Pak. J. Bot., 46(6): 2121-2130, 2014.

GENETIC POPULATION STRUCTURE OF THE DESERT SHRUB SPECIES *LYCIUM RUTHENICUM* INFERRED FROM CHLOROPLAST DNA

HAIKUI CHEN[1,2¶], LIYAN ZENG[3¶], TAKAHIRO YONEZAWA[2], XIAN REN[1*] AND YANG ZHONG[2,4*]

[1]*School of Life Sciences, Beifang University of Nationalities, Yinchuan, 750021, China*
[2]*School of Life Sciences, Fudan University, Shanghai, 200438,China*
[3]*Shanghai Public Health Clinical Center, Fudan University, Shanghai, 201508, China*
[4]*Institute of Biodiversity Science, University of Tibet, Lhasa, 850000, China*
[*]*Corresponding author's e-mail: renxian57@163.com ; yangzhong@fudan.edu.cn; Tel.: 86-021-65644436*
[¶] *These authors contributed equally to this work*

Abstract

Lycium ruthenicum (Solananeae), a spiny shrub mostly distributed in the desert regions of north and northwest China, has been shown to exhibit high tolerance to the extreme environment. In this study, the phylogeography and evolutionary history of *L. ruthenicum* were examined, on the basis of 80 individuals from eight populations. Using the sequence variations of two spacer regions of chloroplast DNA (trnH-psbA and rps16-trnK) , the absence of a geographic component in the chloroplast DNA genetic structure was identified (G_{ST} = 0.351, N_{ST} = 0.304, $N_{ST}<G_{ST}$), which was consisted with the result of SAMOVA, suggesting weak phylogeographic structure of this species. Phylogenetic and network analyses showed that a total of 10 haplotypes identified in the present study clustered into two clades, in which clade I harbored the ancestral haplotypes that inferred two independent glacial refugia in the middle of Qaidam Basin and the western Inner Mongolia. The existence of regional evolutionary differences was supported by GENETREE, which revealed that one of the population in Qaidam Basin and the two populations in Tarim Basin had experienced rapid expansion, and the other populations retained relatively stable population size during the Pleistocene . Given the results of long-term gene flow and pairwise differences, strong gene flow was insufficient to reduce the genetic differentiation among populations or within populations, probably due to the genetic composition containing a common haplotype and the high number of private haplotypes fixed for most of the population. The divergence times of different lineages were consistent with the rapid uplift phases of the Qinghai-Tibetan Plateau and the initiation and expansion of deserts in northern China, suggesting that the origin and evolution of *L. ruthenicum* were strongly influenced by Quaternary environment changes.

Key words: *Lycium ruthenicum*, phylogeography, haplotype, Qinghai-Tibetan Plateau, Quaternary.

Introduction

The rugged topography of northern China is arid and semi-arid with a cold and dry continental climate. It is characterized by thirteen deserts (13% of the land mass of China) (Sheehy, 1992). The desertification background in northern China occurred due to the climate, sandy soils and water shortage (Chen & Tang, 2005). For example, the arid regions of China experienced significant climatic fluctuations during the Late Quaternary. These fluctuations are suggested by an abundance of geomorphologic, lacustrine, pedologic and geochemical evidence, alongside the presence of faunal and floral fossils (Yang *et al*., 2011). The Quaternary climate oscillations had a profound impact on evolutionary history and spatial distribution of species with which genetic diversity either vanished or generated (Hewitt, 2000).

In recent years, analyses addressing phylogeography aspects of plants in northern China have made some progress and provided important molecular evidence to understand plant diversity and distribution patterns in northern China during the Quaternary glacial, and can be summarized as follows.

Guo *et al*., (2010) suggested that independent glacial refugia of trees with few haplotype fixations were maintained for the duration of the following glacial ages with the stream of migration of trees from northwestern China to southern refugia during the Last Glacial Maximum (LGM). The haplotypes retained in Inner Mongolia during the LGM might have had an expansion, resulting in a wider distribution extending to north-western China after the glacial period. Therefore, the low levels of genetic diversity within each region after regional recolonization were shown at the end of the LGM. Distinct regional genetic differentiation was formed due to the enlarged deserts during the middle Pleistocene, but regional postglacial expansions in Gansu and Qinghai were derived from a common refugium without disturbances from the mountains and large deserts.

However, Ma *et al*., (2012) opined that some small shrub populations survived and developed during glacial ages, more ancient haplotypes and ~~a~~ high genetic diversity were retained, and significant regional divergence existed due to the geographic barriers. The haplotypes of these disjointed regional populations were gained from independent glacial refugia that gave rise to the Quaternary environmental changes in each region. Postglacially colonized populations were also expected to have low genetic variation, but no postglacial expansions of the shrub populations in northern China were detected.

Therefore, different plant distributions in northern China indicated complex and various evolutionary history corresponding the complex geologic history and climatic oscillation in this region. At present, a small number of species have been studied in desert areas of northern China. More phylogeographic studies of different species sources with different habitats are needed to better understand genetic differentiation and structure of the extant species affected by the Quaternary glacial climate.

Lycium ruthenicum (Solananeae), mostly distributed in the desert regions of north and northwest China, has small, fleshy leaves, zigzagged bending of stems, internodes with short thorns, and purple black berries, containing four aliphatic acids, Vitamin C, crude saccharides, 17 amino acids, 13 kind of minerals, and18 compounds from the essential oil (Kuang & Lu, 1978; Yao *et al*., 2011). The species is used as a traditional Chinese medicine for the treatment of cardiopyretic disease and gynecological diseases (Gan *et al*., 1997). In addition, *L. ruthenicum* is a spiny shrub with a higher tolerance to salt, drought, strong wind and the cold, As such, it is an important constructive and dominant species in the saline desert vegetation communities (Xi *et al*., 2003; He *et al*., 2011). Therefore, *L. ruthenicum* plays an important role in medical care and restoring the desert ecosystem.

Previous genetic studies based on RAPD data revealed phylogenetic relationships between *L. ruthenicum* and other *Lycium* species (Yin *et al*., 2005). In addition, previous researches mostly were focused on phytochemistry, biological activities, physiological ecology, karyotype analysis, and morphological evolution of this species (Zhang & Zhang, 2004; Chen *et al*., 2008; He *et al*., 2011; Yao *et al*., 2011). Despite the great importance of economic value and ecological restoration of *L. ruthenicum*, its population genetics and evolutionary history were still relatively unknown. Molecular markers (e.g. chloroplast DNA, SSR, ISSR and ITS) are useful tools for evaluating genetic diversity and population structure (Gilani *et al*., 2011; Li *et al*., 2012; Shah *et al*., 2013). In this study, the genetic variation, population structure, phylogeography of *L. ruthenicum* from eight populations distributed in the north of China were investigated using 2 chloroplast DNA spacers (trnH-psbA and rps16-trnK). Several aims were to address the following: (1) examine the chloroplast DNA variation and populations genetic differentiation of this species in the north of China; (2) determine the phylogeographic structure of this species affected by the Quaternary glacial stage; (3) determine the demographic dynamics events of this species in different regions associated to Quaternary climatic fluctuations; (4) provide sound knowledge about its glacial refugia in accordance with other species recently reported in the north of China.

