African Journal of Biotechnology Vol. 11(49), pp. 11038-11044, 19 June, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB12.675
ISSN 1684–5315 ©2012 Academic Journals

Full Length Research Paper

Genetic diversity and population structure of *Caragana microphylla* Lam. based on analysis of inter-simple sequence repeat markers

B. Xu, H. W. Gao*, Z. Wang and J. W. Lu

Institute of Animal Science, Chinese Academy of Agricultural Science, 2-Yuan-Ming-Yuan West Rd., Haidian District, Beijing 100094, China.

Accepted 14 February, 2012

Caragana microphylla **Lam. is a long-lived shrub species in the semi-arid, arid and desert regions. To determine the genetic diversity and population structure of *C. microphylla* Lam., 17 wild populations from the central and eastern part of Inner Mongolia were analyzed by inter-simple sequence repeat. 18 primers produced 296 bands across a total of 510 individuals. A high percentage of polymorphic loci was observed at species level (PPB = 81.4%). Based on analysis of molecular variance, 74.99% of the genetic variation of *C. microphylla* Lam. was found within population, 7% difference between regions and 15.2% among collection sites within regions. Cluster analyses showed that 17 populations are most arranged in the same cluster by geographic location. An indirect estimate of the G_{ST}-derived N_m value (N_m = 1.8921) indicate that gene flow is high among 17 populations. No significant correlation (r^2 = 0.13) between genetic and geographic distance was detected. Results of this study suggest that *C. microphylla* Lam. has a high genetic variability and potential as a source of variation for breeding programs.**

Key words: Inter-simple sequence repeat (ISSR), *Caragana microphylla* Lam., genetic structure.

INTRODUCTION

Caragana microphylla Lam. is an important shrub species in the semi-arid, arid and desert regions. It has a potential utilization as a pioneer leguminous shrub species for vegetation re-establishment in North China which is rich in wildlife resources of this plant. *Caragana* species is one of the Leguminosae Papilionoideae Galega family. It is a diploid with 2n=2x=16 chromosomes in somatic cells, and cross-pollinated by insects. Morphology and anatomy of *C. microphylla* Lam. have been well-documented (Xu, 2009). Around the world, there are more than 100 species distributed in Asia and Europe in arid and semi-arid regions (Li and Qu, 2001). *C. microphylla* Lam. has strong profusion base with more number of branches and constitutes a horizontal and vertical fence, and are relatively dense shrub (Zhang and He, 2005). It plays an important role in organic C sequestration, N accumulation and the hydrologic cycle, and has more ecological importance for vegetative restoration and reversal of desertification (Zhang et al., 2006). *C. microphylla* Lam. can improve the micro-environment of soil, and create condi-tions for other species to be immersed into the enriched biological diversity (Liu and Lin, 2004). As a deep-rooted plant species, *C. microphylla* Lam is recommended as an ideal species for ecological restora-tion in degraded sand-land ecosystems (Li et al., 2004). By using amplified fragment length polymorphism (AFLP) markers, a statistically significant correlation was found between pairwise genetic distance and corresponding geographic distance by the 10 *Caragana korshinskii* populations (Wang et al., 2007). The analyses of wild soybeans are found to have three genetic diversity centers proposed based on the geographical distribution of the number of accessions (Dong et al., 2000). The spatial patterns and environmental correlates and predictors of genetic variation of *Hordeum spontaneum* in

*Corresponding author. E-mail: gaohongwen@263.net. Tel/Fax: +86-10-62894560.

Table 1. Locations and genetic diversity of the 17 *C. microphylla* Lam.

