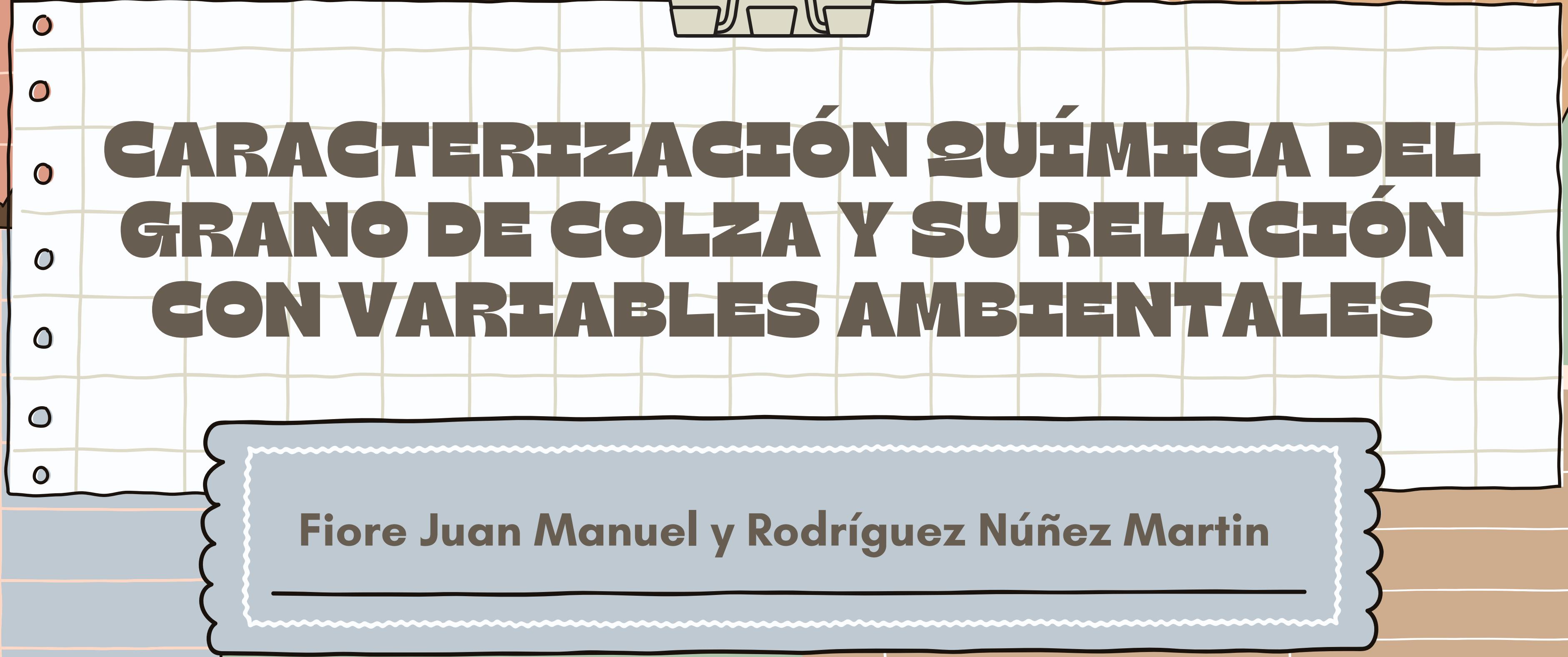




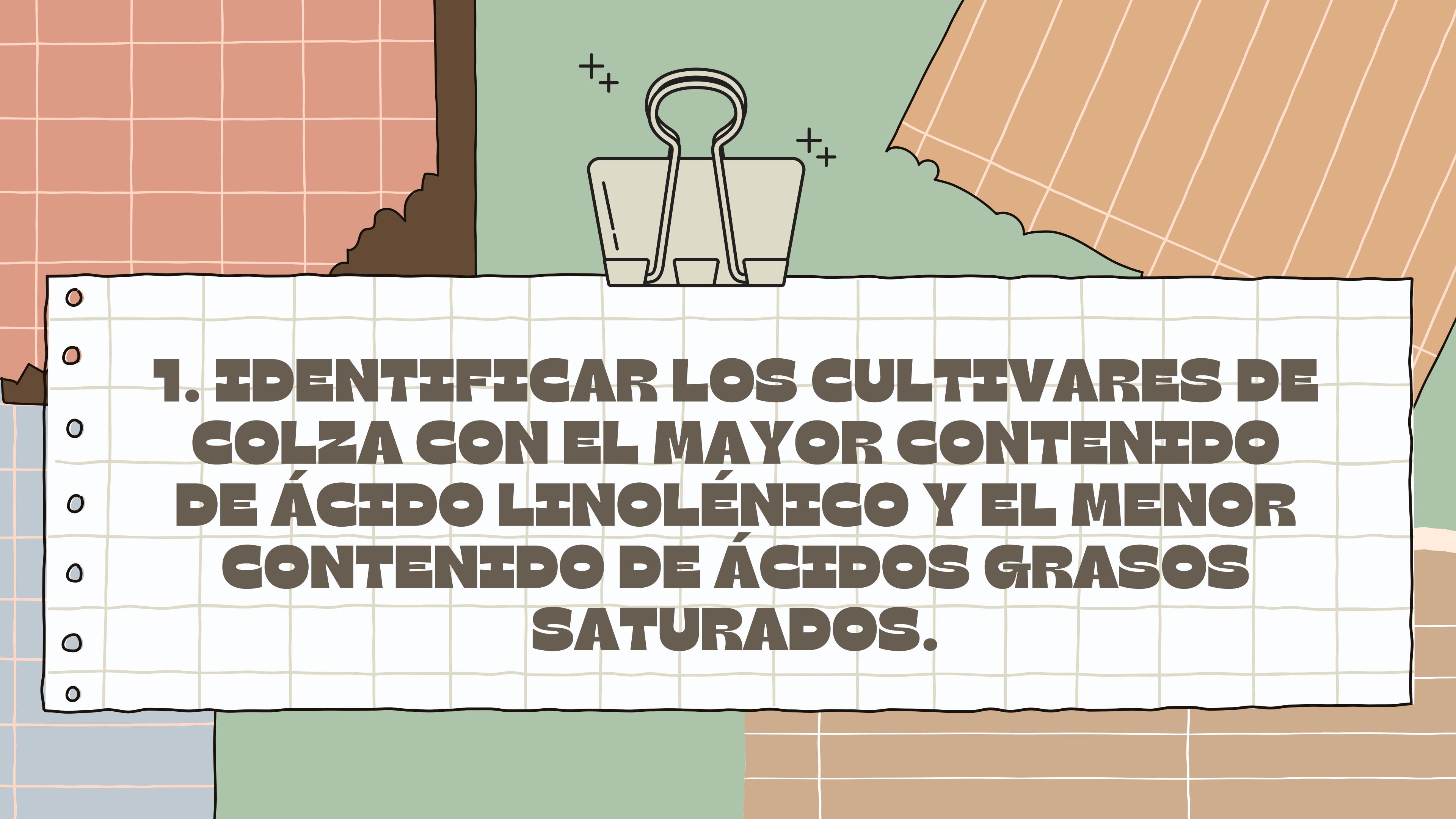
CARACTERIZACIÓN QUÍMICA DEL GRANO DE COLZA Y SU RELACIÓN CON VARIABLES AMBIENTALES



Fiore Juan Manuel y Rodríguez Núñez Martín

OBJETIVOS GENERALES

- Analizar la composición del aceite de colza, evaluando la influencia de factores genéticos y ambientales.
- Realizar recomendaciones para la selección de cultivares en diferentes ambientes.



1. IDENTIFICAR LOS CULTIVARES DE COLZA CON EL MAYOR CONTENIDO DE ÁCIDO LINOLÉNICO Y EL MENOR CONTENIDO DE ÁCIDOS GRASOS SATURADOS.

METODOLOGÍA

ANOVA basado en MLM

Ajuste iterativo

Factores

Efectos

Minimización simultánea de los criterios AIC y BIC

- Cultivar
- Repetición
- Localidad
- Interacción

- Fijos
- Aleatorios
- Función de varianza

Análisis a posteriori

DGC

ABORDAJE

ANOVA basado en MLM

Capacidad controlar la variabilidad que no es considerada en los modelos lineales (OLS).

Control de la homogeneidad de varianza

Buscamos diferencias estadísticamente significativas

Función de varianza

Prueba de hipótesis

$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D$$
$$H_A : \mu_i \neq \mu_j$$

Análisis a posteriori

DGC

Grupos mutuamente excluyentes

ÁCIDO GRASO LINOLÉNICO

Modelo final:

- Efecto fijo: cultivar + repetición
- Efecto aleatorio: localidad
- Función de varianza: cultivar

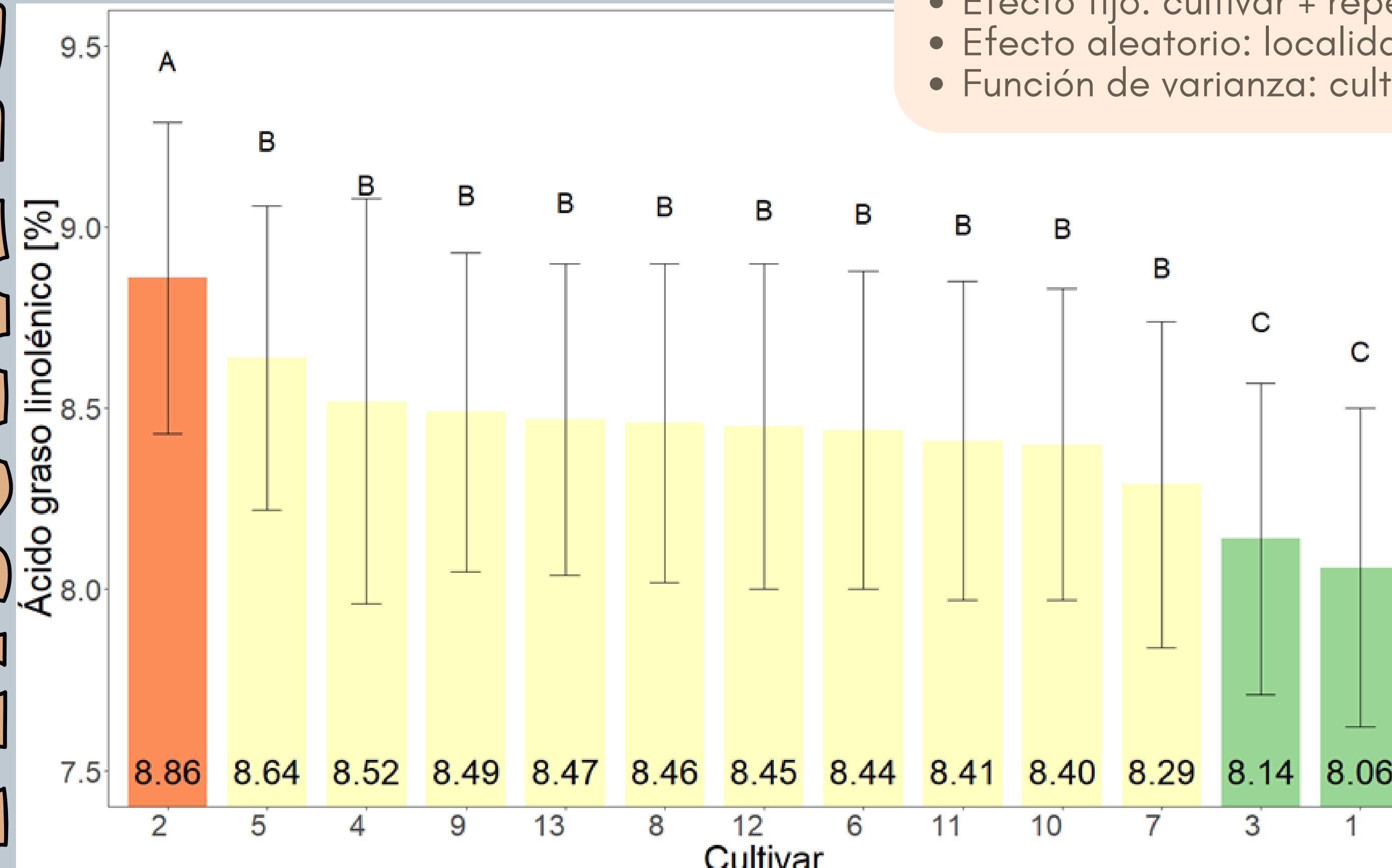


Figura 1: Contenido de ácido graso linolénico en los distintos cultivares de colza evaluados en el estudio y resultados de la prueba a posteriori de DGC.

