

# Análisis de asociación genética y validación de marcadores moleculares en caña de azúcar (*Saccharum* spp.) para la acumulación de sacarosa en ambientes húmedos.

INVESTIGACIÓN DOCTORAL

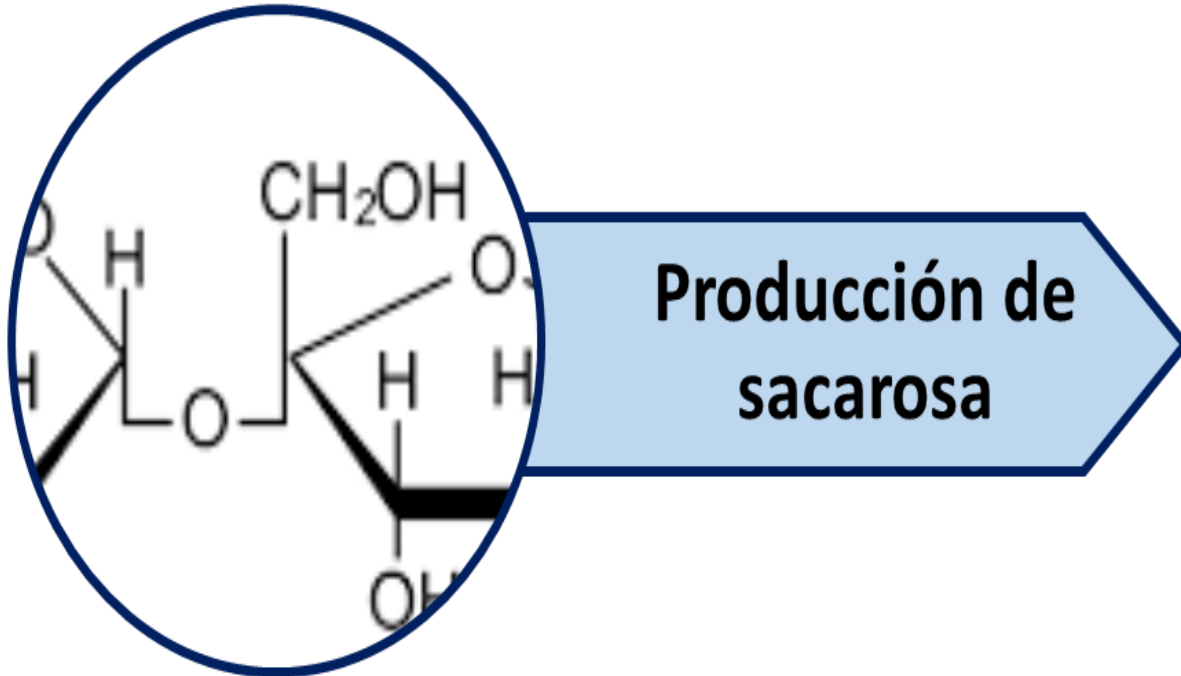
CAROLINA SAAVEDRA DÍAZ



Pontificia Universidad  
**JAVERIANA**  
Cali



# CARACTERÍSTICA



Acumulación de sacarosa en caña de azúcar depende del contenido de humedad del suelo (Viveros-Valens y Cuervo, 2014).



