferred phylogenetic network of the 36 cpDNA haplotypes found. Each haplotype is represented by a circle whose size is proportional to its frequency across all populations. Numbers in brackets on branches indicate the number of mutations between haplotypes. Small solid black circles (*mv*1-13) represent hypothetic unsampled or extinct ancestral haplotypes [Colour figure can be viewed at wileyonlinelibrary.com].
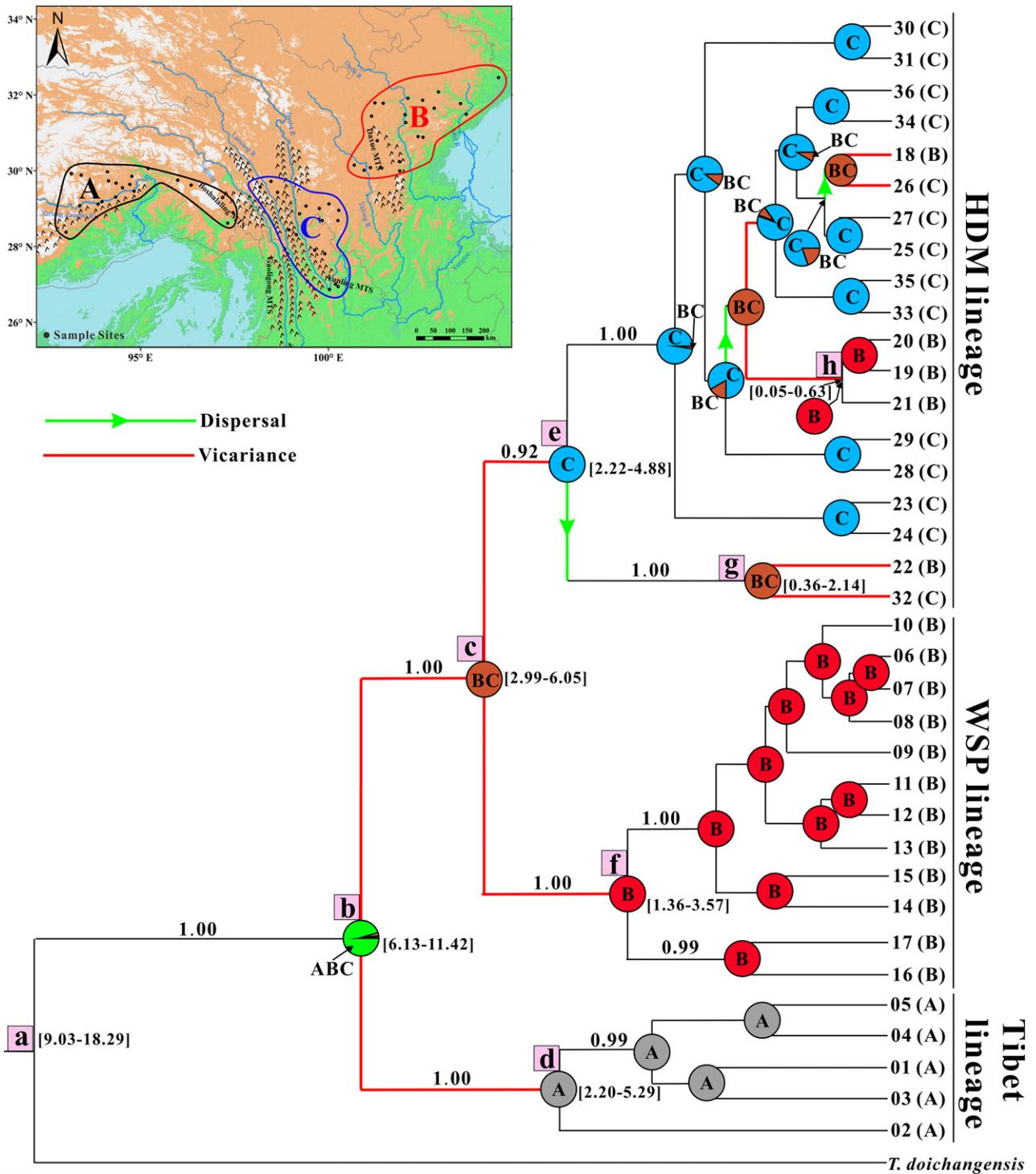
Figure 2 Phylogenetic relationships between chloroplast haplotypes of *Quercus aquifolioides* and their underlying biological processes according to the S-DIVA analysis. Numbers above branches denote posterior probabilities and those shown in brackets indicate the 95% HPD of divergence times (in Ma) of the eight main nodes (a–h). The left top map shows the three range sectors A, B and C used in the S-DIVA analysis. Pie charts indicate proportions of the ancestral ranges. The solid red lines and green lines with arrow indicate events of vicariance and long-distance dispersal respectively [Colour figure can be viewed at wileyonlinelibrary.com].

Gene flow among regions

Historical gene flow among different sectors of the distribution range (Fig. 2) was assessed using the software Migrate-

N (Beerli, 2006). The amount and direction of gene flow was estimated from the nSSR data by calculating the parameters θ (four times effective population size multiplied by mutation rate per site per generation) and M (immigration rate

divided by the mutation rate). We used the continuous Brownian motion model. First, we implemented a pre-run with F_{ST} values to obtain the prior setting of θ and *M*. Then we started five independent MCMC chains with 5,000,000 generations, respectively. We sampled every 100 steps under a constant mutation model, discarding the first 1,000,000 records as burn-in. The mode and 95% highest posterior density were then estimated after checking for data convergence.

