Molecular Phylogenetics and Evolution 68 (2013) 398–409

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

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

Diversification of plant species in arid Northwest China: Species-level phylogeographical history of *Lagochilus* Bunge ex Bentham (Lamiaceae)

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] University of Chinese Academy of Sciences, Beijing 100049, China
[c] Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

ARTICLE INFO

Article history:
Received 3 August 2011
Revised 16 April 2013
Accepted 16 April 2013
Available online 27 April 2013

Keywords:
Phylogeography
Diversification
Lagochilus
Aridification
Northwest China

ABSTRACT

Lagochilus occurs in the arid zones across temperate steppe and desert regions of Northwest China. Cooling with strong dessication in the Pleistocene, along with rapid uplift of mountain ranges peripheral to the Qinghai-Tibet Plateau, appear to have had major impacts on the genetic structure of the flora. To understand the evolutionary history of *Lagochilus* and the divergence related to these past shifts of habitats among these regions, we sequenced the plastid intergenic spacers, *psb*A-*trn*H and *trn*S-*trn*G from populations throughout the known distributions of ten species of the genus. We investigated species-level phylogeographical patterns within *Lagochilus.* Phylogenetic trees were constructed using Neighbor-joining and Bayesian inference. The divergence times of major lineages were estimated with BEAST and IMa. Genetic structure and demographic history were inferred by AMOVA, neutrality tests, mismatch distribution, and Bayesian skyline plot analyses. The results showed that most chloroplast haplotypes were species-specific, and that the phylogeny of *Lagochilus* is geographically structured. The estimated Bayesian chronology and IMa suggested that the main divergence events for species between major eastern and western portions of the Chinese desert occurred at the Plio-/Pleistocene boundary (*ca.* 2.1–2.8 Ma ago), and likely coinciding with the formation of these deserts in Northwest China. The regional demographic expansions, in the western region at *ca.* 0.39 Ma, and in the eastern at *ca.* 0.06 Ma, or across all regions at *ca.* 0.26 Ma, showed the response to aridification accompanied by cooling of the Pleistocene sharply increased aridity in the Chinese deserts, which reflects a major influence of geologic and climatic events on the evolution of species of *Lagochilus*. We suggest that diversification is most likely the result of the past fragmentation due to aridification; the expansion of the range of species along with the deserts was an adaptation to dry and cold environments during the Quaternary.

1. Introduction

Global climate fluctuations, in particular the remarkable climatic oscillations during the Quaternary, instigated repeated cycles of habitat contraction–expansion and latitudinal-altitudinal shifts of species' distributions. These cycles have driven species diversification, as evidenced by phylogeographical studies of many plants and animals, especially in temperate zones of the Northern Hemisphere (Hewitt, 2000; Petit et al., 2003). Phylogeography is a relatively new discipline that examines the spatial arrangements of genetic lineages, and infers phylogeographical patterns, as a powerful tool for investigating processes that determine genetic composition, especially within and among related species (Avise, 2000, 2009). This can untangle historical changes at various spatial and temporal scales in the patterns of gene flow, isolation, demographic expansion, and secondary contact among divergent populations (Hewitt, 2001; Schaal and Olsen, 2000).

Currently, most plant phylogeographical studies in China have focused on the Sino-Japanese Floristic Region of East Asia, which harbors the largest amount of diversity among the world's temperate regions (Myers et al., 2000; Ying et al., 1993), and was the most important glacial refuge for its Tertiary representatives ('relics') throughout Quaternary ice-age cycles (Qiu et al., 2011). Studies have investigated endangered or endemic species, especially concentrated in the areas of the Hengduan Mountains and Qinghai-Tibet Plateau (QTP) (e.g., Cun and Wang, 2010; Jia et al., 2011; Li et al., 2010, 2011; Zhang et al., 2011), which is referred to as the core of the Himalayan hotspot, one of the greatest concentrations of biodiversity in the world, due to its high level of species and generic richness (Myers et al., 2000; Ying et al., 1993).

* Corresponding author. Address: Key Laboratory of Biogeography and Bioresource in Arid Land, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, 818# South Beijing Road, Urumqi 830011, Xinjiang, China. Fax: +86 991 7823151.

E-mail addresses: honghumeng@gmail.com (H.-H. Meng), zhangml@ibcas.ac.cn (M.-L. Zhang).

1055-7903/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2013.04.012

Comparatively, little is known about the effect of Quaternary climatic oscillations on the species of arid Northwest China, although there have been an increasing number of phylogeographical studies that address this issue or focus on other sections of the country. Therefore, few studies have been undertaken concerning the evolutionary history of arid-adapted plant species in these regions. An interesting aspect of arid zones is that the contraction–expansion dynamics of the species that have inhabited them during the past climatic change cycles are expected to differ from the dynamics of species in other areas, relating to the aridity of glacial stages rather than to the presence of ice.

By the early Miocene, central parts of the QTP may have been uplifted to present heights (Wang et al., 2008; Wu et al., 2008), which compelled the retreat of the Tethys Ocean from Central Asia and established the beginning of an arid climate (Guo et al., 2002, 2008; Sun et al., 2010). The size of the QTP was augmented through time by successive uplift of peripheral areas, with probably increasing climatic effects (Clark et al., 2005; Fang et al., 2005; Lu et al., 2004). The cooling of the Pleistocene sharply increased aridity in the Chinese deserts (Ding et al., 2005; Fang et al., 2002). Rapid orogenic movements continued during the Pleistocene in the Tianshan and other mountain ranges related to the QTP which caused local rain shadows and promoted isolation and divergence (Sun and Zhang, 2009).

Over a lengthy period from mid-Tertiary onwards, these changes in and around the QTP made the climate of Northwest China increasingly dry due to the obstruction of sea breezes from the Indian Ocean, and because the East Asian monsoon from the Pacific could not reach so far inland. Consequently, lack of precipitation accelerated the formation of sandy desert, Gobi (rocky desert), and loess (Guo et al., 2002). It notably led to expansion of the Taklimakan Desert of the Tarim Basin, the Gurbantunggut Desert of the Junggar Basin, and Badain Jaran–Tengger deserts to the north of the Hexi Corridor. Although the formation of these deserts had begun much earlier, they increased in aridity and expanded dramatically during the Pleistocene (Yang et al., 2004, 2011). Computer simulations also have suggested that expansion of deserts in China during the Quaternary glacial stages was related to a decrease in the strength of the horizontal winds, enhanced atmospheric subsidence, and reduced moisture transport (Bush et al., 2004). These processes disrupted the migration of various biological groups across arid zones, in comparison to other regions of the world at the same latitude (Tingru, 1983; Zhang et al., 2000). The Pleistocene aridification of Northwest China as a whole and expansion of deserts on large scales (Fang et al., 2002; Sun, 2002b; Yang et al., 2004, 2011), profoundly changed the hydrology and climate of regions as well as playing a significant role in determining the local geographic distribution and evolutionary history of plant species. So far, a growing body of studies, based on pollen cores, fossils, moraine and deposition of loess, have begun to elucidate the possible roles of geology, multiple glaciations and climatic oscillations in shaping the current geo-ecological system occurring across the arid zones of Northwest China (Guan et al., 2011; Owen et al., 2005; Sun, 2002b; Sun and Zhang, 2009; Wu et al., 2002). However, the diversification and demographic history of plant species that span the arid zones in Northwest China are still poorly understood.

The genus *Lagochilus* is highly drought-tolerant, and is also considered as a typical montane plant, found across arid Northwest China. It provides an excellent system for examining the potential influence of geologic and climatic effects on range fluctuations and diversification of species across the temperate steppe and desert regions. *Lagochilus* was hypothesized to have a differentiation center along the coasts of Tethys Ocean (Wu et al., 2003; Wu and Li, 1982), but this did not appear probable because the dessication of interior Asia apparently occurred near the Oligocene-Meiocene boundary (Guo et al., 2002), much earlier than a plausible time of origin for the genus. *Lagochilus* contains 11 described species in China, and most occur in the mountain ranges of the arid zones (Li and Hedge, 1994). Two endemic species, *L. lanatonodus* and *L. xinjiangensis* occur in the Tianshan Mountains. *L. ilicifolius* 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. One dominant species, *L. diacanthophyllus* and other local species, *L. hirtus*, *L. bungei*, and *L. macrodontus* are mainly distributed throughout the Altay Mountains, whereas the other species, *L. grandiflorus*, *L. platyacanthus* and *L. kaschgaricus* occur in the Ili (Yili) Valley and the Karakoram Mountains. Most interesting, *L. ilicifolius* is disjunctive with the other species in China. *Lagochilus* species are entomophilous, perennial, herbaceous plants, and occur in niche habitats along mountains and other sandy areas or edges of deserts. The dispersal capabilities are limited; pollination occurs primarily via insects, while seed dispersal is by gravity. As mentioned, the genus (or perhaps its progenitors) is thought to have evolved from ancient seaside habitats, and most likely underwent rapid speciation during the Quaternary (Wu and Li, 1982). Therefore, understanding the phylogeographical history of *Lagochilus* will provide insights into the diversification of desert plants triggered by the past geologic and/or climatic changes in arid Northwest China.

Molecular markers have been used to study the evolutionary history of many species around the world. Molecular evidence has provided an effective approach, independent of fossil information when detailed reconstruction of the evolutionary process of plant species has been hampered by lack of fossil data, to untangle the evolutionary history of species in phylogeographical studies (Avise, 2000; Comes and Kadereit, 1998). Chloroplast DNA (cpDNA) is maternally inherited in most angiosperm plants and is commonly considered a single, nonrecombinant unit of inheritance (Hewitt, 2000; Liu et al., 2009; Lorenz-Lemke et al., 2010). Most significantly, it is suitable for investigating the phylogeographical processes associated with 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). Genetic variation in the chloroplast genome is often geographically structured, and so it is a useful molecule to study recent dispersal in plants (Petit et al., 1997). Especially using cpDNA to estimate age in general for the problems of recombination and concerted evolution in nuclear markers could be avoided, although phylogeographical inferences based solely on organelle genes may not necessarily be representative of overall genetic aspects of the organisms.

