INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
17–0333/2018/20–4–839–845
DOI: 10.17957/IJAB/15.0573
http://www.fspublishers.org

Full Length Article

Genetic Diversity of Primary Core Kernel-apricot Germplasms using ISSR Markers

Shu-Lin Zhang[1,2†], Han Zhao[1†], Meng-Pei Liu[3], Da-Li Fu[1], Yun-Ling Zhang[1], Ta-Na Wuyun[1] and Gao-Pu Zhu[1*]
[1]*Non-timber Forestry Research and Development Center, Chinese Academy of Forestry, Zhengzhou 450003, P. R. China*
[2]*School of Biology and Food-engineering, Anyang Institute of Technology, Anyang 100091, P. R. China*
[3]*School of Food and Bioengineering, Zhengzhou University of Light Industry, Zhengzhou 450002, P. R. China*
[†]Joint first author
[*]For correspondence: poog502@hotmail.com

Abstract

The aim of this study was to screen the germplasm for desirable traits such as high yield, better kernel quality and strong stress resistance, using collected germplasm resources to establish a core breeding population. Using inter-simple sequence repeat (ISSR) molecular markers, we analyzed 47 initially selected germplasm samples of kernel-apricot with high yield, good quality, and resistance to late spring frost for their nucleotide sequence polymorphism. Sixteen ISSR primers with high polymorphism, clear bands, and good resolution were selected from 100 primers. These primers amplified 126 loci from the test samples, of which 85.7% were polymorphic. The amplified bands were characterized by Shannon index (I) of 0.4167 and Nei's diversity index (H_e) of 0.267; the effective number of alleles (N_e) was 1.4286, genetic identity ranged from 0.4048 to 1.0000, and genetic distance was between 0.0008 and 0.9045. At the genetic similarity coefficient of 0.734, the 47 accessions of kernel-apricot germplasm were divided into three groups. In contrast, principal component analysis resolved these accessions into only two groups, indicating the differences between the two statistical methods in their assessment of the relationships among individual cultivars. The stepwise sampling revealed that the highest effective number of alleles was obtained after four rounds of sampling, whereas Nei's gene diversity and Shannon information indices reached maximum values after five rounds of sampling, although the percentage of polymorphic loci significantly decreased. The number of germplasm resources also dropped to about 27.7% of the total number before sampling, which was no longer representative. In summary, the 16 ISSR polymorphic primers screened and selected in this study, combined with four consecutive stepwise sampling events at minimum distance based on cluster analysis, screened 19 samples whose ratio reached to 40.4% of the whole core germplasm. The high *Ne*, *He*, and *I* of the core germplasms indicated that this strategy produces results that can be used for further kernel-apricot germplasm improvement.

Keywords: *Armerniaca* Mill.; Core germplasm; Kernel-apricot; ISSR; Molecular marker

Introduction

Kernel-apricot is the collective name for the unique kernel-producing genus *Armeniaca* Mill. of the family Rosaceae in China, which includes four species: *A. sibirica*, *A. vulgaris*, *A. cathayana*, and *A. dasycarpa* (Zhang *et al*., 2015; Fu *et al*., 2016; Xu *et al*., 2016). Kernel-apricot is widely distributed in arid and semiarid regions of northern China. Many different varieties have evolved that are highly adapted to adverse conditions such as drought and cold and to barren lands, thus having prominent ecological advantage (Fu *et al*., 2016; Xu *et al*., 2016). At the same time, kernel-apricot plants have remarkable economic value because of their high percentage of fatty acids, unsaturated fatty acids, oleic acid, proteins, amygdalin, dietary fibers, inorganic salts, and many trace elements needed by the human body (Wang, 2013; Rai *et al*., 2016). They have become key economic forest trees in arid and semiarid regions of northern China. With a cultivation area of about 1.35 million hm^2 and still increasing, the country has become the world's largest apricot producer (Xu *et al*., 2016). For long, the many different varieties planted across the regions have caused confusion in nomenclature and augmented the existing problems pertaining to the genetic background and relationship among the cultivars, resulting in slow development of new germplasm. Furthermore, low yield due to low cold hardiness of the seedlings and the phenomenon of "no harvest in nine years out of ten" in main cultivated varieties have become a major bottleneck in the development of kernel-apricot agriculture. Therefore, a survey and collection of local cultivars is imperative to establish a core germplasm and accelerate the screening and breeding of new species with high capacity for cold hardiness.

