CSIRO PUBLISHING

Australian Journal of Botany, 2015, **63**, 467–476
http://dx.doi.org/10.1071/BT14172

Conservation genetics and geographic patterns of genetic variation of the vulnerable officinal herb *Fritillaria walujewii* (Liliaceae)

Zhihao Su[A,D], *Borong Pan*[A], *Stewart C. Sanderson*[B], *Xiaojun Shi*[A] *and Xiaolong Jiang*[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]Shrub Sciences Laboratory, Intermountain Research Station, Forest Service, US Department of Agriculture, Utah 84601, USA.
[C]Shanghai Chenshan Plant Science Research Center, Shanghai Chenshan Botanical Garden, Chinese Academy of Sciences, Shanghai 201602, China.
[D]Corresponding author. Email: suzh@ms.xjb.ac.cn

Abstract. The Chinese herb *Fritillaria walujewii* Regel is an important officinal species that is vulnerable because of over-harvesting. Here, we examined the geographic pattern of genetic variation across the species entire range, to study its evolution process and give implication needed for the conservation. Nine haplotypes were detected on the basis of three chloroplast spacers. The most common haplotype was central in the haplotype network and was distributed widely from the Yili Valley to the eastern Tianshan Mountains. Genetic variation primarily occurred among populations and SAMOVA groups and the analysis of genetic structure showed a significant correlation between genetic and geographical distance. The fragmented distribution of *F. walujewii* in deep valleys may cause gene-flow barriers among distant populations and, along with genetic drift, has caused high genetic structure in the species. We identified Xinyuan County as the centre of diversification of *F. walujewii*, and speculated that populations in the eastern Tianshan Mountains were colonised from the Yili Valley. In relation to conservation management, we identified Xinyuan and Zhaosu County as having a high degree of genetic diversity and these should be the areas of the greatest focus for conservation.

Additional keywords: conservation implications, genetic diversity, genetic structure.

Received 3 July 2014, accepted 19 April 2015, published online 21 May 2015

Introduction

Phylogeography probes the processes that influence the geographical distribution of genealogical lineages (Avise and Walker 1998; Avise 2000). It can make a valuable contribution to conservation biology by providing information for endangered species, such as genetic diversity, population structure and evolutionary history, which are essential to the development of broad-scale conservation strategies and planning of management actions (Avise *et al*. 1987; Moritz 1994; Pope *et al*. 1998; Osborne *et al*. 2000). For example, phylogeographic studies can help in identification of an evolutionarily significant unit (ESU), a phylogenetic unit that is reciprocally monophyletic with its sister clade at a cytoplasmic locus and shows significant divergence of allele frequencies (Ryder 1986; Fraser and Bernatchez 2001). Phylogeographic studies can also identify genetically highly variable populations as the management units for a conservation plan (Ge *et al*. 2011).

The extensive and magnificent Tianshan Mountains in China span 1700 km from west to east, and 250–350 km from south to north (Wei and Hu 1990), separating the Junggar Basin in the north from the Tarim Basin in the south. Because of the abundant precipitation in the form of snow, the Tianshan Mountains have been described as a wet island in the arid north-western region of China (Hu 1998). There are many medicinal plant species distributed in the mountains, such as *Ferula sinkiangensis* K.M. Shen, *Saussurea involucrate* (Kar. et Kir) Sch.-Bip., and *Fritillaria walujewii* Regel. Recently, because of heavy harvesting in the wild, the number of the medicinal plants in the Tianshan Mountains has declined rapidly (Hu 1998), including populations of *F. walujewii* that have become small and isolated. Population-genetic patterns and evolutionary histories of plant species, especially those that are threatened, need to be understood to inform management plans and establish effective conservation strategies.

Fritillaria walujewii is a perennial herb and belongs to the family Liliaceae. Its natural distribution is primarily in Tokkuztara County and Zhaosu County, and other occurrences are found west along the Tianshan Mountains towards Kazakhstan (Wang and Tang 1980). Its natural habitat is *Picea* forests, thickets, meadows and steppes between the altitudes of 1000 and 2600 m above sea level. *Fritillaria walujewii* is a traditional Chinese medicinal plant with extremely important

pharmacological value because the bulb can be used to reduce fever and relieve bronchial symptoms. The species has been listed as vulnerable in the list of rare endangered endemic higher plants in Xinjiang Province (Wang and Tang 1980; Yin *et al*. 2006).

In plants, chloroplast DNA (cpDNA) is known to evolve slowly, with low recombination and mutation rates (Li and Fu 1997; Comes and Kadereit 1998). The maternally inherited cpDNA lineages in natural populations can be used to trace the evolutionary history of a species and often display distinct geographic distributions (Avise 2000). In addition, past and present population dynamics can be also inferred from ecological niche modelling (ENM; Phillips *et al*. 2006; Zhang *et al*. 2013). Here, we probed the evolution history in *F. walujewii*, through combining molecular phylogeography and ENM.

We sampled populations across the extent of the Tianshan Mountains, and assayed genetic variation in three cpDNA spacers (*psb*A–*trn*H, *rps*16 and *trn*S–*trn*G). Specifically, we addressed the following questions: (1) what is the level of genetic variation within and among populations; (2) iIs there a phylogeographic pattern in the species; and (3) are there genetic hotspots that may be significant for conservation?

Materials and methods

Plant materials

In total, 235 individuals from 18 populations of the species were sampled, covering almost the entire geographical range of its distribution. Twelve populations were sampled in the Yili Valley (Populations 1–12) and six in the eastern Tianshan Mountains (Populations 13–18); within the valley, Populations 4–7 were sampled from Zhaosu County, and populations 8–12 were from Xinyuan County. In total, 8–16 individuals were collected per population. Fresh leaves were gathered from each individual and dried in silica gel. We also collected samples of the species *F. pallidiflora* Schrenk for use as an outgroup for the network analysis.

