Analysis of Genetic Diversity in *Prunus Sibirica* L. in Inner Mongolia Using SCoT Molecular Markers

Habuer
Inner Mongolia Agricultural University
Sarula
Inner Mongolia Agricultural University
Wang Zi Yuan
Inner Mongolia Agricultural University
ShuFang
Inner Mongolia Agricultural University
Bai Yu'e (baiyue@imau.edu.cn)
Inner Mongolia Agricultural University

Research Article

Keywords: Prunus sibirica, Provenance, SCoT, Molecular marker, Genetic diversity

Posted Date: July 7th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-652651/v1

Abstract

Population genetic diversity contributes to the protection and utilization of germplasm resources, especially via genetic breeding. In the present study, start codon targeted polymorphism (SCoT) molecular markers were used to study the genetic diversity of 278 individuals from 10 *Prunus sibirica* populations in Inner Mongolia. A total of 289 polymorphic bands were amplified with 23 SCoT primers, showing a polymorphism percentage of 97.94% and an average of 12.6 polymorphic bands per primer. The SCoT21, SCoT32, and SCoT53 primers amplified up to 17 bands, and the polymorphism percentage was 100%. The minimum number of bands amplified by SCoT3 was 9, and the polymorphism percentage was 90%. Therefore, SCoT molecular markers were shown to be highly polymorphic and suitable for genetic diversity studies of *Prunus sibirica* in Inner Mongolia. The analysis of molecular variance (AMOVA) showed that 39% of the observed genetic differentiation occurred among populations and 61% occurred within populations, indicating that the genetic differentiation within populations was greater than that among populations. The results of the unweighted pair-group method with an arithmetic (UPGMA) cluster analysis, principal coordinate analysis (PCoA) and STRUCTURE analysis were basically the same and divided the 278 individuals from the 10 populations into 2 groups. The results indicated that the efficient SCoT molecular marker-based genetic diversity analysis of *Prunus sibirica* in Inner Mongolia can provide a reference for *Prunus sibirica* variety breeding and resource development.

Introduction

Prunus sibirica is a member of the Rosaceae family and is the dominant species on mountain dunes and dry steppes (Yu 1979). *Prunus sibirica* is mainly distributed in Inner Mongolia, Heilongjiang, Jilin, Liaoning, Gansu, Hebei, and Shanxi in China, and the *Prunus sibirica* forest in Inner Mongolia covers 47.44 hectares.

Prunus sibirica shows strong cold and drought resistance and high nutritional and medicinal value (Dong et al. 1991). However, the yield differences in *Prunus sibirica* in China are significant and very unstable because of long-term seed reproduction, delayed breeding, and the influence of climatic factors. In addition, frost can severely reduce the yield of *Prunus sibirica*, greatly restricting its industrialization (Ma et al. 2007; Yao et al. 2007; Jin et al. 2018). *Prunus sibirica* shows self-incompatibility, resulting in large genetic differences among individual genetic resources (Li et al. 2011).

Biodiversity refers to all of the variation among living things on earth and is the basis of survival and biological development (Rao et al. 2002). Genetic diversity is an important component of biological diversity and refers to the genetic differentiation of all living organisms (Lu 2018). Dong et al (2018) studied the diversity of 19 quantitative traits of *Prunus sibirica* from different populations, and the results showed that the coefficient of variation of the economic traits of fruit and kernel were large, and the coefficient of variation of fruit yield was the largest, indicating that *Prunus sibirica* has rich germplasm resources with great potential for breeding good varieties. Wang et al (2019) studied the seed traits of *Prunus sibirica* from 19 populations and showed rich genetic variations in seed traits between 19 populations. Li (2014) studied the genetic diversity of *Prunus sibirica* and *Prunus armeniaca* in North China according to inter-simple sequence repeat (ISSR), sequence-related amplified polymorphism (SRAP), and simple sequence repeat (SSR) analyses. The results showed that the natural population of *Prunus sibirica* in China exhibits a high level of genetic diversity, showing high genetic differentiation within its populations but low genetic differentiation and moderate gene flow among populations.

The initial SCoT marker was a molecular marker based on a single primer amplification reaction proposed by Collard and Mackill (2009) in rice. It was a novel molecular marker of the target gene. The strategy was to conduct genome amplification with a single primer according to the conservative nature of the ATG translation of the flanking

sequences of plant genes; the goal was to reveal the percentage of dominant polymorphic markers in candidate functional gene regions via a procedure with easy operation to identify rich polymorphisms. Compared with random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), SSR, and ISSR markers, SCoT markers have been efficiently produced and linked to traits, making them convenient for use in molecular marker-assisted breeding. SCoT markers have been used in various plant species, such as *Phoenix dactyifera* (Somayeh Saboori et al. 2020), *Mangifera indica* (Li Zhou et al. 2020) and *Diospyros* spp. (Changfei Guana et al. 2020).

Genetic resources for *Prunus sibirica* are extremely abundant and present great developmental potential, and they can provide an important genetic basis for the improvement of Chinese *Prunus sibirica* and the breeding of new varieties. In recent years, research on the genetic diversity of *Prunus sibirica* has mostly been conducted in northeastern and northern China, and *Prunus sibirica* is mainly distributed in Inner Mongolia. Moreover, few studies have been conducted on the genetic diversity of *Prunus sibirica* in Inner Mongolia. Therefore, we used SCoT molecular markers to analyze the genetic diversity of 278 individuals from 10 populations of *Prunus sibirica* in Inner Mongolia. The genetic diversity and genetic structure among populations and individuals of *Prunus sibirica* in Inner Mongolia were revealed. This study provides a scientific basis for the cultivation and exploitation of the abundant genetic resources of *Prunus sibirica* in Inner Mongolia.

Materials And Methods

Plant materials

The materials were obtained from the *Prunus sibirica* Germplasm Resource Garden of the Inner Mongolia Fine Variety Breeding Center (Fig. 1). A total of 278 individuals were collected from 10 populations (Table 1). Fresh young leaves were placed in a bucket containing liquid nitrogen, brought back to the laboratory, and stored in a -80°C freezer for further experiments.

Table 1
Site conditions and sampling information for *Prunus sibirica* populations

Population Abbreviation	Longitude and Latitude	Elevation (m)	Sample size
KSK	117°54'57"/43°26'49"	1,100	46
WJG	111°02'38"/40°75'73"	1,629	23
KYZ	121°47'65"/45°06'08"	270	26
KZH	122°35'67"/42°93'51"	198	38
ZLA	122°89'96"/46°72'32"	628	22
LC	112°50'39"/40°53'15"	1,731	31
BLY	118°66'51"/43°53'44"	1,050	25
WLS	108°77'63"/40°66'26"	2,322	14
AL	120°06'56"/43°87'22"	900	27
ZLT	120°91'16"/44°55'63"	812	26

DNA extraction

Genomic DNA was extracted from fresh leaves using a Plant DNA Extraction Kit (TIANGEN, China). The extracted DNA was tested for quality and purity with a

spectrophotometry at a wavelength of 260/280 nm using a BioPhotometer (Thermo Fisher Scientific, America) and 1% agarose gel electrophoresis (with a nucleic acid stain) and stored in a -20°C freezer for further experiments.

