

OPTIMIZACIÓN DE MAPEO DE LOCI DE CARACTERES CUANTITATIVOS ASOCIADOS A EFICIENCIA DE USO DE NITRÓGENO EN MAÍZ E IDENTIFICACIÓN DE GENES CANDIDATOS

Trabajo Final de Grado
del alumno



**Escuela de Ciencias Agrarias, Naturales y Ambientales.
Universidad Nacional del Noroeste de la Provincia de Buenos Aires.**

Pergamino,.....

OPTIMIZACIÓN DE MAPEO DE LOCI DE CARACTERES CUANTITATIVOS ASOCIADOS A EFICIENCIA DE USO DE NITRÓGENO EN MAÍZ E IDENTIFICACIÓN DE GENES CANDIDATOS

Trabajo Final de Grado

del alumno

GISELA ANDREA MANZELLI

Aprobada por el Tribunal Evaluador

(Nombre y Apellido)
Evaluador

(Nombre y Apellido)
Evaluador

(Nombre y Apellido)
Evaluador

(Antonio Díaz Paleo)
Co-Director

(Cecilia Inés Mandolino)
Directora

**Escuela de Ciencias Agrarias, Naturales y Ambientales,
Universidad Nacional del Noroeste de la Provincia de Buenos Aires**

Pergamino,.....

AGRADECIMIENTOS

A la UNNOBA por la formación académica.

A mi directora Cecilia Inés Mandolino y mi co-director Antonio Díaz Paleo por el apoyo recibido para realizar este trabajo, particularmente a Ceci por su dedicación, orientación y seguimiento continuo.

A los chicos del laboratorio de biotecnología de INTA EEA Pergamino Agostina, Carla, Daniela, Luciano, Marcelo y Viviana por su colaboración y aportes.

A mis compañeros, por todos los momentos compartidos y por su amistad, especialmente a Anto, Emi, Guada, Mati, Rocío, Vivi y Yael por hacerme recordar esta etapa llena de lindos momentos y con tantas anécdotas.

A las amigas que me regaló el paso por la UNNOBA Guada, Male, Mari, Nati y Sole y a mis amigas de la vida por estar siempre presente.

A Gladys, por estar presente con la calidez de la familia durante la estadía en Pergamino.

A toda mi familia, especialmente a mis padres y mi hermano, por el apoyo incondicional de siempre, la confianza y sobre todo por su amor.

ABREVIATURAS

ABA	Ácido abscísico
ADN	Ácido desoxirribonucleico
ASI	Intervalo entre Antesis y emisión de estigmas o <i>Silking</i>
BiomMF	Biomasa a madurez fisiológica
cM	centiMorgan
EUN	Eficiencia de uso de nitrógeno
Exp.	Experimento
FDR	Tasa de falsos descubrimientos, del inglés <i>False Discovery Rate</i>
GS	Glutamina sintetasa
ICN	Índice de cosecha de nitrógeno
INDEL	Combinación de las palabras inglesas <i>insertion – deletion</i>
LOD	Logaritmo de probabilidades de puntuación, del inglés <i>Logarithm Of Odds</i>
MIC	Mapeo por intervalo compuesto
MIS	Mapeo por intervalo simple
Mpb	Mega pares de base
N	Nitrógeno
N/BiomMF	Proporción de nitrógeno en la biomasa a madurez fisiológica
NabsMF	Nitrógeno absorbido a madurez fisiológica
NGP	Número de granos por planta
pb	Pares de bases
PCR	Reacción en cadena de la polimerasa, del inglés <i>Polymerase Chain Reaction</i>
PM	Peso molecular
QTL	<i>Loci</i> de caracteres cuantitativos, del inglés <i>Quantitative Trait Loci</i>
RGP	Rendimiento en grano por planta
RILs	Líneas endocriadas recombinantes, del inglés <i>Recombinant Inbred Lines</i>
SDR	Regiones de distorsión de segregación, del inglés <i>Segregation Distortion Regions</i>
SNP	Polimorfismo en un nucleótido simple, del inglés <i>Single Nucleotide Polymorphism</i>
SSR	Repeticiones de secuencia simple o Microsatélite, del inglés <i>Simple Sequence Repeats</i>

ÍNDICE

ÍNDICE	5
RESUMEN.....	7
INTRODUCCIÓN.....	9
El maíz: características	9
El cultivo de maíz	10
Metabolismo de nitrógeno	11
Eficiencia de uso de nitrógeno	12
Marcadores moleculares	12
Mapeo de QTL	15
Antecedentes del grupo de investigación.....	17
Hipótesis.....	19
Objetivo general	20
Objetivos específicos.....	20
MATERIALES Y MÉTODOS.....	21
Evaluación fenotípica	21
Evaluación genotípica	22
Extracción de ADN.....	22
Cuantificación de ADN	23
Análisis de los marcadores moleculares	24
Búsqueda de marcadores moleculares en base de datos.....	24
Detección de marcadores moleculares: SSR e INDEL	24
Caracterización genotípica de la población de mapeo	27
Análisis de segregación de los marcadores	28
Mapeo de QTL	28
Construcción de los mapas de ligamiento.....	28

Análisis de QTL.....	29
Identificación de genes candidatos	30
RESULTADOS	31
Caracterización genotípica	31
Análisis de segregación de los marcadores	33
Mapeo de QTL	35
Mapas de ligamiento	35
Análisis de QTL.....	36
Identificación de genes candidatos	39
DISCUSIÓN.....	43
Caracterización de la población	43
Mapeo de QTL	44
Mapas de ligamiento	44
Análisis de QTL.....	45
Identificación de genes candidatos	47
CONCLUSIÓN.....	52
BIBLIOGRAFÍA.....	53
ANEXO 1	61
ANEXO 2	62

RESUMEN

La absorción de nitrógeno (N) por las plantas tiene un papel importante en su crecimiento. En consecuencia, los fertilizantes nitrogenados han sido una herramienta útil para aumentar el rendimiento de las plantas cultivadas, pero cuando se usan de forma indiscriminada pueden causar severos daños al medio ambiente y pérdidas económicas. Por lo tanto, el mejoramiento para la adaptación a tipos de estrés como la falta de N es de gran importancia. La mayoría de los caracteres más relevantes para el mejoramiento vegetal son cuantitativos. El mapeo de *loci* de caracteres cuantitativos (QTL) permite identificar regiones en el genoma que explican parte de la variación fenotípica observada.

El objetivo de este trabajo fue aumentar la precisión de la localización de QTL para la eficiencia de uso de nitrógeno (EUN) detectados en los cromosomas 1, 8 y 9 de maíz sobre un mapa genético elaborado a partir de una población de 181 líneas endocriadas recombinantes o RILs (del inglés *Recombinant Inbred Lines*) S₆ derivada del cruzamiento B100xLP2 incrementando el número de marcadores moleculares microsatélites o SSR (del inglés *Simple Sequence Repeats*) e INDEL (combinación de las palabras inglesas *insertion – deletion*).

Para realizar la caracterización genotípica se probaron 149 marcadores SSRs e INDEL en los parentales, de los cuales 40 fueron usados para evaluar la población de RILs. La incorporación de marcadores moleculares nuevos a los mapas de ligamiento de los cromosomas 1, 8 y 9 permitió una mejor representación de los cromosomas y mayor precisión en la localización de los QTL. Para todos los cromosomas se logró incrementar la longitud del mapa con respecto al obtenido previamente.

El mapeo de QTL se realizó mediante mapeo por intervalo compuesto (MIC) y se detectaron 11 QTL en los cromosomas 1, 8 y 9. En el cromosoma 1, se detectaron dos QTL para EUN uno para cada año de ensayo, con intervalos de confianza que se superponen parcialmente, y podría tratarse de un QTL estable en los diferentes

ambientes. En el cromosoma 9, se detectaron tres QTL para EUN, solo uno (EUN-1d) se localizó dentro de la región de interés y la incorporación de marcadores permitió acotar el intervalo del QTL para EUN detectado sobre el cromosoma 9 de 55,3 Mpb a 15,4 Mpb. Los QTL EUN-1d y ICN-1b (índice de cosecha de N) colocalizaron entre los mismos marcadores indicando que se podría tratar de un único QTL. En los cromosomas 8 y 9 se detectaron QTL nuevos para EUN, ICN y rendimiento fuera de la región de interés. Muchos de los QTL detectados coincidieron con los QTL encontrados en la bibliografía.

