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/17/22

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

Original: English
Date: September 7, 2018

USE OF SNP MARKERS FOR SOYBEAN VARIETY PROTECTION PURPOSES IN ARGENTINA

Document prepared by an expert from Argentina

Disclaimer: this document does| not represent UPOV policies or guidance

The presentation prepared by an expert from Argentina on the "Use of SNP markers for soybean variety protection purposes in Argentina" is reproduced in the Annex to this document.

[Annex follows]

ANNEX

USE OF SNP MARKERS FOR SOYBEAN VARIETY PROTECTION PURPOSES IN ARGENTINA

Presentation prepared by an expert from Argentina

Ministerio de Agroindustria
Presidencia de la Nación

Use of SNP markers for soybean variety protection purposes in Argentina

This work was prepared by experts from Argentina

September 10th, 2018

Presented by Dra. Ana Laura Vicario

www.inase.gov.ar @inaseargentina Instituto Nacional de Semillas

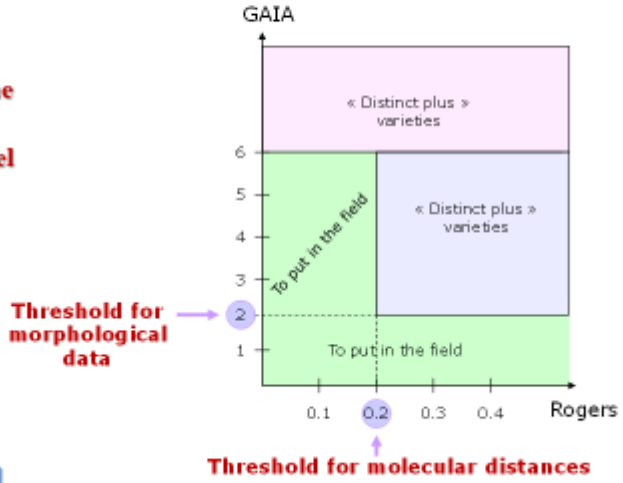
This project is been carried out thanks to a joint initiative between public and private sector, as we had regarded this issue as a problem to be solved together.

We are organized as an ad hoc group whose aim is to deal with variety protection issues based on the use of molecular markers.

INTRODUCTION

Model 2 combines phenotypic and molecular distances in the management of reference collections, reducing the amount of field work without decreasing the quality of the trial.

Based on the GEVES Maize model



OBJECTIVES

A) To create an effective instrument, based on the combination of phenotypic and molecular distances in order to reduce the list of cultivars that should be compared in the field, maintaining the quality of the Distinctness, Uniformity and Stability test (DUS test).

1) to select a set of SNPs markers distributed in the genome that could represent the genetic background and could yield an efficient discrimination power (set "Y"),

2) to determine the molecular distance threshold and the minimum phenotypical distance threshold, which in combination allow to select varieties that require comparison in the field.

B) To select a subset of SNPs for variety identification for:

a) seed trade control and
b) PBR enforcement.

We call this subset "Z" and will be composed by the smallest number of markers that generates a unique DNA profile for each variety.

MATERIALS AND METHODS

Selection of varieties, sample preparation and genotyping

- Varieties analyzed: 858, coming from different breeding programs.
- 378 were provided by the breeding companies and institutions supporting this project. This set of varieties represents those commercialized nowadays.
- 480 varieties were provided by INASE and represent public and some still protected varieties with few or none representation in the seed commerce.
- 31 Duplicates were included in the test as an indication of marker variation in time and within the sample.
- Groups of 100 seeds of each variety were assembled (including duplicates).
- DNA was extracted and samples were sent to an external genotyping service provider for genotyping with the SoySNP6K chip, a subset of the Illumina Infinium Beadchip SoySNP50K developed by the United States Department of Agriculture.