To cite this paper: Zhang, S.L., H. Zhao, M.P. Liu, D.L. Fu, Y.L. Zhang, T.N. Wuyun and G.P. Zhu, 2018. Genetic diversity of primary core kernel-apricot germplasms using ISSR markers. *Int. J. Agric. Biol*., 20: 839−845

Molecular markers are the currently preferred method of collection and evaluation of nucleotide sequence polymorphism in germplasm. This method has been widely used in plants such as cotton (Xu *et al*., 2004), teosinte (Fukunaua *et al*., 2005), cherimoya (Escribano *et al*., 2008), kernel-apricot (Ai *et al*., 2011; Ai *et al*., 2014), wild apricot (Liu *et al*., 2015a) and trifoliate orange (Zhu *et al*., 2015). Similarly, it has also produced remarkable results in evaluating genetic diversity of the genus *Armeniaca* including wild apricot (He *et al*., 2007), *Prunus armeniaca* (Bourguiba *et al*., 2010; Khadivi-khub *et al*., 2015; Cui *et al*., 2016), Siberian apricot (Bao *et al*., 2016), and another five *Armerniaca* species (Fu *et al*., 2016). Simple sequence repeat (SSR) molecular markers were used to study genetic diversity, structure, relationship, and fruit ripening of *A. vulgaris* (He *et al*., 2007; Khadivi-khub *et al*., 2015; Cui *et al*., 2016) and genetic relationship and phylogeny of *A. cathayana*, *A. vulgaris*, *A. sibirica*, *A. dasycarpa*, and *A. mume* (Fu *et al*., 2016). Ai *et al*. (2011, 2014), using sequence-related amplified polymorphism (SRAP) molecular markers, examined the genetic diversity and relationship between sweet kernel-apricot and other species of the genus *Armeniaca* and developed molecular fingerprints. Mir *et al*. (2012) employed random amplified polymorphic (RAPD) molecular markers to study the genetic diversity of 24 accessions of *A. vulgaris*, and Liu *et al*. (2015b) systematically evaluated the genetic diversity of sweet kernel-apricot with SSR and inter simple sequence repeat (ISSR) molecular markers. To construct the core germplasm for kernel-apricot based on molecular markers, Liu *et al*. (2015a) performed a stepwise clustering with allele preferred sampling strategy (APSS) combined with Nei and Li genetic distance (Nei and Li, 1979) and selected 35 most representative germplasms among the 135 samples of *A. vulgaris*. This method overcomes the shortcomings of the phenotypic trait analysis, such as the influence of growing season, variation in ecological adaptability, and human errors in statistics, significantly speeding up the construction of the core collection for the genus *Armeniaca* and conserving large areas of valuable land.

In 2000, our research group created the collection and initiated the preservation and evaluation of the kernel-apricot germplasm from the provinces of Inner Mongolia, Liaoning, Shaanxi, Xinjiang, Shanxi, and Hebei, aimed for breeding of high yield, good quality and species tolerant to late spring frost. More than 1500 germplasms have been collected and preserved since then, and those having a single trait or aggregate traits superior to the existing varieties were selected. To assess their level of DNA diversity and genetic structure and to select the most representative germplasms, the present study analyzed the genetic diversity of primary selected 47 samples of *A. cathayana*, *A. sibirica*, and *A. vulgaris* based on ISSR molecular markers. This selection was based on our prior observation and assessment of the samples for high yield, good quality, and cold hardy phenotypes. However, the core germplasms of *A. vulgaris* have been constructed by Liu *et al*. (2015a) based on ISSR molecular markers, but they have two significant differences with this study. First of all, the least distance stepwise sampling strategy (LDSS) was adopted here for the un-weighted pair-group method with arithmetic means (UPMGA) cluster analysis, *i.e.* one sample was removed every time from the subgroup with the least genetic distance before clustering was performed again, until the genetic diversity could represent the initial collection and enriched representative germplasms were reached in a stepwise fashion. This strategy was different from the APSS method of Liu *et al*. (2015a) based on the number of rare alleles in the germplasms. Secondly, three species of the genus *Armeniaca*, species *A. cathayana*, *A. sibirica* and *A. vulgaris* were selected in this research. This gives our results broader application prospects than Liu *et al*. (2015a) who only used *A. vulgaris*. The report below details our study.

Materials and Methods

Plant Material and DNA Extraction

In the present study, 47 high-quality apricot germplasm samples were initially selected, comprising 39 samples of *A. cathayana*, 8 samples of *A. sibirica*, and 3 samples of *A. vulgaris*. The geographical distribution was as follows: 14 samples were from Hebei, 28 from Henan, 3 from Inner Mongolia, and 1 each from Beijing and Liaoning (Table 1). All the samples were planted in the germplasm nursery of the Non-Timber Forestry Research and Development Center, Chinese Academy of Forestry, Yuanyang County, He'nan, China. In July 2015, new leaves from the top branches of all 47 samples were collected, wrapped in aluminum foil, and placed in cryotubes and transported to the lab in a 10-L liquid nitrogen tank, to be stored at -80°C.

