



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

**INSTITUTO DE BIOLOGÍA
BIOLOGÍA EXPERIMENTAL**

**Estudios de asociación genómica (GWAS) relacionados a la elongación del mesocotilo
en maíz y su validación mediante análisis de ligamiento de genes candidatos**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN CIENCIAS BIOLÓGICAS

PRESENTA:

BIÓL. LEOPOLDO VÁZQUEZ MARCIAL

TUTOR PRINCIPAL DE TESIS: DR. JORGE NIETO SOTELO
JARDÍN BOTÁNICO, INSTITUTO DE BIOLOGÍA, UNAM
COMITÉ TUTOR: DR. LUIS ENRIQUE EGUIARTE FRUNS
INSTITUTO DE ECOLOGÍA, UNAM
DRA. JENY SOLANGE SOTUYO VÁZQUEZ
INSTITUTO DE BIOLOGÍA, UNAM

CIUDAD DE MÉXICO, AGOSTO DE 2018



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Lic. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión del Subcomité por Campo de Conocimiento de Biología Experimental y Biomedicina del Posgrado en Ciencias Biológicas, celebrada el día 12 de marzo de 2018, se aprobó el siguiente jurado para el examen de grado de **MAESTRO EN CIENCIAS** del alumno **VÁZQUEZ MARCIAL LEOPOLDO** con número de cuenta **305300050** con la tesis titulada **"Estudios de asociación genómica (GWAS) relacionados a la elongación del mesocotilo en maíz y su validación mediante análisis de ligamento de genes candidatos"**, realizada bajo la dirección del **DR. JORGE NIETO SOTELO**:

Presidente: DR. LUIS DAVID ALCARÁZ PERAZA
Vocal: DR. ULISES YUNUÉN ROSAS LÓPEZ
Secretario: DRA. JENY SOLANGE SOTUYO VÁZQUEZ
Suplente: DRA. ALEJANDRA CITLALLI MORENO LETELIER
Suplente: DR. MIGUEL ANGEL VALLEJO REYNA

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPIRITU"
Cd. Universitaria, Cd. Mx., a 25 de junio de 2018.


DR. ADOLFO GERARDO NAVARRO SIGÜENZA
COORDINADOR DEL PROGRAMA



c.c.p. Expediente del (la) interesado (a).

Agradecimientos institucionales

Al Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, que a través de su apoyo institucional permite el desarrollo de nuevos científicos mexicanos.

Al consejo Nacional de Ciencia y Tecnología CONACYT por haber aportado el apoyo económico mediante la beca de maestría Reg. Becario 330476 otorgada del año 2014 al 2016.

A los proyectos que permitieron la realización de este trabajo: CONACYT (PN247732) y DGAPA-PAPIIT-UNAM (IG200515) otorgados al Dr. Jorge Nieto Sotelo en colaboración con la Dra. Gladys I. Cassab del Instituto de Biotecnología de la UNAM.

Al Dr. Jorge Nieto Sotelo del Laboratorio de Fisiología Molecular, Jardín Botánico del Instituto de Biología de la Universidad Nacional Autónoma de México (Tutor Principal), por darme la oportunidad de trabajar en su proyecto, por toda la paciencia, confianza y, sobre todo, por compartir conmigo sus conocimientos.

A los miembros del comité tutorial, el Dr. Luis Enrique Eguiarte Fruns y la Dra. Jeny Solange Sotuyo Vázquez, por su apoyo y guía en la realización de este trabajo.

Agradecimientos a título personal

Al Instituto de Biología de la Universidad Nacional Autónoma de México.

Al Jardín Botánico del Instituto de Biología, por permitirme realizar los experimentos en sus instalaciones.

Al Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), por el apoyo otorgado y por permitirme realizar parte de esta investigación en sus instalaciones.

A la Unidad de Servicio Informático del Instituto de Biología por el apoyo brindado.

A los miembros del jurado revisor, el Dr. Luis David Alcaraz Peraza, el Dr. Ulises Yunuen Rosas López, la Dra. Jeny Solange Sotuyo Vázquez, la Dra. Alejandra Citlalli Moreno Letelier y el Dr. Miguel Angel Vallejo Reyna, por haber revisado y aportado sus valiosas ideas para mejorar esta tesis.

Al Dr. Xuecai Zhang del (CIMMYT), por su confianza, dirección y consejo, que permitieron esta investigación.

A mis padres Gabriel Salvador Vázquez y Ofelia Coca, por su apoyo incondicional y ejemplo que me inspira a seguir esforzándome.

A mi hermano Jesús Salvador Vázquez, por su apoyo incondicional, sin su ayuda este trabajo no hubiera sido posible.

A mis compañeros y amigos del Laboratorio de Fisiología Molecular: Idalia Rojas, Luz María Rangel, Miguel Ángel Vallejo, Yajima Osorno, Alberto Cruz, Raymundo Alonso, Daniel Alejandro Martínez, Alejandra Figueroa, Eric Sosa y Jesús Edgardo Gutiérrez, por su ayuda y consejo que hicieron posible este trabajo.

Al personal y alumnos del CIMMYT: Martha Hernández, Yibing Yuan, Nayeli, Faustino, Marcos, Aldo y Eduardo Zilli, por su apoyo y enseñanzas.

ÍNDICE

RESUMEN	1
ABSTRACT	2
INTRODUCCIÓN	3
El cambio climático y el calentamiento global.	3
El cultivo del maíz.	4
La siembra profunda.	5
El mesocotilo.	6
Genética de la variación en la elongación del mesocotilo.	8
Mejoramiento genético.	9
Genotipificación mediante secuenciación (GBS).	12
Análisis de asociación.	13
Mapas de ligamiento.	15
Estudios de asociación genómica (GWAS).	16
ANTECEDENTES	18
HIPÓTESIS	19
OBJETIVOS	19
MÉTODOS	20
Ensayos para evaluar la respuesta de las plántulas a la siembra profunda.	20
Genotipificación por secuenciación (GBS).	22
Estudios de Asociación a Nivel de Todo el Genoma [GWAS] de la elongación del mesocotilo en la colección de líneas de maíz DTMA.	23
Cálculo de los efectos fenotípicos de los SNP asociados a la elongación del mesocotilo en las líneas de maíz DTMA.	24
Otros análisis bioinformáticos llevados a cabo.	24
<i>Análisis de QTL de la variación de la elongación del mesocotilo en las líneas ABQTL-DTP BC₁F_{2:3}</i>	25
RESULTADOS	26
Asociación genómica (GWAS) de la varianza en la elongación del mesocotilo en líneas DTMA.	26
Efectos fenotípicos.	33
Mapeo cromosómico de los SNPs asociados.	36

<i>Loci asociados a la variación de la elongación del mesocotilo mediante QTLs en la población ABQTL- DTP-BC₇F_{2:3}</i>	43
DISCUSIÓN	55
Fenotipos y resultados de la asociación genómica.	55
Genes asociados a metales pesados y su implicación evolutiva.	62
Crecimiento celular.	64
Genes asociados al transporte y metabolismo del nitrógeno.	65
Factores de transcripción.	66
Genes de la literatura que regulan la elongación del mesocotilo.	66
Comparación entre el GWAS y el análisis de QTL de Zhang y colaboradores.	67
Asociación genómica y análisis de QTL en la población DTP.	69
CONCLUSIONES	72
PERSPECTIVAS	72
LITERATURA CITADA	73
Anexos	89

ÍNDICE DE CUADROS

Cuadro 1. SNPs obtenidos mediante el método de genotipificación por secuenciación masiva (GBS) de las líneas endogámicas DTMA y su filtrado. Las primeras dos columnas del lado izquierdo señalan el número total de SNPs por cromosoma; Filtro MAF 0.05) indica el número de SNPs remanentes después de aplicar el filtro de frecuencia mínima del alelo menor 0.05; Filtro CM 150) indica el número de SNPs remanentes después de aplicar a la columna de su costado izquierdo el filtro de conteo mínimo 150. El mapeo por GBS se hizo tomando como referencia la versión B73_V2 del genoma de maíz.	29
Cuadro 2. Cantidad de SNP asociados a la elongación del mesocotilo que fueron deducidos mediante GWAS en cada cromosoma con sus valores p (-log ₁₀ p) de asociación máximos y mínimos.	33
Cuadro 3. Relación de los 51 SNPs asociados a la elongación del mesocotilo en maíces DTMA mediante GWAS. Clave) Cromosoma y posición física del SNPs asociado Cr) ubicación cromosomal y posición del SNP en número de bases; p) probabilidad [-log ₁₀] de asociación; Ma) número de accesiones que presentaron el estado alélico homocigoto mayor; EF Ma) efectos fenotípicos de los estados alélicos homocigotos mayores; Me) número de accesiones	

que presentaron el estado alélico homocigoto menor, EF Me) efectos fenotípicos de los estados alélicos homocigotos menores; He) número de accesiones que presentaron el estado heterocigoto; EF He) efectos fenotípicos de los estados heterocígos, en rojo se encuentran los SNPs que presentaron menos de 25 accesiones con el estado alélico menor.35

Cuadro 4. Listado de los genes candidatos y las funciones conocidas, deducidas de acuerdo a su pertenencia a alguna familia de proteínas (PFAM). Clave) Ubicación cromosomal y posición del SNP en número de bases, Pos) posición relativa del SNP, ya sea Intergénico (inter) o dentro del marco de lectura de un gen o modelo de gen (Marc), $-\log_{10}p$) valor de p en escala \log_{10} , Gen) modelo genético y PFAM) función conocida para la familia de las proteínas.37

Cuadro 5. Relación de los términos ontológicos (Go terms) de las funciones bioquímicas conocidas de los genes asociados a la elongación del mesocotilo en líneas DTMA. Gen) modelo de gen, GO terms) término de ontología genética, y Función) función molecular conocida.40

Cuadro 6. Las diez líneas DTP con los mayores o menores valores promedio de longitud mesocotilo, Nombre) número de accesión n) número de plántulas medidas, D.S.) desviación estándar, Máx) valor máximo de mesocotilo y Min) valor mínimo de mesocotilo.43

Cuadro 7. Filtros aplicados a los SNPs de las líneas DTP.49

Cuadro 8. Construcción de los bloques usados en los análisis de QTL para la población DTP y sus características. SNPs) SNPs usados con frecuencias mayores a 0.05 del alelo menor para formar bloques en cada cromosoma; Filtrado) número de SNPs que pasaron los filtros de calidad; SNPs ligados) el número de SNPs ligados; Bloques) el número de bloques obtenidos; Longitud) la longitud total de los bloques de cada cromosoma en centimorgans.50

Cuadro 9. QTL asociados a la elongación del mesocotilo para la población DTP. QTL) Nombre, número de cromosoma donde se encuentra y número de QTL en ese cromosoma; cM) posición en centimorgan del QTL; Pos inicio) posición física donde inicia el QTL; Pos término) posición física donde termina el QTL; LOD) valor de asociación del QTL; PVE) variación fenotípica explicada y GWAS) nombre los SNP asociados a la elongación del mesocotilo dentro del QTL. En negritas se resalta el valor más alto de asociación.53

Cuadro 10. Coincidencias entre los QTLs para la elongación del mesocotilo propuestos por Zhang y cols *et al.* (Zhang *et al.*, 2012) y los SNPs asociados a la elongación del mesocotilo obtenidos por GWAS en líneas DTMA. QTL) nombre del QTL identificado por Zhang y cols. (2012); Cr) cromosoma; In) posición inicial del QTL de Zhang y cols. (2012); Fin) posición final

del QTL de Zhang y cols. (2012); LOD) valor de asociación de los QTL de Zhang y cols. (2012); Núm) número de SNPs definidos por GWAS en líneas DTMA dentro de los QTLs; SNP) nombre de los SNPs y; p) valor de asociación.54

Cuadro 11. Coincidencias entre los SNPs asociados a la elongación del mesocotilo en la colección DTMA que se encontraron en los QTLs de Zhang y cols. (2012), los QTLs de la población DTP mapeados en este trabajo y la función conocida de los genes asociados.55

ÍNDICE FIGURAS

Figura 1. Esquema de la plántula de maíz. Php) primera hoja plumular, Col) coleoptilo, Rancs) raíces adventicias del nodo coleoptilar, Nc) nodo coleoptilar, Meso) mesocotilo, Car) cariopsis, Rp) raíz primaria y Rs) raíces seminales.7

Figura 2. Esquema que explica el proceso de generación de la población de maíz ABQTL-DTP-BC₁F_{2:3} (DTP) a partir de sus líneas endogámicas parentales y de las cruzas llevadas a cabo (Trachsel *et al.*, 2016).21

Figura 3. Representación esquemática del sistema usado para evaluar la respuesta a la siembra profunda en maíz mediante el uso de tubos de PVC.22

Figura 4. Distribución de las 272 líneas DTMA de los valores promedio del tamaño del mesocotilo y número en cada categoría. Datos obtenidos con los ensayos de respuesta a la siembra profunda en plántulas cultivadas en la oscuridad durante 7 días a 28°C. Datos tomados y regraficados de: Nieto-Sotelo J. *et al.*, 2012; Nieto-Sotelo J. *et al.*, 2013).27

Figura 5. Longitudes promedio de los mesocotilos y desviaciones estándar en los 10 híbridos DTMA con los mesocotilo más largos (verde), en los 10 híbridos con los mesocotilos intermedios (negro) y en los 10 híbridos con los mesocotilos más cortos (rojo). Se realizó ANOVA y un análisis *post hoc* de Tukey. Los valores con letra diferente son estadísticamente diferentes.27

Figura 6. Análisis de los 955,690 SNPs identificados por GBS en la colección de 276 líneas endogámicas de maíz DTMA. A) porcentaje de SNPs con diferente índice de heterocigocidad, B) porcentaje de SNPs con diferentes frecuencias de los alelos menores, C) porcentaje de SNPs con diferentes frecuencias de datos faltantes, D) porcentaje de accesiones con distintos niveles de heterocigocidad y E) porcentaje de accesiones con diferentes frecuencias de datos faltantes.28

Figura 7. Análisis de estructura poblacional de la colección DTMA basado en componentes principales (PC), se muestran los tres primeros componentes principales. Círculos rojos, accesiones con promedio de mesocotilos menores a 6.34 cm (mesocotilo corto); Triángulos grises, accesiones con promedio de mesocotilo mayor a 6.34 cm y menores a 8.85 cm (mesocotilo mediano); círculos verdes, accesiones con promedio de mesocotilo mayor a 8.85 cm (mesocotilo largo).30

Figura 8. Gráfica cuantil-cuantil (QQ) que muestra los valores esperados de p y los valores observados de p para los SNPs usados en el GWAS. La línea azul indica el valor p mínimo designado para aceptar a los SNPs con los mayores niveles de asociación a la variación en la longitud del mesocotilo.31

Figura 9. Número de SNPs y sus valores de asociación p (-log10). La línea azul indica el valor de menos log10 p 3.25, que es el mínimo para considerar a los SNPs fuertemente asociados a la variación de la longitud del mesocotilo de la colección de maíz DTMA.32

Figura 10. Gráfica tipo Manhattan de la asociación MLM para la elongación del mesocotilo en la colección de maíz DTMA. La línea azul señala el valor máximo asociado y la línea roja el valor mínimo aceptado.32

Figura 11. Valores de asociación p con la varianza en la elongación del mesocotilo de la colección DTMA y sus efectos fenotípicos calculados para los alelos mayores y alelos menores de los 63 SNPs asociados. EF Ma) efectos fenotípicos de los alelos mayores, EF Me) efectos fenotípicos alelos menores.34

Figura 12. Categorías de términos ontológicos (*Go terms*) de las funciones generales de los productos de los genes asociados a la variación de la elongación del mesocotilo. Se señala el número de genes asociados en cada categoría.42

Figura 13. Promedios y desviaciones estándar de los mesocotilos de las 173 accesiones de la población DTP.46

Figura 14. Correlaciones de Pearson entre pares de estructuras medidas para las líneas DTP, A) tallo y mesocotilo, B) coleoptilo y mesocotilo, C) primera hoja plumular y mesocotilo, D) raíz primaria y mesocotilo, E) raíz primaria y coleoptilo, F) tallo y coleoptilo, G) primera hoja plumular y coleoptilo, H) tallo y raíz primara, I) primera hoja plumular, J) primera hoja plumular y tallo, ****) $P < 0.0001$48

Figura 15. Características de los SNP obtenidos por GBS y de las líneas DTP, a) porcentaje de SNPs en heterocigosis, b) porcentaje de los SNPs como alelos de frecuencia mínima, c)

porcentaje de SNPs y su proporción de datos faltantes, d) porcentaje de líneas heterocigotas, e) porcentaje de líneas y su proporción de datos faltantes.49

Figura 16. Mapa de los eventos de recombinación de la población DTP. La línea parental 1 (CML491) está señalada en azul y aporta el 0.625 de la variación genética; la línea parental 2 (DTMA 227) está señalada en rojo y aporta el 0.125 de la variación genética; los sitios heterocigotos en verde aportan el 0.25 de la variación genética y los sitios faltantes se señalan en gris.51

Figura 17. Mapas de ligamiento de los SNPs para los 10 cromosomas de maíz de la población DTP. Las líneas indican la posición física de los bloques definida por los SNP. Como referencia del inicio se utilizó 0 y como referencia del término del cromosoma se utilizó el número total de pares de bases del genoma de B73_v2.7. Las regiones en negro denotan la ubicación de una gran concentración de bloques, ya que la escala y la alta concentración de marcadores dificulta su observación. Las posiciones del mapa señalan el punto central de cada bloque.51

Figura 18. Distribución de QTLs y sus valores de asociación LOD para la elongación del mesocotilo en los 10 cromosomas de la población de maíces DTP. Se tomó como referencia el punto medio de las regiones genéticas asociadas La línea roja marca el valor mínimo aceptado como QTL.52

Figura 19. Mapa físico de las regiones genéticas asociadas a la resistencia a la siembra a 10 y 20 cm descritas por Zhang y colaboradores (2012) [rojo], regiones asociadas a la variación de la elongación del mesocotilo en la población DTP [verde] y SNPs asociados a la variación de la elongación del mesocotilo en la colección DTMA [azul].54

Figura 20. Esquema del proceso de evolución de la elongación del mesocotilo, variación fenotípica sometida a la selección de siembra profunda y siembra somera en la domesticación y mejoramiento.60

RESUMEN

La enorme diversidad fenotípica y genética del maíz lo hacen un excelente modelo para estudiar las adaptaciones a la sequía y el calor en plantas. En la Mixteca Alta de Oaxaca, y en otras regiones áridas y semiáridas de México y del suroeste de los Estados Unidos, la siembra de maíz y el crecimiento pre-reproductivo se realizan en condiciones de sequía, con la ayuda de los siguientes factores: 1) terrenos con humedad residual, 2) siembras entre los 15 y 40 cm de profundidad y 3) variedades locales resistentes a la siembra profunda. Para conocer los genes involucrados en la resistencia a la siembra profunda, se analizaron 272 líneas DTMA (*Drought Tolerant Maize for Africa*) y 173 líneas de una población recombinante ABQTL-DTP BC₁F_{2:3} del CIMMYT (Centro Internacional de Mejoramiento de Maíz y Trigo) mediante ensayos de siembra profunda y genotipificación por secuenciación (GBS). La emergencia de las plántulas y la longitud del tallo correlacionaron positivamente con la longitud del mesocotilo. La variación en la elongación del mesocotilo en las líneas DTMA permitió identificar 63 polimorfismos de un solo nucleótido (SNP) asociados a su varianza mediante estudios de asociación genómica (GWAS) ubicados dentro o cerca de 82 marcos de lectura abierta. Los SNPs asociados presentaron las siguientes funciones: catálisis enzimática, transporte de iones metálicos, hidrólisis de proteínas y carbohidratos, interacción con ADN, resistencia a estrés, transporte intracelular e interacción con ARN. En la población ABQTL-DTP BC₁F_{2:3} se encontraron 29 QTLs (Loci de caracteres cuantitativos) mediante análisis de ligamiento de alta densidad. Se encontraron dos coincidencias entre los genes asociados a la elongación del mesocotilo y los QTL en la población ABQTL-DTP BC₁F_{2:3}. Queda pendiente formular modelos que expliquen las interacciones entre los *loci* asociados antes de que se puedan proponer planes de mejoramiento que, mediante el uso de marcadores moleculares, permitan introducir la resistencia a la siembra profunda a variedades de alto rendimiento o a maíces nativos.

ABSTRACT

Because of its large phenotypic and genomic diversity, maize is an excellent model to study adaptations to drought and heat in plants. In arid or semiarid regions of Mexico, such as in the Mixteca Alta of Oaxaca, as well as in the USA southwest, maize sowing and pre-reproductive growth are carried under drought conditions with the help of the following factors: 1) soils with residual moisture, 2) sowing between 15 and 40 centimeters depth, and 3) deep sowing resistant local landraces. To identify genes involved in resistance to deep sowing, we analyzed 272 DTMA lines (Drought Tolerant Maize for Africa) and 173 recombinant ABQTL-DTP BC₁F_{2:3} lines from CIMMYT (Centro Internacional de Mejoramiento de Maíz y Trigo) by performing deep sowing resistance tests in a growth chamber and genotyping by sequencing (GBS). Seedling emergence and stem length positively correlated with mesocotyl length. Variations in mesocotyl length allowed the identification of 63 SNPs (single nucleotide polymorphisms) associated to mesocotyl length variation through genome-wide association studies (GWAS) that mapped inside or close to 63 open reading frames, some of which encode proteins involved in the following functions: enzymatic catalysis, ion metal transport, hydrolysis of proteins and carbohydrates, interaction with DNA, resistance to stress, intracellular transport and interaction with RNA. With the help of the ABQTL-DTP BC₁F_{2:3} population we mapped 29 QTLs associated to mesocotyl elongation using a high density linkage analysis. We found two genes associated to mesocotyl elongation in the DTMA collection that mapped within QTLs (quantitative trait loci) of the ABQTL-DTP BC₁F_{2:3} population. The design of models that explain the interactions between the associated genes remains pending before precise improvement plans using molecular markers are formulated. This should facilitate the introgression of deep planting resistance to high yielding varieties or maize landraces.

INTRODUCCIÓN

El cambio climático y el calentamiento global.

Desde que se tienen registros se ha incrementado la emisión de los gases de efecto invernadero, propiciando un aumento en la temperatura media mundial. A este fenómeno se le ha denominado calentamiento global o cambio climático. El cambio climático hace referencia a cambios en la temperatura, humedad, lluvia, viento y otros fenómenos meteorológicos. Se prevé que el cambio climático tendrá impactos negativos en los sistemas ecológicos, humanos, urbanos, económicos y sociales (Stoker *et al.*, 2013).

Para predecir los posibles cambios ambientales, se han generado escenarios climáticos a partir de registros históricos de temperatura, emisión de gases de efecto invernadero y precipitación. Los escenarios climáticos futuros suponen aumentos constantes en la temperatura media de la Tierra y cambios en los patrones de lluvias y sequías (Burke *et al.*, 2016). Estos escenarios nos obligan a plantear programas de manejo ambiental y políticas públicas para mitigar los efectos del calentamiento global y adaptarnos a la variabilidad climática (Pachauri *et al.*, 2014).

Las predicciones climáticas pueden ser especialmente relevantes para la agricultura. Se estima que debido al calentamiento global, el rendimiento de los principales cultivos disminuirá, lo que aumentará la hambruna y el riesgo alimentario (Lobell y Field, 2007) y los países que no tengan la capacidad de remediar los efectos del calentamiento global serán más afectados (Parry *et al.*, 2004).

A finales del siglo XX se registraron las temperaturas más altas del siglo en zonas tropicales y subtropicales. Se prevé que la temperatura continuará elevándose y que los veranos serán más calientes cada año (Battisti y Naylor, 2009). En África, la sequía y el calor ya han disminuido considerablemente el rendimiento del maíz. Esto tendrá impactos severos en la salud pública y la seguridad alimentaria, ya que en países como Malawi y Zimbabwe, el maíz aporta entre el 40% y el 50% de las calorías y proteínas consumidas (Cairns *et al.*, 2013).

La sequía es un periodo de tiempo más seco, en el que la precipitación pluvial es escasa, disminuye la corriente de los ríos, baja la humedad del suelo y la evapotranspiración es elevada (Dracup *et al.*, 1980). El porcentaje de la tierra en condiciones de sequía en todas sus modalidades se ha incrementado desde inicios del siglo XXI. En el último siglo la sequía extrema se incrementó del 1 al 3 % de la superficie terrestre, en tanto que la sequía severa

pasó de 5 a 10 % y la sequía moderada 20 a 28 %. Se prevé que la sequía en todas sus formas continúe aumentando en el planeta (Burke *et al.*, 2006).

En los últimos 30 años, los efectos negativos del calentamiento global se han incrementado, afectando a los cultivos de maíz, trigo y cebada (Lobell y Field, 2007). El problema más grave que enfrentan los agricultores es la falta de agua que es el resultado de la baja precipitación y/o a la falta de acceso a los cuerpos de agua y/o a la carencia de tecnologías de riego (Rosenweig y Parry, 1994).

Además del calor y la sequía, en el norte de México y el suroeste de los Estados Unidos ha cambiado la estacionalidad. Debido al incremento de la concentración de los gases de efecto invernadero (Cook y Seager, 2013), se vaticina que para finales del siglo XXI en esta región el monzón de verano (julio y agosto), que aporta actualmente hasta el 70% de la precipitación anual, se retrasará paulatinamente y se establecerá en otoño (septiembre-octubre). La modificación de la estacionalidad obligará a los agricultores a tomar medidas de remediación: cambiar las fechas de siembra, usar otros cultivos e implementar nuevas técnicas de cultivo, que requieran menos agua y otros insumos y que tengan menor impacto ecológico (Lobell y Field, 2007).

El cultivo del maíz.

El maíz es una de las especies más cultivadas en todo el mundo y es un elemento imprescindible de la vida social, cultural y económica de México. Su diversidad genética y fenotípica es el producto de la intensa selección realizada por campesinos, agrónomos y mejoradores. La domesticación del maíz fue hace 9,000 años aproximadamente después del último periodo glacial en el suroeste de México. En este periodo de tiempo los cultivos se realizaban en los márgenes de los ríos y a pesar de esto estaban constantemente expuestos a periodos de sequía (Ranere *et al.*, 2009; Piperno *et al.*, 2007).

Las variedades de maíz representan sistemas genéticos complejos que han permitido su cultivo en ambientes poco favorables. La clasificación de las variedades de maíz considera sus características físicas, fisiológicas, genealógicas, genéticas, citológicas y geográficas (Wellhausen *et al.*, 1951; Sánchez *et al.*, 1993). En términos prácticos, las variedades de maíz están definidas por el número de características que permiten definir a una población (Anderson y Cutler, 1942). Algunas de las razas son precursoras de los llamados híbridos o variedades mejoradas. Los híbridos tienden a la uniformidad fenotípica y genética, lo que

facilita su cultivo y la cosecha mecanizada (Muñoz *et al.*, 1998). Estos programas de mejoramiento se han enfocado en aumentar el rendimiento por hectárea. Las variedades más usadas para el mejoramiento en los Estado Unidos son B73 y Mo17 (Mikel y Dudley, 2006).

México es considerado el centro de origen del maíz y uno de sus centros de diversidad, ya que aquí se encuentran en contacto las razas denominadas antiguas que son las variedades de valles altos (Matsuoka *et al.*, 2002) con sus ancestros silvestres, los teocintles (Mangeldorf, 1938). La diversidad genotípica y fenotípica del maíz puede ser atribuida a los siguientes factores: a la gran diversidad de ambientes donde es cultivado, a la dispersión del polen dado que es una especie con polinización cruzada y al intercambio de semilla entre comunidades (Anderson y Cutler, 1942; Wellhausen *et al.*, 1951).

La siembra tradicional del maíz se realiza en pequeñas parcelas y, en la mayoría de los casos, el cultivo es de temporal. Por otro, lado gran parte de la producción de maíz en el mundo se genera en las regiones templadas y es dedicada al consumo animal o a la extracción de materiales industriales. Sin embargo, en México, la mayoría de la producción se utiliza para consumo humano directo (Muñoz *et al.*, 1998).

La siembra profunda.

La siembra profunda se lleva a cabo en regiones áridas y semiáridas de todo el mundo. Se practica en el cultivo del sorgo, la cebada y el trigo. Algunas variedades de sorgo elongan el mesocotilo y pueden emerger de una profundidad de hasta 10 cm (Maiti, 2017). La resistencia a la siembra profunda en la cebada y el trigo se correlaciona con la elongación del coleoptilo (Hoshikawa, 1969). Se han encontrado variedades de cebada provenientes de África que pueden emerger de hasta 18 cm de profundidad y, en la meseta de Loess en China, la variedad de trigo *Hong Mang Mai* [*Triticum aestivum* L.] emerge de una profundidad de hasta 20 cm (Kato, *et al.*, 2011).

La siembra profunda también se practica en el cultivo del maíz. En el sureste de Estados Unidos, la tribu Hopi, perteneciente a los pueblos Navajos, produce maíz en una región con escasa precipitación. Para cultivar maíz utilizan la técnica de siembra profunda. Los Hopi colocan la semilla alrededor de unos 30 cm de profundidad, para que la planta aproveche mejor la humedad residual del suelo. La comparación de la emergencia del maíz navajo (*Hopi*) con la de dos variedades híbridas (*Chinese* y *Boone County White*), mediante pruebas de siembra a diferentes profundidades, permitió encontrar que la elongación exagerada del mesocotilo

facilita la emergencia de las plántulas cuando se siembran profundamente y que no todas las variedades de maíz son aptas para esta técnica de cultivo (Collins, 1914).

En México también practican la siembra profunda de manera tradicional los Mixtecos del altiplano noroccidental del estado de Oaxaca. Estas comunidades han desarrollado un conjunto de variedades de maíz denominadas “de cajete” (Muñoz *et al.*, 2002). Estos maíces de cajete se dividen en ultratardíos, tardíos e intermedios y, de acuerdo al inicio de la temporada de lluvias, los campesinos eligen maíces con diferente precocidad. Además de los maíces de cajete, los mixtecos cultivan maíces precoces seleccionados para mejorar su adaptación a cada microrregión (Muñoz *et al.*, 2002).

La siembra profunda implica la utilización de un paquete tecnológico desarrollado para afrontar la sequía que consta de tres elementos: I) suelos con humedad residual, II) herramientas y técnicas adecuadas que permitan colocar la semilla donde se encuentra la humedad residual y III) variedades de maíz que puedan emerger y soportar la siembra profunda (Muñoz *et al.*, 2002). Por lo tanto, los maíces de cajete aprovechan tanto la humedad residual del suelo como la lluvia. Los agricultores que realizan la siembra profunda utilizan la pala o la coa, escarban el terreno hasta encontrar la humedad residual que puede llegar a una profundidad de 35 cm y colocan de tres a cuatro semillas por punto de siembra. La emergencia de las plántulas demora entre 10 y 15 días, dependiendo de la temperatura y de la profundidad de siembra (Muñoz *et al.*, 2002).

Debido a la variación en la emergencia se han realizado estudios para determinar qué variedades pueden soportar la siembra profunda. Se encontró que las provenientes de valles altos tienen los mesocotilos más largos, mientras que las de altitudes bajas tienden a presentar mesocotilos cortos (Embarcadero, 1971; Vargas, 1966; Cruz-Sánchez *et al.*, datos no publicados; Martínez-Nava., datos no publicados).

El mesocotilo.

En una plántula de maíz el mesocotilo es el internodo del tallo que conecta al escutelo y a la radícula con el nodo coleoptilar y el coleoptilo (Figura 1). Esta estructura se alarga para elevar al coleoptilo, permitiéndole emerger a la superficie (Hayward, 1951).

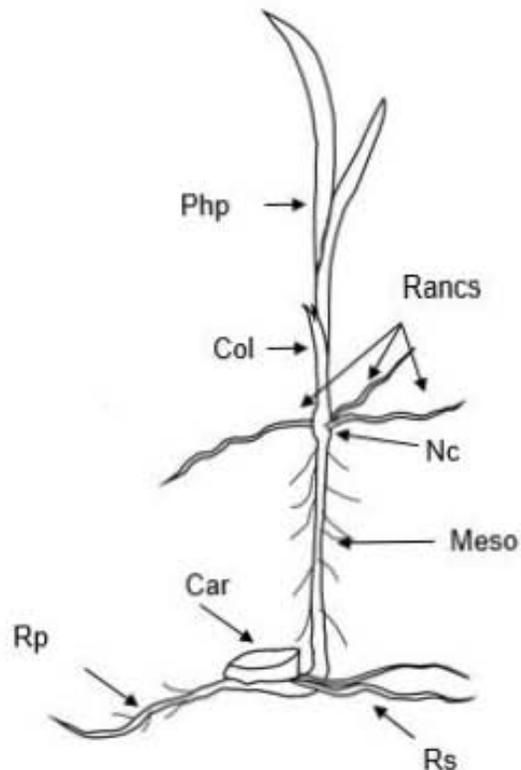


Figura 1. Esquema de la plántula de maíz. Php) primera hoja plumular, Col) coleoptilo, Rancs) raíces adventicias del nodo coleoptilar, Nc) nodo coleoptilar, Meso) mesocotilo, Car) cariopsis, Rp) raíz primaria y Rs) raíces seminales.

El crecimiento del mesocotilo se da por la división y el alargamiento de sus células, que son estimuladas por fitohormonas, principalmente por las auxinas y las giberelinas (Walton y Ray, 1981; Fellner *et al.*, 2006; Dong *et al.*, 2013). Las auxinas regulan el desarrollo de los meristemas, la organogénesis, la floración, el crecimiento celular y el gravitropismo (Muday y DeLong, 2001). La concentración de auxinas en un tejido es controlada por su síntesis y/o por el transporte desde otros tejidos, a través de transportadores específicos (Muday y DeLong, 2001; Dong *et al.*, 2013). Las auxinas promueven la elongación de las células en la base del nodo coleoptilar, aumentando la longitud del mesocotilo (Lino y Carr, 1982) siendo sintetizadas y transportadas desde la punta del coleoptilo. Su transporte es inhibido cuando la punta del coleoptilo recibe la luz solar, provocando irreversiblemente la inhibición del crecimiento del mesocotilo (Walton y Ray, 1981; Fellner *et al.*, 2006). Si las plántulas son expuestas a altas temperaturas, el crecimiento del mesocotilo se inhibe debido a una disminución de la concentración de auxinas en el coleoptilo (Overbeek, 1936).

La elongación del mesocotilo se inhibe por potenciales hídricos menores a -0.3 MP en el sustrato y por altas concentraciones del ácido abscísico, condiciones que promueven el desarrollo de la raíz (Saab *et al.*, 1992). El movimiento del agua dentro de las plantas requiere un gradiente de potencial hídrico entre el sustrato y la planta. En condiciones de bajo potencial hídrico, cuando el suelo tiene menor potencial hídrico que la planta, el transporte de agua hacia la planta disminuye, ocasionando la pérdida de turgencia y deteniendo la elongación de las células (Tang y Boyer, 2001). Por otro lado, la radiación ultra violeta (UVB) activa a la sintasa de óxido nítrico, aumenta a las proteínas de la pared celular y disminuye la actividad de una endoglucanasa. En consecuencia, aumenta el contenido intracelular del óxido nítrico y se inhibe la elongación de las células del mesocotilo (Zhang *et al.*, 2007).

Genética de la variación en la elongación del mesocotilo.

Poco se conoce acerca de la genética involucrada en la amplia variación observada en la elongación del mesocotilo del maíz o de otras gramíneas. Troyer (1997) realizó experimentos utilizando como parental al maíz *Hopi* y determinó que la elongación del mesocotilo es un rasgo altamente heredable que puede ser usado en planes de mejoramiento. Encontró cinco regiones cromosómicas fuertemente asociadas a dicho rasgo distribuidas en los cromosomas 1, 3, 6 y 9 del maíz. Por otro lado Zhang y colaboradores (2012) realizaron análisis de QTLs (*Quantitative Trait Loci*) para identificar las regiones cromosómicas involucradas en la resistencia a la siembra profunda. Utilizando cruza entre dos líneas endogámicas, una con mesocotilos cortos (X178) y la otra con mesocotilos largos (3681-4, desarrollada a partir del maíz *Hopi*), encontraron que la resistencia a la siembra a 10 y 20 cm de profundidad es regulada por 25 QTLs con efectos menores.

