E

BMT/10/16
ORIGINAL: English
DATE: November 9, 2006

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS
GENEVA

WORKING GROUP ON BIOCHEMICAL AND MOLECULAR TECHNIQUES AND DNA PROFILING IN PARTICULAR

Tenth Session
Seoul, November 21 to 23, 2006

A EUROPEAN REFERENCE COLLECTION OF ROSE VARIETIES

Document prepared by experts from Germany, Netherlands and United Kingdom

A EUROPEAN REFERENCE COLLECTION OF ROSE VARIETIES

Ben Vosman[1], Elizabeth Scott[3], Burkhard Spellerberg[4],
Joost Barendrecht[2], Swenja Tams[4], Huw Jones[3] and Danny Esselink[1]

[1] Plant Research International, P.O. Box 16, 6700 AA Wageningen, Netherlands
[2] Dutch Plant Variety Board, Postbus 27, 6710 BA Ede, Netherlands
[3] NIAB, Huntingdon Road, Cambridge CB3 0LE, United Kingdom
[4] Bundessortenamt, Osterfelddamm 80, 30627 Hannover, Germany

Summary

1. An integrated pilot database was constructed containing administrative, morphological and molecular data as well as pictures of each variety. All data for one variety can be shown in one screen. The selected morphological characteristics are all useful for selecting reference varieties, although the degree of usefulness varies according to the type of rose: for example, in greenhouse cut-flower roses, most varieties currently fall into the same flower and plant growth type.

2. The pictures taken for each variety are also considered important, although the composite photo was considered less informative for the cut-flower greenhouse roses, where in comparison to the garden roses there is a more limited variation in the characteristics photographed as far as the non floral parts are concerned.

3. Microsatellite markers have been used to construct a database containing the molecular profiles of approximately 380 varieties. The markers proved to be very informative about the varieties. The database can support and even improve the quality of DUS testing and the quality of protection.

Introduction

4. Rose is the largest ornamental crop and the most important one in many countries. Over 25,000 varieties of modern roses have been described (Cairns, 2000). The first hybrid tea rose was introduced in 1867 and since then more than 10,000 hybrid teas have entered the market. Such large numbers of varieties cause problems in the DUS testing context. A major issue for all countries carrying out DUS tests is the requirement to compare new varieties with an increasing number of existing reference varieties. Clearly, strict adherence to this concept is logistically and financially very difficult in a species such as rose, which is cultivated around the world, although the number of varieties to be considered can be limited by climatic factors, variety type and the continued availability of material. Nevertheless, this still means that many hundreds of existing varieties should be taken into account and the number is increasing all the time, which constantly adds to the costs of testing. In order to reduce these increasing costs and to improve the exchange of information about varieties, a way of managing information about the large number of reference varieties, and of selecting most similar varieties for inclusion in the growing trials is necessary. If, at the same time, access to shared information about existing varieties could be improved, this would improve the efficiency of the DUS testing system. Applicants complete a technical questionnaire, which is submitted to the examination office together with a photograph of the candidate variety.

Candidate varieties must be evaluated against all the relevant possibly similar varieties whose existence is a matter of common knowledge. Thus, a fast and systematic approach for the selection of varieties essential for direct comparison in the growing test is clearly needed to ensure robust results. This selection is currently undertaken in various ways, for example by comparison with an existing, dedicated reference collection, comparison with the collections in public rosariums, literature searches, searches in databases of descriptions and photographs, use of the expertise of the examiners, and the use of outside experts ("walking reference collections"). Having made the selection of similar varieties for a growing trial, these then need to be sourced in case they are not already present in the dedicated reference collection mentioned above. Therefore, the examination office needs to request reference varieties from the breeders. It is important that the examination office can quickly verify the identity of the material submitted. For this aspect of quality assurance, molecular markers are ideally suited, as they are highly discriminating and can be assayed rapidly and relatively cheaply.

5. Several molecular marker systems have been applied to roses. The application of the STMS approach was recently successfully demonstrated (Esselink et al. 2003; Nybom et al 2004; Rusanov et al 2005; Smulders et al 2005).

