Biochemical Systematics and Ecology 39 (2011) 787–796

Contents lists available at ScienceDirect

Biochemical Systematics and Ecology

journal homepage: www.elsevier.com/locate/biochemsyseco

Phylogeography of *Lagochilus ilicifolius* (Lamiaceae) in relation to Quaternary climatic oscillation and aridification in northern China

Hong-Hu Meng [a,b], Ming-Li Zhang [a,c,*]

[a] Key Laboratory of Biogeography and Bioresource in Arid Land, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China
[b] Graduate University, Chinese Academy of Sciences, Beijing 100049, China
[c] State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

ARTICLE INFO

Article history:
Received 25 May 2011
Accepted 16 July 2011
Available online 9 August 2011

Keywords:
Lagochilus ilicifolius
Phylogeography
Northern China
Aridification
Desert expansion

ABSTRACT

Using two chloroplast DNA intergenic spacers, a phylogeographical study of *Lagochilus ilicifolius* with 168 individuals from 14 populations was performed to assess geographical patterns and genetic variation in relation to past climate change. Population structure and history were inferred by nested clade analysis, neutrality tests, divergence time estimates, and mismatch distribution analysis. Twelve haplotypes were identified. Genetic differentiation among populations within groups was also elevated (F_{SC} = 0.772), suggesting restricted gene flow among populations within groups. Divergence time was at approximately 0.71 Ma, consistent with aridification and desert expansion during the middle Pleistocene. The Helan Mountains contained a large proportion of the haplotypes, which implies that the region may be the center of diversification for the species, whereas the Loess Plateau as a dispersal corridor for postglacial re-colonization northward. Climatic oscillation, aridification, and desert expansion within these regions have molded current distribution and biodiversity of *L. ilicifolius*.

1. Introduction

It is now well appreciated that the remarkable climatic fluctuations occurred during the Quaternary have strongly affected the distribution and genetic diversity of most present-day species, through cycles of fragmentation and range advance and retreat, especially in Northern Hemisphere (Hewitt, 2000). The uplift of the Qinghai-Tibetan Plateau (QTP) since the Cenozoic, one of the most important of recent geological events, has modified the global climate and influenced monsoon intensity (An et al., 2001). Continuing rise of this plateau during the Pleistocene enhanced aridity in the Asian interior and enlarged the range of deserts in northern China (Wu, 2002), and thus profoundly changed hydrology and climate of the region. Deserts that developed in the Quaternary appear to have played a significant role in determining the geographic distribution and evolutionary history of plant species.

To the present, most studies on phylogeography of plants in China have focused on the endangered or endemic species (e.g., Ge et al., 2011; Su et al., 2005; Wang and Ge, 2006), and have especially concentrated on species located in the Hengduan Mountains, on the southeast of the QTP. This area is referred to as the core area of the Himalayan hotspot, with one of the greatest concentrations of biodiversity in the world, due to its high level of species and genetic richness (Myers et al., 2000;

* Corresponding author. Key Laboratory of Biogeography and Bioresource in Arid Land, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China. Tel.: +86 991 7885515; fax: +86 991 7885320.
E-mail addresses: honghumeng@gmail.com (H.-H. Meng), zhangml@ibcas.ac.cn (M.-L. Zhang).

0305-1978/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bse.2011.07.015

Ying et al., 1993; Zhang et al., 2010). As a result, few, if any studies have been carried out regarding evolutionary history of species adapted to the arid lands distributed in regions of higher latitude in China. An interesting aspect of these arid areas is that the contraction-expansion dynamics of the species that inhabited them during the past climatic changes are expected to have been different from the dynamics of species in the QTP. Moreover, it seems surprising that, despite increasingly detailed geological reconstructions (Deng et al., 2006; Guo et al., 2002; Li, 1990; Rea et al., 1998; Wu et al., 1992; Yang et al., 2006), but few studies have attempted to document the effect of these events on the evolution of the northern China flora. Generally, detailed reconstruction of the evolutionary history of plant species has been hampered by the lack of fossil data in these regions. Fortunately, molecular evidence has provided an effective approach, independent of fossil information (Avise, 2000; Comes and Kadereit, 1998). This, combined with new analytical methods and recent paleoclimatic and paleogeologic studies, is providing important insight into the range dynamics of organisms during the Quaternary, because of the genetic signatures left in current populations by climatic oscillations (Emerson and Hewitt, 2005; Hewitt, 2004). Molecular markers have been used to study the evolutionary history of numerous species living at high elevations in the Northern Hemisphere (Hewitt, 2000), and chloroplast DNA (cpDNA) has been used in most studies of phylogeography of angiosperm plants, on the merits of it being nonrecombinant and maternally inherited (Cosacov et al., 2010; Lorenz-Lemke et al., 2010; Xu et al., 2010; Zhang et al., 2010). Most importantly, cpDNA is suitable for investigating the phylogeographical processes associated with pollen and seed dispersal, such as range expansion (Cruzan and Templeton, 2000), and the contribution of seed movement to total gene flow (Orive and Asmussen, 2000; Song et al., 2002), although cpDNA evolves slowly relative to other DNA types.

The natural vegetation of northern China is dominated by steppes, with a relatively homogeneous environment (Wu, 1980; Zhang et al., 2010). However, studies on phylogeography have most often been directed at woody plants (Chen et al., 2008; Tian et al., 2009; Wang and Ge, 2006; Xu et al., 2010), and have focused on a few endemics and characteristic components of the mountainous cool-temperate conifer and deciduous forests of this region. Herbaceous plants might have been more sensitive to Quaternary climatic oscillations than woody plants, because of their short life cycles, and would therefore be ideal candidates to illustrate the evolutionary history of the flora. However, there is some lack of knowledge about phylogeographical studies have ever been conducted on herbaceous plants of the northernmost part of China. Therefore, it would be of great interest to study the phylogeography of a herbaceous species with a distribution that covers both arid and montane areas, for a better understanding of the evolution and modern distribution of the vegetation of these regions.