DNA extraction, amplification and sequencing

Total genomic DNA was extracted from silica gel-dried leaf tissue by using a modified 2 × CTAB method (Rogers and Bendich 1985; Doyle and Doyle 1987). The intergenic spacer *trn*H–*psb*A was amplified and sequenced using the primers and protocols of Sang *et al*. (1997), the *trn*S–*trn*G region was amplified and sequenced using the primers and protocols of Shaw *et al*. (2005), and the *rps*16 region was amplified and sequenced using the primers and protocols of Oxelman *et al*. (1997). Amplification products were purified using PCR Product Purification Kits (Shanghai SBS, Biotech Ltd, Shanghai, China), following directions provided by the manufacturer. Sequencing reactions were conducted with the forward or reverse primers of the amplification reactions, using the DYEnamic ET Terminator Kit (Amersham Pharmacia Biotech, Shanghai, China), with an ABIPRISM3730 automatic DNA sequencer from Shanghai Sangon Biological Engineering Technology and Services Co. (Shanghai, China). Electropherograms were edited and assembled using SEQUENCHER 4.8 (Gene Codes, Ann Arbor, MI, USA). Sequences were aligned using the CLUSTAL W program (Thompson *et al*. 1994) and refined by visual inspection, and the alignments were then adjusted manually. Indels were coded as single binary characters (Simmons and Ochoterena 2000). Haplotypes were identified by the program TCS v. 1.21 (Clement *et al*. 2000).

Data analysis

HAPLONST (http://www.pierroton.inra.fr/genetics/labo/Software/index.html) was used to estimate within-population diversity (h_S), total gene diversity (h_T), and genetic differentiation (G_{ST}) at the species level, as well as population subdivision for phylogenetically ordered alleles (N_{ST}). Standard-diversity indices, including haplotype diversity (h; Nei 1987), mean number of pairwise differences (π; Tajima 1983), and nucleotide diversity (π_n; mean number of pairwise differences per site; Nei 1987) were calculated for each location and for groups of locations by using the program ARLEQUIN v.3.01 (Excoffier *et al*. 2005). To test for isolation-by-distance, pairwise estimates of F_{ST} (from ARLEQUIN) and the natural log of geographic distances between locations (calculated in PASSAGE v. 1.1; Rosenberg 2001) were correlated using the program IBD v. 1.52 (Bohonak 2002). A haplotype network of all sequences was constructed using statistical parsimony (Templeton *et al*. 1992) and a maximum connection limit equal to 40 steps was implemented in the program TCS v. 1.21 (Clement *et al*. 2000). We use Bayesian inference (BI), as implemented in MrBayes v. 3.0 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003), to investigate the phylogenetic relationships of the cpDNA haplotypes. In BI analyses, we used Modelltest 3.7 to determine the appropriate nucleotide substitution model (Posada and Crandall 1998), and two separate runs were performed. Each included four chains running for 5 000 000 iterations, with one tree sampled every 100 iterations. The first 25% of the run was treated as burnin and not used for subsequent calculations of tree statistics. A 50% majority rule consensus tree was constructed, and posterior probabilities of nodes were recorded.

To test the spatial genetic structure of cpDNA haplotypes, spatial analysis of molecular variance was performed using the program SAMOVA v.1.0 (Dupanloup *et al*. 2002), so as to define groups of populations (K) that are geographically homogeneous and genetically differentiated from each other, and the analysis was run for $K = 2$–17. Finally, the number of groups maximising the proportion of total genetic variance because of differences among groups of populations (F_{CT}) was retained as the best grouping of populations. On the basis of pairwise differences of the sequences, analysis of molecular variance (AMOVA) was employed to study the genetic structure of the species (Excoffier *et al*. 1992), and the fixation index (F_{ST}) was also estimated. A Mantel test was used to test whether the matrix of genetic distance was significantly correlated with the matrix of geographical distance.

To test for evidence of range expansions, Tajima's D and Fu's F_S statistics were calculated (Tajima 1989; Fu 1997; Jaeger *et al*. 2005; Smith and Farrell 2005). A significant value for D or a significant, large, negative value for F_S may be the result of population expansion (Aris-Brosou and Excoffier 1996; Tajima 1996; Fu 1997). So as to investigate hypotheses of demographic history, the mismatch distribution (MDA) was also calculated.

The shape of the mismatch distribution provides evidence of a sudden population expansion during the history of a species (Slatkin and Hudson 1991; Rogers and Harpending 1992). A unimodal distribution indicates that populations have experienced a recent expansion, and, to test for significance, 10 000 permutations were performed. All expansion tests were implemented in ARLEQUIN v.3.01 (Excoffier *et al*. 2005). If the sudden expansion model was not rejected, we used the relationship $s = 2ut$ to estimate the expansion time (t) (Rogers and Harpending 1992), where s is the total number of mutations, and u is the mutation rate per generation for the whole analysed sequence. The value of u was calculated as $u = 2\,\mu kg$, where μ is the substitution rate per nucleotide site per year ($s\ s^{-1}\ y^{-1}$), k is the average sequence length of the analysed DNA region, and g is the generation time in years. The cpDNA substitution rates for most Angiosperm species have been estimated to vary between 1 and 3×10^{-9} substitutions per site per year ($s\ s^{-1}\ y^{-1}$) (Wolfe *et al*. 1987). Given the uncertainties in these rate values, we used normal distribution priors with a mean of 2×10^{-9} and a standard deviation of 6.080×10^{-10} for cpDNA to cover these rate ranges within the 95% range of the distribution for our estimation of range expansion times (Jia *et al*. 2012). The generation time for this species is 4 years.

On the basis of the geographical distribution of the *Fritillaria walujewii* populations included in the present study (Fig. 1, Table 1), ecological niche modelling was performed in MAXENT, version 3.2.1 (Phillips *et al*. 2006), to construct the present climatic envelope for the species and project it onto two climatic scenarios for the last glacial maximum (LGM). For the past (LGM, ~21 thousand years before present), potential distributions were modelled using the Community Climate System Model (Collins *et al*. 2006) and the Model for Interdisciplinary Research on Climate (Hasumi and Emori 2004), and for the present, models were run using the WorldClim dataset (Hijmans *et al*. 2005). The climatic niche of the species was modelled as a function of six (of 19) BIOCLIM variables screened by principal component analysis. This restricted dataset was used to avoid including highly correlated variables and prevent potential overfitting (Peterson and Nakazawa 2008). To form ENMs, we used the default parameters of MAXENT and the user-selected features, as follows: regularisation multiplier of 3.0, application of a random seed, duplicate-presence records removal, and logistic probabilities used for output (Phillips and Dudík 2008). Model performance was evaluated using the area under the receiver operating characteristic curve (AUC) calculated by MAXENT. We used a jackknife (or 'leave-one-out') procedure to train and test the model. Values over 0.7 indicate good discrimination (Swets 1988).

