Plant Syst Evol
DOI 10.1007/s00606-012-0677-6

ORIGINAL ARTICLE

Phylogeographic and demographic effects of Quaternary climate oscillations in *Hexinia polydichotoma* (Asteraceae) in Tarim Basin and adjacent areas

Zhihao Su · Mingli Zhang · James I. Cohen

Received: 2 February 2012 / Accepted: 12 June 2012

Abstract In order to investigate the genetic diversity and influence of climate oscillations on evolutionary processes of organisms in Northwest China, we selected *Hexinia polydichotoma*, a species endemic to China, and examined the phylogeographic structure and historical factors that influenced the evolutionary history of this species in its entire cover range, Tarim Basin and adjacent areas. In the study, 17 haplotypes were identified in *H. polydichotoma* on the basis of two chloroplast DNA sequences (*trn*H–*psb*A and *ycf*6–*psb*M). Shown in the network, the two common haplotypes, A and D, respectively, mainly distribute along the northern and southern rims of the basin. The analyses of molecular variance analysis suggest that genetic variation primarily occurs among populations, and all populations were subdivided into five groups by SAMOVA. Geographic range expansion along the southern and northern rims of the basin was supported by the significant value for Tajima's *D* and by the unimodal mismatch distribution. It is possible that during the interglacial period of the middle Pleistocene, a large amount of snow and glacial ice melted from the mountains surrounding Tarim Basin. This increased water, the expanding desert, and the dispersal ability of *H. polydichotoma* were important factors driving not only geographic range expansion, but also the current phylogeographic structure of this species. It is possible that during the middle Pleistocene, the climatic fluctuations resulted in expansion and contraction cycles of river systems and oases, and may consequently have caused population fragmentation.

Keywords *Hexinia polydichotoma* · Phylogeography · Long-distance colonization · Habitat fragmentation

Z. Su · M. Zhang (✉)
Key Laboratory of Biogeography and Bioresource in Arid Land, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China
e-mail: zhangml@ibcas.ac.cn

Z. Su
Graduate University, Chinese Academy of Sciences, Beijing 100049, China

M. Zhang
Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

J. I. Cohen
Department of Biology and Chemistry, Texas A&M International University, Laredo, TX 78041, USA

Introduction

Climate oscillations during the Quaternary had a profound impact on the genetic diversity and distribution of plants in the Northern Hemisphere (Hewitt 2000), and in order to understand the impact of these climate shifts, researchers have investigated the phylogeographic history of groups of plants (Hewitt 2004; Taberlet et al. 1998; Soltis et al. 2006). Much research on climatic changes during the Quaternary has focused on the effects of the movement of glaciers on the genetic constitution of groups of plants (Broyles 1998; Wu et al. 2010); however, even if glaciers did not cover an area, such as in northern China, climate shifts may also have had a large impact on the phylogeographic history of species. For example, climatic oscillation is one of the main factors that has molded the current distribution and biodiversity of *Lagochilus ilicifolius* (Meng and Zhang 2011). One region of particular interest for investigating the effect of climatic changes on organisms is Tarim Basin, the largest inland basin in Northwest China, and its adjacent areas, and

Published online: 15 July 2012

this area was not glaciated during the Quaternary (Li 1998). Currently, with the exception of some animal phylogeographical studies (Wu et al. 2011; Zhang et al. 2010), our understanding of the phylogeographic history of this region is poor. Did climatic changes during the Quaternary impact the genetic diversity and distribution of species in this region, and if so, are the patterns congruent with those identified in previous studies or are different patterns identified? In order to understand the pattern and extent of this impact, we studied the phylogeographic history of the plant species *Hexinia polydichotoma* (Ostenf.) H.L. Yang (Asteraceae) in Tarim Basin and adjacent areas.

Tarim Basin is bordered to the north by the Tien Shan Mountains, to the south by the Kunlun and Altun Mountains, to the west by the Pamir Plateau, and to the east, it is connected to the Hexi Corridor. Tarim Basin covers approximately 530,000 km^2, stretching 1,400 km from east to west, and 520 km from north to south (Liu and Qin 2005). The basin includes the world's second largest shifting-sand desert, Taklimakan Desert, which occupies nearly 60 % of the entire basin (Liu and Qin 2005). Glaciers have developed on many of the peaks of the mountains surrounding the basin, and these glaciers play an important role in the ecology of the basin. This is due to the fact that Tarim Basin is an endorheic basin (a closed drainage basin), so melting snow and glacial ice feed the rivers. Given the amount of runoff and groundwater, oases can persist in alluvial-diluvial plains, river deltas, and the edges of alluvial-diluvial fans. This results in a green ring around the rim of the basin (Liu and Qin 2005). Most water in the basin originated from the snow and ice melt in the spring but not via rainfall in the summer; thus, the water volume in the basin changes seasonally. In fact, at the edge of the desert and in riparian areas, oases seasonally fragment (Liu and Qin 2005).

The hot and dry climate of the basin developed in the late Cretaceous, prior to the basin becoming endorheic (Liu and Qin 2005; Mu 1994). The arid environment was exacerbated during the late Tertiary, when the uplift of the Tibetan Plateau altered the atmospheric circulation in the basin (Zheng et al. 2003; Sun et al. 2008). Evidence of the eolianite suggests that the Taklimakan Desert and the adjacent oases initially formed during the middle Pleistocene (Mu 1994), and the rudimentary hydrographic net dates back to the late Miocene (Liu and Qin 2005). During the glacial–interglacial cycles, a serial effect occurred: as the temperature of the basin rose or dropped, the water volume of the river increased or decreased, respectively. This resulted in shifts in the size of the desert and oases (Zhang et al. 2003).