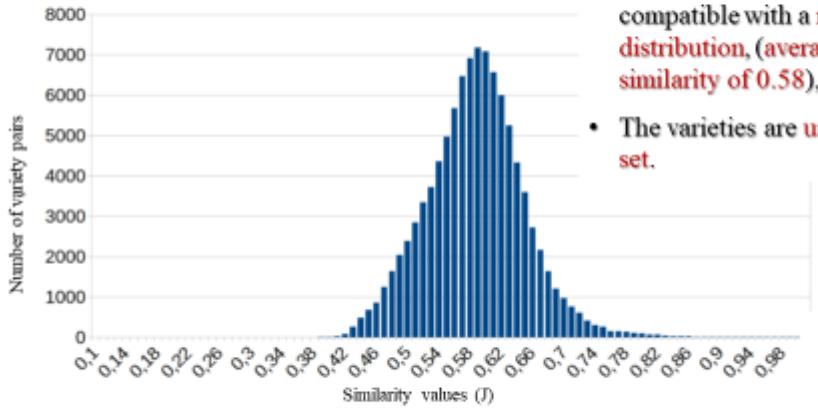
MATERIALS AND METHODS

Morphological characteristics

- 19 relevant characteristics for the distinctness of soybean varieties were selected.
- These characteristics were grouped in three levels of reliability (least reliable, reliable and most reliable) and weightings to each combination of their expression levels were assigned.
- They were used for comparison with molecular processed data and for the determination of the minimum phenotypic threshold.

RESULTS

- Jaccard coefficient (J)
- 6K SNPs.

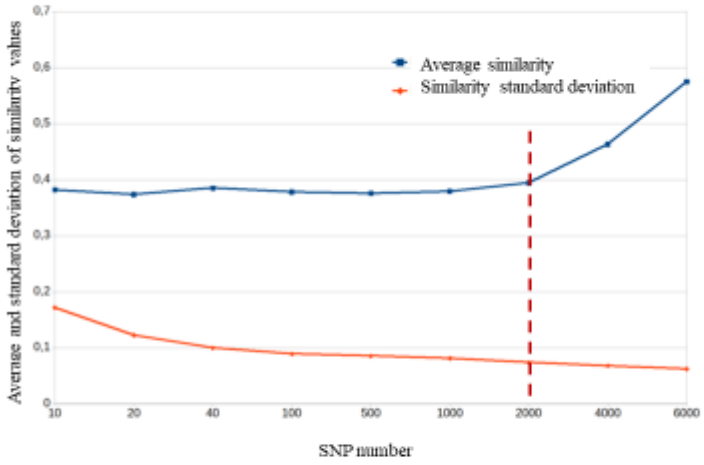


- The distribution is compatible with a normal distribution, (average similarity of 0.58),
- The varieties are unique set.

Histogram of similarities between all pairs of varieties using 6000 SNPs for 858 varieties.

RESULTS

Definition of "Y"

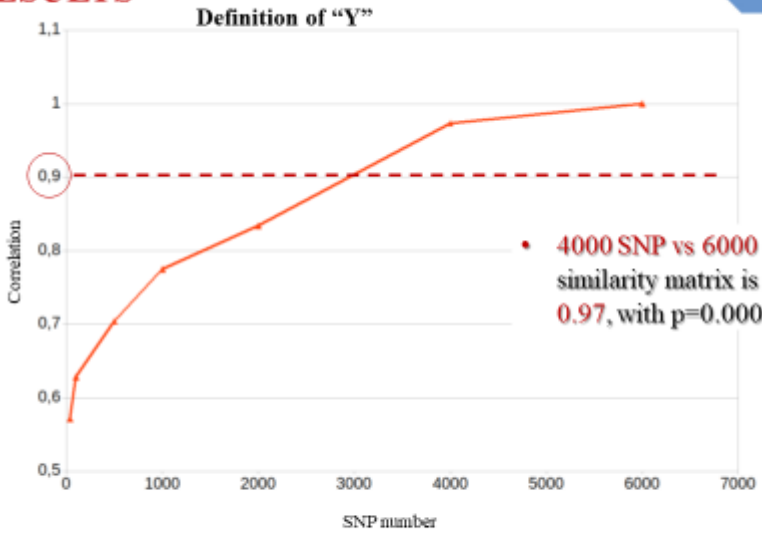


The average of similarity increases when incorporating the less informative SNPs.

The standard deviation, in turn, decreases as these SNPs have less variation.

