New insights into the genetic diversity and species identification of the native apricots in Southern Xinjiang of China

Ling Guo [1,2], Hui Li [1], Zheng-rong Luoa[*]

[1]Key Laboratory of Horticultural Plant Biology, Huazhong Agricultural University, Wuhan 430070, China

[2]College of Plant Science, Tarim University, Alar 843300, China

Corresponding Author: Zheng-rong Luo

E-mail: luozhr@mail.hzau.edu.cn

Genet. Mol. Res. 17 (1): gmr16039874

Received December 24, 2017

Accepted January 18, 2018

Published February 10, 2018

DOI http://dx.doi.org/10.4238/gmr16039874

ABSTRACT. Apricot is a staple stone fruit crop cultivated in Southern Xinjiang of China. This crop is important for the rural communities, as they generate significant employment and income. Here, seventy-eight apricot genotypes, including seventy-six common apricots (Prunus armeniaca L.) and two purple apricots (Prunus dasycarpa Ehrh.), were mainly collected from Aksu, Kashgar, Hetian and Bayingolin. Start Codon Targeted (SCoT) markers and ITS (internal transcribed spacers) sequences were used to investigate the genetic diversity and species identification respectively. Based on POPGENE showed that apricot cultivars from Aksu group exhibited the highest genetically diverse as compared with other groups, cluster analysis of SCoT markers (basede on UPGMA and PCoA method) showed that these apricot cultivars could be divided into three major clusters, which was in agreement with their geographic distribution and pedigrees. It was indicated that there were possible three primary diversity centers in the area: Aksu, Kashgar and Hetian, and also possible introgression among these populations. Furthermore, based on the complete ITS sequences, the phylogenetic analysis showed that P. dasycarpa clustered separated from other the section armeniaca of species. Therefore, it was proposed that P. dasycarpa would be a hybrid species. Our results indicated that SCoT markers are informative and could be used to evaluate genetic diversity of apricots, and ITS could be used effectively to identify P. dasycarpa. These results will provide much more useful information for the native apricots protection and utilization strategies.

KEY WORDS: Prunus armeniaca L., Prunus dasycarpa Ehrh., SCoT markers, ITS, Genetic diversity

INTRODUCTION

Apricot (*Prunus armeniaca* L.; 2n=2x=16) is an important stone fruits, plays a crucial role in the economic and environmental context in the rural communities of Southern Xinjiang in China. It belongs to the section Armeniaca (Lam.) Koch in subgenus Prunophora Focke, genus *Prunus* L., family Rosaceae (Rehder, 1949). According to the morphological classification system, the number of apricot species ranges from 3 to 12. Usually, six distinct species are recognized: common apricot (*P. armeniaca* L.), Briancon apricot (*P. brigantina* Vill.), Manchurian apricot (*P. mandschurica* (Maxim.) Koehne), Japanese apricot (*P. mume* (Sieb.) Sieb. & Zucc.), Siberian apricot (*P. sibirica* L.), and purple or black apricot (*P. dasycarpa* Ehrh.). Besides, cultivated apricot commonly referred to the common apricot (*P. armeniaca* L.) can also be classified into six main eco-geographical groups: the Central Asia group, the Irano-Caucasia group, the European group, the Dzhungar-Zailij group, the North China group and the East China group (Layne et al., 1996).

In China, South of Xinjiang, namely the south area of Tianshan Mountains, including Tarim Basin, Kunlun Mountain Range, as well as Turpan Depression and so on. It mainly includes four regions Aksu, Kashgar, Hetian and Bayingolin in Fig. 1. Owing to the practice of seed propagation and self-incompatibility, South of Xinjiang exhibits abundant apricot resources and highly diversified. It is one of the oldest in the Central Asia group due to its unique location on the historic Land Silk Road (Zhebentyayeva et al., 2012; Zaurov et al., 2013). The most of them are named after a person or a place or according to the fruit characteristics. *P. armeniaca* and *P. dasycarpa* are cultivated widely in South of Xinjiang, especially the *P. armeniaca*, some with the sweet kernels, that vary tremendously in size, shape, color, flavor, glabrous skin and so on. It is difficult to identify species due to the large phenotypic variability and the lack of diagnostic characters. Besides that, some landraces and cultivars are faced to be disappearing due to being replaced by more profitable apricot which meet the market demands. Therefore, the study of the genetic diversity and species identification is very necessary for the protection and utilization of the ancient apricot resources.

