



FACULTAD DE CIENCIAS AGRARIAS

UNIVERSIDAD NACIONAL DE ROSARIO

Caracterización genética mediante mapeo por asociación de la restauración de la androfertilidad en germoplasma templado de maíz (*Zea mays L.*) con citoplasma androestéril CMS-C

ING. AGR. LUCIANA CASTELLUCCIO

TESIS PARA OPTAR AL TÍTULO DE MAGISTER EN GENÉTICA VEGETAL

Director: ING. AGR. MSc. Alejandro Celani

Co-director: DRA. Ileana Gatti

Asesor Técnico: ING. AGR. MSc. Ignacio Torrent

Caracterización genética mediante mapeo por asociación de la restauración de la androfertilidad en germoplasma templado de maíz (*Zea mays* L.) con citoplasma androestéril CMS-C

Luciana Castelluccio

Ingeniera Agrónoma – Universidad Nacional de La Plata

Esta Tesis es presentada como parte de los requisitos para optar al grado académico de Magister en Genética Vegetal, de la Universidad Nacional de Rosario y no ha sido previamente presentada para la obtención de otro título en esta u otra Universidad. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en la República Argentina, durante el período comprendido entre 2019 y 2022, bajo la dirección de Ing. Agr. MSc. Alejandro Celani.

Ciertos detalles metodológicos, genómicos y a nivel de líneas han sido omitidos o generalizados para proteger la propiedad intelectual y los secretos comerciales, y para cumplir con las obligaciones de confidencialidad del patrocinador. Los datos detallados se conservan en registros confidenciales con el socio industrial y están disponibles sujeto a aprobación.

Ing. Agr. Luciana Castelluccio

Ing. Agr. MSc. Alejandro Celani.

Dra. Ileana Gatti

Defendida: 03 de Noviembre de 2026.

AGRADECIMIENTOS

Quisiera agradecer a mi director, codirector y asesor técnico, de quienes fue fundamental su ayuda, apoyo y guía durante todo este tiempo. Me ayudaron y redirigieron incluso cuando por circunstancias externas fue necesario cambiar el tema del proyecto y comenzar otro. Quisiera agradecer a la empresa, al equipo en el que trabajé y a quienes fueron mis líderes durante estos años por permitirme el tiempo y los recursos para poder realizar este trabajo. También a mi familia por su apoyo incondicional.

ÍNDICE

1. INTRODUCCIÓN	11
1.1. Mapeo de QTL y mapeo por asociación	19
1.2. Desequilibrio de ligamiento (DL)	20
1.3. Mapeo por asociación a través de la caracterización de marcadores homogéneamente distribuidos en el genoma (Genome-Wide Association Study: GWAS)	21
1.4. Modelos estadísticos usados para el análisis de asociación en plantas	21
1.5. Modelos estadísticos para el análisis de información fenotípica	22
1.6. Hipótesis	23
2. OBJETIVO	24
3. MATERIALES Y MÉTODOS	25
3.1. Material vegetal	25
3.2. Genotipado	25
3.3. Modelaje de la información genotípica	25
3.3.1. Filtrado de la matriz	25
3.3.2. Estructura genética poblacional: Análisis de componentes principales	25
3.3.3. Parentesco relativo	26
3.3.4. Desequilibrio de Ligamiento	26
3.4. Caracterización de la variabilidad fenotípica en la restauración de la androfertilidad	26
3.4.1. Ensayos a campo	26
3.4.2. Evaluación de la restauración de la androfertilidad	27
3.4.3. Análisis estadístico de la información fenotípica. Índices de Riesgo de ruptura para LPM-H y capacidad de restauración para LPM-M	28
3.5. Mapeo por asociación	32
3.6. Identificación de QTL	32
4. RESULTADOS	34
4.1. Caracterización molecular del panel de estudio	34
4.1.1. Estructura genética poblacional: Análisis de componentes principales	34
4.1.2. Parentesco relativo	38
4.1.3 Desequilibrio de ligamiento	40
4.2 Estudio de la restauración de la androesterilidad	44
4.2.1 Condiciones ambientales	44

4.2.2 Variabilidad fenotípica para la restauración de la androfertilidad en el panel LPM-H	44
4.2.3 Variabilidad fenotípica para la restauración de la androfertilidad en el panel LPM-M	46
4.3 Estudio de mapeo por asociación de la restauración de la androfertilidad con citoplasma androestéril CMS-C	48
4.3.1 Elección del modelo estadístico	48
4.3.2 Detección de QTL- Panel LPM-H	49
4.3.3 Detección de QTL- Panel LPM-M	56
5. DISCUSIÓN	57
6. CONCLUSIONES	65
7. BIBLIOGRAFÍA	67
8. ANEXO	77

ABREVIATURAS Y SÍMBOLOS

ACP: Análisis de componentes principales

AUDPC: Área Integrada Debajo de la curva de progreso de la enfermedad (del inglés; Area Under the Disease Progress Curve)

AUBPC: Área bajo la curva de progreso de ruptura (del inglés; Area Under the Breakage progress Curve)

BLUP: mejor predicción lineal insesgada (del inglés; Best Linear Unbiased Prediction)

CMS: Macho Esterilidad Citoplasmática (del inglés; Cytoplasmic Male Sterility)

CP: Componentes principales

DL: Desequilibrio de Ligamiento

ESA: Efecto de sustitución alélica

FAM: Frecuencia del alelo menor

GMS: Macho esterilidad génica (del inglés; Genetic Male Sterility)

GWAS: Mapeo por asociación a través de la caracterización de marcadores homogéneamente distribuidos en el genoma (del inglés; Genome-Wide Association Study)

IRR: Índice de Riesgo de Ruptura

K: Parentesco relativo

LPM-H: Panel de líneas parentales de maíz a ser utilizadas como hembras en la generación de híbridos.

LPM-M: Panel de líneas parentales de maíz a ser utilizadas como machos en la generación de híbridos.

Matriz_F: Matriz genotípica del conjunto de 251 hembras y 27882 SNPs

Matriz_M: Matriz genotípica del conjunto de 108 machos y 18690 SNPs

MEFF: Número efectivo de pruebas

MLG: Modelo Lineal General

MLM: Modelo Lineal Mixto

QTL: *loci* de caracteres cuantitativos (del inglés; Quantitative Trait Locus)

Rf: Restaurador de la fertilidad

SNP: Marcador molecular de polimorfismo de nucleótido único (Single-Nucleotide Polymorphism)

STR: Esterilidad de panoja

STR90_6: Esterilidad de panoja 6 días posteriores al momento de S90. Nota máxima del valor de ruptura alcanzado

RESUMEN

El uso de macho / andro esterilidad citoplasmática (CMS) tiene importantes ventajas económicas y logísticas en la producción de semilla híbrida de maíz (Fu et al., 2014). Es posible distinguir varios tipos de citoplasma androestériles en función de los genes nucleares específicos que restauran dicha fertilidad, entre ellos, el grupo CMS-C ha sido el más atractivo para la producción híbrida de maíz (Kheyr-Pour et.al., 1981). El objetivo de este trabajo fue caracterizar la variabilidad genética existente en el germoplasma del programa de mejoramiento de maíz híbrido de una empresa de semillas en Argentina para la restauración de la androfertilidad bajo citoplasma estéril CMS-C en un panel de 360 F1 (Donante CMS x líneas a ser usadas como machos y hembras en la generación de híbridos). Las líneas fueron genotipadas con 45.514 SNPs y su respuesta a la restauración de la androfertilidad fue evaluada con una escala de 1 (completamente estéril) a 9 (completamente fértil) en cinco momentos, comenzando con el 50% de estigmas visibles, en cuatro ambientes representativos de producción de semillas durante las campañas 2019/2020 y 2021/2022. La estructura poblacional se evaluó mediante un análisis de componentes principales en el cual las CP uno a diez explicó el 45,26% de la variación entre individuos para LPM-H y el 55,4% de la variación para LPM-M. Los coeficientes de parentesco relativo para la mayoría de las comparaciones entre pares de líneas fueron igual a cero (56,9% LPM-M y 59,9% LPM-H) indicando ningún tipo de relacionamiento o parentesco para estas comparaciones. El desequilibrio de ligamiento, medido como r^2 , estuvo en valores promedio de entre 0,40 y 0,42 en distancias superiores a 1.000.000 pb para ambos grupos heteróticos, indicando un nivel relativamente fuerte de desequilibrio de ligamiento entre los marcadores analizados.

El análisis fenotípico involucró la nota máxima del valor de ruptura alcanzado (STR90_6) para los LPM-M y el cálculo del Área Integrada Debajo del Polígono para LPM-H. Este último permitió integrar mediciones temporales de la restauración de la androfertilidad en un único índice (AUBPC). Las medias de AUBPC para LPM-H mostraron en los cuatro ambientes evaluados valores de 3 a 29, indicando variabilidad fenotípica en el panel estudiado que permitió identificar grupos de genotipos con alta, media y baja ruptura (31,8%, 21,5%, 46% respectivamente). Para LPM-M, las medias de STR90_6 en las cuatro localidades mostraron valores entre 0 y 2 para el 95,7% de los individuos indicando un alto nivel de esterilidad en todos los ambientes y baja variabilidad fenotípica.

El mapeo por asociación a través de la caracterización de marcadores homogéneamente distribuidos en el genoma (GWAS) involucró 27882 SNPs para LPM-H y permitió identificar 50 SNPs en cromosomas 1, 2, 3 y 8 que presentaron asociaciones significativas con la restauración de la androfertilidad ($p < 0,00067$), los cuales fueron agrupados en cinco QTL. Los QTL *Qrf8.1* y *Qrf2.2* reportados en este estudio se ubican en posiciones que ya han sido reportadas en estudios previos de restauración de la androfertilidad, mientras que *Qrf1.1* constituyó un falso positivo y *Qrf3.1* reportado en el presente trabajo en bin 3.11 no pareciera ubicarse en la misma región reportada por Kohls et al., 2011. En conjunto, *Qrf3.1* y *Qrf2.1* podrían representar modificadores de la restauración o loci adicionales no descritos previamente para CMS-C, aunque estudios posteriores mediante otras técnicas, como el ML, serán necesarios para su confirmación.

El GWAS para LPM-M no se reportó en este trabajo debido a las características del panel utilizado: solo el 5% de los individuos presentó restauración de la fertilidad, limitando la capacidad del modelo mixto para establecer relaciones robustas entre genotipo y fenotipo. Por el lado de LPM-H, un estudio más exhaustivo y sobre poblaciones segregantes F2 o F3 de cruza dirigidas será necesario para identificar los genes relacionados a estos QTL identificados, que potencialmente pudieran tener valor en técnicas de selección para el desarrollo de líneas puras CMS promisorias (estériles y estables) en un programa de mejoramiento.

Palabras clave: maíz; androesterilidad citoplasmática; CMS-C; restauración de la androfertilidad; GWAS; SNPs; QTL.

ABSTRACT

Genetic characterization through association mapping of the restoration of androfertility in temperate maize (*Zea mays L.*) germplasm with cytoplasmic male sterility CMS-C

Cytoplasmic male sterility (CMS) is a maternally inherited condition in which a plant is unable to produce functional pollen. Nuclear restoration allows the commercial exploitation of CMS systems in the production of hybrid maize seeds. The objective of this study was to characterize the genetic and phenotypic variation to the restoration of androfertility under sterile cytoplasm CMS-C in a panel of 360 lines related to hybrid corn program (females and males' heterotic groups) developed by the Breeding department of an Industry Partner in Argentina. A genome-wide association study, using 27882 SNPs identified 50 SNPs significantly associated to restoration of androfertility in females' heterotic group ($p < 0.0006$) which were grouped in five QTL. In the male's side, the limited phenotypic variation for restoration of androfertility hindered the identification of robust genotype-phenotype associations. *Qrf8.1* and *Qrf2.2* have correspondence to regions reported by other authors except for *Qrf3.1* and *Qrf2.1*. More detailed studies will be necessary to identify or confirm candidate genes associated with the QTL reported in females.

1. INTRODUCCIÓN

El maíz (*Zea mays*) es una gramínea perteneciente a la familia de las poáceas, originaria de Mesoamérica (Matsuoka et al, 2002), con 10 pares de cromosomas ($2n = 20$) (Mcclintock, 1929). Es una especie monoica de flores unisexuales; con inflorescencias masculinas y femeninas bien diferenciadas en la misma planta. Es el cereal con mayor volumen de producción, con más de 1.200M de toneladas a nivel global, adquiriendo un papel cada vez más importante en los sistemas agroalimentarios. Argentina produce en promedio más de 46M de toneladas y se ubica en el cuarto lugar entre los países con mayor producción, luego de Estados Unidos, China y Brasil (FAOStat, 2023).

La producción mundial de maíz ha experimentado un notable incremento en las últimas décadas, impulsado por la creciente demanda y por una combinación de avances tecnológicos, aumentos en el rendimiento y expansión de la superficie cultivada (Erenstein et al., 2022). Entre los factores que han contribuido a este proceso, el mejoramiento genético ha ocupado un lugar central, especialmente a través del desarrollo de híbridos. En este contexto, el concepto de vigor híbrido o heterosis se relaciona principalmente con las mayores ganancias de rendimiento de los híbridos, debido a su superioridad en biomasa, adaptabilidad, fertilidad y tolerancia al estrés biótico o abiótico en comparación con sus líneas parentales (Birchler, 2016). En términos generales, un híbrido puede definirse como la descendencia resultante de la cruce entre dos individuos genéticamente diferentes (Edera et al., 2019).

La explotación dirigida de la heterosis se ha consolidado como una de las estrategias más exitosas y de mayor impacto económico dentro del fitomejoramiento, particularmente en el cultivo de maíz (Xiao et al., 2021). Sin embargo, la creación de cultivos híbridos no constituye un procedimiento simple, ya que la producción de semillas híbridas en especies que se auto polinizan requiere eliminar el polen en uno de los parentales. En maíz, el esquema de híbridos de cruza simples fue propuesto inicialmente por Shull (1908, 1909) e East (1908); no obstante, los híbridos simples no fueron comercialmente exitosos en ese entonces principalmente debido al alto costo de la producción de semillas. La adopción comercial del maíz híbrido se consolidó posteriormente, cuando Jones (1918) propuso el uso de híbridos dobles, los cuales más tarde fueron desplazados por híbridos de tres vías y, finalmente, por híbridos simples, gracias al desarrollo de líneas puras superiores y de mayor rendimiento.

Como consecuencia de la expansión del uso del vigor híbrido como estrategia de mejoramiento, surgió también la necesidad de implementar mecanismos que facilitaran la producción eficiente

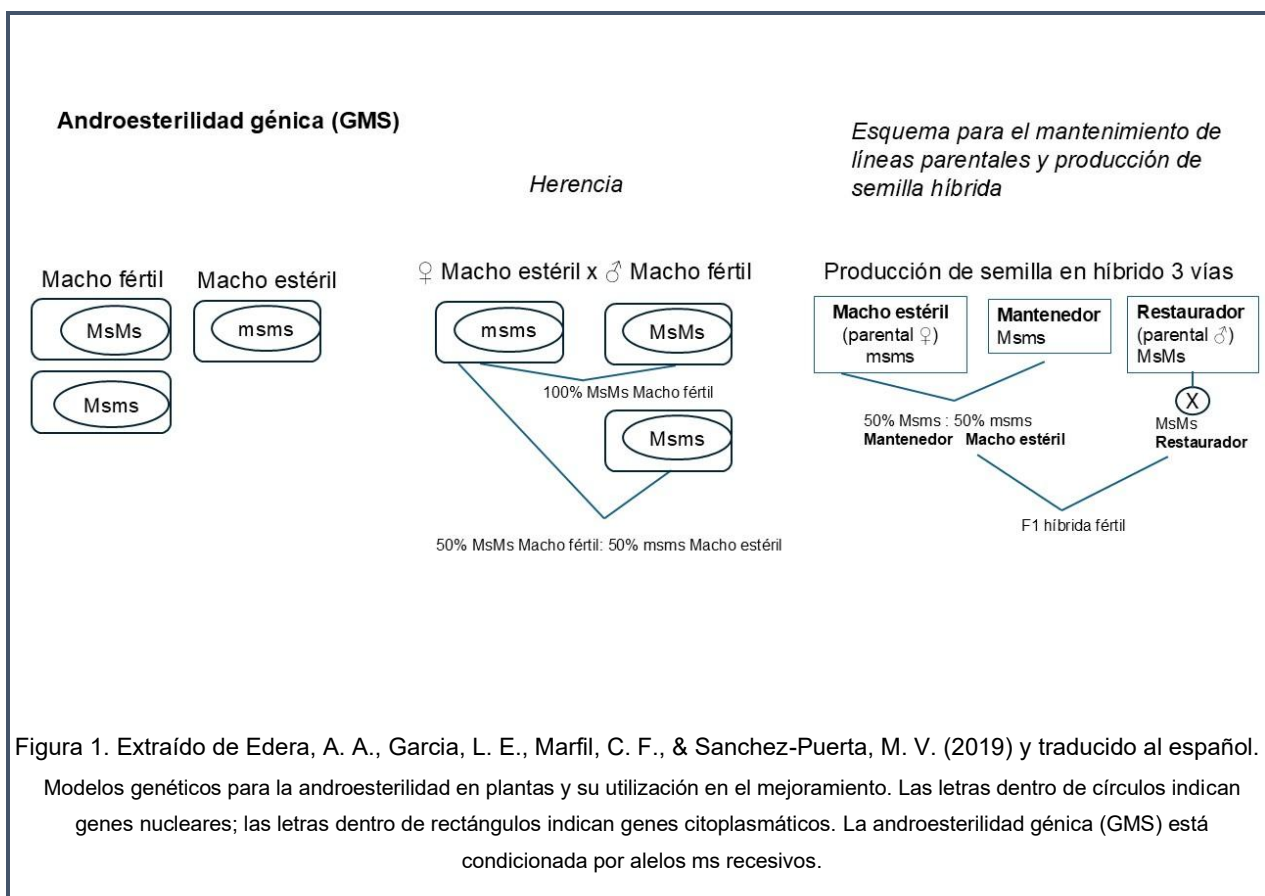
de semilla híbrida, entre ellos la generación de esterilidad masculina (Sage, 1976). La androesterilidad en plantas refiere a la falla en la producción de anteras dehiscentes, polen funcional o gametos masculinos viables (Edera et al., 2019).

La producción tradicional del maíz híbrido involucra la remoción de la inflorescencia masculina o el despanojado del parental que actuará como receptor de polen o “hembra” en la cruza, dejando solamente las panojas del parental que actuará como polinizador o “macho” para fertilizar los estigmas de la hembra. La tarea de despanojado en un lote de producción de semilla híbrida de maíz puede ocasionalmente dejar panojas en las plantas hembra con riesgo de autopolinización y el consecuente impacto en la pureza de la semilla resultante de ese lote. El despanojado mecánico es más rápido y económico frente al manual, aunque las plantas son dañadas durante el proceso, causando una reducción lineal del rendimiento en grano por cada hoja adicional removida con la panoja, que depende cómo y cuándo se realice dicha tarea (Wilhelm et al., 1995; Ghete et al., 2020). Hasta mediados del siglo XX, esta técnica involucró trabajo manual haciéndola costosa, ineficiente y perjudicial para el medio ambiente (Chen y Liu, 2014). En este sentido, el uso de esterilidad masculina reduce el costo de la producción de semillas híbridas evitando la castración y la polinización manual, permitiendo la producción a gran escala de semilla híbrida y la explotación comercial de vigor híbrido (Chen y Liu, 2014).

Dentro del grupo de los métodos alternativos al despanojado para generar esterilidad masculina en maíz se pueden incluir tanto aquellos basados en transgénesis, los que usan esterilidad génica (del inglés Genetic Male Sterility; GMS), los que combinan ambos, así como la macho o andro-esterilidad citoplasmática (del inglés Cytoplasmic Male Sterility; CMS). El sistema de hibridación Round Up (Round Up Hybridization system; RHS) es un método basado en transgénesis que reemplaza el despanojado convencional mediante la aplicación específica del herbicida glifosato en campos de producción de semillas híbridas. Las plantas macho que producen polen son resistentes al glifosato gracias a eventos transgénicos como NK603 o MON88017, mientras que las plantas hembra llevan el evento MON87427 (HAM027), que les impide producir polen viable (Feng et al., 2014). La aplicación se realiza durante la microsporogénesis tardía, en un rango de estadios de cultivo V7 a V10 (Ritchie et al., 1986), utilizando protocolos específicos para asegurar la eficacia.

En GMS, los genes de esterilidad masculina nuclear (*Ms*) controlan la esterilidad sin influencias citoplasmáticas (Edera et al., 2019). En el modelo más simple, hay tres genotipos posibles para el locus nuclear *Ms*, donde el fenotipo masculino estéril depende de alelos *ms* recesivos. Esto genera un patrón de herencia mendeliana, permitiendo que la descendencia de una línea estéril

sea completamente fértil o segregue 50% androestéril y 50% fértil, dependiendo de la homocigosidad de la línea parental (Figura 1). La tecnología de producción de semillas (Seed Production Technology; SPT) combina GMS con transgenes, permitiendo producir semillas de maíz homocigotas para un alelo masculino estéril recesivo. La versión original incluye el alelo dominante *Ms45*, que restaura la fertilidad en plantas homocigotas para el alelo recesivo *ms45*, y expresa un gen de α -amilasa que hace que el polen transgénico sea no viable. También se expresa la proteína DsRed, que permite identificar las semillas transgénicas (Wu et al., 2016). Una segunda versión del SPT utiliza el gen dominante *Ms44*, que permite la producción de líneas endogámicas no productoras de polen. Estas se cruzan con líneas fértiles, resultando en una segregación 1:1 para el alelo *Ms44*, donde la mitad de las plantas serán masculinas estériles, lo que puede aumentar el rendimiento si se produce suficiente polen (Fox et al, 2017). Este sistema también incluye un cassette de miARN para silenciar el gen *Ms44* y restaurar la fertilidad, manteniendo el marcador DsRed y el bloqueador de polen (α -amilasa) (Fox et al, 2017). Este método ha mostrado beneficios en sistemas agrícolas africanos de bajo insumo de nitrógeno (Collinson et al., 2022).



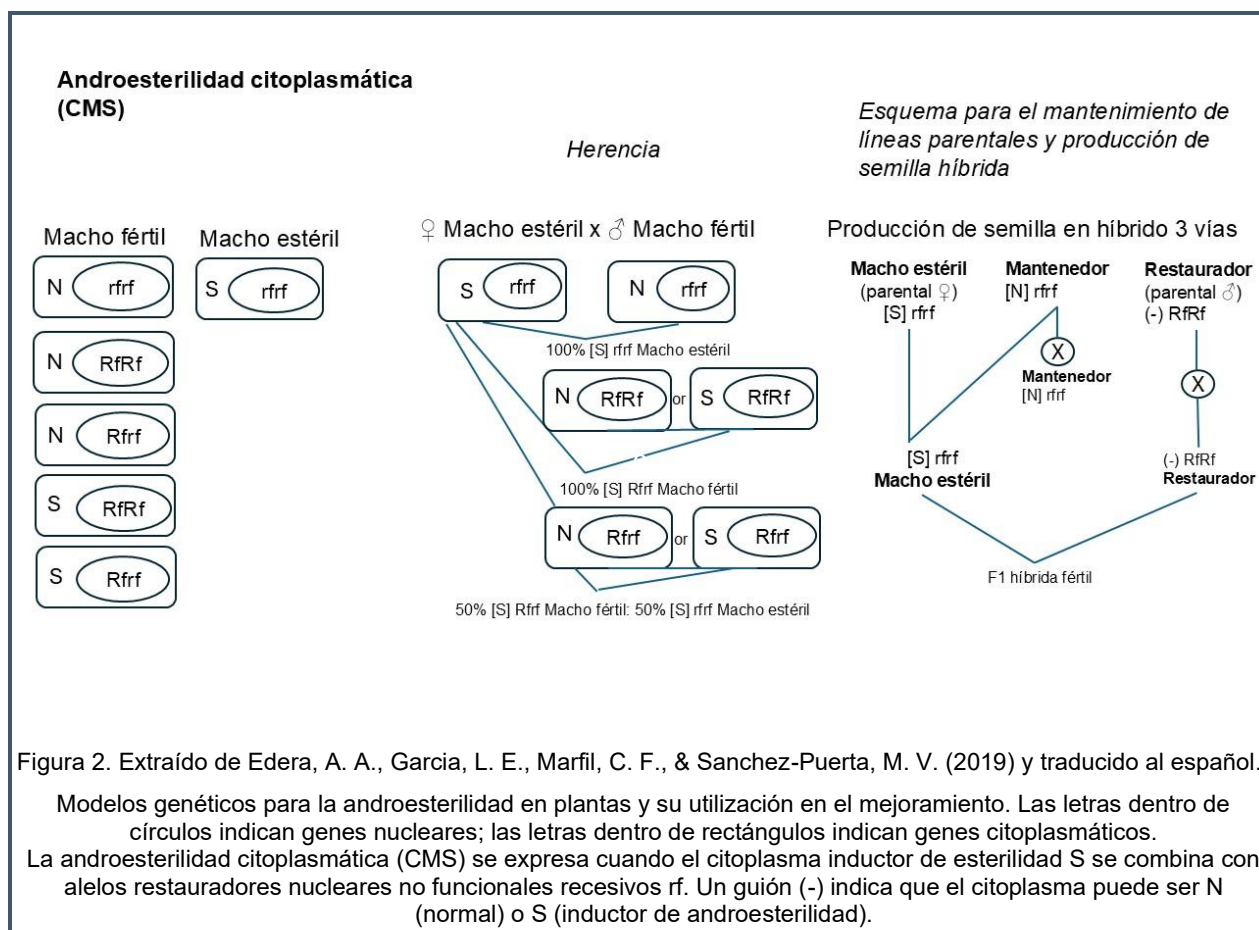
En CMS, la producción de polen no funcional tiene herencia materna y es condicionada por genes citoplasmáticos (mitocondriales) acoplados con genes nucleares (Chen y Liu, 2014). Cuando el gen mitocondrial *atp6c* está presente, la esterilidad CMS-C es conferida dado que la proteína con el mismo nombre muestra interacción con el ATP causando exceso de protones en el espacio intermembrana entre las mitocondrias y esto posteriormente desencadena una explosión de especies reactivas de oxígeno (ROS), muerte celular programada y aborto de polen (Yang et al., 2022). Para que haya restauración de la fertilidad, es necesaria la presencia de genes nucleares restauradores (*Rf*) que intervienen suprimiendo estas disfunciones (Schnable et al., 1998). Estos genes restauradores actúan mediante mecanismos post-transcripcionales, promoviendo la degradación selectiva o el procesamiento del transcrito aberrante *orf355–orf77*. Sin embargo, no es suficiente con la presencia de un *Rf* dominante, ya que su efecto puede verse suprimido o modulado por genes modificadores nucleares o por interacciones epistáticas núcleo-citoplasma (Huang et al., 2013).

