Genetic diversity analysis of *Capparis spinosa* L. populations by using ISSR markers

C. Liu[1]*, G.P. Xue[1,2]*, B. Cheng[1], X. Wang[1], J. He[1], G.H. Liu[1] and W.J. Yang[1]

[1]Xinjiang Institute of Materia Medica, WLMQ, Xinjiang Province, China
[2]Shool of Pharmacy, Shihezi University, Shihezi, Xinjiang Province, China

*These authors contributed equally to this study.
Corresponding authors: C. Liu / W.J. Yang
E-mail: liu_chong02@163.com / wilfred3106@163.com

Genet. Mol. Res. 14 (4): 16476-16483 (2015)
Received July 15, 2015
Accepted September 20, 2015
Published December 9, 2015
DOI http://dx.doi.org/10.4238/2015.December.9.19

ABSTRACT. *Capparis spinosa* L. is an important medicinal species in the Xinjiang Province of China. Ten natural populations of *C. spinosa* from 3 locations in North, Central, and South Xinjiang were studied using morphological trait inter simple sequence repeat (ISSR) molecular markers to assess the genetic diversity and population structure. In this study, the 10 ISSR primers produced 313 amplified DNA fragments, with 52% of fragments being polymorphic. Unweighted pair-group method with arithmetic average (UPGMA) cluster analysis indicated that 10 *C. spinosa* populations were clustered into 3 geographically distinct groups. The Nei gene of *C. spinosa* populations in different regions had Diversity and Shannon's information index ranges of 0.1312-0.2001 and 0.1004-0.1875, respectively. The 362 markers were used to construct the dendrogram based on the UPGMA cluster analysis. The dendrogram indicated that 10 populations of *C. spinosa* were clustered into 3 geographically distinct groups. The results showed these genotypes have high genetic diversity, and can be used for an alternative breeding program.

Key words: *Capparis spinosa* L.; Inter-simple sequence repeat; Genetic diversity

INTRODUCTION

Capparis spinosa L. belongs to the family Capparidaceae. The geographical origin of *C. spinosa* is disputed, with supporters claiming origins in China, India, and central Asian. The plant grows in desert regions of China, including the Gobi. The flower buds of *C. spinosa* have long been employed in culinary and medicinal practices; the buds can be used as a flavoring in cooking or as a diuretic, antihypertensive, poultice and tonic in traditional medicine (Baytop, 1984). In China, *C. spinosa* is mainly distributed in the Xinjiang province. Its fruits have been used in traditional medicine to treat rheumatic arthritis and gout (Fu et al., 2007).

Understanding the level of genetic diversity and the population genetic structure is important for medicinal plant species, because this allows the establishment of effective and efficient conservation practices and can guide choices for their genetic management. Nowadays, it is possible to use several molecular methods to analyze the genetic variability in plant species. One of them, inter-simple sequence repeat polymorphisms (ISSR), have been successfully used for genetic analysis of medicinal plants, as they require no prior knowledge of the DNA sequence and are universally applicable as dominant markers (Sa et al., 2011) for rapid exploratory work on new species. Furthermore, ISSRs have been demonstrated to be useful for the analysis of inter- and/or intra-specific genetic diversity in different Gentianaceae species (Ge et al., 2005; Zhang et al., 2007; Yang et al., 2011; Zheng et al., 2011).

Genetic variation within and among natural populations is crucial for the long-term survival of a species. Especially for medicinal species, an accurate estimate of the genetic variation among or within its populations could be helpful to address its status (Brown et al., 1996; Dávila, et al., 1998; Ge et al., 1999; Camacho et al., 2001), and provide fundamental information in designing conservation programs (Neigel, 2002). In this study, for the first time the population-level genetic diversity of *C. spinosa* was analyzed using ISSR markers to determine the genetic differentiation between populations from different regions in Xinjiang. It will provide invaluable information for future *C. spinosa* L. conservation and management programs.

MATERIAL AND METHODS

Plant materials

Ten populations were sampled from 3 geographically separate locations in North, Central, and South Xinjiang (Figure 1). At each location, samples were randomly collected from 8-10 *C. spinosa* individuals from 10 sites (accessions) separated by a distance of approximately 10 m (Table 1).