# MEGA-AMBIENTES



**Seco-Semiseco**

**Área : 67%**

**Piedemonte**

**Área : 7%**

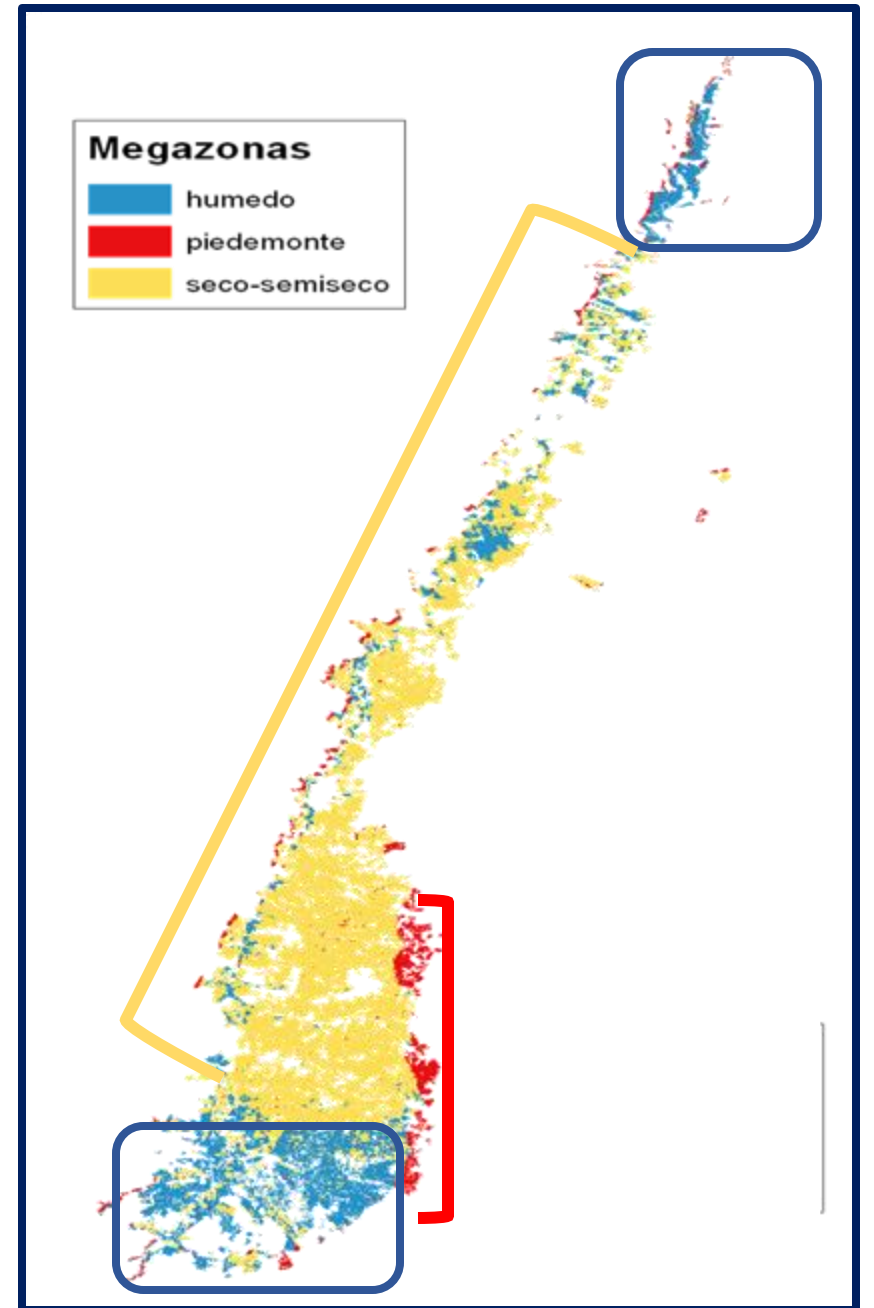
**Húmedo**

**Área : 26%**

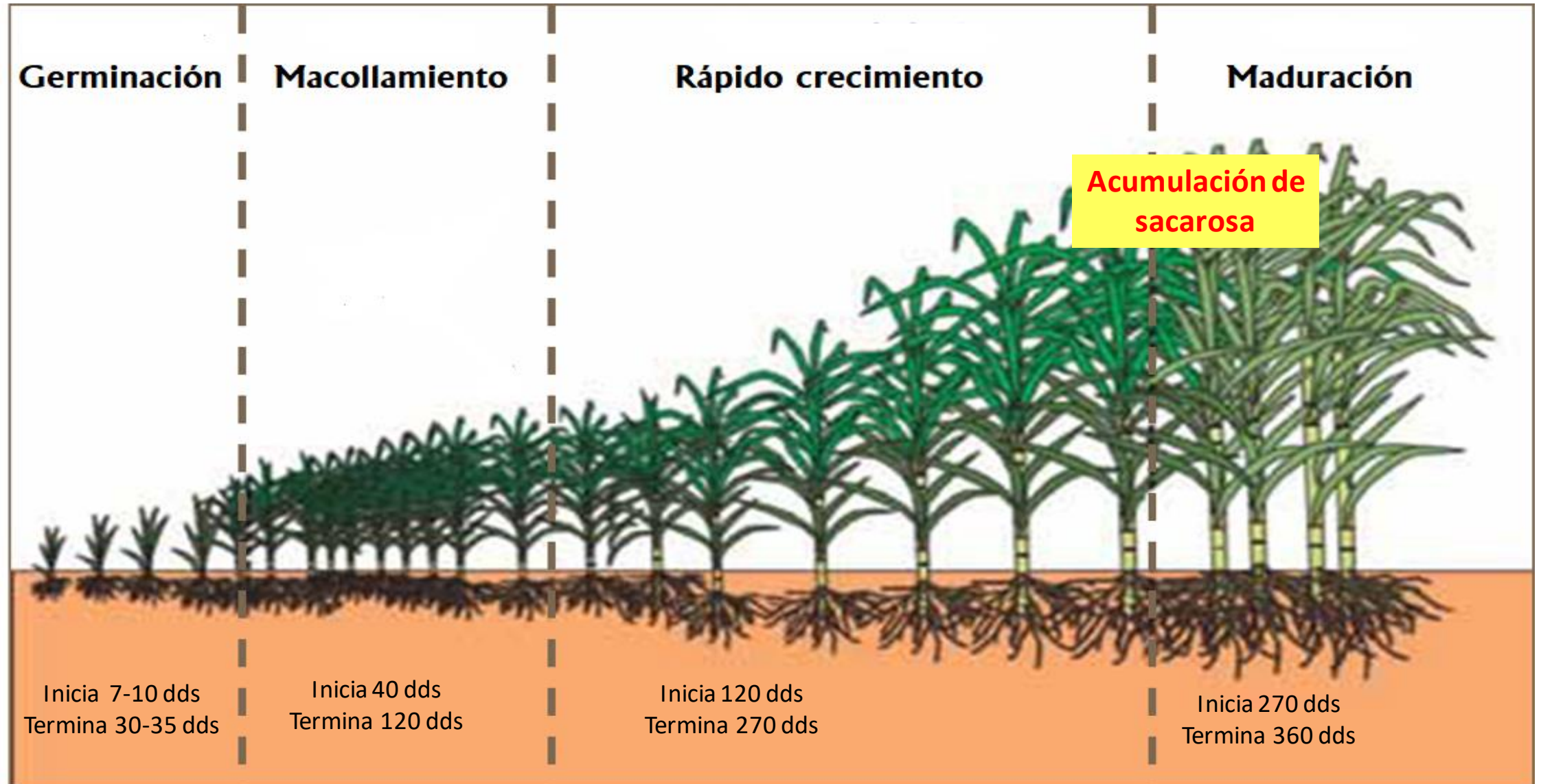


**Megazonas**

-  húmedo
-  piedemonte
-  seco-semiseco

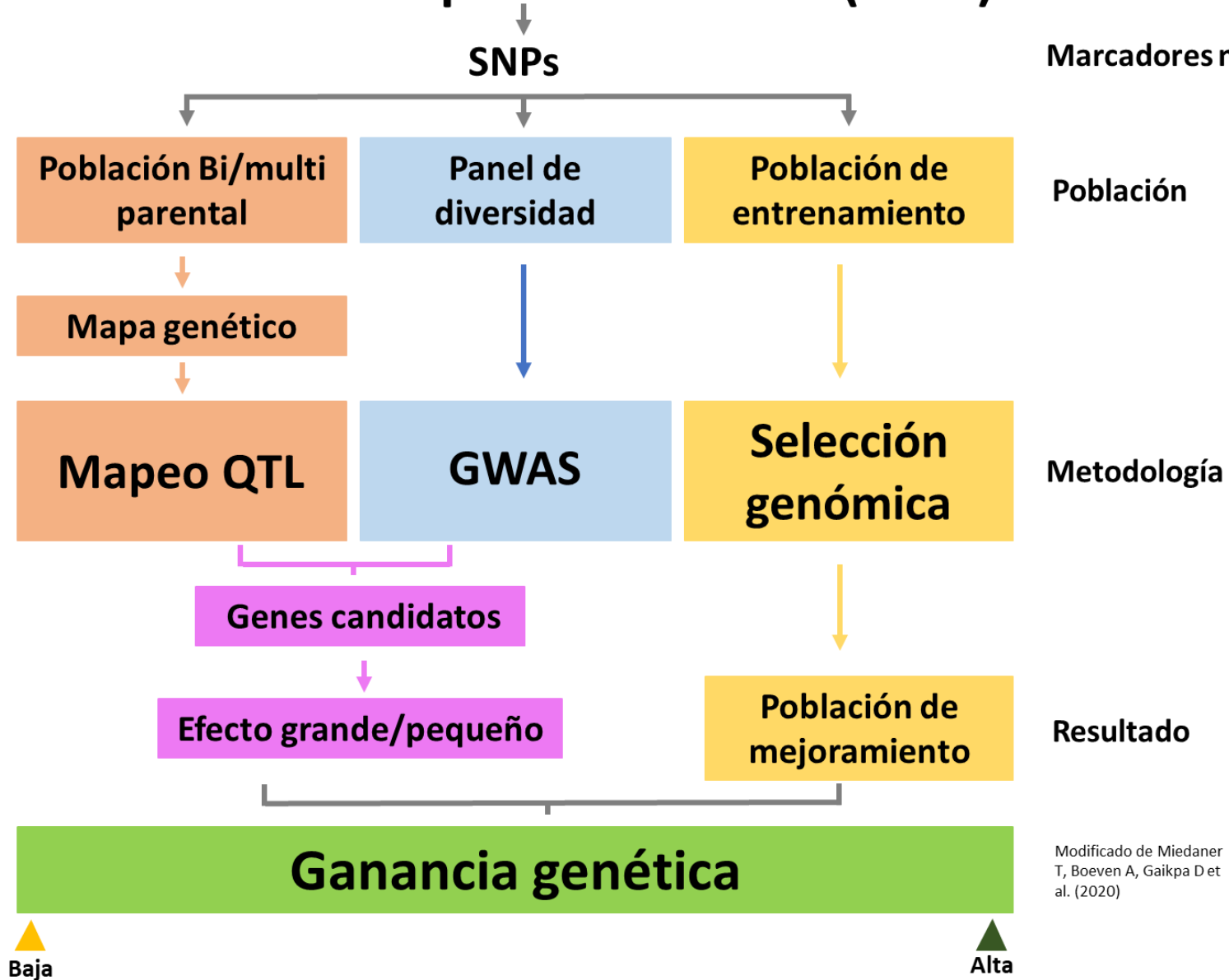


# CICLO DE LA CAÑA DE AZÚCAR



**dds:** días después de siembra

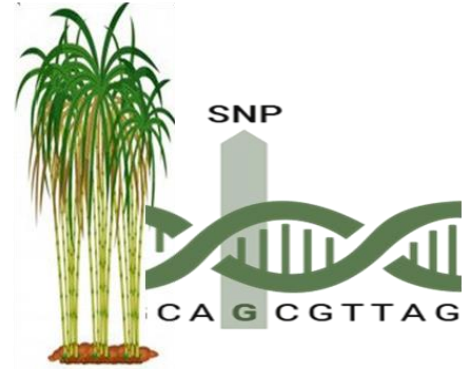
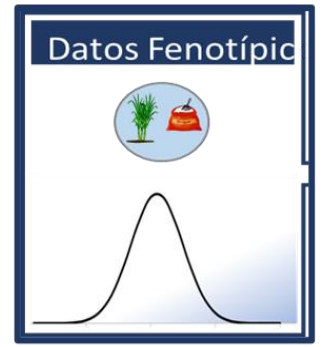
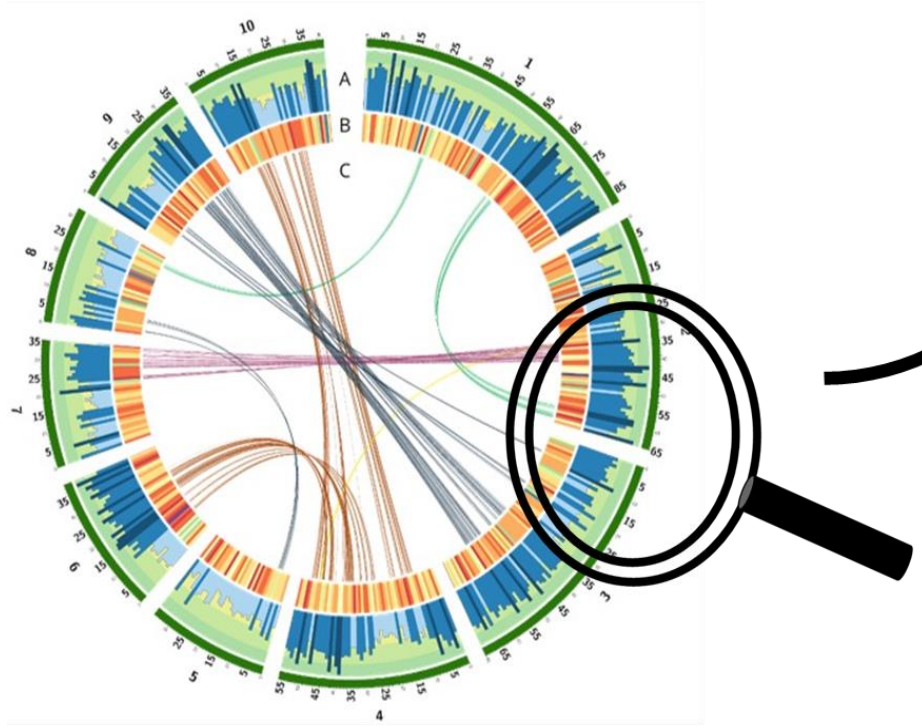
# Selección asistida por marcadores (SAM)

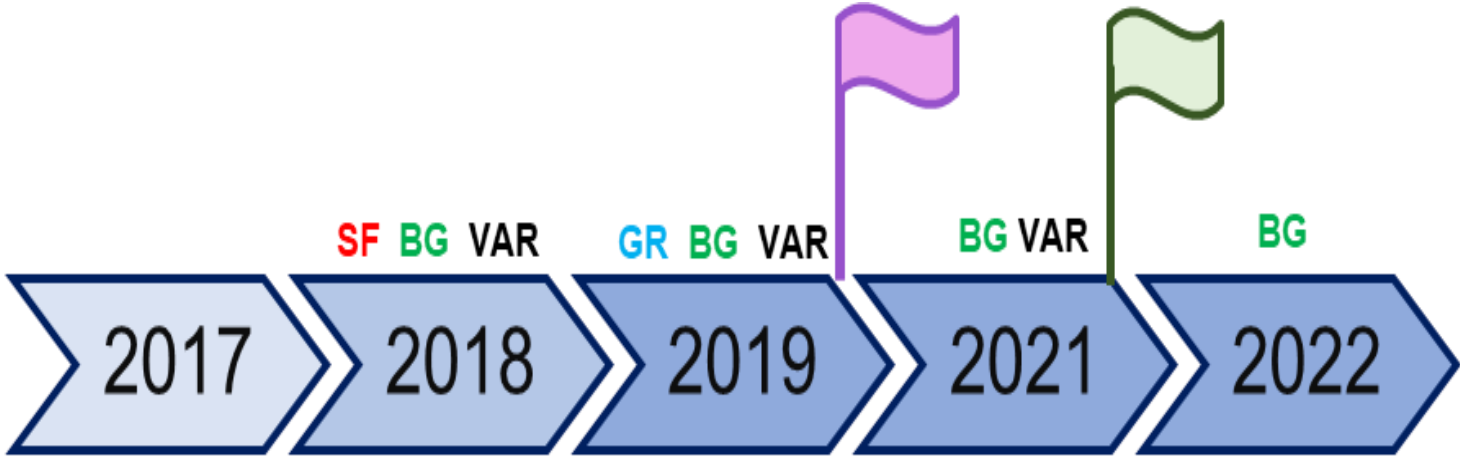
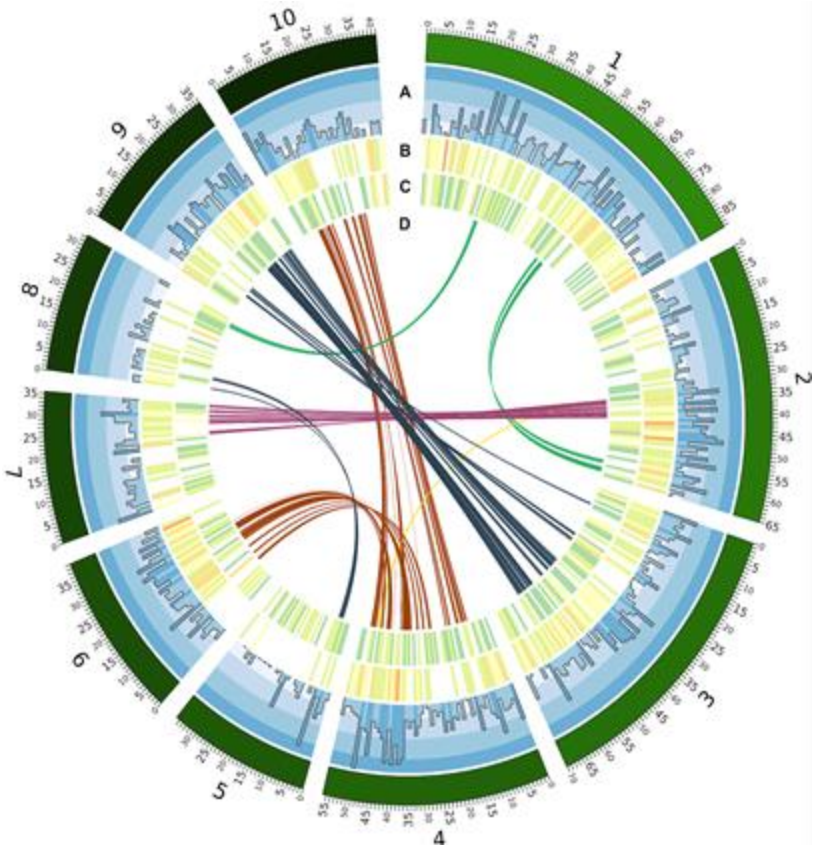


Modificado de Miedaner T, Boeven A, Gaikpa D et al. (2020)

# Estudios de asociación de genoma completo (GWAS)

$$Y = X\beta + Zu + e, \quad \begin{bmatrix} u \\ e \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} K\sigma_a^2 & 0 \\ 0 & \sigma_e^2 \end{bmatrix}\right)$$

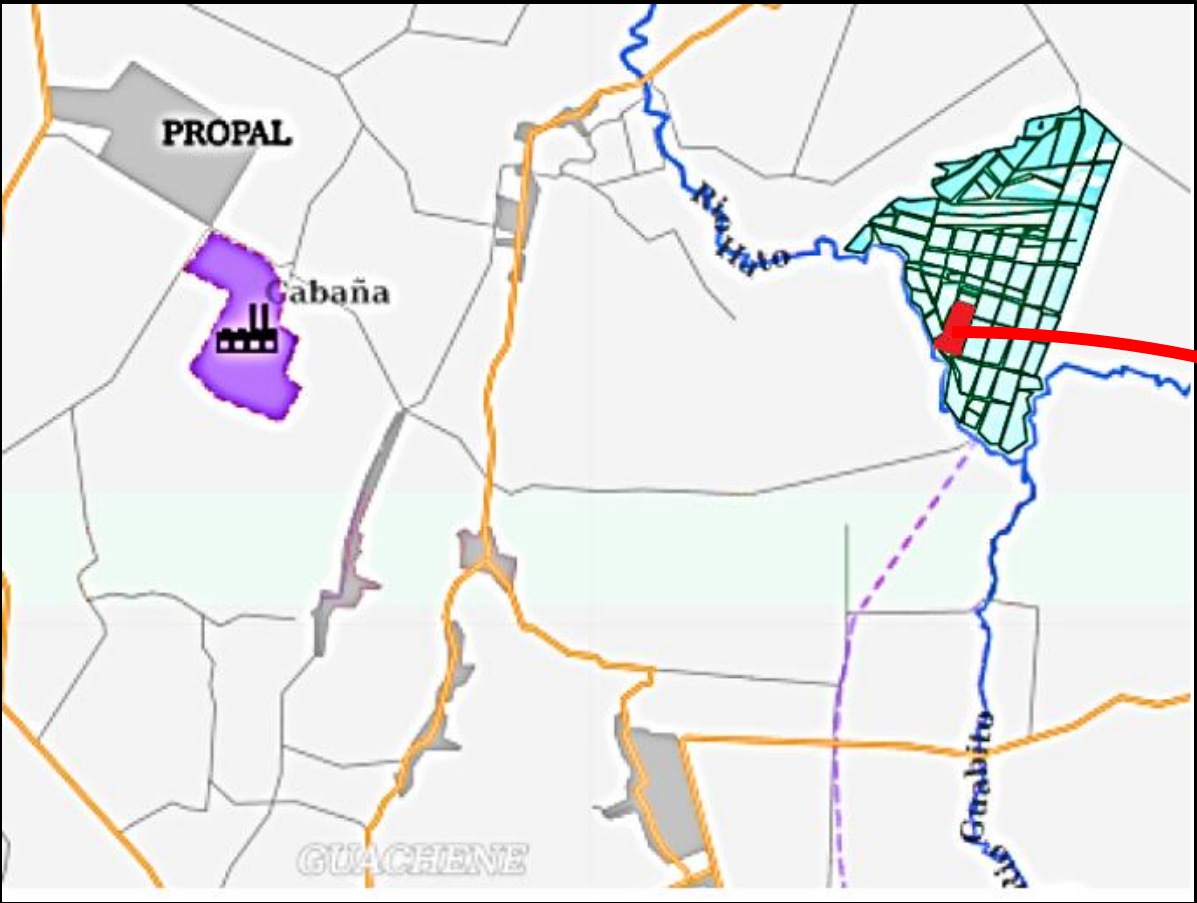
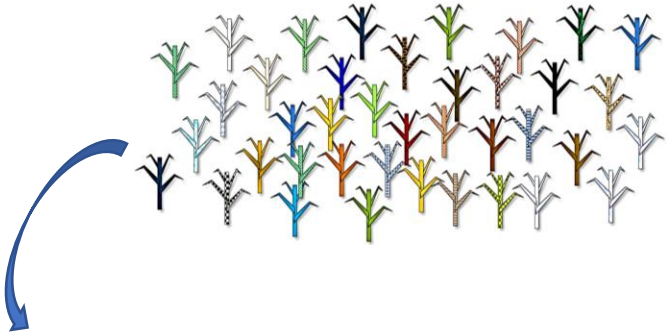




GR: *S. bicolor*  
Software: Gwaspoly, Plink, Tassel  
BG: 19968 SNPs  
Var: Sacarosa (datos históricos)

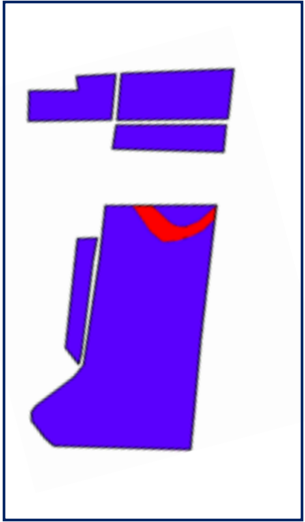
GR: CC 01-1940  
Software: Gwaspoly  
GB: 137889 SNPs  
Var: Sacarosa (%Caña) por mega-ambiente

# LOCALIDAD



Sembrada en el Mega-ambiente húmedo del Valle geográfico del río Cauca, en la hacienda Taula Mejía (Ingenio Cabaña).