Se identificaron 3155 genes dentro de los intervalos de los QTL detectados y se clasificaron según su función. El 67% de los mismos corresponden a proteínas con función putativa, mientras que el 33% restante aún no tienen una función caracterizada. Se identificaron genes relacionados al metabolismo del N en la mayoría de los intervalos de los QTL, como asparagina sintetasa y glutamato sintasa, y genes involucrados en la respuesta al estrés y el transporte de aminoácidos.

La incorporación de marcadores moleculares logró acotar el intervalo de un QTL sobre el cromosoma 9, sin embargo muchos de los intervalos siguen siendo amplios. Un mapeo fino permitiría detectar marcadores más próximos a los QTL de interés, los cuales podrían ser usados para selección asistida por marcadores para mejorar la EUN.

INTRODUCCIÓN

El maíz: características

El maíz (*Zea mays* L.), pertenece a la familia de las *Poáceas* (*Gramíneas*) y es la única especie cultivada del género *Zea*. La planta de maíz es alta, posee un ciclo de vida anual, con abundantes hojas y un sistema radical fibroso. Es una planta diclino monoica que desarrolla inflorescencias con flores de un solo sexo, las que crecen siempre en lugares separados de la planta. La inflorescencia femenina, la cual da lugar a la espiga crece a partir de las yemas axilares en las axilas de las hojas y es la estructura de la planta que almacena reservas. La inflorescencia masculina o panoja se desarrolla en el punto de crecimiento apical en el extremo superior de la planta; esta tiene una espiga central prominente y varias ramificaciones laterales con flores masculinas, las cuales producen abundantes granos de polen (Paliwal *et al.*, 2001).

El maíz es una de las pocas especies de cultivos alimenticios diploides con diez cromosomas.

Existen varias teorías respecto al origen del maíz. Una de ellas plantea que el teosinte (*Zea mays ssp. Parviglumis*) es el antecesor silvestre y que ha participado directamente sobre el origen del maíz cultivado (Beadle, 1980; Doebley *et al.*, 1990; Galinat, 1988), otras plantean su origen a partir de antiguas formas silvestres de maíz (Mangelsdorf, 1986). En relación al centro de origen, la teoría más aceptada es el centro de origen mexicano debido a que allí es donde el maíz y teosinte han coexistido desde la antigüedad y donde ambas especies presentan diversidad muy amplia (Galinat, 1988). Si bien el maíz es originario de América Central, fue introducido en los demás continentes del mundo y las razas con mejor adaptabilidad han sido seleccionadas por los agricultores para los nuevos ambientes dando lugar a nuevas variedades (Prasanna, 2012).

El cultivo de maíz

El maíz es un cultivo de gran importancia económica a nivel mundial, junto con el arroz (*Oriza sativa* L.) y el trigo (*Triticum aestivum* L.). Es utilizado tanto como grano y forraje así como también en la producción farmacéutica y como materia prima para productos químicos orgánicos como el caucho, el etanol y el plástico. Además de su importancia económica, ha sido un modelo clásico de plantas para estudios genéticos durante décadas (Helentjaris, 1988).

Con la creciente diversificación de la demanda y la utilización del maíz, el comercio mundial se ha convertido en una estrategia importante para superar los déficits de producción.

Los rendimientos en ambientes de baja productividad están limitados por una serie de factores, incluyendo estrés abiótico y biótico. La sequía recurrente, los bajos niveles de uso de fertilizantes y la baja adopción de variedades mejoradas contribuyen a los bajos rendimientos. El mejoramiento de caracteres, como tolerancia a la sequía, baja fertilidad de nitrógeno, suelos anegados, calientes o ácidos y la resistencia a enfermedades, insectos y parásitos, pueden reducir significativamente el riesgo asociado con la siembra en seco y promover mejores prácticas de manejo (Shiferaw *et al.*, 2011).

Los aumentos de la producción de cultivos provienen de tres fuentes principales: la expansión de la superficie de siembra, el aumento de la intensidad de los cultivos y la mejora del rendimiento. El incremento en los rendimientos ha sido el factor más importante tanto en los países desarrollados como en aquellos en desarrollo, donde representó el 70 por ciento del aumento de la producción.

La mayor parte del crecimiento futuro de la producción de cultivos se logrará mediante mejores rendimientos. Uno de los medios principales que tienen los agricultores para aumentar los rendimientos es la aplicación de más fertilizantes (Harrison, 2002).

Metabolismo de nitrógeno

Los fertilizantes nitrogenados han contribuido de manera significativa al aumento en la producción de alimentos. El nitrógeno (N) aportado por estos fertilizantes es el principal elemento mineral absorbido por las plantas e interviene en procesos fisiológicos esenciales para su crecimiento y desarrollo. Algunas plantas pueden utilizar el nitrógeno atmosférico, a través de su asociación con microorganismos, pero la mayoría de los cultivos depende del suministro externo de N vía mineralización de la materia orgánica y adición de fertilizantes, para completar su ciclo de crecimiento.

Aunque el uso de fertilizantes nitrogenados ha sido uno de los principales promotores del aumento en el rendimiento en la agricultura actual, su utilización implica riesgos de contaminación ambiental, por la lixiviación del N no absorbido hacia aguas subterráneas. Además, su aplicación excesiva conlleva a la degradación de la calidad del suelo por salinidad y un incremento en los costos de producción. De esta manera, los fertilizantes se han convertido en una tecnología que puede propiciar grandes beneficios, cuando se utilizan de manera adecuada pero que al mismo tiempo pueden causar severos daños al medio ambiente y pérdidas económicas cuando se usa de forma indiscriminada (Cárdenas-Navarro *et al.*, 2004). El N disponible en el suelo está presente típicamente en forma de nitrato y amonio, los cuales son muy móviles y las plantas cultivadas sólo son capaces de utilizar el 30-40% del N aplicado. Por lo tanto, más de 60% del N del suelo se pierde a través una combinación de lixiviación hacia aguas subterráneas, escurrimiento superficial, desnitrificación, volatilización y consumo microbiano. De este modo, la utilización excesiva implica riesgos de contaminación ambiental y degradación de la calidad del suelo (Raun & Johnson, 1999).

Hay dos momentos en el ciclo de vida de las plantas en los cuales el uso de N es importante. Durante la etapa vegetativa, las hojas jóvenes en desarrollo y las raíces se comportan como sumideros para la absorción de N inorgánico; sintetizan y almacenan los

aminoácidos a través de la vía de asimilación de nitrato. Estos aminoácidos son utilizados en la síntesis de proteínas y enzimas involucradas en diferentes vías bioquímicas y en la maquinaria fotosintética durante el crecimiento y el desarrollo de las plantas. Durante la etapa reproductiva, el aumento de compuestos nitrogenados es necesario para una óptima floración y llenado de grano. En esta etapa, tanto la asimilación de N como la removilización se vuelven críticas y las hojas y los brotes actúan como la fuente que proporciona aminoácidos a los órganos reproductores y de almacenamiento (Kant *et al.*, 2011).

Eficiencia de uso de nitrógeno

La EUN se define como la cantidad de materia seca producida por unidad de N disponible en el suelo. En el maíz esta eficiencia depende de numerosos procesos, que incluyen la absorción de N por las raíces, la translocación a las hojas y la removilización de N desde hojas y tallos hacia los granos. Por tal motivo, el mejoramiento genético en EUN en maíz depende del mejoramiento en uno o más de estos procesos (Agrama, 2006).

La mejora genética del cultivo de maíz por una mayor EUN permitiría el logro de mejores rendimientos con menor aplicación de fertilizantes en suelos pobres en el nutriente y/o una mayor respuesta por unidad de N aplicado como fertilizante (Arellano *et al.*, 1996).

Marcadores moleculares

El objetivo del mejoramiento vegetal ha sido, desde sus comienzos, seleccionar genotipos superiores a partir de la identificación de fenotipos superiores. El proceso de caracterización y/o selección se vuelve más eficiente a través del uso de marcadores

genéticos, definidos como caracteres que presentan polimorfismo o variabilidad experimentalmente detectable entre individuos y un tipo de herencia predecible según las leyes de Mendel. Esta variación puede considerarse a diferentes niveles biológicos, desde cambios fenotípicos heredables significativos (marcador morfológico) hasta la variación de un solo nucleótido de ADN (marcador molecular) (Martínez *et al.*, 2010).

Los marcadores moleculares tienen varias ventajas sobre los marcadores fenotípicos tradicionales. Ellos permiten mejorar la eficiencia del mejoramiento convencional, mediante la realización de la selección indirecta del rasgo de interés a través de marcadores moleculares ligados a ese rasgo (Mohan *et al.*, 1997).