Results

Sequence analysis

The aligned sequence length was 303 base pairs (bp) for the *trn*H–*psb*A spacer, 833 bp for the *rps*16 spacer, and 686 bp for the *trn*S–*trn*G spacer. In total, eight informative characters were found in the aligned sequence data, including four nucleotide substitutions (Positions 1124, 1153, 1316, 1479) and four indels (Positions 397, 1271, 1723, 1724). Of the 235 sampled individuals from 18 populations, nine haplotypes (A–I), in total, were identified (Table 1). GenBank accession numbers of the cpDNA sequences are KJ956409–KJ956423.

Haplotype geographical distribution and relationships

The geographic distribution of cpDNA haplotypes, along with the frequency of haplotypes in each population, is presented in Fig. 1 and Table 1. Haplotype A was more widespread than were the others; it was distributed along the Tianshan Mountains in the Yili Valley and extended to the eastern Tianshan Mountains beyond the valley. Haplotype B was also widespread in the valley. In addition, there were some rare haplotypes, such as C, D, E, and F in the valley. Haplotype G, and other rare haplotypes including H and I, were absent from the Yili Valley and were restricted to the eastern Tianshan Mountains. The haplotype network (Fig. 2) showed a star-like structure.

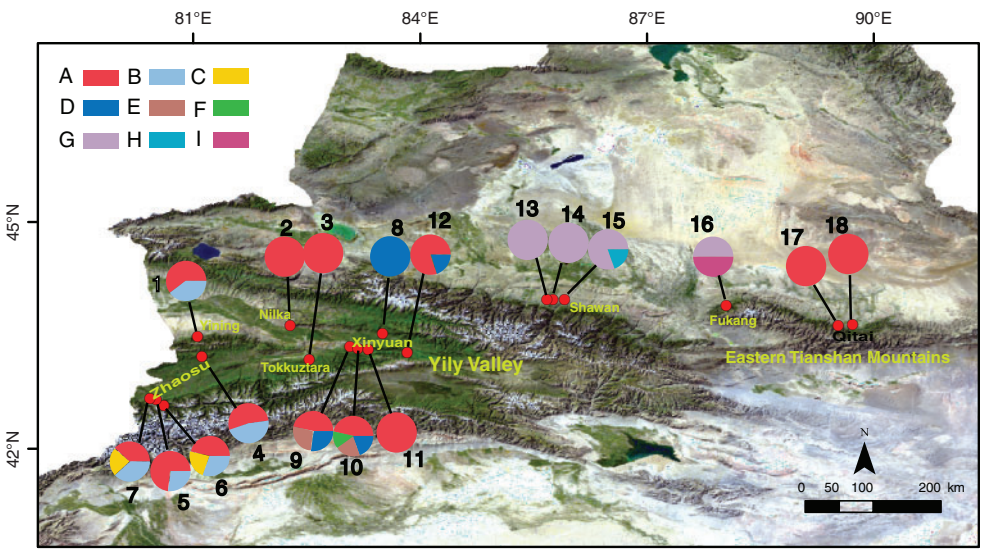

Fig. 1. Geographical distribution of *Fritillaria walujewii* in China. Population numbers correspond to those in Table 1, haplotypes to those in Table 2.

Table 1. Details of sample locations, sample size and haplotype frequencies for 18 populations of *Fritillaria walujewii*
Values in parentheses represent the number of the haplotypes

Region	Number	County	Location	Latitude (N)	Longitude (E)	Altitude (m)	Haplotype
Yili Valley	1	Yining	Gulja	43°29′	81°08′	1744	A (9), B (6)
	2	Nilka	Black hill	43°38′	82°21′	1140	A (13)
	3	Tokkuztara	Talimu	43°11′	82°36′	1400	A (11)
	4	Zhaosu	Stud farm	43°13′	81°11′	2187	A (8), B (7)
	5		Xiata	42°39′	80°35′	2019	A (8), B (3)
	6		Xiata gorge	42°35′	80°41′	2204	A (6), B (4), C (3)
	7		Suolan	42°40′	80°33′	2041	A (5), B (5), C (3)
	8	Xinyuan	Wild fruit forest	43°21′	83°34′	1670	D (8)
	9		Six commune	43°21′	83°13′	1691	A (7), D (4), E (4)
	10		Qiapu river A	43°20′	83°15′	1699	A (7), D (3), E (3), F (2)
	11		Qiapu river B	43°19′	83°18′	1703	A (13)
	12		Nalat Grassland	43°16′	83°54′	1863	A (8), D (2)
Eastern Tianshan	13	Shawan	Shawan A	43°58′	85°53′	1217	G (15)
	14		Shawan B	43°59′	85°52′	1437	G (12)
	15		Shawan C	43°58′	85°50′	1170	G (8), H (2)
	16	Fukang	Tianchi	43°54′	88°07′	1587	G (8), I (8)
	17	Qitai	Tangfangmen	43°38′	89°37′	1630	A (15)
	18		Banjiegou	43°39′	89°44′	1516	A (15)

Table 2. Nine haplotypes of *Fritillaria walujewii* recognised on the basis of three chloroplast DNA sequences (*trn*H–*psb*A, *rps*16, *trn*S–*trn*G)
△, GTAT; ♢, TA

Sequence position	3	1	1	1	1	1	1	1
	9	1	1	2	3	4	7	7
	7	2	5	7	1	7	2	2
		4	3	1	6	9	3	4
Haplotype								
A	–	C	G	–	G	T	T	T
B	–	C	G	–	G	G	T	T
C	–	C	C	–	G	G	T	T
D	–	C	G	–	G	T	T	–
E	–	C	G	–	A	T	T	T
F	–	C	G	♢	G	T	T	T
G	–	T	G	–	G	T	T	–
H	△	C	G	–	G	T	–	–
I	–	T	G	–	G	G	T	–

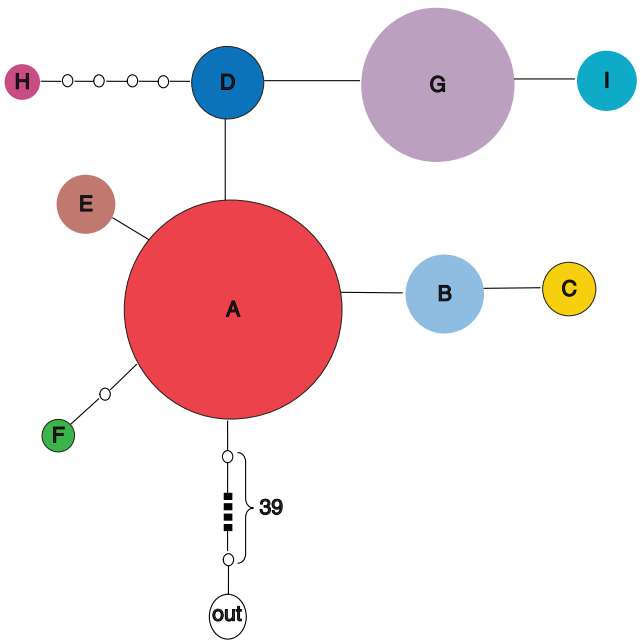

Fig. 2. Cp haplotype network of *Fritillaria walujewii* constructed under the criterion of statistical parsimony. The circle size is proportional to haplotype frequencies. The number of inferred steps between haplotypes is shown near the corresponding branch section. The blank dots represent the missing or inferred haplotypes.