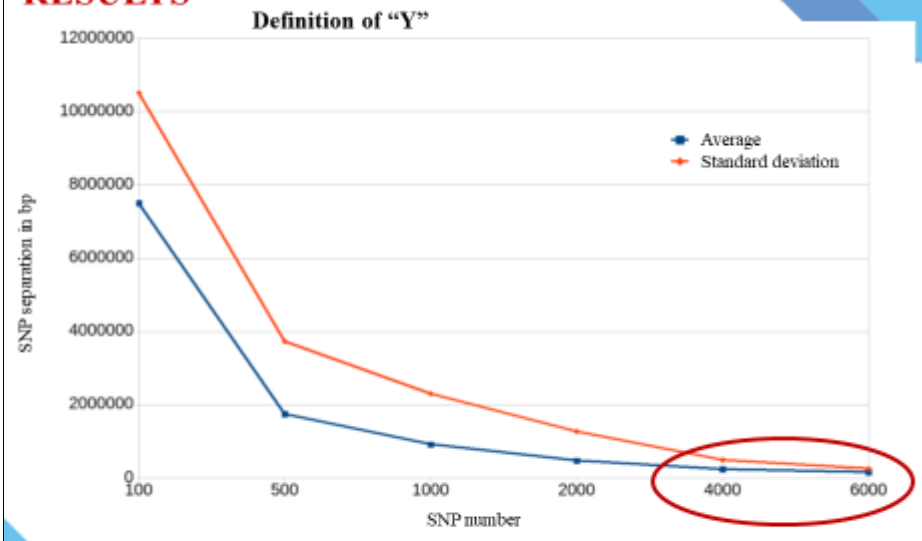
Evolution of the average and standard deviation of the similarities among pairs according to the number of SNPs used for the calculation.

RESULTS



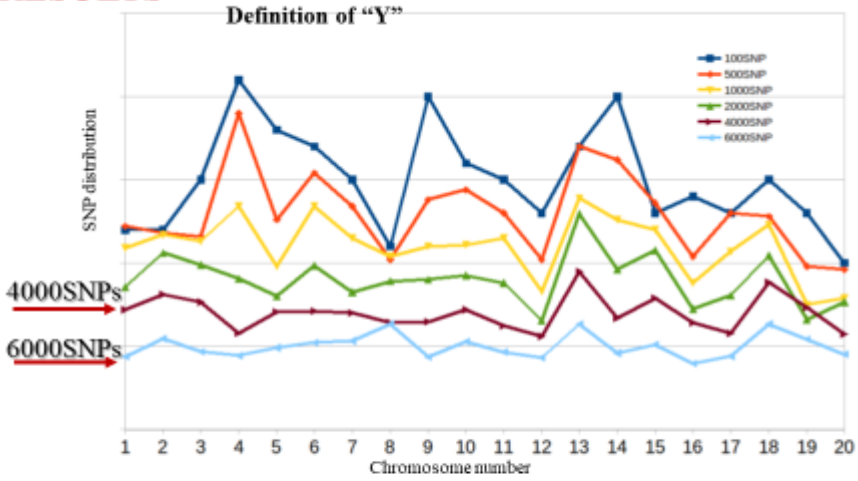
Correlations, using Mantel test, for the similarity matrix of 6000 SNPs against 100, 500, 1000, 2000, 4000 and 6000 SNPs.

RESULTS



Average and standard deviation of the SNPs separation according to the number of SNPs used.

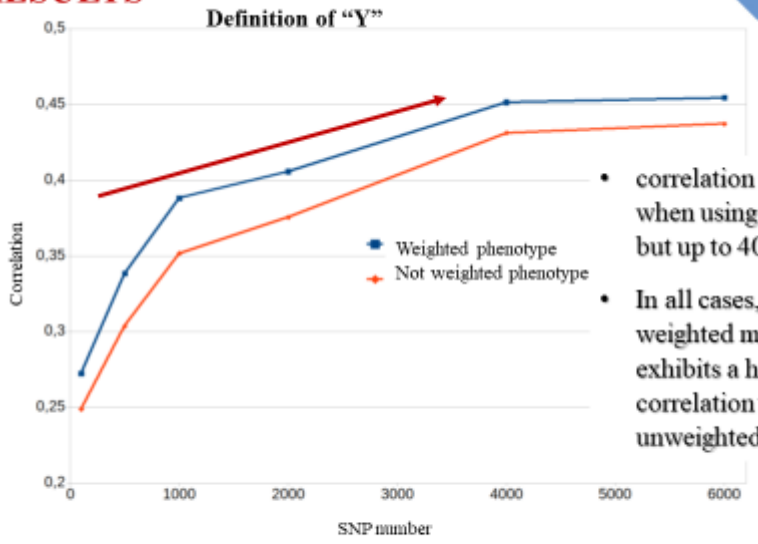
RESULTS



Distribution of the number of SNPs per chromosome (normalized for the chromosome length).

The graph represents the distribution equity for different number of markers throughout the chromosomes.

RESULTS



- correlation increases when using more SNPs, but up to 4000SNPs.
- In all cases, the weighted matrix exhibits a higher correlation than the unweighted one.