En poblaciones naturales comúnmente hay dos mitotipos alternativos en una sola población, uno normal (generalmente designado N) y el inductor de esterilidad masculina (designado S). En el trabajo de Edera et al. (2019), se describe cómo el mitotipo S interactúa con un par de alelos nucleares: un restaurador de la fertilidad (si es dominante, por lo general se designa como *Rf*) y un no restaurador (si es recesivo usualmente designado alelo *rf*). En el modelo genético más simple, seis combinaciones mitotipo-genotipo son posibles, de las cuales solo una conduce a un fenotipo masculino estéril (Figura 2). La descendencia de la línea estéril masculina (línea femenina) podría ser completamente androestéril, completamente fértil, o segregar 50% androestéril: 50% fértil, dependiendo si el parental fértil masculino es homocigota recesivo, homocigota dominante o heterocigota para el locus nuclear restaurador de la fertilidad, respectivamente (Edera et al., 2019). El valor de mejoramiento de CMS depende del manejo de tres líneas diferentes:

- i. macho estéril (progenitor femenino, receptor del polen o hembra)
- ii. mantenedor,
- iii. restaurador (progenitor masculino, polinizador o macho).

La línea androestéril se perpetúa a través de cruces con la línea mantenedora, que es isogénica y difiere solamente en la presencia del citoplasma N. En contraste con GMS, la cruce entre el macho estéril y la línea mantenedora produce sólo descendientes androestériles. Finalmente,

para aquellos cultivos cuyas semillas se cosechan y comercializan, la fertilidad masculina debe ser restaurada en los híbridos F1. La línea restauradora tiene alelos dominantes restauradores de la fertilidad Rf y produce híbridos F1 fértiles. Como el citoplasma se hereda por vía materna, el mitotipo de la línea restauradora es irrelevante (Figura 2).



En relación con los tipos de citoplasma estériles reportados, destacan 3 fuentes principales en maíz, T (Texas) (Rogers and Edwardson, 1952), S (USDA) (Jones et al., 1957), y C (Charrúa) (Beckett, 1971) clasificados en función de reacciones diferenciales en la restauración de la fertilidad (Huang, 2012). El uso de CMS-T en la producción de semilla de maíz en Estados Unidos fue tan exitoso que, durante 1970, más del 85% del maíz producido llevaba CMS-T. Su uso fue discontinuado abruptamente cuando ocurrió la epidemia de Tizón de la hoja causada por el hongo *Bipolaris maydis* (también conocido como *Cochliobolus heterostrophus* en su estado teleomorfo), que fue especialmente virulento en plantas que contenían CMS-T (Laughnan et al., 1983). En cuanto a la estabilidad de esterilidad (aspecto clave para la producción de semillas), se han

comparado genotipos con los 3 tipos de CMS y aquellos que llevaron el tipo T fueron los más estables en los distintos ambientes, mientras que los de citoplasma del tipo S a menudo mostraban restauración parcial de la fertilidad (Weider et al., 2009). En el tipo citoplasma tipo S, dado que el mecanismo de esterilidad ocurre después de la microsporogénesis (Lee et al., 1980), se puede liberar polen no viable, y una planta completamente estéril puede presentar extrusión de anteras (Duvick, 1965). En otras palabras, el CMS-S es de tipo gametofítico, lo que significa que la restauración ocurre en el gameto en lugar de en las células del tejido diploide del esporofito. Por lo tanto, algunos gametos de una planta heterocigota para un factor de restauración serán viables, mientras que otros no lo serán. Por consecuencia, la extrusión puede generar la impresión de que una planta es fértil, lo que puede llevar a los productores a despanojar, reduciendo así el rendimiento en la producción de semilla híbrida y adicionalmente obtener híbridos con una fracción de polen de viabilidad reducida. Por otro lado, CMS-T y CMS-C corresponden a tipos esporofíticos de esterilidad masculina citoplasmática. En estos sistemas, cuando los factores de restauración de la fertilidad están presentes en la planta, aun en estado heterocigoto, se produce polen viable. A diferencia del tipo CMS-S, la esterilidad en CMS-T y CMS-C se manifiesta de manera más evidente, ya que las anteras no se extruyen, y las plantas restauradas producen polen en forma equivalente a su versión fértil. Debido a esta mayor claridad fenotípica y a la restauración efectiva de la fertilidad, los tipos esporofíticos han sido preferidos en los programas de mejoramiento. Sin embargo, la susceptibilidad de CMS-T a *Bipolaris maydis* limitó su utilización, por lo que CMS-C pasó a constituir el tipo de CMS más extendido en maíz (Jaqueth et al., 2020). A modo de síntesis, en la Tabla 1 se resumen las principales características, ventajas y limitaciones de los sistemas CMS-S, CMS-T y CMS-C.

A pesar de estas ventajas, la adopción masiva del sistema CMS-C ha estado restringida por diversas limitaciones. Si bien este sistema representa una alternativa eficiente para la producción de semilla híbrida, su restauración de la fertilidad aún no se comprende por completo, y su uso no resulta eficaz en una amplia proporción del germoplasma comercial de maíz. Esta eficacia depende de la presencia o ausencia de genes nucleares restauradores específicos para cada tipo de citoplasma, así como también depende de si ese germoplasma será utilizado como parental femenino o masculino. Además, en determinadas condiciones ambientales, la CMS puede no mantener completamente la esterilidad masculina (Weider et al., 2009), lo que puede ocasionar cierta producción de polen en el parental femenino CMS y, en consecuencia, reducir la pureza y la productividad de los lotes de semilla híbrida (Wu et al., 2016). En maíz, los mecanismos de aborto y restauración de la fertilidad en CMS-C se asocian con la expresión de

genes mitocondriales quiméricos, compuestos por segmentos cortos derivados de distintas regiones del genoma mitocondrial que se ensamblan para formar nuevos genes codificantes de proteínas (Gabay-Laughnan et al., 2004).

La fertilidad en CMS-C puede ser restaurada por múltiples genes restauradores. *Rf4* es un gen restaurador dominante que es capaz de restaurar completamente la fertilidad en CMS-C, mapeado en el cromosoma 8 (Kheyr- Pour et al., 1981). Un segundo gen dominante restaurador de la fertilidad, el *Rf5* se mapeó en cromosoma 5 y es inhibido por *Rf-I*, mapeado en cromosoma 7. *Rf-I* no inhibe la capacidad de restauración de *Rf4*. Un gen restaurador adicional fue descubierto llamado Rf-A619 (Yongming et al., 2016).

Las influencias ambientales en los patrones de esterilidad fueron investigadas en varias especies. En maíz, la estabilidad de la androesterilidad puede verse afectada por las condiciones ambientales, particularmente por la temperatura y la humedad relativa. En términos generales, altas humedades relativas y bajas temperaturas se han asociado con incrementos en la fertilidad o con la ruptura de la esterilidad, mientras que temperaturas elevadas y fotoperiodos cortos tienden a favorecer el mantenimiento de la esterilidad masculina (Weider et al., 2009; Bueckmann et al., 2014, 2016). Asimismo, se ha propuesto que estos efectos ambientales podrían estar vinculados con la sensibilidad de los genes restauradores de la fertilidad *Rf* a señales previas a antesis, lo que afectaría su activación o supresión (Weider et al., 2009; Bueckmann et al., 2014). En conjunto, estos antecedentes indican que la expresión de la androesterilidad en maíz no depende exclusivamente del fondo genético, sino también de su interacción con el ambiente.

Tabla 1. Comparación de sistemas CMS en maíz

Tipo de citoplasma	Momento donde se interrumpe la formación de polen	Fenotipo asociado a la esterilidad a campo	Estabilidad de la esterilidad	Estado de uso actual
CMS-S	Gametogénesis: El colapso se manifiesta en el polen/microsporo tardío.	Puede mostrar esterilidad incompleta o variable, las anteras pueden extruirse y liberar polen no viable complicando la pureza de un lote de producción.	Baja a media; es el más inestable de los tres. Su esterilidad puede verse afectada por fondo genético y el ambiente.	Uso limitado en producción comercial por su inestabilidad y por quiebres de esterilidad en algunos fondos nucleares o ambientes.
CMS-C	Esporogénesis: La alteración ocurre principalmente durante el desarrollo postmeiótico del microsporo, asociada a fallo del tapete y degradación del microsporo hacia la fase de gran vacuola/microsporo tardío.	En general presenta antera poco o nada exerta, con ausencia de liberación de polen	Media a alta, pero no absoluta. En comparaciones recientes suele ser más estable que CMS-S, aunque revisiones indican que CMS-C y CMS-S no son tan estables como CMS-T.	Es el más usado actualmente entre los sistemas citoplasmáticos clásicos para producción de híbridos, aunque no es el único sistema de esterilidad empleado en la práctica moderna.

Tabla 1 (Continuación). Comparación de sistemas CMS en maíz

Tipo de citoplasma	Momento donde se interrumpe la formación de polen	Fenotipo asociado a la esterilidad a campo	Estabilidad de la esterilidad	Estado de uso actual
CMS-T	Esporogénesis: El defecto se asocia al tapete y al desarrollo postmeiótico del polen.	Suele verse como panojas sin anteras exertas o sin liberación de polen; históricamente fue muy eficiente para evitar despanojado.	Alta, pero su gran problema no fue el quiebre de esterilidad sino la vulnerabilidad fitosanitaria. Revisiones comparativas indican que CMS-T es más estable que CMS-C y CMS-S.	Prácticamente en desuso por su histórica susceptibilidad a tizón foliar del sur por <i>Bipolaris maydis</i> raza T.

1.1. Mapeo de QTL y mapeo por asociación

A partir del uso de marcadores moleculares a través de los años, se han encontrado extensos ejemplos de cómo la variación fenotípica de caracteres de importancia evolutiva y agronómica está determinada por múltiples QTL, la interacción entre éstos, el ambiente y las interacciones QTL-ambiente.

Los análisis de ligamiento en familias o poblaciones biparentales y, más recientemente, los análisis de asociación son las dos estrategias más utilizadas a la hora de disectar caracteres de base genética compleja. Ambos análisis se basan en la coheredabilidad de los polimorfismos funcionales o variantes alélicas vecinas. En los análisis de ligamiento tradicionales, son pocos los eventos de recombinación que llegan a producirse, por ejemplo, en los descendientes de un cruzamiento biparental, resultando en un mapeo de baja resolución. En el mapeo por asociación, en cambio, la recombinación y la diversidad genética natural se analizan para lograr mapeo de alta resolución (Zhu et al., 2008). Es decir, esta estrategia tiene el potencial de alcanzar resolución a nivel de genes individuales, dado que hace uso de los numerosos eventos de recombinación ocurridos en el seno de las poblaciones a lo largo de su historia, permitiendo así detectar aquellas

formas alélicas involucradas en la variación de un determinado carácter fenotípico (Nordborg y Tavaré, 2002). En el caso de las plantas, las poblaciones analizadas pueden ser verdaderas poblaciones naturales, colecciones de germoplasma, líneas o variedades elite de un programa de mejoramiento (Breseghello y Sorrells, 2006). Los estudios de mapeo por asociación se presentan, entonces, como una herramienta novedosa de gran potencial para el mejoramiento agronómico, cuya aplicación involucra la interacción de disciplinas tales como genómica, genética estadística, genética de poblaciones, biología molecular y bioinformática.

1.2. Desequilibrio de ligamiento (DL)

El DL se refiere a asociaciones no aleatorias entre los alelos de distintos loci y constituye la base conceptual de cualquier estudio de mapeo por asociación (Yu et al., 2006). La distancia sobre la cual el DL persiste va a determinar el número y la densidad de marcadores, así como el diseño experimental que se ajuste mejor para el mapeo por asociación. Debido a ello, es importante conocer en primera instancia el DL y determinar el alcance de este en la especie de interés (Ersoz et al., 2007).

Las dos medidas comúnmente usadas para cuantificar DL son el coeficiente de desequilibrio estandarizado (D') (Lewontin, 1964) y la correlación de frecuencias alélicas al cuadrado (r^2) (Hill y Weir, 1994), siendo esta última la más empleada por presentar menor varianza en muestras pequeñas. El análisis del DL ha sido materia de estudio de numerosos grupos de investigación desde hace una década, demostrando que la variación que existe entre las distintas especies es significativa. Una amplia variedad de mecanismos, que pueden operar simultáneamente, afecta el DL, siendo la mutación, la recombinación, el sistema de apareamiento, la estructura críptica de la población y la distribución de frecuencias alélicas, los factores más importantes (Flint-Garcia et al., 2003). La mutación genera el material de partida para la creación de sitios polimórficos que estarán en DL. En cambio, la recombinación es el principal factor que provoca una disminución del DL intra-cromosómico, y la segregación independiente disminuye el DL inter-cromosómico (Flint-Garcia et al., 2003). El sistema de apareamiento también influye los niveles de DL, por ejemplo, en especies de reproducción autógena o de autopolinización el DL tiende a ser mayor que en especies de polinización cruzada (Flint-Garcia et al., 2003).

Una de las restricciones más importantes para la aplicación del mapeo por asociación en plantas cultivadas, es la presencia de estructuración críptica en la población debido a factores como la domesticación y el mejoramiento (Thornsberry et al., 2001; Wright y Gaut, 2005 & Rafalski, 2010).

El DL aumenta en una población estructurada cuando las frecuencias alélicas para dos loci difieren entre las subpoblaciones que la integran, independientemente del ligamiento que exista entre ellos. Un caso extremo de estructuración es el de una población “mezcla” creada a partir de dos poblaciones naturales, provocando la mezcla de cromosomas de diferentes linajes y con diferentes frecuencias alélicas (Pritchard y Przeworski, 2001). Frecuentemente esto provoca que se observe DL significativo aun cuando los loci considerados no estén físicamente ligados. Sin embargo, este DL desaparece, generalmente, luego de unas pocas generaciones de apareamiento aleatorio (Pritchard y Przeworski, 2001). La presencia de estructura en la población utilizada en mapeo por asociación puede llevar a asociaciones espurias, si ésta no es tomada en cuenta en el análisis (Oraguzie et al., 2007).

1.3. Mapeo por asociación a través de la caracterización de marcadores homogéneamente distribuidos en el genoma (Genome-Wide Association Study: GWAS)

Esta modalidad involucra una exploración exhaustiva del genoma para la localización de variación genética causativa. Un gran número de marcadores uniformemente distribuidos en el genoma (10.000 a 100.000) son analizados para encontrar asociaciones con varios caracteres complejos, sin requerir ningún tipo de información previa acerca del control genético del carácter. En el enfoque de GWAS, las colecciones de los genotipos exhiben un desequilibrio de ligamiento mucho más limitado entre pares de marcadores vecinos (Hyten et al., 2007). Como resultado, se requiere un número mucho mayor de marcadores para proporcionar cobertura adecuada de todo el genoma.

1.4. Modelos estadísticos usados para el análisis de asociación en plantas

El éxito final en el mapeo por asociación dependerá en la posibilidad de separar los eventos de DL reales, de los causados por otros factores. Es por ello por lo que, dado que la estructura poblacional es una causa importante de detección de asociaciones espurias, la mayoría de los análisis la incorporan. Los modelos más simples de asociación son los modelos lineales, que asignan un factor de efecto fijo a los marcadores a evaluar y un factor de efecto aleatorio a la presencia de estructura poblacional (Bresnahan y Sorrells, 2006). Algunos estudios incorporan

la información de las poblaciones ideales obtenidas mediante el método bayesiano implementado en el programa STRUCTURE (Pritchard et al., 2000), y la utilizan para asignar genotipos dentro de los análisis de asociación marcador-fenotipo (Remington et al., 2001, Simko et al., 2004b, Skøt et al., 2005). El análisis de componentes principales (ACP) también puede incluirse como covariable en el modelo de asociación en lugar de las subpoblaciones obtenidas de STRUCTURE (Price et al., 2006). ACP se basa en la reducción de dimensionalidad para extraer los valores propios y vectores de todos los marcadores genéticos. Los gráficos de pares de componentes principales y su tendencia se utilizan para indicar la estructura relativa de la población (Torkamaneh et al., 2022).

Por otro lado, la matriz de parentesco relativo (K) permite obtener un coeficiente que se define como la probabilidad de que dos alelos homólogos extraídos de cada uno de dos individuos sean idénticos por descendencia (IBD), y es una medida clásica de parentesco (Speed y Balding, 2015).

Finalmente, varios modelos incorporan ambos datos, tanto de estructura poblacional como las relaciones de parentesco entre individuos (K), en los denominados modelos lineales mixtos de asociación (Yu et al., 2006).

1.5. Modelos estadísticos para el análisis de información fenotípica

Para estas aproximaciones, los modelos lineales mixtos presentan utilidad en la interpretación de la información (Schandry, 2017) así como para la obtención de los BLUPs para cada genotipo, habiéndose ya usado en la evaluación de resistencias a enfermedades para diferentes cultivos y patógenos (Khanal et al., 2022; Zakieh et al., 2023). En las mediciones que se realizan sobre genotipos en ensayos de campo es frecuente encontrar datos desbalanceados debido a pérdidas de parcelas, números de repeticiones desiguales o la presencia de múltiples ambientes con distinta información disponible. En estos casos, el uso de la mejor predicción lineal insesgada o BLUPs (Best Linear Unbiased Predictions) resulta especialmente ventajoso frente a la mejor estimación lineal insesgada (BLUEs), ya que los BLUPs incorporan la estructura de varianzas y covarianzas del modelo mixto o aleatorio y aplican un proceso de “encogimiento” (shrinkage) que reduce tanto el sesgo como la varianza de las predicciones (Henderson, 1975; Henderson, 1984). En el contexto del mejoramiento vegetal, diversos autores han demostrado que los BLUPs ofrecen estimaciones más precisas y robustas que las medias ajustadas, particularmente cuando

los datos son desbalanceados o incompletos (Piepho et al., 2008; Möhring y Piepho, 2009). Por ello, el enfoque basado en modelos mixtos/aleatorios y BLUPs se ha convertido en la metodología estándar para el análisis de ensayos multi-ambiente y la predicción del valor genético de genotipos (Gilmour et al., 1997; Cullis et al., 2006).

1.6. Hipótesis

Existen regiones genéticas nucleares asociadas a la restauración de la androesterilidad citoplasmática para CMS-C, las cuales pueden ser identificables mediante un estudio de mapeo por asociación que involucre sets de marcadores moleculares y datos fenotípicos en germoplasma de maíz templado.

2. OBJETIVO

El objetivo de esta tesis fue caracterizar la variabilidad genética existente en el germoplasma del programa de mejoramiento de maíz de una empresa de la industria semillera de Argentina para la restauración de la androfertilidad bajo citoplasma estéril CMS-C. Los objetivos específicos fueron:

- (i) Caracterizar la variabilidad fenotípica en la restauración de la androfertilidad para el germoplasma estudiado;
- (ii) Caracterizar la estructura genética de la población estudiada;
- (iii) Identificar mediante la aplicación de la metodología GWAS regiones genéticas asociadas a la restauración de la androfertilidad en el germoplasma estudiado.

3. MATERIALES Y MÉTODOS

3.1. Material vegetal

El panel de estudio del proyecto fue constituido por 360 líneas puras desarrolladas por el Programa de Mejoramiento Templado de una empresa de la industria semillera de Argentina (251 líneas de maíz a ser utilizadas como hembras en la generación de híbridos (LPM-H) y 109 líneas de maíz a ser utilizadas como machos en la generación de híbridos (LPM-M)). Estas líneas representan fielmente la diversidad genética presente en ambos grupos heteróticos de este programa de Mejoramiento para dicho programa. Para la caracterización de la restauración de la fertilidad CMS-C en germoplasma de Argentina estas líneas se cruzaron con un donante estéril conocido con ausencia de *Rf* (CMS-C *rfrf*) y se obtuvieron 360 F1 (primera generación filial) CMS.

3.2. Genotipado

Las líneas incluidas en el presente estudio fueron genotipadas con 45.514 marcadores del tipo SNP utilizando la plataforma MaizeSNP50 BeadChip (Illumina Inc., San Diego, CA, USA).

3.3. Modelaje de la información genotípica

3.3.1. Filtrado de la matriz

El filtrado de la matriz de datos se realizó en diferentes pasos y por grupo heterótico utilizando el paquete ASRGenomics (Gezan et al., 2022), mediante el cual se removieron marcadores con el 20% o más de datos faltantes, se descartaron aquellos SNPs con frecuencia del alelo menor (FAM) por debajo del 5% (17629 marcadores para el panel hembras, 26821 marcadores para el panel machos) y aquellos cuya heterocigosidad resultara mayor al 95%.

3.3.2. Estructura genética poblacional: Análisis de componentes principales

Se realizó un análisis de componentes principales (ACP) a partir del conjunto de 27882 y 18690 SNPs (Matriz_F y Matriz_M, respectivamente) con el software TASSEL v5 (Bradbury et al., 2007), y las primeras diez componentes principales (CP) fueron graficadas mediante diagramas de dispersión con el objetivo de identificar patrones de agrupamiento entre individuos. Se optó trabajar con diez componentes principales en función de la proporción de la variancia explicada para ambos grupos heteróticos.

3.3.3. Parentesco relativo

Se obtuvieron dos matrices de parentesco relativo (K) para el conjunto Matriz_F y Matriz_M utilizando el software Tassel v5 (Bradbury et al., 2007).

3.3.4. Desequilibrio de Ligamiento

Para el presente estudio, se calculó el desequilibrio de ligamiento (DL) entre marcadores genéticos con el objetivo de caracterizar la arquitectura genética de la población, evaluar la extensión y el decaimiento del DL, estimar la resolución esperada del estudio y contribuir a la identificación y localización de loci asociados a rasgos de interés en el análisis de asociación del genoma completo (GWAS).

El cálculo del DL se realizó utilizando la Matriz_M y la Matriz_F junto con el análisis del coeficiente de determinación r^2 en el software TASSEL v5 (Bradbury et al., 2007) como medida del DL, que cuantifica la fuerza de la asociación entre pares de SNPs. Los valores de r^2 fueron calculados para cada cromosoma y graficados contra la distancia genética (pb).

3.4. Caracterización de la variabilidad fenotípica en la restauración de la androfertilidad

3.4.1. Ensayos a campo

Se sembraron 200 híbridos CMS durante la campaña 19-20 en parcelas de 1 surco de 5 m de largo, distanciados a 52 cm, en una repetición, con un diseño en bloques completos aleatorizados en dos ambientes. En la campaña 21-22 se sembraron 206 híbridos CMS en dos ambientes (Tabla 2), con el mismo diseño experimental, pero con dos repeticiones. De los híbridos evaluados en el segundo año, 46 fueron genotipos que ya habían sido evaluados en 19-20.

Tabla 2. Detalle de los sitios de testeo

Localidad	Campaña	Código	Fecha de siembra	Fecha Promedio de Floración
Buenos Aires Norte A	1	1 ^a	27-Nov-19	5-Feb-20
Buenos Aires Norte B	1	1B	11-Nov-19	16-Jan-20
Buenos Aires Norte A	2	2 ^a	18-Nov-21	24-Ene-22
Buenos Aires Costa	2	2B	3-Nov-21	18-Ene-22

La siembra se realizó en condiciones de densidad óptima dirigidas a la producción de semillas híbridas y los ensayos se gestionaron bajo condiciones de fertilidad y humedad no limitantes (velando mantener el suelo a capacidad de campo) dado que la producción de semilla del programa templado bajo estudio se realiza 100% bajo riego, por lo que el interés es trabajar sobre condiciones representativas del ambiente objetivo. Por último, las aplicaciones de herbicidas e insecticidas se realizaron según necesidad para mantener condiciones óptimas de cultivo.

3.4.2. Evaluación de la restauración de la androfertilidad

Las evaluaciones de esterilidad en la panoja fueron realizadas en todos los ambientes con una escala de 1 (completamente estéril, sin anteras) a 9 (100% de plantas con al menos una antera visible liberando polen) en cinco momentos (STR50, STR90, STR90_3, STR90_6, STR90_9) detallados en la Tabla 3 y 4, estandarizada y consistente como práctica común de la industria de semillas.

Durante ambas campañas, se evaluó la fecha de floración de cada híbrido en dos momentos: (i) cuando el 50% de las plantas de las parcelas presentaron algún estigma visible (S50D); (ii) cuando el 90% de las plantas de las parcelas presentaron algún estigma visible (S90D).