DNA extraction

Genomic DNA was extracted from young leaves following the CTAB (Cetyl Trimethyl Ammonium Bromide) Method. DNA concentration and quality was assessed on 1% agarose gel using a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

ISSR amplifications

ISSR markers were produced by PCR using the genomic DNA and ISSR primers. PCRs

were conducted using 22 primers (Table 2) to characterize the 94 *Capparis spinosa* L. samples. The 20-μL mixture contained 10 ng template DNA, 2.0 μL 10X PCR buffer, 1.0 μL（0.1 mM）dNTPs (Tiangen, Beijing, China), 2% formamide, 100 nM each primer, 1.5 U Taq polymerase (Tiangen, Beijing, China), and double-distilled water. Amplifications were performed using a BioRAD C1000 Thermal Cycler (Applied Biosystems,California, America) with the following PCR program: 5 min initial denaturing at 94°C, 40 cycles of 94°C for 45 s, 1 min for annealing at the primer-specific melting temperature, and 72°C for 90 s, followed by a final extension of 5 min at 72°C. First, the PCR products were analyzed by electrophoresis on 1.0% agarose gel with 0.5X TBE buffer. All products were analyzed on 1.5% (p/v) agarose gels stained with loading buffer. A 100-bp DNA ladder was used as a size marker.

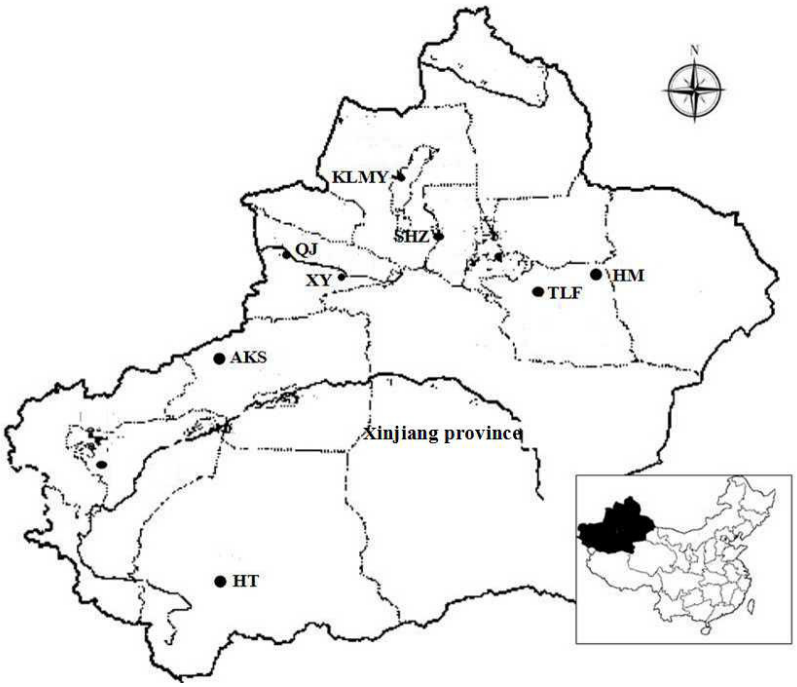

Figure 1. Map showing the distributions of *Capparis spinosa* and locations of ten populations sampled from Xinjiang.

Table 1. Populations of *Capparis spinosa* surveyed.

Location	Population name	Longitude	Latitude	Altitude (m)
Urumqi	WLMQ	87°33'23" E	43°47'18" N	890.6
Qianjin Pastureland	QJ	80°87'97" E	44°05'45" N	812.3
Kelamayi	KLMY	84°45'35" E	44°55'88' N	412.1
Shihezi	SHZ	85°89'68" E	44°73'09" N	490.7
Xinyuan	XY	83°25'23" E	43°18'19" N	816.5
hami	HM	81°32'84" E	43°92'41" N	790.3
Tulufan	TLF	89°11'60" E	42°51'63" N	75.5
Akesu	AKS	75°89'72" E	39°89'37" N	1088.2
Kashi	KS	76°02'13" E	39°75'11" N	1283.3
Hetian	HT	75°80'29" E	39°53'08" N	1275.4

Data analysis

The percentage of polymorphic loci (P), effective numbers of alleles (N_E), gene diversity (expected heterozygosity, H_E), Shannon's diversity index (I), and unbiased genetic distances according to Nei (1978), were calculated using PopGene32 (Yeh et al., 1999) for both markers. ISSRs are dominant markers, with each band representing the phenotype at a single biallelic locus. Only bands that could be unambiguously scored were used in the analysis. ISSR amplified bands were scored for band presence (1) or absence (0), and a binary qualitative data matrix was formed. The obtained genetic distance matrix was used to perform the cluster analysis and construct the unweighted pair-group method with arithmetic average (UPGMA) dendrogram using Power Marker software (Liu and Muse, 2005).

