African Journal of Biotechnology Vol. 10(40), pp. 7756-7765, 1 August, 2011
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.172
ISSN 1684–5315 © 2011 Academic Journals

Full Length Research Paper

Analysis of genetic diversity and construction of core collection of local mulberry varieties from Shanxi Province based on ISSR marker

Zhang Lin[1,2,3], Zhao Wei guo[2], Chen Jun bai[3], Huang Yong[3,4], Shen Xing jia[2,4], Liu Li[2] and Qiang Sheng[1]*

[1]Weed Research Laboratory, Life Sciences College of Nanjing Agricultural University, Nanjing 210095, China.
[2]The Sericultural Research Institute, Chinese Academy of Agricultural Sciences, Zhen jiang 212018, China.
[3]Biological and Environmental Engineering College, Jiangsu University of Science and Technology, Zhen jiang 212018, China.
[4]The Key Laboratory of Genetic Improvement of Silkworm and Mulberry, Ministry of Agriculture, Zhenjiang 212018, China.

Accepted 3 June, 2011

Genetic diversity of 73 local mulberry varieties from Shanxi Province were screened using ISSR markers, with I5 primers combinations selected for their reproducibility and polymorphism. 129 bands were amplified, of which 115 bands showed polymorphism and the ratio of polymorphism bands was 89.15%. Nei's genetic similarity coefficients ranged from 0.5891 to 0.9457 with an average of 0.7674. The observed number of alleles of each loci, effective number of alleles of each loci, Nei's gene diversity, Shannon's information index were 1.8915, 1.4771, 0.2780 and 0.4197, respectively. Clustering results showed that the 73 varieties could be divided into three different groups and nine subgroups. By using stepwise clustering and random methods and the modified heuristic algorithm, 21 core collections were constructed and the ratio of core collection was 28.77%. The result of t-test to the parameters (the number effective of alleles, Nei's genetic diversity index and Shannon's information index) showed that there was not significant difference between the core collection and initial sample with the exception of the number of observed alleles, that is, the core collection could well represent the initial sample.

Key words: Mulberry, germplasm resource, genetic diversity, ISSR, cluster analysis, core collection.

INTRODUCTION

The cultivated mulberry in China could be divided into 8 different eco-types and the local varieties of mulberry in Shanxi province is one of the 8 mulberry eco-types, becoming the most important component of mulberry gene bank of China. Through natural and artificial selection, the local varieties of mulberry in Shanxi province had adapted to the natural environment and formed the unique botanic and biological characteristics (Pan et al., 2000).

In contrast to morphology-based markers, DNA-based markers are unaffected by environment, detectable at all stages of development and ubiquitous in number covering the entire genome. ISSR molecular markers of simple and reproducible benefits have been used in cultivar identification and genome mapping, genetic distance analysis and population genetics studies.

Vijayan et al. (2004) led to the genetic diversity analysis on the Indian wild species of mulberry with 17 ISSR markers. Awasthi et al. (2004) identified the relationship of 15 mulberry species with six ISSR markers, Zhao et al. (2005, 2006a, b, c, 2007, 2008) studied on genetic diversity and phylogeny of cultivated and wild mulberry, diploid and homologous triploid mulberry, breeding variety, Feng Wei Sang and different ecotypes with ISSR markers. Prasanta et al. (2008) analyzed the genetic

*Corresponding author. E-mail: wri@njau.edu.cn. Tel: +86 511 85616570. Fax: +86 511 85628183.

variability and association of ISSR markers with some biochemical traits in mulberry (*Morus* spp.) genetic resources available in India. Huang et al. (2008) analyzed the genetic relationship of local varieties of *Morus alba L*. from Shandong province. Zhang et al. (2010) analyzed the genetic relationship of local varieties of mulberry from the lower area of Yellow River based on ISSR marker.

In 1984, the core collection construction was first proposed by Australian scholar Frankel, that is, with the small amount of genetic resources sample, to maximize the representative of the diversity of the main genetic resources. The aim of core collection construction is to give priority to the evaluation and utilization of core collection and to improve the management of genetic resource bank.

The traditional construction methods of core collection were based on the morphology or isozyme marker. Because morphology marker could be influenced by environmental factors easily and the results would not be accurate. DNA marker is rapid, accurate, efficient and not influenced by environmental factors, so it is an efficient method to construct core collection. Currently, by using the molecular marker, RFLP, RAPD, SSR, AFLP and so on, the construction of core collection of small germplasm samples has been reported: Sun et al. (2001) used RFLP to construct the core collection of common wild rice (*Oryza rufipogon* Griff.) and Asian cultivated rice (*Oryza sativa* L.). Hintum et al. (1994) used different DNA markers to construct the core collection of European spring barley (*Hordeum vulgare* S.) and compared the core collection to initial collection. Shen et al. (2001) used SSR to construct and evaluate the core collection of 120 collections of Yuannan local rice (*O. sativa* L.). Skroch et al. (1998) used RAPD to construct core collection of Mexico common soybean. Liu et al. (2006) used SSR and AFLP to construct and evaluate the core collection of 110 collections of pomelos (*Citrus grandis* Osbeck); the results showed that the core collection can well represent the initial collection.

In order to facilitate the preservation and further evaluation of germplasm of mulberry and in order to promote the management of the national mulberry germplasm gene bank in China, this research studied the diversity of 73 local varieties from Shanxi province with ISSR molecular marker, obtained the clustering UPGMA charts and constructed core collection using stepwise clustering and random sampling method based on the clustering results and clustering charts and finally evaluated as well the core collection using the related parameters of genetic diversity.