For the present study, *H. polydichotoma* was investigated in a phylogeographic context. *Hexinia* is a monotypic genus endemic to China (Ling and Shi 1997; Zhao and Zhu 2003). It belongs to Lactucinae Less. in Cichorieae (Ling and Shi 1997), and preliminarily data suggest that it is derived from *Chondrilla*, a genus of Lactucinae Less. in Cichorieae of Asteraceae of the xerophilic Tethys flora (Ling and Shi 1997; Yang 1992). *H. polydichotoma* is a perennial and drought-tolerant herb, and grows in the gullies of the Gobi desert and areas surrounding the Taklimakan Desert (Fig. 1). Given its restricted distribution to Tarim Basin and adjacent areas, *H. polydichotoma* is a useful species to address phylogeographic patterns in the area.

As a consequence, on the basis of 249 individuals belonging to 16 populations of *H. polydichotoma*, which were collected across the entire range of the species, we chose two chloroplast DNA (cpDNA) spacers (*psb*A–*trn*H, *ycf*6–*psb*M) to sequence in order to conduct a phylogeographic study. In plants, cpDNA is thought to evolve slowly, with low recombination and mutation rates (Li and Fu 1997; Comes and Kadereit 1998). Without the confounding effect of biparental inheritance, the maternally inherited cpDNA lineages in natural populations can often better trace the evolutionary history of the species and display distinct geographic distributions (Avise 2000). A series of non-coding regions of cpDNA have been used successfully in phylogeographic studies (Wu et al. 2010; Sosa et al. 2009), and evidence also suggests that cpDNA sequence lineages are effective in revealing postglacial expansion patterns in plants (Guo et al. 2010; Vidal-Russell et al. 2011). The main objectives of the present study are to determine (1) the hierarchical structure of genetic variation, and (2) the manner in which Quaternary climate oscillations shaped the phylogeographic and genetic structure of the study species.

Materials and methods

Sampling

A total of 249 individuals from 16 populations of *H. polydichotoma* were included in the present study. This sampling covers the entire geographic range of the species, which includes Tarim Basin as well as its adjacent areas, Turpan Basin and Hexi corridor. Of the sampled populations, 13 are from Tarim Basin, one is from Turpan Basin, and 2 are from the Hexi Corridor. In each population, 10–20 individuals were collected, and from each, fresh stems and leaves were gathered and dried in silica gel. We selected two closely related species of *Chondrilla*, *Chondrilla brevirostris* and *Chondrilla ambigua*, as the outgroups of the phylogenic study.

DNA extraction, amplification, and sequencing

Total genomic DNA was extracted from silica-gel-dried leaf tissue using a modified 2× CTAB method (Rogers and

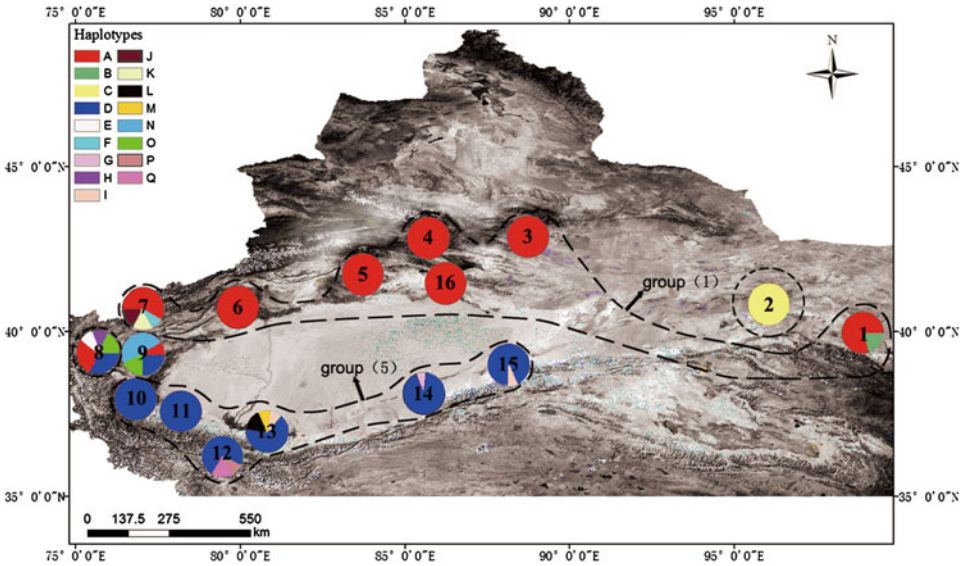

Fig. 1 Geographical distribution of *Hexinia polydichotoma* in China. Population numbers correspond to those in Table 2; haplotypes to those in Table 1. *Groups* (*1*) and (*5*) are also shown in the figure

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), and the *ycf*6–*psb*M spacer was amplified and sequenced using the primers and protocols of Demesure et al. (1995). Amplification products were purified using the PCR product purification kit (Shanghai SBS, Biotech Ltd., China), following protocols provided by the manufacturer. The forward and reverse primers of the amplification reactions were employed in the sequencing reactions, which were conducted with the DYEnamic ET Terminator Kit (Amersham Pharmacia Biotech). Sequencing occurred at the Shanghai Sangon Biological Engineering Technology & Services Co., Ltd. (Shanghai, China), with the use of an ABI-PRISM 3730 automatic DNA analyzer. Electropherograms were edited and assembled using SEQUENCHER 4.1 (Gene Codes, Ann Arbor, MI, USA). Sequences were aligned with CLUSTAL W (Thompson et al. 1994) and refined by visual inspection.