Mediante genética clásica se han identificado dos genes cuya mutación afecta directamente la elongación del mesocotilo en maíz. Los mutantes *elm1* (mesocotilo elongado) no responden a la luz roja ni a la roja lejana y acumulan grandes cantidades de fitocromo A, porque no sintetizan el cromóforo responsable de esta respuesta, interrumpiendo así las rutas de señalización y procesamiento reguladas por la luz. En luz blanca, estos mutantes elongan más el mesocotilo y además inician la floración antes que el silvestre (Sawers *et al.*, 2002). El producto del gen *Elm1* es una fotocromobilina sintasa: ferredoxina oxidoreductasa involucrada en la síntesis del cromóforo de los seis fitocromos del maíz (Alexandrov *et al.*, 2009). Otro de los genes directamente involucrados en la elongación del mesocotilo es *ZmLA1*. Los mutantes

de este gen presentan las siguientes características: mesocotilos alargados, tallo postrado, no son geotrópicos y tienen problemas en la floración y en la formación del fruto. Este gen expresa una proteína con función desconocida que está involucrada en el transporte y la señalización de las auxinas y de allí sus efectos pleiotrópicos sobre la planta (Dong *et al.*, 2013).

Mejoramiento genético.

Los planes de mejoramiento emplean poblaciones generadas a partir de cruzas controladas. Los parentales son elegidos por tener características de valor agronómico o por generar vigor híbrido al cruzarse con otro individuo (Varshney *et al.*, 2015; Meng *et al.*, 2015). La parte más importante del mejoramiento es la correcta selección de los parentales que permitan obtener el fenotipo deseado. Los procesos evolutivos que han actuado sobre las variedades silvestres y las razas han aumentado la variación de las características fenotípicas de la especie que podrían ayudar en los planes de mejoramiento. Para que una característica se considere de valor agronómico debe ser fácilmente heredable, medible cualitativa o cuantitativamente y, preferentemente, no estar regulada por procesos epistáticos.

Las variedades silvestres y las razas son una fuente de diversidad genética pero pueden tener características inapropiadas para el cultivo intensivo y, por ello, es necesario emplear herramientas que nos permitan facilitar la introgresión genética del material deseado en variedades anteriormente mejoradas (Collard *et al.*, 2005). Por esta razón es necesario conocer las regiones génicas y alelos involucrados en la regulación de los caracteres deseados con la finalidad de reducir el tiempo y costo invertidos en el mejoramiento genético (Mohan *et al.*, 1997). Para esto se utilizan líneas élite como receptores, ya que tienen características favorables para el cultivo extensivo. El reconocimiento y el mapeo de los QTL permiten estimar cuántos individuos y generaciones son necesarias para el mejoramiento, ya sea por la introgresión de material genético o por la segregación de caracteres dentro de una población (Mohan *et al.*, 1997). Para realizar la introgresión genética es necesario conocer el tamaño del segmento del donador, la posición y la cantidad de genoma que podría arrastrar que puede no ser favorable o entrar en conflicto con alelos previamente introgresados (Lin *et al.*, 2014) y se deben identificar los donantes con el carácter o fenotipo deseado. Por otro lado, los materiales de élite receptores deben ser varios y pertenecer a diferentes grupos heteróticos para promover el vigor híbrido al realizar las cruzas y culminar con la creación de nuevas variedades (Ndhela *et al.*, 2015).

Para reconocer los genes asociados a un carácter, se emplean los QTLs que utilizan técnicas de biología molecular (Mohan *et al.*, 1997). A través de las técnicas moleculares se han encontrado QTLs asociados al rendimiento, a la calidad del cultivo y a la resistencia a algunas enfermedades (Collard *et al.*, 2005; Zila *et al.*, 2014). Estas características, conocidas como rasgos cuantitativos, generalmente son procesos multifactoriales o complejos y parte de su variación puede ser explicada genéticamente (Mohan *et al.*, 1997). Es importante que los rasgos a mapear no sean afectados por factores ambientales y que nos ayuden a distinguir entre grupos e individuos (Collard *et al.*, 2005). Se espera que los QTL presenten efectos sobre el fenotipo, sin importar el trasfondo genético, independientemente de cuántos genes estén involucrados en la regulación de un carácter.

Para controlar la variación genética y fenotípica los mejoradores utilizan poblaciones biparentales, cruzas controladas, retrocruzas o técnicas de haploidía (Tanksley y Nelson, 1996). Estas técnicas permiten observar la segregación de los genes responsables de la variación de los rasgos de interés agronómicos siempre y cuando existan suficientes eventos de recombinación que rompan el desequilibrio de ligamiento (DL) entre los marcadores asociados y no asociados. La recombinación es el intercambio de material genético entre cromosomas. En los eucariontes diploides el proceso de recombinación se da en la meiosis y éste proceso asegura que los gametos incluyan material genético de ambos parentales y que la descendencia tenga información genética de los cuatro abuelos (Clancy, 2008). El DL es la asociación no azarosa entre dos marcadores o entre dos genes ocasionada por la deriva génica y/o la estructura poblacional (Soto-Cerda y Cloutier, 2012). Algunas veces el DL es confundido con el ligamiento, que se refiere a la unión de dos *loci* debida a una conexión física en el cromosoma, característica que se mantiene por mucho tiempo (Flint-Garcia *et al.*, 2003). Es necesario tener marcadores distribuidos uniformemente en el genoma, ya que, si se encuentran a una distancia considerable del gen responsable de la variación, podrían presentarse eventos de recombinación y se corre el riesgo de obtener falsos positivos y que los marcadores asociados al fenotipo de interés no sean válidos para otras poblaciones (Soto-Cerda y Cloutier, 2012).

Algunos de los marcadores moleculares que se han empleado en el reconocimiento de QTLs son: los polimorfismos de tamaño de fragmentos de restricción que son generados por enzimas de restricción (RFLP, *Restriction Fragment Length Polymorphism*), los microsatélites (SSR) y los polimorfismos de un solo nucleótido (SNPs, *Single Nucleotide Polymorphism*)

(Collard *et al.*, 2005; Zhang *et al.*, 2012; Bush y Moore, 2012). La identificación de polimorfismos por RFLP es una técnica que consiste en la digestión, con endonucleasas de restricción específicas del ADN genómico para generar fragmentos que van de 500 a 2,000 pares de bases que se separan mediante la técnica de electroforesis y las bandas obtenidas son transferidas a un filtro de nitrocelulosa para realizar un *Southern blot* que finalmente permite el reconocimiento, por sondas específicas de gen, del tamaño y el número de los fragmentos reconocidos (Williams, 1989). Para el maíz se han realizado mapas de ligamiento hasta con 338 RFLPs distribuidos en los diez cromosomas (Helentjaris, 1987). En la población IBM (*Intermediated B73 x Mo17*) Syn4 se obtuvo un mapa de ligamiento de 5,917 cM utilizando 190 *RFLsP*. La población IBM Syn4 es utilizada para el análisis de QTLs, para elaborar mapas de alta densidad y en genómica comparada. La población IBM Syn4 es tomada como referencia porque está compuesta por la cruce simple de (B73 x Mo17) y cinco ciclos de autocruzas con el propósito de aumentar el número de eventos de recombinación y con ello la resolución de los mapas de ligamiento (Lee *et al.*, 2002). Las variedades B73 y Mo17 son las líneas endogámicas mejor estudiadas y pertenecen a grupos heteróticos diferentes. Estas características han posibilitado que la población IBM Syn4 sirva como referencia genética de mapeo de QTLs (Liu *et al.*, 2015). Los SSR permiten el reconocimiento de los polimorfismos con base en el número de repeticiones de las regiones repetitivas del genoma. Para la identificación de estos marcadores se diseñan oligonucleótidos que se adhieren a las regiones que rodean estas secuencias repetitivas y se amplifican por PCR. Se han utilizado 748 SSRs en la población IBM con la que se contruyó un mapa de ligamiento de 4,906 cM (Sharopova *et al.*, 2002). Los marcadores RFLP y SSR tienen un alto costo y el uso de la técnica de PCR con oligonucleótidos poco específicos afecta la reproductibilidad de estos análisis.

Los SNPs normalmente tienen dos formas comunes dentro de una población. Su frecuencia alélica está dada en términos de la frecuencia alélica menor (MAF, *Minor allele Frequency*) que se refiere al número de individuos que presentan el segundo alelo con mayor frecuencia (Bush y Moore, 2012). Gracias al surgimiento de las nuevas técnicas de secuenciación masiva, se han realizado monitoreos parciales del ADN genómico de poblaciones que reconocen una gran cantidad de SNPs, que son útiles para identificar a los genes asociados a los rasgos de interés (Elshire *et al.*, 2011; Korte y Farlow, 2013). Con las técnicas de secuenciación masiva se obtuvo un mapa de ligamiento de alta calidad de la

población IBM Syn10 de 11,198.5 cM, utilizando 1,151,856 SNPs de alta calidad que fueron reunidos en 6,632 grupos de recombinación [*Recombination Bins*] (Liu *et al.*, 2015).

Para localizar los genes asociados a un rasgo se utilizan los mapas de asociación, que pueden ser mapas de ligamiento o mapas de GWAS (estudios de asociación a nivel de todo el genoma o *Genome-Wide Association Studies*). Los mapas de ligamiento dependen de la segregación de la diversidad alélica en los descendientes provenientes de la cruce de sus parentales. En una población con pocos eventos de recombinación el DL se mantiene, impidiendo que los marcadores segreguen de forma independiente reduciendo la resolución del mapa de ligamiento (Semagn *et al.*, 2006). Por otro lado, la asociación genómica depende de la diversidad genética y fenotípica de la población, así como de las distancias genéticas entre las accesiones (Nemri *et al.*, 2010; Korte y Farlow, 2013). Los análisis de asociación genómica pueden ser complementarios a los mapas de ligamiento (Soto-Cerda y Cloutier, 2012). La finalidad de localizar los QTLs es tener herramientas suficientes que ayuden a realizar la selección asistida por marcadores moleculares (MAS, *Marker Assisted Selection*) (Semagn *et al.*, 2006; Varshney., 2015).

Genotipificación mediante secuenciación (GBS).

Las técnicas de secuenciación de nueva generación, también conocidas como de secuenciación masiva, permiten el descubrimiento de SNPs y ayudan a comprender la diversidad genotípica de las especies a través de sus diferencias genéticas. Los SNPs pueden ser usados tanto para la construcción de mapas de ligamiento como para el análisis de asociación genómica (Elshire *et al.*, 2011; Mascher *et al.*, 2013).

La secuenciación masiva puede ser aplicada al análisis de genomas pequeños o grandes. Debido a la complejidad de algunos genomas (por ejemplo: el humano o el maíz), se han implementado técnicas que permiten reducir la información de secuencias repetitivas y transposones, enfocándose en las regiones codificantes. La técnica de Genotipificación Mediante Secuenciación (GBS, *Genotyping By Sequencing*) emplea enzimas de restricción sensibles a la metilación del ADN. Estas enzimas no cortan las secuencias de ADN metiladas, esquivando primordialmente a las regiones inactivas o repetidas del genoma. El genoma es recortado en pequeños fragmentos que van de los 170 a los 350 pares de bases. Una vez obtenidas las secuencias de los extremos, éstas se alinean, usando un genoma de referencia

para realizar el llamado de los SNPs, que es el paso donde se reconocen los polimorfismos que hay en una población.

Uno de los principales problemas de la técnica GBS es que genera secuencias cortas, que pueden alinearse incorrectamente si el genoma de referencia no está bien ensamblado. Además, esta técnica tiene una baja profundidad de secuenciación y una menor calidad de secuenciación que la técnica de Sanger (Elshire *et al.*, 2011; Glaubitz *et al.*, 2014). Con las secuencias alineadas al genoma de referencia se realiza el llamado de los SNPs (Deschamps *et al.*, 2012; Glaubitz *et al.*, 2014) mediante algoritmos que realizan el llamado de los SNPs considerando los posibles errores. Existe una gran cantidad de *software* libre para realizar el llamado como: *GATK*, *SAMtools*, *SOAP*, *SNPAAMapper*, *TOMP*, *WEEP* y *Atlas* (Kumar, *et al.*, 2014). El llamado de los SNPs depende de filtros aplicados que retiran las secuencias de baja calidad o poco confiables mediante cálculos probabilísticos.

Si bien la técnica de GBS permite obtener una gran cantidad de SNPs distribuidos a lo largo de todo el genoma, hay que tener en cuenta que no necesariamente están distribuidos uniformemente en los cromosomas (Elshire *et al.*, 2011). Para el caso del maíz, se pueden obtener cerca de un millón de SNPs por individuo. Cabe aclarar que estos marcadores aún deben pasar varios filtros antes de ser utilizados en los estudios de asociación genómica y en la generación de mapas de ligamiento de alta densidad, con la finalidad de tener datos de alta calidad (Chen *et al.*, 2014; Varshney *et al.*, 2015).

Análisis de asociación.

La calidad de los mapas de asociación para identificar los QTLs depende del tamaño poblacional y del número de cruzas que se hayan realizado. Estos factores tienen un efecto directo sobre el número de eventos de recombinación. Además, se necesita tener suficientes marcadores moleculares con la finalidad de reconocer todos los eventos de recombinación. La calidad de los mapas de asociación se ve afectada por el DL, la estructura poblacional y las relaciones parentales, así como por la naturaleza del rasgo a estudiar, ya que algunas características pueden ser reguladas por un solo gen o por varios genes con pequeños efectos fenotípicos (Korte y Farlow, 2013).

En los análisis de asociación se busca que los marcadores se encuentren correlacionados al fenotipo de interés, ya sea porque están ligados o se encuentran en DL con el gen responsable de la variación fenotípica. El DL decrece dentro de una población debido a

la recombinación y por la distancia física entre genes, ya que, a mayor distancia física, es más probable que haya recombinación. Los eventos de recombinación no son homogéneos para todo el cromosoma. En los brazos se observa mayor recombinación que en los centrómeros, que son las partes más conservadas del genoma (Schnable *et al.*, 2012). Además, la recombinación depende de la cantidad de genes y del contenido de guanina y citosina en sus secuencias (Dawson *et al.*, 2002).

El DL depende de la especie, de la población y de las regiones génicas que se estén analizando. Se han observado diferencias entre los sistemas de apareamiento. Por ejemplo: las especies que se autofecundan tienden a la homocigosis y el DL es mayor que en las especies de fecundación cruzada (ya que la recombinación es menos eficiente al aumentar el DL). Para el caso del maíz, la diversidad genética juega un papel importante, debido a que se ha observado que el ligamiento entre genes comienza a decaer a una kilobase de distancia en las razas, mientras que, en las líneas endogámicas, el decaimiento comienza a dos kilobases o más (Flint-Garcia *et al.*, 2003; Jung *et al.*, 2004). Por ello, si se trabaja con líneas altamente endogámicas y se quiere romper el DL, se deben realizar varias cruza o aumentar el tamaño de la población recombinante (Collar *et al.*, 2005)

La estructura poblacional tiene un fuerte efecto sobre todo el genoma de las poblaciones y se da cuando no se cumple el apareamiento aleatorio en una población, ocasionado por la existencia de dos o más subpoblaciones que a su vez generan una desviación del equilibrio de Hardy-Weinberg (Rasmus y Slatkin, 2013). La estructura poblacional es un problema para las asociaciones, debido a que los genes se encuentran ligados por efecto de factores demográficos. Si hay varias poblaciones y solo se muestrea una se puede encontrar que dos alelos estén fuertemente asociados entre ellos. Sin embargo, en las otra poblaciones no se encuentra la misma asociación (Soto-cerda y Cloutier, 2012). La estructura poblacional aumenta el descubrimiento de falsos positivos y disminuye la capacidad de encontrar los marcadores realmente asociados a los fenotipos (Cardon y Palmer, 2003; Slatkin, 2008). Para corregir estos problemas, es importante utilizar modelos que incluyan datos de estructura poblacional y que puedan dar información para distinguir los marcadores que estén realmente asociados al fenotipo (Pritchard *et al.*, 2000; Thornsberry *et al.*, 2001).

Mapas de ligamiento.

Los mapas de ligamiento buscan la asociación de los marcadores con el fenotipo a través de la distancia genética entre marcadores y se calcula a partir de las frecuencias de recombinación (Collard *et al.*, 2005). La resolución de los mapas de ligamiento depende de la cantidad, distribución de los marcadores y el tamaño de la población. A mayor cantidad de marcadores y de individuos, será mayor la precisión al detectar los eventos de recombinación y, por lo tanto, las regiones asociadas a los rasgos de interés. Los mapas de ligamiento realizados con RFLP y SSR presentan regiones pobremente representadas y, debido a la distancia física entre marcadores, no es posible obtener información completa ni precisa de la localización y el número de QTLs. Por ello es necesario el uso de las técnicas de secuenciación de nueva generación que aumentan la densidad de marcadores en las secuencias (Yu *et al.*, 2011). Sin embargo, el uso de una alta densidad de SNPs en pequeñas regiones puede no ser lo más apropiado para los análisis de QTLs, ya que al tener muchos marcadores en regiones pequeñas se generan marcadores con información idéntica y en algunos casos hay muchos datos faltantes que dificultan el análisis y disminuyen la resolución de los mapas de ligamiento (Li *et al.*, 2015). Para solucionar este problema, se ha planteado trabajar con modelos que permitan reconocer los eventos de recombinación y haplotipos, mediante la formación de bloques o *Bins* que reúnen la información de los SNPs vecinos y la condensan en una sola región genómica. (Huang *et al.*, 2009; Xie *et al.*, 2010; Chen *et al.*, 2014; Liu *et al.*, 2015). El nivel de recombinación genómica en maíz se ve afectado por el origen de las variedades; en cada célula de origen meiotico se observan de 24 a 50 eventos de recombinación (Li *et al.*, 2014).

Los mapas de ligamiento indican la distancia genética que hay entre dos marcadores. Ésta es calculada mediante funciones de ligamiento que convierten las fracciones de recombinación en centimorgans (cM) (Collard *et al.*, 2005). No existe una relación directa entre los mapas de ligamiento y los mapas físicos, ya que los marcadores genéticos fuertemente ligados pueden estar muy alejados en términos de pares de bases de distancia, pero estar a pocos cM entre ellos (Semang, 2006). Esto se debe a las diferencias en las frecuencias de recombinación dentro de los cromosomas y a que, mientras más eventos de recombinación puedan observarse, mayor será la posibilidad de reconocer regiones génicas asociadas al fenotipo de interés (Semagn *et al.*, 2006).

La relación entre fenotipos y genotipos se obtiene al comparar los mapas de ligamiento y las variantes alélicas. Los métodos estadísticos más usados son la asociación a un solo marcador (*single-marker analysis*), el mapeo de los intervalos vecinos (*simple interval mapping*) y el mapeo de intervalos compuestos (*composite interval mapping*) (Collard *et al.*, 2005; Zhang *et al.*, 2012; Meng *et al.*, 2015). Estos métodos calculan la posibilidad de que los marcadores no estén asociados por azar (*odds ratios*) con el fenotipo. Esta relación se expresa en forma logarítmica y es llamada probabilidad logarítmica con base 10 (LOD, *Logarithm base 10 of ODDs*) (Rish, 1992). Los valores de LOD generalmente aceptados como apropiados van de 2 a 3 unidades (Collard *et al.*, 2005). Al ser una relación logarítmica, los valores mayores a 3 indican que el ligamiento es 1,000 veces mayor que el no ligamiento (Rish, 1992)

Estudios de asociación genómica (GWAS).

Los estudios de asociación a nivel de todo el genoma (GWAS) se utilizan para medir y analizar las variaciones en las secuencias del ADN y asociarlas a un fenotipo. Actualmente se aplican a la medicina humana, a la farmacología, al mejoramiento de cultivos y en el análisis de procesos evolutivos (Bush y Moore, 2012; Soto-Cerda y Cloutier, 2012). El GWAS utiliza una gran cantidad de SNPs que pueden ser obtenidos mediante técnicas de secuenciación de nueva generación, como el GBS entre otras (Soto-Cerda y Cloutier, 2012; Varshney *et al.*, 2015).

Los GWAS son una serie de pruebas estadísticas para cada *locus*. Los SNPs son tratados como eventos independientes y se correlacionan al fenotipo mediante regresiones lineales, utilizando los modelos lineales generales (GLM, *General Linear Model*), que son un análisis de varianza similar a la regresión lineal con predictores categóricos. Para aplicar este modelo se necesitan tres supuestos, 1) el fenotipo tiene una distribución normal, 2) la varianza de cada grupo es la misma y 3) los grupos son independientes (Bush y Moore, 2012). Los GLM realizan una corrección tomando en cuenta la estructura poblacional generada a partir de los genotipos y consideran falsos positivos a los SNPs asociados por eventos de migración, ancestría y selección artificial. Además, existe el modelo lineal mixto (MLM, *Mixed Linear Model*), que realiza correcciones incorporando una matriz de relaciones parentales (*kinship*) (Zhang *et al.*, 2010; Soto-Cerda y Cloutier 2012; Hoffman, 2013).

Los GWAS detectan los genes candidatos y sus efectos fenotípicos, independientemente de la complejidad genética e interacciones que regulan la variación del

rasgo de interés. Algunos rasgos pueden estar regulados por pocos genes, con grandes efectos sobre el fenotipo o por muchos genes con pequeños efectos sobre el fenotipo (Korte y Farlow, 2013). Debido a la complejidad de algunas características, se han planteado diferentes modelos de interacciones génicas para explicar los cambios fenotípicos. Los modelos más usados son el de dominancia y el aditivo. Cuando una variación genotípica (SNP) modifica directamente al fenotipo, se considera SNP funcional o causal, mientras que si es de forma indirecta, el SNP se considera que se encuentra asociado al SNP causal (Bush y Moore, 2012).

ANTECEDENTES

Nieto-Sotelo y colaboradores (2012; 2013) estudiaron la elongación de mesocotilo y de otras estructuras de la plántula (coleoptilo, primera hoja plumular, raíz primaria, raíces seminales y raíces adventicias del nodo coleoptilar) de 284 híbridos DTMA (*Drought Tolerant Maize for Africa*), que son líneas endogámicas desarrolladas por el CIMMYT provenientes de varios programas de mejoramiento que incluyen gran parte de la variación genética de las variedades tropicales y subtropicales. Las DTMA son una colección de medio hermanos generada al cruzar cada una de las líneas endogámicas DTMA con la línea endogámica *tester* CML312. La cruce con este parental favorece el vigor híbrido, lo que permite una fenotipificación confiable sin los efectos de la depresión génica por la endogamia de los parentales. Esto no representa un problema, ya que al tener todas las líneas un padre común su variación fenotípica se atribuye a la contribución genética del otro parental (Cairns *et al.*, 2013). Nieto-Sotelo y colaboradores (2012; 2013) realizaron ensayos de laboratorio en condiciones controladas en una cámara de crecimiento durante 6.5 días de cultivo, a 28°C, en oscuridad y con iluminación durante lapsos breves usando luz verde de seguridad para regar a las plántulas cada 48 hrs. Se tomaron estas precauciones para que las plántulas no inhibieran la elongación del mesocotilo que pudiese ser provocada por la luz blanca o la roja/roja lejana, ni por la falta de agua. Los ensayos en la cámara de germinación demostraron que los híbridos DTMA presentan diferencias en la elongación del mesocotilo, oscilando entre los 3 y 11 cm de largo en promedio para cada accesión.

Nieto Sotelo y colaboradores (2012; 2013) también evaluaron el rendimiento de grano (T/ha) en campo de cuatro líneas DTMA, en parcelas sembradas a diferentes profundidades (5, 16, 24 y 34 cm de profundidad) en condiciones de sequía. Concluyeron que la emergencia de los híbridos DTMA en condiciones de siembra profunda es muy pobre. No obstante, observaron que el rendimiento de grano de las plantas sembradas a 34 cm y sometidas a sequía era similar al obtenido con plantas sembradas a 5 cm de profundidad bajo riego (cerca de 10 T/ha); valor muy superior al obtenido bajo sequía en parcelas de siembras someras [5, 16 o 24 cm] (Nieto-Sotelo *et al.*, 2012; Nieto-Sotelo *et al* 2013).

En el presente trabajo se busca reconocer los *loci* asociados a la variación de la elongación del mesocotilo de la colección DTMA, utilizando los datos fenotípicos obtenidos mediante experimentos en la cámara de germinación por Nieto-sotelo y colaboradores (2012; 2013) y los datos genotípicos (SNPs) de la misma colección obtenidos por el CIMMYT (Gowda

et al, 2015). Se realizó un GWAS y se compararon los *loci* asociados con los *loci* asociados a la variación de la elongación del mesocotilo de la población recombinante ABQTL-DTP-BC₁F_{2:3} llevando a cabo un análisis de QTL de alta densidad de SNPs de esta población obtenidos por GBS.

HIPÓTESIS

La arquitectura genética que regula la variación de la elongación del mesocotilo en la colección de DTMA permitirá identificar a los SNPs asociados a este rasgo mediante análisis de asociación a nivel de todo el genoma y parte de estos *loci* asociados se encuentran regulando la variación de la elongación del mesocotilo de la población recombinante ABQTL-DTP-BC₁F_{2:3}.

OBJETIVOS

Identificar los SNPs asociados a la variación de la elongación del mesocotilo en la colección de maíces DTMA y los QTL asociados a la variación de la elongación del mesocotilo en la población recombinante ABQTL-DTP-BC₁F_{2:3}.

Los objetivos particulares:

- a. Obtener los SNPs fuertemente asociados a la variación de la elongación del mesocotilo en la colección DTMA.
- b. Inferir los genes candidatos asociados a la variación de la elongación del mesocotilo en la colección DTMA.
- c. Calcular el efecto fenotípico de los alelos de los SNPs asociados a la variación de la elongación del mesocotilo en la colección DTMA.
- d. Obtener los QTL fuertemente asociados a la variación de la elongación del mesocotilo en la población recombinante ABQTL-DTP-BC₁F_{2:3}.
- e. Evaluar posibles coincidencias entre los SNPs asociados a la variación de la elongación del mesocotilo en la colección DTMA identificados por GWAS y los *loci* asociados en la población recombinante ABQTL-DTP-BC₁F_{2:3}.

MÉTODOS

Ensayos para evaluar la respuesta de las plántulas a la siembra profunda.

Para realizar el GWAS se usaron como valores fenotípicos los promedios de elongación del mesocotilo de cada uno de los 272 híbridos de la colección DTMA, que fueron obtenidos previamente por Nieto-Sotelo y colaboradores (2012; 2013). Para esto realizaron mediciones de 10,837 plántulas de maíz (**Anexo I**). Para realizar el análisis de QTLs de alta densidad se fenotipificó la población ABQTL-DTP BC₁F_{2:3} (referida como DTP en este estudio). La población DTP está formada por 173 accesiones que fueron donadas por el Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT) de El Batán, Estado de México. La colección DTP ha sido caracterizada y estudiada previamente para determinar la resistencia a virus (Trachsel *et al.*, 2016). La población DTP resulta de la cruce de los parentales CML491 (parental 1) y DTMA 227 (parental 2). Las líneas F₁ se cruzaron con CML491 (retrocruza), para generar la BC₁F₁. Esta progenie se autofertilizó para obtener la generación BC₁F₂ y se generaron 173 líneas, que fueron secuenciadas por GBS. Finalmente se cruzaron las líneas BC₁F₂ con el *tester* CML503 de diferente grupo heterótico para generar la población BC₁F_{2:3}. Con estos híbridos se realizó la fenotipificación (Figura 2). Para esta población se ha realizado un mapa de ligamiento con 1,956.92 cM con un promedio de densidad de marcadores de 1.55 cM (Trachsel *et al.*, 2016).

Los ensayos para determinar la longitud del mesocotilo de la colección DTP se realizaron por triplicado de acuerdo a un protocolo previamente establecido y validado (Nieto-Sotelo *et al.*, 2012). Se usaron 15 semillas de maíz por bioensayo. Los 15 granos de maíz fueron sumergidos en 20 ml de una disolución de hipoclorito de sodio comercial al 5.1% (Los Patitos ® Blanqueadora Mexicana S.A. de C.V.) durante 10 minutos. Posteriormente se enjuagaron con agua corriente y fueron colocados sobre una capa de 20 cm de turba rubia pura (*peat moss*) de las marcas *Terracult*® o *Hawita*® (HAWITA GRUPPE GmbH); éstas turbas tiene un grado de descomposición bajo y medio, no contienen cal ni fertilizantes añadidos y tienen un pH (H₂O) de 3.0- 4.0. No se recomienda usar otras marcas que comercializan la turba rubia mezclada o adicionada con otros sustratos (ejemplo: la marca *Sunshine*®), ya que estos componentes adicionales promueven un crecimiento más robusto de toda la plántula y enmascaran la heredabilidad del tamaño del mesocotilo en el ensayo de la siembra profunda (Nieto-Sotelo *et al.*, datos no publicados). La turba fue humedecida y homogeneizada con 750 ml de agua corriente dentro de tubos de PVC hidráulico de tres pulgadas de diámetro y 40 cm

de altura. No se recomienda usar tubos de PVC de grado sanitario que, aunque son más baratos, liberan sustancias tóxicas (Kaley *et al.*, 2016) que pueden afectar el crecimiento de las plántulas.

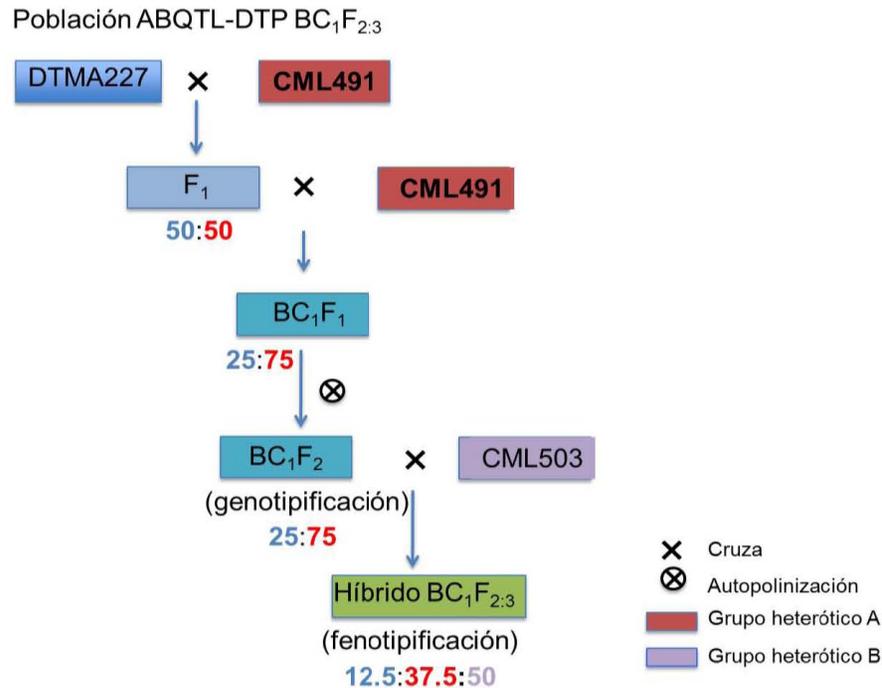


Figura 2. Esquema que explica el proceso de generación de la población de maíz ABQTL-DTP-BC₁F_{2,3} (DTP) a partir de sus líneas endogámicas parentales y de las cruza llevadas a cabo (Trachsel *et al.* 2016).

Las semillas se cubrieron con una capa de 16 cm de vermiculita mediana de la marca Agrolita® (Accimin, S.A. de C.V.) la cual fue regada con 500 ml de agua (Figura 3). Los ensayos se realizaron en la oscuridad dentro de una cámara de germinación a una temperatura constante de 28° C. Los tubos se regaron con 100 ml de agua corriente cada 48 horas, tomando la precaución de usar únicamente luz verde (como medida de seguridad) al entrar a la cámara de germinación. La luz de seguridad usada fue una bombilla *led* de color verde de 5 Watts (lámpara LD1018). Al cabo de 7 días de incubación, se extrajeron las plántulas del sustrato y se evaluó la: longitud de la raíz primaria, longitud del tallo (mesocotilo + coleoptilo + primera hoja plumular), longitud del mesocotilo, longitud del coleoptilo, longitud de la primera hoja plumular y el número de raíces asociadas al nodo coleoptilar (Figura 1).

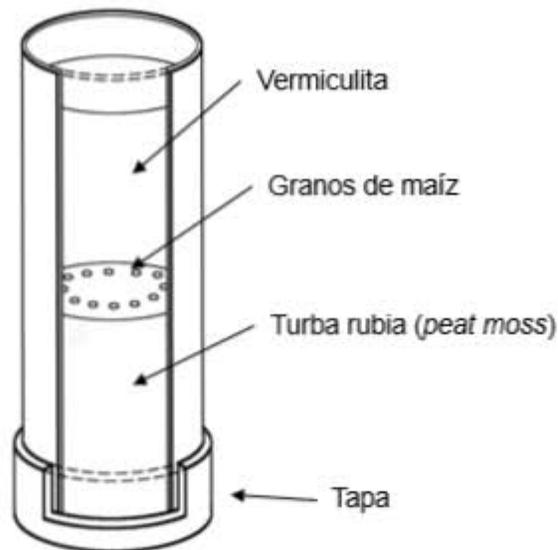


Figura 3. Representación esquemática del sistema usado para evaluar la respuesta a la siembra profunda en maíz mediante el uso de tubos de PVC.

Genotipificación por secuenciación (GBS).

La genotipificación de las líneas DTMA y DTP fue realizada en el Instituto de Biotecnología de la Universidad de Cornell, NY, Estados Unidos, a solicitud del CIMMYT (Cairns *et al.*, 2013). Se tomaron muestras de hoja de 12 plantas de cada línea endogámica DTMA, se mezclaron (*Bulk*) y se extrajo el DNA total de cada mezcla por el protocolo de CTAB (CIMMYT, 2005). La secuenciación se realizó mediante la técnica de genotipificación por secuenciación [GBS, *Genotyping By Sequencing*] usando la enzima de restricción *ApeK1* que es sensible a metilación. Para el reconocimiento de las secuencias obtenidas se utilizaron adaptadores con índices (códigos de barras) que se ligaron a cada fragmento generado y que identifican al ADN de cada línea DTMA (Elshire *et al.*, 2011). Las bibliotecas GBS se combinaron en 96-plex y fueron secuenciadas con un aparato Illumina Hiseq2000 (Gowda *et al.*, 2015; Trachsel *et al.*, 2016; Cao *et al.*, 2017). El tratamiento de los datos obtenidos se realizó de acuerdo al *pipeline Tassel-GBS*. Este protocolo solo acepta lecturas de más de 64 pares de bases sin contar los códigos de identificación. Las secuencias se apilan en posiciones físicas con relación al genoma de referencia de maíz B73_v2.7. Este protocolo, diseñado para funcionar con una profundidad de secuenciación baja (de 0.5-3x), es lo suficientemente flexible para aceptar datos con cobertura alta 20x (Glaubitz *et al.*, 2014). Estos datos genotípicos fueron facilitados por el Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT) de El Batán, Estado

de México. El llamado de los SNPs rindió casi un millón de SNPs por línea endogámica, ya fuese de la colección DTMA o de la población recombinante DTP.

Estudios de Asociación a Nivel de Todo el Genoma (GWAS) de la elongación del mesocotilo en la colección de líneas de maíz DTMA.

Los datos fenotípicos para la elongación del mesocotilo de las líneas DTMA fueron obtenidos en un estudio previo (Nieto-Sotelo *et al.*, 2012). Se utilizó como fenotipo al promedio de la longitud del mesocotilo de cada híbrido DTMA.

A los datos crudos de la lista de SNPs de cada línea endogámica DTMA (955,690 SNPs por línea) se les aplicó un filtro de frecuencia para el alelo menor (MAF, *Minor Allele Frequency*) de 0.05, retirando los SNPs con alelos de frecuencia muy baja o que podrían atribuirse a errores de secuenciación (Deschamps *et al.*, 2012). A los SNPs utilizados para el GWAS se les aplicó un segundo filtro que retiró los que presentaron un exceso de datos faltantes en las accesiones, mediante el filtro de conteo mínimo (MC, *Minimum Count*) que consideró únicamente a los SNPs presentes en por lo menos 150 líneas, quedando 335,931 SNPs por cada una de las 272 líneas endogámicas, que representan el 35.15% de los SNPs obtenidos por GBS. Para conocer si hay estructura genética se realizó el análisis de componentes principales de las DTMA utilizando el paquete GAPIT del lenguaje R (Lipka *et al.*, 2012).

