

Mapeo asociativo para virus en líneas endocriadas de maíz

Fabietti, Verónica A. y Gribaudo, Eliana M.

Tutor: Cecilia Bruno.

Co-tutor: María Eugenia Videla.



ÍNDICE DE CONTENIDOS

I.	INTRODUCCIÓN	6
II.	OBJETIVO.....	8
III.	MATERIALES Y MÉTODOS	9
	A. DESCRIPCIÓN DE LA BASE DE DATOS EXPERIMENTAL.....	9
	> Mapeo asociativo.....	9
	> Análisis de conglomerados.....	10
	> Análisis de la varianza	11
IV.	RESULTADOS	13
	A. DESCRIPTIVA DE MARCADORES MOLECULARES	13
	B. MAPEO ASOCIATIVO	14
	C. ANÁLISIS DE CONGLOMERADOS	17
	D. ANÁLISIS DE LA VARIANZA.....	18
V.	CONCLUSIÓN	21
VI.	REFERENCIAS BIBLIOGRÁFICAS.....	22

ÍNDICE DE TABLAS Y FIGURAS

TABLA 1: Medidas descriptivas de base de datos genéticos de 62 líneas estabilizadas de maíz genotipadas con 569 marcadores moleculares tipo SNP.	13
TABLA 2: Clasificación de la carga viral según los límites inferior y superior y su frecuencia absoluta y relativa.	14
TABLA 3: Resultados obtenidos luego de aplicar el modelo lineal general de análisis de la varianza para los marcadores significativos del modelo QK.	19
TABLA 4: AIC obtenido de cada modelo de análisis de varianza realizado para cada modelo de mapeo asociativo en función de los marcadores significativos.....	20
GRÁFICO 1: Comparación de gráficos de valor-p de 6 modelos de mapeo asociativo para 520 marcadores de tipo snp de 42 líneas de maíz estabilizadas.	16
GRÁFICO 2: Distribución empírica de los marcadores significativos de cada modelo de mapeo asociativo empleado.....	17
GRÁFICO 3: Dendograma obtenido del análisis de conglomerados correspondiente a los datos obtenidos del modelo QK	18

AGRADECIMIENTOS

Queremos agradecer a nuestra familia, amigos por acompañarnos en este proceso y a la Ing. Agr. Cecilia Bruno y a María Eugenia Videla por ponerse a nuestra disposición y comprender las dificultades que tuvimos ajenas a la realización del trabajo.

TÍTULO

Mapeo asociativo para virus en líneas endocriadas de maíz.

RESUMEN

Los virus son, después de los hongos, los patógenos de plantas de mayor importancia económica por el daño directo que las enfermedades causan en los cultivos. El conocimiento de la biología y ecología del virus y de los aspectos básicos de su epidemiología son esenciales para la toma de decisiones sobre las estrategias a utilizar en cada caso. Una de las formas de hacer mejoramiento es a través de los marcadores moleculares y la técnica de mapeo asociativo es un enfoque desarrollado para detectar genes de interés en plantas.

PALABRAS CLAVES

Enfermedades – Modelos de asociación - Resistencia o tolerancia

I. INTRODUCCIÓN

El maíz es susceptible a enfermedades, que en alguna forma afectan el normal desarrollo de las plantas. Las enfermedades son favorecidas por las condiciones ambientales, el tipo de suelo, la susceptibilidad de los materiales y, en el caso de las enfermedades de origen viral, por las condiciones que favorezcan la migración, establecimiento y supervivencia de los insectos vectores. De acuerdo con los diferentes estudios, las enfermedades virales, se presentan en las primeras etapas de desarrollo del cultivo, y su importancia está relacionada con el grado de severidad, la incidencia y la época de infección (Varón de Agudelo y Sarria Villa, 2007).

Las enfermedades virales no son un problema nuevo en el cultivo del maíz y son un factor limitante para su producción. El control de los virus es complejo, ya que sus poblaciones son variables y, a diferencia de otros patógenos de plantas, no se dispone de productos fitosanitarios que puedan utilizarse para su control directo. De ahí que su manejo pasa por la integración de medidas dirigidas a reducir las fuentes de infecciones primarias, y una vez presente la infección en el cultivo, reducir la dispersión secundaria y/o el efecto en las plantas (Samá y Orta, 2017).

Entre los virus y mollicutes que afectan al maíz, tres producen enfermedades de importancia económica en Argentina. Ellos son Mal de Río Cuarto virus (MRCV), Spiroplasma kunkelii Whitcomb (Spiroplasma del Achaparramiento o Corn stunt spiroplasma) y Maize dwarf mosaic virus (MDMV: Virus del mosaico enanizante del maíz). En cuanto a importancia, se destaca el Mal de Río Cuarto, que provoca importantes pérdidas en nuestro país. Prevalece la necesidad de un tratamiento integrado de este problema fitosanitario, que tenga en cuenta tanto al cultivo, los reservorios naturales del virus y su insecto vector, como al medio ambiente y al hombre. Dentro de este contexto, es importante señalar que algunas acciones tomadas por el hombre en la cadena del maíz han colaborado a desencadenar importantes epidemias. Son muy importantes las pautas de manejo, entre las que se destacan selección de híbridos adecuados, fechas de siembras apropiadas y -eventualmente- aplicación de tratamientos de insecticidas sistémicos a la semilla. La zona denominada "endémica" del Mal de Río Cuarto se extiende por el sur de la provincia de Córdoba, sudoeste de Santa Fe, oeste de Buenos Aires, norte de La Pampa y este de San Luis, ya que en esta región -normalmente todos los años- hay presencia de la virosis con mayor o menor intensidad en el cultivo de maíz. En el "cinturón maicero tradicional", se presentan casos esporádicos de la enfermedad y en años de epidemia se han cuantificado pérdidas importantes (Lenardón, 2007). La gravedad de la enfermedad se incrementa en cultivos sembrados tardíamente