Materials and Methods

Population sampling: A total of 80 *L. ruthenicum* individuals were collected from 8 populations nearly covering the entire range of the species, Among those individuals, 9 were from Nuomuhong (NMH, Qinghai), 10 were from Geermu (GEM , Qinghai), 11 were from Minqin (MQ, Gansu), 10 were from Dunhuang (DH, Gansu), 10 were from Anxi (AX, Gansu), 11 were from Ejinaqi (EJNQ, Inner Mongolia) , 10 were from Tianshan (TS, Xingjiang), and 9 were from Yanji (YJ, Xingjiang) (Fig. 1, Table 1). Individuals were chosen randomly about 50 m apart based on the natural size of each population. Healthy leaves were dried directly with silica gel. Dried leaves were preserved in silica gel until DNA extraction. The vouchers of the sampled population were deposited in herbarium of Beifang University of Nationalities.

DNA extraction, PCR and sequence: Genomic DNAs were extracted from the powdered tissue following a CTAB procedure and gel quantified (Doyle & Doyle, 1987). The trnH-psbA and rps16-trnK intergenic spacer of the chloroplast DNA was amplified using pairs of universal primers (Shaw *et al*., 2005, 2007). The polymerase chain reactions (PCRs) were performed in a total volume of 50 µL containing 2µl 10-40 ng template DNA, 5 µl 10 × PCR reaction buffer, 4µl dNTP mix (2.5mM), 2µl of each primer (10 mM), and 2 U of Taq polymerase (TaKaRa, Kyoto, Japan). The thremocycling was performed with first cycle of denaturation at 95°C for 3 min, then 35 cycles of denaturationat 94°C for 30 sec, annealing at 50°C for 30 sec, and extension at 72°C for 1 min 30 sec, plus a final followed extension at 72°C for 10 min. PCR products were sequenced using ABI3730XL sequencer (Applied Biosystems, Foster City, California, USA).

Table 1. Population locations, numbers of sample size and site coordinate of *Lycium ruthenicum*, the estimate of haplotype diversity (h) and nucleotide diversity (π) within populations on chloroplast DNA sequences.

Population	Code	Site coordinate	Elevation	Sample size (N)	Number of haplotypes	Polymorphic sites (S)	h ± SD	Π ± SD
Mountain group				19	4	5	0.661±0.084	0.00118±0.00028
Geermu	GEM	94°56'E 36°24'N	2790m	9	3	3	0.667±0.132	0.00128±0.00035
Nuomuhong	NMH	96°28'E 36°25'N	2852m	10	3	5	0.378±0.181	0.00096±0.00050
Eastern group				42	5	7	0.643±0.062	0.00154±0.00021
Minqin	MQ	103°40'E 38°56'N	1312m	11	3	3	0.564±0.134	0.00157±0.00035
Ejinaqi	EJNQ	102°02'E 41°55'N	932m	11	3	4	0.345±0.172	0.00101±0.00051
Dunhuang	DH	94°50'E 40°22'N	1047m	10	1	0	0.000±0.000	0.00000±0.00000
Anxi	AX	95°02'E 40°31'N	1073m	10	3	3	0.600±0.131	0.00108±0.00048
Western group				19	5	3	0.696±0.077	0.00091±0.00017
Tianshan	TS	86°14'E 42°26'N	1313m	10	3	3	0.600±0.131	0.00083±0.00030
Yanji	YJ	86°32'E 41°56'N	1040m	9	4	2	0.806±0.089	0.00101±0.00015
overall				80	10	9	0.733±0.041	0.00141±0.00015

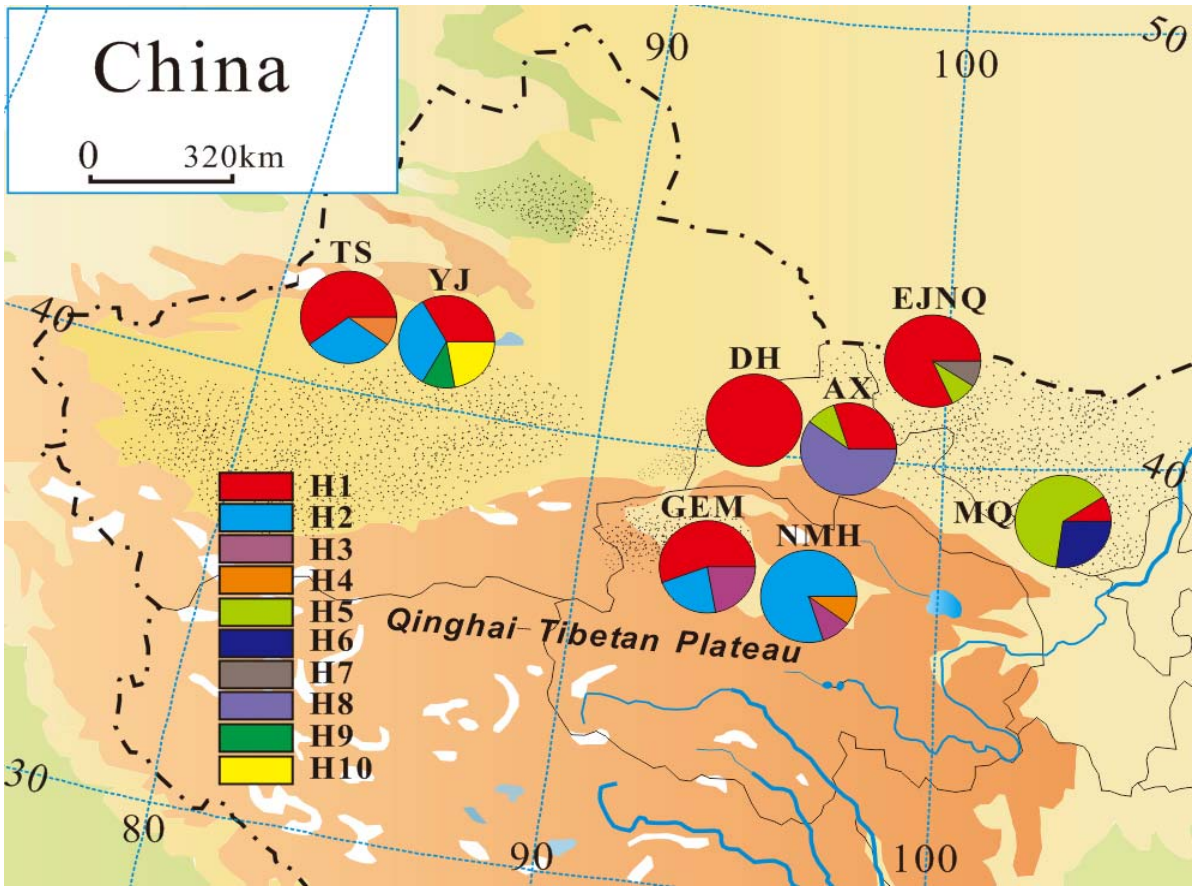

Fig. 1. Map showing population locations of *Lycium ruthenicum*. See Table 1 for the detailed information of populations.