El análisis de asociación genómica se realizó con la plataforma TASSEL 5. 2. 4 (*Trait Analysis by Association, Evolution and Likange*) desarrollada por el laboratorio *Maize Genetics and Diversity* del Dr. Edward Buckler de la Universidad de Cornell, en los Estados Unidos (Tian *et al.*, 2010). El GWAS se realizó por el método de modelos lineales mixtos (*MLM, Mixed linear model*); esta prueba realiza correcciones tomando en cuenta la estructura poblacional y las relaciones de parentesco, para lo cual se realizaron análisis de parentesco (*kinship*) (Bradbury *et al.*, 2007; Zhang *et al.*, 2010) y de componentes principales (Cardon y Palmer 2003; Flint-Garcia *et al.*, 2003; Song *et al.*, 2015).

Para facilitar el trabajo computacional, los genotipos fueron transformados a números y se realizó una imputación numérica, que evaluó los cinco vecinos más cercanos. Para esto se consideró la distancia euclidiana para incorporar los datos faltantes (Hron *et al.*, 2010). El análisis de componentes principales fue de tres componentes por el método *scale-IBS*, como se muestra en el manual de usuarios de *Tassel 5 Source* [<https://bitbucket.org/tasseladmin/tassel-5-source/wiki/Home/>] (Bradbury *et al.*, 2007).

Los resultados del GWAS se analizaron con un gráfico probabilístico normal o *QQ Plot Normal*, que compara la distribución observada de los valores de p con una distribución normal o esperada para obtener el valor de significancia de los SNPs candidatos (Schweder y Spjøtvoll 1982; Ehret, 2010). Finalmente, se generó una gráfica tipo *Manhattan* para observar los valores p de los SNPs asociados a la elongación del mesocotilo.

Cálculo de los efectos fenotípicos de los SNP asociados a la elongación del mesocotilo en las líneas de maíz DTMA.

Se calcularon los efectos fenotípicos de los SNPs asociados para cada uno de sus estados alélicos. Para esto, se restó el promedio de las 272 líneas de la colección DTMA (que fue de 7.06 cm) al promedio de la elongación del mesocotilo de las líneas que presentaron los estados alélicos homocigotos mayores, homocigotos menores o heterocigotos de cada SNP asociado. Se realizaron correlaciones de Pearson entre los efectos fenotípicos y el número de accesiones que presentaron los estados alélicos homocigoto menor o heterocigoto, para observar si existía una correlación directa entre el número de individuos y el efecto fenotípico calculado.

Los SNP asociados fueron localizados en los cromosomas del maíz usando la base de datos del servidor *MaizeGDB* [<https://www.maizegdb.org/>] (Lawrence *et al.*, 2004). Primero se descargaron las secuencias alrededor de los SNP asociados consultando la versión RefGen_v2 del genoma de B73, usado como referencia para la genotipificación mediante GBS, con ayuda del programa Artemis (Rutherford *et al.*, 2000). Se extrajeron secuencias de 101 pares de bases, 50 pares de bases río arriba y 50 pares de bases río abajo, tomando como punto central al SNP asociado (**Anexo II**). Posteriormente, las secuencias se compararon con la versión RefGen_v3 del genoma de B73 (Schnable *et al.*, 2009) en la base de datos del *MaizeGDB* utilizando la herramienta *Blast* [*Basic Local Alignment Search Tool*] (Altschul *et al.*, 1990).

Otros análisis bioinformáticos realizados.

Los datos transcriptómicos de cada uno de los genes más cercanos a los SNP asociados a la elongación del mesocotilo en las líneas DTMA se obtuvieron consultando el Atlas de expresión de maíz (Sekhon *et al.*, 2011). Para conocer la función de los genes candidatos se obtuvieron los términos ontológicos, denominados *Go terms*, de la base de datos *Phytozome* [*The plant genomics resource*; <https://phytozome.jgi.doe.gov/pz/portal.html>] (Goodstein, 2012). Los *Go*

terms se obtienen mediante un análisis de enriquecimiento que predice la función de una lista de genes mediante alineamientos de las secuencias en árboles filogenéticos (Muruganujan *et al.*, 2013). Se encontraron las funciones moleculares de los genes asociados, los procesos biológicos en los que intervienen y su localización subcelular.

Para conocer si hay correspondencia genética entre las regiones asociadas a la elongación del mesocotilo encontradas en la colección DTMA, y los QTL descritos por Zhang y colaboradores (2012), se comparó la posición física de los QTL - que están referidos como Bins -, mapeados en el genoma de referencia B73 RefGen v2 [en el portal del MaizeGDB] (Andorf *et al.*, 2010) con la posición física de los SNP candidatos.

Análisis de QTL de la variación de la elongación del mesocotilo en las líneas ABQTL-DTP BC1F2:3.

Se realizó el análisis de QTLs de la variación de la elongación del mesocotilo en la población recombinante DTP utilizando datos genotípicos y fenotípicos que fueron obtenidos por métodos similares a los señalados para las DTMA. La población DTP corresponde a las frecuencias genotípicas de una población P₁BC₁F₂ con una contribución genotípica del parental 1 de 0.625, del parental 2 de 0.125 y una heterocigosis máxima esperada de 0.25. A los datos genotípicos se les retiraron los SNPs con heterocigosis mayor a 0.25, por ser mayor a la esperada (Meng *et al.*, 2015). Se aplicó el filtro de frecuencia para el alelo menor MAF de 0.05. Los SNPs fueron asociados y codificados de acuerdo a las dos líneas parentales. Estos SNPs fueron definidos como procedentes del parental 1, procedentes del parental 2, heterocigotos o como datos faltantes (Collard *et al.*, 2005; Meng *et al.*, 2015).

Las secuencias codificadas fueron utilizadas para la formación de mapas de ligamiento de alta densidad, mediante el procedimiento de ventanas corredizas (*sliding-window approach*). Con este método se busca reconocer los puntos de recombinación utilizando polimorfismos (Huang *et al.*, 2009). Para tener datos de alta calidad se retiraron los SNPs que tuvieran una heterocigosis mayor al 25%, más del 50% de datos faltantes, frecuencia del alelo menor < 10% y frecuencia del alélica menor > 25%. Se retiraron las líneas que presentaran una heterocigosis mayor al 25% y más del 45% de datos faltantes. Para retirar los SNPs no ligados o aislados poco informativos, se realizaron comparaciones entre SNPs con una ventana de 6 SNP vecinos, un umbral de similitud entre marcadores del 60% y un umbral de similitud entre SNP vecinos del 80%. Una vez retirados los SNP no ligados, se

formaron los bloques (*bins*). Para el claro reconocimiento de los eventos de recombinación se utilizó un umbral de similitud del 98% que por lo menos presentaran 3 SNPs por bloque, los SNPs que no cumplieron con estas características fueron retirados. Debido a que se detectó una gran cantidad de datos faltantes en los bloques, se realizó una imputación con los siguientes parámetros: en las regiones heterocigotas se imputaron las que presentaran menos del 50% de bloques faltantes y en las regiones homocigotas las que presentaran menos del 75%. Con los bloques obtenidos se realizó el mapa de recombinación y se usaron como marcadores (genotipos) para realizar los análisis de QTL (Huang *et al.*, 2009; Chen *et al.*, 2014; Liu *et al.*, 2015;).

Los mapas de ligamiento se realizaron con el software *QTL IciMapping* (Meng *et al.*, 2015). La función de mapeo utilizada fue la de Haldane, la cual no considera la interferencia entre los eventos de recombinación (Collard *et al.*, 2005). Debido a que se utilizaron SNPs previamente asociados a una posición física para formar los bloques, se mantuvo el orden y se agruparon de acuerdo al cromosoma de origen. Para graficar los mapas físicos se utilizó el software *Biomercator V4-2* (Sosnowski *et al.*, 2012).

Utilizando la información de cada mapa de ligamiento y la longitud promedio de los mesocotilos de las líneas DTP, se realizaron los análisis de QTL. El modelo utilizado fue *Simple Inclusive Interval Mapping* (ICIM) que requiere de dos pasos: en el primero se identifican las regiones significativas y en el segundo se realiza el mapeo por intervalos comparando los fenotipos con las regiones anteriormente identificadas (Meng *et al.*, 2015). Los bloques con valores de asociación mayores a 2 fueron considerados como *loci* fuertemente asociados. Estos fueron comparados con los SNP asociados obtenidos por GWAS con la colección de DTMA y con los QTL asociados a la resistencia a siembras a 10 y 20 cm de profundidad encontrados por Zhang y colaboradores (Zhang *et al.*, 2012).

RESULTADOS

Asociación genómica (GWAS) de la varianza en la elongación del mesocotilo en líneas DTMA. Los datos fenotípicos utilizados para realizar los estudios de asociación genómica [GWAS] fueron obtenidos por Nieto Sotelo y colaboradores (2012; 2013). El promedio de la longitud del mesocotilo, de la colección de 272 híbridos DTMA, fue de 7.06 cm, con una desviación estándar de 1.46 cm. La prueba de Kolmogorov-Smirnov mostró que el promedio del largo de mesocotilo para los híbridos DTMA se ajusta a una distribución normal con un valor de $P =$

0.1670 (Figura 4). El valor menor promedio de longitud dentro de la colección DTMA fue de 3.83 cm, en tanto que el máximo fue de 11.36 cm. Se realizó un análisis de ANOVA utilizando los valores extremos y los medios de la colección DTMA para observar las diferencias entre estos grupos. Para ésto se eligieron los 10 híbridos DTMA con los valores promedio de longitud de mesocotilo más largos, los 10 híbridos DTMA con valores intermedios y los 10 híbridos DTMA con valores más cortos. Se encontró que estos tres grupos son significativamente diferentes entre sí, con un valor de p de <0.0001 (Figura 5).

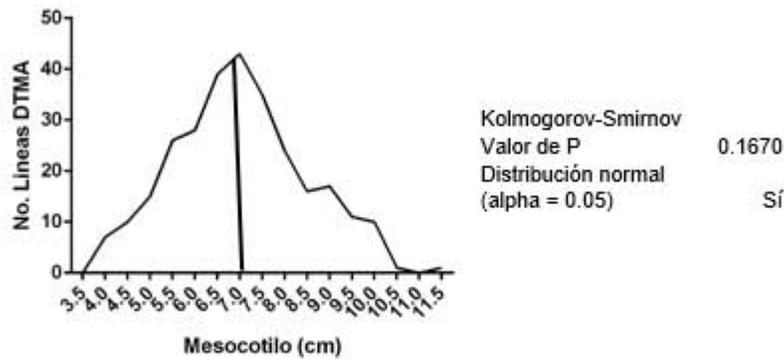


Figura 4. Distribución de los valores promedio del tamaño del mesocotilo de las 272 líneas DTMA y la cantidad de líneas en cada categoría. Datos obtenidos con los ensayos de respuesta a la siembra profunda en plántulas cultivadas en la oscuridad durante 7 días a 28°C. Datos tomados y regráficos de: (Nieto-Sotelo J. *et al.*, 2012; 2013).

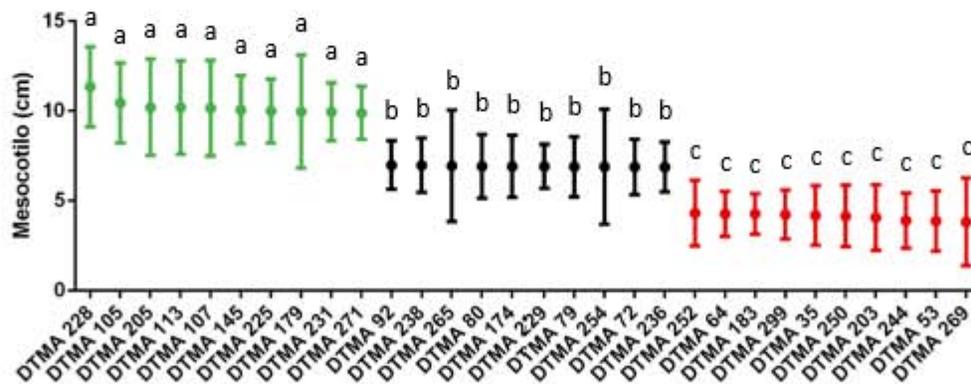


Figura 5. Longitudes promedio y desviaciones estándar de los mesocotilos en los 10 híbridos DTMA con los mesocotilos más largos (verde), en los 10 híbridos con los mesocotilos intermedios (negro) y en los 10 híbridos con los mesocotilos más cortos (rojo). Se realizó ANOVA y un análisis *post hoc* de Tukey. Los valores con letra diferente son estadísticamente diferentes.

La secuenciación de la colección de las líneas endogámicas DTMA mediante GBS rindió 955,690 SNPs por línea endogámica DTMA, distribuidos en los diez cromosomas del maíz.

Además, se incluyó el "cromosoma 0" que corresponde a las regiones no mapeadas en el genoma de referencia B73_v2.7 en donde se alojaron 570 SNPs. Mediante la aplicación de filtros de calidad se retiraron los SNPs con frecuencias alélicas menores a 0.05 y con datos faltantes mayores al 40%, por ser SNPs raros y poco informativos (Crossa *et al.*, 2013). Se mantuvieron los marcadores que estuvieran presentes en por lo menos 150 accesiones, quedando un total de 335,391 SNPs que representan el 35.15% del total de 955,690 SNPs (Cuadro 1). Se utilizó un filtro poco exigente [150/272] que permitiera analizar la mayor cantidad de SNPs y que retirara únicamente los datos con baja calidad o con muchos datos faltantes (Soto-Cerda y Cloutier, 2012).

En promedio, cada cromosoma mostró 33,589 SNPs. El cromosoma 1 presentó la mayor cantidad de SNPs (52,279), mientras que el cromosoma 10 presentó la menor cantidad (23,170 SNPs). En el "cromosoma 0" se ubicaron solo 37 SNPs de alta calidad (Cuadro 1). Los SNPs mostraron una frecuencia máxima de heterocigosidad del 0.2, distribuyéndose el 90% de los SNPs en el rango de heterocigosidad del 0 al 0.1. El rango de las frecuencias de los alelos homocigos menores de los SNPs osciló entre 0.05 y 0.5 (Figura 6).

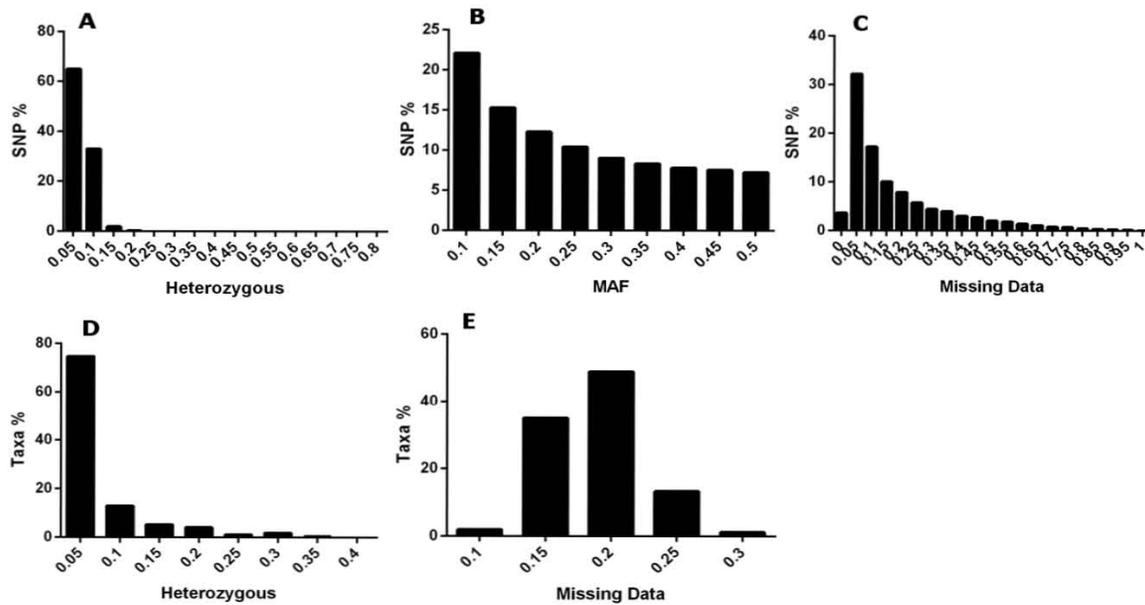


Figura 6. Análisis de los 955,690 SNPs identificados por GBS en la colección de 276 líneas endogámicas de maíz DTMA. A) porcentaje de SNPs con diferente índice de heterocigosidad, B) porcentaje de SNPs con diferentes frecuencias de los alelos menores, C) porcentaje de SNPs con diferentes frecuencias de datos faltantes, D) porcentaje de accesiones con distintos niveles de heterocigosidad y E) porcentaje de accesiones con diferentes frecuencias de datos faltantes.

Cuadro 1. SNPs obtenidos mediante el método de genotipificación por secuenciación masiva (GBS) de las líneas endogámicas DTMA y su filtrado. Las primeras dos columnas del lado izquierdo señalan el número total de SNPs por cromosoma; Filtro MAF 0.05) indica el número de SNPs remanentes después de aplicar el filtro de frecuencia mínima del alelo menor 0.05; Filtro CM 150) indica el número de SNPs remanentes después de aplicar a la columna de su costado izquierdo el filtro de conteo mínimo 150. El mapeo por GBS se hizo tomando como referencia la versión B73_V2 del genoma de maíz.

Cromosoma	SNPs	Filtro MAF 0.05	Filtro CM 150
0	570	76	37
1	148,752	57,212	52,279
2	115,173	46,325	41,802
3	108,224	42,451	38,588
4	94,726	34,216	31,132
5	110,328	43,664	39,821
6	76,475	29,729	27,123
7	80,517	31,218	28,373
8	81,431	31,199	28,332
9	72,368	27,919	25,274
10	67,126	25,815	23,170
Total	955,690	369,824	335,931
%	100	38.69	35.15

El análisis de componentes principales (PCA) mostró que la estructura poblacional en la colección DTMA es amplia, como anteriormente se había reportado para la misma colección (Gowda *et al.*, 2015). Para evaluar si la distribución fenotípica de los valores promedio de los mesocotilos de la colección DTMA respondía a la estructura poblacional, las accesiones fueron categorizadas en tres grandes grupos: las de promedios de mesocotilo largos [8.86 a 11.36 cm], las de promedios de mesocotilo medianos [6.35 cm a 8.85 cm] y las de promedios de mesocotilo cortos [3.83 cm a 6.34 cm]. Salvo una pequeña rama del PCA, los tres agrupamientos se distribuyeron dentro de todo el ámbito tridimensional del PCA (Figura 7), siendo la excepción fue el grupo de mesotilos cortos, ausentes en una rama de pocos elementos.

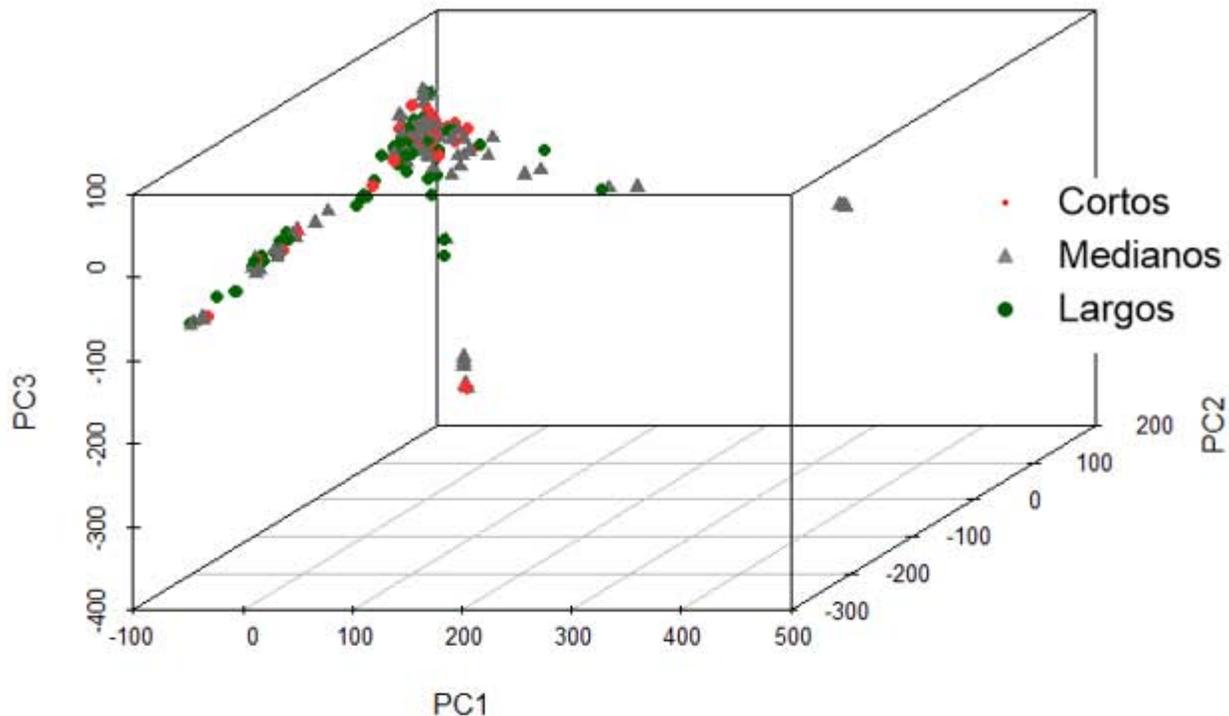


Figura 7. Análisis de estructura poblacional de la colección DTMA basado en componentes principales (PC), se muestran los tres primeros componentes principales. Círculos rojos accesiones con promedio de mesocotilo menores a 6.34 cm (mesocotilo corto), Triángulos grises accesiones con promedio de mesocotilo mayor a 6.34 cm y menores a 8.85 cm (mesocotilo mediano), círculos verdes accesiones con promedio de mesocotilo mayor a 8.85 cm (mesocotilo largo).

La asociación genómica se realizó por el método de Modelos Lineales Mixtos (Zhang *et al.*, 2010) para evitar las asociaciones espurias causadas por la estructura poblacional (Pritchard *et al.*, 2000; Zhou *et al.*, 2010; Gowda *et al.*, 2015). Se realizó la corrección de Bonferroni a una $\alpha = 0.05$ dividida entre el número de pruebas (Holm, 1978) [$0.05/(272 \text{ líneas} * 335931 \text{ SNPs})$] que dió como resultado un valor mínimo de significancia de $p = 2.01 \times 10^{-12}$ o $-\log_{10}$ de $p = 11.69$. No se encontraron SNPs con valores de $-\log_{10}$ de p que cumplieran con este requisito, por lo que no se tomó en cuenta la corrección de Bonferroni para determinar los valores de significancia, ya que es conocido que no siempre es la prueba adecuada para identificar a los SNPs fuertemente asociados para explicar la variación fenotípica (Perneger, 1998). Por otro lado se realizó el gráfico cuantil-cuantil (*quantile-quantile plot*, *QQ plot*) para evaluar la normalidad de los resultados del GWAS. Esta prueba indicó que los valores esperados de p [$-\log_{10}p$] muestran una tendencia lineal y que los valores de $-\log_{10}$ de p observados son menores que los valores de asociación esperados, desviándose de la

normalidad a partir del valor p observado de 1.5 (Figura 8). A partir de este valor mínimo se encontraron 90 SNPs asociados (Figura 9).

Debido a la alta densidad genómica de los SNPs usados en el GWAS, en algunos casos se encontraron varios SNPs fuertemente asociados alejados entre sí a menos de 500 pares de bases. Éstos marcadores fueron tratados como haplotipos y se tomó en cuenta el SNP con el mayor valor de $-\log_{10}p$. Se encontraron 63 SNPs fuertemente asociados. Los cromosomas 4, 5, y 7 presentaron valores p de asociación mayor a 4 (Cuadro 2). El valor p máximo de asociación se encontró en el cromosoma 7, que fue de 4.43, en tanto que el valor mínimo necesario para incluir a los diez cromosomas fue de 3.47 y se encontró en el cromosoma 1 (Figura 11).

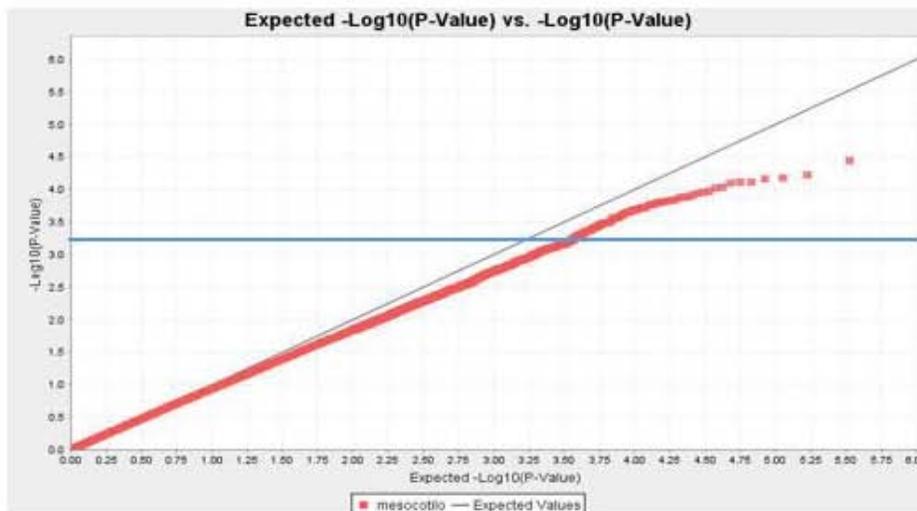


Figura 8. Gráfica cuantil-cuantil (QQ) que muestra los valores esperados de p y los valores observados de p para los SNPs usados en el GWAS. La línea azul indica el valor p mínimo designado para aceptar a los SNPs con los mayores niveles de asociación a la variación en la longitud del mesocotilo.

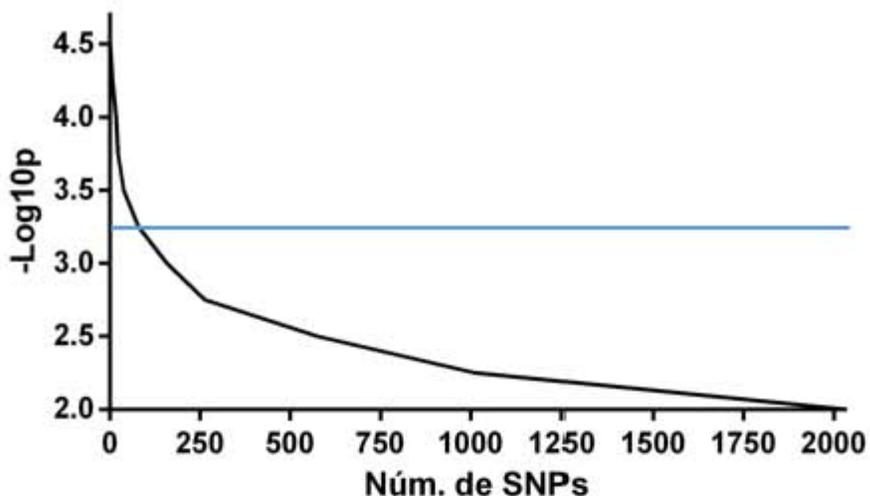


Figura 9. Número de SNPs y sus valores de asociación p ($-\log_{10}$). La línea azul indica el valor de menos \log_{10} p 3.25, que fué el umbral mínimo para considerar a los SNPs fuertemente asociados a la variación de la longitud del mesocotilo en la colección de maíz DTMA.

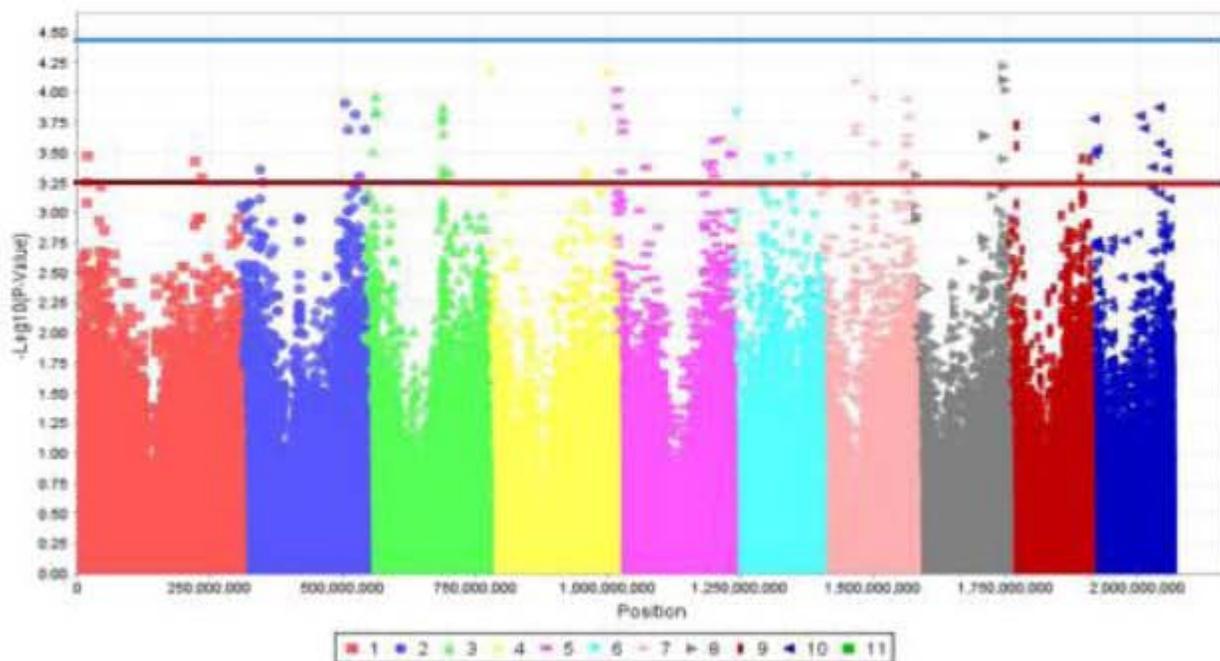


Figura 10. Gráfica tipo *Manhattan* de la asociación MLM para la elongación del mesocotilo en la colección de maíz DTMA. La línea azul señala el valor máximo asociado y la línea roja el valor mínimo aceptado.

Cuadro 2. Cantidad de SNP asociados a la elongación del mesocotilo que fueron deducidos mediante GWAS en cada cromosoma con sus valores p ($-\log_{10}p$) de asociación máximos y mínimos.

Cromosoma	Núm. SNP asociados	p máx.
1	4	3.47
2	7	3.91
3	6	3.96
4	4	4.16
5	10	4.02
6	5	3.83
7	9	4.43
8	6	4.22
9	4	3.72
10	8	3.8
Total	63	4.43

Efectos fenotípicos.

Para estimar el efecto de las combinaciones alélicas en la variación de la longitud del mesocotilo, se realizó el análisis de efectos fenotípicos para cada uno de los 63 SNPs asociados. No se consideraron haplotipos, ni los vecinos muy cercanos al SNP con mayor valor p cuyos valores de asociación fuesen mayores al umbral mínimo de 3.25, ya que la técnica GBS no identifica a todos los SNPs posibles del genoma del maíz, lo que obstaculiza identificar cabalmente a los SNPs causales y a sus haplotipos (Korte y Farlow, 2013; Lin *et al.*, 2014).

Se encontraron SNPs con efectos positivos y negativos, debido a que la asociación genómica reconoce marcadores que explican la varianza en el fenotipo, lo que permite diferenciar a los SNPs que estimulan de los que inhiben la elongación del mesocotilo (Figura 12). Los alelos con mayores efectos fenotípicos fueron tanto los heterocigotos como los homocigotos del alelo menor (Cuadro 3).

Un análisis de correlación entre el efecto fenotípico y el número de accesiones homocigotas menores mostró que a mayor número de accesiones hay un menor efecto fenotípico (**Figura Suplementaria 1**). Por lo anterior, los SNPs asociados cuyo número de accesiones en esas condiciones fuese menor a 25 que, de acuerdo a nuestro análisis, era indicativo de baja confiabilidad estadística. Los valores p de asociación y los efectos

fenotípicos de los estados homocigotos menores no correlacionaron, lo que indica que los valores p altos no conllevan necesariamente a un efecto fenotípico elevado (Figura 12).

Se obtuvieron 51 SNPs asociados, cuyos efectos fenotípicos como homocigotos menores fueron confiables, 27 presentaron efectos positivos en la elongación del mesocotilo. Los SNPs de efectos mayores fueron: S5_11923821, S4_226459824, S8_131049015, S7_160595701 y S2_217651275, con efectos fenotípicos entre 0.96 y 0.7 centímetros más altos que el promedio de las DTMA.

Se encontraron 23 SNPs, cuyos efectos fenotípicos como homocigotos menores fueron confiables que provocaron efectos negativos sobre la elongación del mesocotilo, siendo los de mayor efecto S8_167282998, S7_5093341, S2_38552471, S10_4751416 y S5_1852022, con valores entre -0.91 y -0.79 centímetros menores que el promedio de las DTMA. El SNP S6_10617654 fue neutral para el estado alélico menor lo que indica que el estado heterocigoto tiene el mayor efecto sobre el fenotipo; sin embargo, este solo estuvo presente en 12 accesiones (Cuadro 3).

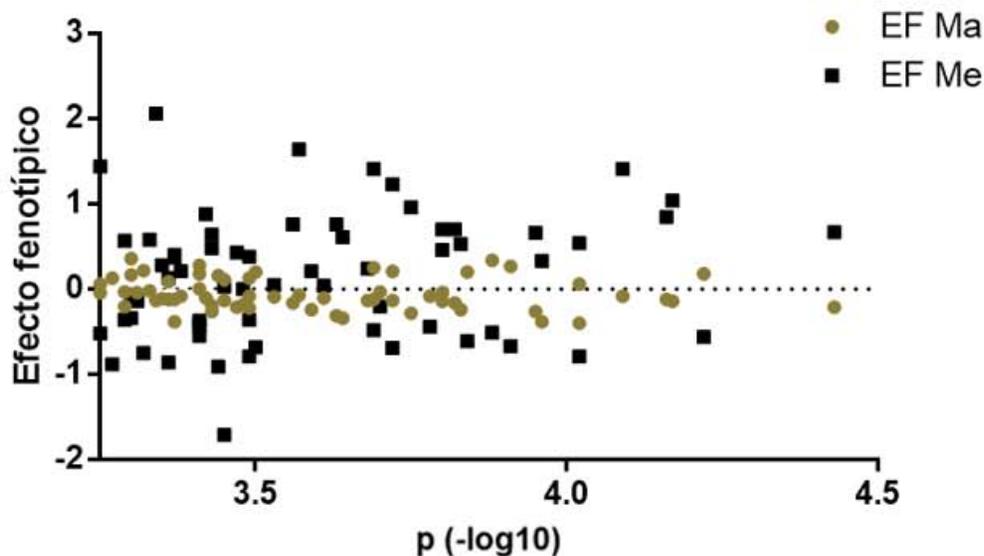


Figura 11. Valores de asociación p y efectos fenotípicos correspondientes en la elongación del mesocotilo de la colección DTMA tanto para los alelos mayores como para los alelos menores. Se presentan los datos de los 63 SNP asociados. EF Ma) efectos fenotípicos de los alelos mayores, EF Me) efectos fenotípicos de los alelos menores.

Cuadro 3 (inicio). Relación de los 63 SNPs asociados a la elongación del mesocotilo en maíces DTMA mediante GWAS. Clave) Cromosoma y posición física del SNPs asociado Cr) ubicación cromosomal; p) probabilidad [-log10] de asociación; Ma) número de accesiones que presentaron el estado alélico homocigoto mayor; EF Ma) efectos fenotípicos de los estados alélicos homocigotos mayores; Me) número de accesiones que presentaron el estado alélico homocigoto menor, EF Me) efectos fenotípicos de los estados alélicos homocigotos menores; He) número de accesiones que presentaron el estado heterocigoto; EF He) efectos fenotípicos de los estados heterocígos, en rojo se encuentran los SNPs que presentaron menos de 25 accesiones con el estado alélico menor.