PCR Reactions

Total DNA was extracted from the 47 samples with a Plant DNA out Column (Tiandz, Inc., Beijing, China) plant DNA extraction kit. Its purity and concentration were analyzed using 1% agarose gel electrophoresis and a UV spectrometer. The DNA samples were all diluted with Tris/EDTA buffer and kept at -20°C for future analysis. Sixteen primers that produced clear reproducible bands with high degree of polymorphism were selected from 100 ISSR primers provided by the Columbia University and used for PCR amplification of the 47 DNA samples. PCR reactions were conducted in total volume of 15 μL, which included 7.5 μL of 2× Taq PCR MasterMix (TaKaRa, Shenyang, China), 6.5 μL of ddH_2O, and 0.5 μL (5 $\mu molL^{-1}$) primer. Amplification conditions for ISSR were as follows: initial denaturation at 94°C for 4 min; 40 cycles of denaturation at 94°C for 30 s, annealing at 50–60°C for 30 s, and extension at 72°C for 30 s; and a final extension at 72°C for 10 min. The PCR product was kept at 4°C. Agarose gel

electrophoresis was performed on 1.5% gel with 100 V for 40 min (DYY-12, Liuyi, Beijing, China), and the bends were imaged and analyzed with a gel imaging system (Gel Doc 2000, Bio-Rad, Hercules, CA, USA).

Data Analysis

ISSR markers are dominant markers. The amplification products of different samples were scored as either 1 or 0 according to the presence or absence of the band on agarose gel. These numbers were compiled into an ISSR data matrix consisting of 0 and 1 suitable for statistical analysis. Using POPGENE 32 software (Yeh *et al*., 1999) the following genetic diversity indices were obtained for the amplification products: the number of polymorphic loci (P_n), percentage of polymorphic loci (*P*), Shannon information index (*I*), Nei's gene diversity (H_e), effective number of alleles (N_e), genetic identity (G_i), genetic distance (*D*), and genetic similarity (G_s). Cluster analysis was conducted using the UPGMA method in NTsys 2.10e cluster analysis software (Rohlf, 2004).

Construction and Evaluation of Primary Core Germplasms

After the cluster analysis of the germplasm with NTsys 2.10e software, the core collection was constructed using the LDSS by resampling of the dendrogram based on genetic similarity coefficients (Xu *et al*., 2004; Escribano *et al*., 2008) according to modifications proposed by Liu *et al*. (2015a). Specifically, using the UPGMA dendrogram based on the 47 accessions, the cluster analysis was repeated on a subset of samples that was created by randomly removing one accession from a subgroup with higher genetic similarity coefficient. This procedure was repeated until the core germplasm that satisfied the selection criteria was selected. Each sampling was based on the cluster analysis of the preceding sampling. After each sampling, POPGENE 32 was used to analyze the genetic diversity of the amplified ISSR data of the established core collection and to obtain multiple evaluations about the representativeness of the primary core collection.

Results

Levels of Polymorphism Inferred by ISSR Method

Results indicated that out of the 100 ISSR primers, 16 primers that were selected for amplification of 47 kernel-apricot samples produced clear, highly polymorphic, and highly reproducible bands (Table 2). The amplification of the 47 germplasm samples using these 16 primers resulted in a total of 126 bands—mostly 10 bands per sample (primer UBC835 and UBC 836) and less than 5 bands per sample (primer UBC846). The length of the amplified fragments ranged from 500 to 2200 bp. Primers UBC827, UBC836, UBC841, UBC847, and UBC848 amplified with 100% success the polymorphic bands. Of the 126 bands, 108 (86.7%) were reproducible and polymorphic. The collected kernel-apricot germplasm from northern China was thus confirmed to be highly polymorphic and representative at the molecular level.

Analysis of Genetic Diversity of Kernel-apricot Germplasm

The 47 samples were highly diverse at the molecular level and highly representative of the kernel-apricot communities (Table 3). The average Shannon index (*I*) of the amplified bands was 0.4167, the average Nei's diversity index (*He*) was 0.267, and the average effective number of alleles (*Ne*) was 1.4286. The genetic identity (*Gi*) between the samples was 0.4048–1.0000 and the genetic distance (*D*) was between 0.0008 and 0.9045; the greatest genetic distance was observed between Ps89 from Xinxiang, Henan, and Pas35 from Huailai, Hebei, and the highest genetic consistency was registered between Ps108 and Ps143.

Cluster Analysis based on ISSR Sequence Data

ISSR amplification of 47 kernel-apricot accessions using 16 random primers produced 126 polymorphic bands, which were analyzed with NTsys 2.10e software. In the constructed similarity coefficient matrix, G_S ranged from 0.72 to 1.000. According to the dendrogram based on the ISSR sequences, the 47 samples were clustered into three groups at G_S= 0.734 (Fig. 1). The first group comprised Ps15 and Ps89 from Xinxiang, Henan; the second group included Pas01 and Pas02 from Zhangjiakou, Hebei, and Pas35 from Huailai, Hebei, while the remaining 42 accessions formed the third group.

Principal Component Analysis based on ISSR Sequence Data

Principal component analysis was conducted on the data obtained with ISSR markers. The results showed that the first principal component contributed 10.02% to the variance, the second contributed 8.56%, and the third 6.32%; the cumulative contribution was 24.9%. Using the values of the principal components of each germplasm sample, scatter plots for the first and second principal components were constructed (Fig. 2).

