

Bioinformática en Mejoramiento de cultivos:

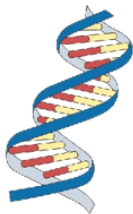
Selección Genómica sobre papás tetraploides

Luis Garreta

Corporación Colombiana de investigación agropecuaria - AGROSAVIA

November 14, 2020

Genotype and Phenotype

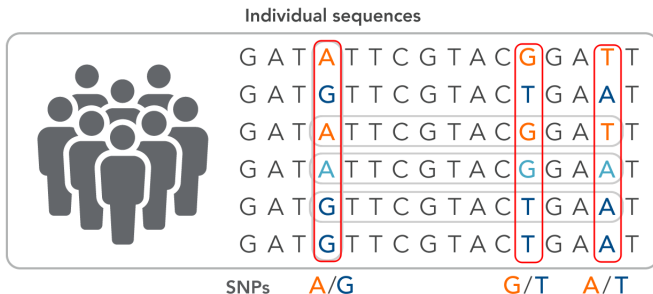


Genotypes are the genetic make-up of an individual⁵.

Phenotypes are the physical traits and characteristics of an individual and are influenced by their genotype and the environment⁶.

Genetic Markers: SNPs

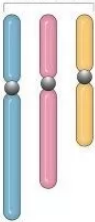
- ▶ **SNP**: Single nucleotide polymorphisms
- ▶ It is a variation in a single nucleotide in the DNA sequence
- ▶ The most common type of genetic **variation** among organisms.



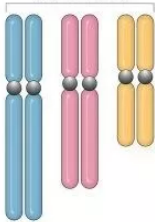
Polypoidy

Humans

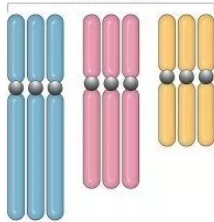
Haploid (N)



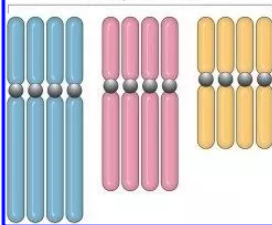
Diploid (2N)



Triploid (3N)

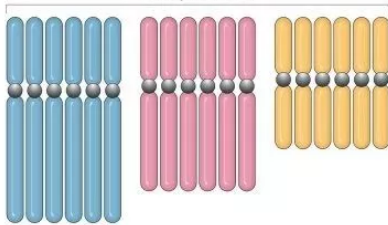


Tetraploid (4N)

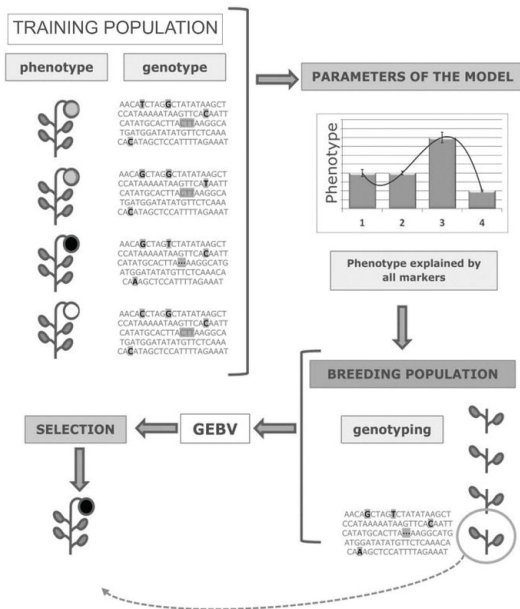


Plants

Hexaploid (6N)



Genomic selection strategy in plant breeding



Basic Model for Genomic Selection

$$Y_i = g(x_i) + e_i$$

Where:

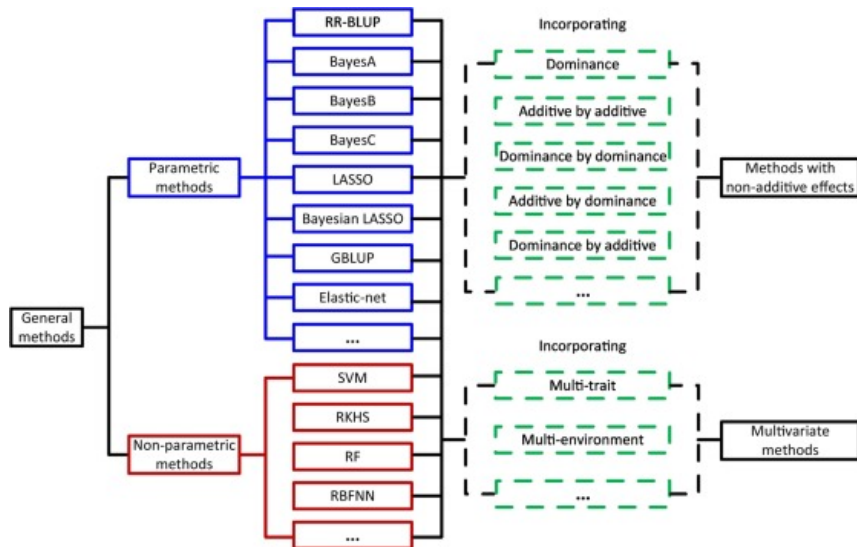
Y_i : An observed phenotype of individual i

x_i : Vector of SNPs genotypes on individual i

$g(x_i)$: Function relating genotypes to phenotypes

e_i : Error or residual term

Evaluation of Genomic Selection Methods



Genomic Datasets for Tetraploid Potato

Fenotipo

Name	tuber shape
A97066-42	4.71
ACBrador	3.59
AdirondackBlue	4.07
AF2291-10	2.81
AF2376-5	2.81
AF2574-1	3.73
AllBlue	4.73
AlpineRusset	4.85
Alturas	4.46
Andover	2.54
A096141-3	5
A096160-3	5
A096164-1	5
Atlantic	2.41
AWN86514-2	4.46
B1829-5	2.68
BannockRusset	4.86
BeaconChipper	2.54
Blintje	4.07

example KEY_RESIZE 187 rows

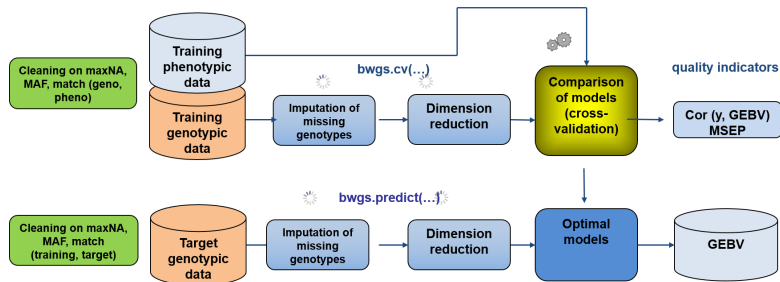
Genotipo tetraploide

Marker	ACBrador	AF2291-10	AF2376-5	AF2574-1	
c2_41437	AAGG	AAAA	AAAG	AAAG	...
c2_24258	GGGG	GGGG	AGGG	AAGG	...
c2_21332	TTCC	TTCC	TTCC	TTTC	...
c2_21320	TTTT	TTTT	TTTT	TTTT	...
c2_21318	NA	NA	CCCC	CCCC	...
c2_21314	TTTC	TTTT	TTTC	TTTC	...
c2_4430	NA	AAGG	NA	AAAG	...
c2_4437	TTCC	TTCC	TTCC	TTTC	...
c2_6674	AAGG	AAGG	AAGG	AGGG	...
c1_16362	GGGG	GGGG	GGGG	GGGG	...
c2_53375	GGGG	GGGG	GGGG	GGGG	...
c2_53376	TTTT	TTTT	TTTT	TTTT	...

Genotipo diploide

Marker	A96104-2	A97066-42	ACBrador	ADCP195284	
c2_41437	GA	GA	GA	AA	
c2_24258	GA	GA	GG	GG	
c2_21332	TC	TC	TC	CC	
c2_21320	TT	TT	TT	TT	
c2_21318	CC	TC	NA	NA	
c2_21314	TC	TC	TC	NA	
c2_4430	GA	GA	NA	GA	
c2_4437	TC	TC	TC	CC	
c2_6674	GA	GA	GA	GA	
c1_16362	GG	GG	GG	GG	
c2_53375	GG	GG	GG	NA	
c2_53376	TT	TT	TT	TT	
c2_48820	NA	TG	TG	TT	
c2_48821	TC	TC	TT	TT	

Workflow for Genomic Selection



Fenotipos seleccionados (características)

	Nombre	Código
1	Color primario brote	CPB
2	Color secundario brote	CSB
3	Color pulpa	CP
4	Color primario tubérculo	CPT
5	Color secundario tubérculo	CST
6	Color baya	CB
7	Color primario flor	CPF
8	Color tallo	CT