Clave	Cr	p	Ma	EF Ma	Me	EF Me	He	EF He
S7_162027150	7	4.43	125	-0.21	55	0.67	7	-1.06
S8_166105182	8	4.22	192	0.18	47	-0.56	12	-1.05
S4_4737438	4	4.17	174	-0.14	20	1.04	8	1.62
S4_226459824	4	4.16	220	-0.12	37	0.85	8	-1.32
S7_63900887	7	4.09	229	-0.08	21	1.41	2	-0.41
S5_1852022	5	4.02	155	0.06	33	-0.79	2	0.98
S8_170878813	8	4.02	122	-0.4	92	0.54	21	-0.3
S3_19260465	3	3.96	128	-0.38	118	0.33	19	0.56
S7_99089950	7	3.95	196	-0.26	56	0.66	17	0.65
S2_198509433	2	3.91	134	0.27	60	-0.67	6	-0.86
S3_146363360	3	3.88	130	0.34	116	-0.51	22	0.55
S3_21368942	3	3.84	174	0.2	72	-0.61	20	0.3
S6_8146003	6	3.83	172	-0.24	69	0.53	21	-0.14
S2_217651275	2	3.82	188	-0.16	42	0.7	4	-1.7
S10_91959067	10	3.8	215	-0.14	34	0.46	4	2.67
S7_165369648	7	3.8	225	-0.04	14	0.7	0	0
S10_5458621	10	3.78	207	-0.08	33	-0.44	24	0.97
S5_11923821	5	3.75	124	-0.28	27	0.96	5	1.6
S9_15357629	9	3.72	152	0.21	63	-0.69	11	0.82
S7_64048287	7	3.72	221	-0.13	18	1.23	3	1.62
S10_98272538	10	3.7	232	-0.03	31	-0.2	6	2.34
S4_174903108	4	3.69	153	0.25	92	-0.48	18	0.68
S2_236282499	2	3.69	223	-0.12	21	1.41	7	0
S2_204476331	2	3.68	183	-0.13	45	0.24	10	-1.43
S8_131240915	8	3.64	124	-0.34	81	0.61	12	0.28
S8_131049015	8	3.63	172	-0.31	49	0.76	9	0.29
S5_193547933	5	3.61	196	-0.1	53	0.04	6	2.34
S5_183326603	5	3.59	114	-0.24	50	0.21	5	0.73
S10_127809396	10	3.57	242	-0.07	14	1.64	7	-1.15
S7_160595701	7	3.56	201	-0.16	26	0.76	5	2.34
S10_10586901	10	3.53	230	-0.09	12	0.05	5	2.86
S3_13726673	3	3.5	160	0.2	47	-0.68	7	0.37
S5_215289316	5	3.49	165	-0.22	84	0.38	21	0.1
S10_4751416	10	3.49	187	0.13	26	-0.79	7	-1.59

Cuadro 3 (término).

Clave	Cr	p	Ma	EF Ma	Me	EF Me	He	EF He
S10_140743010	10	3.49	143	-0.08	77	-0.36	16	1.36
S6_106176548	6	3.48	153	-0.18	89	0	12	1.53
S1_14103918	1	3.47	181	-0.21	66	0.43	23	0.16
S9_139410281	9	3.45	189	-0.13	63	0.03	14	1.63
S6_71803064	6	3.45	194	0.11	11	-1.71	16	0.53
S8_167282998	8	3.44	180	0.16	40	-0.91	14	0.45
S1_216533939	1	3.43	156	-0.19	43	0.64	8	-0.91
S9_153271004	9	3.43	175	-0.26	45	0.48	13	0.99
S5_182909034	5	3.42	241	-0.1	17	0.88	4	2.4
S5_170164202	5	3.41	135	0.18	111	-0.42	24	0.83
S6_72860333	6	3.41	93	0.28	93	-0.37	9	1.01
S7_157194565	7	3.41	242	0.005	16	-0.55	6	2.35
S10_115518234	10	3.38	184	-0.08	28	0.21	4	2.89
S5_56428968	5	3.37	121	-0.38	82	0.4	14	0.48
S7_153568928	7	3.37	216	-0.12	48	0.37	3	1.59
S2_38552471	2	3.36	206	0.09	37	-0.86	8	0.95
S5_178037831	5	3.36	191	-0.12	54	0.13	14	1.33
S4_183699508	4	3.35	172	-0.11	86	0.28	10	-0.9
S5_8266072	5	3.34	193	-0.13	8	2.06	7	0.3
S3_158354184	3	3.33	198	-0.02	29	0.58	4	0.19
S3_146026612	3	3.32	188	0.22	64	-0.75	20	0.32
S8_2763004	8	3.31	170	-0.04	56	-0.14	20	1.09
S2_225857203	2	3.3	127	0.17	29	-0.07	10	-1.98
S6_139454544	6	3.3	126	0.36	118	-0.34	24	-0.35
S1_229707717	1	3.29	202	-0.2	51	0.57	16	0.82
S9_135860028	9	3.29	209	-0.03	38	-0.36	11	1.76
S7_5093341	7	3.27	205	0.13	42	-0.88	22	0.46
S2_44378881	2	3.25	200	0.06	43	-0.52	6	1.76
S1_12854004	1	3.25	239	-0.043	13	1.44	11	-0.62

Mapeo cromosómico de los SNPs asociados.

Los SNPs asociados fueron mapeados físicamente en el genoma de referencia RefGen B73_v3 usando la base de datos del *Maize GDB*. De los 63 SNPs asociados, 44 se encontraron dentro del marco de lectura abierto de un gen, ya sea en genes que codifican para una proteína, en elementos transponibles y en genes de baja confianza. Los genes de baja confianza tienen un alto nivel de similitud con los genes homologos pero presentan diferencias significativas en la región codificante o en la longitud del gen y no hay evidencia de su expresión (Schnable, 2010). Se encontraron 19 SNPs en zonas intergénicas. La mayoría de

los SNPs asociados se encontraron dentro de un marco de lectura, este resultado puede atribuirse a que la secuenciación por GBS utiliza enzimas de restricción que son sensibles a metilación y por ende aumenta la probabilidad de secuenciar genes con marcos de lectura abiertos.

Para el análisis posterior de los SNPs intergénicos se consideraron a sus dos genes vecinos más cercanos [el primer gen río arriba y el primero río abajo]. Con este procedimiento se incluyeron a otros 12 genes como asociados a la elongación del mesocotilo.

En total se obtuvieron 82 genes candidatos (Cuadro 4). Dependiendo del caso, se anotó la función del gen homólogo más cercano, ya sea de arroz (*Oriza sativa*) o, en su defecto, *Arabidopsis thaliana*. Un total de 21 genes asociados codifican proteínas que no tienen función conocida en maíz o cuyos modelos de gen son poco estudiados. Pese a esto, 61 genes revelaron una función conocida o mostraron dominios conocidos que los ubican en familias de proteínas de función conocida.

Las principales funciones de las proteínas codificadas por los genes asociados fueron el transporte intracelular de lípidos, la regulación de la transcripción, la transducción de señales, el desarrollo vegetal, la edición de ARN, la unión a calmodulina, la inhibición de proteasas, el transporte de amonio, el transporte de metales pesados, la respuesta a las auxinas, la ubiquitinación, así como actividades enzimáticas como lipasas o fosfatasa.

Cuadro 4 (inicio). Listado de los genes candidatos y las funciones conocidas, deducidas de acuerdo a su pertenencia a alguna familia de proteínas (PFAM). Clave) Ubicación cromosomal y posición del SNP en número de bases, Pos) posición relativa del SNP, ya sea Intergénico (inter) o dentro del marco de lectura de un gen o modelo de gen (Marc), $-\log_{10}p$) valor de p en escala log10, Gen) modelo genético y PFAM) función conocida para la familia de las proteínas.

Clave	Pos	$-\log_{10}p$	Gen	Pfam
S7_162027150	Inter	4.43	AC192356.3_FG005	Gen de baja confianza
			GRMZM2G549820	Gen de baja confianza
S8_166105182	Marc	4.22	GRMZM2G335218	Transportador de amonio
S4_4737438	Marc	4.17	GRMZM5G845021	Adenilato cinasa
S4_226459824	Marc	4.16	GRMZM2G174673	Dominio asociado a metales pesados
S7_63900887	Inter	4.09	GRMZM2G060644	Elemento transponible
			GRMZM2G071808	Transposasa 24
S5_1852022	Marc	4.02	GRMZM2G138421	Proteína parecida a las que se une al promotor de <i>Squamosa</i>
S8_170878813	Marc	4.02	GRMZM2G069106	Dedos de Zinc con dominio Ran
S3_19260465	Marc	3.96	GRMZM2G044697	Peptidasa M50
S7_99089950	Marc	3.95	GRMZM2G002465	<i>Oryza sativa</i> : proteína expresada

Cuadro 4 (continuación).

Clave	Pos	-log10p	Gen	Pfam
S2_198509433	Inter	3.91	GRMZM2G037873	Elemento transponible
			GRMZM2G037875	MOSC: Sulfurasa cofactor de molibdeno
S3_146363360	Inter	3.88	GRMZM5G839699	Gen de baja confianza
			GRMZM5G813403	<i>Arabidopsis thaliana</i> : Superfamilia de las cupredoxinas
S3_21368942	Marc	3.84	GRMZM2G126749	Proteína parecida a las que se unen a la calmodulina
S6_8146003	Marc	3.83	GRMZM2G136700	<i>Oryza sativa</i> : proteína expresada
S2_217651275	Marc	3.82	GRMZM2G097277	Alfa-L-arabinofuranosidasa dominio C-terminal
S7_165369648	Inter	3.8	GRMZM2G111224	Factor Von Willebrand tipo A
			GRMZM2G549006	Elemento transponible
S10_91959067	Marc	3.8	GRMZM2G158141	<i>Oryza sativa</i> : dedos de zinc tipo C ₂ H ₂
S10_5458621	Marc	3.78	GRMZM2G044423	<i>Oryza sativa</i> : proteína expresada
S5_11923821	Inter	3.75	GRMZM2G57747	Gen de baja confianza
			GRMZM2G157166	<i>Oryza sativa</i> : proteína de unión al oxisterol supuestamente expresada
S7_64048287	Inter	3.72	GRMZM2G072000	Gen de baja confianza
			AC198118.4_FG001	Gen de baja confianza
S9_15357629	Marc	3.72	GRMZM2G328154	Gen de baja confianza
S10_98272538	Marc	3.7	GRMZM2G093022	Citidina y desoxicitidilato deaminasa, región de unión a zinc
S2_236282499	Marc	3.69	GRMZM2G402239	Duf1719
S4_174903108	Marc	3.69	GRMZM2G039263	HSP60, proteína de choque de calor
S2_204476331	Marc	3.68	GRMZM2G163233	Hélice-vuelta-hélice; V-Atpasa, sub-unidad H
S8_131240915	Inter	3.64	GRMZM2G001648	Codifica para una proteína desconocida
			GRMZM2G054584	Gen de baja confianza
S8_131049015	Inter	3.63	GRMZM2G117007	Dedos de zinc tipo C-x8-C-x5-C-x3-H
			GRMZM2G119527	Ferroportina 1
S5_193547933	Marc	3.61	GRMZM2G165530	MFS_1 proteínas facilitadoras principales
S5_183326603	Marc	3.59	GRMZM2G014750	Dominio La, motivo de unión al ARN, reconocimiento de motivos de ARN
S10_127809396	Marc	3.57	GRMZM2G066162	Glicosido hidrolasa de la familia 9
S7_160595701	Inter	3.56	GRMZM2G058087	TIP49 C-terminal
			GRMZM2G517577	Gen de baja confianza
S10_10586901	Marc	3.53	GRMZM2G026231	Proteína de endomembrana 70
S3_13726673	Inter	3.5	GRMZM2G048022	<i>Oryza sativa</i> : proteína expresada
			GRMZM2G047815	GPI-GlcNAc complejo transferasa, PIG-H componente
S5_215289316	Marc	3.49	GRMZM5G806449	Piridina disulfuro oxidoreductasa
S10_4751416	Inter	3.49	GRMZM2G031261	PSK, proteína precursora de fitosulfoquina
			GRMZM2G031164	Elemento transponible
S10_140743010	Marc	3.49	AC208110.2_FG007	<i>Oryza sativa</i> : Bn15D1B proteína específica de la semilla
S6_106176548	Marc	3.48	GRMZM2G071638	DUF212, familia divergente PAP2

Cuadro 4 (término).

Clave	Pos	-log10p	Gen	Pfam
S1_14103918	Marc	3.47	GRMZM5G850455	Tryp_alpha_amyl, proteína de transferencia de lípidos en plantas
S6_71803064	Inter	3.45	GRMZM2G164779	Elemento transponible
			GRMZM2G581506	Elemento transponible
S9_139410281	Inter	3.45	GRMZM2G165357	Epimerasa NADH dehidrogenasa
			GRMZM2G057467	DUF 568
S8_167282998	Marc	3.44	GRMZM2G177324	Peptidasa_C48, Ulp1 familia de proteasas, dominio catalítico C-terminal
S1_216533939	Inter	3.43	GRMZM2G124835	Gen de baja confianza
			GRMZM2G497425	Gen de baja confianza
S9_153271004	Inter	3.43	GRMZM2G360025	Proteína de unión a zeína
			GRMZM2G178106	Glicosil hidrolasa 35
S5_182909034	Inter	3.42	GRMZM2G546415	Gen de baja confianza
			GRMZM2G106680	<i>Oryza sativa</i> : proteína de unión a calcio en mitocondria
S5_170164202	Marc	3.41	GRMZM2G071484	Dominio caja U
S6_72860333	Marc	3.41	GRMZM2G152777	Lipasa_GDSL, acilhidrolasa
S7_157194565	Inter	3.41	GRMZM2G031323	Proteína de unión al ADN parecida a MYB
			GRMZM2G036291	Elemento transponible
S10_115518234	Marc	3.38	GRMZM2G179005	Histona H2A, H2B, H3 y H4
S5_56428968	Marc	3.37	GRMZM2G099191	ATPasa transportadora de Cadmio/Zinc
S7_153568928	Marc	3.37	GRMZM2G150791	<i>Arabidopsis thaliana</i> : repetición de ankirina
S2_38552471	Marc	3.36	GRMZM2G165844	Unión a AMP, NADH dehidrogenasa, catalasa
S5_178037831	Marc	3.36	GRMZM2G095782	tRNA sintasa de pseudouridina
S4_183699508	Marc	3.35	GRMZM2G088847	MIF, factor de inhibición de migración de macrófagos
S5_8266072	Inter	3.34	GRMZM2G548459	Gen de baja confianza
			GRMZM2G010801	<i>Arabidopsis thaliana</i> : metiltransferasa
S3_146026612	Marc	3.32	GRMZM2G047457	Proteína transportadora de la mitocondria
S8_2763004	Marc	3.31	GRMZM2G312419	MYb 60 factor de transcripción
S2_225857203	Marc	3.3	GRMZM2G433731	<i>Oryza sativa</i> : Repetición rica en leucinas
S6_139454544	Marc	3.3	GRMZM2G423863	Codifica para una proteína desconocida
S3_158354184	Marc	3.33	GRMZM2G139617	Dominio DEK C-terminal
S1_229707717	Marc	3.29	AC199893.4_FG002	Codifica para una proteína desconocida
S9_135860028	Marc	3.29	GRMZM2G500857	Gen de baja confianza
S7_5093341	Marc	3.27	GRMZM2G011800	<i>Arabidopsis thaliana</i> : miosina
S1_12854004	Marc	3.25	GRMZM2G136081	Parecida a PDDEXK de función desconocida
S2_44378881	Marc	3.25	GRMZM2G015844	<i>Oryza sativa</i> : receptor de esteroides nucleares

Para conocer los niveles de los transcritos de los genes candidatos durante el desarrollo del maíz, se consultó la base de datos *MaizeGDB* (Sekhon *et al.*, 2011). Desafortunadamente, en esta base de datos no se consideró al mesocotilo como una estructura independiente,

siendo analizado como parte de la “plántula” y de la “hoja”. Los genes candidatos en general tuvieron mayor expresión en la “hoja” y en la raíz (Sekhon *et al.*, 2011). También, se encontraron genes activos en la cariósida, las anteras, el coleoptilo y el embrión. En el **Anexo III** se muestran las estructuras de la planta de maíz donde los transcritos de cada gen asociado a la elongación del mesocotilo son más abundantes.

Para realizar una interpretación biológica de los genes candidatos, sus funciones se relacionaron con *GO terms*. Se encontraron *GO terms* para 37 de los 82 genes asociados, algunos de los genes candidatos mostraron más de un *GO term* asociado (Cuadro 5).

Cuadro 5 (inicio). Relación de los términos ontológicos (*Go terms*) de las funciones bioquímicas conocidas de los genes asociados a la elongación del mesocotilo en líneas DTMA. Gen) modelo de gen, *GO terms*) término de ontología genética, y Función) función molecular conocida.

Clave	Gen	GO. terms	Función
S8_166105182	GRMZM2G335218	GO:0008519 GO:0015696 GO:0016020 GO:0072488	Transporte de amonio entre membranas celulares
S4_4737438	GRMZM5G845021	GO:0000103 GO:0004020 GO:0005524	Catálisis de la reacción ATP+adenilato cinasa= ADP + 3 -fosfoadenosina 5´fosfosulfato
S4_226459824	GRMZM2G174673	GO:0030001 GO:0046872	Transporte de iones metálicos fuera y dentro de las células
S5_1852022	GRMZM2G138421	GO:0003677 GO:0005515 GO:0005634	Unión a ADN
S8_170878813	GRMZM2G069106	GO:0008270	Interactúa de forma no covalente con iones de zinc
S3_19260465	GRMZM2G044697	GO:0004222 GO:0006508	Proteólisis, cataliza la hidrólisis de los enlaces peptídicos en cadenas de polipéptidos
S2_198509433	GRMZM2G037875	GO:0003824 GO:0030151 GO:0030170	Interactúa de forma no covalente con molibdeno y vitamina B6
S3_21368942	GRMZM2G126749	GO:0005516 GO:0006950	Proteína que interactua con calmodulina y está involucrada en la respuesta a estres
S2_217651275	GRMZM2G097277	GO:0046373 GO:0046556	Catálisis de la hidrólisis de alfa-L-arabinofuranósido
S10_91959067	GRMZM2G158141	GO:0003950 GO:0046872	Catálisis de la reacción de: NAD+ADP-ribosiltransferasa
S10_98272538	GRMZM2G093022	GO:0008270	Unión no covalente a iones de zinc
S4_174903108	GRMZM2G039263	GO:0005524 GO:0006457 GO:0042026 GO:0051082	Chaperona, ayuda a recuperar la actividad de proteínas mal plegadas

Cuadro 5 (continuación).

Clave	Gen	GO. terms	Función
S2_204476331	GRMZM2G163233	GO:0000221 GO:0005488 GO:0015991 GO:0016820 GO:0046961	Transporte de protones en un gradiente electroquímico utilizando ATP, transporte de sustancias a través de la membrana mediante la catálisis de la hidrólisis del ácido anhídrido
S8_131049015	GRMZM2G117007	GO:0046872	Interactúa de forma no covalente con iones metálicos
S8_131049015	GRMZM2G119527	GO:0005381 GO:0016021 GO:0034755	Transporte de iones de hierro a través de la membrana celular
S5_193547933	GRMZM2G165530	GO:0005215 GO:0016021 GO:0055085	Transporte de sustancias (iones, macromoléculas y moléculas)
S5_183326603	GRMZM2G014750	GO:0000166 GO:0003723 GO:0006396 GO:0030529	Unión a ARN, participa en su maduración
S10_127809396	GRMZM2G066162	GO:0003824 GO:0004553 GO:0005975	Catálisis de la hidrólisis del enlace glucosídico
S7_160595701	GRMZM2G058087	GO:0003678 GO:0005524 GO:0043141	Función de helicasa
S10_10586901	GRMZM2G026231	GO:0016021	Componente de la membrana celular
S3_13726673	GRMZM2G047815	GO:0017176	Actividad de fosfatidil-inositol N-acetylglucosaminil-transferasa
S5_215289316	GRMZM5G806449	GO:0004362 GO:0006749 GO:0016491 GO:0050660 GO:0050661	Proteína involucrada en la catálisis de glutatión y en la reacción de óxido-reducción de FADH
S10_4751416	GRMZM2G031261	GO:0005576 GO:0008083 GO:0008283	Factor de crecimiento celular y proliferación celular
S1_14103918	GRMZM5G850455	GO:0006869 GO:0008289	Función de transporte de lípidos dentro y entre células
S9_139410281	GRMZM2G165357	GO:0003854 GO:0006694 GO:0016616 GO:0055114 GO:0055114	Catálisis de la reacción de síntesis de esteroides
S8_167282998	GRMZM2G177324	GO:0006508 GO:0008234	Hidrólisis de proteínas
S9_153271004	GRMZM2G178106	GO:0004553 GO:0005975	Hidrólisis del enlace glucosídico
S5_182909034	GRMZM2G106680	GO:0005509	Unión a iones de calcio
S5_170164202	GRMZM2G071484	GO:0004842 GO:0005488 GO:0016567	Catálisis de la transferencia de ubiquitina de una proteína a otra
S6_72860333	GRMZM2G152777	GO:0016788	Hidrólisis de los enlaces éster
S7_157194565	GRMZM2G031323	GO:0003677	Unión a ADN
S10_115518234	GRMZM2G179005	GO:0000786 GO:0003677 GO:0046982	Formación de heterodímeros

Cuadro 5 (término).

Clave	Gen	GO. terms	Función
S5_56428968	GRMZM2G099191	GO:0008270 GO:0030001	Transporte de iones metálicos fuera y dentro de las células
S2_38552471	GRMZM2G165844	GO:0006979	Respuesta a estrés oxidativo,
S5_178037831	GRMZM2G095782	GO:0001522	Síntesis de pseudouridina
S5_8266072	GRMZM2G010801	GO:0008168	Metilación de ARN
S8_2763004	GRMZM2G312419	GO:0003677	Unión a ADN
S3_158354184	GRMZM2G139617	GO:0003677	Unión a ADN

Para evaluar si los genes asociados a la variación de la elongación del mesocotilo presentaban un patrón funcional, se formaron 10 grupos de acuerdo a su función general. La función que tuvo más representantes fue la catálisis enzimática, seguida por el transporte de iones metálicos y la hidrólisis de carbohidratos y proteínas (Figura 12),

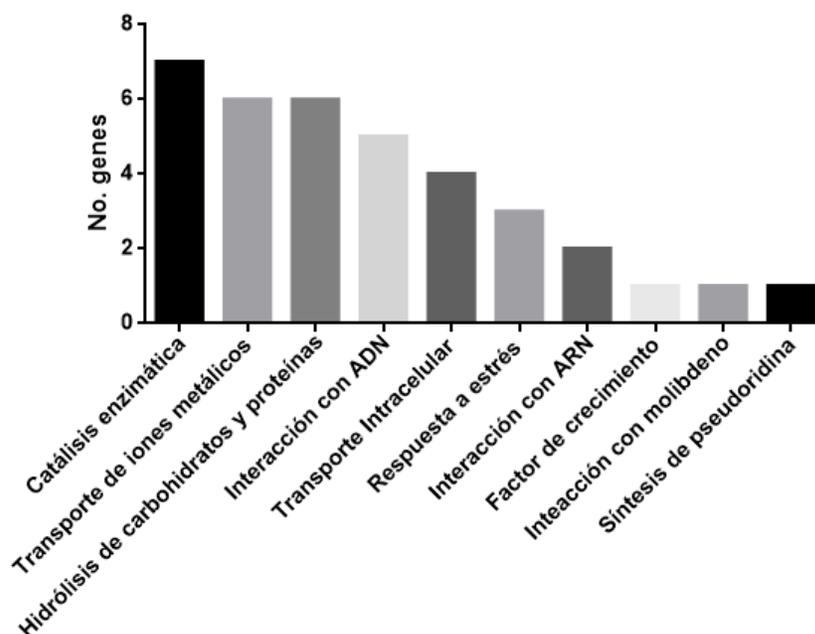


Figura 12. Categorías de términos ontológicos (*Go terms*) de las funciones generales de los productos de los genes asociados a la variación de la elongación del mesocotilo. Se señala el número de genes asociados en cada categoría.

Loci asociados a la variación de la elongación del mesocotilo mediante QTLs en la población ABQTL- DTP-BC₁F_{2:3}.

El análisis de ligamiento es una herramienta que nos permite conocer la segregación de caracteres y los alelos involucrados en la variación fenotípica (QTL). Este método nos permite evaluar si una variación alélica se encuentra dentro de una población, es heredada y si se encuentra involucrada en la regulación del mismo carácter. Además de ser una prueba experimental, es un método sin sesgos, que permite comparar las regiones asociadas de los resultados obtenidos por el GWAS en la colección de las DTMA.

Se evaluó el crecimiento de todas las estructuras de un número total de 5,491 plántulas (el mesocotilo, la primera hoja plumular, la raíz primaria y el coleoptilo) en la población DTP de maíz, constituida por 173 accesiones. El promedio de la longitud del mesocotilo en la población fue de 10.49 cm. Este valor promedio es 3 centímetros más largo que el promedio de la colección DTMA. La accesión con el promedio máximo fue DTP148 con 14.18 cm. El promedio mínimo fue de 6.30 cm en la accesión DTP144 (Cuadro 6).

Las líneas DTP mostraron variación en la longitud del mesocotilo entre y dentro de las accesiones (Figura 13). Se realizaron análisis de correlación de Pearson entre el mesocotilo, el coleoptilo, la raíz primaria, el tallo y la primera hoja plumular. Todas las correlaciones fueron significativas (Figura 14). Las correlaciones que tuvieron el valor de r^2 más alto fueron la primera hoja plumular y el tallo; el tallo y el mesocotilo; la primera hoja plumular y el mesocotilo.

Cuadro 6. Las diez líneas DTP con los mayores o menores valores promedio de longitud del mesocotilo, Nombre) nombre de la accesión n) número de plántulas medidas, d.s.) desviación estándar, Máx) valor máximo de mesocotilo y Mín) valor mínimo de mesocotilo.

Nombre	n	Promedio (cm)	d.s. (cm)	Máx (cm)	Mín (cm)
DTP148	43	14.198	2.093	18.4	4.5
DTP162	31	13.894	2.372	17.2	4.7
DTP154	18	13.478	2.142	15.5	6.5
DTP167	36	13.342	1.958	16.4	8.4
DTP105	14	13.336	3.619	15.6	1.1
DTP14	20	7.23	2.529	11.8	3
DTP106	13	7.1	1.718	10.5	4.7
DTP90	15	7.047	3.542	13.1	1.3
DTP26	21	6.857	2.248	10.5	1.8
DTP144	33	6.306	2.47	10.6	2.2

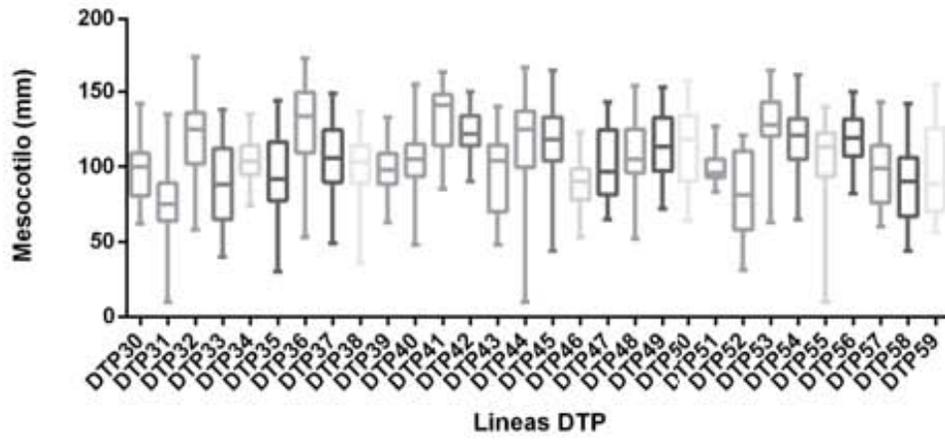
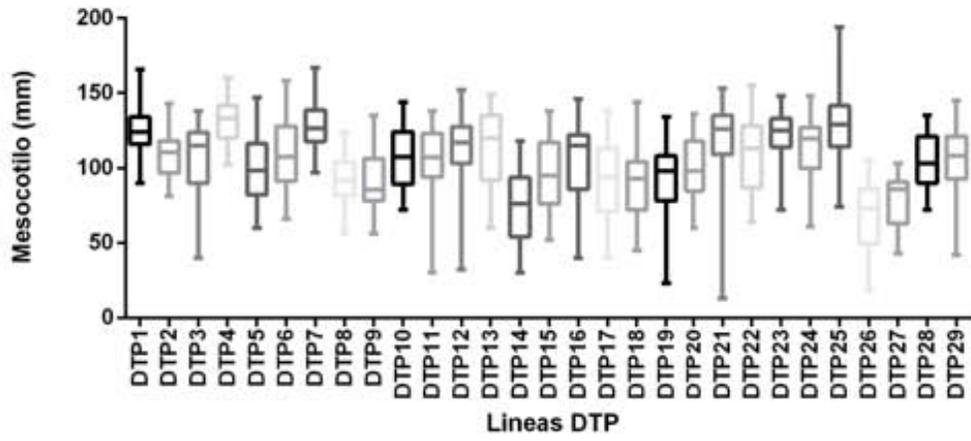


Figura 13 (inicio)

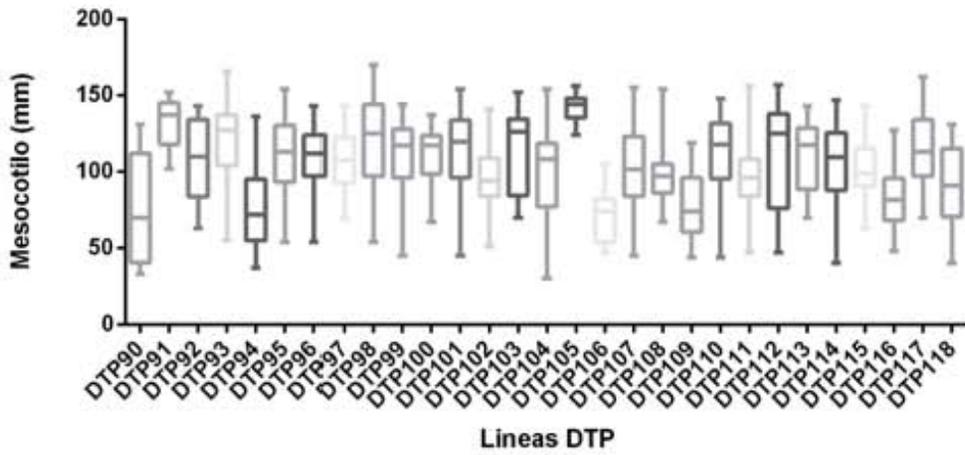
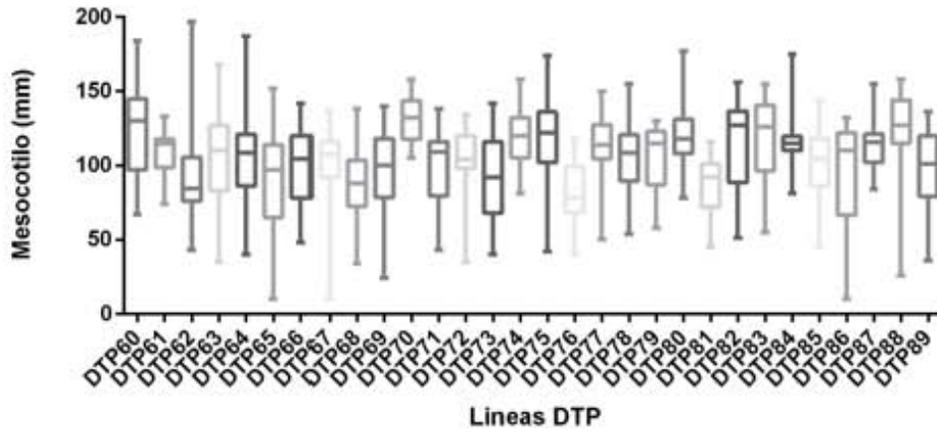


Figura 13 (continuación)

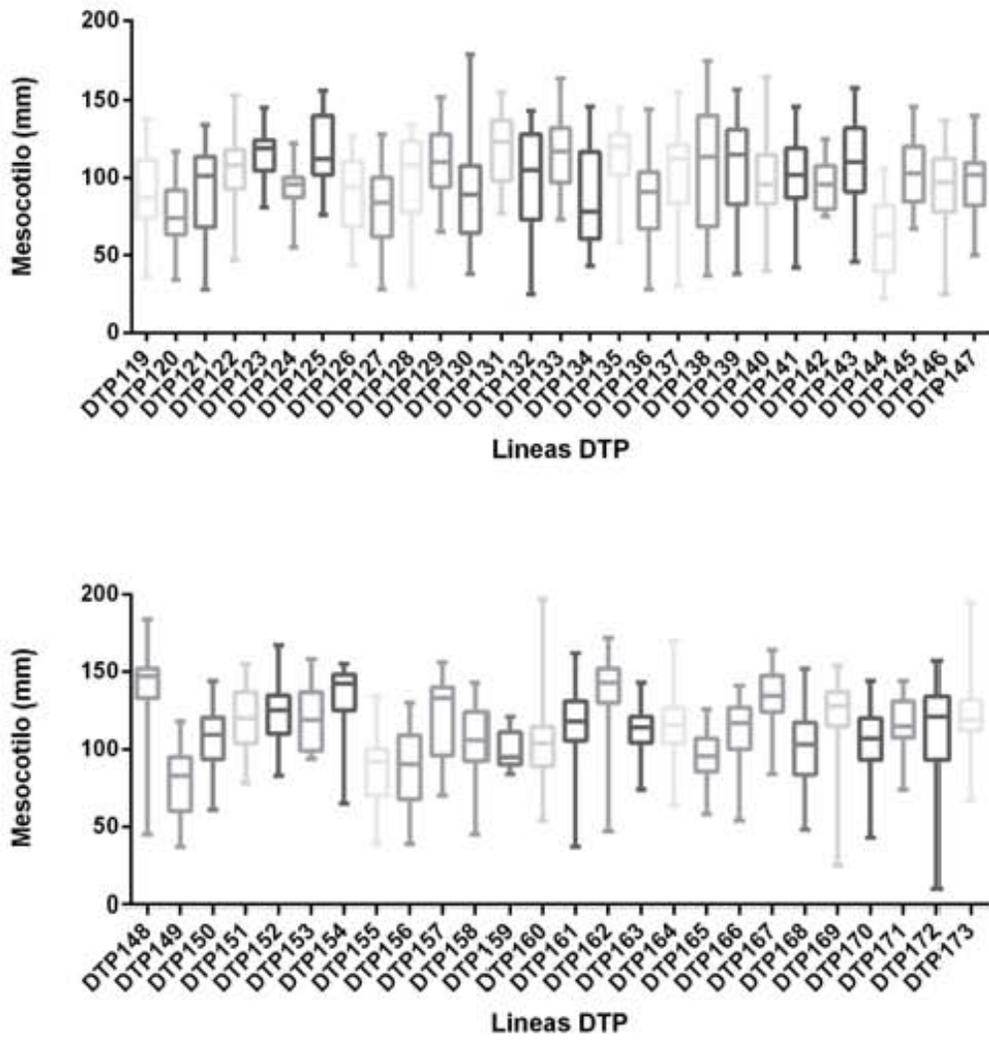


Figura 13 (término). Promedios y desviaciones estándar de los mesocotilos de las 173 accesiones de la población DTP.