Population	Origin	Longitude	Latitude	Altitude	AMT	AMP	H	I	PPB (%)
P1	Eqi, Innermogolia	49°07′	120°07′	660	-0.2	316	0.316	0.462	89.55
P2	Cogang, Innermogolia	49°11′	118°14′	610	-0.5	270	0.276	0.406	82.12
P3	Galabuer, Innermogolia	49°25′	118°25′	598	-0.5	270	0.295	0.435	88.44
P4	Wangongzhen, Innermogolia	49°08′	118°51′	618	-1.75	345	0.288	0.422	83.61
P5	Wugunoer, Innermogolia	49°06′	119°15′	678	-1.75	345	0.323	0.475	94.01
P6	Baogedewula, Innermogolia	48°35′	117°17′	572	-0.5	250	0.296	0.436	87.70
P7	Xinyouqi, Innermogolia	48°48′	116°59′	549	-0.5	250	0.258	0.384	82.49
P8	Nanshatozi, Innermogolia	43°35′	122°17′	182	5.9	375	0.288	0.426	87.70
P9	Gahaimiao, Innermogolia	44°47′	121°15′	239	5.9	341.8	0.262	0.388	82.12
P10	Dadianzi, Innermogolia	43°55′	121°12′	242	5.9	341.8	0.205	0.299	64.04
P11	Siziwang, Innermogolia	42°05′	111°54′	1464	3	310	0.192	0.289	73.74
P12	Zhengxiangbai, Innermogolia	42°36′	115°12′	1226	6.8	363	0.235	0.354	78.77
P13	Duolun, Innermogolia	42°15′	116°13′	1387	1.6	385	0.252	0.368	73.94
P14	Huade, Innermogolia	42°11′	114°41′	1278	1.6	330	0.240	0.361	82.49
P15	Xianghuang, Innermogolia	42°12′	113°50′	1328	2.5	260	0.234	0.352	79.14
P16	Chayouhou, Innermogolia	41°27′	113°03′	1420	3.4	292	0.246	0.370	82.12
P17	Xiwu, Innermogolia	44°30′	117°03′	1040	1	345	0.196	0.296	67.44
Mean							0.259	0.384	81.14

AMT, annual mean temperature; AMP, annual mean precipitation; H, Nei's (1973) gene diversity; I, Shannon's information index; PPB, percent of polymorphic locus.

Turkey, indicate that genetic variation in wild barley populations is not only common, but also at least partly, adaptive (Nevo et al., 1986).

Some PCR-based marker systems are widely used to study genetic diversity in Caragana plant species, such like random amplified polymorphic DNA (RAPD), AFLP and simple sequence repeats (SSRs). Inter-simple sequence repeat (ISSR) markers with low cost and low labor requirement but with high reliability have been developed to measure genetic diversity on plant. Like other PCR-based markers, they are rapid and require only small amount of the template DNA, and it seems to have the reproducibility of SSR's and the usefulness of RAPD's. It is applied to the study of genetic relationships and phylogenetic analysis of various plant species.

In the present study, the genetic diversity of 17 wild *C. microphylla* Lam. populations was examined using ISSR markers. The major objectives are to determine genetic diversity; quantify the amount and distribution of genetic variation within and among populations; and assess genetic structure for the native *C. microphylla* Lam. in China.

MATERIALS AND METHODS

Population sampling

A total of 510 individuals from 17 *C. microphylla* Lam. populations were collected in the summer of 2006 and 2007 from Inner Mongolia for this study (Table 1 and Figure 1). There were 30 individuals for each population. Each individual was selected by random and contains at least 100 seeds. 5 to 10 seeds from each individual were sown in the greenhouse and the leaves of 3-weeks seeding were harvested as one sample. All the leaves were stored at -70°C until they were used for DNA preparation.

DNA extraction

Total genomic DNA was isolated following the modification of the SDS "miniextraction" protocol developed (Edwards, 1991) which involved an additional phenol-chloroform extraction, alcohol precipitations and acetate salt rinse. The DNA sample was resuspended in 1× TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), and diluted to 50 ng/µl and stored at 4°C or -20°C for polymerase chain reaction.

ISSR analysis

Eighteen ISSR primers were selected from a total of 100 different SSR primers (The University of British Columbia) after being tested. Each 25 µl amplification reaction consists of 1.5 µl Mg^{2+} (25 mM), 0.16 µl dNTPs (40 mM), 0.12 µl Taq polymerase (5 U/µl)10 mM and 50 ng of genomic DNA. All the PCR amplifications were performed on a Mastercycler Gradient PCR (Eppendorf Research, Inc.) and consist of a denaturing step of 5 min at 94°C, followed by 40 cycles, each including 45 s at 94°C for denaturation, 1 min for annealing, 1.5 min at 72°C for extension and final extension at 72°C for 10 min. Depending on the G/C composition, the annealing temperature varied from 47 to 58°C.