The third approach used the program Structure, version 2.3.4 (Pritchard et al., 2000), which identifies clusters of related individuals from multilocus genotypes. Individuals were assigned (probabilistically) to a cluster or jointly to two or more clusters if their haplotypes indicated that they are admixed; each cluster is characterized by a set of allele frequencies at each locus (Pritchard et al., 2000). To choose the best number of genetic clusters (K), multiple values were tested (from 1 to 7) using a length of burning period of 10,000 steps and 10 repetitions. The results were analyzed using the on-line tool, Structure Harvester (Earl et al., 2012), which implements the method of Evanno et al., 2005 to detect the true number of clusters in a non-homogeneous sample of individuals.

RESULTS

The amplification of the ISSR fragments in the 94 individuals, analyzed with ten primers, yielded 313 unambiguous and reproducible electrophoretic bands ranging from 5 to 15 bands for each of the primers, with an average of 9.3 bands per primer (Table 2). The estimates of genetic diversity in *C. spinosa* demonstrated a remarkable level of genetic variation among populations (PPL = 76.13%, Nei's gene diversity (H_E) = 0.2878); the average genetic diversity at the population level was H_E = 0.1686. The percentages of polymorphic loci for a single population ranged from 20.64% (XY) to 47.87% (TLF) with an average of 26.96% in the Tianshan Mountain populations (Table 3). According to the source of grow into WLMQ, QJ, KLMY, SHZ, XY and HM, TLF and AKS, KS, HT, The three group of Xinjiang, detection of genetic diversity of each area level. The results showed that the Nei gene of *C. spinosa* populations in different regions had a diversity (H) and Shannon's information index (I) range, respectively, of 0.1312-0.2001 and 0.1004-0.2855.The highest H_E and *I* value were for TLF populations; the lowest values were for the WLMQ population. The average Shannon's indices showed a strong correlation (Pearson's r = 0.999) with gene diversity. The values of gene diversity and Shannon's index showed a similar trend to that of the percentages of polymorphic loci.

Cluster analysis of *C. spinosa* populations

The neighbor-joining cladogram based on these genetic distances clustered the populations according to their geographical regions of origin (Figure 2). Three clusters of populations of *C. spinosa* from the Tianshan Mountains were obtained in three cladograms with high bootstrap values; the first group consisted of 5 populations from WLMQ, QJ, KLMY and SHZ,

all located in the northern part of the region. The second cluster was formed by AKS, KS and HT, the three populations located in the southern region. The third cluster included HM and TLF, the two populations located in the eastern portion of the sampled region. As can be seen from the clustering results, the geographic distribution of the 10 populations and the germplasm, which had a shorter geographical distance, tended to be in the same category. In general, the 10 *C. spinosa* populations had small genetic distance and high genetic similarity.

Table 2. Sequences and numbers of bands for 16 primers.

Primer	Sequence (5'-3')	Polymorphic amplicons	Polymorphism %	Ta (°C)
UBC807	AGA GAG AGA GAG AGA GT	11	91.2	55.2
UBC808	AGA GAG AGA GAG AGA GC	9	100.0	57.2
UBC811	GAG AGA GAG AGA GAG AC	6	100.0	54.2
UBC813	CTC TCT CTC TCT CTC TT	8	100.0	54.2
UBC815	CTC TCT CTC TCT CTC TG	8	100.0	50.1
UBC826	ACA CAC ACA CAC ACA CC	8	100.0	55.3
UBC844	CTC TCT CTC TCT CTC TRC	11	100.0	53.4
UBC888	BDB CAC ACA CAC ACA CA	11	100.0	55.0
UBC897	CCG ACT CGA GNN NNN NAT GTG G	11	100.0	56.8
UBC899	CAT GGT GTT GGT CAT TGT TCC A	10	100.0	56.8

R = (A, G), B = (C, G, T), N = (A, G, C, T), D = (A, G, T).

Table 3. Genetic variation in populations of *Capparis spinosa* detected by ISSR markers.

Geographic region	Population name	Sample size	N	PPL	N_A	N_E	H_E	I
North Xinjiang (north of Tianshan Mountains)	WLMQ	10	23	22.55	1.4225	1.223	0.1312	0.1004
	QJ	10	24	26.81	1.4681	1.2875	0.1649	0.1453
	KLMY	10	31	23.19	1.5319	1.25	0.1533	0.1385
	SHZ	10	33	22.13	1.5213	1.2955	0.1735	0.1615
	XY	8	29	20.64	1.6064	1.3397	0.1549	0.1017
East Xinjiang (The Eastern of Tianshan Mountains)	HM	9	52	42.13	1.5213	1.3091	0.1783	0.2675
	TLF	11	64	47.87	1.4787	1.2591	0.2001	0.2855
South Xinjiang (The Southern of Tianshan Mountains)	AKS	10	30	26.81	1.4681	1.2744	0.1594	0.1992
	KS	8	34	29.57	1.5957	1.3136	0.1884	0.1869
	HT	8	42	32.87	1.4787	1.3231	0.1817	0.1769
	Mean values		36.2	29.56	1.509	1.288	0.1686	0.1763
	Group			76.13	1.967	1.5091	0.2878	0.3614
	Total	94	362					

N = number of polymorphic loci; PPL = percentage of polymorphic loci; N_A = observed mean number of alleles per locus; N_E = effective mean number of alleles per locus; I = Shannon's information index; H_E = Nei's gene diversity.