MATERIALS AND METHODS

Plant materials

The 73 local mulberry varieties used in this study were obtained from the national mulberry gene bank of the sericultural research institute, Chinese Academy of Agricultural Sciences (CAAS), Zhenjiang, Jiangsu Province, China. The county of origin and the number of the tested varieties are listed in Table 1.

DNA isolation

Total DNA was extracted from approximately 1.5 g of young leaves with the modified CTAB method (Zhao et al., 2000). The genomic DNA was quantified on 0.8% agarose gels and the samples were stored at -20°C for ISSR analysis.

ISSR amplification, separation and visualization

Twenty-two (22) ISSR markers (synthesized by Shanghai Bioasia Technology Co. Ltd., China) were screened using DNA samples from five varieties: Hong Ge Lu (73), Da Jing Sang (11), Ling Gu Da Ye (52), Jin Cheng Bai Pi Sang (56) and Yang Cheng Hei Ge Lu (61). Amplifications for screened primers and DNA samples were conducted independently for two to three times with the same procedure to verify the reproducibility and consistency of the ISSR markers. 15 primers were chosen out from 22 for ISSR analysis of genetic diversity based on their reproducible producing bands (Table 2).

PCR reactions were carried out in a volume of l5 µl containing 10 ng of total DNA, 10 × PCR buffer (200 mmol/l Tris-HCl pH 8.4, 2.5 mmol/l, 500 mmol/l KCl), 0.25 mmol/l of each dNTP, 6 pmol/l of each primer and 1 U of Taq DNA polymerase. The optimum annealing temperature was determined for each primer. PCR cycling conditions for all mulberry varieties (Flexigene thermal cycler) were: 2 min initial denaturation step (94°C); 36 cycles of 40 s at 94°C, 45 s at each specific annealing temperature and 90 s at 72°C; 7 min at 72°C. DNA fragments amplified were separated in 2.2% agarose gels at 90 W for 4 h in 1 X TBE buffer (100 mmol/l Tris-borate, pH 8. 0.2 mmol/l EDTA). The gel was dyed with ethidium bromide, visualized under ultra-violet light and photographed using a Kodak Digital Science 1D - EDAS 120 computerized gel analysis system. Molecular sizes of the amplified fragments were roughly estimated using a 2000 bp ladder (TaKaRa Dalian Biotechnology Co., Ltd., China).

Diversity data analysis

DNA banding patterns generated by ISSR were scored for the present (1) or the absent (0) of each amplified band and all ISSR assays were repeated twice and only distinct, reproducible, well-resolved bands were scored. Calculate the number of all PCR amplification bands and polymorphic bands per primer and evaluate the percentage of polymorphism. Nei'S (Nei and Li, 1979) gene diversity, Shannon's information index, genetic similarity, genetic distance estimated by Nei'S coefficient between pairs and dendrograms based upon the unweighted pair group method with arithmetical averages (UPGMA) were analyzed using Popgene software, version 3.5.

Construction method of core collection

A dendrogram of all 73 local varieties of mulberry based on the genetic similarity coefficient was generated by UPGMA cluster method. According to the clustering results and dendrogram, we used stepwise clustering and random sampling method to construct core collection, that is, according to a dendrogram, one accession of each group with two accessions of similar genetic variation was randomly chosen for next cluster, the accession went into next cluster if there was only one accession in a group. The sample from the first cluster was clustered and chosen again in the same way.

Table 1. Varieties of the local mulberries' origins and names.

Number	Variety	Origin
1	Nan He No.7	Shanxiqinshui county
2	Bai Ge Lu No.1	Shanxiyangcheng county
3	Bai Ge Lu No.2	Shanxigaoping county
4	Nan He No.24	Shanxiqinshui county
5	Ling Lu Shan Sang	Shanxilingchuan county
6	Wang Chuan Ge Lu	Shanxiwangchuan county
7	Hong Ge Lu	Shanxiyangcheng county
8	Yan Shan Sang	Shanxiyangcheng county
9	Jin Niu Er Sang	Shanxiyangcheng county
10	Rui Ma No.1	Shanxilingchuan county
11	Da Jin Sang	Shanxiyangcheng county
12	Qin Zhuang Sang	Shanxigaoping county
13	Hei Lu Sang	Shanxijincheng county
14	Multi-Friuts Sang	Shanxiwanrong county
15	Jin Bai Sang	Shanxi province
16	Jin Luo Sang	Shanxizhongyang county
17	Jin Hong Pi Wo Sang	Shanxiwanrong county
18	Chang Tiao Huang Lu	Shanxigaoping county
19	Yan Shan Sang No.7	Shanxiyangcheng county
20	Hong Yan Sang	Shanxiyuci county
21	Li Ye Sang	Shanxigaoping county
22	Lv Nai Nai Sang	Shanxiqinshui county
23	Yang Sang No.2	Shanxiyangcheng county
24	Wu Zhi Sang	Shanxigaoping county
25	Zhang Zhuang No.5	Shanxiqinshui county
26	Da Jing Sang	Shanxiyangcheng county
27	Huang Ke Sang	Shanxijincheng county
28	Ge Mo Sang	Shanxijincheng county
29	Xian Yi No.5	Shanxiyangcheng county
30	He Kou No.2	Shanxiqinshui county
31	Jin Da Hei Sang	Shanxiyangcheng county
32	Hong Ya Sang No.1	Shanxiwanrong county
33	Duan Shi No.1	Shanxiqinshui county
34	Xiao Hei Lian	Shanxigaoping county
35	Hong Ya Sang No.2	Shanxiwanrong county
36	Da Hei Lian	Shanxigaoping county
37	Zhong Yang No.1	Shanxizhongyang county
38	Zhong Yang No.2	Shanxizhongyang county
39	Zhong Yang No.3	Shanxizhongyang county
40	Zhong Yang No.4	Shanxizhongyang county
41	Liulin No.3hao	Shanxiliulin county
42	Liulin No.2hao	Shanxiliulin county
43	Shi Ye Shan Sang	Shanxiyangcheng county
44	Bai Guo Sang	Shanxiqinshui county
45	Zhang Zhuang No.4	Shanxilingchuan county
46	Che Pi Sang	Shanxilingchuan county
47	Ling Lu Sang	Shanxilingchuan county
48	Ling Chuan No.9	Shanxilingchuan county
49	Rui Ma No.2	Shanxilingchuan county
50	Ling Gu Sang	Shanxilingchuan county
51	Ling Qin Sang	Shanxilingchuan county

Table 1. continues.