The PCR products were separated on 1.5% agarose gel in a Tris-acetate-EDTA buffer and 100 bp ladder was used to label and determine the size of the ISSR bands. It was stained with ethidium bromide and visualized under UV using a gel documentation system. The polymorphic DNA bands were scored as present (1), or absent (0) for each DNA sample, excluding the smeared and weak ones by visual inspection.

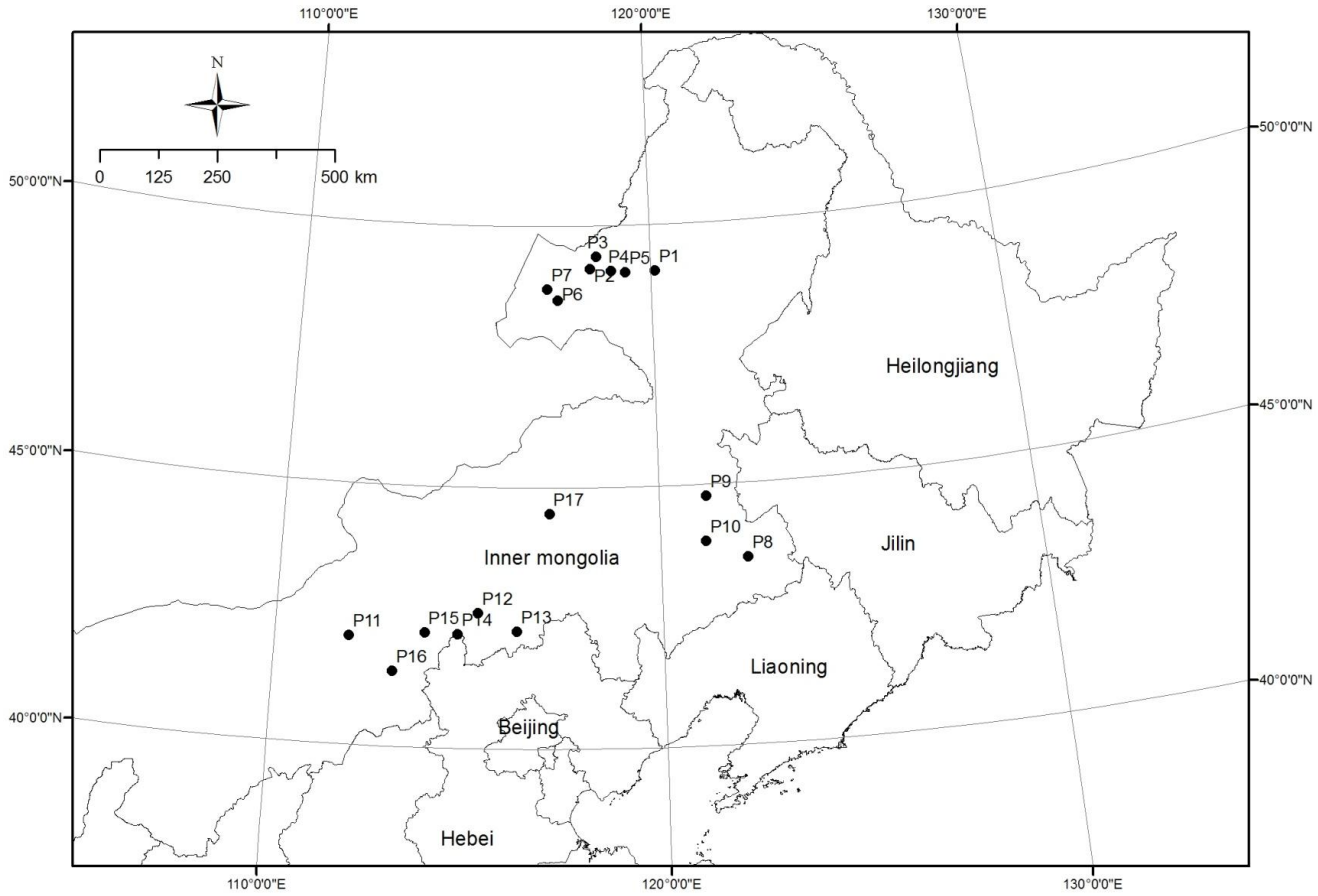

Figure 1. Distribution of the 17 populations of *C. microphylla* Lam. in this study.

