Pak. J. Bot., 52(4): 1315-1322, 2020. DOI: http://dx.doi.org/10.30848/PJB2020-4(17)

GENETIC VARIATION AND POPULATION STRUCTURE OF *ROSA ROXBURGHII* BY EST-BASED AND GENOMIC SSR MARKERS

MIN LU[1], HUAISHAN ZHANG[1], HUAMING AN[1*] AND WEN ZHOU[2]

[1]Agricultural College, Guizhou University, Guizhou Engineering Research Center for Fruit Crops, Guiyang 550025, PR China
[2]Guizhou Hongcaijunong Investment Corporation
[*]Corresponding author's email: anhuaming@hotmail.com

Abstract

Rosa roxburghii Tratt. is a well-known commercial horticultural plant in China with medicinal and nutritional values. The wild germplasm of this rose species is distributed mainly in southwest China, but the population is decreasing due to continuous exploitation, habitat destruction, and fragmentation. Therefore, assessing its genetic variation and population structure is important for effective conservation. In this study, 494 individuals from 29 natural populations from southwest China were analyzed using 10 EST-SSRs and 10 genomic-SSRs primer pairs. One hundred and fourteen different alleles were detected, with an average of 4.1 (for EST-SSR) and 7.3 (for Genomic-SSR) alleles per locus. The overall expected and observed heterozygosity values were 0.728 and 0.852, respectively, and the Shannon's information index for each of the loci was 1.532. Our results revealed that genomic SSR markers were more polymorphic than EST-SSR markers. A moderately high level of genetic variation was maintained in the *R. roxburghii* population in Southwest China. AMOVA analyses showed that 93.10% of the genetic variation occurred within populations. Among the populations, the gene flow estimate of 3.373 had a high value. These *R. roxburghii* accessions were classified into two clusters, which did not follow the geographical distribution. Population ZJ, which showed the highest level of genetic variation and displayed a highly mixed genetic structure, was the most suitable for protection by an *in situ* conservation strategy.

Key words: *Rosa roxburghii* Tratt; Genetic variation; Population structure; Simple sequence repeat.

Introduction

Rosa roxburghii Tratt. (2n = 2x = 14), is a perennial rosebush native to China and widely distributed in the southwestern provinces, mainly in Guizhou Province and sporadically in adjacent provinces, such as Sichuan, Chongqing, Hunan, Guangxi, Hubei, and Shaanxi (Fan & Xia, 1997). Nationwide research interest was aroused because of the wide range of medicinal and nutritional components in the leaves of this plant as well as in its fruits, including ascorbate (AsA), polysaccharides, and flavonoids (Yan *et al*., 2015a; Li *et al*., 2017; Xu *et al*., 2018; Wang *et al*., 2018). A series of health care products has been developed for clinical applications and the economic cultivation area of this species in China was spread on at least 140,000 hectares until 2020.

Collection of wild germplasm for selection of elite genotypes of this plant, based on fruit characteristics and growth vigor, started in China in the early 1980s (Zhu *et al*., 1984), but to date, only one cultivar has been identified (Fan *et al*., 2011). The breeding industry and sustainable development of *R. roxburghii* are based on its abundant wild genetic resources. Any loss of genetic variation would result in a loss of potential application value for humans. However, various environmental stress factors, such as continuous exploitation, habitat destruction, and fragmentation, caused by human activities have caused substantial resource depletion. The number of populations has been decreasing, and thus conservation and protection measures are critically and urgently needed. For effective conservation, the assessment of genetic variation and population structure is imperative.

Simple sequence repeat (SSR) markers are powerful tools for genetic variation and population structure studies (Yang *et al*., 2018; Urrestarazu *et al*., 2018; Mokhtari *et al*., 2018; Hu *et al*., 2018; Guo & Gao, 2017). EST-SSR markers have been developed and applied to *R. roxburghii* genotypes by Yan *et al*., (2015a, b), Zhang *et al*., (2017), and Lu *et al*., (2017). These authors found that *R. roxburghii* had moderate to high diversity; moreover, compared with other populations, it had a more frequent gene exchange and smaller Nei 's genetic distance. Nevertheless, the statistical sampling was limited to 12 populations in Guizhou Province and no data were available from other important provinces in Southwest China. EST-SSRs were found to be superior to genomic SSRs in terms of transferability and was suitable for cross-species amplifications (Zhang *et al*., 2014). However, these researchers often amplified a lower level of polymorphism. Therefore, in the present study, we aimed to (1) analyze the genetic variations among the representatives of *R. roxburghii* populations and (2) determine the genetic structure of the populations based on a combination of EST-SSR and genomic SSR.

Materials and Methods

Population sampling: In this study, 494 individuals from 29 naturally occurring *R. roxburghii* populations were collected in the southwest of China from 2014 to 2017 (Table 1, Fig. 1). A minimum distance of 20 m between individuals was maintained to avoid sampling of similar genotypes. Healthy young leaves were collected and dried in silica gel.