Tabla 3. Escala de esterilidad

Escala	Descripción
1	Completamente estéril (sin anteras).
2	Hasta 3 o 4 plantas con pocas anteras visibles en sus panojas liberando polen.
3	Desde 3 o 4 plantas hasta un 25% de plantas con al menos una antera visible liberando polen.
4	25-50% de plantas con al menos una antera visible liberando polen.
5	50-60% de plantas con al menos una antera visible liberando polen.
6	60-70% de plantas con al menos una antera visible liberando polen.
7	70-80% de plantas con al menos una antera visible liberando polen.
8	80-90% de plantas con al menos una antera visible liberando polen.
9	100% de plantas con al menos una antera visible liberando polen.

Tabla 4. Momentos de evaluación

Momento	Descripción
STR50	Esterilidad de panoja en S50 (cuando el 50% plantas muestran algún estigma visible)
STR90	Esterilidad de panoja en S90 (cuando el 90% de las plantas muestran algún estigma visible)
STR90_3	Esterilidad de panoja 3 días posteriores al momento de S90
STR90_6	Esterilidad de panoja 6 días posteriores al momento de S90
STR90_9	Esterilidad de panoja 9 días posteriores al momento de S90

3.4.3. Análisis estadístico de la información fenotípica. Índices de Riesgo de ruptura para LPM-H y capacidad de restauración para LPM-M

3.4.3.1. LPM-H

Para la caracterización de líneas en cuanto a su ruptura tradicionalmente se asigna la puntuación STR de 1 a 9 acumulativa en distintos intervalos temporales relacionados con la emisión de estigmas, comenzando cuando el 50% de las plantas exhiben estigmas visibles y continuando hasta seis días después de que el 90% de las plantas los presenten como se detalla en la Tabla 3 y 4 de la sección 3.4.2. Esta metodología asegura que las puntuaciones sean acumulativas, donde una puntuación subsiguiente no puede ser inferior a la anterior, y la puntuación final se determina en la última evaluación. Sin embargo, este enfoque presenta limitaciones, especialmente en la pérdida de resolución temporal para la comparación de genotipos. Dicha resolución temporal es importante desde el punto de vista de producción de semillas, ya que una línea que presenta ruptura temprana es más riesgosa que una que presente ruptura tardía por la probabilidad de autofecundación.

Para superar estas limitaciones y refinar la caracterización, se propuso en este trabajo utilizar un nuevo enfoque, el Área Integrada Debajo del Polígono, inspirado en el concepto utilizado para la progresión de enfermedades (Area Under the Disease Progress Curve: AUDPC). Un consolidado de aplicaciones y variantes del cálculo pueden encontrarse en la revisión de Jeger & Viljanen (2001). Para este caso concreto, se calculó el área bajo la curva de progreso de ruptura o STR y se lo denominó al nuevo índice como “Area Under the Breakage Progress Curve” o AUBPC, ya que una progresión más rápida y agresiva de la ruptura de esterilidad tiene un impacto

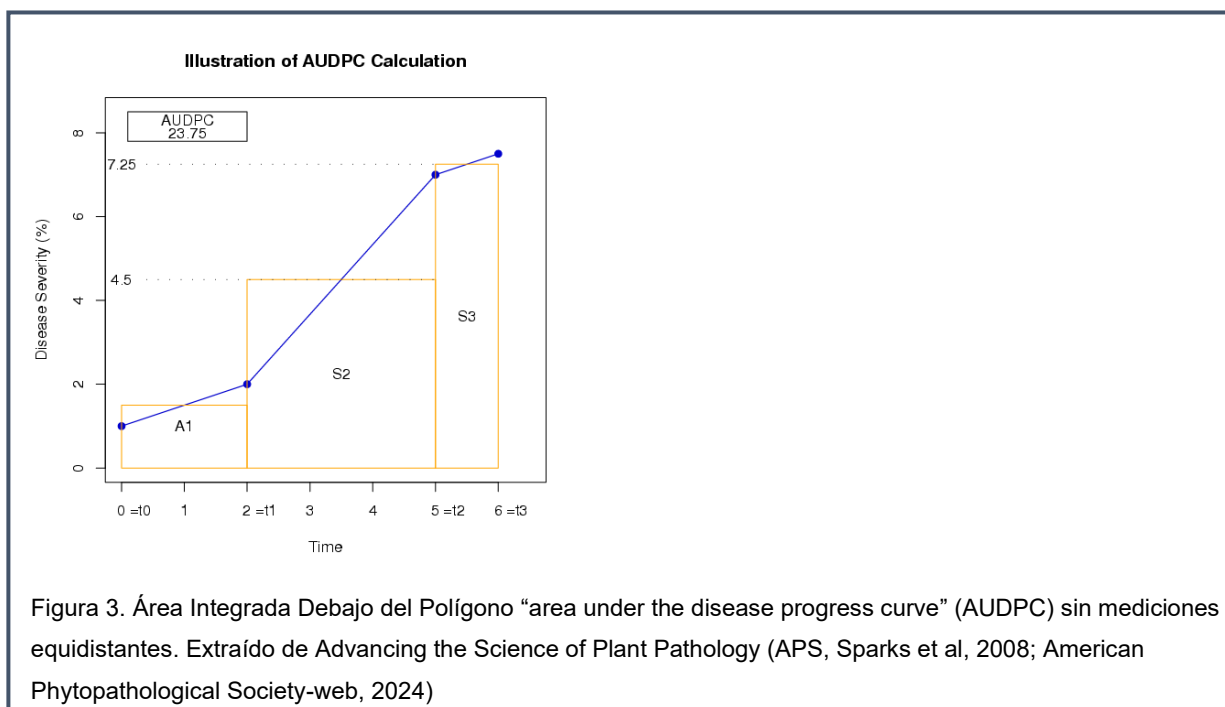
incremental en el riesgo de contaminación / autofecundación con consecuente reducción de la pureza genética de la semilla resultante. El método de cálculo se detalla a continuación:

$$\sum_{i=1}^n \left[\frac{(Y_{i+1} + Y_i)}{2} \right] [X_{i+1} - X_i]$$

Donde:

- Y_i :Ruptura en la parcela (STR) en la observación i
- t_i :tiempo (días) en la observación i
- n : número total de observaciones

En este caso los valores tienen un rango desde 1 (valor 1 en todos los momentos en que se recolectó la nota) a 27 (Valor de 9 en todos los momentos en que se recolectó la nota). Una referencia gráfica de este concepto puede encontrarse en la Figura 3 debajo:



El cálculo del AUBPC se puede expresar como la suma de las severidades observadas en cada medición, teniendo en cuenta los valores intermedios y progresivos (al ser una nota acumulativa). Al suponer que las mediciones se realizan en intervalos de tiempo constantes, la fórmula se puede simplificar, permitiendo la factorización del intervalo de tiempo constante. Por lo tanto, el AUBPC se puede representar como el producto de dicho intervalo por la suma de los promedios de las mediciones adyacentes.

A diferencia de estudios donde el AUDPC es usado para evaluar heredabilidad y comparar genotipos en enfermedades de progresión variable (Jeger & Viljanen, 2001; Veturi et al., 2012), el AUBPC en este estudio se enfoca en la magnitud de la ruptura de esterilidad durante un periodo crítico (ventana de emisión de estigmas).

Adicionalmente, una tercera metodología de caracterización fue propuesta como la nota máxima del valor de ruptura alcanzado (STR90_6) para el caso de LPM-M donde la relación temporal con la emisión de estigmas no es tan crítica ya que presenta importancia para la restauración del híbrido comercial pero no es relevante desde el punto de vista de la producción de semillas a campo.

La alternativa a este índice (AUBPC) que fue considerada, fue la utilización de un índice de Riesgo de Ruptura o IRR en donde se analizó ponderar de manera diferencial si la ruptura ocurría temprana o tardíamente relativo al emisión de estigmas del cultivo. Por ejemplo, ponderar a una ruptura en S50, con el doble de riesgo de autofecundación que si ocurriese en S90. Esta propuesta se basa en una interpretación no lineal del aumento del riesgo en el tiempo. Dicha alternativa fue desechada frente al AUBPC por la complejidad que implica crear un modelo eco fisiológico que estime el volumen de polen potencial por estigma expuesto en la ventana de floración o la arbitrariedad necesaria para definirlo en ausencia de parámetros eco fisiológicos.

3.4.3.2. LPM-M

Para este grupo heterótico, la variable de interés fue directamente el STR en S90_6, ya que el objetivo de evaluación es entender cuál es la proporción de restauración potencial con mayor independencia del tiempo relativo a la emisión de estigmas frente a las líneas utilizadas como hembras en la generación de híbridos.

3.4.3.3. Modelo estadístico de análisis

Se calcularon los BLUPs para las dos variables respuesta basados en un modelo generalizado con distribución gamma, con genotipo, ambiente, e interacción genotipo*ambiente como efectos aleatorios.

$$Y_{ijk} = \mu + L_i + R_j + B_{k(j)} + L \times R_{ij} + \varepsilon_{ijk}$$

donde Y_{ijk} es la respuesta de cada genotipo al riesgo de ruptura del i-ésimo genotipo en la k-ésima repetición del j-ésimo ambiente, μ es la media general, L_i es el efecto aleatorio del i-ésimo genotipo, R_j es el efecto aleatorio del j-ésimo ambiente, $B_{k(j)}$ es el efecto aleatorio de la repetición anidada al ambiente, $L \times R_{ij}$ es el efecto de la interacción entre el i-ésimo genotipo y el j-ésimo ambiente y E es el error residual para el i-ésimo genotipo en la en la k-ésima repetición del j-ésimo ambiente.

De la implementación de este modelo se obtuvieron los BLUPs para AUBPC y STR90_6 del efecto de los genotipos y ambientes que se utilizaron para el análisis de GWAS. El análisis estadístico se llevó a cabo de manera diferenciada para cada grupo heterótico. En el caso de LPM-H, se seleccionó como variable respuesta AUBPC previamente descrito dada su relevancia en términos de riesgo de autofecundación. Por otro lado, en LPM-M la variable respuesta evaluada fue STR90_6 ya que el objetivo fue evaluar su capacidad de restauración potencial en el híbrido sin un fuerte condicionamiento relacionado a la dinámica de emisión de estigmas. Adicionalmente, se utilizaron las medias ajustadas del modelo anterior (BLUEs) para la inferencia de las variables respuesta de los genotipos en el análisis fenotípico descriptivo. Para el GWAS se utilizó un modelo mixto con todos los componentes como aleatorios (genotipo, ambiente y $G \times E$). En este caso la interacción $G \times E$ se modeló como un componente de varianza, lo cual permite absorber la variación adicional sin necesidad de estimar cada combinación explícitamente, logrando así una convergencia adecuada. Por esta razón, en los análisis fenotípicos se discuten únicamente los efectos aditivos, mientras que en el GWAS se utilizaron BLUPs que ya incorporan y corrigen por los efectos de ambiente y de la interacción $G \times E$.

Los análisis descriptos en esta sección fueron realizados con el software estadístico Jamovi Version 2.3 (The jamovi project, 2022) y se utilizó el LSD de Fisher corregido por Bonferroni para la comparación de medias (Fisher 1935; Bonferroni 1936).

3.5. Mapeo por asociación

Con el objetivo de evaluar qué modelo utilizar en el mapeo por asociación se compararon dos modelos estadísticos en el software TASSEL v5 (Bradbury et al., 2007), utilizando como información fenotípica los BLUPs obtenidos de AUBPC y la Matriz_F como información genotípica para el LPM-H, y los BLUPs obtenidos de STR90_6 y Matriz_M para el caso de LPM-M. Los diferentes modelos comparados fueron: 1) Modelo Lineal General (MLG) de efectos fijos + Análisis de Componentes Principales (ACP), 2) Modelo Lineal Mixto (MLM) con la inclusión de la matriz de parentesco entre individuos calculada previamente como efecto aleatorio (Kinship) y la estructura poblacional calculada mediante análisis de componentes principales (ACP). La comparación entre modelos se realizó mediante gráficos cuantil-cuantil (Q-Q plot) realizados en el software TASSEL v5 (Bradbury et al., 2007), que permitieron evaluar la desviación de los valores de probabilidad (valores p) observados para la asociación entre SNPs y la restauración de la androfertilidad de los valores p esperados bajo la hipótesis de ausencia de asociaciones entre los marcadores y la variable.

Para controlar la tasa de error tipo I en el análisis GWAS se empleó la corrección de Bonferroni ajustada por el número efectivo de pruebas (MEFF), estimado mediante un procedimiento de depuración por desequilibrio de ligamiento (denominado en inglés LD pruning). Este método reconoce la correlación entre marcadores genéticos debida al desequilibrio de ligamiento, evitando la sobrecorrección asociada a la Bonferroni tradicional que asume independencia total entre pruebas (Nyholt, 2004; Gao et al., 2008). La estimación de MEFF permite un umbral de significancia más realista y aumenta la potencia estadística del estudio, minimizando el riesgo de falsos positivos sin perder sensibilidad (Duggal et al., 2008).

El modelo elegido fue utilizado para la detección mediante el mapeo por asociación de los loci que rigen la respuesta a la restauración de la androfertilidad, siendo identificados como estadísticamente asociados a la variable respuesta aquellos SNPs con asociaciones de valor $p < 0,0006757$ o valores LOD equivalentes a $-\log_{10}(p) > 3,170262$ (LPM-H) y $p < 0,001111111$ o valores LOD equivalentes a $-\log_{10}(p) > 2,954243$ (LPM-M).

3.6. Identificación de QTL

Para la clasificación de SNPs en QTL se siguieron criterios ya reportados para usos equivalentes (Wang et al., 2018; Zhang et al., 2020 y Tang et al., 2024), los cuales incluyen: (i) identificación

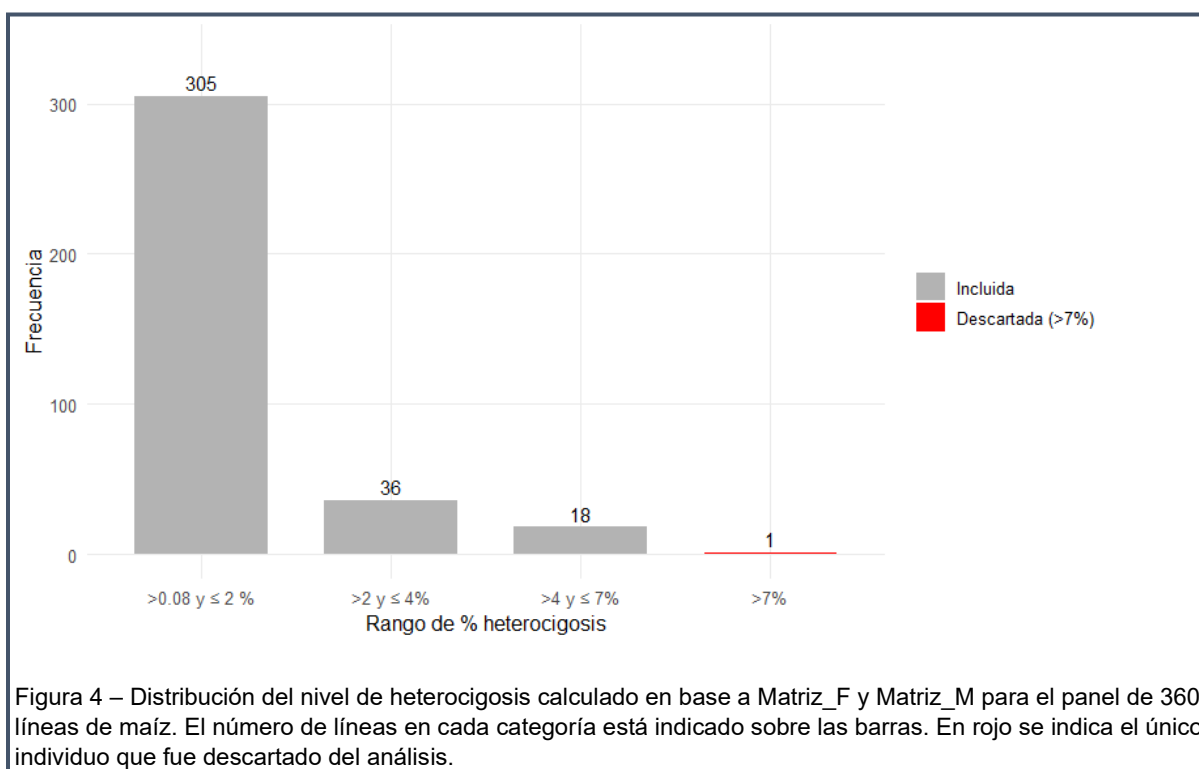
de SNPs con significancia estadística; (ii) agrupamiento de SNPs próximos física/genéticamente (LD, distancia); (iii) efecto de sustitución y frecuencias alélicas; y (iv) estimación del porcentaje de varianza explicada por el SNP. Una vez clasificados, los QTL se denominaron con el prefijo “*QRl*” y “*QRm*” para los identificados en el grupo de LPM-H y LPM-M respectivamente, seguido del cromosoma y orden por bin (Wei et al, 2009). Se seleccionó el SNP de menor valor p dentro de los marcadores para cada QTL.

Adicionalmente, con el fin de complementar la interpretación biológica de los QTL detectados se realizó una caracterización preliminar de las regiones asociadas. Los genes contenidos dentro de los intervalos de los QTL se identificaron utilizando la anotación génica de la referencia de maíz B73 RefGen_v4 (AGPv4; Zm-B73-REFERENCE-GRAMENE-4.0), descargada desde Gramene. Para aquellos QTL definidos por múltiples SNPs significativos, se consideró el intervalo físico comprendido entre el primer y el último marcador de la región. En los QTL representados por un único SNP, el intervalo físico aproximado se definió utilizando una ventana de $\pm 1,9$ Mb alrededor de dicho SNP, en función del decaimiento del desequilibrio de ligamiento observado en LPM-H. Posteriormente, se evaluó como eran los alelos de cada una de las líneas estudiadas para estos QTL y se evaluó el polimorfismo con el objetivo de determinar la consistencia de su variación a través del panel estudiado.

4. RESULTADOS

4.1. Caracterización molecular del panel de estudio

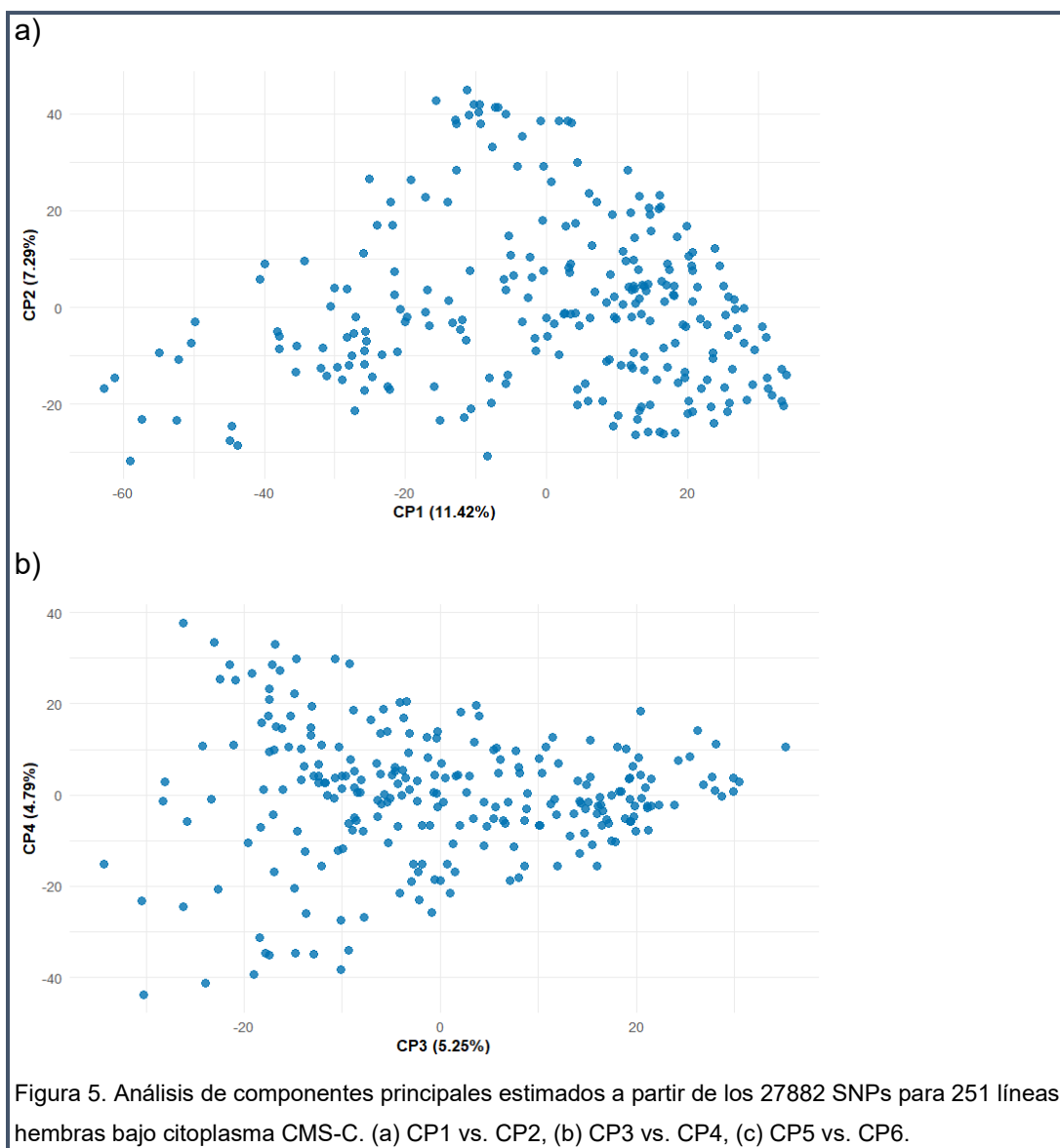
Como resultado del filtrado de la matriz de datos en diferentes pasos y por grupo heterótico se obtuvieron dos matrices genotípicas, un primer conjunto para LPM-H de 27882 SNPs (Matriz_F), y un conjunto para LPM-M de 18690 SNPs (Matriz_M). Adicionalmente, se calculó el porcentaje de heterocigosis para cada una de las líneas evaluadas (Fig. 4). Una sola línea macho fue descartada de los análisis de caracterización molecular y MA dado que presentaba un nivel de heterocigosis del 22% (Fig. 4). Las 359 líneas restantes conformaron el panel en base al cual se realizaron estos análisis (251 genotipos de LPM-H y 108 genotipos de LPM-M).

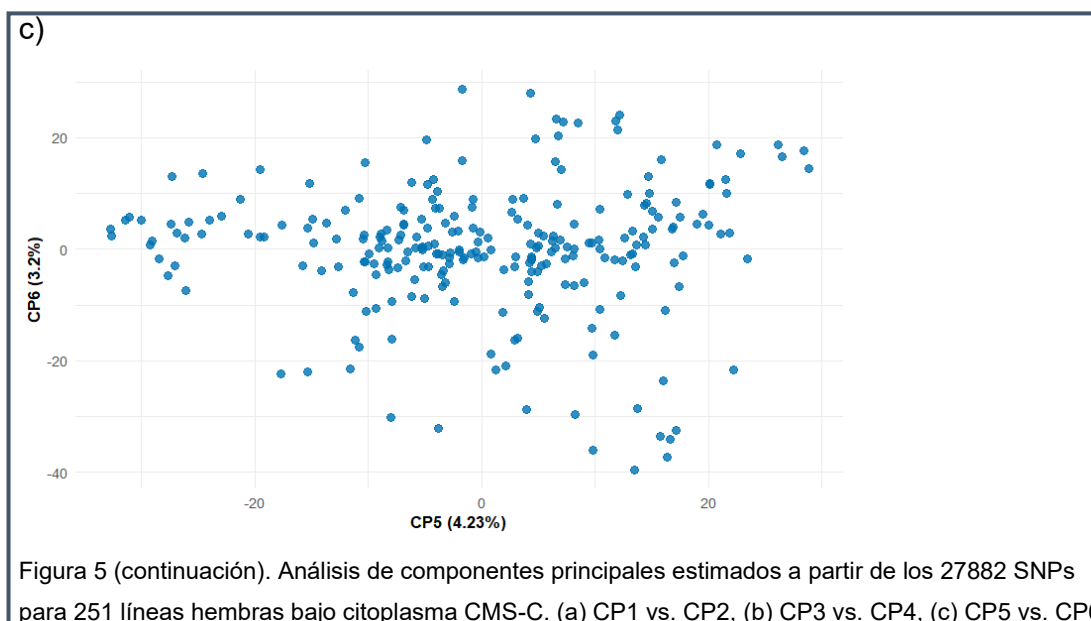


4.1.1. Estructura genética poblacional: Análisis de componentes principales

La estructura poblacional se evaluó mediante un análisis de componentes principales (PCA) utilizando los 27882 marcadores SNP para el caso de Matriz_F y 18690 SNPs para Matriz_M.

En el grupo de las hembras, las CP uno a diez explicó 11,4%, 7,2%, 5,2%, 4,7%, 4,2%, 3,2%, 2,5%, 2,2%, 2,2 y 2% de la variación entre individuos respectivamente, y en conjunto el 45,26% de esta variación (Figura 5a, b y c).