Genetic diversity and genetic structure

Spatial genetic analysis of cpDNA haplotypes using SAMOVA indicated that F_{CT} increased to a maximal value of 0.7276 when K (the number of groups) was raised from $K=2$ to $K=3$. The grouping pattern of populations corresponding to $K=3$ is as follows: (1) Populations 1–7 and 9–12 from the Yili Valley, and 17–18 from the eastern Tianshan Mountains; (2) Population 8 in the Yili Valley; (3) Populations 13–16 from the eastern Tianshan Mountains. Within-population gene diversity (h_S) was 0.308 (s.e. 0.0712), and total gene diversity (h_T) was 0.693 (s.e. 0.0729). Differentiation among populations was moderate ($G_{ST}=0.555$, s.e. 0.1006), indicating some population structure. N_{ST} was 0.648 (s.e. 0.0781), being significantly higher than G_{ST} as shown by the U-test ($U=1.40$, $P<0.01$), indicating significant phylogeographic structure. The AMOVA results provided evidence that 65.02% ($P<0.001$) of the total variation can be explained by differences among populations. When populations were grouped according to geographical region, AMOVA results demonstrated that 72.76% ($P<0.001$) of the total variation occurred among the groups (Table 3). Among all the populations, Population 10 had the highest haplotype diversity, mean number of pairwise differences, and nucleotide diversity. In groups subdivided by SAMOVA, Group

1 had the highest haplotype diversity, mean number of pairwise differences, and nucleotide diversity (Table 4). Mantel's test showed that there was significant correlation between genetic and geographical distance ($r = 0.525$, $P < 0.0001$).

Demography of the groups in Fritillaria walujewii

Demographic analysis of groups and total individuals showed that Group 1 (including Populations 1–7, 9–12, 17–18) experienced range expansion in the past. Range expansion in Group 1 is supported by significant results of Fu's F_S, along with unimodal distributions for the shapes of the mismatch distribution (Fig. 3, Table 5). The time of the geographic range expansion of *F. walujewii* is estimated to have occurred at ~10 600 years ago, which is consistent with the deglaciation period in the last glacial episode (Shi *et al*. 2005).

Table 3. Results of analysis of molecular variance for 18 populations of *Fritillaria walujewii*, based on chloroplast DNA sequence data
*$P < 0.001$

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation (%)
Among populations	17	98.356	0.4263	65.02*
Within populations (1, 3–7, 17, 18) vs (8) vs (2, 13–16)	217	49.772	0.2294	34.98
Among groups	2	79.311	0.8212	72.76*
Among populations within groups	15	19.045	0.0781	6.92*
Within populations	217	49.772	0.2294	20.32*

Phylogenetic analysis

The best nucleotide-substitution model selected by Akaike information criterion (AIC) was HKY + G. In the BI tree, *F. walujewii* is resolved as monophyletic, and this relationship is well supported (1.00 posterior probabilities). Haplotype D is sister to the rest of the haplotypes of *F. walujewii*. Most of the phylogeny is not well supported, but one clade, including Haplotypes I, B and C, received posterior probabilities of >0.8 (Fig. 4).

Past and present distribution of F. walujewii

The test AUC for the ENM was very high (0.999), and the potential range of the species (Fig. 5*a*) was a good representation of the species current distribution, except for areas of high habitat suitability north of the Yili Valley where the species is absent (Fig. 1). Range expansion to the present is shown with both CCSM and MIROC LGM climate models, although there is a considerable discrepancy of potential range during the LGM between the two climate models.

Discussion

Genetic variation of three noncoding spacers of cpDNA in Fritillaria walujewii

Total gene diversity in *F. walujewii* (H_T = 0.693) was high compared with other alpine plants of Liliaceae, such as *Allium przewalskianum* (H_T = 0.612) (Wu *et al*. 2010*a*). The Tianshan Mountains in China are vast and span from west to east 1700 km and from south to north 350 km (Wei and Hu 1990). The largest distance between our sampled populations is 750 km. Such a large regional range of *F. walujewii* may account for the high level of

Table 4. Measures of haplotype diversity, mean number of pairwise differences and nucleotide diversity within sampled locations and SAMOVA groups of *Fritillaria walujewii*, based on three sequences

Parameter	N_{ind}	Haplotype diversity ± s.d.	Mean number of pairwise differences ± s.d.	Nucleotide diversity ± s.d.
Population				
Population 1	15	0.5143 ± 0.0690	0.5143 ± 0.4582	0.0857 ± 0.0856
Population 2	13	0.0000	0.0000	0.0000
Population 3	11	0.0000	0.0000	0.0000
Population 4	15	0.5333 ± 0.0515	0.5333 ± 0.4688	0.0889 ± 0.0876
Population 5	11	0.4364 ± 0.1333	0.4364 ± 0.4218	0.0727 ± 0.0793
Population 6	13	0.6923 ± 0.0750	0.9231 ± 0.6796	0.1538 ± 0.1273
Population 7	13	0.7051 ± 0.0640	0.8974 ± 0.6665	0.1496 ± 0.1249
Population 8	8	0.0000	0.0000	0.0000
Population 9	15	0.6857 ± 0.0683	0.8381 ± 0.6306	0.1397 ± 0.1178
Population 10	15	0.7333 ± 0.0843	1.1810 ± 0.8020	0.1476 ± 0.1124
Population 11	13	0.0000	0.0000	0.0000
Population 12	10	0.3556 ± 0.1591	0.3556 ± 0.3753	0.0593 ± 0.0707
Population 13	15	0.0000	0.0000	0.0000
Population 14	12	0.0000	0.0000	0.0000
Population 15	10	0.3556 ± 0.1591	2.1333 ± 1.2932	0.2370 ± 0.1625
Population 16	16	0.5333 ± 0.0456	0.5333 ± 0.4672	0.1067 ± 0.1047
Population 17	15	0.0000	0.0000	0.0000
Population 18	15	0.0000	0.0000	0.0000
SAMOVA group				
Group 1		0.4603 ± 0.0429	0.5835 ± 0.4723	0.0729 ± 0.0653
Group 2		0.0000	0.0000	0.0000
Group 3		0.3237 ± 0.0740	0.7054 ± 0.5401	0.0784 ± 0.0666

genetic diversity in the species. The varied habitats occupied by *F. walujewii* may harbour locally adapted haplotypes as a consequence of differential geology and topography across the species broad range.

