BMT/14/10 Add.
ORIGINAL: English
DATE: November 27, 2014

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

Geneva

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

Fourteenth Session
Seoul, Republic of Korea, November 10 to 13, 2014

ADDENDUM TO DOCUMENT BMT/14/10

THE USE OF MOLECULAR MARKERS (SNP) FOR MAIZE DUS TESTING

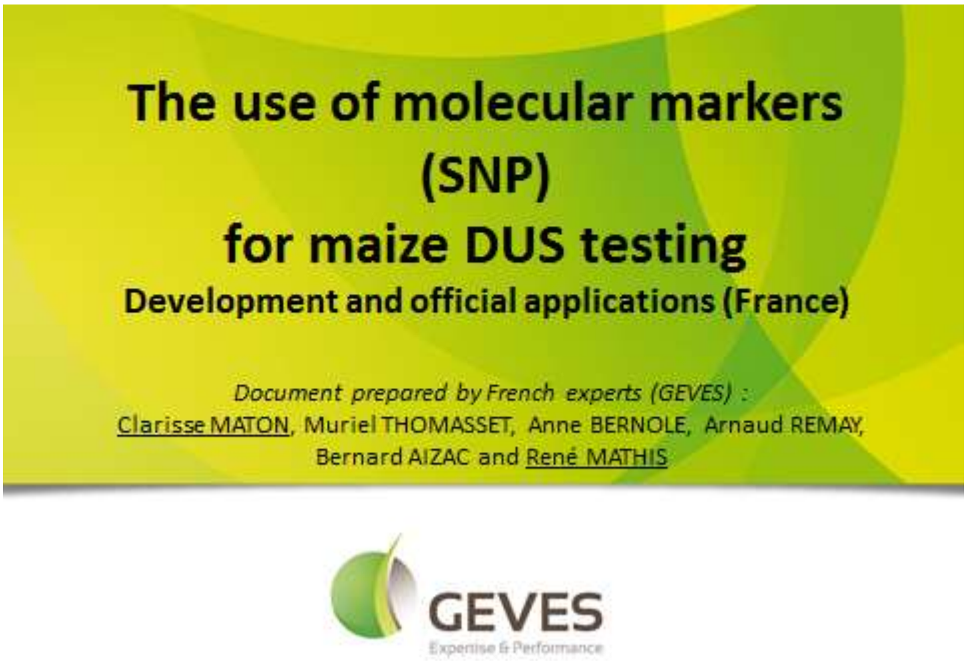
Document prepared by experts from France

The Annex to this document contains a copy of a presentation "The use of molecular markers (SNP) for maize DUS testing" made at the fourteenth session of the Working Group on Biochemical and Molecular Techniques and DNA-Profiling in particular (BMT).

Clarisse MATON, Muriel THOMASSET, Anne BERNOLE, Arnaud REMAY, Bernard AIZAC and René MATHIS
Groupe d'Etude et de contrôle des Variétés Et des Semences (GEVES), France

[Annex follows]

ANNEX



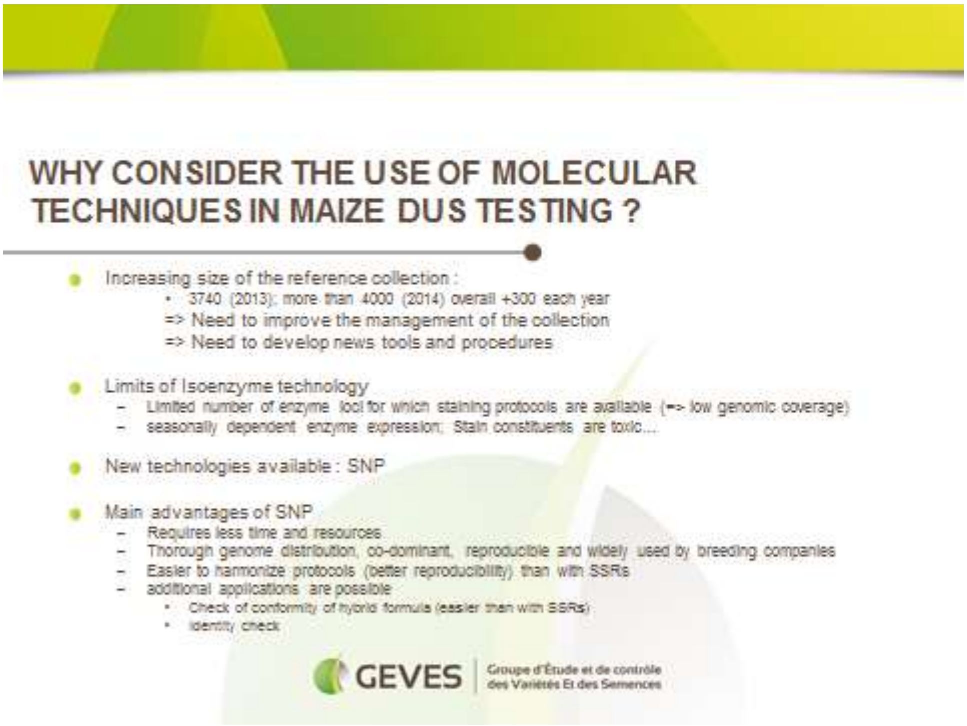
The use of molecular markers (SNP) for maize DUS testing

Development and official applications (France)

Document prepared by French experts (GEVES) :

Clarisse MATON, Muriel THOMASSET, Anne BERNOLE, Arnaud REMAY,
Bernard AIZAC and René MATHIS