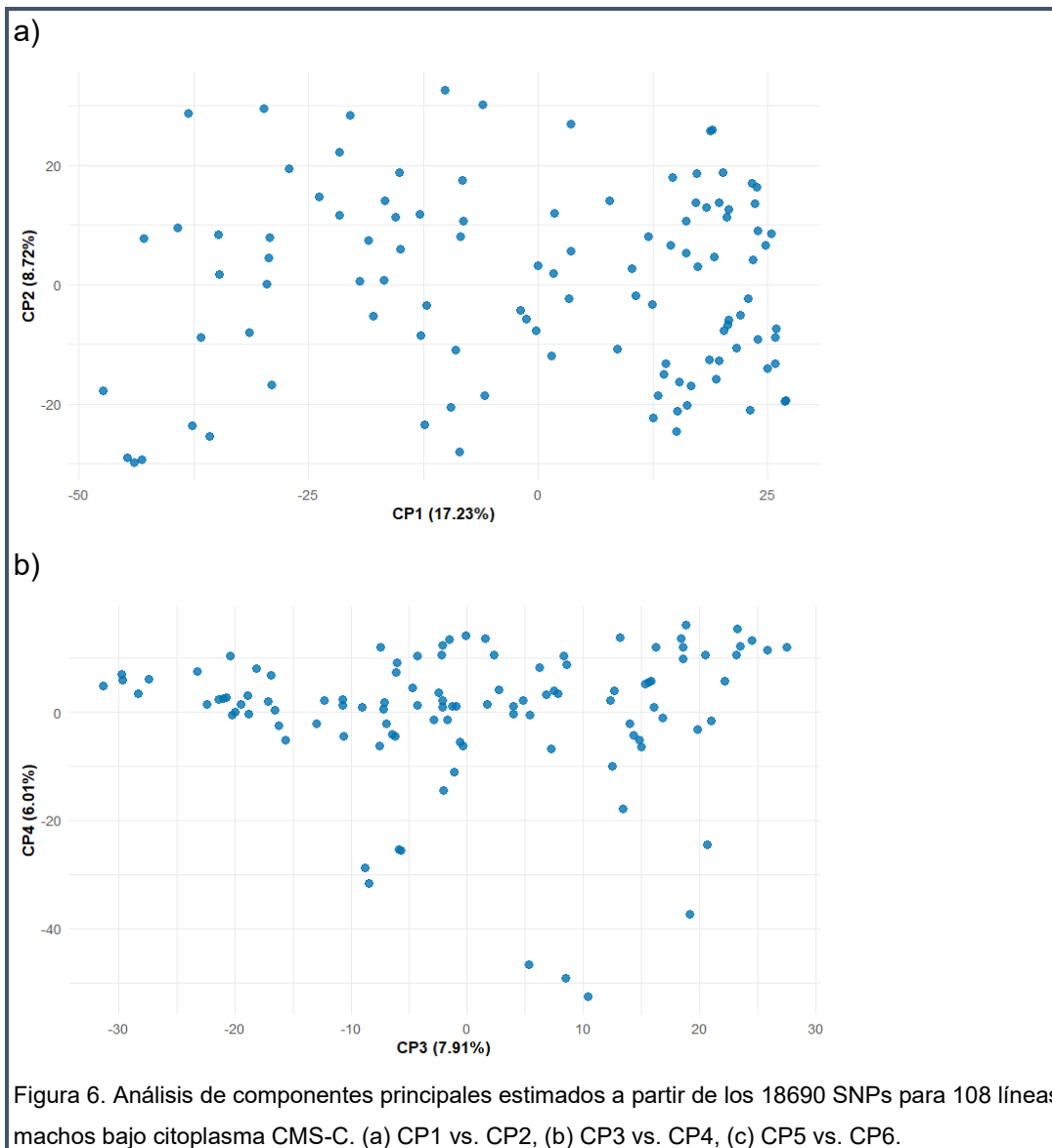


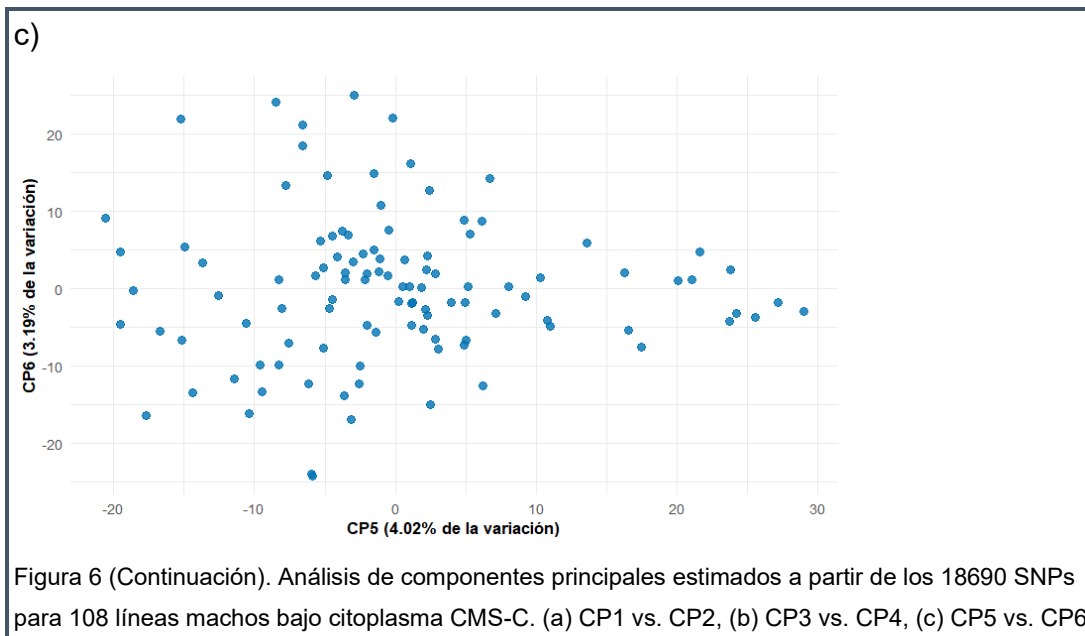


En el grupo de los machos, las CP uno a diez explicó 17,2%, 8,7%, 7,9%, 6,0%, 4,0%, 3,1%, 2,3%, 2,2%, 1,9 y 1,8% de la variación entre individuos respectivamente, y en conjunto el 55,4% de esta variación. Para ambos grupos heteróticos, la distribución continua de los individuos en el plano CP1–CP2 indica la ausencia de subpoblaciones discretas y sugiere una estructura genética gradual, característica del germoplasma templado derivado de programas de mejoramiento con base genética compartida (Figura 6 a, b y c).

No se observaron agrupamientos definidos ni outliers extremos, lo que confirma la homogeneidad relativa del conjunto de líneas.

Los componentes principales fueron incorporados como covariables en el modelo de asociación para controlar la estructura poblacional y reducir la probabilidad de asociaciones espurias (Price et al., 2006).



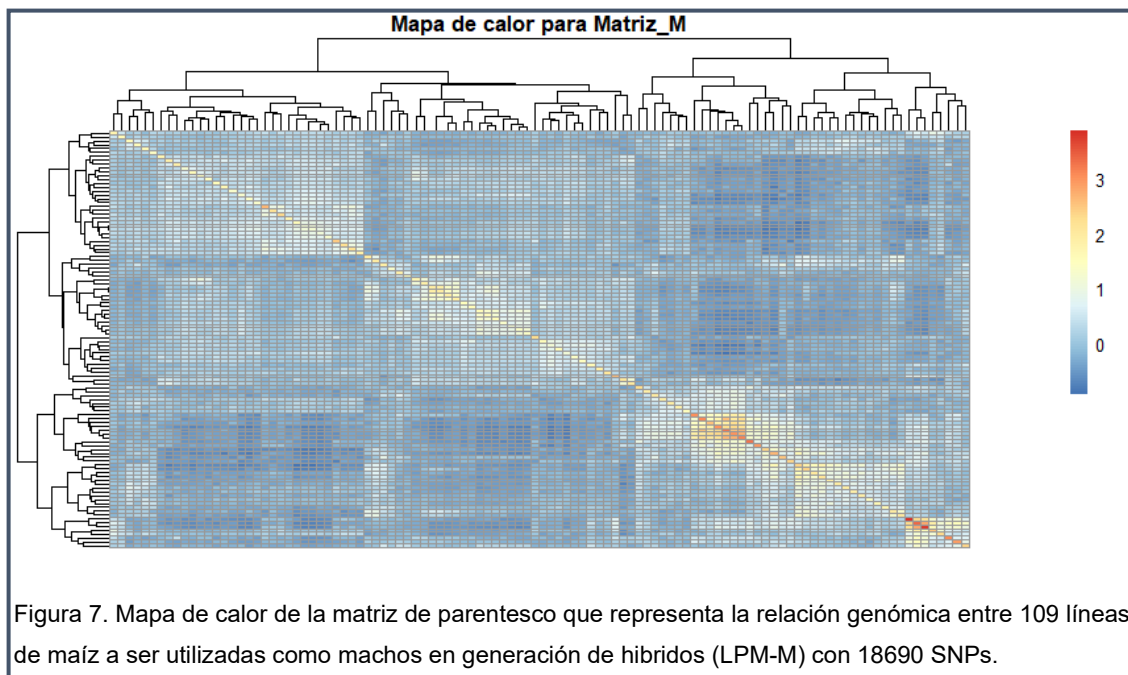


4.1.2. Parentesco relativo

Para el panel LPM-M estudiado, los coeficientes de parentesco relativo para la mayoría de las comparaciones entre pares de líneas (56,9%) fue igual a cero, indicando ningún tipo de relacionamiento o parentesco para estas comparaciones (Tabla 5). El 24% de las comparaciones se ubicaron entre K de 0,1 y 0,5, y un 4% por encima de 1,5, agrupadas en el mapa de calor con colores más intensos, sugiriendo que estas líneas son genéticamente similares y pueden compartir un ancestro común reciente (Figura 7).

Tabla 5. distribución de los coeficientes de parentesco relativo (K) para LPM-M calculados a partir de Matriz_M

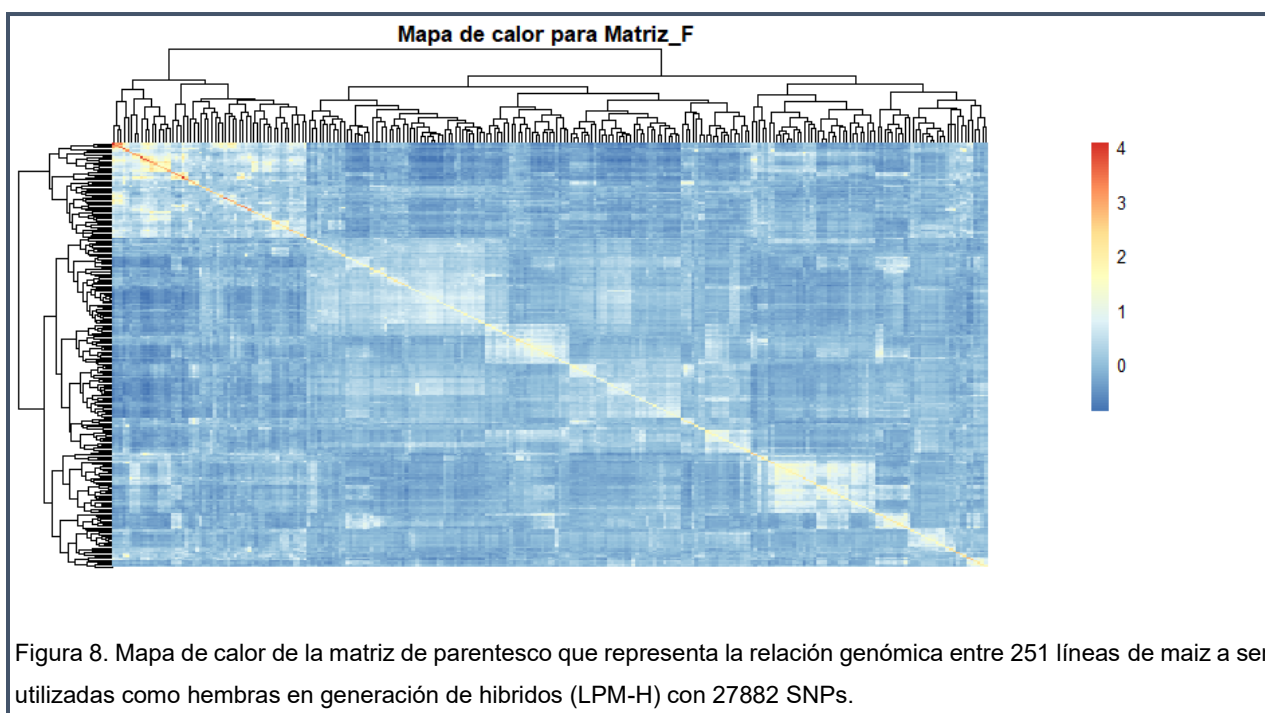
Parentesco relativo	Frecuencia de comparaciones	Frecuencia relativa al total
$x \leq 0$	3292	0,570
$0 < x \leq 0,1$	526	0,091
$0,1 < x \leq 0,5$	1388	0,240
$0,5 < x \leq 1$	455	0,079
$1 < x \leq 1,5$	90	0,016
$1,5 < x \leq 2$	14	0,002
$2 < x \leq 3$	13	0,002



Para el panel de LPM-H estudiado, los coeficientes de parentesco relativo (K) para el 59,9% de las comparaciones entre pares de líneas fue igual a cero, el 21% de las comparaciones se ubicaron con K entre 0,1 y 0,5 y se observó un 2% por encima de 1,5 (Tabla 6). Se observó una sola comparación que presentó un coeficiente de K entre 3 y 4 que no había sido observado para el panel LPM-M. Esta información también se puede interpretar de forma visual en el mapa de calor construido para esta matriz (Figura 8)

Tabla 6. Distribución de los coeficientes de parentesco relativo (K) para LPM-H calculados a partir de Matriz_F

Parentesco relativo	Frecuencia de comparaciones	Frecuencia relativa al total
$x \leq 0$	18803	0,599
$0 < x \leq 0,1$	3518	0,112
$0,1 < x \leq 0,5$	6735	0,215
$0,5 < x \leq 1$	1958	0,062
$1 < x \leq 1,5$	296	0,009
$1,5 < x \leq 2$	45	0,001
$2 < x \leq 3$	19	0,001
$3 < x \leq 4$	1	0,000



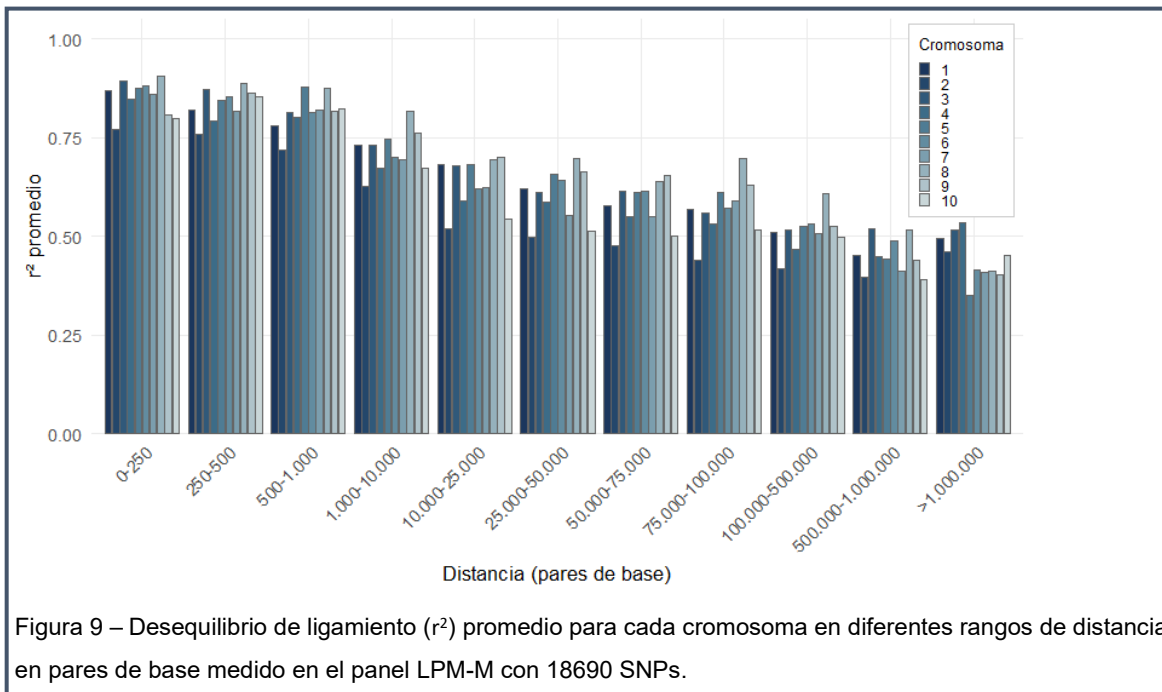
4.1.3 Desequilibrio de ligamiento

El DL para la Matriz_M a través de todo el genoma, medido como r^2 , disminuyó con el aumento de la distancia física entre marcadores, siendo este patrón similar en los diez cromosomas con

poca variación (Tabla 7, Fig. 9,). El DL promedio entre SNPs cercanos, a distancias menores de 250 pb, fluctuó entre 0,77 (cromosoma 2) y 0,90 (cromosoma 8), y la mitad de este valor inicial se observó entre 500000 pb y 1000000 pb. En distancias superiores a 1.000.000 pb, los valores de DL promedio estuvieron entre 0,35 (cromosoma 5) y 0,53 (cromosoma 4), indicando un nivel relativamente fuerte de desequilibrio de ligamiento entre los marcadores analizados.

Tabla 7. Desequilibrios de ligamiento promedio (r^2) a través de los cromosomas y a rangos de distancia entre marcadores en pares de base medido en el panel LPM-M.

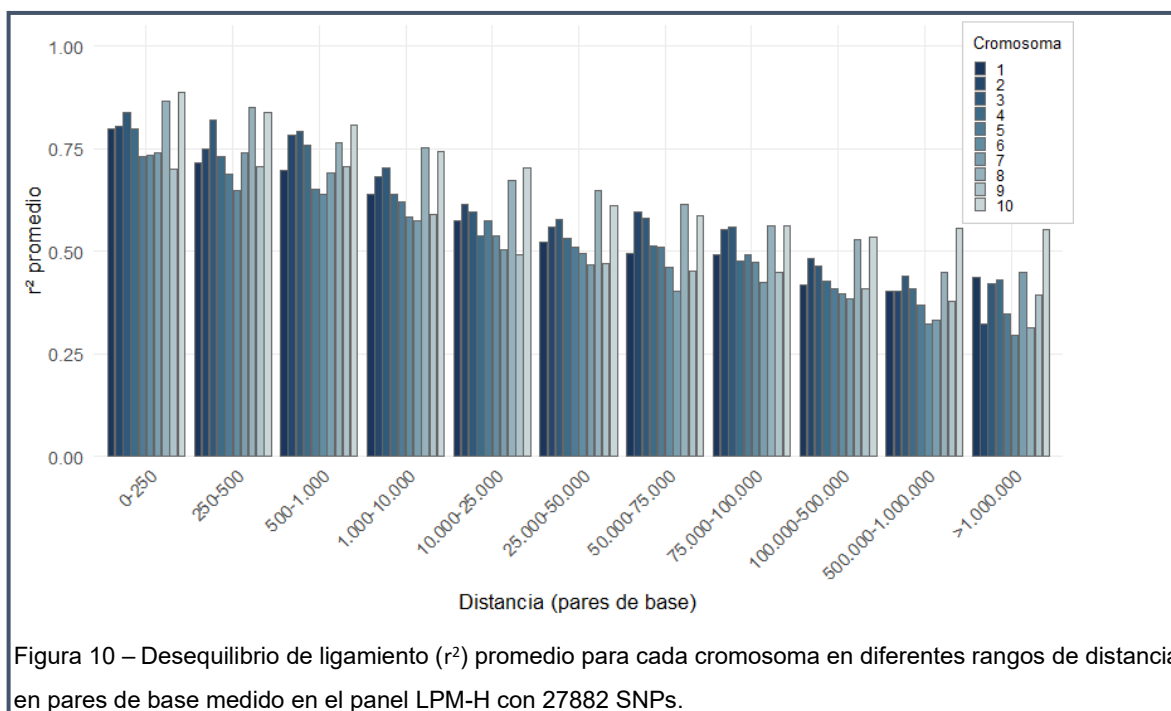
Distancia (pb)	N	Promedio r^2	D,E, r^2
0-250	1009	0,85	0,29
250-500	1170	0,82	0,31
500-1,000	1957	0,80	0,32
1,000-10,000	16422	0,71	0,36
10,000-25,000	11559	0,63	0,38
25,000-50,000	14049	0,59	0,38
50,000-75,000	13341	0,57	0,38
75,000-100,000	13072	0,56	0,38
100,000-500,000	185161	0,50	0,37
500,000-1,000,000	171551	0,45	0,35
1,000,000-10,000,000	446303	0,46	0,35
>10,000,000	46180	0,38	0,34



El DL promedio entre SNPs cercanos para la Matriz_F, a distancias menores de 250 pb, fluctuó entre 0,88 (cromosoma 10) y 0,70 (cromosoma 9), y la mitad de este valor inicial se observó entre 500.000 pb y 1.000.000 pb (Tabla 8, Figura 10). En distancias superiores a 1.000.000 pb, los valores de DL promedio estuvieron entre 0,55 (cromosoma 10) y 0,29 (cromosoma 6). Al igual que en el panel de LPM-M, se observó un nivel relativamente fuerte de desequilibrio de ligamiento entre los marcadores analizados.

Tabla 8. Desequilibrios de ligamiento promedio (r^2) a través de los cromosomas y a rangos de distancia entre marcadores en pares de base medido en el panel LPM-H.

Distancia (pb)	N	Promedio r^2	D.E. r^2
0-250	1785	0,78	0,33
250-500	1897	0,74	0,36
500-1,000	3332	0,72	0,36
1,000-10,000	28504	0,64	0,38
10,000-25,000	20196	0,57	0,38
25,000-50,000	26432	0,53	0,37
50,000-75,000	26203	0,51	0,37
75,000-100,000	25084	0,50	0,37
100,000-500,000	343884	0,43	0,35
500,000-1,000,000	291462	0,39	0,33
1,000,000-10,000,000	571980	0,40	0,33
>10,000,000	36976	0,41	0,37



4.2 Estudio de la restauración de la androesterilidad

4.2.1 Condiciones ambientales

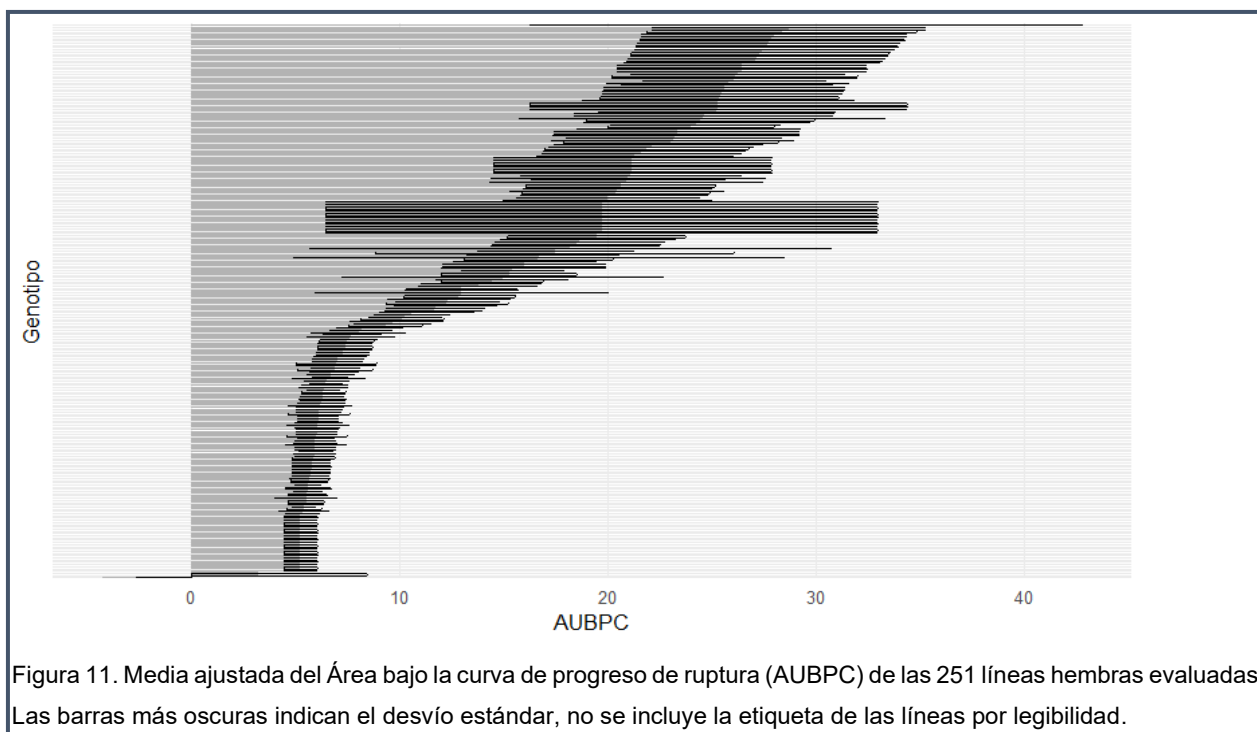
Las condiciones ambientales de temperatura y humedad en los sitios de testeo se describen en la tabla 9. Los ensayos estuvieron 100% irrigados, como en todos los ambientes de producción de semillas de maíz.