Table 1. Description of the 29 sampling sites of wild *R. roxburghii.*

Population ID	Sampling No.	Sampling localities	Altitude(m)	Latitude	Longitude
AL	24	Anlong county (Guizhou province)	1258-1334	25°11'29"-25°14'5"	105°19'23"-105°29'14"
BJ	8	Bijie City (Guizhou province)	1529-1886	27°07'2"-27°18'27"	105°04'5"-105°18'56"
DF	10	Dafang County (Guizhou province)	1400-1416	27°02'59"-27°04'16"	105°41'6"-105°41'10"
DS	10	Dushan County (Guizhou province)	885-917	25°20'5"-25°27'4"	107°52'7"-107°54'11"
FQ	14	Fuquan city (Guizhou province)	957-1000	26°48'2"-26°48'48"	107°35'0"-107°36'10"
HS	20	Huishui county (Guizhou province)	946-1032	25°44'7"-26°04'35"	106°34'26"-106°59'24"
HX	36	Huaxi district (Guizhou province)	1000-1474	26°20'9"-26°41'53"	106°43'11"-106°82'56"
LD	14	Luodian county (Guizhou province)	836-934	25°37'8"-25°43'58"	106°32'12"-106°40'20"
LZ	10	Liuzhi district (Guizhou province)	1214-1638	26°11'29"-26°29'40"	105°17'16"-105°32'5"
PB	40	Pingba county (Guizhou province)	1243-1400	26°21'40"-26°41'34"	106°12'37"-106°33'36"
PX	22	Pan county (Guizhou province)	1530-1780	25°10'50"-25°54'53"	104°32'12"-104°50'14"
QL	12	Qinglong county (Guizhou province)	1499-1583	25°81'51"-25°83'29"	105°19'45"105°21'54"
QX	26	Qianxi County (Guizhou province)	1202-1264	26°57'46"-27°01'27"	106°00'2"-106°06'27"
SC	8	Shuicheng County (Guizhou province)	1698-1796	26°47'32"26°55'41"	104°52'43"-104°55'46"
SQ	12	Shiqian county (Guizhou province)	692-892	27°28'48"-27°35'30"	108°01'14"-108°12'24"
SY	20	Suiyang county (Guizhou province)	734-1405	27°53'23"-29°54'21"	106°58'51"-107°19'59"
XY	36	Xingyi city (Guizhou province)	1528-1920	24°54'36"-25°57'6"	104°43'45"-104°50'17"
ZJ	22	Zhijin County (Guizhou province)	1198-1700	26°22'7"-26°35'39"	105°41'55"-105°44'58"
ZY	24	Zunyi city (Guizhou province)	778-981	27°26'53"-27°58'53"	106°54'39"-107°18'32"
ZT	10	Zhaotong City (Yunnan province)	1380-1391	27°39'5"-27°45'14"	104°48'8"-104°55'11"
JL	10	Junlian County (Sichuan province)	1076-1202	27°51'24"-27°53'15"	104°33'13"-104°35'4"
MN	22	Mianning County (Sichuan province)	1769-1828	28°28'50"-28°35'44"	102°04'5"-102°15'28"
xy	12	Xuyong County (Sichuan province)	1050-1155	27°50'54"-28°12'33"	105°16'39"-105°39'4"
CQ	12	Chongqing City	518-540	30°15'37"-30°15'46"	107°43'2"-107°45'13"
GZ	10	Guzhang County (Hunan province)	250-267	28°40'1"-28°41'3"	109°45'0"-109°46'31"
HY	12	Huayuan County (Hunan province)	522-570	28°26'3"-28°28'0"	109°25'21"-109°29'40"
LC	14	Lichuan City (Hubei province)	963-1088	30°4'35"-30°16'54"	108°36'12"-108°56'13"
NZ	12	Nanzheng County (Shanxi province)	528-586	32°56'43"-32°59'0"	106°45'41"-106°48'27"
MX	12	Mian County (Shanxi province)	547-602	33°6'38"-33°7'27"	106°28'33"-106°31'32"

DNA extraction: A modified cetyltrimethylammonium bromide (CTAB) method (Porebski *et al*., 1997) was used to extract total genomic DNA. The DNA was dissolved in TE buffer (Tris-HCl and EDTA, pH = 8.0) and stored at -70°C for subsequent analyses.

PCR amplification and SSR detection: Twenty SSR loci (Table 2) had been developed in previous studies; ten EST-SSR loci were selected from transcriptome data (Yan *et al*., 2015a, b), and ten genomic-SSR from genome survey sequencing (Lu *et al*., 2016). PCR was conducted in a 20-μL volume containing 1 × Taq PCR Master Mix (KT201, Tiangenbiotech, Beijing, China), 50 ng of DNA template, and 0.25 μM of each primer. The following PCR amplification conditions (Table 2) were used: 3 min at 94°C; followed by 35 cycles of 94°C for 40 s, 49°C–56°C for 40 s, and 72°C for 1 min; with a final extension step for 10 min at 72°C. The PCR products were separated on 8% denaturing polyacrylamide gels stained using the silver method.

Data analyses: The statistics such as the number of different alleles (Na), number of effective alleles (Ne), observed heterozygosity (Ho), expected heterozygosity (He), Shannon's index (I), Nei's genetic variation (Nei's), and the percentage of polymorphic loci (P) were determined by POPGENE version 1.32 (Yeh *et al*., 1999). Principal coordinate analysis (PCoA), analysis of molecular variance (AMOVA) with 1,000 permutations among the populations, genetic differentiation (Fst), gene flow (Nm), and Mantel test (Mantel, 1967) with 9,999 permutations were performed using GenALEx 6.502 (Peakall & Smouse, 2012). Mantel test was implemented to identify significant association between Nei's genetic distance (Nei, 1973) and geographic distance. Based on an admixture model, population structure was assessed by STRUCTURE software v. 2.34. Models were tested ranging from 1 to 30 for K-values, with 100,000 MCMC (Markov Chain Monte Carlo) iterations and thirty independent runs each. With the best K-value selected according to the Evanno Test (Evanno *et al*., 2005; Zhang *et al*., 2014), the most likely number of clusters was chosen by plotting the LnP(D) values against Δk values.

Results

Properties of the SSR markers: In this study, 20 microsatellite regions (10 EST-SSRs and 10 genomic-SSRs) that produced clear polymorphic fragments were evaluated in the 494 studied materials. A total number of 164 alleles (Na = 8.2, Table 2) were amplified, 44 of which belonged to 10 EST-SSR loci (Na = 4.4, Table 2) and 120 to 10 genomic-SSR loci (Na = 12, Table 2. The effective number of alleles per locus (Ne) ranged from 1.921 (Primer 19) to 8.803 (Primer 023) with an average value of 3.0148 and 5.7939 for EST-SSR and genomic SSR, respectively. The value of the genetic variation, calculated according to the method of Nei's 1973 (Nei's) for EST-SSRs, ranged from 0.480 (Primer 19) to 0.755 (Primer 7) with an average value of 0.639; for genomic SSRs, it ranged from 0.707 (Primer 014) to 0.886 (Primer 023) with an average value of 0.8139. Meanwhile, the mean Shannon's index (I), expected heterozygosity (He) and observed heterozygosity (Ho) of genomic-SSRs were all much higher than for EST-SSRs (Table 2), indicating the higher polymorphism at the 10 genomic-SSRs loci.

