Indian J. Genet., 79(3) 583-593 (2019)
DOI: 10.31742/IJGPB.79.3.8

Understanding genetic diversity, structure and population differentiation in selected wild species and cultivated Indian and exotic rose varieties based on microsatellite allele frequencies

Aparna Veluru*, K. V. Bhat[1], T. Janakiram[2], K. V. Prasad[3], D. V. S. Raju[3], C. Bharadwaj[4], Gayacharan[1], Kanwar P. Singh, Namita and Sapna Panwar

ICAR-Indian Agricultural Research Institute, New Delhi 110 012; [1]ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi 110 012; [2]Assistant Director General, Horticulture, KAB-II, Indian Council of Agricultural Research, New Delhi 110 012; [3]ICAR-Directorate of Floricultural Research, Pune 411 005, Maharashtra; [4]Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi 110 012

(Received: April 2019; Revised: July 2019; Accepted: August 2019)

Abstract

Roses are the most important commercial ornamental plants grown for flowers, perfumery and nutraceutical compounds. Commercially cultivated roses (*Rosa* × *hybrida* L.) are complex interspecific hybrids probably derived from 8-10 wild species among the large diversity of 130-200 species in genus *Rosa*. Wild germplasm is a primary source of variability and plays a major role in improving existing varieties by broadening their genetic base. In the present investigation, we have utilized the previously identified SSR primers for studying the diversity among 148 selected rose genotypes, including wild species and cultivated varieties of Indian and exotic origin. A total of 88 alleles was scored using 30 polymorphic loci; they produced average 2.9±1 alleles per locus. Polymorphism information content (PIC) values for different SSR loci ranged from 0.08 to 0.8 with a mean value of 0.5±0.2. The neighbor-joining tree generated based on Nei's (1978) genetic distance values grouped the population into three major clusters. Cluster-I & II consists of all modern rose cultivars (*Rosa* × *hybrida* L.) originated from India and cluster-III consists of all exotic cultivars, wild species and a few cultivars from India. STRUCTURE analysis based on microsatellite allelic data, partitioned the total rose genotypes into four different sub-populations with some individual genotypes having genomic admixture. Population subdivision estimates, F_{ST} between different subpopulations ranged from 0.01-0.15 indicates low to moderate level of divergence existing among the rose cultivars and germplasm. Population differentiation in rose cultivars and wild species corresponds to their geographical origin and lineages. Analysis of molecular variance (AMOVA) results revealed that 83.12 % of the variance was accounted for by within sub-groups followed by significant levels of variation among the populations (10.42%) and least variance (6.46%) was noticed among individuals within groups.

Key words: Rose, Indian roses, wild species, modern cultivars, population structure, genetic diversity, microsatellite markers

Introduction

Rose is commercially important ornamental plant and belongs to family Rosaceae and genus *Rosa*. The genus *Rosa* contains 130-200 wild species and more than 24,000 cultivated varieties (Roberts et al. 2003). Most of the wild rose species are native to Asia and Europe and few of them are also found in North America and North Africa (Gudin 2000; Wissemann 2003). In India, more than 25 wild species (Tejaswini and Prakash 2005) of *Rosa* are found. Ploidy level in roses varies from 2n=2x=14 to 2n=10x=70 and the basic chromosome number X=7 (Hurst 1925; Roberts et al. 2008; Jian et al. 2010). Even though a number of large species exists in genus *Rosa*, only 8-10 species have contributed largely for development of modern rose cultivars (Crespel and Mouchotte 2003; Zlesak 2006) and there exists immense scope for improvement of cultivated types due to huge variability in the genus. Characterization, variability and genetic relationships existing between the varieties and wild species will help to identify the genetic differences and breeding behavior of the genotypes and information generated will be useful for further improvement of existing varieties.

The characterization and diversity of the genotypes can be studied using a variety of markers including physiological, biochemical and genetic markers. A marker based on genomic information

*Corresponding author's e-mail: aparna.cpcri@gmail.com
Published by the Indian Society of Genetics & Plant Breeding, A-Block, F2, First Floor, NASC Complex, IARI P.O., Pusa Campus, New Delhi 110 012; Online management by www.isgpb.org; indianjournals.com

provides more reliable information with high accuracy. Rose is one of the highly studied ornamental crops due to its high economic importance. Genome sequence information was also available in some of the rose cultivars and species belonging to genus *Rosa* (Foucher et al. 2015; Lu et al. 2016; Nakamura et al. 2018). Previous authors had used different marker systems such as Random Amplified Polymorphic (RAPD), Amplified Fragment Length Polymorphism (AFLP), Restricted Fragment Length Polymorphism (RFLP), Inter Simple Sequence Repeats-ISSR, Simple Sequence Repeats (SSR) and Sequence Tagged Microsatellite Site (STMS) for analyzing genetic diversity and characterization of cultivated and wild rose genotypes belongs to particular geographical origin (Wu et al. 2000; De Cock et al. 2008; Samiei et al. 2010; Azeem et al. 2012; Yang et al. 2017) or specific group or subclass of roses (Gardes et al. 2005; Scariot et al. 2006; Kiani et al. 2010; Alsemaan et al. 2011; Akond et al. 2012; Jiang and Zang, 2017). Very few studies have been conducted to examine the genetic diversity in Indian roses (Rai et al. 2015; Panwar et al. 2015; Prasad et al. 2006; Mohapatra and Rout 2005) and investigations comparing the cultivated Indian rose population with native wild roses and exotic cultivated types are rather scarce. The present experiment was therefore, planned to analyse the diversity, genetic structure and level of population differentiation in a large number of Indian cultivated types with respect to the exotic cultivars and wild native species.

Materials and methods

Plant material

In the present investigation, 148 rose genotypes were used including 109 Indian bred modern roses, 20 wild species found in India and 19 exotic modern cultivars (Table 1). These genotypes were collected from the rose garden at the Division of Floriculture and Landscape Architecture, ICAR-Indian Agricultural Research Institute, New Delhi, India.