Data analysis: The chloroplast DNA sequences were aligned with the program Clustal X 1.81 (Thompson *et al*., 1997) after combining DNA sequences of the trnH-psbA and rps16-trnK regions. To examine genetic diversity based on chloroplast DNA, haplotype diversity (h) (Nei & Tajima, 1983) and nucleotide diversity (π) (Watterson, 1975) were quantified using DnaSP (Version 5) (Librado & Rozas, 2009). The parameters of population diversity, the total genetic diversity (H_T) and the average within-population diversity (HS) were calculated using HAPLONST (Pons & Petit, 1996).

To assess the spatial population structure, a comparison of the hierarchical genetic structure between and within populations was undertaken by the analysis of molecular variance (AMOVA) using ARLEQUIN 3.1 (Excoffier *et al*., 2005). The statistical significances were tested by performing bootstrapping (10,000 permutations) (Efron & Gong, 1983). The best grouping of populations (K) was determined using the program SAMOVA 1.0 (Dupanloup *et al*., 2002) based on the largest F_{CT} index (Wright, 1978). The program was run for 100 random initial conditions from K= 2-7 under an iterative simulated annealing process. Wright's F_{ST} (Weir & Cockerham, 1984) were estimated for evaluating genetic variation for all population pairs with significance using 1,000 permutations in ARLEQUIN 3.1 (Excoffier *et al*., 2005). Sequential Bonferroni correction was used to adjust the significance levels of multiple tests (Rice, 1989).

To evaluate phylogeographic structure, two indices of differentiation (G_{ST} and N_{ST}) were estimated with 1000 permutations using HAPLONST (Pons & Petit, 1996). If the value of N_{ST} is significantly larger than the value of G_{ST}, the overall population differentiation presents geographic structure (Pons & Petit 1996).

To visualize the relationships among the chloroplast DNA haplotypes, a median-joining network (Bandelt *et al*., 1999) of haplotypes was constructed with indels (gaps) treated as missing and evenly weighted with the other mutations using Network 4.610 (http://www.fluxus-engineering.com/sharenet.htm).

To reconstruct the phylogeny and estimate divergence times of chloroplast DNA lineages, maximum likelihood (ML) tree was inferred using MEGA 5.0 (Tamura *et al*., 2011) with the GTR model. *L. americanum* and *L. chinense* were used as outgroups (Tu *et al*., 2008, 2010). Reliabilities of branching patterns of the ML tree were evaluated by calculating bootstrap probability (Felsenstein, 1985) with 1000 replication. Based on this ML tree topology, the divergence time applying the strict clock model (LRT: p-value=0.51) was estimated using the BASEML program of PAML 4.2 (Yang, 2007) with GTR+G model. The evolutionary rates of $1.0\text{-}3.0\times10^{-9}$/site/year (Wolfe *et al*., 1986) was applied for the chloroplast DNA.

To estimate the population demographic trends, coalescent analysis for *L. ruthenicum* chloroplast DNA

haplotypes was carried out using GENETREE 9.0 (Griffiths & Tavaré 1994, http://www.stats.ox.ac.uk/~griff/software.html). The exponential growth rate β_{ML} (=*N*e*b: *N*e is effective population size and b is exponential population growth rate) and θ_{ML} (=*N*e*μ: μ is mutation rate/ sequence/ generation) were estimated by the maximum likelihood method. In using this coalescent approach, it is useful to reformulate subdivided populations with or without population growth and estimate effective population size that going backwards at every point in time from a current size as $Ne(t) = Ne^{-bt}$, where t is the time from the present generation, and Ne is the present effective population size, One unit of time corresponds to 2*N*e generations. According to the filed observations, the generation interval of this species is two years, the chloroplast DNA mutation rate (u) of $1.0\text{-}3.0 \times 10^{-9}$/site/year (Wolfe *et al*., 1987) was used for the estimation.

To evaluate migrant events, the mutation-scaled long-term migration rate (M) between populations and the mutation-scaled effective population size parameter (θ) were calculated using MIGRATE 3.2.16 (Beerli & Felsenstein, 1999, 2001). Posterior probability distribution of parameters were estimated by Markov chain Monte Carlo (MCMC) simulations based on coalescence theory. In order to infer the historical gene flow associated with the populations, three groups were divided according to regional position, including western group (TS and YJ), mountain group (NMH and GEM), and eastern group (MQ, DH, AX and EJNQ). A MCMC was run for ten short chains and three long chains with 10000 and 100000 generations recorded every 20 reconstructed generations, respectively, after an initial burn-in period of 10000 iterations. MIGRATE was run three with different random number seeds to provide confidence in the estimates for M and θ.

Results

The lengths of the aligned sequences of the trnH-psbA and rps16-trnK regions of L. ruthenic*um* were 490 bp and 582 bp, respectively. Among the 80 individuals (8 populations) surveyed, four trnH-psbA and six rps16-trnK haplotypes were identified, which have been deposited in GenBank (accession numbers JN865235-JN865240 and JN865241-JN865244). Combining the sequences of these two regions resulted in 10 different haplotypes (H1-H10) based on 8 nucleotide substitutions and a inversion with lengths of 29 bp in the trnH-psbA region (Table 2).

Nucleotide diversity (π) of *L. ruthenicum* was 0.00141 with high level of haplotype diversity (0.733) (Table 1). Nucleotide diversity (π) ranged from 0.000 (DH) to 0.00157 (MQ), and haplotype diversity (h) ranged from 0.000 (DH) to 0.806 (YJ). The total genetic diversity (H_T=0.762) was much higher than the average within-population diversity (H_S=0.495). H1 was the most common haplotype shared by 37 individual, and found in all populations except for the NMH population (Fig. 1). In addition, H6-H10 was private haplotype and occurred only in one population. The highest numbers of haplotype and private haplotype were found in the YJ population, while the DH population was fixed at a single haplotype (Fig. 1).

No clear phylogeographic structure was detected (G_{ST}=0.351, N_{ST}=0.304), suggesting a lack of correlation between haplotypes and geographic distribution ($N_{ST}<G_{ST}$) (Pons & Petit 1996). In the results of the SAMOVA analyses by increasing K value from 2 to 7 revealed that the F_{CT} value increased from K=2 (F_{CT}=-0.18485), then reach the greatest value when K=4 (F_{CT}=0.27356), and dropped fluctuantly with each increasing K value (Table 3). The separation of sampling areas were identified that GEM vs. NMH vs. MQ vs. the others. However, genetic structure was insignificant at three levels based on the apportionment of genetic variation distributed among geographical groups (27.36%, $p>0.05$), among populations within groups (11.47%, $p<0.05$), and within populations (61.17%, $p<0.05$). In addition, there was one grouping containing 9-K populations and K-1 groupings each containing a single population. Thus, no significant evidence of the population subdivision and geographical barrier of *L. ruthenicum* was found using SAMOVA. Analyses of molecular variance (AMOVA) with only two levels of variation (i.e. within and among populations) suggested that most of the variation was found within populations (67.79%), while significant variation distributed among populations (32.21%, $p<0.05$) was detected.

Table 2. Variable sites of the aligned sequences of two chloroplast DNA fragments in the 10 haplotypes (H1–H10) of *Lycium ruthenicu*m. A inversion in position 395 of the trnH-psbA intergenic spacer, identified by Symbol a and b.

