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/17/8 Add.

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

Original: English
Date: October 18, 2018

ADDENDUM TO TEST OF THE POTENTIAL USE OF SNPS MARKERS ON OILSEED RAPE VARIETIES

Document prepared by an expert from France

Disclaimer: this document does not represent UPOV policies or guidance

The Annex to this document contains a copy of a presentation on “Test of the potential use of SNPs markers on oilseed rape varieties”, prepared by an expert from France, 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

TEST OF THE POTENTIAL USE OF SNPS MARKERS ON OILSEED RAPE VARIETIES

Presentation prepared by an expert from France

TEST OF THE POTENTIAL USE OF SNPS MARKERS ON OILSEED RAPE VARIETIES

→ BMT September 10th, 2018
Montevideo, Uruguay

Document prepared by French experts (GEVES)

Pre-Project aims

- **Objective:**

TEST OF THE POTENTIAL USE OF SNPS MARKERS ON OILSEED RAPE VARIETIES

① Selection and validation of a suitable set of SNP markers

② Determination of the analysis matrix
i.e. Individual seed, seed bulk, individual seedling…

GEVES Expertise & Performance
Coordinator

NIAB

Animal & Plant Health Agency

Co-funded by

CPVO Community Plant Variety Office

Molecular & data analysis conducted by NIAB and GEVES

Project structure

- Phase 1: Marker selection/ varieties choice / DNA preparation
- Phase 2: Amplification test/ marker screening:
  - test of 500 selected SNP on 23 individual plants: 275 SNP per lab (50 SNP in common)
- Phase 3: Working in bulk and size of the bulk
  - Selection of 90 SNP => 50 SNP per lab (10 in common)
  - Work on 4 varieties => individual genotyping of 32 plant per varieties + DNA bulk
  - According to previous results => genotyping of leaves and/or seed bulk? Size?
- Phase 4: Variety screening
  - Method chosen in phase 3 applied to the 19 remaining varieties
- Phase 5: Data analysis and review

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

3

Project structure

- **Phase 1: Marker selection/ varieties choice / DNA preparation** ✓
- Phase 2: Amplification test/ marker screening:
  - test of 500 selected SNP on 23 individual plants: 275 SNP per lab (50 SNP in common)
- Phase 3: Working in bulk and size of the bulk
  - Selection of 90 SNP => 50 SNP per lab (10 in common)
  - Work on 4 varieties => individual genotyping of 32 plant per varieties + DNA bulk
  - According to previous results => genotyping of leaves and/or seed bulk? Size?
- Phase 4: Variety screening
  - Method chosen in phase 3 applied to the 19 remaining varieties
- Phase 5: Data analysis and review

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

4

Phase 1 - Marker selection, primer design/ varieties selection

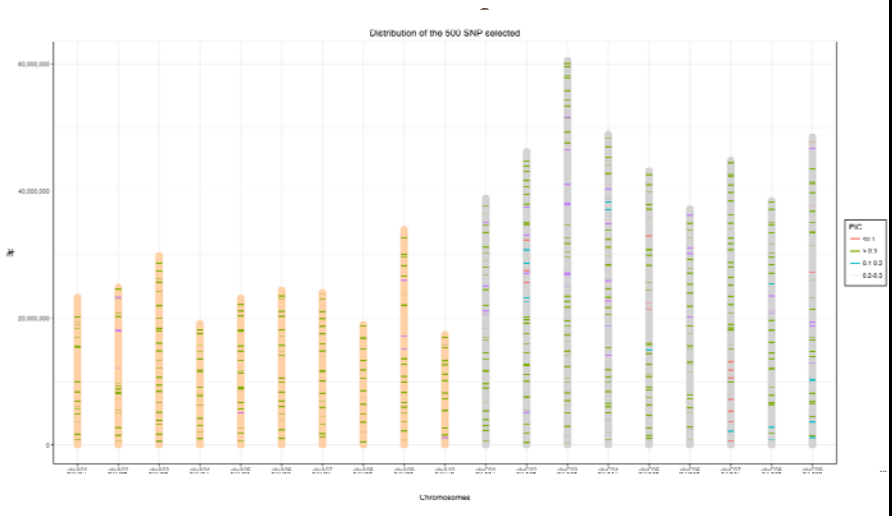
- **Milestone: selection of a set of SNP markers, synthesis of molecular primers and distribution of the primers to the two laboratories**
  - Based on publication of Clarke *et al.* 2016[1] and genotypes data (source : INRA)
  - 500 SNP set choice based on:
    - Homogenous genome coverage
    - Non multilocus (genome map)
    - High PIC (Polymorphism Information Content)
    - Suitable flanking sequence for KASPar design

