DNA MARKER TIGHTLY LINKED TO THE DOUBLE
FLOWER LOCUS IN *ROSA* × *DAMASCENA* MILL. F.
TRIGINTIPETALA SUITABLE FOR MARKER ASSISTED
BREEDING

Krasimir Rusanov[#], Mila Rusanova, Natasha Kovacheva*,
Ivan Atanassov

Received on September 26, 2021

Presented by A. Atanassov, Member of BAS, on September 29, 2021

Abstract

The present study reports the development of a DNA marker tightly linked to the double flower (DF) phenotype in *R. × damascena*, suitable for marker assisted breeding. The development of this marker is based on previous identification of genome region associated with the DF phenotype in *R. × damascena* and recent studies demonstrating involvement of the AP2/TOE gene in the formation of DF in several species from the Rosaceae family. We identified a polymorphic region named RdDF upstream of the AP2/TOE gene in *R. × damascena* and used the sequence data for designing PCR primers. Segregation analysis of the four alleles of RdDF in a population of 104 plants obtained after self-pollination of *R. × damascena* showed that one of the alleles (RdDF-B, 203 bp) is tightly linked to the DF phenotype. Analysis of the segregation of two SSR markers located at a distance of approximately 1578 kb and 1129 kb on both sides of RdDF, showed that walking away from the developed RdDF marker results in the observation of simple flower recombinant plants, supporting association of RdDF with the DF phenotype. The application of the developed RdDF marker for marker assisted selection (MAS) and breeding programmes involving *R. × damascena*, as well as further study of the molecular base of DF formation in this species are discussed.

Key words: *R. × damascena*, double flower, AP2/TOE gene, DNA marker, marker assisted breeding

[#]Corresponding author.

This study was supported by the National Science Fund at the Ministry of Education and Science, Bulgaria, under grant DN 06/2.

DOI:10.7546/CRABS.2022.02.17

Introduction. Nearly the entire worldwide production of rose essential oil is based on industrial cultivation of the single tetraploid rose variety *R. × damascena* Mill. f. *trigintipetala* [1–3]. The main part of the distilled rose oil originates from the volatile compounds accumulated in the epidermal tissue of the rose flower petals [4,5]. This makes the double flower (DF) phenotype indispensable for breeding of oil-bearing roses involving this rose variety.

Previous studies related to the development of genetic maps in roses and mapping traits of interest pointed that DF is a dominant monogenic trait and the locus which determines the trait designated as *Blfo* is located on chromosome 3 in diploid roses [6,7]. Dominant allele of this locus is responsible for the major switch between simple flowers and the much more desired semi-double and double flowers. Recently, we demonstrated that the same locus also controls the trait in the tetraploid *R. × damascena* [8]. In our previous study, we showed that the DF phenotype in this species is determined by a single allele of a gene located in the vicinity of the SSR locus Rh58. Although one of the four alleles of Rh58 showed linked inheritance with the DF trait, a recombination between the actual gene determining the trait and Rh58 was observed and in 8% of the cases, the presence of the allele was not associated with the DF trait [8]. The latter pointed out that further efforts are necessary to develop a DNA marker which could be used for breeding without a compromise of a potential recombination.

A recent study suggested that the DF trait in *R. chinensis* cv. ‘Old Blush’ and several modern roses based on this variety is determined by the presence of a transposable element (TE) in one of the alleles of the AP2/TOE gene which is physically located in the region of the *Blfo* locus in *R. chinensis* [9,10]. According to the authors, the integration of TE in intron 8 of AP2/TOE has caused a premature stop codon and the loss of exon 9 and exon 10 causing the formation of a truncated protein. The integration of the TE has also caused the loss of a miR172 recognition site which consequently leads to increased availability of AP2 transcripts. Disruption of the miR172 recognition site caused by a deletion in the last exon and 3' UTR of AP2/TOE in *R. rugosa* was found to be the cause for the DF phenotype in this rose species [11].

Here we studied the sequence of the intergenic region between the AP2/TOE gene and the 60S ribosomal protein L4 gene in *R. × damascena* as a source for polymorphisms and development of a tightly linked DNA marker to the DF phenotype in *R. × damascena* which can be directly applied in marker assisted breeding programmes. We further discuss the potential presence of a TE in the AP2/TOE gene in *R. × damascena* as a cause for the DF phenotype in this rose species.