SCoT-PCR

The 80 primer sequences were selected by referring to Collard and Mackill (2009) and were synthesized by Shanghai Sangon Biological Engineering Technology and Services. A total of 23 highly polymorphic and repeatable primers were selected from 80 SCoT primers to evaluate the selected *Prunus sibirica* accessions (Table 2).

Table 2
Screened SCoT primer sequence information

Code	Primer name	Primer Sequences (5'- 3')	GC%	Annealing Tm (°C)
1	SCoT3	CAACAATGGCTACCACCG	55.6	52
2	SCoT4	CAACAATGGCTACCACCT	50.0	50
3	SCoT10	CAACAATGGCTACCAGCC	55.6	55
4	SCoT12	ACGACATGGCGACCAACG	61.1	54
5	SCoT13	ACGACATGGCGACCATCG	61.1	54
6	SCoT21	ACGACATGGCGACCCACA	61.1	54
7	SCoT23	CACCATGGCTACCACCAG	61.1	54
8	SCoT25	ACCATGGCTACCACCGGG	66.7	53
9	SCoT27	ACCATGGCTACCACCGTG	61.1	54
10	SCoT29	CCATGGCTACCACCGGCC	72.2	51
11	SCoT30	CCATGGCTACCACCGGCG	72.2	51
12	SCoT31	CCATGGCTACCACCGCCT	66.7	53
13	SCoT32	CCATGGCTACCACCGCAC	66.7	53
14	SCoT34	ACCATGGCTACCACCGCA	61.1	54
15	SCoT43	CAATGGCTACCACCGCAG	61.1	54
16	SCoT44	CAATGGCTACCATTAGCC	50.0	50
17	SCoT45	ACAATGGCTACCACTGAC	50.0	53
18	SCoT46	ACAATGGCTACCACTGAG	50.0	50
19	SCoT47	ACAATGGCTACCACTGCC	55.6	52
20	SCoT53	ACAATGGCTACCACCGAC	55.6	52
21	SCoT58	ACAATGGCTACCACTAGG	50.0	50
22	SCoT71	CCATGGCTACCACCGCCG	72.2	51
23	SCoT72	CCATGGCTACCACCGCCC	72.2	51

A 96-well gradient PCR instrument was used for the amplification reactions. PCR was performed in a 20 μL volume containing 10 μL of 2× Taq Mix, 1.6 μL (30 ng/μL) of genomic DNA, 0.8 μL (10 mol·L^{-1}) of each primer, and 7.6 μL of ddH_2O. SCoT-PCR amplification was conducted with initial denaturation for 5 min at 95 ℃, followed by 40 cycles of 45 s at 94℃, 45 s at 50 to 54℃ and 2 min at 72℃, with a final extension of 7 min at 72℃ and holding at 4℃. The PCR products were separated in a 1.5% agarose gel using 1 × TBE running buffer, stained, and photographed, and the records were preserved.

Data analysis

SCoT amplicons were scored in a binary matrix as present (1) or absent (0) for each sample based on the corresponding standard size. Vague bands that could not be easily detected were not scored. POPGEN32 software was used to calculate the total number of bands (NPB) and the number of polymorphous bands (PPB) obtained with the SCoT primers to analyze the genetic diversity index of *Prunus sibirica* and to calculate Nei's gene diversity index (H), Shannon's information index (I), the number of alleles (Na), the effective number of alleles (Ne), total genetic diversity (Ht), population genetic diversity (Hs), the coefficient of genetic differentiation (Gst), and gene flow (Nm). Ntsys-2.0 software was used to calculate Nei's genetic distances from the binary (0, 1) matrices obtained, and the unweighted pair-group method with arithmetic means (UPGMA) method performed in the SHAN program was used for the cluster analysis of the 10 populations to construct the cluster relationship tree diagram. MEGA7 software was used for the cluster analysis of the 278 individuals based on Nei's genetic similarity matrix obtained via the arithmetic mean (UPGMA) method. An analysis of molecular variance (AMOVA) and GenAlEx software were used for the analysis of genetic differentiation among and within populations. A principal coordinate analysis (PCoA) was performed according to the binary matrix using GenAlEx software. STRUCTURE software was used to determine the genetic structure of the studied population, and K was tested from 1 to 10 with ten replicates. IBD software was used for Mantel test to analyze the correlation between genetic distance and geographic distance and altitude .

Results

SCoT polymorphism analysis

In the SCoT analysis, 23 SCoT primers were used for marker amplification in 278 *Prunus sibirica* individuals. A total of 292 bands that could be scored were produced, among which 289 bands were polymorphic, with a mean of 12.6 polymorphic bands per primer. The polymorphism percentage was 98.87%. Among the 23 SCoT primers, SCoT21, SCoT32, and SCoT53 amplified a maximum of 17 bands, and the polymorphism percentage was 100%. The number of bands amplified by the SCoT 25 primer was at least 9, and the polymorphism percentage was 90%. The 23 SCoT primers produced Na values ranging from 1.9000 to 2.0000, with a mean of 1.9803; Ne values ranging from 1.2848 to 1.5955, with a mean of 1.4252; H values ranging from 0.1808 to 0.3399, with a mean of 0.2571; and I values ranging from 0.2969 to 0.5049, with a mean of 0.3989 (Table 3).

Table 3
Polymorphism information obtained from SCoT primer amplification in *Prunus sibirica*

Primer	Total loci	NPB	PPB (%)	Na	Ne	H	I
SCoT3	9	9	100	2.0000	1.4959	0.3021	0.4660
SCoT4	11	11	100	2.0000	1.4514	0.2851	0.4453
SCoT10	11	11	100	2.0000	1.5080	0.3192	0.4970
SCoT12	11	11	100	2.0000	1.3490	0.2189	0.3528
SCoT13	16	16	100	2.0000	1.2848	0.1836	0.3050
SCoT21	17	17	100	2.0000	1.4044	0.2448	0.3797
SCoT23	11	11	100	2.0000	1.5643	0.3250	0.4858
SCoT25	10	9	90	1.9000	1.4931	0.2832	0.4288
SCoT27	12	12	100	2.0000	1.5101	0.3005	0.4557
SCoT29	11	11	100	2.0000	1.3379	0.2084	0.3313
SCoT30	12	11	91.67	1.9167	1.3590	0.2223	0.3527
SCoT31	10	10	100	2.0000	1.5955	0.3399	0.5049
SCoT32	17	17	100	2.0000	1.4511	0.2691	0.4331
SCoT34	14	14	100	2.0000	1.4565	0.2826	0.4358
SCoT43	11	10	90.91	1.9091	1.2851	0.1808	0.2969
SCoT44	14	14	100	2.0000	1.4857	0.2761	0.4186
SCoT45	15	15	100	2.0000	1.3031	0.1991	0.3258
SCoT46	12	12	100	2.0000	1.4191	0.2373	0.3622
SCoT47	13	13	100	2.0000	1.3864	0.2401	0.3845
SCoT53	17	17	100	2.0000	1.3779	0.2317	0.3656
SCoT58	10	10	100	2.0000	1.5572	0.3348	0.5045
SCoT71	12	12	100	2.0000	1.4638	0.2705	0.4163
SCoT72	16	16	100	2.0000	1.5146	0.3085	0.4664
Mean	12.7	12.6	98.87	1.9803	1.4252	0.2571	0.3989
Total	292	289					