Tabla 9. Detalle de temperatura media (°C) y humedad relativa (%) en los sitios de testeo

Localidad	Campaña	Código	Periodo vegetativo - Temperatura promedio (°C)	Periodo Floración - Temperatura promedio (°C)	Periodo vegetativo - HR promedio (%)	Periodo Floración - HR promedio (%)
Buenos Aires Norte A	1	1A	22,03	21,37	70,44	78,31
Buenos Aires Norte B	1	1B	21,88	23,76	62,02	81,44
Buenos Aires Norte A	2	2A	23,71	23,7	62,32	78,89
Buenos Aires Costa	2	2B	19,35	22,19	72,28	68,66

4.2.2 Variabilidad fenotípica para la restauración de la androfertilidad en el panel LPM-H

Las medias de AUBPC mostraron en los cuatro ambientes evaluados valores de 3 a 29, indicando variabilidad fenotípica en el panel estudiado que permitió identificar grupos de genotipos con alta, media y baja ruptura. El 31,8% de las líneas presentaron AUBPC mayores a 20, identificadas como alta ruptura, el 21,5 % presentó AUBPC entre 10 y 20 (media ruptura) y el 46% del total de panel presentó AUBPC entre 3 y 10 (baja ruptura) (Figura 11).

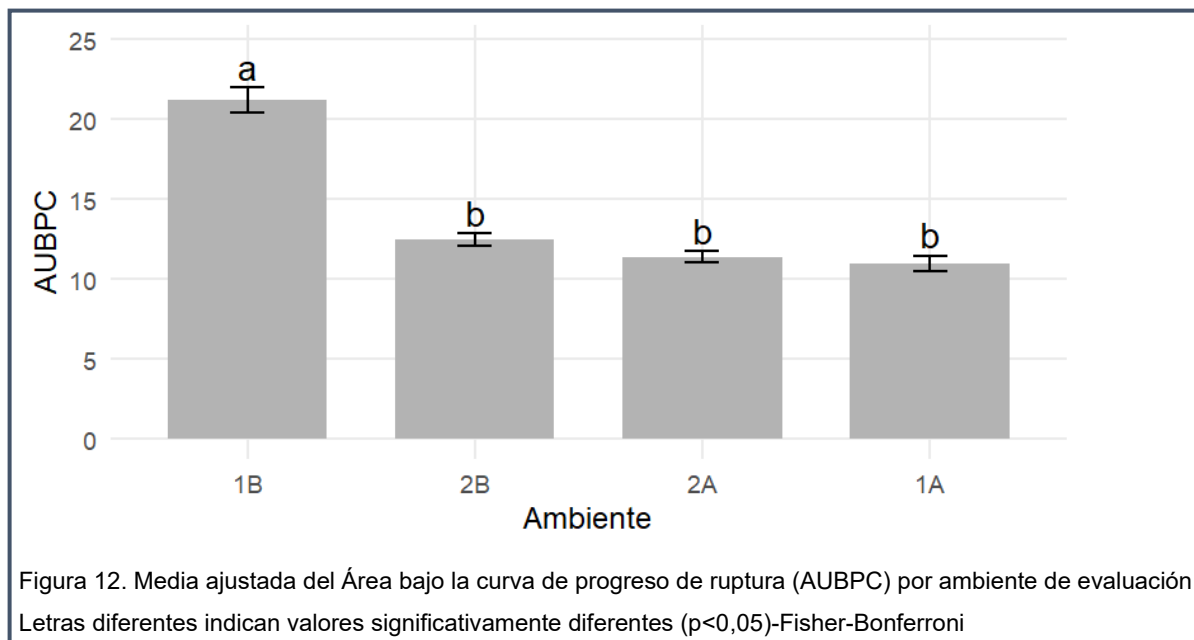


El AUBPC por ambiente presentó valores medios de 21,2 para el ambiente 1B, 11 para 1A, 12,5 para 2B y 11,4 para 2A (Tabla 10). 1B fue el ambiente donde más restauración de la fertilidad se evaluó diferenciándose significativamente de las otras tres localidades, donde las medias observadas no presentaron diferencias significativas entre sí (Figura 12).

Tabla 10. AUBPC medio por ambiente

Ambiente	Media ajustada	D,E	IIC 95%	ISC 95%
1B	21,2	0,783	19,7	22,8
1A	11	0,487	10	11,9
2B	12,5	0,401	11,7	13,3
2A	11,4	0,385	10,7	12,2

D,E: desvío estándar, IIC 95%: intervalo inferior de confianza 95%, ISC 95%: intervalo superior de confianza 95%



4.2.3 Variabilidad fenotípica para la restauración de la androfertilidad en el panel LPM-M

Una tercera metodología de caracterización fue propuesta para el caso de este panel como la nota máxima del valor de ruptura alcanzado (STR90_6). En este grupo heterótico, la relación temporal con la emisión de estigmas no es tan relevante desde el punto de vista de la producción de semillas a campo, sino que su importancia radica en la capacidad de restauración en el híbrido comercial.

Las medias de STR90_6 en las cuatro localidades mostraron valores entre 0 y 2 para el 95,7% del panel estudiado (Figura 13) y apenas un 4,62% de los individuos presentaron notas mayores a 7,8 indicando ausencia de restauración en todos los ambientes (Tabla 11). Las diferencias en las medias ajustadas por ambiente de evaluación pueden observarse en la Figura 14.

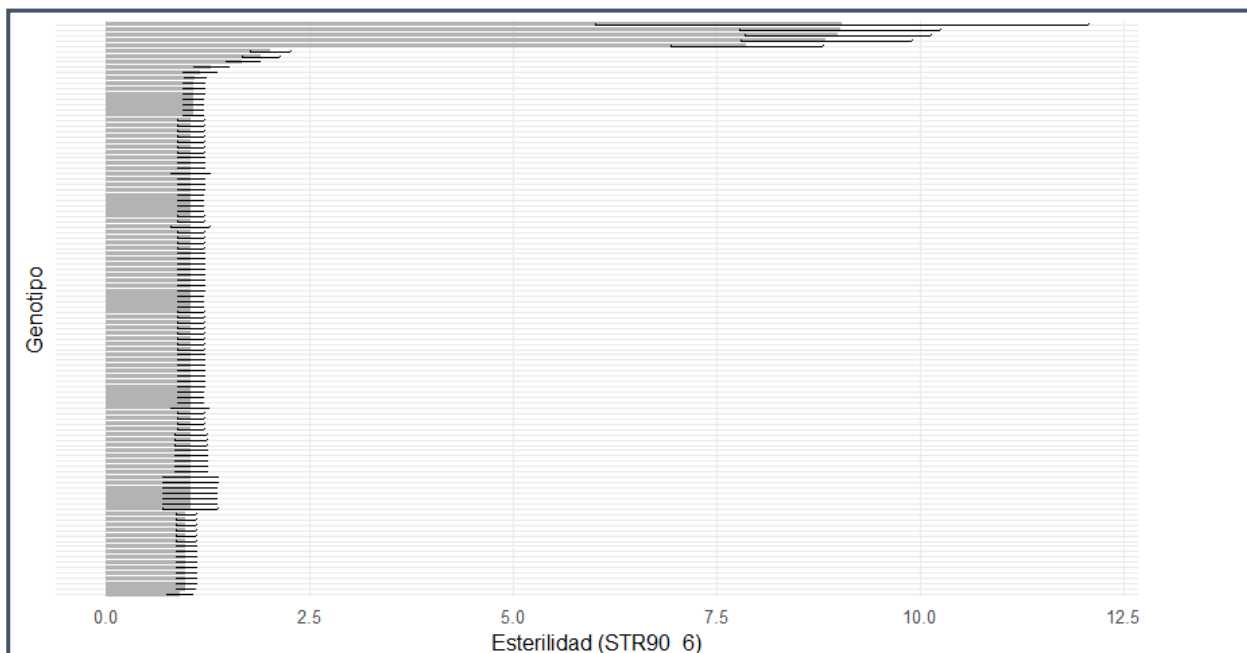


Figura 13. Media ajustada de la Nota máxima del valor de ruptura alcanzado (STR90_6) del panel LPM-M evaluado. Las barras más oscuras indican el desvío estándar, no se incluye la etiqueta de las líneas por legibilidad.

Tabla 11. STR90_6 medio por ambiente

Ambiente	Media ajustada	D.E	IIC 95%	ISC 95%
1B	1,32	0,0557	1,21	1,43
1A	1,31	0,0568	1,2	1,42
2B	1,58	0,0704	1,44	1,72
2A	1,46	0,0662	1,33	1,59

D,E: desvío estándar, IIC 95%: intervalo inferior de confianza 95%, ISC 95%: intervalo inferior de confianza 95%

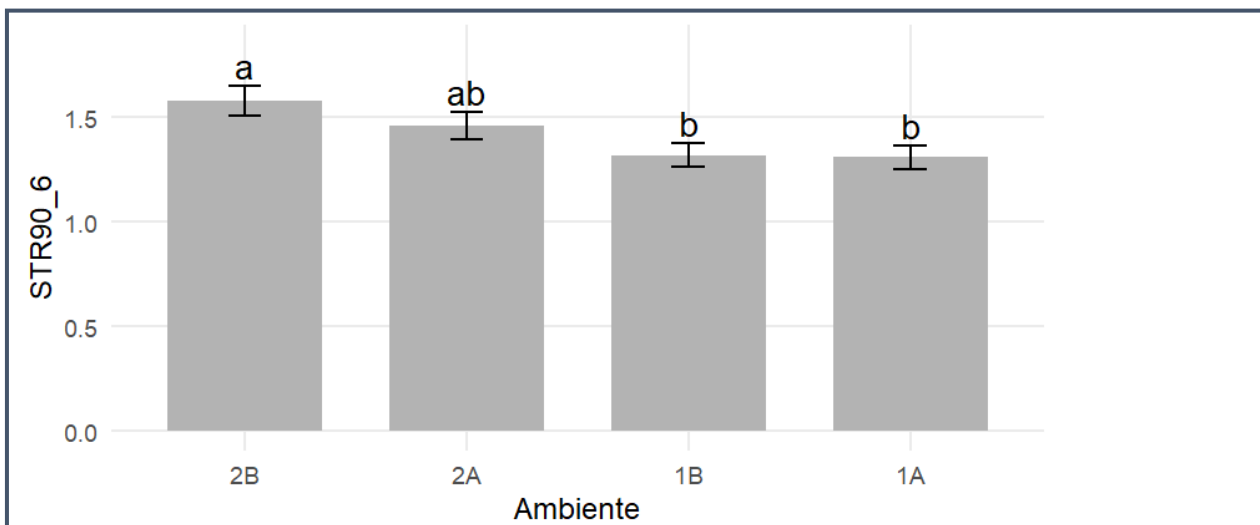


Figura 14. Media ajustada de la Nota máxima del valor de ruptura alcanzado (STR90_6) por ambiente de evaluación. Letras diferentes indican valores significativamente diferentes ($p < 0,05$)-Fisher-Bonferroni

4.3 Estudio de mapeo por asociación de la restauración de la androfertilidad con citoplasma androestéril CMS-C

4.3.1 Elección del modelo estadístico

Los modelos estadísticos de mapeo por asociación se llevaron a cabo en el software TASSEL v5 (Bradbury et al, 2007) con el objetivo de encontrar cuál de ellos permitía controlar mejor el error tipo 1, observado a través de la distribución de valores p. En TASSEL, los modelos GLM + ACP y MLM+ACP+K se compararon para las variables respuesta AUBPC en el caso de LPM-H y STR90_6 en el caso de LPM-M (Figura 15)

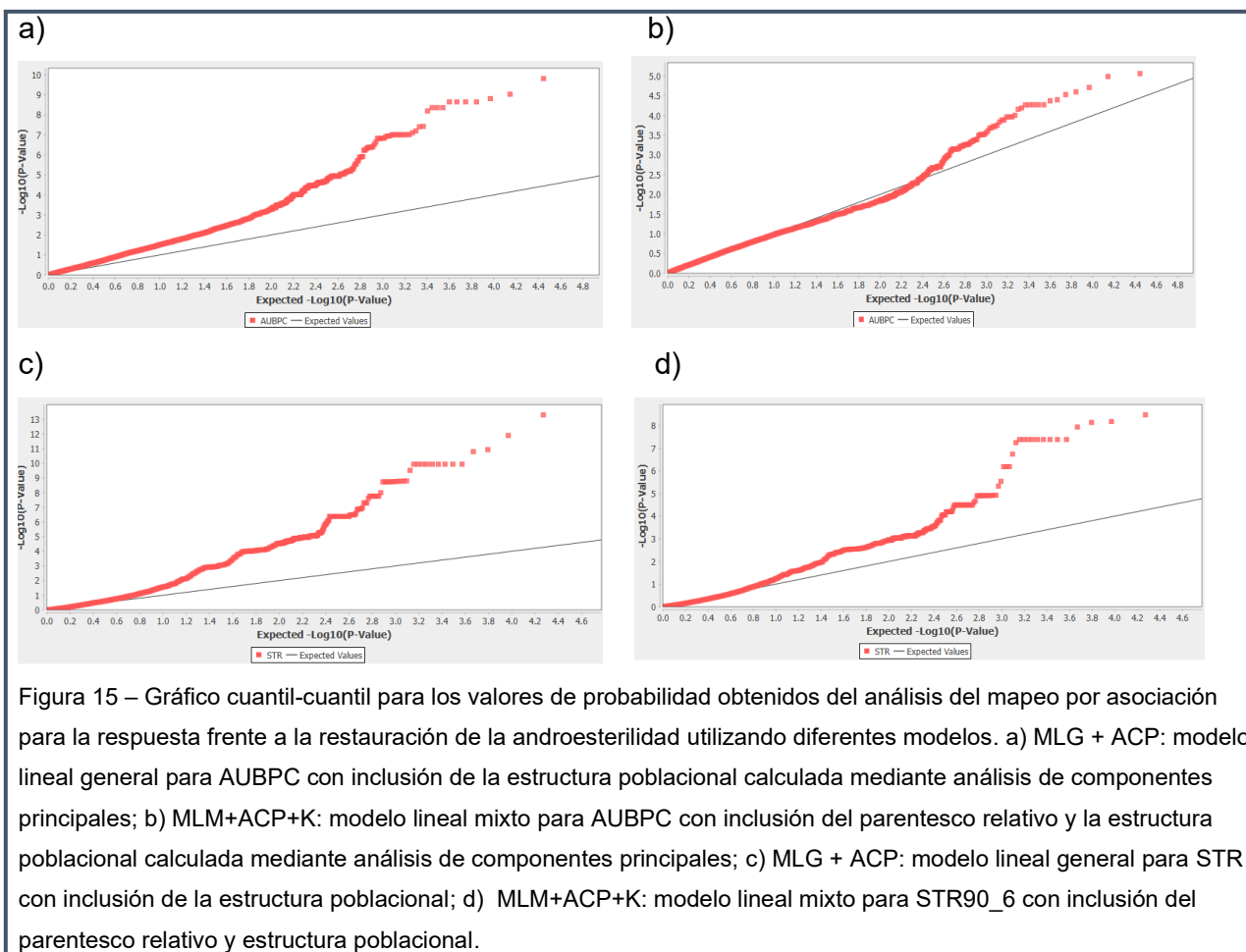


Figura 15 – Gráfico cuantil-cuantil para los valores de probabilidad obtenidos del análisis del mapeo por asociación para la respuesta frente a la restauración de la androesterilidad utilizando diferentes modelos. a) MLG + ACP: modelo lineal general para AUBPC con inclusión de la estructura poblacional calculada mediante análisis de componentes principales; b) MLM+ACP+K: modelo lineal mixto para AUBPC con inclusión del parentesco relativo y la estructura poblacional calculada mediante análisis de componentes principales; c) MLG + ACP: modelo lineal general para STR con inclusión de la estructura poblacional; d) MLM+ACP+K: modelo lineal mixto para STR90_6 con inclusión del parentesco relativo y estructura poblacional.

La distribución de los valores p en el MLG, se alejaron bastante de la distribución esperada en ausencia de asociaciones entre marcadores y la variable estudiada, especialmente para la variable respuesta AUBPC. El MLM con control de estructura poblacional o parentesco relativo resultó con un mejor ajuste para valores p por encima del nivel de significancia estadística establecido tanto LPM-H (valores $p < 0,0006756757$ o $\text{LOD} > -3,170262$) como para LPM-M (valores $p < 0,0011111111$ o $\text{LOD} > -2,954243$ (machos) y fue el modelo seleccionado para ser utilizado en el mapeo por asociación.

4.3.2 Detección de QTL- Panel LPM-H

50 SNPs fueron identificados en los cromosomas 1, 2, 3 y 8 presentando asociaciones significativas con la variable respuesta AUBPC ($p < 0,000675675$) (Figura 15).

Se definió como alelo alternativo al alelo menor, y se calculó el efecto de sustitución alélica (ESA) como $(\text{efecto_ALT} - \text{efecto_REF})/2$, fluctuando los valores para estos SNPs entre -0,24 y 0,36. Adicionalmente, la proporción de la variancia explicada para estos SNPs fue del 4,80% al 9,75% a través de los ambientes evaluados. Estos 50 SNPs con asociaciones estadísticamente significativas a AUBPC fueron agrupados según lo explicado en la sección materiales y métodos (el criterio de distancia física utilizado fue de 1,900,000 pb) en cinco QTL: *Qrf8.1*, *Qrf2.1*, *Qrf2.2*, *Qrf1.1*, *Qrf3.1* (Tabla 12)

Tabla 12. SNPs y QTL identificados en el mapeo por asociación para la respuesta a la androesterilidad en LPM-H.

SNP	QTL	Cr	Pos (pb)	Bin	Valor p	r ²	ESA	Alelo de ref	n
SNP94417	<i>Qrf1.1</i>	1	15,3XX,XXX	1.01	3,21E-04	0,067	-1,82E-01	G	186
SNP24981		2	24,4XX,XXX	2.02	2,99E-04	0,068	1,70E-01	G	208
SNP39209	<i>Qrf2.1</i>	2	24,4 XX,XXX	2.02	2,11E-04	0,071	1,76E-01	G	209
SNP59138		2	29,4XX,XXX	2.02	5,30E-04	0,063	1,65E-01	G	152
SNP26252		2	234,1XX,XXX	2.12	5,06E-04	0,063	1,44E-01	A	174
SNP74697		2	234,1XX,XXX	2.12	5,47E-04	0,050	1,46E-01	G	182
SNP36022		2	235,8XX,XXX	2.12	5,56E-04	0,063	1,52E-01	A	167
SNP42479		2	235,8XX,XXX	2.12	2,51E-05	0,090	1,70E-01	A	168
SNP86723		2	235,8XX,XXX	2.12	1,30E-04	0,061	1,60E-01	G	185
SNP51098		2	235,8XX,XXX	2.12	1,30E-04	0,061	1,60E-01	G	185
SNP48712		2	235,9XX,XXX	2.12	9,88E-05	0,078	1,51E-01	A	165
SNP39212		2	236,0XX,XXX	2.12	6,71E-04	0,048	1,38E-01	A	183
SNP86273	<i>Qrf2.2</i>	2	236,3XX,XXX	2.12	1,02E-05	0,098	1,45E-01	G	161
SNP59963		2	236,4XX,XXX	2.12	1,75E-04	0,059	1,53E-01	A	183
SNP88382		2	236,4XX,XXX	2.12	4,22E-05	0,085	1,64E-01	G	170
SNP41990		2	236,6XX,XXX	2.12	2,95E-05	0,073	1,72E-01	A	178
SNP83151		2	236,6XX,XXX	2.12	1,09E-04	0,063	1,49E-01	A	170
SNP46676		2	236,6XX,XXX	2.12	5,42E-05	0,068	1,58E-01	G	171
SNP41413		2	236,9XX,XXX	2.12	6,02E-04	0,062	1,41E-01	A	146

Referencias: Cr: Cromosoma, Pos (pb): posición en pares de base- la información específica se encuentra en el material confidencial del patrocinador-, Bin: asignados como aproximación basada en intervalos físicos de 20 Mb por bin en B73 RefGen v4, r²: variación explicada, ESA: Efecto de sustitución alélica, Alelo de ref: Alelo en mayor proporción en el cálculo de ESA, n: número de individuos con el alelo indicado.

Tabla 12 (continuación). SNPs y QTL identificados en el mapeo por asociación para la respuesta a la androesterilidad en LPM-H.

SNP	QTL	Cr	Pos (pb)	Bin	Valor p	r ²	ESA	Alelo de ref	n
SNP88637	Qrf3.1	2	236,9XX,XXX	2.12	4,37E-04	0,065	1,38E-01	A	142
SNP35055		3	203,5XX,XXX	3.11	3,95E-04	0,052	-2,49E-01	A	235
SNP72008		8	329,XXX	8.01	5,32E-05	0,068	1,93E-01	A	220
SNP48226		8	332,XXX	8.01	5,32E-05	0,068	1,93E-01	A	220
SNP54009		8	333,XXX	8.01	5,32E-05	0,068	1,93E-01	A	220
SNP100636		8	334,XXX	8.01	5,32E-05	0,068	1,93E-01	A	220
SNP71624	Qrf8.1	8	429,XXX	8.01	4,15E-04	0,052	1,82E-01	G	226
SNP24941		8	467,XXX	8.01	4,15E-04	0,052	3,64E-01	A	226
SNP93928		8	811,XXX	8.01	2,72E-04	0,055	3,59E-01	A	224
SNP39986		8	956,XXX	8.01	8,66E-06	0,084	2,27E-01	G	223
SNP24812		8	991,XXX	8.01	6,93E-05	0,066	1,83E-01	G	218
SNP48578		8	1,0XX,XXX	8.01	1,46E-04	0,060	1,97E-01	G	226
SNP10719		8	1,2XX,XXX	8.01	1,94E-05	0,077	2,04E-01	G	220
SNP59462		8	1,5XX,XXX	8.01	3,99E-05	0,071	2,00E-01	A	221
SNP29287		8	1,7XX,XXX	8.01	1,88E-04	0,058	1,73E-01	A	216
SNP34764		8	1,8XX,XXX	8.01	5,46E-04	0,050	-1,54E-01	G	213
SNP38343		8	1,8XX,XXX	8.01	1,88E-04	0,058	1,73E-01	G	216
SNP36264		8	1,9XX,XXX	8.01	5,46E-04	0,050	1,54E-01	G	213
SNP51962		8	1,9XX,XXX	8.01	6,03E-04	0,049	1,56E-01	A	213
SNP17276		8	1,9XX,XXX	8.01	6,03E-04	0,049	1,56E-01	G	213
SNP97745		8	2,0XX,XXX	8.01	3,04E-04	0,054	1,62E-01	A	212
SNP25848		8	2,1XX,XXX	8.01	3,04E-04	0,054	1,62E-01	A	212
SNP71373		8	2,1XX,XXX	8.01	3,04E-04	0,054	1,62E-01	G	212
SNP36418		8	2,4XX,XXX	8.01	2,05E-04	0,058	1,76E-01	A	216
SNP26121		8	2,5XX,XXX	8.01	5,77E-04	0,049	1,74E-01	A	220
SNP36238		8	2,5XX,XXX	8.01	2,48E-04	0,056	1,77E-01	A	217
SNP64182		8	2,6XX,XXX	8.01	4,56E-04	0,051	1,79E-01	G	222
SNP14578		8	3,1XX,XXX	8.01	4,72E-04	0,064	1,86E-01	A	222
SNP17911		8	3,2XX,XXX	8.01	1,08E-04	0,063	1,92E-01	A	220
SNP71805		8	3,2XX,XXX	8.01	1,08E-04	0,063	1,92E-01	G	220
SNP23125		8	3,2XX,XXX	8.01	6,46E-05	0,067	1,92E-01	A	221

Referencias: Cr: Cromosoma, Pos (pb): posición en pares de base- la información específica se encuentra en el material confidencial del patrocinador-, Bin: asignados como aproximación basada en intervalos físicos de 20 Mb por bin en B73 RefGen v4, r²: variación explicada, ESA: Efecto de sustitución alélica, Alelo de ref: Alelo en mayor proporción en el cálculo de ESA, n: número de individuos con el alelo indicado.