Suelo con baja capacidad de drenaje



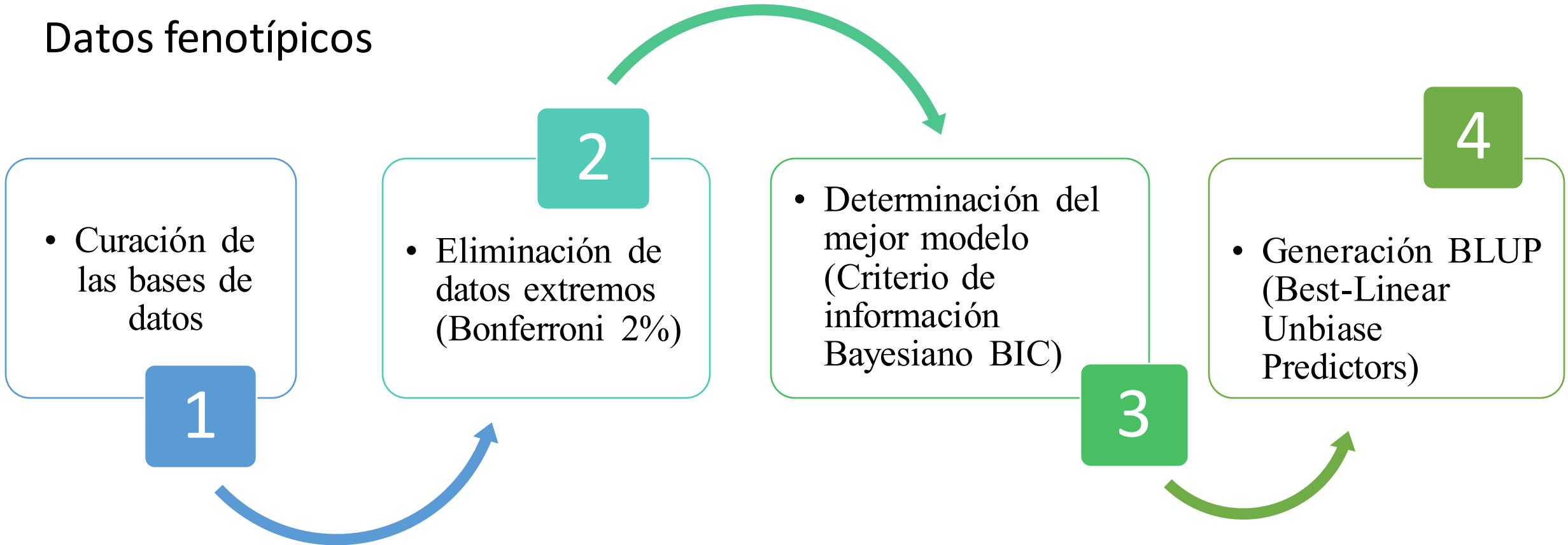
Suerte 77

**Grupo de mejoramiento. Cenicaña**



# ANÁLISIS DE DATOS FENOTÍPICOS

Datos fenotípicos



# OBTENCIÓN DE DATOS GENOTÍPICOS

## Colecta de tejido foliar



3 gr de Hoja TVD-1

## Extracción de ADN



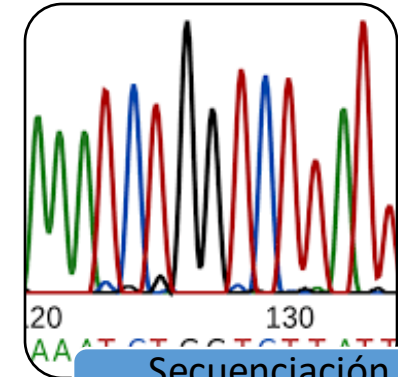
Fenol-cloroformo  
Dellaporta et al.  
1983

## Cuantificación de ADN



Espectrofotómetro  
Nanodrop 2000

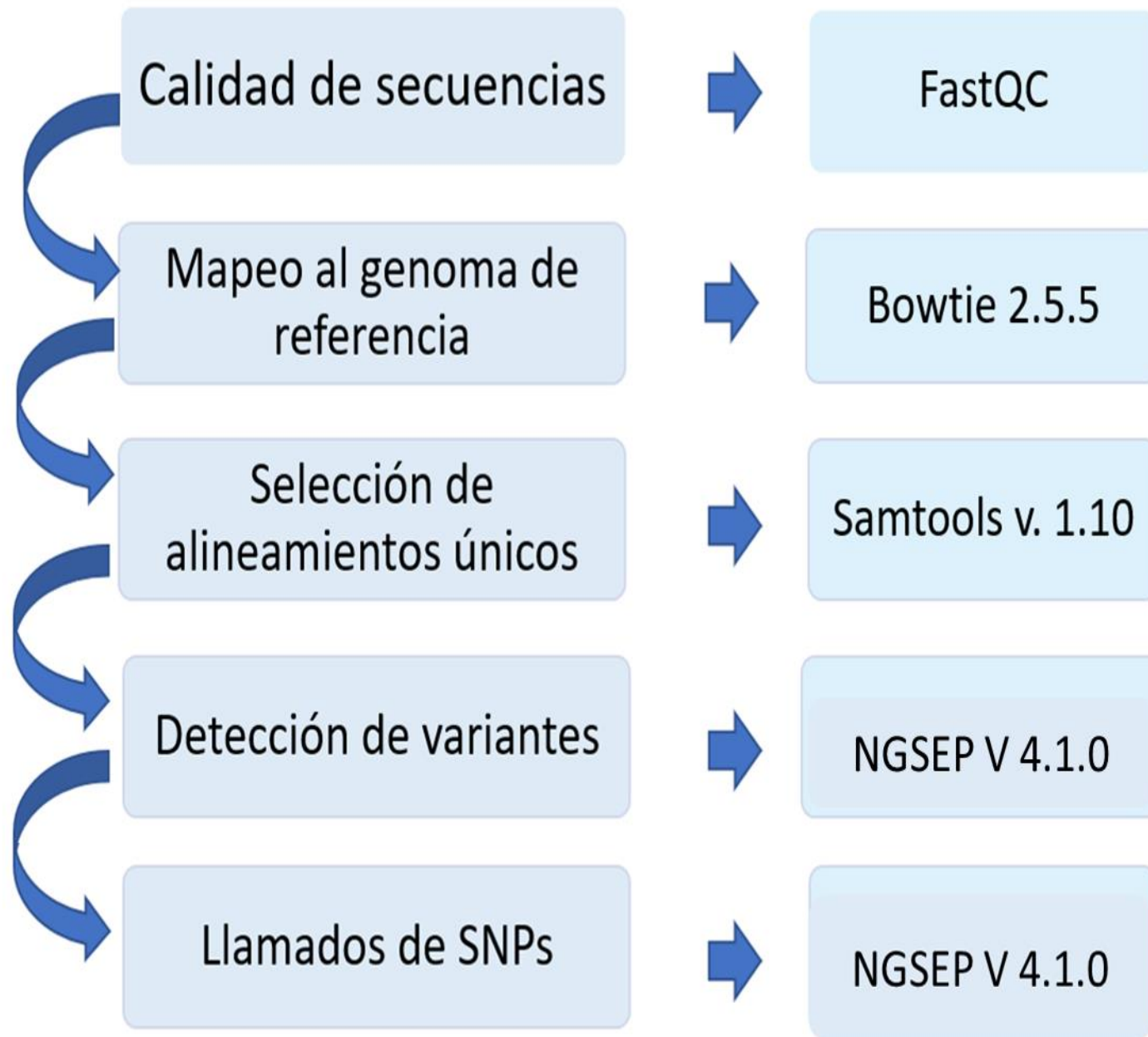
## Secuenciación de ADN



Secuenciación  
completa del  
genoma (WGS)

# Flujo Bioinformático

Datos Genotípicos



Identificación  
de datos  
extremos

Corrección *Bonferroni* 2%  
3 Datos extremos: 10 meses.  
24 Datos extremos: 13 meses.

Selección del  
modelo (BIC)

**Amb:** Plantilla, Soca I y Soca II  
**Rep:** Repetición  
**B:** Bloque  
**Gen:** Genotipo  
**BIC:** Criterio de información Bayesiano

	Amb	Gen	Rep(Amb)	B(Rep Amb)	Gen*Amb	Heterogeneidad en varianza residual	BIC	
							10 meses	13 meses
13 Meses →	X	X	X	X	X	si	5314.8	4300.1*
	X	X	X	X	X	No	5312.5	4445.1
	X	X		X	X	si	5330.1	4301
	X	X		X	X	No	5327.4	4444.7
	X	X	X		X	si	5330.1	4305.7
	X	X	X		X	No	5330.5	4459.2
	X	X			X	si	5401	4319.3
	X	X			X	No	5395.7	4477.6
	X	X	X	X		si	5310.6	4306.8
10 Meses →	X	X	X	X		No	5308.4*	4525.6
	X	X		X		si	5326.1	4307.2
	X	X		X		No	5323.5	4525
	X	X	X			si	5328.3	4312
	X	X	X			No	5325.4	4533.9
	X	X				si	5401	4323.7
	X	X				No	5395.7	4544.7

# Selección de modelo

$$Y_{ijkl} = C_i + R_{j(i)} + B_{k(ij)} + G_l + \varepsilon_{ijkl}$$

$$Y_{ijkl} = C_i + R_{j(i)} + B_{k(ij)} + G_l + I_{il} + \varepsilon_{ijkl}$$

					BIC	
Amb	Gen	Rep(Amb)	B(Rep Amb)	Gen*Amb	10 meses	13 meses
X	X	X	X	X	5314.8	4300.1*
X	X	X	X	X	5312.5	4445.1
X	X	X	X	X	5320.1	4301
X	X	X	X	X	5327.4	4444.7
X	X	X	X	X	5330.1	4305.7
X	X	X	X	X	5330.5	4459.2
X	X	X	X	X	5401	4319.3
X	X	X	X	X	5477.6	4477.6
X	X	X	X	X	5310.6	4306.8
X	X	X	X	X	5308.4*	4525.6
X	X	X	X	X	5326.1	4307.2
X	X	X	X	X	5323.5	4525
X	X	X	X	X	5328.3	4312
X	X	X	X	X	5325.4	4533.9
X	X	X	X	X	5301	4323.7
X	X	X	X	X	5395.7	4544.7



Generación BLUP (Best-Linear Unbiase Predictors)

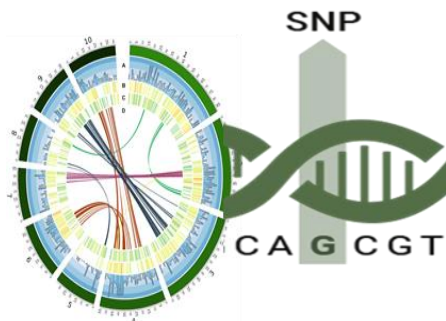
# Análisis de asociación de genoma completo (GWAS)

Datos fenotípicos



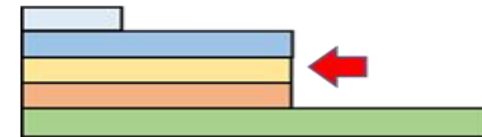
220 genotipos evaluados

Datos genotípicos



137889 marcadores (SNPs)