Molecular variability and demographic analysis

To test the spatial genetic structure of cpDNA haplotypes, spatial analysis of molecular variance was performed using program SAMOVA v.1.0 (Dupanloup et al. 2002) in order to define groups of populations (K) that are geographically homogeneous and genetically differentiated from each other. The analysis was run for $K = 2$–15, starting from 100 random initial conditions for each run. Finally, the number of groups that maximizes the proportion of total genetic variance because of differences among groups of populations (F_{CT}) was retained as the best grouping of populations.

HAPLONST (http://www.pierroton.inra.fr/genetics/labo/Software/index.html) was used to estimate within-population diversity (h_S) and total gene diversity (h_T). Based on pairwise differences of the sequences, analyses of molecular variance (AMOVA) were employed to study the genetic structure of the species (Excoffier et al. 1992). 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 (Fu 1997; Aris-Brosou and Excoffier 1996; Tajima 1996). In order to investigate, in another manner, hypotheses of demographic history, the mismatch distribution (MDA) also was 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 multimodal distribution usually suggests that populations are at demographic equilibrium, while a unimodal distribution indicates that populations have experienced a recent expansion. All expansion tests were implemented in ARLEQUIN v.3.01 (Excoffier et al. 2005), and in order to test for significance, 10,000 permutations were performed. If the sudden expansion model was not rejected, we used the relationship $\tau = 2ut$ to estimate the expansion time (t) (Rogers and Harpending 1992), where τ is the total number of mutations, and u is the mutation rate per generation for the whole analyzed 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/y), k is the average sequence length of the analyzed DNA region, and g is the generation time in years. The nucleotide substitution rate of *H. polydichotoma* had not been estimated; therefore, we used a range for the estimated rate of nucleotide substitution (Wolfe et al. 1987), with 1.0×10^{-9} s/s/y as the lower limit and 3.0×10^{-9} s/s/y as the upper limit. Because

Table 1 Seventeen haplotypes of *Hexinia polydichotoma* recognized on basis of two chloroplast DNA sequences, *trn*H–*psb*A and *ycf*6–*psb*M

Haplotype	Sequence position																		
	33	56	65	81	110	172	222	336	363	370	382	385	386	443	513	799	876	966	984
A	T	A	A	T	G	★	–	A	C	A	–	–	–	A	C	–	▼	A	T
B	T	A	–	T	G	★	–	A	C	A	–	–	–	A	C	–	▼	A	T
C	T	A	A	T	C	★	–	A	C	A	–	–	–	A	C	–	▼	A	T
D	T	A	A	T	G	★	–	A	C	A	–	–	–	A	C	–	▽	A	T
E	T	A	A	T	G	★	–	A	C	A	–	–	–	A	T	–	▽	A	T
F	G	A	A	T	G	★	–	C	C	A	–	–	–	A	C	–	▼	A	T
G	T	A	A	T	G	–	–	A	C	A	–	–	–	A	C	–	▽	A	T
H	T	A	A	T	G	★	–	A	C	A	–	–	–	A	C	–	▽	A	G
I	T	A	A	T	G	★	–	A	C	A	–	C	△	A	C	–	▽	A	T
J	T	G	A	T	G	★	–	A	C	C	–	–	–	A	C	–	▽	G	T
K	T	G	A	T	G	★	–	A	C	C	–	–	–	A	C	–	▼	G	T
L	T	A	A	G	G	★	◇	A	C	A	–	–	–	A	C	–	▽	A	T
M	T	A	A	T	G	★	◇	A	C	A	–	–	–	A	C	–	▽	A	T
N	T	A	A	T	G	★	–	A	C	A	–	–	–	C	C	–	▼	A	T
O	T	A	A	T	G	★	–	A	C	A	◆	T	△	A	C	□	▽	A	T
P	T	A	A	T	G	★	–	A	C	A	–	–	–	A	C	–	▽	A	T
Q	T	A	A	T	G	★	–	A	T	A	–	–	–	A	C	–	▽	A	T

★, AAATAACAA; ◇, TATAAATTTG; ◆, TTA; △, TTATTACTTT; □, GCTAAT; ▼, TTTC; ▽, GAAA

plants of *H. polydichotoma* are annuals, the generation time for this desert species is one year.

Phylogenetic and phylogeographic analysis

Phylogenetic relationships of the cpDNA haplotypes were investigated with the use of maximum parsimony (MP), as implemented in PAUP* version 4.0b10 (Swofford 2002), and Bayesian inference (BI), as implemented in MrBayes v. 3.0 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). For MP analyses, the following search strategy was conducted: 100 random additions of sequences, with the tree-bisection–reconnection (TBR) branch swapping, and MULTREES, COLLAPSE, and STEEPEST DESCENT options switched on. Characters were weighted equally, and states were treated as unordered. In the MP and BI analysis, each indel was treated as a single mutation event and coded as a substitution (Simmons and Ochoterena 2000). Support values were calculated using 1,000 bootstrap replicates.

We used Modeltest 3.7 to determine the appropriate nucleotide substitution model (Posada and Crandall 1998) to use in BI analyses. For BI analyses, 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. This percentage was determined via observation of the log-likelihood plots. A 50 % majority rule consensus tree was constructed, and posterior probabilities of nodes were recorded. Using the median-joining method, implemented in the Network 4600 program (Bandelt et al. 1999), genealogical relationships among all haplotypes were estimated.