Ecological niche modelling

We used MAXENT (Phillips *et al.*, 2006) to hindcast potential past distributions of *Q. aquifolioides* by relating its modern distribution records and bioclimatic variables. We used data on precise geographical locations obtained from our own sampling and from records of *Q. aquifolioides* in the Chinese Virtual Herbarium (http://www.cvh.org.cn/) to characterize the species' geographical distribution. A total of 80 records spread across the distribution range were used for the analysis. Bioclimatic variables for current conditions at these sites were downloaded from WorldClim (Hijmans *et al.*, 2005). Last Glacial Maximum (*c.* 21 ka) data were obtained from two general circulation models (GCM) simulations provided by the Palaeoclimate Modelling Intercomparison Project (http://pmip2.lsce.ipsl.fr/): the Community Climate System Model (CCSM) (Collins *et al.*, 2006) and the Model for Interdisciplinary Research on Climate (MIROC) (Hasumi & Emori, 2004). Data for the Last Interglacial (LIG, *c.* 130 ka) were drawn from Otto-Bliesner *et al.* (2006). The Pliocene (*c.* 3 Ma) data were obtained from their original author Dr Dan Lunt (Bristol University). We used resolutions of 2.5-arc minutes for the present and LGM climatic scenario(s), 30-arc seconds for the LIG scenario and 5-arc minute for the Pliocene scenario, respectively. Four ecologically relevant bioclimatic variables were retained after eliminating highly correlated variables (i.e. $r \geq 0.75$) (Hijmans *et al.*, 2005): maximum temperature of the warmest month, mean temperature of the coldest quarter, precipitation of the wettest quarter, and precipitation of the warmest quarter. Model validation was performed using MAXENT default settings with 20 independent replicates of cross-validation procedures.

RESULTS

Phylogenetic relationships and geographical distribution of cpDNA haplotypes

We sequenced a total of 2,709 bp from the four cpDNA fragments and constructed a matrix of combined sequences for the 590 individuals. The chloroplast fragment sequences have been deposited in GenBank (KF171124–KF171181 and KT625425–KT625438). A total of 36 haplotypes with 98 substitutions and 13 indels were detected (see Appendix S2). The phylogenetic network revealed three distinct lineages with adjacent but non-overlapping geographical ranges

(Fig. 1b,c). Haplotypes 1–5 composed a lineage of exclusively Tibetan distribution. Haplotypes 6–17 formed a second lineage (hereafter termed WSP) that was distributed through the West Sichuan Plateau. Haplotypes 18–36 composed a third lineage (hereafter HDM) distributed through the Hengduan Mountain massif (an orogenic complex that includes the Gaoligong, Nushan and Yunling Mountains; see Fig. 1a) and the Daxue Mountains. This area corresponds to the centre of the Hengduanshan Biodiversity Hotspot.

Surprisingly, the Tibetan lineage was more closely related with the geographically more distant WSP lineage than with the interjacent HDM lineage (Fig. 1c). The relationship was still distant, however, as we observed no less than 38 mutation steps from the nearest WSP lineage relative to the most common Tibetan haplotype. The WSP lineage was more closely related with the HDM lineage, as only 13 mutation steps separated the nearest WSP relative from the core of the HDM network.

The three lineages varied greatly in their internal structure; while the Tibetan lineage included a few remarkably distinct haplotypes, the HDM lineage was characterized by a notably rich and complete star-like structure with most haplotypes differing only by few mutation steps from each other (Fig. 1c). The structure of the WSP lineage was intermediate between the other two, with many haplotypes being closely related while some others differed by several mutation steps from their nearest neighbour in the network. Private haplotypes were common in the populations of HDM (10 of 24) and WSP (7 of 14) but markedly scarcer in Tibet (2 of 20). Non-private haplotypes tended to be more geographically restricted in the HDM area than in WSP, where haplotypes 6 and 7 were relatively widespread. At the other extreme, Tibetan populations were largely dominated by a single haplotype (1) except for the easternmost part, where populations were fixed for two further haplotypes (2 and 3).

Divergence times and historical biogeographical inference

The phylogenetic tree constructed by BEAST clustered the 36 haplotypes into three strongly supported clades (Fig. 2) that were almost identical with the three lineages of the haplotype network (see Fig. 1b). The divergence time of the outgroup species *T. doichangensis* was estimated at 13.2 (95% HPD: 9.0–18.3) Ma (node a in Fig. 2). The divergence time of the early diverging Tibetan lineage (node b) was estimated at 8.6 (6.1–11.4) Ma and that of the WSP and HDM lineage (node c) at 4.4 (3.0–6.1) Ma. All major divergence events within each of the three major lineages (nodes d–h) were assigned to the early and mid-Quaternary.

The S-DIVA analysis classified the divergence of the three major lineages and a few minor splits as vicariance events. It also identified three dispersal events of HDM lineage populations from range sector C into B (Fig. 2), reflecting a consistent secondary northward expansion of the HDM lineage after its initial establishment.

Diversity, differentiation and gene flow at cpDNA and nSSR markers

Overall cpDNA diversity was high (mean $H_d = 0.89$, mean $\pi = 8.15 \times 10^{-3}$; see Appendix S1) although it decreased from HDM through WSP to Tibet. Genetic differentiation was remarkably high ($G_{ST} = 0.93$ and $N_{ST} = 0.98$), and a significant range-wide phylogeographical structure was observed (Table 1). The AMOVA on cpDNA data revealed that 74.1% of the overall variation was distributed among lineages, 23.9% among populations within lineages and only 2.0% within populations (Table 2).