Data analysis

The percentage of polymorphic loci (P) and Shannon's information index of diversity (I) were analyzed using POPGENE 1.31 (Yeh et al., 1999). The amount of gene flow among these populations was estimated as $N_m = (1/G_{ST} - 1)/4$. The analysis of molecular variance (AMOVA) was conducted to calculate variance components and their significance levels for variation between plants from Hunshandake sand, Horqin sand land and Hulunbuir sand land among populations within a region, and within populations using GenAlex 6.2.Nei's genetic distances (D) were calculated for each pairwise combination among populations (Nei, 1978). A principle coordinate analysis was also performed for displaying the population differentiation using the NTSYS-pc software (Rohlf, 1997). Cluster analysis using the Neighbor–Joining method was performed using NTSYS-pc. A matrix of geographical distances among populations was obtained by the use of Google earth. To test for a correlation between genetic distances and geographical distances among populations, a Mantel test was performed (Smouse et al., 1986).

RESULTS

Polymorphism and genetic diversity of *C. microphylla* Lam. populations

18 ISSR primers amplified a total of 269 bands for the 510 surveyed individuals. Of these bands, 217 (81.14%) are polymorphic. The total number of bands vary from 11 (UBC884 and UBC888) to 21 (UBC835) with a mean of 15 bands per primer (Table 3). The number of polymorphic bands ranged from 9 (UBC884 and UBC888) to 17 (UBC835) with a mean of 12 bands per primer. The value of Nei's gene diversity (H) was 0.259 and Shannon's information index (I) was 0.384 at species level (Table 1). Within populations, the PPB varied from 64.04 to 94.01%, and the mean value of H ranged from 0.192 (P11) to 0.323 (P5). The value of I shows similar trends, ranging from 0.289 (P11) to 0.475 (P5). The results show that P5 had the richest genetic diversity and P11 had the lowest. Partial correlation between eco-environmental factors and genetic diversity index (H, I and PPB) shows significant correlation from latitude and AMT, which means that impact from some environmental factors is evident (date not shown).

Genetic differentiation among and within populations

To estimate the distribution of genetic diversity of *C. microphylla* Lam., POPGENE 1.31 software was used to

Table 2. Constitution of gene diversity at all polymorphic loci of seventeen populations.

Parameter	Sample size	Ht	Hs	G_{ST}	N_m
Mean	510	0.3365	0.2798	0.2091	1.8912
S.D.		0.0073	0.0094		

Nm = estimate of gene flow from GST: Nm = 0.5 (1 - GST)/GST.

Table 3. Numbers of ISSR generated with 18 selective ISSR primer pairs of *C. microphylla* Lam.

Primer	Total number of band	Number of polymorphism band	Polymorphism (%)
U810	18	13	72.22
U811	15	11	73.33
U817	16	13	81.25
U825	19	14	73.68
U827	16	14	87.50
U834	16	13	81.25
U835	21	17	80.95
U841	17	13	76.47
U846	13	11	84.62
U847	17	13	76.47
U849	12	10	83.33
U860	13	12	92.31
U868	13	11	84.62
U884	11	9	81.82
U888	11	9	81.82
U889	13	11	84.62
U890	14	11	78.57
U891	14	12	85.71
Mean	14.9	12.1	
Sum	269	217	81.14

Table 4. Analysis of molecular variance (AMOVA) for the 17 populations of *C. microphylla* Lam. from China.

Source of variation	df	Sum of squares	Variance component	Percentage of total variation	*p* value
Among regions	2	1695.9	5.9	9.8	<0.001
Among populations	14	4892.1	11.5	19.1	<0.001
Within populations	493	13549.1	42.7	71.1	<0.001
Total	509	20136.9			

analyze genetic differentiation from ISSR date (Table 2). Nei's genetic differentiation (G_{ST}) was calculated to be 0.2091. That is, 20.91% of gene differentiation occurred among population and 79.09% within populations. It shows that higher level of genetic differentiation is relative to within populations. The gene flow number (N_m) is 1.8912, which indicates higher gene exchange between populations.

In addition, for the purpose of detecting the relationship among and within populations, molecular variance is examined using AMOVA based on ISSR banding patterns (Table 4). The variance component found within populations was 71.1, 19.1% was found among populations and 9.8% was found within regions. It is similar to the result of G_{ST}.