ÁCIDOS GRASOS SATURADOS

Modelo final:

- Efecto fijo: cultivar + repetición
- Efecto aleatorio: localidad

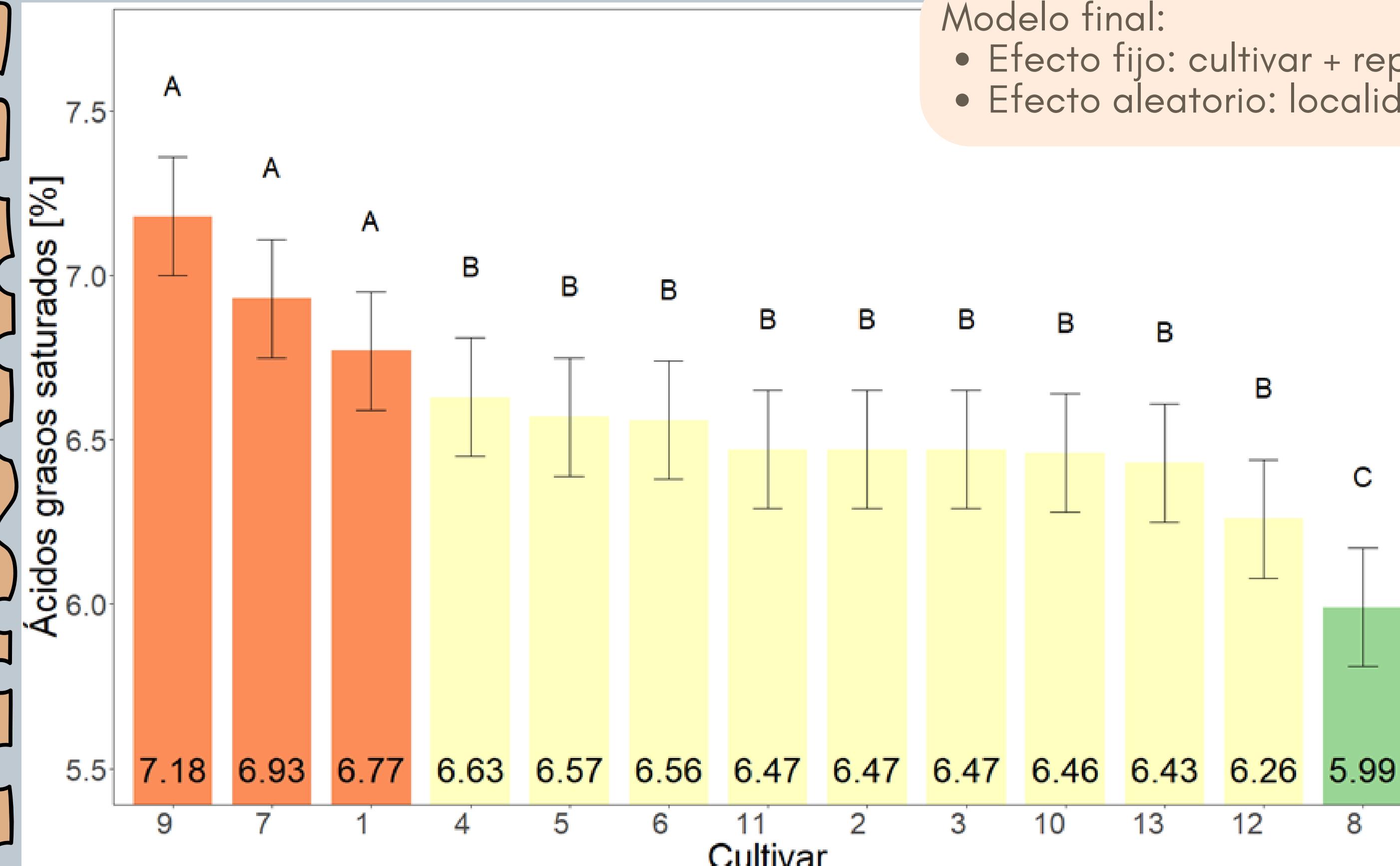


Figura 2: Contenido de ácidos grasos saturados en los distintos cultivares de colza evaluados en el estudio y resultados de la prueba a posteriori de DGC.

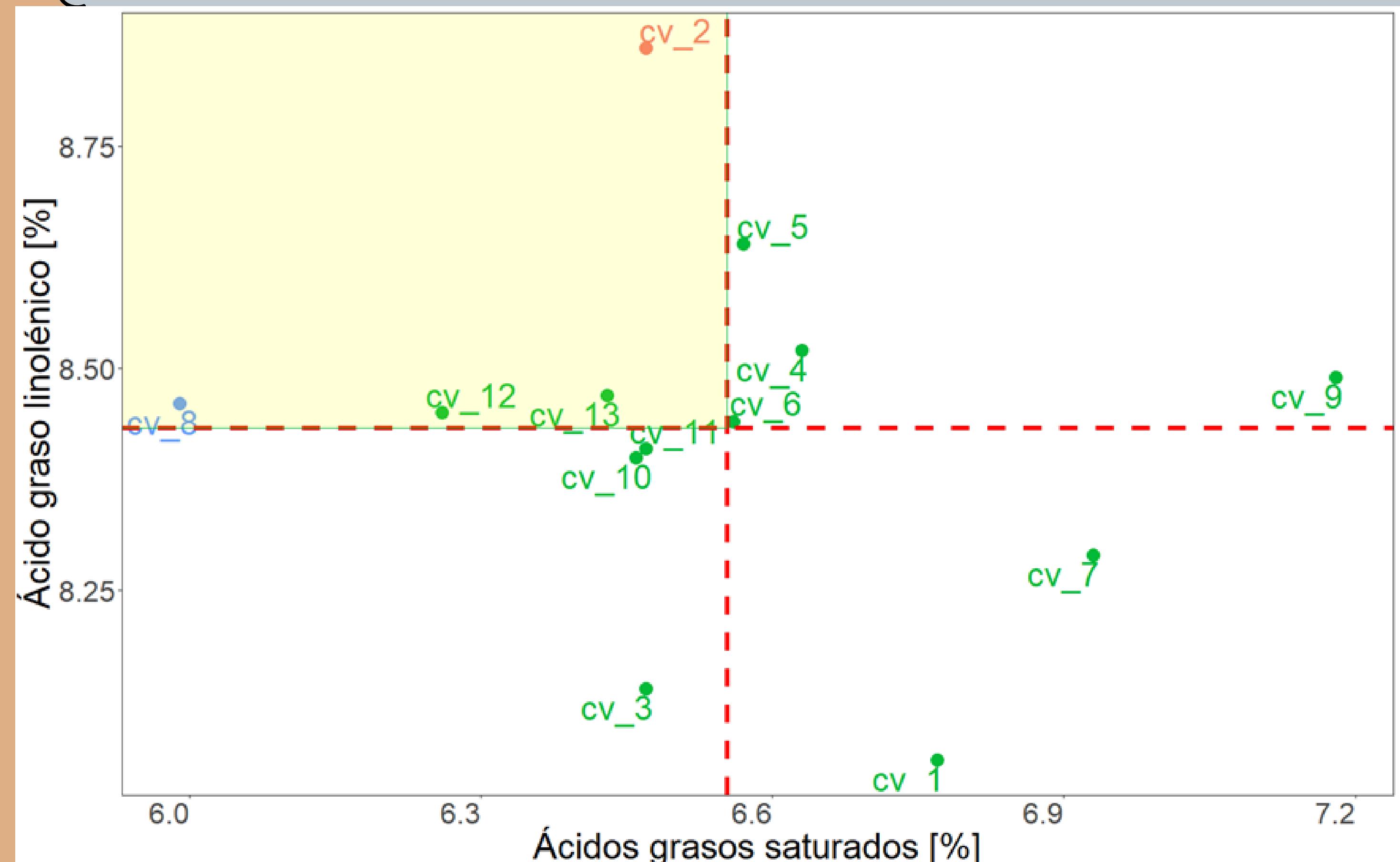
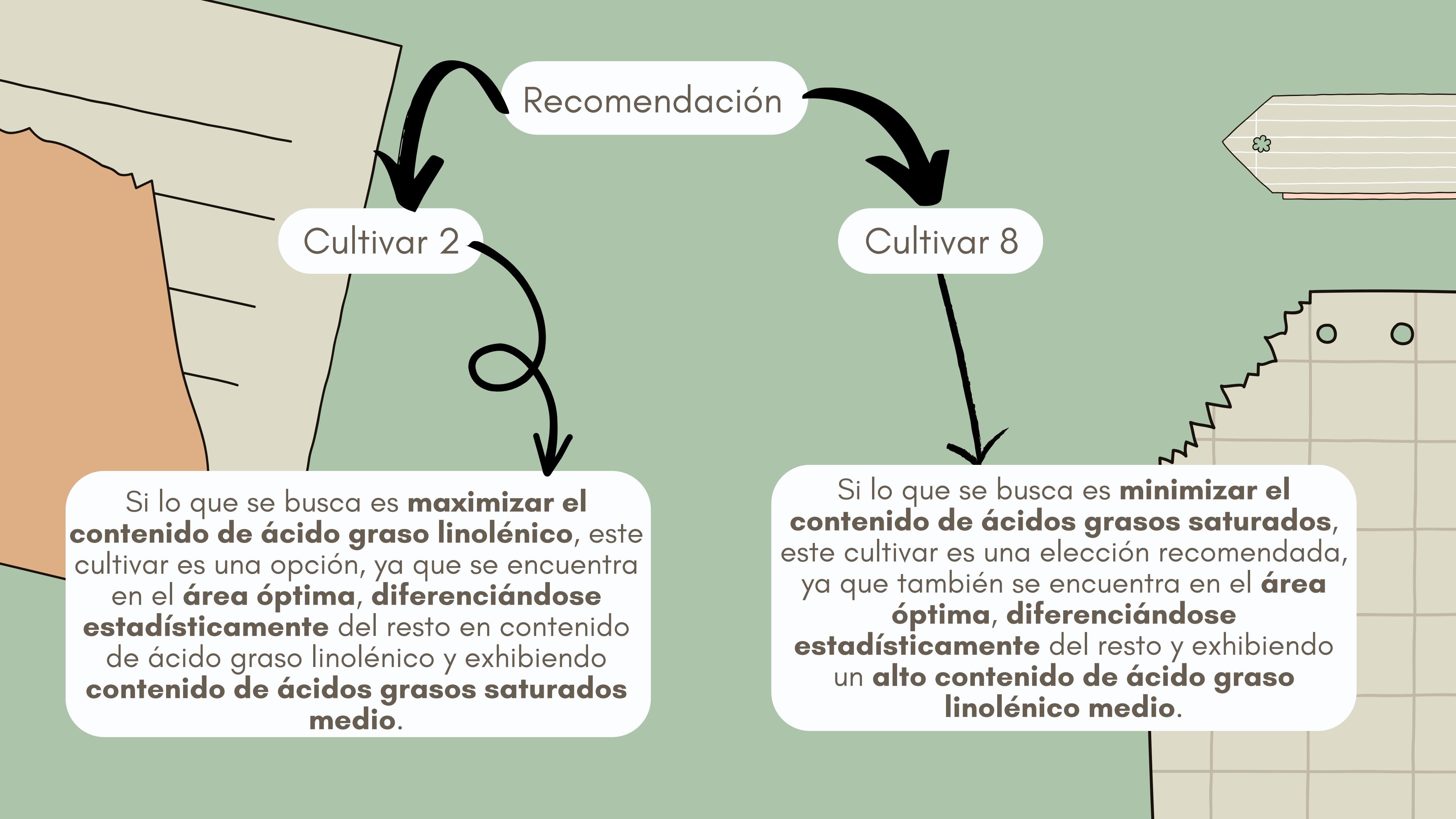


Figura 3: Análisis conjunto de las componentes de contenido de ácido graso linolénico y ácidos grasos saturados mediante un diagrama de dispersión. El área de resultados óptima se encuentra coloreada, además de los cultivares que exhibieron diferencias significativas en la prueba a posteriori que lo distingue como el óptimo.



2. REALIZAR AGRUPACIONES DE CULTIVARES BASADAS EN LA SIMILITUD DE PERFILES DE ÁCIDOS GRASOS, TOCOFEROLES E ÍNDICE DE ESTABILIDAD.

METODOLOGÍA

¿Número óptimo de grupos ?

- Índice de Silhouette
- Índice de Calinski-Harabasz

Agrupamiento

Algoritmo de clustering jerárquico

Dendrograma

Algoritmo de clustering no jerárquico

Análisis de componentes principales

Biplot

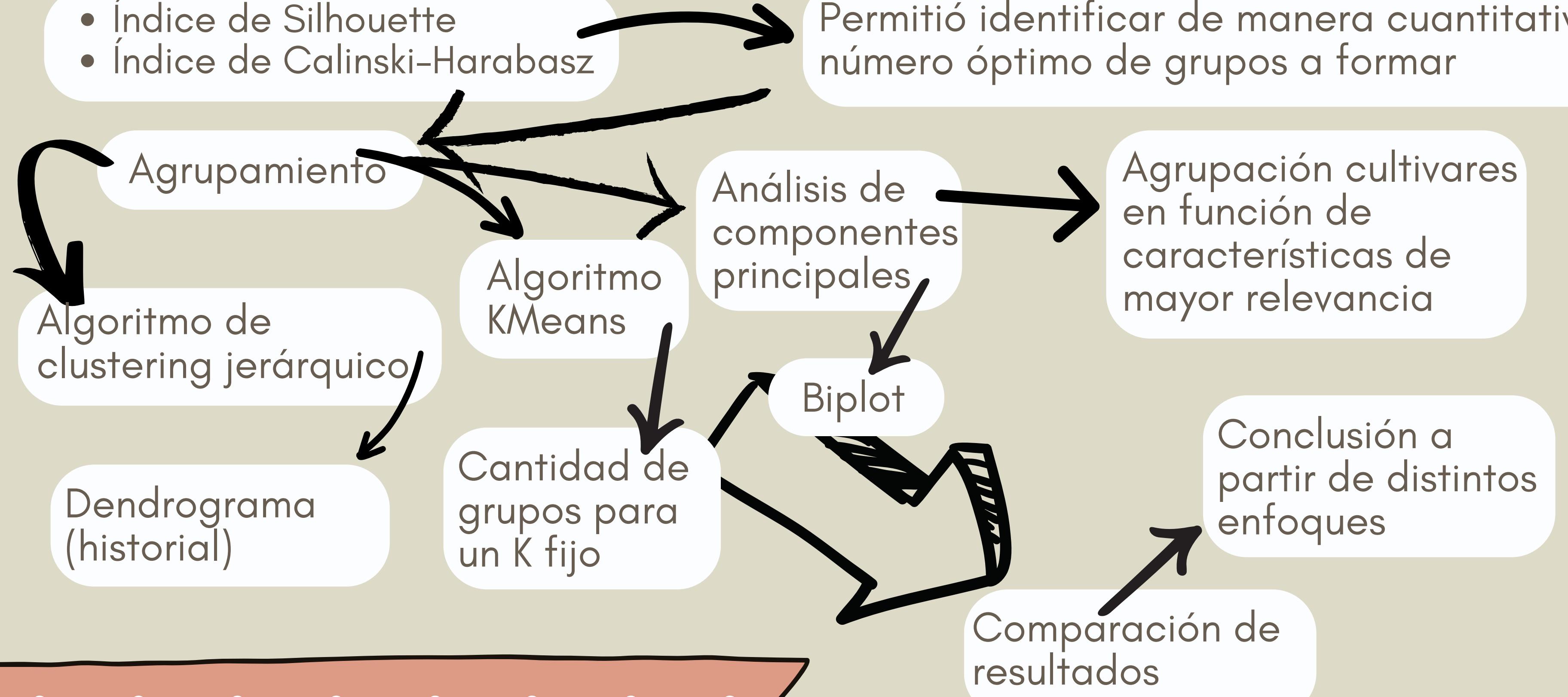
- Método de Ward
- Dist euclidea

Algoritmo KMeans

Comparación de resultados

ABORDAJE

- Índice de Silhouette
 - Índice de Calinski-Harabasz
- Permitió identificar de manera cuantitativa número óptimo de grupos a formar



ANÁLISIS INICIAL

Análisis conglomerado jerárquico sin tener en cuenta efecto ambiente ni repetición.