Materials and methods. Plant material, flower morphology and genomic DNA extraction. Plant material from *Rosa × damascena* Mill. f. *trigintipetala* as well as from 104 plants obtained after its self-pollination was collected in June 2019. At least three fully developed flowers corresponding to

phase 6 according to RUSANOV et al. [12] were collected from each plant and used for determination of the flower type. Flowers with 6 ± 1 petals, $7 < 20$ petals and > 20 petals were scored as simple, semi-double and double, respectively. Genomic DNA (gDNA) was purified from leaves using a modified CTAB protocol [13].

Cloning, sequencing and analysis of polymorphic regions of the intergenic spacer between the AP2/TOE and 60S ribosomal protein L4 genes in *R. × damascena*. The primer pair 60SAP2F and 60SAP2R was used for PCR amplification of the intergenic region between the genes AP2/TOE and 60S ribosomal protein L4 in *R. × damascena* (Table 1, Fig. 1b). The PCR reaction included 50 ng of gDNA from *R. × damascena*, 10 µL of 5× Phusion buffer (New England Biolabs), 1 µL 10 mM dNTPs, 25 pmol from each primer, 0.5 µL of Phusion® High-Fidelity DNA Polymerase (New England Biolabs) and 31.5 µL of Milli-Q® water. The PCR conditions were: 30 s at 98 °C, then 35 cycles of 10 s at 98 °C, 30 s at 59.8 °C, 2 min 10 s at 72 °C, and final elongation at 72 °C for 10 min. The obtained single PCR product was purified following electrophoresis in 1% agarose and cloned in pJET1.2/blunt vector (Thermo). The obtained plasmids were transformed in *E. coli* Dh5a. Plasmids from individual colonies were used for sequencing with pJET Fw, pJET Rev as well as internal primers. Sequencing was performed by MacroGen Inc. PCR primers for the polymorphic locus RdDF (Table 1) were designed using Primer 3 Plus [14]. PCR amplification of the RdDF locus was performed using 50 ng of gDNA, 10 pmol of RdDFF labelled with FAM, 10 pmol of RdDFR, 10 µL of 2× HS Mix (Bioline) and 6 µL of Milli-Q® water. The PCR conditions were: 94 °C for 5 min, then 30 cycles of 30 s at 94 °C, 30 s at 54 °C, 30 s at 72 °C and final elongation for 7 min at 72 °C.

T a b l e 1

PCR primers used in the current study. Uppercase letters indicate the 5' M13 tail for SSR primers

Primer name	Primer sequence
60SAP2F	5' agagttcgatgtcttccaaa 3'
60SAP2R	5' tcgggagtgtagaagttgag 3'
RdDFF	5' atttctgtgcatgaaaatc 3'
RdDFR	5' cgtttgtttacgtgagtga 3'
AP2R	5' ccatcattaagacttgcat 3'
RdSSR49F	5' GTAAAACGACGGCCAGTttcctctttaccttcagacg 3'
RdSSR49R	5' cagacaccaagcctgtat 3'
RdSSR154F	5' GTAAAACGACGGCCAGTaatttctgaagagccacatc 3'
RdSSR154R	5' aaagcttgtagaagccaga 3'
M13F	5' gtaaaacgacggccagt 3'
pJET Fw	5' cgactcactataggagagcggc 3'
pJET Rev	5' aagaacatcgattttcatggcag 3'

Fig. 1. Schematic representation of: a) the AP2/TOE gene in *R. chinensis* cv. 'Old Blush'. Grey rectangles represent exons. Smaller black rectangles represent 5' and 3' UTR. The position of TE in one of the two alleles of the gene in *R. chinensis* is marked above intron 8, b) the intergenic region between the 60S ribosomal protein L4 and AP2/TOE genes in *R. × chinensis*, c) the position of SSR markers RdSSR154 and RdSSR49 located on both sides of the RdDF locus on chromosome 3 in the genome of *R. chinensis*. Distances are present in kilobase pairs (kb). Primer positions are designated with arrows and their names according to Table 1

Fragment analysis was performed on a 3130 Genetic Analyzer (Thermo) using 36 cm long capillaries, POP-7™ polymer (Thermo) and GeneScan™ 500 LIZ™ as a size standard (Thermo).