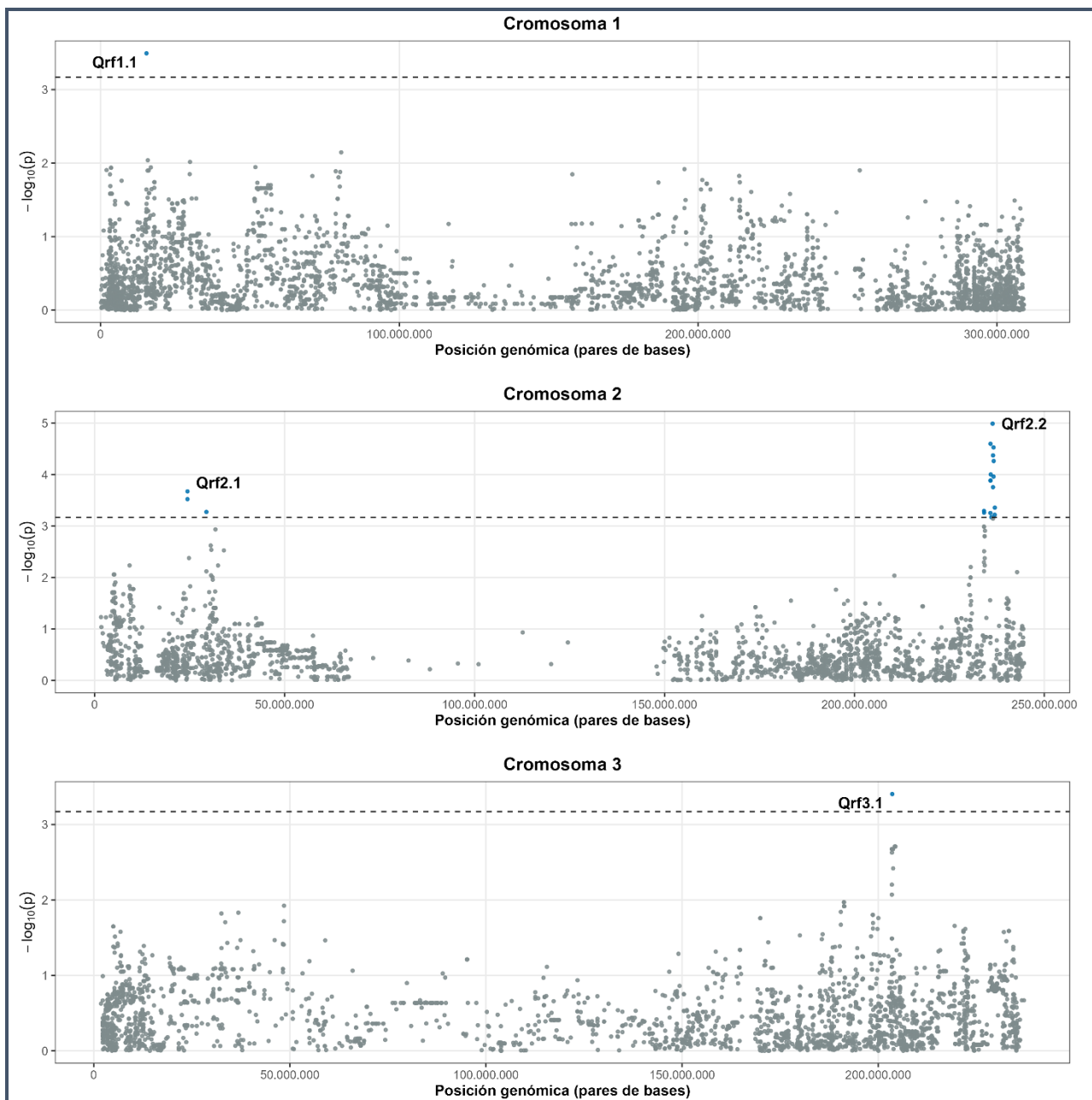
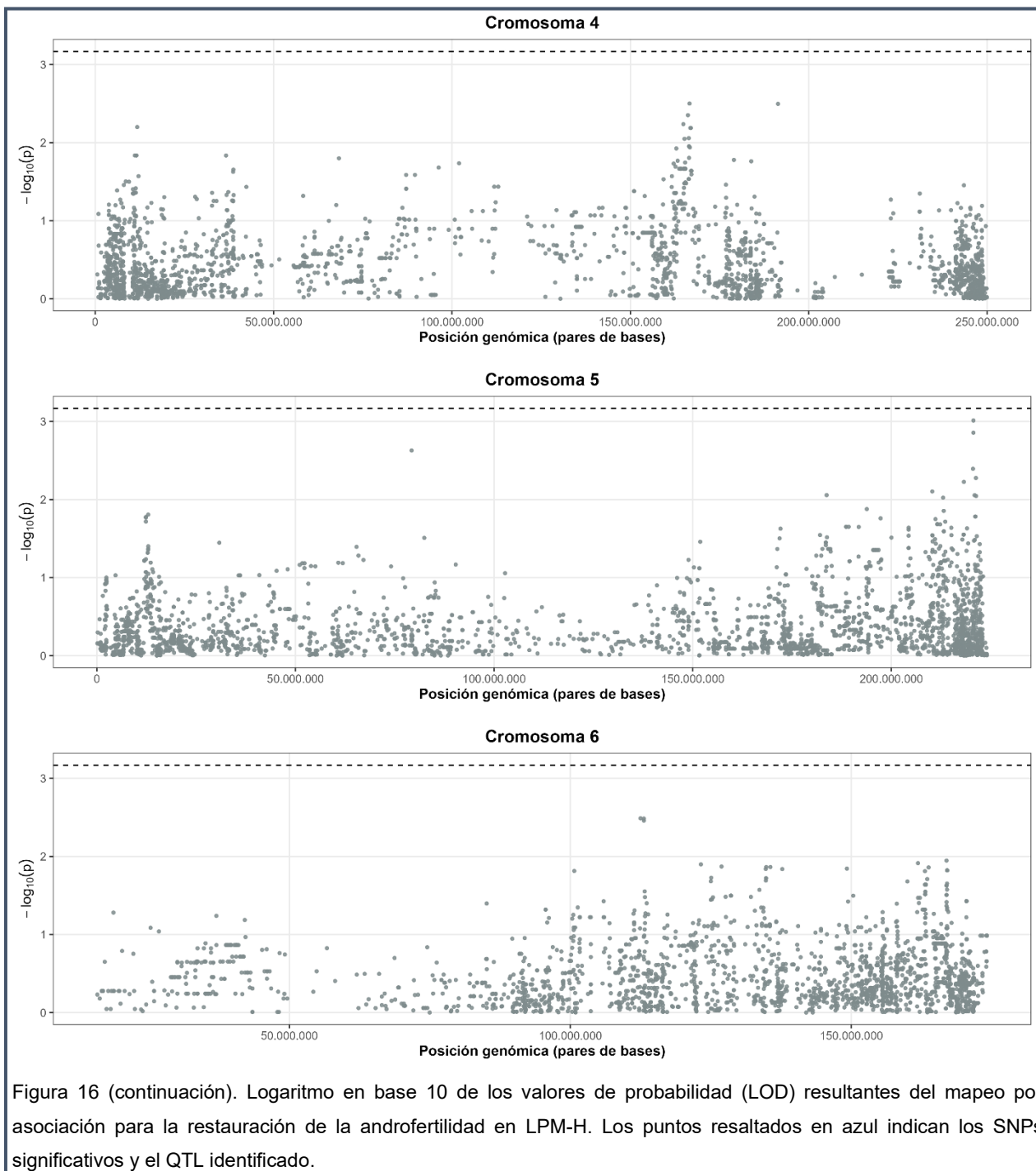
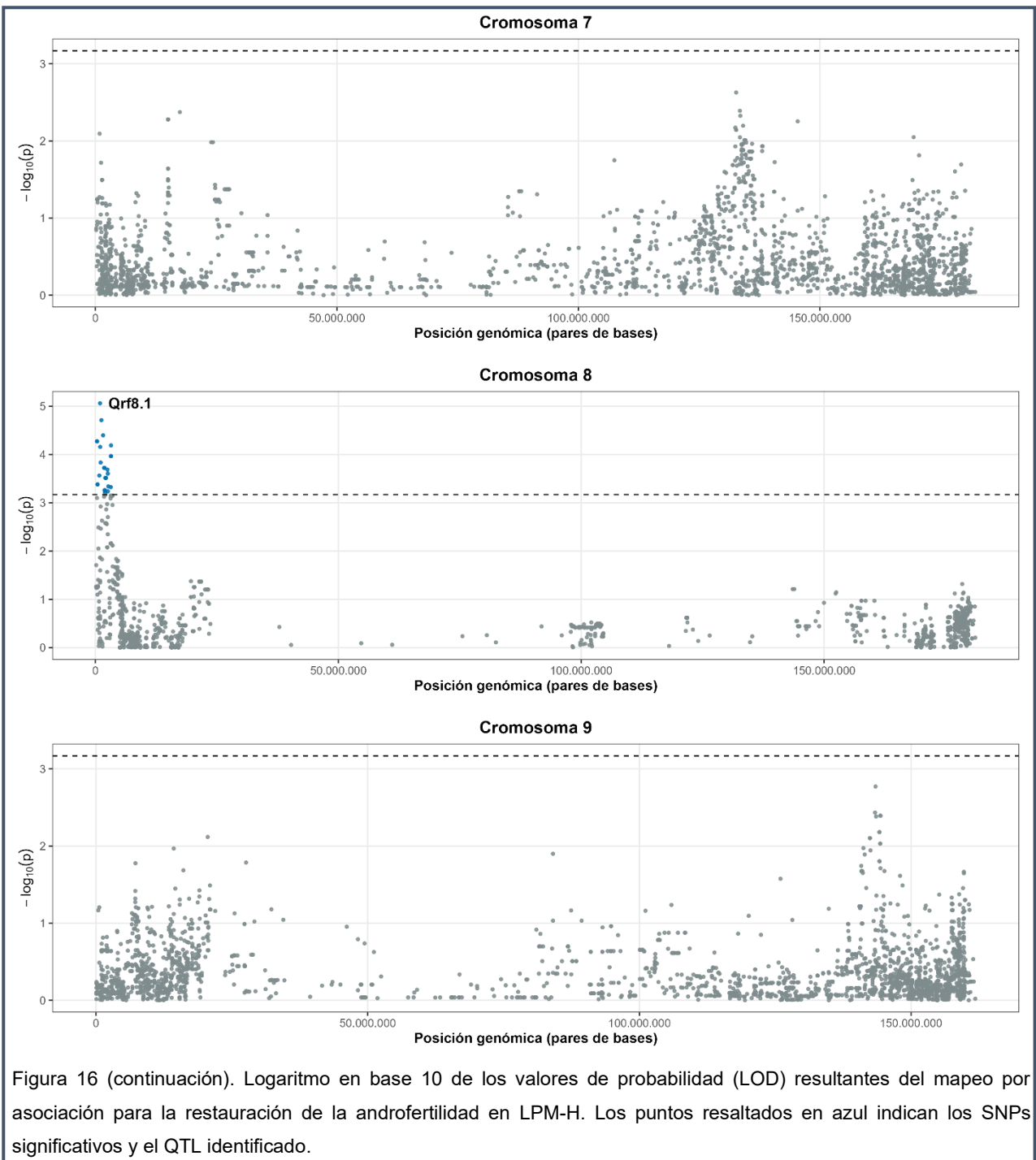
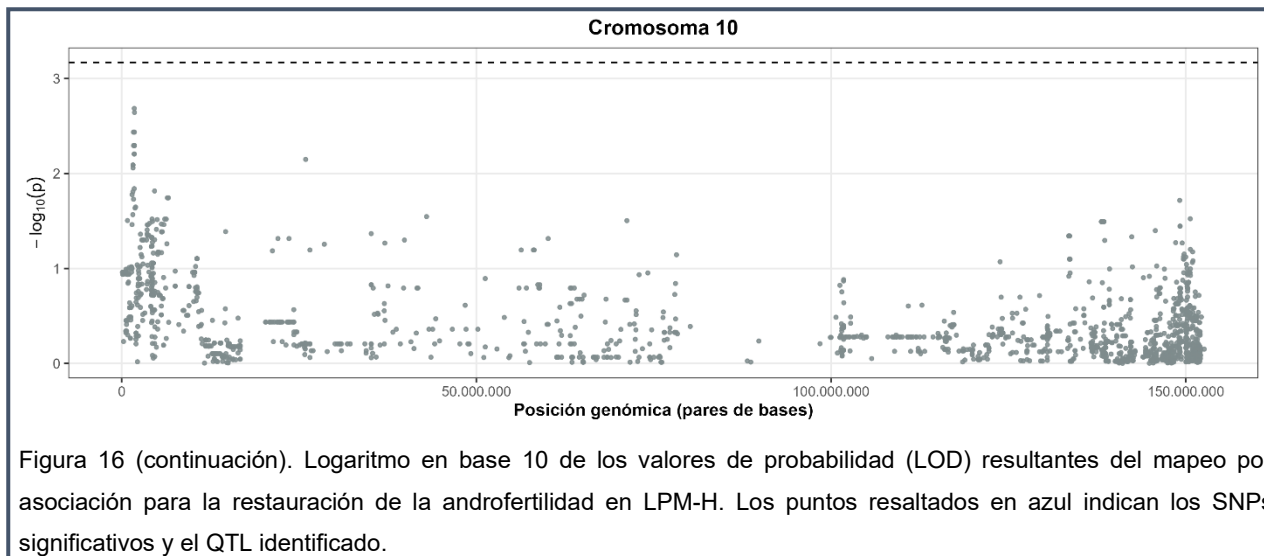


Figura 16. Logaritmo en base 10 de los valores de probabilidad (LOD) resultantes del mapeo por asociación para la restauración de la androfertilidad en LPM-H. Los puntos resaltados en azul indican los SNPs significativos y el QTL identificado.







Se realizó una caracterización preliminar de las regiones asociadas de los QTL detectados con el fin de complementar la interpretación biológica, incluyendo su intervalo físico, el número de genes reportados y la eventual presencia de genes restauradores conocidos (Tabla 13).

Tabla 13. Caracterización de los QTL asociados con la restauración de la androfertilidad en CMS-C.

QTL	Cr/Bin	Intervalo físico (Mb)	Tamaño (Mb)	Nº de genes	Genes restauradores reportados	Referencias
<i>Qrf1.1</i>	1/1.01	13,4-17,2	3,8	104	-	-
<i>Qrf2.1</i>	2/2.02	24-29	5	131	-	-
<i>Qrf2.2</i>	2/2.12	234,1-236,9	2,8	85	<i>Rfm1</i> , <i>Rf12</i> (<i>Zm00001d007531</i>)	Chen y Chen (1992); Yang et al., (2010); Zhang et al., (2023)
<i>Qrf3.1</i>	3/3.11	201,1-204,9	3,8	104	Región cercana a QTL de restauración parcial (bin 3.06)	Kohls et al. (2011)
<i>Qrf8.1</i>	8/8.01	0,33-3,22	2,89	66	<i>Rf4</i>	Kheyr-Pour et al., (1981); Sisco et al., (1991)

Adicionalmente, se verificó el estado alélico para cada uno de los cinco QTL en el panel LPM-H (Anexo, Tabla Suplementaria 1) y se observó que el 51% de las líneas presentó el alelo que confiere alta ruptura o fertilidad para *Qrf1.1*, 75% para *Qrf2.1*, 79% para *Qrf2.2*, 56% para *Qrf3.1* y el 100% de las líneas resultaron fértiles cuando presentaron el alelo que confiere alta ruptura o fertilidad para *Qrf8.1*. (Tabla 14).

Tabla 14. Estado alélico determinado por el SNP de mayor significancia en cada uno de los cinco QTL determinados en el mapeo por asociación y fenotipo expresado como ratio para el panel LPM-H, E: estado alélico que confiere el mejor comportamiento (baja ruptura), R: estado alélico alternativo (alta ruptura o fértil).

QTL	Estado alélico	Ratio Alta Ruptura	Ratio Media Ruptura	Ratio Baja Ruptura
<i>Qrf1.1</i>	E	32%	22%	46%
	R	30%	21%	49%
<i>Qrf2.1</i>	E	26%	23%	51%
	R	63%	12%	24%
<i>Qrf2.2</i>	E	21%	18%	61%
	R	51%	28%	22%
<i>Qrf3.1</i>	E	13%	6%	81%
	R	33%	23%	44%
<i>Qrf8.1</i>	E	0%	0%	100%
	R	90%	10%	0%

4.3.3 Detección de QTL- Panel LPM-M

Se identificaron SNPs con asociaciones significativas con la variable respuesta STR90_6 ($p < 0,0011$), distribuidos en los cromosomas 1, 2, 3, 4, 6, 7, 8, 9 y 10. De acuerdo con el criterio de agrupamiento descrito en la sección Materiales y métodos, estos SNPs fueron agrupados en 36 QTL. Sin embargo, la escasa variabilidad fenotípica observada en este panel, evidenciada por una frecuencia de restauración de sólo 4,63%, no permitió realizar inferencias robustas a partir del análisis GWAS. En consecuencia, se decidió excluir estos resultados del presente trabajo.

5. DISCUSIÓN

La caracterización de la estructura genética de la población es un paso esencial previo al análisis de mapeo por asociación, dado que la presencia de subestructuras o relaciones de parentesco no controladas puede inducir asociaciones erróneas entre marcadores y rasgos fenotípicos (Price et al., 2006; Yu et al., 2006). En este estudio, el análisis de componentes principales (ACP) se utilizó para describir y considerar la estructura poblacional en los paneles LPM-M y LPM-H, permitiendo de esta manera resumir la variación genética entre individuos en un conjunto reducido de ejes ortogonales (CP).

Usualmente, en la mayoría de los trabajos de GWAS el número de CPs es del orden de uno a diez, pero en muchos estudios pueden utilizarse más (Grinde et al., 2024). En este estudio, los diez primeros componentes principales explicaron el 45,26 % de la variación genética total en LPM-H y el 55,4 % en LPM-M. En ambos paneles, el ACP mostró que las líneas no se agrupan en subpoblaciones definidas, sino que exhiben una distribución continua. La ausencia de agrupamientos nítidos refuerza la adecuación del modelo mixto empleado en el GWAS, que combina componentes principales y matriz de parentesco para reducir las asociaciones espurias (Yu et al., 2006).

Por otro lado, se calculó el coeficiente de parentesco relativo (K) estimado a partir de los datos genotípicos el cual reflejó el grado de relación genética entre los individuos dentro de cada panel. En el panel LPM-M, el 56,9% de las comparaciones entre pares de líneas presentaron un coeficiente de parentesco igual a cero, indicando ausencia de relación genética detectable entre ellas. Un 24% de las comparaciones mostró valores de K entre 0,1 y 0,5, lo cual representa relaciones de parentesco débiles a moderadas, y solo un 4% de las comparaciones superó valores de 1,5, reflejando parentescos más estrechos. En el panel LPM-H, el 59,9% de las comparaciones entre pares de líneas presentaron valores de K iguales a cero, el 21% valores intermedios (0,1–0,5) y un 2% valores superiores a 1,5. Estos resultados sugieren que ambos paneles presentan una base genética con predominio de individuos no emparentados o con relaciones distantes. Este tipo de distribución es deseable en estudios de asociación, ya que reduce la probabilidad de detectar asociaciones debidas al parentesco en lugar de a una relación causal entre marcador y rasgo (Yu et al., 2006; Stich et al., 2008). No obstante, la presencia de un pequeño porcentaje de comparaciones con valores altos de K indica que algunos individuos podrían compartir segmentos genómicos idénticos por ascendencia reciente, lo que justifica la

inclusión del término de parentesco dentro del modelo mixto (K + PCA) que fue utilizado en el GWAS.

Por su parte, el patrón de DL observado en el panel LPM-H estudiado mostró valores elevados de r^2 a distancias cortas ($r^2=0,78$ en 0–250 pb) y un decaimiento lento con la distancia física, manteniéndose en niveles intermedios a distancias grandes ($r^2 \approx 0,39$ – $0,40$ entre 0,5–10 Mb). Un valor de r^2 cercano a 1 indica un fuerte desequilibrio de ligamiento (los alelos casi siempre se heredarían juntos), mientras que un valor cercano a 0 sugiere una asociación aleatoria entre los alelos (los alelos se heredarían de manera independiente). Un patrón similar se observó en LPM-M, lo cual suele ser consistente con lo reportado para paneles de líneas endogámicas de maíz y podría deberse a la estructura poblacional y a la presencia de haplotipos conservados entre líneas con ancestros comunes (Flint-Garcia et al., 2003). Ambos patrones de DL tienen implicancias sobre el mapeo por asociación, por un lado, el poder de detección se incrementa al permitir que algunos marcadores identifiquen variantes dentro del mismo haplotipo, y por otro, la resolución se reduce dado que probablemente varios SNPs significativos en el mismo cromosoma y cercanos no sean asociaciones independientes lo cual dificulta la delimitación precisa del locus causal (Yu et al., 2006; Truntzler et al., 2012; Korte y Farlow, 2013).

En el contexto de este trabajo, ambos patrones de DL observados confirman que el GWAS fue realizado en un rango de resolución compatible con la extensión de los bloques de desequilibrio, lo que permite interpretar los QTL detectados como regiones genómicas candidatas, aunque no necesariamente con resolución génica. En futuros trabajos, la incorporación de datos de secuenciación o genotipado de alta densidad podría permitir un mapeo detallado de las regiones detectadas y una identificación más precisa de los genes involucrados.

En este estudio, para evaluar la restauración de la androesterilidad, un panel representativo de líneas del programa de mejoramiento de una empresa de la industria de semillas en Argentina fue cruzado con un tester estéril conocido para obtener las F1 en la siguiente generación que permitió realizar la evaluación fenotípica en cuatro ambientes distintos.

La utilización de F1 como metodología para evaluar genes restauradores de la fertilidad presenta una serie de limitaciones, tales como la dificultad en la identificación de restauradores mayores y menores y de interacciones epistáticas, así como la imposibilidad de observar la segregación que ocurriría en poblaciones F2 limitando el estudio de la herencia de los genes involucrados. A su vez presenta ventajas como la facilidad de fenotipar muchos individuos en una sola generación

con el consiguiente ahorro de tiempo y recursos. El uso de las F1 como método de evaluación en este estudio presentó ventajas dado que fue posible realizarlo en una generación para fenotipar muchos genotipos sin necesidad de crear la línea pura CMS, con el consecuente ahorro de tiempo y recursos, priorizando líneas más promisorias (estériles y estables). Sin embargo, para que haya restauración de la fertilidad no es suficiente con la presencia de un *Rf* dominante, ya que su efecto puede verse suprimido o modulado por genes modificadores nucleares o por interacciones epistáticas núcleo-citoplasma (Huang et al., 2013) que pueden pasar desapercibidos en generaciones F1. Adicionalmente, en una F1 si el parental fértil aporta un gen restaurador dominante, la fertilidad se restaura completamente, pero no se puede diferenciar si la restauración es por un solo gen mayor o por la suma de varios genes menores. Para mapear genes restauradores y cuantificar su efecto con mayor resolución, sería preferible trabajar con poblaciones segregantes (F2, retrocruzas, líneas recombinantes), donde se puede asociar el fenotipo con el genotipo de manera más asertiva (Jaqueth et al., 2020). Estos puntos expuestos determinan, en el marco de este trabajo, las limitaciones identificadas en la utilización de las F1-CMS como panel de estudio.

Por otro lado, si bien la escala tradicional -que asigna una puntuación acumulativa de 1 a 9 en distintos intervalos temporales relacionados con la emisión de estigmas- utilizada en la evaluación logró captar variabilidad fenotípica en cada uno de los ambientes de evaluación, el análisis posterior de la información fenotípica incluyó un enfoque novedoso en el cual se incorporó la ponderación temporal del riesgo de ruptura. La ventaja principal de la aproximación sobre la variable respuesta desarrollada en este trabajo radica en la integración de medidas temporalmente consecutivas en un solo índice para una comparación genotípica y selección dentro del programa de mejoramiento de líneas a ser utilizadas como hembras más promisorias. En el presente estudio, se observó, en términos relativos, una alta variabilidad fenotípica en el panel LPM-H representado por grupos de genotipos que exhibieron ruptura baja, media y alta (46%, 21,5 % y 31,8% del panel respectivamente) a través de los ambientes, lo que permitió identificar marcadores significativamente asociados en este panel. En contraste, para el panel LPM-M se observó una baja variabilidad fenotípica relativa dado que solo cinco líneas presentaron restauración de la fertilidad mientras que el resto permaneció completamente estéril. Esta escasa representación de genotipos restauradores impidió capturar con precisión la variación alélica necesaria para establecer relaciones robustas entre genotipo y fenotipo (Korte y Farlow, 2013).

Respecto al mapeo por asociación, según lo reportado por diversos autores, la efectividad en los análisis de GWAS para el análisis de caracteres complejos depende de múltiples premisas que deben ser tenidas en cuenta *ex ante*, por un lado, los datos fenotípicos debieran seguir distribuciones normales o aproximadamente normales dado que las distribuciones no normales pueden inflar el error de tipo II y consecuentemente reducir el poder estadístico en el mapeo. El tamaño de la población también es de importancia, así como su estructura, el desequilibrio de ligamiento que es clave para definir bloques de haplotipos, así como la densidad de los marcadores para asegurar la cobertura de cada bloque de haplotipo, evitando así la pérdida de asociaciones potenciales. La frecuencia alélica también es otro factor importante, si bien la mayoría de los estudios de GWAS se enfocan únicamente en variantes comunes (es decir, aquellas con una frecuencia alélica $\geq 5\%$), ignorando las variantes raras con una frecuencia del alelo menor (MAF) $< 5\%$ (como es el caso de este trabajo), estas variantes raras podrían explicar una parte importante de la variación fenotípica, aunque su baja frecuencia puede conducir a una falta de poder de resolución. Asimismo, el modelo estadístico elegido para identificar marcadores genéticos que estén significativamente asociados con la variación en el fenotipo es otro de los factores importantes en el éxito de los GWAS. Los primeros modelos de GWAS desarrollados fueron los modelos lineales generales (MLG), los cuales incluían la estructura poblacional para reducir el sobreajuste y falsos positivos en las regiones genómicas. Sin embargo, estos modelos basados en MLG tienen poca potencia para considerar las relaciones de parentesco entre los individuos, por lo cual se desarrollaron nuevos métodos de GWAS basados en modelos lineales mixtos (MLM) que permitieron incluir simultáneamente en los análisis la estructura poblacional o análisis de componentes principales como efecto fijo y la matriz de parentesco se ajusta como una estructura de varianza-covarianza aleatoria entre los genotipos (Torkamaneh et al., 2022). En este estudio, este enfoque de MLM+ACP+K fue el que demostró un mejor ajuste para el análisis de LPM-H.

Adicionalmente, otro parámetro importante en los estudios de MA es el nivel de significancia establecido para declarar un marcador como estadísticamente asociado al carácter. En este sentido, si bien en la bibliografía está ampliamente difundido el método de Bonferroni como método de corrección para ajustar el nivel de significancia, este método asume que todas las pruebas son independientes lo cual representa una limitación para aplicarlo en este estudio, dado que en presencia de DL elevado observado en ambos paneles Bonferroni se vuelve demasiado conservador al ignorar la dependencia entre marcadores aumentando el error de tipo II. Por ello,

se aplicó un umbral basado en el número efectivo de pruebas (MEFF), estimado mediante depuración por desequilibrio de ligamiento (LD Pruning), resultando en un nivel de corte de $p < 0,0006756757$ para el panel LPM-H y $p < 0,0011111111$ para el panel LPM-M. La elección de este enfoque está alineada con recomendaciones metodológicas que proponen ajustar por la dependencia entre pruebas reconociendo entonces la relación entre marcadores, y mejorar en consecuencia la potencia relativa respecto a Bonferroni en poblaciones como la de este estudio (Nyholt, 2004; Li y Ji, 2005; Gao et al., 2008).