SSR marker analysis

The total genomic DNA was extracted from freshly collected young leaves of selected genotypes using the CTAB method (Murray and Thompson 1980) with minor modifications. A total of 50 SSRs were selected from the literature and synthesized for analysis (Yan et al. 2005; Kimura et al. 2006; Jowkar et al. 2009). Out of those, 30 markers exhibited polymorphisms among all selected genotypes. Primer amplification was carried out in 25 µL reaction volume consists 2 µL of template DNA (20 ng/µL), 2 µL of F/R primer mix (15 pM), 0.33 µL of Taq DNA polymerase (5 U/µL) (Genei, India) 2.5 µL Taq DNA buffer (10X), 2µL of dNTPs (20 mM) and 16.17 µL sterile water. The PCR temperature profile was 94^{o}C for 6 min, followed by 35 cycles of 94^{o}C for 1 min, annealing at 48^{o}C to 55^{o}C (depend upon primer annealing temperature) for 1 min and 72^{o}C for 2 min and a final extension at 72^{o}C for 10 min before cooling it to 4^{o}C. Amplified products were resolved in 3% agarose gel using 1X TAE buffer. A 100-bp DNA ladder used as a size standard and Alpha Ease FC 4.0 software (Yeh, 1999) was used for scoring the amplified bands.

Statistical analysis

The polymorphic information content (PIC) for individual loci were calculated as described by Botstein et al. (1980) and modified by Anderson et al. (1993). The PIC formula followed is $PIC = 1 - \sum_{j=1}^{n} P_{ij}^2$, where, P_{ij} is the frequency of the j^{th} allele for the i^{th} marker, and summed over n alleles. The PIC value indicates the genetic variation and also discriminatory power of a marker. Total SSR genotypic data was used for creating weighted neighbour joining tree based on simple matching (SM) dissimilarity matrix using DARwin version 6.0.10 software (Perrier and Jacquemoud-Collet, 2006).

The multi-locus SSR genotyping data generated for all rose genotypes was analyzed using STRUCTURE 2.3.4 program (Prichard et al. 2000). The program assigns the individual genotypes to distinct populations and also illustrates hybrid zones, migrants and admixed individuals. To obtain the optimal number of genetic clusters (best K) (number of distinct populations, the method proposed by Evanno et al. (2005) was used. STRUCTURE program was run for 10 replications of each K (2-15). A burn-in period length of 10,000 with 100,000 Markov Chain Monte Carlo (MCMC) replications after burn-in was used. The K value obtained was 4, in which rose genotypes from Indian origin were divided into two distinct populations while exotic and rose species grouped into distinct populations. Based on this information individual genotypes were assigned to their respective populations for downstream analyses. Various genetic diversity parameters among and within populations were estimated using PopGene Version 1.32 software (Yeh et al. 2000). The Nei's (1972, 1978) genetic distance matrix obtained was used for generating

neighbor-joining tree using NTSYSpc 2.11V software (Rohlf, 2000). The SSR genotypic data with individuals designated to four different populations was further used for Analysis of Molecular Variation (AMOVA) using Arlequin 31 software (Exoffier et al. 2005). AMOVA partitions total variation into sources of variation based on gene frequencies, taking into account the number of mutational differences between molecular haplotypes. Fixation indices (Weir and Cockerham 1984) and pair-wise F_{ST}'s to depict short-term genetic distances between populations were also computed. These F_{ST}'s are transformed to linearize the distance with population divergence time (Reynolds et al. 1983; Slatkin 1995). Ewens-Watterson test for neutrality of microsatellite markers was also performed (Manly, 1985). R ver. 3.5.1 (Rcmd function) embedded with Arlequin was used to generate graphical illustrations of the results.

Results

A total of 88 alleles were identified among 148 rose genotypes using 30 polymorphic SSR loci. Polymorphic Information Content (PIC) values for SSRs used, varied from 0.08 (Rh50) to 0.8 (Rh58) with a mean value of 0.5±0.2. Number of alleles (na) in different loci ranged from 2 to 6 with an average value of 2.9±1 alleles per locus. At the same time effective number of alleles (ne) in different loci varied from 1 (Rh50) to 3.8 (Rh79, Rh58) with an average of 2.3±0.9 per locus. The 'ne' value indicates the number of equally frequent alleles required to achieve the same expected heterozygosity as in the studied population. Average heterozygosity (H_{avg}) value for different loci ranged from very low 0.07 (Rh50) to very high 0.72 (Rh58) with the mean value of 0.5±0.2 per locus. The gene diversity (He) indicated the high level of variability among the individuals from each group of population. For the exotic rose cultivars and wild species, differences between observed heterozygosity (H_o) and expected heterozygosity (H_e) were high, indicating that the allelic frequencies are not in Hardy Weinberg Equilibrium (H.W.E) (Table 2). This might be due to high level of inbreeding attributing to isolation, vegetative propagation, geographical distances or interspecific crossing barriers.

Similarly, many individual loci were found in the disequilibrium state. Highest and lowest gene flow (Nm) values for different loci were 67.9 (ABRII/Rpu4) and 0.5 (ABRII/Rpu4) and mean Nm value identified for 30 SSR was 2.1 (Table 3). The microsatellite loci used in this study were also tested for understanding their segregation behavior (random or non-random). Ewens-Watterson test (based on the infinite allele's model) identified five non-neutral microsatellite markers (Rh79, Rh58, Rh77, RhAB28 and ABRII/Rpu11). F value (sum of squares of allelic frequency) of these non-neutral markers lied outside upper and lower limit of 95% confidence region (Table 4) this indicate that the loci may be linked to traits where active selection is operating, however, this might be also the result of population bottleneck or this effect may have come due to species specific phenotypic differences. Further joint distribution of *FST* and (heterozygosity within populations)/(1- *FST*) resulted in the identification of loci under selection pressure at 1% confidence level (i.e. Rh80, ABRII/Rpu7 and ABRII/Rpu11) (Fig. 1).

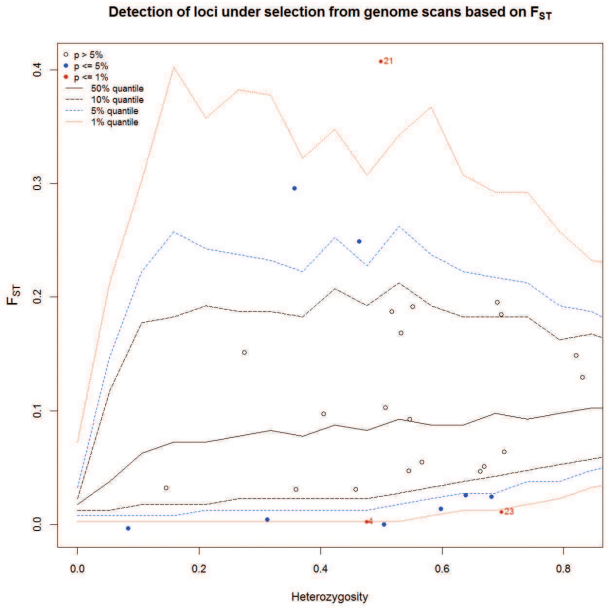

Fig. 1. Joint distribution of F_{ST} and heterozygosity for the detection of loci potentially under selection. F_{ST} are and (heterozygosity within the population)/ (1- F_{ST}) are plotted for each loci. Dashed lines are one-sided confidence interval limits obtained from simulated data. Loci under selection at 5% confidence level are shown as filled blue circles; loci under selection at 1% level are shown as red filled circle. The loci Rh80 (4), ABRII/Rpu7 (21) and ABRII/Rpu11 (23) were found significant at 1% level