Notable variabilidad asociada a los factores no contemplados

Variabilidad entre repeticiones considerable

Tendencias comunes a lo largo de los distintos ambientes

Cómo se agrupaban los cultivares dentro de cada ambiente.

Análisis condicionado por ambiente.

Análisis promedios del perfil químico para cada cultivar, agrupando las repeticiones, y condicionado por ambiente.

- cultivar 1 y cultivar 3
- cultivar 10 y cultivar 2
- cultivar 7 y cultivar 9
- cultivar 11 y cultivar 5

ABORDAJE

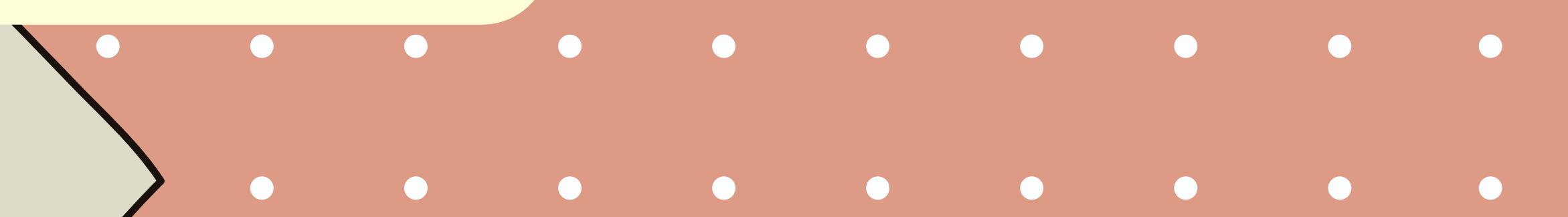
Notable variabilidad asociada
a localidad y repetición

MLM con efectos aleatorios

Variabilidad principalmente
asociada a cultivar

¿Cómo podemos reducirla?

Análisis sobre
residuos



¿NÚMERO ÓPTIMO DE GRUPOS?

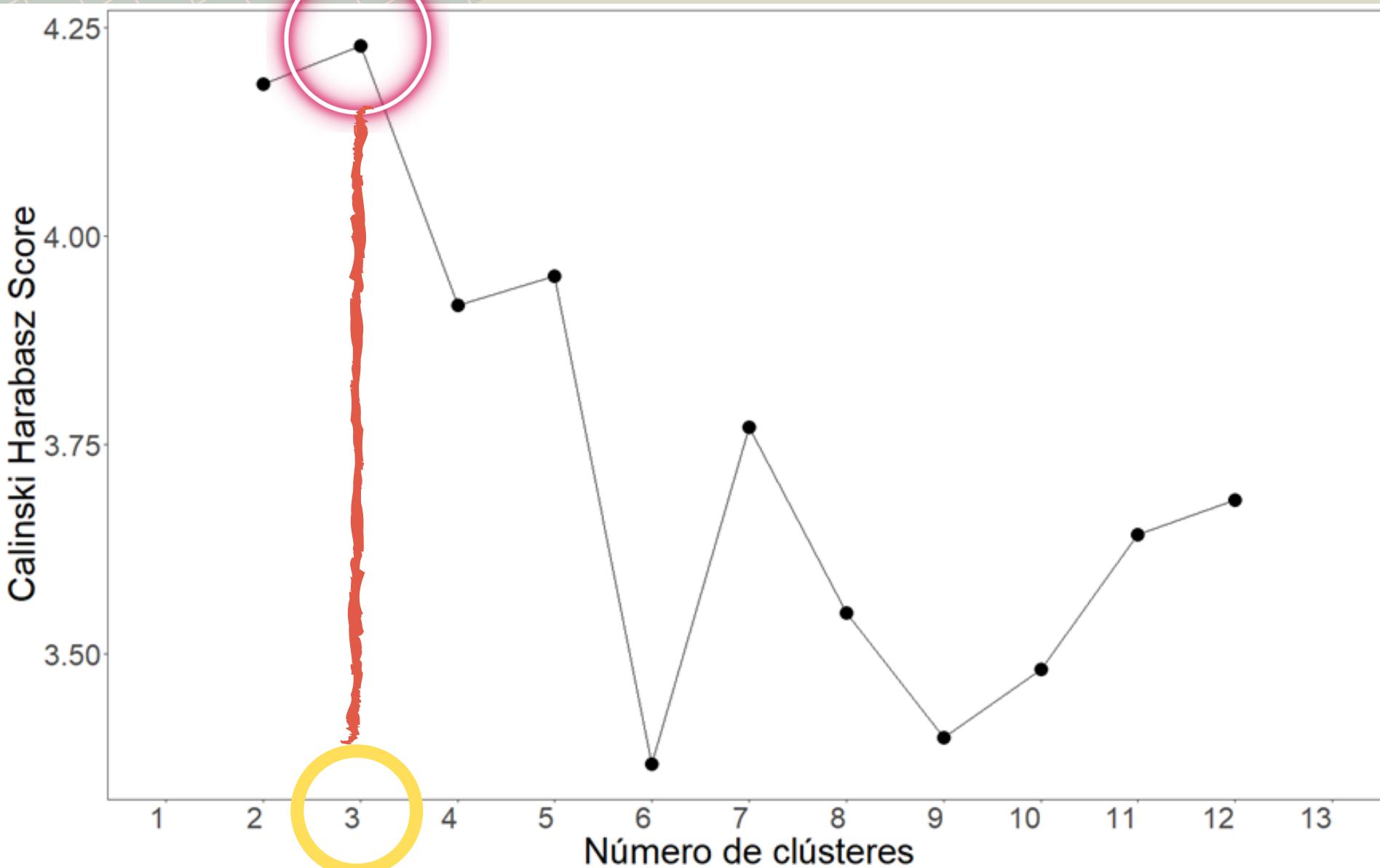


Figura: Variación del índice de Calinski Harabasz con el número de clústeres.

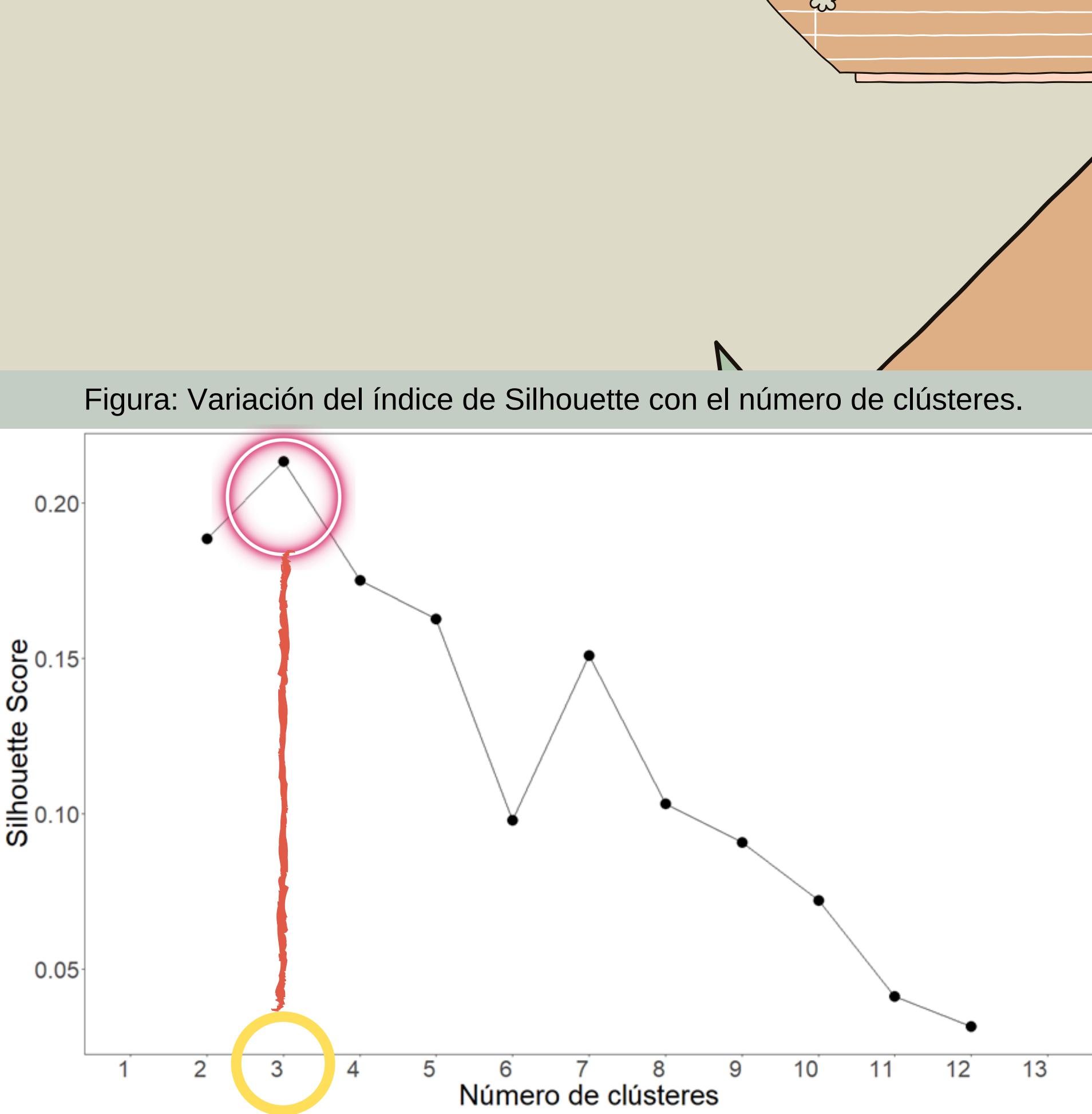


Figura: Variación del índice de Silhouette con el número de clústeres.

CLUSTERING JERÁRQUICO

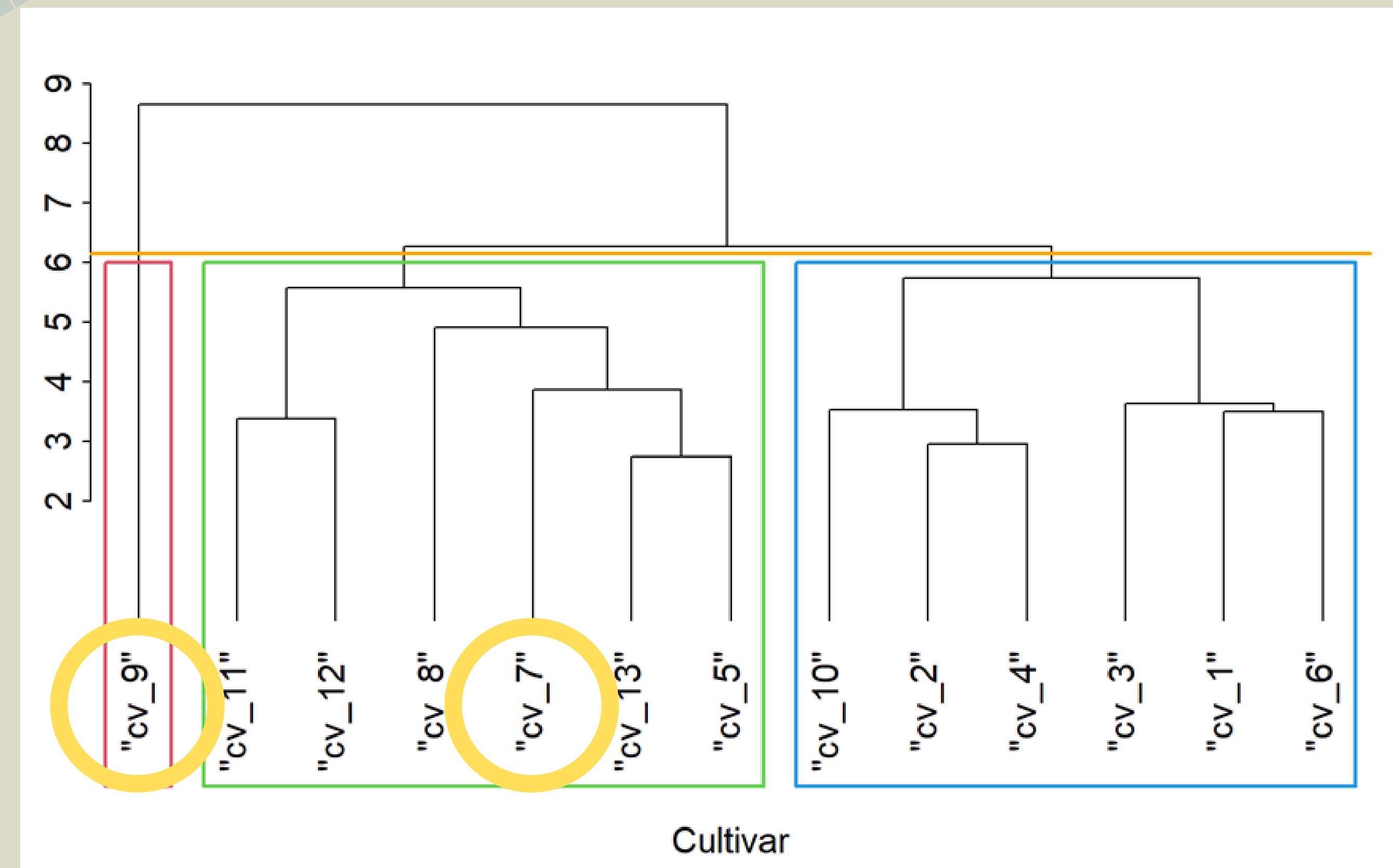


Figura 4: Dendrograma de agrupamiento jerárquico obtenido con la métrica euclídea y el algoritmo de agrupamiento completo. Los cuadrados indican los agrupamientos de los cultivares al considerar 3 grupos.