Los marcadores moleculares surgen de diferentes clases de mutaciones de ADN tales como mutaciones puntuales, inserciones, deleciones o errores en la replicación de ADN repetido en tándem (Paterson, 1996). Estos marcadores son selectivamente neutros porque normalmente están localizados en regiones no codificantes. A diferencia de los marcadores morfológicos y bioquímicos, los marcadores de ADN son prácticamente ilimitados en número y no se ven afectados por factores ambientales y/o la etapa de desarrollo de la planta (Winter & Kahl, 1995). Lo más frecuente es que las diferencias alélicas presentes en los marcadores moleculares se resuelvan por visualización directa de productos de PCR (del inglés *Polymerase Chain Reaction*) o reacción en cadena de la polimerasa, en geles utilizando la técnica de electroforesis. Son particularmente útiles si revelan diferencias entre individuos. Estos marcadores se llaman marcadores polimórficos, mientras que los marcadores que no discriminan entre genotipos se llaman marcadores monomórficos. Los marcadores polimórficos también pueden describirse como dominantes o codominantes, basándose en si los marcadores pueden discriminar entre individuos homocigotas y heterocigotas (Collard *et al.*, 2005).

Los microsatélites o SSR son regiones hipervariables del genoma que contienen arreglos de secuencias simples en tándem de mono, di, tri, tetra o pentanucleótidos que

se repiten entre 10 y 100 veces. Los marcadores microsatélites, se encuentran distribuidos por todo el genoma de la mayoría de las especies eucariotas. Estos varían de acuerdo al número de unidades repetidas en una región definida del genoma. Las secuencias de nucleótidos que flanquean la repetición se utilizan para diseñar cebadores para amplificar las secuencias por PCR. Este tipo de marcadores son altamente reproducibles, muy útiles para la detección rápida y precisa de *loci* polimórficos y la información podría ser utilizada para el desarrollo de un mapa genético (Mohan *et al.*, 1997).

Estos marcadores se resuelven por electroforesis en geles de poliacrilamida en condiciones desnaturalizantes, mediante tinción con plata o en geles de agarosa. Si se dispone de un secuenciador automático, se pueden resolver utilizando cebadores marcados con fluoróforos, posibilitando un análisis automatizado.

La base genética del polimorfismo detectado en microsatélites se basa en la variabilidad del número de repeticiones en tándem y, consecuentemente, en el tamaño del microsatélite amplificado en individuos de una especie. Estas diferencias son originadas durante la replicación del ADN debido a fallas en la acción de la ADN polimerasa durante el copiado de una región repetida donde incorpora o elimina repeticiones. Otro mecanismo responsable de la variación es el entrecruzamiento desigual entre cromosomas homólogos. En este caso se generan alelos con diferencias mayores en el número de repeticiones (Cervigni *et al.*, 2010).

Los INDEL son la forma más abundante de variación estructural en todos los genomas. Son una fuente importante de marcadores moleculares debido a la alta densidad de ocurrencia, bajo costo y facilidad de genotipificación. El genoma de maíz es rico en polimorfismo por INDEL (Liu *et al.*, 2015). Estos marcadores pueden producirse por deslizamiento de las cadenas nucleotídicas durante la replicación o por entrecruzamiento desigual (Levinson & Gutman, 1987). Los INDEL pueden ser utilizados

para el mapeo genético y el polimorfismo INDEL puede analizarse fácilmente como diferencias en longitud de los productos de PCR (Bhatramakki *et al.*, 2002).

Mapeo de QTL

La mayoría de los caracteres de importancia agronómica son heredados cuantitativamente. Los caracteres cuantitativos son controlados por varios genes cada uno de los cuales, en general, contribuyen al carácter con un pequeño efecto, estando ampliamente influenciados por factores ambientales (Xu, 2010). Las regiones genómicas que comprenden genes que controlan la expresión de un carácter cuantitativo se conocen como *loci* de caracteres cuantitativos o QTL (del inglés *Quantitative Trait Loci*) (Boopathi, 2013).

Uno de los principales usos de los marcadores moleculares en la investigación agrícola ha sido la construcción de mapas de ligamiento en especies cultivadas. Los mapas de ligamiento de marcadores moleculares se utilizan para la identificación de regiones cromosómicas que contienen QTL asociados a los caracteres cuantitativos. El proceso de construcción de mapas de ligamiento y la realización del análisis de QTL para identificar regiones genómicas asociadas con caracteres se conoce como mapeo de QTL (Collard *et al.*, 2005).

El mapeo de QTL se basa en el principio de que algunos genes y marcadores están ligados y cosegregan en la progenie. Los genes y marcadores que están fuertemente ligados serán transmitidos juntos de los padres a la progenie con más frecuencia que los genes y marcadores que se encuentran más separados (Paterson, 1996).

El mapeo de QTL involucra cuatro pasos: la selección y/o desarrollo de una población de mapeo, la evaluación fenotípica de la población para el carácter de interés, la caracterización genotípica de la población utilizando marcadores moleculares y la

identificación de marcadores moleculares ligados a los caracteres de interés usando programas estadísticos (Semagn *et al.*, 2010).

Los métodos para la detección de QTL son mapeo de marca simple, mapeo por intervalo simple (MIS) y mapeo por intervalo compuesto (MIC). El mapeo de marca simple es el método más sencillo para detectar QTL asociados con marcadores individuales. Los métodos estadísticos utilizados incluyen una prueba t (*t-test*), análisis de varianza (ANOVA) y regresión lineal. Este método no requiere un mapa de ligamiento y puede realizarse con programas estadísticos comunes. Sin embargo, la principal desventaja del mapeo por marca simple es que la estimación del efecto del QTL es altamente dependiente de la distancia genética con el marcador. Cuanto más lejos se encuentre el QTL del marcador es menos probable que se detecte debido a que puede existir recombinación entre el marcador y el QTL.

El mapeo por intervalo simple utiliza mapas de ligamiento y analiza los intervalos entre pares adyacentes de marcadores ligados a lo largo de los cromosomas, en lugar de analizar marcadores individuales, y se considera estadísticamente más poderoso que el mapeo de marca simple ya que es posible compensar la recombinación entre los marcadores y el QTL, aumentando la probabilidad de detectar estadísticamente el mismo..

El mapeo por intervalo compuesto combina mapeo de intervalo con regresión lineal e incluye marcadores genéticos adicionales en el modelo estadístico además del par adyacente de marcadores ligados para la asignación del intervalo, lo que resulta más preciso para la detección de QTL.

El mapeo de QTL por intervalo simple y por intervalo compuesto se realiza con programas computacionales específicos de mapeo de QTL (Collard *et al.*, 2005; Tanksley, 1993).

El análisis de la presencia de QTL generalmente se lleva a cabo usando mapeo de intervalo. En el procedimiento básico del análisis de QTL se calcula un test estadístico en muchas posiciones en el mapa de ligamiento para comprobar si es probable que un QTL esté presente en el intervalo o no. Los resultados de las pruebas se expresan como valores LOD (del inglés *Logarithm Of Odds*) o logaritmo de probabilidades de puntuación, que comparan la evaluación de la función de probabilidad bajo la hipótesis nula (ausencia de QTL) con la hipótesis alternativa (un QTL en la posición de prueba). Los perfiles LOD se utilizan para identificar la posición más probable para un QTL en relación con el mapa de ligamiento, que es la posición donde se obtiene el valor LOD más alto. Se determina un valor LOD umbral, el cual debe ser superado para determinar la significancia de un QTL. Generalmente se utiliza un valor LOD de 3 como el umbral de significación. Un valor LOD mayor a 3 implica que la hipótesis nula es rechazada. Este valor corresponde a una relación de probabilidades de 1.000 a 1, es decir, de cada 1.000 análisis existe la posibilidad de 1 de obtener un falso positivo (Boopathi, 2013; Doerge, 2002; Van Ooijen, 1999).

