Phylogeography and genetic structure of a Tertiary relict tree species, *Tapiscia sinensis* (Tapisciaceae): implications for conservation

Jinju Zhang¹, Zuozhou Li¹, Peter W. Fritsch², Hua Tian¹, Aihong Yang¹ and Xiaohong Yao^{1,*}

¹Key Laboratory of Plant Germplasm Enhancement and Speciality Agriculture, Wuhan Botanical Garden, The Chinese Academy of Sciences, Wuhan 430074, Hubei, China and ²Department of Botany, California Academy of Sciences, 55 Music Concourse Drive, San Francisco, CA 94118, USA

* For correspondence. E-mail yaohx@wbcas.cn

Received: 8 February 2015 Returned for revision: 14 May 2015 Accepted: 5 June 2015 Published electronically: 17 July 2015

• **Background and Aims** The phylogeography of plant species in sub-tropical China remains largely unclear. This study used *Tapiscia sinensis*, an endemic and endangered tree species widely but disjunctly distributed in sub-tropical China, as a model to reveal the patterns of genetic diversity and phylogeographical history of Tertiary relict plant species in this region. The implications of the results are discussed in relation to its conservation management.

• **Methods** Samples were taken from 24 populations covering the natural geographical distribution of *T. sinensis*. Genetic structure was investigated by analysis of molecular variance (AMOVA) and spatial analysis of molecular variance (SAMOVA). Phylogenetic relationships among haplotypes were constructed with maximum parsimony and haplotype network methods. Historical population expansion events were tested with pairwise mismatch distribution analysis and neutrality tests. Species potential range was deduced by ecological niche modelling (ENM).

• **Key Results** A low level of genetic diversity was detected at the population level. A high level of genetic differentiation and a significant phylogeographical structure were revealed. The mean divergence time of the haplotypes was approx. 1.33 million years ago. Recent range expansion in this species is suggested by a star-like haplotype network and by the results from the mismatch distribution analysis and neutrality tests.

• **Conclusions** Climatic oscillations during the Pleistocene have had pronounced effects on the extant distribution of *Tapiscia* relative to the Last Glacial Maximum (LGM). Spatial patterns of molecular variation and ENM suggest that *T. sinensis* may have retreated in south-western and central China and colonized eastern China prior to the LGM. Multiple montane refugia for *T. sinense* existing during the LGM are inferred in central and western China. The populations adjacent to or within these refugia of *T. sinense* should be given high priority in the development of conservation policies and management strategies for this endangered species.

Key words: Conservation, ecological niche modelling, multiple refugia, Pleistocene climate change, Tertiary relict tree species, local range expansion, sub-tropical China, *Tapiscia sinensis*, Staphyleaceae, Tapisciaceae.

INTRODUCTION

Pleistocene climatic oscillations have profoundly shaped the current geographical ranges and patterns of genetic variation among populations of many animal and plant species (Hewitt, 2000). Phylogeography, the study of the spatial and temporal history of a species, is a powerful tool for inferring post-glacial migration and potential glacial refugia, and is especially useful for the study of plants that lack a fossil pollen record (Avise, 2000). The locations of glacial refugia and post-glacial migration routes of species currently found in Europe and North America have been extensively documented with fossil and genetic evidence (e.g. Taberlet *et al.*, 1998; Abbott *et al.*, 2000; Petit *et al.*, 2005). In China, a number of phylogeographical studies have focused on the Qinghai-Tibetan Plateau as a glacial refugium (Qiu *et al.*, 2011; Liu *et al.*, 2012). In contrast, relatively few such studies have been conducted in plant species in sub-tropical China, an area harbouring many ancient endemic genera and families (Ying *et al.*, 1993; López-Pujol *et al.*, 2011; Qiu *et al.*, 2011; Liu *et al.*, 2012).

The area between the Qinling Mountains-Huai River (at approx. 34°N) and the tropical South ($\leq 22^\circ\text{N}$) is generally referred to as ‘sub-tropical China’ and consists mainly of hills and low mountains (Qiu *et al.*, 2011). Although sub-tropical China experienced severe climatic oscillations during Pleistocene glacial and interglacial cycles, large areas of this region remained unglaciated during the Last Glacial Maximum (LGM) (Axelrod *et al.*, 1996). Hence, sub-tropical China is considered a Pleistocene refugium for large numbers of ancient and relict species (e.g. Wang *et al.*, 2009). Moreover, owing largely to its great geographical heterogeneity and unique sub-tropical monsoonal climate, the region is a global centre of biodiversity, with many endemic and often endangered species (Ying *et al.*, 1993).

Because of the high physiographic and climatic complexity of sub-tropical China, it has been hypothesized that multiple glacial refugia existed in this region throughout the Pleistocene (Qiu *et al.*, 2011). This hypothesis has been supported by palaeoecological data and recent phylogeographical studies based on

plastid DNA markers (e.g. Zhou and Momohara, 2005; Gong et al., 2008; Wang et al., 2009). Compared with areas that have been colonized after the retreat of glaciers, refugial regions are expected to harbour unique haplotypes and maintain high levels of genetic diversity, and genetic differentiation among glacial refugia should be high owing to the random fixation of alleles (Hewitt, 2000). Although several phylogeographical studies have focused on the range dynamics of species in sub-tropical China and adjacent areas (e.g. Gong et al., 2008; Wang et al., 2009), the various species studied have largely non-overlapping distributional ranges, and the lack of detection of a consistent phylogeographical pattern may be due to species sampling artefact. More case studies on the phylogeography of Tertiary relict plant species in sub-tropical China are needed to test the existence and locations of glacial refugia during the LGM and to investigate the post-glacial migration routes of their descendants (Qiu et al., 2011).

The flowering plant genus *Tapiscia* (Tapisciaceae) is known from more than ten fossil species from East Asia, Europe and North America in the Eocene approx. 60 million years ago (Ma). *Tapiscia* became extinct in North America in the Oligocene, and in Europe during the Pleistocene interglacial period (Mai, 1980; Manchester, 1988). Fossil data indicate the presence of *Tapiscia pseudosinensis* in the northern part of its current range (e.g. Shandong) during the mid Miocene (10.64–18.87 Ma; Manchester, 1988). Currently, it comprises only two species, *T. sinensis* and *T. yunnanensis* (Li et al., 2008). *Tapiscia sinensis* is a geographically widespread and endangered woody species mainly distributed in the south-western and central sub-tropical mountains of China between 250 and 2200 m elevation (Fu and Jin, 1992). Currently, the natural habitat of *T. sinensis* at low elevation has been largely destroyed due to deforestation and reclamation; thus, natural populations of *T. sinensis* are small and rare. *Tapiscia sinensis* has been listed in the red list as an endangered plant species (Fu and Jin, 1992).