Table 2. Information of 20 pairs of SSR primers developed from *R. roxburghii*.

Primer No.	Marker type	Primer sequences (5′→3′)	Repeat motif	Anneal temperature (℃)	Expected size (bp)	Na	Ne	I	Ho	He	Nei's
04	EST-SSR	F TTGTGGTTATAGTTCAGCCCCTA R CAATCACGGAAATCATACATTCA	CCT (3×5)	52	95	4	3.017	1.211	0.972	0.670	0.669
07		F AATATACACGAACAACAACCATCG R GGAACATGACCCCTTTTCTTATT	CTACA (5×4)	53	147	10	4.089	1.706	0.717	0.757	0.755
14		F GTGGATGTGTCAAATCTAATGGC R GGAGGGAAGAAGTAGTGAAGAACA	CT (2×6)	52.5	152	6	3.812	1.478	0.976	0.739	0.738
17		F GATGCTTTTCATTCTGCTTCAAC R ATTTTTACCGTACTCTGGGTGCT	TC (2×6)	54.3	143	2	1.938	0.677	0.822	0.485	0.484
19		F ATTTGTTTTTCGTTTTTCTTCCC R TTTGGTCATTCATTCTTCTCCTC	CAT (3×5)	51.9	128	2	1.921	0.673	0.502	0.481	0.480
24		F GAGTCATGTTGAATGATATTGGC R TTTCCTCTTTTCTTCTTTTTCCC	TGGA (4×13)	50.7	131	2	1.960	0.683	0.486	0.491	0.490
56		F AGCAATAGAGAACTAGAATATCAGAGA R TTTTCCCCCTTCAGTTGTAAGAT	GT (2×10)	52.6	116	4	3.919	1.376	0.976	0.746	0.745
58		F TGAGAACCAAAAGAAACGACTACA R GTCACAAATTCACGAGCCATATT	AT (2×8)	51.9	146	4	2.620	1.116	0.927	0.620	0.618
65		F GAGAGAGTAGCCTTTGATGAGGA R GAATGTCGTCAGGGAGCAGTAG	CTC (3×7)	55.9	146	6	3.065	1.276	0.559	0.675	0.674
70		F TGATTGGTAGTGGAAGCGATATT R ACACATCAGTCCTTCAACAGTCC	TCA (3×5)	53.4	108	4	3.807	1.361	0.571	0.739	0.737
	Average					4.4	3.015	1.156	0.751	0.640	0.639
014	Genomic-SSR	F CCTCTGGTAGTCTGGTACAACCT R ATCTCTTCCCTCTTAGCCGC	GAGAA (5×5)	55.3	184	9	3.409	1.435	0.976	0.708	0.707
016		F TTCTGTCGACCTTAGGGGTG R CAAGGCCTCTGAAACTTGGA	ATAA (4×6)	52.6	189	13	5.900	1.940	0.972	0.832	0.831
017		F ACAGCCATGATGATGAACCA R CAACAAATATCCCTCCGCAT	AAT (3×12)	49.6	234	14	7.948	2.253	1.000	0.876	0.874
019		F GTAAAATGGGGCAAAGGGAT R TCCCACCCAATCATGACATT	TTTC (4×5)	49.6	271	10	6.235	1.975	0.956	0.841	0.840
022		F AGATGGGCGATCTAGGCTTT R CTTAGAACCGCGACCCACTA	GA (2×36)	52.6	257	6	4.407	1.636	0.968	0.775	0.773
023		F TCTAAGGCGCCAGATTAAGG R TCTTTCTGCAATCGGAGGTT	TCCA (4×7)	50.6	226	15	8.803	2.370	0.773	0.888	0.886
026		F GGGCTCATCTCTAATCACGG R GCTTTCAGCTGGAATCGAAC	TTC (3×11)	52.6	212	9	4.259	1.641	0.927	0.767	0.765
027		F CTCTAACCGGTTGCAGTGCT R AGGGAAGGTTGTCATTTGTGA	CTT (3×18)	52	275	9	5.598	1.837	1.000	0.823	0.821
028		F CGGGAGATTTTGCAGGTTAG R CGGAGCCTTCAATTTCATTC	GAGCAG (6×9)	50.6	271	18	4.935	1.881	0.968	0.799	0.797
029		F GGAGGAGAATGTGGAGGTGA R CAAGACGGCATTCTTGACAT	AAT (3×11)	51.6	280	17	6.445	2.118	1.000	0.847	0.845
	Average					12	5.794	1.909	0.954	0.816	0.814
	Mean					8.2	4.404	1.532	0.852	0.728	0.726

N_a = Observed number of alleles, N_e = Effective number of alleles, I = Shannon information index, H_o = Observed heterozygosity, H_e = Expected heterozygosity, Nei's Nei's = Gene diversity

Table 3. Frequency of EST-SSRs in the *R. roxburghi*i fruit transcriptome.