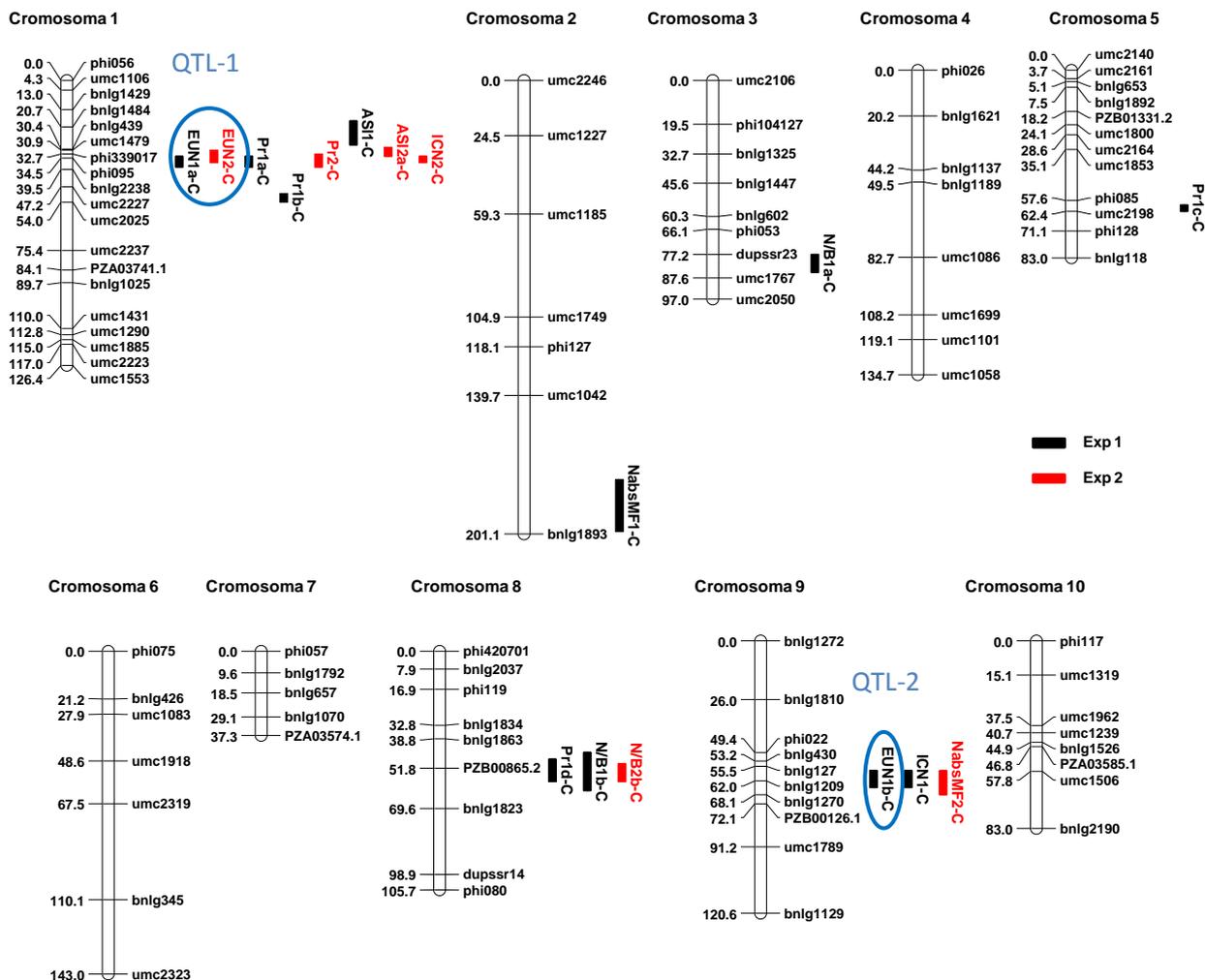
Antecedentes del grupo de investigación

En trabajos anteriores realizados en el Instituto Nacional de Tecnología Agropecuaria (INTA) – Estación Experimental (EEA) Pergamino – se ha logrado identificar 17 QTL asociados a EUN y caracteres relacionados (Figura 1) en una población de o RILs de maíz derivada del cruzamiento de las líneas B100 (dentado, norteamericano) y LP2 (flint, caribeño-argentino) (Mandolino, 2015). El fenotipado de la población de las RILs se realizó a través de dos años de ensayos y se usaron 94 marcadores moleculares neutros SSR en el genotipado. La mayoría de los QTL detectados se localizaron sobre los cromosomas 1, 8 y 9 de maíz. Los QTL para diferentes caracteres compartieron, al menos en parte, el intervalo entre marcadores flanqueantes, lo que podría deberse a la existencia

de un único QTL con efectos pleiotrópicos. Para el carácter EUN en particular, se identificaron dos QTL, uno en el cromosoma 1 (QTL-1) y otro en el cromosoma 9 (QTL-2). El QTL-1 se expresó en los dos años de experimentación indicando que dicho QTL fue estable a través de los ambientes y comparte el intervalo con QTL detectados para otros caracteres asociados a EUN como son el índice de cosecha de N (ICN) y el porcentaje de proteína en grano. El QTL-2 sobre el cromosoma 9, sólo se expresó en uno de los ambientes de experimentación y compartió intervalo con un QTL para ICN. Por último, sobre el cromosoma 8 se detectaron QTL para proporción de N en la biomasa y porcentaje de proteína.

La distancia media resultante entre marcadores en el mapa genético generado fue de 12 cM (centiMorgan) y la localización de los QTL dentro de intervalos de confianza muy amplios fue poco precisa. Consecuentemente se demarcaron extensas zonas físicas en las cuales pueden existir centenares de genes. La saturación de dichas regiones con un mayor número de marcadores posibilitará la búsqueda de genes candidatos. El descubrimiento y/o validación de genes que controlan la EUN y sus componentes, junto con el desarrollo de marcadores moleculares podrían contribuir al proceso de mejoramiento de la EUN a través de la selección asistida por marcadores.

Figura 1. Localización de los QTL en los 10 cromosomas de maíz detectados por MIC. El tamaño de las barras es proporcional a la longitud del intervalo de confianza de cada QTL en cM. Las barras negras corresponden a los QTL detectados con los datos del experimento 1 y las barras rojas corresponden a los QTL detectados con los datos del experimento 2. Los QTL detectados para EUN se remarcan en celeste.



ASI: intervalo antes-silking; EUN: eficiencia de uso de N; ICN: índice de cosecha de N; NabsMF: N absorbido a madurez fisiológica; N/B: proporción de N en la biomasa a madurez fisiológica; Pr: porcentaje de proteína en grano.

Hipótesis

Existe variabilidad alélica detectable a través de marcadores SSRs e INDEL dentro de las zonas adyacentes a los QTL identificados para EUN en la población de RILs derivada del cruzamiento B100xLP2.

Objetivo general

Aumentar la precisión de la localización de QTL para EUN detectados en los cromosomas 1, 8 y 9 de maíz sobre el mapa genético elaborado a partir de la población de 181 RILs derivada del cruzamiento B100xLP2 aumentando el número de marcadores ensayados.

Objetivos específicos

- Identificar marcadores SSR públicos y marcadores INDEL polimórficos dentro de los intervalos correspondientes a los QTL detectados.
- Desarrollar marcadores SSR polimórficos nuevos dentro de los intervalos correspondientes a los QTL detectados.
- Ubicar los marcadores polimórficos en un mapa genético a partir de la población de 181 RILs mencionada.
- Analizar la asociación entre los marcadores y los caracteres relacionados a la EUN.
- Buscar genes candidatos dentro de los nuevos intervalos.

MATERIALES Y MÉTODOS

El material genético evaluado consistió en una población de mapeo de 181 RILs derivada del cruzamiento B100xLP2.

Las RILs se obtuvieron mediante autofecundaciones sucesivas en estado S_6 con polinización controlada. La línea B100 deriva de material amarillo dentado norteamericano y LP2 de germoplasma semiflint de origen caribeño-argentino, siendo la textura de grano de B100 semidentado a dentado y LP2 semiflint (endosperma semi cristalino) (Mandolino, 2015).

Evaluación fenotípica

Las RILs junto con los parentales fueron evaluadas fenotípicamente para EUN y caracteres relacionados durante dos años de ensayos en el campo experimental de INTA EEA Pergamino (Mandolino, 2015). Dichos datos fenotípicos fueron utilizados para la optimización del análisis de QTL.

La evaluación fenotípica se realizó en dos ensayos en condiciones contrastantes de N en el suelo:

- Exp. 1: Campaña 2009/2010 donde la disponibilidad de N en el suelo fue buena.
- Exp. 2: Campaña 2011/2012 donde el N en el suelo fue escaso.

Los caracteres evaluados fueron: RGP: rendimiento en grano por planta ($\text{g grano planta}^{-1}$), NabsMF: nitrógeno absorbido a madurez fisiológica (g N planta^{-1}), EUN: eficiencia de uso de nitrógeno (g grano g N^{-1}), ICN: índice de cosecha de nitrógeno ($\text{g N grano g N biomasa}^{-1}$), Proteína: porcentaje de proteína en grano, ASI: intervalo entre antesis y emisión de estigmas o *silking* (días), BiomMF: biomasa a madurez fisiológica (g planta^{-1}) y N/BiomMF: proporción de nitrógeno en la biomasa a madurez fisiológica.