Here, we used an integration of molecular data with paleoclimatic and paleogeologic methods to examine the phylogeographical patterns of *Lagochilus ilicifolius*, a perennial herb belonging to the family Lamiaceae, ranging from northern China to parts of Mongolia and Russia (Li and Hedge, 1994). As with most plants of this family, *L. ilicifolius* is entomophilous, up to 10–30 cm in height, and flowers from early June till August, setting fruits in September or October. In China, the species mainly covers Inner Mongolia, Ningxia, Shaanxi and Gansu, a distribution spanning desert steppe and grassland, as well as sandy areas, and thickets on gentle slopes in semidesert. In the study reported herein, we investigate variability in sequence of two chloroplast cpDNA spacers in samples of natural populations of *L. ilicifolius*. This analysis enables us to characterize (1) the genetic structure of *L. ilicifolius* populations in this area according to cpDNA sequence variation. (2) the phylogeographical pattern and demography, to determine how these were affected by desert expansion associated with climate changes of the late Quaternary.

2. Materials and methods

2.1. Plant materials

A total of 168 individuals of *L. ilicifolius* from 14 natural localities, covering most of the distribution range, were collected during the field investigation and collections (Table 1 and Fig. 1). The latitude, longitude and altitude of each sampling site were

Table 1
Sample locations and cpDNA haplotypes for 14 populations of *Lagochilus ilicifolius* in northern China.

Population location	Code	Latitude (N)	Longitude (E)	Altitude (m)	n	cpDNA haplotype
Sunite, Inner Mongolia	SNT	42°46′07″	112°44′09″	1096	12	H7
Alxa, Inner Mongolia	ALS	38°47′53″	105°37′14″	1429	12	H11, H12
Ruqigou, Ningxia	RQG	38°57′51″	106°14′24″	1437	12	H1
Helan, Ningxia	HL	38°43′05″	105°59′35″	1441	12	H10
Lingwu, Ningxia	DW	38°07′13″	106°50′36″	1391	12	H8, H9
Yanchi, Ningxia	YC	37°44′31″	107°16′02″	1419	12	H7
Tongxin, Ningxia	XMG	37°07′26″	106°26′27″	1521	12	H2, H5
Qingtongxia, Ningxia	NSS	37°47′43″	106°03′05″	1410	12	H1, H4
Hongquan, Ningxia	HQ	37°14′21″	105°12′47"	1796	12	H3
Dingbian, Shaanxi	MGQ	37°24′35″	107°44′32″	1712	12	H7
Hongliugou, Shaanxi	YJS	37°27′17″	107°18′52"	1491	12	H6, H7
Jingtai, Gansu	JT	37°14′51″	104°05′01″	1601	12	H1
Jingyuan, Gansu	QWS	36°53′56″	104°55′14″	1842	12	H2
Huanxian, Gansu	TSP	37°05′30″	106°47′41″	1559	12	H2

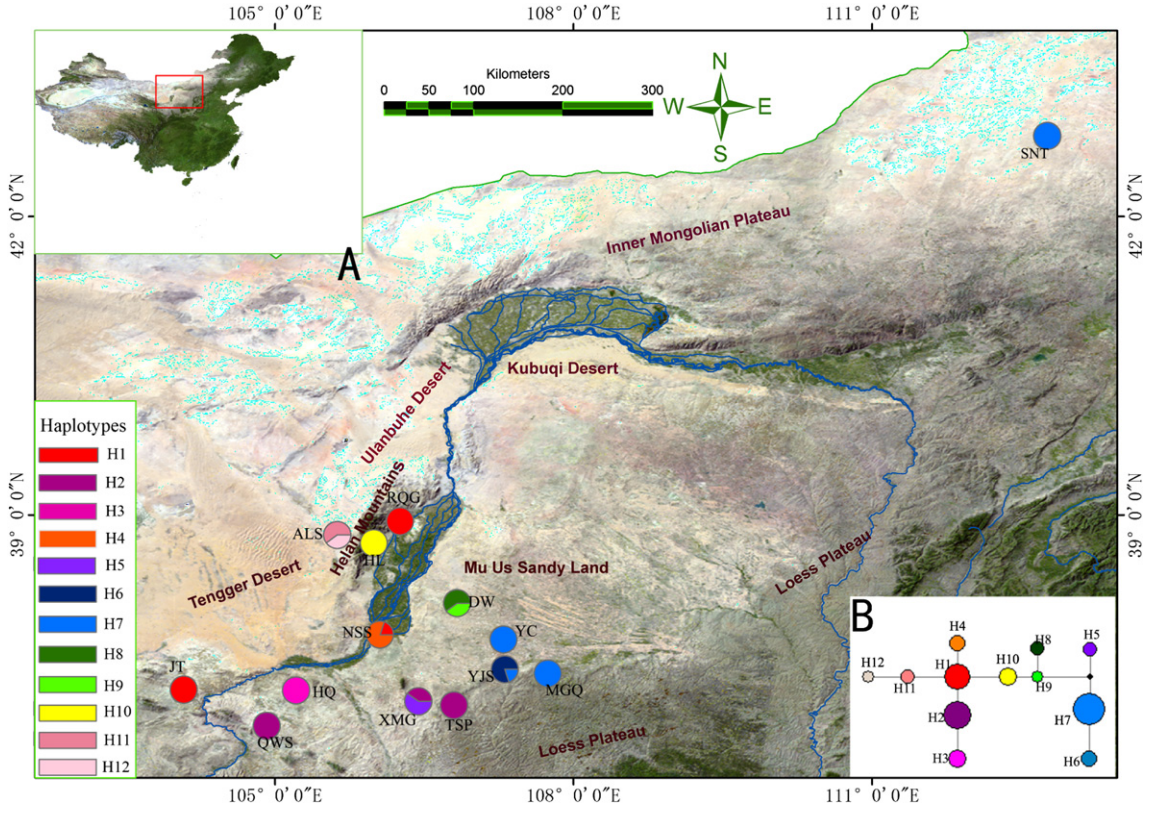

Fig. 1. Geographical distribution and genealogical relationships of the 12 cpDNA haplotypes recovered from *Lagochilus ilicifolius* populations from northern China. (A) Pie charts reflect the frequency of occurrence of each haplotype in each population. Haplotype colors correspond to those shown in panel (B). (B) Median-joining network linking the 12 haplotypes. Haplotypes are designated by numbers, and circle sizes are proportional to haplotype frequencies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

recorded using a global positioning system (GPS) receiver. Sampled individuals were separated by at least 50 m to avoid the collection of clones or close relatives. Young and healthy leaves were sampled randomly and quickly dried with silica gel in the field and stored frozen until DNA extraction. Voucher specimens from each sampling site were deposited in the Herbarium of Xinjiang Institute of Ecology and Geography, Chinese Academy of Science (XJBI). From the viewpoint of morphology, the genera *Lagochilus* and *Panzeria* are derived from the genus *Leonurus* (Wu et al., 2003; Wu and Li, 1982). Thus, *Panzeria alashanica* Kupr.(outgroup_1) and *Leonurus turkestanicus* V. Krecz. et Kupr. (outgroup_2) were used as outgroups in the analysis.