Neighbour-Joining tree constructed based on Nei's genetic distance values grouped the all characterized genotypes into three major clusters (Fig. 2). Major cluster-I & II consist of all modern cultivated varieties of Indian origin, whereas major cluster-III was separated with wild species, exotic cultivars and some Indian origin cultivars. Exotic types and wild species

Table 1. List of rose genotypes used for characterization and diversity analysis

ID	Cultivar Name	Pop. ID*	ID	Cultivar Name	Pop. ID*	ID	Cultivar Name	Pop. ID*
Indian cultivars				Ranjana	1	101	Suchitra	3
1	Abhisarika	1	52	Ratnaar	1	102	Surdas	1
2	Anurag	3	53	Sahasradhara	3	103	Suryakiran	1
3	Arjun	1	54	Shanti Pal	1	104	Suryodaya	3
4	Arka Parimala	1	55	Shreyasi	1	105	Tarang	3
5	Aruna	1	56	Sir C. V. Raman	2	106	Pusa Veerangana	2
6	Ashwini	1	57	Soma	2	107	Pusa Urmil	2
7	Bhim	1	58	Sugandha	2	108	Pusa Muskan	2
8	Century Two Seedling	1	59	Surabhi	2	109	Rose Sherbet	1
9	Chambe-di-Kali	1	60	Surekha	2		**Exotic cultivars**	
10	Chitra	1	61	Surkhab	2	110	Blue Moon	3
11	Dil-Ki-Rani	1	62	Jawahar	2	111	Bonne Nuit	3
12	Dr B. P. Pal	1	63	Shiloz Mukherjee	2	112	Brandy	3
13	Dr Benjamin Pal	1	64	Jawani	2	113	Century Two	3
14	Dr Bharat Ram	1	65	Indian Princess	2	114	Christian Dior	3
15	Dr M.S. Randhawa	1	66	Akash Sundari	2	115	Double Delight	3
16	Dr R. R. Pal	1	67	Delhi White Puff	2	116	Eiffel tower	3
17	Dulhan	1	68	Delhi Pink Puff	2	117	Elle	3
18	Eiffel Tower × Queen Elizabeth	1	69	Anitha	2	118	Fragrant Lace	3
19	Ganga	1	70	Arunima	2	119	Fragrant Plum	3
20	Golden Afternoon	1	71	Banjaran	2	120	Karen Blixen	3
21	Haseena	1	72	Chingari	2	121	Melody Perfume	3
22	Homage	1	73	Deepak	1	122	Memorial day	3
23	Lalima	1	74	Delhi Brightness	1	123	Midas touch	3
24	Lalmakhmal	1	75	Delhi Princess	1	124	Oklahoma	3
25	Madhosh	1	76	Dr. S.S. Bhatnagar	3	125	Papa Meilland	3
26	Maharana	1	77	Himangini	2	126	Perfume de French Comete	3
27	Mother Teresa	1	78	Jantar Mantar	1	127	Sweet Afton	3
28	Mridula	1	79	Krishna	2	128	Velvet Fragrance	3
29	Mrinalini	1	80	Lahar	2	**Wild Species**		
30	Ms K.B. Sharma	1	81	Loree	2	129	*R. indica*× *R. nitida*	4
31	Nayika	1	82	Madhura	2	130	*R. slancensis*	4
32	Nehru Centenary	1	83	Manmatha	2	131	*R. indica major*	4
33	Nurjahan	1	84	Manasi		132	*R. brunonii*	4
34	Pink Montezuma	1	85	Mohini	1	133	*R. macrophylla*	4
35	Preyasi	1	86	Navneet	1	134	*R. wichuraiana*	4
36	Priyadharshini	1	87	Neelambari	2	135	*R. moschata*	4
37	Pusa Ajay	1	88	Prema	2	136	*R. dumalis*	4
38	Pusa Aruna	1	89	Punchu	2	137	*R. multiflora*	4
39	Pusa Bahadur	2	90	Pusa Abhishek	2	138	*R. lutea*	4
40	Pusa Garima	1	91	Pusa Baramasi	2	139	*R. glutinosa*	4
41	Pusa Mansij	1	92	Pusa Gaurav	2	140	*Rose* spp (Nepal)	4
42	Pusa Mohit	1	93	Pusa Komal	2	141	Himroz	4
43	Pusa Priya	1	94	Pusa Manhar	2	142	Jwala	4
44	Pusa Shatabdi	1	95	Pusa Pitambar	2	143	*R. bourboniana*	4
45	Pusa Sonara	1	96	Rupali	2	144	*R. chinensis viridiflora*	4
46	Raja Ram Mohan Roy	1	97	Sabnam	2	145	Dr Huey	4
47	Raja S. S. Nalagarh	1	98	Sadabahar	1	146	Kakinada rose	4
48	Raj Kumari	1	99	Shola	3	147	Ranisahiba	4
49	Raktagandha	1	100	Sindhoor	3	148	*R. tomentosa*	4
50	Raktima	1						

ID = Identity; *Population ID is given to genotypes based on the structure analysis

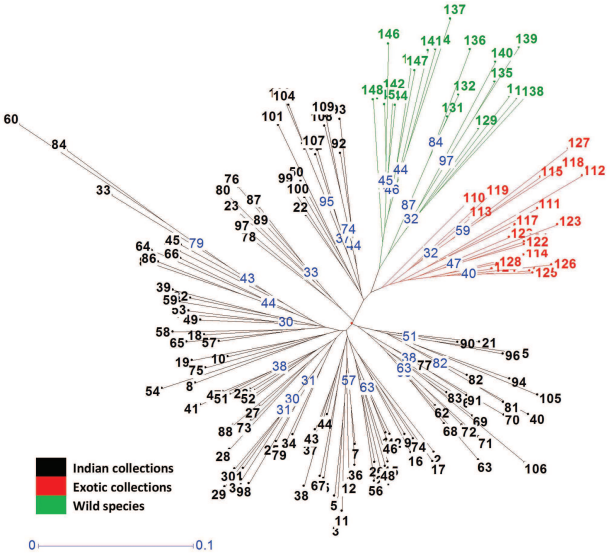

Fig. 2. Neighbour-joining tree of 148 rose genotypes: based on weighted function and simple matching coefficient dissimilarity matrix. The numerical values representing the different genotypes are the cultivar ID given in the Table 1. Numerical values in blue colour represent the bootstrap value of the respective nodes

