UPOV

International Union for the Protection of New Varieties of Plants

E

Working Group on Biochemical and Molecular Techniques and DNA-Profiling in Particular

BMT/17/15 Add.

Seventeenth Session
Montevideo, Uruguay, September 10 to 13, 2018

Original: English
Date: September 13, 2018

ADDENDUM TO A DNA DATABASE FOR ROSE: DEVELOPMENT AND VALIDATION OF A SNP MARKER SET

Document prepared by experts from the Netherlands, the International Rose Breeders Association (IRBA) and the International Community of Breeders of Asexually Reproduced Ornamental and Fruit-Tree Varieties (CIOPORA)

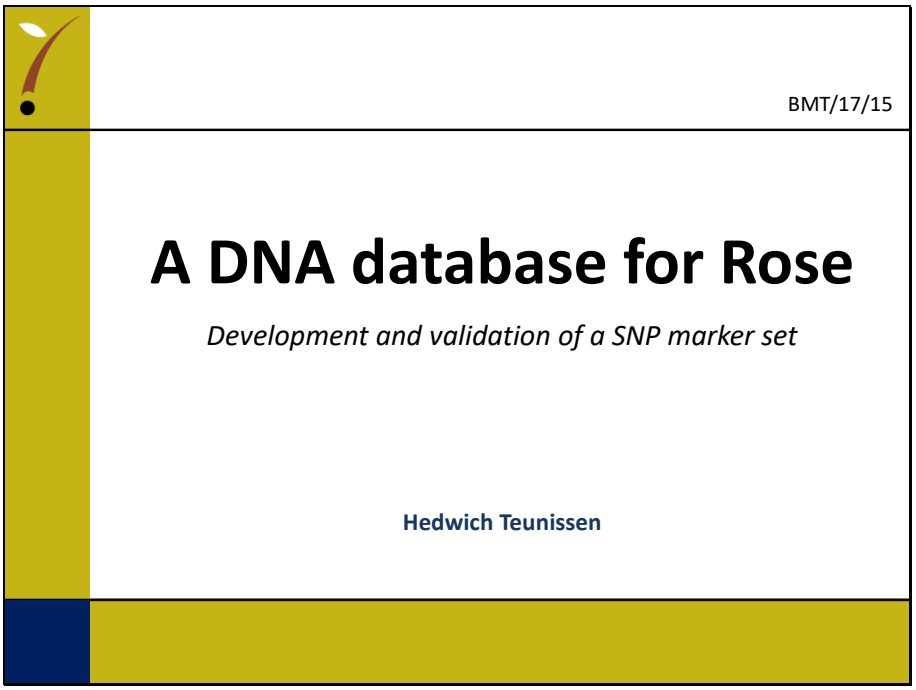
The Annex to this document contains a copy of a presentation on "A DNA database for Rose: Development and validation of a SNP marker set", prepared by experts from the Netherlands, the International Rose Breeders Association (IRBA) and the International Community of Breeders of Asexually Reproduced Ornamental and Fruit-Tree Varieties (CIOPORA), which was made at the seventeenth session of the Working Group on Biochemical and Molecular Techniques and DNA-Profiling in Particular (BMT).

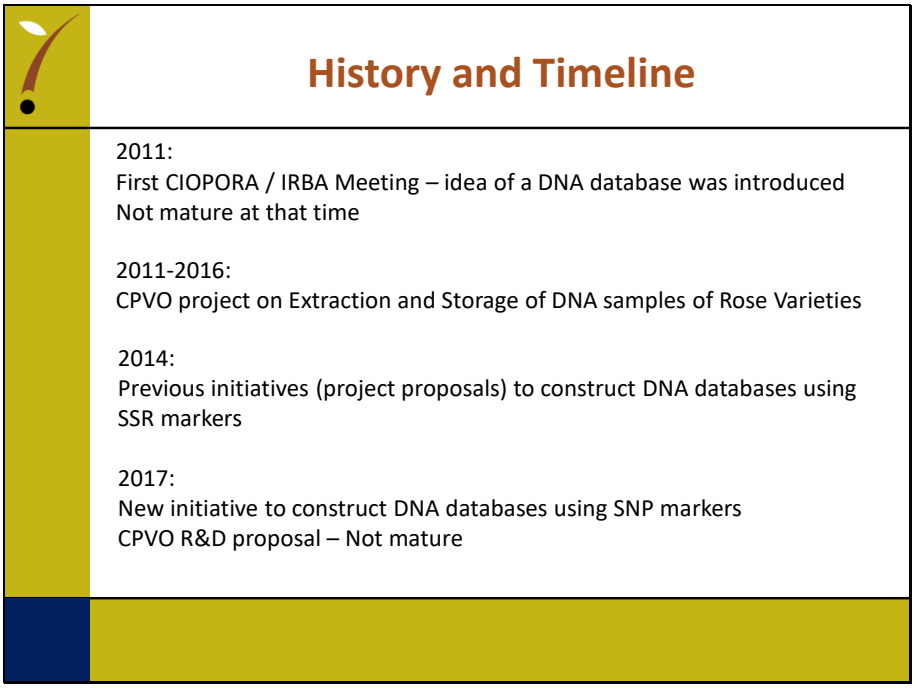
[Annex follows]