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

[1] Clarke et al., 2016. A high-density SNP genotyping array for *Brassica napus* and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. **Theor. Appl. Genet**. 129, 1887–1899.

5

Phase 1 - Selection of a set of SNP markers



6

Phase 1 - Primer design

- **Milestone: selection of a set of SNP markers, synthesis of molecular primers and distribution of the primers to the two laboratories**
  - Primer sequences & design

  - Synthesis by IDT-DNA (private provider)
  - Divided SNP between FR and UK

[1] Clarke et al., 2016. A high-density SNP genotyping array for *Brassica napus* and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. **Theor. Appl. Genet**. 129, 1887–1899.

Phase 1 - Varieties selection

- **Milestone: selection of varieties to be tested and material requirements**
  - Selection of 23 varieties:
    - inbred lines
    - hybrids and their parental components
    - homogeneous or heterogeneous based on field and molecular biology observations

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

Phase 1 - Milestone: Growing of the materiel

- To work on the same material: one growing site



One tray per variety


Harvesting and sampling after 3 weeks of germination

9

Phase 1 - Milestone: Sampling and bulking

- Preparation of the DNA for phase 2, 3 and 4
  - On individual leaves, bulks of DNA, bulks of leaves and bulks of seeds

→ Only one DNA extraction

10

Phase 1 - Milestone: DNA extraction

- 1st DNA extraction carried out in one lab :
  - Phase 2: individual leaf (one leaf per variety) on the 23 varieties
  - Phase 3: 32 individual plant for each of the 4 varieties selected + leaf bulk + seed bulk

→ One DNA extraction => to be divided between the 2 labs

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

11

Project structure

- Phase 1: Marker selection/ varieties choice / DNA preparation √
- **Phase 2: Amplification test/ marker screening:** √
  - test of 500 selected SNP on 23 individual plants: 275 SNP per lab (50 SNP in common)
- Phase 3: Working in bulk and size of the bulk
  - Selection of 90 SNP => 50 SNP per lab (10 in common)
  - Work on 4 varieties => individual genotyping of 32 plant per varieties + DNA bulk
  - According to previous results => genotyping of leaves and/or seed bulk? Size?
- Phase 4: Variety screening
  - Method chosen in phase 3 applied to the 19 remaining varieties
- Phase 5: Data analysis and review

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

12

Phase 2 - Amplification test/markers screening

- **Milestone: amplification test to create a valid set of SNPs to be deployed by the partners**
  - Amplification test on the 500 SNP (selected *in silico*).
  - One leaf sample of each of the 23 varieties to be genotyped.
  - Harmonisation and standardisation of the PCR & reading conditions.