Haplotype	trnH-psbA				rps16-trnK					
	252	303	395		101	302	308	363	509	537
H1	A	T	a		T	G	T	G	G	C
H2	C	T	a		T	G	T	G	G	C
H3	C	T	a		T	T	G	G	G	C
H4	A	T	a		T	G	T	A	A	C
H5	C	T	b		T	G	T	G	G	C
H6	C	T	a		G	G	T	G	G	C
H7	C	T	a		T	T	T	G	G	T
H8	A	G	a		T	G	T	G	G	C
H9	C	T	a		T	G	T	A	G	C
H10	A	T	a		T	G	T	A	G	C

a:GCACCTTCTTGATAGAACAAGAAAATGAT

ḥ:ATCATTTTCTTGTTCTATCAAGAAGGTGC

Table 3. Fixation indices corresponding to groups of populations inferred by SAMOVA for *Lycium ruthenicum* populations tested for the chloroplast DNA.

Group	Population groupings	F_{CT}	P
2	(GEM) (NMH, MQ, EJNQ, DH, AX, TS, YJ)	-0.18485	0.74878
3	(GEM) (NMH) (MQ, EJNQ, DH, AX, TS, YJ)	-0.05048	0.58553
4	(GEM) (NMH) (MQ) (EJNQ, DH, AX, TS, YJ)	0.27356	0.06549
5	(GEM) (NMH) (MQ) (EJNQ) (DH, AX, TS, YJ)	0.18152	0.20821
6	(GEM) (NMH) (MQ) (EJNQ) (TS) (DH, AX, YJ)	0.07467	0.42229
7	(GEM) (NMH) (MQ) (EJNQ) (TS) (YJ) (DH, AX)	0.10485	0.10485

Table 4 Pairwise F_{ST} estimates based on chloroplast DNA haplotypes for eight populations of *Lycium ruthenicu*m. Bold F_{ST} estimates indicated significant P values after Bonferroni correction (minimum adjusted alpha = 0.00178). All other estimates were no significant at this level.

	GEM	NMH	MQ	EJNQ	DH	AX	TS	YJ
GEM	0.0000							
NMH	0.1047	0.0000						
MQ	**0.3550**	**0.2813**	0.0000					
EJNQ	0.0386	**0.3366**	**0.4525**	0.0000				
DH	0.2683	**0.6154**	**0.6780**	0.0296	0.0000			
AX	0.2949	**0.4965**	**0.5524**	0.2480	0.4167	0.0000		
TS	0.0225	0.2350	**0.4355**	-0.0287	0.1333	0.2742	0.0000	
YJ	0.0607	0.1518	**0.3884**	0.1021	0.3403	**0.3332**	-0.0166	0.000

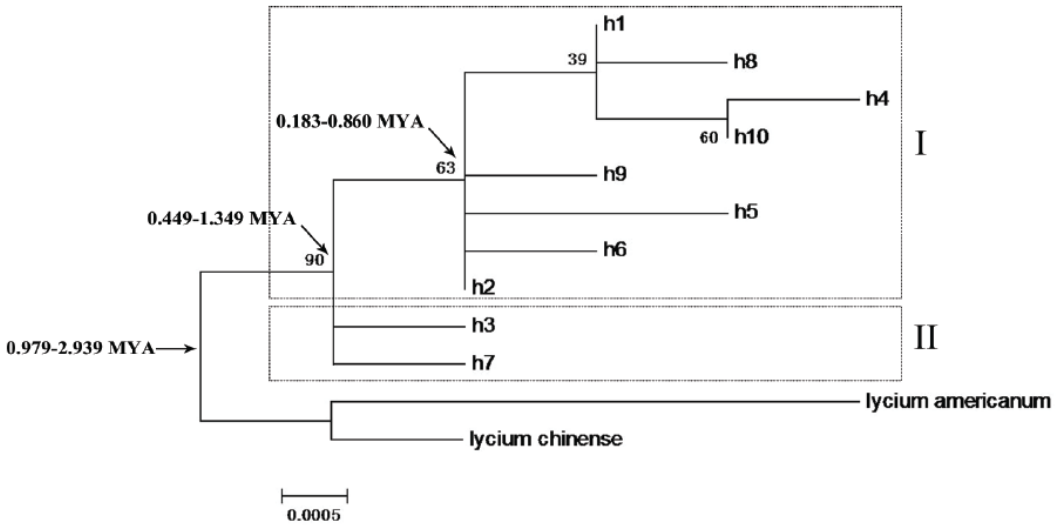

Fig. 2. Maximum-likelihood tree of chloroplast DNA haplotypes of *Lycium ruthenicum.* Numbers at nodes indicate bootstrap values. Divergent times at nodes were estimated by the global molecular clock model.

Pairwise FST values indicated high levels of genetic differentiation among population pairs and statistical significance was detected in 11 out of 28 comparisons after sequential Bonferroni correction (minimum adjusted alpha=0.00178, Table 4), indicating that the gene flow is currently restricted between them. Specifically, the high levels and significance of differentiation were observed between MQ and the other sites and ranged from 0.2813 to 0.6780 (Table 4), suggesting that it was more differentiated from other regions. The pairwise FST value of the adjacent sites between GEM and NMH was higher (0.1047), whereas this value of the geographically distant sites between GEM and TS was low (0.0225) with exceptions.

The ML tree on the basis of combined chloroplast DNA trnH-psbA and rps16-trnK intergenic spacer regions was shown in Fig. 2. *L. americanum* (accession numbers: EU742440, EU742372) and *L. chinense* (accession numbers: HQ216168, HQ216007) was used as outgroup. *L. ruthenicum* were devided into two major clades of I and II. Clade I containing eight haplotypes was observed in all sampled populations while clade II with two haplotypes (H3 and H7) was found in two mountain populations (NMH and

GEM) and one Inner Mongolia population (EJNQ). However, only four haplotypes from clade I occurred in more than two populations (Fig. 1). In the Median-joining network based on the mutational steps between haplotypes, two major clades were similarly identified with two mutational steps apart (Fig. 3). Haplotype 3 and 7 were considered to be the ancestral haplotypes when *L. americanum* and *L.* chinense were selected as outgroups.

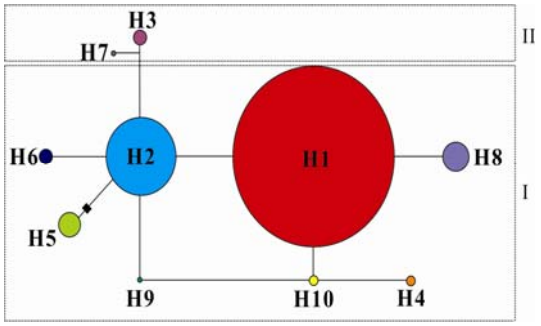

Fig. 3. Median-joining network constructed from the 10 haplotypes in our sample. Haplotype number is shown on the circles, and mutation positions are indicated on the branches linking two haplotypes. Circle areas are proportional to the frequency of the haplotypes. Black squares represent a inversion.