Compared with total gene diversity, within-population gene diversity was relatively lower, resulting in a moderate level of differentiation among populations. The moderate level of differentiation among populations or groups was also supported by the Mantel test and AMOVA analyses. Numerous geographic barriers within the species range are likely to promote vicariant processes and account for the moderate genetic differentiation among population groups shown here. The Tianshan Range in China contains more than 20 east–west mountains and valleys, and the altitude of the main mountains exceeds 4000 m (Wei and Hu 1990) and the Yili Vally is subdivided into five main valleys (Zhang 2006). The multiple deep river valleys of variable dimensions may obstruct gene flow and increase genetic differentiation among populations.

Phylogeographic patterns in Fritillaria walujewii

Climate oscillation during the Quarternary is usually considered an important factor influencing the current geographical distribution patterns and population-genetic structures (Hewitt 2004). Previous palynological and phylogenetic research on plant responses to Quaternary climate change has generally agreed that the fundamental response of most temperate plant taxa in the northern hemisphere to Quaternary climate changes was migration (Forester *et al*. 2013), compared with persistence as a major response in temperate species in southern hemisphere systems (Byrne 2008). The climate niche modelling showed that the potential present range approximately includes the existing locations where the species is found, but also includes areas of high habitat suitability north of the Yili Valley, from which the species is currently absent. We speculate this was caused by recent land development, or is due to long-term barriers to dispersal over historical time. The different results of CCSM and MIROC models (Fig. 5*b*, *c*) might be caused by the different data used by them; usually, MIROC model will give a wider range of predictions (Su and Zhang 2013). Although there is discordance, the predicted range of the species at the LGM was significantly smaller than present in both models. During the LGM, the climate was cooler and drier than present, and the harsh environment would be expected to force species to contract to habit with suitable conditions, and restrict the distribution to a smaller range.

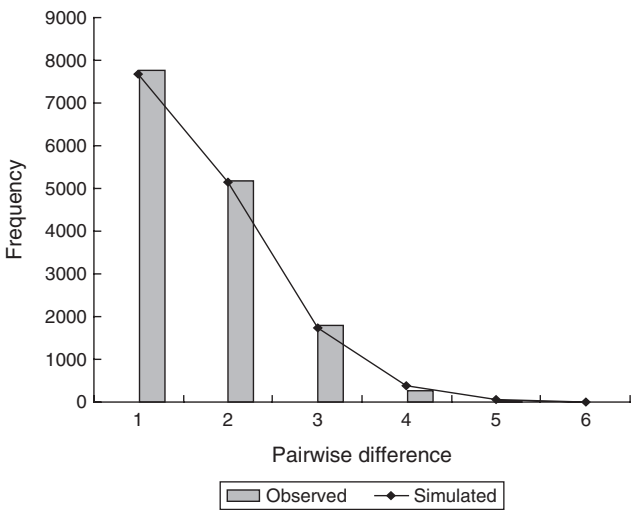

Fig. 3. Mismatch distribution analysis for chloroplast DNA data for Group 1 that includes Populations 1–7, 9–12, 17–18 (SSD = 0.0000, $P = 0.87$).

We interpreted the patterns of diversity as locations with high levels of genetic variation and unique haplotypes as being possible sites of refugia or as a centre of diversification (Taberlet and Cheddadi 2002), and locations with low levels of genetic variation as possible sites of population expansion (Fehlberg and Ranker 2009). In comparison of genetic diversity among the populations, populations in Xinyuan County have the highest haplotype diversity, and contain the common Haplotype A that is at the centre of the network. Thus, we suggest Xinyuan region as a possible centre of diversity or a refugium for *F. walujewii*. This hypothesis is consistent with the LGM refugial areas inferred from ENMs. Compared with current distribution, locations in Xinyuan County are persistent in current and past LGM-modelled areas compared with locations in Zhaosu County, Yining County and the eastern Tianshan Mountains that are absent from the distribution modelled at the LGM. This implies that Xinyuan Country might be a refugial area, and other areas might be colonisation zones as a result of post-glacial warming or *F. walujewii* may have persisted in small populations in localised areas of suitable habitat with reduced diversity and differentiation.

Although there is no palynological or macro-fossil evidence, there are signatures of range expansion in the significant value of Fu's F_S, as well as a unimodel mismatch distribution, that support our speculation about the colonisation of the species. The Zhaosu and Yining counties did not appear to retain suitable habitat at the LGM on the basis of ENM, suggesting that the species contracted from the area and recolonised during warm and wet climate in the post-glacial period that is likely to have provided suitable

Table 5. Results of neutrality tests and mismatch distribution analysis for two groups and total individuals in *Fritillaria walujewii*

Group (1) includes populations 1–7, 9–12, 17–18, and group (3) includes populations 13–16. τ, time in number of generations elapsed since the sudden expansion episode; Hrag, the Harpending's raggedness index; SSD, sum of squared deviations

Group	τ	SSD (*P*-value)	Hrag (*P*-value)	Tajiam's *D* (*P*-value)	Fu 's F_S (*P*-value)
Group 1	0.62	0.0004 (0.7)	0.1074 (0.44)	–0.2629 (0.45)	–1.6185 (0.248)
Group 3	3.00	0.0103 (0.36)	0.2510 (0.56)	–0.4209 (0.318)	1.1930 (0.714)
Total	1.738	0.0063 (0.62)	0.0414 (0.87)	0.1693 (0.647)	–0.9944 (0.34)

