

Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot

Jie Liu^{1,2,3}, Michael Möller^{1,4}, Jim Provan⁵, Lian-Ming Gao¹, Ram Chandra Poudel^{1,2,3} and De-Zhu Li^{1,2}

¹Key Laboratory of Biodiversity and Biogeography, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China; ²Plant Germplasm and Genomics Center, Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China; ³University of Chinese Academy of Sciences, Beijing 100049, China; ⁴Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, UK; ⁵School of Biological Sciences, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL, UK

Authors for correspondence:

Lian-Ming Gao

Tel: +86 871 65223505

Email: gaolm@mail.kib.ac.cn

De-Zhu Li

Tel: +86 871 65223503

Email: dzl@mail.kib.ac.cn

Received: 6 March 2013

Accepted: 17 April 2013

New Phytologist (2013) 199: 1093–1108

doi: 10.1111/nph.12336

Key words: ecological differentiation, Himalaya-Hengduan Mountains region, population demography, speciation, species distribution modeling, *Taxus wallichiana*.

Summary

- The interplay of orographic uplift and climatic changes in the Himalaya-Hengduan Mountains region (HHM) have had a key role in speciation and population demography. To gain further insight into these processes, we investigated their effects on *Taxus wallichiana* by combining molecular phylogeography and species distribution modeling.
- Molecular data were obtained from 43 populations of *T. wallichiana*. Nineteen climatic variables were analyzed alongside genetic discontinuities. Species distribution modeling was carried out to predict potential past distribution ranges.
- Two distinct lineages were identified, which diverged c. 4.2 (2.0–6.5) million years ago (Ma), a timescale that corresponds well with the recent uplift of the Qinghai-Tibet Plateau and subsequent climatic changes of the region. Correlations with climatic variables also suggest that ecological factors may have further reinforced the separation of the two lineages. Both lineages experienced population expansion during the last glaciation.
- The high genetic divergence, long-term isolation and ecological differentiation suggest a scenario of cryptic speciation in *T. wallichiana* associated with geological and climatic changes in the HHM. Our findings also challenge the notion of general population 'contraction' during the last glaciation in the HHM.

Introduction

The origin and evolution of biodiversity is strongly linked with many historical and ecological factors. These can include geological and/or climatic processes such as continental drift, the uplift of mountain chains, and climatic oscillations associated with ice ages. The interaction of these processes can create new ecological niches and thus provide opportunities for speciation. Indeed, divergent selection and adaptation associated with different habitats or ecological niches are increasingly viewed as a major cause of speciation in plants (Rieseberg & Burke, 2001; Funk *et al.*, 2006; Harmon *et al.*, 2008). A range of ecological factors have been proposed in driving speciation, including climatic variables such as precipitation and temperature (see Givnish, 2010; Keller & Seehausen, 2012; Wagner *et al.*, 2012), and differential climatic adaptation has been implicated in cases of incipient or completed speciation events (Nosil *et al.*, 2005; Lowry *et al.*, 2008).

The Himalaya-Hengduan Mountains (HHM) region is considered a key biodiversity hotspot (Myers *et al.*, 2000). The region was formed by the collision of the Indian continent with the Eurasian plate 40–50 million yr ago (Ma) (Yin & Harrison, 2000), and the consequent rise of the Himalaya and the Qinghai-Tibet Plateau

(QTP). Despite some controversy around the exact timing of the formation of the QTP (e.g. Spicer *et al.*, 2003), episodes of uplifts continued throughout the late Pliocene (c. 3 Ma) and well into the Quaternary (c. 2.6 Ma; Li & Fang, 1999). The rise of the QTP modified global and East Asian climate dramatically (Ruddiman & Kutzbach, 1989; Shi *et al.*, 1998), triggering and intensifying the Asian monsoon, which in turn profoundly influenced biological processes in the QTP (Li & Fang, 1999). Present atmospheric circulation patterns over the QTP and surrounding areas are characterized by the Indian and East Asian monsoon in summer and the Westerlies in winter (An *et al.*, 2001; Yao *et al.*, 2012b).

The HHM region, which extends along the southern and southeastern fringe of the QTP (Fig. 1), is a region of extreme elevational changes within relatively short distances, with a corresponding diversity of habitats. In these areas, the wet summer monsoons are blocked by North to South oriented mountain ranges, such as the Gaoligong Mountains and Mekong-Salween Divide, resulting in the formation of a longitudinal moisture gradient (Fig. 1). This area shows a profound ecological heterogeneity and species diversity, which has been linked to the geological and climatic effects associated with the uplift of the QTP. More recently, climatic fluctuations during the Quaternary (2.6 Ma –

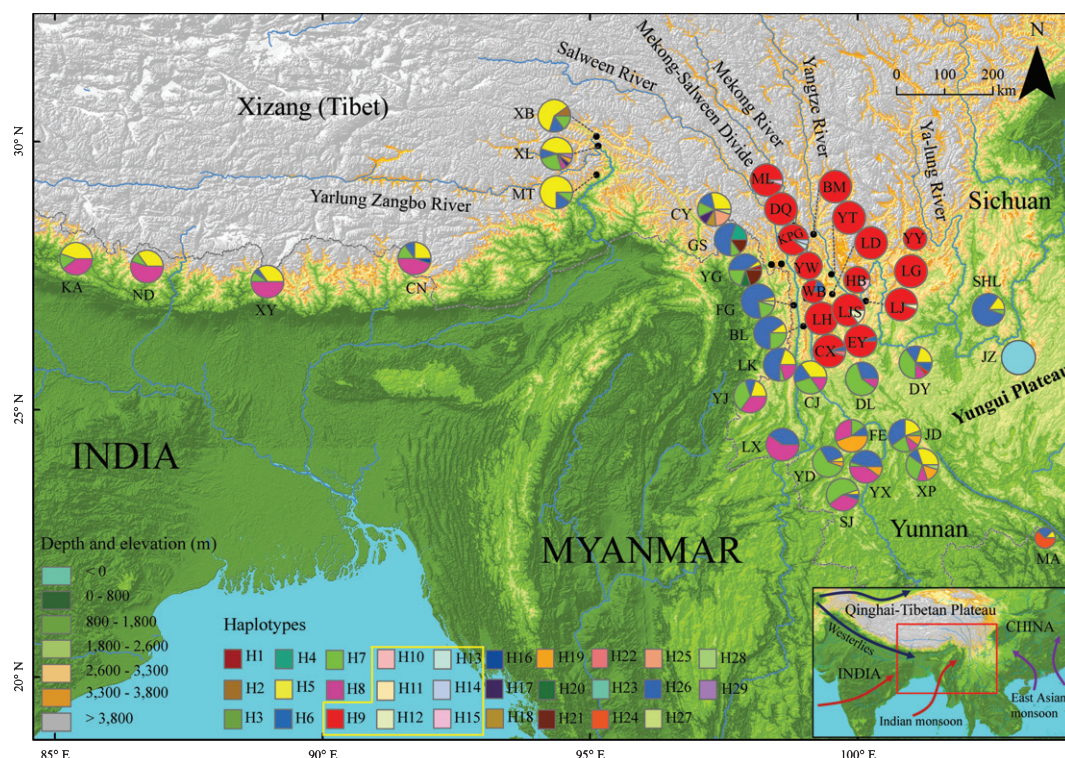


Fig. 1 Geographic locations of the 43 *Taxus wallichiana* populations analyzed in the present study and the distribution of the 29 cpDNA haplotypes detected (see Table 1 for population codes; boxed haplotypes represent those found in the HM lineage). Pie chart size corresponds to the sample size of each population. Present atmospheric circulation patterns over the Qinghai-Tibet Plateau (QTP) and surroundings were shown in the lower-right corner (Adapted from An *et al.*, 2001 and Yao *et al.*, 2012b).

present) might have further driven allopatric divergence and speciation, as populations contracted into glacial refugia during the four major periods of glaciations in the area (e.g. Wang *et al.*, 2010; Xu *et al.*, 2010; Jia *et al.*, 2012). Such climatic oscillations had profound effects on population demography, causing range shifts or extinctions, as well as possibly driving local adaptation (Davis & Shaw, 2001; Hewitt, 2004).

With such a complex geological, climatic and ecological diversity, the HHM regions are ideal for studying the effects of different factors on species diversification and evolution. Previous phylogenetic and biogeographical studies in the QTP and neighboring regions have focused on species-level diversification resulting from the uplift of the QTP (e.g. Liu *et al.*, 2002, 2006; Wang *et al.*, 2009b). Phylogeographical analyses at intraspecific levels have suggested that the divergence and demography of populations were also profoundly affected by the rise of the QTP and the Quaternary climatic oscillations in this area (e.g. Zhang *et al.*, 2005; Yang *et al.*, 2008, 2012; Wang *et al.*, 2010; Xu *et al.*, 2010; Jia *et al.*, 2011, 2012; Li *et al.*, 2011), with deep intraspecific divergences formed over a range of timescales (e.g. Jia *et al.*, 2012; Li *et al.*, 2011; Wang *et al.*, 2009a). Among these studies, however, very few (Cun & Wang, 2010; Opgenoorth *et al.*, 2010; Xu *et al.*, 2010) have focused specifically on the HHM region because of its remoteness and inaccessibility, and, consequently, the mechanisms of the interplay between geological and climatic events in driving speciation and evolution remain understudied.

Taxus wallichiana Zucc. var. *wallichiana*, as defined in the Flora of China (Fu *et al.*, 1999), was recently raised to species

level (Farjon, 2010). It included two genetically closely related lineages, designated *T. wallichiana* proper and the Hengduan type (Liu *et al.*, 2011b). The species is confined to the East Himalaya and the Yunnan plateau region (Gao *et al.*, 2007; Poudel *et al.*, 2012; Fig. 1), where it grows as scattered understory trees in subtropical forest between 2000 and 3500 m altitude. Thus, this species is ideal to study the interplay of the geological and climatic factors on population divergence and demographic history in the HHM region.

While most previous studies in the QTP and/or HHM region used genetic discontinuities to demonstrate incipient or cryptic speciation events, they were explained as a result of geological and presumed associated habitat and climatic variations, but ecological data were not explicitly examined to support these processes. Here, we use an integrative approach by combining genetic approaches with climatic and paleoenvironmental data to elucidate the speciation and demographic history of *T. wallichiana*, and determine the role of a range of climatic variables in this scenario. The specific objectives were: to quantify the genetic differences between the two lineages of *T. wallichiana*; to date the divergence between the two lineages and relate their age to geological and climatic events; to determine in detail the main ecological (climatic) factors separating the two lineages; and to unravel the demographic history of *T. wallichiana*. The outcome of these analyses will contribute to a better understanding of how environmental factors interacted with the orogeny of the QTP, contributing to the present-day abundance of species diversity in the HHM region.