- Out of the 500 SNP
  - 466 amplified & 454 were polymorphic
  - 34 no amplification & 12 monomorphic → ~ 10%

→ Good conversion rate between *Illumina* and *KASP* technology

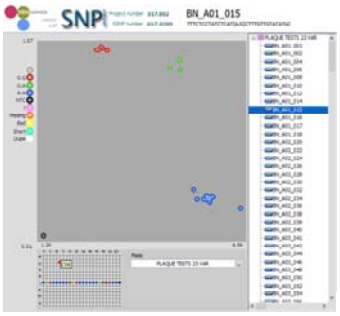
GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

13

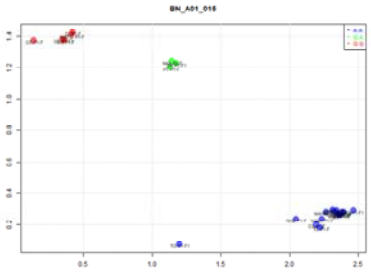
Phase 2 - Amplification test/markers screening

Comparison of the results for 50 SNP in common:

- Identical calling and clustering



FR lab



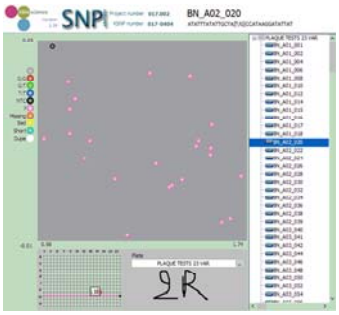
UK lab

14

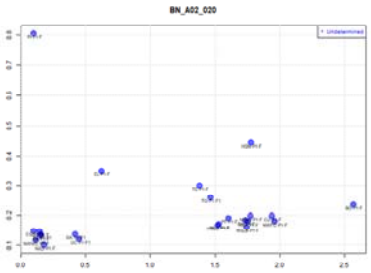
Phase 2 - Amplification test/markers screening

Comparison of the results for 50 SNP in common:

- Amplification problem /not working



FR lab



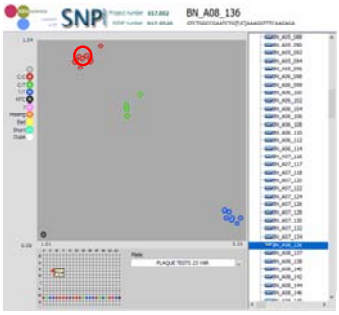
UK lab

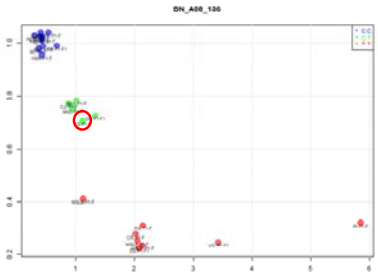
Phase 2 - Amplification test/markers screening

Comparison of the results for 50 SNP in common:

- From 2300 data points generated between the 2 labs, not including missing data, results differed in approximately 6.5%

- Genotyping error

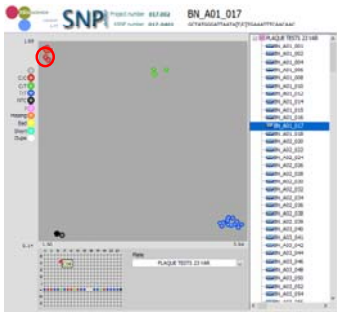


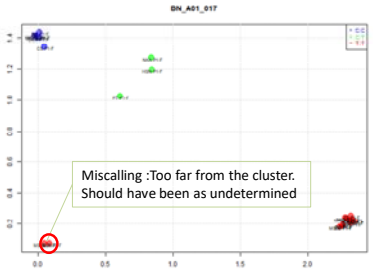


Phase 2 - Amplification test/markers screening

Comparison of the results for 50 SNP in common:

- But mostly miscalling





FR lab

UK lab

→ After reviewing the 50 SNP in common we had less than 2% of differences

Phase 2 - Amplification test/markers screening

- Two labs working together in different locations, using different resources and software