6. This project aimed to produce a pilot database of rose varieties that would be available to the Community Plant Variety Office (CPVO) testing stations. The database contains not only the molecular profiles of varieties, but also photographs and information on the most important morphological characteristics. Because the database covers different types of rose tested for plant breeders' rights (PBR) and the same set of molecular markers will be used for all varieties, an additional benefit will be the ability to easily cross-check applications for cut-flower glasshouse varieties against outdoor garden varieties, in those cases when varieties are mutations. Currently there is no quick way to do this, so the first part of the distinctness assessment rests on the declaration of the breeder as to the use of the variety - a situation that cannot be checked by the testing station until the first growing trial. After this, further comparative tests may be necessary. Furthermore, it enables an easier exchange of information between testing stations about contemporary varieties in test, both for CPVO and national PBR. At the moment this is based on morphological data and where necessary photographs, both of which might be affected by environmental conditions and hence the information requires careful interpretation by the examiners.

7. This database will improve the quality assurance role outlined above and expand the data available on the range of varieties taken into consideration in each country, effectively improving the management of the reference collections. In addition, as the database is constructed on the basis of molecular profiles produced on the material submitted for PBR, the breeders will have a very effective tool (identification label) for tracing potential infringements - the database can be used to assist in the quicker technical verification of varieties after a grant of PBR. All of this will lead to better possibilities for enforcing PBR and technical verification.

Construction of the database

8. For the construction of an integrated pilot database, several choices needed to be made. These concerned the morphological characteristics to be included, the markers and pictures to be used and the database structure and format. All of these issues are discussed below.

(a) Selection of morphological descriptors to be included in the database

9. During the first two project meetings morphological descriptors were selected from CPVO/TQ-EN-011 to be included in the database. Selection was determined by the robustness of the descriptor, as well as its usefulness for selecting varieties for comparison. It is important that these characteristics are part of the Technical Guideline (TG/11/7) and the Technical Questionnaire (TQ), which will ensure that all descriptors used are scored by the DUS stations and that breeders use them for describing their candidate variety.

10. The following set of descriptors was selected:

4.1 Origin
5.2 Flower: type
5.3 Flower: diameter
5.4 Flower color group
5.5. Plant growth type
7.2.1. Special conditions: group

11. It was concluded that the characteristics 5.4 and 5.5 will be included in the new TG/11/8.

(b) Selection of molecular markers

12. During the first year of the project, Plant Research International (PRI) and NIAB genotyped a selected set of 23 garden rose varieties using 24 STMS markers. From these markers a subset was selected for database building. Criteria used for selection included (1) level of polymorphism, (2) robustness, (3) ease of scoring and as far as possible (4) distribution over the genome. Details on the selected markers can be found in Table 1. The markers RhP50, RhP518 and RhAB73 proved to be useful for garden roses but did not fit the selection criteria for glasshouse roses. Instead the markers RhM405, RhAB15 and RhO507 are included in the core set for genotyping glasshouse roses.

Table 1: Characteristics of the selected set of microsatellite markers for genotyping garden and glasshouse roses. "Selected for" indicates whether the marker is selected for use in garden roses (gr), glasshouse roses (ht) or both (gr/ht).

STMS	Linkage Group	No. of alleles in 23 varieties	No. of allele phenotypes	Selected for	Scoring quality
RhO517	1	5	14	gr/ht	1
RhEO506	2	12	19	gr/ht	1
RhD221	4	8	12	gr/ht	1
RhE2b	6	7	12	gr/ht	1
RhB303	Unknown	6	14	gr/ht	1
RhP519	Unknown	7	15	gr/ht	1
RhAB40	4	11	18	gr/ht	1
RhD201	Unknown	7	10	gr/ht	1
RhAB22	6	12	15	gr/ht	1
RhP50	3	11	13	Gr	1
RhP518	5	7	15	Gr	1
RhAB73	7	9	18	Gr	1
RhM405	Unknown	5	13	Ht	1
RhAB15	2	10	5	Ht	1
RhO507	4	14	18	Ht	1

(c) Selection of photo format

13. During the first meeting of the partners it was agreed that two pictures would be made:

(a) Flower from the top, on grey background (figure 1)
(b) A composite photo containing open flower from top and bottom, a bud, and a leaf (figure 2).

14. A ruler was included, marked in centimeters. Photos were labeled with CPVO number and national number. All photos were made in jpg format. For greenhouse and garden roses approximately 100 varieties were photographed for the database.