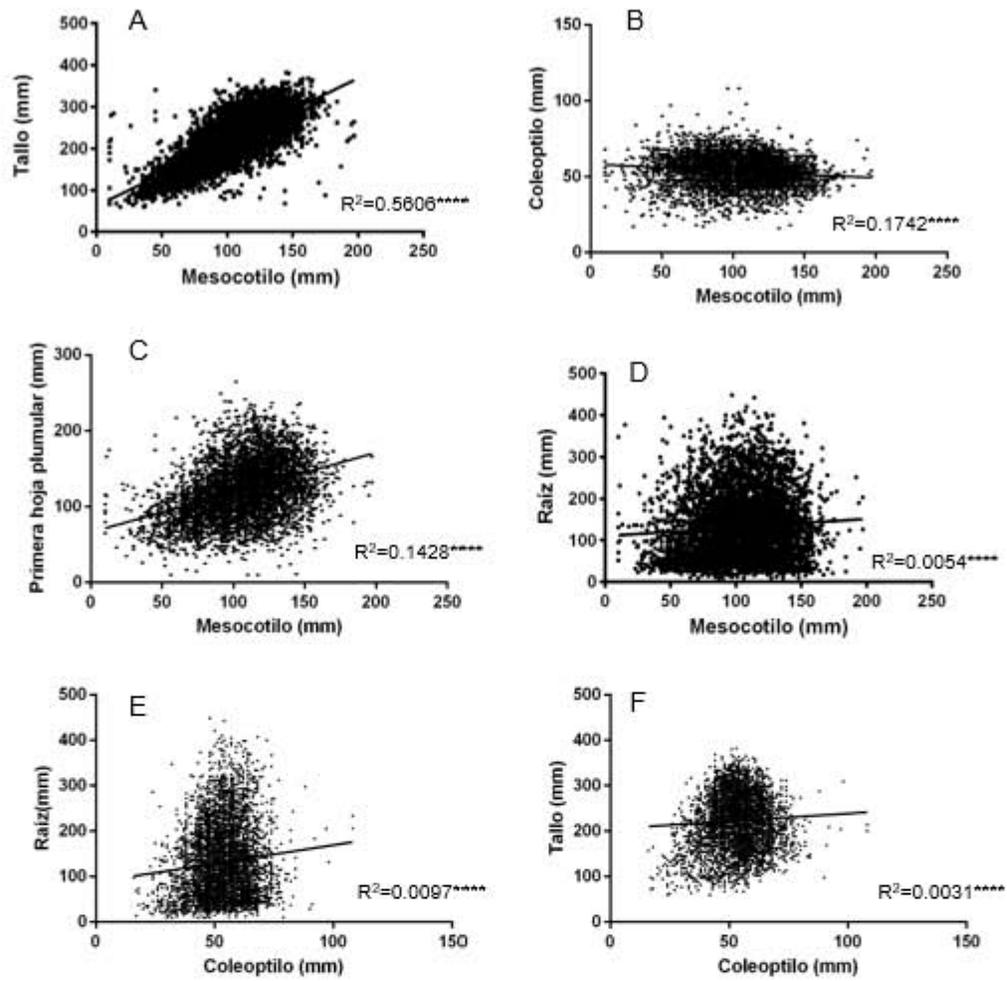


Figura 14 (inicio).

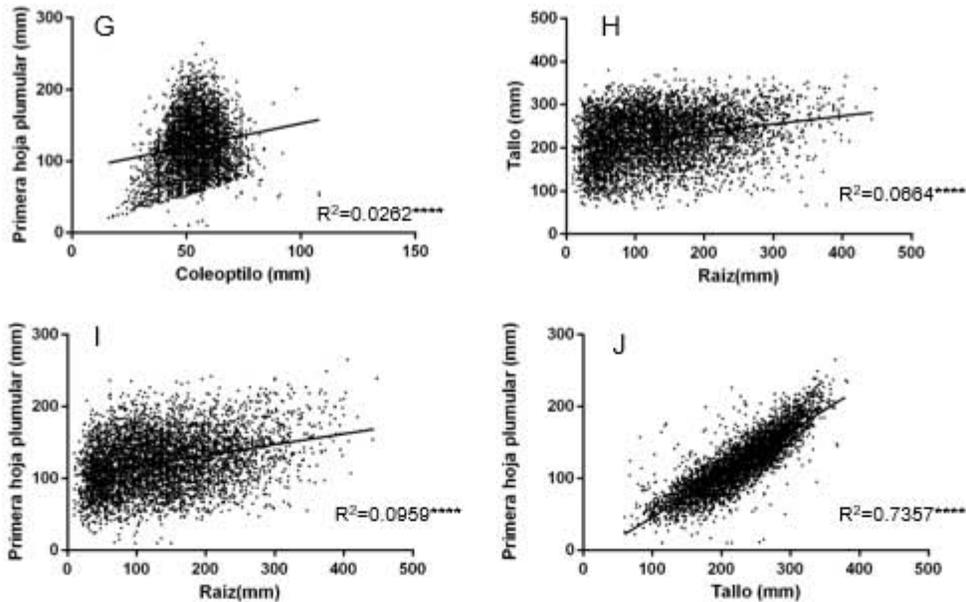


Figura 14 (término). Correlaciones de Pearson entre pares de estructuras medidas para las líneas DTP, A) tallo y mesocotilo, B) coleoptilo mesocotilo, C) primera hoja plumular y mesocotilo, D) raíz primaria y mesocotilo, E) raíz primaria y coleoptilo, F) tallo y coleoptilo, G) primera hoja plumular y coleoptilo, H) tallo y raíz primaria, I) primera hoja plumular, J) primera hoja plumular y tallo, ****) $P < 0.0001$.

Antes de iniciar el análisis de ligamiento de los SNPs asociados a la variación de la elongación del mesocotilo, se retiraron los SNPs de la población DTP que no cumplieran criterios mínimos de calidad. Éstos parámetros fueron determinados al realizar el análisis general de las secuencias de las líneas DTP. Se eliminaron los SNP con frecuencia alélica menor a 0.05, quedando 100,021 SNPs. Si bien el 70% de los SNPs presentó una proporción de heterocigosis menor al 0.2, el valor máximo de heterocigosis encontrado fue de 0.45 (Cuadro 7). Las frecuencias alélicas menores (MAF) se distribuyeron entre 0.2 y 0.35. Más del 60% de los SNPs contenían datos faltantes. La heterocigocidad de las líneas presentó un máximo de 0.25. Como algunas líneas presentaron más del 90% de datos faltantes, fueron retiradas de los siguientes análisis (Figura 15).

Para conocer si es posible observar la segregación de los SNPs encontrados en las DTMA asociados a la variación de la elongación del mesocotilo, se evaluó la presencia y las frecuencias alélicas de éstos en los datos genotípicos de la población DTP. De los 63 SNP asociados en la colección DTMA, sólo 18 SNPs se detectaron como tales en la población DTP (**Anexo IV**). Por lo tanto, éste análisis de ligamiento permitió evaluar solo una pequeña proporción de los SNP asociados a la variación en la elongación del mesocotilo.

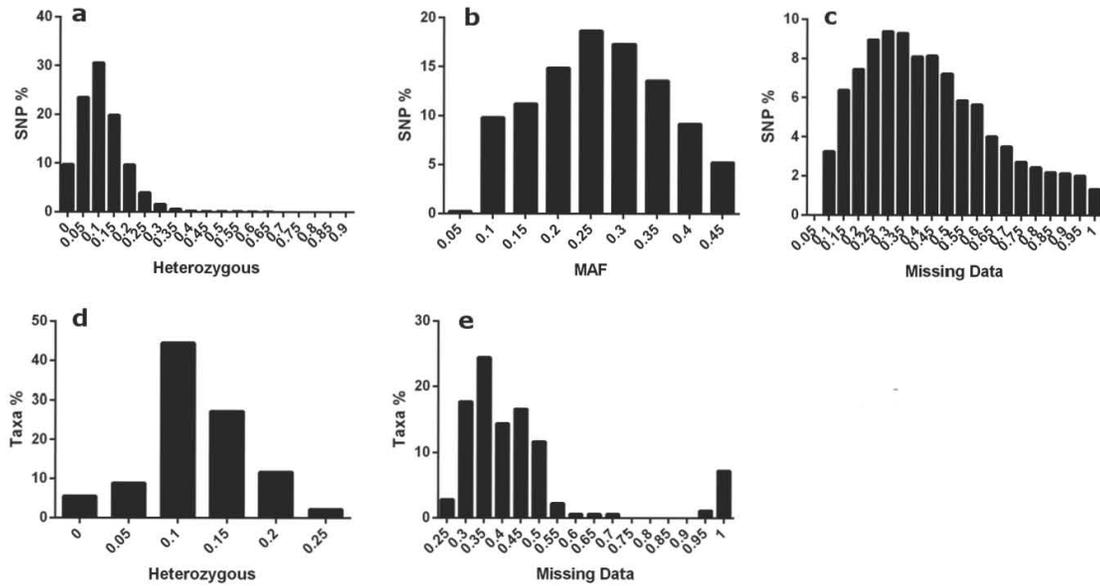


Figura 15. Características de los SNP obtenidos por GBS y de las líneas DTP, a) porcentaje de SNPs en heterocigosis, b) porcentaje de los SNPs como alelos de frecuencia mínima, c) porcentaje de SNPs y su proporción de datos faltantes, d) porcentaje de líneas heterocigotas, e) porcentaje de líneas y su proporción de datos faltantes.

Cuadro 7. Filtros aplicados a los SNPs de las líneas DTP.

Parámetros:	Valor
Número inicial de SNPs	100,021
Número inicial de líneas	180
Filtros aplicado a los SNPs:	
Heterocigosis máxima permitida	0.25
Máximo de datos faltantes	0.50
Mínimo de MAF	0.10
Máximo de MAF	0.25
Filtros aplicado a las líneas:	
Heterocigosis máxima permitida	0.25
Máximo de datos faltantes	0.45
Resultados:	
Número final de SNPs	30,598
Número final de líneas	160

En este análisis, el cromosoma 5 tuvo el mayor número de SNPs de alta calidad (4,659 SNPs), mientras que el cromosoma 9 mostró el menor número de SNPs de alta calidad (sólo 703 SNPs). Mediante el procedimiento de ventanas deslizantes se generaron 2,229 bloques (Cuadro 10). Con los bloques obtenidos se generó el mapa de recombinación con base en la contribución genética de cada parental en la población DTP para cada accesión (Figura 16). Hay regiones donde solo se observa la contribución de un solo parental. Hubo mayor contribución genética del parental 1 (CML 491) que del parental 2 (DTMA 227), como era lo esperado para una población BC₁F₂. El mapa de recombinación se utilizó como genotipo para la construcción del mapa de ligamiento (Figura 17), presentando una longitud total de 14,471.61 centimorgans (Cuadro 8).

Cuadro 8. Construcción de los bloques usados por cromosoma en los análisis de QTL para la población DTP y sus características. SNPs) SNPs usados con frecuencias mayores a 0.05 del alelo menor para formar bloques en cada cromosoma; Filtrado) número de SNPs que pasaron los filtros de calidad; SNPs ligados) el número de SNPs ligados Bloques) el número de bloques obtenidos; Longitud) la longitud total de los bloques de cada cromosoma en centimorgans.

Cromosoma	SNPs	Filtrado	SNPs ligados	Bloques	Longitud (cM)
1	14,961	4,391	4,375	333	2,069.02
2	11,944	3,246	3,244	246	1,869.63
3	13,512	2,966	2,892	197	1,866.19
4	9,436	3,938	3,931	278	1,742.25
5	11,711	4,659	4,656	355	1,699.3
6	7,449	1,412	1,409	108	589.88
7	8,672	3,237	3,211	229	1,589.82
8	8,590	3,012	2,977	193	1,530.06
9	7,291	703	701	55	342.65
10	6,455	3,034	3,033	236	1,172.81
Total	100,021	30,598	30,428	2,229	14,471.61

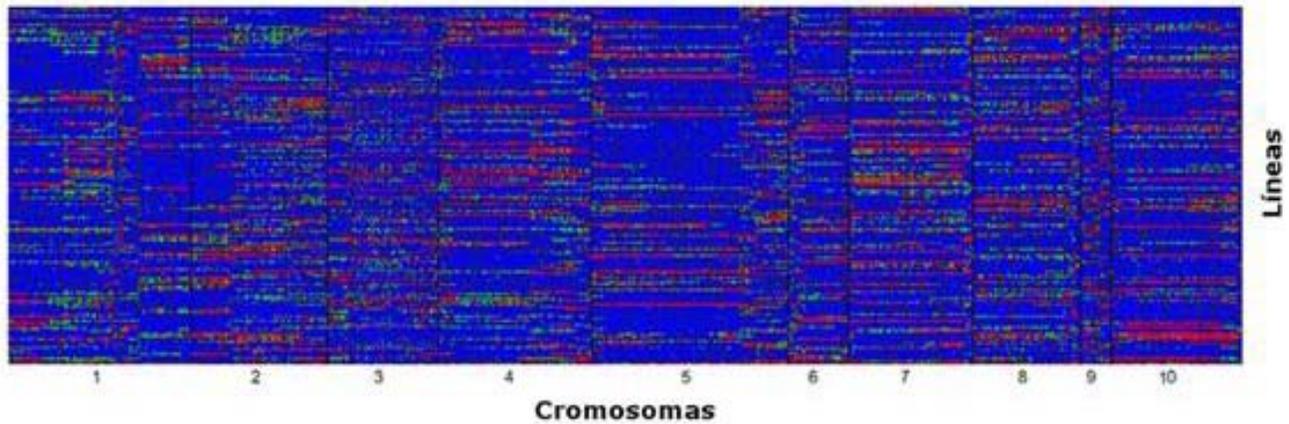


Figura 16. Mapa de los eventos de recombinación de la población DTP. La línea parental 1 (CML491) está señalada en azul y aporta el 0.625 de la variación genética, la línea parental 2 (DTMA 227) está señalada en rojo y aporta el 0.125 de la variación genética, los sitios heterocigotos señalados en verde aportan el 0.25 de la variación genética. Los sitios faltantes se indican en gris.



Figura 17. Mapas de ligamiento de los SNPs para los 10 cromosomas de maíz de la población DTP. Las líneas indican la posición física de los bloques definida por los SNP. Como referencia del inicio se utilizó 0 y como referencia del término del cromosoma se utilizó el número total de pares de bases del genoma de B73_v2.7. Las regiones en negro denotan la ubicación de una gran concentración de bloques, ya que la escala y la alta concentración de marcadores dificulta su observación. Las posiciones del mapa señalan el punto central de cada bloque.

Con el mapa de ligamiento, se realizó el análisis de QTLs utilizando el modelo ICIM. Los bloques que presentaron valores de LOD mayores a 2 se consideraron significativos para la elongación del mesocotilo (Figura 18). Se encontraron 29 QTLs asociados a la elongación del mesocotilo que fueron denominados DTPMES-c-n [donde c indica el número de cromosoma donde se encontró el QTL y n el número de QTL de ese cromosoma]. El valor máximo de LOD fue de 4.44 para el QTL DTPMES-2-7 (Cuadro 9). El SNP S9_135860028 se encontró dentro del QTL DTPMES-9-1 con un valor LOD de 2.02; éste SNP se encuentra dentro del marco de lectura del gen GRMZM2G500857 que es un modelo de baja confianza. El SNPs S10_10586901 se ubicó dentro del QTL DTPMES-10-6 con un valor LOD de 2.17; está asociado al gen GRMZM2G026231 que codifica una proteína endomembranal de 70 kDa.

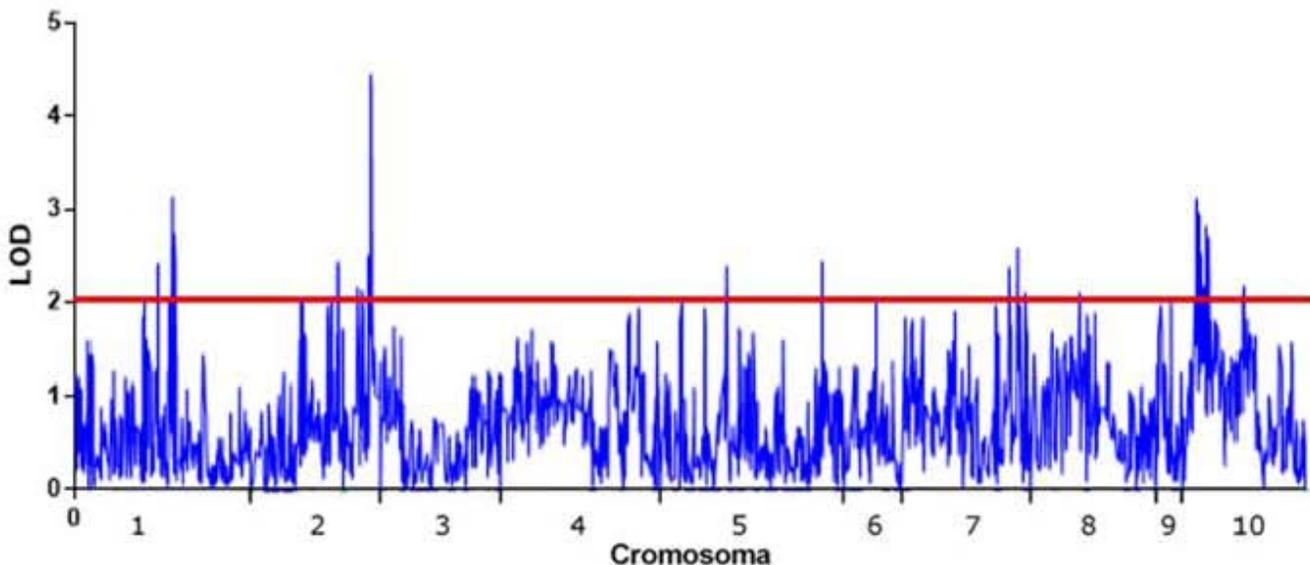


Figura 18. Distribución de QTLs y sus valores de asociación LOD para la elongación del mesocotilo en los 10 cromosomas de la población de maíces DTP. Se tomó como referencia el punto medio de las regiones genéticas asociadas. La línea roja marca el valor mínimo aceptado como QTL.

Cuadro 9. QTL asociados a la elongación del mesocotilo para la población DTP. QTL) Nombre, número de cromosoma donde se encuentra y número de QTL en ése cromosoma; cM) posición en centimorgan del QTL; Pos inicio) posición física donde inicia el QTL; Pos término) posición física donde termina el QTL; LOD) valor de asociación del QTL; PVE) variación fenotípica explicada y GWAS) nombre los SNP asociados a la elongación del mesocotilo dentro del QTL. En negritas se resalta el valor más alto de asociación.

QTL	cM	Pos inicio	Pos término	LOD	PVE	GWAS
DTPMES-1-1	1098	31765462	31848706.5	2.04	5.36	
DTPMES-1-2	1323	38850990.5	39141100	2.42	6.48	
DTPMES-1-3	1501	51169515	60149829	2.02	6.80	
DTPMES-1-4	1560	96239938	115058996	3.13	9.65	
DTPMES-1-5	1586	177627132	209306234	2.73	9.52	
DTPMES-2-1	3681	33466342	36822734.5	2.01	5.41	
DTPMES-2-2	4159	173947459	174339821	2.05	7.05	
DTPMES-2-3	4256	184746288	185691731	2.44	6.36	
DTPMES-2-4	4577	222814950	223975790	2.16	7.86	
DTPMES-2-5	4654	226197906	228804456	2.13	7.24	
DTPMES-2-6	4757	232367336	232381215	2.06	5.82	
DTPMES-2-7	4795	232877129	233677685	4.45	12.07	
DTPMES-5-1	9844	11591636	11594203.5	2.00	5.57	
DTPMES-5-2	10572	38577815.5	39820777.5	2.39	6.11	
DTPMES-5-3	12128	211204012	211301570	2.44	6.55	
DTPMES-7-1	15152	161347567	161685546	2.37	8.33	
DTPMES-7-2	15288	163790534	163935907	2.58	8.85	
DTPMES-7-3	15412	170251445	170496483	2.10	7.59	
DTPMES-8-1	16301	91861481.5	94710507.5	2.11	5.49	
DTPMES-9-1	17793	135834877	136019935	2.02	5.31	S9_135860028
DTPMES-10-1	18183	5725138	5759594	2.02	5.42	
DTPMES-10-2	18208	6249604.5	7585277	3.11	8.82	
DTPMES-10-3	18254	8093592	9399945.5	2.95	7.63	
DTPMES-10-4	18279	9532310.5	9594339.5	2.52	7.22	
DTPMES-10-5	18287	9598121.5	10245519.5	2.21	5.90	
DTPMES-10-6	18330	10474918	10621192.5	2.17	5.80	S10_10586901
DTPMES-10-7	18362	11315417.5	11602583.5	2.82	10.24	
DTPMES-10-8	18404	13335108.5	14522410	2.69	7.88	
DTPMES-10-9	18954	87132368	89420444	2.17	5.71	

Los SNPs asociados a la elongación del mesocotilo en la colección DTMA que coincidieron con los QTL reportados anteriormente por Zhang y colaboradores (2012) fueron: S1_14103918, S1_14103918, S4_226459824, S6_139454544, S6_8146003, S6_71803064, S6_72860333, S7_160595701 y S7_162027150 (Cuadro 10). Ninguno de estos coincidió con los QTL encontrados en la población DTP (Figura 20).

Cuadro 10. Coincidencias entre los QTLs para la elongación del mesocotilo propuestos por Zhang *et al.* (2012) y los SNPs asociados a la elongación del mesocotilo obtenidos por GWAS en líneas DTMA. QTL) nombre del QTL identificado por Zhang y cols. (2012); Cr) cromosoma; In) posición inicial del QTL de Zhang y cols. (2012); Fin) posición final del QTL de Zhang y cols. (2012); LOD) valor de asociación de los QTL de Zhang y cols. (2012); Núm) número de SNPs definidos por GWAS en líneas DTMA dentro de los QTLs; SNP) nombre de los SNPs y; p) valor de asociación.

QTL	Cr	In	Fin	LOD	SNP	p
qMES10-1	1	14078343	27918358	3.92	S1_14103918	3.47
qMES20-1	1	14078343	27918358	4.75	S1_14103918	3.47
qMES20-4	4	199459775	231903992	4.06	S4_226459824	4.16
qMES10-6-2	6	109488480	151640124	5.28	S6_139454544	3.30
qMES20-6-1	6	1	89145137	7.92	S6_8146003	3.83
					S6_71803064	3.45
					S6_72860333	3.41
qMES20-7	7	160198093	163184932	3.77	S7_160595701	3.56
					S7_162027150	4.43

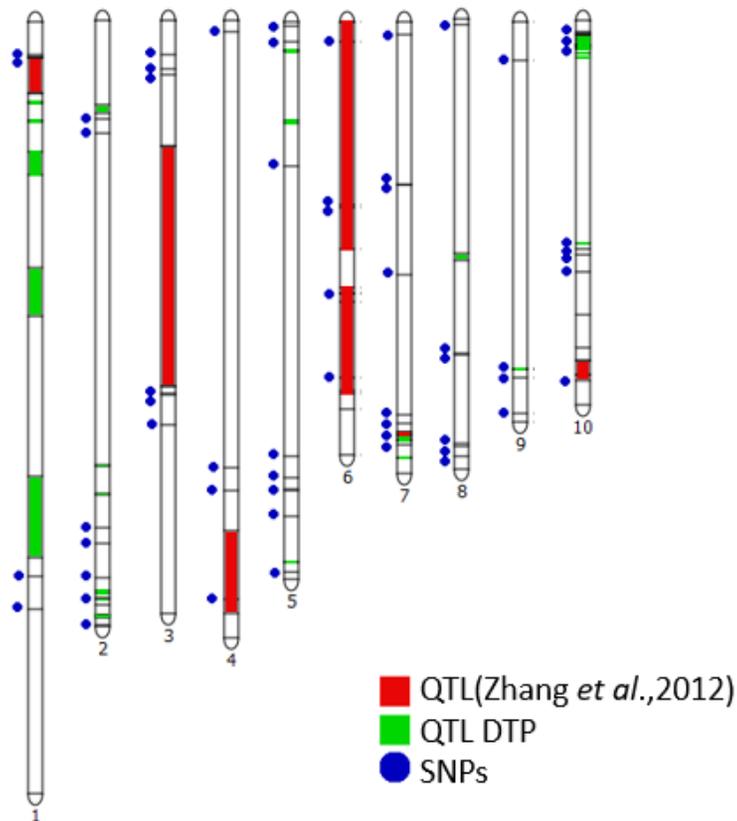


Figura 19. Mapa físico de las regiones genéticas asociadas a la resistencia a la siembra a 10 y 20 cm descritas por Zhang y colaboradores (2012) [rojo], regiones asociadas a la variación de la elongación del mesocotilo en la población DTP [verde] y SNPs asociados a la variación de la elongación del mesocotilo en la colección DTMA [azul].

Cuadro 11. Coincidencias entre los SNPs asociados a la elongación del mesocotilo en la colección DTMA que se encontraron en los QTLs de Zhang y cols. (2012), los QTLs de la población DTP mapeados en este trabajo y la función conocida de los genes asociados.

SNP obtenido por GWAS	p del SNP	QTL de Zhang y cols.	QTL deducido con la población DTP	Función del gen asociado al SNP
S1_14103918	3.47	qMES20-1		Tryp_alpha_amyl, proteína de transferencia de lípidos en plantas
S4_226459824	4.16	qMES20-4		Dominio asociado a metales pesados
S6_8146003	3.83	qMES20-6-1		<i>Oryza sativa</i> : proteína expresada
S6_72860333	3.41	qMES20-6-1		Lipasa_GDSL, acilhidrolasa
S6_139454544	3.29	qMES10-6-2		Codifica para una proteína desconocida
S7_160595701	3.56	qMES20-7		TIP49 C-terminal
S7_162027150	4.43	qMES20-7		Gen de baja confianza
S9_135860028	3.29		DTPMES-9-1	Gen de baja confianza
S10_10586901	3.53		DTPMES-10-6	Proteína de endomembrana 70

DISCUSIÓN

Fenotipos y resultados de la asociación genómica.

a) Patrones de elongación del mesocotilo.

Los valores promedio máximos de las accesiones DTMA y DTP para la elongación del mesocotilo fueron de 11.35 cm y 14.19 cm respectivamente. En comparación con las variedades criollas o razas de maíz de la Mixteca alta de Oaxaca, de la Faja Volcánica Trans-Mexicana o del SO de los Estados Unidos, estos mesocotilos serían considerados de cortos a medianos, como veremos más adelante (Collins, 1914; Vargas 1966; Embarcadero, 1971; Martínez Nava *et al.*, datos no publicados; Cruz Sánchez *et al.*, datos no publicados).

A pesar de no contar con los mesocotilos más largos de la especie *Zea mays* ssp. *mays*, los análisis fenotípicos del material genético usado en este trabajo detectaron suficiente diversidad fenotípica como para realizar de manera confiable GWAS y el mapeo de QTLs. El aumento en la elongación del mesocotilo de la población DTP, respecto a lo observado en la colección DTMA, podría atribuirse a la incorporación de nuevos alelos y/o al vigor híbrido generado al realizar las cruces de la línea DTMA 227 con los *testers* CML491 y CML503 (Figura 1). Como no se realizaron ensayos de siembra profunda para estos *testers*, no se puede saber

si el incremento en la elongación del mesocotilo fue debido a la introgresión de alelos con efectos fenotípicos positivos [que promueven los mesocotilos largos] o es un efecto de interacciones alélicas debidas al vigor híbrido.

En otras variedades de maíz se han encontrado valores promedio de la elongación del mesocotilo que son mayores a los 20 cm (Collins, 1914; Vargas 1966; Embarcadero, 1971; Martínez Nava *et al.*, datos no publicados; Cruz Sánchez *et al.*, datos no publicados). Embarcadero (1971) encontró que, de entre 25 variedades mexicanas de maíz analizadas, la raza cacahuacintle fue la de mayor valor promedio (32 cm) en la longitud del mesocotilo, mientras que la raza olotillo mostró el menor valor promedio (6.1 cm). También encontró que los maíces de los valles altos presentaron mesocotilos más largos que los provenientes de valles bajos.

Los ancestros silvestres (teocintles) tienen una amplia gama en la longitud de sus mesocotilos, variando en cada población entre cortos y largos, lo cual sugiere que la elongación del mesocotilo es un carácter ancestral que está bajo presión de selección (Cruz Sánchez *et al.*, datos no publicados). La baja elongación del mesocotilo para la colección DTMA y la población DTP podría deberse a que éstas provienen de variedades tropicales y subtropicales (Cairns *et al.*, 2013), variedades que históricamente no han sido seleccionadas para tener mesocotilos largos pues, de acuerdo a un estudio amplio de manejo agronómico tradicional, la siembra profunda se practica únicamente en regiones de valles altos y correlaciona fuertemente con los mesocotilos largos (Martínez Nava *et al.*, datos no publicados; Nieto-Sotelo *et al.*, 2018). Por lo tanto, se puede suponer que, de manera individual, una línea dada de la colección DTMA o de la población DTP no reúne en su genoma a todos los alelos o las combinaciones genéticas necesarias para tener mesocotilos largos, debido a su segregación histórica.

Como primera hipótesis se puede suponer que la colección de DTMA posee en su conjunto los alelos necesarios para generar mesocotilos largos y que ninguna de las líneas los tiene apilados. Una manera de someter a prueba experimental esta hipótesis es mediante cruces y selección de las líneas endogámicas DTMA que permitan acumular alelos con efecto positivo y disminuir alelos con efectos negativos. Si la hipótesis es correcta, las líneas derivadas podrían aumentar la elongación del mesocotilo en relación a sus parentales y, ya seleccionadas, podrían ser útiles en esquemas de siembra profunda. En apoyo a ésta idea se observaron las frecuencias alélicas de 18 de los 63 SNP asociados a la elongación del

mesocotilo en la población biparental DTP (Anexo IV). Cabe resaltar que las DTP tiene como progenitor a una de las líneas endogámicas DTMA (DTMA 227). De igual forma, el análisis de efectos fenotípicos de las DTMA demostró claramente que ninguna reúne todos los SNPs asociados a la elongación del mesocotilo (**Figura Suplementaria 2**).

b) SNPs en la colección DTMA y en la población DTP.

La secuenciación masiva del genoma mediante la técnica GBS ha probado ser útil para encontrar las variaciones genéticas de varias especies (Elshire *et al.*, 2011). Esta técnica nos permitió detectar casi un millón de SNPs para cada accesión de la colección DTMA y de la población DTP. La diversidad genética en éstas dos poblaciones es muy amplia, lo que permite usarlas para llevar a cabo análisis de asociación (Gowda *et al.*, 2015; Trachsel *et al.*, 2016).

La principal diferencia entre las DTMA y las DTP fue la cantidad de marcadores de alta calidad obtenidos para el análisis. Para la asociación genómica con las DTMA se utilizaron un poco más de 300,000 SNPs, mientras que para el mapeo de QTL con las DTP se usaron 30,428 SNPs, diez veces menos que en las DTMA; esto se debió a que la cantidad de SNPs con frecuencias alélicas menores a 0.05 en las DTP fue de casi 900,000 (Cuadro 10). Esto indica que en las DTP hay una gran cantidad de sitios con poca variación genética, ocasionada por una reducción de su diversidad genética, ya que las DTP son una población con solo dos líneas parentales y con 2 generaciones de recombinación, en comparación con las DTMA, que forman una colección de maíces con diferentes orígenes que han divergido entre sí a lo largo de cientos y quizá miles de generaciones (Cairns *et al.*, 2013). También es posible que esta gran cantidad de SNPs de baja frecuencia en las líneas DTP sean errores de secuenciación.

La colección de las DTMA ha sido utilizada para encontrar a los genes asociados a la resistencia a la enfermedad letal de la necrosis del maíz o MLND [*The Maize Lethal Necrosis Disease*] que es ocasionada por dos virus: el de la mancha clorótica del maíz [*Maize chlorotic mottle*] y el del mosaico de la caña de azúcar [*Sugarcane mosaic virus*]. Este es un carácter regulado por muchos genes con pequeños efectos fenotípicos (Gowda *et al.*, 2015), lo que nos indica que dentro de la colección han ocurrido suficientes eventos de recombinación que permiten identificar de manera detallada *loci* asociados a rasgos asociados a muchos genes con pequeños efectos fenotípicos. En el estudio de la resistencia a MLND se emplearon 264,000 SNPs que presentaron menos del 5% de datos faltantes, aproximadamente 100,000 SNPs menos de los que se analizaron en ésta tesis. Ésta diferencia entre estudios se debe a

que los filtros aplicados, fueron diferentes. Estos cambios se realizaron con la idea de mantener el mayor número posible de SNPs de alta calidad para alcanzar una mayor cobertura del genoma. Durante el desarrollo del GWAS se observó que la longitud del mesocotilo no correlaciona con la estructura poblacional de la colección DTMA, lo cual indica que ha habido suficientes eventos de recombinación como para romper el desequilibrio de ligamiento. Aún así, se aplicó una corrección de estructura poblacional utilizando los tres primeros componentes principales para disminuir al mínimo posible la probabilidad de asociación de falsos positivos (Bradbury *et al.*, 2007; Gowda *et al.*, 2015).

Los resultados del GWAS (Cuadro 3) corroboraron estudios previos (Troyer 1997; Zhang *et al.*, 2012) en el sentido de que la elongación del mesocotilo es un carácter regulado por muchos genes con pequeños efectos fenotípicos y es heredable. Se encontraron seis regiones cromosómicas donde los SNPs asociados al mesocotilo obtenidos por GWAS y los QTLs de Zhang y colaboradores (Zhang *et al.*, 2012) coinciden (Figura 19 y Cuadro 11). Cabe resaltar que los QTL encontrados por estos autores corresponden a regiones genómicas enormes y no se puede determinar todavía si estas coincidencias corresponden exactamente a los mismos genes o alelos (Figura 19). En las DTMA la elongación del mesocotilo resulta un carácter fuertemente asociado a 82 genes candidatos (Cuadro 4). Esta cantidad de genes asociados y la falta de estructura genética en la colección DTMA, nos indica que la elongación del mesocotilo es un carácter que ha tenido una historia evolutiva compleja.

Se ha propuesto que la evolución del maíz es un proceso que involucra dos grandes pasos, la domesticación de los ancestros silvestres para generar las razas y el mejoramiento de las razas que dio origen a las variedades mejoradas (Hufford *et al.*, 2012). Estos eventos evolutivos suponen una disminución de diversidad genética a causa de procesos de selección y cuellos de botella en cada paso.

Como se indicó anteriormente, la siembra profunda es una técnica agrícola que selecciona a las variedades de maíz que tienen la capacidad de emerger a altas profundidades y se esperaría que los genes que regulan esta característica estén bajo una fuerte selección. Sin embargo, está no es una técnica aplicada por todos los agricultores y por tanto existe una gran diversidad de tamaños de mesocotilos en las razas (Collins, 1914; Vargas 1966; Embarcadero, 1971; Muñoz, 2002; Martínez Nava *et al.*, datos no publicados; Cruz Sánchez *et al.*, datos no publicados; Nieto-Sotelo *et al.*, 2018a).

Para el caso de las variedades mejoradas, se ha observado que estas presentan mesocotilos entre medianos y cortos, ya que la mayoría de las variedades mejoradas tienen un origen tropical y subtropical que son maíces de mesocotilos cortos (Nieto Sotelo *et al.* 2013). De seguir esta tendencia, se esperaría una disminución de la diversidad genética en los genes asociados a la elongación del mesocotilo en la domesticación y el mejoramiento. Para conocer si los genes asociados a la elongación del mesocotilo están bajo estos procesos selectivos, se compararon con los genes asociados a la domesticación y al mejoramiento del maíz propuestos por Hufford y colaboradores (2012), quienes compararon el nivel de variación genética entre tres grupos: silvestres, domesticados y mejorados. Utilizaron 72 muestras: 17 ancestros silvestres, 14 *Zea mays ssp parviglumis*, dos de *Zea mays ssp. mexicana* y uno de *Tripsacum dactyloides var. meridionalis*, también incluyeron a 23 razas, como Bolita (de Nochixtlán, Oaxaca), Havasupai (de Supai, Arizona) y Santo Domingo (de Santa Domingo Pueblo, Nuevo México), que podrían tener mesocotilos largos, y 35 variedades mejoradas de diferentes orígenes, posiblemente de mesocotilos cortos. Ya que las variedades silvestres pueden presentar mesocotilos cortos, medianos y largos, al igual que las razas nativas de maíz (Cruz Sánchez *et al*, datos no publicados; Nieto-Sotelo *et al.*, 2018a), al comparar los genotipos de las variedades silvestres con las razas criollas, no se deberían encontrar diferencias significativas, ya que la suma de toda la diversidad genética que hay entre los maíces domesticados podría encubrir estos cuellos de botella. Para poder encontrar estas asociaciones, se necesitaría analizar el universo de los maíces como subpoblaciones separadas por el tamaño de sus mesocotilos [cortos, medianos o largos] (Figura 20).

La capacidad de detectar los cambios de diversidad genética de los genes asociados a la elongación del mesocotilo es limitada en el estudio de Hufford y colaboradores (Hufford *et al.*, 2012), porque en todos los niveles podrían estar incluidas variedades con mesocotilos cortos, medianos y largos. Además, la elongación del mesocotilo es un carácter multigénico con pequeños efectos fenotípicos, que podría estar bajo interacciones génicas y cuyos alelos podrían distribuirse de acuerdo al origen de las variedades de maíz disminuyendo drásticamente la presión de selección para una sola región genética. Por otro lado Es posible que la baja cantidad (*chip* de 50,000 SNPs) de SNPs usados en el trabajo de Hufford y cols. (2012) haya omitido una gran cantidad de SNPs asociados a la elongación del mesocotilo. Pese a esto, se encontraron genes coincidentes que han jugado un papel relevante durante la evolución del maíz.