Table 1 Genetic diversity within populations of *Taxus wallichiana* based on chloroplast haplotype and nuclear microsatellite DNA data

Population code	cpDNA <i>trnL-trnF</i>	Nuclear microsatellite							
	Haplotypes (number of individuals)	H_P	H_d	$\pi (\times 10^{-2})$	N_A	N_P	A_R	H_O	H_E
EH group									
KA	H5(9), H7(3), H8(8)	0	0.647	0.222	20	0	2.05	0.21	0.29
ND	H5(7), H7(2), H8(11)	0	0.595	0.204	32	2	2.77	0.37	0.43
XY	H5(7), H6(1), H7(2), H8(10)	0	0.647	0.204	26	0	2.38	0.27	0.36
CN	H5(5), H6(2), H7(3), H8(9), H16(1)	1	0.737	0.206	33	1	2.86	0.41	0.48
MT	H5(15), H6(3), H7(2)	0	0.426	0.120	35	2	3.15	0.40	0.50
XL	H5(9), H6(2), H7(5), H8(1), H17(1), H19(1), H29(1)	1	0.753	0.242	35	0	3.17	0.48	0.51
XB	H2(2) , H5(12), H6(3), H7(3)	1	0.616	0.164	35	0	2.99	0.35	0.46
CY	H5(6), H6(3), H7(3), H17(2) , H18(2) , H25(4), H27(1)	2	0.862	0.294	39	1	3.39	0.52	0.55
GS	H4(4) , H6(11), H7(2), H21(3)	1	0.658	0.211	40	1	3.53	0.36	0.55
YG	H1(1) , H5(1), H6(8), H7(4), H20(2) , H21(4)	2	0.784	0.159	42	0	3.61	0.43	0.53
FG	H5(2), H6(16), H7(4), H27(1)	0	0.502	0.116	35	0	3.19	0.45	0.50
BL	H5(2), H6(13), H7(5)	0	0.532	0.103	32	0	3.00	0.42	0.47
LK	H5(4), H6(11), H7(1), H8(4)	0	0.647	0.136	28	0	2.80	0.44	0.45
CJ	H5(7), H6(4), H7(6), H8(3)	0	0.763	0.138	39	0	3.37	0.51	0.56
YJ	H5(4), H6(2), H7(7), H8(7)	0	0.742	0.199	30	0	2.90	0.40	0.50
DL	H6(6), H7(12), H8(2)	0	0.568	0.083	34	0	3.22	0.53	0.53
DY	H5(4), H6(3), H7(8), H8(2), H9(1), H26(2)	0	0.795	0.268	39	0	3.34	0.46	0.51
SHL	H5(3), H6(16), H7(1)	0	0.353	0.078	29	0	2.76	0.44	0.45
JZ	H23(20)	1	0.000	0.000	32	0	3.05	0.45	0.47
LX	H6(8), H8(12)	0	0.505	0.060	24	0	2.46	0.36	0.38
YD	H5(1), H6(5), H7(11), H19(1)	0	0.575	0.102	35	1	3.29	0.48	0.54
FE	H6(2), H7(3), H8(6), H19(9)	0	0.711	0.147	32	0	3.03	0.43	0.47
YX	H6(9), H7(1), H8(8), H19(2)	0	0.658	0.095	32	1	3.19	0.49	0.53
SJ	H5(1), H7(11), H8(7), H26(1)	0	0.600	0.157	28	0	2.72	0.44	0.42
JD	H3(1) , H5(4), H6(7), H7(4), H8(2), H19(2),	1	0.816	0.196	33	0	3.14	0.52	0.53
XP	H5(6), H6(2), H7(6), H8(2), H19(3), H28(1)	1	0.816	0.235	30	0	2.98	0.44	0.51
MA	H5(1), H6(2), H24(5)	1	0.607	0.192	26	0	2.79	0.52	0.43
HM group									
ML	H9(19), H15(1)	1	0.100	0.012	35	0	2.85	0.28	0.36
DQ	H9(20)	0	0.000	0.000	33	1	2.87	0.28	0.42
BM	H9(20)	0	0.000	0.000	38	0	2.96	0.28	0.37
KPG	H9(16), H13(1) , H14(1)	2	0.216	0.027	41	0	3.47	0.36	0.47
YT	H9(20)	0	0.000	0.000	37	0	2.97	0.29	0.38
YW	H9(14)	0	0.000	0.000	32	0	2.91	0.27	0.41
LD	H9(20)	0	0.000	0.000	37	0	2.89	0.29	0.36
LH	H9(20)	0	0.000	0.000	40	0	3.28	0.32	0.41
LJS	H9(17), H12(3)	1	0.268	0.033	38	0	3.16	0.32	0.47
LJ	H9(19), H11(1)	1	0.100	0.012	36	0	3.24	0.36	0.46
HB	H9(11), H10(2)	1	0.282	0.034	41	1	3.58	0.41	0.46
LG	H9(20)	0	0.000	0.000	34	1	3.10	0.34	0.45
YY	H9(10)	0	0.000	0.000	29	0	2.92	0.37	0.39
EY	H6(1), H9(19)	0	0.100	0.085	32	0	2.86	0.24	0.39
WB	H6(2), H9(8)	0	0.356	0.304	38	0	3.79	0.39	0.49
CX	H6(1), H9(18), H22(1)	1	0.195	0.174	41	0	3.43	0.36	0.51
Mean		0.442	0.626	0.160	33.88	0.28	3.06	0.39	0.46

Private haplotypes are given in bold print.

N , number of individuals per population; H_P , number of private haplotypes; H_d , haplotype diversity; π , nucleotide diversity; N_A , number of total alleles; A_R , alleles richness; N_P , number of private alleles; H_O , observed heterozygosity; H_E , expected heterozygosity; EH group, East Himalaya to the Yunnan Plateau region; HM group, South Hengduan Mountains region.

Materials and Methods

Sample collection and DNA isolation

A total of 815 individuals were sampled from 43 populations of *Taxus wallichiana* Zucc. for this study, with 16 populations

originating from the South Hengduan Mountains region (HM group) and 27 from the East Himalaya to the Yunnan Plateau region (EH group) (Fig. 1; Supporting Information Table S1). Eight to 23 individuals were collected per population (average 19 plants per population). Young and healthy leaves were collected and dried immediately in silica gel for DNA extraction. Voucher

specimens of each sampled individual are deposited at the Herbarium of the Kunming Institute of Botany, Chinese Academy of Sciences (KUN). *Taxus brevifolia* was used as an outgroup for the phylogenetic analyses of the cpDNA haplotypes, based on a previous study (Hao *et al.*, 2008).

Total genomic DNA was isolated using a modified CTAB method as described by Liu & Gao (2011). The quality and quantity of DNA was determined on 1% TAE agarose gels and with a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). The DNA was diluted to a final concentration of 30–50 ng µl⁻¹.

Chloroplast DNA sequencing

The plastid *trnL* (UAA)-*trnF* (GAA) region (*trnL-trnF*) was amplified for all accessions with primers 'c' and 'f' described in Taberlet *et al.* (1991). The 25 µl PCR mix contained 2.5 µl 10× PCR buffer, 2.5 µl MgCl₂ (25 mM), 2.0 µl dNTP mixture (2.5 mM), 0.3 µl each primer (10 µM), 0.2 µl Taq polymerase (5 U µl⁻¹) (TaKaRa, Dalian, China), 1 µl template DNA (c. 30–50 ng genomic DNA), and finally 16.2 µl distilled deionized water. PCRs were performed on a GeneAmp PCR System 9700 thermal cycler (Perkin Elmer, Foster City, CA, USA) with the following profile: initial denaturation at 94°C for 3 min, followed by 30 cycles of 95°C for 1 min, 50°C for 1.5 min, 72°C for 2 min, and a final step of 72°C for 10 min.

PCR products were visualized on 1% TAE agarose gels and then purified using a Sangon Purification Kit (Sangon, Shanghai, China). The purified products were used for bi-directional sequencing using the PCR primers and the PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Foster City, CA, USA). The products were run on an ABI 3730 xl DNA Sequencer. All newly acquired sequences have been submitted to GenBank (accession numbers in Table S2; *T. brevifolia*, JQ406921).

Microsatellite genotyping

The nuclear genetic diversity of all 815 accessions was assessed using the following nine microsatellite loci: TA116, TG111, TG06, TG34, TG47, TC04, Tax86, TW01, TS03 (Liu *et al.*, 2011a). Forward primers were 5'-end fluorescently labelled using either FAM or TAMRA (Applied Biosystems). PCR was carried out on a GeneAmp PCR System 9700 thermal cycler (Perkin Elmer) following the protocols described in Liu *et al.* (2011a). PCR products were separated on an ABI 3730 xl DNA Sequencer and individuals genotyped with GENEMAPPER v3.2 (Applied Biosystems).

Chloroplast DNA analysis

The chromatograms of each *trnL-trnF* sequence were assembled with SEQMAN (DNASTar Inc., Madison, WI, USA). Consensus sequences were aligned in CLUSTAL X v2.0 (Larkin *et al.*, 2007) and subsequently manually adjusted where necessary. All sequences were assigned to different haplotypes using DNASP

v5.10 (Librado & Rozas, 2009). All indels were treated as single mutation events. A median-joining haplotype network was constructed using the program NETWORK v4.5.1 (available at <http://www.fluxus-engineering.com>) with MP criterion.

The spatial genetic structuring of the chloroplast haplotypes was analyzed using a spatial analysis of molecular variance analysis (SAMOVA v1.0; Dupanloup *et al.*, 2002). Molecular diversity indices, including the number of haplotypes, haplotype diversity (H_d) and nucleotide diversity (π), were estimated using ARLEQUIN v3.5.1 (Excoffier & Lischer, 2010).

The existence of phylogeographic structure was tested following Pons & Petit (1996). Two estimates of population diversity (H_S , H_T) and two of differentiation (G_{ST} , N_{ST}) were obtained using PERMUT v1.0 (<http://www.pierroton.inra.fr/genetics/labo/Software/>) with 1000 permutations. An analysis of molecular variance (AMOVA) was performed to partition variation within and among defined groups and populations using ARLEQUIN.

Neutrality tests with Tajima's D and Fu's F_S were implemented in ARLEQUIN to test for deviations from neutrality. In an attempt to further infer demographic processes, pairwise mismatch distributions were carried out using ARLEQUIN to infer the demographic history of *T. wallichiana* at the lineage level. The sum of squared deviation (SSD) and Harpending's raggedness index (HRI) between observed and expected mismatch distributions were used as test statistics. Where the null hypothesis was not rejected, the formula $t = \tau/2u$ was used to estimate the age of expansion (t), where $u = \mu kg$, and μ the mutation rate of the sequence in substitutions per site per year (s/s/y), k the number of nucleotides and g the generation time in years. Because no evolutionary rates for gymnosperms are available for the *trnL-trnF* region, we obtained the average rate of *trnL-trnF* for Taxaceae (see next paragraph) using BEAST. For the generation time g , a value of 25 yr was taken according to Wang *et al.* (2006).

trnL-trnF mutation rate and lineage divergence time

In order to estimate the *trnL-trnF* mutation rate of Taxaceae, we first estimated the divergence time of Taxaceae and related genera. Sequences of two cpDNA regions (*rbcl* and *matK*) were retrieved from GenBank or newly generated as described in Liu *et al.* (2011b), for 66 samples representing 39 genera of extant gymnosperms (dataset I: Table S3). The fossil record of modern Taxaceae is generally scarce (Hao *et al.*, 2008). Thus, only one fossil *Marskea jurassica* (Harris, 1976) was chosen as the earliest known fossil for Taxaceae s.s., designated 'K' with an age of 165 Ma (Florin, 1958) (Table S4; Fig. S2). In addition, ten well-identified fossils from other gymnosperms were used to calibrate the tree following Mao *et al.* (2012) (Table S4). We used the age estimate for Taxaceae and related genera from dataset I, and dataset II (which consisted of *trnL-trnF* sequences of 39 species of Taxaceae plus four outgroups, *Cunninghamia lanceolata* var. *konishii*, *Cupressus gigantea*, *Juniperus taxifolia*, *Taiwania cryptomerioides*, retrieved from GenBank, Table S3), to estimate the *trnL-trnF* mutation rate for Taxaceae. With the *trnL-trnF* mutation rate, we estimated the divergence time of the two

lineages of *T. wallichiana*, using dataset III (comprising the 29 haplotypes of *T. wallichiana* and the outgroup *T. brevifolia*).