Here, we infer the gene tree, the timing and extent of the range fluctuations from cpDNA variation in *Lagochilus* plants as well as the paleoclimatic and paleogeologic records in arid Northwest China to determine species-level phylogeographical history. Specifically, the analyses enabled us to test (1) the phylogenetic relationships between them and the times at which they diverged; (2) the phylogeographical patterns and demography, in order to determine how they were affected by cycles of range expansion in association with climatic changes during the Quaternary.

2. Materials and methods

2.1. Taxon and sample collection

For the cpDNA survey, a total of 386 individuals were sampled at 35 localities representing 10 species (Table 1, Fig. 1) throughout the geographic distribution of *Lagochilus* in China. Morphological data provides evidence that the genera *Lagochilus* and *Panzeria* are both derived from *Leonurus* (Wu et al., 2003; Wu and Li,

Table 1
Details of the *Lagochilus* populations used in the study, sample sizes and cpDNA haplotypes observed. The abbreviations of IM, NX, SX, GS and XJ denote Inner Mongolia, Ningxia, Shaanxi, Gansu and Xinjiang, respectively.

Species	Collection site and code	Geographical coordinates, altitude	*n*	cpDNA haplotype
L. ilicifolius	01. Sonid, IM/SNT	42°46′07″N/112°44′09″E,1096m	12	H7
	02. Alxa, IM/ALS	38°47′53″N/105°37′14″E,1429m	12	H11, H12
	03. Ruqigou, NX/RQG	38°57′51″N/106°14′24″E,1437m	12	H1
	04. Helan Mts, NX/HL	38°43′05″N/105°59′35″E,1441m	12	H10
	05. Lingwu, NX/DW	38°07′13″N/106°50′36″E,1391m	12	H8, H9
	06.Yanchi, NX/YC	37°44′31″N/107°16′02″E,1419m	12	H7
	07. Tongxin, NX/XMG	37°07′26″N/106°26′27″E,1521m	12	H2, H5
	08. Qingtongxia, NX/NSS	37°47′43″N/106°03′05″E,1410m	12	H1, H4
	09. Hongquan, NX/HQ	37°14′21″N/105°12′47″E,1796m	12	H3
	10. Dingbian, SX/MGQ	37°24′35″N/107°44′32″E,1712m	12	H7
	11. Hongliugou, SX/YJS	37°27′17″N/107°18′52″E,1491m	12	H6, H7
	12. Jingtai, GS/JT	37°14′51″N/104°05′01″E,1601m	12	H1
	13. Jingyuan, GS/QWS	36°53′56″N 104°55′14″E,1842m	12	H2
	14. Huanxian, GS/TSP	37°05′30″N/106°47′41″E,1559m	12	H2
L. lanatonodus	15. Urumqi, XJ/WLB	43°37′45″N/87°43′44″E,1121m	12	H22, H25
	16. Barkol, XJ/BLK	43°33′18″N/93°03′57″E,2271m	12	H22, H24
	17. Yiwu, XJ/YW	43°15′32″N/94°41′55″E,1838m	10	H22, H23
	18. Sayram lake, XJ/SLM	44°34′03″N/81°21′43″E,2221m	10	H22, H25, H26
	19. Hejing, XJ/BLT	42°45′43″N/86°19′31″E,1840m	13	H22, H26
L. diacanthophyllus	20. Qinghe, XJ/QH	46°41′15″N/90°23′08″E,1482m	11	H13, H14, H15
	21. Fuyun, XJ/FY	47°03′21″N/89°30′09″E,1211m	12	H13, H15, H16
	22. Altay, XJ/ALT	47°51′20″N/88°05′41″E,1054m	12	H13, H15, H17
	23. Jeminay, XJ/JMN	47°23′48N/85°56′12″E,1284m	12	H18,
	24. Hoboksar, XJ/HF	46°48′42″N/85°48′47″E,1275m	10	H13, H18, H19
	25. Wenquan, XJ/WQ	45°01′04″N/81°00′46″E,1421m	12	H18, H20, H21
	26. Bortala, XJ/BL	44°58′03″N/81°52′32″E, 824m	10	H18, H20
L. grandiflorus	27. Tekesi, XJ/TKS	43°09′17″N/81°57′19″E,1449m	11	H13, H28
L. platyacanthus	28. Wuqia, XJ/JG	39°49′04″N/74°06′39″E,2730m	12	H13, H32
L. kaschgaricus	29. Artux, XJ/ATS	40°00′39″N/76°08′20″E,1889m	11	H33, H34
	30. Wuqia, XJ/BY	39°49′08″N/75°37′16″E,2045m	13	H34
L. xinjiangensis	31. Urumqi, XJ/CWP	43°32′26″N/88°03′44″E,1323m	6	H22, H27
L. macrodontus	32. Karamay, XJ/KS	45°39′47″N/84°51′44″E, 679m	13	H13, H31
L. hirtus	33. Fuhai, XJ/FH	46°44′30″N/88°08′39″E, 736m	4	H29
	34. Hoboksar, XJ/HF	46°46′22″N/85°45′11″E,1265m	8	H29
L. bungei	35. Fuyun, XJ/FY	46°59′12″N/89°46′20″E,1351m	4	H30

1982). Thus, *Panzeria alashanica* Kupr. and *Leonurus turkestanicus* V. Krecz. et Kupr. were used as outgroups in the analyses. Species were identified using the key in *Flora of China* (Li and Hedge, 1994). The latitude, longitude and altitude of each sampling site were recorded using GPS. 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 in silica gel in the field then subsequently stored frozen until DNA was extracted. Voucher specimens from each sampling site, and the outgroup taxa, were deposited in the XJBI (Herbarium of Xinjiang Institute of Ecology and Geography, Chinese Academy of Science, Urumqi, Xinjiang).

2.2. Laboratory protocols

Total genomic DNA was extracted from silica-gel-dried leaves tissue using a modified cetyltrimethylammonium bromide (CTAB) protocol (Cullings, 1992). Preliminary screening of the genetypes of DNA sample were scored using 7 cpDNA segments (*atp*B-*rbc*L, *rpo*B-*trn*C, *rps*12-*rpl*20, *rps*16, *trn*L-*trn*F, *trn*S-*trn*G and *psb*A-*trn*H) and 1 nuclear gene segment ITS (rRNA internal transcribed spacers) from representative samples of the species. We found that only the cpDNA intergenic spacers *trn*S-*trn*G (Hamilton, 1999) and *psb*A-*trn*H (Sang et al., 1997) showed higher levels of variation among the surveyed loci, and were befitting to amplify and sequence in the analyses, however, the other primers found no, or extremely low levels of diversity.

Thus, large-scale screening of haplotypes variation was then performed on all individuals' DNA samples in populations. Polymerase chain reactions (PCR) were performed in a total volume of 30 μl reactions. The PCR mixtures contained of 1.5 μl of 10 × PCR reaction buffer, 1.5 μl of 25 mM $MgCl_2$, 1.2 μl of each primer at 50 ng/μl, 2.4 μl of 2.5 mM dNTP solution in an equimolar ratio, 0.6 μl of *Taq* DNA-polymerase and 2 μl of genomic DNA at 5 ng/μl. The PCR amplifications were performed as follows: an initial denaturation step 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 examined with gel electrophoresis using a 0.8% agarose gel in a 0.5 × TAE (pH 8.3) buffer, and then stained with EB (ethidium bromide) to confirm single products. PCR products were purified using the QIAquick Gel Extraction Kit. The same primers were used for sequencing in Sangon Biotech (Shanghai) Co., Ltd., China.

The DNA sequences were edited using SeqMan (Lasergene, DNASTAR Inc., Madison, Wisconsin, USA), and consensus sequences were obtained from each individual. Multiple sequence alignment was carried out in Clustal X 1.81 (Thompson et al., 1997), and refined and adjusted manually.

2.3. Data analyses

2.3.1. Sequence diversity

Basic sequence statistics and molecular indices, such as number of haplotypes, polymorphic sites, haplotype diversities (h) and nucleotide (π) diversities as well as the analysis of molecular variance (AMOVA) among the major regions were obtained with ARLEQUIN version 3.1 (Excoffier et al., 2005). The relationships among haplotypes from all populations were also estimated with the Network version 4.6.0.0 program using the Median-Joining