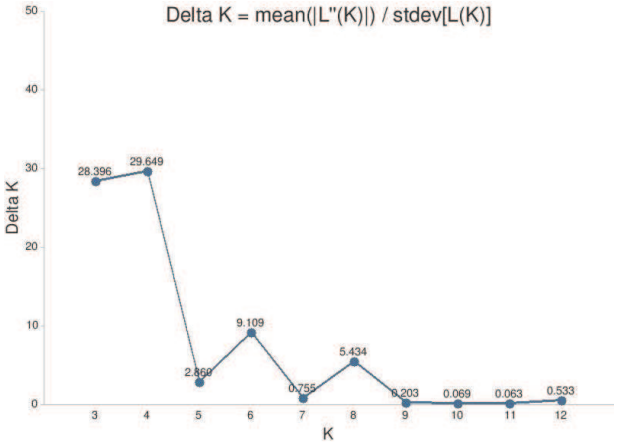

Fig. 3. Best K was determined by Evanno et al. 2005 method. K value is plotted against delta K. K value at the highest delta K is selected as best K which is 4

different clusters of dendrogram which displayed the genetic similarities and dissimilarities among them (Fig. 1) and disclosed the genetic relation of each genotype with another.

A Bayesian model approach for population

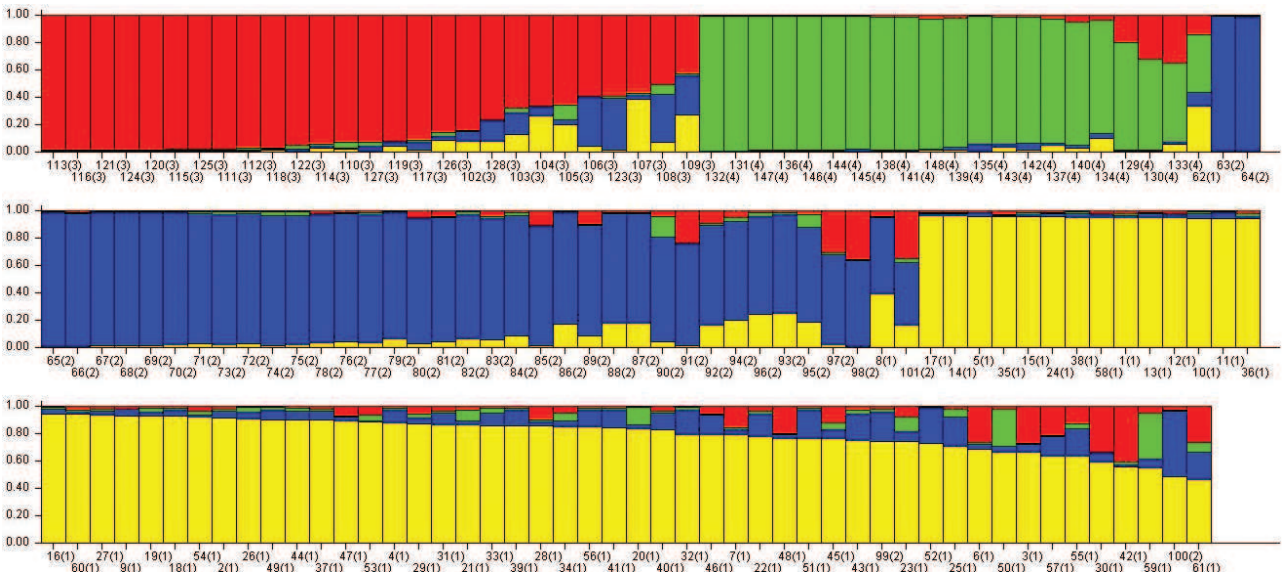

Fig. 4. Bar diagram for 148 rose accessions arranged based on inferred ancestry at K=4; Colour codes represents the four subpopulations. Values in the left indicate the membership coefficient (Q). Proportions of colours in each bar indicate the allelic affiliation with respect to the sub-populations. The numerical values representing the different genotypes are the cultivar ID given in Table 1. Numbers in parenthesis indicate the designated population for the genotype based on the structure analysis

were further separated into different subclusters of major cluster-III. The dendrogram generated based on weighted function and simple matching dissimilarity matrix clearly illustrated the 148 rose genotypes in

structure revealed the highest log likelihood value at K=4 (Fig. 3) indicated four distinct genetic groups or subpopulations, where Indian cultivars were separated into two distinct subpopulations (subpopulations-I and

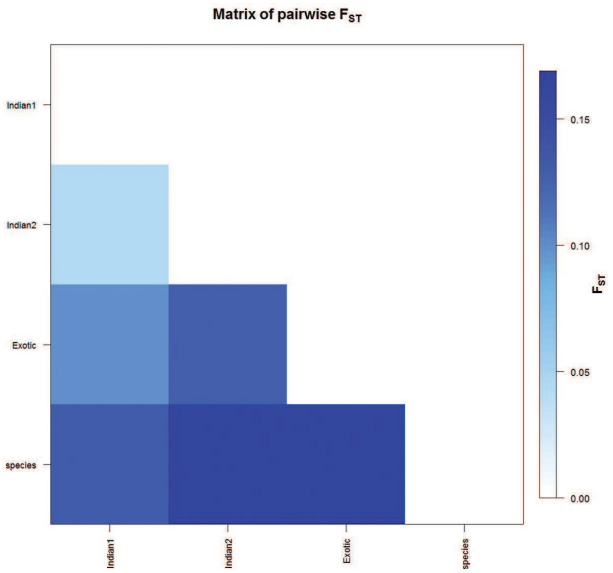

Fig. 5. Matrix of Pair wise F_{ST} of four identified subpopulations of roses

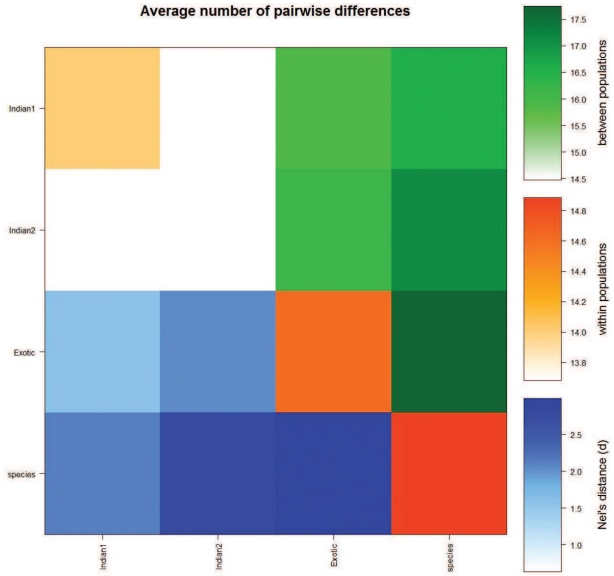

Fig. 6. Matrix for different genetic distance indices for an average number of pair wise differences (ð) within and between populations. Orange on diagonal represents ð within populations; green above diagonal represent $ð_{xy}$ between pairs of populations; blue below diagonal indicates a net number of allelic differences between populations