GEVES
Expertise & Performance



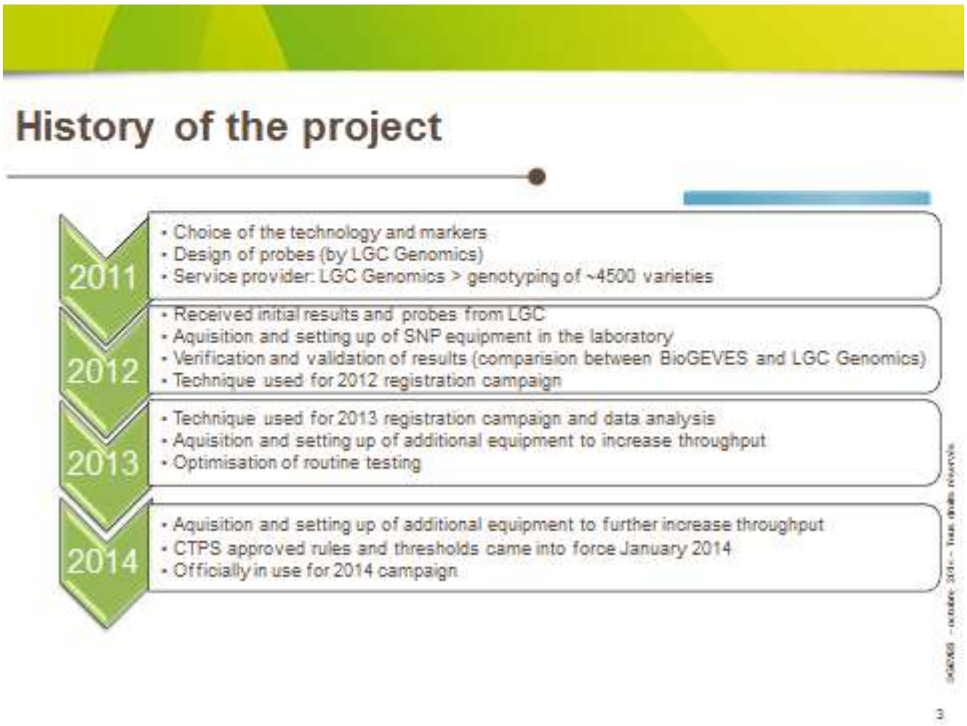
WHY CONSIDER THE USE OF MOLECULAR TECHNIQUES IN MAIZE DUS TESTING ?

- Increasing size of the reference collection :
    - 3740 (2013); more than 4000 (2014) overall +300 each year

  => Need to improve the management of the collection

  => Need to develop news tools and procedures
- Limits of Isoenzyme technology
    - Limited number of enzyme loci for which staining protocols are available (=> low genomic coverage)
    - seasonally dependent enzyme expression; Stain constituents are toxic…
- New technologies available : SNP
- Main advantages of SNP
    - Requires less time and resources
    - Thorough genome distribution, co-dominant, reproducible and widely used by breeding companies
    - Easier to harmonize protocols (better reproducibility) than with SSRs
    - additional applications are possible
        - Check of conformity of hybrid formula (easier than with SSRs)
        - identity check

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



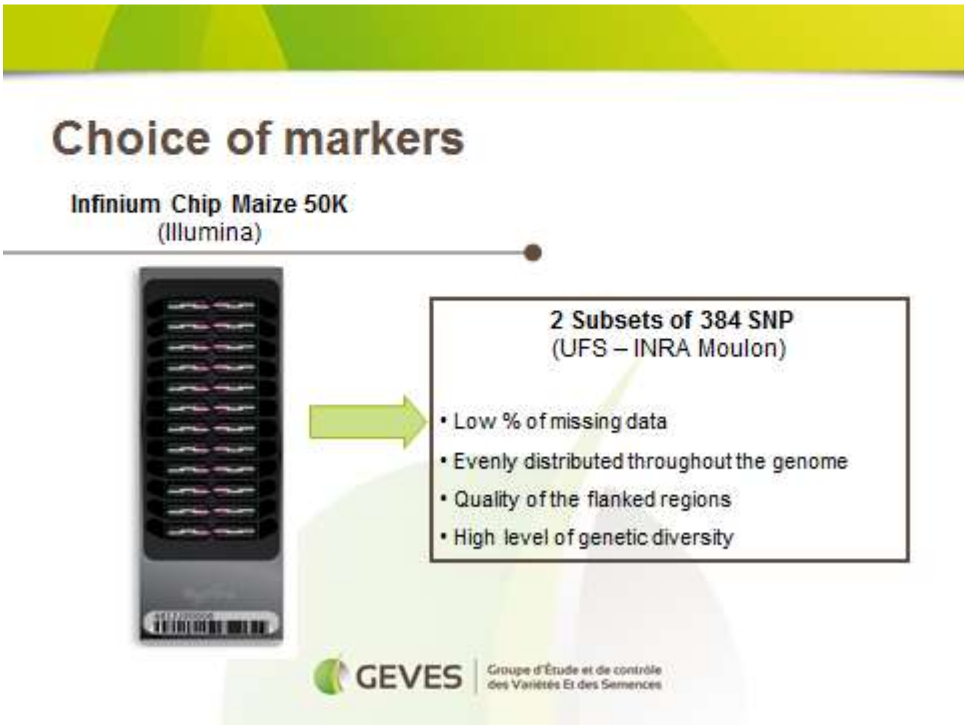
History of the project

2011
- Choice of the technology and markers
- Design of probes (by LGC Genomics)
- Service provider: LGC Genomics > genotyping of ~4500 varieties

2012
- Received initial results and probes from LGC
- Aquisition and setting up of SNP equipment in the laboratory
- Verification and validation of results (comparision between BioGEVES and LGC Genomics)
- Technique used for 2012 registration campaign

2013
- Technique used for 2013 registration campaign and data analysis
- Aquisition and setting up of additional equipment to increase throughput
- Optimisation of routine testing

2014
- Aquisition and setting up of additional equipment to further increase throughput
- CTPS approved rules and thresholds came into force January 2014
- Officially in use for 2014 campaign