Development and analysis of SSR markers based on the *R. chinensis* genome sequence. The genomic sequence of chromosome 3 from *R. chinensis* was used for extraction of SSR sequences on both sides of the RdDF locus with the help of SciRoKo 3.4 [15]. SSR primers were designed using Primer 3 Plus [14]. The amplification of RdSSR49 and RdSSR154 was performed using a two-step PCR protocol. The first step PCR included 50 ng of gDNA, 10 pmol of each SSR primer, 10 µL of 2× AmpliTaq Gold® PCR mix (Thermo) and 6 µL of Milli-Q® water. The PCR conditions were: 10 min at 95 °C, then 30 cycles of 15 s at 95 °C, 15 s at 50 °C, 30 s at 72 °C and final elongation for 7 min at 72 °C. The second step PCR included 2 µL from the first PCR reaction as DNA template, 10 pmol of M13 primer 5'-labelled with FAM, 10 pmol of the reverse SSR primer, 10 µL of 2× AmpliTaq Gold® PCR mix (Thermo), 6 µL of Milli-Q® water and the same conditions as for the first step PCR. The PCR reaction was used for fragment analysis as described earlier.

PCR amplification and fragment analysis of the region spanning the RdDF marker and exon 10 of the AP2/TOE gene. The primer pair consisting of RdDFF and AP2R (Table 1) was used for amplification of the region spanning the RdDF marker and exon 10 of AP2/TOE in *R. × damascena*. The PCR reaction included 50 ng of gDNA, 10 pmol of each forward and reverse primer, 10 µL of 2× HS Mix (Bioline) and 6 µL of Milli-Q® water. The PCR conditions were: 94 °C for 5 min, then 35 cycles of 30 s at 94 °C, 30 s at 54 °C, 2 min at 72 °C and final elongation for 7 min at 72 °C. The obtained PCR product with approximate size of 3 kb was purified following electrophoresis in 0.8% agarose gel using GeneJET Gel Extraction Kit (Thermo). The purified product was used for PCR amplification with the primer pair RdDFF and RdDFR (Table 1). The PCR reaction was performed as described earlier for gDNA amplification of the same locus as 1 µL of purified PCR product was used as a template. Fragment analysis was performed as described earlier.

Results and discussion. Development of a tightly linked marker to the locus controlling the double flower phenotype in *R. × damascena*. We utilized the intergenic region upstream of the AP2/TOE gene in search of a polymorphic DNA marker and used the sequence of the AP2/TOE gene (RchiOBHmChr3g0468481) and the 60S ribosomal protein L4 gene (RchiOBHmChr3g0468471) from the *Rosa chinensis* cv. 'Old Blush' genome for designing primers (60SAP2F and 60SAP2R, Table 1, Fig. 1b) and cloning of the intergenic region between the two genes from *R. × damascena*. The obtained PCR product showed approximate size of 4 kb as estimated by agarose gel electrophoresis, which was in compliance with the size of the same region in *R. chinensis*. The PCR product was cloned in plasmid and transformed in *E. coli*. Sequencing of

Fig. 2. Electropherogram of PCR product of the RdDF locus from: a) the DF genotype of *Rosa* $\times$ *damascena* Mill. f. *Trigintipetala*, and b) a simple flower genotype obtained after self-pollination of *R.* $\times$ *damascena*. Red arrow shows the 207 bp allele which is tightly linked to the double flower phenotype in *R.* $\times$ *damascena*

the cloned product from 16 randomly chosen single colonies revealed at least three different alleles of the intergenic region based on the observed multiple polymorphisms due to indels ranging from 1 to 21 nucleotides as well as SNPs (data not shown). A primer pair (RdDFF and RdDFR) was designed for PCR amplification of one of the target polymorphic regions from *R.* $\times$ *damascena* (Table 1, Fig. 1a, b). Fragment analysis of the obtained PCR product revealed a total of four different alleles (Fig. 2) with sizes of 203, 207, 219 and 223 bp. Analysis of the segregation of the four alleles in a population of 104 plants obtained after self-pollination of *R.* $\times$ *damascena* showed that the 207 bp allele is linked to the DF phenotype in all analysed plants (Fig. 3). We also studied two SSR markers positioned on both sides of the analysed locus at a distance of 1578 kb (RdSSR154)