Results

Sequence analysis

The aligned sequence length for the *trn*H–*psb*A spacer is 530 base pairs (bp) and for the *ycf*6–*psb*M spacer is 600 bp. A total of 19 informative characters were found in the aligned sequence data: 12 nucleotide substitutions (positions 33, 56, 81, 110, 336, 363, 370, 443, 513, 876–879, 966, and 984) and seven indels (positions 65, 172–180, 222–231, 382–384, 385, 386–395, and 799–804). Of the 249 sampled individuals from 16 populations, a total of 17 haplotypes (A–Q) were identified (Table 1). GenBank accession numbers of the cpDNA sequences are JX183915-JX183936.

Phylogenetic analysis

The best nucleotide substitution model selected by AIC was HKY + G. The MP and BI phylogenetic analyses

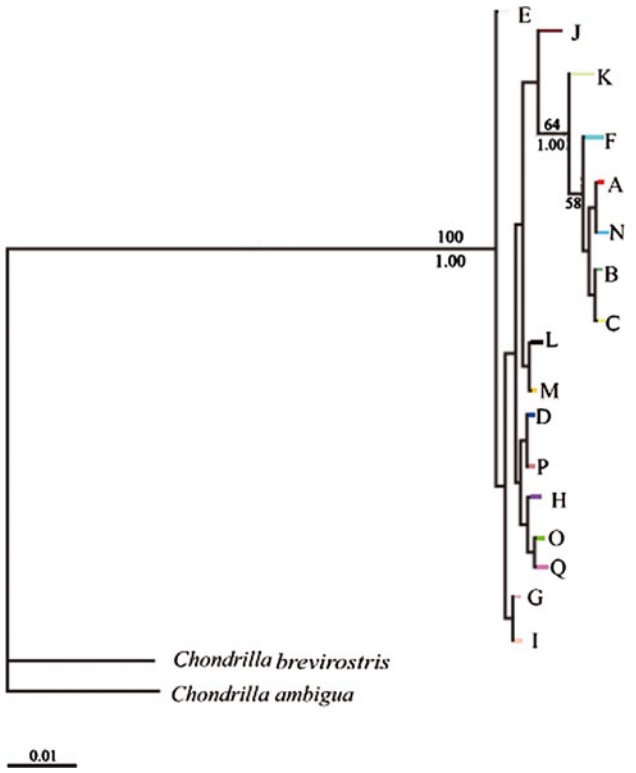

Fig. 2 Phylogenetic relationships of 17 haplotypes *Hexinia polydichotoma* and related species. *Numbers above/below branches* are support values (maximum parsimony bootstrap values >50 %/ posterior probabilities >0.90)

produced the same consensus tree (Fig. 2). In this tree, *Hexinia* is resolved as monophyletic, and this relationship is well supported (100 % bootstrap support and 1.00 posterior probabilities). Haplotype E is sister to the rest of the haplotypes of *Hexinia*. Most of the phylogeny is not well supported, but one clade received bootstrap values >50 %. This is the clade that include A, B, C, F, K, and N. Haplotypes A, B, C, and N have a greater number of mutations than the other haplotypes (Fig. 2).

Haplotype geographical distribution and relationships

The geographic distribution of haplotypes, along with the frequency of haplotypes in each population, is presented in both Fig. 1 and Table 2. In the network, haplotypes A and D are more widespread than the other haplotypes. Haplotype A is widespread along the western rim of the basin, northern rim of the basin, and Hexi Corridor. In contrast, haplotype D is widespread along the western and southern rims of the basin. Most of the individuals of haplotype A, together with that of haplotypes F, K, and J, are distributed along the northern edge of the basin, and most individuals of haplotype D, along with that of haplotypes I, G, Q, M, and L, are distributed across the southern edge of the basin. The western rim of the basin houses a small number of individuals of haplotypes A and D together with haplotypes N, H, E, and O.

The structure of the network demonstrates that some haplotypes may have direct or more distant relationships with others. For example, haplotypes B, C, and N are connected to haplotype A by one substitution, while others are connected to it by more steps. Indeed, haplotype K is connected to A by three mutations, haplotype J is connected to K via four mutations, and haplotype F is connected to A by two mutations. Haplotypes E, H, and Q all connect to haplotype D by one step, while others, such as G, L, M, and P, are connected to D by a greater number of steps. Haplotypes J and A are both connected to D via three and four mutations, respectively, resulting in a ring structure among haplotypes A, K, J, and D.

Genetic diversity and genetic structure

Spatial genetic analysis of cpDNA haplotypes using SAMOVA indicated that F_{CT} increased to a maximal value of 0.6006 when K (the number of groups) was raised from $K = 2$ to $K = 5$. The grouping pattern of populations corresponding to $K = 5$ is: (1) populations 1, 3–7, and 16, belonging to the northern rim of Tarim Basin and the Hexi Corridor; (2) population 2, belonging to the Hexi Corridor; (3) population 8, belonging to the western rim of the basin; (4) population 9, belonging to the western rim of the basin; and (5) populations 10–15, belonging to the southern rim of the basin. Within-population gene diversity (h_S) was 0.266 (SE 0.0817), and total gene diversity (h_T) was 0.739 (SE 0.0544). The AMOVA results provide evidence that 51.72 % ($P < 0.001$) of the total variation can be explained by differences among populations, and 48.28 % ($P < 0.001$) of the total variation occurs within the populations. When populations were grouped according to geographical region, AMOVA results demonstrated that 60.06 % ($P < 0.001$) of the total variation occurred among the groups and only 0.03 % occurred among populations within groups (Table 3).