Population genetic diversity and genetic differentiation

The results of the genetic diversity and genetic differentiation analyses in the 10 populations of *Prunus sibirica* in Inner Mongolia showed that NPB ranged from 110 to 180, with a mean of 156.4 bands; PPB ranged from 37.67 to 61.64, with a mean of 53.59 polymorphisms; Na ranged from 1.3934 to 1.6230, with a mean of 1.5455; Ne ranged from 1.1956 to 1.3130, with a mean of 1.2767; H ranged from 0.1181 to 0.1871, with a mean of 0.1662; and I ranged from 0.1812 to 0.2859, with a mean of 0.2540. Among the 278 individuals, the overall PPB, Na, Ne, H, and I total

values were 97.94%, 1.9803, 1.4252, 0.2571, and 0.3989, respectively (Table 4), implying the presence of considerable genetic diversity among the *Prunus sibirica* accessions. The genetic diversity was determined to be highest among the KYZ population, with an H value of 0.1871 and a PPB of 61.64%. The second-highest H value of 0.1848 was observed for KZH, for which PPB was 60.62%. In contrast, the genetic diversity was found to be lowest in WJG, with an H value of 0.1181 and a PPB of 37.67%.

Table 4
Genetic diversity indices of *Prunus sibirica* populations

Population	No. of Samples	NPB	PPB (%)	Na	Ne	H	I
KSK	46	145	49.65	1.5082	1.2654	0.1584	0.2409
WJG	23	110	37.67	1.3934	1.1956	0.1181	0.1812
KYZ	26	180	61.64	1.6230	1.3130	0.1871	0.2859
KZH	38	177	60.62	1.6131	1.3099	0.1848	0.2790
ZLA	22	162	55.48	1.5639	1.2872	0.1746	0.2698
LC	30	156	53.42	1.5443	1.2705	0.1613	0.2462
BLY	26	169	57.87	1.5869	1.3067	0.1834	0.2793
WLS	14	153	52.39	1.5344	1.2711	0.1636	0.2504
AL	27	156	53.42	1.5443	1.2654	0.1612	0.2479
ZLT	26	156	53.42	1.5443	1.2818	0.1699	0.2591
Mean	27.8	156.4	53.56	1.5455	1.2767	0.1662	0.2540
Total	278	286	97.94	1.9803	1.4252	0.2571	0.3989

The 278 individuals showed an Ht of 0.2598, with a variance of 0.0294; Hs of 0.1662, with variance of 0.0122; Gst of 0.3601, where the mean showed a high Gst value (Gst > 0.25); and Nm of 0.8884 (Nm < 1), with the mean showing low gene flow (Table 5). The AMOVA showed 9 degrees of freedom among the populations and 267 degrees of freedom within the populations. The sum of squares of the variance among the populations was 4,622.89, and that within the populations was 7,385.47. The among-population variance was 17.61 on average and the within-population variance was 27.56; 39% of the genetic variation occurred among the populations, and 61% occurred within the populations (Table 6).

Table 5
Genetic differentiation within *Prunus sibirica* populations

	Total	Ht	Hs	Gst	Nm
Mean	278	0.2598	0.1662	0.3601	0.8884
StDev		0.0294	0.0122		

Table 6
Analysis of molecular variance (AMOVA) of *Prunus sibirica* populations

Source of Variance	df	Sum of squares	Variance component	Percentage variance (%)	P value
Among populations	9	4,622.89	17.61	39	< 0.001
Within populations	267	7,385.47	27.56	61	< 0.001
Total	278	12,008.37	45.22		

Analysis of the among-population genetic structure

POPGEN32 was used to analyze the genetic distance and genetic similarity of the 10 *Prunus sibirica* populations in Inner Mongolia. The analyses of Nei's genetic similarity and Nei's genetic distance showed that the population genetic distance ranged from 0.0311 to 0.2486 while the genetic similarity ranged from 0.7799 to 0.9694. The genetic similarity between WJG and KSK was the highest, at 0.9694, and their genetic distance was the lowest, at 0.0311. The lowest genetic similarity between ZLT and WJG was 0.7799, and the greatest genetic distance was 0.2486 (Table 7). Based on Nei's genetic distance, the UPGMA method was used for cluster analysis. The results showed that the 10 populations were divided into two major groups at a genetic distance threshold of 0.22 (Fig. 2). Group I included KSK, WJG, LC, KYZ, KZH, and ZLA. Group II included BLY, AL, WLS, and ZLT. When the genetic distance was less than 0.17, group I was divided into subgroup I and subgroup II. Subgroup I included KSK, while subgroup II included WJG, LC, KYZ, KZH, and ZLA. Group II was divided into subgroup III and subgroup IV. Subgroup III included BLY, AL, and WLS, while subgroup IV included ZLT.

Table 7
Genetic similarity (above diagonal) and genetic distance (below diagonal)among populations

Population	KSK	WJG	KYZ	KZH	ZLA	LC	BLY	WLS	AL	ZLT
KSK		0.9694	0.9087	0.8870	0.8800	0.8522	0.8660	0.8337	0.8074	0.8008
WJG	0.0311		0.9027	0.8731	0.8620	0.8317	0.8507	0.8117	0.7837	0.7799
KYZ	0.0958	0.1024		0.9484	0.9235	0.8759	0.8838	0.8495	0.8156	0.8114
KZH	0.1199	0.1357	0.0530		0.9422	0.9078	0.8996	0.8641	0.8442	0.8219
ZLA	0.1279	0.1485	0.0796	0.0595		0.9578	0.9268	0.8973	0.8804	0.8733
LC	0.1599	0.1843	0.1325	0.0968	0.0430		0.9506	0.8960	0.8898	0.8917
BLY	0.1438	0.1617	0.1235	0.1058	0.0760	0.0507		0.9190	0.8914	0.8935
WLS	0.1819	0.2086	0.1631	0.1460	0.1084	0.1098	0.0845		0.9373	0.9135
AL	0.2139	0.2438	0.2038	0.1693	0.1274	0.1167	0.1150	0.0647		0.9582
ZLT	0.2221	0.2486	0.2089	0.1962	0.1354	0.1146	0.1126	0.0905	0.0427	

Analysis of genetic structure within populations

Based on Nei's genetic similarity, MEGA7 software was used to perform the clustering analysis of 278 individuals (Fig. 3), with the PCoA performed using GenAlEx software (Fig. 4). The results showed that the 278 individuals were divided into two groups. The UPGMA clustering analysis showed that the 278 individuals were divided into group I and group II when the genetic similarity was 0.17. Group I included 144 individuals, while group II included 134 individuals.