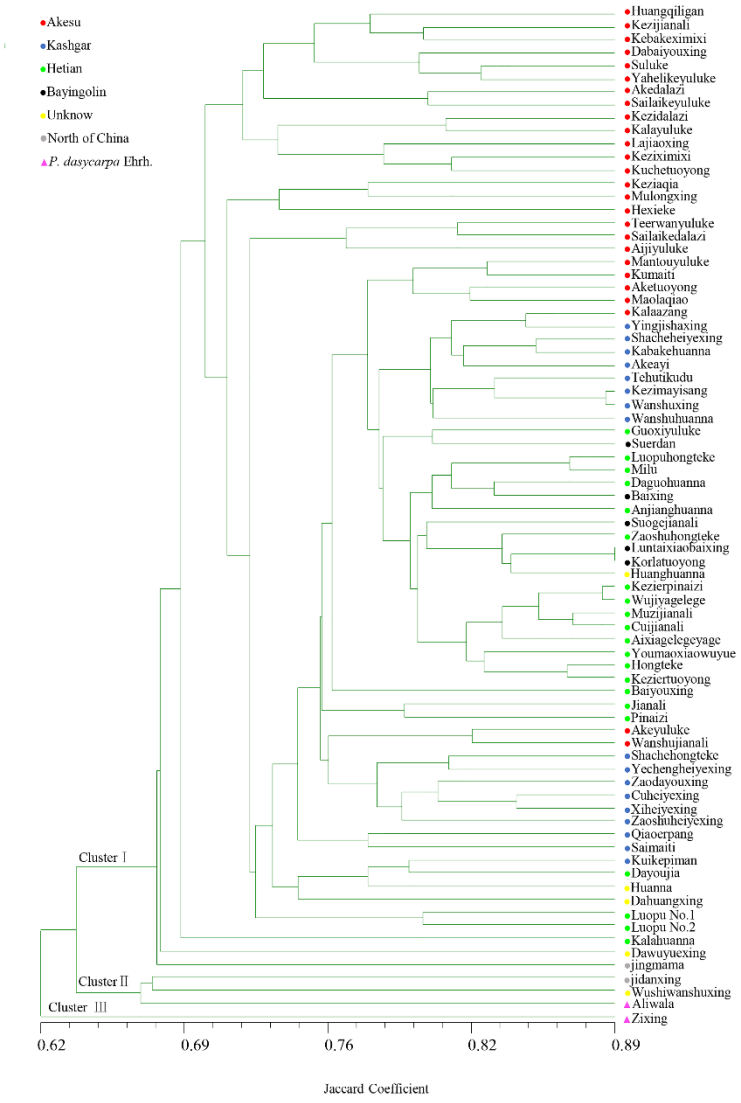

Fig. 1. Geographical distribution of the local apricots (*Prunus armeniaca* L.) and (*Prunus dasycarpa* Ehrh.) in Southern Xinjiang (Aksu, Kashgar, Hetian and Bayingolin) of China.

Understanding the genetic background and phylogenetic relationship of the native apricots is important to conservation and utilization strategies. Many molecular markers have been used to evaluate phylogenetic relationship and genetic diversity, such as SSR (Hormaza, 2002; Khadivi-Khub et al., 2015), ISSR (Li et al., 2013; Li et al., 2014), AFLP (Hagen et al., 2002), SRAP (Li et al., 2014). A part of the apricot resources in Xinjiang has been analyzed by using SSR (He et al., 2007), AFLP (Yuan et al., 2007) and ISSR (Li et al., 2013; Liu et al., 2016). In recent years, the start codon targeted (SCoT) polymorphism technique appears to be widely used for it provides more genetic information (Collard & Mackill., 2009). SCoT markers have been successfully applied on *Dimocarpus longan* (Chen et al., 2010), *Vitis vinifera* (Guo et al., 2012), and *Myrica rubra* (Chen & Liu., 2014). Nevertheless, it has not been used to analyze the genetic diversities of apricots. Besides, many studies have shown that the sequences of Internal transcribed spacer (ITS) region appear to be effective in the identification of classification *Prunus* species (Lee et al., 2001; Shi et al., 2013; Zhao et al., 2016). Thus, in this study, ITS sequence is used to identify *P. armeniaca* and *P. dasycarpa*.

The aim of the present study was to assess the genetic diversity and relationships of 74 accessions of *P. armeniaca* and 2 accessions of *P. dasycarpa* from Southern Xinjiang China by SCoT markers. Meanwhile, ITS sequence as molecular markers to classify the local *P. armeniaca* and *P. dasycarpa*.

Table 1. List of 78 apricot genotypes and their sources in this study

1	Huangqiligan	Aksu, Xinjiang	40	Wanshuhuanna	Kashgar, Xinjiang
2	Kalaazang	Aksu, Xinjiang	41	Yingjishaxing	Kashgar, Xinjiang
3	Dabaiyouxing	Aksu, Xinjiang	42	Kuikepiman	Kashgar, Xinjiang
4	Akedalazi	Aksu, Xinjiang	43	Saimaiti	Kashgar, Xinjiang
5	Kezijianali	Aksu, Xinjiang	44	Luopu No.1	Hetian, Xinjiang
6	Kebakeximixi	Aksu, Xinjiang	45	Luopu No.2	Hetian, Xinjiang
7	Kezidalazi	Aksu, Xinjiang	46	Luopuhongteke	Hetian, Xinjiang
8	Keziaqia	Aksu, Xinjiang	47	Milu	Hetian, Xinjiang
9	Suluke	Aksu, Xinjiang	48	Kalahuanna	Hetian, Xinjiang
10	Yahelikeyuluke	Aksu, Xinjiang	49	Anjianghuanna	Hetian, Xinjiang
11	Kalayuluke	Aksu, Xinjiang	50	Daguohuanna	Hetian, Xinjiang
12	Sailaikeyuluke	Aksu, Xinjiang	51	Muzijianali	Hetian, Xinjiang
13	Mantouyuluke	Aksu, Xinjiang	52	Jianali	Hetian, Xinjiang
14	Lajiaoxing	Aksu, Xinjiang	53	Pinaizi	Hetian, Xinjiang
15	Keziximixi	Aksu, Xinjiang	54	Kezierpinaizi	Hetian, Xinjiang
16	Kuchetuoyong	Aksu, Xinjiang	55	Zaoshuhongteke	Hetian, Xinjiang
17	Aketuoyong	Aksu, Xinjiang	56	Guoxiyuluke	Hetian, Xinjiang
18	Hexieke	Aksu, Xinjiang	57	Dayoujia	Hetian, Xinjiang
19	Kumaiti	Aksu, Xinjiang	58	Baiyouxing	Hetian, Xinjiang
20	Teerwanyuluke	Aksu, Xinjiang	59	Hongteke	Hetian, Xinjiang
21	Sailaikedalazi	Aksu, Xinjiang	60	Youmaoxiaowuyue	Hetian, Xinjiang
22	Aijiyuluke	Aksu, Xinjiang	61	Keziertuoyong	Hetian, Xinjiang
23	Mulongxing	Aksu, Xinjiang	62	Wujiyagekeli	Hetian, Xinjiang
24	Maolaqiao	Aksu, Xinjiang	63	Aixiagelegeyage	Hetian, Xinjiang
25	Akeyuluke	Aksu, Xinjiang	64	Cuijianali	Hetian, Xinjiang
26	Wanshujianali	Aksu, Xinjiang	65	Korlatuoyong	Bayingolin, Xinjiang
27	Shacheheiyexing	Kashgar, Xinjiang	66	Suerdan	Bayingolin, Xinjiang
28	Shachehongteke	Kashgar, Xinjiang	67	Suogejianali	Bayingolin, Xinjiang
29	Kabakehuanna	Kashgar, Xinjiang	68	Baixing	Bayingolin, Xinjiang
30	Akeayi	Kashgar, Xinjiang	69	Luntaixiaobaixing	Bayingolin, Xinjiang
31	Zaodayouxing	Kashgar, Xinjiang	70	Dawuyuexing	Unknow
32	Yechengheiyexing	Kashgar, Xinjiang	71	Dahuangxing	Unknow
33	Zaoshuheiyexing	Kashgar, Xinjiang	72	Huanna	Unknow
34	Cuheiyexing	Kashgar, Xinjiang	73	Huanghuanna	Unknow
35	Xiheiyexing	Kashgar, Xinjiang	74	Wushiwanshuxing	Unknow
36	Qiaoerpang	Kashgar, Xinjiang	75	Jingmama	North China（GanSu）
37	Tehutikudu	Kashgar, Xinjiang	76	Jidanxing	North China（NingXia）
38	Kezimayisang	Kashgar, Xinjiang	77	Aliwala*	Aksu , Xinjiang
39	Wanshuxing	Kashgar, Xinjiang	78	Zixing*	Aksu , Xinjiang