II). Subpopulation-III consists of all exotic cultivars along with five Indian cultivars. Wild species were separated from all cultivated types and identified separately in subpopulation-IV (Table 1). Population structure of 148 rose accessions based on their inferred ancestry was shown in Fig. 4. Matrix (Fig. 6) for the average number of pair wise differences (π) within and between populations, displayed the highest variability within wild species followed by exotic collections whereas lowest variability was noticed in subpopulation-II. Variability between different population groups was concerned, higher genetic variability and allelic differences were noticed between exotic collections and wild rose species followed by wild species and Indian population group-II and least differences were noticed between population groups consists Indian cultivars (I&II). The results were further supported by genetic variability parameters calculated among four subpopulations (Table 2).

Analysis of Molecular Variance (AMOVA) was used to study genetic differentiation among populations. The analysis allocates the total variability into the sources of variation i.e. population subdivisions. The maximum variance was identified within individuals (83.12%) followed by variation among the populations (10.42%) and least variance (6.46%) was identified among individuals within populations (Table 5). The AMOVA result indicated that the population differentiation was moderate as the maximum genetic variation lies within the individuals rather than between populations. Another measure of population differentiation due to genetic structure i.e. F_{ST} (fixation index) also showed the moderate level of differentiation between the populations of rose cultivars including wild species. Pair-wise genetic differentiation

Table 2. Summary of mean values for genic variation and heterozygosity parameters in different subpopulations

Population type	na	ne	I	Ho	He	Nei
Indian Population-I	2.8±0.8	2.07±0.7	0.77±0.32	0.46±0.28	0.53±0.18	0.46±0.18
Indian population-II	2.9±1.09	2.04±0.7	0.75±0.35	0.46±0.29	0.46±0.2	0.42±0.19
Exotic cultivars	2.8±0.8	2.12±0.74	0.79±0.29	0.40±0.25	0.49±0.16	0.48±0.16
Wild species	2.9±1.05	2.27±0.88	0.83±0.39	0.37±0.26	0.49±0.22	0.48±0.21

Na=observed number of alleles; ne=effective number of alleles, I=Shannon's Information Index; Ho=Observed Heterozygosity; He= Expected Heterozygosity; Nei= Nei's Expected heterozygosity

Table 3. Details of Polymorphic SSR markers and their descriptive statistical parameters for overall rose population

Locus	PIC	na	Ne	I	Ho	He	Nei	Havg	Nm*
Rh79	0.79	4	3.8	1.35	0.25	0.75	0.73	0.67	1.4
RhAB9	0.66	3	2.8	1.06	0.3	0.66	0.64	0.57	1.6
Rh48	0.52	6	3.5	1.44	0.70	0.74	0.72	0.47	1.0
Rh80	0.47	2	1.9	0.66	0.35	0.48	0.47	0.46	28.3
Rh96	0.51	3	2.4	0.98	0.20	0.59	0.58	0.43	1.0
Rh50	0.08	1	10	0.00	0.00	0.00	0.00	0.07	11.4
Rh58	0.80	4	3.8	1.35	0.75	0.75	0.73	0.72	2.0
RhABT12	0.45	3	20	0.77	0.40	0.51	0.5	0.44	6.5
Rh65	0.33	2	1.3	0.42	0.00	0.26	0.26	0.26	0.7
Rh78	0.56	4	3.5	1.31	0.50	0.73	0.72	0.55	3.2
Rh77	0.69	4	3.6	1.34	0.75	0.74	0.73	0.66	3.1
Rh93	0.49	2	1.5	0.53	0.45	0.36	0.35	0.44	1.8
RhAB38	0.63	4	30	1.21	0.55	0.68	0.66	0.62	5.9
Rh60	0.48	2	1.2	0.27	0.15	0.14	0.14	0.36	1.0
Rh85	0.53	3	1.9	0.83	0.60	0.50	0.48	0.49	2.7
Rh98	0.43	4	3.6	1.31	0.35	0.74	0.72	0.39	0.7
Rh72	0.65	3	20	0.85	0.30	0.50	0.49	0.53	0.8
Rh73	0.65	3	2.1	0.82	0.30	0.53	0.52	0.61	4.5
RhAB28	0.66	3	2.5	0.98	0.65	0.62	0.6	0.61	4.2
ABRIIRpu4	0.50	2	20	0.69	0.90	0.51	0.5	0.5	67.9
ABRIIRpu7	0.44	2	1.9	0.67	0.00	0.49	0.48	0.31	0.5
ABRIIRpu10	0.26	3	1.2	0.39	0.00	0.19	0.19	0.26	1.6
ABRIIRpu11	0.70	4	3.1	1.24	0.55	0.70	0.68	0.68	10.0
ABRIIRpu12	0.54	2	1.9	0.66	0.55	0.48	0.47	0.53	4.5
ABRIIRpu32	0.14	2	1.1	0.2	0.10	0.10	0.1	0.14	6.3
ABRIIRpu33	0.35	2	1.2	0.33	0.20	0.18	0.18	0.33	5.2
ABRIIRpu36	0.31	2	1.3	0.38	0.15	0.22	0.22	0.31	13.4
RA013	0.39	2	20	0.69	0.05	0.51	0.5	0.4	2.4
RA023	0.59	3	2.9	1.08	0.55	0.67	0.65	0.6	9.2
RA043a	0.67	4	2.1	1.01	0.60	0.54	0.53	0.65	6.8
Mean±Std.Dev.	0.5±0.2	2.9±1	2.3±0.9	0.8±0.4	0.4±0.3	0.5±0.2	0.5±0.2	0.5±0.2	2.1

PIC=polymorphic information content; na=observed number of alleles; ne=effective number of alleles; I= Shannon's information index; Ho=observed heterozygosity; He=Expected heterozygosity, Nei=Nei's expected heterozygosity; Havg= Average heterozygosity;
* Nm = Gene flow estimated from Fst = 0.25(1 - Fst)/Fst.

values (F_{ST}) among four subpopulations ranged from 0.01 to 0.15 (Fig. 4).