- Aditivo
- General
- 1 dominante (A>R y R>A)
- 2 dominante (A>R y R>A)
- 3 dominante (A>R y R>A)
- 4 dominante (A>R y R>A)
- 5 dominante (A>R y R>A)



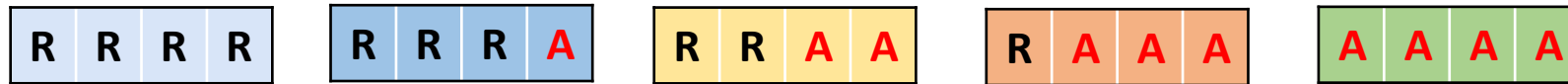
**GWASpoly** 

# Modelos: Análisis de asociación de genoma completo (GWAS)

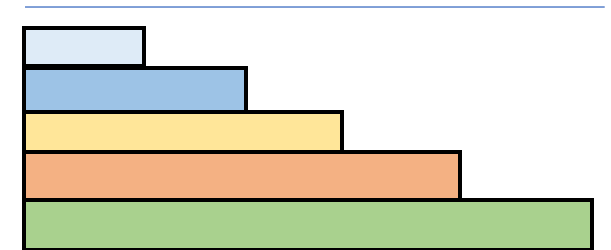
**R** Alelo de referencia

**A** Alelo Alternativo

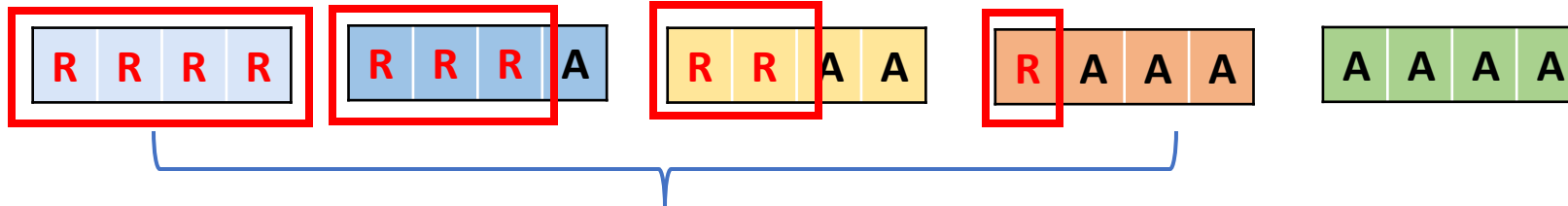
## Aditivo



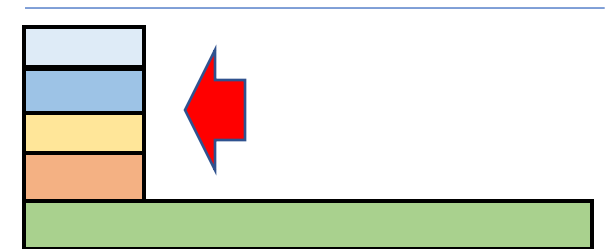
Su efecto es diferente e incrementa al aumentar el número de alelos alternativos



## Simplex dominante (R>A)

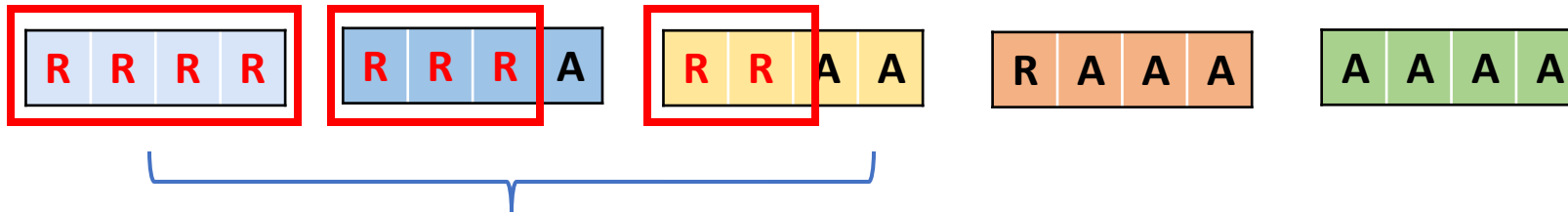


La presencia de al menos **un alelo referencia**, genera que se presente el mismo efecto que el homocigoto



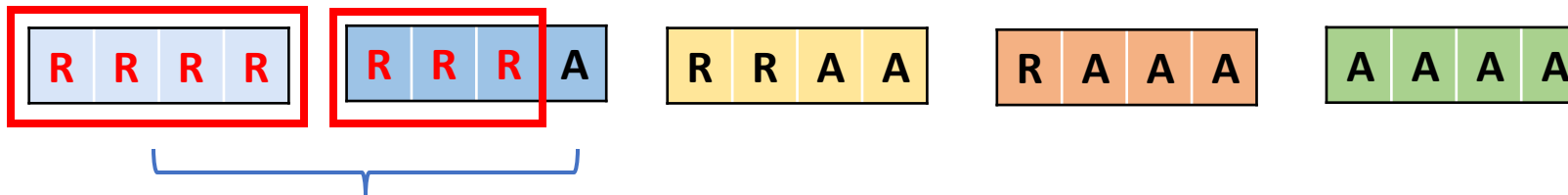
# Modelos: Análisis de asociación de genoma completo (GWAS)

## Duplex dominante (R>A)



La presencia de al menos **dos alelos de referencia**, genera que se presente el mismo efecto que el homocigoto

## Triplex dominante (R>A)

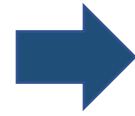


La presencia de al menos **tres alelos de referencia**, genera que se presente el mismo efecto que el homocigoto



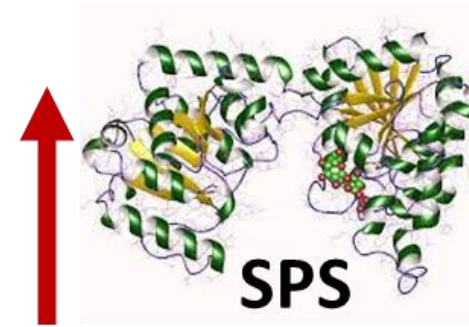
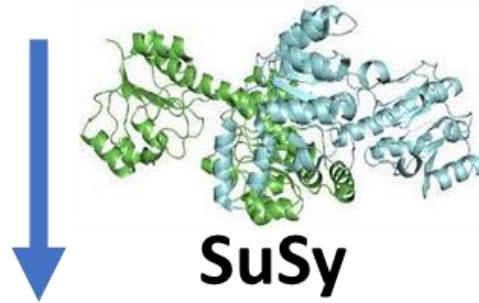
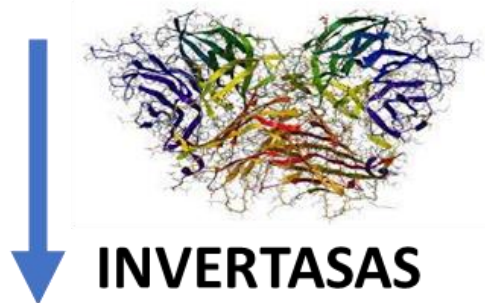
# Análisis fenotípico

Acumulación a los 10 meses fue **1.73%** menor a la acumulación a los 13 meses.



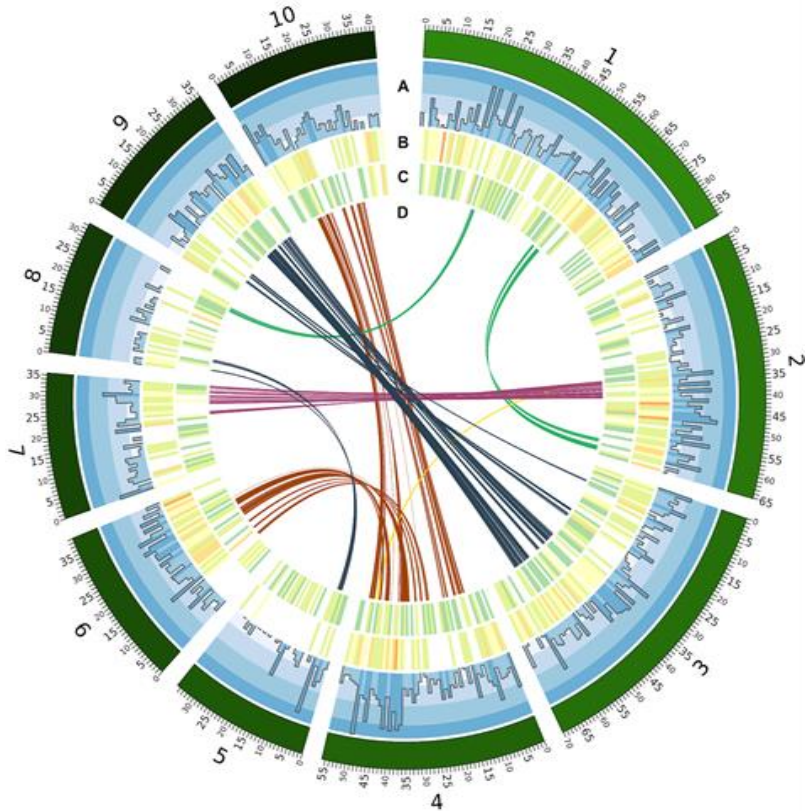
Etiopia (Urgesa & Keyata, 2021)  
Egipto (Ahmed & Awadalla, 2016)  
Brasil (Marin et al., 2021)

**A mayor edad,  
mayor contenido  
de sacarosa**



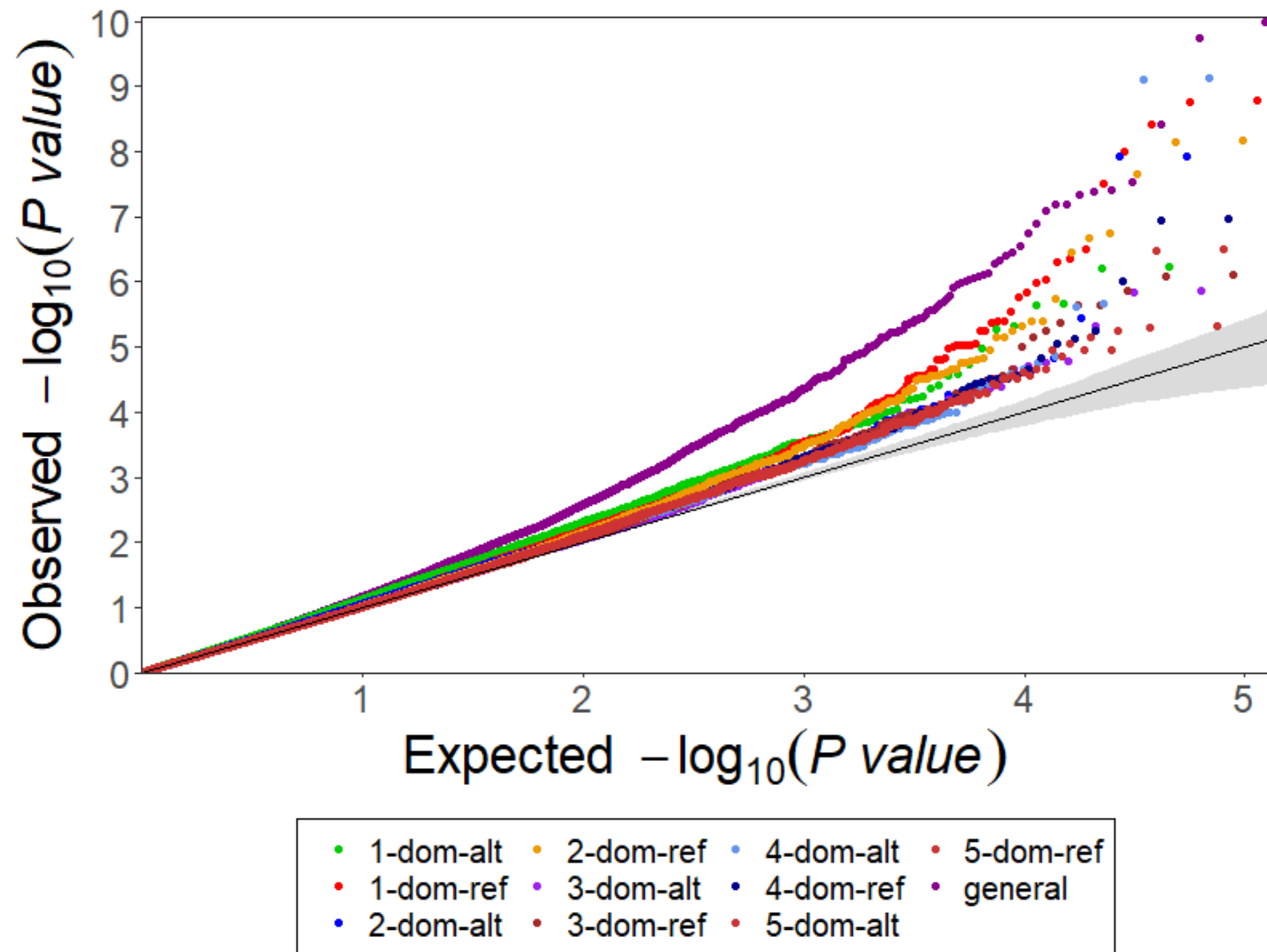
**Cambios  
fisiológicos**

# Análisis de asociación



Modelo	10 meses	13 Meses
additive	0	0
general	215	137
1-dom-alt	2	5
1-dom-ref	4	24
2-dom-alt	0	1
2-dom-ref	11	8
3-dom-alt	0	0
3-dom-ref	0	4
4-dom-alt	2	1
4-dom-ref	1	2
5-dom-alt	2	1
5-dom-ref	0	0
<b>Total</b>	<b>237</b>	<b>183</b>

**420 marcadores**



**Presencia de falsos positivos**



Filtros:

Evaluación de los marcadores de acuerdo con el valor Score.