52	Ling Gu Da Ye	Shanxilingchuan county
53	Nan He No.19	Shanxiqinshui county
54	Jin Cheng Huang Lu Tou No.1	Shanxijincheng county
55	Ling Chuan No.23	Shanxilingchuan county
56	Jin Ceng Bai Pi Sang	Shanxijincheng county
57	Bai Ge Lu No.3	Shanxiyangcheng county
58	Hei Ge Lu No.1	Shanxiyangcheng county
59	Jin Hei Ge Lu	Shanxiyangcheng county
60	Hei Ge Lu No.5	Shanxiyangcheng county
61	Yang Cheng Hei Ge Lu	Shanxiyangcheng county
62	Hei Ge Lu No.4	Shanxiyangcheng county
63	Jin Cheng Bai Ge Lu	Shanxijincheng county
64	Yang Cheng Bai Ge Lu	Shanxiyangcheng county
65	Ding Dian Bai Ge Lu	Shanxiyangcheng county
66	Jin Cheng Bai Ge Lu No.1	Shanxiyangcheng county
67	Heng He Hong Ge Lu	Shanxiyangcheng county
68	Ge Lu Sang	Shanxiqinshui county
69	Huang Ge Lu No.2	Shanxiqinshui county
70	Nan He No.26	Shanxiqinshui county
71	He Kou No.23	Shanxiqinshui county
72	Yang Cheng Huang Ge Lu	Shanxiyangcheng county
73	Huang Ge Lu No.1	Shanxiqinshui county

Table 2. List of primers, amplification conditions and polymorphism of ISSR markers used.

Primer name	Sequence (3′-5′)	Tm(℃)	Annealing temperature (℃)	Number of amplified band	Number of polymorphic band	Polymorphic percentage
ISSR01	$(GTGC)_4$	61.80	54	5	5	100.00
ISSR02	$(GAG)_4GC$	55.87	54	7	4	57.14
ISSR03	$(CT)_8TG$	55.02	56	8	8	100.00
ISSR04	$(AG)_8TA$	52.74	56	10	9	90.00
ISSR06	$(GA)_6GG$	50.01	56	11	11	100.00
ISSR08	$(CT)_8AC$	50.02	56	9	8	88.89
ISSR10	$(GACA)_4$	51.55	57	12	12	100.00
ISSR13	$(CT)_8GC$	57.30	57	9	9	100.00
ISSR14	$(AG)_8TC$	55.02	55	9	7	77.78
ISSR16	$(CT)_6GC$	50.01	52	9	8	88.89
ISSR17	$(AG)_6TC$	47.09	44	8	7	87.50
ISSR18	$(CT)_6TG$	47.09	42	8	6	75.00
ISSR19	$(AG)_6TA$	44.16	42	7	6	85.71
ISSR20	$(CT)_6CC$	50.01	56	9	9	100.00
ISSR22	$(CT)_6AC$	47.09	42	8	6	75.00
Total				129	115	
Mean				8.6	7.7	89.15

When the sample number meets the designed standard, the cluster was stopped and the core collection could be constructed by these accessions. The difference in genetic diversity between core collection and initial samples was measured by t-test for means, coefficient of variation and range (Hu et al., 2000, 2001).

According to the methods earlier, 62 varieties were selected out of the 73 varieties after the first round of the cluster, 47 varieties were selected out of the 62 varieties after the second round of the cluster, 30 varieties were selected out of the 47 varieties after the third round of the cluster, 21 varieties were selected out of the 30 varieties after the fourth round of the cluster, 18 varieties were selected out of the 21 varieties after the fifth round of the cluster, 15

varieties were selected out of the 18 varieties after the sixth round of the cluster. Finally, six sample groups, composed of 62, 47, 30, 21, 18 and 15 were obtained successively and were numbered as follows: Group i, ii, iii, iv, v, vi

Group i was composed of 62 collections, the local mulberry varieties number:
1,2,4,5,6,7,8,9,10,11,12,13,14,16,18,19,20,22,23,24,25,26,27,28,29,30,31,32,34,35,36,37,38,39,40,42,43,44,45,47,48,49,50,51,52,53,54,55,56,57,59,60,61,62,63,64,66,67,68,70,71,72.

Group ii was composed of 47 collections, the local mulberry varieties number:
1,4,6,7,8,9,10,11,12,13,16,19,20,23,24,25,27,29,30,32,34,35,36,37,38,39,42,43,44,45,47,48,50,51,52,54,55,56,59,60,62,63,66,67,70,71,72.

Group iii was composed of 30 collections, the local mulberry varieties number:
1,6,7,9,11,12,16,20,24,25,29,32,36,37,39,42,44,47,50,54,56,59,60,62,63,66,67,70,71,72.

Group iv was composed of 21 collections, the local mulberry varieties number:
1,9,11,16,20,24,25,29,32,36,39,44,47,54,59,62,66,67,70,71,72.

Group v was composed of 18 collections, the local mulberry varieties number:
1,9,11,16,20,24,29,32,36,39,44,47,54,59,66,67,71,72.