2.2. DNA extraction, PCR amplification, and sequencing

Total genomic DNA was isolated from silica gel-dried leaf tissue using a modified cetyl trimethyl ammonium bromide (CTAB) protocol as described by Cullings (Cullings, 1992). A preliminary screen for variation in cpDNA used seven available universal primers pairs, including one nuclear region, ITS (rRNA internal transcribed spacers), from representative samples of the species. However, except for the cpDNA spacers *trn*S-*trn*G (Hamilton, 1999) and *psb*A-*trn*H (Sang et al., 1997), which showed the highest levels of variation among the surveyed loci, and were befitting to amplify and sequence in the analysis, the other primers (*atp*B-*rbc*L, *rpo*B-*trn*C, *rps*12-*rp*l20, *rps*16, *trn*L-*trn*f and ITS) found no, or extremely low levels of, diversity.

Thus, large-scale screening of haplotype variation was then performed on all individuals and populations. Polymerase chain reactions (PCR) were carried out in a total volume of 30 μl, containing 1.5 μl of 10 × PCR reaction buffer (Takara, Japan), 1.5 μl of 25 mM $MgCl_2$, 1.2 μl of each primer (Shanghai Sangon, China) at 50 ng/μl, 2.4 μl of 2.5 mM dNTP solution in an equimolar ratio, 0.6 μl of *Taq* DNA-polymerase (5U/μl, Takara, Japan) and 2 μl of genomic DNA at 5 ng/μl. The protocol for amplification consisted of an initial hotstart at 95 °C for 2 min, followed by 30 cycle of denaturing at 94 °C for 30 s, annealing at 52 °C for 30 s, extension at 72 °C for 90 s, and a final extension at 72 °C for 10 min. PCR products were electrophoresed using a 0.8% agarose gel in a 0.5 × TAE (pH 8.3) buffer, then stained with EB (ethidium bromide) to confirm single products, and PCR products were purified using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA). These were sequenced using an ABI Prism 3770 Genetic Analyzer (Shanghai Sanggon Biological Engineering Technology & Service, Shanghai, China).

The DNA sequences were edited using SeqMan (Lasergene, DNASTAR Inc., Madison, Wisconsin, USA), and consensus sequences were obtained for each individual. Multiple sequence alignment was carried out in Clustal X 1.81 (Thompson et al., 1997), refined and adjusted manually. All sequences, representing all haplotypes generated from this study, were deposited in NCBI GenBank (accessions numbers JN375738–JN375749 for *trnS-trnG* and JN375724–JN375735 for *psbA-trnH*; outgroup of *P. alashanica* accessions numbers JN375751 for *trnS-trnG*, accessions numbers JN375737 for *psbA-trnH*; outgroup of *L. turkestanicus* accessions numbers JN375750 for *trnS-trnG*, accessions numbers JN375736 for *psbA-trnH*).

2.3. Sequence analysis

2.3.1. Molecular diversity and population genetic structure

Population genetic diversity based on cpDNA was quantified using indices of polymorphic sites (S), nucleotide diversity (π), and haplotype diversity (H_d), calculated using the DnaSP 4.0 program (Rozas et al., 2003). A haplotype network was constructed using Network 4.6.0.0, followed by the median joining (MJ) algorithm to resolve intermediate nodes (Bandelt et al., 1999).

Based on the resulting network, and in order to elucidate the possible history of populations, the statistical significance of the association of phylogenetic and geographical positions was assessed with a permutational contingency test (1000 resamples), using the program Geodis 2.5 (Posada et al., 2000). Two statistics were calculated: 1) clade distance, *Dc*, a measure of the geographical spread of a clade, and 2) nested clade distance, *Dn*, a measure of the geographical distribution of a clade relative to other clades in the same higher level nesting category. These measures of geographical distribution were used to infer historical processes (Templeton et al., 1995). Although nested clade analysis (NCA) has been recently challenged (Knowles and Maddison, 2002), and debates have been lasting, in the present study we have used this analysis to generate a phylogeographical hypothesis.

To investigate hierarchical levels of population structure, an analyses of molecular variance (AMOVA) was performed, considering genetic distances between haplotypes and their frequencies among the defined populations, using ARLEQUIN 3.1 (Excoffier et al., 2005), with statistical significance determined by 1000 permutations. This analysis was based on the geographical division of the population. Locations with high levels of genetic variation and unique haplotypes were examined as possible sites of refugia, or as a diversification center for the species, whereas locations with low levels of genetic variation were examined as possible sites of recent colonization (Taberlet and Cheddadi, 2002).

Phylogeographical signals in the defined regions were inferred by testing $N_{ST} > G_{ST}$ with 1000 replicates using Permut 2.0 (Burban et al., 1999; Pons and Petit, 1996). This program was used for a test of whether N_{ST} was larger than G_{ST}, which would be indicative of phylogeographical structure (i.e., a situation when closely related haplotypes are more often found in the same area than less closely related haplotypes).

2.3.2. Demographic history analyses

Tests of neutrality were investigated using Tajima's D (Tajima, 1989) as well as Fu and Li's D^* test (Fu and Li, 1993) to detect historical demographic expansions. Significant values of Tajima's D and Fu's F_S indicate an excess frequency of mutations relative to expectation under the standard neutral model (i.e., strict selective neutrality of variants, and constant population size). The significance of both values was calculated from 1000 simulated samples using a coalescent algorithm. In addition, a 'mismatch distribution' analysis was used to distinguish between models invoking past exponential growth versus historical population stasis, performed using the DnaSP program (Rozas and Rozas, 1999). We also calculated the raggedness index, r, which measures the smoothness of observed mismatch curves (Harpending, 1994), and tested for significant departure from unimodal. A multimodal distribution of differences between haplotypes is usually found in samples drawn from populations at demographic equilibrium, whereas the distribution is usually unimodal in populations having passed through a recent demographic expansion.