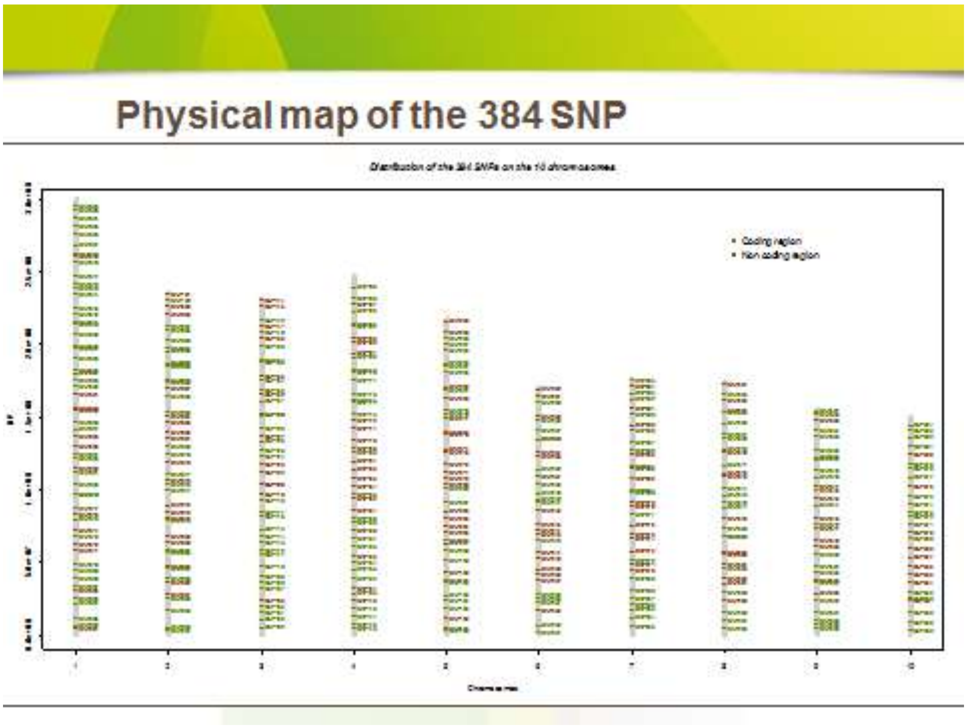
Choice of markers

Infinium Chip Maize 50K
(Illumina)

2 Subsets of 384 SNP
(UFS – INRA Moulon)

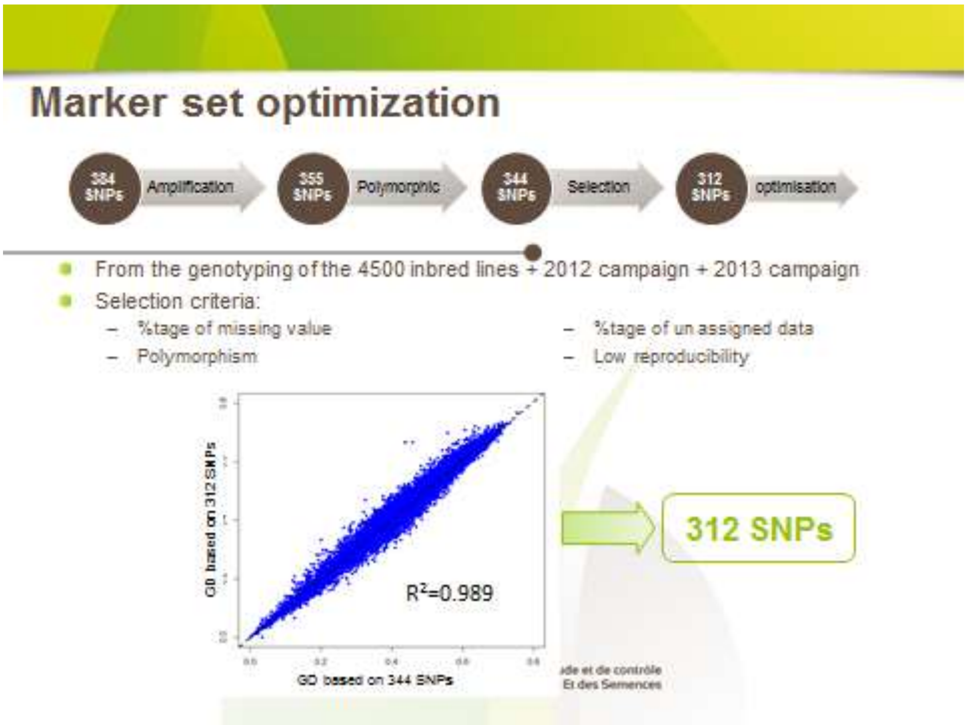
- Low % of missing data
- Evenly distributed throughout the genome
- Quality of the flanked regions
- High level of genetic diversity

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



Physical map of the 384 SNP

Distribution of the 384 SNPs on the 10 chromosomes



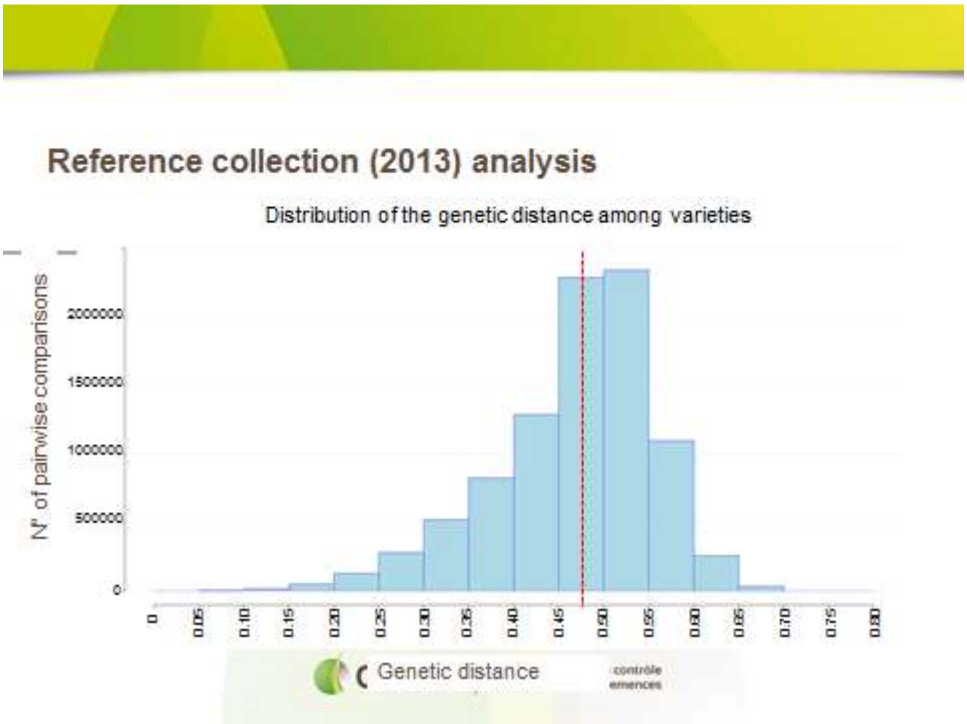
Marker set optimization

384 SNPs → Amplification → 355 SNPs → Polymorphic → 344 SNPs → Selection → 312 SNPs → optimisation

- From the genotyping of the 4500 inbred lines + 2012 campaign + 2013 campaign
- Selection criteria:
  - %tage of missing value
  - Polymorphism
  - %tage of un assigned data
  - Low reproducibility