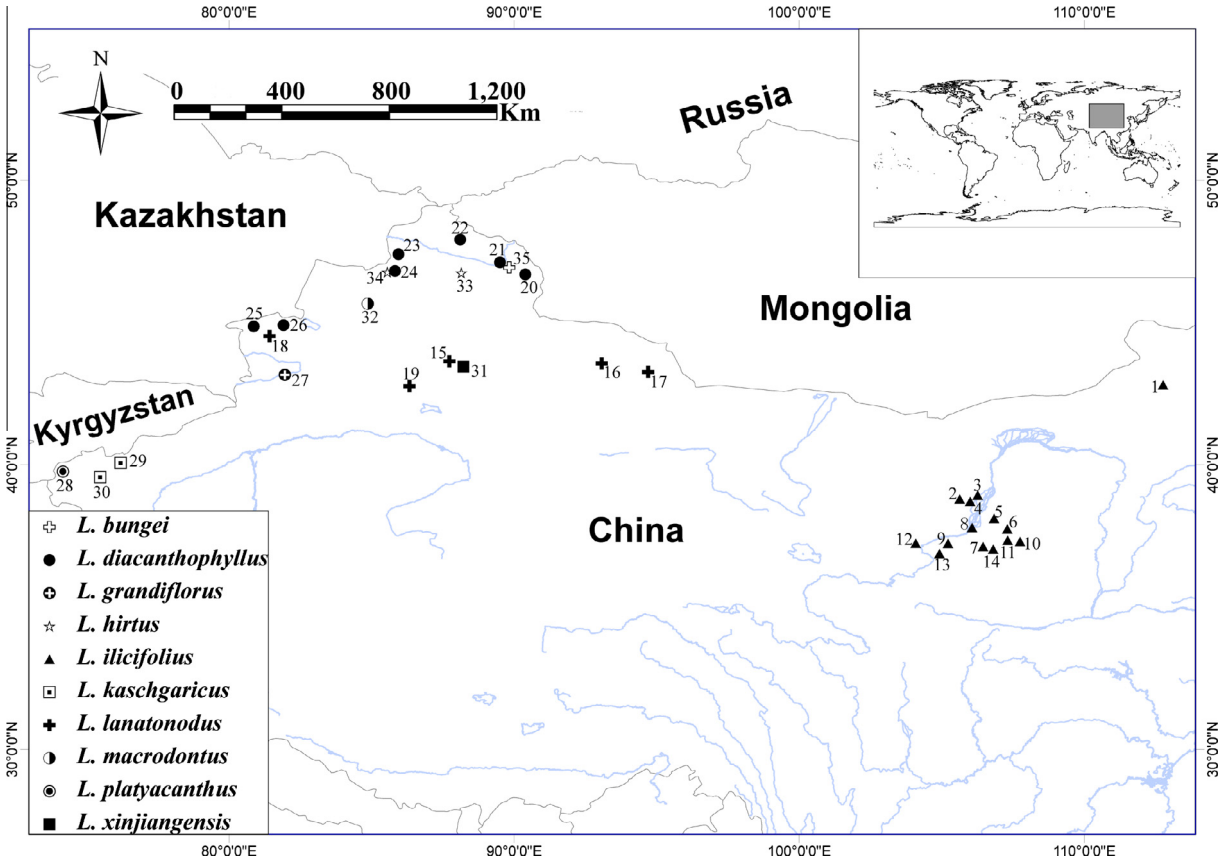

Fig. 1. Map of sample sites for ten species of *Lagochilus* in arid Northwest China. Location details are given in Table 1, and locality numbers correspond to those in Table 1.

method (Bandelt et al., 1999). The network represents genealogical relationships among haplotypes, even in case of shallow genetic divergence, which tree-building methods (e.g., NJ and BI) do not always detect. In these analyses, all sequences, representing all haplotypes and outgroups were deposited in NCBI GenBank under accession numbers JX014468-JX014489, JX014502-JX014523 and JN375724-JN375751.

2.3.2. Molecular variability population genetic structure

The geographical structure of genetic variation was assessed by AMOVA using ARLEQUIN version 3.1 (Excoffier et al., 2005). In addition, a spatial analysis of molecular variance also implemented in the program SAMOVA version 1.0 (Dupanloup et al., 2002) to define partitions of sampling sites that are maximally differentiated from each other without any a priori assumption about population structure. This program implements a simulated annealing approach to gather geographically homogenous populations to match the user defined number of groups (K), F_{CT} value was given for every calculation. We performed these analyses for the range of $2 \leqslant K \leqslant 20$, starting from 100 random initial conditions for each run. Finally, the number of groups that maximizes the proportion of total genetic variance among groups of populations (F_{CT}) was retained as the best grouping of populations. So, we chose the number of groups with the highest F_{CT} value.

HAPLONST was used to estimate within population diversity (h_S) and total gene diversity (h_T). A pattern of isolation by distance (IBD) was also assessed by testing the correlation between the matrix of pairwise population differentiation statistics (F_{ST}) values and the matrix of geographical distances between pairs of populations using a Mantel test (10,000 permutations) implemented in the MANTEL version 2.0 package (Liedloff, 1999).

2.3.3. Phylogenetic analyses and estimate of divergence time

The phylogenetic relationships and divergence time analyses among all haplotypes found in all species of *Lagochilus* were inferred as follows, using *Panzeria alashanica* and *Leonurus turkestanicus* as outgroups.

Our preliminary sequence analyses found poly-A/T regions and small inversions to be highly variable and homoplasious as described previously (Kelchner, 2000; Kim and Lee, 2005); thus, they were not included in further analyses. The two linked cpDNA intergenic spacers were concatenated and analyzed as one sequence in this study.

The phylogeny reconstruction based on the combined sequence matrix was performed in both Neighbor-joining (NJ) and Bayesian inference (BI). The NJ analysis was conducted with MEGA version 4.0 (Tamura et al., 2007), incorporating Kimura's 2-parameter model of DNA evolution (Kimura, 1980). To evaluate clades support, 1000 bootstrap replicates (Felsenstein, 1985) were performed using fast heuristic search and tree bisection-reconnection (TBR) branch swapping. BI analysis was conducted using MrBayes version 3.1.2 (Ronquist and Huelsenbeck, 2003). The best substitution model for BI analysis was identified using the Akaike Information Criterion (AIC) (Kelchner and Thomas, 2007), which is determined via Modeltest version 3.7 (Posada and Crandall, 1998). The Markov chain Monte Carlo (MCMC) algorithm was run for 2,000,000 generations, with four incrementally heated chains. The analysis involved starting from a random tree and sampling every 1000 generations. The first 10% of trees were treated as burn-in, and discarded. The remaining trees were used to construct a Bayesian consensus tree.

Approximate divergence times were estimated using a Bayesian approach with the software BEAST version 1.5.4 (Drummond and

Rambaut, 2007). BEAUti was used to set criteria for the analysis. We used the AIC estimated by Modeltest version 3.7 (Posada and Crandall, 1998) and the Bayes Factor (BF) calculated by Tracer version 1.5 to check for convergence of MCMC and adequate effective sample sizes (ESSs) (>200) after the first 20% of generations had been discarded as burn-in. The MCMC simulation was run for 10,000,000 generations, and trees were sampled every 1000 generations. In addition, the settings used a Yule tree prior, an HKY substitution model with four gamma categories and the uncorrelated log-normal relaxed clock. The final joint sample was used to estimate the maximum clade credibility tree using TreeAnnotator version 1.5.2 (part of the BEAST package) with a burn-in of 1000 trees. Final trees were evaluated and edited in FigTree version 1.3.1. Statistical support for the clades was determined by assessing the Bayesian posterior probability. Substitution rates and the 95% highest posterior densities (HPDs) were determined with Tracer version 1.5 in combined runs. The divergence times are given as the mean and the 95% HPDs in millions of years. The 95% HPDs intervals defines the precision of estimation.

An Isolation-with-Migration (IM) model of population structure was fitted to the two major groups of the East and West region populations with the software IMa2 (Hey, 2010), aiming at estimation of divergence time between the 14 populations in the East and 21 populations in the West. To obtain reasonable and consistent results, ten independent runs of IMa analyses were performed, each with a different random seed. All analyses were conducted under the HKY model as recommended for cpDNA. We checked effective sample sizes throughout the run and compared results to assess convergence. The burn-in period was set to 100,000 iterations. We chose mutation models and at least three runs with ESS > 50 were compared to ensure convergence. We only trusted estimates whose posterior distribution dropped to zero within the prior intervals investigated. Finally, the Poisson model of mutation implemented in coalescent analyses accounted for randomness of the mutation process, and thus potentially for its non-clock behavior.

In this study, we used published nucleotide substitution rates of 0.28% Myr^{-1} in accordance with the substitution rate of *psb*A-*trn*H + *trn*S-*trn*G spacers in *Petunia* (Lorenz-Lemke et al., 2010), 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.

2.3.4. Demographic analyses

Demographic history was initially explored using mismatch distribution analysis (MDA), which represented the frequency distribution of pair-wise nucleotide differences among all haplotypes in the population(s) (Rogers and Harpending, 1992). A population in demographic equilibrium is characterized by multi-model mismatch distributions, whereas populations that have experienced recent demographic expansion should show smooth uni-model distributions. We conducted this analysis to test for demographic patterns in the major regional groups.

Demographic expansion was also inferred from three parameters: θ_0, θ_1 (θ before and after the population growth) and τ (time since expansion measured in units of mutational time) (Rogers, 1995). These parameters (θ_0, θ_1 and τ) were estimated by a general nonlinear least-squares approach using ARLEQUIN. The values of τ were transformed to estimate the real time since expansion with the equation $\tau = 2ut$, where u is mutation rate for the whole sequence under study, and t is the time since expansion in generations. The value u was calculated from the formula $u = 2\mu k$, where μ is the mutation rate per nucleotide and k is the number of nucleotides assayed (k = 1102 in this study). In addition, we used Tajima's D test (Tajima, 1989) and Fu's F_S test (Fu, 1997) as implemented in ARLEQUIN to detect evidence of a recent demographic expansion within each inferred biogeographical region. The significance of deviation from value was tested with 10,000 bootstrap replicates.

However, MDAs and neutrality tests sometimes cannot take full advantage of historical signals within DNA data because they rely solely on segregating sites and haplotype patterns (Fitzpatrick et al., 2009). To obtain estimates of changes in demographic growth over the history of major areas, the historical demographic dynamics of *Lagochilus* were inferred from Bayesian skyline plot (BSP) analyses using BEAST (Drummond et al., 2005). The BSP analyses are preferred because multiple loci are used to estimate effective population size through time. Linear and stepwise models were explored using an uncorrelated lognormal relaxed clock. Runs consisted of 20,000,000 generations, with trees sampled every 1000 generations. The BSP was visualized in the program Tracer, which summarizes the posterior distribution of population size over time.