For all three datasets, phylogenetic analyses were carried out using BEAST v1.6.2 (Drummond & Rambaut, 2007) with a GTR + G + I substitution model, selected by MODELTEST v3.6 (Posada & Crandall, 1998), and an uncorrelated lognormal relaxed clock (Drummond *et al.*, 2002). A Speciation: birth to death process was specified as tree prior. The uniform distribution was chosen for dataset I, and normal distribution was selected for datasets II and III, respectively, as suggested in the BEAST manual. Each BMCMC was run for 70 million generations, sampling every 7000th generation, and discarding the first 10% as burn-in. All these analysis were carried out on the Oslo Biportal (<http://www.biportal.uio.no/>). Convergence was checked using TRACER v1.5 (Available from <http://beast.bio.ed.ac.uk/Tracer>). Trees were visualized with TREEANOTATOR v1.6.2 in BEAST.

Microsatellite data analysis

For all nine loci, MICROCHECKER v2.2.3 (Van Oosterhout *et al.*, 2004) was used to detect the presence of null alleles and genotyping errors such as large allele dropout or stuttering using 1000 randomizations. Tests for departures from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were performed in each population as well as a globally unified population using FSTAT v2.9.3 (Goudet, 2001). Significance levels were adjusted using the sequential Bonferroni correction for multiple comparisons (Rice, 1989).

Genetic diversity indices (total number of alleles, N_A ; observed heterozygosity, H_O ; expected heterozygosity, H_E) were calculated with GENALEX v6.4 (Peakall & Smouse, 2006) for each population, locus and group. FSTAT was used to estimate allelic richness (A_R) at population, locus and group levels.

We estimated global F_{ST} (Weir & Cockerham, 1984) and population-pairwise levels of F_{ST} using ARLEQUIN. In addition, population differentiation F_{ST} was also estimated for each locus across all populations using FSTAT. The standardized genetic differentiation (G'_{ST} ; Hedrick, 2005) was calculated with SMOGD (Crawford, 2010). G'_{ST} is a more suitable measure of differentiation than traditional measures (e.g. F_{ST} and G_{ST}) for highly polymorphic markers such as microsatellite (Hedrick, 2005; Crawford, 2010). Partitioning of total genetic variation within and among populations was further analyzed by AMOVA with 1000 permutations using ARLEQUIN.

Starting from individual genotypes, the number of genetic clusters was inferred using the program STRUCTURE v2.3 (Pritchard *et al.*, 2000). We applied the admixture model and the option of correlated allele frequencies between populations as recommended by Falush *et al.* (2003). The simulation was run with the number of clusters (K) from one to twenty, and was repeated twenty times for each K . Each run comprised of a burn-in period of 10^5 iterations, followed by 10^5 Markov chain Monte Carlo (MCMC) steps. The most likely value for the number of populations (K) was determined by log-likelihood value, and the second-order rate of change in values of K (ΔK) method suggested by Evanno *et al.* (2005).

BARRIER v2.2 (Manni *et al.*, 2004) was used to compute the Monmonier's maximum-difference algorithm for identifying biogeographic boundaries or areas exhibiting the largest genetic discontinuities between population pairs. The robustness of these boundaries was assessed by running BARRIER on 1000 Nei's distance matrices, produced by bootstrapping over loci with MICROSATELLITE ANALYSER (MSA) v4.05 (Dieringer & Schlötterer, 2003).

Climatic data analysis

Based on our population geo-referenced locality data (Table S1), we used recent (*c.* 1950–2000) data for 19 BIOCLIM variables (Hijmans *et al.*, 2005), which summarized temperature and precipitation variables (Table S5), to identify climatic factors potentially associated with the genetic structure and divergence between the EH and HM populations. Because the landscape of the HHM region is topographically heterogeneous, for each set of coordinates we extracted monthly temperature and precipitation values from the high-resolution global gridded climate layer data with a grid size of 30", which corresponds to 1 km² at the equator. We used DIVA GIS (Hijmans *et al.*, 2001) to extract the climatic variables and calculated the mean values and standard errors with MICROSOFT EXCEL for each population and a two-tailed *t*-test on the EH and HM groups separately. We further conducted paired two-tailed *t*-tests on pairs of populations (EH and HM) along the westerly and southerly gene flow barrier between EH and HM populations in order to characterize the strength of ecological factors in the barrier. The pairs were arranged by shortest distance across the barrier or by matching altitudes where possible to eliminate the effects of altitude on temperature *per se* (Table S6).

Species distribution modeling (SDM)

SDM was carried out to determine suitable present and past climate envelopes for the HM and EH lineages. Models were generated for the Last Glacial Maximum (LGM; *c.* 21 ka), based on the Community Climate System Model (CCSM), and the Last Interglacial (LIG; *c.* 120 ka). Species occurrence data were obtained from the Global Biodiversity Information Facility (GBIF; <http://data.gbif.org>) and personal observations, totaling 67 and 31 occurrences for the EH and HM lineages, respectively. Niche models based on the 19 BIOCLIM variables in the WorldClim dataset (Hijmans *et al.*, 2005) were generated at 2.5 min resolution using the MAXENT v3.2 software package (Phillips *et al.*, 2006) with the default parameters for convergence threshold (10^{-5}) and number of iterations (500). Duplicate records from the same locality were removed to reduce the effects of spatial autocorrelation.

Results

Chloroplast DNA sequencing data

The lengths of the cpDNA *trnL-trnF* sequences ranged from 812 to 912 bp and the aligned matrix was 914 bp in length. The

alignment contained 20 polymorphic sites and 8 indels (Table S2). A total of 29 chloroplast haplotypes was identified among the 815 accessions (Fig. 1; Table 1 and Table S2).

The BEAST-derived tree and haplotype network revealed similar relationships (Fig. 2). The 29 haplotypes clustered into two lineages, one including 22 haplotypes from the EH lineage, and the other consisting of seven haplotypes from the HM lineage (Fig. 2a). The haplotype network showed two clades, also corresponding to the EH and HM lineages (Fig. 2b; Table 1). Of the 22 haplotypes found in the EH lineage, 13 were private. Within this lineage, haplotypes H5, H6, H7 and H8 were most frequent, and were widely distributed across the EH populations. The remaining eighteen haplotypes were linked to the four main haplotypes in a 'star-like' network. Of the seven haplotypes detected in the HM lineage, haplotype H9 was most abundant and predominant in all populations within the HM lineage (Fig. 1; Table 1). The other six were private, and connected to H9 by single steps in a 'star-like' network (Fig. 2b). Four populations (i.e. WB, CX, EY and DY, Fig. 1; Table 1) possessed haplotypes from both the EH and HM lineages.

Molecular genetic diversity indices H_d and π for each population are summarized in Table 1; H_d ranged from 0.000 to 0.862, and π from 0.000 to 0.294. At the species level, H_T was estimated as 0.809 (± 0.034), H_S as 0.430 (± 0.046) and π as 0.00117. At the group level, the EH group exhibited higher genetic diversity (0.824, 0.625 and 0.00106 for H_T , H_S and π , respectively) compared to the HM group (0.100, 0.095 and 0.00043; Table 2).

The AMOVA revealed that 61.8% of genetic variation was partitioned among groups, 8% among populations within groups, and 30.2% within populations (Table 3). The coefficient of differentiation measured over the 43 populations was for $G_{ST} = 0.49$ and $N_{ST} = 0.72$ (Table 2). The permutation test indicated that N_{ST} was significantly higher than G_{ST} ($P < 0.01$), pointing to a strong phylogeographical structure for the *T. wallichiana* cpDNA haplotypes. In the SAMOVA

analysis, with $K = 2$, two groups were well defined, corresponding to the EH and HM groups, although the Φ_{CT} value was not the highest. The Φ_{CT} values changed little with increasing number of groups (K), and was highest when $K = 4$, when populations JZ and MA, with a high proportion of private haplotypes, were each separated as independent groups from the EH group.

The mismatch distribution analyses gave unimodal graphs for both the EH and HM lineage. A sudden population expansion is inferred by a near-perfect fit of the observed mismatch distribution with the expected (Fig. S1a–d), indicated by nonsignificant SSD and HRI statistics (Table 4). Values for the Fu's F_S and Tajima's D tests were significantly different from zero for the HM lineage, but not for the EH lineage (Table 4). Based on the value of τ , the demographical and spatial expansion times were calculated as 60.4 (32.2–92.3) thousand years ago (ka) and 56.2 (23.9–85.4) ka, respectively, for the EH lineage respectively, and 85.8 (19.9–91.5) ka and 6.8 (0–20.7) ka, respectively, for the HM lineage (Table 4).

trnL-trnF mutation rate and lineage divergence time

The BEAST-derived cpDNA (*rbcl* and *matK*) tree of dataset I is shown in Fig. S2. The crown age of Taxaceae is *c.* 210.5 Ma (95% highest posterior density, HPD: 173.6–247.4 Ma) (Table S7). The mean ages of other main genera for Taxaceae were dated as 23.9–30.6 Ma (Table S7). Based on these node ages, the BEAST analysis provided an average substitution rate for *trnL-trnF* of 8.08×10^{-10} substitutions/site/year (s/s/y) from dataset II. Using this rate and dataset III, the most common ancestor of the EH and HM lineages was estimated to date back to *c.* 4.2 (95% HPD: 2.0–6.5) Ma (Fig. 2a), suggesting a Pliocene split between the two lineages. The crown ages of the EH and HM lineages were estimated with *c.* 2.4 (95% HPD: 1.2–3.8) Ma and 2.1 (95% HPD: 0.7–3.6) Ma, respectively, indicating a Pleistocene diversification of the haplotypes within each lineage.

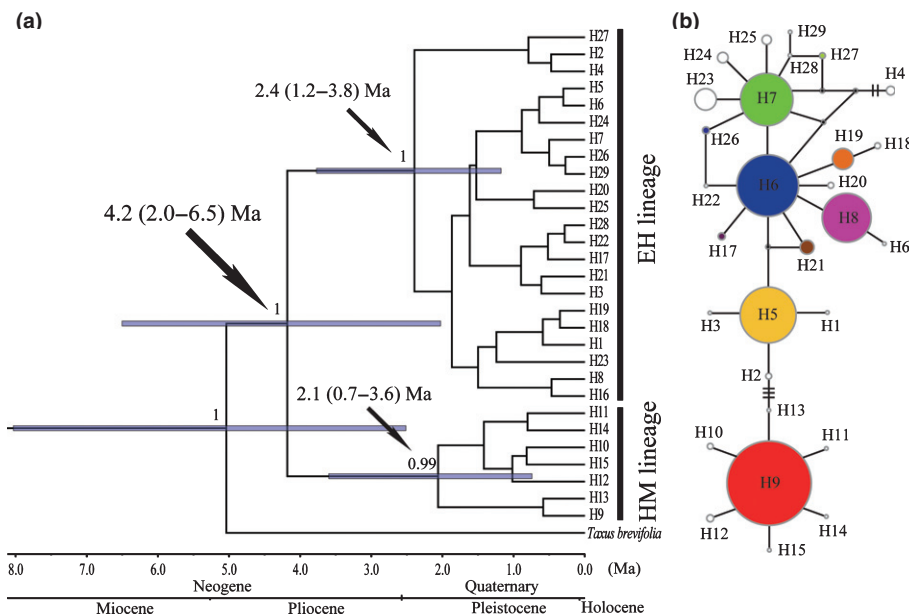


Fig. 2 Bayesian tree (a) with *Taxus brevifolia* as outgroup and MP median-joining network (b) based on the 29 cpDNA haplotypes. Numbers above branches in the Bayesian tree are posterior probability values. The size of circles in the network correspond to the frequency of each haplotype. Slashes across network branches indicate the number of mutational steps. The remaining branches represent single mutational steps. Colored haplotypes are shared by two or more populations. HM lineage, South Hengduan Mountains region; EH lineage, East Himalaya to the Yunnan Plateau region; Ma, million years ago.