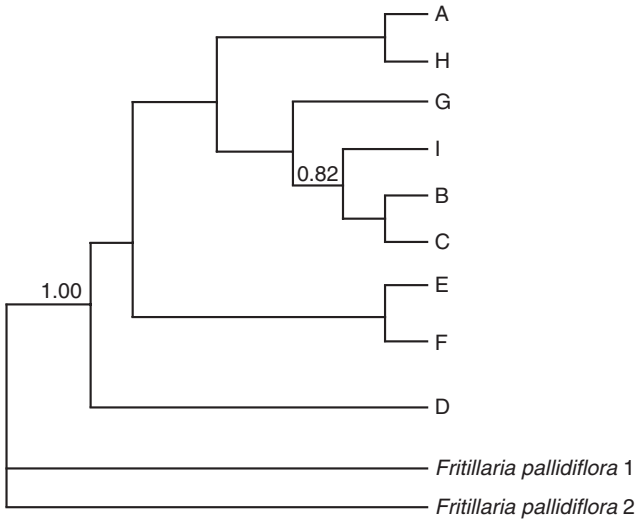

Fig. 4. Phylogenetic relationships of nine haplotypes of *Fritillaria walujewii* and related species. Numbers above branches are support values (posterior probabilities of >0.80).

conditions for the species to expand out from a refugium. However, populations in this area showed moderate diversity, which was greater than would be expected with a simple model of recent expansion. Nevertheless, range expansions can show more complex patterns because various forms of dispersal can affect colonisation, leading to allele surfing and patchy distributions of haplotypes (REFS). Spatial expansions can generate allele-frequency gradients, promote the surfing of some variant into newly occupied territories, and induce the structuring of newly colonised areas into distinct sectors of low genetic diversity (Excoffier and Ray 2008; Excoffier *et al.* 2009). Thus, the haplotype diversity in the Zhaosu and Yining counties may be due to such patterns of colonisation. The patterns of diversity to the east of the proposed refugium in Xinyuan County are also complex, because populations in the nearer Shawan and Fukang counties shared no haplotypes with the Xinyuan County, whereas populations in the further eastern Tianshan Mountains were fixed for the common Haplotype A. Although the ENM showed that there was no suitable habitat at the LGM in the area of the sampled populations in the Shawan county and the eastern Tianshan Mountains, it did show occurrence of some suitable habitat in the intervening area and the population at Fukang. The presence of differentiated haplotypes in Shawan and Fukang counties may have been the result of complex patterns of colonisation, leading to fixation of rare haplotypes at the front of an expansion wave from a refugium in Xinyuan County, or may have resulted from colonisation from populations that persisted in areas of suitable habitat around Fukang.

Similarly, the presence of the common Haplotype A in the eastern Tianshan Mountains may have been the result of long-distance colonisation from the refugium in Xinyuan County, or could have resulted from colonisation from populations that persisted in areas of suitable habitat in between. A variant such as Haplotype A might surf on the wave of advance of the refugium range expansion, reaching high frequencies, spreading to the other suitable habitats in the Yili Valley, and spreading

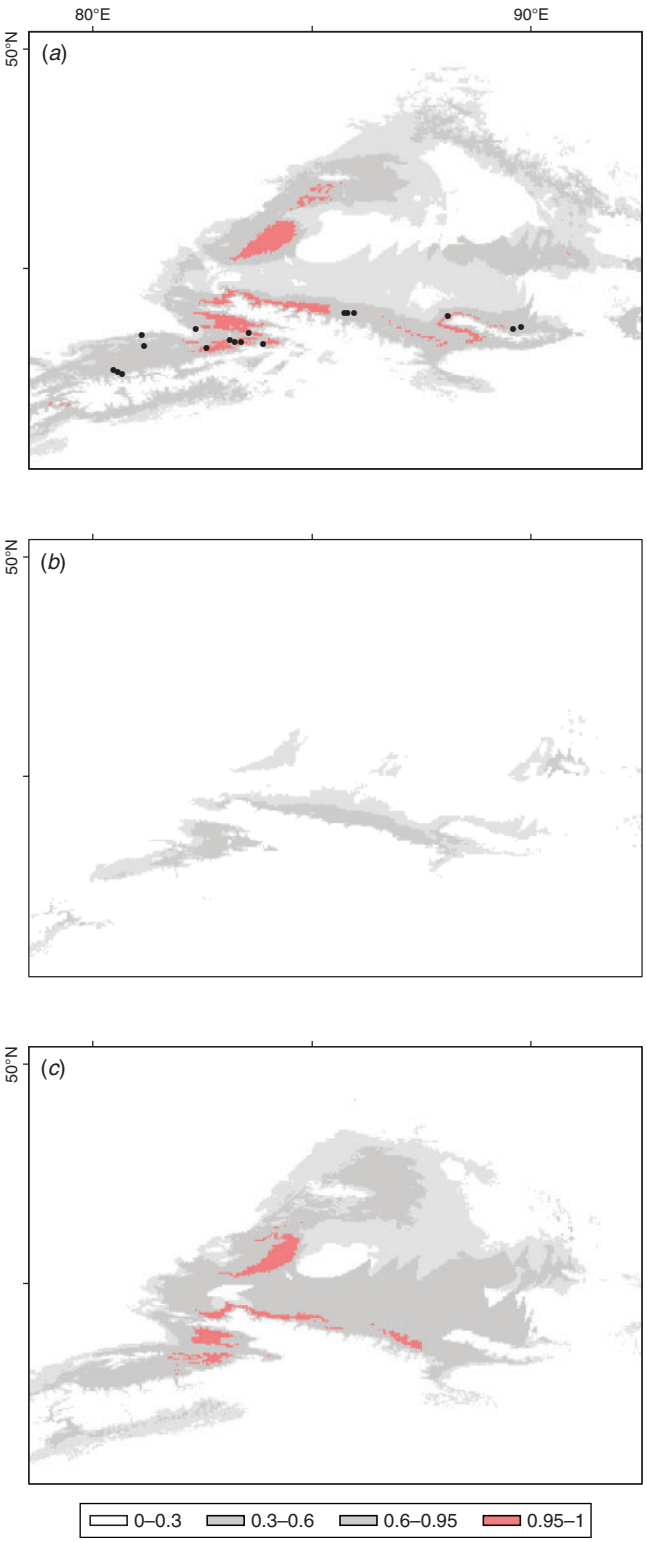

Fig. 5. Maps depicting potential distribution as occurrence probability of the *Fritillaria walujewii* in the Tianshan Mountains during (*a*) present-day and (*b*) last glacial maximum (LGM), based on the Community Climate System Model. (*c*) LGM, based on Model for Interdisciplinary Research on Climate models as derived from the ecological niche models (ENMs) in MAXENT.

to the eastern Tianshan Mountains from the Yili Valley. During the multiple processes of colonisation, genetically distinct populations from dispersed refugia would meet, and intermingled isolates of the genetic variations were found (Nichols and Hewitt 1994; Mellick *et al*. 2012). Some populations of *F. walujewii* that have mixed distributions of haplotypes, such as Populations 1, 4–7 and 12, might be an example of this colonisation pattern.