All eight nSSR loci were highly variable with 10–28 alleles per locus across the 58 populations. Diversity estimates decreased consistently from HDM through WSP to Tibet (Table 1; repeated-measures ANOVA: $F \geq 3.577$; d.f. = 2, 55; $P \leq 0.05$). Contrary to the cpDNA data, by far most of the overall variation at nSSR loci (92.6%; $F_{ST} = 0.07$ in Table 2) occurred within populations, with the rest being roughly equally distributed among populations within lineages and among lineages (Table 2).

The STRUCTURE analysis produced a clear maximum value of ΔK for $K = 2$, suggesting the existence of two major clusters in the data set (Fig. 3b). All populations of the WSP and HDM lineages were primarily assigned to cluster I (green), whereas the populations in the Tibetan lineage appeared mostly in cluster II (red). However, numerous individuals of the Tibetan populations BM, BMZ, CY, CYX and CBG – that is, those located nearest to the westernmost populations of the HDM lineage – also had notable fractions assigned to cluster I (Fig. 3a).

The Migrate-n analysis produced θ and M values greater than zero (Table 3). θ values did not vary among range sectors, whereas scaled immigration rates (M) revealed the

Table 1 Genetic diversity estimates for cpDNA and nuclear microsatellite markers (nSSR) in the investigated *Quercus aquifolioides* populations.

	cpDNA					nSSR				
	N	H_T (se)	H_S (se)	G_{ST} (se)	N_{ST} (se)	N	H_O	H_E	A_R	A_P
Tibet lineage	212	0.57 (0.12)	0.02 (0.02)	0.96 (0.03)	0.95 (0.04)	380	0.71[b]	0.65[b]	3.23[b]	0.02[b]
WSP lineage	157	0.84 (0.05)	0.12 (0.05)	0.86 (0.06)	0.84 (0.06)	158	0.74[ab]	0.74[a]	3.87[a]	0.05[ab]
HDM lineage	221	0.92 (0.03)	0.08 (0.03)	0.92 (0.03)	0.92 (0.04)	421	0.75[a]	0.75[a]	3.87[a]	0.06[a]
Total	590	0.93 (0.02)	0.07 (0.02)	0.93 (0.02)	0.98 (0.01)*	959	0.74	0.71	3.65	–

N, number of individuals; H_T, total genetic diversity; H_S, genetic diversity within populations; G_{ST}, interpopulation differentiation; N_{ST}, interpopulation differentiation taking similarities between haplotypes into account; *, N_{ST} differs from G_{ST} at $P \leq 0.05$; H_O and H_E, observed and expected heterozygosity, respectively; A_R and A_P, allelic richness and private allelic richness, respectively; Superscript letters indicate significant differences according to Tukey's HSD tests ($P < 0.05$).

Table 2 Hierarchical analyses of molecular variance (AMOVA) of *Quercus aquifolioides* populations based on cpDNA polymorphisms and nuclear microsatellite allele frequencies. A global analysis is followed by three lineage-specific analyses (see Fig. 1 for lineages and their geographical distributions).

	cpDNA			nSSR		
Source of variation	d.f.	Percentage of variation (%)	Fixation indices	d.f.	Percentage of variation (%)	Fixation indices
Total populations						
Among lineages	2	74.1	$F_{CT} = 0.74$	2	3.3	$F_{CT} = 0.03$
Among populations within lineages	55	23.9	$F_{SC} = 0.92$	55	4.2	$F_{SC} = 0.04$
Within populations	532	2.0	$F_{ST} = 0.98$	1860	92.6	$F_{ST} = 0.07$
Tibet lineage						
Among populations	19	95.0	$F_{ST} = 0.95$	19	5.0	$F_{ST} = 0.05$
Within populations	192	5.0		740	95.0	
WSP lineage						
Among populations	13	88.5	$F_{ST} = 0.89$	13	3.5	$F_{ST} = 0.04$
Within populations	143	11.5		302	96.5	
HDM lineage						
Among populations	23	92.1	$F_{ST} = 0.92$	23	3.9	$F_{ST} = 0.04$
Within populations	197	7.9		818	96.1	

d.f., degree of freedom.
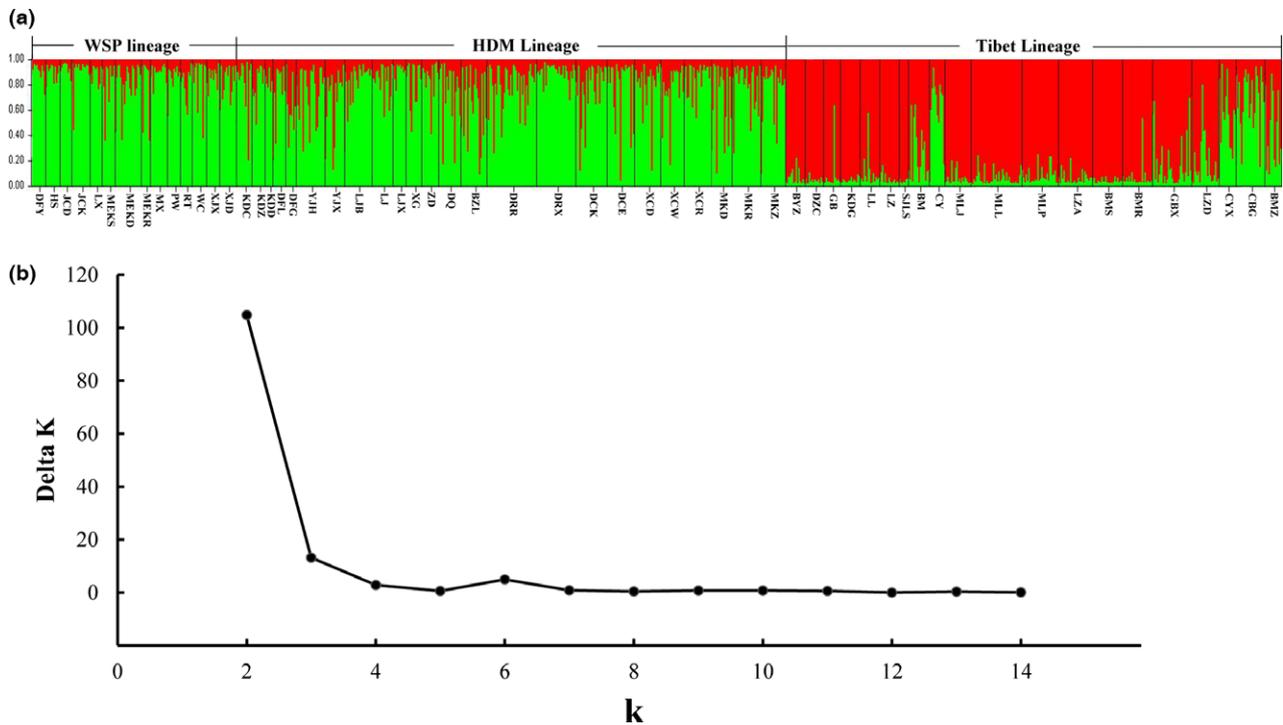
Figure 3 STRUCTURE clustering results for 959 *Quercus aquifolioides* individuals from 58 populations based on variation at eight nuclear microsatellite loci. (a) Histogram of individual assignments. Black lines separate different populations. The major lineages are indicated on top and population codes below the histogram. See Fig. 1(b) for their respective locations. (b) The magnitude of ΔK as a function of K suggests the existence of two major clusters as the most likely scenario [Colour figure can be viewed at wileyonlinelibrary.com].