Group vi was composed of 15 collections, the local mulberry varieties number: 1,9,16,20,29,32,36,44,47,54,59,66,67,71,72.

Core collection data analysis

With the original data (0 and 1 composed matrix) gained from ISSR marker, we calculated the number of polymorphic loci, the percentage of polymorphic loci, the number of observed alleles, the number of effective alleles, Nei's gene diversity and Shannon's information index of the samples by PopGene32 software and did t-test by SPSS 13.0 software (Liu et al., 2006).

RESULTS

Levels of polymorphism revealed by ISSR-PCR markers

From prescreening assays with five mulberry varieties using 22 ISSR primers, 15 markers generated bright amplification products and polymorphisms and were used in further analysis (Table 2). A total of 129 reliable fragments were obtained. The number of bands per primer ranged from 5 to 12 with an average of each primer amplified 8.6 bands. Among them, 115 bands were polymorphic, accounting for 89.15%. The number of polymorphic bands per primer ranged from 4 to 12 with the average number of bands per primer being 7.7. The results of PCR amplification are given in Figure 1.

Genetic variation and cluster analysis of local varieties mulberry from Shanxi Province

Using the data from all PCR amplification bands shown by 15 ISSR markers, the genetic similarity matrix among all sources used in this study was obtained by multivariate analysis using Nei's coefficient. Similarity coefficients ranged from 0.5891 to 0.9457 with an average of 0.7674. The highest genetic similarity coefficient (0. 9457) was found between Bai Ge Lu No.1 hao(2) and Bai Ge Lu No.2(3),indicating that they are closely related. The lowest genetic similarity coefficient (0.5891) was found between Ge Mo Sang(28) and He Kou No.23(71), indicating that they are relatively remote in relationship. Evenly, each loci owned that the observed number of alleles, effective number of alleles, Nei's gene diversity, Shannon's information index was 1.8915, 1.4771, 0.2780 and 0.4197, respectively. List of genetic diversity information are given in Table 3. The observed number of alleles of each loci, effective number of alleles of each loci, Nei's gene diversity, Shannon's information index were 1.8915, 1.4771, 0.2780 and 0.4197, respectively.

A dendrogram was obtained by UPGMA method using the total number of amplified fragments of the 15 ISSR primers. Clustering results showed that the tested varieties could be divided into three different groups (49 mulberry cultivars were clustered into Group i, 23 mulberry cultivars were clustered into Group ii, only 1 mulberry cultivar Jin Newer Sang was clustered into Group iii) and nine subgroups obviously (A,B,C,D,E F,G,H and I) (Figure 2).

Group I was divided into five subgroups: A,B,C,D and E. Subgroup A consisted of 11 mulberry cultivars: Ling Lu Shan Sang(5),Black-green Sang(13), Multi-friuts Sang(14), Jin Bai Sang(15), Jin Luo Sang(16), Nan He No.7(1), Bai Ge Lu No.1(2), Bai Ge Lu No.2(3), Nan He No.24(4), Rui Ma No.1(10), Nan He No.19(53). Subgroup B consisted of 13 mulberry cultivars: Shi Ye Shan Sang(43), Che Pi Sang(46), Ling Lu Sang(47), Ling Gu Sang(50), Ling Qin Sang(51), Zhong Yang No.1(37), Zhong Yang No.3(38), Zhong Yang No.5(39), Zhong Yang No.4(40), Liu Lin No.3(41), Liu Lin No.2(42), Zhang Zhuang No.4(45), Rui Ma No.2(49). Subgroup C consisted of 13 mulberry cultivars: Yan Shan Sang(8), Qin Zhuang Sang(12), Jin Hong Pi Wo Sang(17), Chang Tiao Huang(18), Hong Yan Sang(20), Li Ye Sang(21), Lv Nai Nai Sang(22), Wu Zhi Sang(24), Da Jing Sang(26), Ge Mo Sang(28), Yang Shan Sang No.7(19), Yang Sang No.2(23), Zhang Zhuang No.5(25). Subgroup D consisted of 8 mulberry cultivars: Huang Ke Sang(27), Jin Da Hei Sang(31), Hong Ya Sang No.1(32), Duan Shi No.1(33), Xiao Hei Lian(34), Da Hei Lian(36), Xian Yi No.5(29), He Kou No.2(30). Subgroup E consisted of 4 mulberry cultivars: Wang Chuan Ge Lu(6), Hong Ge Lu(7), Da Jin Sang(11), Bai Guo Sang(44).

Group ii was divided into five subgroups: F, G, H and I cluster groups. Subgroup F consisted of 2 mulberry cultivars: Hong Ya Sang No.2(35), Nan He No.26(70). Subgroup G consisted of 3 mulberry cultivars: Ling Chuan No.9(48), Heng He Hong Ge Lu(67), He Kou No.23(71).Subgroup H consisted of 4 mulberry cultivars: Ling Gu Da Ye(52), Ling Chuan No.23(55), Jin Cheng Bai Pi Sang(56), Jin Cheng Huang Lu Tou No.1(54). Subgroup I consisted of 14 mulberry cultivars: Bai Ge Lu