On the basis of this, we suggest the following the migration route for the species: populations distributed in the Yili Valley extended outwards to the eastern Tianshan Mountains, and during the immigration, some rare haplotypes such as G, H, and I evolved and were fixed in locations along the route.

Conservation implications for Fritillaria walujewii

Genetic drift is expected to occur in small populations, especially when gene flow between populations is restricted (Frankham *et al*. 2002); consequently, this may cause a reduction of genetic diversity. Endangered species with fragmented distributions are particularly susceptible to genetic drift and may suffer a loss of genetic diversity (Fischer and Matthies 1997; Reed and Frankham 2003). The maintenance of genetic diversity is a critical issue in the conservation and management for long-term survival of threatened species (Frankel 1983). Owing to human overexploitation, populations of *F. walujewii* have become small and isolated, so reservation regions covering major populations with high genetic variation should be established. Populations in Xinyuan County and Zhaosu County have a higher level of genetic diversity than do other populations, and include the main genetic diversity of *F. walujewii*; so, these two counties should be the areas of greatest focus when considering strategies of conservation management for the recovery of this important officinal species.

Acknowledgements

This research was supported by grants from the Western Doctor Project (XBBS201306) of Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences; National Natural Science Foundation of China (31400561); the Young Doctor Science and Technology Talents Training Project (2013731026) of Xinjiang Uygur Autonomous Region. We thank Huiliang Liu, Xiaobin Zhou, Xiaoshan Kang at the Xinjiang Institute of Ecology and Geography, CAS, for their kind help in sampling in the wild.

References

Aris-Brosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Molecular Biology and Evolution* **13**, 494–504. doi:10.1093/oxfordjournals.molbev.a025610

Avise JC (2000) 'Phylogeography: the history and formation of species.' (Harvard University Press: Cambridge, MA)

Avise JC, Walker D (1998) Pleistocene phylogeographic effects on avian populations and the speciation process. *Proceedings of the Royal Society of London* **265**, 457–463. doi:10.1098/rspb.1998.0317

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**, 489–522.

Bohonak AJ (2002) IBD (isolation by distance): a program for analyses of isolation by distance. *The Journal of Heredity* **93**, 153–154. doi:10.1093/jhered/93.2.153

Byrne M (2008) Evidence for multiple refugia at different time scales during Pleistocene climatic oscillations in southern Australia inferred from phylogeography. *Quaternary Science Reviews* **27**, 2576–2585. doi:10.1016/j.quascirev.2008.08.032

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657–1659. doi:10.1046/j.1365-294x.2000.01020.x

Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JJ, Henderson TB, Kiehl JT, Large WG, McKenna DS, Santer BD, Smith RD (2006) The Community Climate System Model version 3 (CCSM3). *Journal of Climate* **19**, 2122–2143. doi:10.1175/JCLI3761.1

Comes HP, Kadereit JW (1998) The effect of quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science* **3**, 432–438. doi:10.1016/S1360-1385(98)01327-2

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochemical Bulletin* **19**, 11–15.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**, 2571–2581. doi:10.1046/j.1365-294X.2002.01650.x

Excoffier L, Ray N (2008) Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution* **23**, 347–351. doi:10.1016/j.tree.2008.04.004

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: applications to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* **1**, 47–50.

Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annual Review of Ecology Evolution and Systematics* **40**, 481–501. doi:10.1146/annurev.ecolsys.39.110707.173414

Fehlberg SD, Ranker TA (2009) Evolutionary history and phylogeography of *Encelia farinosa* (Asteraceae) from the Sonoran, Mojave, and Peninsular Deserts. *Molecular Phylogenetics and Evolution* **50**, 326–335. doi:10.1016/j.ympev.2008.11.011

Fischer M, Matthies D (1997) Mating structure and inbreeding and outbreeding depression in the rare plant *Gentianella germanica* (Gentianaceae). *American Journal of Botany* **84**, 1685–1692. doi:10.2307/2446466

Forester BR, DeChaine EG, Bunn AG (2013) Integrating ensemble species distribution modelling and statistical phylogeography to inform projections of climate change impacts on species distributions. *Diversity & Distributions* **19**, 1480–1495. doi:10.1111/ddi.12098

Frankel OH (1983) The place of management in conservation. In 'Genetics and conservation'. (Eds CM Schonewald-Cox, SM Chambers, B MacBryde, L Thomas L) pp. 1–14. (Benjamin-Cummings: Menlo Park, CA)

Frankham R, Ballou JD, Briscoe DA (2002) 'Introduction to conservation genetics.' (Cambridge University Press: Cambridge, UK)

Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology* **10**, 2741–2752. doi:10.1046/j.1365-294X.2001.t01-1-01411.x

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking, and background selection. *Genetics* **147**, 915–925.

Ge XJ, Hwang CC, Liu ZH, Huang CC, Huang WH, Hung KH, Wang WK, Chiang TY (2011) Conservation genetics and phylogeography of endangered and endemic shrub *Tetraena mongolica* (Zygophyllaceae) in inner Mongolia, China. *BMC Genetics* **12**, 1. doi:10.1186/1471-2156-12-1

Hasumi H, Emori S (2004) 'K-1 coupled GCM (MIROC) description.' (Center for Climate System Research, University of Tokyo: Tokyo)

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of The Royal Society of London Series B, Biologic* **359**, 183–195. doi:10.1098/rstb.2003.1388

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**, 1965–1978. doi:10.1002/joc.1276

Hu RJ (1998) 'Physical geography of the Tianshan Mountains in China.' (China Environmental Science Press: Beijing)

Huelsenbeck JP, Ronquist F (2001) MrBayes: a program for the Bayesian inference of phylogeny. *Bioinformatics* **17**, 754–755. doi:10.1093/bioinformatics/17.8.754