Evaluación de marcadores utilizando  $R^2$

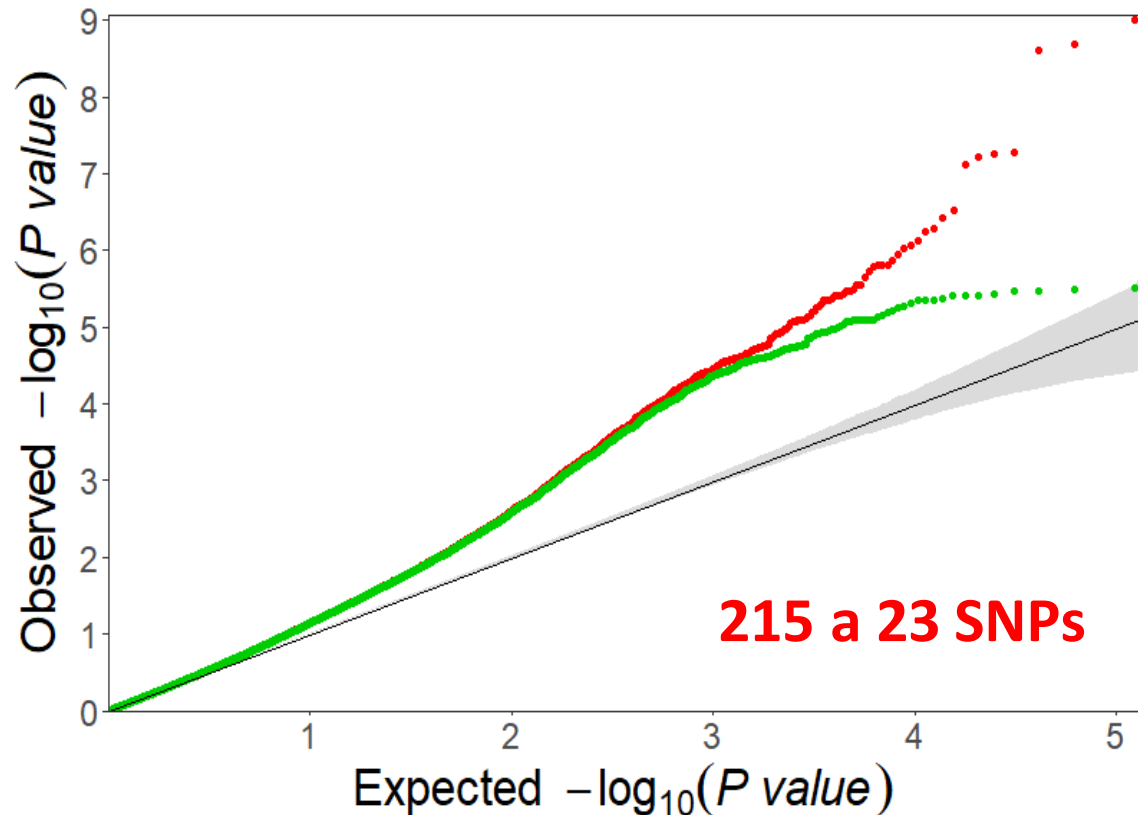
# Análisis de asociación

Característica	Modelo	Marcador	Cromosoma	Posición	Ref	Alt	Score
meses_13	general	contig_36908_12883	contig_36908	12883	C	T	9.75
meses_13	4-dom-alt	3_57539943	3	57539943	A	G	9.12
meses_13	1-dom-ref	4_55115204	4	55115204	C	G	8.76
meses_13	1-dom-ref	contig_50499_5072	contig_50499	5072	C	T	8.43
meses_13	general	contig_44982_8900	contig_44982	8900	C	T	8.41
meses_13	2-dom-ref	2_28319736	2	28319736	C	T	8.14
meses_13	1-dom-ref	8_13879677	8	13879677	A	G	8
meses_13	2-dom-alt	6_38041421	6	38041421	C	T	7.92
meses_13	2-dom-ref	contig_40813_18644	contig_40813	18644	A	G	7.65
meses_13	general	1_86961880	1	86961880	A	G	7.55
meses_13	1-dom-ref	3_1957031	3	1957031	A	G	7.51
meses_13	general	9_33845426	9	33845426	C	T	7.42
meses_13	general	contig_36570_19433	contig_36570	19433	C	T	7.38
meses_13	general	1_112769	1	112769	C	G	7.34
meses_13	general	8_13879675	8	13879675	C	T	7.2
meses_13	general	1_61256982	1	61256982	C	T	7.18
meses_13	general	4_55115204	4	55115204	C	G	7.09
meses_13	4-dom-ref	contig_39315_31775	contig_39315	31775	A	G	6.95

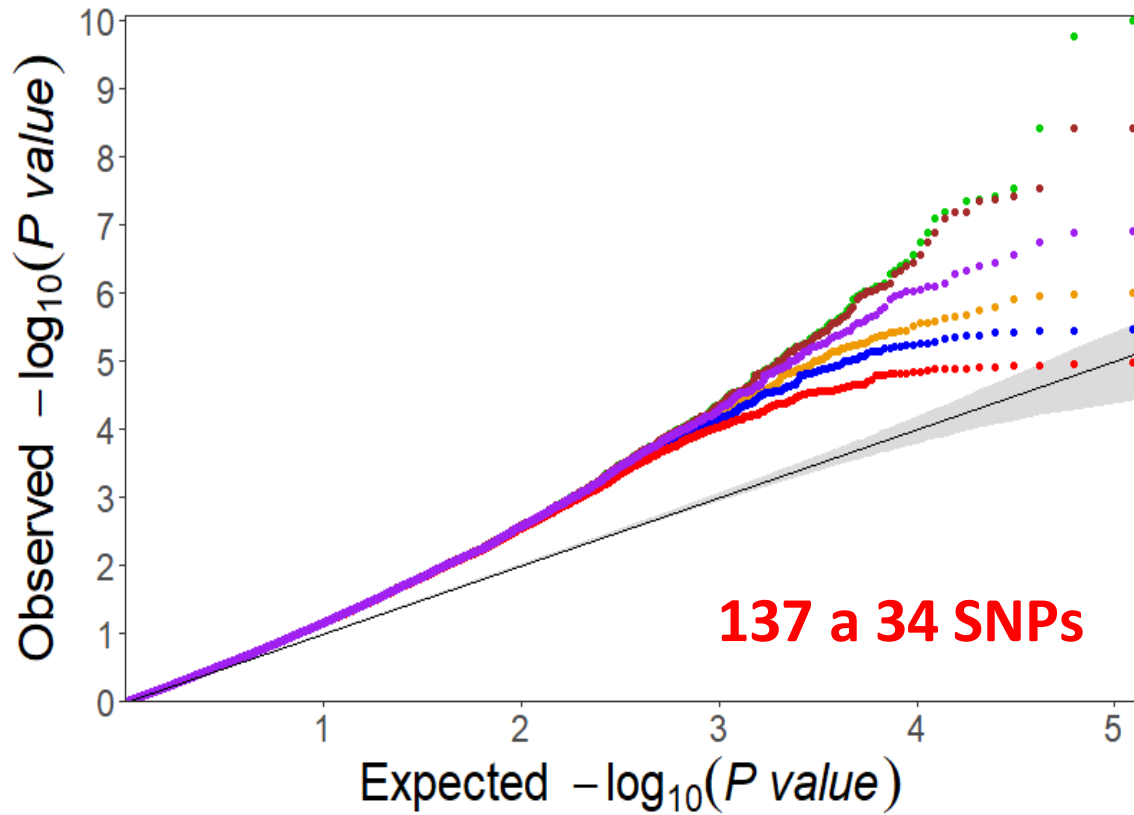


# Marcadores asociados: Modelo general

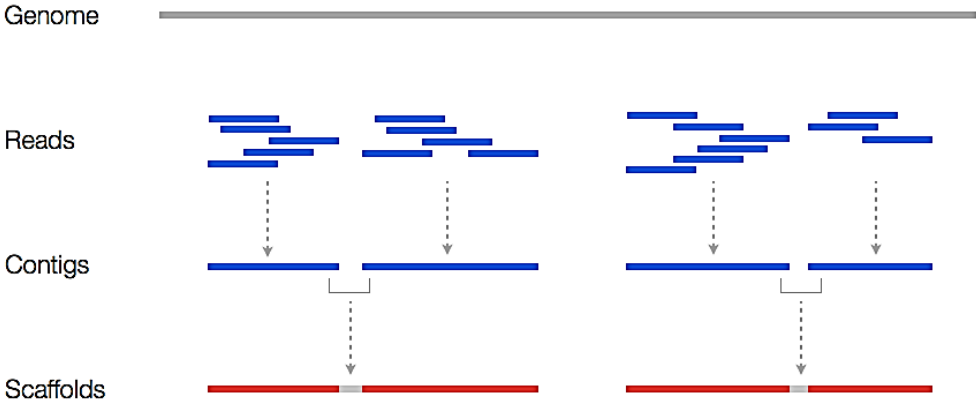
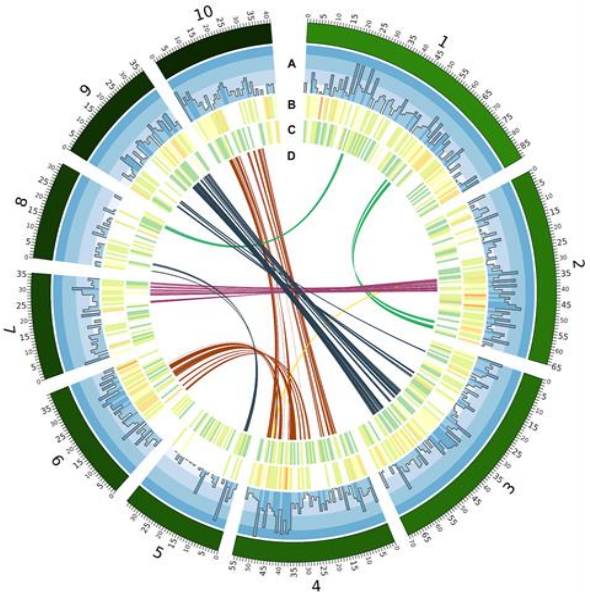
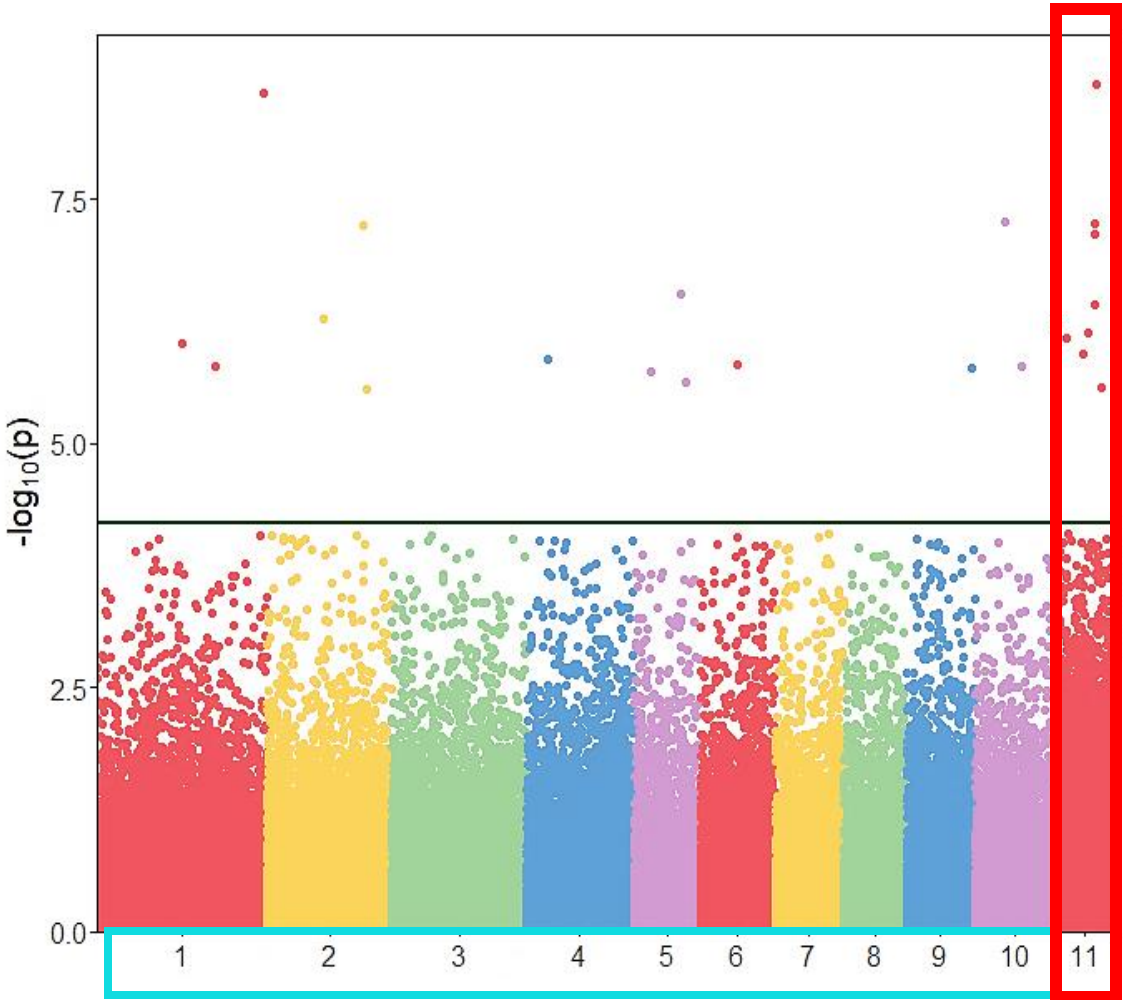
10 Meses