*" Noted that two materials belong to *P. dasycarpa* Ehrh.

MATERIALS AND METHODS

Plant materials and DNA extraction

A total of 78 genotypes were used for the diversity analysis with SCoT markers, and 10 of the genotypes were selected for the phylogenetic relationship with ITS sequences. Among them, 76 apricots are originated from Southern Xinjiang (Fig. 1), the other two cultivars come from the North China group. All of them from the National Field GenBank for Particular Fruit Tree of Xinjiang (Luntai County, Bayingolin). Details of sample code and further information were summarized in Table 1. Young leaves were dried by silica gel. Total genomic DNA was extracted using CTAB method with minor modifications (Doyle, 1990). DNA were checked by agarose gel and ultra-micro UV spectrophotometer (Syngene, USA).

PCR amplification of SCoT markers

Polymerase chain reaction (PCR) amplification of SCoT molecular markers was carried out according to the previous study (Guo et al., 2012). PCR cycling conditions were: pre-denaturation at 95 °C for 3 min, followed with 30 cycles of denaturation at 95 °C for 15 s, annealing at 55 °C for 60 s, and extension at 68 °C for 60 s, and a final extension at 72 °C for 5 min. The amplified products were separated by 1.5% agarose electrophoresis in 1×TBE buffer, photographed with a Gel Documentation System (Syngene, USA).

Amplification and sequencing of ITS

The ITS region (ITS1, 5.8S nuclear ribosomal RNA gene and ITS2) was amplified with the primer pair ITS5 (5'-GGA AGT AAA AGT CGT AAC AAG G-3') and ITS4 (5'- TCC TCC GCT TAT TGA TAT GC-3') (White et

al., 1990). Polymerase chain reaction (PCR) amplification of ITS region was carried out according to the previous study (Tang et al., 2015). The ITS region was determined by directly sequencing the amplified products of the common apricot, while it was identified by cloning strategies for the PCR products of two purple apricots. For the cloning, the purified amplicons were ligated into the pMD19-T vector (TaKaRa, Japan) and transformed into *Escherichia coli* DH5α competent cells. At least 5 positive clones from each amplicon were sequenced with the universal primers. All the sequences have been deposited to GenBank (https://www.ncbi.nlm.nih.gov/genbank) with the accession numbers KX890449 to KX890460 (Supplementary Table 1).

Data analysis of SCoT markers

The strong and well-separated bands were selected for the scoring and further recorded as either 1 (Present) or 0 (Absent). To measure the effectiveness of the SCoT markers, the capacity of the primers in distinguishing genotypes was evaluated by calculating the resolving power (Rp) for the SCoT primer. Four parameters were calculated, including the total number of bands (TNB), the number of polymorphic bands (NPB), the percentage of polymorphic bands (PPB), the polymorphic information content (PIC). The effective multiplex ratio (EMR) was defined as the number of polymorphic markers generated per assay, and the diversity index (DI) was calculated as the average PIC value. Genetic diversity was evaluated by the program POPGENE 1.32 (Yeh et al., 1997). Various other indicators were also calculated, including the percentage of polymorphic bands (PPB), the number of polymorphic loci (Np), the effective number of alleles per locus (Ne), the observed number of alleles per locus (Na), Nei's gene diversity (*H*), and Shannon's information index (*I*) (Nei, 1973). Genetic similarity between accessions was calculated based on the Jaccard coefficient using the SIMQUAL subprogram of the NTSYS-PC Version 2.10e (Exeter Software, Setauket, NY, USA) (Rohlf, 2000). Cluster analysis was performed using the unweighted pair group method with arithmetic mean (UPGMA) in the SAHN subprogram. Principal coordinate analysis (PCoA) was performed based on the Nei's genetic distance using the NTSYS--PC software package.

Sequence alignment and phylogenetic analysis

The complete ITS sequences acquired per accession were aligned and analyzed with MEGA 6.0 (Tamura et al., 2013). The Clustal-W software (Larkin et al., 2007) was applied for multiple sequences alignment. Other related sequences were downloaded from GenBank (Supplementary Table 1).