ANNEX

A DNA DATABASE FOR ROSE: DEVELOPMENT AND VALIDATION OF A SNP MARKER SET

Presentation prepared by experts from the Netherlands, the International Rose Breeders Association (IRBA) and the International Community of Breeders of Asexually Reproduced Ornamental and Fruit-Tree Varieties (CIOPORA)





Title and scope of the project

Title:

Development and validation of a SNP set for cut rose

Ultimate goal:
- construction of a database
- implementation of DNA in DUS testing of cut rose (management of reference collection)

First step: this project

Dutch Board for Plant Varieties & Naktuinbouw

Project Partners

- Naktuinbouw
- Rose Breeders (IRBA / CIOPORA)
  (10 international breeding companies)

Why Cut Rose?

- Model for ornamental crops with a complex polyploid genome.
- Intensively studied in scientific world (also previous work by Ben Vosman)
- Relevant data on molecular markers and DNA sequences in public domain
- Recently, two papers describing two reference genome sequences (Hibrand Saint-Oyant et al., 2018 and Raymond et al., 2018)

Why Cut Rose?

- World wide breeding activities and trade. No restriction to regions or countries
- Number of existing varieties (common knowledge) is high
- Very important ornamental crop with many new applications each year
- There is a need to manage the variety collection to
  - Safely exclude comparing varieties from the trial
  - Reduce workload and save costs
  - Avoid an extra year of testing

Challenges and Risks

- No full overview and availability of common knowledge:
  - ✓ Rapid development of new varieties
  - ✓ Global character of breeding
- Phytosanitary restriction of importing living reference material:
  - ✓ Risk for plant health
  - ✓ Very costly
- Stability of the variety description
  - ✓ Phenotype is influenced by environment
  - ✓ Risk on wrong comparing varieties based on photos and TQ
  - ✓ DUS trials will take more than 1 year

Benefits and Applications

DNA profiles in databases overcome all risks

<u>Future perspective</u>

More reliable decision on Distinctness: objective comparison of genetic distance for each application with all varieties in the database

Fits with approved UPOV model 2 ‘combining phenotypic and molecular distances in the management of variety collections’

Project Objectives and Tasks

- **Develop Cut Rose-specific SNP panel**
  - Identification of cut rose-specific SNP markers
  - Selection of SNP markers based on quality criteria and performance
  - Test and Optimization of the genotyping method, discard badly performing SNPs
  - Genotyping method validation
  - Fit for Purpose Validation of SNP marker set
  - Publication of the results

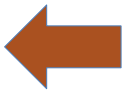
Identification of Cut Rose specific SNPs

Strategies for the Identification of SNP
(genetic diversity in training set)

€ Select SNPs from publically available genome sequences that represent genetic diversity

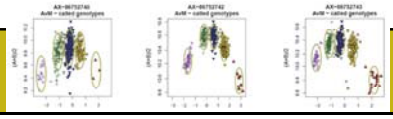
68,983 SNPs

€€ Screen commercial available SNP arrays to obtain SNPs from our selection of varieties



€€€ Re-sequence selection of varieties to obtain SNPs

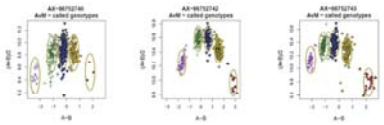
From tetraploid cut rose and garden rose varieties

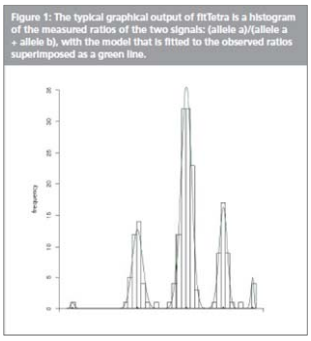


Identification of Cut Rose specific SNPs

Quality criteria for SNP selection

- ✓ Allele frequencies / dosage
- ✓ PIC value (highly discriminative)
- ✓ Single locus
- ✓ Equal distribution
- ✓ Whole genome covered
- ✓ Easy to score





Identification of Cut Rose specific SNPs

- Define **training set** of varieties that is representative for broad genetic diversity. Focus on cut rose (70%), other types (30%). Species used as rootstock are excluded (not relevant)
- This set is used for the identification of SNPs

- Define **test set** of varieties that represent difficult samples in respect to DUS and controls (e.g. same variety different origins; mutants and/or derived varieties; seedlings and parents).
- This set is used for the validation