2.3.3. Estimate of divergence time

Based on substitution rates suggested for cpDNA spacers (Alsos et al., 2005) and estimates from a comprehensive study (Wolfe et al., 1987), at sites not under selective pressure, rates of divergence for cpDNA are of the order of 0.0011–0.0029 substitutions per synonymous site per Myr (0.22–0.58%). We therefore assumed substitution rates of 0.28% Myr^{-1} to estimate divergence times among haplotypes of the studied populations, in accordance with the substitution rate estimated for the *psb*A-*trn*H + *trn*S-*trn*G spacers in *Petunia* (Lorenz-Lemke et al., 2010). Neighbor-joining (NJ) analysis based on Kimura's (Kimura, 1980) 2-parameter distance was performed using the software MEGA 4.0 (Tamura et al., 2007). To evaluate clade support, 1000 replicates of bootstrap analysis (Felsenstein, 1985) were performed using fast heuristic search and TBR branch swapping. We used published nucleotide substitution rates because there are neither fossil records nor specific substitution rates available by which to calibrate a molecular clock. Although these estimates are provisional and should be interpreted with caution, they provide approximations that allow us to hypothesize possible scenarios under which lineages would have diverged.

3. Results

3.1. Population and phylogeographical analysis

The *psb*A-*trn*H alignment was 330 bp long and the *trn*S-*trn*G alignment was 678 bp. These were concatenated and treated as one sequence in all analyses. Average nucleotide frequency for the portion of cpDNA region sampled was 38.10% (A), 33.33%

(T), 9.52% (G), and 19.05% (C). The concatenated chloroplast spacer is A/T rich with an average content of 71.43%, which is consistent with the nucleotide composition of most noncoding spacers and pseudogenes, because of low functional constraint (Li, 1997). Based on the concatenated sequence, a total of 15 polymorphic sites (5 nucleotide substitutions and 10 indels) were detected (Table 2), and 12 haplotypes were identified. The haplotype frequencies in the 14 investigated populations of *L. ilicifolius* are listed in Table 1, not including intermediate haplotypes inferred from the analysis (Fig. 1 B).

Genetic diversity analysis of *L. ilicifolius* revealed high diversity ($h_T = 0.925 \pm 0.0374$, and $v_T = 0.988 \pm 0.1017$), and significant phylogeographical signal ($N_{ST} > G_{ST}$) between populations ($G_{ST} = 0.799 \pm 0.0724$ and $N_{ST} = 0.911 \pm 0.0608$). When G_{ST} was compared with the estimate of genetic differentiation taking into account haplotype similarities (N_{ST}), the difference between the two estimates was significant ($P = 0.003$). As estimated with the program DnaSP, high levels of haplotype diversity ($H_d = 0.8824$) and low levels of nucleotide diversity ($\pi = 0.0016$) were detected within the whole species.

The haplotype distribution is shown in Fig. 1 A, and phylogenetic relationships among haplotypes are presented in Fig. 1 B. The most frequent and widespread haplotypes (H1, H2 and H7) of the sampled populations were interspersed in Helan Mountains area, the Loess Plateau and Inner Mongolian plateau, and H7 was also distributed from the Loess Plateau to the Inner Mongolian plateau. H1 took the central position in the network, and H3 and H10 were private. H10 was restricted to the Helan Mountains, and H3 was found in the southwest population of the Helan Mountains, HQ. The haplotypes H1and H4 were shared in one population, NSS; H2, H5; H6, H7; H8, H9; H11, H12 were shared in populations XMG, YJS, DW, and ALS, respectively. Populations with the lowest haplotype diversity were found along the southeastern margin of Mu Us Sandy Land and along a northern strip of the Loess Plateau, ranging to the Inner Mongolian Plateau, where the majority of populations were monomorphic. The highest haplotype diversity was found in populations along the Helan Mountains and adjacent regions, from the southwest ridge (JT) to the north (RQG). The geographic distribution of all 12 haplotypes is very asymmetrical (Fig. 1 A).

Nested clade analysis (NCA) (Fig. 2) shows a significant relationship between genetic and geographical distributions in *L. ilicifolius*. The topologies of NCA are congruent with the topologies resulting from Network 4.6.0.0 (Fig. 1 B). Clade 3-1 included H1, H2, H3, H4, H11 and H12, Clade 3-2 included H5, H6 and H7, and Clade 3-3 included H8, H9 and H10.

The AMOVA revealed that only 17.56% of the variation was explained by differences among geographical groups, which was not significant, whereas inter- and intra- population differences explained 63.61% and 18.83% of the variation, respectively (Table 3). The F_{SC} value obtained by AMOVA also indicated high differentiation among the populations within groups ($F_{SC} = 0.772$, $P < 0.001$).

3.2. Demographic history analyses

Demographic analysis revealed evidence of contiguous range expansion in the total cladogram, particularly in populations along the edge of the Loess Plateau (Clade 3-2) in the species-level analysis. Restricted gene flow was found for populations along the ridge of the Helan Mountains, with isolation by distance, probably related to the complex geomorphology.

To test the hypothesis of population expansion of *L. ilicifolius*, we calculated the frequency distribution of pairwise nucleotide differences among individual haplotypes. The mismatch distribution for the species was clearly unimodal, and consistent with a model of sudden range expansion (Fig. 3). This was also supported by a low value of *r*, Harpending's measure of raggedness ($r = 0.0377$, $p = 0.5535$).

Neutrality tests were performed to determine whether the *trn*S-*trn*G and *psb*A-*trn*H loci were subject to selection or evolved neutrally, determined on the concatenated sequence. Significant *D* values can be due to factors such as bottlenecks

Table 2
Distributions of variable locus of each haplotype for *Lagochilus ilicifolius* in the *trn*S-*trn*G(678bp) and *psb*A-*trn*H(330bp) regions.