En el presente estudio, el mapeo por asociación demostró ser una herramienta efectiva que permitió identificar cinco QTL asociados a la restauración de la androesterilidad en el panel LPM-H estudiado a partir de los 50 SNPs con asociaciones significativas identificados en cromosomas 1, 2, 3 y 8.

La posición de estos QTL fue comparada con estudios previos, observando como resultado que algunos de los QTL detectados en este estudio coinciden con regiones reportadas en estudios anteriores de mapeo por asociación o secuenciación, mientras que otros se encuentran en regiones que no fueron reportadas previamente.

El QTL denominado en este trabajo como *Qrf8.1* coincide con intervalos de confianza o se ubica en regiones cercanas del mismo bin en el brazo corto del cromosoma 8 anteriormente reportados (Kheyr-Pour et al., 1981; Sisco et al., 1991). Este gen mayor denominado *Rf4* es un gen restaurador dominante capaz de restaurar completamente la fertilidad en CMS-C que posteriormente fue clonado empleando una población de mapeo F3 de gran tamaño (Jaqueth et al., 2020).

En el presente estudio, este QTL denominado *Qrf8.1* explicó el 8,4% de la variación en el panel LPM-H y se observó que cuando está presente, la fertilidad fue completamente restaurada para todos los genotipos de este panel CMS-C testeadas en este estudio.

Por otro lado, *Qrf2.1* y *Qrf2.2* fueron identificados en este estudio en bins 2.02 y 2.12 explicando el 7,1% y el 9,8% de la variación respectivamente. En el fenotipo evaluado, se verificó la respuesta a la restauración para estos dos QTL en presencia de *Qrf8.1* en estado alélico “E” y se observó restauración parcial en ambos casos (64% y 74%). El brazo largo del cromosoma 2 incluye regiones ricas en genes restauradores para CMS, como el *Rfm1* (≈ 234 Mb en v4), reportado por Chen y Chen (1992) y refinado por Yang et al., (2010), en una posición muy similar a la encontrada

en este estudio. Algunos autores mencionan a *Rfm1* como un restaurador menor o como el “primer gen identificado antes del mapeo molecular”, reconociendo que fue el antecedente histórico del estudio de CMS-C en maíz. Zhang et al. (2023) identificaron y mapearon en bin 2.09 un nuevo gen restaurador mayor para CMS-C, denominado *Rf12*; apenas 0,064 Mb desde el borde izquierdo de *Qrf2.2* reportado en este estudio. El brazo largo del cromosoma 2 es de especial interés dado que contiene genes restauradores asociados con los tres sistemas principales de androesterilidad citoplasmática de maíz: CMS-C, CMS-T y CMS-S (Kohls et al., 2011).

Adicionalmente, fue reportado en bin 3.06 del cromosoma 3 un QTL de efecto menor que involucró restauración parcial de fertilidad (Kohls et al., 2011). En este estudio, el QTL identificado se denominó *Qrf3.1* en bin 3.11. Dado que estas distancias son muy superiores a la caída del DL para el panel estudiado, sería poco probable que este QTL detectado corresponda a la misma región que el reportado por Kohls et al., 2011, aunque estudios posteriores mediante otras técnicas, como el ML, serán necesarios para su confirmación.

Finalmente, en lo que respecta al panel de hembras, *Qrf1.1* fue el último de los QTL que arrojó una asociación significativa a la variable respuesta AUBPC en el estudio de GWAS. Este QTL fue clasificado como falso positivo ya que no evidenció diferencias en el fenotipo, dado que la ratio de baja y alta ruptura para el panel de hembras se ubicó en 46% para estado alélico “E” y 49% para estado alélico “R”, no siendo posible identificar proporciones altas de restauración que lo identifiquen como un QTL de interés para la restauración de la androesterilidad.

Por todo lo antedicho, el análisis de asociación genómica (GWAS) permitió identificar múltiples marcadores significativamente asociados con la restauración de la androesterilidad citoplasmática CMS-C en el panel LPM-H. Asimismo, las regiones genómicas detectadas mostraron coincidencia con QTLs previamente reportados para genes restauradores de fertilidad previamente descritos, validando el enfoque aplicado y respaldando la consistencia de las asociaciones obtenidas en el material femenino.

En contraste, las características del panel LPM-M no permitieron realizar un análisis de asociación válido genotipo-fenotipo: de las 108 líneas evaluadas, únicamente cinco presentaron restauración de la fertilidad, mientras que el resto resultaron completamente estériles.

En estudios de GWAS, una distribución fenotípica tan desbalanceada reduce de manera drástica el poder estadístico para detectar asociaciones genuinas (Korte y Farlow, 2013). La baja

variabilidad fenotípica disminuye la varianza explicable por los marcadores, impidiendo detectar efectos alélicos causales reales. Además, la escasa frecuencia de alelos restauradores en el panel limita la capacidad del modelo mixto para establecer relaciones robustas entre genotipo y fenotipo, incluso cuando se corrige la estructura poblacional mediante componentes principales o matriz de parentesco como fue el caso de este estudio (Yu et al., 2006).

Desde un punto de vista biológico, los resultados pueden interpretarse considerando que el panel LPM-M probablemente no porta alelos restauradores activos para el citoplasma CMS-C, o bien que dichos alelos se encuentran en baja frecuencia o no segregan en esta población. Esto coincide con el hecho de que la restauración del sistema CMS-C sea propuesto como un carácter nuclear condicionado por genes dominantes específicos cuya presencia depende del origen genético del germoplasma (Tang et al., 2001). En consecuencia, la falta de restauración observada en la mayoría de las líneas macho sugiere que los genes restauradores efectivos para este sistema citoplasmático no se encuentran representados en dicho panel, lo cual limita la posibilidad de detectar asociaciones significativas. En estos casos, el bajo número de genotipos portadores del fenotipo alternativo (restaurado) tiende a generar que la metodología no sea capaz de discernir señales estadísticamente significativas por azar o ruido estructural (Kaler y Purcell, 2019). Por lo antedicho, los resultados de GWAS en el panel LPM-M fueron excluidos del trabajo.

Adicionalmente, es importante destacar que este patrón de restauración contrastante observado entre los paneles LPM-H y LPM-M resulta llamativo para un programa de mejoramiento con CMS-C operativo. Lo esperable sería que las líneas a ser utilizadas como hembras no sean portadoras de genes restauradores, y que las líneas a ser utilizadas como machos contaran con genes *Rf* capaces de asegurar la restauración de la fertilidad en el híbrido comercial. Por lo antedicho, resulta evidente que, en este programa de mejoramiento, el esquema CMS-C aún no se encuentra completamente consolidado, por lo que resulta necesario implementar estrategias específicas en ambos paneles. En el panel de líneas a ser utilizadas como hembras, será necesario avanzar en la selección en contra de alelos restauradores; en este sentido, los marcadores identificados en este estudio, particularmente el SNP asociado a *Qrf8.1*, debido a su mayor efecto mayor y a su clara asociación con el fenotipo de restauración, constituyen herramientas potencialmente útiles para la selección asistida en contra de dichos alelos, con el objetivo de desarrollar líneas CMS más estables y con menor riesgo de ruptura de la androesterilidad.

En el panel de líneas a ser utilizadas como machos, en cambio, será necesario definir e implementar una estrategia de restauración de la fertilidad como línea de trabajo futuro, mediante el desarrollo o la incorporación de líneas macho restauradoras.

En ausencia de líneas macho completamente restauradoras, otra alternativa para la producción comercial de híbridos bajo CMS consiste en el uso de mezclas de componentes fértiles y estériles. La proporción entre ambos componentes no puede definirse de manera arbitraria, sino que debe establecerse experimentalmente, a partir de ensayos que permitan determinar la cantidad mínima de componente fértil necesaria para asegurar una polinización adecuada, considerando la producción de polen, la sincronía floral, la aptitud polinizadora y la estabilidad del sistema en distintos ambientes.

En conjunto, los resultados del presente trabajo refuerzan la utilidad del enfoque GWAS para la detección de loci asociados con la restauración de la androesterilidad CMS-C en maíz cuando se dispone de paneles con adecuada variación genética y fenotípica, como se evidenció en el panel LPM-H. Por otro lado, el panel LPM-M ilustra cómo la falta de variabilidad fenotípica y la baja frecuencia de alelos restauradores pueden limitar la capacidad del GWAS para detectar asociaciones biológicamente relevantes. Estos hallazgos resaltan la importancia de diseñar paneles balanceados y representativos, y de complementar el análisis de asociación con estudios de validación funcional y mapeo fino para confirmar la identidad y efecto de los loci candidatos.

6. CONCLUSIONES

En el presente estudio, se caracterizó la estructura genética de dos grupos heteróticos de un programa de mejoramiento de maíz híbrido a través de la implementación de distintas técnicas como análisis de componentes principales, cálculo de coeficiente de parentesco relativo y desequilibrio de ligamiento. Para la mayoría de las comparaciones entre líneas, el coeficiente de parentesco fue bajo, indicando falta de estratificación marcada en ambas poblaciones y el DL presentó una caída de 1,900,000 pb para LPM-H y 1,232,454 pb para LPM-M indicando una buena potencia para el conjunto de 27,882 y 18,690 SNPs utilizados en este mapeo por asociación.

Adicionalmente, se caracterizó la variabilidad fenotípica en la restauración de la androfertilidad a través de un método novedoso no reportado previamente para un uso análogo como el AUBPC, proporcionando una visión integral de la progresión de las mediciones de esterilidad especialmente útil en un contexto de producción de semillas.

Por lo tanto, se puede confirmar la hipótesis de que existen regiones genéticas nucleares asociadas a la restauración de la androesterilidad citoplasmática CMS-C y estas, a su vez, pueden ser identificables mediante un estudio de mapeo por asociación que involucre sets de marcadores moleculares y datos fenotípicos en germoplasma de maíz templado.

El panel de asociación de LPM-H presentó un grado de variabilidad fenotípica para la restauración de la androesterilidad que permitió la detección de zonas del genoma asociadas al fenotipo objetivo, no así el panel de LPM-M que exhibió una baja frecuencia de alelos restauradores. En el caso de LPM-H, la implementación del GWAS con el modelo MLM+ACP+K permitió estudiar las regiones que regulan la restauración de la androesterilidad, y se detectaron cinco QTL que explicaron entre el 5,2% y el 9,8% de la variación. La ubicación de la mayoría de ellos coincide con lo reportado en la bibliografía o con regiones similares, a excepción de *Qrf1.1* que constituyó un falso positivo. Otros QTL asociados con la restauración de la fertilidad en el sistema CMS-C no fueron detectados en el presente estudio, probablemente debido a un número insuficiente de individuos en los grupos contrastantes, a la forma de caracterización fenotípica o a las limitaciones de detección propias del GWAS para este tipo estudio. Asimismo, es importante remarcar que el marco de estudio corresponde a F1s, y por lo tanto la necesidad de realizar mapeo sobre F2 y/o F3 de cruas dirigidas específicamente para tal fin en el caso de requerir mayor resolución en el mapeo de asociación.

Adicionalmente, como se discutió previamente, es importante mencionar que este trabajo debe complementarse con una validación de los resultados entre las F1s y las líneas puras. Esto se debe a que el genotipo, en interacción con el ambiente tiene impacto en el intervalo anthesis-silking (ASI) (Bolaños y Edmeades, 1996; Silva et al., 2022) por lo que características no estrictamente relacionadas con los genes restauradores podrían impactar en la probabilidad de ruptura durante a la ventana temporal de emisión de estigmas en un genotipo CMS.

Adicionalmente, es importante tener en consideración que la restauración de la fertilidad en los sistemas CMS está controlada no solo por genes nucleares, sino también por genes mitocondriales. Esta característica podría influir en la eficiencia de detección de QTL nucleares (Zheng et al., 2020). Para trabajos futuros, sería fundamental identificar y localizar más genes restauradores para el desarrollo de líneas androestériles y restauradoras estables en maíz.

7. BIBLIOGRAFÍA

American Phytopathological Society. (2024). Area under the disease progress curve (AUDPC). APSnet.

Beckett, J. B. (1971). Classification of male-sterile cytoplasm in maize (*Zea mays* L.). *Crop Science*, 11, 724–727.

Birchler, J. A. (2016). Hybrid vigour characterized. *Nature*, 537, 620.

Bolaños, J., & Edmeades, G. O. (1996). The importance of the anthesis–silking interval in breeding for drought tolerance in tropical maize. *Field Crops Research*, 48, 65–80.

Bonferroni, C. (1936). *Teoria statistica delle classi e calcolo delle probabilità*. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze, 8, 3–62.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23, 2633–2635.

Breseghello, F., & Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat using single nucleotide polymorphisms. *The Plant Genome*, 46, 1665–1677.

Bueckmann, H., Schiemann, J., & Thiele, K. (2014). Influence of air temperature on the stability of cytoplasmic male sterility (CMS) in maize (*Zea mays* L.). *Agricultural Sciences*, 5, 571–581.

Bueckmann, H., Thiele, K., & Schiemann, J. (2016). Influence of soil moisture and air temperature on the stability of cytoplasmic male sterility (CMS) in maize (*Zea mays* L.). *Agricultural Sciences*, 7, 70–81.

Chen, L., & Liu, Y. (2014). Male sterility and fertility restoration in crops. *Annual Review of Plant Biology*, 65, 579–606.

Chen, S. J., & Chen, W. C. (1992). The chromosomal location of restoring genes of C-cytoplasmic male-sterile in maize. *Acta Agriculturae Universitatis Henanensis*, 2, 125–130.

Collinson, S., Hamdziripi, E., De Groote, H., Ndegwa, M., Cairns, J. E., Albertsen, M., Ligeyo, D., Mashingaidze, K., & Olsen, M. S. (2022). Incorporating male sterility increases hybrid maize yield in low input African farming systems. *Communications Biology*, 5(1), 729.

Cullis, B. R., Smith, A. B., & Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics*, 11, 381–393.

Duggal, P., Gillanders, E. M., Holmes, T. N., & Bailey-Wilson, J. E. (2008). Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genomics*, 9, 516.

Duvick, D. N. (1965). Cytoplasmic pollen sterility in corn. *Advances in Genetics*, 13, 1–56.

East, E. M. (1908). Inbreeding in corn. In *Connecticut Agricultural Experiment Station Report 1907* (pp. 419–428).

Edera, A., Garcia, L., Marfil, C. F., & Sanchez-Puerta, M. V. (2019). Male sterility and somatic hybridization in plant breeding. *Revista de la Facultad de Ciencias Agrarias, UNCuyo*, 51(2), 475–486.

Erenstein, O., Jaleta, M., Sonder, K., Mottaleb, K., & Prasanna, B. M. (2022). Global maize production, consumption and trade: Trends and R&D implications. *Food Security*, 14(5), 1295–1319.

Ersoz, E. S., Yu, J., & Buckler, E. S. (2007). Applications of linkage disequilibrium and association mapping in crop plants. In C. Kole (Ed.), *Genomics-Assisted Crop Improvement: Vol. 1. Genomics Approaches and Platforms* (pp. 97–119). Springer, Dordrecht.

FAOStat. (2023). *FAO Stat*. FAO, Roma.

Feng, P. C. C., Qi, Y., Chiu, T., Stoecker, M. A., Schuster, C. L., Johnson, S. C., Fonseca, A. E., & Huang, J. (2014). Improving hybrid seed production in corn with glyphosate-mediated male sterility. *Pest Management Science*, 70, 212–218.

Fisher, R. A. (1935). *The design of experiments*. Oliver & Boyd, Edinburgh.

Flint-Garcia, S. A., Thornsberry, J. M., & Buckler, E. S. (2003). Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology*, 54, 357–374.

Fox, T., DeBruin, J., Haug Collet, K., Trimnell, M., Clapp, J., Leonard, A., Li, B., Scolaro, E., Collinson, S., Glassman, K., Wang, Y., Thilmony, R., Gengenbach, B., Zhang, Z., Clemente, T., & Albertsen, M. (2017). A single point mutation in Ms44 results in dominant male sterility and improves nitrogen use efficiency in maize. *Plant Biotechnology Journal*, 15(8), 1067–1079.

Fu, D., Xiao, M., Hayward, A., Fu, Y., Liu, G., Jiang, G., & Zhang, H. (2014). Utilization of crop heterosis: A review. *Euphytica*, 197(2), 161–173.

Gabay-Laughnan, S., Laughnan, J. R., & Cisar, G. (2004). Molecular-genetic characterization of CMS-S restorer-of-fertility alleles identified in Mexican maize and teosinte. *Genetics*, 166, 959–970.

Gao, X., Starmer, J., & Martin, E. R. (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genetic Epidemiology*, 32(4), 361–369.

Gezan, S.A., de Oliveira, A.A., Galli, G., and Murray, D. (2022). *ASRgenomics: An R package with Complementary Genomic Functions*. Version 1.1.0. VSN International, Hemel Hempstead, United Kingdom.

Ghețe, A. B., Haș, V., Vidican, R., Copândeian, A., Ranta, O., Moldovan, C. M., & Duda, M. M. (2020). Influence of detasseling methods on seed yield of some parent inbred lines of Turda maize hybrids. *Agronomy*, 10(5), 729.

Gilmour, A. R., Cullis, B. R., & Verbyla, A. P. (1997). Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agricultural, Biological, and Environmental Statistics*, 2, 269–293.

Gramene. *Zea mays* B73 RefGen_v4 (Zm-B73-REFERENCE-GRAMENE-4.0). Disponible en el portal Gramene.

Grinde, K. E., Browning, B. L., Reiner, A. P., Thornton, T. A., & Browning, S. R. (2024). Adjusting for principal components can induce spurious associations in genome-wide association studies in admixed populations. *PLOS Genetics*, 20(4), e1011242.

Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 31, 423–447.

Henderson, C. R. (1984). *Applications of linear models in animal breeding*. University of Guelph.

Hill, W. G., & Weir, B. S. (1994). Maximum-likelihood estimation of gene frequency disequilibrium. *American Journal of Human Genetics*, 54, 705–714.

Huang, J., Xiang, J., Liu, T., Rong, J., Wang, Y., Lu, Q., Tang, W., Wen, W., & Cao, M. (2012). Expression characterization of genes for CMS-C in maize. *Protoplasma*, 249, 1119–1127.

Huang, W., Xu, C., Marand, A. P., & Jiang, J. (2013). Molecular and epigenetic mechanisms of cytoplasmic male sterility in plants. *Journal of Experimental Botany*, 64(3), 1063–1076.

Hyten, D. L., Choi, I.-Y., Song, Q., Shoemaker, R. C., Nelson, R. L., Costa, J. M., Specht, J. E., & Cregan, P. B. (2007). Highly variable patterns of linkage disequilibrium in multiple soybean populations. *Genetics*, 175(4), 1937–1944.

Jaqueth, J. S., Hou, Z., Zheng, P., et al. (2020). Fertility restoration of maize CMS-C altered by a single amino acid substitution within the Rf4 bHLH transcription factor. *The Plant Journal*, 101, 101–111.

Jeger, M., & Viljanen-Rollinson, S. (2001). The use of the area under the disease-progress curve (AUDPC) to assess quantitative disease resistance in crop cultivars. *Theoretical and Applied Genetics*, 102, 32–40.

Jones, D.F. (1918). The effects of inbreeding and crossbreeding upon development. *Connecticut Agricultural Experiment Station Bulletin*, 207, 5–100.

Jones, D. F., Stinson, H. T. J., & Khoo, U. (1957). Pollen restoring genes. *Connecticut Agricultural Experiment Station Bulletin*, 610, 1–43.

Kaler, A. S., & Purcell, L. C. (2019). Estimation of a significance threshold for genome-wide association studies in soybean. *Crop Science*, 59(2), 652–659.

Khanal, A., Adhikari, P., Kaiser, C., Lipka, A. E., Jamann, T. M., & Mideros, S. X. (2022). Genetic mapping of sorghum resistance to an Illinois isolate of *Colletotrichum sublineola*. *The Plant Genome*, 15, e20243.

Kheyr-Pour, A., Gracen, V. E., & Everett, H. L. (1981). Genetics of fertility restoration in the C-group of cytoplasmic male sterility in maize. *Genetics*, 98, 379–388.

Kohls, S., Stamp, P., Knaak, C., & Messmer, R. (2011). QTL involved in the partial restoration of male fertility of C-type cytoplasmic male sterility in maize. *Theoretical and Applied Genetics*, 123(2), 327–338.

Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods*, 9, 29.

Laughnan, J., & Gabay-Laughnan, S. (1983). Cytoplasmic male sterility in maize. *Annual Review of Genetics*, 17(1), 27–4.

Lee, S. J., Earle, E. D., & Gracen, V. E. (1980). The cytology of pollen abortion in S cytoplasmic male-sterile corn anthers. *American Journal of Botany*, 67, 237–245.

Lewontin, R. C. (1964). The interaction of selection and linkage, I, General considerations; heterotic models. *Genetics*, 49, 49–67.

Li, J., & Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, 95, 221–227.

Matsuoka, Y., Vigouroux, Y., Goodman, M. M., Sanchez, G., Buckler, E., & Doebley, J. A. (2002). A single domestication for maize is shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences of the United States of America*, 99(9), 6080–6084.

McClintock, B. (1929). Chromosome morphology in *Zea mays*. *Science*, 69, 629–629.

Möhring, J., & Piepho, H. P. (2009). Comparison of BLUE and BLUP in genomic selection. *Crop Science*, 49(6), 1977–1984.

Nordborg, M., & Tavaré, S. (2002). Linkage disequilibrium: What history has to tell us. *Trends in Genetics*, 18(2), 83–90.

Nyholt, D. R. (2004). A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other, *American Journal of Human Genetics*, 74(4), 765–769.

Oraguzie, N. C., Rikkerink, E. H. A., Gardiner, S. E., & De Silva, H. N. (Eds.). (2007). *Association mapping in plants*. Springer.

Piepho, H. P., Möhring, J., Melchinger, A. E., & Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica*, 161, 209–228.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8), 904–909.

Pritchard, J. K., & Przeworski, M. (2001). Linkage disequilibrium in humans: Models and data, *American Journal of Human Genetics*, 69, 1–14.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959.

Rafalski, A. (2010). Association genetics in crop improvement. *Current Opinion in Plant Biology*, 13, 174–180.

Remington, D. L., Thornsberry, J. M., Matsuoka, Y., Wilson, L. M., & Whitt, S. R. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences of the USA*, 98, 11479–11484.

Ritchie, S. W., Hanway, J. J., & Benson, G. O. (1986). How a corn plant develops. Iowa State University, Special Report 48.

Rogers, J. S., & Edwardson, J. R. (1952). The utilization of cytoplasmic male-sterile inbreds in the production of corn hybrids. *Agronomy Journal*, 44, 8–13.

Sage, G. C. M. (1976). Nucleo-cytoplasmic relationships in wheat. *Advances in Agronomy*, 28, 265–298.

Schandry, N. (2017). A practical guide to visualization and statistical analysis of *R. solanacearum* infection data using R. *Frontiers in Plant Science*, 8.

Schnable, P. S., Wise, R. P., & Pring, D. R. (1998). The molecular basis of cytoplasmic male sterility and fertility restoration. *Trends in Plant Science*, 3, 175–180.

Shull, G. H. (1908). The composition of a field of maize. *American Breeders' Association Reports*, 4, 296–301.

Shull, G. H. (1909). A pure-line method of corn breeding. *American Breeders' Association Reports*, 5, 51–59.

Silva, P. C., Sánchez, A. C., Opazo, M. A., Mardones, L. A., & Acevedo, E. A. (2022). Grain yield, anthesis-silking interval, and phenotypic plasticity in response to changing environments: Evaluation in temperate maize hybrids. *Field Crops Research*, 285, 108583.

Sisco, P. H., Wernsman, E. A., & Dudley, J. W. (1991). Nuclear restoration of C-type cytoplasmic male sterility in maize. *Crop Science*, 31(3), 646–651.

Simko, I., Haynes, K. G., Ewing, E. E., Costanzo, S., Christ, B. J., & Jones, R. W. (2004b). Mapping genes for resistance to *Verticillium albo-atrum* in tetraploid and diploid potato populations using haplotype association tests and genetic linkage analysis. *Molecular Genetics and Genomics*, 271(4), 522–531.

Skøt, L. L. (2005). An association mapping approach to identify flowering time genes in natural populations of *Lolium perenne* (L.). *Molecular Breeding*, 15, 233–245.

Sparks, A., Esker, P., Bates, M., Dall Acqua, W., Guo, Z., Segovia, V., Silwal, S. D., Tolos, S., & Garrett, K. (2008). Disease progress over time. *Plant Health Instructor*.

Speed, D., & Balding, D. J. (2015). Relatedness in the post-genomic era: Is it still useful? *Nature Reviews Genetics*, 16, 33–44.

Stich, B., Mohring, J., Piepho, H.-P., Heckenberger, M., Buckler, E. S., & Melchinger, A. E. (2008). Comparison of mixed-model approaches for association mapping. *Genetics*, 178(3), 1745–1754.

Tang, J., Liu, Z., Chen, W., Li, X., & Xiao, Y. (2001). The SSR markers of the main restorer genes for CMS-C cytoplasmic male sterility in maize. *Scientia Agricultura Sinica*, 34, 592–596.