Principal coordinate analysis

A principal coordinate analysis based on the Jaccard values was undertaken (Figure 2); the first principal coordinate accounts for 30.29% of the variation, the second for 17.09% and the third for 15.80%. Although, there was considerable overlapping between samples

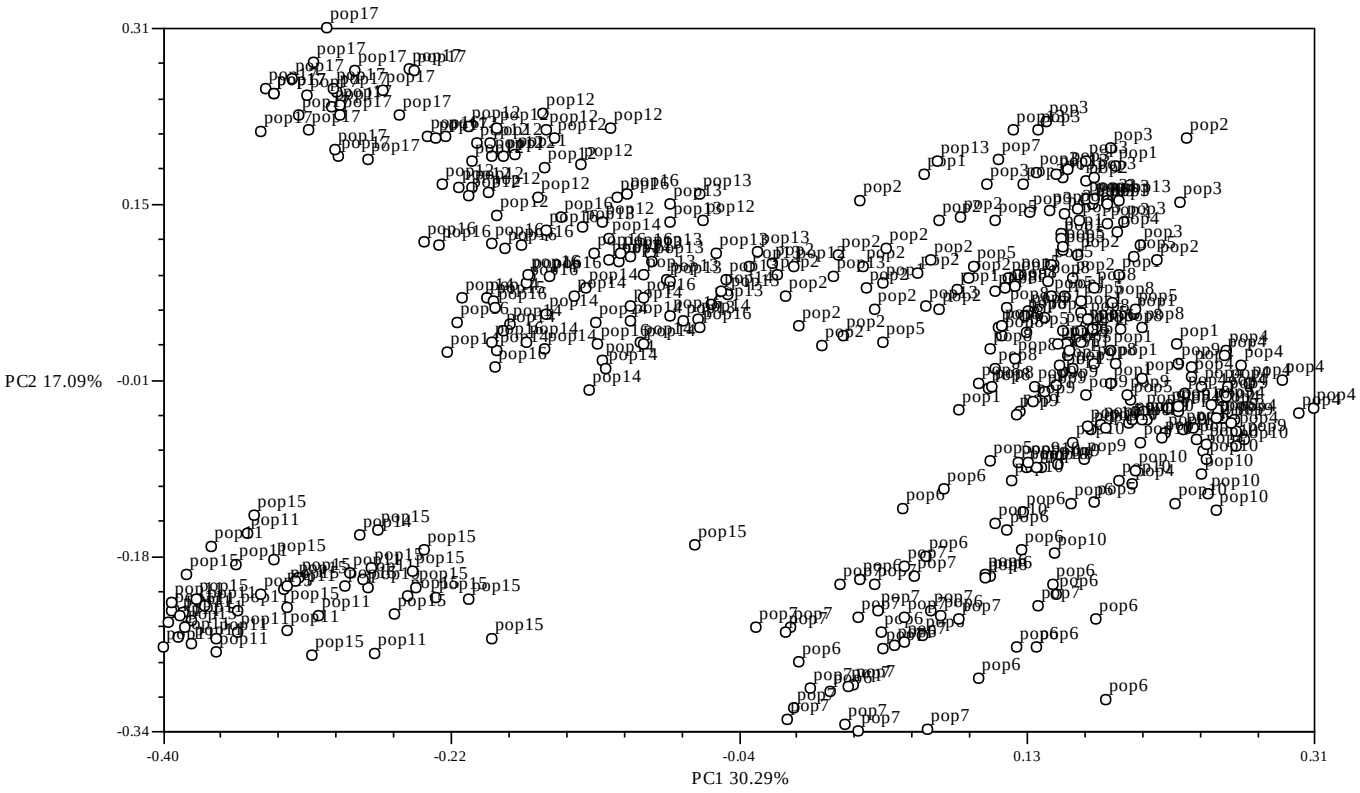

Figure 2. Dispersion graph from principal coordinate analysis of amplified fragment length polymorphism data of 408 individuals of *C. microphylla* Lam.

from different populations, some patterns could still be discerned. P11 and P15 populations were mostly found in the lower and right side of the diagram, and P6 and P7 populations were also found in the lower and right side of the diagram. P12, P13, P14 and P17 populations were found in the upper and right side of the diagram. The other populations were mostly found in the down and right side of the diagram. This diagram clearly suggests that there is a relationship between geographical locations and the ISSR data.