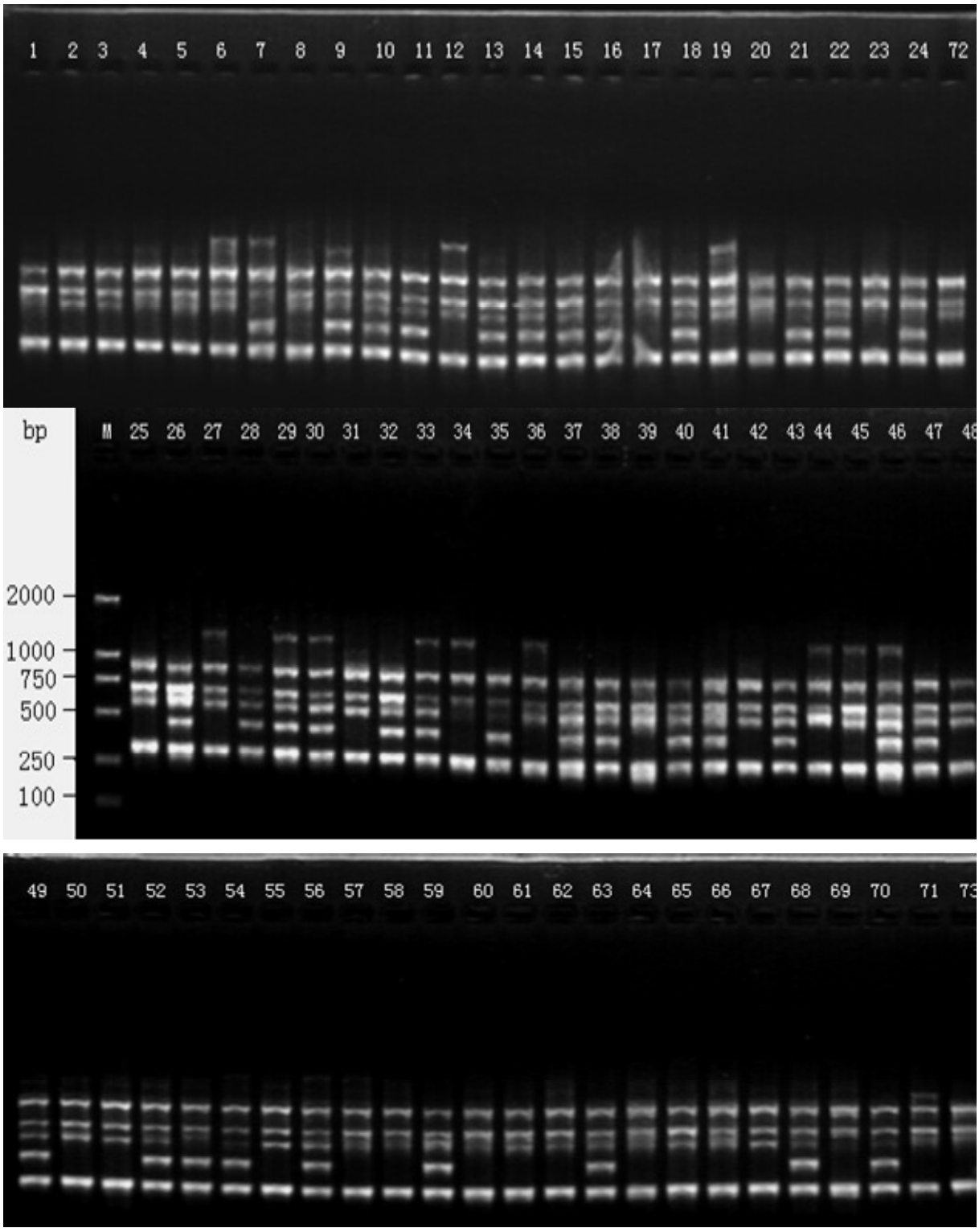

Figure 1. Electrophoretic pattern of 73 mulberry varieties amplified by primer ISSR02. The numbers in the figure are the same as those listed in Table 1. M is the DNA marker (DL2000).

No.3(57), Hei Ge Lu No.1(58), Jin Hei Ge Lu(59), Yang Cheng Hei Ge Lu(61), Hei Ge Lu No.4(62), Jin Cheng Bai Ge Lu(63),Yang Cheng Bai Ge Lu(64), Ding Dian Bai Ge Lu(65), Ge Lu Sang(68),Yang Cheng Huang Ge Lu(72),Hei Ge Lu No.5(60), Jin Cheng Bai Ge Lu No.1(66), Huang Ge Lu No.2(69), Huang Ge Lu No.1(73).

Table 3. List of genetic diversity index.

Parameter	Observed the number of alleles (NA)	Effective number of alleles (NE)	Nei's gene diversity (H)	Shannon's information index (I)
Mean	1.8915	1.4771	0.2780	0.4197