Tabla de Colores RHS Purpura-Azul

RHS	sRGB				CIE Lab D65 / 10°			CIE LCh D65 / 10°		
	Out of RGB	R	G	B	L	a	b	L	C	h
N57A		215	45	98	49	67	11	49	68	10
N57B		219	51	110	50	67	7	50	68	6
N57C		228	92	145	58	58	-2	58	58	358
N57D		231	121	164	65	48	-4	64	48	355
58A		171	65	97	43	47	4	43	47	5
58B		218	72	114	53	60	8	53	61	7
58C		235	93	132	59	58	6	59	59	7
58D		240	138	167	69	43	1	69	43	1
59A		123	49	68	31	35	4	31	35	7
59B		143	46	76	35	43	4	35	44	5
59C		163	54	91	40	48	3	40	48	3
59D		199	82	122	52	51	1	51	51	1
60A		154	45	75	36	47	8	36	48	9
60B		164	45	84	38	51	5	38	52	6
60C		172	59	101	43	50	1	42	50	0
60D		199	73	122	50	55	-1	50	55	359
61A		155	46	86	37	49	1	37	49	2
61B		186	50	99	44	58	4	44	58	3
61C		211	69	120	51	60	2	51	60	2
61D		244	112	153	64	55	2	64	55	2
62A		249	144	181	72	44	-2	72	44	357
62B		245	161	194	75	36	-5	75	36	352
62C		247	179	207	80	29	-5	80	29	350
62D		247	205	220	86	17	-3	86	18	351

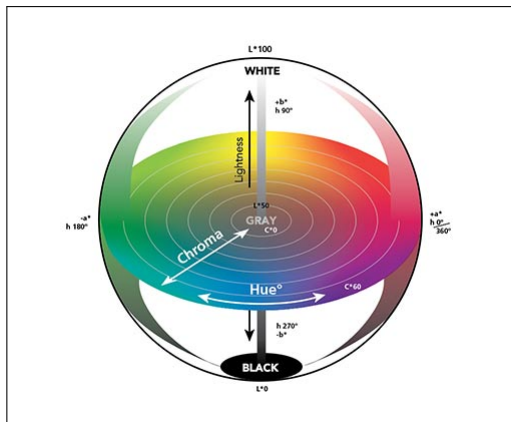


Espacio de
Colores
CIE LCh

Espacio de colores LCH

Componentes:

- ▶ L: Luminosity
 - ▶ Luminosidad
 - ▶ De Negro a Blanco)
- ▶ C: Chroma
 - ▶ Saturación
 - ▶ Vector
- ▶ H: Hue
 - ▶ Tono
 - ▶ Angulo