ACP + BIPLOT

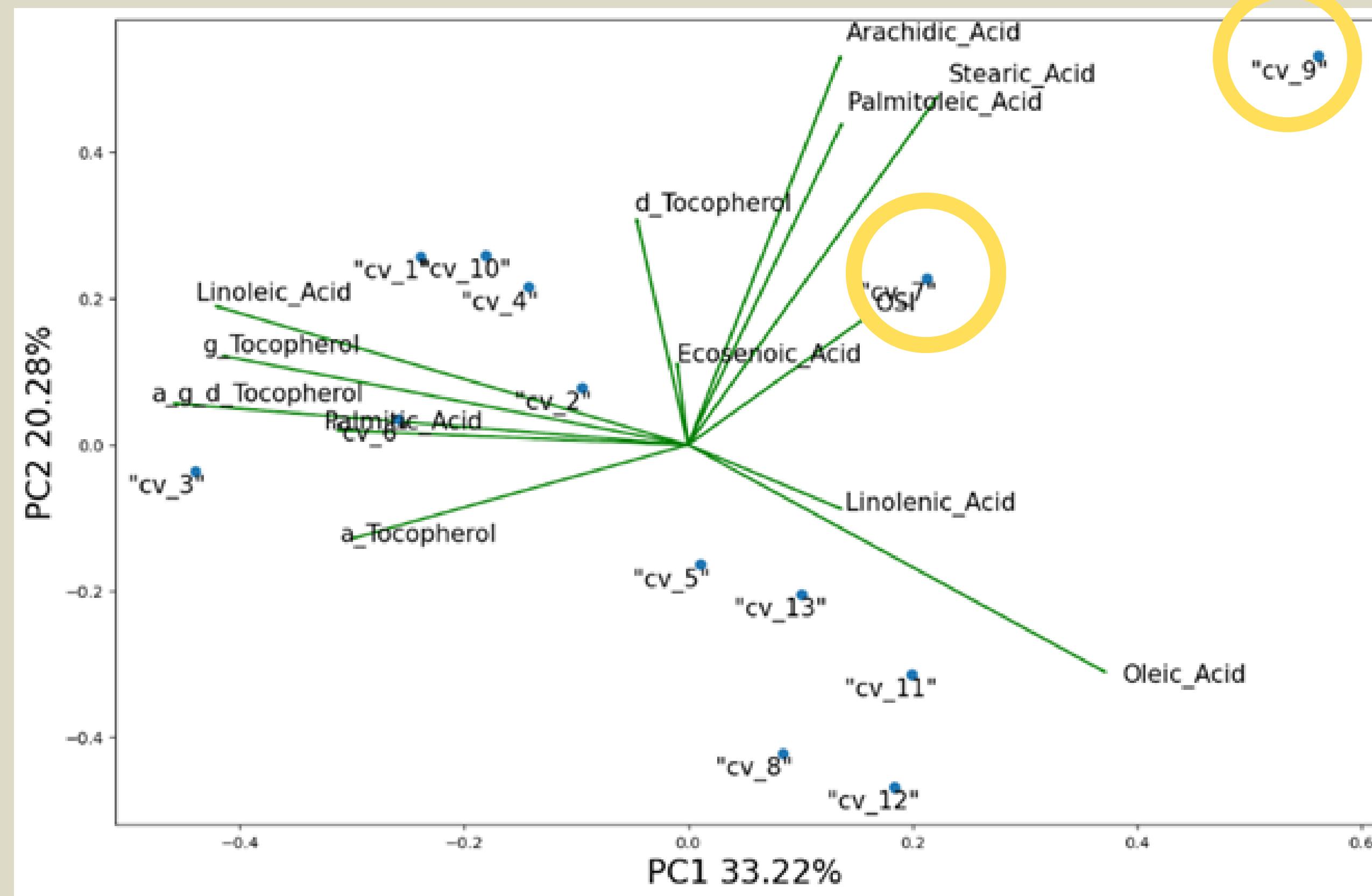
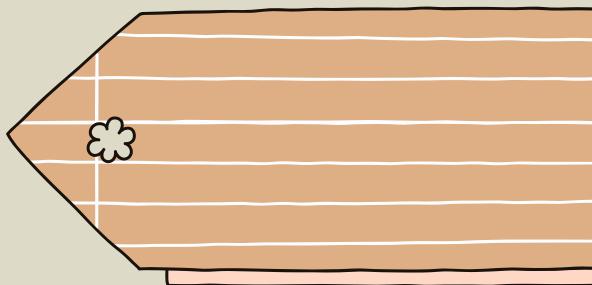


Figura 5: Gráfico biplot resultante de un Análisis de Coordenadas Principales a partir las componentes que definen el perfil químico de los cultivares de Colza evaluados.

Y ENTONCES, ¿QUIÉN AGRUPAMOS?



Análisis inicial

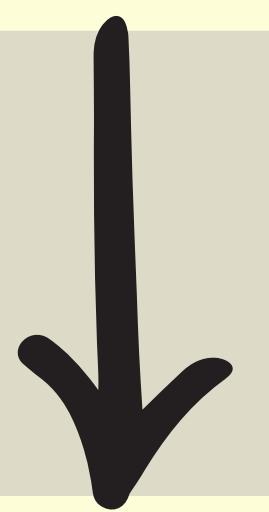


- cultivares 1 y 3
- cultivares 10 y 2
- cultivares 7 y 9
- cultivares 11 y 5

K-means

- Grupo 1: 7 y 9
- Grupo 2: 11, 12, 8, 13 y 5
- Grupo 3: 10, 2, 4, 3, 1 y 6

Jerárquico

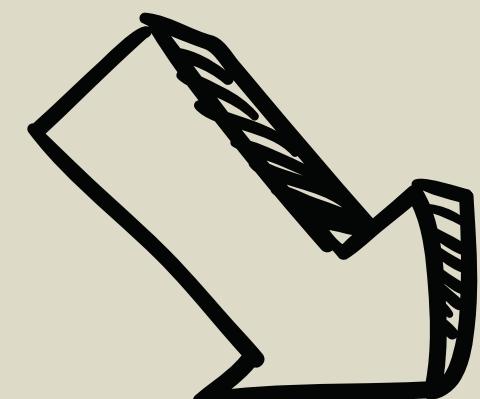


- Grupo 1: 9
- Grupo 2: 7, 11, 12, 8, 13 y 5
- Grupo 3: 10, 2, 4, 3, 1 y 6

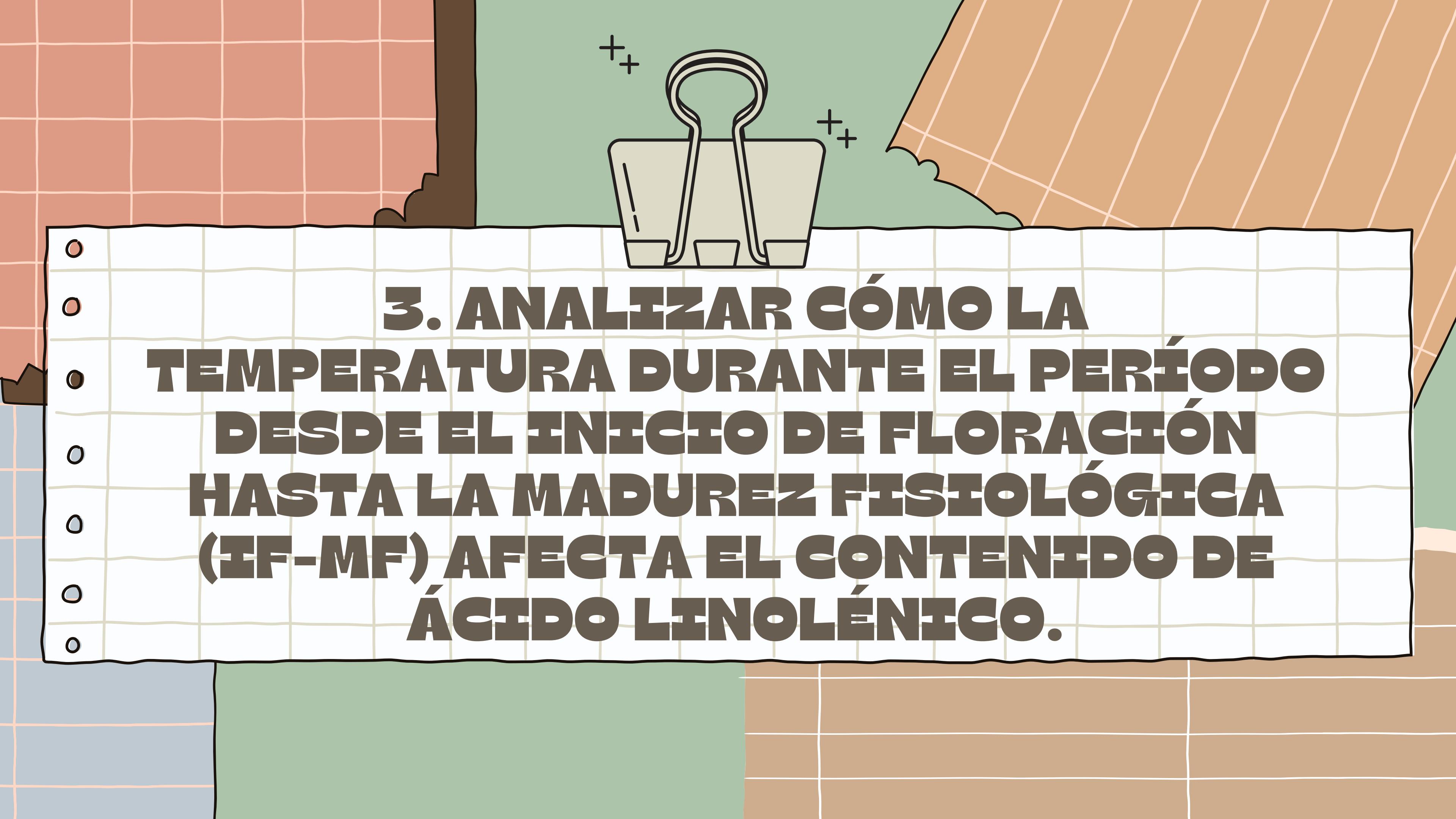
ACP + BI PLOT

- Grupo 1: 9
- Grupo 2: 11, 12, 8, 13 y 5
- Grupo 3: 10, 2, 4, 3, 1 y 6

(7) Función
de la
proyección.



A partir del análisis conjunto de los resultados de los distintos enfoques, detectamos que es más pertinente conservar el cultivar 7 conjuntamente con los cultivares 11, 12, 8, 13 y 5; a pesar de esto, es el cultivar que mayor similitud guarda con el número 9.



3. ANALIZAR CÓMO LA TEMPERATURA DURANTE EL PERÍODO DESDE EL INICIO DE FLORACIÓN HASTA LA MADUREZ FISIOLÓGICA (IF-MF) AFECTA EL CONTENIDO DE ÁCIDO LINOLÉNICO.

METODOLOGÍA

Análisis basado en un modelo de regresión lineal múltiple

Ajuste iterativo

Chequeo supuestos:

- Normalidad (QQ-plot)
- Homogeneidad de varianzas (Residuos vs predichos)
- Residuos parciales

Análisis de correlación

Criterio de selección:

- Minimización
 - AIC
 - BIC
 - MSE
- Maximización
 - R^2

SIMULTANEAEMENTE

Evitamos multicolinealidad entre variables ambientales

ABORDAJE

Análisis basado en un modelo de regresión lineal múltiple

Prueba de hipótesis

Significancia estadística de la relación entre las variables

Análisis de correlación

Análisis de residuos parciales

Descontar la influencia de las demás variables

Evitamos multicolinealidad entre variables ambientales

$$\begin{aligned} H_0 &: \beta = \beta_0 \\ H_1 &: \beta \neq \beta_0 \end{aligned}$$

RESIDUOS PARCIALES

Acido linolenico = $b_0 + b_1 * \text{Tmax} + b_2 * \text{Precip.} + b_3 * \text{Evap.}$

Coeficiente	Estimación	p-valor
b_0	14,64	<0,0001
b_1	-0,28	<0,0001
b_2	0,01	<0,0001
b_3	0,00002	0,9487