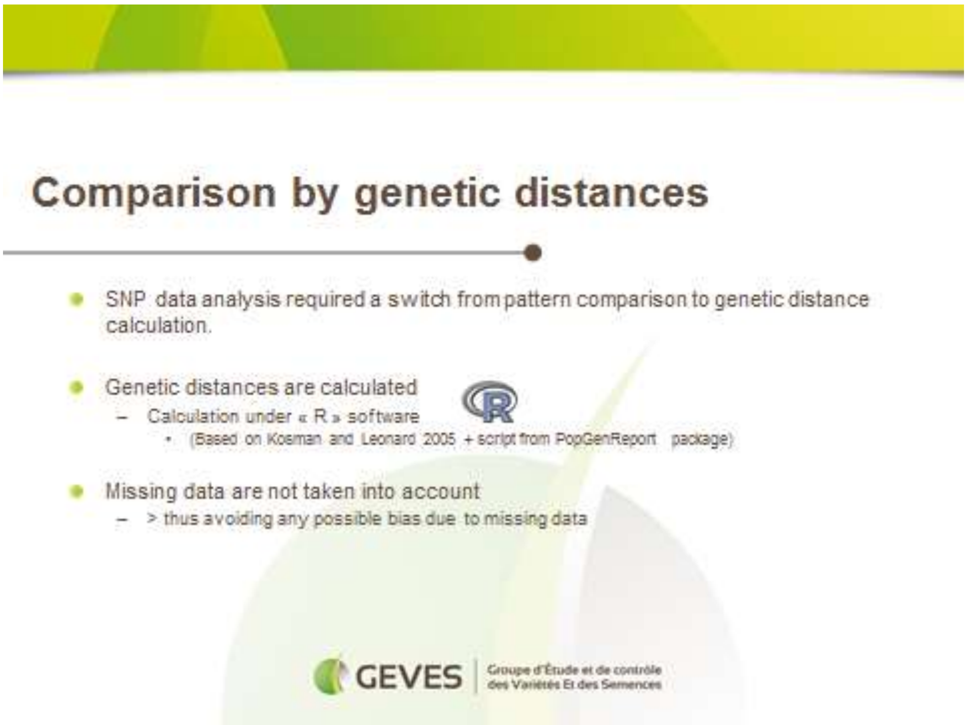
R^2=0.989

312 SNPs



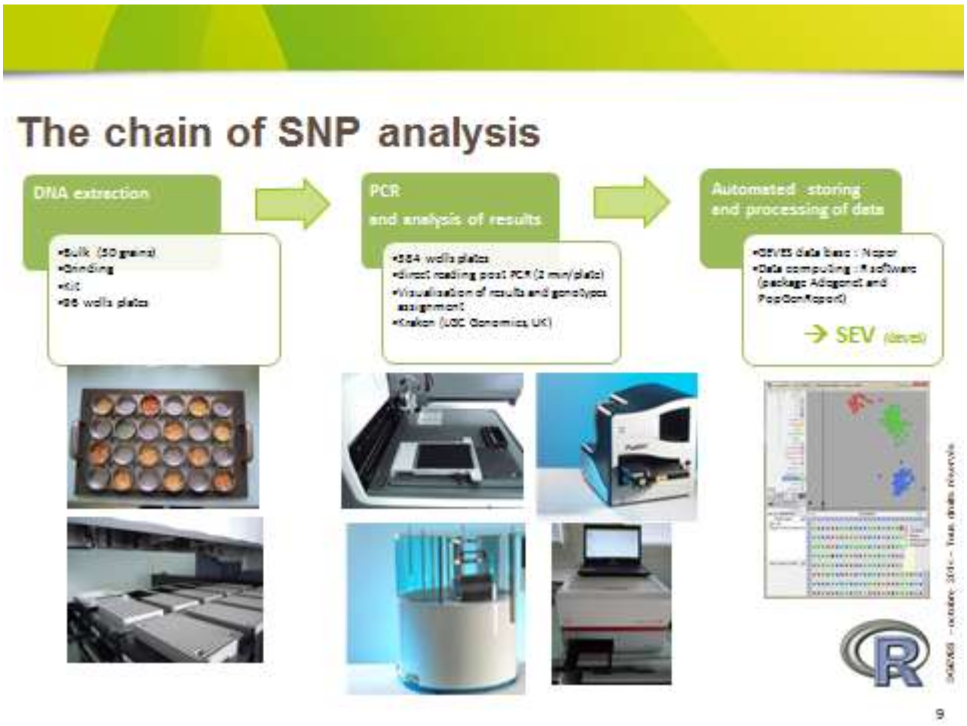
Reference collection (2013) analysis

Distribution of the genetic distance among varieties



Comparison by genetic distances

- SNP data analysis required a switch from pattern comparison to genetic distance calculation.
- Genetic distances are calculated
  - Calculation under « R » software
    - (Based on Kosman and Leonard 2005 + script from PopGenReport package)
- Missing data are not taken into account
  - > thus avoiding any possible bias due to missing data



The chain of SNP analysis

DNA extraction

- Bulk (50 grains)
- Grinding
- Kit
- 96 wells plates

PCR and analysis of results

- 384 wells plates
- direct reading post PCR (2 min/plate)
- Visualisation of results and genotypes assignment
- Kraken (LGC Genomics, UK)