Genetic distance and genetic identity among populations

To further elucidate the gene differentiation between populations, unbiased measure of Nei's genetic distance was evaluated (Table 5). Nei's genetic distance ranged from 0.1077 to 0.4902 with an average of 0.253. The largest genetic difference occurred between P10 and P17. Furthermore, relationships of 17 wild populations were revealed by cluster analysis based on Nei's genetic distance using the UPGMA method (Figure 3). Most populations were arranged in the same cluster by geographic location. To further clarify the relationships between populations, the Mantel's test was carried out between Nei's genetic distance and pairwise geographic distances. The matrix correlation (r^2) was 0.1251, and showed no significant correlation, which means that 12.51% genetic variance can be explained by geographic distance. The relationship between genetic diversity parameters and geographic and ecological factors of *C. microphylla* Lam. showed no significance.

DISCUSSION

Studies on genetic diversity of *C. microphylla* Lam. were undertaken by other researchers using other different molecular markers such as AFLP. AFLP is widely used nowadays because of its rapidity, numerous polymorphisms and reproducibility but tedious procedures and high expenses are its two major disadvantages. In the present study, ISSR markers were used to assess the pattern of genetic diversity in 17 wild populations of *C. microphylla* Lam. All individuals tested produced different ISSR profiles. Polymorphic bands (269), with percentage (PPBs = 81.14%), the average value of Nei's gene diversity (H = 0.259) and Shannon's information index (I = 0.384) at species level was generated by 18 ISSR primers. As compared to a previous study (Song, 2005), ISSR-PCR appears to produce reliable and highly polymorphic band profiles in *C. microphylla* Lam., which is also found in other plants (Fang and Roose, 1997;

Table 5. Nei's (1978) unbiased genetic distance among 17 *C. microphylla* Lam populations.

Pop ID	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17
P1	0																
P2	0.1698	0															
P3	0.1497	0.1781	0														
P4	0.1077	0.1858	0.1637	0													
P5	0.1313	0.1681	0.1275	0.1237	0												
P6	0.174	0.182	0.1806	0.1645	0.1312	0											
P7	0.2388	0.2071	0.2476	0.2297	0.2077	0.0923	0										
P8	0.1388	0.1897	0.147	0.1419	0.1182	0.1663	0.2333	0									
P9	0.1585	0.2308	0.1789	0.1684	0.1733	0.1859	0.2214	0.1644	0								
P10	0.2238	0.3289	0.2641	0.2081	0.2731	0.2993	0.3047	0.2659	0.2072	0							
P11	0.3825	0.3833	0.4047	0.403	0.3873	0.2986	0.2806	0.3634	0.3858	0.4074	0						
P12	0.3189	0.2659	0.302	0.3248	0.2589	0.2967	0.2965	0.3211	0.3421	0.354	0.3151	0					
P13	0.2342	0.2092	0.202	0.2385	0.2079	0.2048	0.2792	0.2018	0.2451	0.2871	0.3392	0.2388	0				
P14	0.2638	0.2508	0.251	0.2558	0.2411	0.213	0.2424	0.2537	0.2596	0.2946	0.2474	0.2386	0.174	0			
P15	0.2653	0.2891	0.2995	0.2851	0.2689	0.2289	0.2502	0.2461	0.3043	0.3647	0.1108	0.3235	0.2648	0.2502	0		
P16	0.2147	0.2716	0.213	0.2227	0.2035	0.2308	0.2823	0.2478	0.2456	0.3446	0.2773	0.2821	0.2499	0.2114	0.1931	0	
P17	0.2993	0.2976	0.344	0.3751	0.327	0.3247	0.3243	0.306	0.3726	0.4902	0.3113	0.2646	0.3165	0.298	0.32	0.316	0