The principal component analysis resolved two groups within the 47 samples of kernel-apricot germplasm (Fig. 2). These results did not corroborate the results of the cluster analysis, although the grouping of Ps15 with Ps89, both from Xinxiang, Henan, and Pas01 and Pas02 from Zhangjiakou, Hebei, with Pas35 from Huailai, Hebei, was consistent for both methods. This difference in grouping between the two analytical methods is explained by the fact that cluster analysis is more informative about the relationship between closely related data, whereas principal component analysis provides more information about the aggregation into different groups.

Table 1: Samples origin and species in the study

Code	Species and individuals	Origin	Code	Species and individuals	Origin
1	Pas22/ *A. cathayana*	Huairou, Beijing	25	Zhongren No. 1 / *A. cathayana*	Luoyang, Henan
2	Pas35/ *A. cathayana*	Huailai, Hebei	26	Baiyubian / *A. cathayana*	Sanmenxia, Henan
3	Pas01/ *A. cathayana*	Zhangjiakou, Hebei	27	Fengren / *A. cathayana*	Sanmenxia, Henan
4	Pas02/ *A. cathayana*	Zhangjiakou, Hebei	28	Guoren / *A. cathayana*	Sanmenxia, Henan
5	Pas06/ *A. cathayana*	Zhangjiakou, Hebei	29	Longwangmiao / *A. cathayana*	Sanmenxia, Henan
6	Pas08/ *A. cathayana*	Zhangjiakou, Hebei	30	Sanganqi / *A. cathayana*	Sanmenxia, Henan
7	Pas144/ *A. cathayana*	Zhangjiakou, Hebei	31	Weixuan No. 1 / *A. cathayana*	Sanmenxia, Henan
8	Pas31/ *A. cathayana*	Zhangjiakou, Hebei	32	Youren / *A. cathayana*	Sanmenxia, Henan
9	Pas40/ *A. cathayana*	Zhangjiakou, Hebei	33	Yiwofeng / *A. cathayana*	Sanmenxia, Henan
10	Ps123 / *A. vulgaris*	Zhangjiakou, Hebei	34	Yangshaoxing / *A. vulgaris*	Sanmenxia, Henan
11	Pas03/ *A. cathayana*	Zhulu, Hebei	35	Ps108 / *A. sibirica*	Xinxiang, Henan
12	Pas04/ *A. cathayana*	Zhulu, Hebei	36	Ps143 / *A. sibirica*	Xinxiang, Henan
13	Pas05/ *A. cathayana*	Zhulu, Hebei	37	Ps15 / *A. sibirica*	Xinxiang, Henan
14	Pas07/ *A. cathayana*	Zhulu, Hebei	38	Ps46 / *A. sibirica*	Xinxiang, Henan
15	Pas09/ *A. cathayana*	Zhulu, Hebei	39	Ps89 / *A. sibirica*	Xinxiang, Henan
16	Pas5-5/ *A. cathayana*	Zhulu, Hebei	40	Boke No. 1 / *A. cathayana*	Yuanyang, Henan
17	Pas132/ *A. cathayana*	Luoyang, Henan	41	Xin No. 4 / *A. cathayana*	Yuanyang, Henan
18	Pas23/ *A. cathayana*	Luoyang, Henan	42	Youyi / *A. cathayana*	Yuanyang, Henan
19	Pas32/ *A. cathayana*	Luoyang, Henan	43	Kaitexing / *A. vulgaris*	Yuanyang, Henan
20	Pas44/ *A. cathayana*	Luoyang, Henan	44	Pm05 / *A. cathayana*	Chaoyang, Liaoning
21	Pas45/ *A. cathayana*	Luoyang, Henan	45	Ps02 / *A. sibirica*	Huhehaote, Inner Mongolia
22	Pas46/ *A. cathayana*	Luoyang, Henan	46	Ps156 / *A. sibirica*	Huhehaote, Inner Mongolia
23	Pas49/ *A. cathayana*	Luoyang, Henan	47	Ps53 / *A. sibirica*	Huhehaote, Inner Mongolia
24	Pas50/ *A. cathayana*	Luoyang, Henan			

Table 2: ISSR primers and amplification polymorphism of PCR reaction

Primer number	Primer's sequence (5'→3')	Total of amplified band	Polymorphic band	*P* / %
BC823	TCTCTCTCTCTCTCTCC	7	5	71.4
UBC825	ACACACACACACACACT	8	6	75
UBC827	ACACACACACACACACG	9	9	100
UBC834	AGAGAGAGAGAGAGAGYT	7	6	85.7
UBC835	AGAGAGAGAGAGAGAGYC	10	7	70
UBC836	AGAGAGAGAGAGAGAGYA	10	10	100
UBC840	GAGAGAGAGAGAGAGAYT	7	6	85.7
UBC841	GAGAGAGAGAGAGAGAYC	9	9	100
UBC842	GAGAGAGAGAGAGAGAYG	7	5	71.4
UBC846	CACACACACACACACART	5	4	80
UBC847	CACACACACACACACARC	7	7	100
UBC848	CACACACACACACACARG	9	9	100
UBC850	GTGTGTGTGTGTGTGTYC	8	6	75
UBC853	TCTCTCTCTCTCTCTCRT	8	6	75
UBC856	ACACACACACACACACYA	8	7	87.5
UBC860	TGTGTGTGTGTGTGTGRA	7	6	85.7