Table 2 Genetic diversity and genetic differentiation of 43 populations of *Taxus wallichiana* at the species and group levels

Groups	cpDNA <i>trnL-trnF</i>					Nuclear microsatellite				
	H_S	H_T	$\pi (\times 10^{-2})$	G_{ST}	N_{ST}	H_S	H_T	G_{ST}	G'_{ST}	F_{ST}
Total	0.430	0.809	0.117	0.469	0.768*	0.47	0.60	0.21	0.42	0.32
EH	0.625	0.824	0.106	0.242	0.275 ^{ns}	0.49	0.55	0.11	0.24	0.11
HM	0.095	0.100	0.043	0.043	0.061 ^{ns}	0.44	0.48	0.07	0.16	0.08

H_S , mean genetic diversity within populations; H_T , total genetic diversity; π , nucleotide diversity; F_{ST} , among population differentiation; G_{ST} , genetic differentiation use only of the allelic frequencies; N_{ST} , genetic differentiation considered similarities between the haplotypes; G'_{ST} , standardized measure of genetic differentiation; ns, not significant. HM group, South Hengduan Mountains region; EH group, East Himalaya to the Yunnan Plateau region.

* N_{ST} is significantly different from G_{ST} ($P < 0.01$).

Table 3 Results of the analysis of molecular variance (AMOVA) for chloroplast DNA data and nuclear microsatellite data of 43 *Taxus wallichiana* populations

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	Fixation index (F_{ST})
cpDNA <i>trnL-trnF</i>					
Among groups	1	3452.29	9.24	61.8	
Among populations within groups	41	1116.89	1.20	8.0	
Within populations	772	3480.06	4.51	30.2	
Total	814	8049.24	14.95		0.70*
Nuclear microsatellite					
Among groups	1	561.06	0.75	24.0	
Among populations within groups	41	457.17	0.24	7.8	
Within populations	1573	3333.09	2.12	68.3	
Total	1615	4351.32	3.11		0.32*

Significance: *, $P < 0.001$.

d.f., degree of freedom.

Table 4 Results of the mismatch distribution analysis and neutrality tests of the two *Taxus wallichiana* lineages

Populations	Expansion types	Tau	Time (t) since expansion began, in ka	SSD	P -value	HRI	P -value	D	P -value	F_S	P -value
EH lineage	Demographic expansion	2.100	60.4 (32.2–92.3)	0.00551	0.310	0.038	0.580	−0.868	0.203	−6.782	0.050
	Spatial expansion	2.026	56.2 (23.9–85.4)	0.00468	0.350	0.038	0.580				
HM lineage	Demographic expansion	3.000	85.8 (19.9–91.5)	0.00002	0.250	0.768	0.770	−1.817	0.002	−13.018	0.000
	Spatial expansion	0.200	6.8 (0–20.7)	0.00003	0.130	0.768	0.870				

EH lineage, East Himalaya to the Yunnan Plateau region; HM lineage, South Hengduan Mountains region; SSD, sum of squared deviation under expansion model; HRI, Harpending's raggedness index; D , Tajima's D test statistic; F_S , Fu's F_S test statistic.

Microsatellite data

For the nuclear microsatellite markers, no evidence of scoring errors due to large allele dropout or stuttering at any locus was indicated. For the nine microsatellite loci used here, only 4.4% population-locus combinations deviated significantly from HWE after sequential Bonferroni correction (corrected $\alpha = 0.00013$, $P < 0.01$). There was no evidence for linkage disequilibrium.

Genetic diversity indices are summarized for each locus (Table 5) and population (Table 1), respectively. The number of alleles per locus ranged from 6 to 15. Mean H_S and H_T among loci were 0.47 and 0.60, respectively. At the population level, N_A and A_R ranged from 20 to 42 and 2.05 to 3.79, H_O and H_E ranged from 0.21 to 0.53 and 0.29 to 0.56, respectively (Table 1).

Population differentiation was significant for all loci ($P < 0.001$; Table 5), with the average F_{ST} for multilocus estimates being 0.22. The standardized genetic differentiation, G'_{ST} was higher than F_{ST} across all loci ($G'_{ST} = 0.42$ Table 5). The AMOVA results indicated significant genetic differentiation ($F_{ST} = 0.32$, $P < 0.001$), with 24% of the variation partitioned among groups, only 7.8% of the variation among populations within groups, and the rest of 68.3% variation within populations (Table 3). The F_{ST} was also significant among the populations within the two groups. The population genetic differentiation within the EH group was $F_{ST} = 0.11$ ($P < 0.001$), and 0.08 ($P < 0.001$) among populations of the HM group (Table 2).

The STRUCTURE analysis suggested Log-likelihood values of the data increased considerably when raising $K = 1$ –2 (Fig. 3a), after

Table 5 Summary of the nine microsatellite loci used to study the population genetics among 43 populations of *Taxus wallichiana*

Locus	Size range (bp)	N_A	H_O	H_E	H_S	H_T	F_{ST}	G'_{ST}
TA116	191–218	9	0.30	0.35	0.36	0.60	0.42	0.64
Tax86	142–176	15	0.26	0.30	0.31	0.34	0.11	0.12
TC04	132–148	9	0.48	0.48	0.49	0.69	0.31	0.57
TG06	176–202	12	0.31	0.52	0.54	0.63	0.17	0.32
TG34	254–266	7	0.44	0.48	0.49	0.54	0.10	0.16
TG47	333–393	12	0.44	0.57	0.59	0.75	0.25	0.55
TG111	99–115	9	0.53	0.56	0.58	0.81	0.30	0.68
TS03	259–293	6	0.15	0.16	0.17	0.19	0.13	0.13
TW01	376–406	15	0.60	0.71	0.74	0.87	0.18	0.60
Mean	—	10.44	0.39	0.46	0.47	0.60	0.22	0.42

N_A , number of alleles; H_O , observed heterozygosity; H_E , expected heterozygosity; H_S , mean genetic diversity within populations; H_T , total genetic diversity; F_{ST} , among population differentiation; G'_{ST} , standardized measure of genetic differentiation.

which the $\ln Pr(X/K)$ did not increase significantly. Based on the second-order rate of change in values of K (ΔK), the most likely number of genetic clusters for the complete dataset was also estimated at 2 (Fig. 3b). These confirmed the existence of two distinct genetic units corresponding to the EH and HM lineages, with the exception of only a few individuals (Fig. 3c). Four populations (CX, WB, EY and DY) located on the geographical boundary between the two groups included individuals assigned to both lineages. The genetic discontinuities between the EH and HM groups were also identified by BARRIER (Fig. 4). Strong genetic barriers between

the groups were detected that were supported by high bootstrap values of 92–99.8%. However, no clear genetic barriers were detected among populations within the two groups (Fig. 4).

Climatic data

Out of the 19 bioclimatic variables, all but four (bio2: mean diurnal temperature range, bio3: isothermality, bio7: temperature annual range, bio15: precipitation seasonality), showed significant differences in mean values between populations of the HM and EH group (Table 6 and Table S8). The variables bio1 (annual mean temperature: HM = 9.2°C vs EH = 11.5°C), bio9 (mean temperature of driest quarter: HM = 3.1°C vs EH = 5.8°C), bio11 (mean temperature of coldest quarter: HM = 2.8°C vs EH = 5.4°C), and bio12 (annual precipitation: HM = 960 mm vs EH = 1271 mm) exhibited the highest significance levels of differences of the means between the two groups ($P < 0.001$), followed by bio6 (min temperature of coldest month) and four rainfall variables, bio13 (precipitation of wettest month), bio 16 (precipitation of wettest quarter), bio18 (precipitation of warmest quarter), and bio 19 (precipitation of coldest quarter) ($P = 0.001$).

Pairing populations of the two regions along the entire gene flow barrier showed statistically highly significant differences for 17 of the 19 BioClim variables (Table 6). Splitting the pairs into westerly and southerly barriers showed 16 significant variables for the first, and five variables for the latter which were all temperature variables, with bio17 (precipitation of driest quarter) and bio19 (precipitation of coldest quarter) having highly significant

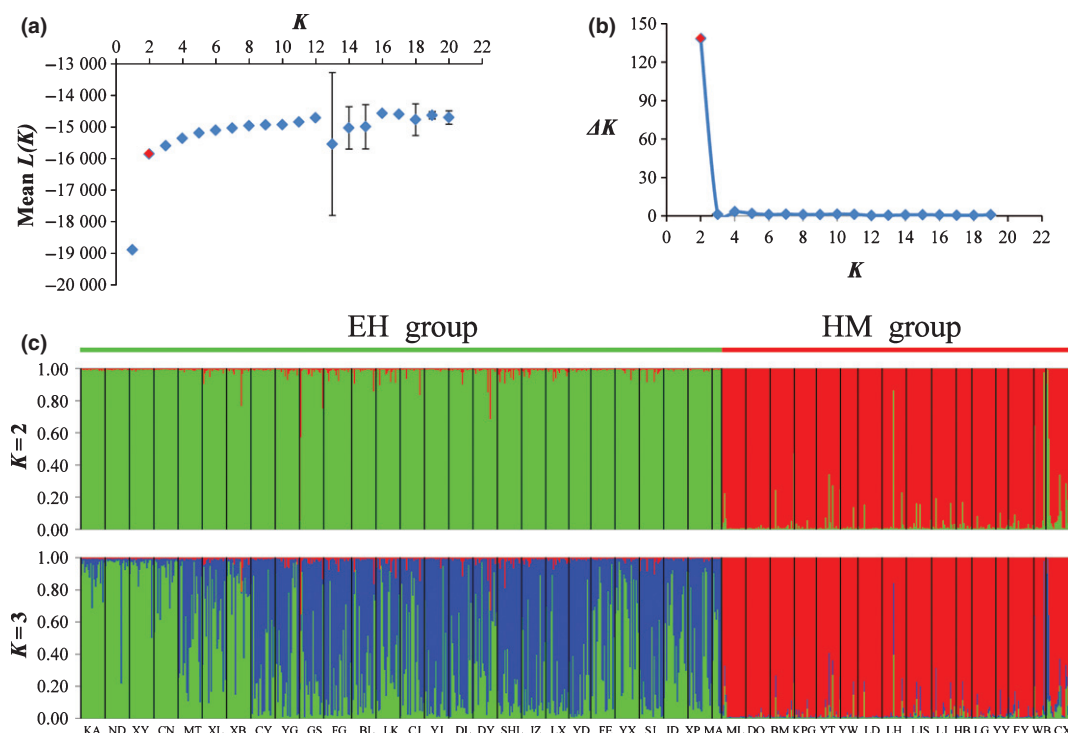


Fig. 3 Bayesian inference analysis of nuclear microsatellite data for the determination of the number of clusters (K) for the *Taxus wallichiana* populations analyzed. (a) The mean posterior probability of the data for $K = 1$ – 20 (20 replicates), the standard deviation of each mean $L(K)$ value was given respectively, the red rhombus indicated $K = 2$; (b) distribution of ΔK , the red rhombus indicated $K = 2$; (c) results for two and three clusters as detected by STRUCTURE. HM group, South Hengduan Mountains region; EH group, East Himalaya to the Yunnan Plateau region.