En los experimentos en condiciones contrastantes de N en el suelo se encontró diferencia significativa entre los parentales B100 y LP2 cuando el N fue escaso, siendo LP2 significativamente superior a B100 para los caracteres número de granos por planta (NGP) y EUN , mientras que B100 fue significativamente superior a LP2 para N absorbido, rendimiento y porcentaje de proteína (Mandolino, 2015).

Evaluación genotípica

Extracción de ADN

La extracción de ADN del germoplasma (líneas parentales y 181 RILs) se realizó a partir de hojas de plantas adultas de maíz. Se tomaron muestras de cinco plantas de cada genotipo. El tejido vegetal se liofilizó y se molió de manera conjunta para obtener bulks. Para cada bulk, se extrajo el ADN genómico utilizando el método de Kleinhofs *et al.* (1993). Para ello, al material vegetal molido se agregó 600µl de buffer de extracción (Tabla 1) por muestra, se incubó en baño de agua a 65°C durante 30 minutos. Luego se realizaron dos agregados de cloroformo para remover las proteínas. En cada uno, se agregó 750 µl de cloroformo mezclando suavemente mediante inversiones durante 10 minutos, se centrifugó a 12.000 rpm durante 30 minutos y se transfirió el sobrenadante a un microtubo nuevo. Se precipitó el ADN con isopropanol frío, agregando el 60% del volumen y mezclando suavemente. Luego se centrifugó 10 minutos a 12.000 rpm (revoluciones por minuto) y se descartó el sobrenadante. Para el lavado del pellet se agregó 500µl de etanol 70%, se centrifugó y se descartó el sobrenadante (dos veces). El pellet se dejó secar hasta la evaporación total del etanol. Por último se resuspendió el ADN con 100µl de buffer TE 1X: 10mM Tris (Tris-hidroximetil-aminometano) pH 8 y 1mM EDTA (Ácido etildiaminotetraacético).

Tabla 1. Componentes del buffer de extracción.

Componentes	Concentración final	1X
1M TrisClH (pH 8)	100mM	75µl
5M NaCl	700mM	105µl
0.5M EDTA (pH 8)	50mM	75µl
*SDS 20%	2%	75µl
*CTAB 10%	1%	75µl
H ₂ O		345µl
Volumen total		750µl

*Agregado después de calentar 10 minutos a 65°C.

TrisClH: Tris-hidroximetil-aminometano (pH=8, ajustado con ácido clorhídrico); NaCl: Cloruro de sodio; EDTA: Ácido etildiaminotetraacético; SDS: Dodecilsulfato sódico; CTAB: Bromuro de hexadeciltrimetilamonio.

Cuantificación de ADN

El ADN extraído se cuantificó en geles de agarosa al 0,8% teñido con bromuro de etidio (0,1µg/ml). En cada pocillo del gel se sembró un volumen de 2µl de muestra con 6 µl de azul de bromofenol (BPB). La electroforesis se llevó a cabo a un voltaje constante de 150 Volts/cm aproximadamente durante 30 minutos en buffer TAE 1X (40mM Tris, 5mM de ácido acético glacial (NaOAc) y 0,7mM EDTA, en un volumen final de 100ml de agua destilada y ajustado a pH 8.0 con ácido acético glacial). La visualización del ADN se realizó con luz ultravioleta en un transiluminador UVP Modelo M-20. La concentración de ADN se determinó visualmente y como referencia se utilizaron muestras de ADN con concentraciones conocidas.

Análisis de los marcadores moleculares

Búsqueda de marcadores moleculares en base de datos

Se seleccionaron 21 marcadores SSRs públicos para acortar intervalos localizados entre los marcadores flanqueantes a los QTL en los cromosomas 1, 8 y 9 previamente detectados asociados a la EUN. Para ello, se utilizó la base de datos pública Maize Gene Data Base (MGDB) <http://www.maizegdb.org/> y se obtuvo información de la ubicación de los SSRs en los cromosomas, el tipo y composición de la repetición y la secuencia de los oligonucleótidos.

Se seleccionaron 116 marcadores INDEL dentro de la región de los QTL, según Settles *et al.* (2014), con diferencias en la longitud de pb (pares de bases) mayor a 7 para que pudieran ser identificados fácilmente en geles de poliacrilamida.

Se buscaron SSRs nuevos dentro de la región de los QTL, utilizando el programa WebSat (Martins *et al.*, 2009) que permite identificar secuencias repetidas en tándem. Se seleccionaron regiones sucesivas de hasta 150.000 pb dentro del QTL en la base de datos Gramene <http://gramene.org/> y se diseñaron 12 cebadores correspondientes con la aplicación Primer Blast de la base de datos del NCBI <http://ncbi.nlm.nih.gov/>. Los SSRs diseñados se detallan en el Anexo 1.

Detección de marcadores moleculares: SSR e INDEL

Se evaluaron las líneas parentales B100 y LP2 con marcadores SSRs e INDEL en las siguientes regiones cromosómicas donde se identificaron los QTL asociados a EUN del genoma de maíz: cromosoma 1 entre los bin 1.03-1.05, cromosoma 8 entre los bin 8.03-8.04 y cromosoma 9 entre los bin 9.03-9.05.

Los SSRs e INDEL se amplificaron mediante la técnica de PCR, en placas de 96 pocillos en un volumen de reacción de 13µl. La mezcla de reacción consistió en 30ng de

ADN, 250nM de cada oligonucleótido (cebador *forward* y *reverse*), 200 μ M de cada uno de los desoxinucleótidos trifosfato (dNTPs), 1,5mM de cloruro de magnesio ($MgCl_2$), 0,5 μ l de Taq polimerasa, buffer de PCR 1X y agua bidestilada estéril para llegar al volumen final de reacción.

Las reacciones de amplificación se llevaron a cabo utilizando un programa de ciclado del tipo *touchdown* (Don *et al.*, 1991):

- Desnaturalización inicial: 94°C durante 5 minutos;
- 10 ciclos:
 - Desnaturalización: 94°C durante 30 segundos;
 - Hibridación: 65-55°C (decrece 1°C/ciclo) durante 30 segundos;
 - Extensión: 72°C durante 30 segundos;
- 25 ciclos:
 - Desnaturalización: 94°C durante 30 segundos;
 - Hibridación: 55°C durante 30 segundos;
 - Extensión: 72°C durante 30 segundos;
- Extensión final: 72°C durante 10 minutos.

Los productos de PCR se separaron en geles desnaturalizantes de poliacrilamida: acrilamida/bis-acrilamida (19/1) al 6% m/v y urea 7M. Para la polimerización de los geles se usó TEMED (tetrametiletilendiamina) al 0,01% y persulfato de amonio en TBE 1X al 0,035%. El buffer TBE se preparó con Tris Base 0,45M; ácido bórico 0,45M y EDTA 2mM; ajustándose a pH 8.

Los vidrios recibieron un tratamiento con la finalidad de adherir el gel a uno sólo de ellos. Para esto, se trató el vidrio con solución adherente (3 μ l *Bind Silane*, 5 μ l ácido acético glacial 0,5%, 950 μ l etanol) y la cuba se trató con solución hidrofóbica (*Rain X*[®]). La cuba y el vidrio fueron ensamblados usando espaciadores de 0,4 mm de espesor y posteriormente se introdujo entre ellos 60ml de la solución de poliacrilamida 6% con 35 μ l

de TEMED y 350µl de persulfato de amonio 10%. Se dejó polimerizar aproximadamente durante dos horas. Se emplearon peines de 97 calles.

El gel fue sometido a una pre-corrída en buffer TBE de dos concentraciones distintas (0,5X arriba y 1X abajo). Previo a la siembra, se agregó a las muestras amplificadas buffer de carga (200µl EDTA 0,5M, 9,6ml formamida y 300µl colorante-azul bromofenol y xilen cianol) y se desnaturalizaron por calentamiento a 94°C durante 5 minutos y enfriando posteriormente en hielo. El volumen de siembra fue de 2 µl por calle.

Los geles se corrieron en cubas de electroforesis a una potencia constante de 55W aproximadamente una hora y media.

Posteriormente, los patrones de banda se visualizaron mediante tinción con nitrato de plata (AgNO_3) siguiendo el protocolo para secuenciación de Promega (2007).