Table 3: Genetic diveristy index in core germplasms of kernel-apricot by different sampling grade

Sampling grade	Sampling amount	Observed number of alleles	N_e	H_e	I	P_n	P
First sampling	47	1.9365±0.2448	1.4286±0.3156	0.267±0.1537	0.4167±0.2016	118	93.65
Second sampling	36	1.9286±0.2586	1.4314±0.3211	0.2707± 0.1544	0.4191±0.2028	117	92.86
Third sampling	28	1.9127±0.2834	1.4396±0.3224	0.2714 ±0.1557	0.4213±0.2053	115	91.27
Fourth sampling	19	1.9048±0.2947	1.4564±0.3191	0.2804±0.1545	0.4324±0.2050	114	90.48
Fifth sampling	13	1.8413±0.3669	1.4494±0.3289	0.2737±0.1650	0.4185±0.2264	106	84.13

N_e: effective number of alleles; H_e: Nei's gene diversity; I: Shannon information index; P_n: the number of polymorphic loci; P: percentage of polymorphic loci

Construction of Core Germplasms in Kernel-apricot and Data Analysis

The 47 accessions were first subjected to cluster analysis, and LDSS was used to construct the core collection based on the results of the cluster analysis (Fig. 2). The resampling was repeated five times and the level of genetic diversity was obtained each time. As the number of germplasm sampled decreased, the changes in genetic diversity parameters were generally small. The N_e changed little and it was largest after the fourth resampling, H_e and I increased slightly and reached their maximum after the fifth resampling, and P decreased most significantly after the fifth sampling.