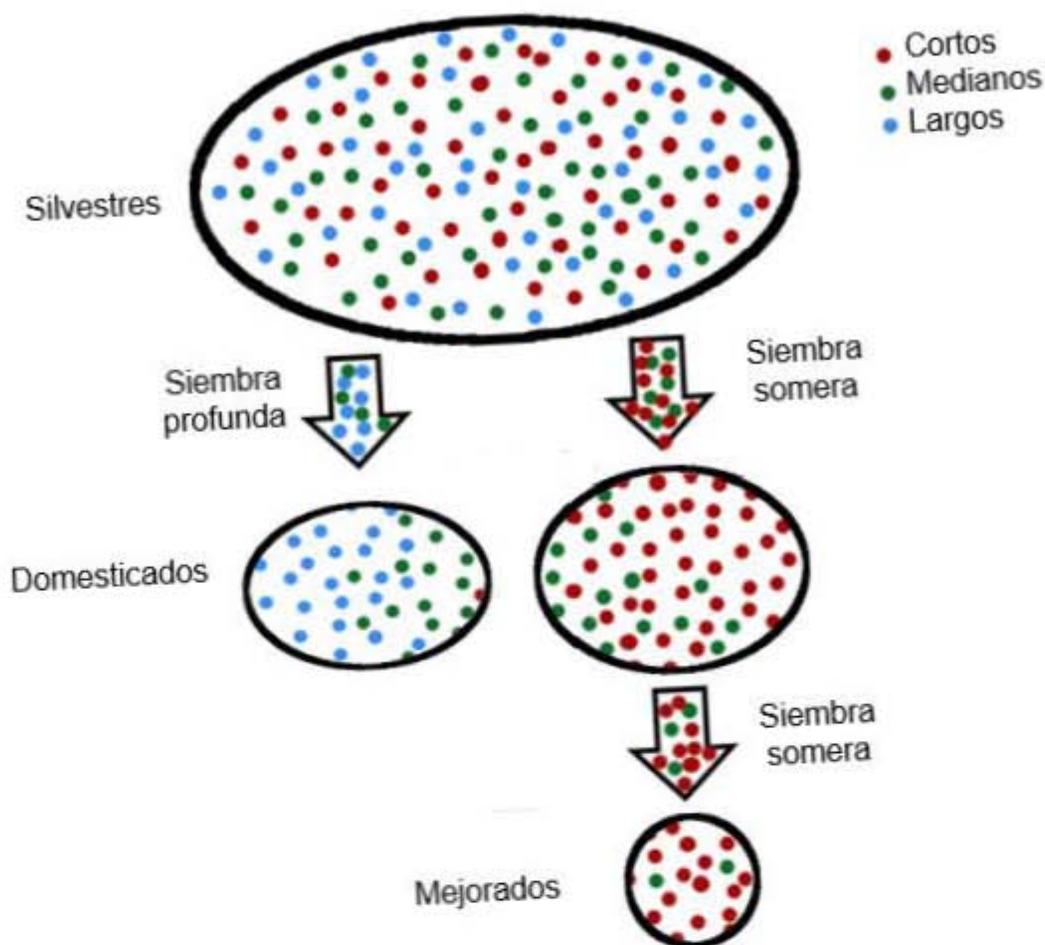


Figura 20. Esquema del proceso de evolución de la elongación del mesocotilo, variación fenotípica sometida a la selección de siembra profunda y siembra somera en la domesticación y mejoramiento.

Al comparar los genes asociados a la elongación del mesocotilo con los genes propuestos por Hufford se encontraron dos genes coincidentes, uno como candidato del proceso de domesticación y otro como candidato del mejoramiento. El gen GRMZM2G014750 se encontró asociado a la domesticación y codifica a un factor de transcripción de función desconocida en maíz; la inactivación de su ortólogo en *A. thaliana* hipertrofia al núcleo celular generando un mal desarrollo embrionario (Fleurdépine *et al.*, 2007). Por otra parte, el gen GRMZM2G119527 está asociado al mejoramiento y codifica a un transportador de iones de hierro a través de la membrana celular (Andorf *et al.*, 2016). Este gen podría estar involucrado en la adaptación del maíz a diferentes tipos de suelo, como se explicará más adelante.

La raza Palomero Toluqueño presenta mesocotilos largos y se cultiva tradicionalmente en el eje neovolcánico de México, por lo cual podría tener los alelos que hemos asociado para la variación en la elongación del mesocotilo (Cruz Sánchez *et al.*, datos no publicados; Nieto-Sotelo *et al.*, 2018). Para localizar los alelos de los SNP asociados a la variación de la elongación del mesocotilo en las DTMA (**Anexo V**), en el genoma del Palomero Toluqueño, se interrogó la base de datos de maíz GDB con secuencias genómicas de 101 bases compuestas por el SNP de interés y las 50 bases en ambos flancos(<https://www.maizegdb.org/>; Andorf *et al.*, 2016; Vielle-Calzada *et al.*, 2009).

De los 63 SNP asociados a la variación de la elongación del mesocotilo en la colección DTMA, se encontraron 36 en el genoma del Palomero Toluqueño, de estos, ocho se encontraron homocigotos para el alelo menor en la colección DTMA (**Anexo V**). Por otra parte, 27 SNPs asociados a la elongación del mesocotilo no fueron localizados en el genoma del Palomero Toluqueño, mientras que 28 SNPs coincidieron con alelos mayores. Otro dato que aporta elementos importantes para apreciar mejor la relevancia de estas 8 coincidencias es el dato de que los alelos de frecuencia menor de los 63 SNP asociados a la variación de la elongación del mesocotilo en la colección DTMA están ausentes en B73, el genoma de referencia para definir los SNPs en las DTMA. Al respecto se debe destacar el hecho de que B73 produce mesocotilos cortos con un promedio de 3.2 cm (Nieto-Sotelo *et al.* 2018b). Debido al desconocimiento de el tamaño del mesocotilo en respuesta a siembra profunda de la accesión de Palomero Toluqueño empleada para ensamblar el genoma de esta raza, no es posible evaluar la relevancia de los 8 alelos menores asociados a la elongación del mesocotilo. Un análisis reciente de otra accesión de Palomero Toluqueño indicó que desarrolla mesocotilos medianos [12 cm] (Nieto-Sotelo *et al.*, 2018b). Será importante analizar con cuidado los genomas de más razas de maíz con fenotipos conocidos (resistencia o sensibilidad a la siembra profunda; mesocotilos cortos o largos) para evaluar la correlación entre estos rasgos y la presencia de los alelos menores de los 63 SNPs asociados a la elongación del mesocotilo.

Los SNPs asociados que presentaron el alelo menor fueron: S3_13726673, que se encuentra en una región intergénica cercana al gen GRMZM2G048022 que no tiene función conocida y el gen GRMZM2G047815, que codifica para el complejo de transferasa GPI-GlcNAc. S3_158354184 asociado al gen GRMZM2G139617, que codifica para una proteína con el dominio DEK C-terminal. S5_1852022 que se encuentra dentro del marco de lectura del gen GRMZM2G138421 que codifica para una proteína que se une al promotor de *Squamosa*.

S5_183326603, que se encuentra dentro del marco de lectura del gen GRMZM2G014750 que codifica una proteína con el dominio La que se une a RNAs. S8_167282998, que se encuentra dentro del marco de lectura del gen GRMZM2G177324 que codifica para una peptidasa. S8_170878813, en el gen GRMZM2G069106 que codifica un factor de transcripción. S10_140743010, que se encuentra en el gen AC208110.2_FG007 del cual no se conoce su función en maíz, pero se ha descrito que codifica a una proteína específica de la cariósida de *Oryza sativa* (**Anexo III**). El SNPs S8_2763004, se encuentra dentro del marco de lectura del gen GRMZM2G312419 que está asociado a un factor de transcripción tipo Myb 60; los genes de esta familia están involucrados en el desarrollo y en la respuesta a estrés de las plantas (Ambawat *et al.*, 2013).

Estos SNPs podrían estar directamente relacionados a las variedades con mesocotilos largos y sus genes candidatos son buenas opciones para realizar estudios funcionales profundos que nos permitan entender cómo estas variaciones genéticas repercuten en la variación de la elongación del mesocotilo.

Genes asociados a metales pesados y su implicación evolutiva.

El análisis de asociación genómica indicó que el SNP con el mayor valor de asociación fue S7_162027150, ubicado en una zona intergénica. Los genes más próximos río arriba y río abajo, son modelos de gen de función desconocida.

Entre las principales funciones (pfam y GO terms) reportadas para los 63 genes asociados a la variación de la elongación del mesocotilo en las DTMA, se encontraron genes relacionados con interacciones con metales como: el transporte de metales pesados (GRMZM2G174673), la actividad de metaloendopeptidasa (GRMZM2G044697), la cupredoxina (GRMZM5G813403), la unión a iones de zinc (GRMZM2G093022), el transporte de hierro a través de la membrana celular (GRMZM2G119527), la unión a iones de calcio (GRMZM2G106680) y el transporte de cadmio-zinc (GRMZM2G099191).

Una de las posibles razones de encontrar esta relación entre el largo del mesocotilo y el transporte de metales, puede ser atribuida al tipo de suelo donde se realiza la siembra profunda en México. Actualmente esta práctica se lleva a cabo en la región Mixteca alta de Oaxaca y en la Faja Volcánica Trans-Mexicana (Martínez Nava *et al.*, datos no publicados; Cruz Sánchez *et al.*, datos no publicados). Estos suelos son derivados de rocas y cenizas volcánicas y pertenecen al orden de los andosoles y vertisoles. Son suelos ligeros, con alta

capacidad de retención de agua, nutrientes y fósforo (Aguilera, 1989) acumulan complejos organometálicos como alofanos (aluminosilicatos) y ferrihidrita (Shoji *et al.* 1996; Takahashi y Shoji *et al.*, 2002). Las altas concentraciones de metales como el aluminio, cobre, zinc, hierro, cadmio y magnesio, pueden ser tóxicas para las plantas, pues afectan el desarrollo de la raíz y dañan otros órganos debido a que generan especies reactivas de oxígeno. En los suelos con mayor contenido de arcilla y en la temporada de sequía es mayor la concentración y disponibilidad de metales pesados (Hernández *et al.*, 2011). Para poder sobrevivir en este ambiente, las plantas han evolucionado estrategias que reducen la toxicidad de los metales pesados como: 1) mantener una baja concentración de metales pesados en la parte aérea, aumentado la concentración en la raíz o acumulándolo en zonas poco sensibles, 2) aumentando la producción de compuestos quelantes y transportadores de metales pesados, y 3) activando las defensas al estrés oxidativo (Rao *et al.*, 2016). Por otro lado, el maíz es uno de los cultivos con mayor resistencia a altas concentraciones de aluminio (Baran, 2013) y es el cultivo más afectado por deficiencia de aluminio (Alloway, 2008), por lo que es posible cultivarlo en lugares con altas concentraciones de aluminio, como en los suelos de origen volcánico.

También se ha postulado que la resistencia a los metales pesados es un carácter relacionado a los procesos de domesticación y mejoramiento del maíz (Vielle-Calzada *et al.*, 2009). Esta relación fue encontrada al comparar el genoma del Palomero Toluqueño con el de B73. En el mencionado estudio se utilizó como genoma de referencia al Palomero Toluqueño que, como se había mencionado antes, es una de las variedades con mesocotilos más largos. La accesión de Palomero Toluqueño utilizada provino de San Lorenzo Teotuitlán, Edo. De Méx., ubicado a una altura de 2,453 msn. Esta localidad pertenece a la Faja Volcánica Trans-Mexicana y, por ende, éste Palomero Toluqueño crece en suelos con las características antes discutidas. Ocho alelos de frecuencia menor asociados a la variación de la elongación del mesocotilo en las DTMA se encontraron en el genoma del Palomero Toluqueño. Queda por analizar esta accesión (de San Lorenzo Teotuitlán, Edo. de Méx.) mediante ensayos de siembra profunda con la finalidad de saber cuál es la longitud del mesocotilo.

También queda en deuda el estudio de la relación que tiene la elongación del mesocotilo con la resistencia a ambientes ricos en metales pesados, aunque postulamos que es un carácter secundario que correlaciona con el hecho de que las razas nativas de siembra profunda se siembran en este tipo de suelos, que son diferentes de las de siembra somera.

Se puede añadir que la siembra profunda se realiza en campos no irrigados con poca disponibilidad de agua. A pesar de que la planta aprovecha la humedad residual, esta podría no ser óptima generando estrés hídrico en la planta, lo que a su vez provocaría el aumento de la toxicidad tanto de los metales pesados como de las especies reactivas de oxígeno al interaccionar entre sí de manera sinérgica, lo que en cierto momento podría frenar el crecimiento de la plántula y el mesocotilo (Saab *et al.*, 1992; Zhang *et al.*, 2007). Además, debemos tomar en cuenta que los suelos donde se realiza esta práctica tienen una alta concentración de metales que aumentan su disponibilidad en condiciones de sequía (Hernandez *et al.*, 2011), lo que implica que para tener mesocotilos largos las plántulas necesitan tener mecanismos que les ayuden a soportar el estrés hídrico y la toxicidad de algunos iones, de tal modo que no interrumpan su crecimiento. Apoyando la idea de la resistencia al estrés también se encontraron los siguientes genes asociados con las siguientes funciones: GRMZM2G126749 que codifica para una proteína que interactúa con la cadmodulina, el gen GRMZM2G039263 que codifica para una proteína HSP60 que tiene la función de chaperona que ayuda al correcto plegamiento proteico y el gen GRMZM2G165844 que tiene la función de respuesta al estrés oxidativo. Las posibles interacciones génicas entre los genes asociados al transporte de metales y los genes asociados a la resistencia al estrés podrían ser puntos clave para la generación de variedades adaptadas a la siembra profunda y todo lo que conlleva esta práctica, por ello es necesario conocer la variación genética responsable de esta resistencia en las variedades adaptadas a la siembra profunda.

Crecimiento celular.

El SNP S3_19260465 se encuentra en el marco de lectura del gen GRMZM2G044697 que codifica para una familia de peptidasas, que pertenecen al clan M50 de metaloendopeptidasas, que tienen un ligando de zinc y están involucradas en la proteólisis intramembranal e intervienen en la regulación de la expresión génica de la síntesis del colesterol (Barrett, 2013). Un subgrupo de estas proteínas son las metaloproteinasas de la matriz celular también conocidas como [MMPs], que están involucradas en el remodelamiento de la matriz extracelular (Marino y Funk, 2012). No se conoce la función de estas proteínas en maíz, pero se encontró que *Pta1-MMP* en *Pinus taeda* tienen su mayor expresión durante la germinación de la semilla y durante la primera etapa del crecimiento de la plántula interviniendo en la división y expansión celulares (Ratnaparkhe *et al.*, 2009). Para el caso de *Arabidopsis thaliana*,

se ha encontrado que los genes *MMP* tienen una expresión diferencial en los tejidos y en las etapas de crecimiento incrementándose durante el desarrollo de la semilla y durante la germinación. (Marino y Funk, 2012).

El gen GRMZM2G044697 tiene una mayor expresión en la plántula, en las primeras hojas plumulares y se expresa tanto en la semilla como en la raíz (Sekhon *et al.*, 2011; Stelpflug *et al.*, 2015). Claramente esta proteína está involucrada en las primeras etapas del desarrollo de la plántula. Por otro lado, su asociación con la variación en la elongación del mesocotilo podría atribuirse a que esta proteína está involucrada en el alargamiento de las células de la plántula, mediante modificaciones en la matriz extracelular, probablemente asociados a cambios en los niveles de expresión dando como resultado variación en la elongación del mesocotilo. El gen GRMZM2G031261 codifica para una proteína PSK, proteína precursora de la fitosulfoquina, la cual se sabe que promueve la proliferación y el crecimiento celulares (Sauter, 2015).

Genes involucrados en el transporte y metabolismo del nitrógeno.

El SNP S8_166105182 tiene un valor de asociación de p de 4.22 y un efecto fenotípico para el alelo menor de -0.56 cm y se encontró asociado al gen GRMZM2G335218, que tiene la función de transportar nitrógeno, principalmente amonio y urea, a través de la membrana celular. El nitrógeno es obtenido del suelo por la raíz y transportado a toda la plántula (Ludewig *et al.*, 2007); su mayor concentración en el suelo se encuentra en la capa superficial (Benintende *et al.*, 2008) entre los 0 y 30 centímetros de profundidad (Cristóbal *et al.*, 2011). Esta función es coherente con la posición de la plántula por debajo de los 30 cm, cuando se realiza la siembra profunda. En estas condiciones se esperaría una menor concentración de nitrógeno al de un ambiente somero, por lo que se necesitan variantes alélicas que contribuyan a una mejor captación de nitrógeno y mejoren la adaptación a éste manejo agronómico. Aunado a esto, este rasgo agronómico podría usarse para mejorar la capacidad de las plantas de crecer en suelos pobres en nitrógeno independientemente de la profundidad de siembra. Siguiendo la misma línea de pensamiento, los maíces que no están completamente adaptados a la siembra profunda podrían aumentar su crecimiento y emergencia si se incrementara la concentración de nitrógeno en el suelo antes del ciclo agrícola o cuando se realiza la siembra.

El SNP S2_198509433 se encontró en una región intergénica; el gen más cercano río abajo es un elemento transponible y el gen más cercano río arriba es un acarreador de

molibdeno (GRMZM2G037875). El molibdeno es uno de los micronutrientes con menor concentración en las plantas. Forma parte de la nitrato reductasa y la nitrogenasa y regula el óxido nítrico intracelular (Giles *et al.* 2014). Está involucrado en los últimos pasos de la síntesis de ácido abscísico [ABA] y de las auxinas [IAA] (Kaiser *et al.* 2005). El déficit de molibdeno genera una baja asimilación de nitrógeno, provocando un aumento de nitrato en los tejidos de las plantas. En las DTMA se encontró que el alelo con frecuencia menor está asociado a mesocotilos más cortos. Esto podría atribuirse a que este alelo disminuye la capacidad de transporte de molibdeno en las plántulas, interfiriendo en la absorción y aprovechamiento del nitrógeno en los tejidos y reduciendo el crecimiento del mesocotilo. También se podría hipotetizar que las líneas con este alelo tienden a producir menos auxinas y/o ácido abscísico, con el consecuente desbalance hormonal que podría influir en la elongación del mesocotilo (Saab *et al.*, 1992).

Factores de transcripción.

El SNP S5_1852022 tiene un valor de p de 4.02 y se encuentra en el marco de lectura del gen GRMZM2G138421 también llamado ZmSPL18, que codifica para el factor de transcripción ZmSBP6 que se une al promotor de *Squamosa* (SBP). En maíz hay 31 de estos factores de transcripción y están subdivididos en 6 grupos. El gen ZmSPL18 pertenece al quinto grupo, está compuesto por 11 exones y mide 2,886 pb; éste grupo, además de tener dos sitios de unión de zinc, como todos los miembros de esta familia, tiene tres o cuatro dominios de anquirina [ANK o Ank-2] (Mao *et al.*, 2016). Estos dominios se encuentran en proteínas transmembranales que están asociados a la respuesta a la sequía y al estrés oxidativo (Sharma y Pandey, 2016). Este gen tiene una alta expresión en todos los tejidos de la planta durante todo su ciclo de vida (Sekhon *et al.*, 2011) y su expresión varía en respuesta a diferentes factores de estrés: la sequía reduce su expresión, mientras que el frío, la hipersalinidad y el ABA aumentan su expresión (Mao *et al.*, 2016). Aunque la función específica de este gen es aún desconocida, es claro que se trata de un gen esencial y que la planta al encontrarse en condiciones de sequía limita su expresión, probablemente interrumpiendo el desarrollo para poder hacer frente a este estrés. La relación de este gen con la elongación del mesocotilo podría deberse a que la siembra profunda se lleva a cabo en condiciones limitadas de agua lo que generaría una disminución en su expresión, limitando la elongación del mesocotilo. Es probable que ciertos alelos no sean tan sensibles a la sequía y continúen

expresando este gen promoviendo la elongación del mesocotilo. En las DTMA, el efecto fenotípico del SNP S5_1852022 tiene un efecto del alelo menor negativo de -0.79 cm, este alelo podría estar relacionado a que las DTMA tengan una hipersensibilidad a la sequía y limiten la elongación del mesocotilo en condiciones de siembra profunda.

Genes de la literatura que regulan la elongación del mesocotilo.

Es interesante que, pese a que se contó con SNPs dentro de los marcos de lectura de los genes *ZmLA1* y *Elm1*, el GWAS mostró valores $-\log_{10}$ de p de asociación menores al mínimo aceptado como fuertemente asociados a la variación de la elongación del mesocotilo, lo cual indica que pueden estar involucrados pero no son los que más contribuyen en esta variación. ¿Cómo explicar esta omisión tan importante en el GWAS? Los fuertes efectos pleiotrópicos de las dos mutaciones conocidas en *ZmLA1* podrían ayudar a entender este resultado, ya que además de desarrollar mesocotilos sumamente largos en la oscuridad, presentan un tallo postrado, debido a que su respuesta gravitropica es más lenta, y tienen problemas de floración. Aún no se conoce la función específica de la proteína codificada por este gen. Sin embargo, se ha observado que la falta de este transcrito se relaciona fuertemente con el transporte de las auxinas. Además, las mutaciones en *ZmLA1* tienen fuertes efectos pleiotrópicos, ya que afectan la expresión de 931 genes (Dong *et al.*, 2013). Los grandes efectos pleiotrópicos de las mutaciones en *ZmLA1* podrían explicar su nula selección en respuesta a la siembra profunda, pues reducen fuertemente el vigor y desarrollo de la planta.

En cuanto a los mutantes *elm1*, éstos presentan mesocotilos muy largos en condiciones de luz, aunque en oscuridad su tamaño es similar al silvestre. Además tienen severas deficiencias en la percepción y en la respuesta a la luz roja y roja lejana (Sawers *et al.*, 2002). La posible razón de que no se encontrasen alelos fuertemente asociados de *elm1* a la elongación del mesocotilo por GWAS en los DTMA, se podría deber a que la fenotipificación de los híbridos DTMA fue realizada en la oscuridad, lo cual encubre el fenotipo de los mutantes *elm1*. Sin embargo, al igual que *ZmLA1*, las variaciones en este gen causan efectos fuertemente pleiotrópicos que impiden su selección agronómica.

Comparación entre el GWAS y el análisis de QTL de Zhang y colaboradores.

Los SNPs asociados a la variación de la elongación del mesocotilo del presente trabajo fueron comparados con los QTLs de resistencia a la siembra profunda reportados por Zhang y

colaboradores. (2012). Para esto se consideraron las posiciones físicas de los *bins* reportados en la base de datos del MaizeGDB. Se tomaron en cuenta tanto los QTL asociados a siembras a 10 cm como a 20 cm de profundidad. Esta comparación no pudo ser muy fina, ya que los QTLs de Zhang y cols. comprenden grandes segmentos del genoma que, en algunos casos, ocupan más de 100,000 pares de bases (Cuadro 12).

Se encontraron 8 SNPs con valores $p \geq 3.25$ dentro de 6 QTLs de Zhang y colaboradores: qMES10-1, qMES20-1, qMES20-4, qMES10-6-2, qMES20-6-1 y qMES20-7. En el caso de otros 6 QTLs: qMES10-3, qMES20-3, qMES10-6-1, qMES20-6-2, qMES10-10 y qMES20-10 no se encontraron SNPs con valores de $p \geq 3.25$ (Cuadro 12). Tres SNPs asociados a mesocotilo coincidieron en el QTL qMes20-6-1, mientras que otros dos SNPs coincidieron con qMES20-7. Mediante esta comparación se ratificó la importancia de 9 de los 63 SNPs propuestos; estos son: S7_162027150, que es un SNP intergénico que presentó el valor más alto de p y está flanqueado por dos genes de baja confianza; S1_14103918, que se encuentra dentro de dos QTL y en el marco de lectura del gen GRMZM5G850455 con la función de transportador de lípidos; S4_226459824, en el gen GRMZM2G174673 asociado al transporte de metales pesados; S6_139454544, con el modelo GRMZM2G423863 de función desconocida; S6_8146003, en el gen GRMZM2G136700 de función desconocida; S6_71803064, que es elemento transponible; S6_139454544, en el gen GRMZM2G423863 sin caracterizar; S7_160595701, en el gen GRMZM2G058087 con la función de chaperona involucrada en la regulación del ADN. Las coincidencias observadas son en cierto modo burdas debido al gran tamaño de las regiones génicas utilizadas para definir a los QTLs en el trabajo de Zhang y colaboradores (2012). Además, debemos tomar en cuenta que desconocemos las distancias genéticas de la población utilizada por Zhang y colaboradores (2012) y si comparten los alelos menores de los SNPs mencionados de la colección DTMA.

Aunado a lo anterior, el trabajo de Zhang y cols. no caracterizó de manera específica QTLs relacionados a mesocotilos largos sino a resistencia a siembras de 10 y 20 cm de profundidad. Sabemos que este tipo de resistencia puede depender, ya sea de la elongación exagerada del mesocotilo o del crecimiento de la plúmula bajo el suelo previa ruptura del coleptilo (Martínez Nava, datos no publicados; Cruz Sánchez, datos no publicados; Collins, 1914).

Asociación genómica y análisis de QTL en la población DTP.

Para conocer si las regiones asociadas a la variación de la elongación del mesocotilo encontradas en la colección de líneas DTMA son las mismas para otras poblaciones, se recurrió al análisis de QTL usando una población recombinante BC₁F_{2:3} para poder evaluar la correlación entre la segregación del carácter y la segregación alélica. Se eligió el análisis de QTL sobre la asociación genómica, debido a que las poblaciones biparentales con pocas cruzas han experimentado pocos eventos de recombinación, no tienen una gran diversidad genética y, por el contrario, tienen una fuerte estructura genética, lo que genera falsos positivos si se realiza GWAS (Li *et al.*, 2015).

En éste trabajo se utilizaron miles de SNPs obtenidos por GBS, lo que aumentó la cantidad de marcadores por región muchos de ellos fuertemente ligados entre sí, con estos marcadores se pudieron inferir los genotipos parentales a través del protocolo de ventanas deslizantes y la formación de bloques (Xie *et al.*, 2010). Esta técnica resultó ser muy efectiva para reconstruir los eventos de recombinación, generar mapas de recombinación y obtener las distancias genéticas entre los marcadores (Cuadro 8). Con los bloques se generó un mapa de ligamiento con 2,229 bloques con un tamaño total de 14,471 centimorgans. El mapa modelo donde se han podido reconocer la mayor cantidad de eventos de recombinación pertenece a la población IBM, en este se han encontrado 6,618 bloques de recombinación con un tamaño de 11,198 centimorgans (Liu *et al.*, 2015). El mapa aquí propuesto tuvo un tercio de los bloques de recombinación encontrados en la población IBM. Una explicación de estas diferencias es que la población IBM ha tenido más cruzas y su tamaño poblacional es mayor, lo que implica un mayor número de eventos de recombinación con lo que se obtienen más bloques. Además, la población DTP se formó al realizar la cruce de la DTMA 227 con el *tester* CML491, una cruce de la F₁ con el *tester* CML491 y una autocruza que al finalizar nos deja con sólo el 25% de la variación genotípica proveniente de la DTMA 227 y el otro 75% del material genético proveniente del *tester* CML491 (Figura 2 y 17). Esto disminuye la probabilidad de obtener eventos de recombinación de algunas regiones genéticas.

Con el mapa de ligamiento se realizó el análisis de QTL para la variación en la elongación del mesocotilo. Se encontraron 30 QTLs para la elongación del mesocotilo con valores de asociación LOD entre 2.002 y 4.45 (Cuadro 12). Estos valores son bajos en comparación con los valores LOD reportados por Zhang y cols. (2012) que van de 3.77 a 14.21, en su ensayo de 20 cm de profundidad, y entre 3.92 y 11.57, en su ensayo de 10 cm de

profundidad. La diferencia de estos valores podría atribuirse a las diferencias en las líneas parentales. En el trabajo de Zhang y cols. (2012) una de las líneas parentales era derivada del maíz *Hopi*, que es una variedad que ha sido adaptada a la siembra profunda y presenta mesocotilos sumamente largos (Collins, 1914). Por el contrario, las líneas parentales de las DTP nunca fueron seleccionados para esta práctica.

No obstante los bajos valores LOD de la población DTP, se han podido identificar 42 QTLs que controlan el vigor de la plántula y la resistencia a la sequía (Trachsel *et al.* 2016) con valores que oscilan entre 2.0 y 5.0. La complejidad de los caracteres estudiados en las poblaciones DTP bastaría para explicar la obtención de muchos QTL con efectos menores sobre el fenotipo (Trachsel *et al.* 2016). Sin embargo, una hipótesis más razonable sería postular que esto se debe al material genético usado, es decir, contar con líneas progenitoras con caracteres contrastantes y extremos podría ayudar a detectar QTL con mayores valores LOD que expliquen mayores efectos fenotípicos. Este no fue el caso de la población DTP para el estudio de la elongación del mesocotilo, pues la DTMA 227 y la CM L491, sus progenitoras, no generaron líneas con mesocotilos muy largos al hibridarse con CML503.

Para evaluar si las regiones asociadas a la elongación del mesocotilo son las mismas para dos poblaciones, se compararon los resultados de los dos enfoques metodológicos de asociación empleados. Este experimento es limitado, debido a que los datos obtenidos mediante GWAS con las 272 líneas DTMA engloban, en conjunto, una mayor diversidad genética que la de la población DTP que proviene de dos líneas parentales, cuya desventaja radica en que solo se pueden evaluar los efectos de unos cuantos alelos y solo se pudieron evaluar 18 SNPs fuertemente asociados (**Anexo IV**).

Los SNPs y los QTL asociados a la elongación del mesocotilo difirieron en la cantidad de regiones asociadas. A partir de la colección DTMA se encontraron 63 SNPs asociados mediante GWAS; de estos, 18 SNPs eran evaluables en el análisis de la población DTP. Pese a estas limitaciones, se pudo corroborar la importancia de un SNP, S9_135860028 que se encuentra en el marco de lectura del gen GRMZM2G500857 que es un gen de baja confianza y S10_10586901 que está dentro del marco de lectura de un gen con la función de una proteína componente de la membrana celular, desgraciadamente no se conoce más acerca de estos genes y no se puede interpretar de forma precisa cómo afectan a la variación de la elongación del mesocotilo en la colección DTMA y DTP. La menor diversidad genética en la población DTP permite sugerir que se realicen más análisis de QTL con otras líneas parentales DTMA.

Sin embargo, los costos económicos y de tiempo de esta validación podrían ser una clara limitante. Al comparar los QTLs de elongación de mesocotilo de las DTP con el trabajo de Zhang y cols. (2012) no se encontraron regiones coincidentes.

El reconocimiento de los genes asociados fue un primer paso para explicar la variación en la elongación del mesocotilo. Sin embargo, es necesario, aplicar otras técnicas para validar la efectividad de la asociación y descartar falsos positivos.

Debido a que los SNPs con mayor valor de asociación pueden no ser los causales de la variación fenotípica, será muy importante resecuenciarlos en una cantidad considerable de líneas DTMA, ya que la secuenciación GBS sólo reconoce parte los SNPs del genoma de maíz. En jitomate se encontró que el gen *SIMYB12* está asociado al color rosa de algunas variedades. Para descubrir los SNPs causales se realizó una re-secuenciación de este gen, utilizando 20 accesiones de tomates rojos y 20 accesiones de tomates rosas, lo que permitió reconocer una delección de 603 pb para la mayoría de las accesiones y la introducción de codones de paro para las demás accesiones. Estos alelos disminuyeron la expresión del gen *SIMYB12* (Lin *et al*, 2014). Este es un ejemplo del poder predictivo de la asociación genómica, pero requirió del uso de otras técnicas que pudieran validar y reconocer las variaciones genéticas causales de los cambios fenotípicos.

CONCLUSIONES

La variación de la elongación del mesocotilo en las DTMA y DTP es una característica regulada por muchos genes con pocos efectos fenotípicos. La estructura genética de la variación de la elongación del mesocotilo nos permitió reconocer los loci (SNPs) asociados y proponer 82 genes candidatos que regulan este rasgo en la DTMA y en la población DTP se encontraron 29 QTLs. Las principales funciones de los genes candidatos estuvieron involucradas con: actividad catalítica, hidrólisis de proteínas y carbohidratos, transporte de iones metálicos, transporte intracelular e interacción con ADN. Al comparar las regiones asociadas a la elongación del mesocotilo entre las DTMA y las DTP, se lograron reconocer dos regiones coincidentes, pero no hay suficiente información que nos permita conocer cuál es la función de los genes asociados.

PERSPECTIVAS

Este trabajo nos permitió reconocer loci asociados a la variación de la elongación del mesocotilo pero aún es necesario determinar si hay relaciones de aditividad y/o epistasis de los SNPs fuertemente asociados. Además, tendrán que generar modelos que puedan predecir las interacciones entre los genes asociados y permitan seleccionar aquellos alelos que al apilarlos generen plantas resistentes a la siembra profunda con mesocotilos más largos. Será igualmente importante utilizar los SNPs asociados a la elongación del mesocotilo como referencia para realizar la re-secuenciación de los genes asociados e identificar los SNPs o haplotipos causales de la variación, mediante una nueva asociación genómica y observar las frecuencias alélicas de los SNPs asociados en poblaciones nativas de maíz, de siembra somera y siembra profunda.

LITERATURA CITADA

Aguilera, H. N. (1989). *Tratado de edafología de México*. México, Facultad de Ciencias.

Alexandrov, N., Brover, V., Freidin, S., Troukhan, M., Tatarinova, T., Zhang, H., Swaller, T., Lu, Y., Bouck, J., Flavell, R., Feldmann, K. (2009) Insights into corn genes derived from large-scale cDNA sequencing. *Plant Mol Biol* 69(1-2):179-194.

Alloway, B., (2008). *Zinc in Soils and Crop Nutrition*. Brussels. Belgium: IZA and IFA.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, y Lipman, D. (1990). Basic Local Alignment Search Tool. *Journal of molecular biology*, 215(3):403–10.

Ambawat, S., Sharma, P., Yadav, N. R., & Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. *Physiology and Molecular Biology of Plants*, 19(3), 307–321.

Anderson, E., y Cutler H. C. (1942). Races of Zea Mays: I, Their Recognition and Classification. *Annals of the Missouri Botanical Garden*, 29:69–88.

Andorf, C. M., Lawrence, C. J., Harper, L. C., Schaeffer, M. L., Campbell, D. A., y Sen, T. Z. (2010). The Locus Lookup tool at MaizeGDB: identification of genomic regions in maize by integrating sequence information with physical and genetic maps. *Bioinformatics* 26(3), 434–436.

Barrett, A. (Ed). (2013). *Handbook of Proteolytic Enzymes*. San Diego. USA: Academic Press.

Baran, A. (2013). Assessment of Zea Mays Sensitivity to Toxic Content of Zinc in Soil. *Polish Journal of Environmental Studies*, 22(1):77–83.

Battisti, D. S., y Naylor, R. L. (2009). Historical Warnings of Future Food Insecurity with Unprecedented Seasonal Heat. *Science*, 323(5911), 240–244.

Benintende, M. C., J. J. De Battista, S. M. Benintende, M. F. Saluzzio, C. Muller y Sterren, M. A. (2008). Estimación del aporte de nitrógeno del suelo para la fertilización racional de cultivos. *Ciencia, Docencia y Tecnología* 19: 141-174.

Burke, E. J., Brown, S. J., y Christidis, N. (2006). Modeling the Recent Evolution of Global Drought and Projections for the Twenty-First Century with the Hadley Centre Climate Model. *Journal of Hydrometeorology*, 7(5), 1113–1125.

Bush, W., y Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *PLoS Computational Biology*, 8(12).

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., y Buckler, E. S. (2007). TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. *Bioinformatics* 23(19):2633–35.

Cairns, J. E., Crossa, J., Zaidi, P. H., Grudloyma, P., Sanchez, C., Luis Araus, J., y Atlin, G. N. (2013). Identification of drought, heat, and combined drought and heat tolerant donors in maize. *Crop Science*, 53(4), 1335–1346.