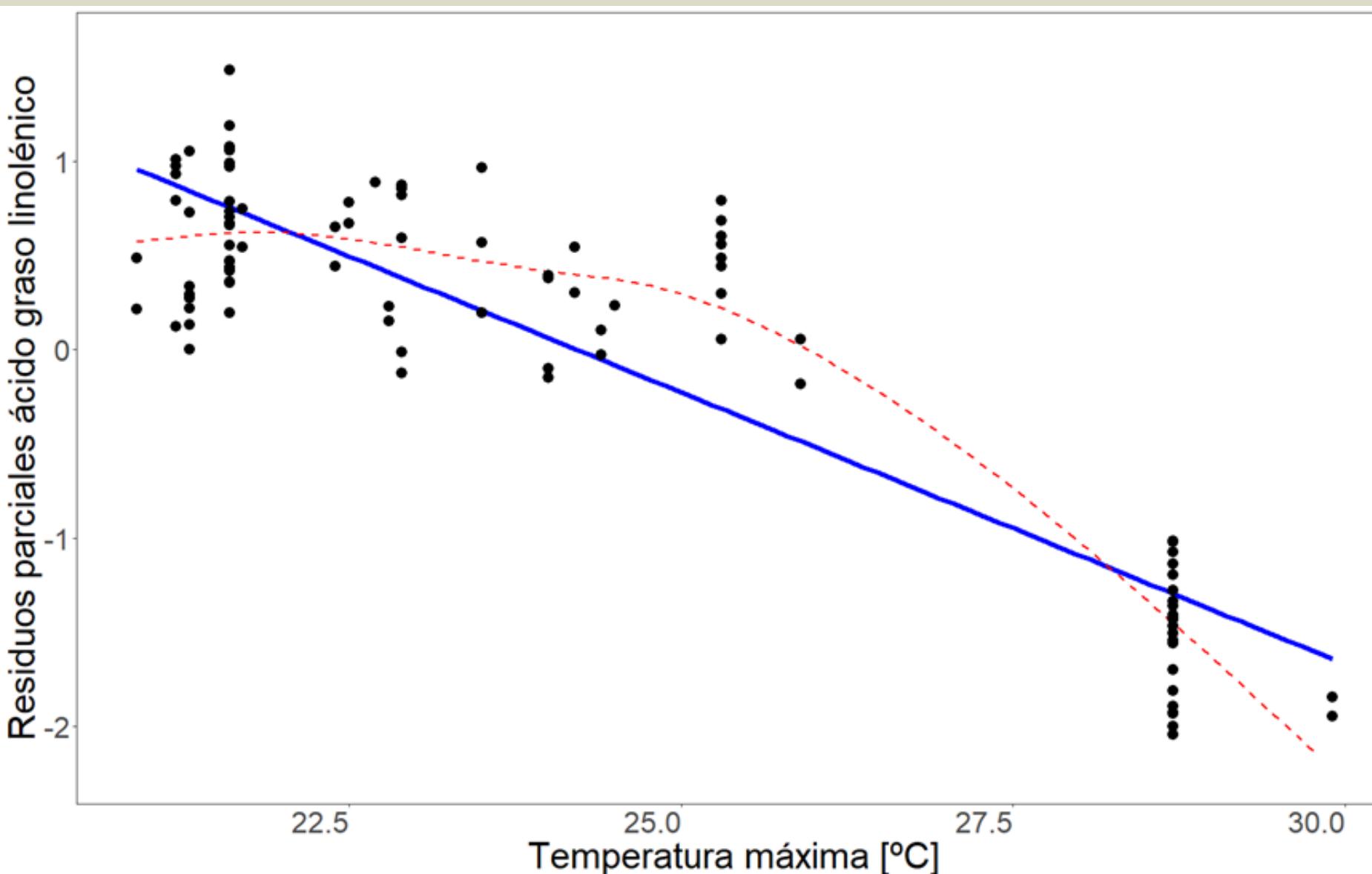
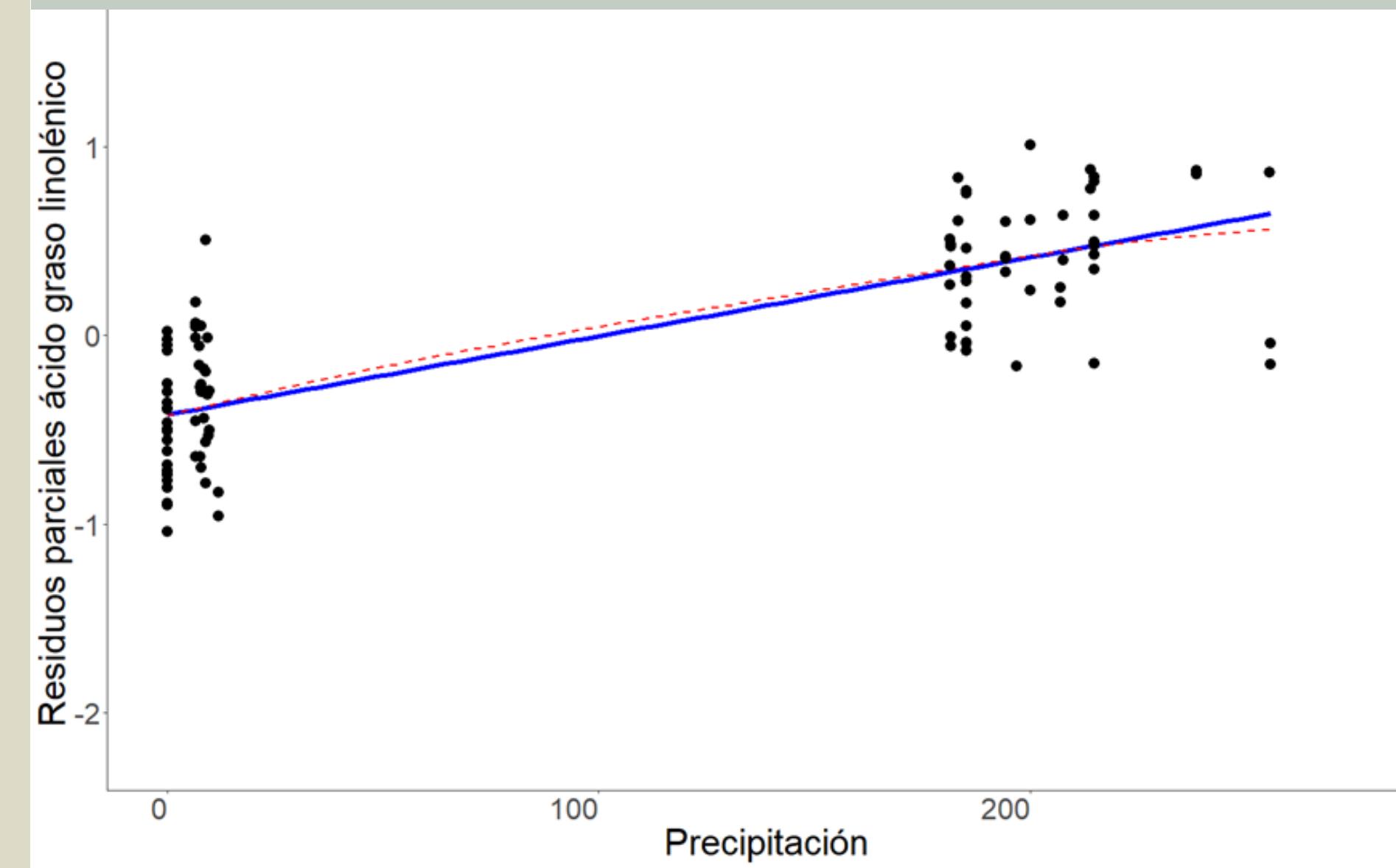


Figura 6: Gráfico de residuos parciales de ácido graso linolénico contra temperatura máxima

- AIC = 120,51
- BIC = 133,58
- R^2 = 0,67
- MSE = 0,21

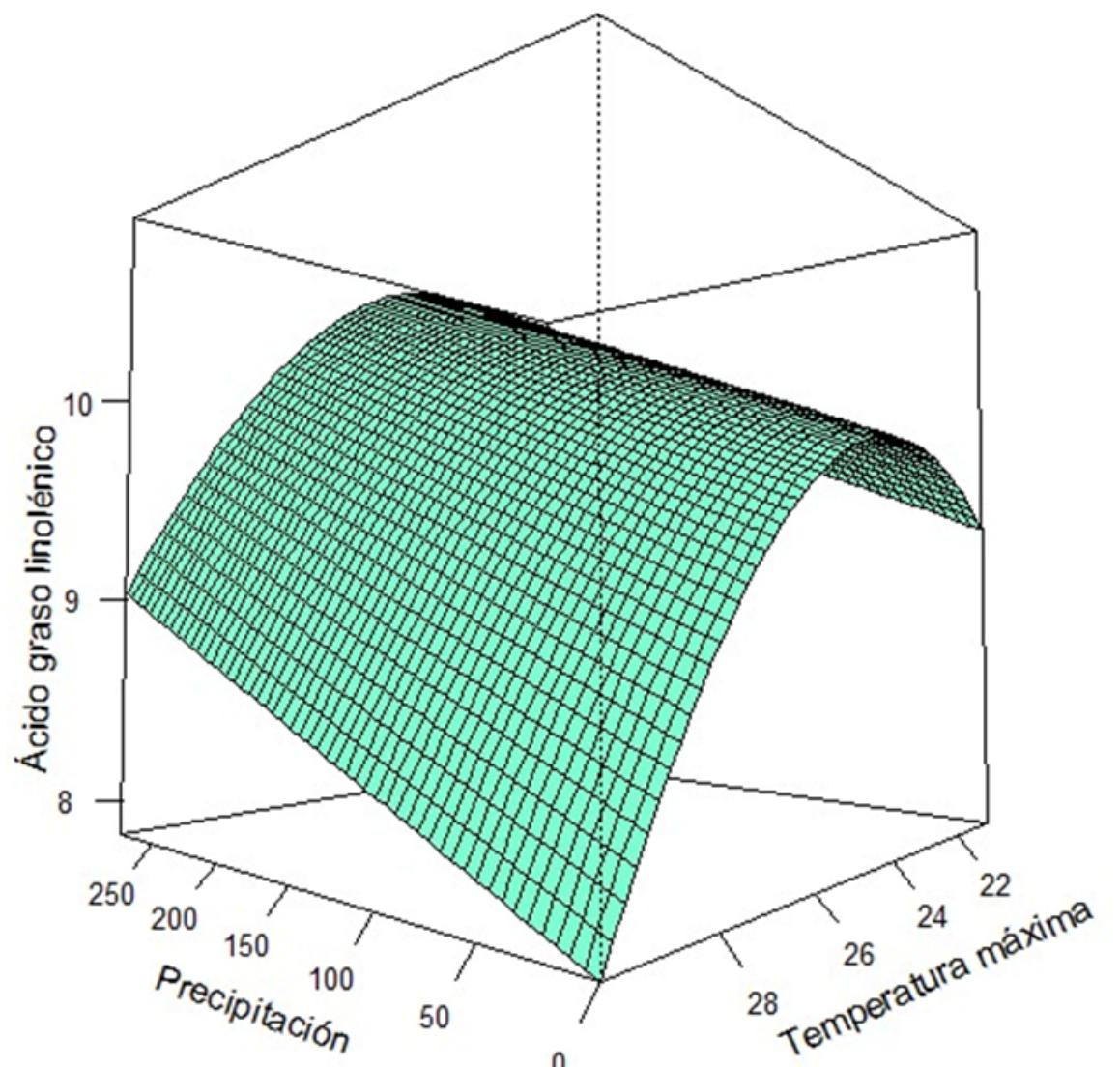
Figura 7: Gráfico de residuos parciales de ácido graso linolénico contra precipitación.



¿CUÁL FUE EL MEJOR MODELO?

Acido linolenico = b0 + b1 * Tmax + b2 * Tmax^2 + b3 * Precip.

Coeficiente	Estimación	p-valor
b0	-18,33	<0,0001
b1	2,37	<0,0001
b2	-0,05	<0,0001
b3	0,0043	<0,0001



- AIC = 64,01
BIC = 76,88
 R^2 = 0,81
MSE = 0,12

Figura 12: Vista en planta del gráfico 12 con isolíneas de rendimiento.

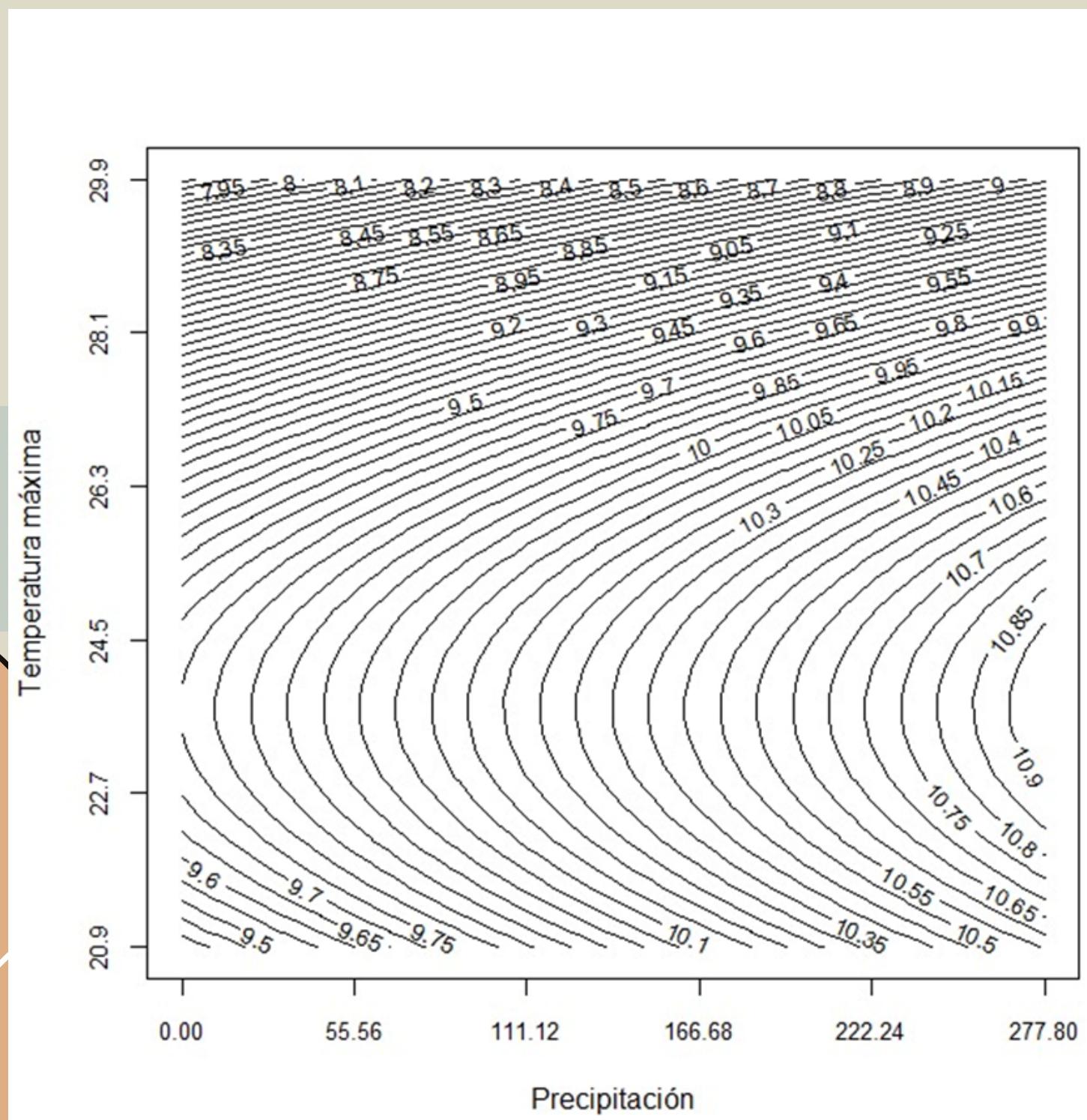


Figura 11: Gráfico de superficie y contorno para la relación de vinculación entre la temperatura máxima, la precipitación y el contenido de ácido graso linolénico.