Table 3 Historical gene flow as estimated by Migrate-n among the three range sectors of *Quercus aquifolioides* indicated in Fig. 2.

| | | M (m/μ) | | |
	θ	A→	B→	C→
A	**1.7** [0.9–3.3]		**11.2** [7.0–20.7]	**20.2** [15.7–30.3]
B	**1.3** [0.6–2.9]	**48.8** [44.0–61.0]		**55.2** [50.0–69.7]
C	**1.8** [1.0–3.7]	**46.8** [42.7–57.7]	**24.2** [20.0–34.0]	

The mode of the posterior distribution is shown in bold and the values in square brackets give the 95% credibility interval; θ, $4Ne\mu$; →, source populations; M, mutation-scaled immigration rate; m, immigration rate; μ, mutation rate.

existence of highly asymmetric historical gene flow between all three sectors. Gene movements occurred predominantly from sector A into sectors B (48.8 vs. 11.2) and C (46.8 vs. 20.2) as well as from sector C into sector B (55.2 vs. 24.2).

Niche modelling

The projected extant distribution of *Q. aquifolioides* was similar to the species' actual distribution (Fig. 4a), resulting in a high AUC (0.95 ± 0.03, mean ± SD). The potential range of *Q. aquifolioides* shrinked and shifted somewhat southwards during the LGM (Fig. 4b,c). It was, however, most restricted during the LIG, suggesting that this warm period would have provided the least favourable climatic conditions for the species (Fig. 4d). The modelling exercise also showed a large area with high climatic suitability for *Q. aquifolioides*

during the Pliocene (Fig. 4e) that was largely congruent with the current distribution (albeit more prominent, especially in Tibet).