Ecological niche modelling (ENM) has become an increasingly popular tool for predicting the changes in species distribution in response to past or future climatic events (e.g. Qi et al., 2012; Liu et al., 2013). In this study, seven nuclear microsatellites and two plastid DNA intergenic spacer regions (*psbA-trnH* and *rpl32-trnL*) were utilized as molecular markers to construct phylogeographical patterns of 24 populations across the geographical range of *T. sinensis*. Moreover, ENM was utilized to characterize the potential geographical range at present and at the LGM. We address the following questions. (1) What is the degree of genetic diversity within the extant populations of *T. sinensis*? (2) Where were refugia, if any, located, and did multiple refugia exist across the geographical range of the species during the LGM? (3) What is the demographic history of *T. sinensis* populations in sub-tropical China? The data should also be valuable to help formulate appropriate strategies for the long-term conservation and management of *T. sinensis* and other plant species distributed in sub-tropical China.

MATERIALS AND METHODS

Study species

Tapiscia sinensis ($2n = 30$) is morphologically androdioecious but functionally dioecious, with male and hermaphroditic

individuals (Li et al., 2008). Adults of this species reach a height of 25 m. Flowering occurs from late May to early June. The flowers remain open for 4–5 d and can be visited by multiple insect pollinators, including members of Apidae and Syrphidae. The breeding system is suggested to be predominantly outcrossed (Lü and Liu, 2010). The drupaceous fruit is sub-globose to oval, 7–18 mm in diameter and borne on a short thick peduncle (Li et al., 2008). Seeds are dispersed mainly by gravity and occasionally by ants.

Sample collection

In this study, (1–) 6–30 individuals per population were sampled from 24 populations covering the natural geographical distribution of *T. sinensis*. In addition, one population from *T. yunnanensis* was sampled as the outgroup (Fig. 1; Table 1). Four to five fresh leaves were collected from each individual, dried in silica gel and stored at -20°C .

Genotyping and sequencing

Whole genomic DNA was extracted from leaf tissue using the cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1987). Seven nuclear simple sequence repeat (SSR) markers (*TS1*, *TS2*, *TS3*, *TS8*, *TS9*, *TS10* and *TS11*) developed by Zhang et al. (2013) were used for genotyping the 508 individuals. Details regarding genotyping the microsatellites are described in Zhang et al. (2013).

Seven non-coding plastid DNA regions were amplified for four individuals from four distant populations of *T. sinensis* with the following primer pairs: *trnL-c/trnL-d* and *trnL-e/trnL-f* (Taberlet et al., 1991), *psbA/trnH* (Sang et al., 1997), *trnS/trnG*, *rpl20/5'-rps12* (Hamilton, 1999), *rpl32-trnL* and *ndhF-rpl32* (Shaw et al., 2007). Of the seven primer pairs screened, *trnL-c/trnL-d*, *trnS/trnG*, *rpl20/5'-rps12*, *ndhF-rpl32*, *rpl32-trnL* and *psbA-trnH* produced clear single bands. Of these, *rpl32-trnL* and *psbA-trnH* were found to be the most variable and were therefore selected for a large-scale survey of haplotype variation in *T. sinensis*. The *psbA-trnH* and *rpl32-trnL* regions of the plastid genome were amplified and sequenced in accordance with the protocols in Sang et al. (1997) and Shaw et al. (2007), respectively. The *psbA-trnH* and *rpl32-trnL* sequences were deposited in GenBank under the accession numbers KJ881419–KJ881423 and KJ881424–KJ881434, respectively.

Microsatellite data analysis

The mean number of alleles per locus (A), observed heterozygosity (H_O), expected heterozygosity (H_E) and fixation indices (F_{IS}) for each population, and total expected heterozygosity (H_T) were estimated with FSTAT software (Goudet, 2001). Departure from Hardy–Weinberg expectations and linkage disequilibrium between microsatellites were tested by Fisher's exact tests (GENEPOP, version 4.0.10; Rousset, 2008) within each population and for each locus. Sequential Bonferroni correction was used to judge significance levels for all tests with an initial α level of 0.05 (Rice, 1989). A test for null alleles was

FIG. 1. Map showing the geographical distribution of the plastid DNA haplotypes among populations of *Tapiscia sinensis*.

performed with MICRO-CHECKER 2.2 (Van Oosterhout *et al.*, 2004).

The individual-based population assignment test was used to examine population structure with the program STRUCTURE (Pritchard *et al.*, 2000). The range of clusters (K) was predefined from 1 to 30. For each K -value, five independent runs were performed to ensure consistent results using the admixture model and such that the allele frequencies are correlated within populations, with a burn-in of 10^5 steps and a run length of 10^5 steps. To determine the optimal K -value, the posterior probability [$\ln P(D)$] was calculated for each mean value of K by using the mean log-likelihood of K . The most appropriate K was also predicted from plots of ad hoc posterior probability models of ΔK (Evanno *et al.*, 2005). Individuals were assigned to clusters for selected K -values, and individual membership likelihoods were calculated. Partitioning genetic differentiation within and among populations was further performed by analysis of molecular variance (AMOVA) with Arlequin 3.0 (Excoffier *et al.*, 2005) for both data sets [nuclear SSRs (nSSRs) and plastid DNA]. Significance was determined with 10 000 permutations.

The possible loss of genetic variation through bottlenecks, when populations have experienced a recent demographic

decline, was tested by using the program Bottleneck (Piry *et al.*, 1999). The Wilcoxon test was used to test the significance of heterozygote excess under the infinite allele model (IAM), the stepwise mutation model (SMM) and the two-phase model (TPM), as recommended by Piry *et al.* (1999).

A Mantel test (Mantel, 1967) was used to examine the relationship between genetic (pairwise F_{ST}) and geographical distances employing the program IBD (isolation by distance; Bohonak, 2002) for both data sets (nSSRs and plastid DNA). The analysis was based on 10 000 randomizations of the data. Confidence limits of the relationship were based on 10 000 bootstrap re-samples of the data.