⇒ **Results similar after harmonization but still 2%**

- To keep in mind=> step of cross-checking and harmonization is mandatory

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

Project structure

- Phase 1: Marker selection/ varieties choice / DNA preparation √
- Phase 2: Amplification test/ marker screening: √
  - test of 500 selected SNP on 23 individual plants: 275 SNP per lab (50 SNP in common)
- **Phase 3: Working in bulk and size of the bulk** √
  - Selection of 90 SNP => 50 SNP per lab (10 in common)
  - Work on 4 varieties => individual genotyping of 32 plant per varieties + DNA bulk
  - According to previous results => genotyping of leaves and/or seed bulk? Size?
- Phase 4: Variety screening
  - Method chosen in phase 3 applied to the 19 remaining varieties
- Phase 5: Data analysis and review

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

19

Phase 3 - Working in Bulk and size of the Bulks

- Test on 4 varieties

Type	Comment
LINE	Maintainer line
HYB	Cytoplasmic male sterility hybrid
LINE	Variety heterogenous in the field (on leaves characteristics and on fertility)
LINE	Variety which is heterogenous on electrophoresis but uniform in the field

20

Project structure

- Phase 1: Marker selection/ varieties choice / DNA preparation √
- Phase 2: Amplification test/ marker screening: √
  - test of 500 selected SNP on 23 individual plants: 275 SNP per lab (50 SNP in common)
- **Phase 3: Working in bulk and size of the bulk** √
  - **Selection of 90 SNP => 45 SNP per lab (10 in common)**
  - Work on 4 varieties => individual genotyping of 32 plant per varieties + DNA bulk
  - According to previous results => genotyping of leaves and/or seed bulk? Size?
- Phase 4: Variety screening
  - Method chosen in phase 3 applied to the 19 remaining varieties
- Phase 5: Data analysis and review

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

21

Project structure

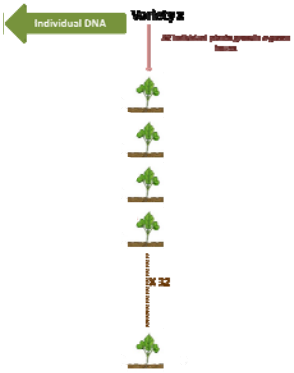
- Phase 1: Marker selection/ varieties choice / DNA preparation √
- Phase 2: Amplification test/ marker screening: √
  - test of 500 selected SNP on 23 individual plants: 275 SNP per lab (50 SNP in common)
- **Phase 3: Working in bulk and size of the bulk**
  - Selection of 90 SNP => 50 SNP per lab (10 in common) √
  - **Work on 4 varieties => individual genotyping of 32 plant per varieties + DNA bulk**
  - According to previous results => genotyping of leaves and/or seed bulk? Size?
- Phase 4: Variety screening
  - Method chosen in phase 3 applied to the 19 remaining varieties
- Phase 5: Data analysis and review

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

22

Phase 3 - Working in Bulk and size of the Bulks

- **Milestone: scoring DNA bulks**
- Genotyping of individual seedlings & bulks of DNA: comparison of allele frequency
  - 32 individual seedlings
  - Bulks of DNA mixing 8, 16, 24 & 32
  - 3 independent random samplings for each bulk



23

Phase 3 - Working in Bulk and size of the Bulks

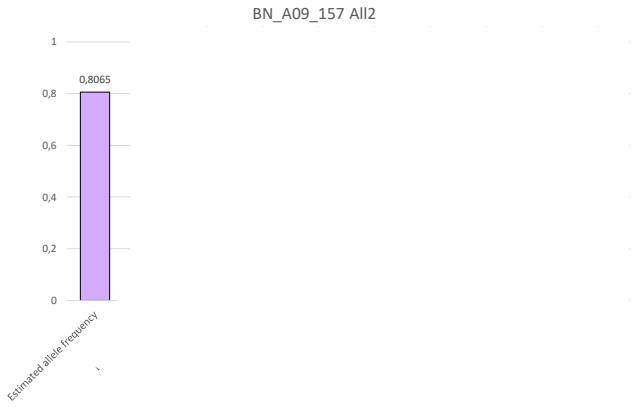
- **Milestone: scoring DNA bulks**

- Data analysis based on **allele frequency**
  - **For individual plants**: profiles of individual seedlings were used to compute observed allele frequencies for each variety
  - **For DNA bulk** :
    - For homozygous individuals A/A => freq Allele1 =1 and freq Allele2= 0
    - For heterozygous individuals A/B => freq Allele1 =0.5 and freq Allele2= 0.5
  - **Comparisons per variety of the allele frequency and bulk calling**