13 Meses

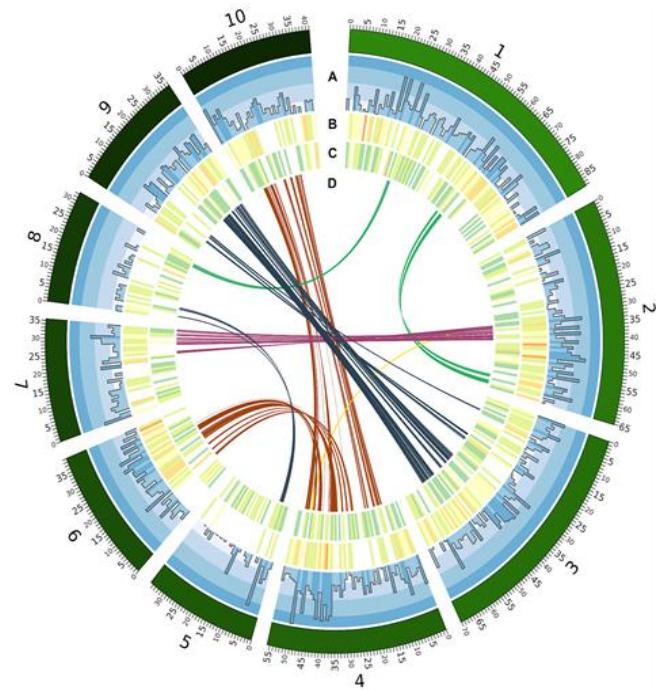


# Marcadores asociados: Modelo general





# Análisis de asociación



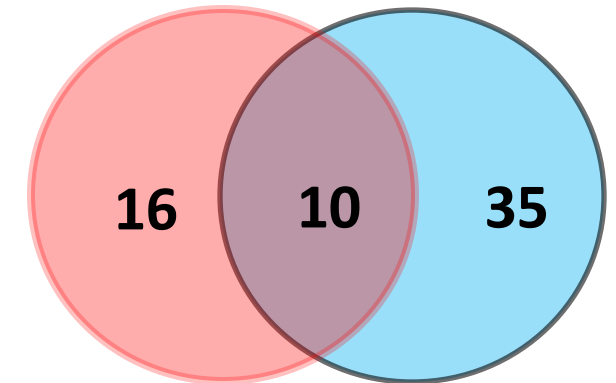
Modelo	10 meses	13 Meses
additive	0	0
general	23	34
1-dom-alt	2	5
1-dom-ref	4	12
2-dom-alt	0	1
2-dom-ref	6	8
3-dom-alt	0	0
3-dom-ref	0	4
4-dom-alt	2	1
4-dom-ref	1	2
5-dom-alt	2	1
5-dom-ref	0	0
<b>Total</b>	<b>56</b>	<b>68</b>

**Filtrados por QQplot**

Marcadores filtrado

$R^2 \geq 1\%$

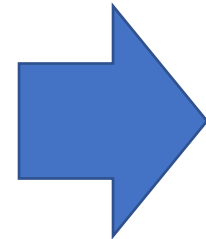
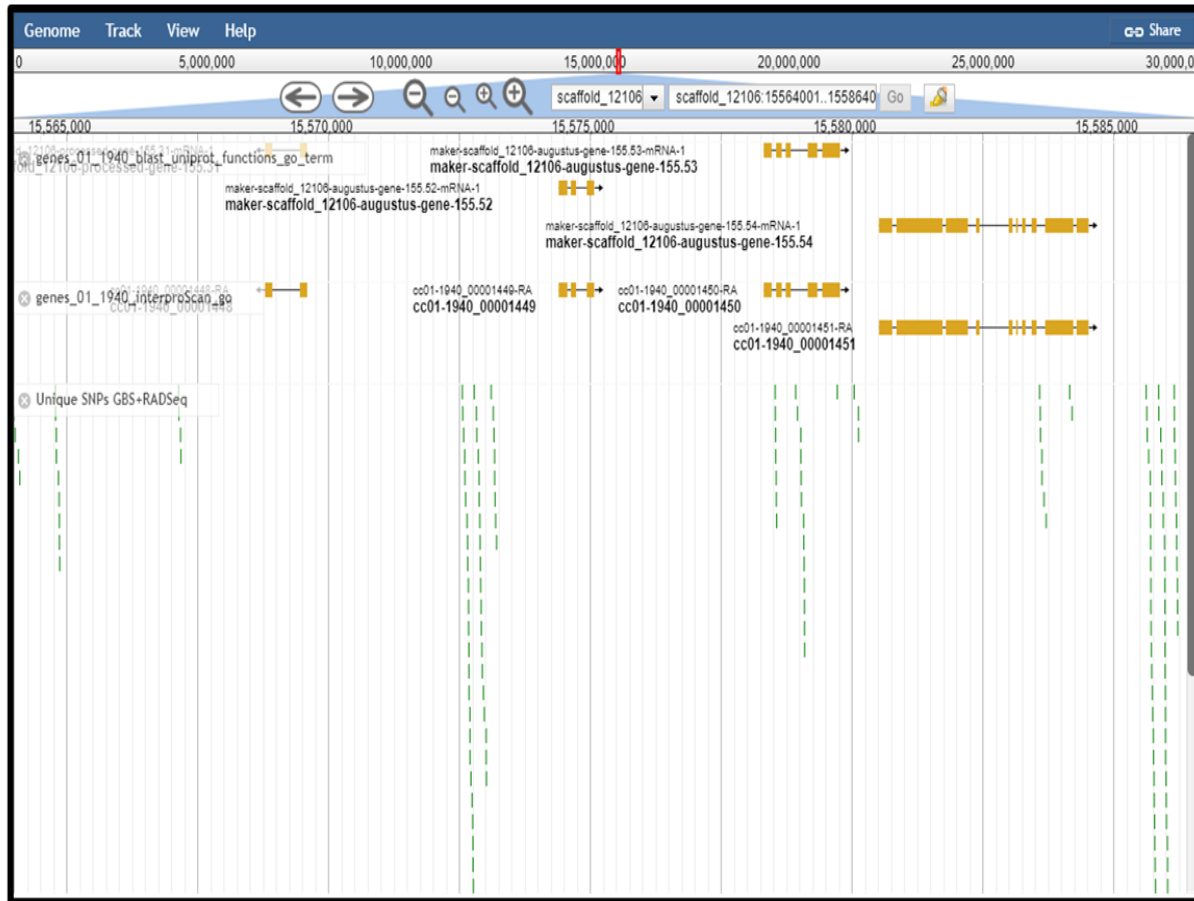
10 meses 13 meses



**61 marcadores**



# Genes candidatos



Sacarosa

4234 genes

Total

2647 genes

Con función

1406 genes

Plantas

62 genes  
Candidatos

- ❖ Transpotadores
- ❖ Glucosidasas
- ❖ glycosyltransferase
- ❖ UDP-glycosyltransferase
- ❖ transcription factor
- ❖ Auxina
- ❖ Etileno
- ❖ glucosyltransferase
- ❖ trehalose
- ❖ Translocador
- ❖ Hexoquinasa



- **PLT5**: Polyol transporter 5 (*A. thaliana*)

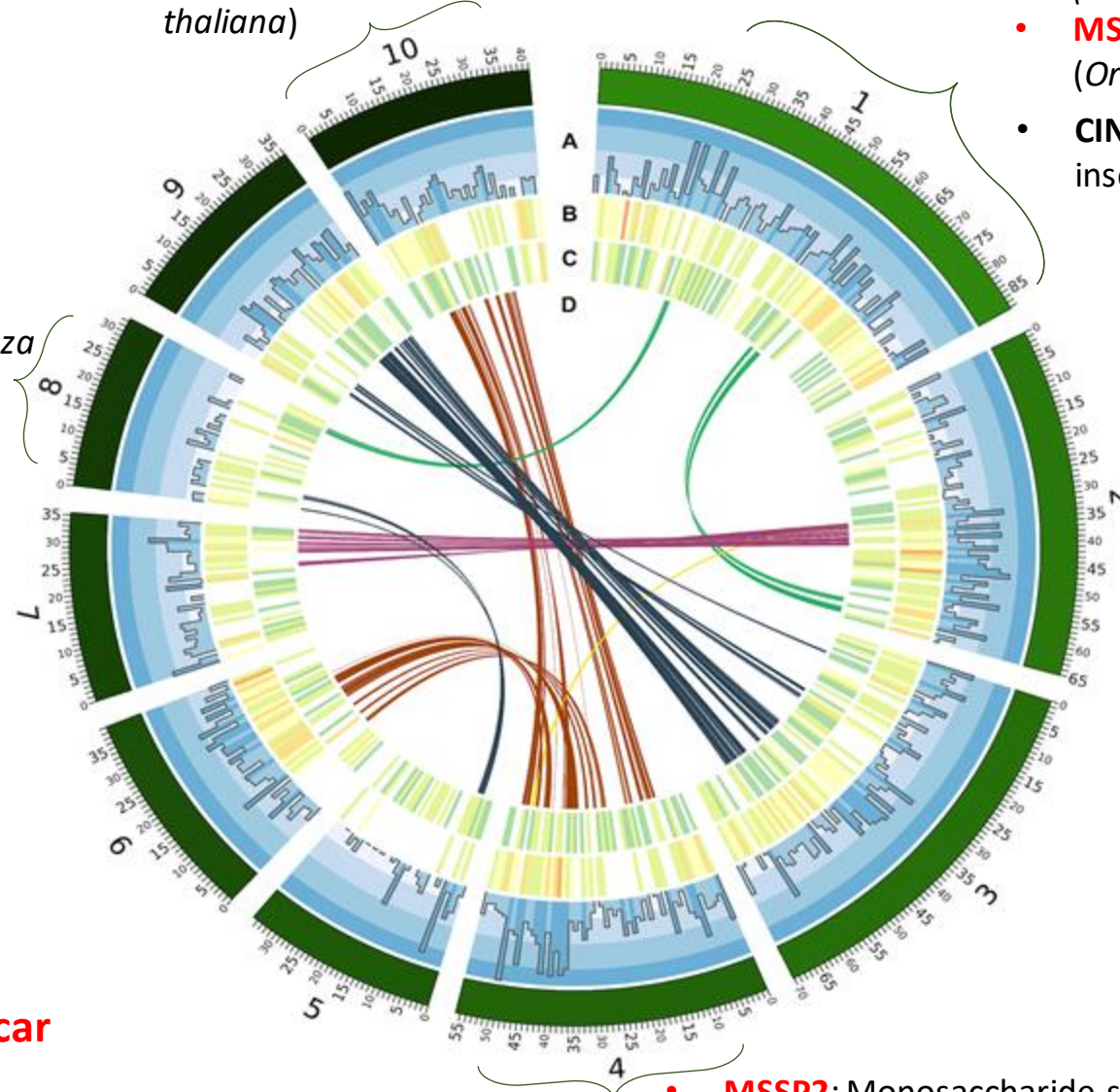
- **MST2**: Sugar transport protein MST2 (*Oryza sativa subsp. japonica*)
- **MST1**: Sugar transport protein MST1 (*Oryza sativa subsp. japonica*)
- **CIN3**: Beta-fructofuranosidase 2C insoluble isoenzyme 3

- **HXK3**: Hexokinase-3 (*Oryza sativa subsp. japonica*)

- Probable plastidic **glucose transporter 1** (*A. thaliana*)
- **CSTLP2**: CMP-sialic acid transporter 2 (*Oryza sativa subsp. japonica*)