La temperatura y la precipitación afectan significativamente el contenido de ácido linolénico desde el inicio de floración hasta la madurez fisiológica (IF-MF).

Las condiciones climáticas óptimas para alcanzar el contenido máximo de ácido graso linolénico corresponden a **precipitaciones más elevadas**, específicamente registradas en la base de datos con un valor de **277.8**, y **temperaturas máximas que son iguales o inferiores a 23.5**. En estas condiciones, se logra un **contenido de ácido graso linolénico que puede alcanzar hasta 10.9**.

Estos hallazgos indican que la interacción entre estas variables climáticas específicas puede ser determinante para maximizar el contenido deseado de ácido graso linolénico.

4.

- EVALUAR LA CONTRIBUCIÓN RELATIVA DE LOS FACTORES (CULTIVAR O AMBIENTE) EN LA VARIACIÓN DE CADA COMPONENTE DEL GRANO.
- DETERMINAR SI EXISTE INTERACCIÓN GENOTIPO \times AMBIENTE.
- RECOMENDAR CULTIVARES EN UN CONTEXTO GENERAL O ESPECÍFICO POR SITIO.

METODOLOGÍA

Ajuste de MLM

Iterativo

- Fijo (Interacción)
- Aleatorio
- Función de varianza

Criterio de selección:

- Minimización simultánea
 - AIC
 - BIC

Efectos aleatorios: cultivar, ambiente y su interacción.

$$y = G + E + GxE$$

BLUPs de cultivar
+ Estabilidad

AMMI

Recomendación
general

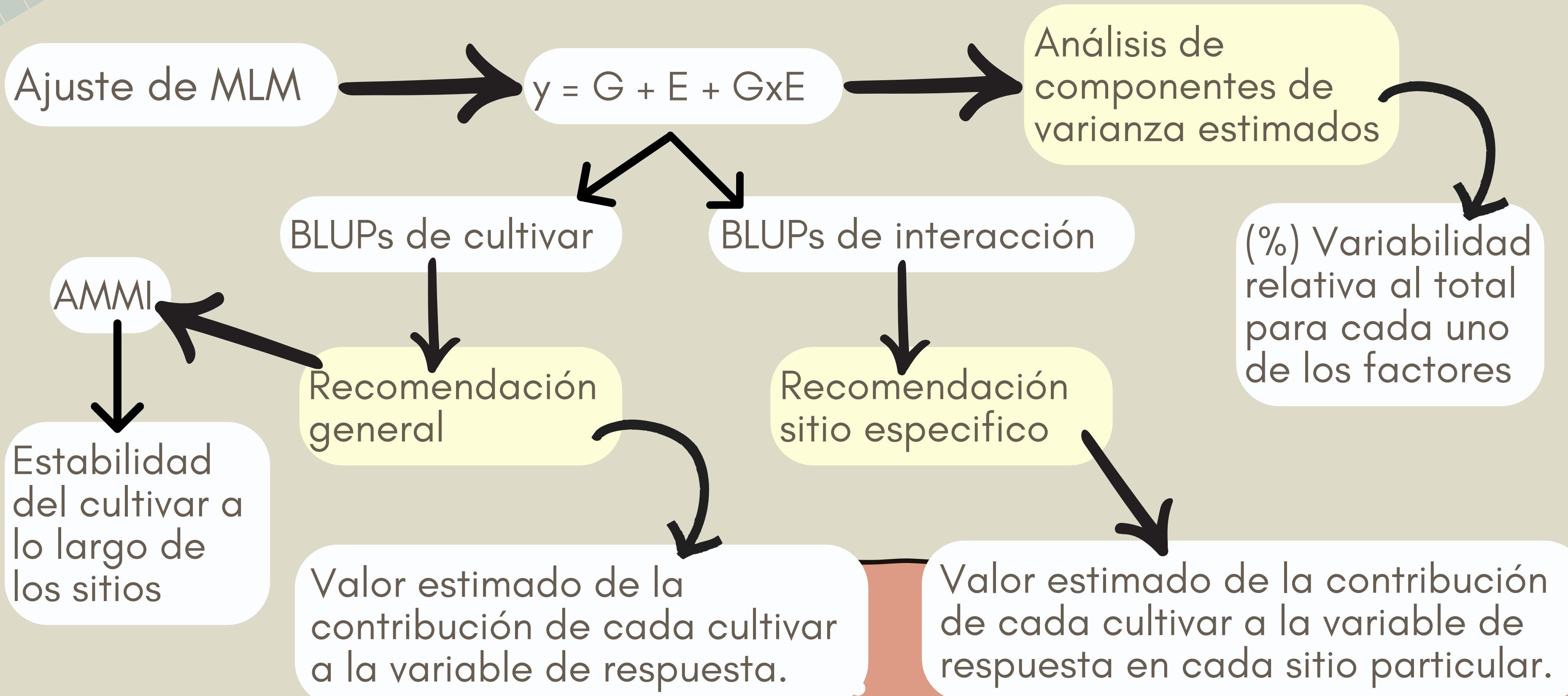
Análisis de
componentes de
varianza estimados

BLUPs de interacción

Recomendación
sitio específico

Interacción ¿es
significativa?

ABORDAJE



CONTRIBUCIÓN DE CADA FACTOR EN CADA CARACTERÍSTICA DEL PERFIL

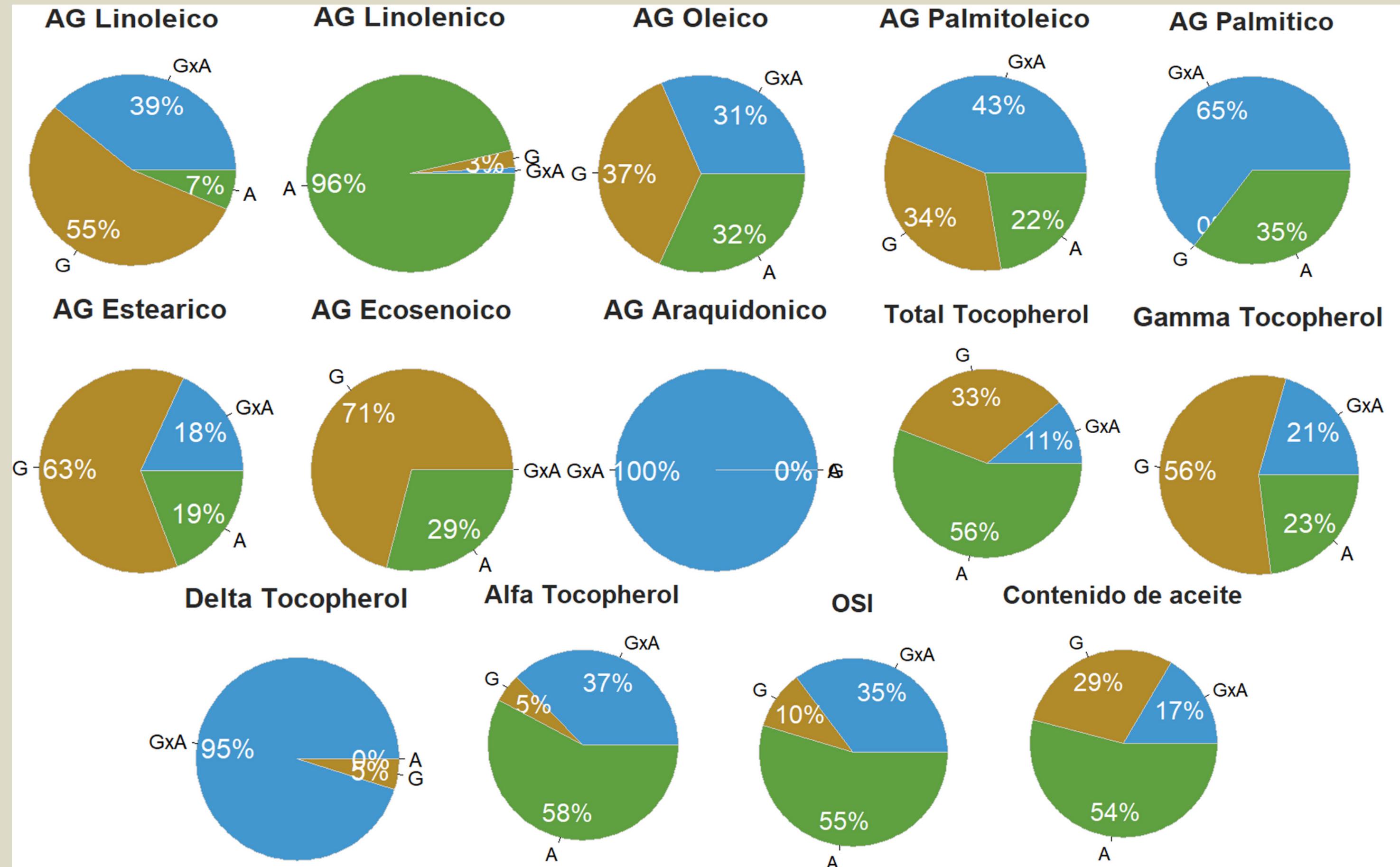


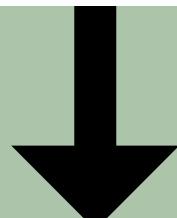
Figura 9: Gráfico de sectores de la proporción de varianza para cada componente del perfil explicada por G (marrón), A (verde) y GxA (azul).

LA INTERACCIÓN ¿ES SIGNIFICATIVA?

Compuesto	p valor interacción (Localidad*Cultivar)
Tocopherol α	0,0003
Tocopherol γ	0,0055
Tocopherol δ	<0,0001
Tocopherol αγδ	0,0617
Ácido Palmitico	0,0022
Ácido Palmitoleico	0,2816
Ácido Estearico	0,0735
Ácido Linoleico	0,0232
Ácido Linolenico	0,0261
Ácido Arachidico	0,0855
Ácido Ecosenoico	0,1581
OSI	0,0259
Contenido de aceite	0,1745

METODOS PARA LA SELECCION DEL OS CULTIVARES EN SENTIDO AMPLIO

Selección Multicriterio



- Criterios del perfil:
- AG Linolenico
 - Oleico
 - Tot tocoferoles
 - OSI
 - AG Insaturados
 - Cont. de aceite

Ajuste de MLM con matriz de BLUPs

Ajuste de un AMMI y generación de un índice de estabilidad

Compuesto	CP1 %
AG Linolenico	42.0
Oleico	52.8
tot. tocoferoles	50.7
OSI	72.0
cont. de aceite	47.9

Recomendaciones

Para aceite de mesa se recomienda el cultivar 2 por su perfil de AG y estabilidad.

Para un uso mas de fritura se recomienda el cultivar 5 con un alto OSI y AG insat

CONTRIBUCIÓN DE CADA FACTOR EN CADA CARACTÉRISTICA DEL PERFIL

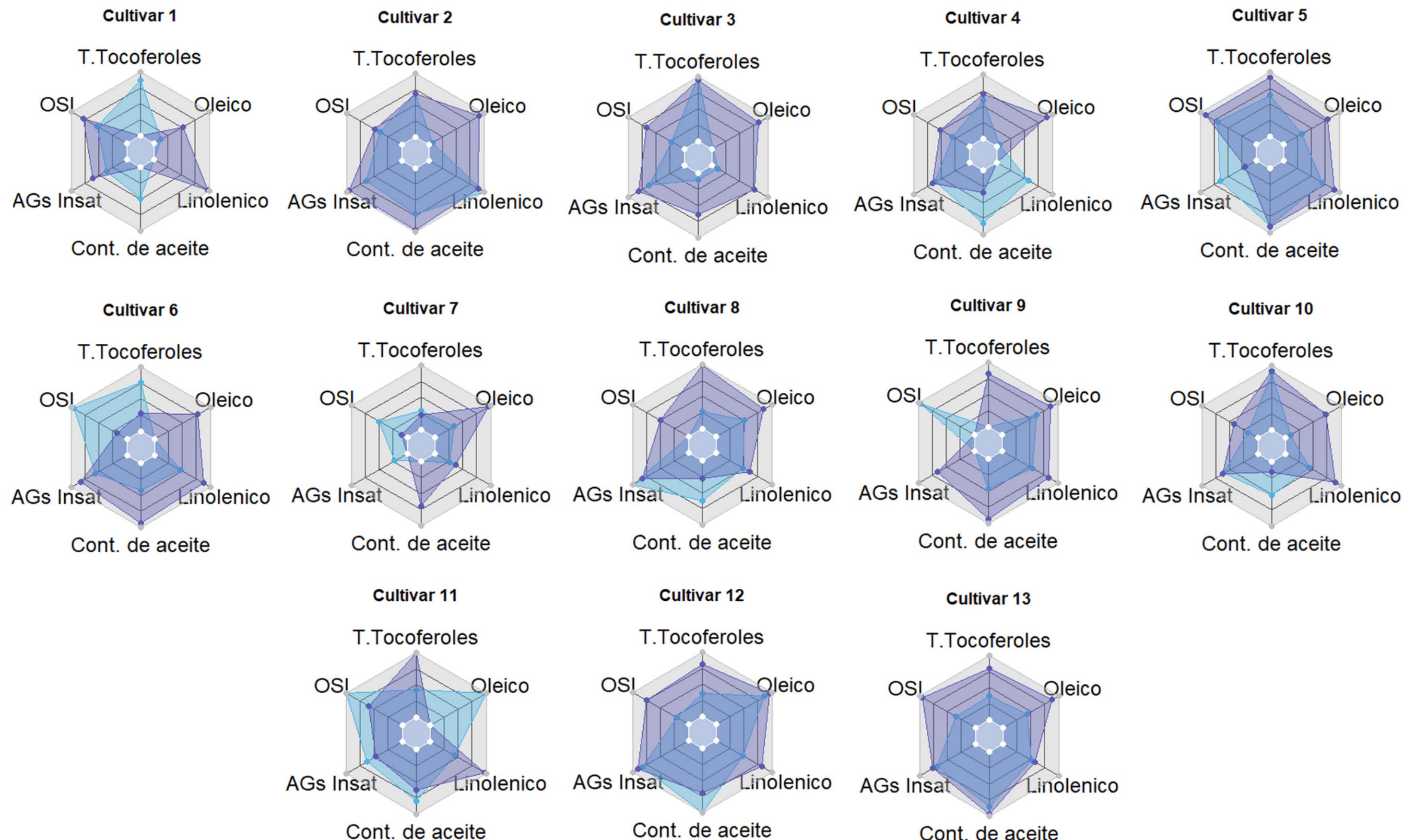
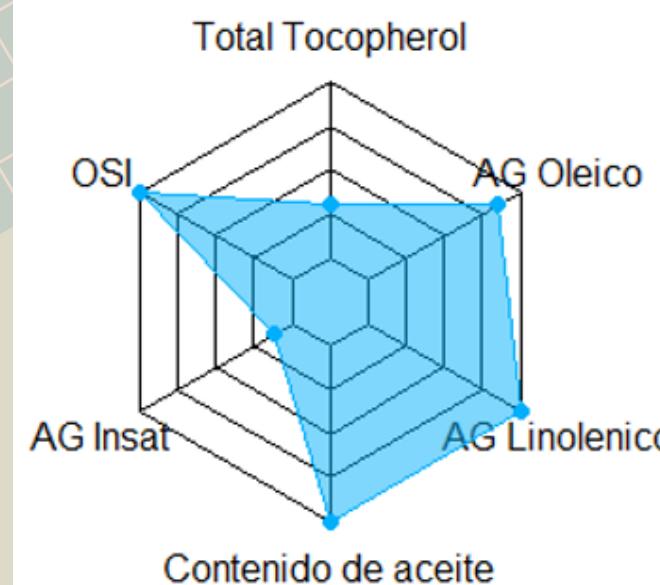