DISCUSSION

The eastern Himalayas are one of the tectonically most active regions of the world. The rapid uplift of the Hengduan Mountains has converted the area into the 'largest "evolutionary front" of the world's temperate zone' (López-Pujol *et al.*, 2011). Our study underpins that *Q. aquifolioides* was present in the region during much of the upfolding process. The species is hence an ideal witness of the physiographic, climatic and biological processes that have generated one of the principal biodiversity hotspots outside the tropics.
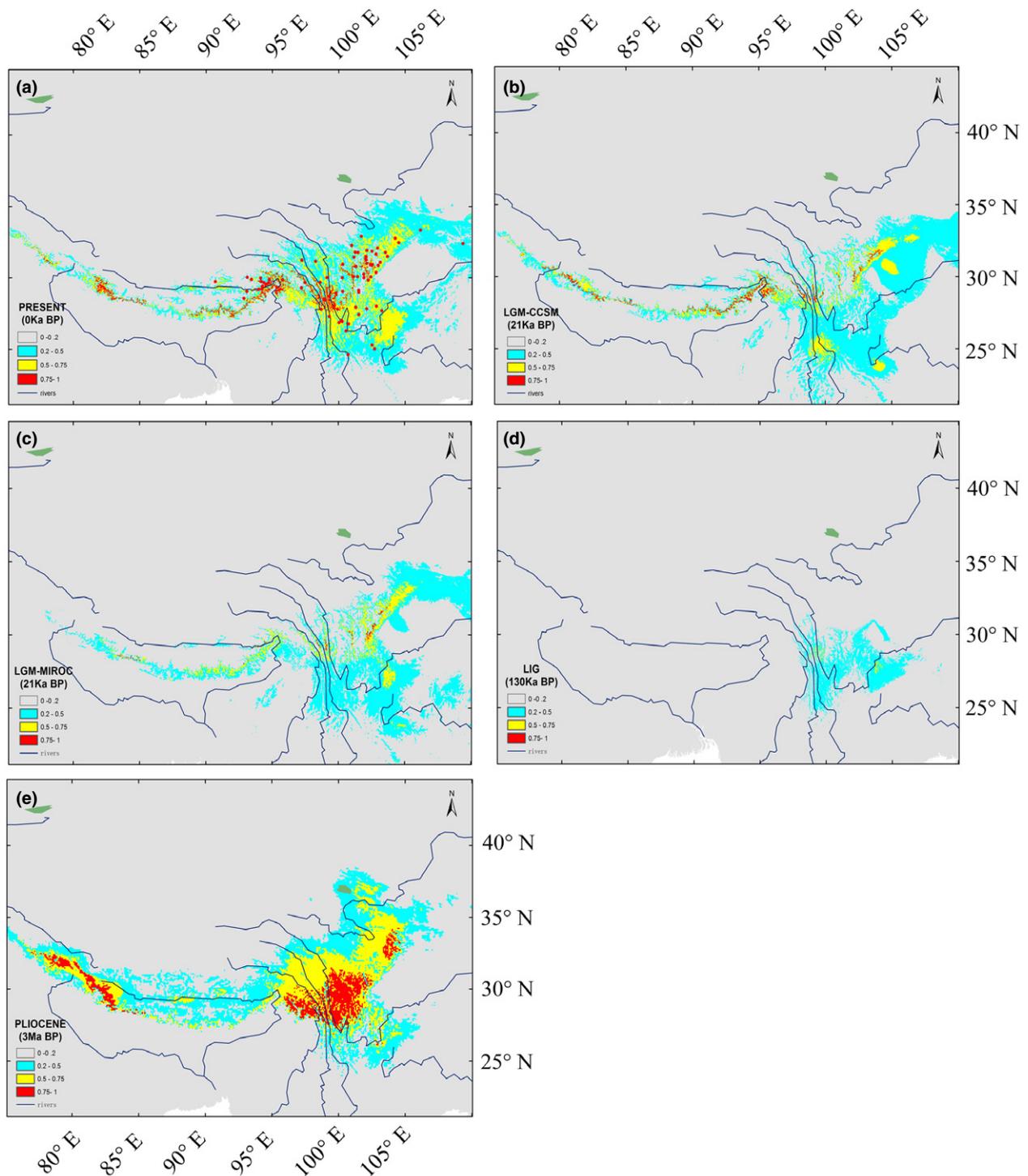
Figure 4 Modelled potential distribution of *Quercus aquifolioides* at present (a); during the Last Glacial Maximum (21 ka) under the CCSM (b) and MICRO (c) general circulation model, respectively; during the Last Interglacial (130 ka) (d); and during the Pliocene (*c.* 3 Ma) (e). Red dots in panel (a) indicate the occurrence data of *Q. aquifolioides* on which the model is based [Colour figure can be viewed at wileyonlinelibrary.com].

Miocene to Pliocene divergence of the three major lineages

We detected three distinct genetic lineages with adjacent but non-overlapping distributions. The Tibetan populations

formed the first diverging lineage with an estimated divergence time of roughly 8.6 (95% HPD: 6.1–11.4): Ma. The timing of this earliest lineage split in *Q. aquifolioides* occurred relatively early compared to other woody species from the study area [*Cupressus chengiana*: 9 Ma (Xu *et al.*,

2010), *Tetracentron sinense*: 9 Ma (Sun *et al.*, 2014), *Picea likiangensis*: 5.9 Ma (Li *et al.*, 2013), *Taxus wallichiana*: 4.2 Ma (Liu *et al.*, 2013), *Hippophaë tibetana*: 3.2 Ma (Wang *et al.*, 2010)]. Importantly, it dates back to an initial stage in the uplift of the mountain ranges that harbour today the Hengduanshan Hotspot. Although the lineage split might have happened at low elevation, it appears more likely that it would have occurred while *Q. aquifolioides* was colonizing its present-day distribution range from higher parts of the QTP situated further to the west or north-west. The species would subsequently have gone extinct in these areas as the regional climate changed (Zhou *et al.*, 2007; Favre *et al.*, 2015; Renner, 2016). This scenario of a late Miocene migration of *Q. aquifolium* from the east-central QTP into its present-day distribution range would explain our surprising discovery that the Tibetan *Q. aquifolium* populations are not most closely related with the neighbouring HDM populations but instead with those from the distant western Sichuan Plateau. It is further supported by the Migrate-n analysis and, significantly, by two Miocene and Pliocene records of the fossil oak taxon *Q. preguyavifolia* – morphologically equivalent probable ancestor of the extant *Q. aquifolioides* (Zhou *et al.*, 2007) – in central and southern Tibet respectively (see Fig. 1a and Appendix S1). Similar cases of evolutionary splits going along with – and probably driven by – colonization processes have been described from other tree species including oaks (Petit *et al.*, 2005) and have apparently been a common phenomenon in our study area (Favre *et al.*, 2015). Note that S-DIVA classified the split of the two lineages as vicariance event; we interpret this apparent contradiction as a consequence of the old age of the colonization process, as S-DIVA is known for its tendency to misidentify ancient dispersal as vicariance (Yu *et al.*, 2015).

We estimated the split of the WSP and the HDM lineage to have occurred around 4.4 (3.0–6.1) Ma. This would be consistent with the existence of Pliocene fossil records of *Q. preguyavifolia* at three sites located near the current south-western range border of our target species (LanPing, YongPing and ErYuan, see Fig. 1a and Appendix S1). Our estimate coincides with a period of intense uplift of the Hengduan Mountain massif (Sun *et al.*, 2011; Favre *et al.*, 2015), suggesting that the split could similarly have occurred during the colonization of the newly available terrain by the species. The notably star-like cpDNA haplotype structure of the HDM lineage suggests indeed that its current distribution range could have been colonized by a single ancient haplotype that diversified subsequently during a slowly succeeding species expansion across the Hengduan and Daxue Mountains. Such a scenario is supported by both the S-DIVA and the Migrate-n analyses: The first indicated that the extant Daxue Mountain populations could stem from a (northward) dispersal event; the second indicated notable historical gene flow from range sector C (which includes most of the HDM lineage) into sector B (which includes the Daxue Mountains). Our hypothesis represents a refined deep-time perspective of what Qiu *et al.* (2011, p. 240) called 'the predominant biotic response of temperate plant species to Quaternary environmental change' in the region: ancient colonization, followed by vicariance and population isolation with limited spatial demographic expansion. One can only speculate whether the mountain uplift and the resulting extreme topographic heterogeneity of the area have actually driven haplotype diversification in the HDM lineage during its slow expansion (acting as 'cradle') or whether they have only helped conserve the resulting elevated haplotype diversity (acting as 'museum').