y sometidos a estrés. Esta enfermedad se presenta al final del ciclo del cultivo y para su correcta identificación deben usarse técnicas de laboratorio.

Una de las formas de hacer mejoramiento es a través de los marcadores moleculares. Un marcador genético o marcador molecular es un segmento de ADN con una ubicación física identificable (locus) en un cromosoma y cuya herencia genética se puede rastrear. Los marcadores se usan para el mapeo genético como el primer paso para encontrar la posición e identidad de un gen. Permiten evidenciar variaciones (polimorfismos) en la secuencia del ADN entre dos individuos, modifiquen éstas o no su fenotipo. Funcionan como señaladores de diferentes regiones del genoma (Franco, 2016). Los marcadores moleculares tienen la ventaja de no variar con el ambiente y pueden ser observados en etapas tempranas del ciclo de vida de un individuo.

La técnica de mapeo asociativo es un enfoque desarrollado para detectar genes de interés en plantas (Jannink et al., 2001; Kraakman et al., 2004; von Zitzewitz et al., 2011). El mapeo asociativo (MA) o mapeo por desequilibrio de ligamiento (LD) es una herramienta analítica para asociar fenotipos (características observadas de los organismos) a genotipos (constitución genética del organismo). Se realiza una examinación de numerosas variantes genéticas comunes en una población de mapeo para identificar si alguna variante está asociada con una característica fenotípica del individuo que interese mapear sobre el genoma.

II. OBJETIVO

Identificar la relación existente entre las líneas de maíz (genotipo) y la expresión de los virus en cada una de ellas (fenotipo).

III. MATERIALES Y MÉTODOS

A. DESCRIPCIÓN DE LA BASE DE DATOS EXPERIMENTAL

La base de datos genética cuenta con 61 individuos, líneas estables de maíz, genotipadas por 569 marcadores moleculares del tipo SNP (polimorfismo de un solo nucleótido). Los marcadores fueron codificados con 1 que indica polimorfismo (es decir que hubo una mutación) y con 0 que indican que coincide con el genoma de referencia. La variable fenotípica es el síntoma que manifiesta la línea de maíz ante el ataque de virus. Se midió en porcentaje de plantas enfermas de una parcela, realizándose dos repeticiones, y se estimó la relación entre las diferentes líneas de maíz y el contenido de virus de las mismas.

En primer lugar, se realizó una caracterización de los marcadores a través de medidas descriptivas, luego se aplicaron seis modelos de asociación, se continuó con un análisis multivariado (análisis de conglomerados) y un análisis de varianza (modelo general, modelos lineales generales y mixtos y modelo lineal generalizado mixto).

> Mapeo asociativo

El mapeo asociativo permite encontrar lugares específicos del genoma relacionados con la variación de un carácter fenotípico. Es una práctica difundida en el mejoramiento de especies vegetales ya que no necesita la utilización de poblaciones provenientes de cruzamientos controlados. Sin embargo, se ha detectado que, en poblaciones estructuradas genéticamente, el número de falsos positivos en la asociación marcador-carácter puede aumentar significativamente (Peña Malavera, Gutiérrez y Balzarini, 2014).

La población de mapeo en estudio está formada por 42 individuos y 520 marcadores, el carácter fenotípico está representado por el porcentaje de virus. Se utilizaron diferentes estrategias de modelado teniendo en cuenta la estructura genética en el mapeo de asociación, los cuales son:

- a. **Modelo SCE** (Modelo sin corrección por estructura): un modelo lineal general de regresión con efectos de marcadores en el vector de coeficientes de regresión y sin explicitar de ninguna manera (ni en la parte fija ni en la parte aleatoria del modelo) la existencia de estructura en los datos.
- b. **Modelo P y Modelo Q:**
 - Modelo P: Modelo con corrección por estructura estimada a través de un Análisis Multivariado de Componentes Principales.
 - Modelo Q: Modelo con corrección por estructura estimada a través de la matriz de relaciones genéticas Q.

El modelo de mapeo asociativo presenta por un lado el componente asociado al efecto de los marcadores moleculares y por otro un componente asociado a la estructura genética poblacional (EGP), ambos tratados como de efectos fijos. En este modelo son usadas alguna de dos importantes matrices 1. La matriz Q proveniente del análisis de estructura realizado con el programa STRUCTURE, que presenta las probabilidades de pertenencia a las poblaciones en estudio o 2. La matriz P de componentes principales significativas según el estadístico de Tracy-Widom (1994).