Automated storing and processing of data

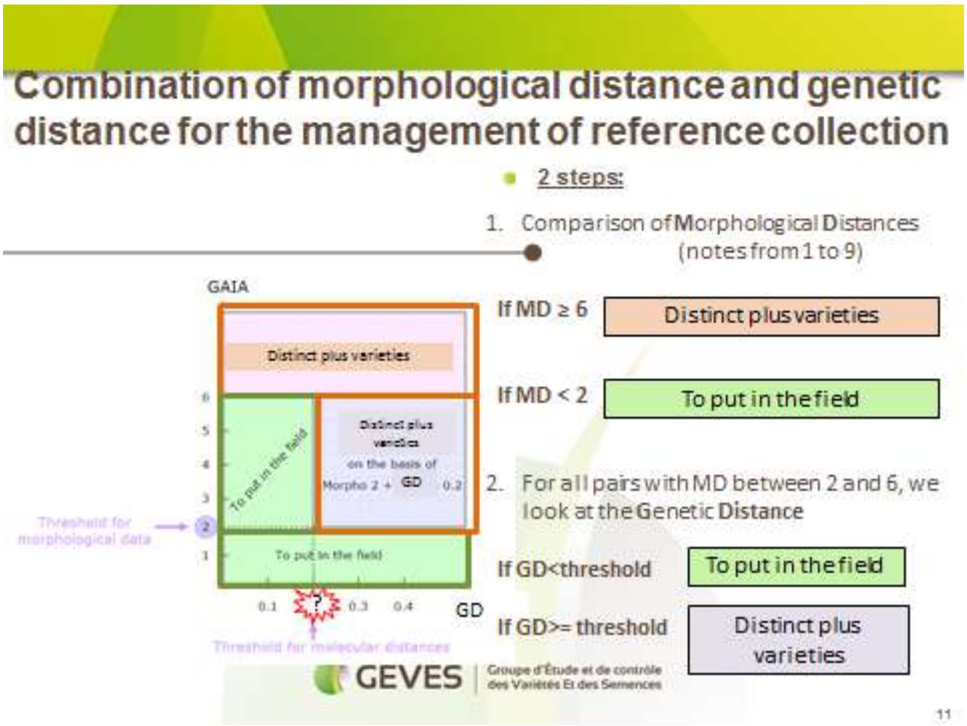
- GEVES data base : Nopor
- Data computing : R software (package Adegenet and PopGenReport)

→ SEV (geves)

PRACTICAL APPLICATIONS FOR MAIZE DUS TESTING AT GEVES

- Management of the reference collection by combining morphological distance and genetic distance
- Verification of new samples / Renewal of reference material
- Checking hybrid conformity

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



Combination of morphological distance and genetic distance for the management of reference collection

- 2 steps:

1. Comparison of Morphological Distances (notes from 1 to 9)

If MD ≥ 6 — Distinct plus varieties

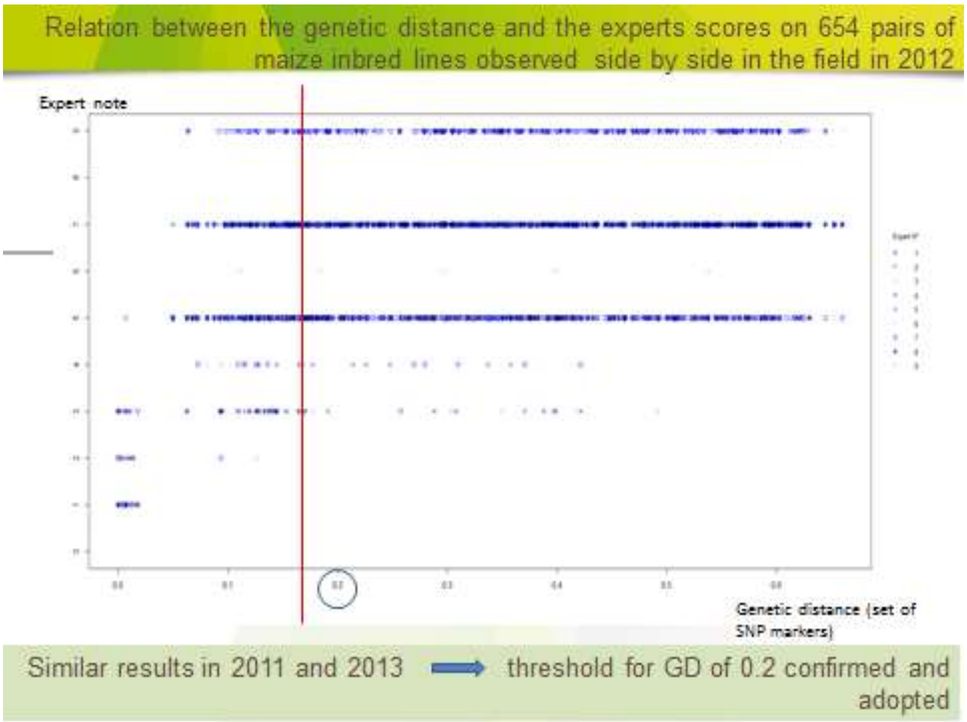
If MD < 2 — To put in the field

2. For all pairs with MD between 2 and 6, we look at the Genetic Distance

If GD<threshold — To put in the field

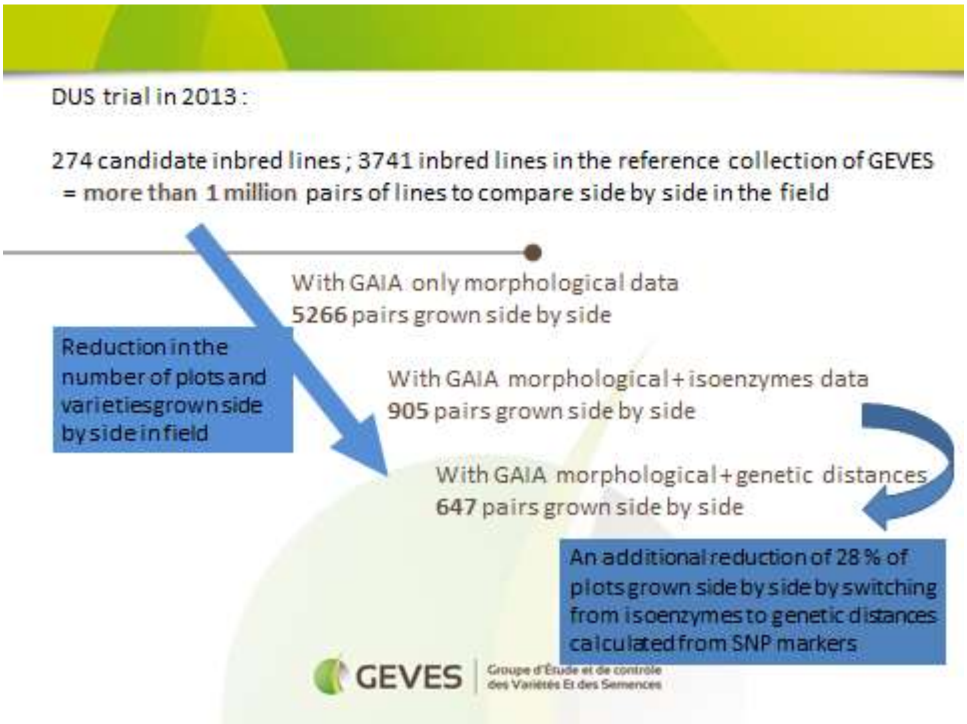
If GD>= threshold — Distinct plus varieties

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



Relation between the genetic distance and the experts scores on 654 pairs of maize inbred lines observed side by side in the field in 2012

Similar results in 2011 and 2013 ⟹ threshold for GD of 0.2 confirmed and adopted



DUS trial in 2013 :

274 candidate inbred lines ; 3741 inbred lines in the reference collection of GEVES
= more than 1 million pairs of lines to compare side by side in the field

With GAIA only morphological data
5266 pairs grown side by side

Reduction in the number of plots and varieties grown side by side in field

With GAIA morphological + isoenzymes data
905 pairs grown side by side

With GAIA morphological + genetic distances
647 pairs grown side by side

An additional reduction of 28 % of plots grown side by side by switching from isoenzymes to genetic distances calculated from SNP markers

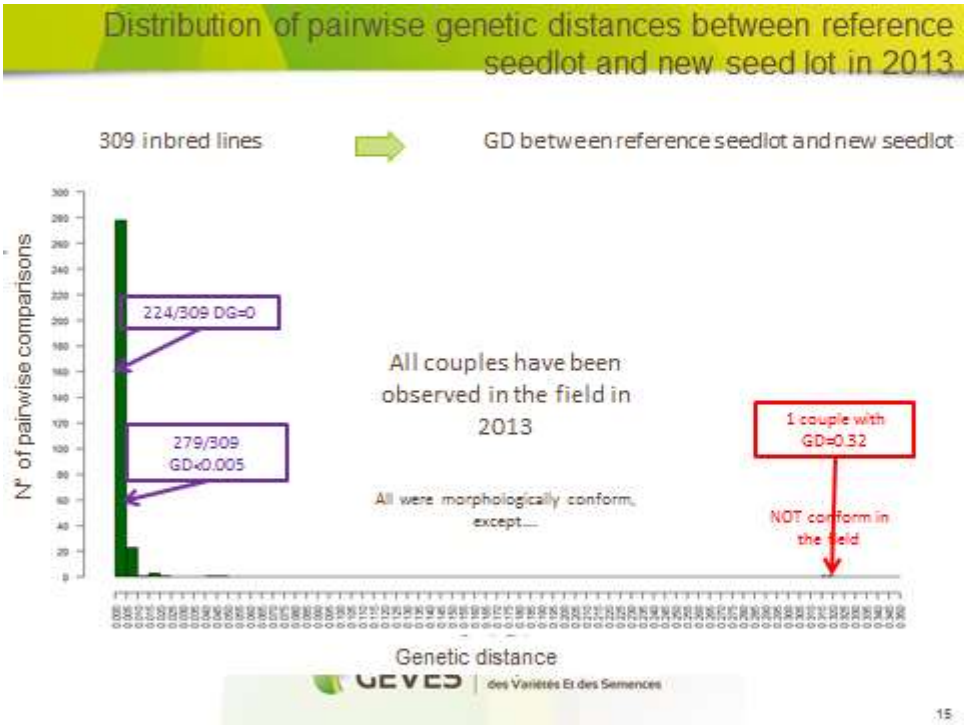
GEVES | Groupe d'Étude et de contrôle des Variétés Et des Semences

Renewal of reference material

- We maintain a viable reference collection of seeds
- Seeds easily stored, when needed new samples requested from the **maintainer**, which are compared to original sample in a field trial.
- But sometimes the old reference material does not germinate, or we don't have enough seeds left.
  - In this case how to ensure that the identity of the new sample is the same as the reference sample?
- > We compare with the descriptions in the database…
- … But using molecular techniques can be much more efficient !

Calculation of Genetic Distances between the reference sample and the new sample received from the maintainer

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



Distribution of pairwise genetic distances between reference seedlot and new seed lot in 2013