- **Transportador de azúcar**
- **Invertasa**
- **Hexoquinasa**

- **MSSP2**: Monosaccharide-sensing protein 2 (*A. thaliana*)
- **SUT4**: Sucrose transport protein SUT4 (*Oryza sativa subsp. japonica*)
- **SUS6**: Sucrose synthase 6 (*Oryza sativa subsp. japonica*)



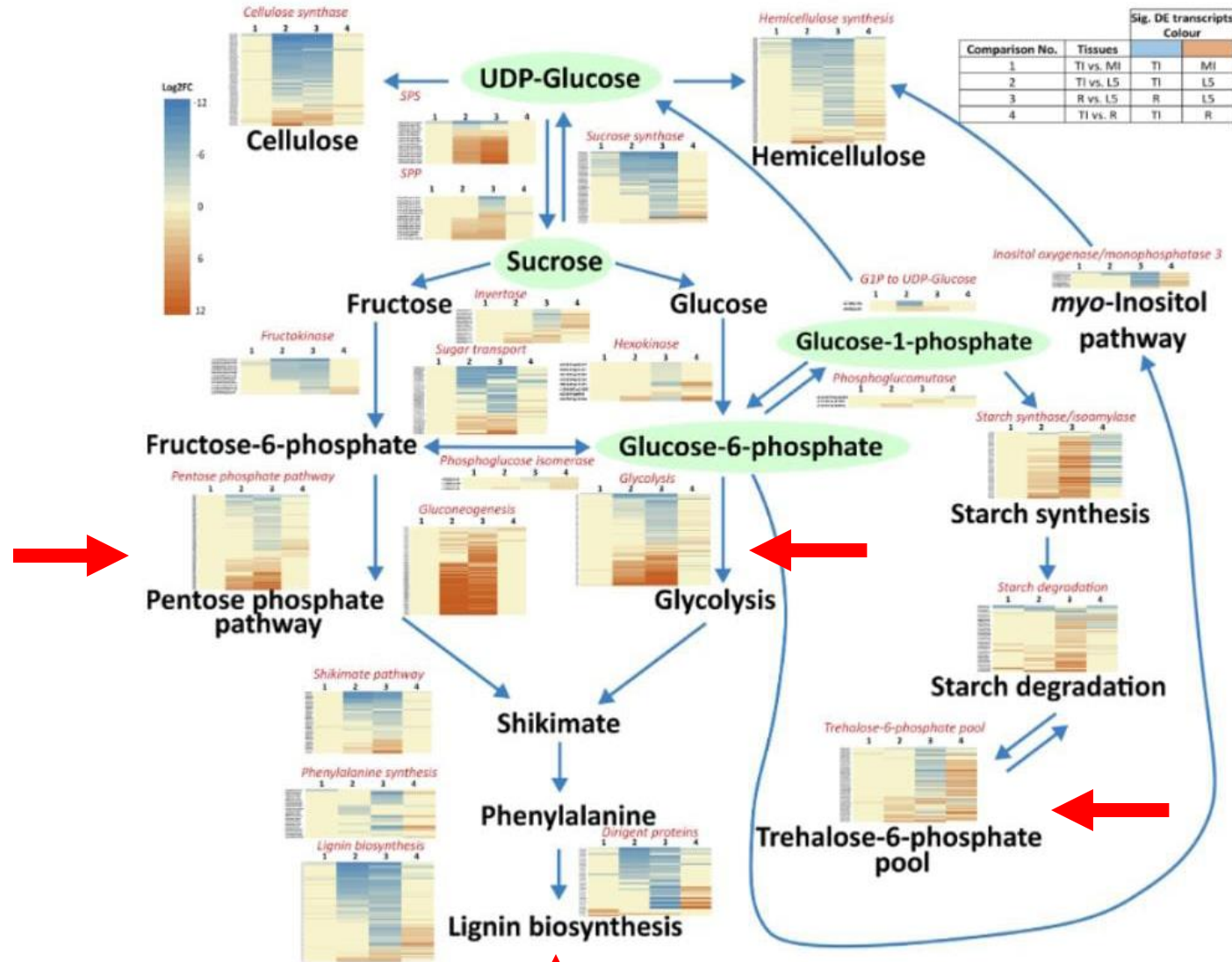
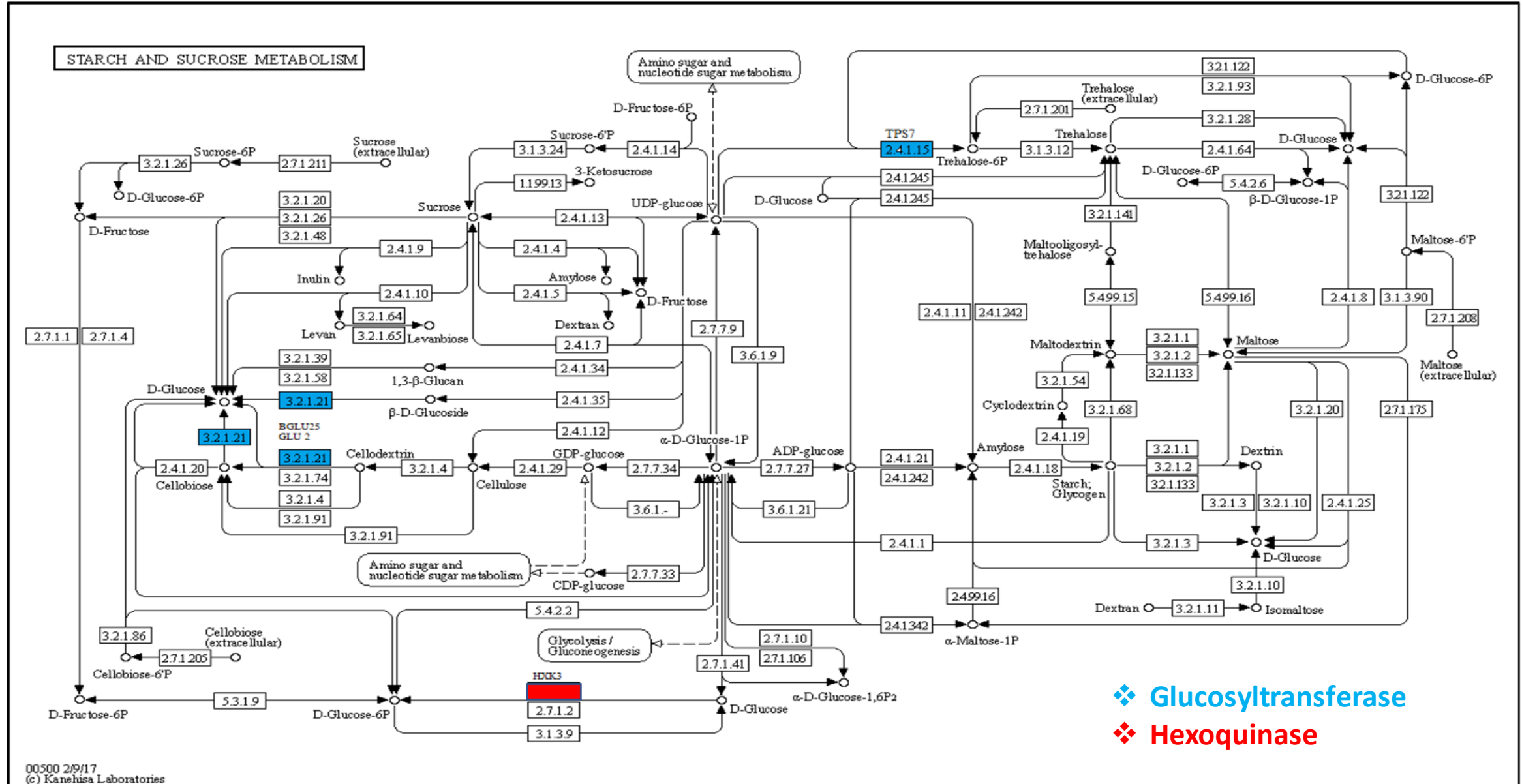
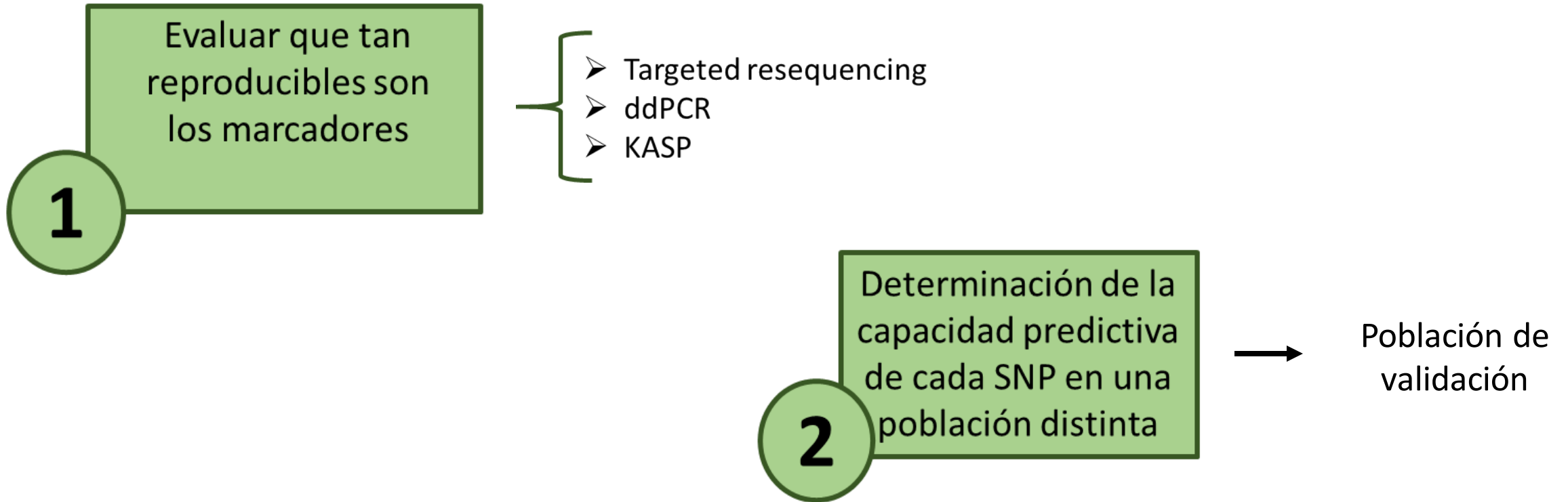


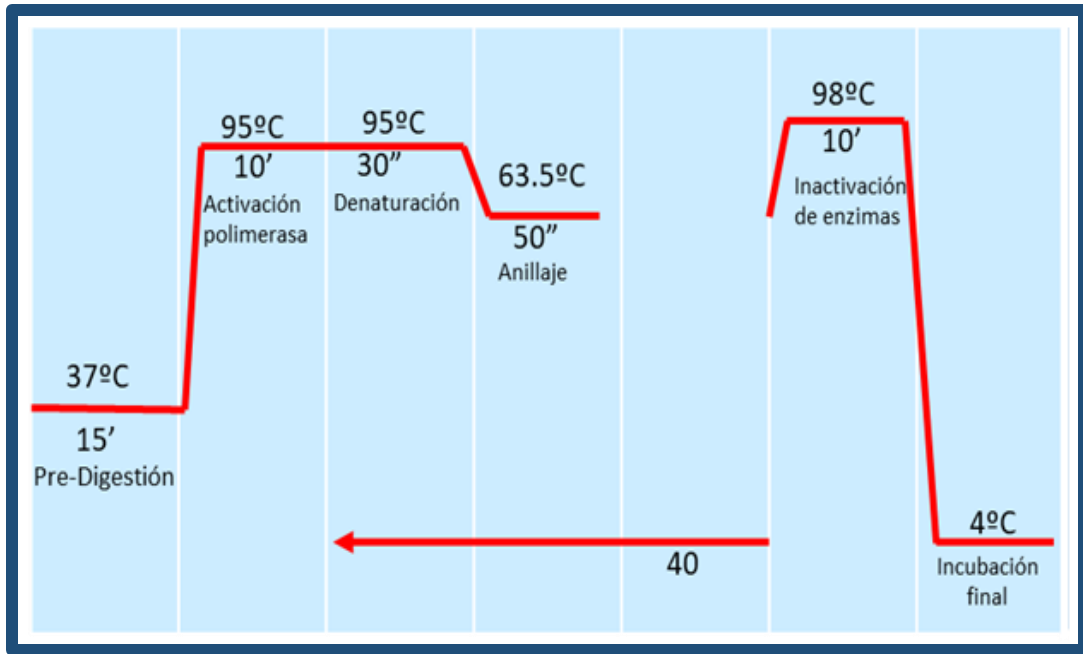
Fig. 5. Heatmaps of differentially expressed genes within the major pathways of carbon partitioning. Blue or red colour indicates transcripts with significantly higher expression within the tissue corresponding to the colour (colour legend shown within figure). Only transcripts that were significantly differentially expressed (FDR p-value  $\leq 0.01$ ) are plotted. The plotted  $\log_2FC$  values were derived from DGE analysis in CLC. Abbreviations, TI: top internode; MI: middle internode; 5th visible dewlap leaf: L5; R: root. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

