Microsatellite analysis of Damask rose (Rosa damascena Mill.) accessions from various regions in Iran reveals multiple genotypes

Alireza Babaei1, Seyed Reza Tabaei-Aghdaei2, Morteza Khosh-Khui3, Reza Omidbaigi1, Mohammad Reza Naghavi4, Gerhard D Esselink5 and Marinus JM Smulders*5

Address: 1Department of Horticultural Science, Faculty of Agriculture, Tarbiat Modares University, P.O. Box 14115-365, Tehran, Iran, 2Biotechnology Research Department of Natural Resources, Research Institute of Forests and Rangelands, P.O. Box 13185-116, Tehran, Iran, 3Department of Horticultural Science, Faculty of Agriculture, Shiraz University, Shiraz, Iran, 4Department of Plant Breeding, Faculty of Agriculture, University of Tehran, Tehran, Iran and 5Plant Research International, Wageningen UR, P.O. Box 16, 6700 AA Wageningen, The Netherlands

Email: Alireza Babaei - arbabaei@modares.ac.ir; Seyed Reza Tabaei-Aghdaei - tabaei@rifr-ac.ir; Morteza Khosh-Khui - mkhoshkhui@yahoo.com; Reza Omidbaigi - romidbaigi@yahoo.com; Mohammad Reza Naghavi - mnaghavi@ut.ac.ir; Gerhard D Esselink - danny.esselink@wur.nl; Marinus JM Smulders* - rene.smulders@wur.nl

* Corresponding author

Abstract

Background: Damask roses (Rosa damascena Mill.) are mainly used for essential oil production. Previous studies have indicated that all production material in Bulgaria and Turkey consists of only one genotype. Nine polymorphic microsatellite markers were used to analyze the genetic diversity of 40 accessions of R. damascena collected across major and minor rose oil production areas in Iran.

Results: All microsatellite markers showed a high level of polymorphism (5–15 alleles per microsatellite marker, with an average of 9.11 alleles per locus). Cluster analysis of genetic similarities revealed that these microsatellites identified a total of nine different genotypes. The genotype from Isfahan province, which is the major production area, was by far the most common genotype (27/40 accessions). It was identical to the Bulgarian genotype. Other genotypes (each represented by 1–4 accessions) were collected from minor production areas in several provinces, notably in the mountainous Northwest of Iran.

Conclusion: This is the first study that uncovered genetic diversity within Damask rose. Our results will guide new collection activities to establish larger collections and manage the Iranian Damask rose genetic resources. The genotypes identified here may be directly useful for breeding.

Background

There are almost 200 species and more than 18000 cultivars in the genus Rosa [1]. They are mostly shrubs, distributed in the temperate zones of the Northern hemisphere [2]. One of the important Rosa species is Rosa damascena Mill., which is commercially used for essential oil production and cultivated as garden rose [3]. In recent years, antioxidant, antibacterial and antimicrobial activities of R. damascena essential oil have been demonstrated [4-7]. Three recent studies on molecular analyses of genetic diversity of Rosa damascena Mill. with RAPD, AFLP and SSR markers did not show any polymorphism among R.
**Results**

**Microsatellite analysis**

In this study 40 accessions of *Rosa damascena* (Table 1) that showed a high level of phenotypic and oil content variation were analyzed with nine microsatellite markers. All markers detected polymorphisms among the samples. The number of alleles ranged from 5 to 15 with an average of 9.11 (Table 2). Using the MAC-PR method, we determined the allelic configurations at six loci (RhP519, RhB303, RHEO506, RhD221, RhP50, RhE2b) for all investigated accessions (Table 3).

**Genotype identification**

Cluster analysis resulted in grouping of the 40 accessions into nine distinct genotypes (Fig. 1). The main group consisted of 27 landraces that showed the same microsatellite profile. This group included all accessions from the main rose oil production sites of Damask rose in Iran. The pattern of this group was identical to that of an accession from Bulgarian production areas. Rusanov et al. showed that all Bulgarian Damask roses are this genotype [3].