Plastid DNA data analysis

Haplotype diversity (h ; Nei, 1987) and nucleotide diversity (π ; Tajima, 1983) were estimated for each population with DnaSP version 5.0 (Librado and Rozas, 2009). A median-joining network was constructed with the Network software package (V4.5.1.6; <http://www.fluxus-engineering.com>). Phylogenetic relationships among the plastid DNA haplotypes

TABLE 1. Collection sites, sample sizes (n) and estimates of genetic diversity of the 24 *Tapiscia sinensis* populations and one *Tapiscia yunnanensis* population (YN) from sub-tropical China

Code	Population locality	Mountain region	Altitude	Latitude (N)	Longitude (E)	SSR analysis					Plastid DNA analysis			
						n	A	H _O	H _E	F _{IS}	n	Haplotype (Sample no.)	h	$\pi \times 10^{-3}$
TMS	Lin'an, Zhejiang	Tianmu Mts.	596–630	30°19'	119°26'	7	3.000	0.510	0.459	–0.111	7	H14	0	0
BYY	Tonglu, Zhejiang	Tianmu Mts.	441–680	29°41'	119°44'	20	4.500	0.529	0.515	–0.026	13	H14	0	0
QLF	Jixi, Anhui	Tianmu Mts.	711–818	30°08'	118°51'	6	2.571	0.476	0.474	–0.005	6	H14	0	0
WFS	Shucheng, Anhui	Dabie Mts.	802–1207	31°03'	116°33'	30	3.285	0.305	0.363	0.160***	15	H6	0	0
LS	Jiujiang, Jiangxi	Mufu Mts.	1100–1234	29°32'	115°59'	25	2.714	0.280	0.304	0.080	15	H7	0	0
YF	Yifeng, Jiangxi	Luoxiao Mts.	347–379	28°32'	114°33'	30	3.667	0.581	0.583	0.004	15	H11	0	0
JGS	Jinggangshan, Jiangxi	Luoxiao Mts.	850–1104	26°38'	114°16'	20	3.571	0.650	0.584	–0.113	15	H11	0	0
WYS	Wuyishan, Fujian	Wuyi Mts.	901–1083	27°51'	117°43'	30	4.000	0.414	0.497	0.167***	12	H6	0	0
XDS	Lichuan, Hubei	Wuling Mts.	1202–1532	29°58'	108°58'	16	4.289	0.554	0.639	0.133**	15	H2(6), H3(9)	0.514	3.1
HHB	Wufeng, Hubei	Wuling Mts.	1073–1235	30°06'	110°34'	11	4.571	0.571	0.673	0.151***	9	H3(3), H4(6)	0.500	5.1
DLL	Yichang, Hubei	Wuling Mts.	1497–1606	31°05'	110°55'	16	3.571	0.348	0.525	0.337***	13	H5	0	0
FJS	Jiangkou, Guizhou	Wuling Mts.	883–1269	28°01'	108°48'	30	5.571	0.678	0.345	0.123**	15	H4	0	0
ZP	Zhenping, Shannxi	Daba Mts.	864–950	31°46'	109°29'	9	4.428	0.556	0.707	0.215**	9	H3(7), H4(2)	0.389	4.0
SN	Shennongjia, Hubei	Daba Mts.	1026–1603	31°38'	110°25'	26	4.571	0.566	0.692	0.182**	8	H12	0	0
LJ	Nanjiang, Sichuan	Daba Mts.	1181	32°39'	106°54'	—	—	—	—	—	1	H4	0	0
JFS	Nanchuan, Chongqing	Dalou Mts.	1341–1455	29°04'	107°19'	16	3.429	0.619	0.671	0.078	14	H4	0	0
LTS	Zhengfeng, Guizhou	Yungui Plateau	1500–1628	25°23'	105°29'	30	4.143	0.486	0.562	0.135**	14	H13	0	0
XWS	Xifeng, Guizhou	Yungui Plateau	1300–1408	27°07'	106°37'	30	5.000	0.533	0.659	0.191***	15	H4	0	0
LGS	Leishan, Guizhou	Yungui Plateau	1372–1783	26°22'	108°11'	30	4.143	0.538	0.629	0.144***	15	H4	0	0
DG	Daguan, Yunnan	Yungui Plateau	1618–1895	28°15'	103°55'	30	5.143	0.633	0.640	0.010	15	H8(9), H9(6)	0.514	5.3
DJY	Duijiangyan, Sichuan	Min Mts.	701–1240	31°00'	103°36'	30	5.000	0.574	0.624	0.076	15	H1	0	0
EM	Emeishan, Sichuan	Qionglai Mts.	815–1660	29°35'	103°22'	30	5.429	0.624	0.675	0.076	14	H4	0	0
GX	Longsheng, Guangxi	Nanling Mts.	794–970	25°37'	109°54'	6	3.143	0.571	0.562	–0.017	6	H4	0	0
SHS	Yongzhou, Hunan	Nanling Mts.	1012–1230	26°26'	111°01'	30	2.571	0.205	0.223	0.081	15	H10	0	0
Mean						—	4.013	0.513	0.548	—	—	—	—	—
YN	Wenshan, Yunnan	Yungui Plateau	1344–1566	23°21'	104°26'	—	—	—	—	—	6	H4(5), H15(1)	0.333	3.4

A, average number of alleles per locus; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, within-population coefficient of inbreeding; h, haplotype diversity; π , nucleotide diversity.

Levels of significance for F_{IS} of each population were determined after 1000 permutations: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

were reconstructed by maximum parsimony in Paup 4.0b10 (Swofford, 2000) with *T. yunnanensis* employed as outgroup, and with random addition of sequences (1000 replicates) and TBR branch swapping selected. Bootstraps were generated by using full heuristic searches (1000 bootstrap replicates). In this analysis, indels were treated as single mutational events.

Phylogeographical structure was inferred with PERMUT (Pons and Petit, 1996) from 1000 random permutations of haplotypes across populations. A significantly higher N_{ST} than G_{ST} usually indicates the presence of phylogeographical structure.

A spatial analysis of molecular variance (SAMOVA) was used to identify clusters of genetically similar populations. Geographically homogeneous populations are clustered into a user-defined number of groups (K) with a simulated annealing approach to maximize the proportion of total genetic variance observed between groups (F_{CT}). SAMOVA was computed by SAMOVA 1.0 (Dupanloup et al., 2002) with 100 simulated annealing processes for each value of K from $K = 2$ to $K = 18$.