The analysis of individual multilocus genotypes of the 94 samples using the Structure algorithm showed the best clustering solution for K = 3. Figure 3 shows that all individuals of the Tianshan Mountain *C. Spinosa* populations were assigned to the three remaining clusters: i) individuals from TLF and HM (eastern populations), ii) individuals from AKS, KS and HT (southern populations), and iii) individuals from XY, QJ, KLMY, SHZ and WLMQ (northern populations). Individuals showing probabilities of assignment to more than one cluster were observed in all of the Tianshan Mountains populations, revealing that there has been some gene flow between clusters. The proportions of individuals that have at least a 5% probability of assignment to another cluster were 23.1% for the eastern population group, 29.2% for the northern population group, and 10.1% for the southern population group.

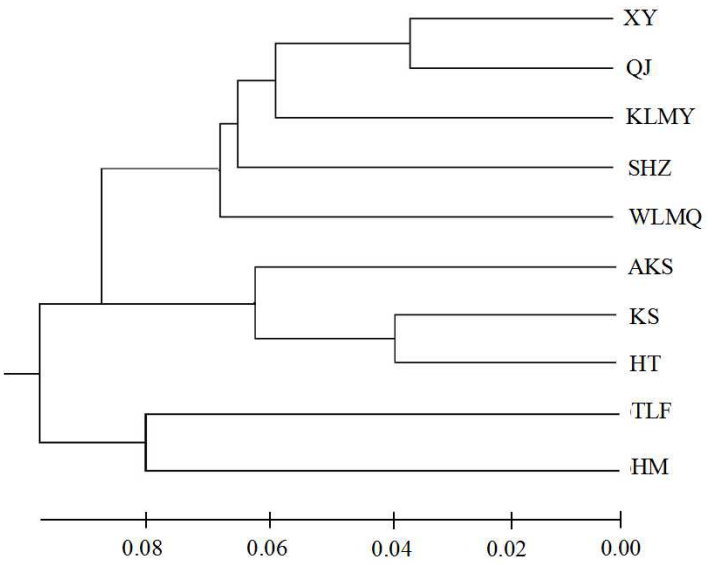

Figure 2. UPGMA clustering of ten *Capparis spinosa* populations based on Nei's genetic distance.

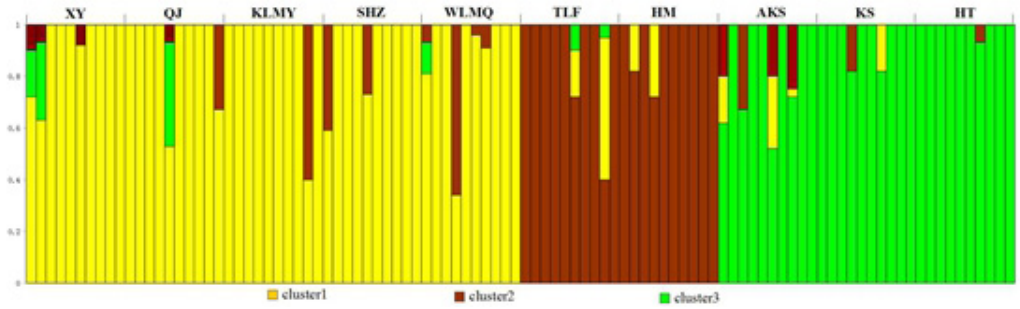

Figure 3. Genetic relationships among the 10 *Capparis spinosa* populations estimated using the STRUCTURE program based on ISSR data. The model with K = 3 showed the highest ΔK value.