24

Phase 3 - Working in Bulk and size of the Bulks

- **Milestone: scoring DNA bulks**



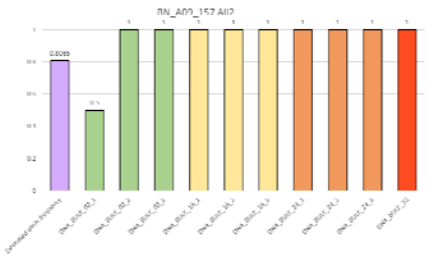
25

Phase 3 - Working in Bulk and size of the Bulks

- **Milestone: scoring DNA bulks**
- Very little differences between repetitions of bulk sizes except for bulk of 8 DNA
- No differences between bulk sizes (except for bulk of 8)

  => ↘ size of bulks ↗ sampling variance
- Minor differences in term of allele frequency between individual genotyping and bulk
- Undetected allele in bulk due to small frequency

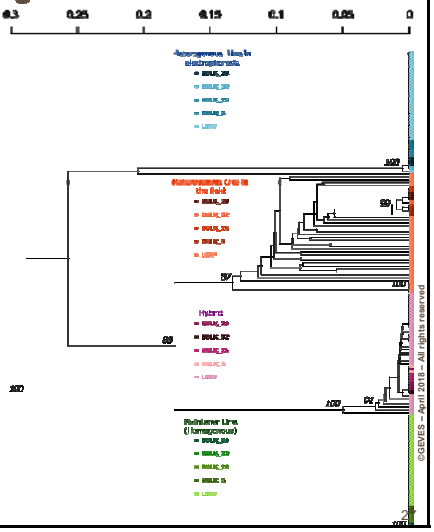
→ Keep bulk of 16, 24 & 32 for next step of phase 3



26

Phase 3 - Working in Bulk and size of the Bulks

- **Milestone: Individual genotyping and intra varietal variability**
- The dendrogram based on genetic distances among the four varieties



Project structure

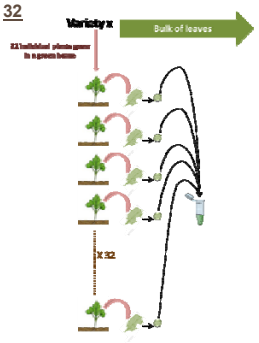
- Phase 1: Marker selection/ varieties choice / DNA preparation √
- Phase 2: Amplification test/ marker screening: √
  - test of 500 selected SNP on 23 individual plants: 275 SNP per lab (50 SNP in common)
- **Phase 3: Working in bulk and size of the bulk** √
  - Selection of 90 SNP => 50 SNP per lab (10 in common)
  - Work on 4 varieties => individual genotyping of 32 plant per varieties + DNA bulk
  - **According to previous results => genotyping of leaves and/or seed bulk? Size?**
- Phase 4: Variety screening
  - Method chosen in phase 3 applied to the 19 remaining varieties
- Phase 5: Data analysis and review

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

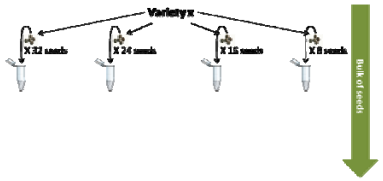
Phase 3 - Working in Bulk and size of the Bulks

- **Milestone: Optimum size and type of bulk**
- Genotyping of leaves and seeds bulks