The other genotypes that we identified in the cluster analyses were present in much smaller numbers. Some genotypes were unique (accessions from Tehran, Guilan, Kermanshah, Qom provinces and one accession from Fars province); others were present as two or four accessions (Fig. 1a and Table 1). The unique accessions were from mountainous and remote areas in the Northwest of Iran where roses are cultivated on small scale. In addition, the accessions from the humid area near the Caspian Sea were different from all other accessions as well.

The accessions from Fars province formed two distinct clusters in the dendrogram. They are from an environmentally very distinct region, far to the South of Iran. One of these samples was hexaploid, while all other samples were tetraploid, as expected for *R. damascena*.

As expected, the absolute magnitude of genetic distances based on codominant scoring is much smaller than that of dominant scores, as more alleles are shared, but the topologies of the trees (Figure 1a and Figure 1b) are largely comparable for those samples that were not too genetically distant.

**Discussion**

It seems that for commercial rose production only one and the same genotype is used in several countries. This makes it likely that also in Turkey this genotype is being used for large-scale production, but this remains to be confirmed as samples from Turkey were not included in the study of Rusanov et al. [3] nor in the present study. Except one plant, all genotypes identified here were tetraploid, consistent with the general literature. One plant was hexaploid. At this moment, we do not know whether this is the first of more hexaploid *R. damascena* plants. It may be misclassified, but cuttings from all plants have been evaluated by several experienced taxonomists after cultivation for 2–3 years in a common garden.

The genetic distances among accessions were not correlated with geographical distances among their places of origins (not shown). Clearly, a larger sample of genotypes will be necessary to determine whether there is some relationship with geographical distance, whether there is isolation of populations due to barriers in gene flow, or whether different climatic conditions lead to differentiation within the species.

In MAC-PR analysis we determined the allelic configuration based on six loci, because in the other three loci, not all alleles were present in plants in completely heterozygous configurations, which is necessary to be able to accurately determine the relative amplification of each allele [16]. Genotype G_II and G_III differ by only one allele at locus RHEO506. This is surprising as genotypes in roses are usually identical (due to vegetative propagation).
or very different (due to segregation of alleles from the heterozygous parents) [17]. Remarkably, this small difference is confirmed in the MAC-PR analysis, as no differences were found in allele frequencies at the other five loci. Although this does not completely rule out that the two plants are close relatives, a mutation leading to an allele that is one repeat longer is a more likely possibility. Genotype G_III was from Qom, which borders the three provinces in which genotype G_II was found.

Table 1: Geographical origins of Iranian Damask rose accessions

| Origin site no. | Province(s) included | Accession name | Climate<sup>a</sup> | Genotype<sup>b</sup> |
|----------------|----------------------|---------------|---------------------|---------------------|
| Os1            | Isfahan              | Isf01         | Cool temperate – semi arid | G_I               |
|                |                      | Isf02         |                     | G_I               |
|                |                      | Isf03         |                     | G_I               |
|                |                      | Isf04         |                     | G_I               |
|                |                      | Isf05         |                     | G_I               |
|                |                      | Isf06         |                     | G_I               |
|                |                      | Isf07         |                     | G_I               |
|                |                      | Isf08         |                     | G_I               |
|                |                      | Isf09         |                     | G_I               |
|                |                      | Isf10         |                     | G_I               |
| Os2            | East & West Azarbayjan, Ardabil | EastAzar    | Cool temperate – semi arid | G_II              |
|                |                      | WestAzar     |                     | G_V               |
|                |                      | Ardabil      |                     | G_V               |
| Os3            | Kermanshah, Eilam    | Kermanshah   | Temperate – semi humid | G_VII             |
|                |                      | Eilam        |                     | G_V               |
| Os4            | Tehran, Markazi      | Tehran       | Cool temperate – semi arid | G_VI              |
|                |                      | Arak         |                     | G_I               |
| Os5            | Chaharmahall, Kohkiloie, Lorestan | Chaharmahall | Temperate – semi arid | G_I               |
|                |                      | Kohkiloie    |                     | G_I               |
|                |                      | Lorestan     |                     | G_I               |
| Os6            | Razavi Khorasan, South Khorasan | Khor01      | Temperate – arid    | G_I               |
|                |                      | Khor02       |                     | G_I               |
| Os7            | Khoozestan, Hormozgan, Baloochastan | Khooz     | Warm – arid        | G_I                |
|                |                      | Hormoz      |                     | G_I               |
|                |                      | Baloochastan |                     | G_I               |
| Os8            | Zanjan, Qazvin       | Zanjan       | Cool temperate – semi arid | G_II              |
|                |                      | Qazvin       |                     | G_II              |
| Os9            | Semnan, Qom          | Semnan01     | Warm temperate – arid | G_I               |
|                |                      | Semnan02     |                     | G_I               |
|                |                      | Qom          |                     | G_I               |
| Os10           | Fars, Kerman         | Fars01       | Temperate – semi arid | G_I               |
|                |                      | Fars02       |                     | G_I               |
|                |                      | Kerman       |                     | G_I               |
| Os11           | Kurdistan, Hamedan   | Kurdistan    | Cool – semi arid   | G_I               |
|                |                      | Hamedan      |                     | G_I               |
| Os12           | Guilan, Mazandaran, Golestan | Guilan    | Temperate – humid  | G_VIII             |
|                |                      | Mazan        |                     | G_IV               |
|                |                      | Golestan     |                     | G_IV               |
| Os13           | Yazd                 | Yazd01       | Warm temperate – arid | G_I                |
|                |                      | Yazd02       |                     | G_I               |