3. Results

3.1. Sequence characteristics and within-population genetic diversity

Sequence analyses of *Lagochilus* resulted in an aligned fragment length of 375 base pairs (bp) of *psb*A-*trn*H and 727 bp of *trn*S-*trn*G intergenic spacers. The concatenated chloroplast spacer is A/T rich, with an average content of 68.49%, and these data are consistent with the nucleotide composition of most noncoding spacers and pseudogenes, because of low functional constraints (Li, 1997). In relation to base substitutions and length variation character states, there are 37 variable sites, including 2 singleton variable sites, and 35 parsimony informative sites. From all variable sites a total of 34 haplotypes were identified from the *Lagochilus* plants analyzed (Table 1 and Fig. 2). Table 3 shows basic genetic data and statistics for all species in the major regions that were studied.

3.2. Population genetic and phylogeographical structure

Spatial genetic analyses of cpDNA haplotypes in 35 populations using SAMOVA indicated that F_{CT} increased to a maximal value of 0.90963 when K = 2 (K, the number of groups). When K was varied from 2 to 6, the values indicated smooth trends, but for K from 7 to 20, the values show a downturn. For K = 2, 3, 4, 5, 6, 7, the values of F_{CT} = 0.90963, 0.90844, 0.90166, 0.90861; 0.90333, 0.89622, respectively. So the division by SAMOVA of the chlorotypes of all the 35 sampled populations into two groups is appropriate. The grouping pattern of populations corresponding to K = 2 is: (1) populations 1–14, belonging to the East region of the arid zones in Northwest China, including Inner Mongolia, Shaanxi, Gansu and Ningxia; (2) populations 15–35 belonging to the west region of the arid zones in Northwest China, mainly in Xinjiang, the arid center of Central Asia. Differentiation among population based on cpDNA variation, HAPLONST indicated that within-population gene diversity (h_S) was 0.321 (SE 0.0238), and total gene diversity (h_T) was 0.952 (SE 0.0562).

The results of the AMOVA are presented in Table 2. Most of the total variation (71.24%, $P < 0.001$) was explained by differences among populations. There was a significant (46.91%, $P < 0.001$) variation between regions (east and west), suggesting that the two separately distributed groups have a strong genetic differentiation. Results of the Mantel test showed a significant correlation between the pairwise estimates of F_{ST} and the natural logarithm of the geographical distances (r = 0.542; P = 0.0001), indicating high genetic structure, and that genetic differentiation and geographical distance were significantly correlated, implying strong isolation by distance (IBD).

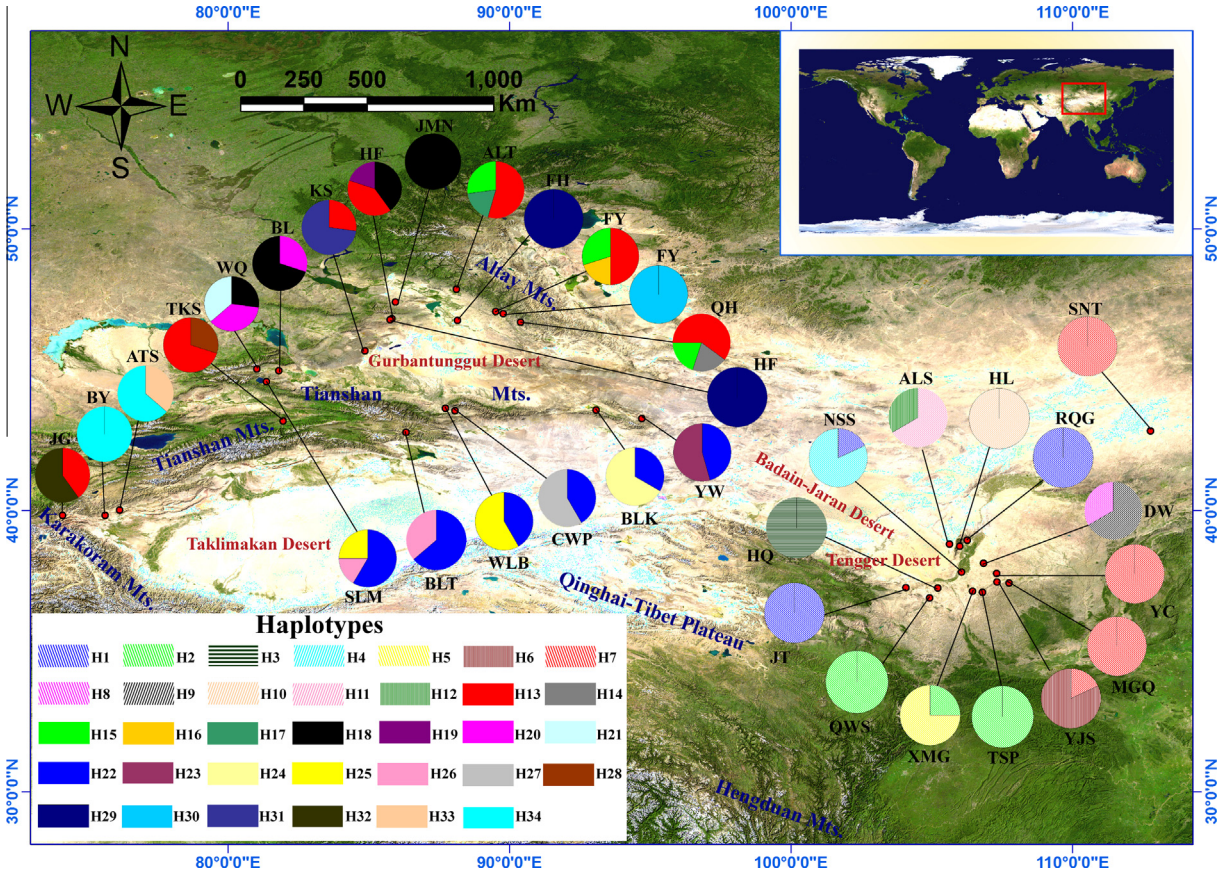

Fig. 2. Geographical distribution of 34 cpDNA haplotypes recovered from *Lagochilus* populations from Northwest China. The pie charts reflect the frequency of haplotype occurrence in each population. Haplotype colours correspond to those shown in panel.

Table 2
Results of the analysis of molecular variance for 35 populations of *Lagochilus* grouped in two geographical regions (East region and West region) based on cpDNA *trn*S-*trn*G and *psb*A-*trn*H sequence data.

Source of variation	d.f.	Sum of squares	Variance of component	Percentage of variance (%)
Among populations (total)	34	425.339	1.49419	71.24**
Within populations (1–14) vs (15–35)	361	301.948	0.91745	28.76**
Among groups	1	117.138	0.86559	46.91**
Among populations within groups	33	308.201	0.01822	4.05
Within populations	361	301.948	0.91745	49.04**

** $p < 0.001$.

3.3. Phylogenetic and genealogical relationships of cpDNA haplotypes

BI and NJ analyses yielded essentially identical topologies, but only the NJ tree is presented here, for higher resolution. Relationships among all identified haplotypes (Fig. 3A) were very similar with the topology presented by a haplotypic network (Fig. 3B). The phylogram showed the most basal divergence of all haplotypes is presented in Fig. 3A. The phylogenetic relationships of ingroup haplotypes of *Lagochilus* were well-resolved and two main haplotype groups can be recognized on the tree that all clades with high bootstrap values (BP) in the NJ analysis and high posterior probability (PP) values in the BI analysis. Two main clades can be recognized on the phylogeny, an East region clade and a West region clade. The East region clade includes *L. ilicifolius* which is restricted in Inner Mongolia, Ningxia, Shaanxi and Gansu. The West region clade contains the haplotypes of the other species, *L. diacanthophyllus, L. hirtus*, *L. bungei*, *L. macrodontus, L. kaschgaricus, L. platyacanthus* and *L. grandiflorus*, and two endemic species (*L. lanatonodus* and *L. xinjiangensis*) which are restricted in Xinjiang. In the analyses of phylogenetic and taxonomic relationships of chloroplast haplotypes, all are not species-specific. Often more than one haplotype was found in a single species, whereas several closely related species often share the same haplotype (Table 1 and Fig. 2). For instance, 12 haplotypes (H1 to H12) were found in *L. ilicifolius*, 9 haplotypes in *L. diacanthophyllus*, and 5 haplotypes in *L. lanatonodus*. On the other hand, one haplotype presents in closely related species. For instance, *L. macrodontus*, *L. platyacanthus*, *L. grandiflorus*, and *L. diacanthophyllus* share the same haplotype (H13). Given these information, species of *Lagochilus* in China are not reciprocally monophyletic. The haplotypic network for all species is more similar to an intraspecific topology with two lineages than an interspecific topology, and some haplotypes from *L. hirtus*, *L. bungei* and *L. kaschgaricus* are dispersed in the network. The detailed information concerning each species is provided in Table 1, Figs. 2 and 3.