Results of Identification

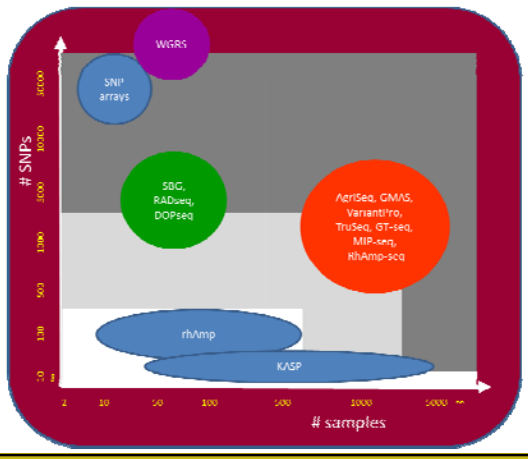
First selection of SNPs (~384)

- On performance based on Quality Criteria
- On performance on Distinctness of varieties in training set

Choice on SNP detection technology

- Inventory and benchmark of genotyping technologies in 2017
- Promising Genotyping technologies:
  - ✓SNP Genotyping by Multiplexed Targeted Amplicon Sequencing Approach (e.g. GT-Seq, MIP-seq)
  - ✓Single SNP assays - End-Point measurements (e.g. RhAmp, KASP)

Inventory BMT for genotyping



SNP genotyping by multiplexed targeted amplicon sequencing

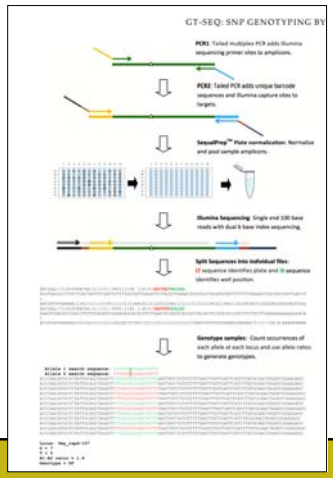
- **GT seq**

Genotyping-in-thousands by sequencing

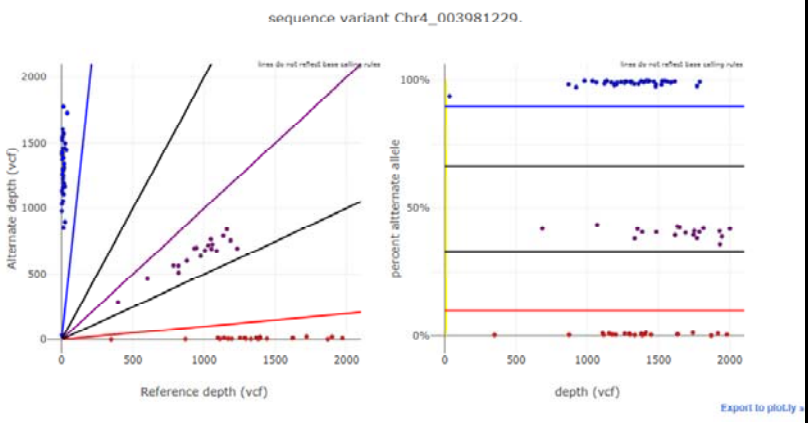
Cost effective SNP genotyping method based on multiplexed targeted amplicon sequencing

Combining multiple loci and multiple samples in 1 sequencing run (even samples of different crops and experiments)

Open source method



First results are very promissing



Data for Rubus

Optimization of the method

- PCR components and reaction conditions
- Primer-dimers (inefficient amplification)
- PCR duplicates (inefficient amplification and hamper allele dosage determination)
- Normalisation
- Multiplex factor (# samples x # SNPs)
- Sequence coverage needed for reliable dosage calling?
- Bioinformatic pipelines

Method validation

- After optimization the protocol is fixed
- Test how well the method performs:
  - Accuracy
  - Reproducibility
  - Repeatability
  - Robustness

Fit for Purpose Validation

- Validation of selected SNPs using defind varieties from the **test set** (samples that should be distinct and samples that should not be distinct)
  - ✓ Genetically very similar varieties or lines
  - ✓ Parents and off-spring
  - ✓ Genetically close but morphologically distinct varieties
  - ✓ Series of mutants e.g.
- Blind tests with new sets of varieties, replica's and controls
- Select final SNPs set

Publically available SNP set

- Publish the results in peer reviewed scientific paper. SNP set is publically available. Also for other DUS testing authorities to use.
- The choice on SNP detection method is open for every lab
- Validation between labs is required

Memorandom of Understanding

- Breeders provide well defined plant material for both training set and test set
- Naktuinbouw performs the laboratory work
- Screening of the Axiom Array is outsourced to service provider in NL
- All partners have signed the MoU: legal and practical guidance during the project

State of Affairs

- Training set of 164 varieties and 28 control samples obtained from 9 breeding companies
- DNA was extracted (is challenging when old leaf material was provided)
- DNA was send to service provider. Currently screening of Axiom Array takes place

Quality in Horticulture

[End of Annex and of document]