Demographic analyses

Significant results of Tajima's D, along with unimodal distributions for the shapes of the mismatch distribution, suggest range expansions along the northern and southern rims of the basin (Fig. 4; Table 4). Based on the range of the cpDNA substitution rate, a haplotype sequence length of 1,130 bp, and 1-year generation time, the time of the geographic range expansion of *H. polydichotoma* is estimated to have occurred between 0.22 and 0.66 million years ago (Mya).

Table 2 Details of sample locations, sample size, and haplotype frequencies for 16 populations of *Hexinia polydichotoma*

Number	Location	Latitude (N)	Longitude (E)	Altitude (m)	Haplotype
1	Jinta	39°58′	98°55′	1,242	A(13), B(2)
2	Guazhou	40°50′	96°03′	1,219	C(10)
3	Tuokexun	42°52′	88°44′	0	A(15)
4	Kuerle	41°52′	85°42′	918	A(15)
5	Kuche	41°44′	83°43′	1,060	A(16)
6	Akesu	40°44′	79°55′	1,063	A(15)
7	Atushi	39°44′	76°12′	1,295	A(10), F(1), J(2), K(2)
8	Shule	39°23′	75°58′	1,290	A(6), D(6), E(2), H(2) O(4)
9	Yuepuhu	39°06′	77°02′	1,183	A(3), D(5), N(8) O(4)
10	Yecheng	37°57′	77°13′	1,320	D(15)
11	Pishan	37°35′	78°11′	1,424	D(15)
12	Hetian	37°01′	80°18′	1,378	D(12), P(1), Q(2)
13	Cele	36°58′	80°49′	1,407	D(13), I(1), L(2), M(1)
14	Qiemo	38°08′	85°34′	1,333	D(14), G(1)
15	Ruoqiang	39°00′	88°09′	889	D(18), I(1)
16	Yuli	41°28′	86°14′	888	A(12)

Table 3 Results of analysis of molecular variance for 16 populations of *Hexinia polydichotoma* based on chloroplast DNA sequence data

Source of variation	*df*	Sum of squares	Variance components	Percentage of variation (%)
Among populations	15	246.406	0.9975	51.72*
Within populations	233	216.911	0.9310	48.28*
(1, 3–7, 16) vs. (2) vs. (8, 9) vs. (10–15)				
Among groups	3	231.732	1.4032	59.64*
Among populations within groups	12	14.674	0.0186	0.79
Within populations	233	216.911	0.9310	39.57*

* $P < 0.001$

Table 4 Results of neutrality tests and mismatch distribution analysis for two groups

Group	τ	SSD (*P* value)	Hrag (*P* value)	Tajiam's *D* (*P* value)	Fu's F_S (*P* value)
Populations 1, 3–7, 16	3.00	0.0065 (0.11)	0.6969 (0.78)	−1.8827 (0.003)	−1.6892 (0.17)
Populations 10–15	3.00	0.0151 (0.17)	0.6237 (0.66)	−1.2213 (0.045)	0.045 (0.546)

One group includes populations 1, 3–7, and 16, and the other comprises populations 10–15

τ time in number of generations elapsed since the sudden expansion episode, *Hrag* the Harpending's raggedness index, *SSD* sum of squared deviations

Discussion

Genetic diversity and genetic structure of *H. polydichotoma*

In *Hexinia*, there is low genetic differentiation among within-group populations, which suggests an absence of reproductive isolation among populations within groups of the study species. Two factors help explain the lack of clear genetic structure among populations within groups. First, the fruits of *H. polydichotoma* are abundant, small, and light; consequently, they can be dispersed far from the mother plant. This long dispersal distance will lead to limited genetic structure among populations (Palmé et al. 2003; Jones et al. 2006). Second, the terrain along the rim of the basin tends to be flat; therefore, unobstructed gene exchange may occur among populations. Despite these reasons for low genetic differentiation among particular groups of population, there is moderately high genetic differentiation among the five identified groups of *Hexinia*

populations. Given that one of the largest groups is on the northern edge of the basin, while the other is one the southern edge (Fig. 1), this suggests that the vast desert area did obstruct gene flow and affect the phylogeographic history of the species.

Allopatric divergence

The results of the network analyses are, in general, congruent with those from the phylogenetic analyses. For example, the clade that includes haplotype A, B, C, F, J, K, and N is present in both the network and the phylogenetic tree. Haplotypes of this clade are mainly distributed along the northern rim of the basin and are apparently differentiated from those along the southern rim (Fig. 3). A high portion of total genetic variation (60.06 %) exists among the five identified groups (Fig. 1). The genetic divergence between the two large groups on the northern and southern rims of the basin could have resulted from the vast area of the Taklimakan Desert separating them, resulting in a lack of gene flow between the populations on either side of the basin.

Along with the geographic separation of the populations, the dry climate of the region also could have played a role in restricted gene flow. The annual average rainfall in Tarim Basin is 50–70 mm, and the annual average evaporation is ca. 1,500 mm. Due to the small amount of rainfall and the high amount of evaporation, the basin is the driest region in China (Liu and Qin 2005). Mu (1994) has hypothesized that the dry climate in the basin could date back to the late Cretaceous. Since the Pliocene, the uplift of the Tibetan Plateau intensely changed the atmospheric circulation in the area, causing widespread aridification across the entire basin (Zheng et al. 2003; Sun et al. 2008; Wang et al. 1996). Aridification and the formation of the desert likely obstructed gene flow between various groups of populations of the southern and northern areas of Tarim Basin. The processes involved in restricting gene flow among populations of *H. polydichotoma* also have been reported in a number of other plant phylogeographic studies (Bittkau and Comes 2005; Cornman and Arnold 2007; Ge et al. 2011).