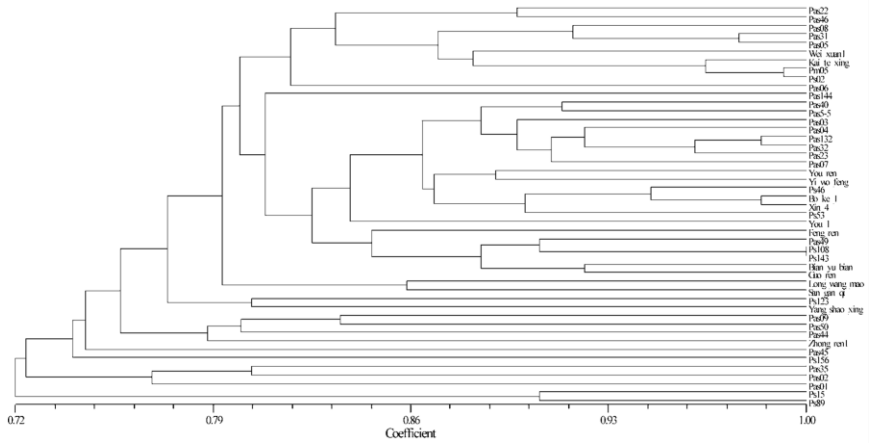

Fig. 1: Dendrogram of 47 kernel-apricot samples based on UPGMA method

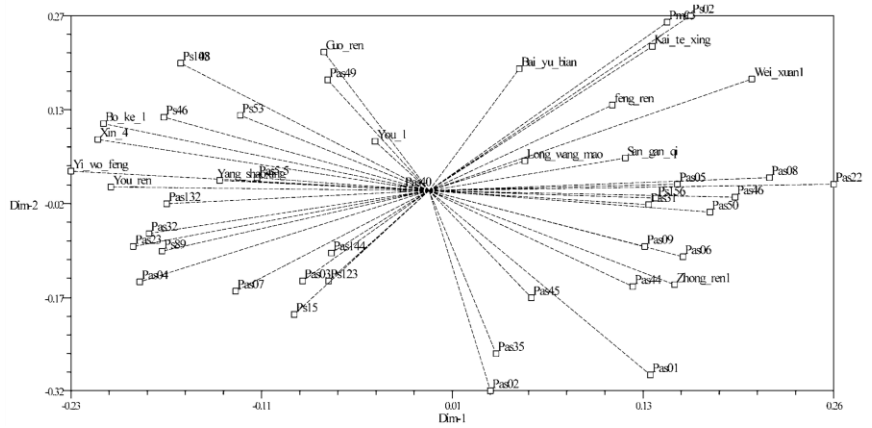

Fig. 2: The analysis of principal component based on ISSR method

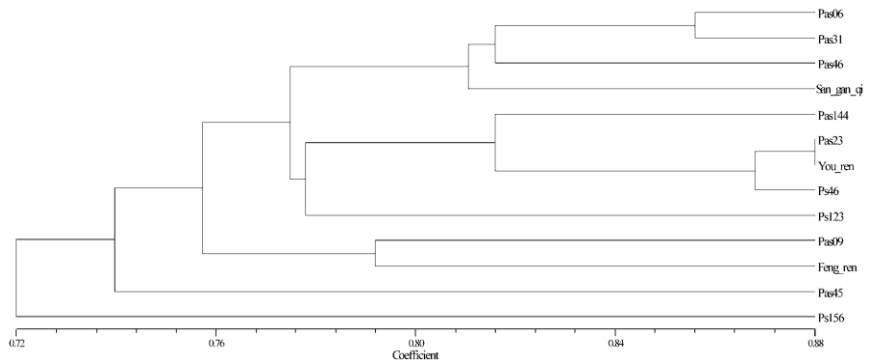

Fig. 3: The cluster analysis at the fifth sampling grade

After the five resampling events, the number of samples in the established germplasm collection was reduced to 27.7% of the original (Fig. 3). Based on these results, the fourth resampling was selected as the primary core collection, the number of samples after the fourth resampling reached 40.4% of the original number and was considered representative of the original germplasm collection.

Discussion

The molecular markers allow rapid evaluation of collected germplasm resources at molecular level. The differences at the molecular level are determined by comparison of the collected germplasm (unknown traits) with that of known varieties when the morphology is similar (Liu *et al*., 2014, 2015b; Zhang *et al*., 2015), and the early evaluation of

beneficial traits (yield, quality, stress tolerance, etc.) was efficient, and the process of aggregation of beneficial traits is short. The molecular marker techniques, unlike phenotypic analysis, were not limited by growing season, they have great potential and have been widely applied because of their efficient identification of differences between DNA sequences of different samples. In this experiment, 16 ISSR primers were used for amplification of 47 germplasm samples of kernel-apricot characterized by high yield, good quality, and resistance to late spring frost. A total of 126 distinct bands were produced, 93.65% of which were polymorphic, indicating a good diversity at molecular level. Ai *et al*. (2014), in their study on genetic diversity of 24 kernel-apricot varieties using SRAP markers, produced 280 bands using 15 pairs of primers, of which 241 (85.34%) were polymorphic. This shows that, despite the different methods employed, a high level of polymorphism is displayed in the DNA of various kernel-apricot germplasm.

The relatively high level of genetic diversity among individual kernel-apricot plants promotes good ecological adaptability of the species. He *et al*. (2007), using SSR markers, reported that high level of diversity in wild *Prunus armeniaca* population has been present for a long time; the genetic variation among individuals within populations reaches 83.6–86.3%. These estimates were confirmed by Li *et al*. (2016) using SRAP markers. In the present experiment, the N_e (1.4286), H_e, (0.267) and I (0.4167) values of the 47 germplasm samples indicated their rich genetic diversity. Moreover, the genetic differences among individuals were also significant. The greatest genetic distance (0.9045) was observed between Ps89 from Xinxiang, Henan, and Pas35 from Huailai, Hebei, although the species in Henan was introduced from Zhuolu and Zhangjiakou in Hebei in 1970s (Liu *et al*., 2015b), and therefore, the two accessions were expected to be genetically closely related. Moreover, Liu *et al*. (2014) reported variations in multiple phenotypic traits, such as leaf and flower color and kernel size, during their sampling and analysis of kernel-apricot plants in Zhuolu, indicating a hybridization event between populations in that region. The occurrence of the greatest genetic distance between Ps89 and Pas35 from Huailai in the present study was in accord with their morphological classification. These results indicate that the rich genetic diversity among individual kernel-apricot cultivars is the basis of their high ecological adaptability, while insect pollination may have accelerated the mixing of individual gene pools (He *et al*., 2007).

Principal component analysis and cluster analysis revealed that geographical origin has little influence on the genetic relationship within the kernel-apricot germplasm. The genetic similarity coefficients between 'Katy' apricot from Yuanyang, Henan, and Pm05 from Chaoyang, Liaoning, and between 'Katy' and Ps02 from Hohhot, Inner Mongolia, were 0.9603 and 0.9682 (Fig. 1), respectively, resolving them as members of the same cluster at molecular level, despite their apparently different geographical origins and ecological environments. This shows that germplasm resources from the different regions could be genetically related at the molecular level. In contrast, germplasm from the same region was only remotely related genetically; for example, genetic similarity coefficient between nine samples from Luoyang, Henan, that belong to Pas132 and Pas50 was 0.7222, which was obviously smaller than those of other germplasms in this region (Fig. 1). This may be due to the past exchange of kernel-apricot varieties between different regions, which may have resulted in hybridization of cultivars and name confusion, and therefore warrants further study.