Group I included 23 WJG individuals, 46 KSK individuals, 26 KYZ individuals, 38 KZH individuals, and 11 ZLT individuals. Group II included 27 AL individuals, 14 WLS individuals, 26 ZLT individuals, 25 BLY individuals, 31 LC individuals, and 11 ZLA individuals. When the genetic similarity was ≤ 0.14, group I and group II were divided into four subgroups. A total of 23 WJG individuals and 46 KSK individuals were clustered into subgroup I; 26 KYZ individuals, 38 KZH individuals and 11 ZLA individuals were divided into subgroup II; 27 AL individuals, 12 ZLT individuals, and 9 WLS individuals were clustered into subgroup III; and 14 ZLT individuals, 5 WLS individuals, 11 ZLA individuals, 26 BLY individuals, and 30 LC individuals were included in subgroup IV. The PCoA results showed that the 278 *Prunus sibirica* individuals were divided into groups I and II. These results were relative to those of the UPGMA clustering analysis of similarity. The two groups were divided into four subgroups. Group I included all individuals in the KSK, WJG, KZH, and KYZ populations and some individuals from ZLA. Group II included all individuals from AL, WLS, ZLT, BLY, and LC and some individuals from ZLA.

To further explore the genetic relationships within populations of *Prunus sibirica*, the population structure of 278 *Prunus sibirica* individuals was evaluated using STRUCTURE software. ΔK values computed for all classes indicated a strong signal for K = 2 (Fig. 5),and K = 2 values provided the most rational arrangement of *Prunus sibirica* in different regions. A total of 278 *Prunus sibirica* individuals were divided by the STRUCTURE analysis into 2 groups (Fig. 6). Group I included all individuals in the KSK, WJG, KZH, and KYZ populations and some individuals from ZLA. Group II included all individuals from AL, WLS, ZLT, BLY, and LC and some individuals from ZLA. This result is similar to the clusters identified using UPGMA and PCoA.

Discussion

In this study, a genetic diversity analysis was performed using SCoT markers to assess phylogenetic relationships among 278 *Prunus sibirica* individuals from 10 different populations in Inner Mongolia. Twenty-three SCoT primers were used for the amplification of genomic DNA from *Prunus sibirica*. A total of 292 clear bands were obtained, 289 of which were polymorphic. The average number of amplified bands per primer was 12.7, with a polymorphism percentage of 98.87%. In a previous study, researchers used SRAP (Ai et al. 2011), AFLP (Wang 2008), RAPD (Lu 2008), and ISSR (Liu et al. 2011; Duan et al. 2010; Li et al. 2009; Liu et al. 2007) molecular markers to analyze the genetic diversity of *Prunus sibirica* in different regions, and the results showed that the polymorphism ranged from 58–90%. In contrast, the SCoT molecular markers were found to be more polymorphic than other molecular markers. Therefore, SCoT markers may be suitable for the study of genetic diversity according to molecular markers in *Prunus sibirica*. A previous study of the genetic diversity of *Prunus armeniaca* (Li 2014) indicated H, I, and polymorphism percentage values of H = 0.1720, I = 0.2576 and PPB = 50.07%; *Sibirica* (Ma 2014) indicated H, I, and polymorphism percentage values of H = 0.2318, I = 0.3530 and PPB = 76.74%. However, the H, I, and polymorphism percentage values obtained in the present study were H = 0.2571, I = 0.3989 and PPB = 97.94%, respectively. We showed that the genetic diversity of the *Prunus sibirica* populations in Inner Mongolia is higher than that of *Prunus armeniaca* and *Sibirica*. This result was consistent with the conclusions of other scholars that the genetic diversity of outbred species is higher (Hamrick 1989; Zheng et al. 2008).

As a perennial wild resource, *Prunus sibirica* is widely distributed in Inner Mongolia and exhibits a large distribution area, long-term evolution and considerable diffusion as well as ecological diversity, leading to rich genetic diversity. According to Wright (1972), when the F_{ST} (G_{ST}, genetic differentiation coefficient) value is 0 to 0.05, the genetic differentiation within populations is low; when the F_{ST} value is 0.05 to 0.15, the genetic differentiation within populations is moderate; when the F_{ST} value is 0.15 to 0.25, the genetic differentiation within populations is high; and when the F_{ST} value is greater than 0.25, the genetic differentiation within populations is great. In our research, the G_{ST}

was 0.3601. The results showed that the genetic differentiation coefficient of *Prunus sibirica* within populations in Inner Mongolia was relatively high, with among-population genetic differentiation of 39% and a within-population genetic differentiation of 61%. Wang (2019), Liu et al. (2012) and Ma (2013), who used SSR and ISSR markers for *Prunus sibirica* and wild apricot genetic diversity analyses, showed that within-population variation dominated, and the among-population variation was much lower than the within-population variation. which is similar to the results of a previous study in *Prunus sibirica.*

The reasons for the high within-population genetic differentiation of *Prunus sibirica* in Inner Mongolia may include the following: 1. Gene flow is an important factor affecting genetic differentiation in a population when Nm < 1, which can conversely effectively prevent genetic differentiation caused by genetic drift. The Nm value for *Prunus sibirica* in Inner Mongolia was 0.8884, indicating that genetic drift was not the main factor influencing the genetic differentiation of the *Prunus sibirica* populations in Inner Mongolia. Some gene flow exists among the *Prunus sibirica* population in Inner Mongolia, but the Nm intensity was relatively low. 2. The self-incompatibility breeding characteristic of *Prunus sibirica* may eliminate self-pollination as a possible cause of genetic differentiation, and its breeding mainly relies on natural pollination or pollination by visiting insects such as bees, flies and a few butterflies (Liu et al. 2010; Liu 2010). Although pollen can be spread by insects and by wind over long distances, the large distribution area and discontinuities in the population distribution of *Prunus sibirica* are restricted due to geographic isolation or habitat fragmentation (Wu et al. 2015), thus weakening the gene flow within populations of *Prunus sibirica*. 3. Some populations have been destroyed or their habitats have been degraded, resulting in a gradual fragmentation of the population distribution and limiting the gene flow within the populations. Therefore, *Prunus sibirica* in Inner Mongolia shows high within-population genetic differentiation under the influence of multiple factors, such as physical geographic isolation and human activities.