Fig. 3. Box-and-whisker plot showing the distribution of data for number of petals among plants with different allele phenotypes of the RdDF locus in a segregating population of 104 plants obtained by self-pollination of *R. × damascena* Mill. f. *trigintipetala*. A, B, C and D designate the four different alleles of the RdDF locus (A = 203 bp, B = 207 bp, C = 219 bp and D = 223 bp)

and 1129 kb (RdSSR49) from the position of the RdDF marker based on the reference genome of *R. chinensis* (Fig. 1c). Marker RdSSR154 had three alleles in *R. × damascena* with sizes of 189, 195 and 211 bp while marker RdSSR49 also had three alleles with sizes of 269, 297 and 313 bp. Analysis of the segregation of the alleles in the segregating population showed that one of the alleles of both markers (195 bp for RdSSR154 and 313 bp for RdSSR49) is linked to the DF trait. However, recombination events were observed in three plants for marker RdSSR154 and two for RdSSR49 thus pointing that walking in both directions from RdDF along chromosome 3 results in moving away from the actual DF locus. The results further support the tight link of the RdDF region & AP2/TOE gene with the DF formation in *R. × damascena*.

In order to get insight about the possible cause for the DF formation involving the AP2/TOE gene, we amplified large part of the gene region spanning from the RdDF marker in the promoter region to exon 10 of the AP2/TOE gene from the original genotype of *R. × damascena* using primers RdDFF and AP2R (Table 1, Fig. 1a). The same region in the sequenced genome of the diploid *R. chinensis* cv. ‘Old Blush’ includes two alleles as one of the alleles harbours a transposable element (TE) with a size of approximately 10.8 kb inserted in intron 8 (Fig. 1a) [9,10]. Agarose gel electrophoresis revealed a single PCR product with an approximate size of 3 kb from *R. × damascena* which corresponded well to the size of the allele from the *R. chinensis* cv. ‘Old Blush’ not containing the TE (Fig. 1a) [9]. We purified the PCR product from the gel and used it for amplification with the primer pair developed for the RdDF marker. Fragment analysis revealed the same profile with all four RdDF alleles as observed with amplification of *R. × damascena* gDNA. We therefore conclude that a TE of a size similar to that observed in *R. chinensis* the AP2/TOE gene [9] is not present in *R. damascena* Mill. f. *trigintipetala* and is not the cause of the DF phenotype in this rose species. The DF phenotype in *R. × damascena* could be due also to other types of disruption of the expression pattern of the identified RdDF & AP2/TOE allele for dominant DF formation. For example, the small deletion in the mir172 recognition site located within one of the AP2/TOE alleles was reported to be the cause of the DF phenotype in *R. rugosa* and other members of the Rosaceae family [11,16]. Based on the present research, next comparative sequence and expression characterization of the identified RdDF & AP2/TOE allele of *R. × damascena* could reveal the actual cause and mechanism of the DF phenotype in this species.

Conclusions. The current study results in development of an efficient PCR based DNA marker for a single step analysis of all four alleles of the RdDF region & AP2/TOE gene, linked to DF phenotype in *R. × damascena* Mill. f. *trigintipetala*. Characterization of a segregating population from self-pollinated *R. × damascena* demonstrated the tight linkage between one of the alleles of the RdDF marker and the dominant DF phenotype in the analysed plants. Considering the importance of the DF phenotype for rose oil production, the developed RdDF marker could be

directly used as an essential first step for marker assisted pre-selection of seedlings which will develop double flowers in breeding programmes of oil-bearing roses involving *R. × damascena*. Moreover, the results of the study offer a good starting point for further research and understanding of the molecular mechanism of the DF formation in *R. × damascena* through sequence and expression analysis of the identified RdDF & AP2/TOE allele for dominant DF phenotype.