<sup>a</sup> Yearly mean temperature in warm, temperate and cool climates are 15–25°C, 10–15°C and 0–5°C, respectively. Yearly mean rainfalls in semi humid, semi arid and arid climates are 600–1400 mm, 300–600 mm and 100–300 mm, respectively.

<sup>b</sup>Genotypes as identified in this study

Conclusion

Our analysis showed for the first time the existence of multiple genotypes within *Rosa damascena*. We are currently performing an analysis of oil production across several years, in order to determine whether different genotypes also have a qualitative difference in production and/or composition of essential oil. If so, these genotypes may be used to broaden the production of rose oil, and they can also be used as the basis of a breeding program.
As these nine genotypes were found after sampling only 40 large and small production fields, we expect that a more intensive sampling will be valuable in order to find more genetic diversity. For this, we will focus on the areas where we have found the unique genotypes, i.e., the Western and Northern provinces.

**Methods**

**Plant material**

A total of 40 Damask rose accessions were collected from 28 provinces of Iran (Table 1), in order to obtain a good geographical coverage of the country and a good coverage of the 13 different climatic regions that have been identified [13]. Samples were taken from commercial production fields and from small (< 5 ha) or abandoned production fields. All accessions were grown from 2000 onwards in experimental field of the Research Institute of Forests and Rangelands (RIFR), Tehran, Iran. DNA was extracted from fresh young leaves using the Qiagen DNeasy Plant Mini Kit (Westburg, The Netherlands).

**Microsatellite analysis**

A set of nine robust microsatellite markers were selected from Esselink et al. [17] and Yan et al. [15] representing different linkage groups on the genetic map of rose (Table 2). These markers are highly polymorphic in hybrid tea rose [17] and in other Rosa species [18-20], and hence have a high discriminative power to differentiate genotypes. Fluorescently labelled (6FAM, HEX or NED) primer pairs were amplified in three multiplexes using the Qiagen PCR multiplex kit (Westburg, The Netherlands). The PCR program for amplification were as follows: 94°C for 15 min; 30 cycles of 94°C for 30 s, ramp to 50°C (1°C/s), 50°C for 30 s, ramp to 72°C (1°C/s), 72°C for 2 min; and a final elongation step at 72°C for 10 min. Fluorescent amplification products were detected using an ABI Prism 3700 DNA Analyzer (Applied Biosystems) and all samples were genotyped in accordance with reference alleles for each locus as described by Vosman et al. [21], using Genotyper 3.5 NT (Applied Biosystems).