In the present study, the genetic structure of *Prunus sibirica* populations in Inner Mongolia was analyzed. From the clustering diagram, principal component analysis, and structure analysis, we can see that most of the populations are clustered together with a similar geographic distribution, and the Mantel test showed a significant correlation between the genetic distance and geographic distance and between the genetic distance and altitude of the *Prunus sibirica* studied ($r = 0.0426$, $P \leqq 0.4790$;$r = 0.0305$, $P \leqq 0.4260$). Therefore, we hypothesized that a geographic isolation effect exists among populations of *Prunus sibirica* populations in Inner Mongolia, weakening the opportunities for gene flow within the populations and causing the intensification of genetic differentiation within populations. The clustering diagram, principal component analysis and structure analysis of the 278 individuals divided the populations into two groups (I and II). KSK, WJG, KYZ, and KZH were clustered in group I; LC, BLY, WLS, AL, and ZLT were clustered in group II; and only the 22 individuals in the ZLA group were divided between the two groups. A certain degree of gene flow exists among the subgroups, and gene flow also affects the population structure of *Prunus sibirica* (Liu et al 2012). This may occur because of the existence of many mountains and areas of sandy land in Inner Mongolia, and the topography and landforms are very complex. Thus, topographic barrier characteristics are prominent in the regional distribution, resulting in a natural isolation effect on different provenances of species and weakening the gene flow opportunities among populations. Second, habitat fragmentation is also an important factor that affects the composition and structure of the ecosystem and the genetic structure of a species.

Conclusion

This study revealed that *Prunus sibirica* shows high genetic diversity in Inner Mongolia. Therefore, improving the preservation and utilization of specific germplasms of different populations is necessary to avoid the disappearance of a large number of valuable genetic resources. Moreover, strengthening the protection of the genetic resources of *Prunus sibirica* and expanding the collection of *Prunus sibirica* resources to implement transfer protection will be

conducive to better scientific research and resource protection in *Prunus sibirica*, increase the gene flow among populations, and improve the genetic diversity of *Prunus sibirica.* Based on resource collection, genetic diversity analyses, and genetic structure analyses, the breeding of *Prunus sibirica* varieties with high yields and frost resistance is being actively performed to meet the increasing demand for the industrial production of *Prunus sibirica* resources.

Declarations

Acknowledgments This research was supported by the National Science and Technology of China (2013BAD14B0203) and the Inner Mongolia Science and Technology Department of China (2015020553).

Author contribution statement HBR wrote the manuscript, organized the data and designed the research. SRL provided the funding for the experiment. WZY and SF conducted experiments. BYE conceived and designed the research. All authors read and approved the manuscript.

Data Availability Statement: Data is contained within the article.

Conflict of Interest: The authors declare that there is no conflict of interests regarding the publication of this paper.

References

1. Ai PF, Zhen ZJ, Jin ZZ (2011) Genetic diversity and relationships within sweet kernel apricot and related *Armeniaca*, species based on sequence-related amplified polymorphism markers. Biochemical Systematics & Ecology 39(4–6):694-699.
2. Guan CF, Chachar S, Zhang PX, Hua CQ, Wang RZ, Yang,Y (2020) Inter-and intra-specific genetic diversity in diospyros using SCoT and IRAP markers.Horticultural Plant Journal 6 (2): 71-80.
3. Collard BCY, Mackill DJ (2009) Start codon targeted (SCoT) polymorphism: A simple novel DNA marker technique for generating gene-targeted markers in plants. Plant Molecular Biology Reporter 27(1):86-93.
4. Dong SJ, Wang L, Liu LX, Liu MG, Wu YL, Chen JH, Xia ZZ (2018) Studyon quantitative characteristics in Armeniaca sibirica clones from different provenances.Nonwood Forest Research 36(04):1-8.
5. Dong YS, He R, Lin FQ (1991) Study on genetic resources of *Prunus sibirica*. Northern Fruit Trees 1:4-8.
6. Duan W, Yan WD, Wu YTN (2010) Identification and analysis of 14 samples in *Prunus Simonii* by ISSR molecular marker. Journal of Inner Mongolia agricultural university 2:147-153.
7. Hamrick JL (1989) Isozymes and the analysis of genetic structure in plant populations. Isozymes in Plant Biology pp 87-105.
8. Jin L, Liu GQ, Huang SJ, Wu YL, Zhang X (2018) Genetic diversity and fingerprints of 97 *Armeniaca sibirica* clones based on SSR markers. Scientia Silvae Science 54(7):51-61.
9. Li M (2014) Genetic diversity of common apricot and siberian apricot in north China. Journal of Northwest A&F University.
10. Li M, Zhao Z, Yang JA (2011) Classification on genetic resources of *Prunus sibirica* in the Loess Plateau. Journal of Northwest Forestry University 1:8-12.
11. Li M, Zhao Z, Yang JA (2011) Genetic diversity analysis on germplasm of *Prunus sibirica* in different counties in Loess Plateau. Journal of Northwest A&F University 2:143-149+156.
12. Li MM, Cai YL, Qian ZQ (2009) Genetic diversity and differentiation in Chinese sour cherry Prunus pseudocerasus lindl., and its implications for conservation. Genetic Resources and Crop Evolution 56(4):455-464.

