Molecular identification of *Rosa x damascena* growing in Taif region (Saudi Arabia)

Sayed Amer,[1,2] Salih A. Basaid,[1] Esmat Ali[1,3]

[1]Biology Department, Faculty of Science, Taif University, Saudi Arabia; [2]Faculty of Science, Cairo University, Egypt; [3]Horticulture Department, Faculty of Agriculture, Assiut University, Egypt

Abstract

A fragment of 772 bp of the chloroplast maturase K gene was amplified and sequenced for *Rosa x damascena trigintipetala* variety growing in Taif region of Saudi Arabia. The data were aligned with their counterparts of other varieties already found in the Genbank database and were analyzed by maximum-parsimony, neighbor-joining and maximum-likelihood methods and a single rooted tree was executed. *R. x damascena trigintipetala* was paraphyletic where one sample [A] clustered with all varieties while the second [B] was basal. *R. x damascena* was sister to *R. x chinensis semperflorens* with the later being basal. *R. x damascena gori* was basal for all taxa studied. *R. moschata* was inside the clade of *R. x damascena*. Hybridization could be possible among *R. damascena*, *R. chinensis* and *R. moschata*. The genetic distance and tree topology indicated that [A] variety could be originated from *R. moshata* while [B] could be originated from gori or *R. chinensis semperflorens*. We, therefore, may consider that *R. x damascena gori* or *R. chinensis* could be the origin of all nowadays *R. x damascena* varieties.

Introduction

Rosa is a genus of an ornamental plant containing hundreds of species and is one of the most cultivated plants in the world.[1] Roses are economically valuable as they contain essential oils which are a source of perfumes and scents.[2] The researchers have identified significant pharmacological properties of rose extracts as it was used in radioprotective[3] and anti-inflammatory[4] purposes. Compounds like polysaccharides, polyphenols and other secondary metabolites are found with high level in rose plants.[5] They tightly bind with DNA and inhibit its isolation.[6] They co-precipitate with the DNA during its extraction and making it more viscous and also inhibit the polymerase chain reaction (PCR) enzymatic reaction.[7-9] Molecular study on the genetic framework of these plants requires large yield of DNA and therefore, method described by Ibrahim[10] was used for this purpose with few modifications.

The inter-specific relationships within *Rosa* remain controversial. Different names were given to *Rosa* species based on the morphological variants and hybrids.[11] The genetic studies on the genus were primarily based on isozymes[12] and random amplification of polymorphic DNA (RAPD)-PCR.[13,14] Microsatellite analyses for both wild and cultivated roses were also conducted by Baydar and colleagues.[15] The sequences of the internal transcribed spacer (ITS) were used by several researchers.[16-18] The chloroplast *matK* gene[16,19] and the *atpB-rbcL* intergenic spacer[18] were conducted for phylogenetic resolution of the genus and its subgenera, however, the genetic controversial is still standing. These investigations explained partly the extremely low levels of sequence divergence observed across the genus.[18,19] On the other hand DNA-based molecular marker systems are efficient and informative for genetic analysis of roses because DNA polymorphism indicated by these markers is not affected by environmental conditions. Several molecular marker systems, *i.e.,* Restriction Fragment Length Polymorphism (RFLP),[20] RAPD,[21,22] Simple Sequence Repeat (SSR)[23] and Amplified Fragment Length Polymorphism (AFLP)[15] have been used for genetic analysis of rose species, cultivars and rootstocks.

R. x damascena is a famous species within this genus that possesses different varieties of which *R. d. trigintipetala* and *R. d. semperflorens*. *R. x damascena trigintipetala* (the so-called *Ward taifi*) is highly cultivated in Taif governorate and there are different debates about its origin. This Taif-rose is the famous rose cultivated in different areas of Taif city at the western region of Saudi Arabia and is also known as Damask rose. There are two varieties of Damask Rose, such as Autumn Damask rose (*R. x damascena semperflorens*) and Damask rose Kazanlik (*Rosa x damascena trigintipetala*). It is known that Kazanlik in Bulgaria and Isparta in Turkey are the main areas of origin of *R. x damascena* (references cited in El-Assal and colleagues).[24] Moreover, Taif-rose has been suggested to be brought to Taif from Balkans by Turks, who occupied this area in the 14th century. However, El-Assal and colleagues[24] indicated that Taif-rose has closed genetic relations to the gory rose-Syrian cultivated in Syria. It was therefore necessary to conduct PCR amplification and sequencing for *matK* gene in order to identify genetically Taif-rose and construct its relationship to other varieties.