Population and range dynamics after establishment of the major lineages

The three lineages varied greatly in the character and geographical distribution of their haplotypes. The HDM lineage was characterized by very high allelic richness, a pronounced geographical cpDNA structure with many populations fixed for haplotypes of local distribution, and the scarcity of 'missing' or hypothetical haplotypes within the notably star-like network structure. All these phenomena point towards great population stability through an extended period of time. Accordingly, our niche modelling exercise suggests that the Hengduan Mountains would have offered reasonably favourable climatic conditions through all modelled periods. In addition, the range of available topoclimates and microclimates must have been much larger in this extremely heterogeneous landscape than reflected by the niche model. Long-term persistence of numerous populations in the area throughout the Quaternary hence seems quite likely (see also Qiu *et al.*, 2011; Liu *et al.*, 2013).

In contrast to the HDM lineage, the Tibetan lineage was composed of relatively few remarkably distinct cpDNA haplotypes. Three of them occupy most of the area without showing geographical overlap, and private haplotypes are rare. Moreover, Tibetan populations are less genetically diverse than those of the other lineages. We interpret these findings as the outcome of a widespread extinction–recolonization process, where only some haplotypes would have survived in different refugia located in the region. Three of them would subsequently have expanded to fill the present-day Tibetan range. Interestingly, our niche modelling exercise indicated that *Q. aquifolioides* persistence in Tibet would have been most difficult during the Last Interglacial period (rather than during the LGM). This observation suggests that the species' long-term persistence in this area might actually have been more constrained by high evapotranspiration than by low temperatures.

The WSP lineage showed an intermediate behaviour between the other two lineages. Both its diversity and the frequency of private haplotypes were high, yet the existence of several unsampled haplotypes as well as the presence of some more widespread haplotypes (notably 6 and 7) could reflect some extinction and re-colonization dynamics in a relatively recent past. Their impact would, however, have been markedly weaker than in Tibet. It seems noteworthy

that the contact zone between the WSP and the HDM lineage (see Fig. 1b) consists of populations that appear to stem from relatively recent secondary expansions (as indicated by haplotypes 6 and 7 in WSP and by the derived haplotypes 18–22 in HDM), whereas we detected no more traces of the Pliocene colonization event that resulted in the establishment of the HDM lineage.

Disparate patterns of seed and pollen-mediated gene flow

Our combined survey of chloroplast and nuclear DNA revealed a stark contrast between range-wide patterns of seed and pollen-mediated gene flow. The almost complete differentiation and geographical structure of the cpDNA (G_{ST} = 0.93 and N_{ST} = 0.98) underpins that seed-mediated gene flow has been restricted to the few colonization events described above, and that posterior population admixture through seed dispersal has been virtually inexistent. Although marked spatial structuring of cpDNA markers is a widespread phenomenon in oaks and other nut-producing tree species (Xu *et al.*, 2015), our values clearly exceed those of any other range-wide study we are aware of. Importantly, the phenomenon is almost as evident within each lineage (average F_{ST} = 0.92, see Table 2) as among lineages, indicating that strong population divergence also exists within the same mountain ranges and not only along major geographical barriers as commonly outlined (e.g. Qiu *et al.*, 2011; Favre *et al.*, 2015).

The nSSR markers produced a totally different pattern. By far most variation (93%) occurred within populations. Our STRUCTURE analysis was unable to detect any difference between WSP and HDM populations, and even some HDM and Tibetan populations showed clear signs of gene exchange despite being separated by several scores of kilometres and the geographical barrier of the Boshulaling Mountains. The STRUCTURE analysis also revealed that this presumably pollen-mediated gene flow has primarily occurred from HDM towards Tibet (cf. strong presence of cluster I in populations CY, CYX, CBG and BMZ). One possible explanation for this asymmetry is that the Tibetan populations could have been particularly prone to incorporating immigrant genes during their relatively recent demographic expansion (Excoffier *et al.*, 2009). This secondary gene flow is also reflected in our Migrate-n analysis which detected a considerably weaker asymmetry in historical gene flow between range sectors A and C than between sectors A and B.

Conclusion: past and future of a major global hotspot of plant biodiversity

Many of the world's major diversity hotspots are located in areas that harbour very complex landscapes and have experienced gradual but significant climatic changes. This combination spurs evolutionary change while providing niches for long-term persistence, and it may explain why most reports of tree species with genetic structures reflecting Neogene