Nucleotide variable locus																				
*trn*S-*trn*G								*psb*A-*trn*H												
				1	1	3	4	7	7	7	8	9	9	9	9	9	9	9	9	9
	3	3	3	6	6	2	1	1	5	8	2	3	3	3	3	5	5	5	5	5
Haplotype	0	3	4	5	6	5	6	1	0	6	8	1	2	3	4	3	4	5	6	7
H1	A	–	–	A	T	T	G	C	–	C	C	C	A	T	T	C	T	A	A	I*
H2	A	–	–	A	T	T	G	C	–	C	A	C	A	T	T	C	T	A	A	I*
H3	A	–	–	A	C	T	G	C	–	C	A	C	A	T	T	C	T	A	A	I*
H4	A	–	–	A	T	T	G	C	–	T	C	C	A	T	T	C	T	A	A	I*
H5	–	T	A	–	T	–	A	T	–	C	A	C	A	T	T	–	–	–	–	–
H6	A	–	–	–	T	–	A	T	–	C	C	–	–	–	–	–	–	–	–	–
H7	A	–	–	–	T	–	A	T	–	C	C	–	–	–	–	–	–	–	–	–
H8	A	–	–	–	T	–	G	T	C	C	C	C	A	T	T	C	T	A	A	I*
H9	A	–	–	–	T	–	G	T	–	C	C	C	A	T	T	C	T	A	A	I*
H10	A	–	–	A	T	–	G	C	–	C	C	C	A	T	T	C	T	A	A	I*
H11	A	–	–	A	T	T	G	C	–	C	C	C	A	–	–	C	T	–	–	I*
H12	A	–	–	A	T	T	G	C	C	C	C	C	A	–	–	C	T	–	–	I*

–, refers to gap; I* refers to TAGTA.

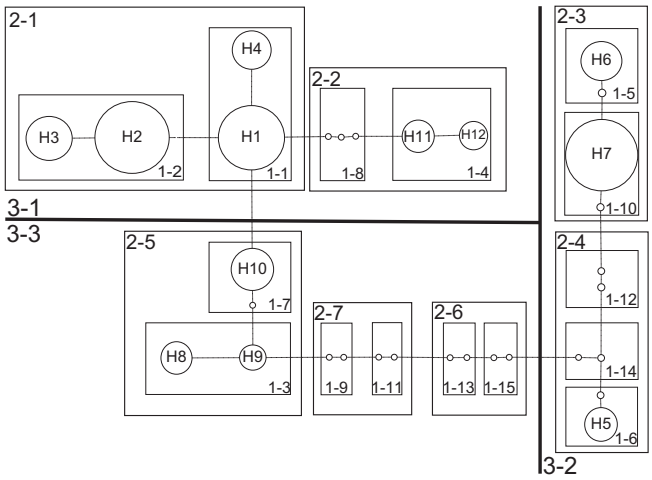

Fig. 2. Nested clade design based on 12 haplotypes of *Lagochilus ilicifolius,* haplotypes are designated by numbers, and missing intermediate haplotypes are indicated by open dots.

and selection (Tajima, 1989), and significant F^* values generally suggest recent demographic expansion (Fu and Li, 1993). Both values were investigated by Tajima's criterion ($D = 1.213$, $P > 0.10$), Fu and Li's tests ($D^* = 1.067$, $0.1 > P > 0.05$; $F^* = 1.308$, $0.1 > P > 0.05$).

3.3. Divergence time estimates

As the NJ and Bayesian methods yielded essentially identical topologies, only the NJ tree is presented here for high resolution. Divergence times estimated in *L. ilicifolius* for the supported clades are shown in Fig. 4. According to these estimates, the lineage of clade 3-1 began to diversify at approximately 0.71 Ma, whereas the lineage of clade 3-3 diverged at approximately 0.34 Ma. And the estimated time of diversification within the lineage of clade 3-2, haplotypes (H5, H6, and H7) along the Loess Plateau was at approximately 0.18 Ma. All of the clades originated almost simultaneously during the middle Pleistocene.

4. Discussion

4.1. Sequence heterogeneity and variation

In this study, the phylogeographical pattern of *L. ilicifolius* populations distributed in northern China was demonstrated, based on variability of the intergenic cpDNA spacer *trn*S-*trn*G and *psb*A-*trn*H. Base substitutions, hypervariable indels, microsatellites and mononucleotide repeats were detected in the cpDNA sequence (Table 2). These types of variation should have different mutation rates, and as a consequence could potentially be useful for inferring phylogeographical events that occurred at different time periods. As in many studies, the coding of length variants as multistate characters might have affected the phylogenetic analysis, as length variants and base substitutions were effectively treated as being equivalent (Muller, 2006). However, the procedure of haplotype network construction should effectively take into account the different mutation rates of these sequence variants (Baenfer et al., 2006). In the current study, we detected a strong correspondence between genealogy and geography. The structure of the haplotype network is strongly congruent with the topologies from the Neighbor-joining tree and NCA analysis (Fig. 1 B, Figs. 2 and 4), indicating that the data were appropriate for phylogenetic and phylogeographic analysis. In comparison with the Bayesian analysis (not shown), the Neighbor-joining tree provided higher resolution, particularly in the relationships of tip clades on the tree topology (Fig. 4).

Table 3
Analysis of molecular variance (AMOVA) for the 14 populations of *Lagochilus ilicifolius* based on cpDNA *trn*S-*trn*G and *psb*A-*trn*H.

Source of variation	d. f.[a]	s. s.[b]	v. c.[c]	p. v.[d]	Fixation Index
Among groups	2	6.971	0.08659 Va	17.56	F_{SC} : 0.77157, $p < 0.001$
Among populations within group	11	18.271	0.31364 Vb	63.61	F_{ST} : 0.81168, $p < 0.001$
Within population	56	5.200	0.09286 Vc	18.83	F_{CT} : 0.17562, $p < 0.001$
Total	69	30.443	0.49309		

[a] freedom.
[b] sum of squares.
[c] variance of component.
[d] percentage of variance.