- c. **Modelo K** (Kinship): Modelo con corrección por estructura usando la matriz de parentesco. Se modela la variación en datos fenotípicos teniendo en cuenta los marcadores como efectos fijos y la matriz de coancestría, matriz “kinship” (K), para indicar relaciones de parentesco entre los efectos aleatorios.
- d. **Modelo PK**: Modelo con corrección por estructura estimada a través de un Análisis Multivariado de Componentes Principales (Matriz P) y usando la matriz de parentesco (K).
- e. **Modelo QK**: Modelo con corrección por estructura usando la matriz de parentesco K y la matriz de relaciones genéticas Q(Structure). (Balzarini et al.2008).

> **Análisis de conglomerados**

El objetivo del análisis de conglomerados, o declasificación no supervisado, es formar grupos tal que los elementos de un grupo sean más parecidos entre sí que con los elementos de otro grupo. Estos métodos se aplican cuando se supone que existe una estructura y se desea identificarla. El agrupamiento de objetos multivariados es frecuentemente utilizado como método exploratorio de datos con la finalidad de obtener mayor conocimiento sobre la estructura de las observaciones y/o variables en estudio. Si bien es cierto que el proceso de agrupamiento conlleva inicialmente a una pérdida de información ya que se sitúan en una misma clase unidades que no son idénticas (solo semejantes), la síntesis de la información disponible sobre las unidades consideradas puede facilitar considerablemente la visualización de relaciones multivariadas de naturaleza compleja. Las técnicas de clasificación basadas en agrupamientos implican la distribución de las unidades de estudio en clases o categorías de manera tal que cada clase (conglomerado) reúne unidades cuya similitud es máxima bajo algún criterio. Es decir, los objetos en un mismo grupo comparten el mayor número permisible de características y los objetos en diferentes grupos tienden a ser distintos.

> **Análisis de la varianza**

El Análisis de la Varianza (ANAVA) permite probar hipótesis referidas a los parámetros de posición (esperanza) de dos o más distribuciones. Para hacer un análisis de la varianza entre los grupos se probarán diferentes modelos. Primero comenzaremos con un análisis simple, viendo si se cumplen o no los supuestos clásicos, luego seguiremos con el modelo mixto, y en caso que el supuesto de normalidad no se cumpla se puede probar con un modelo lineal generalizado (MLG). En nuestro caso la variable respuesta (fenotipo) es la concentración de virus; es una variable continua, la distribución que podría tomar en el marco de MLG, sería una gamma o una normal inversa. Para todos los modelos empleados en el análisis de la varianza los supuestos que deben cumplirse son:

1. **Normalidad** de los términos de error aleatorio
2. **Homogeneidad de varianza** dentro de cada tratamiento
3. **Independencia** entre términos de error aleatorio

En el caso de que algunos de esos supuestos no se cumplan, impactaran sobre el verdadero valor nivel de significancia de la prueba de hipótesis del ANAVA, afectando la calidad de las conclusiones que podamos obtener.

1. **Normalidad:** la técnica más usada es construir un Q-Q plot normal. Se obtiene un diagrama de dispersión en el que, si los residuales son normales y no hay otros defectos en el modelo, los residuos observados se alinean sobre una recta a 45°, ya que correlacionan bien con los residuos esperados bajo el supuesto que la muestra de datos realmente sigue una distribución normal.

2. **Homogeneidad de varianzas:** cuando los términos de error tienen varianzas homogéneas y el modelo explica bien a los datos, el gráfico de dispersión de residuos vs. predichos presentara una nube de puntos sin patrón alguno.

3. **Independencia:** para estudiar la posible falta de independencia entre los errores se puede realizar un gráfico de los residuos según la secuencia en el tiempo o espacio físico en que han sido recolectados los datos, para esto debe conocerse como ha sido el mecanismo de recolección de los datos. Si los residuos aparecen en secuencia de varios valores positivos seguidos de varios valores negativos puede ser un indicio claro de falta de independencia.

Modelos lineales generales y mixtos: son una propuesta de modelación estadística avanzada, que permiten mejorar la calidad del análisis de los factores fijos y factores aleatorios, modelando la variabilidad aleatoria y la correlación de los errores, siendo muy

útiles en el análisis de datos desbalanceado, datos con pseudoréplicas, o datos con algún tipo de estructura jerárquica o de agrupación (Riobóo y Pacheco,2018).

Modelos lineales generalizados mixtos: permiten que la media de una población dependa de un predictor lineal a través de una función de enlace (link function) de tipo no lineal y que la distribución probabilística de la variable respuesta sea cualquiera de la familia exponencial. De este modo, las variables explicativas pueden ser cualquier combinación de variables continuas, variables de clasificación y sus interacciones. La estimación de parámetros en estos modelos se realiza mediante procedimientos de máxima verosimilitud. Hacen posible extender los modelos lineales clásicos de efectos fijos, al incluir efectos aleatorios y predictores BLUPs, para el análisis de datos con distribuciones de la familia exponencial. Además, permiten procesar datos correlacionados al usar efectos aleatorios y estimar sus componentes de varianza asociados al modelo, adicionales a los del error residual. Los procedimientos de estimación usados posibilitan reducir los sesgos cuando los datos están incompletos, desbalanceados o ajustar datos dispersos y permiten además modelar la estructura de los errores en datos provenientes de mediciones longitudinales (Bandera-Fernández y Pérez-Pelea,2018).