- Bulks of leaves 16, 24 & 32



- Bulks of seeds 16, 24 & 32

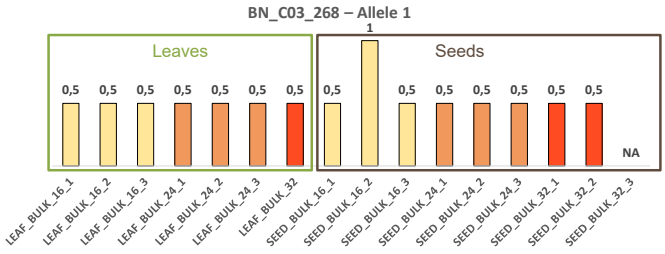


– 3 independent random samplings for each bulk (except the bulk of 32 leaves)

29

Phase 3 - Working in Bulk and size of the Bulks

- **Milestone: Optimum size and type of bulk**
- Size of bulk: no differences between size (16, 24, 32) except for bulk of 16



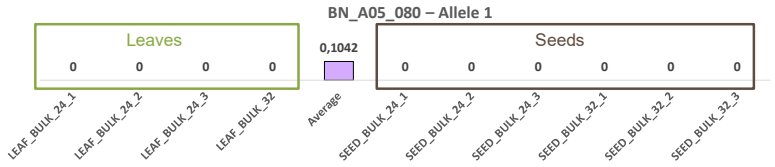
→ Keep bulk of 24 or 32 leaves or seeds

30

Phase 3 - Working in Bulk and size of the Bulks

- **Milestone: Allele dilution/threshold of detection**
- Only alleles above a given threshold will be expected to be scored. The phenomena called allele dilution will be more important when increasing size of bulks

⇒ bulk DNA may mask partially a part of genetic diversity within oilseed rape varieties



- Increasing the size of the bulk did not affect the threshold of detection since different type of bulk gave similar results. 1 allele dose will be detected or not equally in bulk of size 24 or 32

31

Phase 3 - Working in Bulk and size of the Bulks

- **Milestone: Allele dilution/threshold of detection**
- An allele with a frequency <0.05 will never be detected in bulk
- However threshold of detection for allele frequency between 0.05 and 0.3 will be different depending of the SNP.

32

Phase 3 - Working in Bulk and size of the Bulks

- According to our results → bulks of at least 24 seeds
- According to breeders advice → bulks of at least 30 seeds

→ Bulks of 32 seeds

- To confirm and secure results 2 independent samplings of 32 seeds for each of the 23 varieties for phase 4

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

33

Project structure

- Phase 1: Marker selection/ varieties choice / DNA preparation √
- Phase 2: Amplification test/ marker screening: √
  - test of 500 selected SNP on 23 individual plants: 275 SNP per lab (50 SNP in common)
- Phase 3: Working in bulk and size of the bulk √
  - Selection of 90 SNP => 50 SNP per lab (10 in common)
  - Work on 4 varieties => individual genotyping of 32 plant per varieties + DNA bulk
  - According to previous results => genotyping of leaves and/or seed bulk? Size?
- **Phase 4: Variety screening** √
  - Method chosen in phase 3 applied to the 19 remaining varieties
- Phase 5: Data analysis and review

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

34

Phase 4 - Variety screening

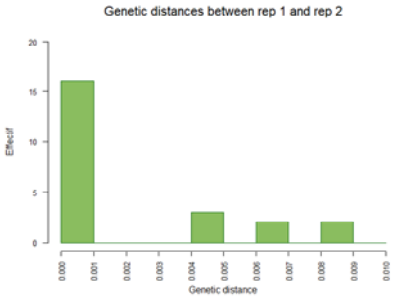
- **Milestone: Allele dilution/threshold of detection**
- 2nd DNA extraction :
  - Extraction on 2 bulks of 32 seeds for each variety
  - Invoice => End of September 2017
- Number of SNP per lab: 236 for FR and 256 for UK (38 in common)