Correspondence: Sayed Amer, Biology Department, Faculty of Science, Taif University, Taif, 888, Saudi Arabia.
Tel.: +96.6559822001 - Fax: +96.6127343414.
E-mail:yasser92us@yahoo.com

Key words: Maturase K; rose varieties; *Rosa damascene*; Saudi Arabia; Taif-rose.

Contributions: SA, laboratory setting and manuscript writing; SB and EF executing idea and experimental design; KA reading and revising the manuscript.

Conflict of interest: the authors declare no potential conflict of interest.

Received for publication: 13 November 2015.
Revision received: 25 December 2015.
Accepted for publication: 25 December 2015.

International Journal of Plant Biology 20165; 7:6307
doi:10.4081/pb.2016.6307

Materials and Methods

Samples collection and DNA extraction

Fresh young leafs of 8 samples of *R. x damascena* were collected from the Botanical Garden at Faculty of Science, Taif University, Taif, Saudi Arabia (5 samples from Taif variety *Rosa x damascena trigintipetala*, 2 samples from *Rosa x damascena gori* and one sample from *Rosa x damascena* French variety). Approximately, 300 mg of the leaf tissues were ground separately to a fine powder under liquid nitrogen with the use of pre-chilled mortar and pestle. The method detailed by Ibrahim[10] (CTAB-PVP- βME) was used to extract DNA form the plant tissues with some modifications. These modifications were that the DNA precipitate was dissolved in 100 µL bi-distilled water. Finally, spin column (BioFlux, Tokyo, Japan) was used to purify the extracted DNA according to the manufacturer instructions.

Polymerase chain reaction experiments

PCR was conducted in a final volume of 25 µL containing 2 µL DNA template, 0.2 µL of 10 Pmolar a forward primer (RosaL: 5`-CACTTATCTTTCGGGAGT-3`), 0.2 µL of 10 Pmolar a reverse primer (RosaH: 5`-GAGTATCTTTTGATGCCAGA-3`), 12.5 µL PCR master mix

(Promega Corporation, Madison, WI) and 10.3 µL autoclaved deionized water. PCR was carried out using a PeX 0.5 thermal Cycler with the cycle sequence at 94°C for 5 min one cycle, followed by 35 cycles each of which consisted of denaturation at 94°C for one min, annealing at 50°C for one min and extension at 72°C for one min. A final strand elongation for one cycle at 72°C was done for additional 5 min. The PCR products were analyzed in 1% agarose gel electrophoresis in TAE buffer (40 mm Tris, 40mM acetic acid and 1 mM EDTA) with ethidium bromide staining. A 100-bp DNA ladder (Biolabs) was used as a molecular marker. PCR products were visualized under UV light and photographed (Figure 1). The products were then purified using spin column (BioFlux, Tokyo) according to the manufacturer instructions.

Phylogenetic analysis

The sequenced *matK* gene fragment of *R. x damascena trigintipetala* and *R. damascena gori* was aligned manually with its counterparts of other related taxa collected from the database (see their accession numbers on the tree) and the aligned data were used for phylogenetic analysis with a total length of 772 bp. The same sequence for *R. californica* was used for rooting the tree. The phylogenetic analyses were done primarily by maximum-likelihood (ML) method with PAUP* 4.0b10[25] by heuristic searches with the nearest-neighbor interchange (NNI) branch swapping and 10 random taxon additions. The general reversible model (GTR+I) and parameters optimized by Modeltest 3.0[26] were used. Bootstrapping replicates were set to 300. Maximum-parsimony (MP) and neighbor joining (NJ) were also conducted with PAUP* 4.0b10 under bootstrapping replications of 10000 to confirm the robustness of the ML tree topology. The distance option of Tamura-Nei was adjusted for executing the neighbor-joining analysis.

Results

No morphological characteristics were observed to differentiate between any of *R. x damascena* samples and all specimens from each variety were identical. The extracted DNA of the studied samples was tested by electrophoreses using 1% agarose and all samples gave positive results (data not shown). The amplified fragments of the *matK* gene were also examined on the agarose gel electrophoreses and the PCR profile was shown in Figure 2.

