

# A practical introduction to Genome-Wide Association Studies applied to plants

Luis Garreta

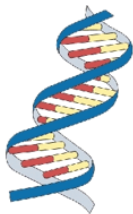
Electiva de Bioinformática  
MAESTRÍA EN INFORMÁTICA BIOMÉDICA  
Universidad del Bosque  
Bogotá-Colombia

August 28, 2020

# Methods

- ▶ Genome-Wide Association Studies (GWAS)
- ▶ Genomic Selection (GS)

# Genotype and Phenotype

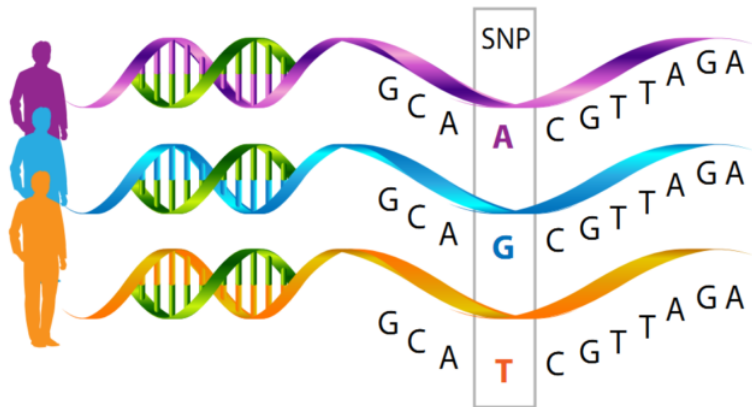


**Genotypes** are the genetic make-up of an individual<sup>5</sup>.

**Phenotypes** are the physical traits and characteristics of an individual and are influenced by their genotype and the environment<sup>6</sup>.

# Single nucleotide polymorphisms (SNPs)

- ▶ The most common type of genetic **variation** among people.
- ▶ It is a substitution of a single nucleotide that occurs at a specific position in the genome



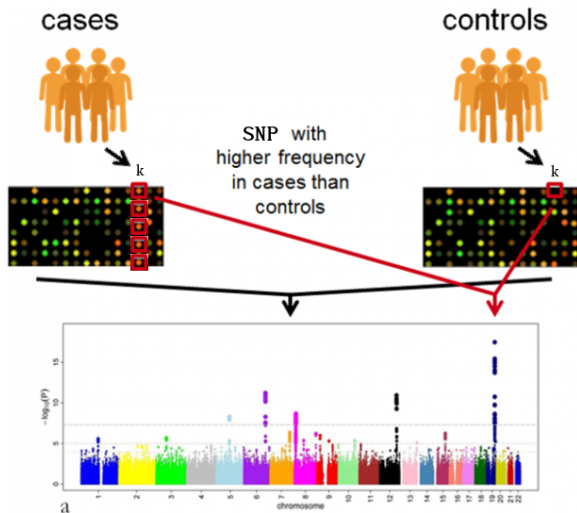
# SNPs in Whole Genomes

<u>Plant</u>	<u>Sequence</u>
1	GAATTCCGCAATGCAGGTTAAGAGCTCTGTGAAAGAGGAAAACGAAAAAC
2	GAATTCCGCAATGCAGGTTAAGAGCTCTGTGAAAGAGGAAAACGAAAAAC
3	GAATTCCGCAATGCAGGTTAAGAGCTCTGTGAAAGAGGAAAACGAAAAAC
4	GAATTCCGCAATGCAGGTTAAGAGCTTTGTGAAAGAGGAAAACGAAAAAC
5	GAATTCCGCAATGCAGGTTAAGAGCTCTGTGAAAGAGGAAAACGAAAAAC
6	GAATTCCGCAATGCAGGTTAAGAGCTCTGTGAAAGAGGAAAACGAAAAAC
7	GAATTCCGCAATGCAGGTTAAGAGCTTTGTGAAAGAGGAAAACGAAAAAC
8	GAATTCCGCAATGCAGGTTAAGAGCTCTGTGAAAGAGGAAAACGAAAAAC

# GWAS: Genome-Wide Association Studies

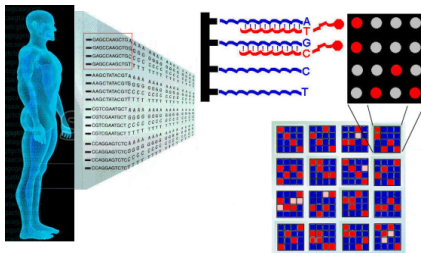
An approach used in genetics research to associate specific genetic variations (SNPs) with particular traits (e.g. diseases, agronomic trait).

GWAS involves scanning the genomes from many different people and looking for SNPs that can be used to predict the presence of a disease.