13. Zhou L, He XH, Yu HX, Chen MY, Fan Y, Zhang XJ，Fang ZB，Luo C (2020) Evaluation of the genetic diversity of mango (*Mangifera indica* L.) seedling germplasm resources and their potential parents with start codon targeted (SCoT) markers.Genet Resour Crop Evol 67:41-58.
14. Liu HB, Wang Z, Liu J (2012) Genetic diversity and genetic structure of *Prunus sibirica* populations in Yan Mountains. Scientia Silvae Science 8:68-74.
15. Liu LM (2010) Primary research on reproductive ecology of *Armeniaca vulgaris* Lam. Xinjiang Agricultural University ,Wulumuqi.
16. Liu MP, Bo DL, Tian M (2011) Analysis on genetic differentiation of *Armeniaca cathayana* F_1 hybrids by ISSR. Journal of Central South University of Forestry and Technology 10:100-104.
17. Liu W, Liu D, Zhang A (2007) Genetic diversity and phylogenetic relationships among plum genetic resources in China assessed with inter-simple sequence repeat markers. Journal of the American Society for Horticultural Science 132(5):619-628.
18. Liu YX, Chen WZ, Liu WS (2010) Palynological study on the origin and systematic evolution of kernel-using apricots. Acta Horticulturae Sinica 9:1377-1387.
19. Lu CY (2018) Genetic diversity based on SSR and its association analysiswith phenotypic traits in *Armeniaca vulgaris*. Shenyang Agricultural University, Shenyang.
20. Lu NC (2008) Genetic resources on the genetic diversity ofsiberian apricot in Jilin. Jilin Agricultural University,Changchun.
21. Ma LJ (2013) Genetic diversitv and genetic structure analysis of wild apricot populations in Northeast China. Jilin Agricultural University,Changchun.
22. Ma YH, Zhao Z, Li KY (2007) Optimization of technology for almond oil extraction by supercritical CO_2. Transactions of The CSAE 4:272-275.
23. Ma LJ, Guo TJ, He DD, Li F, Zhang YB, Ma LX (2014) Genetic diversity analysis of two natural A.*sibirica* populations in northeast China. Journal of Jilin Agricultural University 36(03):300-305.
24. Rao VR, Hodgkin T (2002) Genetic diversity and conservation and utilization of plant genetic resources. Plant Cell, Tissue and Organ Culture 68(1):1-19.
25. Somayeh Saboori, Zahra Noormohammadi, Masoud Sheidai, Seyyed Samih Marashi (2020). SCoT molecular markers and genetic fingerprinting of date palm (Phoenix dactylifera L.) cultivars.Genet Resour Crop Evol 67:73-82.
26. Wang L (2019) Genetic diversity and population genetic structure of *Armeniaca sibirica* clones. Shenyang Agricultural University,Shenyang.
27. Wang L (2008) Using AFLP markers to analyze genetic diversityin *Prunus Sibirica* (L.) Lam. Hebei Agricultural University, Baoding.
28. Wang SX, Hu XY, Fan SQ, Cu YF, Yu HY, Wang LB (2019) Geographical variation analysis on Armeniaca sibirica seed characteristics from different provenances.Nonwood Forest Research 37(01):74-79+86.
29. Wright S (1972) Evolution and the genetics of populations. A treatise in four volumes.Volume 4. Variability within and among natural populations. Journal of Biosocial Science.4(2):253-256.
30. Wu YT, Amo H, Xu J (2015) Population structure of and conservation stpercentagegies for wild pyrus ussuriensis maxim in China. Plos One 10(8):e0133686.
31. Yao ZY, Li KY, Zhao Z (2007) Antioxidant properties of melanin from testae of *Armeniaca vulgaris* var. Ansu. Scientia Silvae Science 10:59-63.Yao ZY, Zhao Z, Shi QH (2007) Isolation of melanin from testae of *Prunus Ameniaca* Linn. Journal of Northwest A&F University 5:120-126.

32. Yu DJ (1979) Chinese fruit tree taxonomy. Beijing Agricultural Press. pp 45-90.
33. Zeng J, Zeng YQ, Zang CH (2008) Genetic diversity in natural populations of *Sorbus Pohuashanensis*. Biodiversity Science 6:562-569.