Finalizada la corrida electroforética, se separó la cuba del vidrio, quedando el gel adherido a él. Se colocó el vidrio con el gel en una solución fijadora compuesta por ácido acético 10% durante 15 minutos en agitación, hasta que el colorante xilencianol no fuera visible, con el fin de prevenir la difusión de los fragmentos de ADN y remover restos de buffer de electroforesis y la urea. Luego se realizó el lavado del gel mediante 3 enjuagues con agua bidestilada de 2 minutos cada uno para eliminar el ácido acético, la urea y el buffer TBE. Se transfirió el gel a una solución de tinción (AgNO_3 0,1% m/v y formaldehído 0,056% m/v) incubándolo durante 30 minutos en agitación suave. Luego se realizó un lavado sumergiendo el vidrio durante 3 segundos en agua bidestilada para remover el exceso de plata e inmediatamente se reveló colocando el vidrio en una solución alcalina de carbonato de sodio 3% m/v, formaldehído 0,056% m/v y tiosulfato de sodio 0,2mg/ml, previamente refrigerada entre 4-10°C. En estas condiciones de pH, los iones de plata depositados sobre los ácidos nucleicos se reducen a iones metálicos por acción del formaldehído. Una vez que se visualizaron nítidamente las bandas se detuvo la reacción

mediante el agregado de ácido acético 10%. Finalmente, el gel se enjuagó dos veces con agua destilada durante 2 minutos y se dejó secar a temperatura ambiente.

La lectura del gel se realizó manualmente utilizando un transiluminador de luz blanca. Se consideró "A" al alelo presente en el padre B100 y "B" al alelo presente en el padre LP2. Las imágenes se almacenaron digitalmente utilizando un Scanner Astra 2400slt.

Los marcadores SSR e INDEL evaluados se clasificaron en:

- Marcadores polimórficos: Revelaron diferencias en la secuencia de ADN de la región amplificada.
- Marcadores monomórficos: No revelaron diferencias en la secuencia de ADN de la región amplificada.
- Marcadores dudosos: El patrón de bandas no pudo identificarse claramente.
- Marcadores sin producto de amplificación.

Los marcadores polimórficos entre las líneas parentales fueron seleccionados y usados para evaluar la población de RILs.

La línea de referencia secuenciada B73 (Schnable *et al.*, 2009) se utilizó como control de peso molecular (PM) de los SSRs e INDEL. A partir de esta información se pudo establecer el PM de los alelos de cada amplificación en B73 y estimar el tamaño probable del alelo de cada SSR en B100 y LP2.

Caracterización genotípica de la población de mapeo

Luego del reconocimiento del polimorfismo de los parentales, se seleccionaron los marcadores cuya detección eficiente y reproducible permitieron ser aplicados para caracterizar genotípicamente la población de 181 RILs S₆ derivada del cruzamiento B100xLP2. A partir de cada muestra de ADN genómico extraído de cada RIL se realizó la amplificación de los marcadores mediante PCR, la separación de los productos de PCR a

través de electroforesis en geles de poliacrilamida y la visualización mediante tinción con nitrato de plata, siguiendo los mismos procedimientos indicados para los genotipos parentales. Se registró el genotipo de cada RIL, asignando “A” a los homocigotas para el alelo proveniente del padre B100, “B” a los homocigotas para el alelo proveniente del padre LP2 y “H” a los heterocigotas.

Análisis de segregación de los marcadores

El análisis de segregación de los marcadores SSR e INDEL se llevó a cabo utilizando el programa GQMol Versión 2008 6.1 (Cruz & Schuster, 2008).

La segregación mendeliana 1:1 para una población de RILs de los marcadores SSR e INDEL fue evaluada mediante la prueba tasa de falsos descubrimientos o FDR (del inglés *False Discovery Rate*). Este método estadístico es adecuado cuando se quiere comparar múltiples hipótesis. En una lista de hipótesis rechazadas el criterio FDR controla la proporción esperada de hipótesis nulas incorrectamente rechazadas (Benjamini & Hochberg, 1995). Se consideró un nivel de significancia global del 5%.

Mapeo de QTL

Construcción de los mapas de ligamiento

El mapa de ligamiento indica la posición y la distancia genética relativa entre marcadores a lo largo del cromosoma a partir de los valores de frecuencia de recombinación observados entre ellos. La construcción de los mapas de ligamiento se llevó a cabo utilizando en programa GQMol Versión 2008 6.1.

Los marcadores que se usaron para la construcción de los mapas de ligamiento fueron los que segregaron según lo esperado. De este modo, para la construcción de los mapas de ligamiento se usaron 77 marcadores.

Las frecuencias de recombinación fueron transformadas en unidades de mapeo (centiMorgan - cM) utilizando la función de mapeo desarrollada por Kosambi (1944), la cual tiene en cuenta la interferencia. La interferencia se relaciona con la imposibilidad física de que adyacente a un punto de entrecruzamiento se produzca otro en la misma meiosis debido a un impedimento físico entre ellos.

El método usado para seleccionar el mejor orden de los marcadores en el mapa fue el de máxima verosimilitud. Los mapas de ligamiento obtenidos fueron graficados con el programa Map Chart 2.2 (Voorrips, 2006).

Análisis de QTL

Se realizó el análisis de asociación entre los marcadores moleculares y el fenotipo.

De los datos fenotípicos evaluados para EUN y caracteres relacionados durante dos años de ensayo (campaña 2009/2010 y 2011/2012), los caracteres utilizados para la optimización del análisis de QTL fueron: porcentaje de proteína en grano (Proteína), índice de cosecha de nitrógeno (ICN), proporción de nitrógeno en biomasa a madurez fisiológica (N/BiomMF), eficiencia de uso de nitrógeno (EUN) y rendimiento en grano por planta (RGP).

La detección de QTL se realizó mediante el programa Win QTL Cartographer V2.5 (Wang *et al.*, 2013) mediante el método de mapeo por intervalo compuesto, para cada año y cada carácter por separado. El análisis se realizó usando el modelo 6 del procedimiento Zmapqtl, con una velocidad de mapeo de 1 cM, un número de marcadores como cofactores de 5 y un tamaño de ventana de 10 cM.

En el método de MIC se estimaron, para cada posición en el genoma, la presencia de un QTL, los efectos de aditividad y la proporción de la variancia fenotípica explicada por el QTL (R^2).

Para identificar la presencia de un QTL el valor de LOD umbral considerado fue 3 y además se utilizó una herramienta del Win QTL Cartographer que localiza los QTL automáticamente. Esta herramienta usa como criterio para considerar la existencia de un QTL una distancia mínima entre dos QTL adyacentes de al menos 5 cM (*Min distance between QTLs*) y un valor de LOD entre el punto más alto y la base del pico de QTL de al menos 1 (*Min LOD from top valley*).

La posición del QTL fue asignada como el punto con valor de LOD máximo y el intervalo de confianza se calculó substrayendo una unidad de LOD a cada lado de la posición de *LOD score* máximo.

Identificación de genes candidatos

Se identificaron posibles genes candidatos dentro de los intervalos de los QTL detectados por el programa Win QTL Cartographer V 2.5. Para ello se determinaron las coordenadas en pb de los marcadores flanqueantes usando la aplicación Locus Lookup en la base de datos MGDB (B73 RefGen_v2/v3 genome sequence). Una vez conocidas la posiciones físicas de los intervalos se identificaron los genes putativos dentro de dichas regiones, basados en la anotación de B73 RefGen_v4 genome sequence, usando la aplicación BioMart en la base de datos Gramene <http://gramene.org/>. La función de cada gen fue confirmada mediante análisis de genes ortólogos en *Arabidopsis thaliana* u *Oryza sativa*. Los genes dentro de cada intervalo fueron clasificados según la función predicha en seis categorías principales basado en el criterio establecido por *Munich Information Center for Protein Sequences* (Ruepp *et al.*, 2004). Pueden existir pequeñas diferencias en los intervalos y los genes candidatos identificados, ya que las aplicaciones Locus Lookup en la base de datos MGDB y BioMart en la base de datos Gramene usan versiones distintas de secuenciación del genoma de maíz.