(d) Selection of database format

15. The prototype database is an Access database containing administrative data, morphological and molecular data, and a link to the photograph, as a (scanned) JPEG image. The database consists of four files: 1. Morphological data; 2. Administrative data; 3. Marker data; 4. Images. All files have an Excel format. Links between the Excel files are made through a "leading number" which is unique for each variety analyzed. Pictures are stored in a separate file and are included in a data interrogation. The four Excel files are linked together in the Access database. All data of one variety are shown in one screen. After setting this up the database was populated with data.

(e) Filling the database

16. In this project we were set to populate the database with at least 200 varieties under evaluation at the testing stations on behalf of the CPVO. In the final database, 400 varieties are included, of which 314 varieties were under evaluation on behalf of the CPVO. Of these

400, morphological data is available for all, with at least one photograph for 215 varieties (193 single pictures and 184 composite pictures) and molecular profiles for 364 varieties.

(f) Standardization between laboratories

17. Protocols routinely used at PRI for rose genotyping were transferred to NIAB. When data obtained at NIAB and PRI for the test set of 23 varieties were compared it became evident that there were several discrepancies between the data produced at each laboratory. The discrepancies were:

(i) Differences in signal intensity resulted in the scoring of a peak in one lab as a marker but did not result in the scoring of the same peak in the other lab. To be scored as a marker, a peak needs to have a certain minimum intensity (i.e. reach a pre-set threshold level). In principle, four different alleles can be detected in one variety. A peak was considered to be an allele when the peak area of the smallest peak was at least 15% of the area of the largest peak. Differences in amplification efficiency resulted in differences in allele calling.

(ii) Discrepancies also arose from missing values, for example some samples gave amplification product within one laboratory and not in the other. This problem is also related to the quality of the DNA obtained. Clearly, DNA extracts from garden roses appeared to be more difficult than from glasshouse roses. The DNA extracted from the garden roses most probably contained substances likely to interfere with the PCR stage of the analytical protocol. Missing data occurred at both laboratories, but the missing data points were not always the same in the two laboratories. The non-coincident missing data cause noise when comparing the two data sets.

(iii) Discrepancies originated form mis-scoring of alleles. This type of error is easily corrected if profiles are also analyzed by a second person and/or when all samples are done in duplicate.

18. From this, it was concluded that it is not currently possible to produce a unified molecular database for roses using data collected in two different laboratories. The major reasons for this are discussed above. Therefore, we chose to produce the molecular database on the data obtained at PRI only.

Discussion and Conclusions

19. An integrated pilot database, containing administrative, morphological and molecular data as well as pictures of each variety (figure 3), was constructed.

20. All items included in the database have been evaluated by the experts of BSA, NIAB and the board for plant varieties in the Netherlands.

Morphological data

Usefulness of the morphological data in the Database

21. It is essential to have some basic morphological categorization to assist in sorting and screening of candidate varieties, particularly when the database grows in size. The selected morphological characteristics are useful for this purpose; although tolerances are of course needed as there can be influences of the environment on TQ characteristics 5.3 and 5.4 and even sometimes on 5.2 and 5.5. In the normal way, experts need to bear this in mind when using the data coming from different testing stations, and check across an appropriate range of groups.

Which data to include in the database

22. It is suggested that the standard characteristics chosen should be included for all types of variety because, even though currently a particular group may seem to consist largely of one type, one cannot pre-judge the future and also the full information is needed to compare data from varieties of different types.

23. Comparison of data supplied by the applicants via the TQ, with data obtained from the testing stations, shows the importance of the examiners' expertise in assessing the varieties in a standardized way. For example, flower diameter showed differences between TQ data provided by the applicant and the observations made by the examiner in 30% of cases, but they were not large and could be expected in a quantitative characteristic where the examiner is applying a standardized system. For inclusion in the database, it seems to be preferable that for TQ 5.2, 5.3, 5.4 and 5.5 only the data from the examination offices are entered into the database, to reduce the possible deviations to a minimum.

Pictures

Usefulness of the pictures

24. The pictures taken from each variety are very important, although the composite photo was considered less informative for the greenhouse roses, because in comparison to the garden roses there is very little variation in the extra characteristics photographed. By contrast, for the garden varieties the composite photo adds very useful information.

25. A point to consider is the labor that is involved in taking the pictures: the time investment in the chain from collecting the leaves to the storage of the pictures should not be underestimated, but can hopefully be balanced against efficiency gains elsewhere.