Moreno et al., 1998; Prevost and Wilkinson, 1999; Yao et al., 2008). It was reported that there are some correlations between genetic diversity and ecogeographic variation (Fahima et al., 2002). In the present study, we showed that the correlations between intrapopulation genetic diversity with latitude and mean annual temperatures are significant. Yang et al. (2006) have reported that Shannon's information index of *Caragana davazamcii* increased along with changes of the habitats from east to west and with lack of average annual rainfall; the arid environment of Korshinsk peashrub is characterized by low precipitation and high annual precipitation fluctuations (Geng, 1986). Wang et al. (2007) showed that there is a significant correlation between intrapopulation genetic diversity of *C. korshinsk* with latitude. Many studies also showed similar results, and relationships between levels of genetic polymorphism and degree of environmental stress are strong (Hedric, 1986; Nevo, 2001). Other ecological factors and the limited selected sampling site may influence these results. Across the species' range, natural populations are found to show a high level of genetic diversity. Across the 17 populations of *C. microphylla* Lam. surveyed for ISSR variation, Nei's genetic differentiation (G_{ST}) indicates a fairly high level of population differentiation within populations. Only small amount of genetic differentiation among populations corresponds to G_{ST} equalling 0.209, which reveals that 79.01% of genetic variation comes from the component within population. It could be explained by frequent natural out-crossing, and seed dispersal which lead to high gene flow (Nm = 1.8912) among different populations. The result is further confirmed by the AMOVA analysis. Of the total genetic diversity, 74.99% of the variance occurred among individuals within populations and 25.01% occurred among populations. Similar results were obtained for *Caragana polourensis* (91.7%) (Liu et al., 2005) and *C. korshinskii* (77.8%) (Wang et al., 2007). A high level of population differentiation may be explained by several factors, including the species' breeding system, genetic drift or geographic isolation of populations (Hogbin and Peakall, 1999). Higher level of gene flow and genetic drift might have influenced the extent of differentiation among *C. microphylla* Lam. Populations. The G_{ST}-derived N_m value of 1.8921 is indicative of a mass of gene flow among natural populations, and this value is actually over the level (approximate $N_m = 1$) needed to counteract

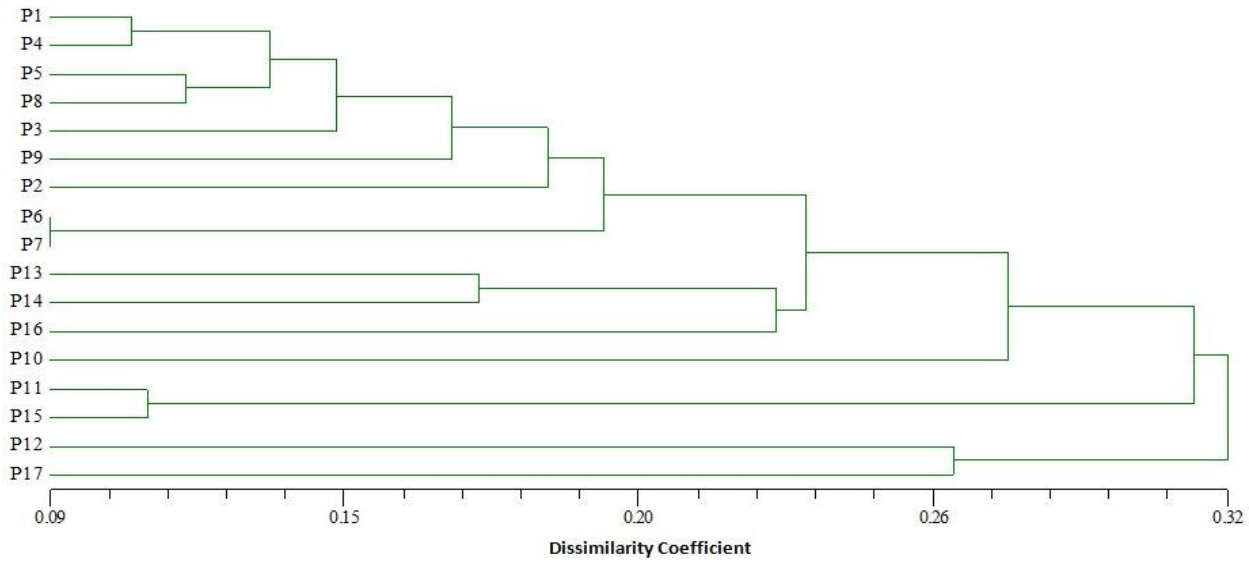

Figure 3. UPGMA-derived dendrogram showing the clustering of the 17 populations of *C. microphylla* Lam. based on Nei's (1978) genetic distance.

genetic drift (Slatkin, 1993).

ACKNOWLEDGEMENTS

Forage germplasm conservation project from ministry of Agriculture of China' and International cooperation project in science and technology(2010DFR30620-2).