IV. RESULTADOS

A. DESCRIPTIVA DE MARCADORES MOLECULARES

En la Tabla 1 se puede observar que la base de datos cuenta con 569 bandas, de las cuales 47 son duplicadas, 2 monomórficas y el 99.65% se trata de polimórficas.

Un marcador molecular monomórfico es invariable en todos los organismos estudiados, en cambio, cuando presenta diferencias en el peso molecular, actividad enzimática, estructura, o sitios de restricción, se dice que es polimórfico (Díaz, 1998).

Tabla 1: Medidas descriptivas de base de datos genéticos de 62 líneas estabilizadas de maíz genotipadas con 569 marcadores moleculares tipo SNP.

Resumen	Cantidad de casos
Muestras	61,00
Muestras duplicadas	19,00
Bandas (número)	569,00
Patrón bandas duplicadas	47,00
Bandas monomórficas	2,00
Bandas polimórficas (%)	99,65
Primers	2,00

En base a lo observado en la Tabla 1, se eliminaron las 19 muestras duplicadas, 47 bandas duplicadas y 2 bandas monomórficas ya que no interfieren en el resultado final, obteniendo así una base de datos final depurada de 42 líneas y 520 marcadores.

Luego se realizó una tabla de frecuencia en la cual se clasificó a la variable virus según la cantidad de virus en función de los límites inferior y superior y sus frecuencias absoluta y relativa. Esta información se puede observar en la Tabla 2.

Tabla 2: Clasificación de la carga viral según los límites inferior y superior y su frecuencia absoluta y relativa.

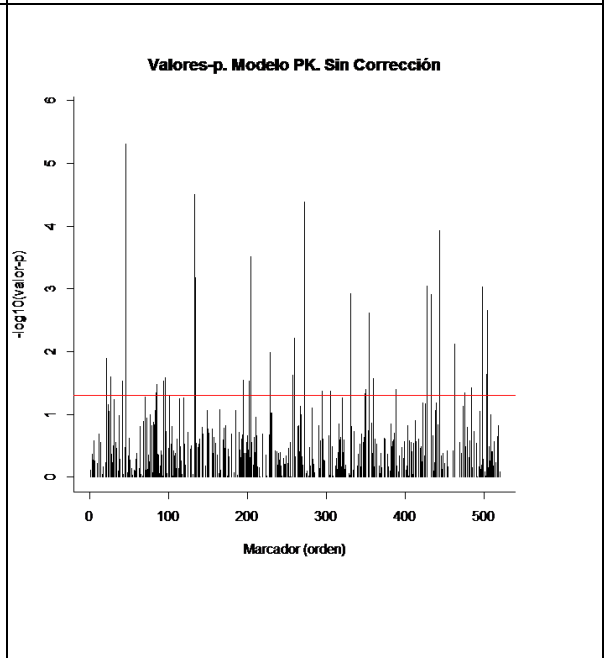
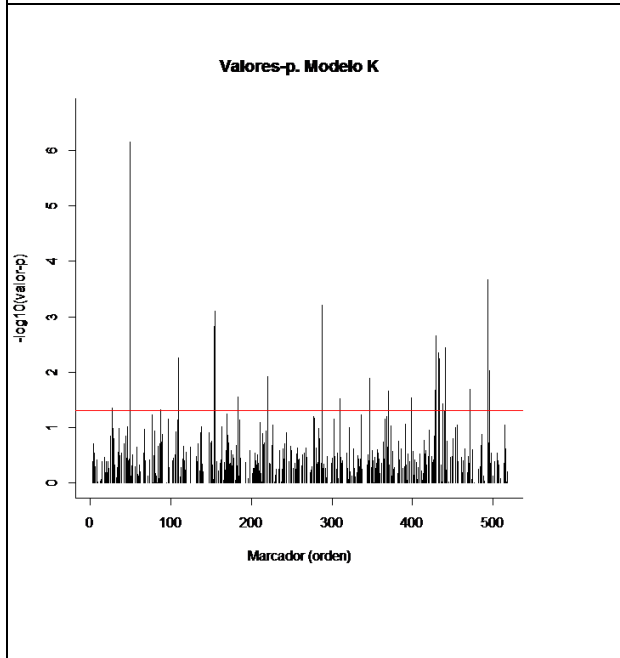
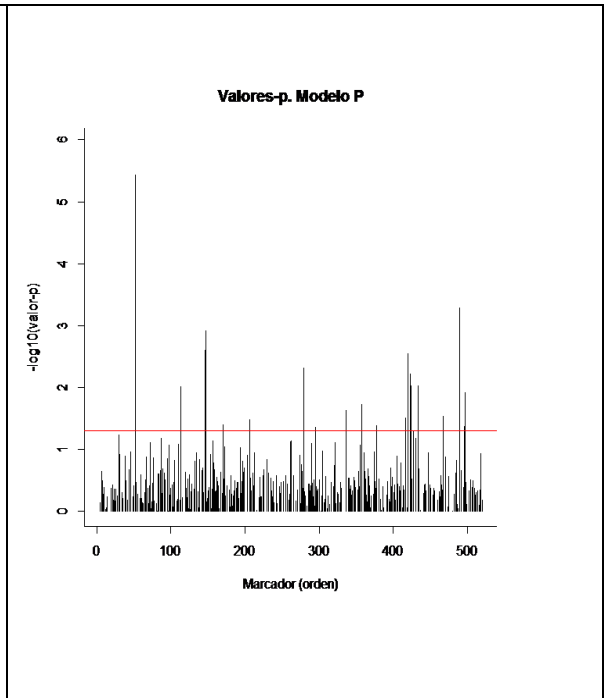
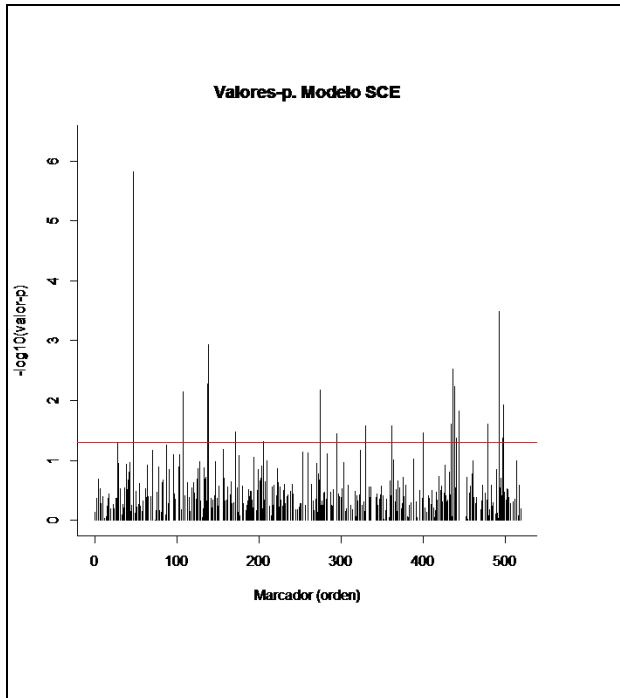
Variable	Clase	LI	LS	MC	FA	FR
Virus	1	[8,08	22,58]	15,33	50	0,82
Virus	2	(22,58	37,07]	29,83	1	0,02
Virus	3	(37,07	51,57]	44,32	1	0,02
Virus	4	(51,57	66,07]	58,82	5	0,08
Virus	5	(66,07	80,56]	73,31	4	0,07