3.4. Molecular dating and historical demography of major lineages

Estimates of divergence times were carried in the program BEAST. The ESS values were over 200 for all the nodes discussed

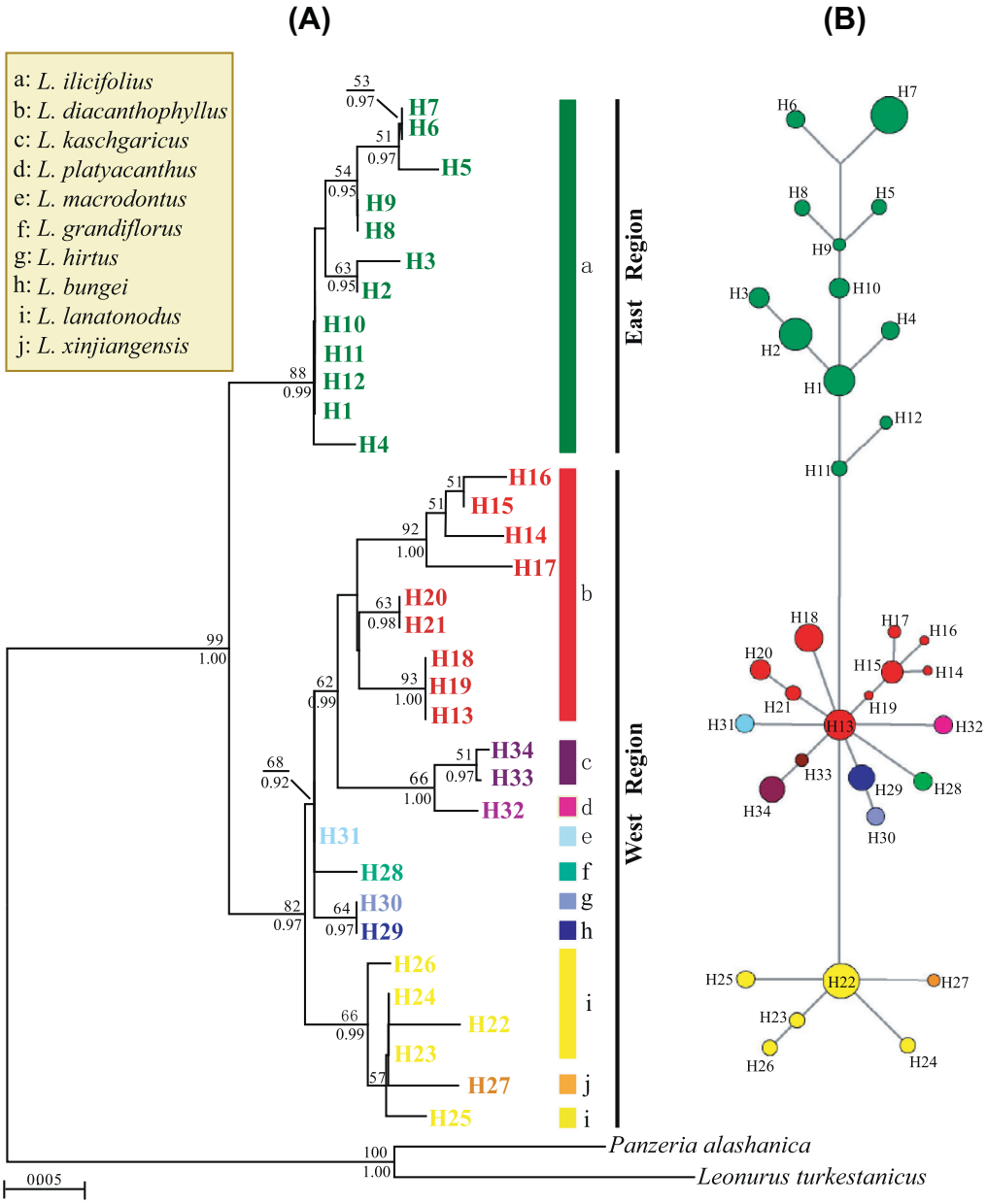

Fig. 3. The evolutionary relationships among cpDNA haplotypes of *Lagochilus*. (A) NJ phylogenetic tree for the 34 cpDNA haplotypes of *Lagochilus*. Numbers above branches are support values from bootstrap resampling/Bayesian inference. (B) Median-joining network. Sizes of the circles are proportional to the overall frequency of the haplotypes in the entire sample of all ten species.

as follows. These age estimates suggested that the separation of *Lagochilus* from its possible ancestor, the genus *Leonurus*, occurred at *ca.* 6.3 Ma (Fig. 4). Subsequently, the East region and West region clades diverged at *ca.* 2.1 Ma; or the IM model analysis estimated the divergence time between the East region and the West region groups at *ca.* 2.78 Ma (HDP95Lo = 1.426, HDP95Hi = 3.998). Detailed information concerning each of the species separated was shown in Fig. 4. These ancestral divergence events occurred between the late Tertiary and the late Pleistocene, which is in accordance with the intensification of aridification in Northwest China resulting from glaciation.

The observed MDAs of *Lagochilus* based on cpDNA revealed similar results with those of the BSP analyses. MDAs of *Lagochilus* were established for the two major sampling areas and all populations, the results of which are presented in Fig. 5. Three MDAs were close to an expected Poisson model, and also Tajima's *D* and Fu's F_S statistics were negative despite not being significant (Table 3), which suggest that overall, east and west regions display unimodal distributions. Combined with the neutrality tests, these indicate a significantly simulated expansion model. Based on the corresponding τ value (Table 3), and the generation time (as a short lived desert plant, the generation time for of *Lagochilus* is possibly 1 year), it was calculated that for the East region, the time of demographic expansion occurred *ca.* 0.06 Ma ago, whereas it was at *ca.* 0.39 Ma ago for the West area, and *ca.* 0.26 Ma ago for all regions together.

Furthermore, BSP analysis of the East region does not show a significant distribution, but does not allow the rejection of the null hypothesis of population expansion. The MDA of the East region is unimodal and significant, consistent with a model of expansion, suggesting historical demographic expansion. Generally, BSP indicated that the study area underwent distinct population expansion

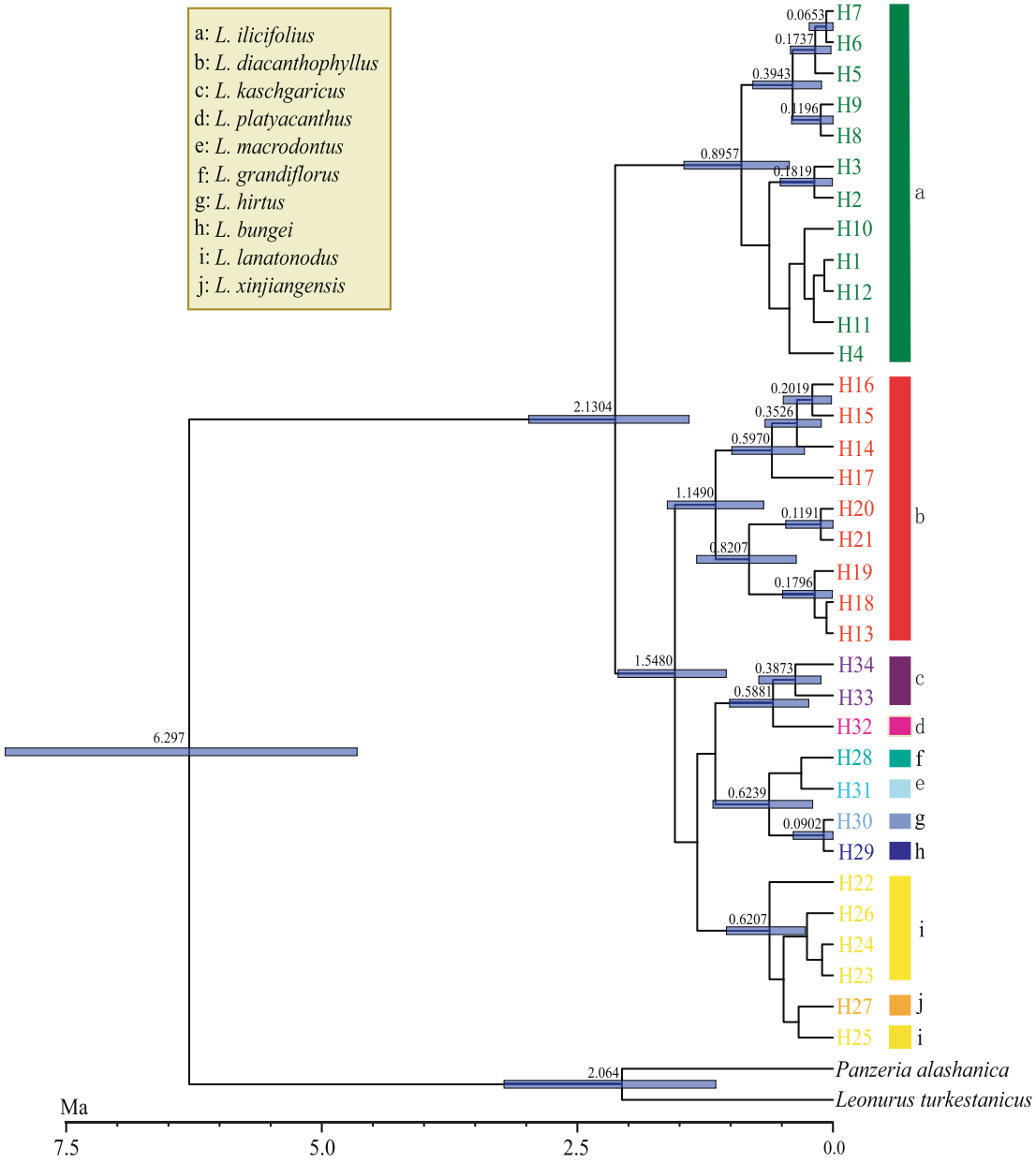

Fig. 4. Bayesian divergence time estimates of *Lagochilus* based on the combined cpDNA sequence data from two plastid gene markers (*psb*A-*trn*H, *trn*S-*trn*G). The blue bars on the nodes indicate 95% posterior credibility intervals.

or shrinkage, and specifically in this study, gave evidence of demographic expansion. According to the BSP analyses and the available value of 0.28% Myr^{-1}, since the *x*-axes in the BSP are in units of substitution per site and can thus be transformed to years before present by dividing using the mutation rate, it appears that all of the regions experienced an expansion beginning at *ca.* 0.3 Ma, whereas for the east and west regions, the corresponding expansion times were at *ca.* 0.08 Ma and *ca.* 0.4 Ma, respectively.