Pairwise mismatch distributions were carried out with DnaSP version 5.0 (Librado and Rozas, 2009) to infer the historical demography of *T. sinensis*. Mismatch distributions for the species were calculated with the expected frequency based on a population growth-decline model. The sum of squared deviations (SSDs) between observed and expected mismatch distribution and the raggedness index (HR_{ag} ; Harpending, 1994) were calculated to test the null hypothesis of spatial expansion using Arlequin ver. 3.0 (Excoffier et al., 2005). Statistical significance of population expansion was further assessed with

three neutrality tests, the Tajima's D (Tajima, 1989), Fu and Li's D^* (Fu and Li, 1993) and Fu's F_S (Fu, 1997) implemented in DnaSP or Arlequin, whereby significantly negative values of the neutrality tests indicate population expansion (Aris-Brosou and Excoffier, 1996; Fu, 1997).

Net pairwise sequence divergence per base pair (d_A), which is proportional to time since divergence (T) of two haplotype lineages, was calculated with MEGA6 (Tamura et al., 2013) under the Kimura-2 model. Divergence time was calculated as $T = d_A/2\mu$, where μ is the rate of nucleotide substitution (Nei and Kumar, 2000). Because a taxon-specific substitution rate has not been reported for non-coding plastid DNA regions of *Tapiscia*, we assumed a substitution rate of 1.52×10^{-5} substitutions per neutral site per year (Wolfe et al., 1987). A conservative range for divergence time was calculated by the use of a 95 % confidence interval (CI).

Ecological niche modelling

Ecological niche modelling was carried out to predict suitable present and past climate envelopes for *T. sinensis* by using the MAXENT v3.2 software package (Phillips et al., 2006), with 80 % of the data used in training and 20 % retained as test points. The geographical range estimate of *T. sinensis* incorporated all sites sampled in this study and 37 other available specimen records with GPS data from the Chinese Virtual Herbarium (Supplementary Data Table S2). Current environmental data layers were constructed for the 19 BIOCLIM

TABLE 2. Results of AMOVA testing of the population genetic structure in *Tapiscia sinensis*

	Source of variation	d.f.	Sum of squares	Variance components	Percentage variation	F_{ST}
SSR	Among populations	22	963.121	0.95357	32.56	0.326*
	Within populations	993	1960.909	1.97473	67.44	
	Total	1015	2924.030	2.92830		
Plastid DNA	Among populations	23	896.612	3.22628	98.77	0.988*
	Within populations	267	10.756	0.04028	1.23	
	Total	290	907.368	3.26656		

* $P < 0.001$, 1000 permutations.

variables (at 2.5 arcmin resolution) in the WorldClim data set (Hijmans *et al.*, 2005), whereas the LGM [approx. 21 thousand years ago (Ka)] climate data were drawn from general circulation model (GCM) simulations from the Community Climate System Model (CCSM) (version 3.0; Collins *et al.*, 2006). The convergence threshold was left at the default value (10^{-5}) and the maximum number of iterations was set at 500. Model performance was evaluated by using the area under the receiver operating characteristic curve (AUC). Values between 0.7 and 0.9 indicate good discrimination (Swets, 1988).

RESULTS

Microsatellite diversity

The micro-checker software detected the presence of null alleles in three loci, and the frequency of null alleles per locus ranges from 0.5 % (TS10) to 1 % (TS3). Across all populations, 130 alleles were identified from the seven microsatellite loci surveyed. The expected heterozygosity (H_E) ranged from 0.304 for population LS to 0.675 for population EM, with the average expected heterozygosity (H_E) per population being 0.548. A high level of microsatellite diversity was observed at the species level ($H_T = 0.816$). The average observed heterozygosity (H_O) is 0.513, and the average number of alleles per locus is 4.013 (Table 1). The estimated level of inbreeding (F_{IS}) for each population ranges from -0.113 to 0.337 . Eleven of 23 populations were observed to deviate from Hardy–Weinberg equilibrium when all loci were combined and a Bonferroni-type correction applied (Table 1). Significant linkage disequilibrium between SSR loci within each population was detected for 73 out of 483 comparisons at the 5 % significance level, but none of the comparisons was significant after a Bonferroni-type correction was applied.

The AMOVA revealed that 32.56 % of the total partitioned genetic variation can be attributed to differences among populations, and 67.44 % to differences among individuals within populations ($P < 0.001$; Table 2). Two methods used for determining the number of clusters present in the data set (K) resulted in different estimates of K . The mean estimated logarithm of probability of the data [$\ln P(D)$] is maximal for $K = 17$. Moreover, the highest peak in ΔK at $K = 9$ indicates that nine genetic clusters were detected (Fig. 2A, B). The majority of multilocus genotypes exhibited high membership proportions in the nine genetic clusters. These clusters almost corresponded well to geographical regions. Individuals from Tianmu, Wuyi, Luoxiao, Min and Qionglai Mountain grouped as different clusters, whereas the six populations from western Hubei and eastern Chongqing (DLL, SN, HHB, ZP, XDS and

JFS) grouped as a single cluster. Populations from Yungui Plateau grouped as three different clusters. Individuals from the Nanling Mountains (GX and SHS) formed two different clusters (Fig. 2C).

The Wilcoxon's statistical test revealed a demographic bottleneck in 11 and five populations under the IAM model and TPM model, respectively. However, only one population of *T. sinensis* has probably experienced a recent decline of population size under the SMM model (see Table 3).

No significant correlation was detected between genetic distances and geographical distances among the 23 natural populations ($r = -0.0451$, $P = 0.80$, Supplementary Data Fig. S1), indicating that gene flow was not geographically limited among *T. sinensis* populations.

Haplotype and nucleotide diversity

Sequence analyses of *T. sinensis* resulted in an aligned fragment length of 388 bp of *psbA-trnH* with 12 nucleotide substitutions. In *rpl32-trnL*, two substitutions and two indels were detected in the 617 bp of the partial intron sequence (5' end). Fourteen haplotypes (H1–H14; Supplementary Data Table S1) were identified from 16 polymorphic sites. The haplotype diversity (h) for each population ranged from 0 to 0.514, and the nucleotide diversity (π) from 0 to 0.0053. Genetic diversity calculated for all populations ($h_T = 0.896$) was higher than the average intrapopulation diversity ($h_S = 0.083$).