population dynamics stem from high-diversity regions of the world (Hampe & Petit, 2007). Although causal relationships between past geological processes and current patterns of biodiversity are inherently difficult to establish, our detailed phylogeographical survey across the Hengduanshan Biodiversity Hotspot provides various interesting insights into the biogeographical processes that could have contributed to its formation: First, the area acted in the late Neogene as a refugium for temperate species that immigrated from the rising and drying central parts of the QTP. Such range displacements have been reported from several other woody species of the region (Wang *et al.*, 2010; Xu *et al.*, 2010; Li *et al.*, 2013), and also have commonly been invoked as a potential catalysator of speciation processes in numerous herbaceous and some woody taxa of the region (reviewed in Wen *et al.*, 2014; Favre *et al.*, 2015; Hughes & Atchison, 2015). Second, remarkably great stability and low levels of extinction characterize populations in the area that harbours the centre of the hotspot (i.e. the Hengduan and Daxue Mountains). Note that *Q. aquifolioides* arrived relatively late in this area and the outlined population stability would have been limited to the latest Neogene and the Quaternary, fully consistent with the historical scenarios of López-Pujol *et al.* (2011) and Sun *et al.* (2011). Third, the haplotype radiation of the HDM lineage mirrors radiations at species level that have been reported for many herbaceous taxa in the area (López-Pujol *et al.*, 2011; Hughes & Atchison, 2015). The timing of this haplotype radiation could fall together with periods of intense radiation in some prominent herbaceous plant genera of the Hengduanshan Hotspot (e.g. *Delphinium*, *Rheum*, *Rhodiola* or *Isodon*; Hughes & Atchison, 2015), although the molecular estimates are burdened with large uncertainty. Fourth, even very rugged mountain ranges have apparently not impeded extensive, presumably pollen-mediated, gene flow that has further helped maintain a remarkably high diversity in *Q. aquifolioides* populations throughout the range. Fifth, more marginal parts of the hotspot and in particular the Tibetan part of the species range have apparently undergone notable extinction–recolonization dynamics during the Late Quaternary. Unfortunately, human activities are likely to significantly accelerate up extinction processes as only c. 8% of the original vegetation are left in the Hengduanshan Biodiversity Hotspot and landscape destruction is ongoing (Myers *et al.*, 2000). Ambitious conservation and management efforts will hence be required that should focus on providing elevational migration routes at local scales in order to conserve as well as possible the species and within-species diversity that characterizes this globally unique region.

REFERENCES

Bandelt, H.J., Forster, P. & Röhl, A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Beerli, P. (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, **22**, 341–345.

Collins, W.D., Bitz, C.M., Blackmon, M.L., Bonan, G.B., Bretherton, C.S., Carton, J.A., Chang, P., Doney, S.C., Hack, J.J., Henderson, T.B., Kiehl, J.T., Large, W.G., McKenna, D.S., Santer, B.D. & Smith, R.D. (2006) The community climate system model version 3 (CCSM3). *Journal of Climate*, **19**, 2122–2143.

Donoghue, M.J. (2008) A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences USA*, **105**, S11549–S11555.

Dow, B.D., Ashley, M.V. & Howe, H.F. (1995) Characterization of highly variable $(GA/CT)_n$ microsatellites in the bur oak, *Quercus macrocarpa*. *Theoretical & Applied Genetics*, **91**, 137–141.

Drummond, A.J. & Rambaut, A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Durand, J., Bodénès, C., Chancerel, E. *et al.* (2010) A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. *BMC Genomics*, **11**, 570.

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

Excoffier, L. & Lischer, H.E. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier, L., Foll, M. & Petit, R.J. (2009) Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 481–501.

Favre, A., Päckert, M., Pauls, S.U., Jähnig, S.C., Uhl, D., Michalak, I. & Muellner-Riehl, A.N. (2015) The role of the uplift of the Qinghai-Tibetan plateau for the evolution of Tibetan biotas. *Biological Reviews*, **90**, 236–253.

Hampe, A. & Petit, R.J. (2007) Ever deeper phylogeographies: trees retain the genetic imprint of Tertiary plate tectonics. *Molecular Ecology*, **16**, 5113–5114.

Hasumi, H. & Emori, S. (2004) *K-1 Coupled GCM (MIROC) description*. K-1 Tech. Rep. 1, Center for Climate System Research (CCSR), University of Tokyo, National Institute for Environmental Studies (NIES), Frontier Research Center for Global Change (FRCGC), 34 pp.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hoorn, C., Wesselingh, F.P., Ter Steege, H. *et al.* (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927–931.

Huang, C.J., Zhang, Y.L. & Bartholomew, R. (1999) Fagaceae. *Flora of China*. Vol 4. (ed. by Z.Y. Wu and P.H. Raven), pp. 314–380. Science Press, Beijing. Missouri Botanical Garden Press, St. Louis.

Hughes, C.E. & Atchison, G.W. (2015) The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytologist*, **207**, 275–282.

Kampfer, S., Lexer, C., Glössl, J. & Steinkellner, H. (1998) Characterization of $(GA)_n$ microsatellite loci from *Quercus robur*. *Hereditas*, **129**, 183–186.

Kanno, M., Yokoyama, J., Suyama, Y., Ohyama, M., Itoh, T. & Suzuki, M. (2004) Geographical distribution of two haplotypes of chloroplast DNA in four oak species (*Quercus*) in Japan. *Journal of Plant Research*, **117**, 311–317.

Li, B.Y. (1987) On the boundaries of the Hengduan Mountains. *Journal of Mountain Research*, **5**, 74–82.

Li, L., Abbott, R.J., Liu, B., Sun, Y., Li, L., Zou, J., Wang, X., Miehe, G. & Liu, J. (2013) Pliocene intraspecific divergence and Plio-Pleistocene range expansions within *Picea likiangensis* (Lijiang spruce), a dominant forest tree of the Qinghai-Tibet Plateau. *Molecular Ecology*, **22**, 5237–5255.

Librado, P. & Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Linder, H.P. (2008) Plant species radiations: where, when, why? *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 3097–3105.