En la Tabla 2, se clasificó la carga viral en 5 clases, destacándose que el 82% de las líneas tiene un rango de virus relativamente bajo y solo el 18% tiene una carga viral por encima de 22,58%.

B. MAPEO ASOCIATIVO

En la aplicación de mapeo asociativo se utilizó para todos los modelos un nivel de significancia de 0,05 y sin corrección por multiplicidad. Con el modelo SCE se obtuvieron 22 marcadores significativos, el modelo P arrojó 21 marcadores significativos, el modelo K alcanzó 23 marcadores significativos, el modelo PK dio 37 marcadores significativos, el modelo Q logró 20 marcadores significativos y el modelo QK obtuvo 34 marcadores significativos. Las diferencias obtenidas en los modelos se deben a que unos modelos son más precisos que otros, dando menor cantidad de falsos positivos. Los falsos positivos son marcadores que aparecen como significativos pero que realmente no lo son.

En el Gráfico 1 se pueden observar los gráficos obtenidos para cada modelo evaluado, indicando los marcadores significativos para cada uno de ellos por encima de la línea roja.



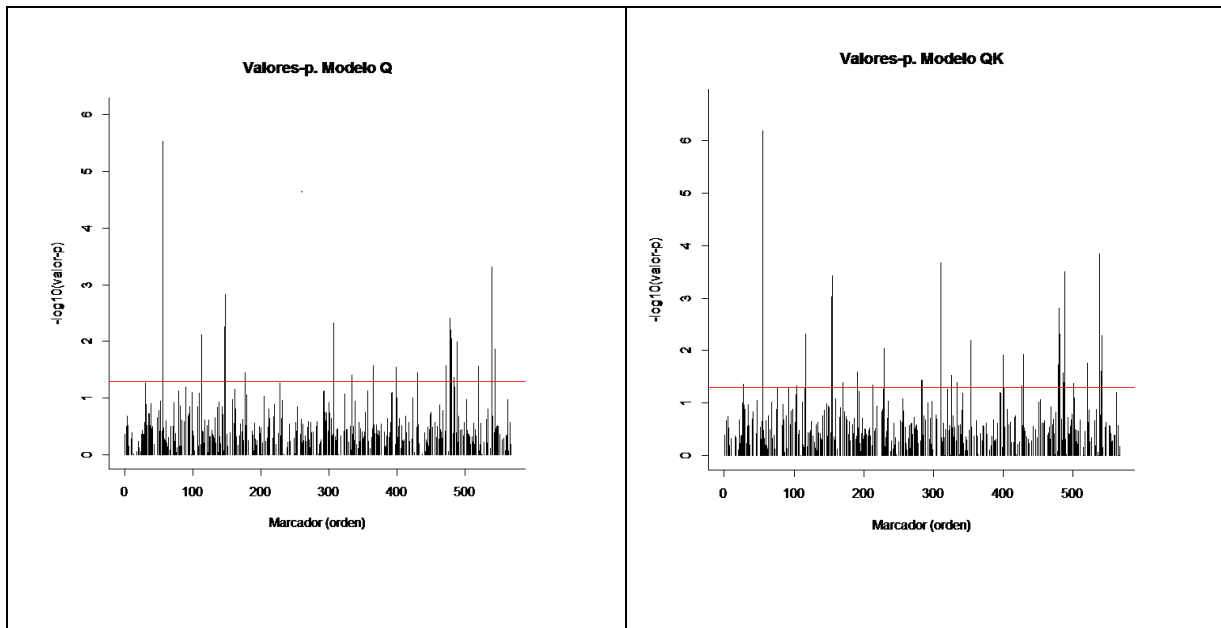


Gráfico 1: Diagramas de valor-p de 6 modelos de mapeo asociativo para 520 marcadores de tipo SNP sobre 42 líneas de maíz estabilizadas.

La línea roja observada en el Gráfico 1 indica el umbral para determinar la significancia del marcador, valores por encima del umbral indica que el marcador está asociado al carácter bajo estudio. Las opciones para calcular el umbral de significancia para disminuir la tasa de falsos positivos son: Li-Ji que estima el número efectivo de pruebas independientes por método propuesto por Li y Ji (Li y Ji, 2005) y FDR la tasa de falsos positivos.