Twenty of the 24 analysed populations were monomorphic, whereas the remaining populations (DG, HHB, XDS and ZP) exhibited two haplotypes (Fig. 1). Of the 14 haplotypes, nine were found in only one population and five in more than one population (Table 1; Fig. 1). The most frequent and widespread haplotype (H4) was found in 28.5 % of all samples and was connected to the outgroup sequence of *T. yunnanensis* by four mutations. Haplotypes H6, H7, H11 and H14 were fixed in the eastern region of the species (Tianmu, Dabie, Wuyi, Mufu and Luoxiao Mountains). Haplotypes H2, H3, H5 and H12 were distributed in the central populations (Daba, Wuling and Dalou Mountains). In the western region (Yungui Plateau, and Min and Wuling Mountains), haplotypes H1, H10 and H13 were found in the sole populations DJY, SHS and LTS, respectively. The two remaining haplotypes (H8 and H9) were found in population DG.

Maximum parsimony analysis resulted in a single tree (length = 18 and consistency index = 1.0). Ten clades with low bootstrap support were identified: one consisting of Hap H2 and H3, one of Hap H6, H7 and H14, one of Hap H8 and H9 and the other haplotypes forming a single clade each (Fig. 3A).

FIG. 2. Bayesian inference of the number of clusters (K) of *Tapiscia sinensis* based on SSR data. (A) Plot of the posterior probability [$\ln P(D)$] for each K (five replicates). (B) The ΔK statistics. (C) Bar plot of STRUCTURE analysis for the model with $K=9$.

TABLE 3. Probabilities from tests (Wilcoxon's) for mutation drift equilibrium in 23 populations of *Tapiscia sinensis* under the infinite allele model (IAM), stepwise mutation model (SMM) and two-phase model (TPM) for microsatellite data

Population	Mutation drift test			Mode shift
	IAM	TPM	SMM	
TMS	0.016*	0.078	0.375	Shifted
BYY	0.023*	0.296	0.375	Normal
QLF	0.023*	0.297	0.938	Normal
WFS	0.023*	0.055	0.578	Shifted
LS	0.078	0.039*	0.578	Normal
YF	0.039*	0.297	1.000	Normal
JGS	0.078	0.438	0.563	Shifted
WYS	0.094	0.813	0.813	Normal
XDS	0.688	0.469	0.078	Normal
HHB	0.008**	0.008**	0.023*	Normal
DLL	0.109	0.688	0.938	Normal
FJS	0.008**	0.031*	0.563	Normal
ZP	0.008**	0.039*	0.055	Shifted
SN	0.313	0.875	0.875	Normal
LJ	—	—	—	—
JFS	0.688	0.843	0.563	Normal
LTS	0.375	0.375	0.938	Normal
XWS	0.375	0.938	0.813	Shifted
LGS	0.023*	0.109	0.938	Normal
DG	0.078	0.469	0.938	Normal
DJY	0.055	0.375	1.000	Normal
EM	0.023*	0.109	1.000	Normal
GX	0.156	0.813	1.000	Normal
SHS	0.008**	0.008**	0.297	Normal

*Significant at $P < 0.05$; **significant at $P < 0.01$.

In the haplotype network, a star-like pattern was recovered. The ancestral-like haplotype H4 lies at the centre of the network, with other derivative haplotypes connected to it independently (Fig. 3B).

Mantel tests revealed no significant association ($r = 0.406$, $P = 1.00$) between interpopulation pairwise geographical distances and their corresponding genetic distance.

Population genetic structure and demographic history in plastid DNA data

As tested by the AMOVA, all *T. sinensis* populations were strongly differentiated ($F_{ST} = 0.988$, $P < 0.0001$; Table 2). The range-wide N_{ST} value of 0.946 was significantly larger than its corresponding G_{ST} (0.907, $P < 0.05$), indicating the existence of marked phylogeographical structure at this scale. In the SAMOVA, the F_{CT} values increased with K , nearly reaching plateau values at $K = 13$ ($F_{CT} = 0.92$). However, F_{CT} values above $K = 10$ did not increase significantly.

The neutrality tests with the entire data set supported the null hypothesis of range expansion (Fu's $F_S = -2.807$, $P < 0.05$; Tajima's $D = -1.594$, $P < 0.05$; Fu and Li's $D^* = -2.562$, $P < 0.05$). The observed mismatch distributions for plastid DNA haplotypes for *T. sinensis* were unimodal, and thus consistent with the bell-shaped curve expected for an expanding population (Fig. 4) (Rogers and Harpending, 1992). The non-significant SSD statistic ($SSD = 0.032$, $P > 0.05$) and the H_{Rag} value ($H_{Rag} = 0.042$, $P > 0.05$) under the spatial expansion models failed to reject the spatial expansion model.

FIG. 3. Haplotype network (A) and phylogenetic relationships (B) of 14 cpDNA haplotypes resolved in *Tapiscia sinensis*. Circle size is proportional to the relative haplotype frequency. Bootstrap values ($>50\%$) are denoted above branches.

FIG. 4. Frequency distributions of pairwise nucleotide differences for plastid DNA sequence data in all populations under the sudden expansion model. The solid line indicates observed values; the dashed line represents expected values.

Mean pairwise divergence ($d_A = 0.00404$) and substitution rate (1.52×10^{-9} per site per year) recovered a mean divergence time of the haplotype of approx. 1.33 Ma (95 % CI 1.02–1.74 Ma).

Ecological niche models

Ecological niche models performed well in predicting the geographical distribution of *T. sinensis* (the test AUC = 0.852 ± 0.047 s.d. across ten replicate runs). The current distributional predictions were good representations of the actual distribution of the species (Fig. 5A). However, the distribution models suggest an expansion in suitable distribution area in the current distribution model of *Tapiscia* in comparison with the paleodistribution model for the LGM (Fig. 5B).

DISCUSSION

Genetic diversity in *Tapiscia sinensis*

At the species level, our result reveals a high level of microsatellite genetic diversity ($H_T = 0.816$) and total haplotype diversity ($h_T = 0.896$) across the 23 populations of *T. sinensis* in sub-tropical China. The genetic diversity at plastid DNA sequences in *T. sinensis* is higher than that in other tree species, such as *Eurycorymbus cavaleriei* ($h_T = 0.834$; Wang *et al.*, 2009) and *Dalbergia nigra* ($h_T = 0.753$; Ribeiro *et al.*, 2011), but lower than that in the herb species, *Ligularia hodgsonii* ($h_T = 0.913$; Wang *et al.*, 2013). In the Tertiary, *Tapiscia* is inferred to have had a wide natural range, covering much of the world (Mai, 1980; Manchester, 1988). However, most *Tapiscia* spp. are extinct due to Pleistocene climatic oscillations. Hence, high-level diversity indices in *T. sinensis* might result from high inherent variability of the ancestral species.