Cao, S., Loladze, A., Yuanj, Y., Wu, Y., Zhang, A., Chen, J., Huestis, G., Cao, J., Chaikam, V., Olsen, M., Prasanna, B., San Vicente, F., Zhang, X. (2017). Genome-Wide Analysis of Tar Spot Complex Resistance in Maize Using Geno-by- Sequencing SNPs and Whole-Genome Prediction. *THE PLANT GENOME*, 10: 1-14.

Cardon, L. R., y Palmer, L. J. (2003). Population stratification and spurious allelic association. *The lancet*, 361(9357), 598–604.

Clancy, S. (2008). Genetic recombination. *Nature Education* 1(1): 40.

CIMMYT applied molecular genetics laboratory (2003) Laboratory Protocols, 3rd ed. pp 7–11. CIM-MYT, Mexico, D.F., Mexico.

Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., y Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, 142(1–2).

Collins, G. N. (1914). Pueblo indian maize breeding. *Journal of Heredity*, 5(6), 255–268.

Cook, B. I., y Seager, R. (2013). The response of the North American Monsoon to increased greenhouse gas forcing. *Journal of Geophysical Research Atmospheres*, 118(4), 1690–1699.

Chen, Z., Wang, B., Dong, X., Liu, H., Ren, L., Chen, J., y Lai, J. (2014). An ultra-high density bin-map for rapid QTL mapping for tassel and ear architecture in a large F₂ maize population. *BMC Genomics*, 15(1).

Cristóbal-Acevedo, D., Álvarez-Sánchez, M. E., Hernández-Acosta, E., y Améndola-Massiotti, R. (2011). Concentración de nitrógeno en suelo por efecto de manejo orgánico y convencional. *Terra Latinoamericana*, 29(3), 325-332.

Crossa, J., Beyene, Y., Kassa, S., Pérez, P., Hickey, J. M., Chen, C., ... y Babu, R. (2013). Genomic prediction in maize breeding populations with genotyping-by-sequencing. *G3: Genes/ Genomes/ Genetics*, 3(11), 1903–26.

Cruz Sánchez A. (2018). La elongación del mesocotilo en respuesta a la siembra profunda en los géneros *Zea* (razas de maíz y teocintles) y *Tripsacum* (Poacea). Tesis de pregrado. Universidad Autónoma Benito Juárez de Oaxaca.

Dawson, E., Abecasis, R., Bumpstead, S., Chen, Y., Hunt, S., Beare, D. M., y Bentley, D. R. (2002). A first-generation linkage disequilibrium map of human chromosome. *Nature*, 22, 418, 544–548.

Deschamps, S., Llaca, V., y May, G. D. (2012). Genotyping-by-Sequencing in Plants. *Biology*, 1(3), 460–483.

Dracup, J. A., Lee, K. S., y Paulson, E. G. (1980). On the definition of droughts. *Water Resources Research*, 16(2), 297–302.

Dong, Z., Jiang, C., Chen, X., Zhang, T., Ding, L., Song, W., y Jin, W. (2013). Maize LAZY1 Mediates Shoot Gravitropism and Inflorescence Development through Regulating Auxin Transport, Auxin Signaling, and Light Response. *Plant Physiology*, 163(3), 1306–1322.

Ehret, G. B. (2010). Genome-Wide Association Studies: Contribution of Genomics to Understanding Blood Pressure and Essential Hypertension. *Current Hypertension Reports*, 12(1), 17–25.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., y Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, 6(5), 1–10.

Embarcadero, A. G. (1971). *Estudio del mesocotilo como característica en la clasificación biosistemática de maíz*. Tesis de pregrado. Universidad Nacional Autónoma de México, México.

Fellner, M., Ford, E. D., y Volkenburgh E. (2006). Development of Erect Leaves in a Modern Maize Hybrid is Associated with Reduced Responsiveness to Auxin and Light of Young Seedlings In Vitro. *Plant Signaling & Behavior*, 1(4), 201–211.

Flint-Garcia, Sherry A., Jeffrey M. Thornsberry, y Buckler E. S. (2003). Structure of Linkage Disequilibrium in Plants. *Annual Review of Plant Biology* 54(1):357–74.

Fleurdépine, S., Deragon, J., Devic, M., Guilleminot, J., y Bousquet-Antonelli, C. (2007). A Bona Fide La Protein Is Required for Embryogenesis in Arabidopsis Thaliana. *Nucleic Acids Research*, 35(10):3306–21.

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., y Buckler, E. S. (2014). TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*, 9(2).

Giles, L. Ruppelt, C., Yang, J., Mendel, R. R., Bittner, F., y Kirk, M. L. (2014). Molybdenum Site Structure of MOSC Family Proteins. *Inorganic chemistry* 53(18):9460–9462.

Goodstein, M. D., Shu, S. Howson, R., Neupane, R., Hayes, R. D., Fazo J., Mitros T., Dirks, W., Hellsten, U., Putnam, N., y Rokhsar, D. S. (2012) Phytozome: a comparative platform for green plant genomics, *Nucleic Acids Research*, 40, (D1) D1178–D1186.

Gowda, M., Das, B., Makumbi, D., Babu, R., Semagn, K., Mahuku, G., y Prasanna, M. B. (2015). Genome-wide association and genomic prediction of resistance to maize lethal necrosis disease in tropical maize germplasm. *Theoretical and Applied Genetics*, 128(10), 1957–1968.

Hayward, H. E. (1938). *The Structure of Economic Plants*. New York, USA: The Macmillan Company.

Helentjaris, T. (1987). A genetic linkage map for maize based on RFLPs. *Trends in genetics*, 3: 217-221.

Hernandez, A., Lopez-Luna J, y Gonzalez, E. (2011). Heavy Metal Quantification in Community Soils Impacted by Mining Activities in the Northern Mountains of Oaxaca, Mexico. *Environmental Science*, 4(11):4650.

Hoffman, G. E. (2013). Correcting for Population Structure and Kinship Using the Linear Mixed Model: Theory and Extensions. *PLoS ONE*, 8(10).

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scan J Stat.* 6:65-70.

Hoshikawa, K. (1969). Underground Organs of the Seedlings and the Systematics of Gramineae. *Botanical Gazette*, 130(3):192–203.

- Hron, K., Templ, M., y Filzmoser, P. (2010). Imputation of missing values for compositional data using classical and robust methods. *Computational Statistics and Data Analysis*, 54(12), 3095–3107.
- Huang, X., Feng, Q., Qian, Q., Zhao, Q., Wang, L., Wang, A., y Han, B. (2009). High-throughput genotyping by whole-genome resequencing. *Genome Research*, 1068–1076.
- Hufford, M. B., Xu, X., van Herwaarden, J., Pyhajarvi, T., Chia, J. M., *et al.* (2012). Comparative population genomics of maize domestication and improvement. *Nature Genetics*, 44, 808-811.
- Jung, M., Ching, A., Bhatramakki, D., Dolan, M., Tingey, S., Morgante, M., y Rafalski, A. (2004). Linkage disequilibrium and sequence diversity in 500-kpb region around the *adh1* locus in elite maize germplasm. *Theoretical and Applied Genetics*, 109: 681-689.
- Kato, F., Araki, M., Miyazawa, Y., Fujii, N., Takeda, K., Suge, H., y Takahashi, H. (2011). Factors responsible for Deep -sowing tolerance in wheat seedlings: varietal differences in cell proliferation and the co-ordinated synchronization of epidermal cell expansion and cortical cell division for the gibberellin-mediated elongation of first internodes. *Annals of Botany*, 108: 439-447.
- Kaiser, B. N., Kate L., Ngaire B, Phillips, T, y Tyerman, S. D. (2005). The Role of Molybdenum in Agricultural Plant Production. *Annals of botany* 96(5):745–54.
- Kaley, K., Carlisle, J., Siegel, D., Salinas, J. (2006). *Health concerns and environmental issues with PVC-containing building materials in green buildings. In: Contractor's report to the integrated waste management board.* Sacramento. California.
- Korte, A., y Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, 9(1), 29.

Kumar, P., Al-Shafai, M., Al-Muftah *et al.*, (2014). Evaluation of SNPs calling using single and multiple-sample calling algorithms by validation against array base genotyping and Mendelian inheritance. *BMC Research Notes*, 7, 1-13.

Lawrence, C. J., Dong, Q., Polacco, M. L., y Seigfried, T. E. (2004). MaizeGDB, the community database for maize genetics and genomics. *Nucleic Acids Research*, 32(10).

Lee, M., Sharopova, N., Beavis, W., Grant, David., Katt, M., Blair, D. y Hallauer A. (2002). Expanding the genetic map of maize with the intermediated B73 x Mo17 (*IBM*) population. *Plant Molecular Biology*, 48: 453-461.

Li, C., Li, Y., Badbury, P., Wu, X., Sni, Y., Song, Y., Zhang, D., Rodgers-Melnick, E., Buckler, E., Zhang, Z., Li, Y., y Wang, T. (2015). Construction of High-Quality Recombination Maps with Low-Coverage Genomic Sequencing for Joint Linkage Analysis in Maize. *BMC Biology* 13(1):78.

Li, X., Li, L. y Yan J. (2014). Dissecting meiotic recombination based on tetrad analysis by single-microspore sequencing in maize. *Nature communications*. 6, 1-9.

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., y Huang, S. (2014). Genomic analyses provide insights into the history of tomato breeding. *Nature Genetics*, 46(11), 1220–1226.

Lino, M., y Carr, D. (1982). Sources of Free IAA in the Mesocotyl of Etiolated Maize Seedlings. *Plant physiology*, 69(5):1109–12.

Lipka, A., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P., Gore, A., Buckler, E. y Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics*. 28: 2397-2399.

Liu, H., Niu, Y., Gonzalez-Portilla, P. J., Zhou, H., Wang, L., Zuo, T., y Pan, G. (2015). An ultra-high-density map as a community resource for discerning the genetic basis of quantitative traits in maize. *BMC Genomics*, 16(1), 1078.

Lobell, D. y Field C. (2007). Global scale climate-crop yield relationships and the impacts of recent warming, *Environmental Research Letters*, 2 (1).

Ludewig, U., Neuha, B., y Dynowski, M. (2007). Molecular Mechanisms of Ammonium Transport and Accumulation in Plants. *FEBS Letters*, 581(12):2301–8.

Maiti, R. (2017). *Horizon of field crops*. Salt Lake City. USA: American Academic Press.

Mangelsdorf, P. C., y Reeves, R. G. (1938). The Origin of Maize. *Proceedings of the National Academy of Sciences of the United States of America*, 24(8), 303–312.

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

Mao, H., Yu, L., Li, Z., Yan, Y., Ham, R., Liu, H., y Ma, M. (2016). Genome-Wide Analysis of the SPL Family Transcription Factors and Their Responses to Abiotic Stresses in Maize. *Plant Gene*, 6:1–12.

Mascher, M., Wu, S., St. Amand, P., Stein, N., y Poland, J. (2013). Application of Genotyping-by-Sequencing on Semiconductor Sequencing Platforms: A Comparison of Genetic and Reference-Based Marker Ordering in Barley. *PLoS ONE*, 8(10), 1–12.

Marino, G., y Funk, C. (2012). Matrix Metalloproteinases in Plants: A Brief Overview. *Physiologia Plantarum*, 145(1):196–202.

Martínez-Nava, D. A. (2018). Comparación morfológica del crecimiento de plántulas de maíces nativos en función de la profundidad de siembra. Tesis de pregrado. Universidad Nacional Autónoma de México, México.

Mikel, M. A., y Dudley, J. W. (2006). Evolution of North American dent corn from public to proprietary germplasm. *Crop Science*, 46(3), 1193–1205.

Meng, L., Huihui L., Luyan Z., y Jiankang, W. (2015). QTL IciMapping: Integrated Software for Genetic Linkage Map Construction and Quantitative Trait Locus Mapping in Biparental Populations. *The Crop Journal*, 3(3):269–83.

Mohan, M., Nair, S., Bhagwat, a, Krishna, T. G., Yano, M., Bhatia, C. R., y Takuji, S. (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding*, 3, 87–103.

Muday, G. K., y DeLong, A. (2001). Polar auxin transport: Controlling where and how much. *Trends in Plant Science*, 6(11), 535–542.

Muñoz, A. O. (1998). Diversidad del maíz en los nichos ecológicos y culturales de México. *Lecturas En Etnobotánica*, 5, 283–296.

Muñoz, A. O., Pérez, G. J., López, P. A., y Salvador, R. J. (2002). Maíz de cajete: agrosistema y resistencia a sequía. In J. P. Viqueira (Ed.), *Antología sobre pequeño riego*, (Vol. III Sistem, pp. 137–161). Colegio de Posgraduados, México: Editores Plaza y Valdez.

Muruganujan, A., Casagrande, J. T., y Thomas, P. D. (2013). Large-scale gene function analysis whit the PANTHER classification system. *Nature protocols*, 8, 1551-1566.

Ndhela, T., Herselman, L., Semagn, K., Magorokosho, C., Mutimaamba, C., y Labuschagne, M. T. (2015). Relationships between heterosis, genetic distances and specific combining ability among CIMMYT and Zimbabwe developed maize inbred lines under stress and optimal conditions. *Euphytica*, 204(3), 635–647.

Nemri, A., Atwell, S., Tarone, A., Huang, Y. S., Zhao, K., Studholme D. J., Nordborg, M., y Jones, J. D. G. (2010). Genome-Wide Survey of Arabidopsis Natural Variation in Downy Mildew Resistance Using Combined Association and Linkage Mapping. *PNAS*, 107(22).

Nieto-Sotelo J., Cassab G.I., Eapen D. *et al.* 2012. Informe Técnico Final del proyecto IMIC-2011-008: "Descripción de caracteres hidrotropicos y del desarrollo y crecimiento del mesocotilo y de las raíces adventicias en razas y líneas de maíz.

Nieto Sotelo J., Cassab G.I., Eapen D. *et al.* 2013. Informe Técnico Final del proyecto MASAGRO IMIC-2012-035: "Descripción de caracteres hidrotropicos y del desarrollo y crecimiento del mesocotilo y de las raíces adventicias en razas y líneas de maíz.

Nieto-Sotelo J., Martínez Nava D.A., Cruz Sánchez C.A. y L. Vazquez Marcial. 2018a. Prácticas agrícolas tradicionales y la conservación de maíces nativos en México. *Sistemas Agroalimentarios Localizados y Prácticas agrícolas tradicionales Hacia una propuesta de política pública para el desarrollo rural*. Ciudad de México, México. Yod Estudio.

Nieto-Sotelo J, Vázquez-Marcial L, Vallejo-Reyna MA, Figueroa-Maya A, Martínez A, Cruz A, Alonso R, Villa JM, Ávila AX, Rojas CI, Aguilar C, Rangel LM, Zhang X, Sawers R, Cassab GI. 2018b. Analyses of genetic variation associated to deep planting resistance in maize revealed genes controlling development, growth, and adaptation to soil conditions. *60th Annual Maize Genetics Conference*, del 22 al 25 de marzo de 2018. Saint Malo, Francia.

Overbeek, V. (1936). Growth Hormone and Mesocotyl Growth. *Plant Physiology* (33):333–40.

Pachauri, R. K., y Meyers, L. A. (eds). (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: IPCC.

Parry, M. L., Rosenzweig, C., Iglesias, A., Livermore, M., y Fischer, G. (2004). Effects of climate change on global food production under SRES emissions and socio-economic scenarios. *Global Environmental Change*, 14(1), 53–67.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal*, 316(7139), 1236–1238.

Piperno, D., Moreno, J., Iriarte, J., Holst, I., Lachinet, J., Jones, A., Ranere, A. y Castanzo, R. (2007). Late Pleistocene and Holocene environmental history of the iguala Valley Central Balsas Watershed of Mexico. *Proceedings of the National Academy of Sciences*, 104 (29) 11874-11881.

Pritchard, J. K., Stephens, M., Rosenberg, N. A., y Donnelly, P. (2000). Association mapping in structured populations. *American Journal of Human Genetics*, 67, 170–181.

Ranere, A., Piperno, D., Holst, I., Dickau, R. y Iriarte, J. (2009). The cultural and chronological Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. *Proceedings of the National Academy of Sciences of the United States of America*, 106(13) 5014-5018.

Rao, N. K. S., Shivashankara, K. S., y Laxman, R. H. (2016). *Abiotic Stress Physiology of Horticultural Crops*. New Delhi: Springer India.

Ratnaparkhe, S. M., Ulrika E., y Flinn, B. S. (2009). Identification and Characterization of a Matrix Metalloproteinase (Pta1-MMP) Expressed during Loblolly Pine (*Pinus Taeda*) Seed Development, Germination Completion, and Early Seedling Establishment. *Planta*, 230(2):339–54.

Rasmus, N., Slatkin, M. An introduction to population genetics theory and applications. Massachusetts U.S.A. Sinauer Associates, Inc Publishers Sunderland.

Rish, N. (1992). Genetic Linkage: Interpreting Lod Scores. *Science*, 225(7), 803–804.

Rosenzweig, C., y Parry, M. L. (1994). Potential impact of climate change on world food supply. *Nature*, 367(6459), 133–138.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., y Rice, P. (2000). Artemis: sequence visualization and annotation. *Bioinformatics* 16(10), 944–945.

Saab, I. N., Sharp, R. E., y Pritchard, J. (1992). Effect of inhibition of abscisic Acid accumulation on the spatial distribution of elongation in the primary root and mesocotyl of maize at low water potentials. *Plant Physiology*, 99(11), 26–33.

Saitou, N., y Nei, M. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), 406–425.

Sánchez, J., Goodman, M., y Rawlings, J. (1993). Appropriated Characters for Racial Classification in Maize. *Economic Botany*, 47, 44-55.

Sauter, M. (2015). Phytosulfokine peptide signaling. *Journal of Experimental Botany*, 66, 5161–5169.

Sawers, R. J. H., Linley, P., Farmer, R., Hanley, N., Costich, D. E., Terry, M., y Brutnell, T. (2002). Elongated mesocotyl1, a Phytochrome-Deficient Mutant of Maize. *PLANT PHYSIOLOGY*, 130(1), 155–163.

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., R. K. Wilson. Et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science*, 326(5956), 1112–5.

Schweder, T y Spjotvoll, E. (1982). Plots of *P*- values to evaluate many tests simultaneously. *Biometrika*, 63 (3), 493-502.

Sekhon, R. S., Lin, H., Childs, K. L., Hansey, C. N., Robin Buell, C., De Leon, N., y Kaeppler, S. M. (2011). Genome-wide atlas of transcription during maize development. *Plant Journal*, 66(4), 553–563.

Semagn, K., Bjørnstad, Å., y Ndjiondjop, M. N. (2006). Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*, 5 (25) (December), 2569–2587.

Sharma, M., y Pandey, G. (2016). Expansion and Function of Repeat Domain Proteins During Stress and Development in Plants. *Frontiers in Plant Science*, 6:1–15.

Sharopova, N., McMullen, M.D., Schultz, L. *et al.* (2002). Development and mapping of SSR markers for maize. *Plant Mol Biol* (2002) 48: 463.

Shoji, S., Masami N., R. A. Dahlgren., y Quantin, P. (1996). Evaluation and Proposed Revisions of Criteria for Andosols in the World Reference Base for Soil Resources. *Soil Science* 161(9):604–15.

Slatkin, M. (2008). Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, 9: 477-485.

Schnable, J., Springer, N. y Freeling, M. (2010). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *PNAS*. 108: 4069-4074.

Song, M., Wei, H., y Storey, J. D. (2015). Testing for Genetic Associations in Arbitrarily Structured Populations. *Nature Genetics*, 47(5):550–54

Sosnowski, O., Charcosset, A., y Joets, J. (2012). Biomecator V3: An upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. *Bioinformatics*, 28(15), 2082–2083.

Soto-Cerda, B. J., y Cloutier, S. (2012). Association mapping in plant genomes. *Genetic Diversity in Plants*, 29–54.

Stoker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex y Midgley, P. M. (eds). (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. New York, USA: Cambridge University Press.

Stelpflug, S., Sekhon, R., Vaillancourt, B., Hirsch, C., Buell, C., De Leon, N., Kaeppler, S., (2015) An expanded maize gene expression atlas based on RNA-sequencing and its use to explore root development. *Plant Genome*, 1-16.

Takahashi, T. y Shoji, S. (2002). Distribution and Classification of Volcanic Ash Soils. *Global Environ. Res*, 6(2):83–97.

Tang, A. y Boyer, J. (2002). Growth-induced water potentials and the growth of maize leaves. *Journal of Experimental Botany*, 53: 489-503.

Tanksley, S. D. y Nelson, J. C. (1996). Advanced Backcross QTL Analysis: A Method for the Simultaneous Discovery and Transfer of Valuable QTLs from Unadapted Germplasm into Elite Breeding Lines. *Theoretical and Applied Genetics*, 191–203.

Thornsberry, J. M., Goodman, M. M., Doebley, J., Kresovich, S., Nielsen, D., y Buckler, E. S. (2001). Dwarf8 polymorphisms associate with variation in flowering time. *Nature Genetics*, 28(3), 286–289.

Tian, F., Bradbury, P. J., Brown, P. J., Hung, H., Sun, Q., Flint-Garcia, S., Torbert R Rocheford James B Holland y Buckler, E. S. (2010). Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nature Genetics*, 43(2), 159–164.

Trachsel, S., Sun, D., Sanvicente, F. M., Zheng, H., Atlin, G. N., Suarez, E. A., y Zhang, X. (2016). Identification of QTL for early vigor and stay-green conferring tolerance to drought in two connected advanced backcross populations in tropical maize (*Zea mays* L.). *PLoS ONE*, 11(3), 1–22.

Troyer, A. F. (1997). The location of genes governing long first internode of corn. *Genetics*, 145(4), 1149–1154.

Vargas, M. (1966). *Influencia de La Longitud de Mesocotilo Coleoptilo En La Emergencia de Plantulas En Maiz*. Tesis de pregrado. Chapingo, México.

- Varshney, R. K., Singh, V. K., Hickey, J. M., Xun, X., Marshall, D. F., Wang, J., y Ribaut, J. M. (2015). Analytical and Decision Support Tools for Genomics-Assisted Breeding. *Trends in Plant Science*, 21, 1–10.
- Vielle-Calzada, J. P. Martínez, O., Hernández-Guzmán, G., et al. (2009). The Palomero Genome Suggests Metal Effects on Domestication. *Science*, 326(5956):1078–1078.
- Walton, J. D., y Ray, P. M. (1981). Evidence for Receptor Function of Auxin Binding Sites in Maize. *Plant Physiology*, 68(6): 1334–1338.
- Wellhausen, E., Roberts, L., Hernández X. E., y Mangelsdorf, P. (1951). Razas de Maíz en México, su origen, características y distribución. *Folleto Técnico*, (5), 237.
- Williams, R. (1989). Restriction fragment length polymorphism (RFLP). *America Journal of Physical Anthropology*, 32: 159-184.
- Yu, H., Xie, W., Wang, J., Xing, Y., Xu, C., Li, X., y Zhang, Q. (2011). Gains in QTL detection using an ultra-high-density SNP map based on population sequencing relative to traditional RFLP/SSR markers. *PLoS ONE*, 6(3): e17595.
- Xie, W., Feng, Q., Yu, H., Huang, X., Zhao, Q., Xing, Y., Yu, S., Bin, H., y Zhang, Q. (2010). Parent-Independent Genotyping for Constructing an Ultrahigh-Density Linkage Map Based on Population Sequencing. *Proceedings of the National Academy of Sciences*, 107(23):10578–83.
- Zhang, M., An, L., Feng, H., Chen, T., Chen, K., Liu, Y., Tang, H., Chang, J., y Wang, X. (2007). The Cascade Mechanisms of Nitric Oxide as a Second Messenger of Ultraviolet B in Inhibiting Mesocotyl Elongations¶. *Photochemistry and Photobiology*, 77(2):219–25.

Zhang, H., Ma, P., Zhao, Z., Zhao, G., Tian, B., Wang, J., y Wang, G. (2012). Mapping QTL controlling maize deep-seeding tolerance-related traits and confirmation of a major QTL for mesocotyl length. *Theoretical and Applied Genetics*, 124(1), 223–232.

Zhang, Z., Ersoz, E., Lai, C., Todhunter, R., Tiwari, H., Bradbury, P. J., Yu, J., Arnett, D. Ordovas, y Buckler. E. S. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics*, 42(4), 355–60.

Zhou, X., Carbonetto, P., Stephens, M., Afterwards, X., Datasets, G., Data, M., Buckler, E. S. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics*, 42(4), 355–60.

Zila, C. T., Ogut, F., Romay, M. C., Gardner, C., Buckler, E. S., y Holland, J. B. (2014). Genome-wide association study of Fusarium ear rot disease in the U.S.A. maize inbred line collection. *BMC Plant Biology*, 14(1), 372.

Anexo I (inicio). Promedios de la elongación del mesocotilo de la colección DTMA; DTMA) nombre de la accesión; Prom) promedio de la accesión; D.S) desviación estándar y n) número de plántulas medidas por accesión.

DTMA	Prom	D.S.	n	DTMA	Prom	D.S.	n	DTMA	Prom	D.S.	n
DTMA 1	8.44	2.36	44	DTMA 103	8.07	1.52	45	DTMA 199	7.47	3.82	45
DTMA 2	4.60	2.44	41	DTMA 104	5.94	1.36	43	DTMA 200	8.14	1.92	45
DTMA 3	7.47	1.42	24	DTMA 105	10.45	2.23	45	DTMA 201	6.79	1.79	41
DTMA 5	7.01	2.29	44	DTMA 106	6.36	2.43	42	DTMA 202	5.97	2.53	42
DTMA 6	8.34	1.40	45	DTMA 107	10.16	2.67	44	DTMA 203	4.08	1.82	24
DTMA 7	6.66	1.70	44	DTMA 108	5.66	1.99	32	DTMA 204	5.85	1.87	46
DTMA 8	6.79	1.35	45	DTMA 109	7.89	1.50	43	DTMA 205	10.21	2.68	43
DTMA 9	7.55	1.63	45	DTMA 110	9.16	2.42	45	DTMA 206	9.24	2.09	20
DTMA 10	7.18	1.66	45	DTMA 111	5.65	2.09	44	DTMA 207	5.76	0.88	24
DTMA 11	7.55	1.36	44	DTMA 112	9.05	2.27	44	DTMA 208	8.20	2.99	42
DTMA 12	7.20	1.61	44	DTMA 113	10.20	2.61	43	DTMA 209	8.53	2.50	41
DTMA 13	7.22	1.51	45	DTMA 114	8.95	2.46	37	DTMA 210	6.35	2.47	45
DTMA 14	7.43	2.36	44	DTMA 115	9.86	1.52	24	DTMA 211	8.01	2.08	42
DTMA 15	6.86	2.00	43	DTMA 116	8.11	1.91	45	DTMA 212	6.30	2.46	42
DTMA 17	6.80	2.10	44	DTMA 117	8.76	1.43	23	DTMA 213	6.35	1.23	43
DTMA 18	7.08	1.73	44	DTMA 118	8.10	1.96	33	DTMA 214	6.38	1.33	34
DTMA 19	5.93	1.98	43	DTMA 119	9.02	2.66	43	DTMA 215	5.00	1.91	43
DTMA 20	9.17	3.36	41	DTMA 120	7.06	2.33	44	DTMA 216	7.61	1.13	24
DTMA 21	8.08	1.96	39	DTMA 121	8.07	1.93	44	DTMA 217	8.22	1.54	21
DTMA 22	6.61	1.72	44	DTMA 122	7.51	1.81	45	DTMA 218	9.39	1.74	45
DTMA 24	6.42	1.39	44	DTMA 123	9.62	2.70	37	DTMA 219	7.42	1.55	42
DTMA 25	5.35	1.12	45	DTMA 124	7.65	2.26	43	DTMA 220	4.74	2.25	44
DTMA 26	4.74	1.69	43	DTMA 125	7.67	1.94	42	DTMA 221	7.82	2.67	44
DTMA 27	9.84	2.12	23	DTMA 126	4.83	1.75	39	DTMA 222	4.84	2.22	34
DTMA 28	5.55	1.23	41	DTMA 127	7.44	1.85	44	DTMA 223	4.91	2.15	43
DTMA 29	8.27	2.63	45	DTMA 128	6.26	2.28	42	DTMA 224	9.74	2.39	45
DTMA 31	5.36	1.59	23	DTMA 129	6.11	2.46	41	DTMA 225	10.00	1.77	23
DTMA 32	5.52	1.37	21	DTMA 131	7.45	2.17	39	DTMA 226	7.31	2.58	36
DTMA 33	6.80	1.53	44	DTMA 132	5.19	2.67	40	DTMA 227	6.60	2.60	44
DTMA 34	6.72	1.38	39	DTMA 133	7.46	3.18	41	DTMA 228	11.36	2.23	45
DTMA 35	4.19	1.66	39	DTMA 134	8.06	2.37	43	DTMA 229	6.92	1.22	45
DTMA 36	7.16	1.72	45	DTMA 135	5.44	2.00	42	DTMA 230	5.59	1.48	42
DTMA 37	5.51	1.60	43	DTMA 136	7.31	2.13	44	DTMA 231	9.96	1.61	44
DTMA 38	7.24	2.12	44	DTMA 137	7.78	2.92	45	DTMA 232	9.19	1.66	44
DTMA 39	6.49	1.55	43	DTMA 138	8.95	2.99	40	DTMA 233	9.48	1.17	24
DTMA 40	7.12	2.03	45	DTMA 139	8.21	3.77	41	DTMA 234	6.57	2.57	45
DTMA 41	5.60	1.58	42	DTMA 140	7.44	3.28	41	DTMA 235	7.52	1.25	24
DTMA 42	5.24	1.30	42	DTMA 141	9.51	2.06	23	DTMA 236	6.88	1.38	42
DTMA 43	6.36	1.73	39	DTMA 142	6.61	2.26	45	DTMA 237	6.76	1.09	23

Anexo I (continuación)

DTMA	Prom	D.S.	n	DTMA	Prom	D.S.	n	DTMA	Prom	D.S.	n
DTMA 44	4.87	1.59	42	DTMA 143	8.75	2.90	44	DTMA 238	6.98	1.51	24
DTMA 45	8.03	2.09	24	DTMA 144	9.25	2.01	21	DTMA 239	8.45	1.23	19
DTMA 47	6.50	1.97	42	DTMA 145	10.07	1.90	23	DTMA 240	7.19	0.61	24
DTMA 48	8.29	1.62	45	DTMA 146	6.62	1.66	45	DTMA 241	7.93	2.05	45
DTMA 49	8.24	2.02	41	DTMA 147	6.67	1.45	39	DTMA 242	6.58	1.74	42
DTMA 50	6.58	2.32	36	DTMA 148	7.14	1.68	24	DTMA 243	5.95	1.39	23
DTMA 51	6.49	2.51	37	DTMA 149	6.00	1.74	43	DTMA 244	3.91	1.52	36
DTMA 52	7.40	1.53	44	DTMA 150	7.67	1.41	24	DTMA 245	4.60	2.51	41
DTMA 53	3.88	1.67	36	DTMA 151	5.99	3.26	40	DTMA 246	6.81	2.86	35
DTMA 54	6.45	1.28	22	DTMA 152	7.68	1.63	45	DTMA 247	7.24	2.46	38
DTMA 55	5.41	1.18	43	DTMA 153	5.26	2.28	44	DTMA 248	8.44	1.34	23
DTMA 56	7.51	1.39	43	DTMA 154	6.05	4.84	42	DTMA 249	5.21	2.28	33
DTMA 57	7.04	2.12	39	DTMA 155	6.85	4.61	44	DTMA 250	4.16	1.71	43
DTMA 58	6.70	2.00	34	DTMA 156	5.75	3.13	45	DTMA 251	8.29	2.47	24
DTMA 59	6.37	2.14	42	DTMA 157	7.37	2.34	45	DTMA 252	4.32	1.83	44
DTMA 60	5.84	2.19	38	DTMA 158	6.05	1.11	45	DTMA 253	5.55	1.62	45
DTMA 61	8.51	1.60	23	DTMA 159	8.43	2.39	39	DTMA 254	6.90	3.19	34
DTMA 62	9.45	2.16	24	DTMA 160	5.86	1.85	39	DTMA 255	6.47	2.31	40
DTMA 63	7.54	2.10	43	DTMA 161	7.17	2.29	37	DTMA 256	9.35	0.97	21
DTMA 64	4.28	1.26	41	DTMA 162	7.75	2.40	45	DTMA 257	9.23	1.39	20
DTMA 65	7.53	1.94	44	DTMA 163	6.23	2.72	44	DTMA 258	6.61	3.26	37
DTMA 66	5.60	1.61	42	DTMA 164	8.93	1.58	21	DTMA 259	6.99	1.71	45
DTMA 67	5.94	1.32	35	DTMA 165	7.12	1.24	23	DTMA 260	7.58	3.27	45
DTMA 68	7.03	1.51	39	DTMA 166	8.20	1.82	23	DTMA 261	9.03	2.05	23
DTMA 69	6.53	1.18	45	DTMA 167	5.94	3.05	38	DTMA 262	5.70	2.20	32
DTMA 70	6.64	1.31	45	DTMA 168	4.83	1.20	44	DTMA 263	8.65	1.65	23
DTMA 71	7.10	1.72	44	DTMA 169	5.45	1.22	44	DTMA 264	5.71	2.68	42
DTMA 72	6.89	1.54	44	DTMA 170	7.89	2.21	40	DTMA 265	6.95	3.11	41
DTMA 73	9.17	2.01	23	DTMA 171	6.03	1.33	44	DTMA 266	5.78	1.80	44
DTMA 74	6.61	1.42	42	DTMA 172	6.78	2.58	37	DTMA 267	8.56	1.39	24
DTMA 75	7.11	2.00	22	DTMA 173	6.20	1.73	41	DTMA 268	5.57	2.20	43
DTMA 77	5.55	1.82	40	DTMA 174	6.92	1.72	17	DTMA 269	3.83	2.43	37
DTMA 78	5.86	2.12	42	DTMA 175	7.21	2.03	39	DTMA 270	6.35	1.46	30
DTMA 79	6.91	1.67	13	DTMA 176	6.81	1.18	24	DTMA 271	9.90	1.47	22
DTMA 80	6.93	1.77	43	DTMA 177	5.73	2.58	41	DTMA 275	9.57	1.24	20
DTMA 81	5.50	1.98	41	DTMA 178	5.02	1.21	44	DTMA 276	5.43	2.32	41
DTMA 82	7.43	1.85	43	DTMA 179	9.97	3.15	43	DTMA 280	8.16	2.33	42
DTMA 83	9.43	1.77	44	DTMA 180	9.66	2.75	42	DTMA 281	7.35	2.34	44
DTMA 84	7.60	1.00	44	DTMA 181	6.58	1.94	45	DTMA 282	4.99	1.92	39
DTMA 85	7.12	1.68	45	DTMA 182	7.55	1.27	22	DTMA 283	8.61	1.05	23
DTMA 86	5.20	1.43	43	DTMA 183	4.28	1.13	41	DTMA 284	7.30	2.04	45
DTMA 87	6.19	1.73	42	DTMA 184	4.98	1.31	39	DTMA 285	9.18	2.27	21

Anexo I (término)

DTMA	Prom	D.S.	n	DTMA	Prom	D.S.	n	DTMA	Prom	D.S.	n
DTMA 88	8.00	1.27	34	DTMA 185	7.08	3.35	42	DTMA 286	5.46	2.60	42
DTMA 89	4.37	1.50	43	DTMA 186	5.44	2.18	41	DTMA 287	5.42	2.29	42
DTMA 90	6.07	0.96	24	DTMA 187	5.78	2.65	40	DTMA 289	5.92	2.99	45
DTMA 91	8.24	2.11	44	DTMA 188	6.34	2.72	36	DTMA 291	5.09	2.42	40
DTMA 92	6.99	1.35	44	DTMA 189	8.33	4.14	45	DTMA 292	6.72	3.62	45
DTMA 93	7.51	2.09	44	DTMA 190	9.41	2.21	22	DTMA 293	6.58	2.40	44
DTMA 94	8.48	1.27	24	DTMA 191	6.30	1.55	42	DTMA 294	7.16	3.60	44
DTMA 95	4.56	1.46	44	DTMA 192	5.09	2.65	42	DTMA 296	4.49	2.15	42
DTMA 96	6.45	1.53	40	DTMA 193	7.38	0.95	23	DTMA 297	7.51	2.20	45
DTMA 98	8.27	1.24	45	DTMA 194	7.57	0.97	24	DTMA 298	6.19	2.31	44
DTMA 99	6.62	1.65	45	DTMA 195	8.03	1.23	24	DTMA 299	4.24	1.36	45
DTMA 100	6.42	1.79	45	DTMA 196	5.85	2.90	45	DTMA 300	5.27	0.98	45
DTMA 101	8.99	1.19	24	DTMA 197	7.94	2.22	44				
DTMA 102	6.65	1.83	42	DTMA 198	6.06	2.00	45				

Anexo II. Secuencias *fasta* alrededor de los 63 SNPs asociados (50 bases río arriba, el SNP asociado y 50 bases río abajo) a la elongación del mesocotilo obtenidas del genoma de referencia (B73_v2.7).