Phylogenetic analyses were performed by using neighbor-joining (NJ) method (Tamura et al., 2013) with the data matrices of ITS region. The sequence divergence between the taxa was calculated by the Kimura 2-parameter (K2P) model (Kimura, 1980), and the gaps or missing data were treated with the pairwise-deletion option. The NJ trees were obtained by using the p-distance model (Kimura, 1980). All runs were done with 1, 000 bootstrap replicates to test the branch support levels (Felsenstein, 1985). In addition, Japanese plum was placed as an outgroup in the analysis.

RESULTS

Polymorphism of SCoT markers

A total of 36 SCoT primers were screened, of which 24 with high reproducibility and abundant polymorphism were selected for genetic diversity studies in Table 2. Altogether, 228 stable and clear bands were obtained, including 213 polymorphic bands. The percentage of polymorphic bands for each marker varied from 70% (SCoT15) to 100% (SCoT1, 4, 12, 13, 14, 15, 17, 19, 21, 24, 30, 31, 32) with an average of 93.1%. The average PIC value was 0.278 and the average Rp-value was 9.986.

Table 2. Summary of the primer sequences, polymorphism and information index based on SCoT markers data

Primer Code	Sequences (5'-3')	TNB	NPB	PPB (%)	PIC	MI	RP
SCoT 1	CAACAATGGCTACCACCA	10	10	100.0	0.300	3.000	10.769
SCoT 2	CAACAATGGCTACCACCC	8	7	87.5	0.224	1.568	8.513
SCoT 3	CAACAATGGCTACCACCG	9	8	88.9	0.377	3.016	7.385
SCoT 4	CAACAATGGCTACCACCT	10	10	100.0	0.315	3.150	10.077
SCoT 11	AAGCAATGGCTACCACCA	9	7	77.8	0.232	1.624	12.974
SCoT 12	ACGACATGGCGACCAACG	11	11	100.0	0.342	3.762	7.436
SCoT 13	ACGACATGGCGACCATCG	9	9	100.0	0.221	1.989	14.564
SCoT 14	ACGACATGGCGACCACGC	6	6	100.0	0.256	1.536	6.923
SCoT 15	ACGACATGGCGACCGCGA	5	5	100.0	0.236	1.180	5.872
SCoT 16	ACCATGGCTACCACCGAC	10	9	90.0	0.381	3.429	9.974
SCoT 17	ACCATGGCTACCACCGAG	5	5	100.0	0.153	0.765	2.846

SCoT 18	ACCATGGCTACCACCGCC	7	6	85.7	0.241	1.446	7.718
SCoT 19	ACCATGGCTACCACCGGC	12	12	100.0	0.353	4.236	12.923
SCoT 21	ACGACATGGCGACCCACA	12	12	100.0	0.287	3.444	10.564
SCoT 23	CACCATGGCTACCACCAG	9	8	88.9	0.285	2.280	9.436
SCoT 24	CACCATGGCTACCACCAT	9	9	100.0	0.346	3.460	9.333
SCoT 29	CCATGGCTACCACCGGCC	11	10	90.9	0.230	2.300	10.154
SCoT 30	CCATGGCTACCACCGGCG	12	12	100.0	0.253	3.031	11.538
SCoT 31	CCATGGCTACCACCGCCT	15	15	100.0	0.263	3.945	17.103
SCoT 32	CCATGGCTACCACCGCAC	11	11	100.0	0.335	3.685	10.359
SCoT 33	CCATGGCTACCACCGCAG	6	5	83.3	0.273	1.365	9.359
SCoT 34	ACCATGGCTACCACCGCA	12	11	91.7	0.284	3.124	11.205
SCoT 35	CATGGCTACCACCGGCCC	10	8	80.0	0.334	2.652	13.359
SCoT 36	GCAACAATGGCTACCACC	10	7	70.0	0.162	1.134	9.282
Min.		5	5	70.0	0.153	0.765	2.846
Max.		15	15	100.0	0.381	4.236	14.564
Average		9.5	8.9	93.1	0.278	2.547	9.986
Total		228	213	-	-	61.121	239.666

TNB: total number of bands; NPB: number of polymorphic bands; PPB: percentage of polymorphic bands; PIC: polymorphism information content; MI: marker index; Rp: resolving power.

The genetic diversity analysis

The genetic diversity of 68 apricot genotypes was assessed by Shannon's information index (*I*), Nei's gene diversity (*H*) and the percentage of polymorphic bands (PPB). At the species level, *H* was 0.2868, *I* was 0.4380 and PPB was 93.42%. At the group level, the Aksu group showed the highest genetic diversity (*H* = 0.2612, *I* = 0.4003, PPB = 85.09%) among the four areas of Southern Xinjiang (Table 3). The diversities from Hetian and Kashgar groups were lower than that from Aksu group while higher than that from Bayingolin group. Based on these parameters, the Bayingolin group showed the lowest genetically diversity among the four groups. Overall, the data collectively illustrated that the apricots exhibited high genetic diversity in the Aksu group.

Table 3. Genetic diversity indexes among the different apricot groups in this study

Population	Sample size	Np	PPB (%)	Na	Ne	*H*	*I*
Aksu	26	194	85.09	1.8509	1.4310	0.2612	0.4003
Kashgar	17	158	69.30	1.6930	1.3801	0.2254	0.3408
Hetian	20	177	77.63	1.7763	1.4309	0.2547	0.3843
Bayingolin	5	88	38.60	1.3860	1.2521	0.1436	0.2126
Species level	68	213	93.42	1.9342	1.4752	0.2868	0.4380

Np: the number of polymorphicloci bands; PPB: the percentage of polymorphic bands; Na: the observed number of alleles per locus; Ne: the effective number of alleles per locus; *H*: Nei's gene diversity; *I*: Shannon's information index.