Divergence times of the *Lycium* species were estimated based on the ML tree topology shown in Fig. 2. *L. ruthenicum* and other two *Lycium* species could be traced back to a common ancestor about 2.939-0.979 MYA. The tMRCA (time of the most recent common ancestor) of *L. ruthenicum* was estimated to be in the range of 1.349-0.449 MYA, and the tMRCA of clade I was estimated to be about 0.860 -0.183MYA (Fig. 2).

Population demographic history of *L. ruthenicum* was estimated using GENETREE. The demographic curves of sequential effective population size and showed recent population expansion for *L. ruthenicum* over Pleistocene (Fig. 4a). The demographic curves of historical effective population size among the population of GEM, NMH, MQ, the Western group, and the Eastern group (excluding MQ population) based on the results of FST analyses indicated that NMH population (θ_{ML} =3.82, β_{ML}= 2.47) and the western group (θ_{ML} =0.78, β_{ML}= 0.645) had experienced rapid expansion in both instances, whereas the population of GEM (θ_{ML} =1.18, β_{ML}= 0.115), MQ(θ_{ML} =1.04, β_{ML}= 0.043), and the eastern group(excluding MQ population, θ_{ML} =1.22, β_{ML}= 0.00098) had maintained a relatively stable population size since this species diverged from the others (Figs. 4b and 4c).

The Bayesian approach estimates of long-term gene flow using MIGRATE indicated all of the migration rate (M) values were non-zero, suggesting the occurrence historic gene exchange between groups or populations. The results indicated the migration rates M were higher from the western group to the mountain group ($M_{W\rightarrow M}$ =14.5, 95%CIs= (6.0, 20.0)) whereas the reciprocal migration rates were lower ($M_{M\rightarrow W}$ =11.5, 95%CIs= (2.2, 20.0)) (Fig. 5a). The migration analysis among group and population based on the results of F_{ST} analyses indicated that the main direction of movement of migrants were upward and northwards from the Xingjiang region (western group) to NMH population (one of the plateau populations) ($M_{W\rightarrow N}$ =13.7, 95%CIs= (4.8, 20.0)) (Fig. 5b), and there was no evidence of asymmetrical gene flow between western group and eastern group, as well as between eastern group and mountain group (Fig. 5), although the pairwise F_{ST} value between MQ and the other populations was higher.

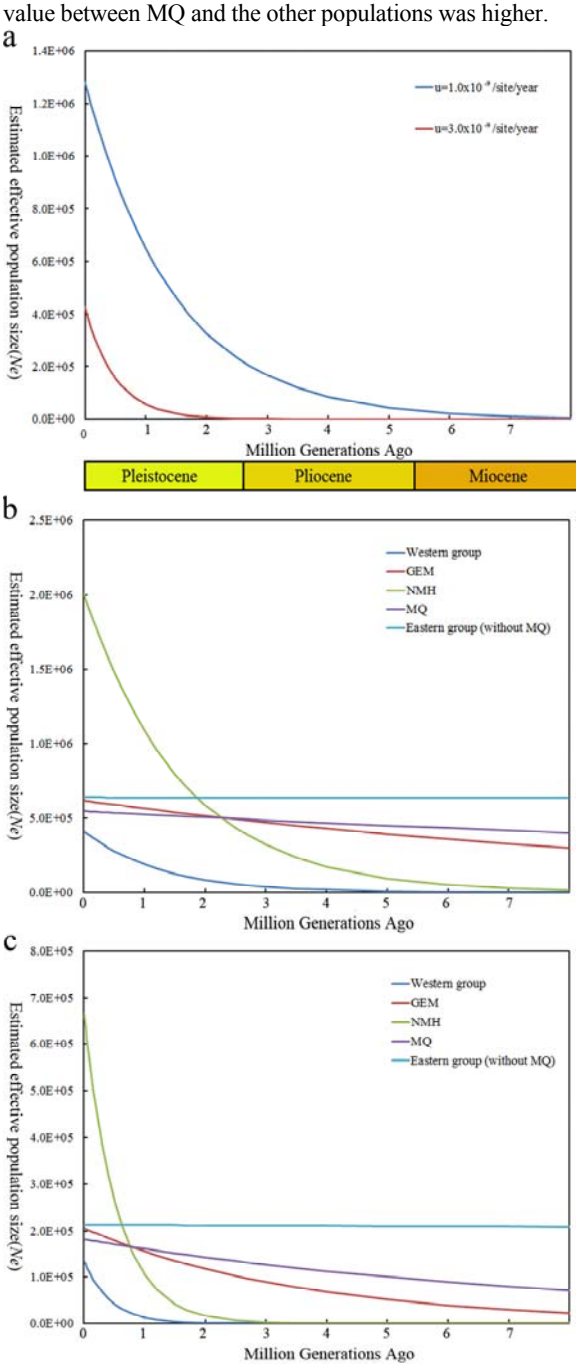

Fig. 4. Population demography pattern of *Lycium ruthenicum* was estimated from chloroplast DNA trnH-psbA and rps16-trnK spacer sequence data by the GENETREE program. The vertical axis in graph represents ancestral *Ne* relative to current population size; horizontal axis is time, measured in millions of years. a: historical population dynamics of this species was estimated using two mutation rate of 1.0×10^{-9}/site/year and 3.0×10^{-9}/site/year. b: The demographic curves of historical effective population size among the GEM, NMH, MQ

population, the Western group, and the Eastern group (excluding MQ population) using a mutation rate of 1.0×10^{-9}/site/year. c: The demographic curves of historical effective population size among the GEM, NMH, MQ population, the Western group, and the Eastern group (excluding MQ population) using a mutation rate of 3.0×10^{-9}.