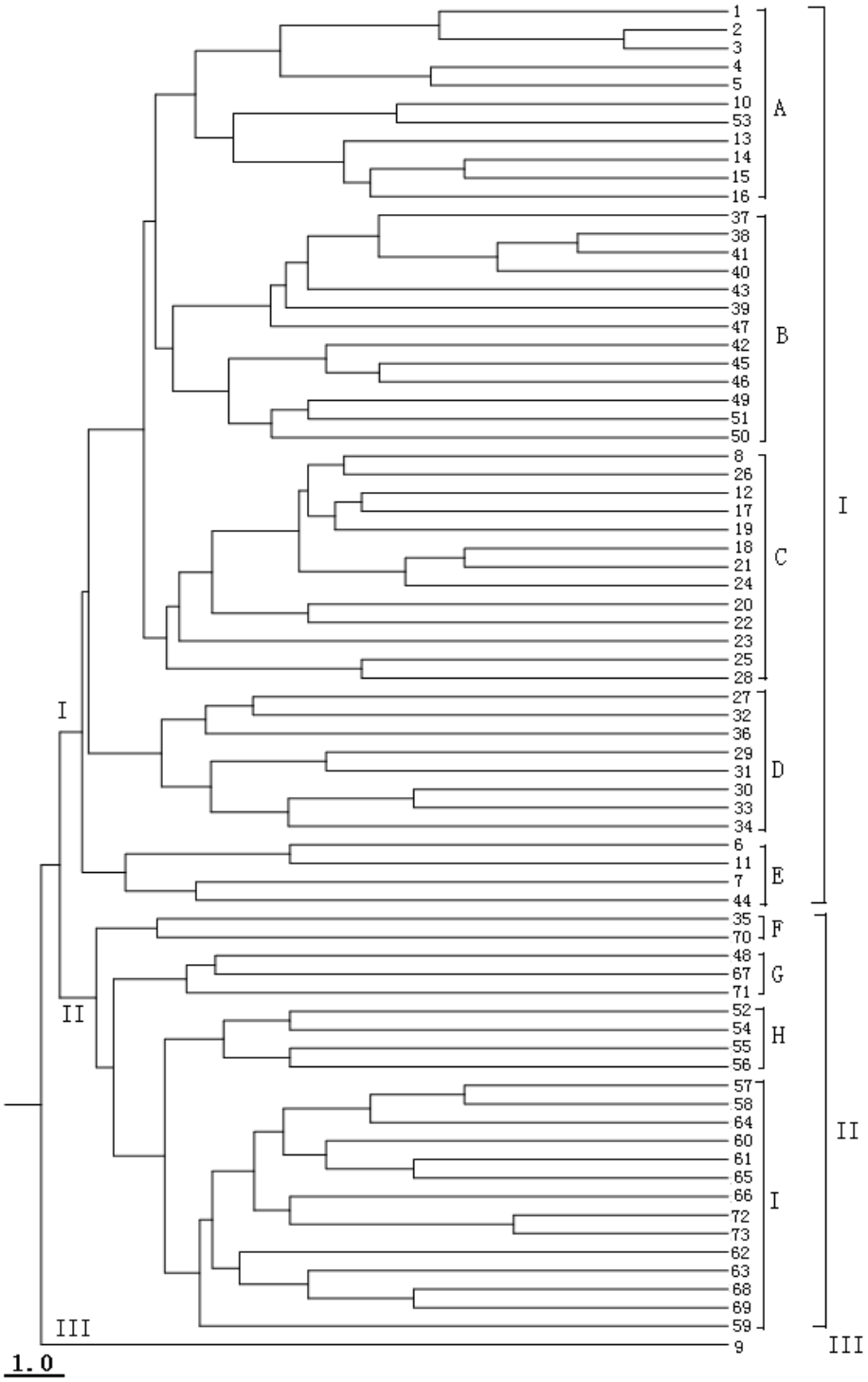

Figure 2. A dendrogram obtained by UPGMA for 73 mulberry cultivars based on ISSR markers. The numbers in the figure are the same as those listed in Table 1.

Table 4. Comparison of genetic diversity among different sampling groups.

Group	AS	PS (%)	NPL	PPL (%)	NA	NE	H	I
Initial I	73	100	115	89.15	1.8915	1.4771	0.2780	0.4197
Group	62	84.93	114	88.37	1.8837	1.4842	0.2818	0.4247
Group i	47	64.38	113	87.60	1.8760	1.4932	0.2871	0.4317
Group ii	30	41.10	111	86.05	1.8605	1.4990	0.2886	0.4325
Group iii	21	28.77	108	83.72	1.8372	1.4990	0.2898	0.4337
Group iv	18	24.66	107	82.95	1.8295	1.4876	0.2853	0.4284
Group v	15	20.55	107	82.95	1.8295	1.4812	0.2842	0.2842
Group vi								

AS= amount of sample; PS= percentage of sample; NPL = number of polymorphic loci; PPL = percentage of polymorphic loci; NA = observed the number of alleles; NE = effective number of alleles; H = Nei's gene diversity; I = Shannon's information index.

Core collection construction

With the stepwise clustering and random sampling method, six primary core collection groups (i, ii, iii, iv, v, vi) were chosen out, which was composed of 62, 47, 30, 21, 18 and 15 collections, respectively. The ratio of primary core collection samples was 84.93, 64.38, 41.10, 28.77, 24.66 and 20.55%.The number of polymorphic loci, the percentage of polymorphic loci, the number of observed alleles, the number of effective alleles, Nei's genetic diversity and Shannon's information index of Group i, ii, iii, iv, v, vi were calculated by PopGene32 software (the results are shown in Table 4). Comparing the genetic data gained with different groups, we discovered that the number of effective alleles, Nei's genetic diversity and Shannon's information index of Group iv which were composed by 21 samples were the highest among all the groups, although, the number of polymorphic loci, percentage of polymorphic loci, number of observed alleles of Group iv were lower than the initial Group and Group i, ii, iii. When the sampling rate falls to 24.66%, some of the molecular marker loci were lost due to sampling. Therefore, sampling rate of 28.77% is the best and Group iv could preserve the original diversity of samples. So, we regarded Group iv which was composed by 21 samples as the core collection. The core collections were as follows: Nan He No.7(1), Jin Niu Er Sang(9), Da Jin Sang (11), Jin Luo Sang (16), Hong Yan Sang (20), Wu Zhi Sang (24), Zhang Zhuang No.5 (25), Xian Yi No.5 (29), Hong Ya Sang No.1 (32), Da Hei Lian (36), Zhong Yang No.3 (39), Bai Guo San(44), Ling Lu Sang(47), Jin Cheng Huang Lu Tou No.1 (54), Jin Hei Ge Lu (59), Hei Ge Lu No.4 (62), Jin Cheng Bai Ge Lu No.1 (66), Heng He Hong Ge Lu(67), Nan He No.26 (70), He Kou No.23 (71), Yang Cheng Huang Ge Lu (72).