The concept of core collection was first proposed by Frankel in 1984 and the research about this concept has made a remarkable progress in the last 30 years. Many researchers are using molecular markers to evaluate the representativeness and genetic diversity of core germplasm, which has become an effective and widely applied method; for example, they were used in studies on teosinte (SSR markers) (Fukunaua *et al*., 2005), cherimoya (SSR markers) (Escribano *et al*., 2008), wild apricot (SRAP markers) (Li *et al*., 2016), etc. Xu *et al*. (2004) showed that simple random sampling and stratified cluster sampling were two important sampling strategies for the construction of core collection; most researchers adopt the latter or combine the two methods to ensure better representativeness of the germplasm. Brown (2011) pointed out that the sampling fraction for the core germplasm of most plants is 5−10%, and Li *et al.* (1999) suggested that the appropriate core collection size for each species should not be determined in a simplistic way, but rather tailored to the specific species according to its genetic diversity. In the present study, we performed the least distance stepwise sampling of 47 kernel-apricot germplasm samples to construct a core collection, with genetic diversity analysis carried out after each sampling. As the number of selected samples reduced, small changes in genetic parameters and changes in some peak values were observed. However, the percentage of polymorphic loci decreased. Based on the above results, four rounds of sampling were performed to select 40% of the original collection as the primary core germplasm. The N_e and I after the four rounds of sampling were 0.2804 and 0.4324, respectively, whereas the P and the number of alleles were above 90.48% and same as before sampling. This is in line with the requirement of 70–80% genetic diversity that core germplasm must satisfy to be representative of the original germplasm (Gu *et al*., 2005) and shows that the germplasm obtained after four consecutive rounds of sampling, which constituted 40% of the initial collection, was highly representative of the genetic diversity of the original population. This was different from the *A. vulgaris* core germplasm constructed by Liu *et al.* (2015a) using stepwise clustering with allele preferred sampling and genetic distance (Nei and Li, 1979), in which the N_e, I, percentage of polymorphic loci, and

allele number were 109.24, 108.31, 98.83 and 99.42%, respectively and 22.96% of the original germplasm was retained. These results indicate the difference between the results and strategies of different methods used for the construction of core germplasm. However, appropriate grouping and sampling methods, as well as sampling fraction are the key factors in the construction of core germplasm (Li *et al.*, 2016) and they should be further researched.

Conclusion

The high genetic diversity is the basis for better adaptability and survivability of species and the driving force for their evolution, breeding, and genetic optimization. The establishment of core germplasm and the collection and preservation of germplasm resources with high genetic diversity help to utilize land resources more effectively and serve as seed bank in future breeding. The 16 ISSR polymorphic primers screened and selected in this study, combined with four consecutive stepwise sampling events at minimum distance based on cluster analysis, screened 19 samples which presented 40.4% of the whole core germplasm, The high N_e, H_e, and I indexes of the core germplasms showed that this strategy is applicable at the molecular level, can effectively establish the core germplasm of kernel-apricot, and be applied to practice.

Acknowledgements

This study was funded by the "Foundation of Henan Educational Committee (15A210016)", "Anyang Institute of Technology (YJJ2015015)" and "National Forest Genetic Resource Platform (2005DKA21003)".

References

Ai, P.F., S. Su and Z.Z. Jin, 2014. Analysis of genetic diversity and development of fingerprinting key among accessions of kernel-using apricot using SRAP markers. *Acta Hortic. Sin.*, 41: 1191−1197

Ai, P.F., Z.J. Zhen and Z.Z. Jin, 2011. Genetic diversity and relationships within sweet kernel apricot and related *Armeniaca* species based on sequence-related amplified polymorphism markers. *Biochem. Syst. Ecol.*, 39: 694–699

Bao, W.Q., T.N. Wuyun, L. Wang, H. Zhao and H.Y. Du, 2016. Genetic diversity and population genetic structure of Siberian apricot in Inner Mongolia. *Acta Bot. Boreal.*, 36: 2182−2191

Bourguiba, H., L. Krichen, J.M. Audergon, K. Bouchaib and T.F. Neila, 2010. Impact of mapped SSR markers on the genetic diversity of apricot (*Prunus armeniaca* L.) in Tunisia. *Plant Mol. Biol. Rep*., 28: 578–587

Brown, A.H.D., 2011. Core collections: A practical approach to genetic resources management. *Genome*, 31: 818–824

Cui, H., X. Liu, S. Zhao, M. Yang and H. Liang, 2016. Genetic diversity of SSR analysis of *Prunus armeniaca* cultivars based on maturity. *North. Hortic.*, 12: 106–111

Escribano, P., M.A. Virule and J.I. Hormaza, 2008. Comparison of different methods to construct a core germplasm collection in woody pernnial species with simple sequence repeat markers. A case study in cherimoya (*Annoacherimola,* Annonaceae), an underutilized subtropical fruit tree species. *Ann. Appl. Biol.*, 153: 26–32

Fu, D., L. Ma, Y. Qin, M.P. Liu, H. Zhao, G.P. Zhu and H. Fu, 2016. Phylogenetic relationships among five species of *Armeniaca* Scop. (*Rosaceae*) using microsatellites (SSRs) and capillary electrophoresis. *J. For. Res.*, 27: 1–7