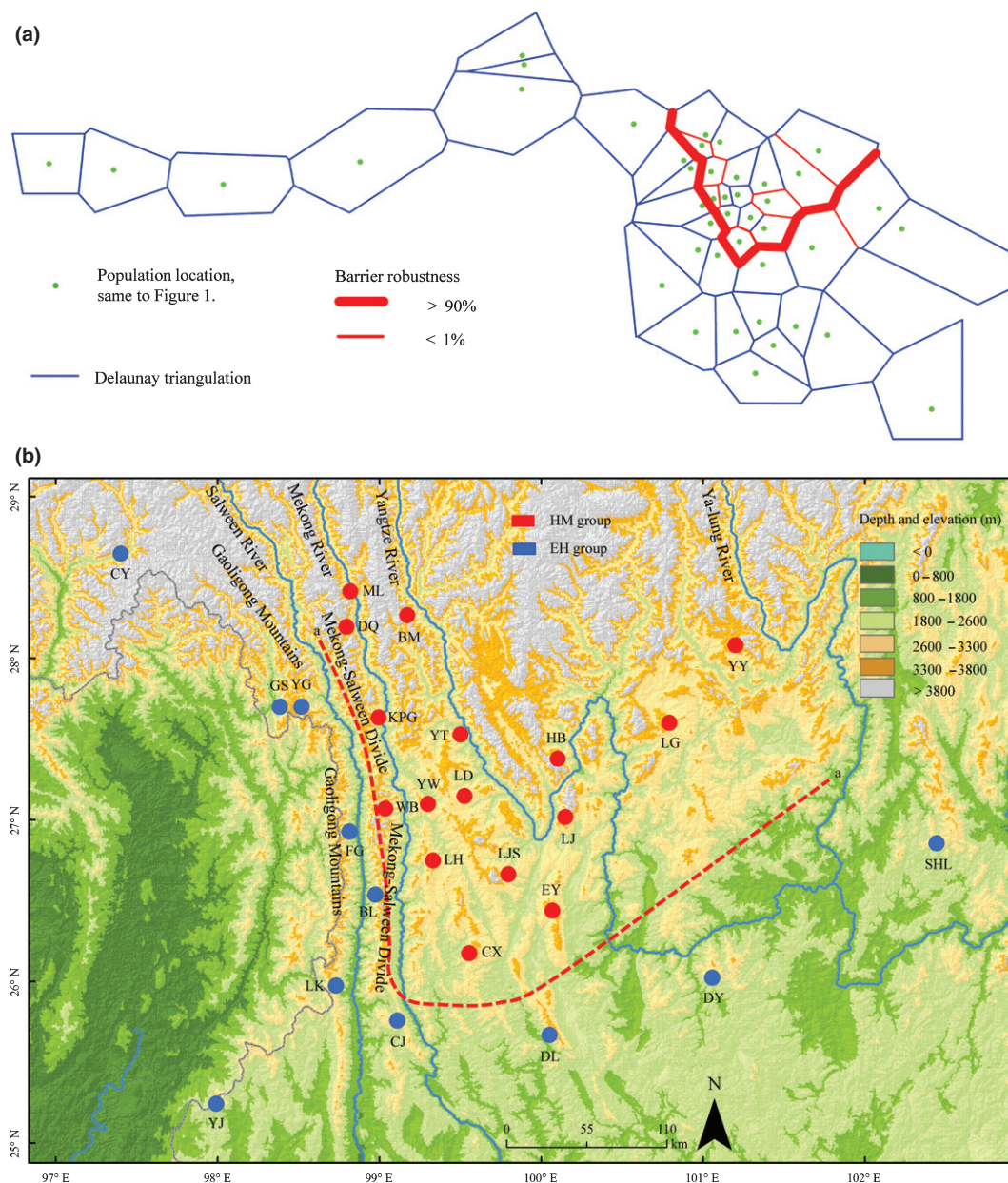


Fig. 4 Results of the BARRIER analysis based on microsatellite data, showing the spatial separation of *Taxus wallichiana* populations. (a) Delaunay triangulation and detected barrier (thick red line) separating the EH and HM populations, with bootstrap values over 1000 replicates using Nei's genetic distances (1983). (b) Geographic location of the genetic barrier, indicated by a dashed red line, as suggested in the BARRIER analysis. Circles indicate the location of the East Himalaya to the Yunnan Plateau (EH) and South Hengduan Mountains (HM) populations, corresponding to the two genetic clusters identified by the NETWORK and STRUCTURE analyses.

differences. There were a few differences in the results when aligning the pairs by altitude or by shortest distance, with the former giving higher significant differences for some variables (Table 6).

Species distribution modeling

For both the EH and HM lineages under all climate scenarios, the area under the receiver operating characteristic curve (AUC) values were ≥ 0.999 , indicating a far better than random prediction. Current distribution predictions were generally good representations of the actual distributions of both lineages, the only exception being the predicted occurrence of the HM lineage in the East Himalaya

and West Sichuan Basin, where this lineage does not occur at present. Palaeodistribution modeling for both lineages indicated a much more restricted distribution during the LIG than at present, with subsequent expansion at the LGM to cover a slightly greater area than that predicted under current climatic conditions (Fig. 5).

Discussion

Cryptic speciation in *Taxus wallichiana*

The strong population genetic structure identified in *T. wallichiana*, for both the cpDNA haplotypes and microsatellite

Table 6 Means, standard errors (SE) and results of *t*-tests on 19 BioClim variables for 43 populations of *Taxus wallichiana*

	Average altitude (m)	Bio1 (°C)	Bio2 (°C)	Bio3 (°C)	Bio4 (SD*100)	Bio5 (°C)	Bio6 (°C)	Bio7 (°C)	Bio8 (°C)	Bio9 (°C)	Bio10 (°C)	Bio11 (°C)	Bio12 (mm)	Bio13 (mm)	Bio14 (mm)	Bio15 (CV)	Bio16 (mm)	Bio17 (mm)	Bio18 (mm)	Bio19 (mm)	
(A) All populations (two-tailed <i>t</i> -test)																					
HM	Mean	2987	9.2	11.3	46.7	495	19.3	-4.9	24.1	14.9	3.1	14.9	2.8	960	222	6.3	96.3	574	28.1	574	29.3
	SE	38	0.4	0.2	0.4	7	0.4	0.5	0.3	0.3	0.4	0.3	0.4	27	6	0.5	2	17.9	2.3	17.9	2.6
EH	Mean	2661	11.5	11.1	47.4	463	21	-2.4	23.4	16.7	5.8	16.7	5.4	1271	281	9.1	94	746.6	40.4	746.6	44.8
	SE	34	0.4	0.2	0.4	12	0.4	0.5	0.5	0.4	0.4	0.4	0.5	54	15	0.8	1.9	36.1	3	36.1	3.3
<i>P</i>	>0.001	>0.001	ns	ns	0.05	0.01	0.001	0.001	ns	>0.001	>0.001	>0.001	>0.001	0.001	0.01	0.01	ns	0.001	0.01	0.001	0.001
(B) Populations paired by altitude (paired two-tailed <i>t</i> -test)																					
HM	Mean	2925	9.5	11.3	46.8	496	19.5	-4.7	24.2	15.1	3.6	15.1	3.1	990	225	6.7	95.2	586.9	29.4	586.9	31.3
	SE	51	0.6	0.3	0.7	10	0.6	0.8	0.5	0.6	0.7	0.6	0.6	39	8	0.8	3	24.2	3.7	24.2	4.2
EH	Mean	2690	11.8	11	47.3	469	21.1	-2.1	23.2	17.1	6	17.1	5.7	1228	266	10.1	91.6	709.6	44.9	709.6	49.1
	SE	86	0.6	0.1	0.6	8	0.6	0.6	0.2	0.6	0.7	0.6	0.5	51	7	0.8	2.7	20.3	4.8	20.3	5.9
<i>P</i>	0.05	0.01	ns	ns	0.001	0.05	0.01	0.01	0.05	0.01	0.01	0.01	0.01	0.01	0.05	0.001	0.05	0.01	0.001	0.01	0.001
(B1) Western barrier ML-CX (paired two-tailed <i>t</i> -test)																					
HM	Mean	2918	9.9	10.8	45.8	497	19.9	-3.7	23.7	15.6	4.3	15.6	3.6	1008	222	8	89.8	579.8	35.8	579.8	38.7
	SE	74	0.7	0.1	0.8	14	0.9	0.7	0.4	0.8	0.8	0.8	0.7	56	13	0.4	2	36.4	2.2	36.4	2.3
EH	Mean	2646	12.5	10.8	46.7	474	21.7	-1.4	23.1	17.8	6.8	17.8	6.3	1309	273	11.3	86.6	736.8	52.8	736.8	59.2
	SE	104	0.7	0.1	0.7	9	0.8	0.7	0.2	0.7	0.8	0.7	0.7	48	8	0.7	1.2	22.9	3.6	22.9	4.2
<i>P</i>	ns	0.05	ns	0.01	0.01	ns	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.01	0.05	0.01	ns	0.05	0.01	0.05	0.01
(B2) Southern barrier EY-YY (paired two-tailed <i>t</i> -test)																					
HM	Mean	2938	8.6	12.4	48.7	494	18.8	-6.6	25.4	14.2	2.1	14.2	2.1	953	231	4	106.1	601	16.7	601	16.7
	SE	59	0.9	0.6	0.1	18	0.7	1.5	1.2	0.8	1	0.8	1	35	5	1	1.7	16.5	4.7	16.5	4.7
EH	Mean	2779	10.5	11.4	48.5	459	19.9	-3.5	23.5	15.6	4.5	15.6	4.5	1067	251	7.7	101.5	655	29	655	29
	SE	167	0.2	0.2	0.7	16	0.5	0.4	0.7	0.3	0.3	0.3	0.3	14	11	1.2	2.7	8.9	5	8.9	5
<i>P</i>	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.05	ns	0.01	0.05	ns	0.001	ns	0.001
(C) Populations paired by shortest distance (paired two-tailed <i>t</i> -test)																					
HM	Mean	2911	9.4	11.5	47.7	484	19.2	-4.9	24.1	14.9	3.4	14.9	3.1	1033	237	6.4	98.1	621.3	29.3	621.3	30.6
	SE	64	0.8	0.4	0.5	9	0.8	1	0.7	0.8	0.9	0.8	0.8	34	3	1	3.1	10.5	4.9	10.5	5.4
EH	Mean	2755	12.6	11.2	48.3	456	21.8	-1.5	23.2	17.6	6.7	17.6	6.5	1184	258	10.4	93.2	689.9	43.3	689.9	45.6
	SE	66	1.3	0.2	0.5	9	1.3	1.2	0.3	1.2	1.4	1.2	1.3	58	8	1.3	3.3	22.8	5.6	22.8	7.1
<i>P</i>	ns	0.01	ns	ns	0.05	0.05	0.05	0.01	ns	0.01	0.01	0.01	0.01	ns	0.05	0.001	0.01	0.05	0.001	0.05	0.01
(C1) Western barrier KPG-CX (paired two-tailed <i>t</i> -test)																					
HM	Mean	2891	10	10.9	46.9	477	19.5	-3.7	23.2	15.4	4.4	15.4	3.8	1093	242	8.3	92.1	636.5	38.8	636.5	41
	SE	110	1.2	0.1	0.6	10	1.3	1.1	0.2	1.2	1.3	1.2	1	23	2	0.5	2.1	8.4	1.7	8.4	2.7
EH	Mean	2736	14.1	11.1	48.1	453	23.1	0.1	23.1	19	8.3	19	8.1	1272	264	12.5	86.9	716	54	716	58
	SE	30	2	0.2	0.8	12	2.1	1.9	0.3	1.8	2.1	1.8	2	76	11	1.3	2	35.3	3.1	35.3	6.7
<i>P</i>	ns	0.05	ns	ns	ns	ns	0.05	0.05	ns	0.05	0.05	0.05	0.05	ns	ns	0.05	ns	ns	0.01	ns	ns
(C2) Southern barrier EY-YY (paired two-tailed <i>t</i> -test)																					
HM	Mean	2938	8.6	12.4	48.7	494	18.8	-6.6	25.4	14.2	2.1	14.2	2.1	953	231	4	106.1	601	16.7	601	16.7
	SE	59	0.9	0.6	0.1	18	0.7	1.5	1.2	0.8	1	0.8	1	35	5	1	1.7	16.5	4.7	16.5	4.7
EH	Mean	2779	10.5	11.4	48.5	459	19.9	-3.5	23.5	15.6	4.5	15.6	4.5	1067	251	7.7	101.5	655	29	655	29
	SE	167	0.2	0.2	0.7	16	0.5	0.4	0.7	0.3	0.3	0.3	0.3	14	11	1.2	2.7	8.9	5	8.9	5
<i>P</i>	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.05	ns	0.01	0.05	ns	0.001	ns	0.001

Entire populations sample (A); populations along the barrier paired by altitude (B); shortest distance (C) (see Table S4 for pairs).

ns, not significant; EH, East Himalaya to the Yunnan Plateau region; HM, South Hengduan Mountains region.