Molecular data

26. Microsatellite markers have been used to construct a database containing the molecular profiles of approx. 380 varieties. The markers proved to be very informative about the varieties. Large numbers of allelic phenotypes (on average 32 for glasshouse varieties and 45 for garden roses per marker) have been detected. As far as we are able to tell, seedling varieties can be distinguished from each other on the basis of DNA profiles, and mutant varieties and mutant groups showed identical patterns. However, there were 2 cases in the garden roses where two varieties were suspected to be mutants of each other based on the DNA analysis, but this was not confirmed by the TQ information. In greenhouse roses,

similar situations were encountered but, in those cases, the mutant nature could be confirmed afterwards. From this it is clear that marker data can add to, and improve, the quality of the DUS work.

27. Based on the frequency with which the different allelic phenotypes occur, it is possible to calculate the chances that two varieties have identical profiles, assuming an independent breeding history. Using the 12 markers, these chances are 10^{-8} and $5x10^{-11}$ for glasshouse and garden roses respectively, which is extremely low. Effectively, this means that when two identical profiles are detected, the chance that the two samples are identical or belong to the same mutant group is almost a 100%. When we combined the data of greenhouse and garden roses it was observed that all of them (excluding the mutants within a given set of garden or greenhouse roses) showed different profiles.

28. A problem was encountered when we tried to merge data gathered in two different laboratories into one database. Several differences in allele calling were observed. These problems were most probably caused by a lower quality of the DNA samples obtained from the garden roses. Clearly, more effort is needed to harmonize the molecular marker analysis between different laboratories, including the development of good protocols for taking and handling leaf samples, extracting DNA, applying clean-up methods as required, PCR of fragments, standardizing allele calling and coordinating allele nomenclatures for databasing. In spite of these difficulties, we have demonstrated that two laboratories can produce substantially equivalent data and that the molecular data produced is useful as a tool for managing reference collections.

Use of the database

29. In the project proposal we identified 5 possible uses of the integrated database

- Characterization and cataloguing of the reference collection
- Pre-screening and selection of appropriate reference varieties
- Exchange of data on current candidate varieties between testing stations
- Strong reduction or replacement of permanent living reference collections at testing stations
- Quality assurance within examination offices (verification of identity/authenticity)

30. The database produced will be helpful for all aspects. However, not every part of the database (morphological data, molecular data and picture) is equally useful for each application.

31. For characterization and cataloguing of the reference collection, all types of data are valuable. In addition to this, one might consider storing a DNA sample as well, for future verification of replacement samples for the trials and to provide the applicant an opportunity to show a firm link with the DUS-tested material in the case of suspected infringement.

32. For pre-screening and selection of appropriate reference varieties, all data types are also useful. In particular, the morphological data and photographs are valuable for this and marker data can be used to identify or confirm mutants and thus possible varieties for comparison. Hence, the database allows the possibility to categorize the reference collection in the sense of storing information on morphology, photographs and molecular data. This information can then be used for screening and the selection of necessary similar varieties to grow in the test.

33. In addition, the database will facilitate the exchange of data between the different testing stations. However, a lot more research will be necessary to standardize sample handling, DNA extraction and scoring for the molecular data. On the morphological side, ring tests will be useful to ensure continued consistency of scoring.

Acknowledgements

Funding for this project was provided by the CPVO, Ministry of Agriculture, Nature and Food quality of the Netherlands, DEFRA (UK) and the Bundessortenamt, Germany.

Literature

Cairns, T (2000) Modern Roses XI, The World Encyclopedia of Roses. Academic Press

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

Nybom, H. , D. Esselink, G. Werlemark & B. Vosman (2004) Microsatellite DNA marker inheritance indicates preferential pairing between two highly homologous genomes in polyploid and hemisexual dogroses, Rosa L. sect. Caninae. Heredity 92: 139-150.