Population ID	Na	Ne	I	Ho	He	Nei's	P (%)
AL	4.700	3.458	1.277	0.863	0.691	0.662	100
BJ	3.650	3.125	1.128	0.863	0.707	0.619	95
DF	3.800	3.076	1.151	0.870	0.707	0.636	100
DS	3.650	3.005	1.122	0.830	0.694	0.625	100
FQ	3.900	3.082	1.119	0.829	0.654	0.608	100
HS	4.200	3.223	1.171	0.810	0.663	0.630	100
HX	5.150	3.571	1.337	0.900	0.701	0.681	100
LD	3.700	2.940	1.103	0.886	0.664	0.616	100
LZ	3.650	2.933	1.101	0.860	0.686	0.617	100
PB	4.800	3.460	1.257	0.879	0.668	0.650	100
PX	4.850	3.613	1.282	0.888	0.686	0.657	100
QL	3.950	3.036	1.119	0.808	0.652	0.597	95
QX	5.300	3.910	1.398	0.850	0.729	0.700	100
SC	3.600	3.084	1.139	0.888	0.729	0.638	100
SQ	3.750	2.935	1.090	0.775	0.649	0.595	100
SY	4.200	3.058	1.174	0.875	0.668	0.635	100
XY	5.100	3.437	1.301	0.844	0.681	0.662	100
ZJ	5.350	3.953	1.402	0.905	0.736	0.702	100
ZY	4.900	3.403	1.296	0.900	0.700	0.671	100
ZT	3.250	2.825	0.998	0.850	0.637	0.573	95
JL	3.900	3.262	1.180	0.850	0.712	0.641	100
MN	4.500	3.178	1.218	0.846	0.676	0.646	100
xy	3.800	3.159	1.144	0.858	0.686	0.629	100
CQ	3.650	2.804	1.056	0.850	0.644	0.590	100
GZ	2.900	2.428	0.844	0.750	0.543	0.489	85
HY	2.700	2.404	0.893	0.908	0.616	0.565	100
LC	3.850	3.218	1.151	0.886	0.685	0.636	100
NZ	2.500	2.037	0.730	0.617	0.485	0.444	90
MX	2.800	2.518	0.930	0.717	0.620	0.568	100
Mean	4.002	3.108	1.142	0.843	0.668	0.620	98.62

N_a = Observed number of alleles, N_e = Effective number of alleles, I = Shannon information index, H_o = Observed heterozygosity, H_e = Expected heterozygosity, Nei's Nei's = Gene diversity, P = Percentage of polymorphic loci

Genetic variation: Twenty nine natural populations were selected to estimate the genetic variation in *R. roxburghii* around the southwest China (Table 3). The population ZJ had the highest genetic variation (Na = 5.350, Ne = 3.953, I = 1.402, He = 0.736, Nei's = 0.702) and the population NZ had the lowest diversity parameters (Na = 2.500, Ne = 2.037, I = 0.730, Ho = 0.617, He = 0.485, Nei's = 0.444). The Ho over 20 loci was highest in population HY (0.908). For all populations, all of the observed heterozygosity (Ho) was higher than expected (He), suggesting a high genetic variability in the populations due to the presence of many equally frequent alleles. The percentages of polymorphic loci were 85% in population GZ, 90% in NZ, 95% in BJ, QL and ZT, and 100% in the rest of the populations. The following average values were established: Na = 4.002, Ne = 3.108, Ho = 0.843, He = 0.668, I = 1.142, Nei's = 0.620, and P = 98.62%. The moderately high genetic variation indices supported the notion that the *R. roxburghii* population in the Southwest China still maintained a high level of genetic variation.

Genetic differentiation: Nei's genetic distance (Nei, 1979) among populations was calculated (Table 4). Out of twenty nine populations, BJ and NZ populations were the most distantly related (1.006), whereas LD and PB populations were the most closely related (0.076). Similarly, the genetic identity was the highest between LD and PB populations (0.927). The average Nei's genetic distance calculated among the 29 populations was 0.373, which was consistent with the high and steady gene flow among the populations (Table 5). The Mantel test results showed no significant correlation between Nei's genetic distance and the geographical distance among the populations (r = 0.027, P = 0.4, Fig. 2).

The AMOVA conducted by GenAlEx 6.5 software revealed that 93.10% (EST-SSR: 92.72%; nSSR: 93.40%) of all genetic differentiation was within populations, whereas 6.90% of all genetic differentiation was among populations. In the total populations, there was a low genetic differentiation (Fst = 0.069) as evidenced by the high level of gene flow estimate (Nm = 3.373) (Table 5).

Population structure: The analysis of the rate of change ΔK over the range of K-values evaluated in STRUCTURE revealed that the optimal number of genetic clusters determined equaled 2 as ΔK reached a maximum value for K = 2, established by the method of Evanno *et al*., (2005) (Fig. 3). Based on our results, we grouped the 29 populations included in this study into two clusters (Fig. 1, 4A). The QX, SC, ZJ, BJ, and DF populations in Guizhou Province together with 10 populations not in Guizhou Province were classified into one group; and the remaining 14 populations in Guizhou Province were categorized into the other group. The genetic structure and relationships were further investigated using principal coordinate analysis (PCoA). The variations expressed on axes 1 and 2 were 28.49 and 18.76%, respectively. Populations clustering by PCoA (Fig. 4B) denoted groupings consistent with STRUCTURE.

Table 4. Nei's unbiased measures of genetic identity and genetic distance based on 20 SSR markers.