REFERENCES

Fahima T, Roder MS, Wendehake K, Kirzhner VM, Nevo E (2002). Microsatellite polymorphism in natural populations of wild emmer wheat, *Triticum dicoccoides*, in Israel. Theor. Appl. Genet., 104: 17-29.

Fang DQ, Roose ML (1997). Identification of closely related citrus cultivars with inter-simple sequence repeat markers. Theor. Appl. Genet., 95: 408-417.

Geng KH (1986). Climate in China desert. Sci. Press. Beijing.

Hedric PW (1986). Genetic polymorphism in heterogeneous environment: A decade later. Annu. Rev. Ecol. Syst., 17:535-566.

Hodgkin PM, Peakall R (1999). Evaluation of the conservation of genetic research to the management of endangered plant *Zieria prostrata*. Conserv. Biol., 13:514-522.

Li J, QU SZ (2001).The study progress on the genus *Caragana* Fabr. Bull. Bot. Res., 21(4):35-38.

Li YG, Li LH, Jiang GM, Niu SL, Liu M (2004). Traits of chlorophyll fluorescence in 99 plant species from the spare-elm grassland in Hunshandak Sandland. Photosynthetica, 42(2): 243-249.

Liu HX, Lin WH (2004). Oligomeric Stilbenes from the Root of *Caragana stenophylla*. Chem. Pharmac. Bull., 52(11): 1339-1341.

Liu MH, Li ZC, Chen HS, Ge XJ (2005). Genetic variation of *Caragana polourensis* (Leguminosae). Guihaia, 25(1): 53-55.

Moreno S, Martin JP, Ortiz J M (1998). Inter-simple sequence repeats PCR for characterize- ation of closely related grapevine germplasm. Euphytica, 101: 117-125.

Nevo E, Beiles A, Kaplan D (1988). Genetic diversity and environmental associations of wild emmer wheat in Turkey. Biomed. Life Sci., 68(3): 203-213.

Nevo E (2001). Evolution of genome–phenome diversity under environmental stress. Proc. Natl. Acad. Sci., 98: 6233-6240.

Prevost A, Wilkinson MJ (1999). A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. Theor. Appl. Genet., 98: 107-112.

Rohlf FJE (1997). NTSYS-pc: Numerical taxonomy and multivariate analysis system, version 2.2. Appl. Biostatistics, New York.

Slatkin M (1993). Isolation by distance in equilibrium and non-equilibrium populations. Evolution, 47: 264-279.

Smouse PE, Long JC, Sokal RR (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Syst. Zool., 35: 627-632.

Song JS, Gao HW, Wang Z, Sun GZ, Jia M (2005). The phenotype variation of three *Caragana* species. Acta Prataculturae Sin., 14(3): 123-130.

Wang Z, Gao HW, Wu YQ, Han JG (2007). Genetic Diversity and Population Structure of *Caragana korshinskii* Revealed by AFLP, 47: 1737-1747.

Xu B, Gao HW (2009). Morphological Variation of *Caragana microphylla* Lam. in the East Part of Inner Mongolia, Bull. Bot. Res. 29(3):22-27.

Xu DH (2000). The genetic diversity of annual wild soybeans grown in China. Biomed. Life Sci., 103(1): 98-103.

Yang MB, Yang J, Yang JY, Zang CX (2006). Genetic analysis based on ISSR of *Caragana davazamcii* Sancz. Populations in different precipitation areas in Erdos Plateau. Plant Physiol. Commun., 42(3): 411-416.

Yao MZ, Chen L, Liang YR (2008). Genetic diversity among tea cultivars from China, Japan and Kenya revealed by ISSR markers and its implication for parental selection in tea breeding programmes. Plant Breed., 127: 166-172.

Yeh FC, Yang RC, Boyle T (1999). POPGENE version 1.32: Microsoft Windows–based freeware for population genetic analysis, quick user guide. Center for International Forestry Research, University of Alberta, Canada.

Zhang H, He H (2005). Preliminary Study on the Dustfall in the Zone Sheltered by Sand-Fixation Forest, arid land geography. Arid Land Geogr., 34(3): 15-19.

Zhang TH, Su YZ, Cui JY, Zhang ZH, Chang XX (2006). A leguminous shrub (*Caragana microphylla*) in semiarid sandy soils of north China. Pedosphere, 16: 319-325.