4. Discussion

4.1. Genetic structure in relation to Lagochilus populations

In this study, the results illustrate significant genetic divergence and a highly structured phylogenetic and phylogeographical signal for populations and species of *Lagochilus* (Fig. 3 and Table 2). In fact, SAMOVA analysis identified two well defined groups corresponding to east and west region populations at every level of divergence. In addition, the assessment of areas where the cpDNA haplotype frequency changes and genetic barriers are more robust, due to the extremely dry environment in the desert (e.g., Bardain–Jaran Desert, Tengger Desert) separating the east and west regions, this must be regarded as a significant barrier. Our results demonstrate that the expansion of intervening deserts may have increased genetic isolation between the east region and west region *Lagochilus* populations. Given that the result of Mantel test suggested a pattern of isolation by distance, this indicates that the vast desert area did obstruct gene flow and affect the phylogeographical history of the species. The important result of our study is that there is significant genetic differentiation among *Lagochilus* populations between east and west region, these populations are geographically structured according to an IBD pattern. The habitats of *Lagochilus* are either in or near deserts of aridity, in which past fragmentation due to

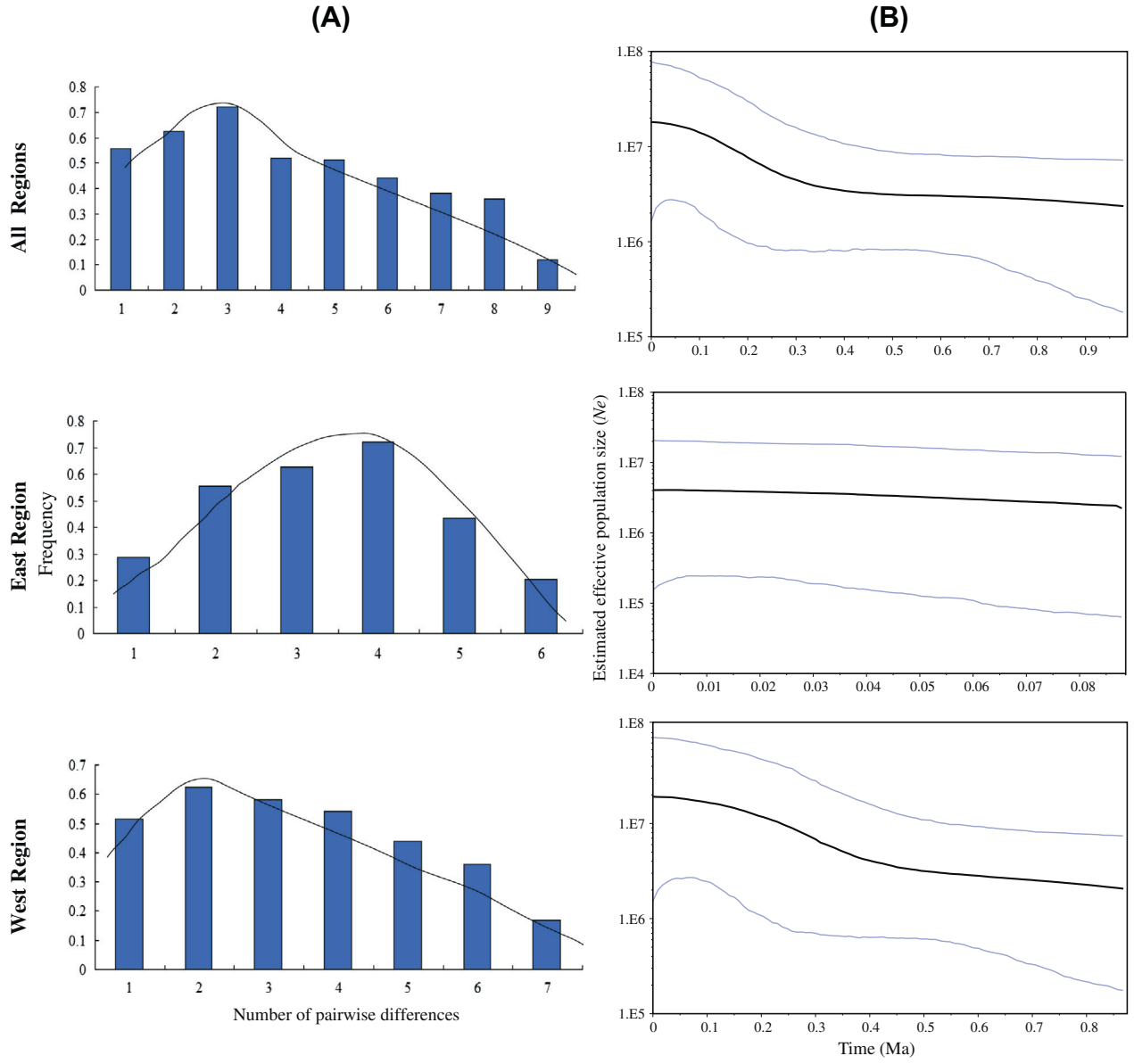

Fig. 5. Historical demography of *Lagochlius* inferred from cpDNA sequences. (A) Pair-wise mismatch distribution analyses (MDAs) for major geographical regions. (B) Bayesian skyline plots (BSPs) for the same regional groups, showing effective population size as a function of time.

Table 3
Molecular diversity indices, neutral tests and demographic estimates from two major regions and all regions of *Lagochilus* plant.

Group	N[a]	h[b]	H_d[c]	π[d]	Tajima's D (P)	Fu's F_S (P)	θ_0[e]	θ_1[f]	τ[g]	SSD[h]
East	144	12	0.882	0.0016	−0.1942 (0.12)	−0.0029 (0.726)	0.00858	14287	0.7465	0.05570
West	242	22	0.905	0.0059	−0.9968 (0.17)	−4.6997 (0.047)	0.04298	23810	4.9341	0.25260
All	386	34	0.933	0.0067	−0.6128 (0.28)	−3.3474 (0.052)	0.07722	42426	3.2091	0.17384

[a] Sample size (N).
[b] Number of haplotypes (h).
[c] Haplotype diversity (H_d).
[d] Nucleotide diversity (π).
[e] Pre-expansion population size (θ_0).
[f] Post-expansion population size (θ_1).
[g] Time in number of generations, elapsed since the sudden expansion episode (τ).
[h] The sum of squared differences (SSD).

aridification might promote the differentiation. Moreover, the result is not surprising because the vegetation is compact in relative moist place, and potential barriers to gene flow are numerous (e.g., lack of water restricts the dispersal ability of the seeds porter, extreme drought makes the plants cannot survive).

Assuming maternal inheritance of cpDNA in most angiosperms and disperses only through seeds, a highly divergent genetic structure is thought to be due to limited seed dispersal (Petit et al., 2003). So the spatial scale of cpDNA genetic differences among populations is a direct estimate of previous seed dispersal that

led to plants being established. In arid Northwest China, oasis or the edge of desert provides habitats and environmental conditions that support thousands of desert plants. However, aridification creates a mosaic landscape in arid zones and fragments, which become the source of future plants populations. Fragmentation restricts the movement of pollen and seed dispersal, modifying the gene flow and altering historical patterns of genetic subdivision. *Lagochilus* are typical montane plants, which genetic structure is largely dependent upon the dispersal ability of the seeds. On the basis of the phylogenetic analysis, we investigated the phylogeographical patterns and population structure of *Lagochilus* in two geographical regions and over all regions respectively. Genetic differentiation among populations within groups was high, as well as among groups. However, differentiation in population was low, which suggested no restricted gene flow within populations, but there is restricted gene flow among populations and groups. In Northwest China, *Lagochilus* grows are under extreme drought condition, including deserts, sandy regions, and semidesert grasslands. In these environments, heavy wind and lack of water during the flowering phase would restrict the dispersal capacity of seed dispersers, and most of seeds dispersal is constrained by gravity (by field observation), likely resulting in most seed transfer being confined to short distances.

4.2. Phylogenetic implications and possible past scenarios

Analyses of the combined cpDNA matrix support the monophyly of two clades, representing an East region and a West region (Fig. 3A), and the phylogenetic relationships among all haplotypes shows an overall congruence with the result of the haplotype network (Fig. 3B). All haplotypes are split into two major groups according to cpDNA data, are subdivided in the chloroplast phylogram as well, and chloroplast haplotypes of the species of these sections belong to well-separated chloroplast lineages of the network. However, our comprehensive sampling showed that assessing the species-level relationships in our study group is complicated by the fact that not all chloroplast haplotypes are species-specific. Incongruence between the species boundaries and the genealogy of their chloroplasts is illustrated by sharing of a single chloroplast haplotype (e.g., H13) among four species. On the other hand, several haplotypes were found within single species (Table 1 and Fig. 2). Despite the fact that chloroplast haplotype distribution does not strictly follow species circumscription, and several haplotypes were found within single species, sharing distantly related haplotypes was not a common phenomenon in *Lagochilus*. Also, an association of chloroplast haplotypes with geographically circumscribed regions rather than with taxonomic boundaries is a phenomenon observed in *L. xinjiangensis* and *L. lanatonodus*. Molecular data, along with information on the morphotaxonomic difference of the calyx teeth (4 or 5 in *L. xinjiangensis* and 5 in *L. lanatonodus*) and the geographic distribution of these two species, indicate inconsistencies with species delimitation. The sharing of haplotypes and lack of reciprocal monophyly of species might be explained by the persistence of ancestral polymorphisms during speciation events and/or exchange of genes by interspecific hybridization. These two species have the same habitat and the same period of flowering, and possible visitation by the same pollinators in the absence of physiological or genetic reproductive barriers, might have enabled hybridization events. Nonetheless, further work on chromosomes and the sequencing of additional genomic regions, especially nuclear gene regions, are needed to make more in-depth conclusions about the hybrid status of these taxa. Thus, the presence of network-based approaches, combined with tree-based analysis methods, perform well in situations where complex relationships among haplotypes exist.