# Genes candidatos

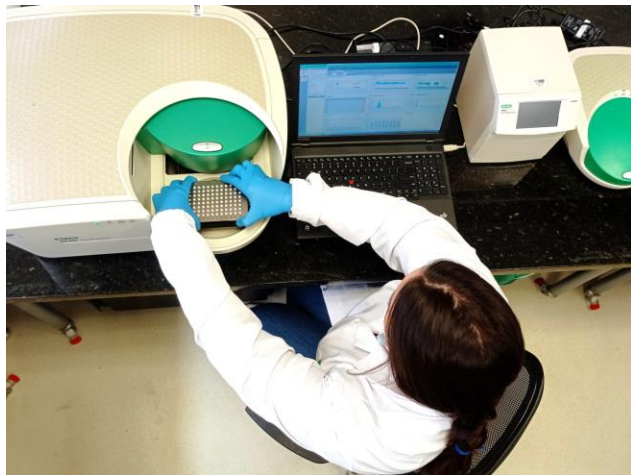


# Validación de marcadores

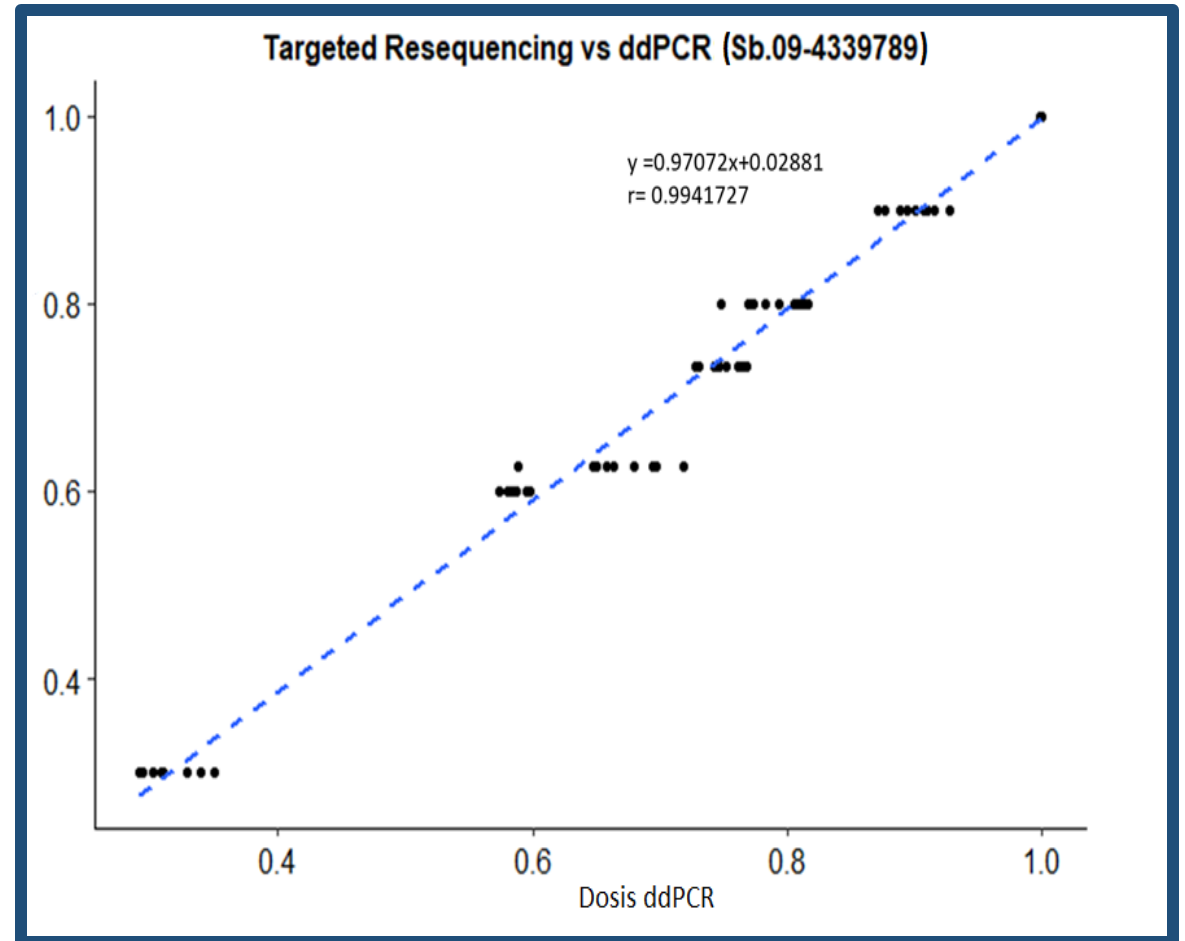




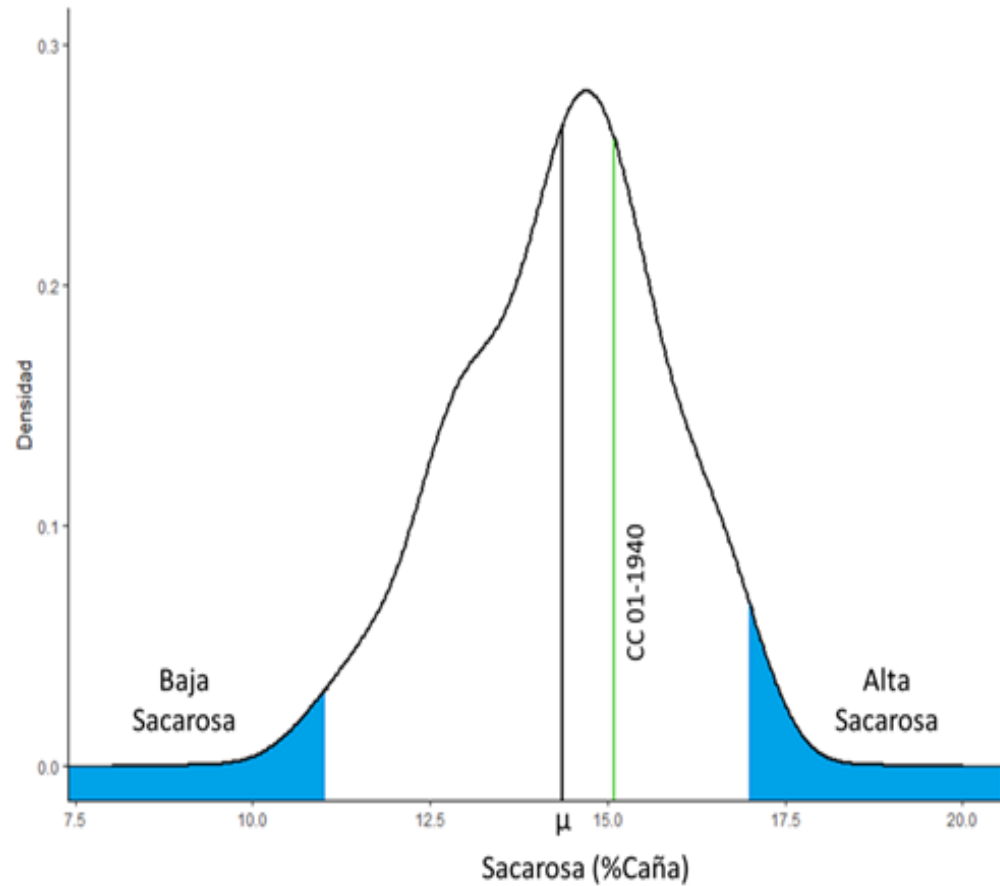
## Estandarización de la técnica ddPCR



## Correlación Targeted Resequencing y ddPCR



# POBLACIÓN DE VALIDACIÓN



Semillero Estado II 2021



Siembra: 13 Octubre 2021  
150 genotipos de alta y baja sacarosa.

## Evaluación agronómica

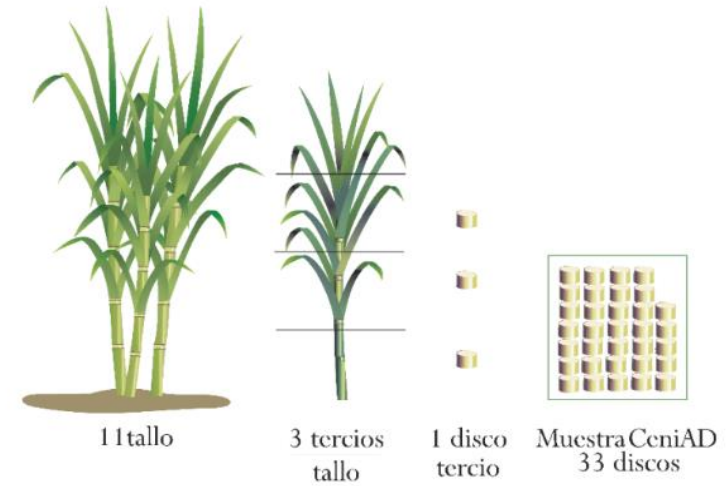
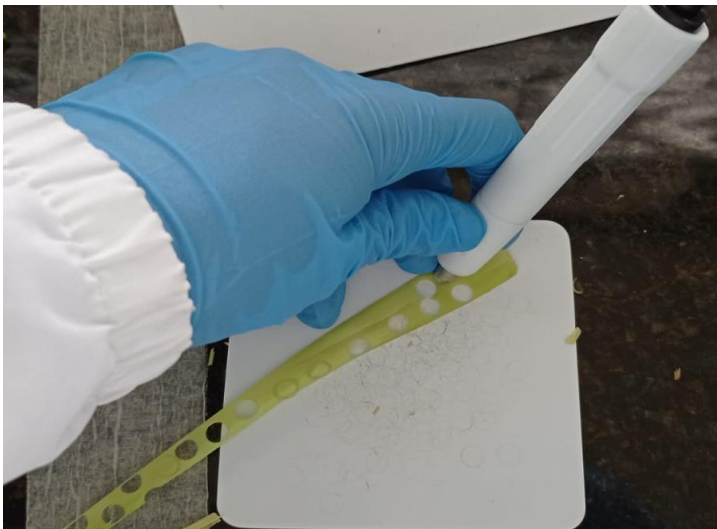


## Evaluación de sacarosa



**NIR portable**

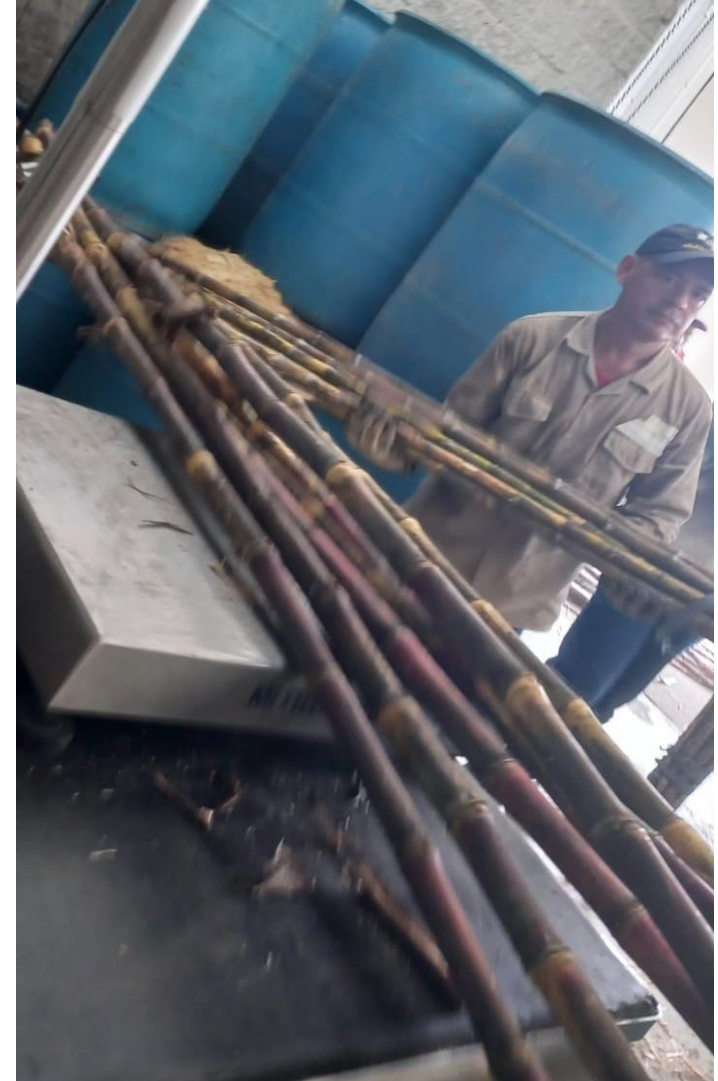
## Colecta y extracción de ADN



Fuente: Tomado de Amaya, Larrahondo, Rangel, Moreno, & Viveros (2001)



## Evaluación de sacarosa





Gracias