Pop ID	AL	BJ	DF	DS	FQ	HS	HX	LD	LZ	PB	PX	QL	QX	SC	SQ	SY	XY	ZJ	ZY	ZT	JL	MN	xy	CQ	GZ	HY	LC	NZ	MX
AL	*	0.645	0.691	0.774	0.783	0.833	0.861	0.833	0.806	0.873	0.890	0.843	0.808	0.727	0.756	0.807	0.800	0.801	0.836	0.602	0.714	0.700	0.761	0.655	0.550	0.598	0.689	0.475	0.547
BJ	0.439	*	0.705	0.648	0.547	0.610	0.714	0.595	0.649	0.662	0.672	0.605	0.769	0.727	0.569	0.642	0.606	0.751	0.677	0.694	0.786	0.774	0.657	0.726	0.527	0.550	0.688	0.366	0.614
DF	0.370	0.350	*	0.656	0.683	0.714	0.733	0.695	0.770	0.709	0.684	0.618	0.829	0.680	0.626	0.722	0.684	0.845	0.781	0.521	0.764	0.731	0.746	0.672	0.645	0.678	0.747	0.473	0.525
DS	0.256	0.434	0.421	*	0.782	0.818	0.830	0.790	0.821	0.808	0.788	0.766	0.748	0.717	0.702	0.779	0.802	0.734	0.782	0.714	0.681	0.671	0.774	0.619	0.650	0.584	0.654	0.485	0.520
FQ	0.245	0.603	0.381	0.246	*	0.847	0.815	0.765	0.823	0.809	0.836	0.766	0.767	0.697	0.811	0.826	0.829	0.766	0.857	0.614	0.610	0.592	0.772	0.552	0.644	0.678	0.711	0.553	0.510
HS	0.182	0.494	0.337	0.201	0.166	*	0.878	0.875	0.817	0.911	0.864	0.804	0.852	0.694	0.745	0.811	0.804	0.799	0.865	0.640	0.688	0.664	0.756	0.620	0.586	0.645	0.651	0.499	0.539
HX	0.150	0.337	0.310	0.186	0.205	0.130	*	0.861	0.858	0.908	0.903	0.882	0.873	0.724	0.811	0.901	0.889	0.854	0.907	0.696	0.750	0.737	0.767	0.631	0.644	0.702	0.690	0.495	0.551
LD	0.183	0.519	0.364	0.236	0.267	0.133	0.149	*	0.792	0.927	0.829	0.809	0.811	0.603	0.721	0.786	0.820	0.831	0.816	0.556	0.687	0.710	0.750	0.607	0.628	0.639	0.628	0.424	0.462
LZ	0.216	0.432	0.261	0.198	0.195	0.202	0.153	0.233	*	0.817	0.852	0.788	0.826	0.693	0.744	0.839	0.777	0.846	0.823	0.616	0.716	0.677	0.815	0.628	0.674	0.628	0.761	0.553	0.503
PB	0.136	0.413	0.343	0.213	0.212	0.093	0.097	0.076	0.202	*	0.881	0.896	0.851	0.701	0.761	0.832	0.843	0.831	0.884	0.608	0.712	0.724	0.741	0.606	0.568	0.682	0.662	0.475	0.504
PX	0.117	0.398	0.380	0.238	0.179	0.146	0.102	0.187	0.160	0.127	*	0.836	0.836	0.692	0.784	0.852	0.816	0.839	0.866	0.627	0.701	0.670	0.742	0.630	0.549	0.665	0.676	0.555	0.582
QL	0.171	0.503	0.481	0.267	0.267	0.219	0.126	0.213	0.238	0.110	0.179	*	0.797	0.677	0.727	0.838	0.807	0.780	0.825	0.591	0.675	0.659	0.697	0.576	0.508	0.601	0.639	0.454	0.425
QX	0.214	0.263	0.187	0.291	0.265	0.160	0.136	0.209	0.191	0.162	0.179	0.227	*	0.732	0.703	0.831	0.783	0.921	0.872	0.669	0.794	0.794	0.783	0.726	0.644	0.698	0.749	0.543	0.644
SC	0.319	0.319	0.386	0.333	0.361	0.365	0.323	0.505	0.367	0.356	0.368	0.389	0.311	*	0.657	0.766	0.718	0.701	0.757	0.626	0.724	0.672	0.717	0.626	0.582	0.652	0.712	0.471	0.557
SQ	0.280	0.564	0.468	0.355	0.210	0.295	0.210	0.327	0.296	0.273	0.243	0.318	0.352	0.421	*	0.743	0.883	0.699	0.848	0.631	0.649	0.647	0.757	0.462	0.674	0.695	0.625	0.564	0.548
SY	0.215	0.444	0.325	0.250	0.191	0.209	0.105	0.240	0.176	0.183	0.160	0.176	0.185	0.267	0.298	*	0.863	0.817	0.878	0.639	0.691	0.655	0.725	0.566	0.610	0.677	0.677	0.476	0.461
XY	0.223	0.501	0.380	0.221	0.187	0.218	0.118	0.198	0.252	0.171	0.204	0.215	0.245	0.331	0.125	0.148	*	0.782	0.876	0.676	0.686	0.686	0.800	0.537	0.699	0.737	0.693	0.521	0.543
ZJ	0.222	0.286	0.168	0.309	0.267	0.225	0.158	0.185	0.167	0.185	0.176	0.249	0.083	0.356	0.358	0.202	0.246	*	0.847	0.602	0.774	0.773	0.775	0.712	0.620	0.723	0.764	0.506	0.571
ZY	0.179	0.390	0.247	0.246	0.154	0.145	0.098	0.203	0.194	0.124	0.144	0.193	0.137	0.278	0.165	0.130	0.132	0.166	*	0.657	0.732	0.706	0.767	0.597	0.687	0.755	0.735	0.565	0.568
ZT	0.507	0.365	0.652	0.338	0.487	0.447	0.362	0.587	0.485	0.498	0.466	0.526	0.402	0.469	0.460	0.448	0.391	0.507	0.420	*	0.710	0.658	0.673	0.611	0.594	0.489	0.585	0.455	0.640
JL	0.336	0.241	0.270	0.384	0.494	0.374	0.288	0.375	0.334	0.339	0.355	0.393	0.231	0.323	0.432	0.370	0.376	0.257	0.312	0.342	*	0.835	0.801	0.836	0.620	0.656	0.771	0.492	0.632
MN	0.356	0.257	0.313	0.400	0.525	0.409	0.305	0.343	0.390	0.323	0.400	0.417	0.231	0.398	0.436	0.423	0.377	0.258	0.348	0.419	0.180	*	0.782	0.732	0.721	0.657	0.694	0.421	0.553
xy	0.274	0.420	0.294	0.257	0.259	0.279	0.266	0.287	0.204	0.300	0.298	0.360	0.245	0.333	0.279	0.321	0.223	0.256	0.265	0.397	0.222	0.246	*	0.717	0.770	0.686	0.796	0.491	0.590
CQ	0.423	0.320	0.397	0.480	0.595	0.478	0.460	0.499	0.465	0.502	0.463	0.552	0.320	0.469	0.772	0.570	0.622	0.340	0.516	0.493	0.180	0.313	0.332	*	0.543	0.561	0.786	0.404	0.556
GZ	0.598	0.641	0.439	0.431	0.441	0.534	0.441	0.465	0.395	0.566	0.600	0.678	0.440	0.541	0.394	0.495	0.358	0.478	0.376	0.521	0.478	0.327	0.262	0.611	*	0.699	0.707	0.475	0.452
HY	0.514	0.598	0.388	0.539	0.389	0.439	0.353	0.447	0.466	0.383	0.408	0.509	0.360	0.428	0.364	0.391	0.305	0.324	0.282	0.716	0.422	0.420	0.376	0.579	0.359	*	0.699	0.429	0.424
LC	0.372	0.374	0.292	0.425	0.341	0.429	0.372	0.466	0.274	0.413	0.392	0.448	0.289	0.340	0.470	0.390	0.366	0.269	0.308	0.537	0.260	0.365	0.228	0.240	0.346	0.358	*	0.507	0.547
NZ	0.745	1.006	0.750	0.724	0.592	0.696	0.703	0.858	0.593	0.744	0.589	0.791	0.612	0.753	0.573	0.742	0.653	0.681	0.570	0.787	0.710	0.866	0.710	0.907	0.745	0.846	0.679	*	0.698
MX	0.603	0.489	0.644	0.655	0.673	0.618	0.596	0.772	0.687	0.685	0.541	0.857	0.440	0.586	0.602	0.775	0.610	0.560	0.565	0.446	0.459	0.592	0.527	0.586	0.793	0.859	0.603	0.360	*