309 inbred lines → GD between reference seedlot and new seedlot

224/309 DG=0

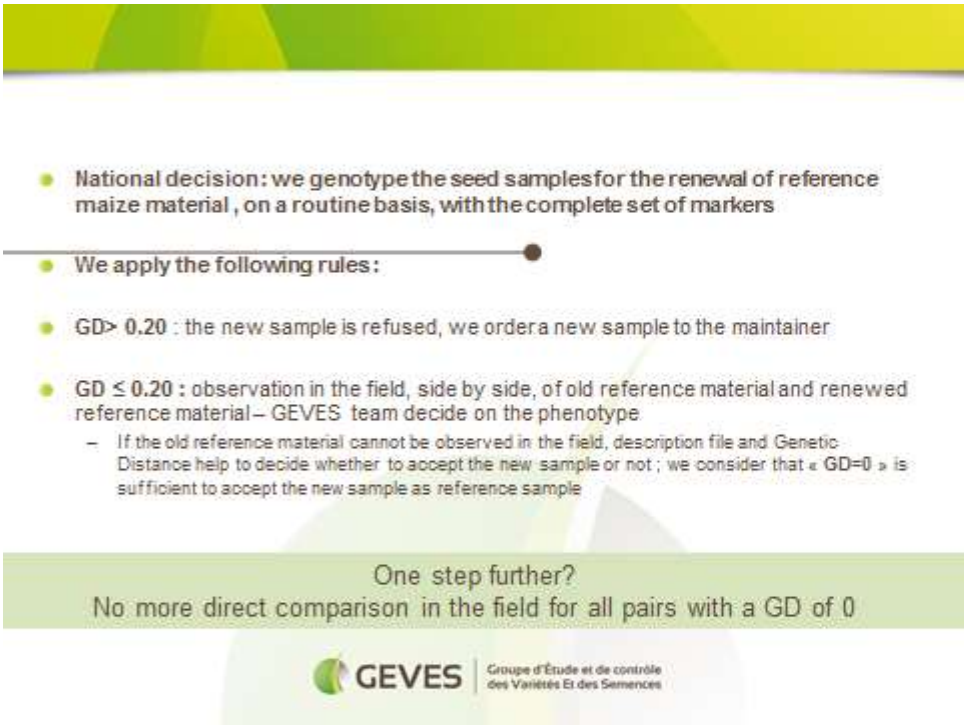
279/309 GD<0.005

All couples have been observed in the field in 2013

All were morphologically conform, except....

1 couple with GD=0.32

NOT conform in the field



- **National decision: we genotype the seed samples for the renewal of reference maize material , on a routine basis, with the complete set of markers**
- **We apply the following rules :**
- **GD> 0.20** : the new sample is refused, we order a new sample to the maintainer
- **GD ≤ 0.20 :** observation in the field, side by side, of old reference material and renewed reference material – GEVES team decide on the phenotype
  - If the old reference material cannot be observed in the field, description file and Genetic Distance help to decide whether to accept the new sample or not ; we consider that « **GD=0** » is sufficient to accept the new sample as reference sample

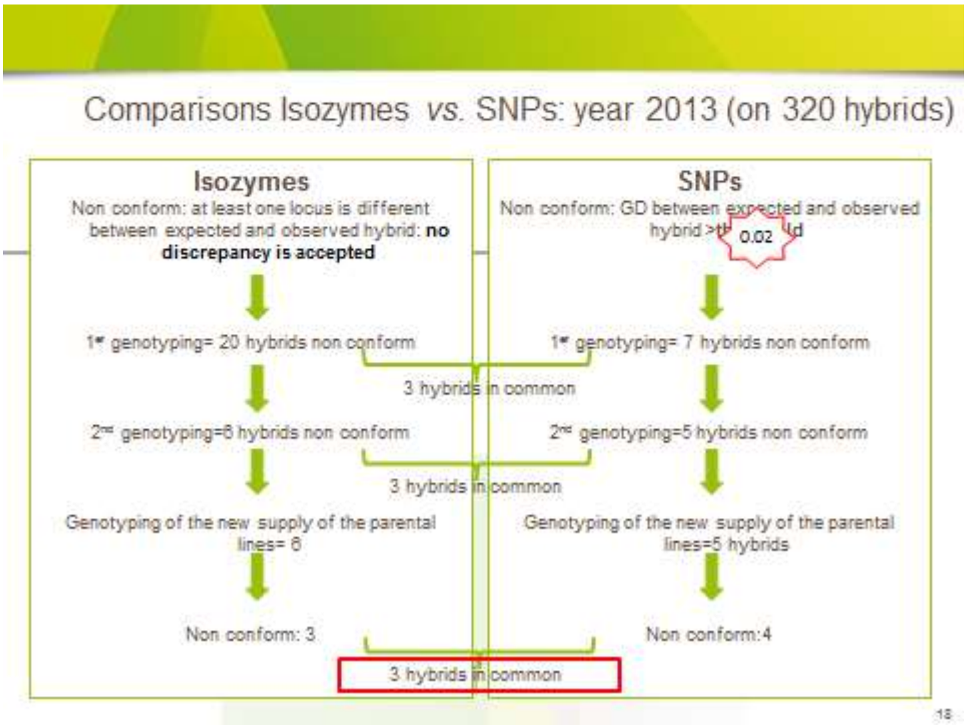
One step further?
No more direct comparison in the field for all pairs with a GD of 0

Checking hybrid conformity

- 2 years of data 2012 and 2013
- Comparison of the results for hybrid conformity : isoenzymes *vs.* SNPs
- With SNPs

  => **calculation of the Genetic Distance between the observed hybrid and the expected hybrid** (expected from the genotype of the parental inbred lines)

  ➡ establishment of a new rule and a threshold for hybrid conformity with SNPs



Comparisons Isozymes *vs.* SNPs: year 2013 (on 320 hybrids)

Isozymes

Non conform: at least one locus is different between expected and observed hybrid: **no discrepancy is accepted**