Genetic relationship of the native apricots

Based on the SCoT data, genetic similarity values ranged from 0.62 to 0.89 in Fig. 2, the lowest similarity coefficient between them indicated that considerable genetic difference existed among all of *P. armeniaca* and 'zixing'. The highest Jaccard coefficient suggested that the 'Luntaixiaobaixing', and the 'Kuerlatuoyong' were of similar genetic composition. The dendrogram clearly revealed that almost all of the studied accessions were organized into three distincts clusters at a coefficient of > 0.62 in Fig. 2. Cluster Ⅰ was the most complex, with the most of common apricot and it was further divided into two mainly subgroups; one subgroup comprised 13 cultivars from Aksu ('Huangqiligan', 'Kezijianali', 'Kebakeximixi', 'Dabaiyouxing', 'Suluke', 'Yahelikeyuluke', 'Akedalazi', 'Sailaikeyuluke', 'Kezidalazi', 'Kalayuluke', 'Lajiaoxing', 'Keziximixi', 'Kuchetouyong'). The other subgroup was composed by 13 Aksu apricots, 17 Kashgar apricots, 20 Hetian apricots, 5 Bayingolin apricots. Meanwhile, cluster Ⅱ included two apricot cultivars and one purple apricot. Cluster Ⅲ, there was only one purple apricot 'Zixing', suggesting its distinct genetic background with other genotypes.

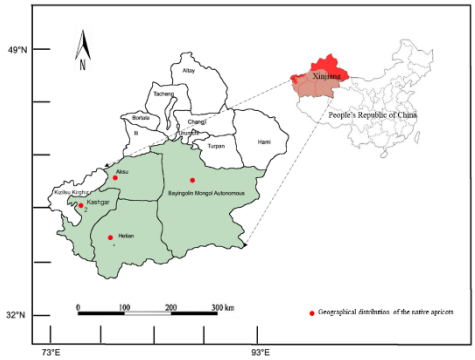

Figure 2. Dendrogram depicting the relationships among 78 apricot samples constructed using UPGMA and based on SCoT markers.

The genetic divergence among the 78 apricot genotypes were further graphically elucidated by theprincipal coordinate analysis (PCoA) scatter plot in Fig. 3. In general, similar results were found as obtained with the UPGMA dendrogram. These results indicate that the possible existing of three centers of diversity in Southern Xinjiang: Clusters Ⅰ (Aksu), Clusters Ⅱ (Kashgar) and Clusters Ⅲ (Hetian). The most of them are consistent with their genetic origin and geographic distribution, as well as the existence of introgression among the populations in Fig. 3. Similarly, the Aksu group displayed the highest genetic variation. The distant 2 members of *P. dasycarpa* and 2 members of the North China group and all members of the rest of the native apricots, and they were distinct from the others.

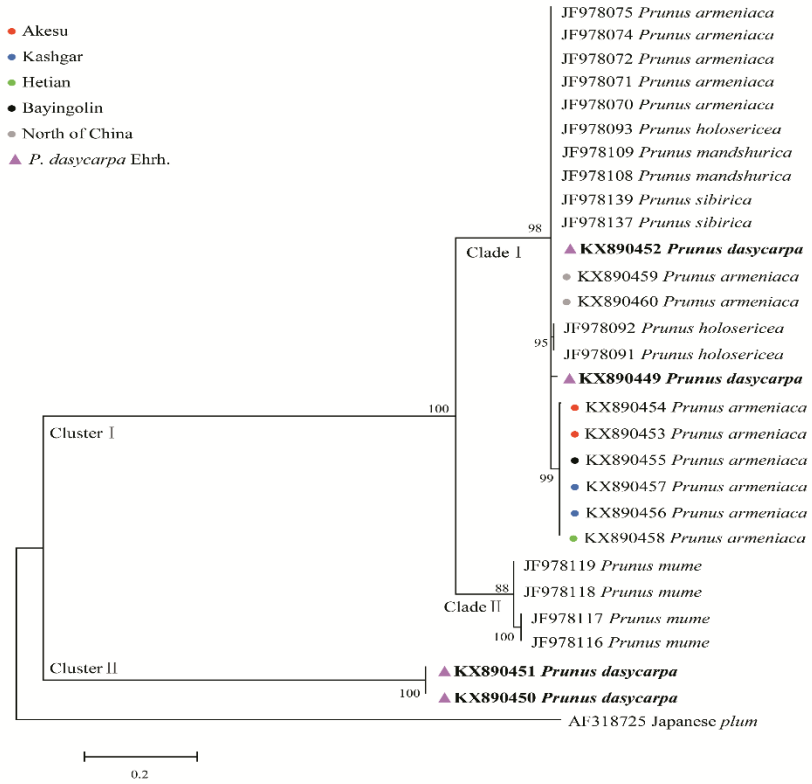

Figure 3. Two-dimensional projection of the PCoA of 78 apricot samples based on SCoT markers along the frist two principal axes.

Sequence and Phylogenic analysis

Sequences of ITS were generated for the 8 investigated accessions of *P. armeniaca* and 2 accessions of *P. dasycarpa* (‘Zixing’ and ‘Aliwala’). The other 17 accessions of five species were obtained from NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank) (Supplementary Table S1). The final alignments consisted of 609 -631 aligned positions of the ITS region, of which 21 were of variable sites, and 15 were parsimony informative. Two sites of the aligned sequences involved gaps. In particular, among the studied two purple apricots, showed a 22-bp indel was also found at 518 in the multiple sequence alignment. Thus, two lengths of sequences (609 bp and 631 bp for both ‘Zixing’ and ‘Aliwala’) were used in the following analysis.

Basing on the NJ trees of complete ITS region in Fig. 4, the studied 27 accessions could be divided into two clusters. Cluster Ⅰ could be splitted into clade Ⅰ and clade Ⅱ. Clade Ⅰ contained all the studied *P. armeniaca*, *P. mandschurica*, *P. holosericea*, *P. sibirica* and *P. dasycarpa* (KX890449 and KX890452). However, Clade Ⅱ comprised all the *P. mume*. Clade Ⅰ and clade Ⅱ were included in a monophyletic group, which meant that they shared the same ancestor. Cluster Ⅱ only contained the ITS sequences of 2 members of *P. dasycarpa* (‘Zixing’KX890450 and ‘Aliwala’KX890451), and they are always clustered separated from other species with 100% bootstrap.