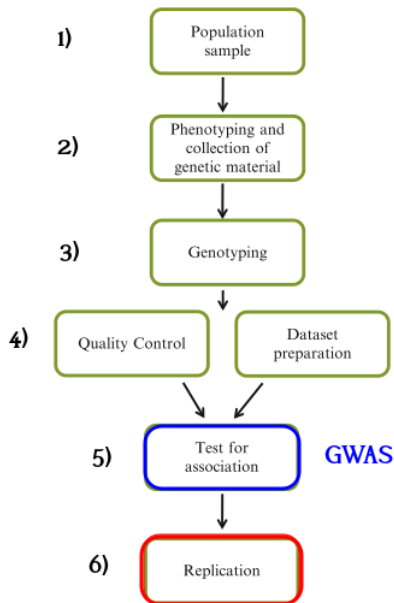


# SNP-chips (DNA microarray)

The DNA microarray is a tool used to determine whether the DNA from a particular individual contains a mutation in genes



# WGAS Process





# GWAS Methods (Methods for test associations)

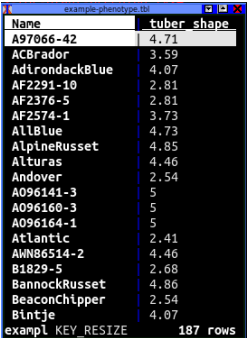
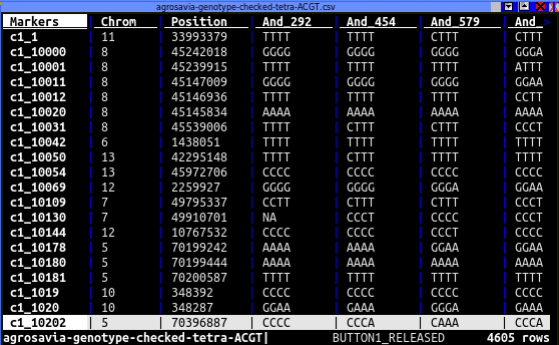
- ▶ Linear Models
- ▶ Linear Regression
- ▶ Logistic Regression
- ▶ Linear Mixed Models
- ▶ Bayesian Mixed Models

# Linear Mixed Model

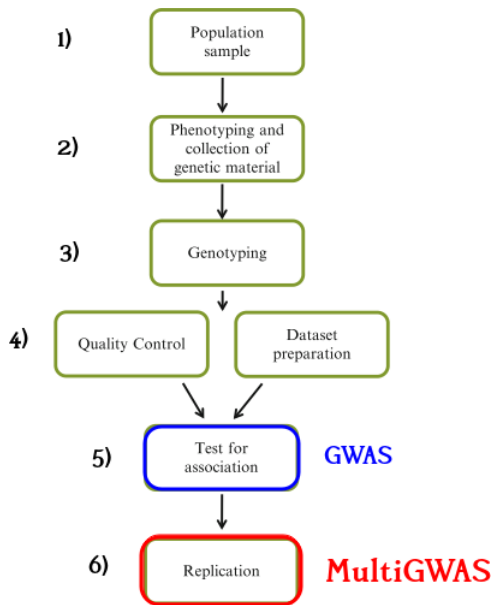
Phenotype	Genotype	Population structure	Unequal relatedness				
↓	↓	↓	↓				
Y =	SNP	+	Q (or PCs)	+	Kinship	+	e
	(fixed effect)		(fixed effect)		(random effect)		
<hr/>							
General Linear Model (GLM)							
<hr/>							
Mixed Linear Model (MLM)							

# Genomic Data Tetraploid Potato

- ▶ Example of tetraploid potato (Four set of chromosomes):
- ▶ Humans are diploids (Father and Mother)

Phenotype	Genotype
	

# MultiGWAS: a tool for integrating multiple GWAS tools



# Popular Tools for GWAS analysis

## 1. **GWASPoly:**

- ▶ Relatively new R library for GWAS in polyploid.

## 2. **PLINK:**

- ▶ Popular Program for GWAS in diploids
- ▶ Mainly used in humans.

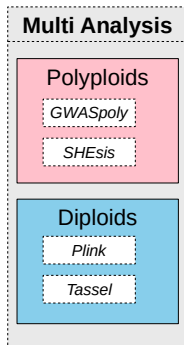
## 3. **TASSEL:**

- ▶ Popular Java Package and program for GWAS in diploids
- ▶ Mainly use in plants.

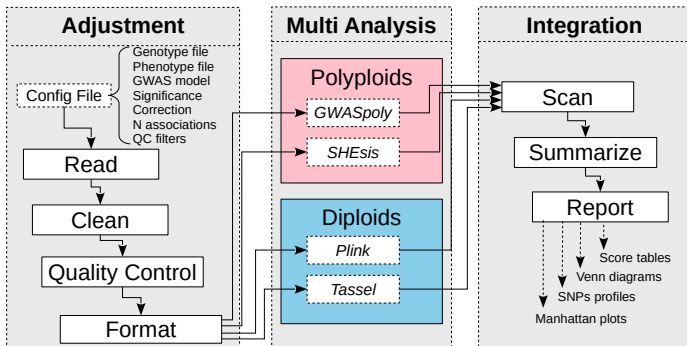
## 4. **SHEsis:**

- ▶ Simple Program for GWAS analysis in polyploids.

# MulitiGWAS: A tools for GWAS analysis integrating four GWAS tools



# MulitiGWAS: A tools for GWAS analysis integrating four GWAS tools



# MultiGWAS outputs

```
file:///home/lg/agrosavia/docs/presentations/  
biomedicine/multiGWAS-report.html
```



# What next?

- ▶ Genomic Selection