Estimated times of divergence indicated that *Lagochilus* was derived from the possible ancestral genus, *Leonurus*, at *ca.* 6.3 Ma; and the sister relationships between east and west regions formed at *ca.* 2.1 Ma according to BEAST, similar to IMa which gave *ca.* 2.8 Ma, when the distribution between east and west clades become disjunctive (Fig. 4). Disjunctions in Eurasia have been outlined within many plant genera (Sun, 2002a), we propose that major factor leading to the vicariance of *Lagochilus* between the east and west regions might be postulated to be the initiation and expansion of the Badain Jaran–Tengger deserts, which followed cooling of the climate in the late Tertiary and early Pleistocene. The divergence times indicate that genetic divergence among most *Lagochilus* species occurred in Pleistocene. Although caution is needed in interpreting these results because of the poor palaeo-geographical data and fossil records, the estimated timescale for the inferred historical events is congruent with known geological and climate scenarios for the geographical areas. The QTP underwent about four or five periods of glaciations between 4.43 and 1.21 Ma (Shi, 2002; Zheng and Rutter, 1998). These palaeoclimatic conditions inferred from glacial landforms indicate the important relationships: the mid-latitude westerly originating from northern Atlantic and Mediterranean was weakened by the uplift of QTP, and Siberian high pressure weakened the Asian monsoon (Xu et al., 2010), which aggravated the dry and cold conditions in arid zones. As in other regions in the Northern Hemisphere, the late Quaternary climate in the Tianshan Mountains was generally cold-dry during glacial times and warm-humid during interglacial times. The successive enhancement of cold and dry conditions during that time period in Northwest China helped expand deserts of this region (e.g., Badain Jaran and Tengger Deserts), which most likely impelled *Lagochilus* to split into two independently evolving clades. The mixed and complex effects of mountains forced rapid genetic divergence, local adaptation, speciation and dispersal of *Lagochilus* plants across the mountain ranges of Northwest China, especially the origin and dispersal of endemic species, *L. xinjiangensis* and *L. lanatonodus*. The East region clade was restricted along the edges of deserts, which intensified in the arid zones of Northwest China *ca.* 0.85 Ma (Guan et al., 2011), and this desertification might have triggered the diversification of *L. ilicifolius* (Meng and Zhang, 2011). All of these are supported by the inference of Bayesian divergence time estimates (Fig. 4). After its origin in the latter Miocene, *Lagochilus* differentiated and evolved rapidly, accompanied by the orogeny of the Tianshan and other ranges associated with the northern QTP (Wu et al., 2003; Wu and Li, 1982).The divergence times estimated are congruent with other paleogeologic and paleoclimatic studies, however, limitations on our abilities to estimate accurate divergence times also limit the robustness of inferences; these inferences should be accepted with caution and, as hypotheses, remain open for testing and refinement by future studies.

4.3. Historical demography of the major Lagochilus lineages

Our analyses have revealed a complex phylogeographical history for *Lagochilus*, with different demographic processes affecting each of the major geographic regions. Whereas inferred phylogeography from the cpDNA has a restricted resolution on the time-span of the Pleistocene, the integrated genealogies and frequency analyses of haplotypes provide a novel insight into the demographic history of the genus.

As with other plants having high levels of haplotype diversity but moderate to low levels of nucleotide diversity, in *Lagochilus* (Table 3), this pattern is mostly attributed to expansion after a period of low effective population size, because rapid population growth or expansion within a short history works against the accumulation of large numbers of mutations (Avise, 2000). Further

support for this can be found in MDAs of the number of nucleotide differences detected across the major three regions (Fig. 5). This pattern is widely considered to be the result of mutation drift disequilibrium caused by explosive population expansion, and the population growth was also likely associated with the aridification in Northwest China following desert expansion.

The all-region grouping contains the greatest nucleotide diversity and average number of pairwise differences across the sequences, a finding that is characteristic of regional population groups with long evolutionary histories and relatively moderate demographic fluctuation based on MDAs and BSP analyses. Based on current patterns of genetic variation, along with the results of demographic analyses, we propose that regional demographic expansion of the species probably resulted from the large deserts that developed during the Pleistocene in Northwest China. These regions were characterized by intense arid conditions due to the presence of the QTP and Pleistocene cooling, which affected glacial movement, desert and loess formation, and propelled the organisms to migrate from the extremely cold and dry environments to warmer and more humid habitats (Zhang et al., 2000). However, drought-tolerant plants, such as *Lagochilus*, might have expanded during such conditions, and therefore these may have driven rapid speciation in the mountains or at the edges of deserts as adaption to the environment during these geological events. For example, the Gurbantunggut Desert expanded during the middle Pleistocene (Ding et al., 2005; Sun et al., 1998). This enlargement perhaps finally extended the demography and distribution areas of populations of *Lagochilus* in Xinjiang. Similarly, the Badain Jaran and Tengger Deserts enlarged greatly during the middle Pleistocene (Guan et al., 2011; Guo et al., 2002; Sun et al., 1998), ultimately isolating the populations of the west region occurring in Xinjiang from those in the East region in Northern China. We suggest that aridification in Northwest China accelerated specific divergence as well as geographic range expansion of a few exceptional species throughout these regions.

5. Conclusions

Our cpDNA phylogeographic reconstruction provides vivid evidences that an overall congruence between the current geographical distribution of *Lagochilus* and effects of geology and climate on evolution in arid zones of Northwest China. It is notable that the primary genetic divergence of *Lagochilus* species occurred between the late Miocene and middle Pleistocene, which was predominantly driven by the recent desert development resulting from aridification in the region inland for the rapid uplifting of QTP. *Lagochilus* diversification and demographic expansion in response to aridification in the Quaternary have also significantly contributed to genetic variation as well as the speciation and dispersal over different regional scales. Comparing with alpine plants, the response of desert plants to the climatic oscillations in Quaternary might not be synchronous with the expansion and contraction of high latitude ice sheets, but rather, expansion of species range mainly depended on the temperature and moisture requirements, especially on drought-enduring ability. Desert, as a dynamic ecosystem, along with interactions between vicariance, dispersal and habitat shift, often plays a significant role in genetic diversity and geographic distribution pattern of species of *Lagochilus*. Our study provides a useful case to understand the general biogeographic history of organisms in arid zones and gains new insight into interspecific phylogeographic patterns of plant species. Principal geological effects generally lead to habitat modifications of major biota within these regions, while climatic changes alter the ecological composition of regional or local landscapes concomitantly, which in turn, impose new selective pressures and/or geographical isolation leading to genetic diversification, adaptation and evolution of desert species in Northwest China.

Acknowledgments

The anonymous reviewers and the editor-in-chief of the journal, Dr. Derek Wildman, are gratefully acknowledged for the valuable suggestions and comments, which helped very much to improve our paper. Also, we would like to thank Dr. Stewart C. Sanderson (Shrub Sciences Laboratory, USDA, Utah, USA) and Dr. James I. Cohen (Department of Biology and Chemistry, Texas A&M International University, USA) for English improvement on the manuscript and valuable suggestions. Dr. Zhi-Hao Su, Dr. Hong-Xiang Zhang (Xinjiang Institute of Ecology and Geography, CAS), Dr. Guo-Dong Li, Dr. Liang-Liang Yue (Kunming Institute of Botany, CAS), Dr. Bin Tian (Lanzhou University) and Dr. Dong-Rui Jia (Institute of Botany, Academy of Sciences of the Czech Republic) for their kindly help about molecular data analyses. This study was financially supported by CAS Important Direction for Knowledge Innovation Project (No. KZCX2-EW-305), and Xinjiang Institute of Ecology and Geography, CAS (No. Y276031).

References

Avise, J.C., 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MA.
Avise, J.C., 2009. Phylogeography: retrospect and prospect. J. Biogeogr. 36, 3–15.
Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.
Bush, A.B.G., Little, E.C., Rokosh, D., White, D., Rutter, N.W., 2004. Investigation of the spatio-temporal variability in Eurasian Late Quaternary loess-paleosol sequences using a coupled atmosphere-ocean general circulation model. Quaternary Sci. Rev. 23, 481–498.
Clark, M.K., House, M.A., Royden, L.H., Whipple, K.X., Burchfiel, B.C., Zhang, X., Tang, W., 2005. Late Cenozoic uplift of southeastern Tibet. Geology 33, 525–528.
Comes, H.P., Kadereit, J.W., 1998. The effect of Quaternary climatic changes on plant distribution and evolution. Trends Plant Sci. 3, 432–438.
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.W., 1992. Design and testing of a plan-specific PCR primer for ecological and evolutionary studies. Mol. Ecol. 1, 233–240.
Cun, Y.Z., Wang, X.Q., 2010. Plant recolonization in the Himalaya from the southeastern Qinghai-Tibetan Plateau: geographical isolation contributed to high population differentiation. Mol. Phylogenet. Evol. 56, 972–982.
Ding, Z.L., Derbyshire, E., Yang, S.L., Sun, J.M., Liu, T.S., 2005. Stepwise expansion of desert environment across northern China in the past 3.5 Ma and implications for monsoon evolution. Earth Planet. Sci. Lett. 237, 45–55.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22, 1185–1192.
Dupanloup, I., Schneider, S., Excoffier, L., 2002. A simulated annealing approach to define the genetic structure of populations. Mol. Ecol. 11, 2571–2581.
Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinf. 1, 47–50.
Fang, X.M., Lu, L.Q., Yang, S.L., Li, J.J., An, Z.S., Jiang, P., Chen, X.L., 2002. Loess in Kunlun Mountains and its implications on desert development and Tibetan Plateau uplift in west China. Sci. China Ser. D Earth Sci. 45, 289–299.
Fang, X.M., Yan, M.D., Van der Voo, R., Rea, D.K., Song, C.H., Parés, J.M., Gao, J.P., Nie, J.S., Dai, S., 2005. Late Cenozoic deformation and uplift of the NE Tibetan Plateau: evidence from high-resolution magnetostratigraphy of the Guide Basin, Qinghai Province, China. Geol. Soc. Am. Bull. 117, 1208–1225.
Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.
Fitzpatrick, S.W., Brasileiro, C.A., Haddad, C.F.B., Zamudio, K.R., 2009. Geographical variation in genetic structure of an Atlantic coastal forest frog reveals regional differences in habitat stability. Mol. Ecol. 18, 2877–2896.
Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915–925.
Guan, Q.Y., Pan, B.T., Li, N., Zhang, J.D., Xue, L.J., 2011. Timing and significance of the initiation of present day deserts in the northeastern Hexi Corridor, China. Palaeogeogr. Palaeocl. Palaeoecol. 306, 70–74.
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.