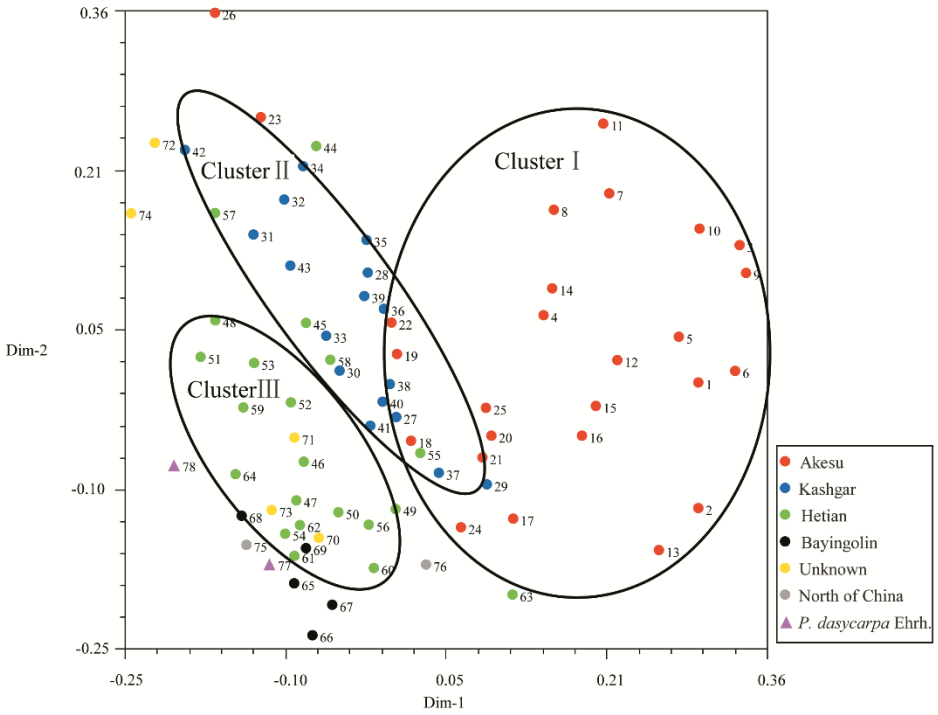

Figure 4. The neighbor-joining tree of the 27 taxa of the ITS region. Numbers above the branches indicate the support level (for branches with ＞50%) determined from 1,000 bootstrap replications.

ITS sequences were more conservative which contained a little informative sites in species level, except for 2 members of *P. dasycarpa* (‘Zixing’ and ‘Aliwala’). According to their origin and ecological distribution, the most of them have similar ITS sequences, such as 6 members of native apricots and 2 accessions of North China apricot and *P. holosericea* and *P. sibirica*, included *P. mume*. In addition, ITS sequences are inherited by parents, and the sequences of parents are often expressed in hybrid progenies (Du et al., 2010; Rauscher et al., 2002). Two different ITS sequences of each purple apricot strongly supported that *P. dasycarpa* is a hybrid species.

DISCUSSION

Polymorphism analysis of SCoT markers and ITS

Since the development of SCoT markers (Collard & Mackill., 2009), It has been successfully employed in genetic studies of other plants and proved to be a very efficient dominant marker (Guo et al., 2012; Chen & Liu, 2014).

In this study, we used SCoT markers to investigate genetic diversities of the native apricots in Southern Xinjiang. Based on SCoT markers, the polymorphism information of native apricot cultivars was 93.42%, which was remarkably higher than AFLP markers (72.70%) in previous study (Yuan et al., 2007). Likewise, SCoT markers exhibited high polymorphism information according to previous reports on other plants (Guo et al., 2012; Chen & Liu, 2014). Also, this is the first study of the apricot genetic diversity in Southern Xinjiang by SCoT markers. SO far, and the results showed that SCoT markers are much more effective for illustrating genetic relationships among apricot cultivars.

In our study, the polymorphism of ITS sequences within these studied species could be successfully detected. Meanwhile, the variation of the ITS data can be used effectively to identify *P. dasycarpa.* As shown the phylogenetic tree in Fig. 4, Clade Ⅰ and clade Ⅱ were sister clades, which meant that these species share the same ancestor. Shi et al. (2013) and Lee et al. (2001) reported the similar phenomenon about *Prunus sensu*. Based on ITS sequences, It was difficult to precisely discriminate all the studied taxa based on the only 21 informative sites in ITS region. Further revealed that ITS sequences were more conservative and seemed to evolve much slower in species level. Many apricot cultivars (included the native apricots and the North China group) possessed remarkably different the phenotypic traits, such as the characteristics of leaf, flower and fruit. The diversity of apricot cultivars is a result for the adaptation of the variational environment in the evolutionary process. Our conclusion agreed with Decroocq et al. (2016), that cultivar apricots likely underwent two independent domestication events, with bottlenecks, from the same wild population.

Genetic diversity and relationship of the native apricot

Southern Xinjiang, due to its special location, one of primary genetic center of apricots, contains the oldest and most highly diversified apricot resources (Vavilov, 1951; Zaurov et al., 2013). Previous studies of apricot in Xinjiang (or Xinjiang Province included) supported that there was a high level of genetic diversity of apricot (Yuan et al., 2007; Zhang et al., 2014; Li et al., 2014). This study demonstrated different parameters to evaluate the genetic diversity of apricots in Southern Xinjiang. Our work, cultivars ‘Cuheiyexing’, ‘Xiheiyexing’ and ‘Zaoshuheiyexing’ are closely related, which were similarly proved by pervious study using ISSR and SRAP markers (Li et al., 2014). Additionally, ‘Luopu No.1’ and ‘Luopu No.2’ were more clearly distinguished from each other compared to previous work (Yuan et al. 2007). Interestingly, ‘Luntaixiaobaixing’, as a well-known cultivar extensively cultivated in Bayingolin, was unable to be differentiated from and ‘Kuerlatuoyong’, we suggesting that they are synonymous cultivars.