Range expansion in Tarim Basin and adjacent areas

If large-scale geographic range expansion occurs, it usually leaves at least one distinct genetic signature: a single, widespread genotype (Hewitt 2000; Avise 1987; Comes and Kadereit 1998). The distribution of genetic variation in Tarim Basin and adjacent areas is consistent with this type of genetic signature. It is apparent that haplotype A is widespread along the northern rim of the basin, while D is

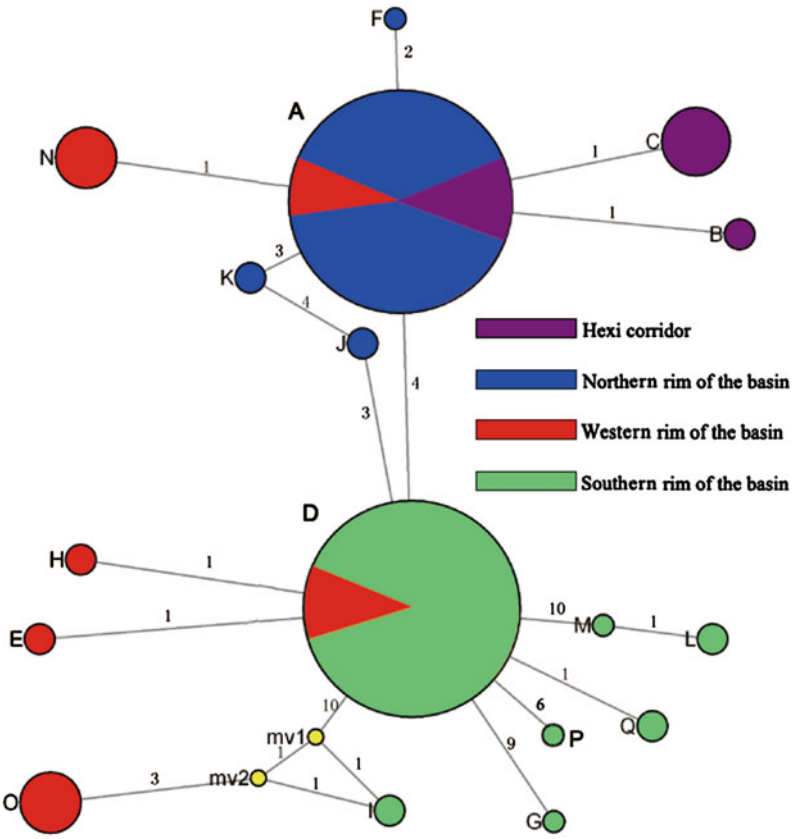

Fig. 3 Haplotype median-joining network of *Hexinia polydichotoma*. The *circle* size is proportional to haplotype frequencies. The number of inferred steps between haplotypes is shown near the corresponding branch section. Haplotypes in the network shown in *different colors* represent the following four geographical areas. The *yellow circles* (*mv1* and *mv2*) represent the missing or inferred haplotypes

common along the southern rim of the basin (Fig. 1). Range expansion was detected in the two large groups [groups (1) and (5)], which are distributed along the northern and southern edges of the basin, respectively, and the range expansion of the two groups was strongly supported by the significant value of Tajima's *D* (Table 4) as well as by the unimodal mismatch distribution (Fig. 4). In group (1), haplotype A is widespread, and haplotypes B and C are both present in the two populations from the Hexi Corridor. According to the phylogenetic analyses (Fig. 2), these latter two haplotypes are derived and sister to the clade the includes haplotypes A and N. Haplotype F, which is found in Atushi, is the sister to the clade that includes haplotypes A, B, C, and N. Haplotypes A and F are only found together in one population, Atushi (7). Given the phylogeny and geographic arrangement of these haplotypes, it is quite possible that haplotype F gave rise to haplotype A, and this latter haplotype migrated east, moving through the Shule Valley and reaching to the Hexi Corridor, to become widespread along the northern edge of the basin. Group (5) along the southern rim of the basin, contains the widespread haplotype D as well as haplotypes I, G, Q, M, and L. The individuals of haplotype D appear to have migrated along the southern rim of the basin. In general, the migration patterns of individuals of *H. polydichotoma* have occurred along the rims of the basin, where the hydrographic net is well-developed and water is most abundant.

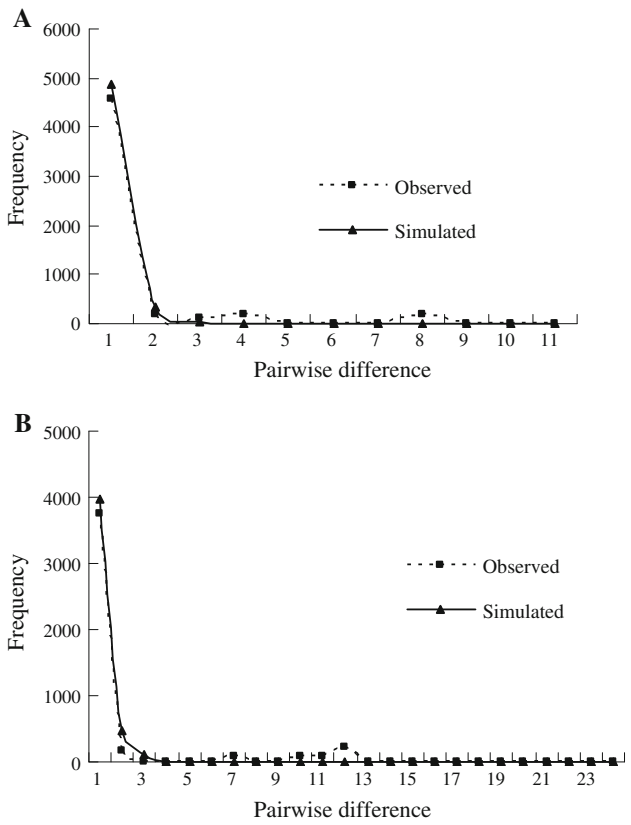

Fig. 4 Mismatch distribution analysis for chloroplast DNA data for **a** group (1) that includes populations 1, 3–7, and 16 (SSD = 0.0065, $P = 0.11$), and **b** group (5) that includes populations 10–15 (SSD = 0.0151, $P = 0.17$)