**MAC-PR and statistical analysis**

The microsatellite DNA allele counting – peak ratios method (MAC-PR), which was developed for the tetraploid hybrid tea rose (Rosa × hybrida L.) varieties by Esselink et al. [16], assigns precise allelic configurations (the actual genotype) based on quantitative values for peak areas provided by the Genotyper software. For each

| Locus | Label | Linkage group | Number of alleles |
|-------|-------|---------------|------------------|
| RhPS19 | 6FAM | n.d. | 6 |
| RhB303 | HEX | n.d. | 11 |
| RhOS17 | NED | 1 | 5 |
| RHEO506 | 6FAM | 2 | 13 |
| RhD221 | HEX | 4 | 7 |
| RhAB73 | NED | 7 | 9 |
| RhPS50 | 6FAM | 3 | 15 |
| RhAB40 | HEX | 4 | 8 |
| RhE2b | NED | 6 | 8 |
| Average | | | 9.11 |

from Debener et al. [14] and Yan et al. [15]

n.d. = not determined

Table 2: Characteristics of the microsatellite markers used.

| Locus | Label | Linkage group | Number of alleles |
|-------|-------|---------------|------------------|
| RhPS19 | 6FAM | n.d. | 6 |
| RhB303 | HEX | n.d. | 11 |
| RhOS17 | NED | 1 | 5 |
| RHEO506 | 6FAM | 2 | 13 |
| RhD221 | HEX | 4 | 7 |
| RhAB73 | NED | 7 | 9 |
| RhPS50 | 6FAM | 3 | 15 |
| RhAB40 | HEX | 4 | 8 |
| RhE2b | NED | 6 | 8 |
| Average | | | 9.11 |

Table 3: Allele configuration of the nine different R. damascena genotypes based on MAC-PR analyses
**Figure 1**

1a UPGMA clustering of Dice genetic similarities based on dominant scores of microsatellite alleles, among all accessions of Damask rose included in this study. Note that 1 (similarity) = genetically identical. 1b UPGMA clustering of genetic distances based on pairwise Fst among the nine unique genotypes, derived from codominant scores of six microsatellite loci. Note that 0 (distance) = genetically identical.
locus, all alleles were analyzed in pairwise combinations in order to determine their copy number in the individual samples. This was accomplished by calculating ratios between the peak areas for two alleles in all samples in which these two alleles occurred together.

Genetic distances were calculated either as Dice similarities on the basis of dominant scoring of individual alleles in NTSYS 2.1 (Applied Biostatistics) or as pairwise Fst of the MAC-PR genotypes using SPAGEDi 1.2 [22]. The use of Dice (Nei & Li) coefficient is more suitable for codominant markers such as SSRs when they are scored dominantly [23,24]. The accessions were clustered using the unweighted pair group method using arithmetic averages (UPGMA) module of NTSYS.

**Authors’ contributions**

SRTA established the Damask rose collection. AB, SRTA, MKK, MRN and RO designed the study. AB selected plant material and performed DNA extraction. AB, GDE and MJMS performed SSR and data analysis. AB, MJMS, GDE and MRN wrote the primary draft. All authors were involved in the final version of the paper.

**Acknowledgements**

The authors would like to thank Yolanda Noordijk for her kind assistance in laboratory procedures. Also, we acknowledge Ivan Atanassov, AgroBio-Institute, Sofia, Bulgaria and Natasha Kovacheva, Institute of Rose and Aromatic Plants, Kazanlak, Bulgaria for providing leaf material of Bulgarian damask roses. Ben Vosman and Paul Arens are greatly acknowledged for their critical comments. This work is partly financed by the Ministry of Science and Technology of Iran (MSRTI) through a travel grant for A. Babaei.

**References**

1. Gudin S: Rose: genetics and breeding. Plant Breed Rev 2000, 17:159-189.

2. Horn WAH: Micropropagation of rose. In Biotechnology in agriculture and forestry Volume 4. Edited by: Bajaj YPS. Springer-Verlag, Berlin; 1992:320-324.

3. Rusanov K, Kovacheva N, Vosman B, Zhang L, Rajapakshe S, Atanassov A, Atanassov I: 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, 111:804-809.