Amongst Southern Xinjiang area, the genetic diversity parameters of Aksu group was the highest, which were consistent with previous study of genetic diversity in Kuche (a county of Aksu) population reaching to the highest genetic diversity among Kashgar and Hetian populations (Yuan et al., 2007), suggesting that Aksu might be an ancestral region. Considering the peculiar natural environment with Tian-Shan Mountains and the Kunlun Mountains surrounded, there have been natural barriers to prevent the spread of apricot germplasm due to surrounding Takla Makan Desert. The practice of seed propagation and self-incompatibility were lead to its differentiation and genetic diversity. Meanwhile, Bayingolin group exhibited the lowest genetic diversity parameters (H = 0.1436, I = 0.2126, PPB = 38.60%). The low genetic diversity parameters may affected by endangered apricot cultivars (smaller sample size). A larger sample size will draw more accurate conclusions. However, the results may provide a reference to further study the genetic diversity of cultivated apricot and breeding.

The identification of apricots species

Many studies on the phylogeny of the section Armeniaca were reported and different conclusions were suggested. So far, this situation reflected the complexity of the section Armeniaca. DNA sequences could provide deep insight into evaluation of genetic polymorphism. Obviously, ITS sequence was one of the most important molecular markers to reveal the phylogenetic of the section Armeniaca. In previous study, *P. dasycarpa* was suggested that it belonged to a hybrid species between common apricot (*P. armeniaca*) and cherry plum (*P. cerasifera*) (Zhebentyayeva et al., 2012). Simultaneously, Hagen used AFLP to study the genetic diversity of apricot species and pointed that *P. dasycarpa* was intermediate between *P. brigantiaca* and *P. mume* (Hagen et al., 2002). Furthermore, such as SSR (Zhang et al., 2014), ISSR and SRAP marker (Li et al., 2014) reported *P. dasycarpa* is the most distant from all the other apricot accessions. In our study, the results showed strongly supports that *P. dasycarpa* was a hybrid. It was proved by ‘Aliwala’ and ‘Zixing’ displayed two of different ITS sequence and the phylogenic relationships (Fig. 4). The individuals of *P. dasycarpa* clustered together and clearly distincted from other related species by both SCoT marker and ITS sequences (in Fig. 2 and Fig. 4). The high level genetic differentiation detected among apricot accessions could be attributed to different geographical regional adaptation of accessions, reproductive mode, hybrid introgressions, or evolutionary history. Obviously,

if more genetypes were included in the analysis, and the more genome sequences or the whole genome information of the section Armeniaca were analyzed, the precise contour of phylogenic would be clear.

In conclusion, SCoT markers successfully evaluated genetic relationships and provided detailed information on genetic diversity of apricot cultivars in Southern Xinjiang China. The most of apricots genetic variability occurred intra-group in Southern Xinjiang and Aksu group displayed relatively primitive diversity center, which exhibited highest genetic diversity compared with other groups. Meanwhile, ITS sequences results was effectively to identify *P. dasycarpa* and suggested it is a hybrid-origin species. Our work provide some clues in the study of controversial species and will be helpful in the germplasm conservation, breeding, and genetic diversity study of apricots in the future.

AUTHOR CONTRIBUTION STATEMENT

Ling Guo conducted experiments, analyzed data and wrote the manuscript and Hui Li conducted experiments. Zheng-rong Luo designed the experiments and revised the manuscript. Two authors have read and approved the manuscript.

ACKNOWLEDGMENT

This research was financially supported by the Natural Science Foundation of China (31760560) and the President of Tarim University of Key Fund Project (TDZKGG201706). At last, thanks for all the materials provided by Mr. Tang Zhang-hu from National Field GenBank.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

Chen FY, Liu JH (2014). Germplasm genetic diversity of *Myrica rubra* in Zhejiang Province studied using inter-primer binding site and start codon-targeted polymorphism markers. Sci Hortic. 170:169-175. https://doi.org/10.1016/j.scienta.2014.03.010

Chen H, He XH, Luo C, Zhu JH, et al. (2010). Analysis on the genetic diversity of 24 longan (*Dimocarpus longan*) accessions by SCoT markers. Acta Horticulturae Sinica. 37(10): 1651-1654.

Collard BCY, Mackill DJ (2009). Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Mol Biol. 27:86-93. https://doi.org/10.1007/s11105-008-0060-5

Decroocq S, Cornille A, Tricon D, Babayeva S, et al. (2016). New insights into the history of domesticated and wild apricots and its contribution to Plum pox virus resistance. Molecular Ecology, 25(19): 4712-4729.

Doyle JJ (1990). Isolation of plant DNA from fresh tissue. Focus. 12: 13-15.