Unambiguous nucleotides of 772 bp of *matK* gene were sequenced for the collected samples. The base compositions of these data were A=31.04%, C=17.50%, G=14.50% and T=36.96%. Among the 772 sites included for tree analysis, 766 were constant and 6 were variables. Five variable sites were parsimony uninformative and only 1 was informative

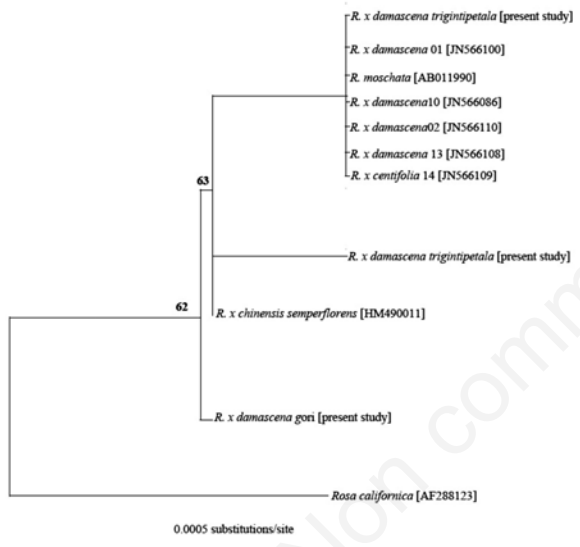

Figure 1. A neighbor-joining tree constructed from 722 bp *matK* gene fragment sequenced in this study. The values at some nodes refer to the bootstrapping of maximum-parsimony, neighbor-joining and maximum-likelihood analyses. The numbers between brackets are the accession numbers of the sequences collected from the Genbank database.

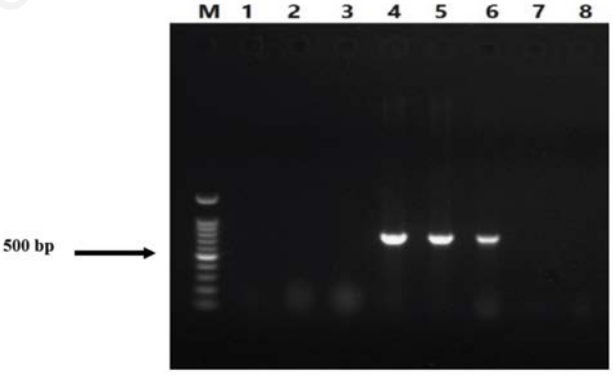

Figure 2. The amplified products of the chloroplast *matK* gene for eight rose samples. R. x damascena trigintipetala (lanes 1-5), *R. x damascena gori* (lane 6) and R. x damascena Finish variety (lanes 7-8). Lane 1 represents 100 bp DNA ladder (M).

Table 1. Pairwise genetic distances among the different varieties of *R. damascena* populations. The distances were calculated from matK data sequenced in this study.

	R. x damascena trigintipetala [A]	*R. x damascena trigintipetala [B]*	*R. x damascena* (other varieties)	*R. x damascena moschata*	*R. x chinensis simperflornes*
R. x damascena trigintipetala [A]	-	-	-	-	-
R. x damascena trigintipetala [B]	0.003	-	-	-	-
R. x damascena (other varieties)	0.000	0.003	-	-	-
R. x damascena moschata	0.000	0.003	0.000	-	-
R. x chinensis simperflornes	0.0013	0.0013	0.0013	0.0013	-
R. x damascena gori	0.0013	0.0013	0.0013	0.0013	0.000

under parsimony criterion. The constructed parsimony tree showed consistency index (CI=1.00), homology index (HI=0.00), retention index (RI=1.00) and rescaled consistency index (RC=1.00).

The ML analysis exhibited a negative log likelihood score –lnL=1049.6. Both MP and NJ methods executed a similar tree topology to that given by ML method. The three methods gave the same bootstrap probabilities at the node clustering all varieties (Figure 1). The best-fit model that explained the datasets was GTR+I exhibiting model parameters as follows: substitution rate matrix R (a)=3.97; R (b)=2.75; R (c)=0.00; R (d)=5.62, R (e)=2.76 and R(f)= 1.00. The proportion of invariable sites (I) was 0.947.

In the NJ tree (Figure 1), a clustering of all varieties in one clade was shown with *R. x damascena trigintipetala* being paraphyletic. One sample of *R. x damascena trigintipetala* [A] clustered with all available varieties while the second sample [B] was basal to this cluster. All varieties of *R. x damascena* showed a sister relationship to *R. x chinensis semperflorens* with the later being basal. *R. x damascena gori* appeared to the basal for all *R. x damascena* studied. *R. moschata* was found to be inside the clade of *R. x damascena*. It seemed, therefore, that there was a possible hybridization among the three taxa of *R. damascena*, *R. chinensis* and *R. moschata*.