Al observar las comparaciones del Gráfico 1 pareciera que todos los modelos son similares en cuanto a los resultados arrojados por lo que creímos conveniente realizar otro análisis adicional. Se llevó a cabo un gráfico de la distribución empírica utilizando como variable el valor-p de cada modelo de mapeo asociativo, el mejor modelo va a ser aquel que se ajuste a la línea de 45°. En el Gráfico 2 se puede ver que el supuesto de normalidad no se cumple debido a que existen desviaciones respecto a la línea de 45°. Pero se puede establecer como el mejor modelo de mapeo asociativo el modelo QK, ya que pareciera ser el que tiene menor desviación respecto a la línea de 45° y es más estable en su trayectoria.

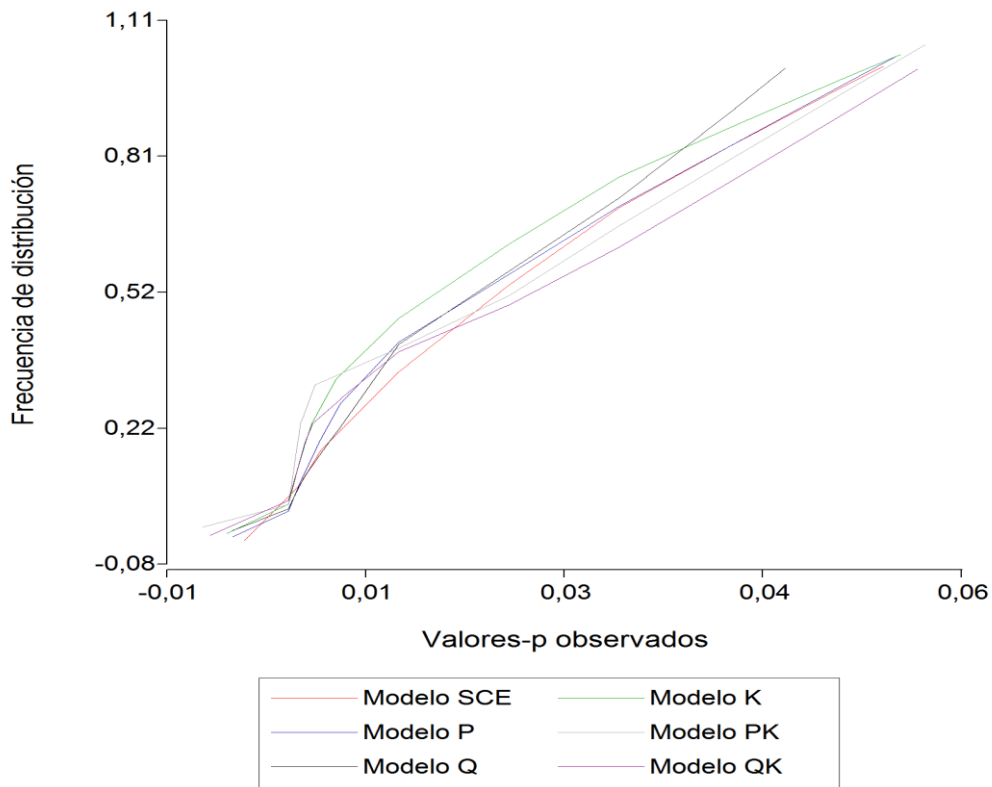


Gráfico 2: Distribución empírica de los marcadores significativos de cada modelo de mapeo asociativo empleado.

C. ANÁLISIS DE CONGLOMERADOS

Luego de aplicar el mapeo asociativo, que nos permitió reconocer los marcadores significativos para cada modelo, se realizó una depuración de los datos. Una vez obtenida la tabla depurada con los marcadores seleccionados, se procedió a realizar un análisis de conglomerados para cada modelo, conglomerando por filas y seleccionando dos grupos de conglomerados. A continuación, se muestra el dendograma obtenido con el modelo seleccionado como el mejor según el gráfico de la distribución empírica, que corresponde al modelo QK.