>S7_162027150_B73V2.7_Chr7:162027100..162027200

TCCTCGATTGACCGCAGCCCTGCATGCAGGCGTTTCGATTCGTGGTGCTACCTAGACTCCAGAACCCTTCC
GAGAAGACCACTTCTCTTCCGCGCGTTTCGC

>S8_166105182_B73V2.7_Chr8:166105132..166105232

ACGTTGTTTAAACCAAACCCCTGATCGCTTATTCCCAACAACAACACCTCCCTCCAGACCCAGACACCGCAA
ACCCTGTGGTCAGACGTTCTGCGTGACGC

>S4_4737438_B73V2.7_Chr4:4737388..4737488

CGACGGTCGAGGTCACCAGCGCGCTGCTGCTGCTGCTGATCCCTGTCCCAACTCTCAGAATTATCAATGT
TGAAATAAAAAAATATATATCTTTTTTTCT

>S4_226459824_B73V2.7_Chr4:226459774..226459874

TGCAATAATATCTGGCTTCCAGGCAAGTGACTGTCCTGTGAGCGAGCCTACTCTGTGCTCAAGATTCAAGA
AGGAAGGCTAGCGGAAGGGTGCGAGTGTG

>S7_63900887_B73V2.7_Chr7:63900837..63900937

GTACACGAAGGCGAGGTCGATGGCGAACGCGACTACTATTGCCTACTGCGACACCACGTGGGTGAGCGTAGTGAGT
CGAGCACGTGACACGGAGGGGC

>S5_1852022_B73V2.7_Chr5:1851972..1852072

AGCACCCCAATTTTCAAGCTGAGCGACCTGCCACCATGCAAGACCGTGCACTGATCCTCGGTTTCGTCAT
CATCGATGACCCTTTTCCTCTTGTCAGTGT

>S8_170878813_B73V2.7_Chr8:170878763..170878863

CATTGTAGTTATAATCACGATTGCTGTTGCAGCTGGTATTGTTCCTAATACACCCGTTATAGGAATTGTT
ACCACTCCAAGCATTATTGTTGTTATAA

>S3_19260465_B73V2.7_Chr3:19260415..19260515

GTTCTCACCGGCGGGTGGCGGCGGGCCGAGAAGCAGCATCCCGTCCCCCGCCACGGAGTCGCCAGAGA
GCAGGAGCTGGTCCGCCGCCGCCACCGCCGCTG

>S7_99089950_B73V2.7_Chr7:99089900..99090000

CACGCGCAGGGCCTCCTCGGCGGTGCTGCAGAGCCTCCTCTCTTTCTCCGTCTTCTCCTTCCCGTCGACAT
GCCTGTCCCGTACCTCCAGCCCGACGCTG

>S2_198509433_B73V2.7_Chr2:198509383..198509483

AAATTAACAGGGCTGCTTGTGCATATTGTCTGAAAGTGTAGATAGTTTCACAGCTTTTTATAGTCCCATGTTCA
GAAGTCGAGATGCATTGCGAAATCGACA

>S3_146363360_B73V2.7_Chr3:146363310..146363410

GCGTTCTGTACTGCACATTTTTAAAGGAAACCAGGACAGGTATCTAACGCCCGTCATTGTTGTGCTCAGACC
TGGTACGTTTTGCAGCTACTCGAGTCATT

>S3_21368942_B73V2.7_Chr3:21368892..21368992

TCGGACACTTACTCTGCAGCAGTGGGACGTCGCTGGCGGCGTGGCCATGGACCATGAAGGGTTCGGCCTC
CTGGAGGCTGTGGCGGTGCTCGTACGCGTCC

>S6_8146003_B73V2.7_Chr6:8145953..8146053

GGGCAGCTCTACGGCTGCGGGCTGCGCGCCTACGGCCTACCTACCAGCGGCGCGTCGTCTTCGCGCGGA
GGTACCTAGCGGCGGCACGTCGTGCTACCTAA

>S2_217651275_B73V2.7_Chr2:217651225..217651325

AGCTTTTGAAAACATATTAGTCGAAGAGGTGTAGACCTGCAGCAAAGATGATGTCTTTTTCAACCATAACTA
AAACATTGCTCAAGCAAGCAAGGAACAT

>S10_91959067_B73V2.7_Chr10:91959017..91959117

ACCGCATCCTCAAGGTGCACAACGCCCGCGCACCGTCCGCCGCTTCGAGGCCTACCGCGACGCCGTGC
GCTCCCGCTGCCGCGCCGTCGTGCGCCGCCGG

>S7_165369648_B73V2.7_Chr7:165369598..165369698
GCACCTGGGATAAATTCTACGGGGTCTCGGTTTACTGGCCGGATCTGGCTATATCATTGGTCGGGATTTT
TTTTACAGCAGCAGAGAACGATGGGCGATG

>S10_5458621_B73V2.7_Chr10:5458571..5458671
GCTGCGGTTCGGTGAGCAACGCGGTGCCGGTGGTGGCGACGAGCGGCAGGCTGTCGTCGTCCTCCCTTCTT
CGGTCACTACAGCTCCAGCTCCAGGAGGCTCT

>S5_11923821_B73V2.7_Chr5:11923771..11923871
GGCGCGTCGCGACCCTGTCTGTCCACGCAACCATCTGCGACATCGGGGGCGTCTCAACGAACCGAGAGC
CTGCTCGGCGCAGCTGCCTGGGGCCTAGTGCC

>S9_15357629_B73V2.7_Chr9:15357579..15357679
CGTAGCATGGTAGCAGTAGTAGCACCTCGCTCGTCCCCTCCCCTGCCAGCCGAGGCACAGGCACGCGTGT
TGCTGTAGCTTTCACCCTGGAGGGTCACTCA

>S7_64048287_B73V2.7_Chr7:64048237..64048337
GGGCGCTAGGCGAGCGGACAGCGAGGGAGCTGCAGAGAGATAGACGCTGAGCGAGAGGAGCAAGGAAGAGAGAG
AGCACGACAGGGAGCAGGGAGCGGC

>S10_98272538_B73V2.7_Chr10:98272488..98272588
GGTTCTCTGACTGTATTAGACGGGCAATGCTGCAATCATTGTTGACCCATCAAGCATGCAAATAGTCGCAA
GGCCACGGACCAGACACACAAGCATGATA

>S4_174903108_B73V2.7_Chr4:174903058..174903158
AGCTCAGTTGCATCTTTTGCAGCATTAAACAGATAGTACCTGTAGGCGCAAGGTCAGGTCATTAATGCTAT
GTTGATTTGCGAAAAAAGTAAACTACAA

>S2_236282499_B73V2.7_Chr2:236282449..236282549
ACGCCGCCAGACGTCTCTCTGCTAGCGTCGGTCATCGGAGTGCACCTGCAGTGCCAGGTCGCGTTGCAGGATGAGGGCTC
CCTGCAGTACCTGACAGCCGG

>S2_204476331_B73V2.7_Chr2:204476281..204476381
TGCTTGGTACCATTACAAAATACAAATAGTTAGTTAATCATGCATTGTGATTTGTGAAGCACAAGAATCATCT
TTATCTGGAGGATGTGTTGGAGGTTTCT

>S8_131240915_B73V2.7_Chr8:131240865..131240965
AGGATCGGAGCATCAGCATCGCGACCTGTCTCTGCCGTGAACTTGCCTACCTGCTAAGTGCCAAGGCCAG
CAACACAAGGCAGCGAGCGACGGACATGGAC

>S8_131049015_B73V2.7_Chr8:131048965..131049065

AGAAGAGTAGAGCTGTACTACTGCTGTAGCAGGCAGGCAGGCAGGAAACCGCGGGGACGCGACGCCTGA
CGCTTGGCAGGTCGGCAGCCCAAGGGCAACGA

>S5_193547933_B73V2.7_Chr5:193547883..193547983

CCTCGGCTGCCGCCGATAAGGGAGGCCTGCCGTCTCTGTCCGACATGGTCGCGCTTCTCACTGGCAGGT
TAAGAAAAAATTCTCTTGTGTATAGCAG

>S5_183326603_B73V2.7_Chr5:183326553..183326653

TGAAGTAGAACTCGACCTGGGGCAACCGAATCAGATACGTACTTTTCAGAGGAATCGCATGGTGGCGGAAA
AAAAAAGCTGCGCATCTGGAGCAACGAGGC

>S10_127809396_B73V2.7_Chr10:127809346..127809446

CTGGTTCGTGCTTCGCTGCAGGTGGGTGCGGCAAAGATCAAGAGCAGGACGCCGGACGACCACTACTGCTGGAACCGGC
CGGAGGACATGGCGTACCCTCG

>S7_160595701_B73V2.7_Chr7:160595651..160595751

CTGGTCTGCTTGTTCATCGGTTATTATCACAGGCAATTCGCGATGCCAGTGCTACCGGTGCTGCATATT
CCTAGTAAAGAAAGAAAGAAAGAAAGAAA

>S10_10586901_B73V2.7_Chr10:10586851..10586951

GCCGCTGCCGGACGAGGGGAGCATGCCGCCCGCGGAGATCTGAGGAGAAGGACACACGAGTGTGCGGGTCCCAGAC
CTTGTTTCAGCGTTCGGGTTGTG

>S3_13726673_B73V2.7_Chr3:13726623..13726723

GTCGATCGATCGGGGACGCAAACGCGTGACGAATGGACGAACCCTAATAACGACTGAAATTAACCAAGAA
TGCTCCTGCTCCTGCACCGCACCGAACTG

>S10_140743010_B73V2.7_Chr10:140742960..140743060

GCAACTACCACAGGAACTTGGATTGGACCAACTTAACAAGAACGCGCATCCTAAAAAAAACACAGTAGTTC
TACTAGCACAGACGAGCAAACCGAAAGAG

>S10_4751416_B73V2.7_Chr10:4751366..4751466

ACCTTCGCCTCCCATCTGCGCTTCGTCCCACTATTTTACTTTCTGATGCATCATTGGCGCCACTTCCTCCCT
AGCTGCTGCATGGGCACTGAGCTGAGCTG

>S5_215289316_B73V2.7_Chr5:215289266..215289366

TTCGCGGCTCTCGCACCGCCGCCTCCTTCGGAGCTAAGGTTTCGTTGCTTCCCCACACCCACCAGCTCCTA
TGCCTGACGTGATGCCTCCTAGCCGTGATA

>S6_106176548_B73V2.7_Chr6:106176498..106176598
TAGGATCTTTTCTTTCTTTTGTGTTTCGTTTTCTGAATGTTGGTCTGGTTTTCTCGGAAGGTTTTGATTGCG
GCGACAGTGGCGAGCGCGATCGGGCAG

>S1_14103918_B73V2.7_Chr1:14103868..14103968
ACGGCGACGACCGACTCCC CGCGCTGCAGGCGACCACCGGACGCGGCGTCTCCTCCTCGGCGGCGAC
CGCCGGTGCTGCACTCGCGGTGCTGCTCGCCGG

>S6_71803064_B73V2.7_Chr6:71803014..71803114
GCTGAGGGTGGTTTACCCCTTGGGGAGGGGTAAGTGTGCTTATAACACCGAGGAAACCAAACGAGCAGCTATACCTTTTG
GAGAATCTTTGTAAAGGCTAC

>S9_139410281_B73V2.7_Chr9:139410231..139410331
GCGCGAGATTCCAACAACACAACGGGTGCCGCGTGCGATCGATGGCGACAGCGCAGCGCGCCCGGACCT
GTGCGCCCGTCCCGCACCCGTCTTTGTTTACT

>S8_167282998_B73V2.7_Chr8:167282948..167283048
CAGATTAGGTTTGGCTGGTGTGAGACTTAGTTTTAGTTTTCTGGGATATGTGACCTGTCTAGTCCTTGACAAT
ATGCTGCCTGTGATATAGTATATATTTCT

>S9_153271004_B73V2.7_Chr9:153270954..153271054
AGCTCGAGCGATAGAGGTGAAGTGCACCATGAGCTGATACTGGTCAAGGGACAGGGGAATGTCATCGTGA
TCAACACCATCTTTGGTGGTGACACACTGCT

>S1_216533939_B73V2.7_Chr1:182908984..182909084
TGCTGCGGCTGTAGCTCATGCGTCCTGAAGATAGCTGGTGTACTATAGCACTCTGTACTTTACAAGGTTAAT
GTTAACCCAAGTACGTCTTTATCAAGTTT

>S5_182909034_B73V2.7_Chr5:216533889..216533989

>S6_72860333_B73V2.7_Chr6:72860283..72860383
CGGCCGATTCTCCGACGGCAAGATCGTCACCGACATCACCGGTGCTAGCCATTCCCATATATATATACAT
GCTATACCACCACGCATGCATTATGCATGG

>S7_157194565_B73V2.7_Chr7:157194515..157194615
GGCAATTTGGGTTGATTATTGACTCGGTGTCACCCAGACACCTGAACGCCATTCCCTAATCAAGCTAGCAACTAACTAA
GCACCGTCAATTAGTTTAAAG

>S5_170164202_B73V2.7_Chr5:170164152..170164252
GCAGCACAAGCAGCAGTTTTGCCACCACACCGACGGCGAGCATCTCCTGCCGCACCGAGGGCGTCCGGCG
AGTGCCTGGCCACTGCGTGGAGGGCGCGCACG

>S10_115518234_B73V2.7_Chr10:115518184..115518284

AATGGCGTAGCACCATTACCCGTCACGCGGATTGGCACTAGCAGCCGAGCAACGCCCGCCGTCCATCTCC
ATCCGACGGCCGACGCCGTGTCCCCAAAAT

>S7_153568928_B73V2.7_Chr7:153568878..153568978

AGGGTGTCCCGCGGGTGGCGGCGCAGGGGAGAGGCCAAGGAGGAGGGGGAGCGGTGCTGAGGACGC
CGCTGCTGGGCGCGGGGAGGCGGGGCCTGGGGTG

>S5_56428968_B73V2.7_Chr5:56428918..56429018

GGGGAGGCAGCGGGCTGCGGCCGGGACTAGTACTAGGTAGCAGTAGTCCAAGTCCGGCGGTCCGGCC
GCGTCTCCGGCGGCGAAGAGGTAGGGGCGAGCG

>S5_178037831_B73V2.7_Chr5:178037781..178037881

TGTCTAGTGTCTACGTACATGTGCTCTCTGACCCATGCTAGATTTATTACCCTCCTCTTTGCAACCATGATGA
TAGCTCTAGTGACGGCAACCATGACAGC

>S2_38552471_B73V2.7_Chr2:38552421..38552521

GATCTGCCAGCGCGTGCGAAACGGCGAGCCACTGCCGGCCTCGTGGCGGGCGTTCCTCGAGCAGTCAGA
CACCAAGATCGACTTGTGAGGCTGTCCGGTCA

>S4_183699508_B73V2.7_Chr4:183699458..183699558

CAGAGGAGATTCCTTTCCCTTTCCGTTAGCAGGATAGAGAGGGAGAGGAGCAGAAGATAGCAGAGGGCCA
GAGGCGCAGAAGGCAGTCGCGGACAGGAGGA

>S5_8266072_B73V2.7_Chr4:8266022..8266122

GAGGGAGGGGGACAGAGAGCCGAGAGAGCCCTGCTCCACCGCGACCATTTTTCTCACCGAAAGGGAGGCTGCCGATGAT
AGAACGGAGGGACCGGATCACC

>S3_158354184_B73V2.7_Chr3:158354134..158354234

CTGGGACTGCCCAACCACTGAACAGCTTCATGCTGTTGTTAGTGGCATTTTGAAGGAAGTTGACTTCAACA
CAGTAAGCAATTGATTACGTTTTGAACAT

>S3_146026612_B73V2.7_Chr3:146026562..146026662

TCCTTGCTCTGGCAGCGTTCAAGGTAGCACTGCAGGATGGCAAGCTTCGCGGTCTCAAAGGCACAGTAGC
CTAGCGCCGCAAAGCACGCGCTGTGCAGCAC

>S8_2763004_B73V2.7_Chr8:2762954..2763054

GCTCTTCCCTGCGCGCGGTTCCATTCCATTCCATCAGACGACAGAGAGAGAGAGGGAGAGAGAGAGA
GAGAGAAAGGCGCGCAATCACCGGCCGTCAG

>S2_225857203_B73V2.7_Chr2:225857153..225857253

TTGATCAACAACACTATATATAACAACAGTATGGGATGATTGAGTGAATGTCTAGGCATGCATGCAATAAATTG
GCTGATCACTTGACCTGAGGACGCTGCC

>S6_139454544_B73V2.7_Chr6:139454494..139454594

AGAGTAAGAAGAACAGCAGGCCGTGCTAGCTAGCTGCCGGACATATAACCACCATGGTCCATGGCTACCGT
CGTACCAAGTGCAGAGGAGAGGAGAAGCGAC

>S1_229707717_B73V2.7_Chr1:229707667..229707767

ATGACAGCAGCAAGGAGGGGATGAAGAAGCGGAGGAGGAAGGACGGGCGGCAGGCTCAAGCTGTCAT
GGCTGGCGTGGCCGTGGTCACGGAGAAGCGAC

>S9_135860028_B73V2.7_Chr9:135859978..135860078

CACTCACGAGTAAGCGAAGAGAAAATGGATGAACACACGGGAGAGGCTGTTTGTGGTTGCGAGATCGTGA
TTCTTTTCAGCAGTTGGATCGGTCACGCGC

>S7_5093341_B73V2.7_Chr7:5093291..5093391

GGATCCGCGCCATTGGCAGGAGCCGAGAGTCCGCTTCAGCGCCTTGTTGGCGCCGTTGCGGGCGGCTG
GGGTGGGGTTCAGCTGCTCCCCGCCAGCGCG

>S1_12854004_B73V2.7_Chr7:12853954..12854054

CAGCGAAGAAAGAACCACGGAGATGCATGATTTGTTGCTCGCGATCCTGTATGATCTGTCTTCGTTCTGGAGAAGCCGCTC
CCACCCTCGATTTGGATTTG

>S2_44378881_B73V2.7_Chr2:44378831..44378931

CTCCCGCTGTCCCGCAAGAACACGAGCAGGAGGCGTGGCCGTGGCAGCCGTTTGACCCGCCCGGAGC
GTTATAAGCTGTTTCTCCGGAACGTCTTGTC

Anexo III (inicio). Partes de la plántula de maíz con mayor abundancia de transcritos de los genes asociados a la elongación del mesocotilo de acuerdo al atlas de transcripción del desarrollo de maíz (Sekhon *et al.* 2011). Clave) ubicación cromosomal y posición del SNP en número de bases; Gen) nombre del gen asociado; Pfam) función conocida para su familia protéica; Expresión) estructura de la plántula donde hay mayor concentración del transcrito.

Clave	Gen	Pfam	Expresión
S7_162027150	AC192356.3_FG005	Gen de baja confianza	
	GRMZM2G549820	Gen de baja confianza	
S8_166105182	GRMZM2G335218	Transportador de amonio	Internodo
S4_4737438	GRMZM5G845021	Adenilil-sulfato kinasas.	Hoja de la plántula
S4_226459824	GRMZM2G174673	Dominio asociado a metales pesados	Mazorca inmadura
S7_63900887	GRMZM2G060644	Elemento transponible	
	GRMZM2G071808	Transposasa 24	Endospermo
S5_1852022	GRMZM2G138421	Proteína que se une al promotor de <i>Squamosa</i>	Semilla
S8_170878813	GRMZM2G069106	Dedo de Zn con dominio Ran	Punta de la hoja
S3_19260465	GRMZM2G044697	Peptidasa M50	Punta de la hoja
S7_99089950	GRMZM2G002465	<i>Oryza sativa</i> : proteína expresada	Internodo
S2_198509433	GRMZM2G037873	Elemento transponible	Raíz primaria
	GRMZM2G037875	MOSC: Sulfurasa cofactor de molibdeno	Hoja
S3_146363360	GRMZM5G839699	Gen de baja confianza	
	GRMZM5G813403	<i>Arabidopsis thaliana</i> : Superfamilia cupredoxinas	Embrión y hoja
S3_21368942	GRMZM2G126749	Proteína parecida a las que se unen a calmodulina	Raíz primaria
S6_8146003	GRMZM2G136700	<i>Oryza sativa</i> : proteína expresada	
S2_217651275	GRMZM2G097277	Alfa-L-arabinofuranosidasa dominio c terminal	Raíz y hoja
S7_165369648	GRMZM2G111224	Factor Von Willebrand tipo A	Embrión y endospermo
	GRMZM2G549006	Elemento transponible	
S10_91959067	GRMZM2G158141	<i>Oryza sativa</i> : C ₂ H ₂ dedos de zinc	Hoja de la plántula
S10_5458621	GRMZM2G044423	<i>Oryza sativa</i> : proteína expresada	
S5_11923821	GRMZM2G57747	Gen de baja confianza	
	GRMZM2G157166	<i>Oryza sativa</i> : proteína supuestamente expresada de unión al oxisterol	Anteras
S7_64048287	GRMZM2G072000	Gen de baja confianza	Germinación
	AC198118.4_FG001	Gen de baja confianza	

Anexo III (continuación).

Clave	Gen	Pfam	Expresión
S9_15357629	GRMZM2G328154	Gen de baja confianza	
S10_98272538	GRMZM2G093022	Citidina y desoxitidilato deaminasa, región de unión a zinc	Coleoptilo y hoja
S2_236282499	GRMZM2G402239	Duf1719	Raíz
S4_174903108	GRMZM2G039263	HSP60, proteína de choque de calor	Endospermo
S2_204476331	GRMZM2G163233	Hélice-vuelta-hélice; V-ATPasa, sub-unidad H	Raíz primaria
S8_131240915	GRMZM2G001648	Codifica para una proteína desconocida	Endospermo
	GRMZM2G054584	Gen de baja confianza	
S8_131049015	GRMZM2G117007	Dedos de zinc tipo C-x8-C-x5-C-x3-H	Hoja
	GRMZM2G119527	Ferroporfina 1	Endospermo
S5_193547933	GRMZM2G165530	MFS_1 proteínas facilitadoras principales	Anteras
S5_183326603	GRMZM2G014750	Dominio La, motivo de unión al ARN, reconocimiento de motivos de ARN	Endospermo
S10_127809396	GRMZM2G066162	Glicosido hidrolasa de la familia 9	Semilla
S7_160595701	GRMZM2G058087	TIP49 C-terminal	Embrión
	GRMZM2G517577	Gen de baja confianza	
S10_10586901	GRMZM2G026231	Proteína de endomembrana 70	Raíz primaria
S3_13726673	GRMZM2G048022	<i>Oryza sativa</i> : proteína expresada	Plántula
	GRMZM2G047815	GPI-GlcNAc complejo transferasa, PIG-H componente	Punta de la hoja
S5_215289316	GRMZM5G806449	Piridina disulfuro oxidoreductasa	Raíz
S10_4751416	GRMZM2G031261	PSK, proteína precursora de fitosulfoquina	
	GRMZM2G031164	Elemento transponible	
S10_140743010	AC208110.2_FG007	<i>Oryza sativa</i> : Bn15D1B proteína específica de la semilla	Hoja de la plántula
S6_106176548	GRMZM2G071638	DUF212, familia divergente PAP2	hoja
S1_14103918	GRMZM5G850455	Tryp_alpha_amyl, proteína de transferencia de lípidos en plantas	Semilla
S6_71803064	GRMZM2G164779	Elemento transponible	
	GRMZM2G581506	Elemento transponible	
S9_139410281	GRMZM2G165357	Epimerasa NADH dehidrogenasa	Raíz
	GRMZM2G057467	DUF 568	
S2_44378881	GRMZM2G015844	<i>Oryza sativa</i> : receptor de esteroides nucleares	Raíz primaria

Anexo III (termino).

Clave	Gen	Pfam	Expresión
S8_167282998	GRMZM2G177324	Peptidasa_C48, Ulp1 familia de proteasas, dominio catalítico C-terminal	Endospermo
S1_216533939	GRMZM2G124835	Gen de baja confianza	
	GRMZM2G497425	Gen de baja confianza	
S9_153271004	GRMZM2G360025	Proteína de unión a zeina	Anteras
	GRMZM2G178106	Glicosil hidrolasa 35	
S5_182909034	GRMZM2G546415	Gen de baja confianza	
	GRMZM2G106680	<i>Oryza sativa</i> : proteína de unión a calcio en mitocondria	Coleoptilo y hoja
S5_170164202	GRMZM2G071484	Dominio caja U	Raíz primaria
S6_72860333	GRMZM2G152777	Lipasa_GDSL, acilhidrolasa	Hoja
S7_157194565	GRMZM2G031323	Proteína parecida a MYB unión de ADN	Hoja
	GRMZM2G036291	Elemento transponible	
S10_115518234	GRMZM2G179005	Histona H2A, H2B, H3 y H4	Embrión
S5_56428968	GRMZM2G099191	ATPasa transportadora de Cadmio/Zinc	Internodo y raíz
S7_153568928	GRMZM2G150791	<i>Arabidopsis thaliana</i> : repetición de ankirina	Punta de la hoja
S2_38552471	GRMZM2G165844	Unión a AMP, NADH dehidrogenasa, catalasa	Mazorca
S5_178037831	GRMZM2G095782	tRNA sintasa de pseudoridina	Punta de la hoja
S4_183699508	GRMZM2G088847	MIF, factor de inhibición de migración de macrófagos	Hoja
S5_8266072	GRMZM2G548459	Gen de baja confianza	
	GRMZM2G010801	<i>Arabidopsis thaliana</i> : Metiltransferasa	Punta de la hoja
S3_146026612	GRMZM2G047457	Proteína transportadora de la mitocondria	
S8_2763004	GRMZM2G312419	MYb 60 factor de transcripción	Raíces de la corona
S2_225857203	GRMZM2G433731	<i>Oryza sativa</i> : Repetición rica en leucinas	Raíz
S6_139454544	GRMZM2G423863	Codifica para una proteína desconocida	Raíz primaria y hoja
S3_158354184	GRMZM2G139617	Dominio DEK C-terminal	Hoja
S1_229707717	AC199893.4_FG002	Codifica para una proteína desconocida	Raíz primaria
S9_135860028	GRMZM2G500857	Gen de baja confianza	Embrión
S7_5093341	GRMZM2G011800	<i>Arabidopsis thaliana</i> : miosina	Endospermo
S1_12854004	GRMZM2G136081	Parecida a PDDEXK función desconocida	Pericarpo

Anexo IV (inicio). Frecuencias alélicas de los 18 SNP asociados a la elongación del mesocotilo en la población DTP incluidos en los análisis de QTL. Clave) ubicación cromosomal y posición del SNP en número de bases; p) probabilidad $[-\log_{10}]$ de asociación; DF) proporción de datos faltantes; Ma) estado alélico homocigoto mayor; Ma pro) proporción del estado alélico homocigoto mayor; Me) estado alélico homocigoto menor; Me pro) proporción del estado alélico homocigoto menor; He) número de accesiones con el estado alélico heterocigoto; He pro) Proporción del estado alélico heterocigoto.

Clave	p	DF	Ma	Ma pro	Me	Me pro	He	He pro
S8_166105182	4.22	0.22	A	0.47	C	0.31	30.00	0.21
S3_19260465	3.96	0.26	C	0.51	T	0.23	16.00	0.12
S7_99089950	3.95	0.13	C	0.69	T	0.18	20.00	0.13
S3_146363360	3.88	0.19	T	0.48	C	0.33	45.00	0.31
S3_21368942	3.84	0.26	A	0.47	G	0.27	23.00	0.17
S6_8146003	3.83	0.31	C	0.47	T	0.23	17.00	0.14
S10_5458621	3.78	0.44	T	0.43	C	0.13	8.00	0.08
S10_10586901	3.53	0.44	T	0.49	A	0.08	3.00	0.03
S5_215289316	3.49	0.22	T	0.59	C	0.19	20.00	0.14
S6_106176548	3.48	0.38	A	0.39	T	0.23	14.00	0.13
S9_139410281	3.45	0.31	G	0.49	A	0.21	15.00	0.12
S8_167282998	3.44	0.17	A	0.48	G	0.35	32.00	0.21
S5_170164202	3.41	0.12	C	0.79	A	0.09	9.00	0.06
S4_183699508	3.35	0.38	C	0.51	G	0.11	6.00	0.05
S3_158354184	3.33	0.01	C	0.64	T	2.00	0.00	0.00
S6_139454544	3.30	0.26	G	0.53	C	0.21	16.00	0.12
S7_5093341	3.27	0.00	G	0.70	+	1.00	1.00	0.01
S2_44378881	3.25	0.28	C	0.41	G	57.00	9.00	0.09

Anexo V (inicio). Alelos asociados a la elongación del mesocotilo encontrados en el genoma de referencia de Palomero Toluqueño. Clave) ubicación cromosomal y posición del SNP en número de bases; Pos) posición relativa del SNP, ya sea Intergénico (inter) o dentro del marco de lectura de un gen o modelo de gen (Marc); B73_V2.7) alelo en el genoma de referencia B73_V2.7; Ma DTMA) alelo mayor en las DTMA; Me DTMA) alelo menor en las DTMA; Palomero T.) alelo en el genoma de referencia Palomero toluqueño. NA no se encontró la región asociada.

Clave	gen	B73_V2.7	Ma DTMA	Me DTMA	Palomero T.
S7_162027150	Inter	C	C	G	C
S8_166105182	Marc	C	C	A	C
S4_4737438	Marc	A	A	G	NA
S4_226459824	Marc	A	A	G	A
S7_63900887	Inter	C	G	A	G
S5_1852022	Marc	A	C	A	A
S8_170878813	Marc	A	G	A	A
S3_19260465	Marc	C	C	T	NA
S7_99089950	Marc	T	T	C	T
S2_198509433	Inter	A	T	A	NA
S3_146363360	Inter	C	C	T	NA
S3_21368942	Marc	A	A	G	A
S6_8146003	Marc	C	T	C	NA
S2_217651275	Marc	G	G	T	NA
S10_91959067	Marc	G	G	A	G
S7_165369648	Inter	T	T	G	NA
S10_5458621	Marc	T	C	T	NA
S5_11923821	Inter	G	C	G	C
S9_15357629	Marc	C	C	A	C
S7_64048287	Inter	G	G	C	NA
S10_98272538	Marc	C	C	G	C
S4_174903108	Marc	G	G	A	NA
S2_236282499	Marc	A	G	C	G
S2_204476331	Marc	T	T	A	NA
S8_131240915	Inter	C	T	C	T
S8_131049015	Inter	G	T	G	NA
S5_193547933	Marc	C	C	T	C
S5_183326603	Marc	G	G	A	A
S10_127809396	Marc	C	G	A	C
S7_160595701	Inter	A	A	G	A
S10_10586901	Marc	A	A	T	NA
S3_13726673	Inter	C	C	T	T
S10_140743010	Marc	C	A	C	C
S10_4751416	Inter	T	T	A	T
S5_215289316	Marc	C	C	T	NA
S6_106176548	Marc	T	T	A	NA

Anexo V (termino).

Clave	gen	B73_V2.7	Ma DTMA	Me DTMA	Palomero T.
S1_14103918	Marc	C	C	G	C
S6_71803064	Inter	A	A	G	G
S9_139410281	Inter	G	G	A	G
S8_167282998	Marc	G	G	A	A
S9_153271004	Inter	A	A	G	NA
S1_216533939	Inter	C	C	A	NA
S5_182909034	Inter	G	G	T	NA
S6_72860333	Marc	C	C	T	NA
S7_157194565	Inter	C	C	T	NA
S5_170164202	Marc	C	C	A	NA
S10_115518234	Marc	A	A	C	A
S7_153568928	Marc	A	A	G	NA
S5_56428968	Marc	A	A	C	A
S5_178037831	Marc	C	C	T	NA
S2_38552471	Marc	C	C	T	C
S4_183699508	Marc	C	C	G	C
S5_8266072	Inter	T	T	G	NA
S3_158354184	Marc	T	C	T	T
S3_146026612	Marc	G	G	A	NA
S8_2763004	Marc	G	G	C	NA
S2_225857203	Marc	C	A	C	NA
S6_139454544	Marc	C	C	G	C
S1_229707717	Marc	C	C	A	C
S9_135860028	Marc	T	T	C	T
S7_5093341	Marc	G	G	A	G
S1_12854004	Marc	T	A	G	T
S2_44378881	Marc	G	G	C	T

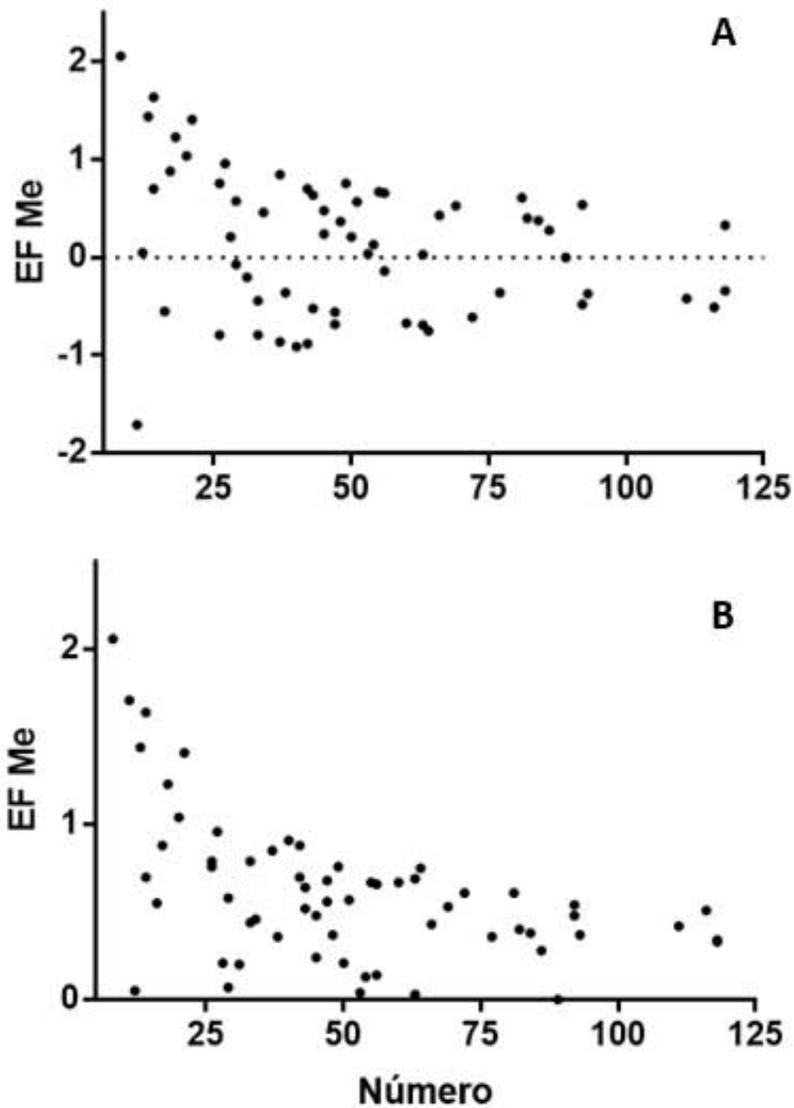


Figura Suplementaria 1. Relación entre el efecto fenotípico del alelo menor de los 63 SNPs fuertemente asociados a la elongación del mesocotilo, obtenidos por GWAS en la colección DTMA, y el número de accesiones (n) conteniendo el alelo menor. A) Valores calculados de los efectos fenotípicos; B) Valores absolutos de los efectos fenotípicos.



Figura Suplementaria 2. Mapa de calor de los efectos fenotípicos de los 63 SNPs asociados a la variación de la elongación del mesocotilo para las 16 DTMA con mesocotilo más largos y las 16 DTMA con los mesocotilos más cortos.