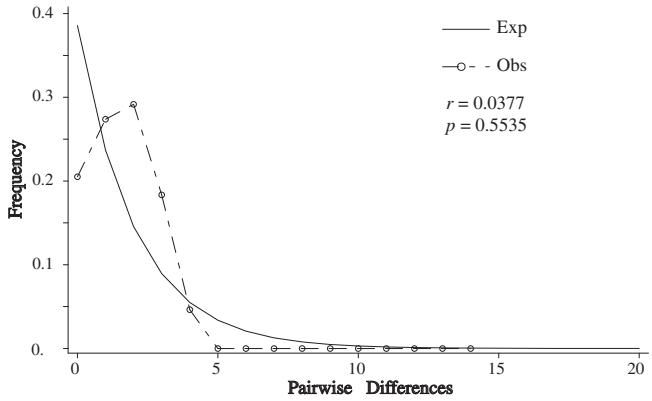

Fig. 3. The mismatch distributions curve of the haplotypes for *Lagochilus ilicifolius* populations in northern China. Expected values under expanding population and constant population models are indicated as solid and dotted lines, respectively (*r*, Ruggedness; *p*, probability).

Analysis of cpDNA intergenic spacers showed a high level of haplotype diversity and low levels of nucleotide diversity within *L. ilicifolius* populations. This is probably associated with particular ecological conditions and a long evolutionary history in the arid lands. Since the *trn*S-*trn*G and *psb*A-*trn*H sequences have relatively few functional constraints and nearly neutral evolution, mutations would have been retained to a considerable extent within each lineage, as has been observed with other noncoding spacers.

4.2. Phylogeographical structure

In this study, we investigated the phylogeographical pattern and population structure of *L. ilicifolius* in northern China. Genetic differentiation among populations within groups was high, as well as among groups (F_{SC} = 0.772), suggesting restricted gene flow. However, within population differentiation was low (F_{CT} = 0.176) (Table 3), probably due to low migration. Gene flow in seed plants may occur either via pollen dispersal prior to fertilization, or by seed transmission. Plants

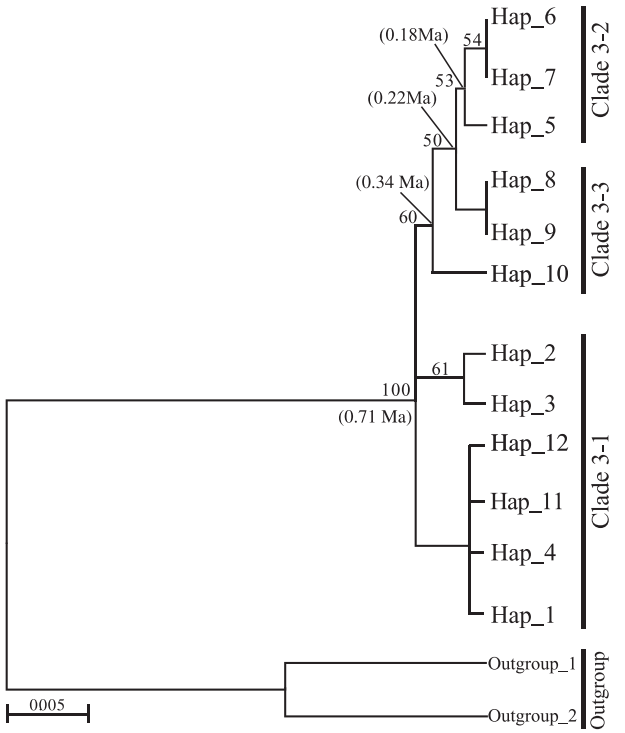

Fig. 4. Neighbor-joining tree of *Lagochilus ilicifolius* based on sequences of the *trn*S-*trn*G and *psb*A-*trn*H intergenic spacers of cpDNA. Numbers above branches are support values from bootstrap resampling, and numbers below branches or given for each node are the estimated divergence times in million years ago. The black bars on the right indicate the corresponding clade number in the nested clade phylogeographic analysis.

of *L. ilicifolius* are cross-pollinated, with a heterotypic flower structure, and as most other species of this family, are primarily pollinated by insects. This mating system should promote the exchange of genes within populations, increase effective population size, and reduce the impact of genetic drift. Pollen and associated gene movements are mediated by vectors that control the dispersal of pollen and genes (Levin and Kerster, 1969). However, the habitats of *L. ilicifolius* in northern China are deserts, sandy lands, and semidesert grasslands, where the arid environment, with heavy wind and lack of water during the blooming phase, would restrict the dispersal capacity of pollinators. In addition, seed dispersal is constrained by gravity, likely resulting in most seed transfer being confined to short distances.

The genealogy of cpDNA in *L. ilicifolius* recovered three main clades in the nested clade analysis (NCA) (Fig. 2), which are congruous with the Network (Fig. 1 B) and the NJ trees (Fig. 4). Most of the populations had a simple genetic composition, with only one haplotype (Fig. 1 A), giving a pattern of most of the genetic variation residing between populations. Overall, NCA provided further insights into the evolutionary history of *L. ilicifolius*, exhibiting a phylogeographical footprint consistent with contiguous range expansion, with the additional suggestion, in clade 3-1, of restricted gene flow with isolation by distance. A possible explanation is that the physiographical heterogeneity of the Helan Mountains area gave rise to geographical and, probably, ecological isolation, and these were responsible for population differentiation along the ridge and adjacent areas. Surrounding the mountains, on the north are the Ulanbuhe and Kubuqi Deserts, on the west and east are the Tengger Desert and Mu Us Sandy Land, respectively; and at the southeast is the Loess Plateau. The vegetation at the edge of the deserts (ALS, NSS and DW), is considered to be highly sensitive and vulnerable to the expansion of deserts resulting from climate change, which may have triggered species differentiation (H1, H4; H8, H9; H11, H12). Interestingly, the Loess Plateau, which presents such an apparently fragmented landform, with hundreds and thousands of hills and gullies, and the Inner Mongolian Plateau, which is considered the flattest plateau in China, are not enough to have served as geographical barriers to *L. ilicifolius*. Instead, populations in these regions (YC, YJS, MGQ and SNT) showed the common haplotype (H7), and so the Loess Plateau appears to have provided an ecological corridor for northward migration for the species during the interglacial phase and desert expansion.