REFERENCES

- [1] RUSANOV K., N. KOVACHEVA, B. VOSMAN, L. ZHANG, S. RAJAPAKSE et al. (2005) 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.*, **111**(4), 804–809.
- [2] BAYDAR N. G., H. BAYDAR, T. DEBENER (2004) Analysis of genetic relationships among *Rosa damascena* plants grown in Turkey by using AFLP and microsatellite markers, *J. Biotechnol.*, **111**(3), 263–267.
- [3] KOVACHEVA N., K. RUSANOV, I. ATANASSOV (2010) Industrial cultivation of oil bearing rose and rose oil production in Bulgaria during 21st century, directions and challenges, *Biotechnol. Biotec. Eq.*, **24**(2), 1793–1798.
- [4] STAIKOV V., G. ZOLOTOVICH (1957) Localization of the essential oil in the flower of *Rosa damascena* Miller, *Bulletin of the Bulgarian Institute of Plant Breeding*, **IV**, 207–215 (in Bulgarian).
- [5] BERGOUNOUX V., J. C. CAISSARD, F. JULLIEN, J. L. MAGNARD, G. SCALLIET et al. (2007) Both the adaxial and abaxial epidermal layers of the rose petal emit volatile scent compounds, *Planta*, **226**(4), 853–866.
- [6] DEBENER T., L. MATTIESCH (1999) Construction of a genetic linkage map for roses using RAPD and AFLP markers, *Theor. Appl. Genet.*, **99**(5), 891–899.
- [7] SPILLER M., M. LINDE, L. HIBRAND-SAINT OYANT, C. J. TSAI, D. H. BYRNE et al. (2011) Towards a unified genetic map for diploid roses, *Theor. Appl. Genet.*, **122**(3), 489–500.
- [8] RUSANOV K., N. KOVACHEVA, M. RUSANOVA, M. LINDE, T. DEBENER et al. (2019) Genetic control of flower petal number in *Rosa × damascena* Mill. f. *trigintipetala*, *Biotechnol. Biotec. Eq.*, **33**(1), 597–604.
- [9] FRANCOIS L., M. VERDNAUD, X. FU, D. RULEMAN, A. DUBOIS et al. (2018) A miR172 target-deficient AP2-like gene correlates with the double flower phenotype in roses, *Sci. Rep.*, **8**(1), 12912.
- [10] RAYMOND O., J. GOUZY, J. JUST, H. BADOVIN, M. VERDNAUD et al. (2018) The *Rosa* genome provides new insights into the domestication of modern roses, *Nat. Genet.*, **50**(6), 772–777.
- [11] GATTOLIN S., M. CIRILLI, S. CHessa, A. STELLA, D. BASSI et al. (2020) Mutations in orthologous PETALOSA TOE-type genes cause a dominant double-flower phenotype in phylogenetically distant eudicots, *J. Exp. Bot.*, **71**(9), 2585–2595.
- [12] RUSANOV K., N. KOVACHEVA, M. RUSANOVA, I. ATANASSOV (2011) Traditional *Rosa damascena* flower harvesting practices evaluated through GC/MS metabolite profiling of flower volatiles, *Food Chem.*, **129**(4), 1851–1859.

- [13] MURRAY M., W. THOMPSON (1980) Rapid isolation of high molecular weight plant DNA, *Nucleic Acids Res.*, **8**(19), 4321–4326.
- [14] UNTERGASSER A., H. NIJVEEN, X. RAO, T. BISSELING, R. GEURTS et al. (2007) Primer3Plus, an enhanced web interface to Primer3, *Nucleic Acids Res.*, **35** (Web Server Issue), W71–4.
- [15] KOFLER R., C. SCHLOTTERER, T. LELLEY (2007) SciRoKo: a new tool for whole genome microsatellite search and investigation, *Bioinformatics*, **23** (13), 1683–5.
- [16] GATTOLIN S., M. CIRILLI, I. PACHECO, A. CIACCIULLI, C. DA SILVA LINGE, et al. (2018) Deletion of the miR172 target site in a TOE-type gene is a strong candidate variant for dominant double-flower trait in Rosaceae, *Plant J.*, **96** (2), 358–371.

Agrobiotechnology Department
Agrobiointitute
Agricultural Academy
8 Dragan Tsankov Blvd
1164 Sofia, Bulgaria
 e-mails: `krusanov@abv.bg`
 `milagradeva@abv.bg`
 `ivan_atanassov@abv.bg`

**Roza Select Ltd*
19 Ivan Vazov St
6139 Cherganovo, Bulgaria
 e-mail: `n.kovatcheva@abv.bg`