Despite a high level of plastid DNA diversity detected at the species level, within-population diversity was low (mean $h_S = 0.083$), especially in view of our large number of samples, and most populations (20 out of 24) are fixed for only one haplotype. The coefficient of genetic differentiation ($G_{ST} = 0.907$) detected in the extant populations of *T. sinensis* was higher than that found in studies of other plants employing plastid DNA markers (average $G_{ST} = 0.637$; Petit *et al.*, 2005). A high degree of genetic differentiation ($F_{ST} = 0.988$) was observed based on plastid DNA data. A low level of genetic variation at the population level in this species was also revealed by microsatellite markers (mean $H_E = 0.548$). Moreover, a high degree of genetic differentiation based on microsatellites has also been found in this species ($F_{ST} = 0.326$). The combination of low genetic diversity and high genetic differentiation has been previously reported for a widespread tree species, *Cercidiphyllum japonicum* (Qi *et al.*, 2012).

FIG. 5. Potential distribution ranges of *Tapiscia* in East Asia at the present (A) and at the Last Glacial Maximum (B) simulated by ecological niche modelling. Warmer colours denote areas with better predicted environmental conditions than cooler colours. Black dots show the extant occurrence points of *Tapiscia* species.

This genetic pattern probably reflects historical demography and biological traits of *T. sinensis*. First, the habitat of plant species in the mountain ranges of sub-tropical China has been fragmented due to past climate change (Fu and Jin, 1992). Populations of particular species consequently experienced a reduction in size and isolation. Theoretically, small populations tend to experience either a genetic bottleneck or genetic drift.

Bottleneck analysis indicated that 11 out of 23 populations have experienced a recent population bottleneck under the assumptions of the IAM. Thus, low genetic diversity within populations and high levels of differentiation among populations observed in the present study are expected. Strong genetic structure resulting from ancient fragmentation has also been documented for other relict tree species occurring across a

similar geographical range to that of *T. sinensis* (e.g. Wang and Ge, 2006). Secondly, the high population differentiation in *T. sinensis* is also likely to be affected by plant life history characteristics, especially the type of breeding system. As in *T. sinensis*, it has been shown that other species with a dioecious or morphologically androdioecious but functionally dioecious breeding system tend to maintain high genetic structure (e.g. Rendell and Ennos, 2003; Wang et al., 2009). Lastly, high levels of genetic differentiation among *T. sinensis* populations may result from restricted gene flow among the studied populations. Species with seeds dispersed by gravity and pollinated by insects tend to exhibit restricted gene flow among populations (e.g. Yao et al., 2007), which is also inferred for *T. sinensis*. With restricted gene flow observed in the fragmented habitat of *T. sinensis*, genetic differentiation between fragments will increase and genetic diversity within fragments will decrease.

Potential refugia and demographic history

The glacial–interglacial cycles during the Quaternary have greatly affected species distribution ranges (Comes and Kadereit, 1998; Hewitt, 2000). Repeated range expansion and contraction of most plants and animals may cause migration and/or extinction of populations, followed by periods of isolation, divergence and subsequent expansion (e.g. Liu et al., 2013). Our dating estimates indicate that the diversification in *T. sinensis* started in the early Pleistocene (1.02–1.74 Ma), suggesting that haplotypes differentiated before the LGM. Therefore, climatic oscillations during the Pleistocene will have had pronounced effects on the recent population history of *T. sinensis* relative to the LGM, as is documented in other plant species in East Asia (Qiu et al., 2011).

According to coalescent theory, the most ancient haplotypes should be located at the interior nodes of a haplotype network, whereas the most recent haplotypes should be at the tips (Posada and Crandall, 2001). In the minimum plastid DNA-spanning network in *T. sinensis*, haplotype H4 was found to be at the interior position, coupled with high frequency. This ancestral haplotype was found to be most widespread across the Yungui Plateau and central China (including the Daba, Shennongjia and Wuling Mountains), suggesting that the species persisted over the longer term in these regions during the LGM. In addition, glacial refugial areas are generally expected to harbour unique haplotypes and a relatively higher level of genetic diversity than unglaciated, newly founded areas (Comes and Kadereit, 1998; Petit et al., 2003). Compared with other regions, a higher level of haplotype diversity observed in the central and western region indicates that the Yungui Plateau and the mountains of central China might be the refugia for *T. sinensis*. The Yungui Plateau and its adjacent regions, as one of the world's biodiversity hotspots (Myers et al., 2000), have long been considered to be an important centre of origin for the East Asiatic flora (Wang, 1992) and a refugium for many plant species in China (Ying et al., 1993). It might furthermore have served as a refugium of ancient elements of Tapisciaceae, because the closest relative of *T. sinensis* (*T. yunnanensis*) also occurs in this area (Li et al., 2008). In addition, the mountains of central China are one of the richest areas for plant diversity in China, including many 'living fossils' such as *Cunninghamia*

lanceolata, *Eucommia ulmoides*, *Metasequoia glyptostroboides* and *Pseudolarix amabilis* (Ying et al., 1993). The great diversity in topography, climate and ecological conditions of both regions makes these regions more likely to be refugia for plant species in general (Qiu et al., 2011). Previous studies based on molecular biogeographical data also indicate that the Yungui Plateau and central China were important refugia for a large number of relict species including *Cathaya argyrophylla* (Pinaceae) (Wang and Ge, 2006) and *Eurycorymbus cavaleriei* (Sapindaceae) (Wang et al., 2009). Moreover, our ENM data also provide firm evidence for the survival of *T. sinensis* in these areas during the LGM.

Plastid haplotype variation across the species exhibits significant phylogeographical structure. In addition, we found a large number of mountain-specific haplotypes. In the western region, ancestral haplotype H1 was specific to the Min Mountains, and H8, H9 and H13 were specific to the Yungui Plateau. For the regions in central China, H2, H3 and H5 are specific to the Wuling Mountains and H6 is fixed in the Shennong Mountains. Thus, the Yungui Plateau and its adjacent regions and central China mountain regions may have been sub-divided into multiple small refugia, consistent with isolation in distinct mountain chains observed in other species (e.g. Wang et al., 2009; López-Pujol et al., 2011; Wang et al., 2013). The multiple-refugia hypothesis is also supported by the STRUCTURE analysis and SAMOVA, i.e. the alignment tests revealed nine clusters almost corresponding to separate geographical regions. Moreover, the 14 plastid DNA haplotypes observed in *T. sinensis* can be divided into ten phylogenetically distinct lineages. Each of the clades (except clade H4) possesses a distinct geographical distribution and is dominated by a particular haplotype.