Du ZY, Yang CF, Chen JM, Guo YH (2010). Identification of hybrids in broad-leaved *Potamogeton* species (Potamogetonaceae) in China using nuclear and chloroplast DNA sequence data. Plant Syst Evol. 287(1-2): 57-63. https://doi.org/10.1007/s00606-010-0289-y

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 14(8): 2611-2620. https://doi.org/10.1111/j.1365-294x.2005.02553.x

Felsenstein J (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution. 783-791. https://doi.org/10.2307/2408678

Guo DL, Zhang JY, Liu CH (2012). Genetic diversity in some grape varieties revealed by SCoT analyses. Mol Biol Rep. 39(5): 5307-5313. https://doi.org/10.1007/s11033-011-1329-6

Hagen LS, Khadari B, Lambert P, Audergen LM (2002). Genetic diversity in apricot revealed by AFLP markers: species and cultivar comparisons. Theor Appl Genet. 105(2-3): 298-305. https://doi.org/10.1007/s00122-002-0910-8

He TM, Chen XS, Xu Z, Gao JS (2007). Using SSR markers to determine the population genetic structure of wild apricot (*Prunus armeniaca* L.) in the Ily Valley of West China. Genet Resour Crop Ev. 54(3): 563-572. https://doi.org/10.1007/s10722-006-0013-5

Hormaza JI (2002). Molecular characterization and similarity relationships among apricot (*Prunus armeniaca* L.) genotypes using simple sequence repeats. Theor Appl Genet. 104(2-3): 321-328. https://doi.org/10.1007/s001220100684

Khadivi-Khub A, Yarahmadi M, Jannatizadeh A, Ebrahimi A (2015). Genetic relationships and diversity of common apricot (*Prunus armeniaca* L.) based on simple sequence repeat (SSR) markers. Biochem Syst Ecol. 61: 366-371. https://doi.org/10.1007/s10722-013-0039-4

Kimura M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 16(2): 111-120. https://doi.org/10.1007/bf01731581

Larkin MA, Blackshields G, Brown NP, Chenna R (2007). Clustal W and Clustal X version 2.0. Bioinformatics. 23(21): 2947-2948. https://doi.org/10.1093/bioinformatics/btm404

Layne REC, Bailey CH, Hough LF (1996). Apricots. In: Fruit breeding, vol. 1: Tree and tropical fruits. Wiley. 79-111.

Lee S, Wen J (2001). A phylogenetic analysis of *Prunus* and the Amygdaloideae (Rosaceae) using ITS sequences of nuclear ribosomal DNA. AM J Bot. 88(1): 150-160. https://doi.org/10.2307/2657135

Li M, Zhao Z, Miao XJ (2013). Genetic variability of wild apricot (*Prunus armeniaca* L.) populations in the Ili Valley as revealed by ISSR markers. Genet Resour Crop Ev. 60(8): 2293-2302. https://doi.org/10.1007/s10722-013-9996-x

Li M, Zhao Z, Miao XJ (2014). Genetic diversity and relationships of apricot cultivars in North China revealed by ISSR and SRAP markers. Sci Hortic. 173: 20-28. https://doi.org/10.1007/s10722-013-9996-x

Liu J, Liao K, Nasir M, Sun Q (2016). Analysis of genetic diversity of the apricot germplasm in the southern region of the Tianshan Mountains in Xinjiang, China using the ISSR technique. Eur J Hortic Sci. 81(1): 37-43. https://doi.org/10.17660/ejhs.2016/81.1.5

Nei M (1973). Analysis of gene diversity in subdivided populations. P Natl Acad Sci USA 70(12), 3321-3323.

Rauscher JT, Doyle JJ, Brown AHD (2002). Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploid complex. Mol Ecol 11(12): 2691-2702. https://doi.org/10.1046/j.1365-294x.2002.01640.x

Rehder A (1949). Manual of cultivated trees and shrubs. Manual of cultivated trees and shrubs. MacMillan, New York, USA.

Rohlf FJ (2000). NTSYS- PC: Numerical Taxonomy System for the PC Exeter Software, Version 2.1. Applied Biostatistics Inc Setauket. USA.

Shi S, Li JL, Sun JH, Yu J, Zhou SL (2013). Phylogeny and classification of *Prunus* sensu lato (Rosaceae). J Integr Plant Biol. 55(11): 1069-1079.

Tang DL, Zhang QL, Yang Y, Luo ZhR (2014). Discriminant analysis of "Jinzaoshi" from persimmon (*Diospyros kaki* Thunb.; Ebenaceae): A comparative study conducted based on morphological as well as ITS and *mat*K sequence analyses. Sci Hortic. 168: 168-174. https://doi.org/10.1016/j.scienta.2014.01.033

Tamura K, Stecher G, Peterson D, Filipski A (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30(12): 2725-2729. https://doi.org/10.1093/molbev/mst197

Vavilov NI (1951). The origin, variation, immunity, and breeding of cultivated plants. Soil Sci. 72: 482.

White TJ, Bruns T, Lee S, Taylor J 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications. 18(1): 315-322. https://doi.org/10.1016/b978-0-12-372180-8.50042-1

Yeh FC, Yang RC, Boyle TB, Ye ZH (1997). POPGENE, the user-friendly shareware for population genetic analysis. Molecular biology and biotechnology center, University of Alberta, Canada.10.

Yuan ZH, Chen XS, He TM, Feng JR (2007). Population genetic structure in apricot (*Prunus armeniaca* L.) cultivars revealed by fluorescent-AFLP markers in southern Xinjiang, China. J Genet Genomics. 34(11): 1037-1047. https://doi.org/10.1016/s1673-8527(07)60117-9

Zaurov DE, Molnar TJ, Eisenman SW, Ford TM (2013). Genetic resources of apricots (*Prunus armeniaca* L.) in Central Asia. HortScience. 48(6): 681-691.

Zhang QP, Liu DC, Liu S, Liu N (2014). Genetic diversity and relationships of common apricot (*Prunus armeniaca* L.) in China based on simple sequence repeat (SSR) markers. Genet Resour Crop Ev. 61(2): 357-368. https://doi.org/10.1007/s10722-013-0039-4

Zhebentyayeva T, Ledbetter C, Burgos L, Llácer G (2012). Apricot. In: Fruit Breeding. Springer. 415-458.

Zhao L, Jiang XW, Zuo YJ, Liu XL (2016). Multiple events of allopolyploidy in the evolution of the Racemose lineages in *Prunus* (Rosaceae) based on integrated evidence from nuclear and plastid data. PLoS One. 11: e0157123. https://doi.org/10.1371/journal.pone.0157123