Our results suggest that *H. polydichotoma* expanded its geographic range between 0.22 and 0.66 Mya, which is during the middle Pleistocene (0.125–0.75 Mya) (Shi et al. 2005). Early in the Pleistocene, deserts were scattered throughout the basin (Shi et al. 2005), but during the late-early or early-middle Pleistocene, the region became drier, resulting in the formation of the Taklimakan Desert and its migratory dunes (Shi et al. 2005). The drifting sand and wind may have provided appropriate conditions for the dispersal of the light seeds of *H. polydichotoma*. Additionally, during the interglacial period in the middle Pleistocene (0.48–0.60 Mya), the climate in the basin was warmer and wetter than during the earlier glacial period (0.60–0.80 Mya), and with the increasing temperatures, a greater amount of runoff from melting snow and glacial ice infused the river systems along the edge of Tarim Basin (Shi et al. 2005). This greater amount of runoff provided favorable conditions for the survival and growth of the migrating seeds of *H. polydichotoma*. The combined actions of these factors allowed for the range expansion of this species in Tarim Basin. Geographic range expansions of this nature also have been reported in species in Europe, North America, and northeast Asia (Hewitt 2000; Broyles 1998; Aizawa et al. 2009).

Habitat fragmentation in Tarim Basin and adjacent areas

Habitat fragmentation was also detected in Tarim Basin. The group along the southern rim of the basin includes individuals of haplotypes I. Haplotype I is present in two populations, Cele (13) and Ruoqiang (15), and the distance between them is 721 km. Consequently, it appears likely that haplotype I might have experienced long-distance dispersal, then became fragmented. Haplotype A spread from the basin to the Hexi Corridor; however, its distribution is not continuous. Population Guazhou, located between the northern rim and the Hexi Corridor, fixed the highly derived haplotype C, so we can infer that after the colonization of the Hexi Corridor, populations experienced fragmentation, and this fragmentation could have resulted in the development of haplotype C. We hypothesize that along the rims of Tarim Basin, habitat fragmentation is a consequence of climate fluctuation throughout the basin. Though the arid climate has dominated the basin since the middle Pleistocene, climatic fluctuations have occurred frequently (Jin et al. 1994; Dong et al. 1997; Zhang and Men 2002). From the middle Pleistocene to the Holocene,

the climate of Tarim Basin has cycled multiple times between dry and humid conditions, resulting in expansion and contraction cycles of river systems and oases (Zhang et al. 2003; Feng et al. 1999; Yang et al. 2002; Luo et al. 2009). These climatic shifts have resulted in multiple rounds of habitat fragmentation, which has provided new opportunities for the colonization of different geographic areas.

Acknowledgments We thank Honghu Meng and Jianfeng Huang at the Xinjiang Institute of Ecology and Geography, CAS, for their help with the experiment. Funding was provided by the CAS Important Direction for Knowledge Innovation Project (no. KZCX2-EW-305) and Xinjiang Institute of Ecology and Geography, CAS.

References

Aizawa M, Yoshimaru H, Saito H, Katsuki T, Kawahara T, Kitamura K, Shi F, Sabirov R, Kaji M (2009) Range-wide genetic structure in a north-east Asian spruce (*Picea jezoensis*) determined using nuclear microsatellite markers. J Biogeogr 36:996–1007

Aris-Brosou S, Excoffier L (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. Mol Biol Evol 13:494–504

Avise JC (1987) Identification and interpretation of mitochondrial DNA stocks in marine species. In: Kumpf H, Nakamura EL (eds) Proceedings of the stock identification workshop. National Oceanographic and Atmospheric Administration, Panama City, FL, pp 105–136

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

Bandelt HJ, Forster P, Röhl A (1999) Median joining networks for inferring intraspecific phylogenies. Mol Biol Evol 14:37–48

Bittkau C, Comes HP (2005) Evolutionary processes in a continental island system: molecular phylogeography of the Aegean *Nigella arvensis* alliance (Ranunculaceae) inferred from chloroplast DNA. Mol Ecol 14:4065–4083

Broyles SB (1998) Postglacial migration and the loss of allozyme variation in Northern population of *Asclepias exaltata* (Asclepiadaceae). Am J Bot 85:1091–1097

Comes HP, Kadereit JW (1998) The effect of quaternary climatic changes on plant distribution and evolution. Trends Plant Sci 3:432–438

Cornman RS, Arnold ML (2007) Phylogeography of *Iris missouriensis* (Iridaceae) based on nuclear and chloroplast markers. Mol Ecol 16:4585–4598

Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 4:129–131

Dong GR, Chen HZ, Wang GY, Li XZ, Shao YJ, Jin J (1997) The evolution of deserts with climatic changes in China since 150 ka BP. Sci China Ser D Earth Sci 40:370–382

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

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

Excoffier L, 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. Evol Bioinforma 1:47–50

Feng Q, Su ZZ, Jin HJ (1999) Desert evolution and climatic changes in the Tarim River basin since 12 ka BP. Sci China Ser D Earth Sci 42:101–112