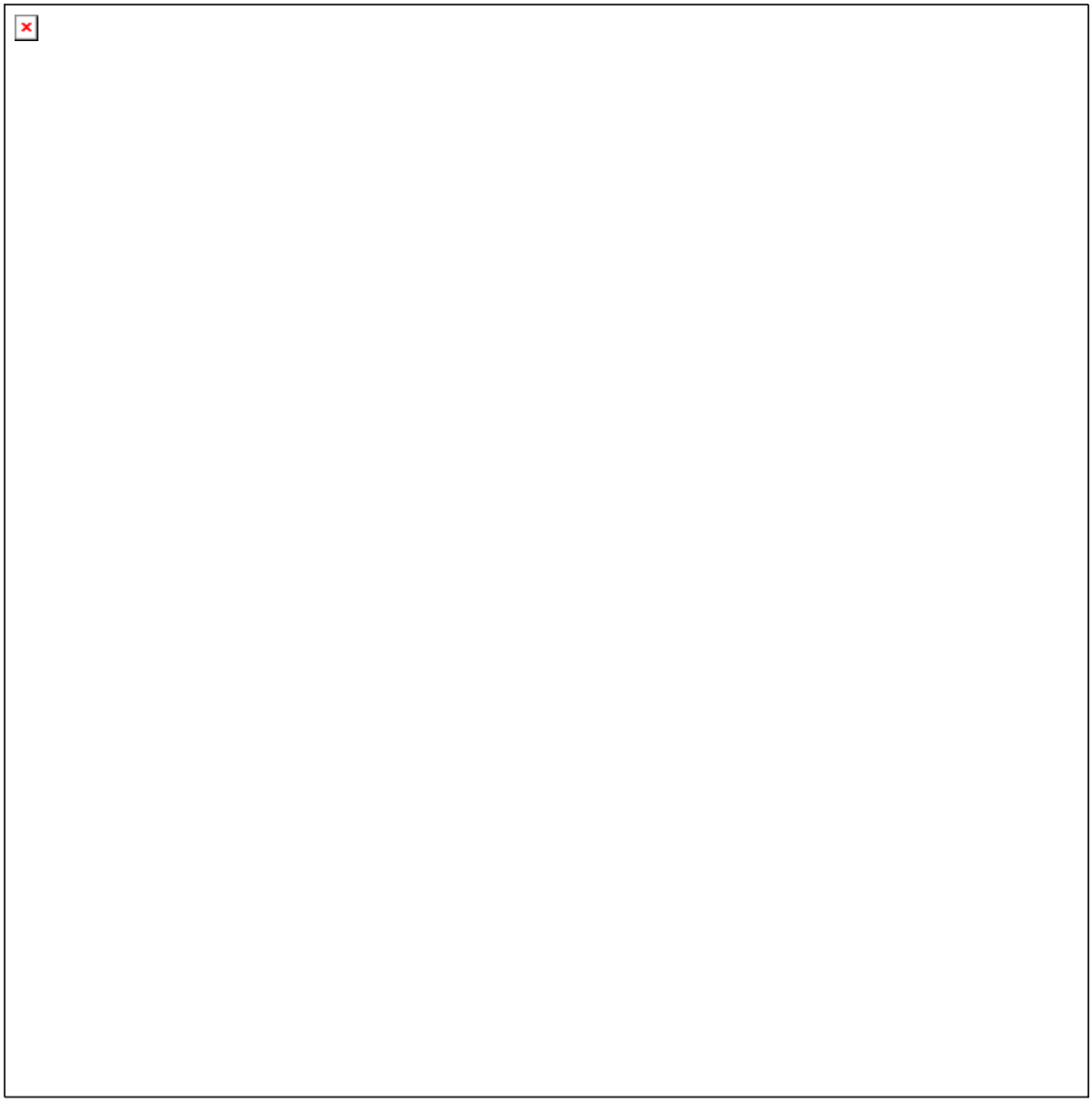

Fig. 5. Posterior probabilities of parameter estimates in *Lycium ruthenicum* from MIGRATE-N. a: The long-term migration rates (M=m/u, where m is the migration rate, and u is the mutation rate, arrows) and the mutation-scaled effective population size (Θ, ovals) of the Mountain group, the Western group, and the Eastern group populations. b: The long-term migration rates (M, arrows) and the mutation-scaled effective population size (Θ, ovals) among the population of GEM, NMH, the Western group, and the Eastern group (excluding MQ population). Values in parenthesis indicate the 95% CIs.

Discussion

The salt-tolerance and drought-resistant *L. ruthenicum*, adapted to the desert habitat in north China showed a high level of haplotype diversity (0.733) and low levels of nucleotide diversity (0.00141). Most polymorphism were probably generated during the adaptation process to specific ecological environment of sand, wind and drought as well as saline and alkaline soil in the arid and semiarid lands., New haplotypes were accumulated by mutations attributed to its long evolutionary history and demographic expansion , but the accumulation of nucleotide sequence diversity were failed to acquire (Avise, 2000), which may due to the low mutation rate in the chloroplast genome. Likewise, this result is similar to the perennial herbaceous plant *Lagochilus ilicifoli*us (h =0.8824, π=0.0016) from semidesert and desert areas in northern China (Meng & Zhang, 2011).

It should be noted that our genetic analyses further suggested that both total and within-population diversities were high in this perennial shrub (H_T=0.762, H_S=0.495). Comparing with the other studied shrub species occurring in the deserts of north China, the gene diversity indices in *L. ruthenicum* are comparable with the rare species *Gymnocarpos przewalskii* mainly restricted to the deserts

of north-western China (H_T=0.903, H_S=0.425; Ma *et al*., 2012), whereas *Juniperus sabin*a which are higher than the small shrub have widespread occurrence in arid rocky and sandy habitats in northern China (H_T=0.577, H_S=0.043; Guo *et al*., 2010). Comparing with the other studied desert species occurring in north China, the genetic diversity indices were relatively low, for example, *L. ilicifoli*us (H_T=0.925, and H_S=0.988; Meng & Zhang, 2011).

The populations of the eastern group had the highest levels of nucleotide diversity (0.00154) in the chloroplast DNA compared with the other two groups. A great number of genetic polymorphisms of the eastern group populations is probably due to the stable population size in the eastern group during a long evolutionary history, and a variety of habitats could be provided under the habitat ranges of the eastern group consisting of Badain Jaran-Tengger Desert are wider than those of the other two groups for *L. ruthenicum*. DH population from the eastern group was fixed a single dominant haplotype H1, which is likely to be attributable to founder and bottleneck effect during the Quaternary climatic oscillation (Hewitt, 2000).

The high genetic differentiation between populations within desert species in north China were suggested in many previous studies, for example, *J. sabina* (N_{ST}=0.980, and G_{ST}=0.926; Guo *et al*., 2010), *L. ilicifolius* (N_{ST}=0.911, and G_{ST}=0.799; Meng & Zhang, 2011), and *G. przewalskii* (N_{ST}=0.752, and G_{ST}=0.547; Ma *et al*., 2012), and the strong phylogeographic structure was evident in those species. However, relatively low estimates of interpopulation differentiation were recorded in *L. ruthenicum* (G_{ST}=0.351, N_{ST}=0.304), suggesting the absence of a geographic component in the chloroplast DNA genetic structure of *L. ruthenicum* populations. This was confirmed by the analyses of SAMOVA (Table 3), and AMOVA, which all failed to reveal strong phylogeographical groupings.

The pairwise comparisons of chloroplast DNA F_{ST} estimates revealed high levels and significance of genetic differentiation between the MQ population and others (F_{ST}=0.2813 - 0.67805, Table 4), possible because recent population bottleneck by chance in the patchy and highly discontinuous distributions of the MQ population due to the MQ population experienced frequently disturbance by human activities (Ma *et al*., 2007).