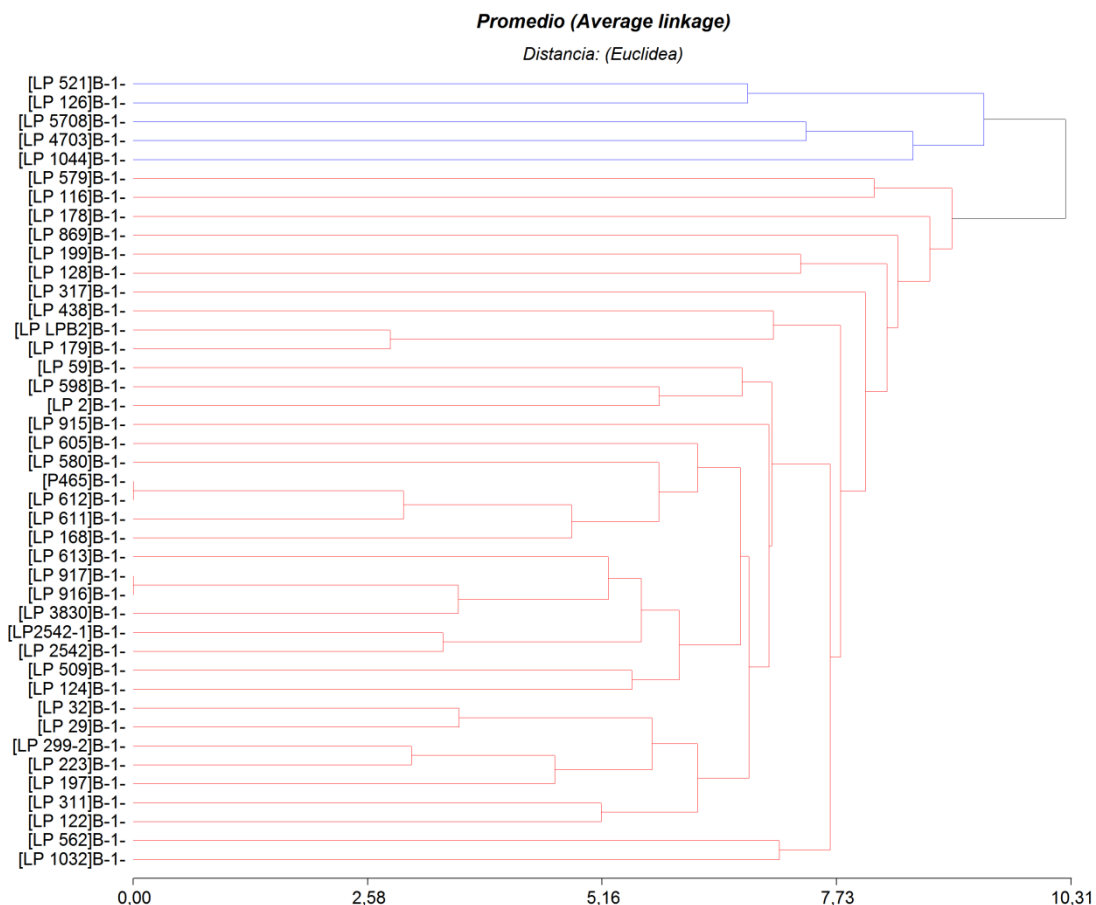


Grafico 3: Dendograma obtenido del análisis de conglomerados correspondiente a los datos obtenidos del modelo QK

El dendograma graficado corresponde a la evolución del agrupamiento en función de la distancia seleccionada identificando los grupos formados con diferentes colores. Observando el Grafico 3 se puede interpretar que hay cinco líneas, dentro del mismo grupo, que se unen a menor distancia que el resto, pudiendo predecir que esas líneas son más susceptibles al ataque de virus.

D. ANÁLISIS DE LA VARIANZA

Otra evaluación que realizamos a la base de datos de 42 líneas estabilizadas de maíz genotipadas 520 marcadores moleculares de tipo SNP es un análisis de la varianza. Utilizamos diferentes modelos de análisis para verificar si se cumplen o no los supuestos (normalidad, homogeneidad de varianza e independencia). Este análisis se realizó por separado según cada modelo de mapeo asociativo con los marcadores significativos arrojados por cada uno de ellos. Primero probamos con el modelo general de análisis de varianza, pero es muy clásico y al ver que no se cumplían los supuestos, implementamos el modelo lineal general y mixto poniendo

como efecto aleatorio a las líneas porque pensamos que podían ser muy distintas. Después, como los residuos seguían dando feos probamos con el modelo lineal generalizado mixto con variable respuesta con distribución gamma y enlace canónico (log) y también usamos como variable respuesta la normal inversa con enlace $1/\mu^2$.

En la Tabla 3 se puede ver el análisis de la varianza clásico realizado para el modelo de mapeo asociativo seleccionado como el mejor que fue el modelo QK. Esas medias muestrales indican que las líneas correspondientes al conglomerado 1 tienen la menor proporción de virus (15,39%) y las líneas del conglomerado 2 tienen mayor cantidad de virus (53,46%). Las letras diferentes indican diferencias significativas entre las medias que se comparan. En el grupo 2 quedaron solo 5 líneas que parecen ser muy susceptibles.