Nei's genetic distance (below diagonal) and genetic identity (above diagonal)

Table 5. Analyses of molecular variation (AMOVA) of *R. roxburghii* populations based on microsatellite data.

	Source of variation	d.f.	Variance components	Percentage of variation	P	Fixtion index
EST-SSR	Among populations	28	0.294	7.28	≤0.001	Fst=0.073
	Within populations	465	3.747	92.72	≤0.001	Nm=3.184
nSSR	Among populations	28	0.342	6.60	≤0.001	Fst=0.066
	Within populations	465	4.834	93.40	≤0.001	Nm=3.538
EST-SSR and nSSR	Among populations	28	0.636	6.90	≤0.001	Fst=0.069
	Within populations	465	8.581	93.10	≤0.001	Nm=3.373

F_{ST} = Variance among coefficient of individual relative to the total variance, Nm = Gene flow

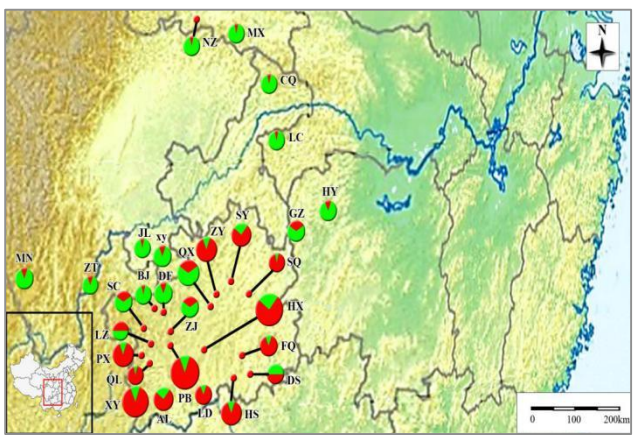

Fig. 1. Description of the 29 sampling sites distribution of wild *R. roxburghii* in Southwest China and proportion of two clusters distribution in each population (K=2; cluster l: green; cluster 2: red).

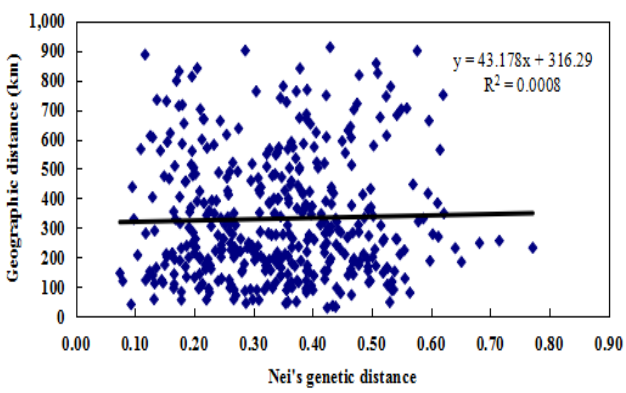

Fig. 2. The correlation between Nei's genetic distance and geographic distance for 29 populations of *R. roxburghii*.

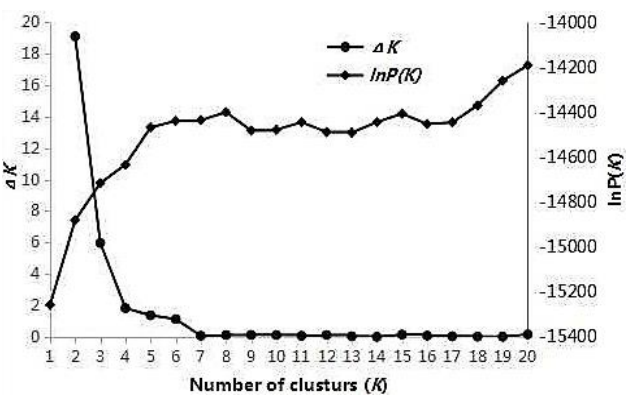

Fig. 3. Modeling of cluster number for *R. roxburghii* using STRUCTURE. Ln P(K) and Delta K were calculated in accordance with the method of Evanno *et al*., (2005).

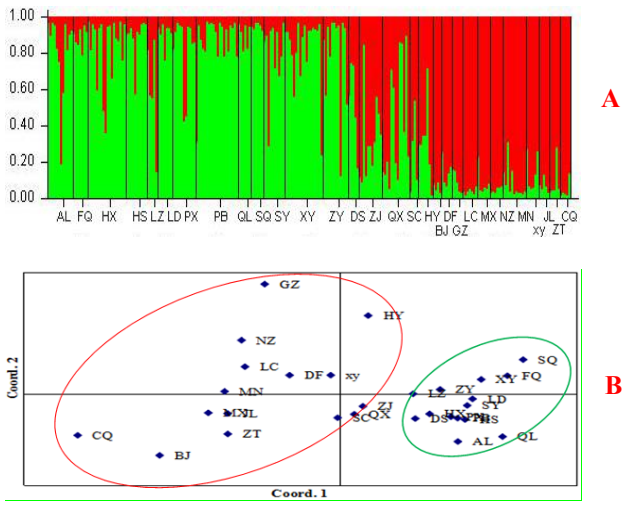

Fig. 4. Genetic structure of 29 populations of *R. roxburghii* in Southwest China. (A): Genetic relationship among the 29 populations revealed by a Bayesian modeling approach under K =2. (B): Principal coordinate analysis of the 29 populations based on the genetic distance.