4. Achuthan CR, Babu BH, Padikkala J: Antioxidant and hepatoprotective effects of Rosa damascena. Pharmaceutical Biology 2003, 41:357-361.

5. Ardogan BC, Baydar H, Kaya S, Demirci M, Ozbazar D, Mumcu E: Antimicrobial activity and chemical composition of some essential oils. Archives of Pharmacal Research 2002, 25:860-864.

6. Basim E, Basim H: Antibacterial activity of Rosa damascena essential oil. Fitoterapia 2003, 74:394-396.

7. Daskan G, Sagdic O, Baydar NG, Baydar H: Antioxidant and antibacterial activities of Rosa damascena flower extracts. Food Sci Technol Int 2004, 10:277-281.

8. Agaoglu Y, Ergul A, Baydar N: Molecular analyses of genetic diversity of oil rose (Rosa damascena Mill.) grown in Isparta (Turkey) region. Biotechnol Biotechnol Eq 2000, 14:16-18.

9. Baydar N, Baydar H, Debener T: Analysis of genetic relationships among Rosa damascena plants grown in Turkey by using AFLP and microsatellite markers. J Biotechnol 2004, 111:263-267.

10. Beales P, Cairns T, Duncan W, Fagan G, Grant W, Grapes K, Harkness P, Hughes K, Mattock J, Ruston D, Sutherland P, Williams T: Botanica’s roses. The encyclopedia of roses. Random House, Australia; 1998.

11. Chevalier A: The Encyclopedia of Medicinal Plants. London, UK: Dorling Kindersley; 1996.

12. Saakoy SG, Rieksta DA: Roses. Zinatne, Riga; 1973. (in Russian)

13. Tabaei-Aghdaei SR, Babaei A, Khosh-Khui M, Jaimand K, Rezsee MB, Assareh MH, Naghavi MR: Morphological and oil content variations amongst Damask rose (Rosa damascena Mill.) landraces from different regions of Iran. Sci Hortic in press.

14. Debener TL, Mattiesch L, Vosman B: A molecular map for roses. Acta Hortic 2001, 547:283-287.

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

16. Esselink GD, Vosman B: Assignment of allelic configurations in polyploids using the MAC-PR (microsatellite DNA allele counting – peak ratios) method. Theor Appl Genet 2004, 109:402-408.

17. Esselink D, Smulders MJM, Vosman B: Identification of cut-rose (Rosa hybridra) and rose varieties using Sequence Tagged Microsatellite markers. Theor Appl Genet 2003, 106:277-286.

18. Nybom H, Esselink GD, Werlemark G, Leus L, Vosman B: Unique genomicsc configuration revealed by microsatellite DNA in polyplody dogroses, Rosa sect. Caninae. J Ecol Bio 2006, 19:635-648.

19. Rusanov K, Kovacheva N, Atanassov A: Microsatellite analysis of oil-bearing roses which do not belong to the Species Rosa damascena Mill. Bulg J Agric Sci 2005, 11:1-9.

20. Vosman B, Esselink D, van Eeuwijk F: The use of microsatellites for identifying putative edv’s in rose. UPOV document BM/18/16 2003.

21. Vosman B, Esselink D, Smulders R: Microsatellite markers for identification and registration of rose varieties. UPOV document BM/1WO/Rose/1/2001.

22. Hardy OJ, Yekemans X: SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes 2002, 2:618-620.

23. Engvist DM, Becker HC: Genetic diversity for alloxymes, AFLPs and RAPDs in resynthesized rape. In Proceedings of the Ninth Meeting of the EUCARPIA Section Biometrics in Plant Breeding: 6–8 July 1994; Wageningen, the Netherlands Edited by: van Ooijen JW, Jansen J. Wageningen: CPRO-DLO; 1994:85-90.

24. Link W, Dickens C, Singh M, Schwall M, Melchinger AE: Genetic diversity in European and Mediterranean fabea bean germ plasm revealed by RAPD markers. Theor Appl Genet 1995, 90:27-32.