Discussion

The microsatellite markers or SSRs have been considered as one of the best markers for assessing the variability at a DNA level, which enables the use of these markers for genetic characterization and diversity analysis (Ben-Meir and Vainstein 1994) also facilitates the breeder for identification of potential parental lines. The SSR markers used in the present study had comparatively high discriminating power, out of 30 markers used in the study, 17 had PIC value > 0.5 and 11 had more than > 0.25 indicating medium to high polymorphism and these markers could provide substantive knowledge for conducting diversity analysis in roses. Several studies on rose with SSR

Table 4. Ewens-Watterson Test for neutrality of microsatellite markers for overall population (Manly, 1985)

Locus	Obs. F	SE*	L95*	U95*
Rh79	0.21	0.03	0.25	0.87
RhAB9	0.34	0.04	0.38	0.99
Rh48	0.48	0.02	0.22	0.82
Rh80	**0.53**	0.03	0.50	0.99
Rh96	0.49	0.04	0.38	0.99
Rh50	0.92	0.03	0.50	0.99
Rh58	0.20	0.03	0.26	0.88
RhABT12	0.55	0.04	0.39	0.99
Rh65	0.67	0.03	0.50	0.99
Rh78	0.44	0.04	0.33	0.96
Rh77	0.31	0.04	0.33	0.96
Rh93	0.51	0.03	0.50	0.99
RhAB38	0.37	0.04	0.33	0.96
Rh60	0.52	0.03	0.50	0.99
Rh85	0.47	0.04	0.39	0.99
Rh98	0.57	0.04	0.32	0.95
Rh72	0.35	0.04	0.33	0.96
Rh73	0.35	0.04	0.39	0.98
RhAB28	**0.34**	0.04	0.39	0.98
ABRIIRpu4	0.50	0.03	0.50	0.99
ABRIIRpu7	**0.56**	0.03	0.51	0.99
ABRIIRpu10	0.74	0.04	0.38	0.98
ABRIIRpu11	**0.30**	0.04	0.32	0.95
ABRIIRpu12	0.46	0.04	0.39	0.98
ABRIIRpu32	0.86	0.03	0.50	0.99
ABRIIRpu33	0.65	0.03	0.50	0.99
ABRIIRpu36	0.69	0.03	0.50	0.99
RA013	0.61	0.04	0.41	0.98
RA023	0.41	0.04	0.39	0.98
RA043a	0.33	0.04	0.33	0.96

*These statistics were calculated using 1000 simulated samplesF-value (sum of square of allelic frequency) highlighted in bold lied outside the lower and upper limit of 99% confidence. This indicates that the respective loci were not neutral but may be linked to some selection traits

markers proved the effectiveness of these markers systems for the characterization and diversity studies (Babaei et al. 2007; Akond et al. 2012; Nadeem et al. 2014). Out of 30 polymorphic microsatellite markers used in the study for characterization three of them were non-neutral in nature, which were actively involved under selection process and linked to some of the rose traits. Markers with non-neutral in nature were able to analyse the functional or adoptive variability exists within the population also helpful for understanding the population inbreeding, evolutionary potential (Kirk and Freeland 2011).

Characterized wild species in the present study exhibited a distant genetic relationship with all modern day cultivars (Fig. 2) and displayed the maximum variability among all population subgroups; results obtained in present study emphasized the importance of wild species in breeding and development of present day cultivated types. Different valuable traits of these wild species can be transferred by making the planned inter crosses between selected wild species and cultivars based on their genetic nearness. Genetic variability was found narrow in the case of Indian cultivated types when compared to the exotic types and wild species. The narrow variability or higher genetic similarity noticed among the Indian rose population could due to their origin from repeated crosses involving few old popular cultivars. At a Nei's genetic distance vale of 0.03, exotic cultivars differentiated from Indian genotypes, explained a slight variability between the collections of different geographical regions. For example, most of the exotic collections used in the experiment were bred in Europe; involvement of native species from their respective regions while breeding of those genotypes also creates a certain degree of separation between the cultivars of different geographical regions.

The SSR markers used in the present assay were also effective in discriminate different modern groups of cultivars up to some extent. Majority of the cultivars used in the study had characteristic features of modern groups of roses such as Hybrid Teas, Floribundas and Grandifloras. Hybrid Tea cultivars were most vigorous growth habit, bear's high centred large flowers on long flower stocks whereas, the second group of cultivars, Floribundas were developed from crosses between Hybrid Teas and Polyanthas, they had shrubby stout plant growth habit and main characteristic feature of this group was flowers appear in large clusters unlike one flower per single stem in Hybrid Teas. The third group Grandifloras were developed from crosses between above two groups Hybrid Teas × Floribundas, Grandifloras types were separated from above both the groups with their intermediate plant growth habit and flowering characteristics. Th dendrogram (Fig. 2), constructed based on available genetic information, classified the cultivars into different sub-clusters, for example sub-cluster-III of major cluster-I consists majority of

Table 5. AMOVA: partitioning of the source of variation among different possible levels of genetic structure

Source of variation	d. f.	Sum of squares	Variance components	Percentage of variation
Among populations	3	194.25	0.82 Va	10.42
Among individuals within populations	144	1093.34	0.51 Vb	6.46
Within individuals	148	972.5	6.57 Vc	83.12
Total	295	2260.09	7.91	
Fixation Indices:	FIS :	0.072		
	FST :	0.104		
	FIT :	0.169		

cultivars with Hybrid Tea characteristics and sub-clusters-I, II and IV of major cluster-I contains cultivars with characteristic features like Floribunda and Grandiflora types, likewise all exotic cultivars used for characterization, had Hybrid Tea characteristics which was also found in single major cluster-II, which explaining the certain degree of genetic separation between different modern groups of rose cultivars. Discriminating power of SSRs for identification Hybrid Teas from the wild types has also been demonstrated in the previous studies (Esselink et al. 2003; Esselink et al. 2004).

To elucidate the genetic relationship among characterized roses, dissimilarity matrix was generated which reveals the pair wise genetic differences for all characterized genotypes. Pair wise dissimilarity values among the rose genotypes ranged from 0.03 to 0.77 indicating significant variability exists among them. The pair wise variability data of rose genotypes generated using SSRs in this study could be highly helpful for the rose breeders for the better selection of parental lines among the species and cultivated types based on genetic similarity and dissimilarity in order to introgress the alleles from wild species and to exploit the hybrid vigour in cultivated types.