Discussion

SSR (or microsatellite) markers have become a preferred choice for assessment of genetic variation of plant germplasm, especially for crops lacking whole-genome sequence information, since they are abundant across genome, highly polymorphic, relatively low cost, and easy to use. We developed 20 polymorphic *R. roxburghii* SSRs, with 10 based on genomic-SSR loci and 10 based on EST-SSR loci in the present study. EST-SSRs are expected to be less polymorphic and more conserved than genomic-SSRs, as reported in pines, barley, and adzuki bean (Liewlaksaneeyanawin *et al*., 2004; Chabane *et al*., 2005; Chen *et al*., 2015) since they are located within genes. As expected, EST-SSR regions showed polymorphism levels lower than those of the genomic-SSR regions (Table 2). Our previous studies showed that EST-SSR markers are informative molecular markers for research on the genetic diversity of *R. roxburghii* (Zhang *et al*., 2017). In the present investigation, both genomic-SSR and EST-SSR markers were used as valuable tools, which revealed moderate to high levels of polymorphism in *R. roxburghii*.

The average number of alleles (8.2) detected here was higher than those reported by Qiu *et al*., (2013), Zhou *et al*., (2011), and Qiu *et al*., (2011), who detected average numbers of alleles of 7.7, 1.2, and 7.2 using 8 *R. longicuspis*, 8 *R. sericea*, and 8 *R. odorata* populations with 14, 11, and 12 SSR markers, respectively.

Meanwhile, our average number of alleles was lower than the ones established by Meng *et al*., (2016), Kelager *et al*., (2013), and Gao *et al*., (2015), found 12.8, 16.7, and 36.3 average numbers of alleles using 27, 32, and 62 *R. odorata* var. *gigantea*, *R. rugosa*, and *R. sericea* populations and 7 SSR markers, respectively. The moderately high genetic variation supports the view that the *R. roxburghii* population in Southwest China still maintains a high level of genetic variation. In contrast to the findings of previous studies of *R. roxburghii* (Zhang *et al*., 2017), we established population genetic variations in Southwest China that was maintained at a fairly high level (Na = 4.002, Ne = 3.108, I = 1.142, Ho = 0.843, He = 0.668, Nei's = 0.620), which was higher than that of *R. roxburghii* germplasm only in Guizhou Province (Na = 3.50, Ne = 2.59, I = 1.05, He = 0.64, Nei's = 0.61). The results of the structure and PCoA analyses were generally consistent; both identified two groups of *R. roxburghii* in Southwest China. The genotypes in Guizhou Province came predominantly from one gene pool, suggesting lower polymorphism.

AMOVA analysis showed that 6.90% of the diversity was among the populations and 93.10% within the populations. According to our obtained Fst index (0.069) and Nm (3.373), these results demonstrate the low differentiation among the populations, which are also in agreement with the other studies (Kelager *et al*., 2013; Zong *et al*., 2014; Kuwi *et al*., 2018). This outcome might indicate that the activities of humans, migratory herbivores, and birds lead to germplasm exchange in different regions. The fact that wild *R. roxburghii* prefer to germinate aside roads and rivers might have largely contributed to such results. In the previous study (Zhang *et al*., 2017), the Mantel test results demonstrated a significant correlation between geographic and genetic distance among 12 populations in Guizhou Province (r = 0.2498, *P* = 0.9512). However, in this study, another seven populations were added from Guizhou Province (BJ, DF, DS, LZ, QX, SC, and ZJ), and ten populations from other provinces in Southwest China (Table 1, Fig. 1). Thus, the Mantel test showed no significant correlation between the geographical distance and Nei's genetic distance among the populations (r = 0.027, *P* = 0.4, Fig. 2). The change detected might have come from the second gene pool and may be due to the discontinuity of the sample (Fig. 1).

Population ZJ was the most suitable for protection under *in situ* conservation strategy. First, population ZJ had the highest level of genetic variation (Table 3). Second, this population represented the two gene pools, which displayed a highly mixed genetic structure (Fig. 1, 4A).

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (Grant No. 31660558), the Science and Technology Department of Guizhou Province (Project No. 20201Y113, 20164016 and 20175788), and the Construction Program of Biology First-class Discipline in Guizhou (GNYL2017009).

References

Chabane, K., G.A. Ablett, G.M. Cordeiro, J. Valkoun and R.J. Henry. 2005. EST versus genomic derived microsatellite markers for genotyping wild and cultivated barley. *Genet. Resour. Crop Evol.*, 52: 903-909.

Chen, H., L. Liu, L. Wang, S. Wang, M.L. Wang and X. Cheng. 2015. Development of SSR markers and assessment of genetic diversity of adzuki bean in the chinese germplasm collection. *Mol. Breeding*, 35: 191.

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

Fan W.G. and G.L. Xia. 1997. Utilization of *Rosa roxburghii* resources and its developing strategy in Guizhou province. *Southwest China J. Agric. Sci.*, 10: 109-115.

Fan, W.G., X.H. Xiang, H.M. An and J.P. Liu. 2011. A new *Rosa roxburghii* cultivar 'Guinong 5'. *Acta Hortic. Sin.*, 38: 1609-1610.

Gao, Y.D., Y. Zhang, X.F. Gao and Z.M. Zhu. 2015. Pleistocene glaciations, demographic expansion and subsequent isolation promoted morphological heterogeneity: a phylogeographic study of the alpine *Rosa sericea* complex (Rosaceae). *Scientific Reports*, 5: 11698.

Guo, L.N. and X.F. Gao. 2017. Genetic diversity and population structure of *Indigofera szechuensis* complex (Fabaceae) based on EST-SSR markers. *Gene*, 624: 26.

Hu, X., P.G. Zheng, B.Y. Ni, X.J. Miao, Z. Zhao and M. Li. 2018. Population genetic diversity and structure analysis of wild apricot (*Prunus armeniaca* L.) revealed by SSR markers in the Tien-Shan Mountains of China. *Pak. J. Bot.*, 50: 609-615.