The pairwise genetic distances among the studied varieties are recorded in Table 1. The distance showed an identity among *R. moschata*, *R. x damascena trigintipetala* [A] and the other varieties of *R. x damascena* (D=0.000). On the other hand, the second variety of *R. x damascena trigintipetala* [B] was similarly distant from the [A] variety, the other *R. x damascena* varieties and *R. moshata* with a distance of 0.003. *R. x damascena gori* and *R. chinensis semperflorens* were distant from all studied varieties with genetic a distance 0.0013.

The genetic distance together with the tree topology indicated, therefore, that one of both *R. x damascena trigintipetala* varieties [A] could be originated from *R. moshata* while the other [B] could be originated from the gori or *R. chinensis semperflorens*. We, therefore, may consider that *R. x damascena trigintipetala* the so-called *Ward taifi* was not the origin of all nowadays *R. x damascena* varieties but the gori or *R. chinensis* could be.

The sequenced matK fragment was aligned with the complete gene sequence (1512 bp) obtained from the Genbank database. This fragment was located between base 75 and base 846 with a length of 772 bp. The sequence showed only 2 base substitutions within *R. x damascena trigintipetala* [B]. One of the two substitutions was a transition within pyrimidines in the third position of the codon where C_{114} was substituted with T_{114}. The second substitution was a transversion, in the second position, of C_{220} to G_{220}. Only the later substitution was non-synonymous recording a change of the amino acid number 74 from leucine into valine at C_{220} G_{220} (Figure 3). This mutation makes *R. x damascena trigintipetala* variety [B] close to both the gori and *R. chinensis semperflorens* while the [A] variety is identical to the other varieties of *R. x damascena*. The number below the letter referred to the corresponding position of the nucleotide inside the complete gene sequence.

```
A  1 XXXXXXXXXXXXXXXXXXXXXXXXXXEYIYALAHDRGLNRSVLLDNVGYDKKSSLLIIKRL  60
B    XXXXXXXXXXXXXXXXXXXXXXXXXXEYIYALAHDRGLNRSVLLDNVGYDKKSSLLIIKRL
C    MEEFQGYLELYRSQQHDFLYPLIFREYIYALAHDRGLNRSVLLDNVGYDKKSSLLIIKRL
                               **********************************

A    ISRMYQQNHFIISLNDSNQNKFFGYNKNLYSQMISEGFAVIVEIPFSLRLVSSLEETETV 120
B    ISRMYQQNHFIISVNDSNQNKFFGYNKNLYSQMISEGFAVIVEIPFSLRLVSSLEETETV
C    ISRMYQQNHFIISLNDSNQNKFFGYNKNLYSQMISEGFAVIVEIPFSLRLVSSLEETETV
     ************:*********************************************
                 ▬

A    KSYNLRSIHSIFPFFEDKFPHLNYASDVLIPYPIHLEILVQTLRYCVKDPSSLHLLRLFL 180
B    KSYNLRSIHSIFPFFEDKFPHLNYASDVLIPYPIHLEILVQTLRYCVKDPSSLHLLRLFL
C    KSYNLRSIHSIFPFFEDKFPHLNYASDVLIPYPIHLEILVQTLRYCVKDPSSLHLLRLFL
     ************************************************************

A    HEYYNWNTLITPKKSIFAKSNQRLFLLLYNSYVCEYESILLFLRNQSNHLRLTSSGILFE 240
B    HEYYNWNTLITPKKSIFAKSNQRLFLLLYNSYVCEYESILLFLRNQSNHLRLTSSGILFE
C    HEYYNWNTLITPKKSIFAKSNQRLFLLLYNSYVCEYESILLFLRNQSNHLRLTSSGILFE
     ************************************************************

A    RIRFYEKIKYPVEEVFANDFPATLWFFKDPFIQYVRYQGKSILLQKKPHTQKKNWGFLGF 300
B    RIRFYEKIKYPVEEVFANDFPATLWFFKDPFIQYVRYQGKSIXXXXXXXXXXXXXXXXXX
C    RIRFYEKIKYPVEEVFANDFPATLWFFKDPFIQYVRYQGKSILASKDTPLLMNKWKYYLV
     ****************************************
```

Figure 3. The aligned amino acid sequences for the two varieties of *R. x damascena trigintipetala* [A] and [B] along with that of *R. x damascena* [C] found in Genbank database. The underlined letter showed the non-synonymous change in the variety [B]. Stars indicate sequence identity. The missed data from both varieties A and B are denoted by X.