Tabla 3: Resultados obtenidos luego de aplicar el modelo lineal general de análisis de la varianza para los marcadores significativos del modelo QK.

Test: LSD Fisher Alfa=0,05 DMS=8,15764

Error:71,7607 gl:40

Conglomerado	Medias	N	E.E	
1	15,39	37	1,39	A
2	53,46	5	3,79	B

Medias con una letra común no son significativamente diferentes ($p > 0,05$)

Realizamos una comparación de los AIC obtenidos de cada modelo de análisis de la varianza para poder establecer cuál es el mejor modelo a pesar de que no se encuentran diferencias significativas entre ellos. El modelo a seleccionar va a ser aquel que tenga menor valor de AIC ya que se considera mejor.

Tabla 4:AIC obtenido de cada modelo de análisis de varianza realizado para cada modelo de mapeo asociativo en función de los marcadores significativos.

	Modelo SCE	Modelo K	Modelo P	Modelo PK	Modelo Q	Modelo QK
Modelo lineales generales y mixtos	310,08	310,08	306,63	328,45	330,10	297,67
ML Generalizado Mixto (gamma)	280,56	280,56	290,81	305,49	307,68	274,89
ML Generalizado Mixto (normal inversa)	275,22	275,22	286,26	295,62	298,68	272,40

AIC menor implica mejor

De lo observado en la Tabla 4 se puede decir que el mejor modelo de asociación es el modelo QK porque tiene el menor valor de AIC y como mejor modelo de análisis de varianza se seleccionó el modelo lineal generalizado mixto con variable respuesta con distribución normal inversa y enlace $1/\mu^2$ ya que posee el valor de AIC más bajo.

Para el análisis de la varianza, el modelo lineal generalizado mixto es más adecuado que los demás modelos empleados porque cumple mejor con los supuestos (normalidad, homogeneidad de varianza e independencia) e incluye efectos aleatorios y predictores de BLUPs para el análisis de los datos, por lo que para establecer cuáles son las líneas susceptibles al ataque de virus, se tuvieron en cuenta los resultados obtenidos de dicho modelo.

V. CONCLUSIÓN

Estos modelos son importantes en el marco del mejoramiento genético vegetal para poder dar una respuesta a la interacción genotipo-fenotipo, dado que la expresión fenotípica depende del ambiente. Poder evaluar a través del genotipo esas asociaciones sería una contribución importante al mejoramiento.

En el presente trabajo se evaluó una base de datos genética de 42 líneas estables de maíz para establecer la susceptibilidad o resistencia de las líneas al ataque de virus. Las siguientes líneas [LP 521]B-1-, [LP 126]B-1-, [LP 5708]B-1-, [LP 4703]B-1- y [LP 1044]B-1- fueron estadísticamente distintas a las demás y son las más susceptibles a virus, mostrando una carga viral de 53,46%.

VI. REFERENCIAS BIBLIOGRÁFICAS

Badiella, L. (2011). Modelos lineales generalizados mixtos: algunos casos prácticos. In Barcelona: X Congreso Galego de Estatística e Investigación de Operacións.

Balzarini M.G., Di Rienzo J.A. InfoGen versión 2018. FCA, Universidad Nacional de Córdoba, Argentina. URL <http://www.info-gen.com.ar>

Balzarini M.G., González L., Tablada M., Casanoves F., Di Rienzo J.A., Robledo C.W. (2008). Infostat. Manual del Usuario, Editorial Brujas, Córdoba, Argentina.

Bandera-Fernández, E., & Pérez-Pelea, L. (2018). Los modelos lineales generalizados mixtos. Su aplicación en el mejoramiento de plantas. *Cultivos Tropicales*, 39(1), 127-133.

Di Rienzo J.A., Casanoves F., Balzarini M.G., González L., Tablada M., Robledo C.W. InfoStat versión 2017. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina. URL <http://www.infostat.com.ar>

Díaz, M. G. C. (1998). Marcadores moleculares: qué son, cómo se obtienen y para qué valen. *Encuentros en la biología*, (49), 3.

Franco, J. L (2016). Instituto Nacional de Investigaciones del Genoma Humano. Definición de marcador genético.

Lenardón, S. (2007). Enfermedades causadas por virus y mollicutes en maíz. INTA – Universidad Nacional de Rio Cuarto.

Peña Malavera, A., Gutiérrez, L., & Balzarini, M. (2014). Componentes principales en mapeo asociativo. *BAG. Journal of basic and applied genetics*, 25(2), 32-40.

Riobóo, L. D., & Pacheco, M. P. (2018). Modelos lineales generales y mixtos en la caracterización de la variable calificación, *Ingeniería Agroindustrial, Uni-norte. Nexa Revista Científica*, 30(2), 84-95.

Samá, M. A. A., & Orta, G. C. (2017). Nuevos retos en el control de las enfermedades virales del maíz. *Vida rural*, (424), 44-50.

Varón de Agudelo, F., & Sarria Villa, G. A. (2007). Enfermedades del maíz y su manejo: compendio ilustrado. Instituto Colombiano Agropecuario-ICA.