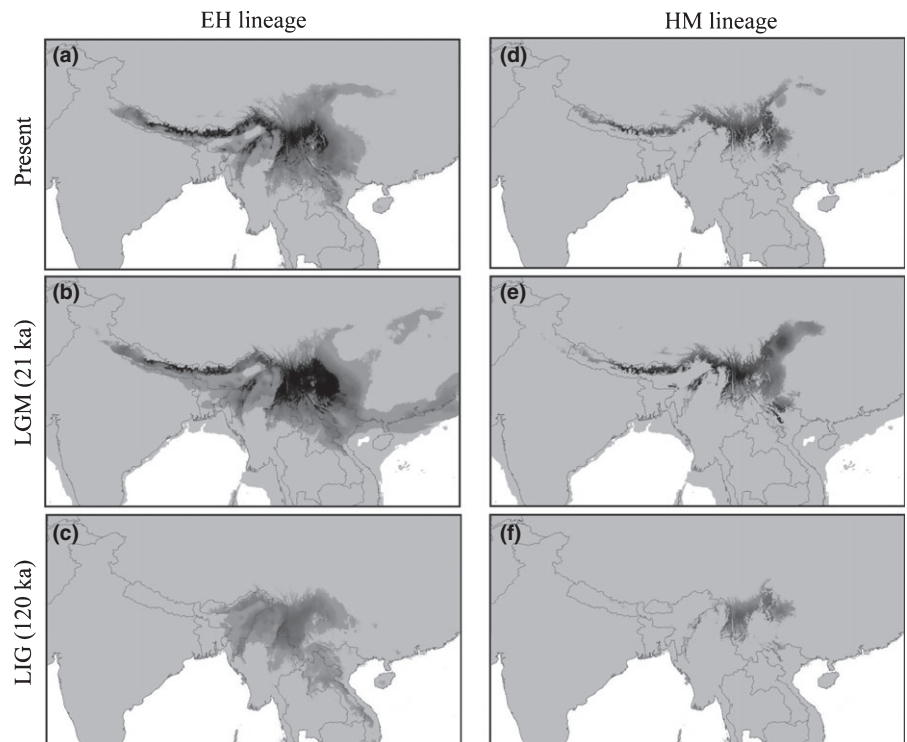


Fig. 5 Potential distribution range of *Taxus wallichiana* in HHM region simulated by Ecological Niche Models using bioclimatic variables. Darker colors show areas with more suitable predicted conditions. LGM, Last Glacial Maximum; LIG, Last Interglacial; EH lineage, East Himalaya to the Yunnan Plateau region; HM lineage, South Hengduan Mountains region.

data indicated that the EH and HM groups represent two distinct lineages (Table 2). The estimated differentiation values (G_{ST}) were larger than the mean genetic differentiation calculated for paternally inherited plastid markers in 37 conifer species (0.469 vs 0.165; Petit *et al.*, 2005) and for biparentally inherited microsatellite loci (0.32 vs 0.24/0.26 for F_{ST}/R_{ST} ; Nybom, 2004), indicating the existence of strong barriers to gene flow between the EH and HM groups. Geographic isolation of populations within species and variation in ecological factors are major precursors to cryptic speciation (Hoskin *et al.*, 2005). The Nushan Mountains, which form a large northern part of the Mekong–Salween divide, appear to act as a topological N–S barrier for *Taxus* to the West of the HM group (Fig. 4), as suggested by previous studies (Gao *et al.*, 2007; Li *et al.*, 2011). It is interesting to note that even though *Taxus* is wind pollinated, pollen does not seem to be able to cross the 4000 m range of the Nushan Mountains. This may be explained by the fact that the plants occur as understory trees, where wind movement is limited, and the dispersal distance of *Taxus* pollen has been shown to be limited to a few meters (Wheeler *et al.*, 1995).

Although the timing of the uplift process that formed the QTP is of longstanding debate, studies suggest that the plateau rose rapidly *c.* 10–8 Ma (Harrison *et al.*, 1992; Royden *et al.*, 2008) or even more recently (since 3.6 Ma; Li & Fang, 1999). Our divergence time estimations suggest that the two lineages diverged in the Pliocene (4.2 (2.0–6.5) Ma), a time that corresponds well with later estimates for the uplift of the QTP and the formation of the Hengduan Mountains (Tapponnier *et al.*, 1990; Shi *et al.*, 1998; Akciz *et al.*, 2008). It is likely that the Nushan Mountains had almost attained their current height by that time,

and that their rapid uplift effectively divided the populations of *T. wallichiana*.

Both lineages mostly comprised reciprocally exclusive haplotypes, along the boundary between the two groups, but five cpDNA haplotypes from the EH region were found in three populations of the HM lineage (EY, WB, CX), and one from the HM region in population DY of the EH lineage (Fig. 1; Table 1). Based on the microsatellite data, only a few of these individuals showed signs of introgression. For instance, the two individuals in population WB possessed EH chloroplast haplotypes and also a nuclear genotype belonging to the EH group (Fig. 3c). Population WB is the most northern mixed population in the HM group that exhibits such admixture, and the presence of migrants here is probably due to rare seed dispersal events across the Nushan Mountains from the West, or a northward migration along the Mekong River from the South. The remaining three populations that exhibit admixture are located at the southern edge of the HM group where no topological barriers to gene flow are existent, and the two lineages remain separated perhaps through niche specialization (see following paragraph). The observed haplotype mixtures indicate that the barriers are 'leaky', and demonstrate the occurrence of recent secondary contacts between the two lineages.

Ecological factors reinforce divergence in *Taxus wallichiana*

The uplift of the QTP and its associated climatic changes has been speculated as a main cause of plant diversification in the QTP area (Cun & Wang, 2010; Xu *et al.*, 2010; Yang *et al.*, 2012). The current distribution of *T. wallichiana* is primarily

determined by the Indian monsoon in summer and Westerlies in winter, with limited influence from the East Asian monsoon (Fig. 1). Because the monsoons were blocked gradually by the N–S mountain ranges in the South Hengduan Mountains region, the climate for the HM lineage populations can be characterized as cold and dry, whereas the distribution range of the EH lineage has a humid and warm climate as a result of the uplift of the QTP. The intensification of the Indian monsoons during the Pliocene (An *et al.*, 2001; Wan *et al.*, 2007; Zhang *et al.*, 2009; Chang *et al.*, 2010), would have increased levels of precipitation in the HHM region. However, the synchronous rise of the Hengduan Mountains, particularly Gaoligong Mountains and Nushan Mountains in this period, obstructed the moist air eastward to the HM lineage distribution range, resulting in differentiation between the EH and HM lineages due to environmental factors, as suggested by paleoclimatic evidence (Kou *et al.*, 2006; Yao *et al.*, 2012a). Such differences were further strengthened by the accelerated uplift phases of QTP in the late Pliocene. Consequently, following geographical isolation by the Nushan Mountains, the two *T. wallichiana* lineages would have been exposed and eventually adapted to their differing environmental conditions. The HM populations occur at higher altitudes (2639–3230 m, average 2987 m) compared to those of the EH lineage (2213–3037 m, average 2661 m), and this is reflected in climatic differences. While both groups experience a dry winter and wet summer, the EH populations are generally associated with a warmer and wetter climate due to their occurrence at lower altitudes and the effects of the Indian monsoon, compared to the colder and drier climate of the HM lineage (i.e. higher altitude) (Table 6). Although no topological barriers are present between the two lineages at the southern edge of the Hengduan Mountains region, significant differences in climatic variables can be detected, particularly those involving winter rainfall, creating an ecological barrier (Table 6).

The ecological differences found between the two *T. wallichiana* lineages studied here seem to represent species-specific characteristics and are apparently sufficient to keep both lineages separated to a high degree, as indicated by the chloroplast and nuclear data. Such ecological niche partitioning will have reinforced the divergence of the two lineages following initial spatial isolation, and may have given rise to some degree of differential adaptation to their respective environmental conditions. This, given sufficient time, might ultimately lead to reproductive isolation (Rieseberg & Burke, 2001; Nosil *et al.*, 2009; Thorpe *et al.*, 2010; Wagner *et al.*, 2012), resulting in the formation of new species.

Unusual demographic history of *Taxus wallichiana*

Climate changes during Pleistocene glacial–interglacial cycles had a dramatic effect on species distribution ranges (Comes & Kadereit, 1998; Hewitt, 2004), causing migration and/or extinction of populations, followed by periods of isolation, divergence and subsequent expansion (Taberlet *et al.*, 1998; Cun & Wang, 2010). Most plant taxa are believed to have shifted latitude or elevation ranges in response to glaciations (Davis & Shaw, 2001). The divergence time estimated for the EH and HM lineages here

suggests that they became separated well before the Quaternary, in the Pliocene. Therefore, climatic oscillations during the Pleistocene will have had pronounced effects on their more recent population history. In contrast with most previously suggested glacial evolutionary scenarios in the QTP and adjacent regions, such as peri-glacial refugia, refugia-within-refugia, and local persistence (e.g. Zhang *et al.*, 2005; Yang *et al.*, 2008; Cun & Wang, 2010; Opgenoorth *et al.*, 2010; Wang *et al.*, 2010), our analyses indicate expansion of *T. wallichiana* populations throughout the most recent glacial period, rather than contraction. For both lineages, the star-like pattern exhibited by the minimum spanning network (Fig. 2) is characteristic of recent population expansion, and the observed pairwise differences of *cpDNA* haplotypes fit a sudden expansion model particularly well (Fig. S1) and produce significantly negative Tajima's *D* and Fu's *F_s* statistics (Table 4). The mismatch distribution analysis suggested different phases of demographic and spatial expansion for the EH and HM lineages (Table 4). Expansion of *T. wallichiana* during the last glaciations was further supported by the SDM analysis, which indicated larger distribution ranges of both lineages at the LGM compared to both the LIG and present (Fig. 5).