Jaeger JR, Riddle BR, Bradford DF (2005) Cryptic neogene vicariance and Quaternary dispersal of the red-spotted toad (*Bufo punctatus*): insights on the evolution of North American warm desert biotas. *Molecular Ecology* **14**, 3033–3048. doi:10.1111/j.1365-294X.2005.02645.x

Jia DR, Abbott RJ, Liu TL, Mao KS, Bartish IV, Liu JQ (2012) Out of the Qinghai–Tibet Plateau: evidence for the origin and dispersal of Eurasian temperate plants from a phylogeographic study of *Hippophaë rhamnoides* (Elaeagnaceae). *New Phytologist* **194**, 1123–1133. doi:10.1111/j.1469-8137.2012.04115.x

Li N, Fu L (1997) Notes on gymnosperms I. Taxonomic treatment of some Chinese conifers. *Novon* **7**, 261–264. doi:10.2307/3391941

Mellick R, Lowe A, Allen C, Hill RS, Rossetto M (2012) Palaeodistribution modelling and genetic evidence highlight differential post-glacial range shifts of a rain forest conifer distributed across a latitudinal gradient. *Journal of Biogeography* **39**, 2292–2302. doi:10.1111/j.1365-2699.2012.02747.x

Moritz C (1994) Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology* **3**, 401–411. doi:10.1111/j.1365-294X.1994.tb00080.x

Nei M (1987) 'Molecular evolutionary genetics.' (Columbia University Press: NewYork)

Nichols RA, Hewitt GM (1994) The genetic consequences of long distance dispersal during colonization. *Heredity* **72**, 312–317. doi:10.1038/hdy.1994.41

Osborne MJ, Norman JA, Christidis L, Murray ND (2000) Genetic distinctness of isolated populations of an endangered marsupial, the mountain pygmy-possum, *Burramys parvus. Molecular Ecology* **9**, 609–613. doi:10.1046/j.1365-294x.2000.00908.x

Oxelman B, Lidén M, Berglund D (1997) Chloroplast *rps16* intron phylogeny of the tribe *Sileneae* (Caryophyllaceae). *Plant Systematics and Evolution* **206**, 393–410. doi:10.1007/BF00987959

Peterson AT, Nakazawa Y (2008) Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography* **17**, 135–144.

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* **31**, 161–175. doi:10.1111/j.0906-7590.2008.5203.x

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231–259. doi:10.1016/j.ecolmodel.2005.03.026

Pope LC, Sharp A, Moritz C (1998) The genetic diversity and distinctiveness of the yellow-footed rock-wallaby *Petrogale xanthopus* (Gray, 1854) in New South Wales Pacific. *Conservation Biology* **4**, 164–169.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818. doi:10.1093/bioinformatics/14.9.817

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. *Conservation Biology* **17**, 230–237. doi:10.1046/j.1523-1739.2003.01236.x

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant-tissues. *Plant Molecular Biology Reporter* **5**, 69–76. doi:10.1007/BF00020088

Rogers A, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**, 552–569.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574. doi:10.1093/bioinformatics/btg180

Rosenberg MS (2001) 'PASSAGE: pattern analysis, spatial statistics and geographic exegesis. Ver. 1.1.' (Department of Biology, Arizona State University: Tucson, AZ)

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology & Evolution* **1**, 9–10. doi:10.1016/0169-5347(86)90059-5

Sang T, Crawford DJ, Stuessy TF (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84**, 1120–1136. doi:10.2307/2446155

Shaw J, Lickey E, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* **92**, 142–166. doi:10.3732/ajb.92.1.142

Shi YF, Cui ZJ, Su Z (2005) 'The quaternary glaciations and environmental variations in China.' (Hebei Science and Technology Publishing House: Hebei, China)

Simmons MP, Ochoterena H (2000) Gaps as characters in sequencebased phylogenetic analyses. *Systematic Biology* **49**, 369–381. doi:10.1093/sysbio/49.2.369

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**, 555–562.

Smith CI, Farrell BD (2005) Range expansions in the flightless longhorn cactus beetles, *Moneilema gigas* and *Moneilema armatum*, in response to Pleistocene climate changes. *Molecular Ecology* **14**, 1025–1044. doi:10.1111/j.1365-294X.2005.02472.x

Su ZH, Zhang ML (2013) Evolutionary response to Quaternary climate aridification and oscillations in northwestern China revealed by chloroplast phylogeography of *Nitraria sphaerocarpa* (Nitrariaceae). *Biological Journal of the Linnean Society. Linnean Society of London* **109**, 757–770. doi:10.1111/bij.12088

Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* **240**, 1285–1293. doi:10.1126/science.3287615

Taberlet P, Cheddadi R (2002) Quaternary refugia and persistence of biodiversity. *Science* **297**, 2009–2010. doi:10.1126/science.297.5589.2009

Tajima F (1983) Evolutionary relationships of DNA sequences in finite populations. *Genetics* **105**, 437–460.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.

Tajima F (1996) The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. *Genetics* **143**, 1457–1465.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**, 619–633.

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal-W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680. doi:10.1093/nar/22.22.4673

Wang FZ, Tang J (1980) Liliaceae. In 'Flora of China, vol 14'. (Eds ZY Wu, PH Raven) p. 102. (Science Press: Beijing)

Wei WS, Hu RJ (1990) Precipitation and climate conditions of Tianshan Mountains. *Arid Land Geography* **13**, 29–36.

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAS. *Proceedings of The National Academy of Sciences, USA* **84**, 9054–9058. doi:10.1073/pnas.84.24.9054

Wu LL, Cui XK, Milne RI, Sun YS, Liu JQ (2010*a*) Multiple autopolyploidizations and range expansion of *Allium przewalskianum* Regel. (Alliaceae) in the Qinghai–Tibetan Plateau. *Molecular Ecology* **19**, 1691–1704. doi:10.1111/j.1365-294X.2010.04613.x

Yin LK, Tan LX, Wang B (2006) 'Rare endanged endemic higher plants in XinJiang of China.' (Science Press: Urumqi, China)

Zhang JM (2006) Studies on the geological structures and characteristic of terrain and landform in Yili river basin. *Journal of Shihezi University* **24**, 442–445.

Zhang HX, Zhang ML, Sanderson SC (2013) Retreating or standing: responses of forest species and steppe species to climate change in arid eastern central Asia. *PLoS One* **8**(4), e61954. doi:10.1371/journal.pone.0061954