Comparison of core collection with initial sample

The core collection reserved 28.77% of initial sample, nevertheless, its retention rate of the number of polymorphic loci, the percentage of polymorphic loci, the number of observed alleles, the number of effective alleles, Nei's genetic diversity and Shannon's information index were 93.91, 93.91, 97.13, 101.48, 104.25 and 103.36 (Table 5), indicating that the core collection could remain the basic structure and the rich genetic diversity of the initial sample.

We did t-test to the parameters of the core collection and initial sample by SPSS software. The results showed that the core collection can well represent the initial sample (Table 6). As seen from Table 6, the variance of effective number of alleles (NE), Nei's gene diversity (H) and Shannon's information index (I) of the core collection were similar to that of the initial sample, the standard deviation of effective number of alleles (NE), Nei's gene diversity (H) and Shannon's information index (I) were not significant at 0.05 levels between the core collection and initial sample, with the exception of observed the number of alleles (NA).

DISCUSSION

China holds over 3000 collections of mulberry germplasm resource, containing 15 species and 4 subspecies. With the amount of mulberry germplasm resource gradually increasing, the conservation, evaluation, research, utilization and management of mulberry would become more and more difficult. The construction study of the core collection was of important significance for the management, utilization, evaluation and identification of germplasm resource. Correct evaluation on genetic similarity of different collections is the premise to construct the core collection; meanwhile, appropriate sampling methods and reasonable percentage were of great importance to construct the core collection. Generally, the sampling percentage was regulated according to the size of the initial collection (Boukema et al., 1997). A low sampling percentage, such as 5 to 10%, was adopted when the size of initial collection is large, whereas a high sampling percentage, such as 20 to 30%,

Table 5. Comparison of the genetic diversity between initial sample and core collection.

Group	Amount of sample	NPL	PPL (%)	NA	NE	H	I
Initial sample(a)	73	115	89.15	1.8915	1.4771	0.2780	0.4197
Core collection(b)	21	108	83.72	1.8372	1.4990	0.2898	0.4337
Percentage of b to a(%)	28.77	93.91	93.91	97.13	101.48	104.25	103.36

NPL = number of polymorphic loci; PPL = percentage of polymorphic loci; NA = observed the number of alleles; NE = effective number of alleles; H = Nei's gene diversity; I = Shannon's information index.

Table 6. T-test results of mean, std dev, difference mean, difference std dev, t value between initial sample and core collection.

Population	Mean	Standard deviation	Difference mean	Difference standard deviation	T value	Sig. (2-tailed)
NA (Initial sample)	1.8915	0.3123				
NA (Core collection)	1.8372	0.3706	0.0543	0.2274	2.710	0.008*
NE (Initial sample)	1.4771	0.3699				
NE (Core collection)	1.4990	0.3634	0.0219	0.1255	1.980	0.050**
H (Initial sample)	0.2780	0.1837				
H (Core collection)	0.2898	0.1818	0.0118	0.0588	2.285	0.024**
I (Initial sample)	0.4197	0.2461				
I (Core collection)	0.4336	0.2472	0.0139	0.0793	1.990	0.049**

NA = observed the number of alleles; NE = effective number of alleles; H = Nei's gene diversity; I = Shannon's information index; *indicates significant difference at 0.05 level between the core collection and initial sample. **indicates no significant difference at 0.05 level between the core collection and initial sample.

was adopted when the size of initial collection is small (Frankel and Brown, 1984). The size of sample in this study is 73, that is, a small sample. When the sampling rate in this study fell to 24.66%, some of the molecular marker loci were lost due to sampling. Therefore, the best sampling rate of 28.77% was obtained, in other words, the core collection construction in this study is in line with common practice in building the core collection when the size of initial collection is small. Chen et al. (2008) established core collection of mulberry germplasm resources from Shandong and Hebei province based on ISSR molecular markers. In this study, the core collection retained the initial 23.91%, the retention rate of core collection in the number of polymorphic loci, the percentage of polymorphic loci, the number of observed alleles, the number of effective alleles, Nei's genetic diversity and Shannon's information index has reached 89.02, 89.03, 95, 102.24, 103.99 and 101.26%. In our study, the core collection retained the initial 28.77%, the retention rate of core collection in the number of polymorphic loci, the percentage of polymorphic loci, the number of observed alleles, the number of effective alleles, Nei's genetic diversity and Shannon's information index has reached 93.91, 93.91, 97.13, 101.48, 104.25 and 103.36%, respectively. As was shown earlier, most of the parameters of the latter were higher than those of the former, that is to say this study created a good and representative core collection.

Although, the local mulberry varieties in Shanxi Province were rich and abundant and distributed in various localities, including Jincheng county, Lingchuan county, Gaoping county, Changzhi county and other places, but Xu proposed that Qinshui county and(or) Yangcheng county were the origin of the main mulberry varieties in Shanxi Province (Xu , 1991). Clustering results of this study was consistent with Xu′s view. The 73 varieties in Shanxi Province were clustered into 3 categories, 9 sub-categories (A,B,C,D,E,F,G,H,I,). From the categories to see, 22 species of Yangcheng county and 13 varieties of Qinshui county distributed in every category of the three categories; from the sub-category to see, all of eight sub-categories (A, B, C, D, E, F, G, I) contain varieties from Yangcheng county and/or Qinshui county, only one sub-category (H) with the exception that all the other counties of Shanxi Province introduced the mulberry from the Yangcheng county and/or Qinshui county in time to come. After the introduction and domestication, the locals might use the mulberry local varieties from Yangcheng county and Qinshui county as female or male breeding material, that is, there was gene flow between each other and the local mulberry varieties all over Shanxi Province had a common blood relationship, the clustering results of this study provides an evidence to Xu′s view " the origin of the main mulberry

varieties in Shanxi Province was Yangcheng county and/or Qinshui county.