RESULTADOS

Caracterización genotípica

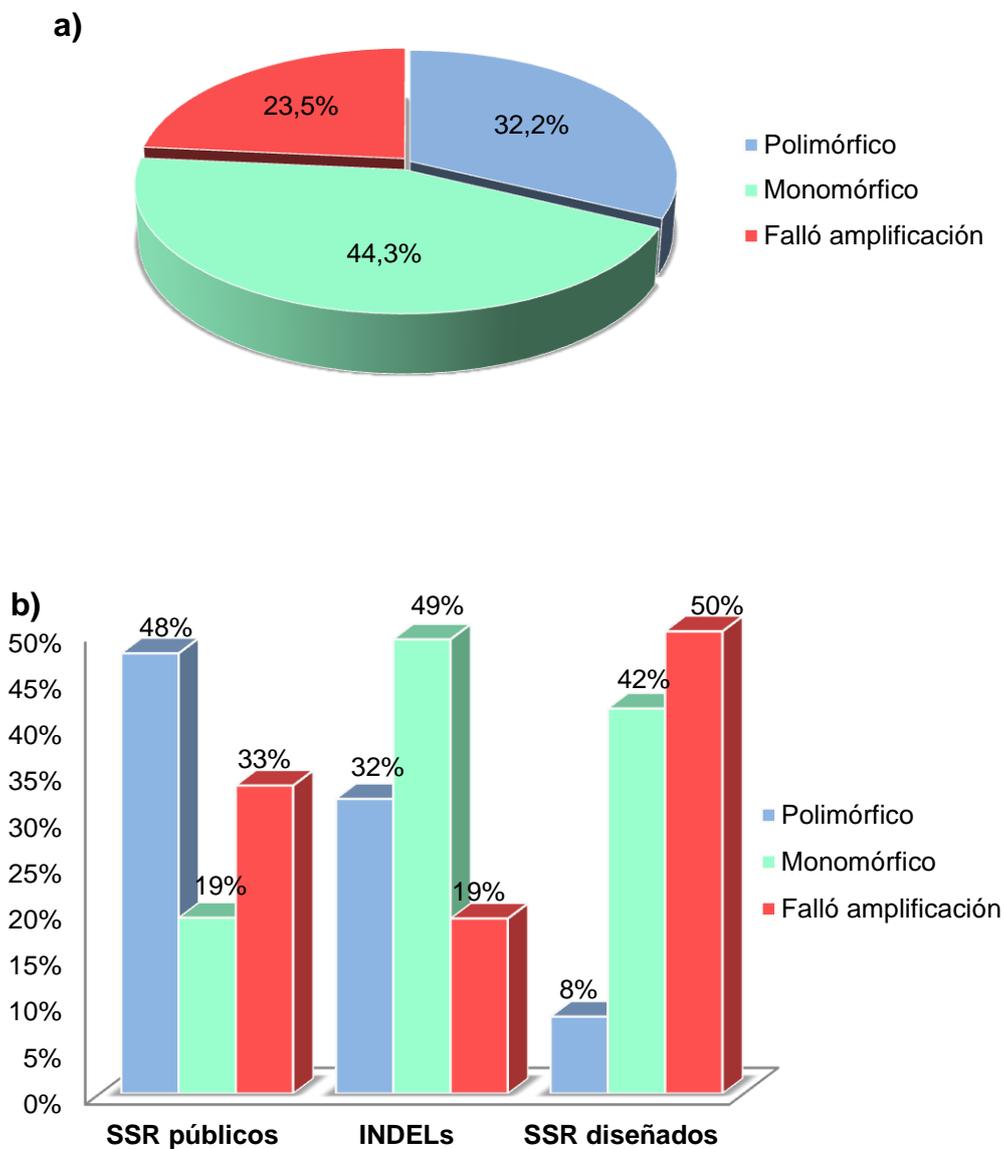
Las líneas parentales B100 y LP2 se probaron con marcadores SSR e INDEL distribuidos en los cromosomas 1, 8 y 9 de maíz, en las regiones cromosómicas donde se identificaron los QTL asociados a EUN. La línea B73 se utilizó como control del PM de los marcadores moleculares ya que se encuentra secuenciada.

En la prueba de polimorfismo se incluyeron las dos líneas parentales y la línea B73, cada una por duplicado. Además, en la etapa de visualización del gel se incluyó el marcador de PM de 25 pb de Invitrogen, para poder diferenciar las bandas de marcadores moleculares de las bandas sombras o espejos, que corresponden a fragmentos espurios o inespecíficos amplificados durante la reacción de PCR.

Se probaron un total de 149 marcadores en las líneas parentales, de los cuales 21 corresponden a SSR públicos, 12 SSR diseñados y 116 INDEL.

En la figura 2 se muestran los porcentajes de cada clase de los distintos marcadores moleculares.

Figura 2. Frecuencia de distribución de los resultados de la prueba de polimorfismo entre las líneas parentales B100 y LP2. a) Frecuencia de distribución para el total de los marcadores usados. b) Frecuencia de distribución para los distintos tipos de marcadores usados.



De los 149 marcadores moleculares que se evaluaron, el 32,2% resultó polimórfico, el 44,3% monomórfico y el 23,5% falló en la amplificación (Figura 2a). Asimismo, se analizó la distribución de los resultados de la prueba de polimorfismo entre los parentales para cada tipo de marcador en particular (Figura 2b). De los marcadores SSR públicos probados casi el 50% resultaron polimórfico entre B100 y LP2. Sin embargo, casi el 50%

de los INDEL evaluados resultaron monomórficos. Para el caso de los SSRs diseñados, el 50% falló en la amplificación y más del 40% fue monomórfico entre los parentales.

Del total de los marcadores probados, en el cromosoma 1 se evaluaron 29 marcadores, de los cuales 8 resultaron ser polimórficos, en el cromosoma 8 se evaluaron 32 marcadores de los cuales 11 fueron polimórficos y en el cromosoma 9 se encontraron 29 marcadores polimórficos de un total de 88 evaluados (Tabla 2).

Tabla 2. Número y porcentaje de marcadores moleculares utilizados en los parentales B100 y LP2.

Cromosoma	Polimórficos		Monomórficos		Falló amplificación		Total	
	Número	Porcentaje	Número	Porcentaje	Número	Porcentaje	Número	Porcentaje
1	8	27,6%	15	51,7%	6	20,7%	29	19,5%
8	11	34,4%	16	50,0%	5	15,6%	32	21,5%
9	29	33,0%	35	39,8%	24	27,3%	88	59,1%
Total	48	32,2%	66	44,3%	35	23,5%	149	100,0%

Los 48 marcadores polimórficos resultantes fueron corridos en la población de 181 RILs. Ocho de los marcadores amplificaron sólo en unos pocos individuos por lo que no se tuvieron en cuenta para realizar el análisis de segregación. Estos fueron N011: *bnlg1626*, Chr1-46232560, Chr1-47774900, Chr8-33195140, Chr9-48269420, Chr9-98560200 y DEL.4697.

Análisis de segregación de los marcadores

Se evaluó la segregación 1:1 de la población RILs, mediante el método estadístico *False Discovery Rate* (FDR), considerando un nivel de significancia global del 5%.

Se evaluó la segregación de 94 marcadores, 40 de ellos corridos en la población durante este trabajo y 54 marcadores SSRs y SNPs (del inglés *Single Nucleotide Polymorphism*) corridos en la población de 181 RILs en estudios anteriores (Mandolino,

2015) ubicados en los cromosomas 1, 8 y 9. Las secuencias de oligonucleótidos de los marcadores moleculares utilizados para el análisis se detallan en el Anexo 2.

De los 94 marcadores, 77 presentaron segregación 1:1. De los 17 marcadores que segregaron de manera distorsiva, siete correspondieron al cromosoma 1, tres al cromosoma 8 y siete al cromosoma 9.

La segregación de los 40 marcadores corridos en la población durante este trabajo se detalla en la Tabla 3. De los 40 marcadores, seis fueron distorsivos. La información del análisis de segregación de los 54 marcadores corridos en trabajos anteriores no se muestra. Para estos, 43 marcadores presentaron segregación 1:1 mientras que 11 segregaron de manera distorsiva (Mandolino, 2015).

La mayoría de los marcadores distorsivos fueron SSRs. Tres INDEL mostraron segregación distorsiva: Chr1-46629720, Chr8-12684300, INS.116034. De los SNPs sólo el PZB01385.3 en el cromosoma 8 tuvo segregación distorsiva.

Tabla 3. Segregación de los marcadores moleculares en la población de RILs.

Marcador	Homocigotas		Homocigotas LP2	Datos perdidos	Prueba de segregación		
	B100	Heterocigotas			X ²	Probabilidad (%)	
Chr1-46628500	72	16	91	2	2,215	13,67	ns
Chr1-46629720	51	6	98	26	14,826	0,0118	*
phi001	85	4	92		0,277	59,8782	ns
Chr1-47770200	69	19	88	5	2,299	12,9427	ns
umc2217	74	25	82		0,410	52,1839	ns
umc2229	46	38	97		18,189	0,002	*
Chr8-12684300	140	1	39	1	56,989	0,000	*
Chr8-13272600	75	19	87		0,889	34,5779	ns
DEL.131996	81	21	79		0,025	87,4367	ns
Chr8-14542640	78	23	78	2	0,000	100,0	ns
Chr8-17067440	81	25	74	1	0,316	57,3943	ns
DEL.132563	75	23	83		0,405	52,4486	ns
DEL.132932	88	19	74		1,210	27,1357	ns
DEL.133392	83	23	74	1	0,516	47,2586	ns
DEL.133721	85	20	74	2	0,761	38,3014	ns
umc1735	82	17	82		0,000	100,0	ns
umc2338	71	43	67		0,116	73,3478	ns

Tabla 3. Continuación.