Discussion and Conclusions

Using the protocol detailed by Ibrahim[10] with some modifications described herein, DNA necessary for molecular genetics and biodiversity applications were successfully extracted from eight different Taif-rose samples containing high levels of ploy phenolics and secondary metabolites. According to the protocol of CITAB mixed with PVP and ME followed by DNA purification kit, a high pure yield of DNA from fresh leaf tissue grinned with liquid nitrogen was obtained. The DNA quantity and purity obtained herein was comparable to that obtained by other methods.[6,7,9] Good DNA quality and quantity was clearly shown. We did not measure the A/A ratio (absorbance at 260/ absorbance at 280) and we went directly to the PCR amplification to ensure the little contamination of proteins, polysaccharides or polyphenolics. This extraction procedure was found suitable for direct PCR amplifications of nuclear or chloroplast DNAs.

In the present procedure, β-mercaptoethanol and high concentration of EDTA in the extraction buffer made some temporary pores in the cytoplasm membrane, eliminating some of soluble polysaccharides, polyphenols and pigments improving the DNA yield of the rose plants and similar ornamental plants.[5,6]

The use of high levels of PVP and β-mercaptoethanol had a considerable effect in preventing oxidization of the secondary metabolites. Considering extraction time, DNA yield and purity and PCR amplifications by this procedure emerged to be the best for molecular diversity analysis of roses. Because different plants can vary considerably in the number and types of produced secondary compounds, it is unlikely that any one technique for isolating contaminant-free nuclear DNA will ever be developed. It is likely that our rose DNA isolation protocol can be used to isolate nuclear DNA from a variety of other plant species.

R. x damascena was believed to be originated in the Middle East as a wild plant and was introduced to Western Europe twice approximately 500 years ago,[27] however, the present results could not support this suggestion. Rusanov and colleagues[28] stated that *R. x damascena trigintipetala* was the common ancestor genotype which has been propagated in Europe and Asia, however, the paraphyly of this variety in the present study disagreed with the authors statement. We may agree with Huxely,[29] who revealed that *R. x damascena* originated partly as a hybrid from *R. moschata.* The author mentioned that *semperflorens* is one variety of *R. x damascena* flowering during autumn. Based on random amplified polymorphic DNA, inter simple sequence repeat and simple sequence repeat markers Taif-rose was shown to have closed genetic relations to the gory rose.[24] These results are in agreement with the present finding where *R. x damascena gori* was shown to be basal in a clade including both varieties of *R. x damascena trigintipetala.*

In conclusion, Taif-rose (*Ward taifi*) or *R. x damascena trigintipetala* is not considered as one homogenous variety but seemed to be two varieties. One variety is originated from *R. moschata* while the other is originated from *R. x. chinensis semperflorens* or from *R. x damascena gori.* The present study is considered a preliminary trial for investigating the genetic variability within *R. x damascena trigintipetala.* However, further investigation is needed to clearly infer the origin of Taif varieties of *R. x damascena trigintipetala*. Such a study must be conducted on different collections of Taif-rose and on more molecular data.