Lippert, P.C., Van Hinsbergen, D.J. & Dupont-Nivet, G. (2014) Early Cretaceous to present latitude of the central proto-Tibetan Plateau: a paleomagnetic synthesis with implications for Cenozoic tectonics, paleogeography, and climate of Asia. *Geological Society of America Special Papers*, **507**, SPE507-01.

Liu, J., Möller, M., Provan, J., Gao, L.M., Poudel, R.C. & Li, D.Z. (2013) Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. *New Phytologist*, **199**, 1093–1108.

López-Pujol, J., Zhang, F.M., Sun, H.Q., Ying, T.S. & Ge, S. (2011) Centres of plant endemism in China: places for survival or for speciation? *Journal of Biogeography*, **38**, 1267–1280.

Matzke, N.J. (2012) Founder-event speciation in BioGeoBEARS package dramatically improves likelihoods and alters parameter inference in dispersal–extinction–cladogenesis (DEC) analyses. *Frontiers of Biogeography*, **4**, 210.

Mulch, A. & Chamberlain, C.P. (2006) Earth science: the rise and growth of Tibet. *Nature*, **439**, 670–671.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Otto-Bliesner, B.L., Marshall, S.J., Overpeck, J.T., Miller, G.H. & Hu, A. (2006) Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science*, **311**, 1751–1753.

Peakall, R. & Smouse, P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, **28**, 2537–2539.

Pennington, R.T., Lavin, M., Särkinen, T., Lewis, G.P., Klitgaard, B.B. & Hughes, C.E. (2010) Contrasting plant diversification histories within the Andean biodiversity hotspot. *Proceedings of the National Academy of Sciences USA*, **107**, 13783–13787.

Petit, R.J. & Hampe, A. (2006) Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 187–214.

Petit, R.J., Hampe, A. & Cheddadi, R. (2005) Climate changes and tree phylogeography in the Mediterranean. *Taxon*, **54**, 877–885.

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

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

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Qian, H. & Ricklefs, R.E. (2000) Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature*, **407**, 180–182.

Qiu, Y.X., Fu, C.X. & Comes, H.P. (2011) Plant molecular phylogeography in China and adjacent regions, tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. *Molecular Phylogenetics and Evolution*, **59**, 225–244.

Renner, S.S. (2016) Available data point to a 4-km-high Tibetan Plateau by 40 Ma, but 100 molecular-clock papers have linked supposed recent uplift to young node ages. *Journal of Biogeography*, doi:10.1111/jbi.12755.

Sechrest, W., Brooks, T.M., da Fonseca, G.A.B., Konstant, W.R., Mittermeier, R.A., Purvis, A., Rylands, A.B. & Gittleman, J.L. (2002) Hotspots and the conservation of evolutionary history. *Proceedings of the National Academy of Sciences USA*, **99**, 2067–2071.

Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E. & Small, R.L. (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142–166.

Sun, B.N., Wu, J.Y., Liu, Y.S.C., Ding, S.T., Li, X.C., Xie, S.P., Yan, D.F. & Lin, Z.C. (2011) Reconstructing Neogene vegetation and climates to infer tectonic uplift in western Yunnan, China. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **304**, 328–336.

Sun, Y., Abbott, R.J., Li, L., Li, L., Zhou, J.B. & Liu, J.Q. (2014) Evolutionary history of Purple cone spruce (*Picea purpurea*) in the Qinghai-Tibet Plateau: homoploid hybrid origin and Pleistocene expansion. *Molecular Ecology*, **23**, 343–359.

Szpiech, Z.A., Jakobsson, M. & Rosenberg, N.A. (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, **24**, 2498–2504.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.

Wang, H., Qiong, L.A., Sun, K., Lu, F., Wang, Y., Song, Z.P., Wu, Q.H., Chen, J.K. & Zhang, W. (2010) Phylogeographic structure of *Hippophae tibetana* (Elaeagnaceae) highlights the highest microrefugia and the rapid uplift of the Qinghai-Tibetan Plateau. *Molecular Ecology*, **19**, 2964–2979.

Wen, J., Zhang, J.Q., Nie, Z.L., Zhong, Y. & Sun, H. (2014) Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Frontiers in Genetics*, **5**, 4.

Xu, J., Deng, M., Jiang, X.L., Westwood, M., Song, Y.G. & Turkington, R. (2015) Phylogeography of *Quercus glauca* (Fagaceae), a dominant tree of East Asian subtropical evergreen forests, based on three chloroplast DNA interspace sequences. *Tree Genetics and Genomes*, **11**, 1–17.

Xu, T., Abbott, R.J., Milne, R.I., Mao, K., Du, F.K., Wu, G., Ciren, Z., Miehe, G. & Liu, J. (2010) Phylogeography and allopatric divergence of cypress species (*Cupressus* L.) in the Qinghai-Tibetan Plateau and adjacent regions. *BMC Evolutionary Biology*, **10**, 194.

Yu, Y., Harris, A.J., Blair, C. & He, X.J. (2015) RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution*, **87**, 46–49.

Zhou, Z., Yang, Q. & Xia, K. (2007) Fossils of *Quercus* sect. *Heterobalanus* can help explain the uplift of the Himalayas. *Chinese Science Bulletin*, **52**, 238–247.

Zhou, Z.K. (1992) Origin, phylogeny and dispersal of *Quercus* from China. *Acta Botanica Yunnanica*, **14**, 227–236.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Sampling and fossil records details.
Appendix S2 Detailed information on cp DNA haplotype.

Appendix S3 Detailed information on nSSR marker selection and genotyping.

BIOSKETCH

Fang K. Du is interested in unravelling the evolutionary history of forest trees. Her newly organized Molecular Ecology lab in BJFU focuses on the phylogeography, molecular evolution and local adaptation on several oak species in China.

Editor: Hans-Peter Comes