The western group, YJ and TS populations is situated in the northern Tarim Basin, in Xinjiang Province, surrounded by Taklamakan Desert and Tian Shan mountain ranges, whereas the mountain group, GEM and NMH population in the middle of the Qaidam Basin on the Qinghai-Tibetan Plateau (QTP) with altitudes of about 2800 m, surrounded by the Kunlun, Altun and Qilian mountains with the elevations exceeding 5000m. Genetic differentiation would be detected between these populations of different regions. However, it is interesting that a noteworthy pattern of high differentiation was found between one mountain population NMH and the other populations (F_{ST}=0.1047-0.6154, Table 4), whereas the pariawise F_{ST} estimates exhibited low levels of genetic differentiation between the other one mountain population GEM and the western group (YJ and TS, F_{ST}=0.0225-0.0607), and one population of the eastern group (EJNQ, F_{ST}=0.0386). According the GENETREE demographic curves for chloroplast DNA, a recent population expansion for *L. ruthenicum* across the Pleistocene (Fig. 4a) was identified, particularly the population size of NMH and the western group experienced rapid growth (Fig. 4b and 4c), whereas the relatively stable population sizes were maintained within the GEM population and the eastern populations. Given the Bayesian approach estimates of long-term gene flow using MIGRATE-N, the evidence of asymmetrical gene exchange, with higher levels of migration were found from the western group to the two mountain populations than the eastern group (MW→ M=14.5 against MW→ E=12.2), whereas the reciprocal migration rate was low within the mountain group and equivalent within the eastern group (M M→W=11.5 against M E→W=12.2) (Fig. 5), suggest the genomic composition of the mountain populations were higher attributable to the genetic material from the western group, although the close genetic similarity and strong gene flow among populations or groups (M>10) might be due to the existence of corridors during glacial maxima. Therefore, given above several factors, the genetic structure would be explained that rapid population growth of NMH population after a period of low effective population size enhances the retention of large numbers of new mutations, even frequent exchange of genes among each populations and immigration from other population for a long time, but still not sufficient to reduce genetic differentiation between NMH population and other populations; from another view, the other mountain population GEM did not undergo rapid expansion, and maintained a relatively stable population sizes in the evolutionary history although this population was geographically close to NMH population, and failed to accumulate or retain new mutations, then the close genetic similarity would be occurred among them after the strong exchange of gene flow.

The analysis of the ML tree and the network suggested 2 potential separate refugia inferred in the mountain populations (NMH and GEM) located on the Qaidam Basin and in the eastern populations due to the ancestral haplotypes, unique genotypes, and high levels of genetic diversity were detected (Comes & Kadereit, 1998; Hewitt, 2000). In the eastern region, the ancestral haplotype (H7) were restricted in the Inner Mongolia population (EJNQ) with low frequency, suggesting the replacement of ancestral haplotypes with new haplotypes after the glacial maximum was perhaps occurred in the eastern populations except EJNQ population. The refugia in the Inner Mongolia are consistent with phylogeographic histories of other desert plant in this region (Ma *et al*., 2012; Guo *et al*., 2010). Meanwhile, the ancestral haplotype (H3) were found in both populations with higher frequency in the Qaidam Basin on the QTP, where it acted as the other refugia for *L. ruthenicu*m, despite the Qaidam Basin experienced extremely harsh and dry events during the late Pleistocene (Jing *et al*., 2001).The inference was consistent with the population demography pattern that NMH population not only survived in glacial refugia, but also maintained the trend of rapid population growth; whereas GEM population retain relatively stable and lager population size since the initial divergence among *L. ruthenicu*m lineages, suggesting there was little

glacier effect in the Qaidamu Basin during the Last Glacial Maximum.

The molecular estimates of divergence times revealed that *L. ruthenicu*m might have originated during the end of the Early Pleistocene/ the Middle Pleistocene (1.349-0.449MYA), as is well consistent with the initial formation of the Badain Jaran Desert (Tan, 1964; Wang, 1990), as well as coincided with geological assumptions of the second major rapid uplift phases of Tibetan Plateau, about 1.1-0.6 MYA (Li & Fang, 1999; Shi *et al*., 1999). The major cladogenetic events between two *L. ruthenicu*m lineages (clade I and clade II) occurred during the middle/later Pleistocene (0.860- 0.183MYA), before the onset of the last glacial maximum (LGM). At this period, the high mountains known as the Himalayas, West Kunlun Mountains had risen to 5,000m, the interior and the north side of the QTP became more and more drought, resulting the enlargement of Taklimakan desert in the northwest side of QTP by the developments and interconnections of a number of scattered desert (Shi *et al*., 1999). Meanwhile, the Tian Shan Mountains underwent uplifting the last two main ages (6.7-0.73 MYA and less than 0.73 MYA; Liu *et al*., 2004), and the Badain Jaran-Tengger Desert was expanding faster than before (Wang, 1990; Sun *et al*., 1998; Yang *et al*., 2006). Therefore, the divergence of this species including all two clades was caused by the progressive uplift of the QTP, acidification, and the expansion of deserts in the north China.

Conclusions

A high level of chloroplast DNA haplotype diversity coupled with low nucleotide diversity were revealed within the desert plant of *L. ruthenicu*m. No clear phylogeographic structure was detected. Compared with other desert species in north China, relatively lower estimates of interpopulation differentiation were observed in this species. Taken together the results of pairwise Fst, gene flow, and population dynamics, the populations from different regions or the same region underwent different evolutionary history that respond to different regional environmental changes, especially in the two mountain populations separated by as little as 170 km in the Qaidam Basin. Two glacial refugia were inferred in the middle of Qaidam Basin and the western Inner Mongolia. Cladogenesis of the identified haplotype lineages were associated with the uplift of the mountain movement, as well as the formation and evolution of desert distributed in/around the plateau during the Pleistocene. Thus, the unique chloroplast phylogeography of *L. ruthenicu*m in the current study is not similar with the phylogeographical studies of the other limited number of desert species in the in north China (Guo *et al*., 2010; Meng & Zhang, 2011; Ma *et al*., 2012), which may be due to the combined effects of the Quaternary glacial climate oscillations and geomorphologic processes during the Quaternary.

Acknowledgments

This work was supported by the scientific research project of Ningxia higher education (2010JY001), Ningxia Higher School Scientific research project (No-NGY20140155), The specimen platform of China (Teaching specimens sub-platform), Development of Beifang university of nationnalities for the national natural science foundation of China (2012QZP09), and the National Natural Science Foundation of China (31200277).

References

Avise, J.C. 2000. Phylogeography: The History and Formation of Species. Cambridge: Harvard University Press.

Bandelt, H.J, Forster, P., Röhl, A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.*, 16: 37-48.

Beerli, P. and J. Felsenstein. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, 152: 763-773.

Beerli, P., J. Felsenstein. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc. Nat. Acad. Sci*., 98: 4563-4568.

Chen, H.K., J.M. Cao, X. Ren, S.P. Huang and S. Zhang. 2008. Karyotype analysis of the chromosome of *Lycium ruthenicum Murr*. *North. Hortic* 7: 207-209.

Chen, Y. and H. Tang. 2005. Desertification in North China: background, anthropogenic impacts and failures in combating it. *Land Degradation & Development*, 16: 367-376.

Comes, H.P. and J.W. Kadereit. 1998. The effect of quaternary climatic changes on plant distribution and evolution. *Trends in Plant Sci.*, 3: 432-438.

Doyle, J.J. and J.L. Doyle. 1987. A rapid isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*, 19: 11-15.

Dupanloup, I., S. Schneider and L. Excoffier. 2002. A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.*, 11: 2571-2581.

Efron, B. and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross validation. *American Statistician*, 37: 36-48.

Excoffier, L., G. Laval and S. Schneider. 2005. Arlequin version 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinformatics*, 1: 47-50.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791.

Gan, Q.M., G.F. Luo, P.Y. Li, Z.D. Drolma, Y.R. Chen and Z.C. Zuo. 1997. The development and utilization of *Lycium ruthenicum* in Tibetan Medicine. *Qinghai Science Technology*, 4: 17-19.