Tang, R., Zhuang, Z., Bian, J., Ren, Z., Ta, W., & Peng, Y. (2024). GWAS and meta-QTL analysis of kernel quality-related traits in maize. *Plants*, 13(19), 2730.

The jamovi project. (2022). jamovi (Version 2.3) [Computer Software]. Recuperado de <https://www.jamovi.org>

Thornsberry, J. M., Goodman, M. M., Doebley, J., Kresovich, S., Nielsen, D., & Buckler, E. S. (2001). Dwarf8 polymorphisms associate with variation in flowering time, *Nature Genetics*, 28, 286–289.

Torkamaneh, D., & Belzile, F. (2022). *Genome-wide association studies* (1st ed.). Humana Press.

Truntzler, M., Ranc, N., Sawkins, M., Nicolas, S. D., Manicacci, D., Lespinasse, D., & Charcosset, A. (2012). Diversity and linkage disequilibrium features in a maize association panel relevant for quantitative trait dissection. *Theoretical and Applied Genetics*, 125, 411–424.

Veturi, Y., Kump, K., Walsh, E., Ott, O., Poland, J., Kolkman, J. M., Balint-Kurti, P. J., & Wisser, R. J. (2012). Multivariate mixed linear model analysis of longitudinal data: An information-rich statistical technique for analyzing plant disease resistance. *Phytopathology*, 102(11), 1016–1025.

Wang, B., Liu, H., Liu, Z., et al., (2018). Identification of minor effect QTLs for plant architecture related traits using super high-density genotyping and large recombinant inbred population in maize (*Zea mays*). *BMC Plant Biology*, 18, 17.

Wei, F., et al., (2009). The physical and genetic framework of the maize B73 genome. *PLoS Genetics*, 5(11), e1000734.

Weider, C., Stamp, P., Christov, N., Hüsken, A., & Foueillassar, X. (2009). Stability of cytoplasmic male sterility in maize under different environmental conditions. *Crop Science*, 49, 77–84.

Wilhelm, W. W., Johnson, B. E., & Schepers, J. S. (1995). Yield, quality, and nitrogen use of inbred corn with varying numbers of leaves during detasseling. *Crop Science*, 35, 209–212.

Wu, Y., Fox, T. W., Trimnell, M. R., Wang, L., Xu, R.-J., Cigan, A. M., Huffman, G. A., Garnaat, C. W., Hershey, H., & Albertsen, M. C. (2016). Development of a novel recessive genetic male sterility system for hybrid seed production in maize and other cross-pollinating crops. *Plant Biotechnology Journal*, 14, 1046–1054.

Wright, S. I., & Gaut, B. S. (2005). Molecular population genetics and the search for adaptive evolution in plants. *New Phytologist*, 165, 595–614.

Xiao, Y., Jiang, S., Cheng, Q. et al., (2021). The genetic mechanism of heterosis utilization in maize improvement. *Genome Biol* 22, 14.

Yang, H., Xue, Y., Li, B., et al., (2022). The chimeric gene *atp6c* confers cytoplasmic male sterility in maize by impairing the assembly of the mitochondrial ATP synthase complex. *Molecular Plant*, 15, 872–886.

Yang, J., Liu, X., Yang, X., Zhang, M., & Li, J. (2010). Molecular mechanisms of cytoplasmic male sterility and fertility restoration in crops. *Plant Cell Reports*, 29, 419–432.

Yongming, L., Zhuofan, Z., Yanli, L., Chuan, L., Jing, W., Boxiao, D., Bing, L., Tao, Q., Wenbing, Z., & Moju, C. (2016). A preliminary identification of *Rf*-A619*, a novel restorer gene for CMS-C in maize (*Zea mays* L.). *PeerJ*, 4, e2719.

Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Holland, J. B., Kresovich, S., & Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, 38(2), 203–208.

Zakieh, M. (2023). Novel methods for disease resistance breeding in winter wheat (Doctoral thesis, Swedish University of Agricultural Sciences). *Acta Universitatis Agriculturae Sueciae* 2023:40.

Zhu, C., Gore, M., Buckler, E. S., & Yu, J. (2008). Status and prospects of association mapping in plants. *The Plant Genome*, 1(1), 5-20.

8. ANEXO

Tabla Suplementaria I. Estado alélico determinado por el SNP de mayor significancia en cada uno de los 5 QTL determinados en el mapeo por asociación y fenotipo para todo el panel LPM-H. E: estado alélico que confiere el mejor comportamiento (baja ruptura), R: estado alélico alternativo, -: estado alélico no disponible.

Código de Línea	Qrf1.1	Qrf2.2	Qrf2.1	Qrf3.1	Qrf8.1	AUBPC	Detalle
F_CMSC30	E	E	E	E	E	-4.28	Baja Ruptura
F_CMSC31	E	E	E	E	E	3.22	Baja Ruptura
F_CMSC32	E	E	E	R	E	3.22	Baja Ruptura
F_CMSC33	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC35	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC37	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC38	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC39	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC40	E	E	E	E	E	5.25	Baja Ruptura
F_CMSC41	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC42	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC43	-	E	E	R	E	5.25	Baja Ruptura
F_CMSC45	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC46	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC49	-	E	E	R	E	5.25	Baja Ruptura
F_CMSC51	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC52	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC53	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC54	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC55	E	E	E	R	E	5.25	Baja Ruptura
F_CMSC57	-	E	E	E	E	5.25	Baja Ruptura
F_CMSC58	E	E	E	E	E	5.25	Baja Ruptura
F_CMSC74	R	E	E	R	E	5.25	Baja Ruptura
F_CMSC75	R	E	E	R	E	5.25	Baja Ruptura
F_CMSC76	R	E	E	E	E	5.25	Baja Ruptura
F_CMSC80	R	E	E	R	E	5.25	Baja Ruptura
F_CMSC89	R	E	E	R	E	5.25	Baja Ruptura
F_CMSC94	R	E	E	R	E	5.25	Baja Ruptura
F_CMSC100	R	E	E	R	E	5.25	Baja Ruptura
F_CMSC59	E	E	R	R	E	5.34	Baja Ruptura
F_CMSC60	E	E	E	E	E	5.39	Baja Ruptura
F_CMSC61	E	E	E	R	E	5.41	Baja Ruptura

Tabla Suplementaria I. (continuación) Estado alélico determinado por el SNP de mayor significancia en cada uno de los 5 QTL determinados en el mapeo por asociación y fenotipo para todo el panel LPM-H. E: estado alélico que confiere el mejor comportamiento (baja ruptura), R: estado alélico alternativo, -: estado alélico no disponible.

Código de Línea	Qrf1.1	Qrf2.2	Qrf2.1	Qrf3.1	Qrf8.1	AUBPC	Detalle
F_CMSC62	E	E	E	E	E	5.41	Baja Ruptura
F_CMSC104	R	E	E	E	E	5.49	Baja Ruptura
F_CMSC64	E	E	E	R	E	5.51	Baja Ruptura
F_CMSC65	E	E	E	R	E	5.51	Baja Ruptura
F_CMSC66	E	E	E	R	E	5.52	Baja Ruptura
F_CMSC68	E	E	E	E	E	5.55	Baja Ruptura
F_CMSC105	R	E	E	R	E	5.55	Baja Ruptura
F_CMSC69	E	E	E	R	E	5.59	Baja Ruptura
F_CMSC70	E	E	E	R	E	5.61	Baja Ruptura
F_CMSC107	R	E	E	R	E	5.61	Baja Ruptura
F_CMSC137	R	E	E	R	E	5.61	Baja Ruptura
F_CMSC73	E	E	E	R	E	5.66	Baja Ruptura
F_CMSC144	R	E	R	R	E	5.66	Baja Ruptura
F_CMSC146	R	E	E	R	E	5.68	Baja Ruptura
F_CMSC77	E	E	E	R	E	5.74	Baja Ruptura
F_CMSC148	R	E	E	R	E	5.74	Baja Ruptura
F_CMSC78	E	E	E	R	E	5.75	Baja Ruptura
F_CMSC79	E	E	E	R	E	5.79	Baja Ruptura
F_CMSC81	E	E	E	R	E	5.79	Baja Ruptura
F_CMSC82	E	E	R	R	E	5.79	Baja Ruptura
F_CMSC151	R	E	E	R	E	5.79	Baja Ruptura
F_CMSC191	-	R	E	R	E	5.79	Baja Ruptura
F_CMSC83	E	E	E	R	E	5.88	Baja Ruptura
F_CMSC84	E	E	E	R	E	5.92	Baja Ruptura
F_CMSC85	E	E	E	R	E	5.93	Baja Ruptura
F_CMSC86	E	E	E	R	E	5.95	Baja Ruptura
F_CMSC87	E	E	E	R	E	5.96	Baja Ruptura
F_CMSC88	E	E	E	R	E	5.96	Baja Ruptura
F_CMSC90	E	E	E	R	E	5.96	Baja Ruptura
F_CMSC157	R	R	R	R	E	5.96	Baja Ruptura
F_CMSC197	R	R	E	R	E	5.96	Baja Ruptura
F_CMSC193	E	R	E	R	E	5.97	Baja Ruptura

Tabla Suplementaria I. (continuación) Estado alélico determinado por el SNP de mayor significancia en cada uno de los 5 QTL determinados en el mapeo por asociación y fenotipo para todo el panel LPM-H. E: estado alélico que confiere el mejor comportamiento (baja ruptura), R: estado alélico alternativo, -: estado alélico no disponible.

Código de Línea	Qrf1.1	Qrf2.2	Qrf2.1	Qrf3.1	Qrf8.1	AUBPC	Detalle
F_CMSC194	E	R	E	R	E	6	Baja Ruptura
F_CMSC91	E	E	E	R	E	6.03	Baja Ruptura
F_CMSC92	E	E	E	R	E	6.05	Baja Ruptura
F_CMSC195	E	R	E	E	E	6.05	Baja Ruptura
F_CMSC93	E	E	E	R	E	6.06	Baja Ruptura
F_CMSC161	R	E	E	R	E	6.07	Baja Ruptura
F_CMSC95	E	E	E	R	E	6.09	Baja Ruptura
F_CMSC96	E	E	E	R	E	6.1	Baja Ruptura
F_CMSC97	E	E	R	R	E	6.1	Baja Ruptura
F_CMSC200	R	R	E	R	E	6.1	Baja Ruptura
F_CMSC98	E	E	E	R	E	6.11	Baja Ruptura
F_CMSC99	E	E	E	R	E	6.13	Baja Ruptura
F_CMSC162	R	E	E	R	E	6.16	Baja Ruptura
F_CMSC101	E	E	E	R	E	6.17	Baja Ruptura
F_CMSC102	E	E	E	R	E	6.19	Baja Ruptura
F_CMSC103	E	R	E	R	E	6.25	Baja Ruptura
F_CMSC177	R	E	E	R	E	6.29	Baja Ruptura
F_CMSC106	E	R	E	R	E	6.3	Baja Ruptura
F_CMSC192	R	E	E	R	E	6.3	Baja Ruptura
F_CMSC196	R	E	E	R	E	6.33	Baja Ruptura
F_CMSC202	R	R	E	R	E	6.33	Baja Ruptura
F_CMSC108	E	E	E	R	E	6.35	Baja Ruptura
F_CMSC198	E	R	E	R	E	6.37	Baja Ruptura
F_CMSC109	E	E	E	R	E	6.39	Baja Ruptura
F_CMSC110	E	E	E	R	E	6.45	Baja Ruptura
F_CMSC111	E	E	E	R	E	6.51	Baja Ruptura
F_CMSC112	E	E	E	R	E	6.58	Baja Ruptura
F_CMSC199	E	R	E	R	E	6.68	Baja Ruptura
F_CMSC113	E	E	E	R	E	6.73	Baja Ruptura
F_CMSC114	E	E	E	R	E	6.86	Baja Ruptura
F_CMSC209	R	R	E	R	E	6.89	Baja Ruptura
F_CMSC115	E	E	E	R	E	6.93	Baja Ruptura

Tabla Suplementaria I. (continuación) Estado alélico determinado por el SNP de mayor significancia en cada uno de los 5 QTL determinados en el mapeo por asociación y fenotipo para todo el panel LPM-H. E: estado alélico que confiere el mejor comportamiento (baja ruptura), R: estado alélico alternativo, -: estado alélico no disponible.

Código de Línea	Qrf1.1	Qrf2.2	Qrf2.1	Qrf3.1	Qrf8.1	AUBPC	Detalle
F_CMSC116	E	R	E	E	E	6.97	Baja Ruptura
F_CMSC117	E	E	E	R	E	6.97	Baja Ruptura
F_CMSC118	E	E	-	R	E	7.03	Baja Ruptura
F_CMSC119	E	E	E	R	E	7.05	Baja Ruptura
F_CMSC201	E	R	R	R	E	7.13	Baja Ruptura
F_CMSC120	E	E	E	R	E	7.28	Baja Ruptura
F_CMSC121	E	E	E	R	E	7.3	Baja Ruptura
F_CMSC122	E	E	R	R	E	7.37	Baja Ruptura
F_CMSC123	E	E	E	R	E	7.39	Baja Ruptura
F_CMSC218	R	R	E	R	E	7.39	Baja Ruptura
F_CMSC124	E	E	R	R	E	7.4	Baja Ruptura
F_CMSC203	-	R	E	E	E	7.46	Baja Ruptura
F_CMSC125	E	E	E	R	E	7.58	Baja Ruptura
F_CMSC126	E	E	E	R	E	7.65	Baja Ruptura
F_CMSC127	E	E	E	R	E	7.69	Baja Ruptura
F_CMSC204	E	R	E	R	E	8.03	Baja Ruptura
F_CMSC128	-	E	R	R	E	8.17	Baja Ruptura
F_CMSC129	E	E	E	R	E	8.6	Baja Ruptura
F_CMSC130	E	E	R	R	E	9.33	Baja Ruptura
F_CMSC131	E	E	E	R	E	9.66	Baja Ruptura
F_CMSC132	E	R	E	R	E	9.88	Baja Ruptura
F_CMSC205	E	R	E	R	E	10.15	Media Ruptura
F_CMSC133	E	E	E	R	E	10.28	Media Ruptura
F_CMSC134	E	E	E	R	E	10.6	Media Ruptura
F_CMSC135	E	R	E	R	E	11.27	Media Ruptura
F_CMSC136	E	E	E	R	E	11.66	Media Ruptura
F_CMSC229	R	E	E	R	E	11.74	Media Ruptura
F_CMSC138	E	E	E	R	E	12.23	Media Ruptura
F_CMSC139	E	E	E	R	E	12.27	Media Ruptura
F_CMSC140	E	E	E	R	E	12.32	Media Ruptura
F_CMSC141	E	E	E	R	E	12.41	Media Ruptura

Tabla Suplementaria I. (continuación) Estado alélico determinado por el SNP de mayor significancia en cada uno de los 5 QTL determinados en el mapeo por asociación y fenotipo para todo el panel LPM-H. E: estado alélico que confiere el mejor comportamiento (baja ruptura), R: estado alélico alternativo, -: estado alélico no disponible.

Código de Línea	Qrf1.1	Qrf2.2	Qrf2.1	Qrf3.1	Qrf8.1	AUBPC	Detalle
F_CMSC142	E	E	E	R	E	12.86	Media Ruptura
F_CMSC143	E	E	E	R	E	12.96	Media Ruptura
F_CMSC206	E	R	E	R	E	12.97	Media Ruptura
F_CMSC230	R	E	E	R	E	12.97	Media Ruptura
F_CMSC145	-	E	E	R	E	12.98	Media Ruptura
F_CMSC233	R	-	E	R	E	13.78	Media Ruptura
F_CMSC207	-	R	E	E	E	13.92	Media Ruptura
F_CMSC208	E	R	E	R	E	14.45	Media Ruptura
F_CMSC147	E	E	E	R	E	14.94	Media Ruptura
F_CMSC234	R	E	E	R	E	14.97	Media Ruptura
F_CMSC149	E	E	E	R	E	15.27	Media Ruptura
F_CMSC150	E	E	E	R	E	15.28	Media Ruptura
F_CMSC249	R	R	E	R	E	15.43	Media Ruptura
F_CMSC210	E	R	R	R	E	15.95	Media Ruptura
F_CMSC20	E	R	E	R	R	16	Media Ruptura
F_CMSC211	E	R	E	R	E	16	Media Ruptura
F_CMSC240	R	E	E	R	E	16	Media Ruptura
F_CMSC152	E	E	E	R	E	16.67	Media Ruptura
F_CMSC212	E	R	E	R	E	16.72	Media Ruptura
F_CMSC153	E	E	E	R	E	16.91	Media Ruptura
F_CMSC228	R	R	R	R	R	17.47	Media Ruptura
F_CMSC154	E	E	E	R	E	17.52	Media Ruptura
F_CMSC213	E	R	E	R	E	18.22	Media Ruptura
F_CMSC155	E	E	E	R	E	18.44	Media Ruptura
F_CMSC214	E	R	E	R	E	18.51	Media Ruptura
F_CMSC156	E	E	R	R	E	18.68	Media Ruptura
F_CMSC28	R	E	R	R	E	19.03	Media Ruptura
F_CMSC158	E	E	E	R	E	19.47	Media Ruptura
F_CMSC159	E	E	E	R	E	19.49	Media Ruptura
F_CMSC10	E	R	E	R	R	19.72	Media Ruptura
F_CMSC34	R	E	E	R	E	19.72	Media Ruptura
F_CMSC36	R	E	R	R	E	19.72	Media Ruptura

Tabla Suplementaria I. (continuación) Estado alélico determinado por el SNP de mayor significancia en cada uno de los 5 QTL determinados en el mapeo por asociación y fenotipo para todo el panel LPM-H. E: estado alélico que confiere el mejor comportamiento (baja ruptura), R: estado alélico alternativo, -: estado alélico no disponible.

Código de Línea	Qrf1.1	Qrf2.2	Qrf2.1	Qrf3.1	Qrf8.1	AUBPC	Detalle
F_CMSC47	R	R	E	R	E	19.72	Media Ruptura
F_CMSC160	E	E	E	R	E	19.72	Media Ruptura
F_CMSC163	E	E	E	R	E	19.72	Media Ruptura
F_CMSC215	E	R	E	R	E	19.72	Media Ruptura
F_CMSC216	E	R	E	R	E	19.72	Media Ruptura
F_CMSC217	E	R	E	R	E	19.72	Media Ruptura
F_CMSC219	E	R	E	R	E	19.72	Media Ruptura
F_CMSC220	E	R	E	R	E	19.72	Media Ruptura
F_CMSC221	E	R	E	R	E	19.72	Media Ruptura
F_CMSC222	E	R	E	R	E	19.72	Media Ruptura
F_CMSC223	E	R	E	R	E	19.72	Media Ruptura
F_CMSC224	E	R	E	R	E	19.72	Media Ruptura
F_CMSC225	E	R	E	R	E	20.01	Alta Ruptura
F_CMSC164	E	E	R	R	E	20.03	Alta Ruptura
F_CMSC165	E	E	E	R	E	20.28	Alta Ruptura
F_CMSC166	E	R	E	R	E	20.4	Alta Ruptura
F_CMSC226	E	R	E	R	E	20.47	Alta Ruptura
F_CMSC167	E	E	E	R	E	20.51	Alta Ruptura
F_CMSC168	E	E	E	R	E	20.63	Alta Ruptura
F_CMSC227	-	R	E	E	E	20.64	Alta Ruptura
F_CMSC48	R	R	E	R	E	20.91	Alta Ruptura
F_CMSC169	E	E	R	R	E	20.98	Alta Ruptura
F_CMSC50	R	R	E	R	E	21	Alta Ruptura
F_CMSC170	E	E	E	R	E	21.12	Alta Ruptura
F_CMSC21	R	R	R	R	R	21.2	Alta Ruptura
F_CMSC56	R	R	E	R	E	21.2	Alta Ruptura
F_CMSC63	R	R	E	R	E	21.2	Alta Ruptura
F_CMSC67	R	R	E	R	E	21.2	Alta Ruptura
F_CMSC171	E	E	E	R	E	21.2	Alta Ruptura
F_CMSC172	E	R	E	R	E	21.2	Alta Ruptura
F_CMSC231	E	R	E	R	E	21.2	Alta Ruptura
F_CMSC232	E	R	E	R	E	21.2	Alta Ruptura

Tabla Suplementaria I. (continuación) Estado alélico determinado por el SNP de mayor significancia en cada uno de los 5 QTL determinados en el mapeo por asociación y fenotipo para todo el panel LPM-H. E: estado alélico que confiere el mejor comportamiento (baja ruptura), R: estado alélico alternativo, -: estado alélico no disponible.

Código de Línea	Qrf1.1	Qrf2.2	Qrf2.1	Qrf3.1	Qrf8.1	AUBPC	Detalle
F_CMSC173	E	E	R	R	E	21.31	Alta Ruptura
F_CMSC235	E	R	E	R	E	21.65	Alta Ruptura
F_CMSC174	-	E	R	R	E	21.77	Alta Ruptura
F_CMSC175	E	R	E	R	E	21.88	Alta Ruptura
F_CMSC176	E	E	E	R	E	22.11	Alta Ruptura
F_CMSC44	R	E	E	R	E	22.47	Alta Ruptura
F_CMSC178	E	E	E	R	E	23.04	Alta Ruptura
F_CMSC236	E	R	E	R	E	23.14	Alta Ruptura
F_CMSC237	E	R	E	R	E	23.21	Alta Ruptura
F_CMSC238	E	R	E	R	E	23.28	Alta Ruptura
F_CMSC22	R	R	E	R	R	23.35	Alta Ruptura
F_CMSC239	E	R	E	R	E	23.35	Alta Ruptura
F_CMSC71	R	R	R	R	E	23.86	Alta Ruptura
F_CMSC241	E	R	E	R	E	24.01	Alta Ruptura
F_CMSC242	E	R	E	R	E	24.24	Alta Ruptura
F_CMSC11	E	E	R	R	R	24.31	Alta Ruptura
F_CMSC243	E	R	E	R	E	24.46	Alta Ruptura
F_CMSC179	E	E	E	R	E	24.57	Alta Ruptura
F_CMSC180	E	E	R	R	E	24.64	Alta Ruptura
F_CMSC244	E	R	E	R	E	24.64	Alta Ruptura
F_CMSC181	E	E	R	R	E	25.27	Alta Ruptura
F_CMSC12	E	E	E	R	R	25.32	Alta Ruptura
F_CMSC13	E	E	E	R	R	25.32	Alta Ruptura
F_CMSC23	R	R	R	R	R	25.32	Alta Ruptura
F_CMSC245	E	R	E	R	E	25.32	Alta Ruptura
F_CMSC182	E	E	R	R	E	25.36	Alta Ruptura
F_CMSC246	E	R	E	E	E	25.38	Alta Ruptura
F_CMSC14	E	E	E	R	R	25.4	Alta Ruptura
F_CMSC24	R	R	E	R	R	25.51	Alta Ruptura
F_CMSC26	E	R	E	R	R	25.51	Alta Ruptura
F_CMSC15	E	E	E	R	R	25.62	Alta Ruptura
F_CMSC183	E	E	R	R	E	25.64	Alta Ruptura

Tabla Suplementaria I. (continuación) Estado alélico determinado por el SNP de mayor significancia en cada uno de los 5 QTL determinados en el mapeo por asociación y fenotipo para todo el panel LPM-H. E: estado alélico que confiere el mejor comportamiento (baja ruptura), R: estado alélico alternativo, -: estado alélico no disponible.

Código de Línea	Qrf1.1	Qrf2.2	Qrf2.1	Qrf3.1	Qrf8.1	AUBPC	Detalle
F_CMSC247	E	R	E	R	E	25.7	Alta Ruptura
F_CMSC248	E	R	E	R	E	25.8	Alta Ruptura
F_CMSC25	R	R	E	R	R	26.1	Alta Ruptura
F_CMSC184	-	E	E	R	E	26.1	Alta Ruptura
F_CMSC72	R	R	R	R	E	26.12	Alta Ruptura
F_CMSC185	E	-	R	R	E	26.25	Alta Ruptura
F_CMSC3	E	E	E	R	R	26.45	Alta Ruptura
F_CMSC186	E	E	R	R	E	26.46	Alta Ruptura
F_CMSC250	E	R	E	R	E	26.46	Alta Ruptura
F_CMSC251	E	R	E	R	E	26.47	Alta Ruptura
F_CMSC1	-	E	R	R	R	26.93	Alta Ruptura
F_CMSC16	R	E	R	R	R	27.09	Alta Ruptura
F_CMSC4	E	R	E	R	R	27.17	Alta Ruptura
F_CMSC17	R	R	E	R	R	27.29	Alta Ruptura
F_CMSC18	E	R	E	R	R	27.29	Alta Ruptura
F_CMSC2	E	E	R	R	R	27.42	Alta Ruptura
F_CMSC6	E	E	R	R	R	27.54	Alta Ruptura
F_CMSC5	E	E	R	R	R	27.65	Alta Ruptura
F_CMSC27	R	-	R	R	R	27.66	Alta Ruptura
F_CMSC188	E	R	E	R	E	27.77	Alta Ruptura
F_CMSC8	E	E	R	R	R	27.9	Alta Ruptura
F_CMSC187	E	R	R	R	E	27.9	Alta Ruptura
F_CMSC189	E	R	E	R	E	28.01	Alta Ruptura
F_CMSC7	E	E	R	R	R	28.02	Alta Ruptura
F_CMSC9	E	E	E	R	R	28.38	Alta Ruptura
F_CMSC19	E	R	R	R	R	28.73	Alta Ruptura
F_CMSC29	E	E	R	R	E	28.73	Alta Ruptura
F_CMSC190	E	R	E	R	E	29.52	Alta Ruptura