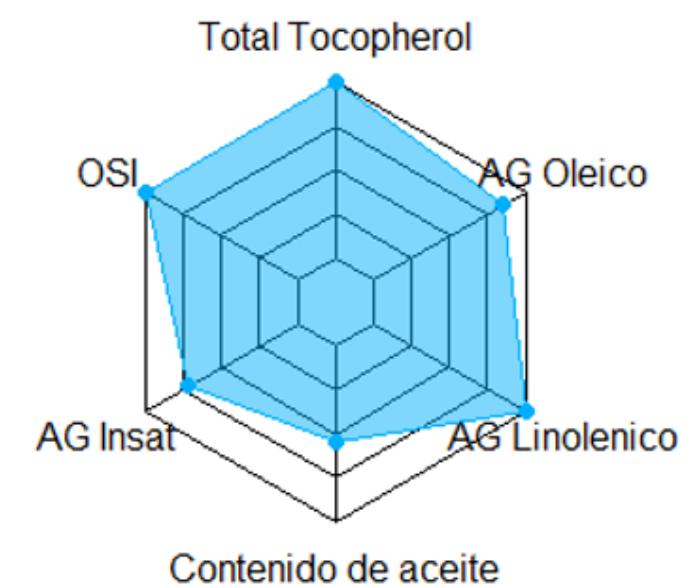
Figura 10: Gráfico de radar para los BLUPs extraídos del factor cultivar, se indica en violeta el índice de estabilidad y en azul el rendimiento para total tocoferoles, concentración de ácido graso linolénico, ácido graso oleico, contenido de aceite, ácidos grasos insaturados y OSI.

RECOMENDACIÓN SITIO ESPECÍFICO

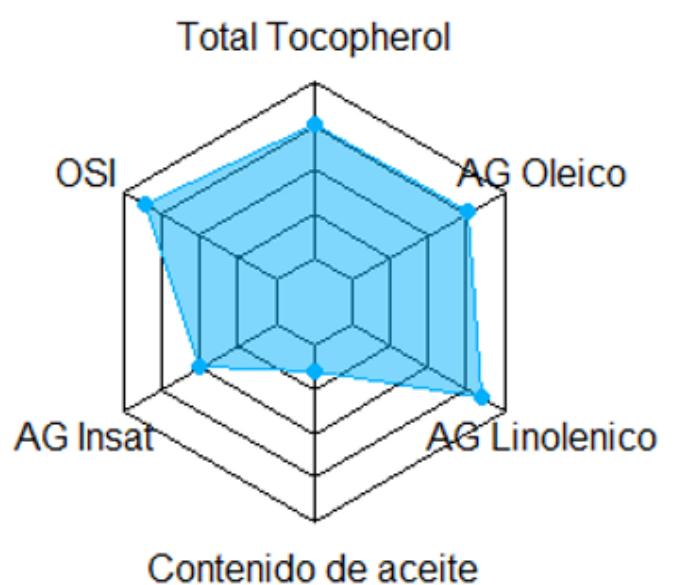
Localidad L1C12



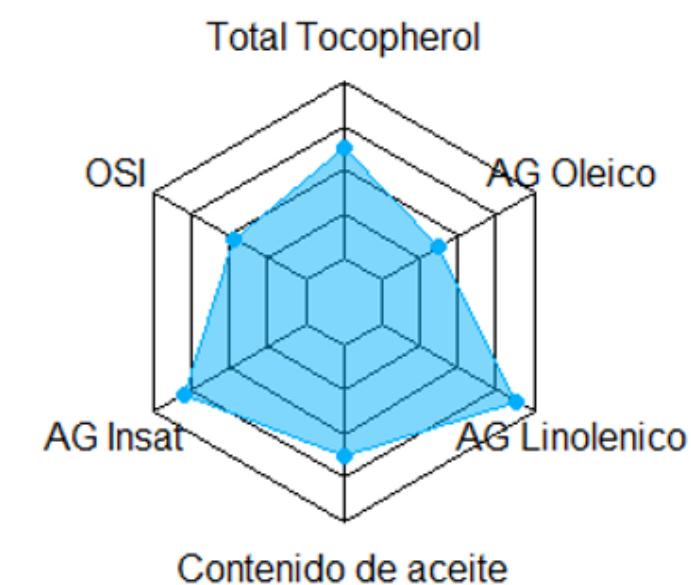
Localidad L2C11



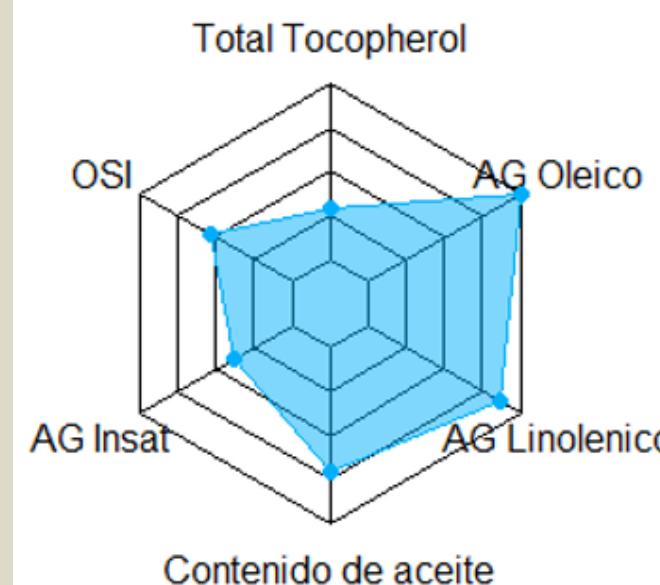
Localidad L3C3



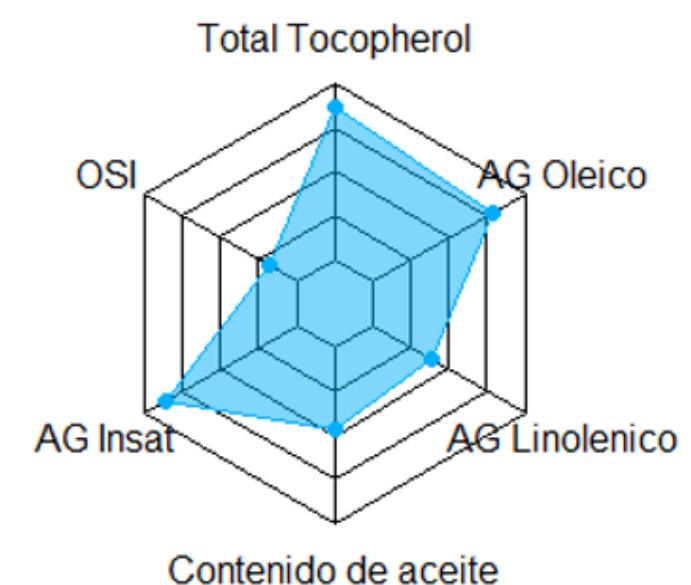
Localidad L4C4



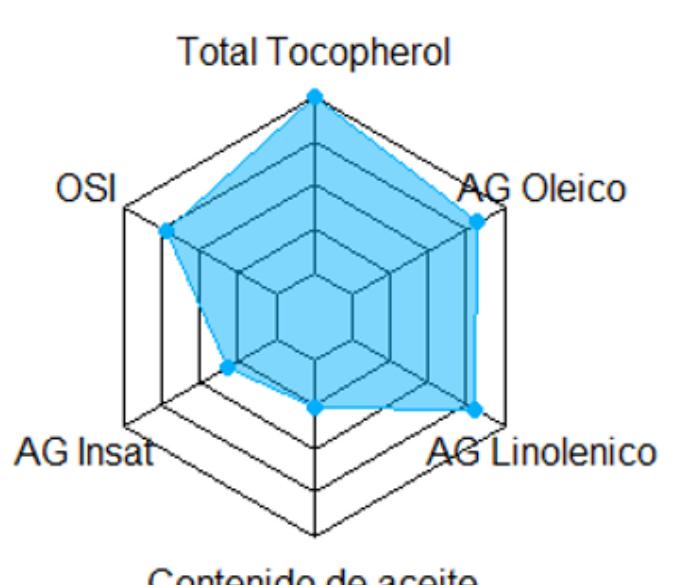
Localidad L1C5



Localidad L2C2



Localidad L3C6



Localidad L4C13

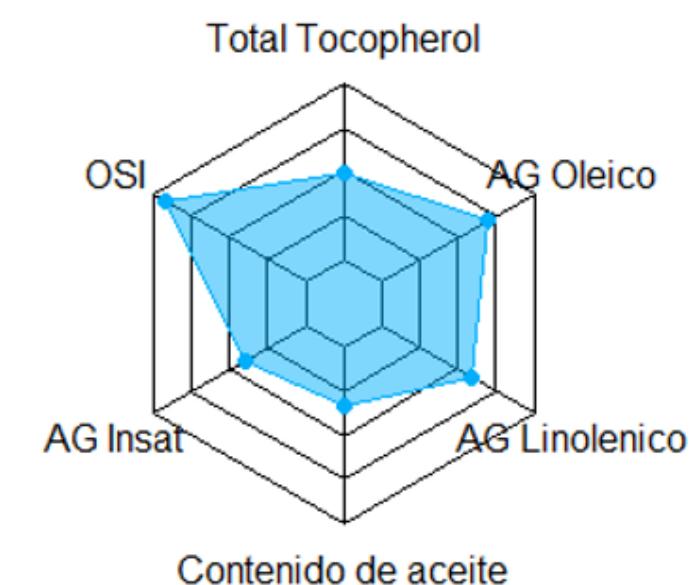
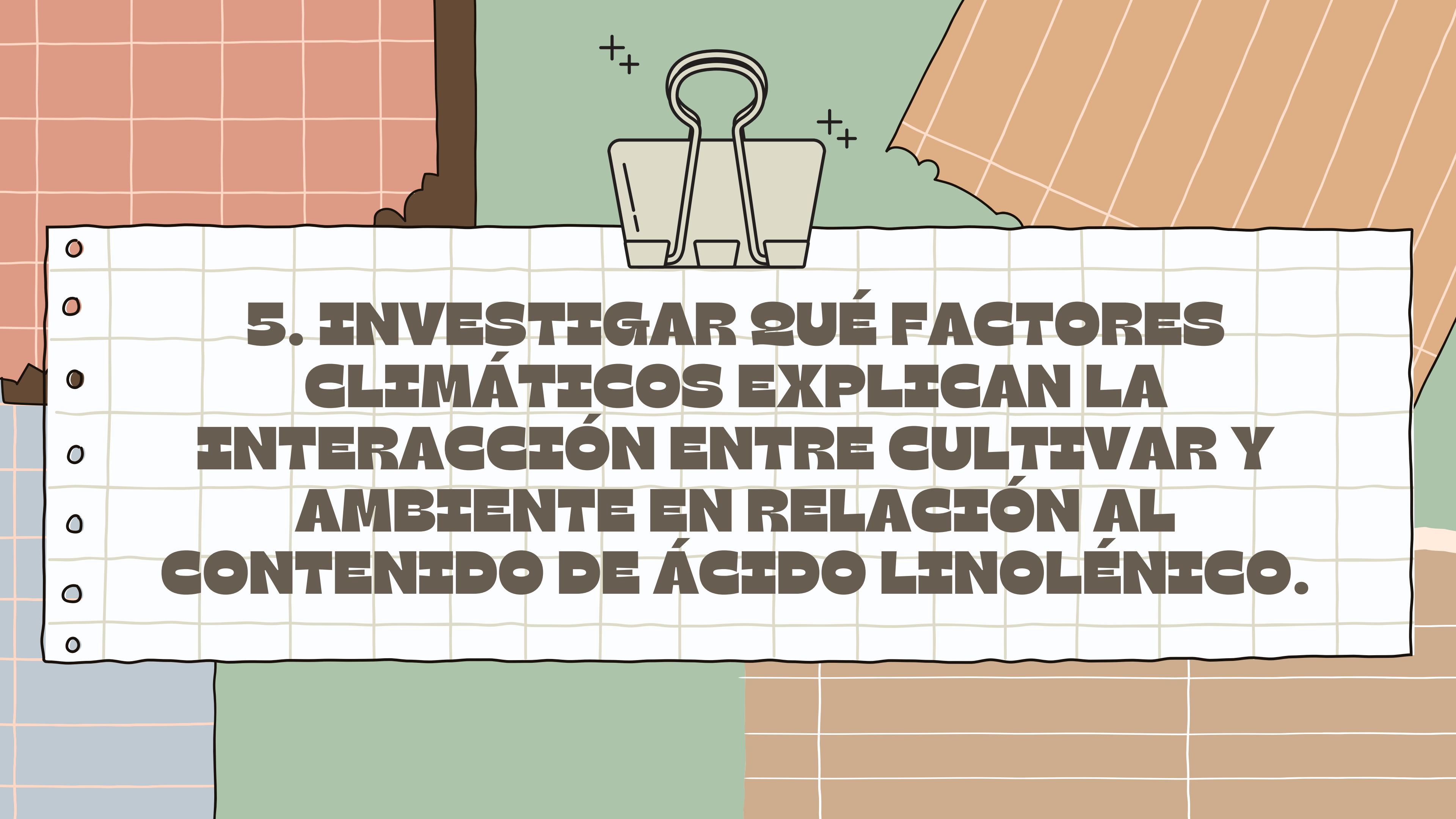


Figura 11: Gráfico de radar para los BLUPs extraídos del factor interacción entre cultivar y localidad, en se indica en violeta el índice de estabilidad y en azul el rendimiento para total tocoferoles, concentración de ácido graso linolénico, ácido graso oleico, contenido de aceite, ácidos grasos insaturados y OSI.