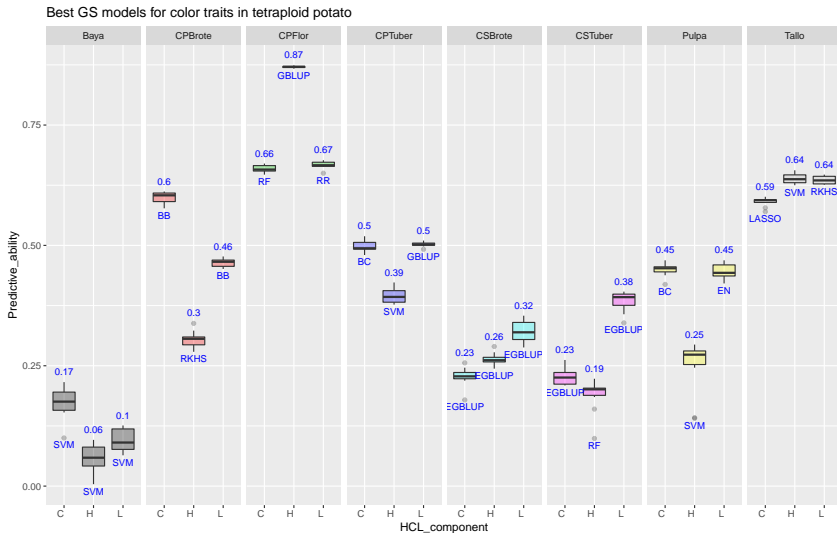


Total de Características a Evaluar

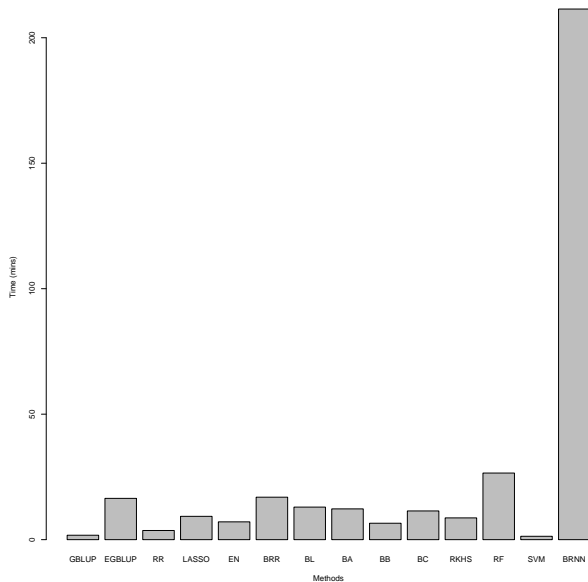
28 Fenotipos a evaluar (**Fenotipo Color** + **Componente LCH**):

Color Primario Brote-H Color Primario Brote-L Color Primario Brote-L	Color Secundario Tubérculo-H Color Secundario Tubérculo-C Color Secundario Tubérculo-L
Color Secudario Brote-H Color Secundario Brote-C Color Secundario Brote-L	Color Primario Flor-H Color Primario Flor-C Color Primario Flor-L
Color Pulpa-H Color Pulpa-C Color Pulpa-L	Color Baya-H Color Baya-C Color Baya-L
Color Primario Tubérculo-H Color Primario Tubérculo-C Color Primario Tubérculo-L	Color Tallo-H Color Tallo-C Color Tallo-L

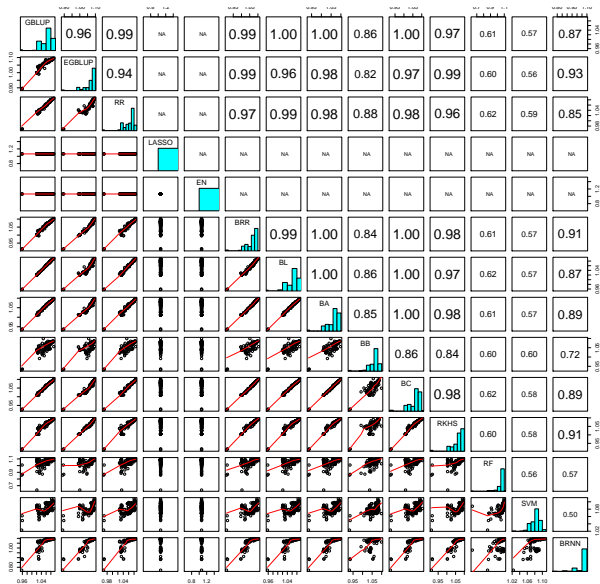
Mejor Modelo de Predicción por Características²



Costo computacional de los Modelos (Tiempo en mins)

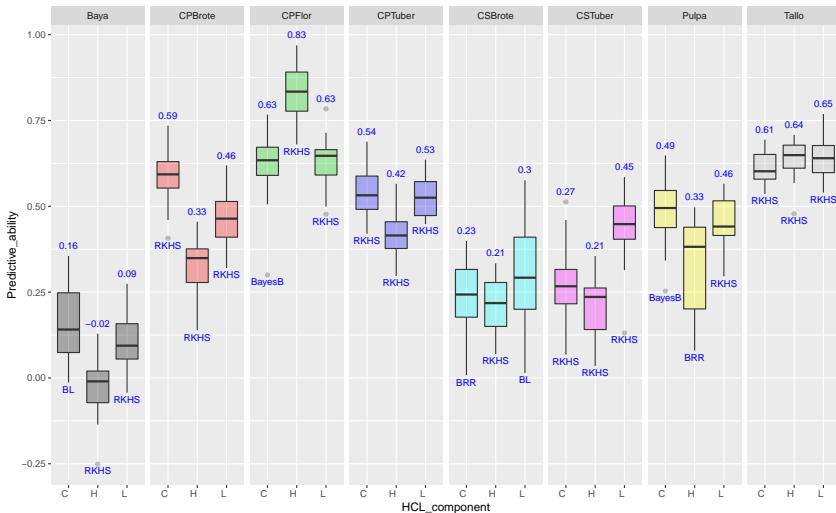


Comparación de Habilidad Predictiva entre los Modelos



Mejor Modelo de Predicción por Características (BGLR for Tetraploids)

Best GS models for color traits in tetraploid potato



Mejor Modelo de Predicción por Características³