Figures

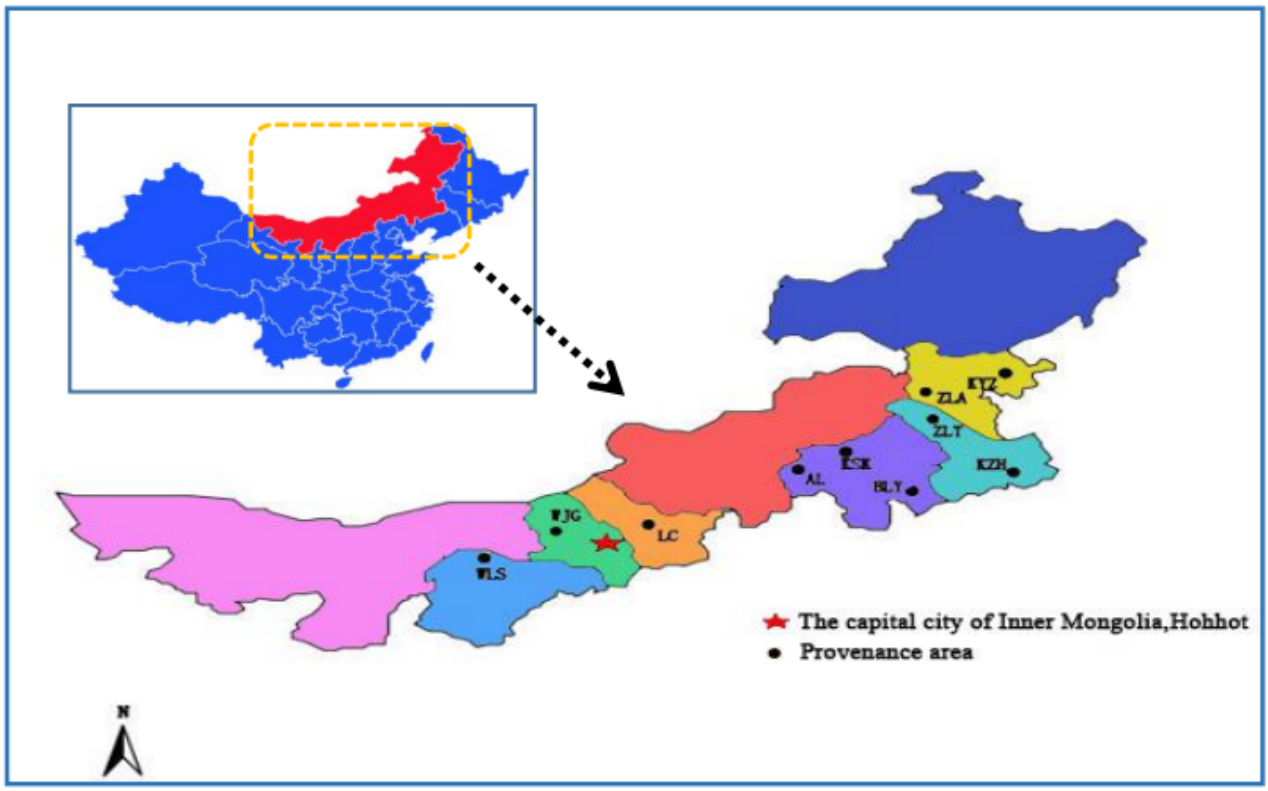

Figure 1

Geographic location map

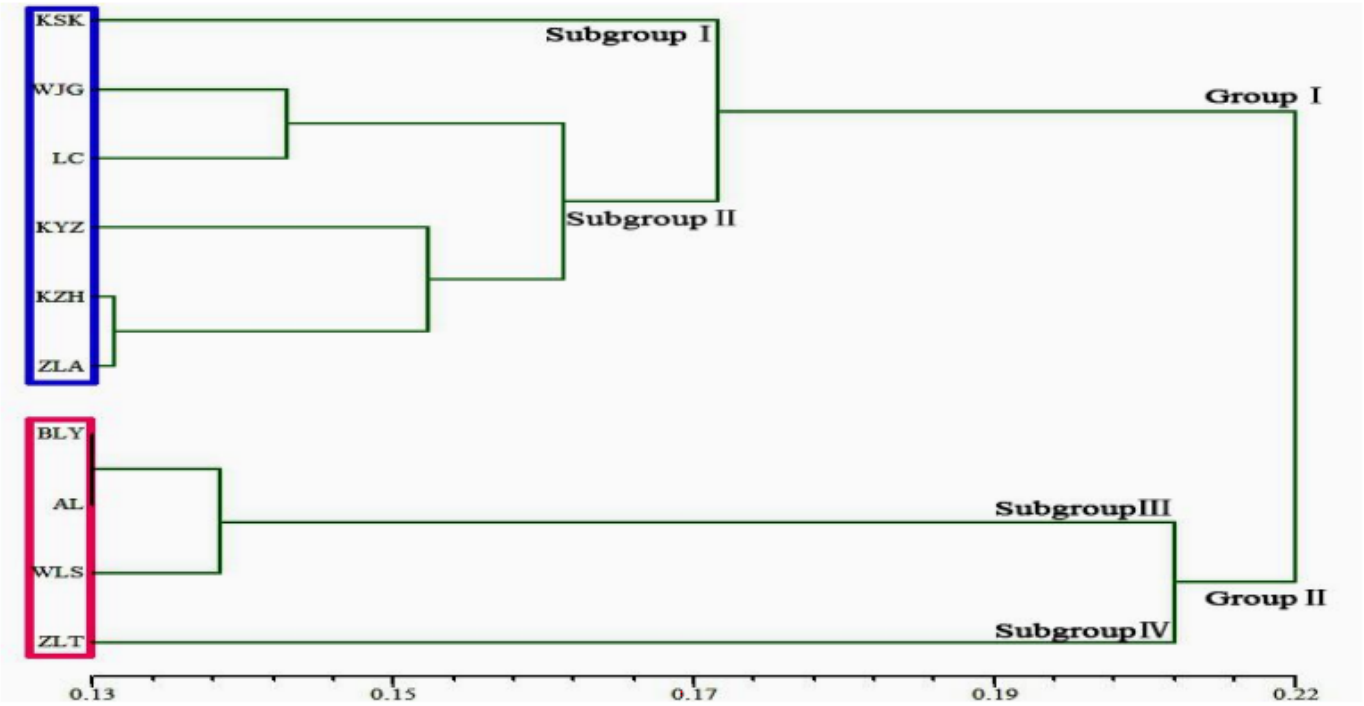

Figure 2

UPGMA dendrogram of Prunus sibirica populations based on Nei's genetic distance

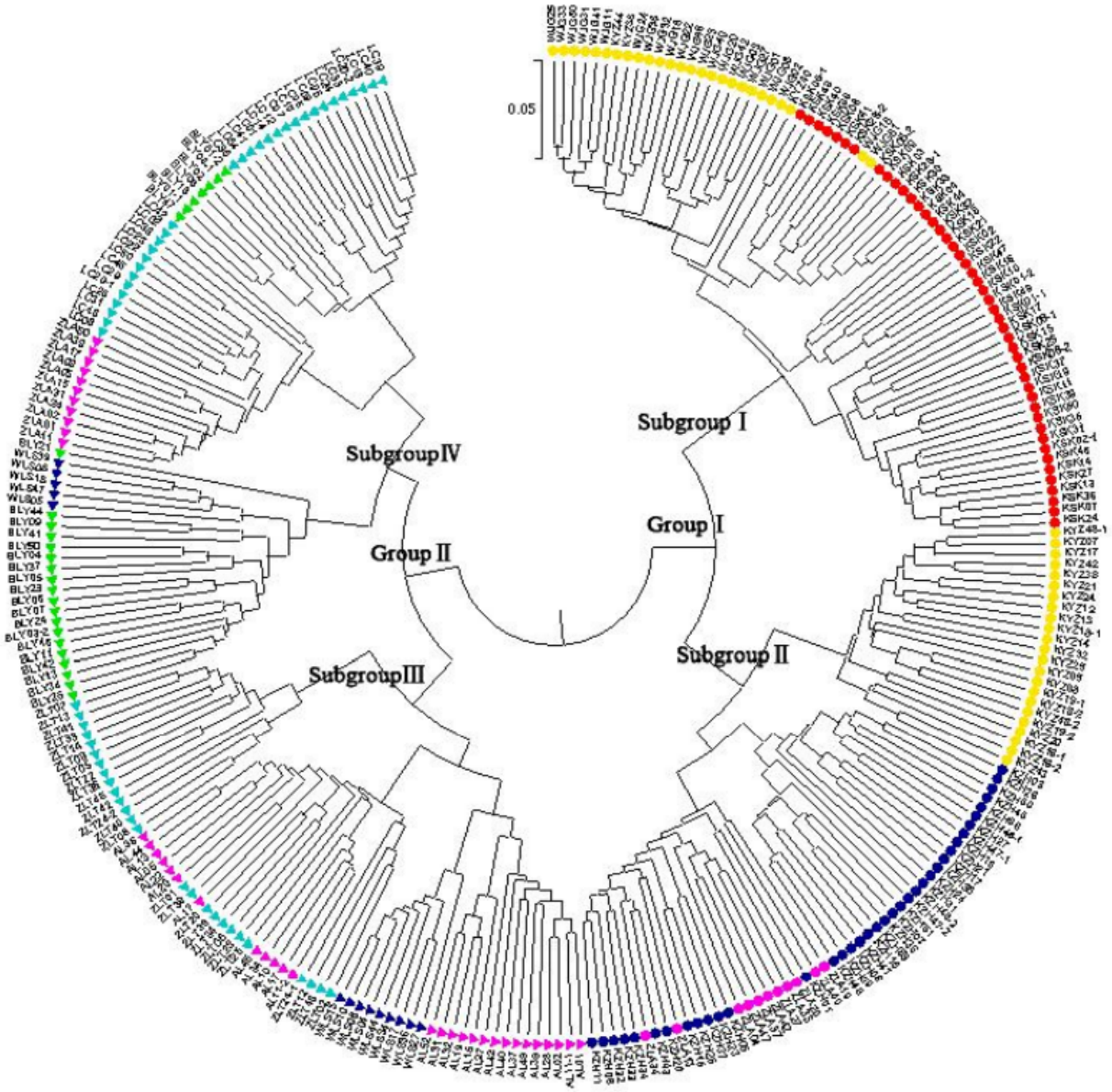

Figure 3

UPGMA dendrogram of Prunus sibirica populations based on the SCoT markers

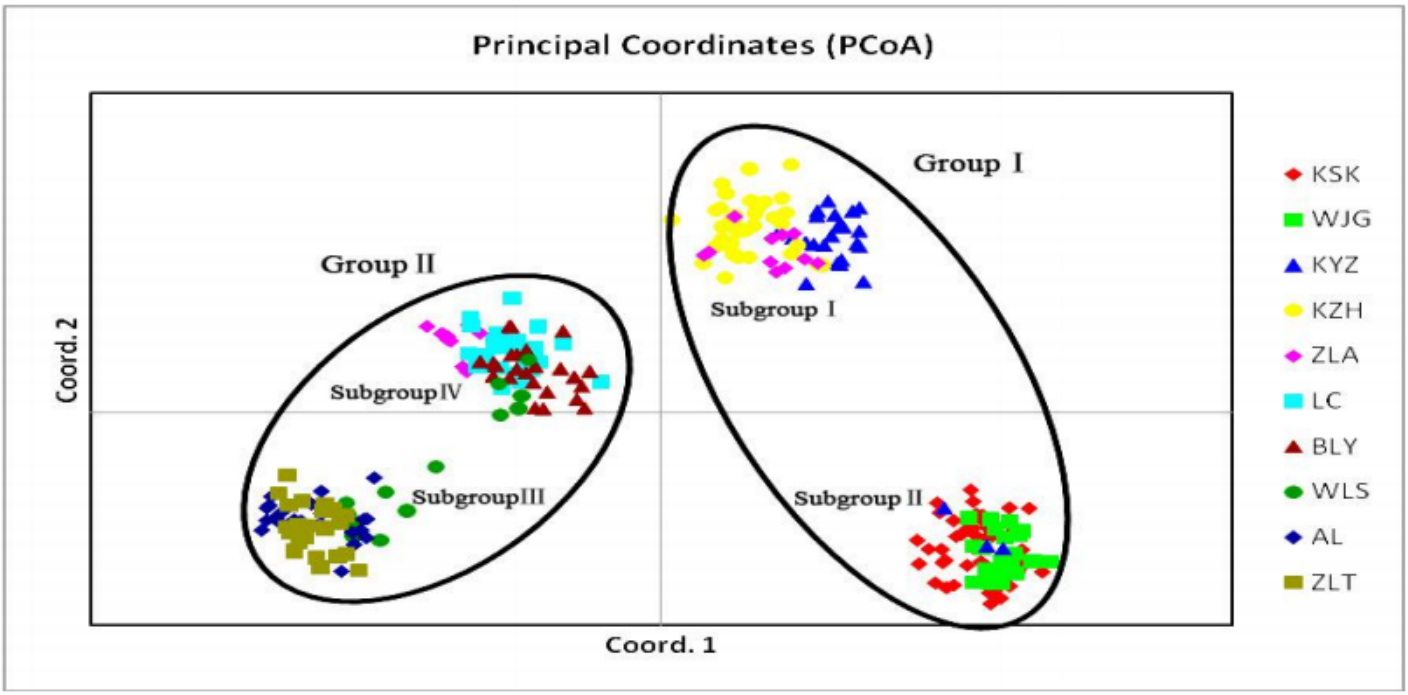