Kelager, A., J.S. Pedersen and H.H. Bruun. 2013. Multiple introductions and no loss of genetic diversity: invasion history of Japanese rose, *Rosa rugosa*, in Europe. *Biol. Invas.*, 15: 1125-1141.

Kuwi, S.O., M. Kyalo, C.K. Mutai, A. Mwilawa, J. Hanson, A. Djikeng and S.R. Ghimire. 2018. Genetic diversity and population structure of *Urochloa* grass accessions from Tanzania using simple sequence repeat (SSR) markers. *Braz. J. Bot.*, 41: 699-709.

Li, L.L., M. Lu and H. An. 2017. Expression profiles of the genes involved in L -ascorbic acid biosynthesis and recycling in *Rosa roxburghii*, leaves of various ages. *Acta Physiol. Plantar.*, 39: 44.

Lu, M., H. An and L. Li. 2016. Genome survey sequencing for the characterization of the genetic background of *Rosa roxburghii* Tratt and leaf ascorbate metabolism genes. *Plos One*, 11: e0147530.

Lu, M., X. Yan, J. Bai, H. Zhang, D. Wang and H. An. 2017. Construction of core collection in wild *Rosa roxburghii* from Guizhou province using EST-SSR markers and fruits quality traits. *Acta Hort. Sin.*, 44: 1486-1495.

Liewlaksaneeyanawin, C., C.E. Ritland, Y.A. El-Kassaby and K. Ritland. 2004. Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. *Theor. Appl. Genet.*, 109: 361-369.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209-220.

Meng, J., S.L. He, D.Z. Li and T.S. Yi. 2016. Nuclear genetic variation of *Rosa odorata*, var. *gigantea*, (Rosaceae): population structure and conservation implications. *Tree Genet. Genom.*, 12: 65.

Mokhtari, N., B.E. Sayed-Tabatabaei, M. Bahar and H. Arabnezhad. 2018. Assessment of genetic diversity and population genetic structure of carthamus, species and Iranian cultivar collection using developed SSR markers. *J. Genet.*, 97: 1-12.

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

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.

Peakall, R. and P.E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics*, 28: 2537-2539.

Porebski, S., L.G. Bailey and B.R. Baum. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol. Biol. Rep.*, 15: 8-15.

Qiu, X.Q., K.X.Tang, H.Y. Jian, Q.G. Wang, S.F. Li, Z. Shao and H. Zhang. 2011. SSR markers based genetic diversity of populations of *Rosa odorata* sweet var.*gigantea* in Yunnan province. *J. Huazhong Agric. Univ.*, 30: 300-304.

Qiu, X.Q., K. Tang, Q.Wang, H. Jian, S. Li, Z. Shao and H. Zhang. 2013. Genetic diversity analysis of populations of *Rosa longicuspis*. *J. Plant Genet. Res.*, 14: 86-91.

Urrestarazu, J., P. Errea, C. Miranda, L.G. Santesteban and A. Pina. 2018. Genetic diversity of Spanish *Prunus domestica* L. germplasm reveals a complex genetic structure underlying. *Plos One*, 13(4): e0195591.

Wang, L., B. Zhang, J. Xiao, Q. Huang, C. Li and X. Fu. 2018. Physicochemical, functional, and biological properties of water-soluble polysaccharides from *Rosa roxburghii* Tratt fruit. *Food Chem.*, 249: 127.

Xu, S.J., F. Zhang, L.J. Wang, M.H. Hao, X.J. Yang, N.N. Li, H.L. Ji and P. Xu. 2018. Flavonoids of *Rosa roxburghii* Tratt offers protection against radiation induced apoptosis and inflammation in mouse thymus. *Apoptosis*, 23: 1-14.

Yan, X., X. Zhang, M. Lu, Y. He and H. An. 2015a. *De novo* sequencing analysis of the *Rosa roxburghii*, fruit transcriptome reveals putative ascorbate biosynthetic genes and EST-SSR markers. *Gene*, 561: 54-62.

Yan, X., M. Lu and H. An. 2015b. Analysis on SSR information in transcriptome and development of molecular markers in *Rosa roxburghii*. *Acta Hort. Sin.*, 42: 341-349.

Yang, X., Z. Yang and H. Li. 2018. Genetic diversity, population genetic structure and protection strategies for *Houpoëa officinalis*, (Magnoliaceae), an endangered Chinese medical plant. *J. Plant Biol.*, 61: 159-168.

Yeh, F.C., R.C. Yang, T.B. Boyle, Z. Ye and J.X. Mao. 1999. POPGENE version 1.32, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada.

Zhang, H.S., X.Q. Yan, M. Lu, D.P. Wang and H.M. An. 2017. Analysis of the genetic diversity of wild *Rosa roxburghii* populations in Guizhou Province based on EST-SSR marker. *Sci. Agri. Sin.*, 50: 1098-1108.

Zhang, M., W. Mao, G. Zhang and F. Wu. 2014. Development and characterization of polymorphic EST-SSR and genomic SSR markers for Tibetan annual wild barley. *PLoS One*, 9: e94881.

Zhou, N. N., H. Zhang, C.M. Yang, G.X. Wang, S.C. Li, Q.L. Shan and K.X Tang. 2011. Genetic diversity analyses of *Rosa sericea* Lindl population based on SSR markers. *Southwest China J. Agric. Scien.*, 24: 1899-1903.

Zhu, W.F., X.H. Xiang, S.X. Yang, Q.Q. Mo and X.F. Gao. 1984. Investigation of chestnut rose germplasms, growth and development, and vitamin C content in deferent ecosystems of Guizhou Province (China). *J. Guizhou Agri. College*, 3: 1-13.

Zong, Y., P. Sun, L. Jing, X. Yue, K. Li and Y. Teng. 2014. Genetic diversity and population structure of seedling populations of *Pyrus pashia*. *Plant Mol. Biol. Rep.*, 32: 644-651.

(Received for publication 26 October 2018)