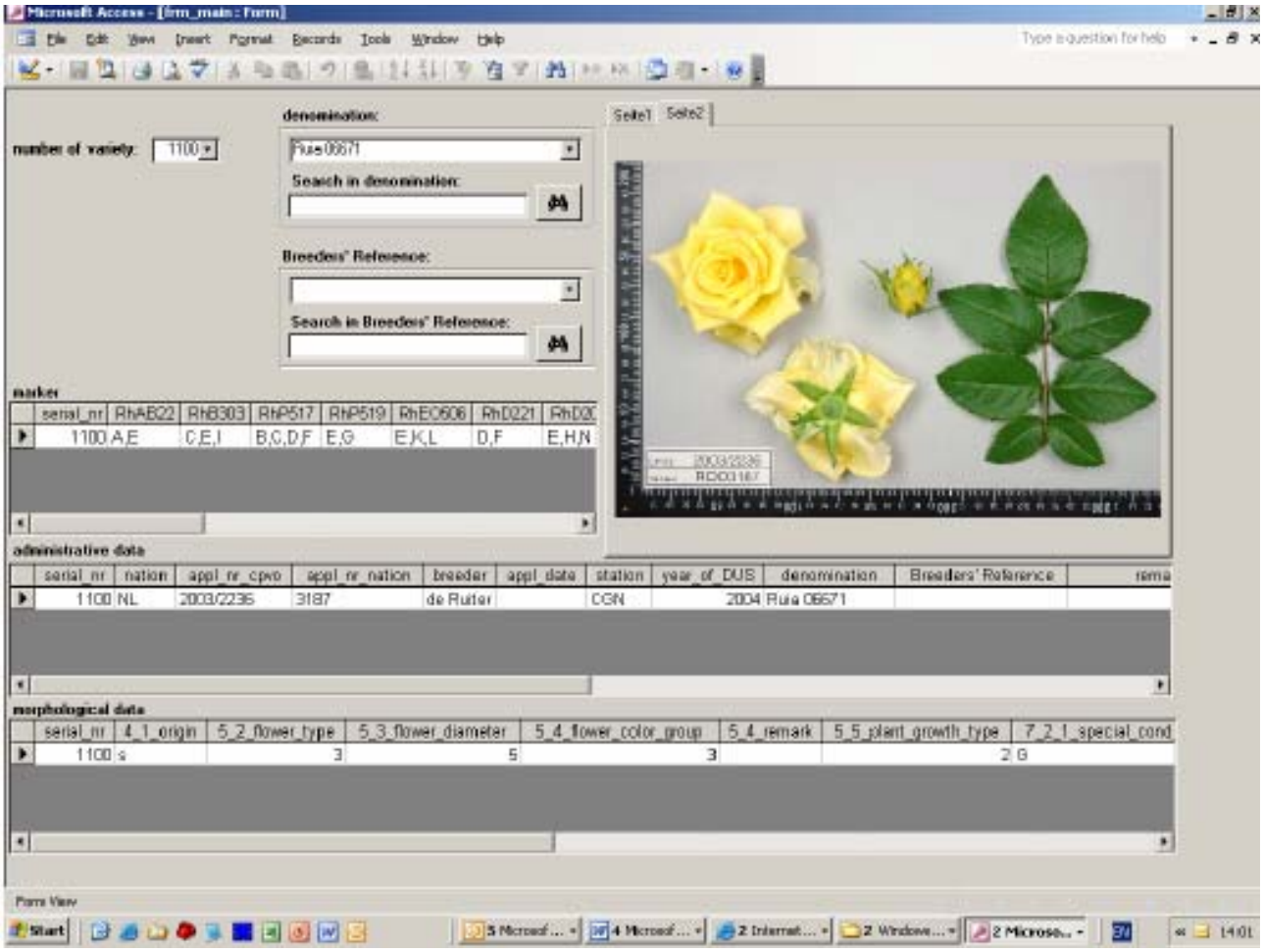
Rusanov, K., N. Kovacheva, B. Vosman, L. Zhang, S. Rajapakse, A. Atanassov, I. Atanassov (2005) Microsatellite analysis of Rosa damascena Mill. accessions reveals genetic similarity between genotypes used for rose oil production and old Damask rose varieties. Theor Appl Genet 111: 804–809

Smulders, M.J.M, D. Esselink, R. E. Voorrips & B. Vosman (2005) Analysis of a database of DNA profiles of 734 hybrid tea rose (*Rosa hybrida*) varieties. Document for UPOV Working Group on Biochemical and Molecular Techniques and DNA-profiling in particular (BMT/9/12).

Figure 1: Flower from the top

Figure 2: Composite photo

Figure 3: Screen showing all information on a variety



[End of document]