References

1. Jabbarzadeh Z, Khosh-Khui M, Salehi H, Saberivand A. Optimization of DNA Extraction for ISSR studies in seven important rose species of Iran. Amer-Euras J Sust Agri 2009;3:639-42.
2. Krussmann G. The complete book of roses. Portland, Oregon: Timber Press; 1981.
3. Akhmadieva AK, Zaichkina SI, Ruzieva RK, Ganassi EF. The protective action of a natural preparation of anthocyanin (pelargonidin-3,5-diglucoside). Radiobiol 1993;33:433-5.
4. Winther K, Rein E, Kharazmi A. The anti-inflammatory properties of rose-hip. Inflammopharmacol 1999;7:63-8.
5. Xu Q, Wen X, Deng X. A simple protocol for isolating genomic DNA from chestnut rose (Rosa roxburghii Tratt) for RFLP and PCR analyses. Plant Mo Biol Rep 2004;22:301a-301g.
6. Dehestani A, Kazemi Tabar SK. A rapid efficient method for DNA isolation from plants with high levels of secondary metabolites. Asian J Plant Sci 2007;6:977-81.
7. Dixit A. A simple and rapid procedure for isolation of amaranthus DNA suitable for fingerprint analysis. Plant Mol Biol Rep 1998;16:1-8.
8. Khanuja PSS, Shasany AK, Darokar MP, Kumar S. Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. Plant Mol Biol Rep 1999;17:1-7.
9. Khan IA, Awan FS, Ahmad A, Khan AA. A modified mini-prep method for economical and rapid extraction of genomic DNA in plants. Plant Mol Biol Rep 2004;22:89a-89e.
10. Ibrahim RIH. A modified CTAB protocol for DNA extraction from young flower petals of some medicinal plant species. Gene Conser 2011;10:165-82.
11. Bruneau A, Starr JR, Joly S. Phylogenetic relationships in the genus rosa: new evidence from chloroplast DNA sequences and an appraisal of current knowledge. Syst Bot 2007;32:366-78.
12. Kim YJ, Byrne D. Interspecific hybrid verification of rosa with isozymes. Hort Sci 1996;31:1207-9.
13. Debener T, Bartels C, Mattiesch L. RAPD analysis of genetic variation between a group of rose cultivars and selected wild rose species. Mol Breed 1996;2:321-7.
14. Jan CH, Byrne DH, Manhart J, Wilson H. Rose germplasm analysis with RAPD markers. Hort Sci 1999;34: 341-5.
15. Baydar NG, Baydar H, Debener T. Analysis of genetic relationships among rosa damascena plants grown in Turkey by using AFLP and microsatellite markers. J Biotechnol 2004;111:263-7.
16. Matsumoto S, Nishio H, Ueda Y, Fukui H. Phylogenetic analyses of the genus rosa: polyphyly of section pimpinellifoliae and origin of rosa 3fortuniana. Lindl Acta Hort 2001;547:357-63.
17. Wu S, Ueda Y, Nishihara S, Matsumoto S. Phylogenetic analysis of Japanese rosa species using DNA sequences of nuclear ribosomal internal transcribed spacers (ITS). J Hort Sci Biotechnol 2001;76:127-32.
18. Wissemann V, Ritz CM. The genus rosa (rosoideae, rosaceae) revisited: molecular analysis of nrITS-1 and atpB-rbcL intergenic spacer (IGS) versus conventional taxonomy. Bot J Linn Soc 2005;147:275-90.
19. Matsumoto S, Kouchi M, Yabuki J, et al. Phylogenetic analyses of the genus rosa using the matK sequence: molecular evidence for the narrow genetic background of modern roses. Scientia Hort 1998;77:73-82.
20. Rajapakse S, Byrne DH, Zhang L, et al. Two genetic linkage maps of tetraploid roses. Theor Applied Genet 2001;103:575-83.
21. Atienza SG, Torres AM, Millan T, Cubero JI. Genetic diversity in rosa as revealed by RAPDs. Agric Conspectus Sci 2005;70:75-85.
22. Kaur N, Sharma RK, Dhyani D, et al. Molecular characterization of interspecific hybrids of scented roses using RAPD markers. Acta Hort 2007;751:175-9.
23. Esselink G, Smulders M, Vosman B. Identification of cut rose (rosa hybrida) and rootstock varieties using robust sequence tagged microsatellite site markers. Theor Applied Genet 2003;106:277-86.
24. El-Assal SE, El-Awady MA, El-Tarras A, Shehab S. Assessing the genetic relationship of Taif rose with some rose genotypes (Rosa sp.) based on random amplified polymorphic DNA, inter simple sequence repeat and simple sequence repeat markers. Amer J Biochem Biotechnol 2014;10:88-98.
25. Swofford DL. PAUP: phylogenetic analysis using maximum parsimony (and other methods). Version 4.0b 10. Sunderland, Massachusetts: Sinauer; 2002.
26. Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. Bioinformatics 1998;14:817-8.
27. Beales P, Cairns T, Duncan W, et al. Botanica's roses: the encyclopedia of Roses. Random House, Australia; 1998.
28. Rusanov K, Kovacheva N, Vosman B, et al. Microsatellite analysis of rosa damascena mill. accessions reveals genetic similarity between genotypes used for rose oil production and old damask rose varieties. Theor Appl Genet 2005;111:804-9.
29. Huxley A. New RHS Dictionary of Gardening. London: Macmillan Publishing Company; 1992.