STRUCTURE analysis results were confirmed results which were obtained with dendrogram analysis (Fig. 3). Four subpopulations identified in STRUCTURE analysis were identical to the genetic clusters of dendrogram and grouping of genotypes among four subpopulations also comparable with each other except for few Indian cultivars which were structured along with exotic cultivars within subpopulation-III (Table 1, Fig. 4) unlike dendrogram where all exotic cultivars grouped in single major cluster (Fig. 2). This could be due to the presence of some of the exotic cultivars within the pedigree of Indian genotypes which were found along with subpopulation-III (exotic cultivars). For example, cultivars Anurag (Sweet Afton × Gulzar), Sahasradhara (sport of Century Two) and Dr S.S. Bhatnagar (Oklahoma × White Christmas) were bred from exotic lines, Sweet Afton, Century Two and Oklahoma.

Genetic variability values calculated among four subpopulations (Table 2) proved the highest variability within wild species could be due to the higher number of alleles found within (na=2.9±1.05; ne=2.27±0.88). Genetic differences between different populations studied revealed highest differences between wild populations and exotic collections; this could be due to the differences in the genetic background of both the population groups due to differences in allelic composition and ploidy levels. Comparatively, low variability was observed between cultivated subpopulations. Lower to moderate level of variability exists among subpopulations also proved by their estimated Fst value range (Fst=0.05-0.15) between four subpopulations (Fig. 4). The level of variability exists among the populations depends upon several factors, including pollen dispersal, germplasm exchange, natural selection, mating system and geographical distribution range (Hamrick and Godt, 1996; Teixeira et al. 2014). The inbreeding depression identified among rose cultivars could be due to its free out crossing nature, which enables the better connectivity among populations and improves the genetic exchange and ultimately contributes to the genetic variability within populations (DeVries and Dubois, 1996; Yang et al. 2016). In addition to this, continuous selection of some of the selective interspecific hybrids and utilization of those selective lines repeatedly for breeding new cultivars were carried out to find out the novel traits. Free exchange of germplasm between different geographical regions for improving native germplasm also contributed to the

genetic uniformity among the cultivated rose genotypes. Comparatively high genetic distance values were noticed between wild subpopulation with remaining three modern groups of roses specifies the potentiality of the wild genotypes for the exploitation of new alleles in creating variability.

Authors' contribution

Conceptualization of research (KVB, TJ, KVP, AV); Designing of the experiments (KVB, DVSR); Contribution of experimental materials (CB, N, SP, KPS); Execution of field/lab experiments and data collection (AV); Analysis of data and interpretation (AV, KVB, GC); Preparation of manuscript (AV, KVB, GC).

Declaration

The authors declare no conflict of interest.

Acknowledgements

Author (AV) is thankful to Indian Council of Agricultural Research for providing fellowship and facilities during the course of research work and also thank, ICAR-NBPGR for providing research facilities and sharing valuable germplasm.

References

Akond M., Jin S. and Wang X. 2012. Molecular characterization of selected wild species and miniature roses based on SSR markers. Sci. Hort., **14**: 789-97.

Alsemaan T., Albatal N., Baydar H. and Almaarri K. 2011. Genetic diversity and qualitative variations of *Rosa damascena* in Syria. Int. J. Agr. Sci., **6**: 429-436.

Anderson J. A., Chirchil G. A., Atrique J. E., Tanksley S. O. and Sorrel M. E. 1993. Optimizing parental selection for genetic linkage maps. Genome., **36**: 181-186.

Azeem S., Khan A. I., Awan F. S., Riaz A. and Bahadur S. 2012. Genetic diversity of rose germplasm in Pakistan characterized by random amplified polymorphic DNA (RAPD) markers. Afr. J. Biotechnol., **11**: 10650-10654.

Babaei A., Tabaei-Aghdaei R., Khosh-Khui M., Omidbaigi R., Naghavi M., Esselink G. D. and Smulders M. J. M. 2007. Microsatellite analysis of Damask rose (*Rosa damascena* Mill.) accessions from various regions in Iran reveals multiple genotypes. BMC Plant Biol., **7**: 7-12.

Ben-Meir H. and Vainstein A. 1994. Assessment of genetic relatedness in roses by DNA fingerprint analysis. Sci. Hort., **58**: 115-121.

Botstein D., White R. L., Skolnick M. and Davis R. W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum Genet., **32**: 314-331.

Crespel L. and Mouchotte J. 2003. Methods of cross-breeding: In: Roberts A., Debener T., Gudin S. (eds) *Encyclopaedia of Rose Science*. Elsevier Science, Oxford, pp. 30-33.

De Cock K. 2008. Genetic diversity of wild roses (*Rosa* spp.) in Europe, with an in-depth morphological study of Flemish populations. PhD thesis, Ghent University.

DeVries D. P. and Dubois L. A. M. 1996. Rose breeding: past, present, prospects. Acta Hortic., **424**: 241-248.

Esselink G. D., Nybom H. and Vosman B. 2004. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting-peak ratios) method. Theor. Appl. Genet., **109**: 402-408.

Esselink G. D., Smulders M. J. M. and Vosman B. 2003. Identification of cut rose (*Rosa hybrida*) and rootstock varieties using robust sequence tagged microsatellite site markers. Theor. Appl. Genet., **106**: 277-286.

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

Exoffier L., Laval G. and Schneider S. 2005. Arlequin version 3.0: An integrated software package for population genetics data analysis, Evol. Bioinform., **1**: 47-50.

Foucher F., Oyant H. S. L. and Hamama L. 2015. Towards the rose genome sequence and its use in research and breeding. Acta Hortic., **1064**: 167-175.

Gardes L., Heizmann R. and Joyaux F. 2005. Molecular typing and history of the province roses horticultural group. Eur. J. Hortic. Sci., **70**: 162-172.

Gudin S.2000. Rose: genetics and breeding. *Plant Breed*. Rev., **17**: 59-189.

Hamrick J. L. and Godt M. J. W. 1996. Effects of life history traits on genetic diversity in plant species. Philos. Trans. R. Soc. Lond. *B.* Biol. Sci., **351**: 1291-129.

Hurst C. C. 1925. Chromosomes and characters in Rosa and their significance in the origin of species. Exper. Genet., **37**: 534-550.

Jian H., Zhang H., Tang K., Li S., Wang Q., Zhang T., Qiu X. and Yan H. 2010. Decaploidy in *Rosa praelucens* by houwer (*Rosaceae*) endemic to Zhongdian Plateau, Yunnan, China. Caryologia, **6**: 162-167.