Star-like patterns in sequence networks are generally considered to be characteristic of recent range expansion (Excoffier et al., 2009). In the present study, the pattern of haplotype distribution is consistent with recent range expansion events. Moreover, the population expansion was supported by the unimodal mismatch distribution and significantly negative Tajima's *D*, Fu and Li's *D** and Fu's *F*_s values. According to the LGM biome map, the refugia of *T. sinensis* in China were covered by warm-temperate evergreen forest prior to the last glacial cycles (Yu et al., 2000; Harrison et al., 2001; Qiu et al., 2011). *Tapiscia sinensis* may have retreated to south-western and central China as a refuge since the Miocene when the global climate began to cool (Zhou and Momohara, 2005). Although our data do not provide accurate information about the migration pathways during the post-glacial period, the presence of haplotype H4 in the Nanling Mountains and derived haplotypes in eastern China suggests that *T. sinensis* probably migrated eastward from the refugial areas in south-western China as the climate warmed. Similarly, the existence of a specific west–east geographical gradient of genetic decline, considered a footprint of eastward range expansion, was detected in *T. sinensis*. The Nanling Mountains functioned as one of the important 'ice-free' corridors for the eastward migration of the East Asiatic flora (Wang, 1992). Such a corridor was inferred to have been responsible for the pattern of genetic variation in *T. sinensis* and thus may also have promoted the eastward spread of *T. sinensis*. Such eastward colonization was also revealed in *Eurycorymbus cavaleriei* (Wang et al., 2009).

The absence of haplotype H4 and no haplotype diversity in eastern China suggest that a rare haplotype was fixed due to a founder effect as the species expanded its range following past fragmentation. However, we cannot exclude the possibility that eastern populations may have been colonized from cryptic refugia at relatively southern latitudes in eastern China. For example, the presence of haplotype H6 in the Wuyi and Dabie Mountains, and haplotype H11 in the Luoxiao Mountains, indicates the occurrence of localized range expansion. Thus, the plastid DNA phylogeographical results from *T. sinensis* suggest a history of ancient fragmentation and localized range expansion from refugia, as is also documented in *Taxus wallichiana* (Gao *et al.*, 2007).

Implications for conservation

From a conservation perspective, the low level of genetic diversity and high genetic differentiation uncovered for *T. sinensis* populations suggest that *in situ* measures should be taken into consideration in the conservation management of this endangered species. Based on high nucleotide diversity levels and haplotype uniqueness, we suggest that *T. sinensis* survived the LGM in multiple isolated refugia, including the Yungui Plateau and the Min, Daba and Wuling Mountains. Thus, *in situ* conservation priorities for this species should target at least the populations of the Yungui Plateau and the mountains in central China. Moreover, several areas (e.g. the Tianmu, Wuyi and Luoxiao Mountains) possessing different haplotypes should also be targeted to maintain the majority of genetic diversity of *T. sinensis*. Although *T. sinensis* is considered to be one of the most endangered plant species in China (Fu and Jin, 1992), it is fortunate that most populations of the species are currently under legal protection in nature reserves. In addition, our field survey work in conjunction with the present study indicates that some populations of this species outside the nature reserves (e.g. populations LJ and ZP) are small and lack regeneration. Moreover, disequilibrium between mutation and genetic drift and heterozygote deficiency were observed in the population ZP, suggesting increased risk of extinction as a result of genetic decline through inbreeding and genetic drift. Hence, *ex situ* conservation measures, such as seed banking and reintroduction, should be implemented urgently to counteract the parlous *in situ* conservation state of the species.

SUPPLEMENTARY DATA

Supplementary data are available online at www.aob.oxfordjournals.org and consist of the following. **Figure S1**: scatter plots of pairwise genetic distances and geographical distances between the 23 *Tapiscia sinensis* populations based on microsatellites and chloroplast DNA data. **Table S1**: description of the 14 *Tapiscia sinensis* haplotypes from the chloroplast DNA fragments combined. **Table S2**: list of available specimen records used in ENM analysis.

ACKNOWLEDGEMENTS

We thank Puxin Gao, Guohua Lan, Tao Deng, Zhiguang Wang and Mingke Yan for their help in collecting samples.

This work was supported in part by the Natural Scientific Foundation of China (31400476, 31270384) and the Director's Fund of Wuhan Botanical Garden, Chinese Academy of Sciences (KSCXZ-EW-J-20).

LITERATURE CITED

- Abbott RJ, Smith LC, Milne RI, *et al.* 2000. Molecular analysis of plant migration and refugia in the Arctic. *Science* **289**: 1343–1346.
- Aris-Brosou S, Excoffier L. 1996. The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Molecular Biology and Evolution* **13**: 494–504.
- Avise JC. 2000. *Phylogeography: the history and formation of species*. Cambridge, MA: Harvard University Press.
- Axelrod DI, Al-Shehbaz I, Raven PH. 1996. History of the modern flora of China. In: A Zhang, S Wu, eds. *Floristic characteristics and diversity of East Asian plants*. New York: Springer, 43–55.
- Bohonak AJ. 2002. IBD (isolation by distance): a program for analyses of isolation by distance. *Journal of Heredity* **93**: 153–154.
- Collins WD, Bitz CM, Blackmon ML, *et al.* 2006. The community climate system model version 3 (CCSM3). *Journal of Climate* **19**: 2122–2143.
- Comes HP, Kadereit JW. 1998. The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science* **3**: 432–438.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.
- 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, Foll M, Petit RJ. 2009. Genetic consequences of range expansions. *Annual Review of Ecology, Evolution and Systematics* **40**: 481–501.
- Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**: 47–50.
- Fu LK, Jin JM. 1992. *China plant red data book – rare and endangered plants*. Beijing: Science Press.
- Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915–925.
- Fu YX, Li WH. 1993. Statistical tests of neutrality of mutations. *Genetics* **133**: 693–709.
- Gao LM, Möller M, Zhang XM, *et al.* 2007. High variation and strong phylogeographic pattern among cpDNA haplotypes in *Taxus wallichiana* (Taxaceae) in China and North Vietnam. *Molecular Ecology* **16**: 4684–4698.
- Gong W, Chen C, Dobeš C, Fu CX, Koch MA. 2008. Phylogeography of a living fossil: Pleistocene glaciations forced *Ginkgo biloba* L. (Ginkgoaceae) into two refuge areas in China with limited subsequent postglacial expansion. *Molecular Phylogenetics and Evolution* **48**: 1094–1105.
- Goudet J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from <http://www.unil.ch/izea/software/fstat.html>.
- Harpending HC. 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* **66**: 591–600.
- Hamilton MB. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology* **8**: 521–522.
- Harrison SP, Yu G, Takahara H, Prentice IC. 2001. Diversity of temperate plants in east Asia. *Nature* **413**: 129–130.
- Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.
- 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.
- Li DZ, Cai J, Wen J. 2008. *Flora of China*, Vol. 11. Beijing: Science Press and St Louis: Missouri Botanical Garden Press.
- Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