Fukunaua, K., J. Hill, Y. Viuouroux, Y. Matsuoka, G.J. Sanchez, K. Liu, E.S. Buckler and J. Doebley, 2005. Genetic diversity and population structure of teosinte. *Genetics*, 196: 2241–2254

Gu, X.F., A.F. Yang, H. Meng and J.R. Zhang, 2005. *In vitro* induction of tetraploid plans from diploid *Zizyphus jujube* Mill. cv. Zhanhua. *Plant Cell Rep.*, 24: 671–676

He, T.M., X.S. Chen, Z. Xu, J.S. Gao, P.J. Lin, W. Liu, Q. Liang and Y. Wu, 2007. Using SSR markers to determine the population genetic structure of wild apricot in West China. *Genet. Resour. Crop Evol.*, 54: 563

Khadivi-Khub, A., M. Yarahmadi, A. Jannatizadeh and A. Ebrahimi, 2015. Genetic relationships and diversity of common apricot (*Prunus armeniaca* L.) based on simple sequence repeat (SSR) markers. *Biochem. Syst. Ecol.*, 61: 366–371

Li, M., X. Hu, X.J. Miao, Z. Xu and Z. Zhao, 2016. Genetic Diversity analysis of wild apricot（*Prunus armeniaca*） populations in the Ili valley as revealed by SRAP markers. *Acta Hortic. Sin.*, 43: 1980–1988

Li, Z., H.L. Zhang, C. Sun and X. Wang, 1999. Status and prospects of core collection in plant germplasm resource. *J. Chin. Agric. Univ.*, 4: 51–62

Liu, J., K. Liao, S.R. Zhao, Q. Cao, Q. Sun and H. Liu, 2015a. The core collection construction of Xinjiang wild apricot based on ISSR molecular markers. *Sci. Agric. Sin.*, 48: 2017–2028

Liu, M.P., H.Y. Du, G.P. Zhu, H. Zhao and T.N. Wuyun, 2015b. Variety identification and genetic relationship analysis of *Armenia cacathayana* in China based on SSR and ISSR markers. *Gene Mol. Res.*, 14: 9722–9729

Liu, M.P., H.Y. Du, G.P. Zhu, H. Zhao and T.N. Wuyun, 2014. Selection of main evaluation indexes of sweet almond-apricot fruit morphology. *J. Cent. South. Univ. For. Technol.*, 34: 38–42

Mir, J.I., S.H. Wani, N. Ahmed, S.H. Wani, M.A. Sheikh, H. Mir, I. Parveen and S. Shah, 2012. Genetic diversity analysis in apricot (*Prunus armeniaca* L.) germplasms using RAPD markers. *Ind. J. Biotechnol.*, 11: 187–190

Nei, M. and W.H. Li, 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA*, 76: 5269–5273

Rai, I., R.K. Bachheti and C.K. Saini, 2016. A review on phytochemical, biological screening and importance of wild apricot (*Prunus armeniaca* L.). *Orient. Pharm. Exp. Med.*, 16: 1–15

Rohlf, J.F., 2004. *NTSYS-pc: 2.11 Numerical Taxonomy and Multivariate Analysis System*. Exeter Software, Setauket, New York, USA

Wang, L.B., 2013. Properties of Manchurian apricot (*Prunus mandshurica* Skv.) and Siberian apricot (*Prunus sibirica* L.) seed kernel oils and evaluation as biodiesel feedstocks. *Ind. Crops Prod.*, 50: 838–843

Xu, H.M., Y.X. Qiu, J. Hu and J.C. Wang, 2004. Methods of constructing core collection of crop germplasm by comparing different genetic distances, cluster methods and sampling strategies. *Acta Agron. Sin.*, 311: 932–936

Xu, M.S., H. Zhao, X.X. Zhou, T.N. Wuyun, F.D. Li and G.P. Zhu, 2016. Responses of photosynthetic physiology and biomass accumulation of sweet kernel apricot (*Prunus armeniaca* × *sibirica*) seedling to soil drought stress in the ancient course of the middle Yellow River. *Taiwan J. For. Sci.*, 31: 271–284

Yeh, F., R. Yang and T. Boyle, 1999. POPGENE, Version 1.32. Mircosoft window based freeware for population genetic analysis. Edmonton: Centre for international Forestry Research, University of Alberta

Zhang, Q.P., W.S. Liu, N. Liu, N. Zhang, Y.P. Zhang, M. Xu and S. Liu, 2015. Classification of kernel-using germplasm based on morphology in apricot. *J. Fruit Sci.*, 32: 385–392

Zhu, S., F. Wang, W. Shen, D. Jiang, Q. Hong and X. Zhao, 2015. Genetic diversity of *Poncirus* and phylogenetic relationships with its relatives revealed by SSR and SNP/InDel markers. *Acta Physiol. Plant.*, 37: 1–11

(Received 24 March 2017; Accepted 26 December 2017)