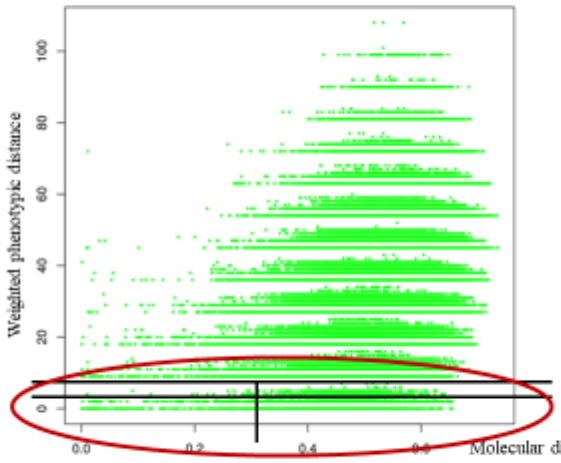
Correlations, using Mantel test, for morphological distance matrices -with and without weighting- against molecular ones (1-J).

RESULTS

The overall conclusion of the data presented up to now, indicates that the SNP subset "Y" is composed by the most polymorphic 4000 SNPs from the 6K Illumina chip.

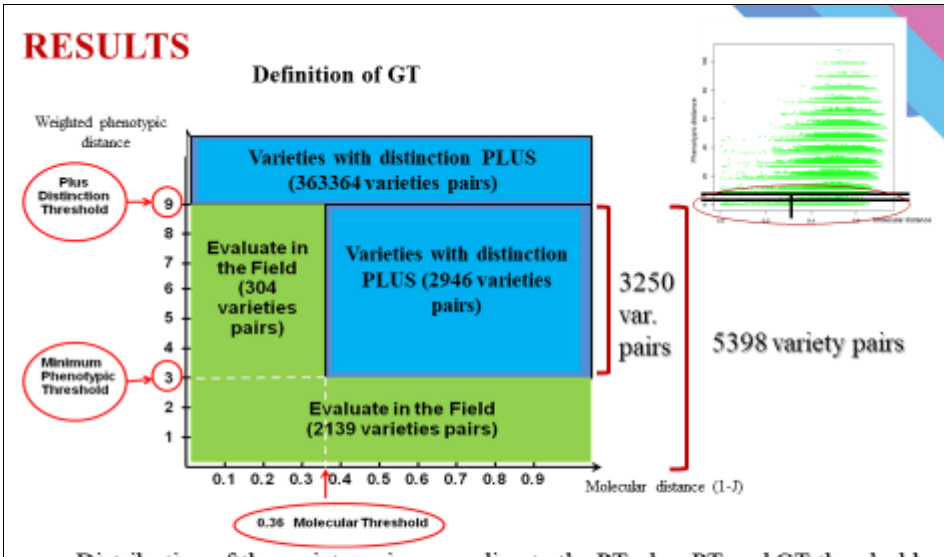
RESULTS

Definition of GT



367653 variety pairs obtained from 858 varieties analyzed.

RESULTS



- Distribution of the variety pairs according to the PT-plus, PT and GT thresholds.
- The GT was selected based on a 3-year field trials where 98% of the pairs gave genetic distances above 0.36.
- **If we would have applied model 2 over the years, we would have saved 55% field work.**

ONGOING ACTIVITIES

- Validate the minimum phenotypic distance threshold and the molecular distance threshold, by testing a subset of varieties that have applied for registration during the last two years.

- The second aim of this work is to calculate a subset of SNPs for variety identification. We call this subset "Z" and will be composed by the smallest number of markers that generates a unique DNA profile for each variety. Those markers will be the most polymorphic markers selected on the criteria that, for a given variety pair, the difference is at least 3 SNPs.

ACKNOWLEDGMENT

Experts from INASE and FCEyN-UBA

Maria Fernanda Dalmau - INASE
María Alicia Loray - INASE
Mariano Alejandro Mangieri - INASE
Ana Laura Vicario - INASE
Marcelo Marti - FCEyN-UBA
German F Burguener - FCEyN-UBA
Jonathan Javier Zaiat - FCEyN-UBA
Raimundo Lavignolle - INASE
Hernando Pecci - INASE
Ignacio Aranciaga - INASE

Experts from Argentine Seed Association (ASA)

Mirta Antongiovanni
Gaspar Malone
Edgardo Gabriel Bresso
Roberto Oscar Enriquez
Claudia Karina Herrmann
Maria Eva González
Andres Pereyra
Mauro Meier
Gustavo Abratti
Julieta Alloatti
Carlos Dezar
Mariana Gallo
Alberto Javier Salvador
Alfredo Paseyro
Juan Erdmann

We thank them all for their collaboration

THANKS FOR YOUR ATTENTION

[End of Annex and of document]