This relatively unusual demographic scenario can be explained partly by the topoclimatic characteristics of the HHM region. It is located at the southern and southeastern fringe of the QTP, and is exposed to the moist and warm air of the Indian monsoon from the South (Figs 1, 5). Because the QTP prevents the southward flow of cold and dry continental air toward the HHM region, the region is characterized by a warmer and wetter climate than the majority of the QTP, which would provide a more stable environment for plant populations, even during glaciations. In addition, the HHM region shows great latitudinal range and elevation differences, which could possibly provide microclimatic refugia for *T. wallichiana* populations during Quaternary climatic oscillations. Such areas would not be as prevalent in the Himalayan regions as in the South Hengduan Mountains region due to the topographic character (Fig. 1). Thus, *T. wallichiana* populations could have migrated downward and southward to track their optimal ecological conditions during the glacial periods and expanded into larger areas of suitable habitats. By contrast, during the LIG, the temperature was at least 5°C higher than at present (Shi *et al.*, 1998). Given that the current temperature in the HHM region decreases by 0.64°C/100 m (Li & Zhang, 2010), and assuming niche conservatism over time, *T. wallichiana* populations would have had to move uphill at least 800 m from their present distributional altitude due to little space for latitudinal shift. The environment in the Himalaya would have been very harsh above 3500 m, and there would be limited suitable habitats for *T. wallichiana* populations, which would ultimately result in population contractions and extinction in this region. Conversely, the topography of the South Hengduan Mountains shows greater potential for latitudinal shift than in the Himalaya region, which could facilitate persistence of *T. wallichiana* populations in this region during the LIG.

The possibility of *T. wallichiana* population expansion during the last glaciation is also supported by paleoclimatic and

paleoenvironmental analysis from this region. There have been four major glaciations in the QTP during the Quaternary, becoming progressively less extensive after the largest Naynayxungla Glaciation (*c.* 0.72–0.5 Ma), and LGM glacier advances were significantly less extensive in this area (Shi *et al.*, 1998; Zheng *et al.*, 2002; Owen, 2009). It is thus likely that the last glaciation only had a limited impact on *T. wallichiana*. Furthermore, paleoclimatic analyses also suggested wetter winters and lower annual mean temperatures than today during the last 40–14.2 ka in the HHM region (Walker, 1986; Shi *et al.*, 2001; Xiao *et al.*, 2009), when *Taxus* still persisted in the South Hengduan Mountains (Jiang *et al.*, 1998). The expansion scenario is further supported by vegetation reconstructions based on pollen fossil data, which indicated that mixed and evergreen subtropical forests were present in SE Tibet during the LGM (e.g. Yadong, XY; Motuo, MT; and Chayu, CY; Fig. 1) (Shi *et al.*, 1998), where populations of *T. wallichiana* are found today. Biome reconstructions on the basis of recent climatology and pollen data also indicate a warm-temperate climate during the LGM for the present-day distribution area of *T. wallichiana* (Harrison *et al.*, 2001).

A scenario of population expansion, or at least population stability, during the last glacial period, as opposed to the general ‘contraction–expansion’ scenario usually described has occasionally been reported in other plants (e.g. Yan *et al.*, 2012) and animals (e.g. Li *et al.*, 2009) in subtropical areas of China. There is also growing evidence for population demographic stability or expansion throughout the LGM elsewhere in a range of different organisms (e.g. King *et al.*, 2009; Marko *et al.*, 2010; Bisconti *et al.* 2011; Cunha *et al.* 2011; Pinheiro *et al.*, 2011; Batalha-Filho *et al.*, 2012).

Conclusions

Our analyses of *Taxus wallichiana*, combining molecular phylogeography and species distribution modeling, strongly suggest that the observed patterns of genetic variation and divergence in *T. wallichiana* are best explained by a combination of both Miocene/Pliocene geological and climatic events, and late Quaternary climatic oscillations. Initial topological constraints, reinforced by subsequent differential ecological (climatic) adaptations have resulted in cryptic speciation, and the formation of two discrete taxonomic entities, a split indicated by both molecular (Gao *et al.*, 2007; Liu *et al.*, 2011b) and morphological analyses (Möller *et al.*, 2007). Furthermore, *T. wallichiana* populations exhibit an unusual demographic history, having expanded their range during the last glacial period. These findings indicate that a combination of orographic and climatic factors has played a fundamental role in promoting diversification and evolution of species in the HHM region, and that these processes may more complex than previously thought.

Acknowledgements

The authors are extremely grateful to a large number of individuals, too numerous to list, who supported our extensive fieldwork and laboratory work. We acknowledge Pete Hollingsworth,

Andrew Lowe, Richard Milne and Jian-Quan Liu for their constructive comments on earlier versions of the manuscript. The University of Oslo Biportal provided computation and software resources. This work was supported by the Key Research Program of the Chinese Academy of Sciences (KSZD-EW-Z-011); the National Natural Science Foundation of China (30700042, 31200182); the Ministry of Science and Technology of China (2012FY110800); and the talent project of Yunnan Province, China (2008YP064). The Royal Botanic Garden Edinburgh is funded by the Rural and Environment Science and Analytical Services division (RESAS) in the Scottish Government.

References

- Akciz S, Burchfiel BC, Crowley JL, Yin J, Chen L. 2008. Geometry, kinematics, and regional significance of the Chong Shan shear zone, Eastern Himalayan Syntaxis, Yunnan, China. *Geosphere* 4: 292–314.
- An ZS, Kutzbach JE, Prell WL, Porter SC. 2001. Evolution of Asian monsoons and phased uplift of the Himalayan–Tibetan plateau since Late Miocene times. *Nature* 411: 62–66.
- Batalha-Filho H, Cabanne GS, Miyaki CY. 2012. Phylogeography of an Atlantic forest passerine reveals demographic stability through the last glacial maximum. *Molecular Phylogenetics and Evolution* 65: 892–902.
- Bisconti R, Canestrelli D, Colangelo P, Nascetti G. 2011. Multiple lines of evidence for demographic and range expansion of a temperate species (*Hyla sarda*) during the last glaciation. *Molecular Ecology* 20: 5313–5327.
- Chang Z, Xiao J, Lü L, Yao H. 2010. Abrupt shifts in the Indian monsoon during the Pliocene marked by high-resolution terrestrial records from the Yuanmou Basin in southwest China. *Journal of Asian Earth Sciences* 37: 166–175.
- Comes HP, Kadereit JW. 1998. The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science* 3: 432–438.
- Crawford NG. 2010. SMOGD: software for the measurement of genetic diversity. *Molecular Ecology Resources* 10: 556–557.
- Cun YZ, Wang XQ. 2010. Plant recolonization in the Himalaya from the southeastern Qinghai–Tibetan Plateau: geographical isolation contributed to high population differentiation. *Molecular Phylogenetics and Evolution* 56: 972–982.
- Cunha RL, Lopes EP, Reis DM, Castilho R. 2011. Genetic structure of *Brachydontes puniceus* populations in Cape Verde archipelago shows signature of expansion during the last glacial maximum. *Journal of Molluscan Studies* 77: 175–181.
- Davis MB, Shaw RG. 2001. Range shifts and adaptive responses to Quaternary climate change. *Science* 292: 673–679.
- Dieringer D, Schlötterer C. 2003. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* 3: 167–169.
- Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W. 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* 161: 1307–1320.
- Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214.
- Dupanloup I, Schneider S, Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* 11: 2571–2581.
- 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 HEL. 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.
- Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.

- Farjon A. 2010. *A handbook of the World's conifers*. Leiden, the Netherlands: Brill Academic Publishers, 983–985.
- Florin R. 1958. On Jurassic taxads and conifers from north-western Europe and eastern Greenland. *Acta Horti Bergiani* 17: 257–410.
- Fu LG, Li N, Mill RR. 1999. Taxaceae. In: Wu ZY, Peter RH, eds. *Floral of China*. Beijing, China: Science Press and St. Louis, MO, USA: Missouri Botanical Garden Press, 89–96.
- Funk DJ, Nosil P, Etges WJ. 2006. Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proceedings of the National Academy of Sciences, USA* 103: 3209–3213.
- Gao LM, Möller M, Zhang XM, Hollingsworth ML, Liu J, Mill RR, Gibby M, Li DZ. 2007. High variation and strong phylogeographic pattern among cpDNA haplotypes in *Taxus wallichiana* (Taxaceae) in China and North Vietnam. *Molecular Ecology* 16: 4684–4698.
- Givnish TJ. 2010. Ecology of plant speciation. *Taxon* 59: 1326–1366.
- Goudet J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). [WWW document] URL <http://www2.unil.ch/popgen/softwares/fstat.htm>. [Accessed on 20 March 2013].
- Hao DC, Xiao PG, Huang BL, Ge GB, Yang L. 2008. Interspecific relationships and origins of Taxaceae and Cephalotaxaceae revealed by partitioned Bayesian analyses of chloroplast and nuclear DNA sequences. *Plant Systematics and Evolution* 276: 89–104.
- Harmon LJ, Melville J, Larson A, Losos JB. 2008. The role of geography and ecological opportunity in the diversification of day geckos (*Phelsuma*). *Systematic Biology* 57: 562–573.
- Harris TM. 1976. Two neglected aspects of fossil conifers. *American Journal of Botany* 63: 902–910.
- Harrison SP, Yu G, Takahara H, Prentice IC. 2001. Diversity of temperate plants in east Asia. *Nature* 413: 129–130.
- Harrison TM, Copeland P, Kidd WSF, Yin A. 1992. Raising Tibet. *Science* 255: 1663–1670.
- Hedrick PW. 2005. A standardized genetic differentiation measure. *Evolution* 59: 1633–1638.
- Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B* 359: 183–195.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.
- Hijmans RJ, Guarino L, Cruz M, Rojas E. 2001. Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter* 127: 15–19.
- Hoskin CJ, Higgie M, McDonald KR, Moritz C. 2005. Reinforcement drives rapid allopatric speciation. *Nature* 437: 1353–1356.
- Jia DR, Abbott RJ, Liu TL, Mao KS, Bartish IV, Liu JQ. 2012. Out of the Qinghai–Tibet Plateau: evidence for the origin and dispersal of Eurasian temperate plants from a phylogeographic study of *Hippophæ rhamnoides* (Elaeagnaceae). *New Phytologist* 194: 1123–1133.
- Jia DR, Liu TL, Wang LY, Zhou DW, Liu JQ. 2011. Evolutionary history of an alpine shrub *Hippophæ tibetana* (Elaeagnaceae): allopatric divergence and regional expansion. *Biological Journal of the Linnean Society* 102: 37–50.
- Jiang X, Wang S, Yang X. 1998. Paleoclimatic and environmental changes over the last 30 000 years in Heqing Basin, Yunnan Province. *Journal of Lake Sciences* 10: 1–7.
- Keller I, Seehausen O. 2012. Thermal adaptation and ecological speciation. *Molecular Ecology* 21: 782–799.
- King MG, Horning ME, Roalson EH. 2009. Range persistence during the last glacial maximum: *Carex macrocephala* was not restricted to glacial refugia. *Molecular Ecology* 18: 4256–4269.
- Kou XY, Ferguson DK, Xu JX, Wang YF, Li CS. 2006. The reconstruction of paleovegetation and paleoclimate in the Late Pliocene of West Yunnan, China. *Climatic Change* 77: 431–448.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al.* 2007. Clustal W and clustal X version 2.0. *Bioinformatics* 23: 2947–2948.
- Li JJ, Fang XM. 1999. Uplift of the Tibetan Plateau and environmental changes. *Chinese Science Bulletin* 44: 2117–2124.
- Li SH, Yeung CKL, Feinstein J, Han L, Le MH, Wang CX, Ding P. 2009. Sailing through the Late Pleistocene: unusual historical demography of an East Asian endemic, the Chinese Hwamei (*Leucodioptron canorum canorum*), during the last glacial period. *Molecular Ecology* 18: 622–633.
- Li Y, Zhai SN, Qiu YX, Guo YP, Ge XJ, Comes HP. 2011. Glacial survival east and west of the 'Mekong-Salween Divide' in the Himalaya-Hengduan Mountains region as revealed by AFLPs and cpDNA sequence variation in *Sinopodophyllum hexandrum* (Berberidaceae). *Molecular Phylogenetics and Evolution* 59: 412–424.
- Li WH, Zhang YG. 2010. *The vertical climate of Hengduan Mountains region and its impact on forest distribution*. Beijing, China: China Meteorological Press.
- Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
- Liu J, Gao LM. 2011. Comparative analysis of three different methods of total DNA extraction used in *Taxus*. *Guohua* 31: 244–249.
- Liu JQ, Gao TG, Chen ZD, Lu AM. 2002. Molecular phylogeny and biogeography of the Qinghai-Tibet Plateau endemic *Nannoglottis* (Asteraceae). *Molecular Phylogenetics and Evolution* 23: 307–325.
- Liu J, Gao LM, Li DZ, Zhang DQ, Möller M. 2011a. Cross-species amplification and development of new microsatellite loci for *Taxus wallichiana* (Taxaceae). *American Journal of Botany* 98: e70–e73.
- Liu J, Möller M, Gao LM, Zhang DQ, Li DZ. 2011b. DNA barcoding for the discrimination of Eurasian yews (*Taxus* L., Taxaceae) and the discovery of cryptic species. *Molecular Ecology Resources* 11: 89–100.
- Liu JQ, Wang YJ, Wang AL, Hideaki O, Abbott RJ. 2006. Radiation and diversification within the *Ligularia-Crematobodium-Parasenecio* complex (Asteraceae) triggered by uplift of the Qinghai-Tibetan Plateau. *Molecular Phylogenetics and Evolution* 38: 31–49.
- Lowry DB, Rockwood RC, Willis JH. 2008. Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*. *Evolution* 62: 2196–2214.
- Manni F, Guérard E, Heyer E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology* 76: 173–190.
- Mao KS, Milne RI, Zhang LB, Peng YL, Liu JQ, Thomas P, Mill RR, Renner SS. 2012. Distribution of living Cupressaceae reflects the breakup of Pangea. *Proceedings of the National Academy of Sciences, USA* 109: 7793–7798.
- Marko PB, Hoffman JM, Emme SA, McGovern TM, Keever CC, Nicole Cox L. 2010. The 'expansion-contraction' model of Pleistocene biogeography: rocky shores suffer a sea change? *Molecular Ecology* 19: 146–169.
- Möller M, Gao LM, Mill RR, Li DZ, Hollingsworth ML, Gibby M. 2007. Morphometric analysis of the *Taxus wallichiana* complex (Taxaceae) based on herbarium material. *Botanical Journal of the Linnean Society* 155: 307–335.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
- Nosil P, Harmon LJ, Seehausen O. 2009. Ecological explanations for (incomplete) speciation. *Trends in Ecology and Evolution* 24: 145–156.
- Nosil P, Vines TH, Funk DJ. 2005. Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* 59: 705–719.
- Nybom H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* 13: 1143–1155.
- Opgenoorth L, Vendramin GG, Mao KS, Mische G, Mische S, Liepelt S, Liu JQ, Ziegenhagen B. 2010. Tree endurance on the Tibetan Plateau marks the world's highest known tree line of the Last Glacial Maximum. *New Phytologist* 185: 332–342.
- Owen LA. 2009. Latest Pleistocene and Holocene glacier fluctuations in the Himalaya and Tibet. *Quaternary Science Reviews* 28: 2150–2164.
- Peakall R, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel: population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.
- Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG. 2005. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* 14: 689–701.
- Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231–259.
- Pinheiro F, de Barros F, Palma-Silva C, Fay MF, Lexer C, Cozzolino S. 2011. Phylogeography and genetic differentiation along the distributional range of