Guo, Z.T., Sun, B., Zhang, Z.S., Peng, S.Z., Xiao, G.Q., Ge, J.Y., Hao, Q.Z., Qiao, Y.S., Liang, M.Y., Liu, J.F., Yin, Q.Z., Wei, J.J., 2008. A major reorganization of Asian climate by the early Miocene. Clim. Past 4, 153–174.
Hamilton, M., 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol. Ecol. 8, 521–523.
Hewitt, G.M., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907–913.
Hewitt, G.M., 2001. Speciation, hybrid zones and phylogeography-or seeing genes in space and time. Mol. Ecol. 10, 537–549.
Hey, J., 2010. Isolation with migration models for more than two populations. Mol. Biol. Evol. 27, 905–920.
Jia, D.R., Liu, T.L., Wang, L.Y., Zhou, D.W., Liu, J.Q., 2011. Evolutionary history of an alpine shrub *Hippophae tibetana* (Elaeagnaceae): allopatric divergence and regional expansion. Biol. J. Linn. Soc. 102, 37–50.
Kelchner, S.A., 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. Ann. Mo. Bot. Gard. 87, 482–498.
Kelchner, S.A., Thomas, M.A., 2007. Model use in phylogenetics: nine key questions. Trends Ecol. Evol. 22, 87–94.
Kim, K.J., Lee, H.L., 2005. Widespread occurrence of small inversions in the chloroplast genomes of land plants. Mol. Cells 19, 104–113.
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.
Li, W.H., 1997. Molecular Evolution Sinauer. Sunderland Press, Mass..
Li, H., Hedge, I., 1994. Lamiaceae. In: Wu, Z.Y., Raven, P.H. (Eds.), Flora of China, vol. 17. Science Press, Missouri Botanical Garden Press, Beijing, St. Louis, pp. 50–299.
Li, Y., Stocks, M., Hemmilä, S., Källman, T., Zhu, H.T., Zhou, Y.F., Chen, J., Liu, J.Q., Lascoux, M., 2010. Demographic histories of four spruce (*Picea*) species of the Qinghai-Tibetan Plateau and neighboring areas inferred from multiple nuclear loci. Mol. Biol. Evol. 27, 1001–1014.
Li, Y., Zhai, S.N., Qiu, Y.X., Guo, Y.P., Ge, X.J., Comes, H.P., 2011. Glacial survival east and west of the 'Mekong-Salween-Divide' in the Himalaya-Hengduan Mountains Region as revealed by AFLP and cpDNA sequence variation in *Sinopodophyllum hexandrum* (Berberidaceae). Mol. Phylogenet. Evol. 59, 412–424.
Liedloff, A., 1999. MANTEL v. 2.0: Mantel Nonparametric Test Calculator. Queensland University of Technology, Brisbane, Australia.
Liu, Y.F., Wang, Y., Huang, H.W., 2009. Species-level phylogeographical history of *Myricaria* plants in the mountain ranges of western China and the origin of *M. laxiflora* in the Three Gorges mountain region. Mol. Ecol. 18, 2700–2712.
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.
Lu, H.Y., Wang, X.Y., An, Z.S., Miao, X.D., Zhu, R.X., Ma, H.Z., Li, Z., Tan, H.B., Wang, X.Y., 2004. Geomorphic evidence of phased uplift of the northeastern Qinghai-Tibet Plateau since 14 million years ago. Sci. China, Ser. D Earth Sci. 47, 822–833.
Meng, H.H., Zhang, M.L., 2011. Phylogeography of *Lagochilus ilicifolius* (Lamiaceae) in relation to Quaternary climatic oscillation and aridification in northern China. Biochem. Syst. Ecol. 39, 787–796.
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.
Owen, L.A., Finkel, R.C., Barnard, P.L., Ma, H.Z., Asahi, K., Caffee, M.W., Derbyshire, E., 2005. Climatic and topographic controls on the style and timing of Late Quaternary glaciation throughout Tibet and the Himalaya defined by 10Be cosmogenic radionuclide surface exposure dating. Quaternary Sci. Rev. 24, 1391–1411.
Petit, R.J., Pineau, E., Demesure, B., Bacilieri, R., Ducousso, A., Kremer, A., 1997. Chloroplast DNA footprints of postglacial recolonization by oaks. Proc. Natl. Acad. Sci. USA 94, 9996–10001.
Petit, R.J., Aguinagalde, I., de Beaulieu, J.L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300, 1563–1565.
Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Qiu, Y.X., Fu, C.X., Comes, H.P., 2011. Plant molecular phylogeography in China and adjacent regions: tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. Mol. Phylogenet. Evol. 59, 225–244.
Rogers, A.R., 1995. Genetic evidence for a Pleistocene population expansion. Evolution 49, 608–615.
Rogers, A.R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9, 552–569.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
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.
Schaal, B.A., Olsen, K.M., 2000. Gene genealogies and population variation in plants. Proc. Natl. Acad. Sci. USA 97, 7024–7029.
Shi, Y.F., 2002. Characteristics of late Quaternary monsoonal glaciation on the Tibetan Plateau and in East Asia. Quatern. Int. 97, 79–91.
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.
Sun, H., 2002a. Tethys retreat and Himalayas-Hengduanshan mountains uplift and their significance on the origin and development of the Sino-Himalayan elements and alpine flora. Acta Bot. Yunnan 24, 273–288.
Sun, J.M., 2002b. Source regions and formation of the loess sediments on the high mountain regions of northwestern China. Quaternary Res. 58, 341–351.
Sun, J.M., Zhang, Z.Q., 2009. Syntectonic growth strata and implications for late Cenozoic tectonic uplift in the northern Tian Shan, China. Tectonophysics 463, 60–68.
Sun, J.M., Ding, Z.L., Liu, T.S., 1998. Desert distributions during the glacial maximum and climatic optimum: example of China. Episodes 21, 28–31.
Sun, J.M., Ye, J., Wu, W.Y., Ni, X.J., Bi, S.D., Zhang, Z.Q., Liu, W.M., Meng, J., 2010. Late Oligocene–Miocene mid-latitude aridification and wind patterns in the Asian interior. Geology 38, 515–518.
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.
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.
Tingru, Z., 1983. The difference of paleogeographic environment in the Chinese Quaternary. Geogr. Sci. 3, 192–205.
Wang, C.S., Zhao, X.X., Liu, Z.F., Lippert, P.C., Graham, S.A., Coe, R.S., Yi, H.S., Zhu, L.D., Liu, S., Li, Y.L., 2008. Constraints on the early uplift history of the Tibetan Plateau. Proc. Natl. Acad. Sci. USA 105, 4987–4992.
Wu, C.Y., Li, H.W., 1982. On the evolution and distribution in Labiatae. Acta Bot. Yunnan 4, 97–118.
Wu, G.J., Pan, B.T., Guan, Q.Y., Gao, H.S., 2002. The maximum glaciation and desert expansion in China during MIS16. J. Glaciol. Geocryol. 24, 544–549.
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, Z.H., Barosh, P.J., Wu, Z.H., Hu, D.G., Zhao, X., Ye, P.S., 2008. Vast early Miocene lakes of the central Tibetan Plateau. Geol. Soc. Am. Bull. 120, 1326–1337.
Xu, X.K., Kleidon, A., Miller, L., Wang, S.Q., Wang, L.Q., Dong, G.C., 2010. Late Quaternary glaciation in the Tianshan and implications for palaeoclimatic change: a review. Boreas 39, 215–232.
Yang, X.P., Rost, K.T., Lehmkuhl, F., Zhu, Z.D., Dodson, J., 2004. The evolution of dry lands in northern China and in the Republic of Mongolia since the Last Glacial Maximum. Quatern. Int. 118, 69–85.
Yang, X.P., Scuderi, L., Paillou, P., Liu, Z.T., Li, H.W., Ren, X.Z., 2011. Quaternary environmental changes in the drylands of China – a critical review. Quatern. Sci. Rev. 30, pp. 3319–3233.
Ying, T.S., Boufford, D.E., Zhang, Y.L., 1993. The Endemic Genera of Seed Plants of China. Science Press, Beijing.
Zhang, D.F., Fengquan, L., Jianmin, B., 2000. Eco-environmental effects of the Qinghai-Tibet Plateau uplift during the Quaternary in China. Environ. Geol. 39, 1352–1358.
Zhang, T.C., Comes, H.P., Sun, H., 2011. Chloroplast phylogeography of *Terminalia franchetii* (Combretaceae) from the eastern Sino-Himalayan region and its correlation with historical river capture events. Mol. Phylogenet. Evol. 60, 1–12.
Zheng, B.X., Rutter, N., 1998. On the problem of Quaternary glaciations, and the extent and patterns of Pleistocene ice cover in the Qinghai-Xizang (Tibet) Plateau. Quatern. Int. 45, 109–122.