- Liu J, Möller M, Provan J, Gao LM, Poude RC, Li DZ. 2013. Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. *New Phytologist* **199**: 1093–1108.
- Liu JQ, Sun YS, Ge XJ, Gao LM, Qiu YX. 2012. Phylogeographic studies of plants in China: advances in the past and directions in the future. *Journal of Systematics and Evolution* **50**: 267–275.
- López-Pujol J, Zhang FM, Sun HQ, Ying TS, Ge S. 2011. Centres of plant endemism in China: places for survival or for speciation? *Journal of Biogeography* **38**: 1267–1280.
- Lü W, Liu WZ. 2010. Pollination biology in androdioecious species *Tapiscia sinensis* (Staphyleaceae). *Chinese Bulletin of Botany* **45**: 713–722.
- Mai DH. 1980. Zur Bedeutung von Relikten in der Florengeschichte. Im: Vent W, ed. *100 Jahre Arboretum (1879–1979)*. Berlin, 281–307.
- Manchester SR. 1988. Fruits and seeds of *Tapiscia* (Staphyleaceae) from the middle Eocene of Oregon, USA. *Tertiary Research* **9**: 59–66.
- Mantel N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**: 209–220.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403**: 53–58.
- Nei M. 1987. *Molecular evolutionary genetics*. New York: Columbia University Press.
- Nei M, Kumar S. 2000. *Molecular evolution and phylogenetics*. New York: Oxford University Press.
- Petit RJ, Aguinagalde I, Beaulieu JL, et al. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**: 1563–1565.
- 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**: 23–259.
- Piry S, Luikart G, Cornuet JM. 1999. BOTTLENECK: a computer program for detecting recent reductions in effective population size using allele frequency data. *Journal of Heredity* **90**: 502–503.
- Pons O, Petit RJ. 1996. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **144**: 1237–1245.
- Posada D, Crandall KA. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution* **16**: 37–45.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Qi XS, Chen C, Comes HP, et al. 2012. Molecular data and ecological niche modelling reveal a highly dynamic evolutionary history of the East Asian Tertiary relict *Cercidiphyllum* (Cercidiphyllaceae). *New Phytologist* **196**: 617–630.
- Qiu YX, Fu CX, Comes HP. 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.
- Rendell S, Ennos RA. 2003. Chloroplast DNA diversity of the dioecious European tree *Ilex aquifolium* L. (English holly). *Molecular Ecology* **12**: 2681–2688.
- Rice WR. 1989. Analysing tables of statistical tests. *Evolution* **43**: 223–225.
- Ribeiro RA, Lemos-Filho JP, Ramos ACS, Lovato MB. 2011. Phylogeography of the endangered rosewood *Dalbergia nigra* (Fabaceae): insights into the evolutionary history and conservation of the Brazilian Atlantic Forest. *Heredity* **106**: 46–57.
- Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**: 552–569.
- Rousset F. 2008. GENEPOP' 007: a complete reimplementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**: 103–106.
- Sang T, Crawford DJ, Stuessy TF. 1997. Chloroplast DNA phylogeny, reticulate evolution and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84**: 1120–1136.
- Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* **94**: 275–288.
- Swets JA. 1988. Measuring the accuracy of diagnostic systems. *Science* **240**: 1285–1293.
- Swofford DL. 2000. *PAUP*. Phylogenetic analysis using parsimony (* and other methods)*, Version 4. Sunderland, MA: Sinauer Associates.
- 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.
- Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**: 437–460.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725–2729.
- 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.
- Wang HW, Ge S. 2006. Phylogeography of the endangered *Cathaya argyrophylla* (Pinaceae) inferred from sequence variation of mitochondrial and nuclear DNA. *Molecular Ecology* **15**: 4109–4122.
- Wang J, Gao P, Kang M, Lowe AJ, Huang H. 2009. Refugia within refugia: the case study of a canopy tree *Eurycorymbus cavaleriei* in subtropical China. *Journal of Biogeography* **36**: 2156–2164.
- Wang JF, Gong X, Chiang YC, Kuroda C. 2013. Phylogenetic patterns and disjunct distribution in *Ligularia hodgsonii* Hook. (Asteraceae). *Journal of Biogeography* **40**: 1741–1754.
- Wang WT. 1992. On some distribution patterns and some migration routes found in the eastern Asiatic region. *Acta Phytotaxonomica Sinica* **30**: 1–24.
- Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences, USA* **84**: 9054–9058.
- Yao XH, Ye QG, Kang M, Huang HW. 2007. Microsatellite analysis reveals interpopulation differentiation and gene flow in the endangered tree *Changiosyrax dolichocarpa* (Styracaceae) with fragmented distribution in central China. *New Phytologist* **176**: 472–480.
- Ying TS, Zhang YL, Boufford DE. 1993. *The endemic genera of seed plants of China*. Beijing: Science Press.
- Yu G, Chen X, Ni J, et al. 2000. Palaeovegetation of China: a pollen data-based synthesis for the mid-Holocene and last glacial maximum. *Journal of Biogeography* **27**: 635–664.
- Zhang PF, Wang ZG, Liu YL, Tian H, Yao XH, Zhang JJ. 2013. Isolation and characterization of 11 polymorphic microsatellite markers in *Tapiscia sinensis* (Staphyleaceae). *Applications in Plant Science* **1**: 1300051.
- Zhou ZK, Momohara A. 2005. Fossil history of some endemic seed plants of East Asia and its phylogeographical significance. *Acta Botanica Yunnanica* **25**: 449–470.