4.3. Divergence times and possible past scenarios

Our results suggest that the initial divergence among *L. ilicifolius* lineages occurred during the middle Pleistocene (c. 0.71 Ma), before the onset of the last glacial maximum (LGM). Quaternary aridification, expansion of the Arctic ice-sheet driven by ongoing global cooling, and progressive uplift of the Himalayan–Tibetan complex during the period enhanced the summer monsoon and brought wetter climates to India and Southeast Asia. However, this moisture could not reach the Asian interior, because the uplifted Himalayan–Tibetan topography blocked airflow from the south (Guo et al., 2002). Meanwhile, uplift of the QTP strengthened the winter monsoon, causing additional mid-latitude to high-latitude Asian interior cooling (An et al., 2001; Rea et al., 1998). As a result, the interior of Asia became increasingly dry and cold during the late Quaternary (Li, 1990), resulting in the expansion of deserts in northern China (Wu et al., 1992; Yang et al., 2006), and ultimately affecting the biota within these regions. In brief, the divergence of all three clades, and including all the haplotypes, estimated at the middle Pleistocene, might have been simultaneous with this desert expansion and aridification, caused by the orogenesis of the Himalayas. These diversifications would have especially affected populations at the edge of deserts (ALS, NSS, and DW et al.). As a natural selective mechanism, ecological factors (e.g., temperature and moisture) might change the genetic structure after a long period of differentiation, reducing ancestral characteristics. This phenomenon is evident in the analysis, and may explain the only moderate frequency of the ancestral haplotype, H1. Most plant communities in these regions migrated southward during middle Pleistocene cold and arid cycles and recolonized their former habitats toward the north when the climate warmed. However, as mentioned, the Ulanbuhe and Kubuqi Deserts are at the north, might have been geographic barriers to the migration of *L. ilicifolius,* as well as the Tengger Desert in the west and Mu Us Sandy Land in the east, causing northward migration of *L. ilicifolius* to be funneled through the Loess Plateau.

Although the divergence time we estimated meshes with the other paleogeologic and paleoclimatic studies cited in the analysis, limitations in our abilities to estimate divergence times more accurately also limit the robustness of our inferences; these inferences should be accepted with caution and, as hypotheses, remain open for testing and refinement by future studies.

5. Conclusions

The phylogeography of *L. ilicifolius* reveals the influence of complex paleogeologic and paleoclimatic events contributing to diversification and evolutionary history that have produced current distributions. For this species, Quaternary climatic fluctuations, with consequent aridification and desert expansion, substantially impacted its diversification, distribution pattern and demography, especially the divergence event which occurred at approximately 0.71 Ma. There is evidence that the Loess Plateau served as a migratory corridor for *L. ilicifolius* in subsequent re-colonization events during the postglacial period. Interaction between desert development and aridification might have shaped genetic differentiation among the geographic units. Due to the limited migratory ability of pollinators, the species displayed restricted gene flow with a pattern of isolation by distance. Our findings demonstrate that deserts are very dynamic ecosystems, with expansion and vicariance, often provide opportunities for shifts in genetic signature and evolutionary trajectory among organisms of these regions.

Acknowledgments

Dr. Stewart C. Sanderson (Shrub Sciences Laboratory, USDA, Utah, USA) is gratefully acknowledged for his English improvement to the manuscript. We also thank the anonymous referees led to valuable improvements in the text. This study is financially supported by CAS Important Direction for Knowledge Innovation Project (No. KZCX2-EW-305), and Xinjiang Institute of Ecology and Geography, Chinese Academy of Science.