Jiang L. and Zang D. 2017. Analysis of genetic relationship in *Rosa rugosa* using conserved DNA-derived polymorphic markers. Biotechnol. Biotechnol. Equip., **32**: 88-94.

Jowkar A., Mardi M., Kermani J. K. M., Kafi M. and Pirseyedi S. M. 2009. Isolation and characterization of microsatellite loci from *Rosa pulverulenta*. Mol. Ecol. Resour., **10**: 1098-1105.

Kiani M. K., Zamanii Z., Khalighi A., Fatahi R. and Byrne D. H. 2010. Microsatellite analysis of Iranian Damask

rose (*Rosa damascena* Mill.) germplasm. Plant Breed., **129**: 551-557.

Kimura T., Nishitani C., Iketani H., Ban Y. and Yanamato Y. 2006. Development of microsatellite markers in rose. Mol. Ecol. Resour., **6**: 210-212.

Kirk H. and Freeland J. R. 2011. Applications and implications of neutral versus non-neutral markers in molecular ecology. Int. J. Mol. Sci., **12**: 3966-3988.

Lu M., An H., Li L. and Jain M. 2016. Genome survey sequencing for the characterization of the genetic background of *Rosa roxburghii* tratt and leaf ascorbate metabolism genes. PLoS One., **11**: 147530.

Mohapatra A. and Rout G. R. 2005. Identification and analysis of genetic variation among rose cultivars using Random Amplified Polymorphic DNA. Z. Naturforsch. B., **60**: 611-617.

Murray M. G. and Thompson W. F. 1980. Rapid isolation of DNA high molecular weight plant DNA. Nucleic Acids Res., **8**: 4321-4325.

Nadeem M., Wang X., Akond M., Awan F. S., Khan A. I., Riaz A. and Younis A. 2014. Hybrid identification, morphological evaluation and genetic diversity analysis of *Rosa* x hybrida by SSR markers. Aust. J. Crop Sci., **8**: 183-190.

Nakamura N., Hirakawa H., Sato S., Otagaki S., Matsumoto S., Tabata S. and Tanaka Y. 2018. Genome structure of *Rosa multiflora*, a wild ancestor of cultivated roses. DNA Res., **25**: 113-121.

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

Panwar S., Singh K. P., Sonah H., Deshmukh P. K., Namita B., Prasad K. V. and Sharma T. R. 2015. Molecular fingerprinting and assessment of genetic diversity in rose. Indian J. Biotechnol., **14**: 518-524.

Perrier X. and Jacquemoud-Collet J. P. 2006. DARwin software, http://darwin.cirad.fr/darwin.

Prasad K. V., Kumar S. and Choudhary M. L. 2006. Molecular characterization of fragrant rose cultivars. Indian J. Hortic., **63**: 229-234.

Prichard J. K., Stephens M. and Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics., **155**: 945-959.

Reynolds J., Weir B. S. and Cockerham C. C. 1983. Estimation for the co-ancestry coefficient: basis for a short-term genetic distance. Genetics, **105**: 767-779.

Roberts A., Debener T. and Gudin S. 2003. Encyclopaedia of rose science, Oxford, Elsevier, Academic Press.

Roberts A., Gladis T. H. and Brumme H. 2008. DNA amounts of Roses (*Rosa* L.) and their use in attributing ploidy levels. Plant Cell Rep., **28**: 61-71.

Rohlf F. J. 2000. NTSYS-PC version 2.02J, Exeter software Setauket, New York, USA.

Rai H., Raju D. V. S., Kumar A. M. B., Janakiram T., Namita S., Krishnan G. and Rana J. C. 2015. Characterization and analysis of genetic diversity among different species of the rose using morphological and molecular markers. Indian J. Agr. Sci., **85**(2): 240-245.

Samiei L., Naderi R., Khalighi A., Shahnejat-Bushehri A. A., Mozaf-Farian V., Esselink G. D., Osaloo S. K. and Smulders M. J. M. 2010. Genetic diversity and genetic similarities between Iranian rose species. J. Hortic. Sci. Biotech., **85**: 231-237.

Scariot V., Akkak A. and Botta R. 2006. Characterization and genetic relationships of wild species and old garden roses based on microsatellite analysis. J. Am. Soc. Hortic. Sci., **131**: 66-73.

Slatkin M. A. 1995. Measure of population subdivision based on microsatellite allele frequencies. Genetics, **139**: 457-462.

Teixeira H., Rodríguez E. S. and Nabais C. 2014. Genetic diversity and differentiation of *Juniperus thurifera* in Spain and Morocco as determined by SSR. PLoS One., **9**: 88996.

Tejaswini and Prakash M. S. 2005. Utilization of wild rose species in India. Acta Hortic., **690**: 91-96.

Weir B. S. and Cockerham C. C. 1984. Estimating F-statistics for the analysis of population structure. Evolution., **38**: 1358-1370.

Wissemann V. 2003. Classification/Conventional taxonomy (wild roses). In: Robert A.V., Debener T., Gudin S. (eds) *Encyclopaedia of Rose Science*. Elsevier, Oxford, pp.111-117.

Wu S., Ueda Y., He H., Nishihara S. and Matsumoto S. 2000. Phylogenetic analysis of Japanese *Rosa* species using matK sequences. Breeding Sci., **50**: 275-282.

Yan Z., Denneboom C., Hattendorf A., Dolstra O., Debener T., Stam P. and Visser P. B. 2005. Construction of an integrated map of rose with AFLP, SSR, PK, RGA, RFLP, SCAR and morphological markers. Theor. Appl. Genet., **110**: 766-777.

Yang S., Guo N. and Hong G. E. 2016. Morphological and AFLP-Based Genetic Diversity in *Rosa platycantha* Population in Eastern Tianshan Mountains of North-western China. Hortic. Plant J., **2**: 55-60.

Yeh F. C., Yang R. C. and Boyle T. B. J. 1999. Power Marker 3.23, Alpha Ease FC 4: Microsoft Window-based free Software for Population Genetic Analysis (ftp://ftp.microsoft.com/Softlib/HPGL.EXE.).

Yeh F. C., Yang R. C., Boyle T. B. J., Ye Z. H. and Mao J. X. 2000. POPGENE 32, Microsoft Windows-based software for population structure analysis (version 1.32). Edmonton, Alberta, Canada: Molecular biology and Biotechnology Centre, University of Alberta.

Zlesak D. C. 2006. *Rosa x hybrida* L. In: Anderson N. O. (ed) Flower Breeding and Genetics: Issues, Challenges and Opportunities for the 21st Century. Springer, The Netherlands, pp. 695-738.