Gilani, S.A., Y.F. Jii, A.K. Chi, Z.K. Shinwari and K. Watanabe. 2011. Ecological consequences, genetic and chemical variations in fragmented populations of a medicinal plant, *Justicia adhatoda* and implications for ITS conservation. *Pak. J. Bot*., 43: 29-37.

Griffiths, R.C. and S. Tavaré. 1994. Simulating probability distributions in the coalescent. *Theory Population Biology*, 46: 131-159.

Guo, Y.P., R. Zhang, C.Y. Chen, D.W. Zhou and J.Q. Liu. 2010. Allopatric divergence and regional range expansion of *Juniperus sabina* in China. *J. Sys. & Evol.,* 48: 153-160.

He, F.L., M. Zhao, J.H. Wang, Q.S. Yu and J.C. Zhang.. 2011. Response to droughty stresses and drought-resistances evaluation of seed germination of four desert vegetation. *Arid Land Geogr.*, 34: 100-106.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. *Nature*, 405: 907-913.

Jing, M.C., Z.C. Sun, D.M. Li, G.L. Yang and P. Yang. 2001. Ilyocypris inermis extinction event and its palaeoenvironment significance in the Qaidam Basin at about 30 ka.B.P. *Geoscience*, 15: 49-57. (In Chinese with English abstract)

Kuang, K. and A. Lu (eds.). 1978. Solanaceae. Flora Reipublicae Popularis Sinicae. 67(1):10.

Li, J.J. and X.M. Fang. 1999. Uplift of the Tibetan Plateau and environmental changes. *Chinese Science Bull*., 44: 2117-2124.

Li, J.M., Z.X. Jin and J.J. Gu. 2012. Genetic isolation by distance in the endangered plant *Sinocalycanthus chinensis* endemic to China. *Pak. J. Bot.,* 44(4): 1275-1280.

Librado, P. and J. Rozas. 2009. DNA SP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25: 1451-1452.

Liu, Y., Z.X. Wang, X. Jin, T. Li and Y. Li. 2004. Evolution, chronology and depositional effect of uplifting in the eastern sector of the Tian shan Mountains. *Acta Geol. Sin*., 78: 319-331.

Ma, S.M., M.L. Zhang and S.C. Sanderson. 2012. Phylogeography of the rare *Gymnocarpos przewalskii* (Caryophyllaceae): indications of multiple glacial refugia in north-western China. *Aust. J. Bot.,* 60: 20-31.

Ma, Y.H., S.Y. Fan, L.H. Zhou, Z.Y. Dong, K.C. Zhang and J.M. Feng. 2007. The temporal change of driving factors during the course of land desertification in arid region of North China: the case of Minqin County. *Environment Geology*, 51: 999-1008.

Meng, H.H. and M.L. Zhang. 2011. Phylogeography of *Lagochilus ilicifolius* (Lamiaceae) in relation to Quaternary climatic oscillation and aridification in northern China. *Biochem. Syst. & Ecol.,* 39: 787-796.

Nei, M. and F. Tajima. 1983. Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. *Genetics*, 105: 207-217.

Pons, O. and R.J. Petit. 1996. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*, 144: 1237-1245.

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution*, 43: 223-225.

Shah, M.S., S.A. Naveed and M. Arif. 2013. Genetic diversity in basmati and non-basmati rice varieties based on microsatellite markers. *Pak. J. Bot*., 45(S1): 423-431.

Shaw, J., E.B. Lickey, E.E. Schilling and R.L. Small. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am. J. Bot.,* 94: 275-288.

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

Sheehy, D.P. 1992. A perspective on desrtification of grazingland ecosystems in North China. *Ambio*, 21: 303-307.

Shi, Y.F., J.J. Li, B.Y. Li, T.D. Yao, S.M. Wang, S.J. Li, Z.J. Cui, F.B. Wang, B.T. Pan, X.M. Fang and Q.S. Zhang. 1999. Uplift of the Qinghai-Xizang (Tibetan) Plateau and East Asia environmental change during late Cenozoic. *Acta Geogr. Sin* 54: 10-20.

Sun, J.M., Z.L. Ding and T.S. Liu. 1998. Desert distributions during the glacial maximum and climatic optimum: example of China. *Episodes*, 21: 28-31.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.*, 28: 2731-2739.

Tan, J.A. 1964. The types of desert in the Alashan Plateau, Inner Mongolia. Geographical Collections, Series 8. *Beijing: Science Press.* (In Chinese)

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin and D.G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, 25: 4876-4882.

Tu, T., M.O. Dillon, H. Sun and J. Wen. 2008. Phylogeny of Nolana (Solanaceae) of the Atacama and Peruvian deserts inferred from sequences of four plastid markers and the nuclear LEAFY second intron. *Mol. Phylogenet. Evol.*, 49: 561-573.

Tu, T., S. Volis, M.O. Dillon, H. Sun and J. Wen. 2010. Dispersals of Hyoscyameae and Mandragoreae (Solanaceae) from the New World to Eurasia in the early Miocene and their biogeographic diversification within Eurasia. *Mol. Phylogenet. Evol.,* 57: 1226-1237.

Wang, T. 1990. Formation and Evolution of Badain Jirin Sandy Desert, China. *J. Desert Res.,* 10: 29-40.

Watterson, G.A. 1975. On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7: 256-276.

Weir, B. and C.C. Cockerham. 1984. Estimating f-statistics for the analysis of population structure. *Evolution*, 38: 1358-1370.

Wolfe, K.H., P.M. Sharp and W.H. Li. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. *Proc. Natl. Acad. Sci*., 84: 9054-9058.

Xi, J.B., F.S. Zhang, D.R. Mao and P. Yan. 2003. The Utilization of halophytes for Traditional medicine in Xngjiang. Rev. *China Agri. Sci. Tech.,* 5: 43-48.

Yang, D., X.M. Fang, G.R. Dong, Z.C. Peng and J.J. Li. 2006. Aeolian deposit evidence for formation and evolution of the Tengger Desert in the north of China since early Pleistocene. *Mar. Geol. Quaternary Geol*, 26: 93-100.

Yang, X.P., L. Scuderi, P. Paillou, Z.T. Liu, H.W. Li and X.Z. Ren. 2011. Quaternary environmental changes in the drylands of China - A critical review. *Quaternary Science Reviews*, 30: 3219-3233.

Yao, X., Y. Peng, L.J. Xua , L. Li, Q.L. Wu and P.G. Xiao. 2011. Phytochemical and Biological Studies of *Lycium* Medicinal Plants. *Chemistry & Biodiversity*, 8: 976-1010.

Yin, X.L., K.T. Fang, Y.Z. Liang, R.N. Wong and A.W.Y. Ha. 2005. Assessing phylogenetic relationships of *Lycium* samples using RAPD and entropy theory. *Acta Pharmacol. Sin.*, 26(10): 1217-1224.

Zhang, Y.C. and J.N. Zhang. 2004. Studies on morphological structure characteristics of leaves of *Lycium Ruthenicum* in Two Salinity Environments. *Journal of Ningxia University*, 25: 365-367.

(Received for publication 20 July 2013)