- the orchid *Epidendrum fulgens*: a Neotropical coastal species not restricted to glacial refugia. *Journal of Biogeography* 38: 1923–1935.
- Pons O, Petit RJ. 1996. Measuring and testing genetic differentiation with ordered vs unordered alleles. *Genetics* 144: 1237–1245.
- Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.
- Poudel RC, Möller M, Gao LM, Ahrends A, Baral SR, Liu J, Thomas P, Li DZ. 2012. Using morphological, molecular and climatic data to delimitate yews along the Hindu Kush-Himalaya and adjacent regions. *PLoS ONE* 7: e46873.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- Rice WR. 1989. Analyzing tables of statistical tests. *Evolution* 43: 223–225.
- Rieseberg LH, Burke JM. 2001. The biological reality of species: gene flow, selection, and collective evolution. *Taxon* 50: 235–255.
- Royden LH, Burchfiel BC, van der Hilst RD. 2008. The geological evolution of the Tibetan plateau. *Science* 321: 1054–1058.
- Ruddiman WF, Kutzbach JE. 1989. Forcing of late Cenozoic Northern Hemisphere climate by plateau uplift in southern Asia and the American West. *Journal of Geophysical Research* 94: 18 409–18 427.
- Shi YF, Li JJ, Li BY. 1998. *Uplift and environmental changes of Qinghai-Xizang (Tibetan) in the Late Cenozoic*. Guangzhou, China: Guangdong Science and Technology Press.
- Shi YF, Yu G, Liu XD, Li BY, Yao TD. 2001. Reconstruction of the 30–40 ka bp enhanced Indian monsoon climate based on geological records from the Tibetan Plateau. *Palaeogeography, Palaeoclimatology, Palaeoecology* 169: 69–83.
- Spicer RA, Harris NBW, Widdowson M, Herman AB, Guo SX, Valdes PJ, Wolfe JA, Kelley SP. 2003. Constant elevation of southern Tibet over the past 15 million years. *Nature* 421: 622–624.
- Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* 7: 453–464.
- Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.
- Tapponnier P, Lacassin R, Leloup PH, Scharer U, Zhong D, Wu H, Liu X, Ji S, Zhang L, Zhong J. 1990. The Ailao Shan/Red River metamorphic belt: tertiary left-lateral shear between Indochina and South China. *Nature* 343: 431–437.
- Thorpe RS, Surget-Groba Y, Johansson H. 2010. Genetic tests for ecological and allopatric speciation in anoles on an island archipelago. *PLoS Genetics* 6: e1000929.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.
- Wagner CE, Harmon LJ, Seehausen O. 2012. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* 487: 366–369.
- Walker D. 1986. Late Pleistocene-early Holocene vegetational and climatic changes in Yunnan Province, southwest China. *Journal of Biogeography* 13: 477–486.
- Wan SM, Li A, Cliff PD, Stuut JBW. 2007. Development of the East Asian monsoon: mineralogical and sedimentologic records in the northern South China Sea since 20 Ma. *Palaeogeography, Palaeoclimatology, Palaeoecology* 254: 561–582.
- Wang LY, Abbott RJ, Zheng W, Chen P, Wang YL, Liu JQ. 2009a. History and evolution of alpine plants endemic to the Qinghai-Tibetan Plateau: *Aconitum gymnanthum* (Ranunculaceae). *Molecular Ecology* 18: 709–721.
- Wang WB, Jiang YB, Wang DM, Jing YB. 2006. Biological and ecological characteristics of *Taxus yunnanensis*. *Journal of West China Forestry Science* 35: 33–39.
- Wang H, Qiong L, Sun K, Lu F, Wang YG, Song ZP, Wu QH, Chen JK, Zhang WJ. 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.
- Wang YJ, Susanna A, Von Raab-Straube E, Milne R, Liu JQ. 2009b. Island-like radiation of *Saussurea* (Asteraceae: Cardueae) triggered by uplifts of the Qinghai-Tibetan Plateau. *Biological Journal of the Linnean Society* 97: 893–903.
- Weir BS, Cockerham CC. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.
- Wheeler NC, Jech KS, Masters SA, Obrien CJ, Timmons DW, Stonecypher RW, Lupkes A. 1995. Genetic variation and parameter estimates in *Taxus brevifolia* (Pacific yew). *Canadian Journal of Forest Research* 25: 1913–1927.
- Xiao XY, Shen J, Wang SM, Xiao HF, Tong GB. 2009. Climate change and evolution of the southwest monsoon revealed by pollen records in the Heqing deep drilling core. *Acta Palaeontologica Sinica* 48: 185–193.
- Xu TT, Abbott RJ, Milne RI, Mao K, Du FK, Wu GL, Ciren ZX, Miehle G, Liu JQ. 2010. Phylogeography and allopatric divergence of cypress species (*Cupressus* L.) in the Qinghai-Tibetan Plateau and adjacent regions. *BMC Evolutionary Biology* 10: 194.
- Yan HF, Zhang CY, Wang FY, Hu CM, Ge XJ, Hao G. 2012. Population expanding with the phalanx model and lineages split by environmental heterogeneity: a case study of *Primula obconica* in subtropical China. *PLoS ONE* 7: e41315.
- Yang FS, Li YF, Ding X, Wang XQ. 2008. Extensive population expansion of *Pedicularis longiflora* (Orobanchaceae) on the Qinghai-Tibetan Plateau and its correlation with the Quaternary climate change. *Molecular Ecology* 17: 5135–5145.
- Yang FS, Qin AL, Li YF, Wang XQ. 2012. Great genetic differentiation among populations of *Meconopsis integrifolia* and its implication for plant speciation in the Qinghai-Tibetan Plateau. *PLoS ONE* 7: e37196.
- Yao YF, Bruch AA, Cheng YM, Mosbrugger V, Wang YF, Li CS. 2012a. Monsoon vs uplift in Southwestern China–Late Pliocene climate in Yuanmou basin, Yunnan. *PLoS ONE* 7: e37760.
- Yao T, Thompson L, Yang W, Yu W, Gao Y, Guo X, Yang X, Duan K, Zhao H, Xu B *et al.* 2012b. Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. *Nature Climate Change* 2: 663–667.
- Yin A, Harrison TM. 2000. Geologic evolution of the Himalayan-Tibetan orogen. *Annual Review of Earth and Planetary Sciences* 28: 211–280.
- Zhang Q, Chiang TY, George M, Liu JQ, Abbott RJ. 2005. Phylogeography of the Qinghai-Tibetan Plateau endemic *Juniperus przewalskii* (Cupressaceae) inferred from chloroplast DNA sequence variation. *Molecular Ecology* 14: 3513–3524.
- Zhang YG, Ji J, Balsam W, Liu L, Chen J. 2009. Mid-Pliocene Asian monsoon intensification and the onset of Northern Hemisphere glaciation. *Geology* 37: 599–602.
- Zheng BX, Xu QQ, Shen YP. 2002. The relationship between climate change and Quaternary glacial cycles on the Qinghai-Tibetan Plateau: review and speculation. *Quaternary International* 97–98: 93–101.
- An Zs, Kutzbach JE, Prell WL, Porter SC. 2001. Evolution of Asian monsoons and phased uplift of the Himalayan-Tibetan plateau since Late Miocene times. *Nature* 411: 62–66.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 Samples of *Taxus wallichiana* used in this study

Table S2 Chloroplast haplotypes detected in the *trnL-trnF* region

Table S3 GenBank accession numbers of all sequences included for date and mutation rate estimate in this study

Table S4 The 11 calibration points used in relaxed molecular clock models for the present studies

Table S5 Codes for the 19 BioClim variables used to define the ecological habitats of *Taxus wallichiana* populations

Table S6 Pairs of *Taxus wallichiana* populations across the gene flow barrier used in the paired *t*-tests

Table S7 Summary of dataset I divergence time estimation results

Table S8 Population level means (and standard errors) for 19 BioClim variables for 43 *Taxus wallichiana* populations, including the results of two-tailed *t*-tests

Fig. S1 Frequency distributions of pairwise nucleotide differences for cpDNA *trnL-trnF* sequence data.

Fig. S2 Divergence times for the Taxaceae based on 39 genera of Gymnosperm.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.



About New Phytologist

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <25 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@ornl.gov)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**