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

35

Phase 4 - Variety screening

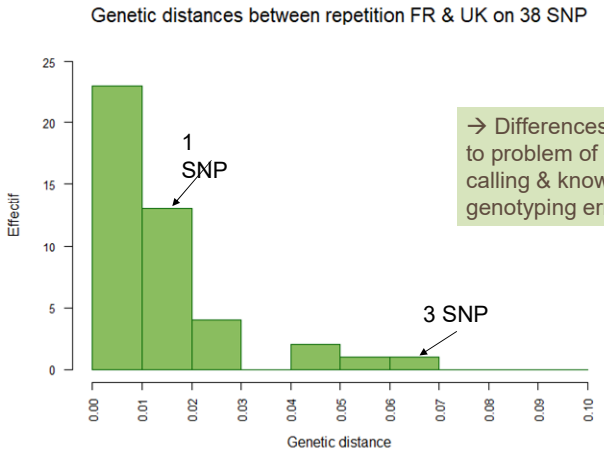
- Results for France
- Genetic distance between 2 repetitions on ~235 SNP always < 0.01=> good reproducibility
- The differences between repetitions rely on a maximum of 3 SNP



→ Bulking strategy give a good representation of the variety profile and is quite reproducible

36

Phase 4 - Variety screening
Comparison of the results for 38 SNP in common



→ Differences due to problem of calling & known genotyping error

37

Project structure

- Phase 1: Marker selection/ varieties choice / DNA preparation √
- Phase 2: Amplification test/ marker screening: √
  - test of 500 selected SNP on 23 individual plants: 275 SNP per lab (50 SNP in common)
- Phase 3: Working in bulk and size of the bulk √
  - Selection of 90 SNP => 50 SNP per lab (10 in common)
  - Work on 4 varieties => individual genotyping of 32 plant per varieties + DNA bulk
  - According to previous results => genotyping of leaves and/or seed bulk? Size?
- Phase 4: Variety screening √
  - Method chosen in phase 3 applied to the 19 remaining varieties
- **Phase 5: Data analysis and review** √

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

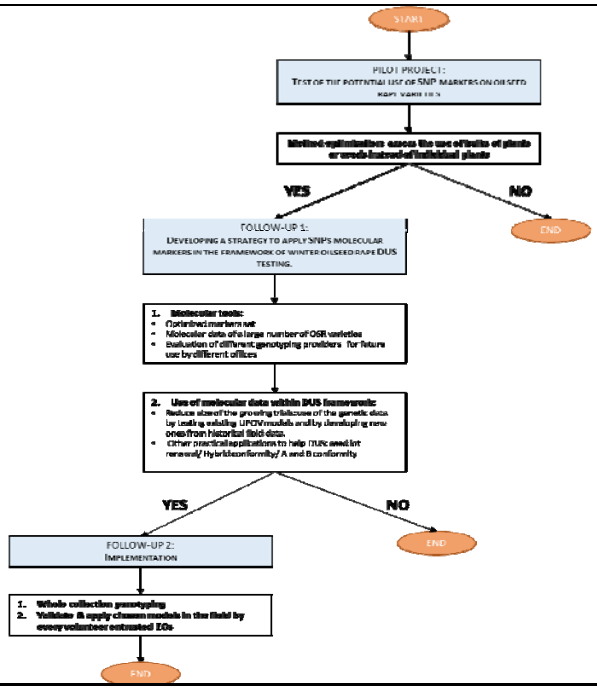
38

Conclusions & deliverables

- Good conversion rate between array and KASP technology
- **Working on bulk of at least 24 seeds but recommended over 30 seeds is a good strategy to describe OSR varieties**
- Marker set of 454 SNP available but can still be improved (PIC on the whole collection/ linkage disequilibrium/hybrids conformity…)
- Good transferability between lab BUT need for harmonisation (analysis and calling => ring test)

GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

39



Thanks for your attention

[End of Annex and of document]