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 Genet 12:1

Guo YP, Zhang R, Chen CY, Zhou DW, Liu JQ (2010) Allopatric divergence and regional range expansion of *Juniperus sabina* in China. J Syst Evol 48:153–160

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. Nature 405:907–913

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Phil Trans R Soc Lond B Biol Sci 359:183–195

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

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. Mol Ecol 14:3033–3048

Jin HL, Dong GR, Jin J, Li BS, Shao YJ (1994) Environmental and climatic changes in the interior of Taklimakan desert since Late Glacial Age. J Desert Res 14:31–37 (in Chinese with English abstract)

Jones ME, Shepherd M, Henry RJ, Delves A (2006) Chloroplast DNA variation and population structure in the widespread forest tree, *Eucalyptus grandis*. Conserv Genet 7:691–703

Li WC (1998) The Chinese Quaternary vegetation and environment. Science Press, Beijing

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

Ling R, Shi Z (1997) Asteraceae. In: Wu ZY, Raven PH (eds) Flora of China, vol 80. Science Press, Beijing

Liu JQ, Qin XG (2005) Evolution of the environmental framework and oasis in the Tarim Basin. Quat Sci 25:533–539 (in Chinese with English abstract)

Luo C, Peng ZC, Yang D, Liu WG, Zhang ZF, He JF, Chou CL (2009) A lacustrine record from Lop Nur, Xinjiang, China: implications for paleoclimate change during Late Pleistocene. J Asian Earth Sci 34:38–45

Meng HH, Zhang ML (2011) Phylogeography of *Lagochilus ilicifolius* (Lamiaceae) in relation to Quaternary climatic oscillation and aridification in northern China. Biochem Syst Ecol 39:787–796

Mu GJ (1994) On the age and evolution of the Taklimakan desert. Arid Land Geogr 17:1–9 (in Chinese with English abstract)

Palmé AE, Semerikov V, Lascoux M (2003) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. Heredity 91:465–474

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant-tissues. Plant Mol Biol 5:69–76

Rogers A, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552–569

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574

Sang T, Crawford DJ, Stuessy TF (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am J Bot 84:1120–1136

Shi YF, Cui ZJ, Su Z (2005) The Quaternary glaciations and environmental variations in China. Hebei Science and Technology Publishing House, Hebei, pp 71–78

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. Syst Biol 49:369–381

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. Mol Ecol 14:1025–1044

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. Mol Ecol 15:4261–4293

Sosa V, Ruiz-Sanchez E, Rodriguez-Gomez FC (2009) Hidden phylogeographic complexity in the Sierra Madre Oriental: the case of the Mexican tulip poppy *Hunnemannia fumariifolia* (Papaveraceae). J Biogeogr 36:18–27

Sun JM, Zhang LY, Deng CL, Zhu RX (2008) Evidence for enhanced aridity in the Tarim Basin of China since 5.3 Ma. Quat Sci Rev 27:1012–1023

Swofford DL (2002) PAUP*: phylogenetic analysis using parsimony (and other methods), version 4.0b10. Sinauer Associates, Sunderland

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. Mol Ecol 7:453–464

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

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 Res 22:4673–4680

Vidal-Russell R, Souto CP, Premoli AC (2011) Multiple Pleistocene refugia in the widespread Patagonian tree *Embothrium coccineum* (Proteaceae). Aust J Bot 59:299–314

Wang Y, Li S, Wang JH, Yan MC (1996) The uplift of the Qinghai-Xizang (Tibetan) Plateau and its effect on the formation and evolution of Chinese deserts. Arid Zone Res 13:20–24 (in Chinese with English abstract)

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAS. P Natl Acad Sci USA 84:9054–9058

Wu LL, Cui XK, Milne RI, Sun YS, Liu JQ (2010) Multiple autopolyploidizations and range expansion of *Allium przewalskianum* Regel. (Alliaceae) in the Qinghai-Tibetan Plateau. Mol Ecol 19:1691–1704

Wu YH, Xia L, Zhang Q, Yang QS, Meng XX (2011) Bidirectional introgressive hybridization between *Lepus capensis* and *Lepus yarkandensis*. Mol Phylogenet Evol 59:545–555

Yang XL (1992) Flora of desert from China. Science Press, Beijing

Yang XP, Zhu ZD, Jaekel D, Owen LA, Han JM (2002) Late Quaternary palaeoenvironment change and landscape evolution along the Keriya River, Xinjiang, China: the relationship between high mountain glaciation and landscape evolution in foreland desert regions. Quat Int 97–98:155–166

Zhang HY, Men GF (2002) Stratigraphic subdivision and climatic change of the Quaternary of the center Taklimakan Desert. Xinjiang Geol 20:256–261 (in Chinese with English abstract)

Zhang H, Wu JW, Zheng QH, Yu YJ (2003) A preliminary study of oasis evolution in the Tarim Basin, Xinjiang. China J Arid Environ 55:545–553

Zhang Q, Xia L, He JB, Fu JZ, Yang QS (2010) Comparison of phylogeographic structure and population history of two Phrynocephalus species in the Tarim Basin and adjacent areas. Mol Phylogenet Evol 57:1091–1104

Zhao YZ, Zhu ZY (2003) The endemic genera of desert region in the centre of Asia. Acta Botanica Yunnanica 25:113–121

Zheng HB, Powell CM, Butcher K, Cao JJ (2003) Late Neogene loess deposition in southern Tarim Basin: tectonic and palaeoenvironmental implications. Tectonophysics 375:49–59