Figure 4

Principal coordinate analysis (PCoA) of Prunus sibirica populations based on SCoT markers

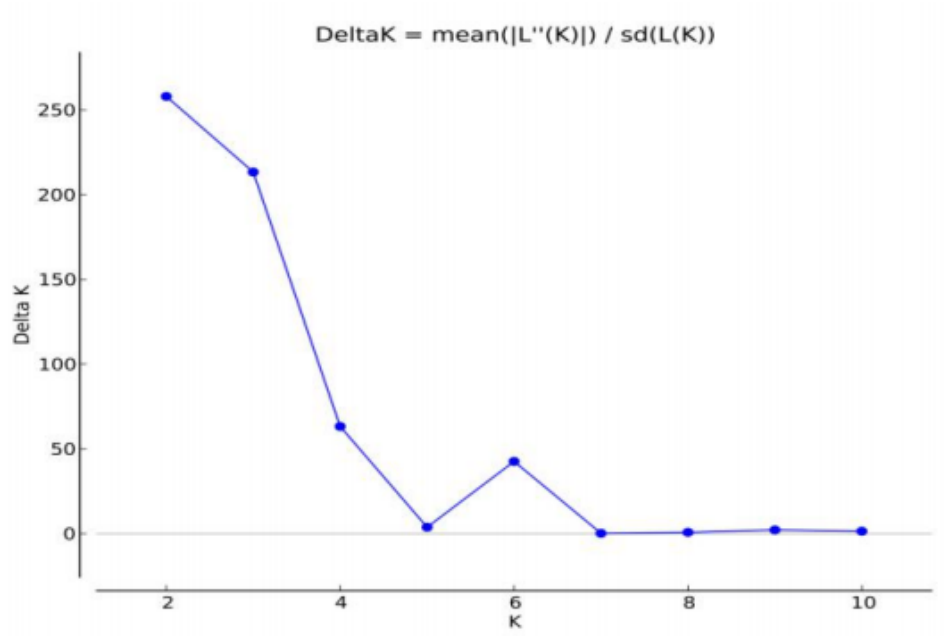

Figure 5

Evanno's test based on delta k among Prunus sibirica cultivars studies

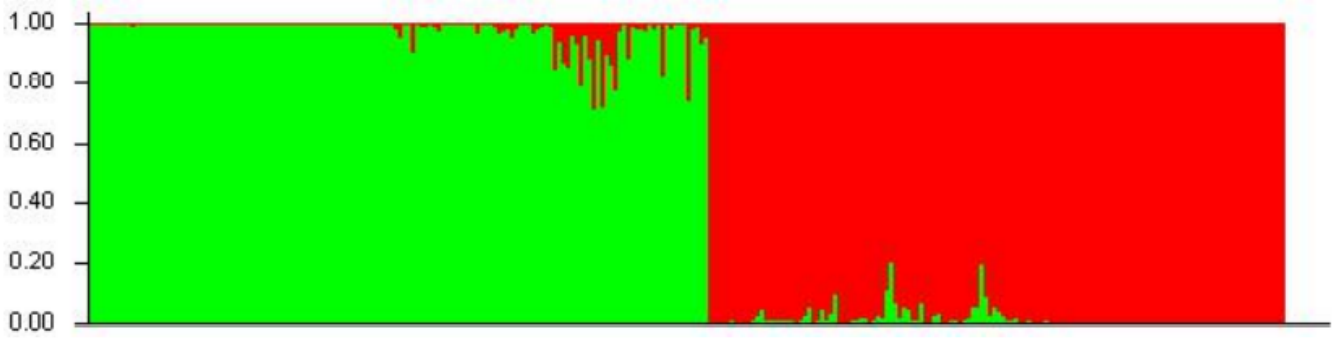

Figure 6

Population structure of the 278 individuals at K = 2. Note: The same color indicates the same group