ACKNOWLEDGEMENTS

We express our hearty thanks to Professor Shi bingkun and Pan yile of the Sericultural Research Institute, Chinese Academy of Agricultural Sciences, China. This work was supported by the mulberry germplasm resources conservation program of the Agriculture Ministry (NB2011-2130135-25).

REFERENCES

Awasthi AK, Nagaraja GM, Naik GV, Kanginakudru S, Thangavelu K, Nagaraju J (2004). Genetic diversity and relationships in mulberry (genus Morus) as revealed by RAPD and ISSR marker assays. BMC Genet. 5: 1-9.

Boukema IW, Hintum van TJL, Astley D (1997). The creation and composition of the *Brassica oleracea* core collection. Plant Genetic Resour. Newslett., 111: 29-32.

Chen JB, Huang Y, Zhang L, Zhao WG, Pan YL (2008). Construction of the core collection of morus multicaulis Perr. germplasm resources from Shandong and Hebei based on ISSR molecular markers. Acta. Sericol. Sin., 34(4): 587-592.

Frankel OH, Brown AHD (1984). Genetic perspectives of germplasm conservation.In: Arber W, Llimensee K, Peacock WJ, Starlinger PS, Genetic Manipulation: Impact on Man and Soctiety. Cambridge: Cambridge Uiversity Press. pp. 161-170.

Hu J, Xu HM, Zhu J (2000). Constructing core collection of crop germplasm by multiple clusters based on genotypic values. J. Biomathematics, (in Chinese). 15(1): 103-109.

Hu J, Xu HM, Zhu J (2001). A method of constructing core collection reserving special germplasm materials. J. Biomathematics, (in Chinese). 16(3): 348-352.

Huang Y, Zhang L, Zhao WG, Shen XJ, Pan YL (2008). ISSR analysis of genetic diversity on 24 local varieties of *Morus alba* Linne. Acta. Sericol. Sin. 34(2): 302-306.

Hintum van TJL (1994). Comparison of marker systems and construction of a core collection in pedigree of European spring barley. Theor. Appl. Genet. 89: 991-997.

Liu Y, Sun ZH, Liu DC (2006). Screening the core collection of pomelo germplasm based on molecular marker. J. Fruit Sci. (in Chinese). 23(3): 339-345.

Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA. 76(10): 5269-5273.

Pan YL (2000). Progress and prospect of germplasm resources and breeding of mulberry. Acta. Sericol. Sin. 26: 1-8.

Prasanta KK, Prem PS, Arvind KA (2008). Genetic variability and association of ISSR markers with some biochemical traits in mulberry (*Morus* spp.) genetic resources available in India[J]. Tree Genet. Genomes, 4(4): 925-938.

Sun SQ, Zeng YW, Li ZC (2001). Molecular markers and their applications for rice core collection in Yunnan. Seed, 119(6): 9-11.

Skroch PW, Nienhuis J, Beebe S, Tohme J, Pedraza F (1998). Comparision of Mexician common bean (*Phaseolus vulgaris* L.) core and reserve germplasm collections. Crop Sci. 38(2): 488-496.

Sun CQ, Li ZC, Wang XK (2001). Studies on evaluation of the genetic diversity of core collection of common wild rice (*Oryza rufipogon* Griff.) and Asian cultivated rice (*Oryza sativa* L.). Acta. Agronomica Sin. (in Chinese), 27(3): 313-318.

Vijayan K, Awasthi AK, Srivastava PP and Saratchandra B (2004). Genetic analysis of Indian mulberry varieties through molecular markers. Hereditas, 141(1): 8-14.

Xu YX (1991), Introduction of the local mulberry varieties in Shanxi province. Shanxi Agric. Sci. 1991(5): 24-25.

Zhang L, Huang Y, Shen XJ, Liu L, Zhao WG, Qiang S (2010). Analysis of genetic relationship of local varieties of *Morus alba* var. multicaulis from the lower area of Yellow River based on ISSR marker. J. Plant Resour. Environ. 19(2): 21-27.

Zhao WG, Pan YL, Huang MR (2000). RAPD analysis for the germplasm resources of genus mulberry. Acta. Sericol. Sin. 26(4): 1-8.

Zhao WG, Miao XX, Huang YP, Pan YL (2005). Analysis of genetic variation between diploids and their homologous tetraploids of mulberry by ISSR. Acta. Sericol. Sin. 31(4): 393-397.

Zhao WG, Wang W, Pan YL (2006a). Genetic variation between feng-wei-sang and its sports as revealed by ISSR markers. Bull. Sericulture, 37(3): 27-29.

Zhao WG, Zhou ZH, Miao XX, Wang SB, Zhang L, Pan Y, Huang YP (2006b). Genetic relatedness among cultivated and wild mulberry as revealed by inter-simple sequence repeat(ISSR)analysis in China. Can. J. Plant Sci. 86: 251-257.

Zhao WG, Zhang JQ, Wang YH, Chen TT, Yin Y, Huang YP, Pan Y, Yang YH (2006c). Analysis of genetic diversity in wild populations of mulberry from western part of northeast China determined by ISSR markers. J. Genet. Mol. Biol. 17(4): 196-203.

Zhao WG, Zhou ZH, Miao XX (2007). A comparison of genetic variation among wild and cultivated Morus Species (Moraceae: Morus) as revealed by ISSR and SSR markers. Biodivers. Conserv. 16(2): 275-290.

Zhao WG, Wang W, Yang YH, Huang YP, Pan YL (2008). Genetic diversity of mulberry local varieties from different ecotype as revealed by ISSR analysis in China. Acta. Sericol. Sin. 34(1): 1-5.