DISCUSSION

C. spinosa is a perennial shrub with a patchy distribution. The populations are often located in distant mountain ranges and are isolated from each other by the Gobi or other deserts. In this sturdy, we report the first study of genetic polymorphism in *C. spinosa* populations using ISSR markers. The high levels of variation among *C. spinosa* populations may be mainly due to gene flow *N*m as low as 0.7093 ($Nm = (1 - G_{ST}) / 2G_{ST}$) When *N*m values are less than 1, it can lead to greater genetic variation in populations because genetic homogenization does not occur readily. In addition, the G_{ST} value of *C. spinosa* populations was 0.4134 which is within the range of G_{ST} values of vegetatively propagated plants (0.410-0.510) (Oscar et al.,2014). Asexual (clonal) reproduction

had a similar effect on the genetic structure of plants as strict self-pollination, which would cause large genetic variation among individuals as well as among populations, leading to a significant genetic differentiation among populations. To identify whether *C. spinosa* were vegetatively propagated plants or not, tools other than G_{ST} values should be used to investigate the reproductive origin. The first cluster, which contains populations WLMQ, QJ, KLMY, SHZ, and XY, had larger genetic similarity and medium-sized fruit. The remaining populations, with the exception of KS and HT, formed another cluster and had large fruit. Genetic differences may have been reflected in the phenotype of fruits; further research of *C. spinosa* genetic variation should be carried out in combination with morphological analyses. *C. spinosa* was discovered in TLF population range and is most widely used as a medicine. From this study, genetic analysis can be utilized to determine suitable growing conditions for planting and provide the basis for resource protection of *C. spinosa* and similar medicinal plants.

ACKNOWLEDGMENTS

Research supported by the National Natural Science Foundation of XinJiang (#2013211A109).

REFERENCES

Baytop T (1999). Therapy with medicinal plants in Turkey (past and present), Nobel Tip Kitabevleri, Istanbul, 2: 142.

Brown SM and Kresovich S (1996). Genome mapping in plants (Paterson AH, ed.). Clandes, New York, 85-93.

Camacho JC and Liston A (2001). Population structure and genetic diversity of *Botrychium pumicola* (Ophioglossaceae) based on inter-simple sequence repeats (ISSR). *Am. J. Bot.* 81: 965-972.

Dávila JA, Sánchez de la Hoz MO, Loarce Y and Ferrer E (1998). The use of random amplified microsatellite polymorphic DNA and coefficients of parentage to determine genetic relationships in barley. *Genome* 41: 477-486.

Earl D and von Holdt B (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4: 359-361.

Evanno G, Regnaut S and Goudet, J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14: 2611-2620.

Fu J, Huang H, Liu J, Pi R, et al. (2007). Tanshinone IIA protects cardiac myocytes against oxidative stress-triggered damage and apoptosis. *Eur. J. Pharmacol.* 568: 213-221.

Ge XJ and Sun M (1999). Reproductive biology and genetic diversity of a cryptoviviparous mangrove *Aegiceras corniculatum* (Myrtinaceae) using allozyme and inter-simple sequence repeat (ISSR) analysis. *Mol. Ecol.* 8: 2061-2069.

Ge XJ, Zhang LB, Yuan YM, Hao G, et al. (2005). Strong genetic differentiation of the east-Himalayan *Megacodon stylophorus* (Gentianaceae) detected by inter-simple sequence repeats (ISSR). *Biodivers. Conserv.* 14: 849-861.

Liu KJ and Muse SV (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21: 2128-2129.

Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.

Neigel JE (2002). Is FST obsolete? *Conserv. Genet.* 3: 167-173.

Oscar GL, Carlos P, Zsuzsanna G, et al. (2014). Genetic variation of the endangered *Gentiana lutea* L. var. *aurantiaca* (Gentianaceae) in populations from the Northwest Iberian Peninsula. *Int. J. Mol. Sci.* 15: 10052-10066.

Pritchard, JK, Stephens M and Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

Sa O, Pereira JA and Baptista P (2011). Optimization of DNA extraction for RAPD and ISSR analysis of *Arbutus unedo* L. leaves. *Int. J. Mol. Sci.* 12: 4156-4164.

Yang LC, Zhou GY and Chen GC (2011). Genetic diversity and population structure of *Swertia tetraptera* (Gentianaceae), an endemic species of Qinghai-Tibetan plateau. *Biochem. Syst. Ecol.* 39: 302-308.

Yeh F, Boyle T, Rongcai Y and Ye Z (1999). Popgene. Version 1.31. Microsoft Window-Based Freeware for Population Genetic Analysis.

Zhang XL, Yuan YM and Ge XJ (2007). Genetic structure and differentiation of *Gentiana atuntsiensis* WW Smith and *G. striolata* TN Ho (Gentianaceae) as revealed by ISSR markers. *Bot. J. Linn. Soc.* 154: 225-232.

Zheng P, Zhang K and Wang Z (2011). Genetic diversity and gentiopicroside content of four *Gentiana* species in China revealed by ISSR and HPLC methods. *Biochem. Syst. Ecol.* 39: 704-710.