5. INVESTIGAR QUÉ FACTORES CLIMÁTICOS EXPLIQUEN LA INTERACCIÓN ENTRE CULTIVAR Y AMBIENTE EN RELACIÓN AL CONTENIDO DE ÁCIDO LINOLÉNICO.

METODOLOGÍA

Ajuste MLM

Efecto aleatorio: cultivar,
ambiente y repetición.

Aislamos interacción
 $G \times E$ en residuos.

Dos enfoques:

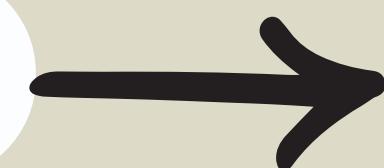
- Modelado estadístico
- Aprendizaje automático

Light GBM, Random
Forest y XGboost
Reg. selección del
mejor ajuste (LOOCV)

Regresión por
mínimos cuadrados
parciales (PLS de sus
siglas en inglés)

ABORDAJE

Residuos



Buscamos explicar la variabilidad de la interacción, no debida a los demás factores.

Aprendizaje automático

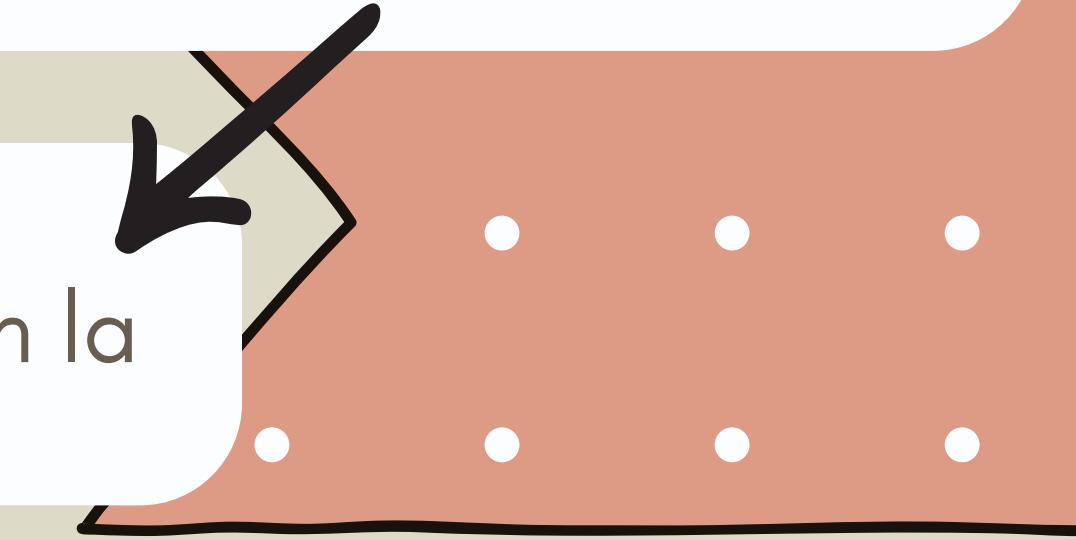
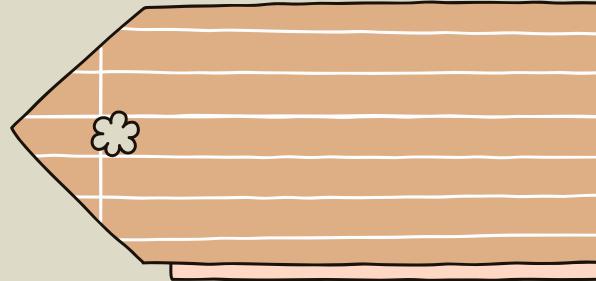
Interpretación de variables de mayor importancia para explicar la interacción

PLS

Permite tener encuentra la correlación existente entre múltiples regresoras (variables climáticas) y respuestas (interacción para los distintos cultivares).

Permutación

Tri-plot: interpretación de variables que más explican la interacción



RESULTADOS

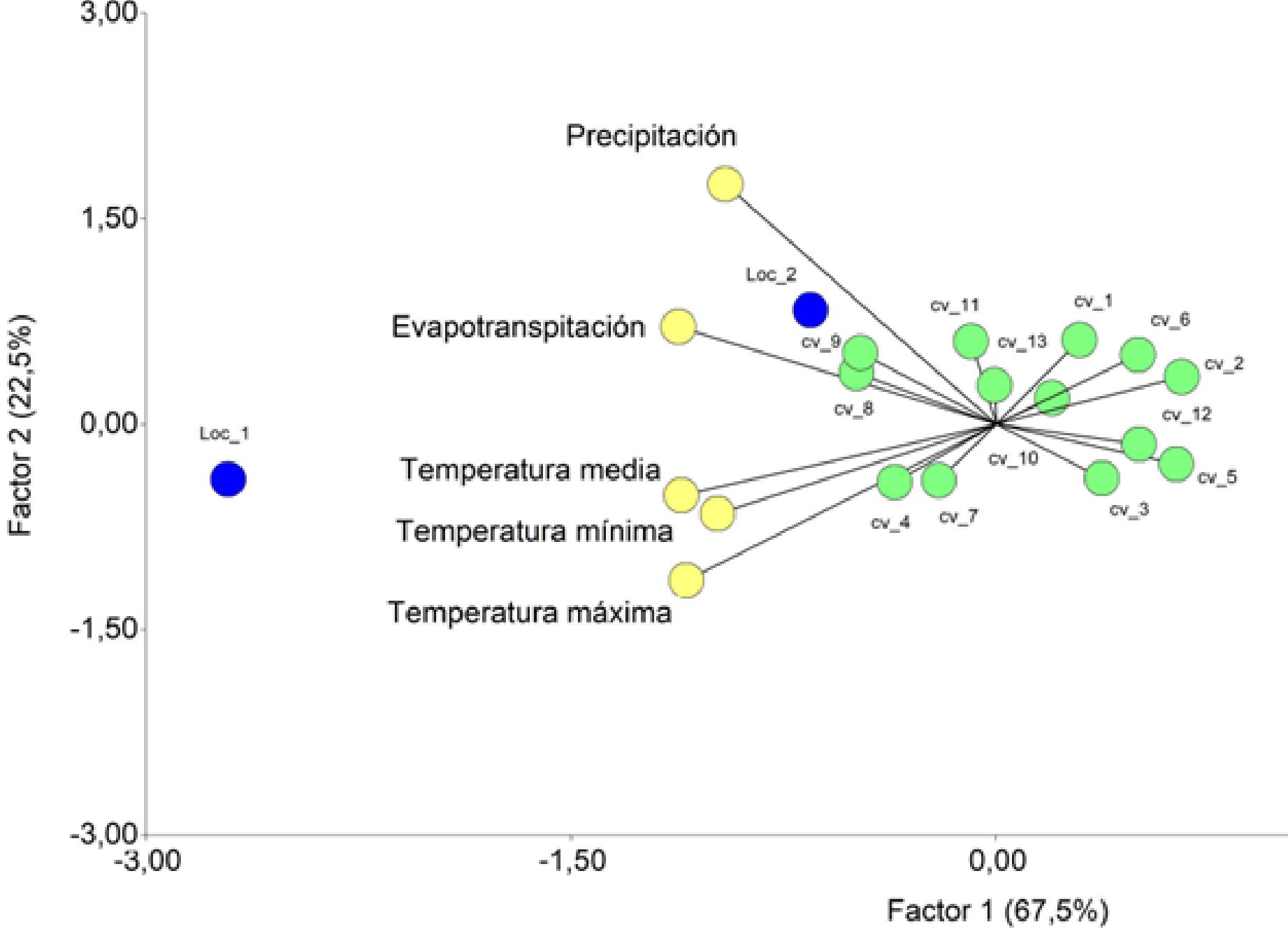
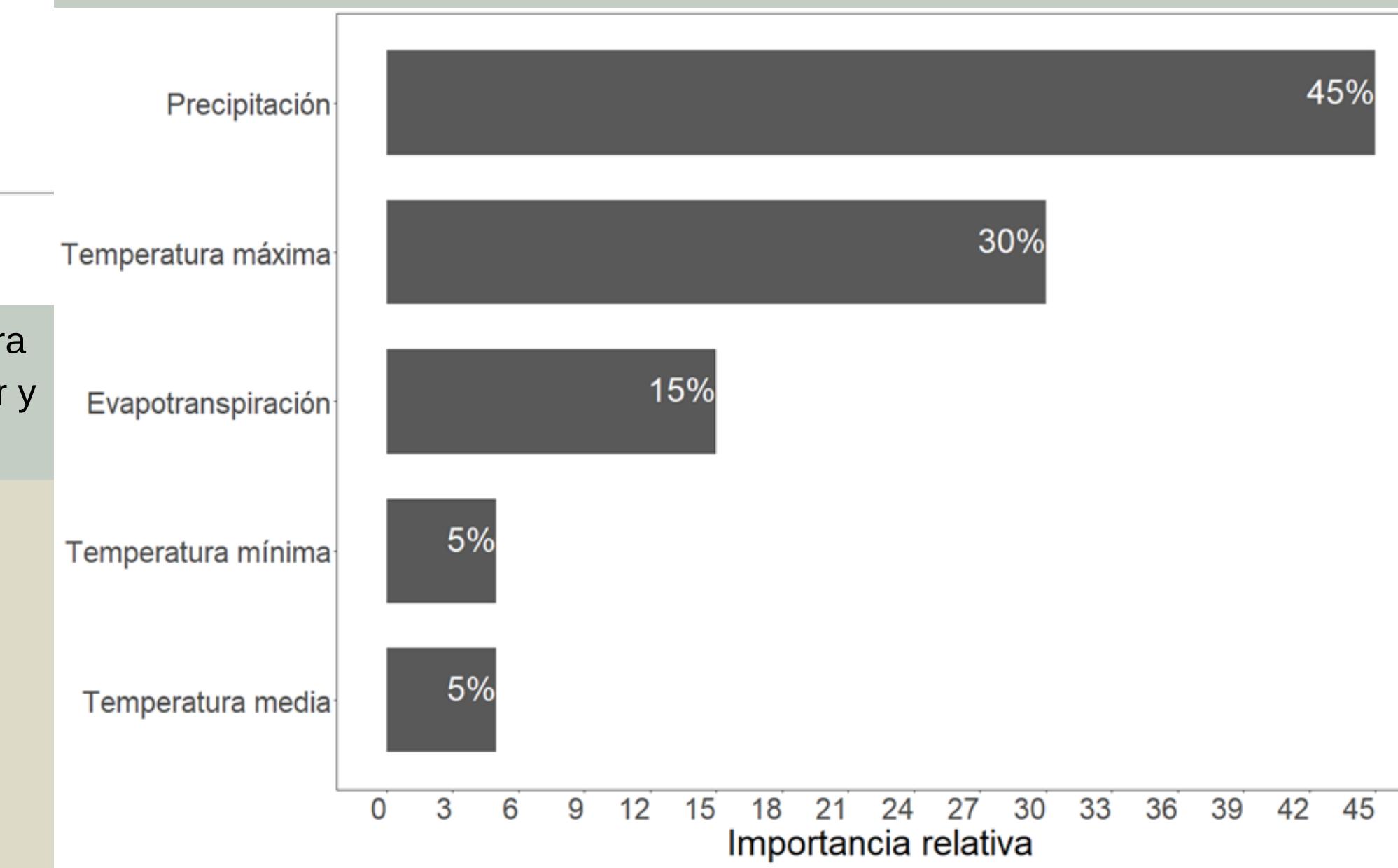


Figura 12: Biplot a partir de los resultados del análisis de PLS que demuestra los pesos de las variables climáticas para explicar la interacción entre cultivar y ambiente.



Figura 13: Importancia de las variables climáticas a partir del enfoque de aprendizaje automático para explicar la interacción entre cultivar y ambiente.



CONCLUSIÓN GENERAL

cultivar 2

Mayor contenido de ácido linolénico

cultivar 8

Menor contenido de ácidos grasos saturados

Influencia crítica de la temperatura y la precipitación

Optimizar los perfiles de ácidos grasos

cultivar 2

Producción de aceite de alta calidad

Estabilidad y rendimientos equilibrados.

cultivar 5

Aplicaciones culinarias, como la fritura

Agrupamientos

Cultivar 9

Cultivares 10, 2, 4, 3, 1 y 6

Cultivares 11, 12, 8, 13, 5 y 7

Localidad 1

cultivar 12

Localidad 2

cultivar 11

Localidad 3

cultivares 3 y 6

Localidad 4

cultivar 4

Muchas gracias