1st genotyping= 20 hybrids non conform

2nd genotyping=6 hybrids non conform

Genotyping of the new supply of the parental lines= 6

Non conform: 3

SNPs

Non conform: GD between expected and observed hybrid >th ld

0.02

1st genotyping= 7 hybrids non conform

2nd genotyping=5 hybrids non conform

Genotyping of the new supply of the parental lines=5 hybrids

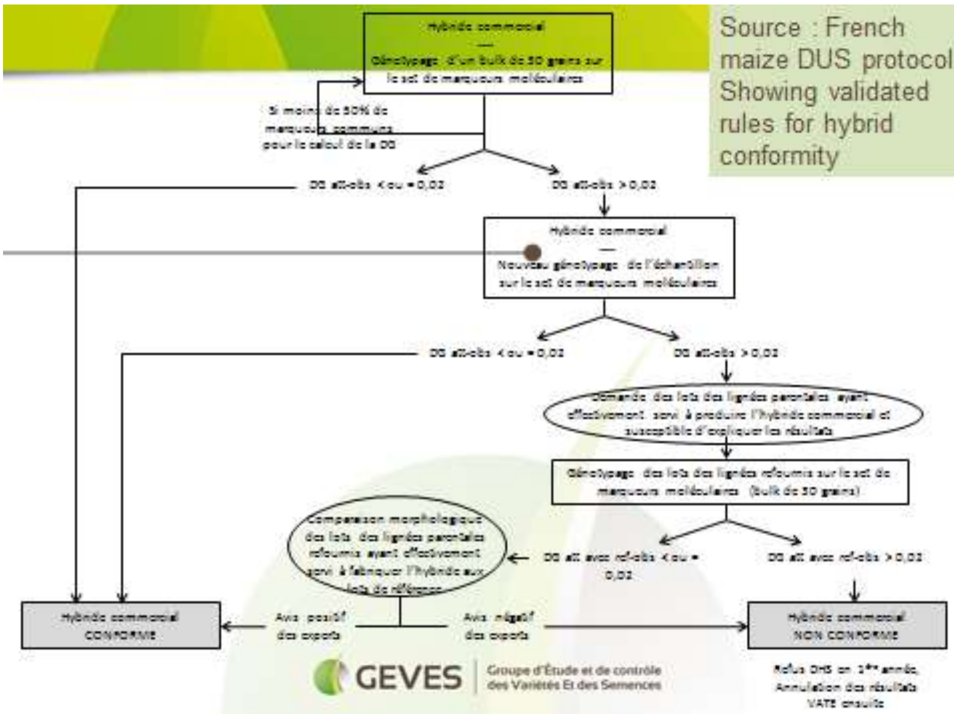
Non conform:4

3 hybrids in common

3 hybrids in common

3 hybrids in common

Source : French maize DUS protocol. Showing validated rules for hybrid conformity



Hybride commercial
—
Génotypage d'un bulk de 50 grains sur le set de marqueurs moléculaires

Si moins de 50% de marqueurs communs pour le calcul de la DG

DG att-obs < ou = 0,02

DG att-obs > 0,02

Hybride commercial
—
Nouveau génotypage de l'échantillon sur le set de marqueurs moléculaires

DG att-obs < ou = 0,02

DG att-obs > 0,02

Demande des lots des lignées parentales ayant effectivement servi à produire l'hybride commercial et susceptible d'expliquer les résultats

Génotypage des lots des lignées refournis sur le set de marqueurs moléculaires (bulk de 50 grains)

DG att avec ref-obs < ou = 0,02

DG att avec ref-obs > 0,02

Comparaison morphologique des lots des lignées parentales refournis ayant effectivement servi à fabriquer l'hybride aux lots de référence

Avis positif des experts

Avis négatif des experts

Hybride commercial
CONFORME

Hybride commercial
NON CONFORME

Refus DHS en 1ère année,
Annulation des résultats
VATE ensuite

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

---

Conclusions

- From the two set of SNP provided by UFS and INRA we have established **an optimized set of SNP**

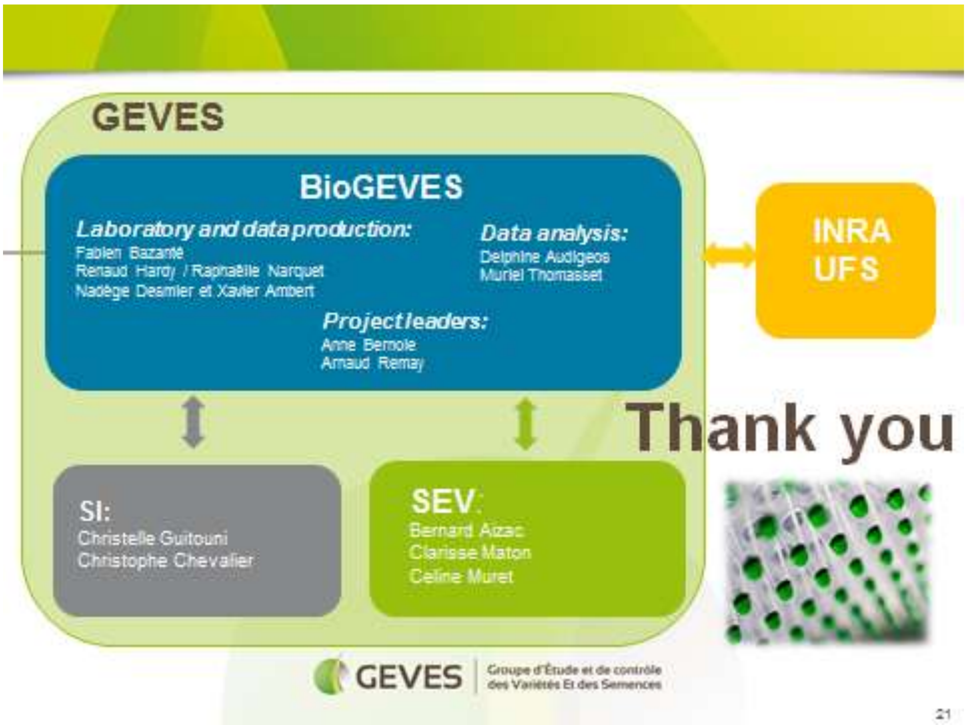
- Tests carried out at GEVES over a three year period **validate** the use of SNP technology
- We have gone through all the steps to achieve **technical validations and official approval**
- **New rules** were adopted at the national official level for DUS testing and this analytic scheme was accepted by the *Community Plant Variety Office of the European Union* (CPVO).

- This approved set of 312 SNP can be used for maize genotyping analysis applied to:
  - (a) management of the reference collection,
  - (b) renewal of reference material;
  - (c) conformity of hybrid formula.

- GEVES has the equipment and the associated processes settled in its laboratory. Since the beginning of this year, SNP technology is **in routine use for maize genotypes** analysis and **electrophoresis is not used anymore** as routine characteristics.

- **Additional benefits** : acquisition / development of transferable skills by engineers an technicians using SNP technology; reduction in time spent in field.
- SNP protocols for **other species** are currently being developed at GEVES.

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



GEVES

BioGEVES

Laboratory and data production:
Fabien Bazanté
Renaud Hardy / Raphaëlle Narquet
Nadège Desmier et Xavier Ambert

Data analysis:
Delphine Audigeos
Muriel Thomasset

Project leaders:
Anne Bernole
Arnaud Remay

SI:
Christelle Guitouni
Christophe Chevalier

SEV:
Bernard Aizac
Clarisse Maton
Celine Muret

INRA
UFS

Thank you

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

[End of annex and of document]