References

Alsos, I., Engelskjon, T., Gielly, L., Taberlet, P., Brochmann, C., 2005. Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species. Mol. Ecol. 14, 2739–2753.
An, Z.S., Kutzbach, J.E., Prell, W.L., Porter, S.C., 2001. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. Nature 411, 62–66.
Avise, J.C., 2000. Phylogeography: the History and Formation of Species. Harvard University Press, Cambridge, MA.
Baenfer, G., Moog, U., Fiala, B., Mohamed, M., Weising, K., Blattner, F.R., 2006. A chloroplast genealogy of myrmecophytic *Macaranga* species (Euphorbiaceae) in Southeast Asia reveals hybridization, vicariance and long-distance dispersals. Mol. Ecol. 15, 4409–4424.
Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.
Burban, C., Petit, R., Carcreff, E., Jactel, H., 1999. Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. Mol. Ecol. 8, 1593–1602.
Chen, K.M., Abbott, R.J., Milne, R.I., Tian, X., Liu, J.Q., 2008. Phylogeography of *Pinus tabulaeformis* Carr. (Pinaceae), a dominant species of coniferous forest in northern China. Mol. Ecol. 17, 4276–4288.
Comes, H.P., Kadereit, J.W., 1998. The effect of Quaternary climatic changes on plant distribution and evolution. Trends Plant Sci. 3, 432–438.
Cosacov, A., Sersic, A.N., Sosa, V., Johnson, L.A., Cocucci, A.A., 2010. Multiple periglacial refugia in the Patagonian steppe and pos-glacial colonization of the Andes: the phylogeography of Calceolaria polyrhiza. J. Biogeogr 37, 1463–1477.
Cruzan, M.B., Templeton, A.R., 2000. Paleoecology and coalescence: phylogeographic analysis of hypotheses from the fossil record. Trends Ecol. Evol. 15, 491–496.
Cullings, K., 1992. Design and testing of a plan-specific PCR primer for ecological and evolutionary studies. Mol. Ecol. 1, 233–240.
Deng, C.L., Shaw, J., Liu, Q.S., Pan, Y.X., Zhu, R.X., 2006. Mineral magnetic variation of the Jingbian loess/paleosol sequence in the northern Loess Plateau of China: implications for Quaternary development of Asian aridification and cooling. Earth. Planet. Sci. Lett. 241, 248–259.
Emerson, B.C., Hewitt, G.M., 2005. Phylogeography. Curr. Biol. 15, 367–371.
Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform. Online 1, 47–50.
Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.
Fu, Y.X., Li, W.H., 1993. Statistical tests of neutrality of mutations. Genetics 133, 693–709.
Ge, X.J., Hwang, C.C., Liu, Z.H., Huang, C.C., Huang, W.H., Hung, K.H., Wang, W.K., Chiang, T.Y., 2011. Conservation genetics and phylogeography of endangered and endemic shrub *Tetraena mongolica* (Zygophyllaceae) in Inner Mongolia, China. BMC Genet. 12, 1–12.
Guo, Z.T., Ruddiman, W.F., Hao, Q.Z., Wu, H.B., Qiao, Y.S., Zhu, R.X., Peng, S.Z., Wei, J.J., Yuan, B.Y., Liu, T.S., 2002. Onset of Asian desertification by 22 Myr ago inferred from loess deposits in China. Nature 416, 159–163.
Hamilton, M., 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol. Ecol. 8, 521–523.
Harpending, H., 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum. Biol. 66, 591–600.
Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907–913.
Hewitt, G., 2004. Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society of London. Ser. B. Bio. Sci. 359, 183–195.
Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.
Knowles, L.L., Maddison, W.P., 2002. Statistical phylogeography. Mol. Ecol. 11, 2623–2635.
Levin, D.A., Kerster, H.W., 1969. The dependence of bee-mediated pollen and gene dispersal upon plant density. Evolution 23, 560–571.
Li, H.W., Hedge, I., 1994. Lamiaceae. In: Wu, Z.Y., Raven, P.H. (Eds.), Flora of China, vol. 17. Science Press/St. Louis: Missouri Botanical Garden Press, Beijing, pp. 50–299.
Li, J.J., 1990. The patterns of enviromental changes since late Pleistocene in Northwestern China. Quatern. Sci. 3, 197–204.
Li, W.H., 1997. Molecular Evolution Sinauer. Sunderland Press (Mass).
Lorenz-Lemke, A.P., Togni, P.D., Mäder, G., Kriedt, R.A., Stehmann, J.R., Salzano, F.M., Bonatto, S.L., Freitas, L.B., 2010. Diversification of plant species in a subtropical region of eastern South American highlands: a phylogeographic perspective on native *Petunia* (Solanaceae). Mol. Ecol. 19, 5240–5251.
Muller, K., 2006. Incorporating information from length-mutational events into phylogenetic analysis. Mol. Phylogenet. Evol. 38, 667–676.
Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.
Orive, M.E., Asmussen, M.A., 2000. The effects of pollen and seed migration on nuclear-dicytoplasmic systems. II. A new method for estimating plant gene flow from joint nuclear-cytoplasmic data. Genetics 155, 833–854.
Pons, O., Petit, R., 1996. Measuring and testing genetic differentiation with ordered versus unordered alleles. Genetics 144, 1237–1245.
Posada, D., Crandall, K., Templeton, A., 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Mol. Ecol. 9, 487–488.
Rea, D.K., Snoeckx, H., Joseph, L.H., 1998. Late Cenozoic eolian deposition in the North Pacific: asian drying, Tibetan uplift, and cooling of the northern hemisphere. Paleoceanography 13, 215–224.
Rozas, J., Rozas, R., 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15, 174–175.
Rozas, J., Sachez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.
Sang, T., Crawford, D.J., Stuessy, T.F., 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am. J. Bot. 84, 1120–1136.
Song, B.H., Wang, X.Q., Wang, X.R., Sun, L.J., Hong, D.Y., Peng, P.H., 2002. Maternal lineages of *Pinus densata*, a diploid hybrid. Mol. Ecol. 11, 1057–1063.
Su, Y.J., Wang, T., Zheng, B., Jiang, Y., Chen, G.P., Ouyang, P.Y., Sun, Y.F., 2005. Genetic differentiation of relictual populations of *Alsophila spinulosa* in southern China inferred from cpDNA *trn*L-F noncoding sequences. Mol. Phylogenet. Evol. 34, 323–333.
Taberlet, P., Cheddadi, R., 2002. Quaternary refugia and persistence of biodiversity. Science 297, 2009–2010.
Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.
Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.
Templeton, A.R., Routman, E., Phillips, C.A., 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, Ambystoma tigrinum. Genetics 140, 767–782.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic. Acids Res. 25, 4876–4882.
Tian, B., Liu, R.R., Wang, L.Y., Qiu, Q., Chen, K.M., Liu, J.Q., 2009. Phylogeographic analyses suggest that a deciduous species (*Ostryopsis davidiana* Decne., Betulaceae) survived in northern China during the Last Glacial Maximum. J. Biogeogr. 36, 2148–2155.
Wang, H.W., Ge, S., 2006. Phylogeography of the endangered *Cathaya argyrophylla* (Pinaceae) inferred from sequence variation of mitochondrial and nuclear DNA. Mol. Ecol. 15, 4109–4122.
Wolfe, K., Li, W.H., Sharp, P.M., 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. P. Natl Acad Sci, USA, 84, 9054–9058.
Wu, C.Y., 1980. Vegetation of China. Science Press, Beijing.
Wu, C.Y., Lu, A.M., Tang, Y.C., Chen, Z.D., Li, D.Z., 2003. The Families and Genera of Angiosperms in China: a Comprehensive Analysis. Science Press, Beijing.
Wu, C.Y., Li, H.W., 1982. On the evolution and distribution in Labiatae. Acta Bot. Yunnan. 4, 97–118.
Wu, G.J., 2002. The maximum glaciation and desert expansion in China during MIS16. J. Glaciol. Geocryol 24, 544–549.
Wu, X., Wang, F., An, Z.S., 1992. The phase and height of Tibetan plateau rise in late Cenozoic [A]. In: Liu, D.S., An, Z.S. (Eds.), Loess·Quaternary Geology·Global Changes, vol 3 [C]. Scientific Publishing House, Beijing, pp. 1–13.
Xu, T., Abbott, R.J., Milne, R.I., Mao, K.S., Du, F.K., Wu, G.L., Ciren, Z.X., Miehe, G., Liu, J.Q., 2010. Phylogeography and allopatric divergence of cypress species (*Cupressus* L.) in the Qinghai-Tibetan Plateau and adjacent regions. BMC Evol. Biol. 10, 194–203.
Yang, D., Fang, X.M., Dong, G.R., Peng, Z.C., Li, J.J., 2006. Aeolian deposit evidence for formation and evolution of the Tengger Desert in the north of China since early Pleistocene. Marine Geol. Quatern. Geol. 26, 93–100.
Ying, T.S., Zhang, Y.L., Boufford, D.E., 1993. The Endemic Genera of Seed Plants of China. Science Press, Beijing.
Zhang, Y.H., Volis, S., Sun, H., 2010. Chloroplast phylogeny and phylogeography of *Stellera chamaejasme* on the Qinghai-Tibet Plateau and in adjacent regions. Mol. Phylogenet. Evol. 57, 1162–1172.