Marcador	Homocigotas		Homocigotas LP2	Datos perdidos	Prueba de segregación		
	B100	Heterocigotas			X ²	Probabilidad (%)	
umc2370	80	25	76		0,103	74,8774	ns
Chr9-27968120	73	16	90	2	1,773	18,301	ns
DEL.144557	71	21	87	2	1,620	20,3056	ns
umc2337	62	17	101	1	9,331	0,2253	*
INS.116020	72	27	82		0,649	42,0345	ns
INS.116034	45	26	106	4	24,642	0,0001	*
DEL.145171	82	14	85		0,054	81,6424	ns
Chr9-45015160	84	17	78	2	0,222	63,7352	ns
DEL.145752	65	21	94	1	5,289	2,1457	ns
Chr9-56424320	79	16	86		0,297	58,5788	ns
DEL.146397	69	34	78		0,551	45,7901	ns
Chr9-74589140	65	23	92	1	4,643	3,1175	ns
Chr9-74592980	68	18	93	2	3,882	4,8807	ns
DEL.147207	72	20	89		1,795	18,0314	ns
Chr9-85234820	93	9	75	4	1,929	16,4915	ns
DEL.148060	71	19	91		2,469	11,6102	ns
DEL.148101	63	18	92	8	5,426	1,9841	ns
umc1191	88	3	89	1	0,006	94,0084	ns
DEL.148523	67	20	93	1	4,225	3,9833	ns
Chr9-98676780	67	16	97	1	5,488	1,915	ns
umc1743	68	6	105	2	7,913	0,4907	*
DEL.148894	70	18	92	1	2,988	8,3902	ns
DEL.149169	64	20	96	1	6,400	1,1412	ns

Nivel crítico: 0,87%; Criterio: FDR; *: Valores por debajo del nivel crítico especificado; ns: no difiere significativamente de la segregación 1:1.

Mapeo de QTL

Mapas de ligamiento

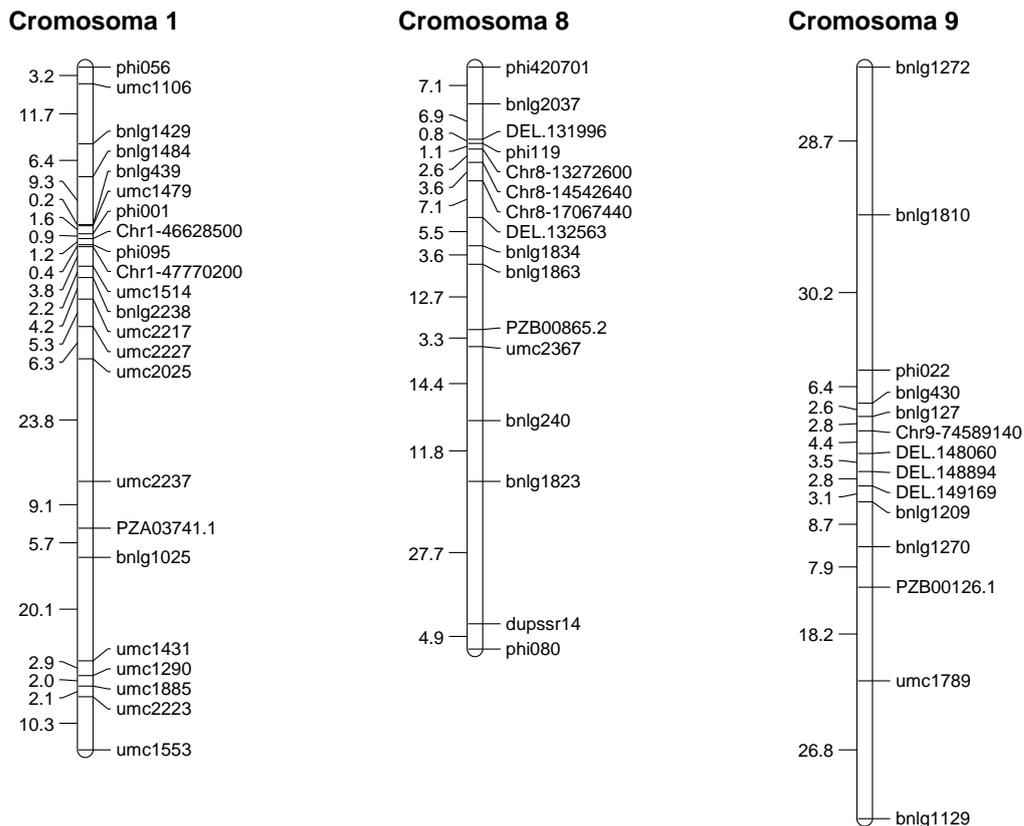
Se construyeron los mapas de ligamiento de los cromosomas 1, 8 y 9 de maíz (Figura 3). De los 77 marcadores con segregación 1:1, 53 fueron usados para la construcción de los mapas, los 24 restantes fueron eliminados por no asociarse con otros marcadores dentro de sus respectivos grupos de ligamiento, como *bnlg504* y *umc1735*, o porque el orden del mapa no coincidió con el orden esperado.

Los marcadores se agruparon en tres grupos de ligamiento, uno por cromosoma. En el cromosoma 1 se cubrió una distancia genética total de 132,7 cM, con una separación

promedio entre marcadores de 5,8 cM. En el cromosoma 8 se cubrió una distancia total de 113,1 cM, con una separación promedio entre marcadores de 7,1 cM. En el cromosoma 9 se cubrió una distancia total de 146,1 cM, con una separación promedio entre marcadores de 10,4 cM.

El orden de los marcadores en el mapa de la población de RILs se correlacionó con lo esperado según la posición en el mapa físico.

Figura 3. Mapas de ligamiento de los cromosomas 1, 8 y 9 de maíz mostrando la localización de los 53 marcadores moleculares basados en la población de RILs. A la derecha de cada cromosoma se indica el nombre de los marcadores y a la izquierda la distancia en cM entre marcadores.



Análisis de QTL

Se detectaron 11 QTL para los caracteres Proteína, ICN, N/BiomMF, EUN y RGP mediante MIC (Tabla 4). La localización de los QTL en los cromosomas detectados por MIC se muestra en la Figura 4.

El porcentaje de variancia fenotípica explicada por los QTL varió entre 5,9 y 12,6%. Para el carácter porcentaje de proteína, sólo se detectó un QTL en el cromosoma 1 (Pr-1) para el primer ensayo (Exp. 1), el cual presentó efecto aditivo negativo. El QTL Pr-1 explicó el 8,4% de la variancia fenotípica.

Para el carácter ICN se identificaron dos QTL, uno en el cromosoma 8 (ICN-1a) para el Exp. 1 y el otro localizado en el cromosoma 9 (ICN-1b) para el Exp. 1. El QTL ICN-1a presentó efecto aditivo negativo, mientras que el QTL ICN-1b presentó efecto aditivo positivo. El QTL ICN-1a explicó el 7,2% de la variancia fenotípica, mientras que el ICN-1b el 12,6%.

Para el carácter N/BiomMF se detectó un QTL en el cromosoma 8 (N/Biom-1) en el Exp. 1, el cual presentó efecto aditivo negativo. El QTL N/Biom-1 explicó el 10,3% de la variancia fenotípica.

Para el carácter EUN se identificaron cinco QTL, cuatro QTL para el Exp. 1: un QTL en el cromosoma 1 (EUN-1a) y tres QTL localizados en el cromosoma 9 (EUN-1b, EUN-1c y EUN-1d) y uno para el segundo ensayo (Exp. 2) en el cromosoma 1 (EUN-2). Todos los QTL presentaron efectos aditivos positivos. Los QTL EUN-1a y EUN-2 se encuentran en el mismo cromosoma y comparten parcialmente el intervalo de confianza.

Para el carácter RGP se identificaron dos QTL, uno en el cromosoma 8 (RGP-1) para el Exp 1 y el otro en el cromosoma 9 (RGP-2) para el Exp 2. El QTL RGP-1 presentó efecto aditivo negativo y el RGP-2 efecto aditivo positivo. El QTL RGP-1 explicó el 8,6% de la variancia fenotípica, mientras que RGP-2 el 6,7%.

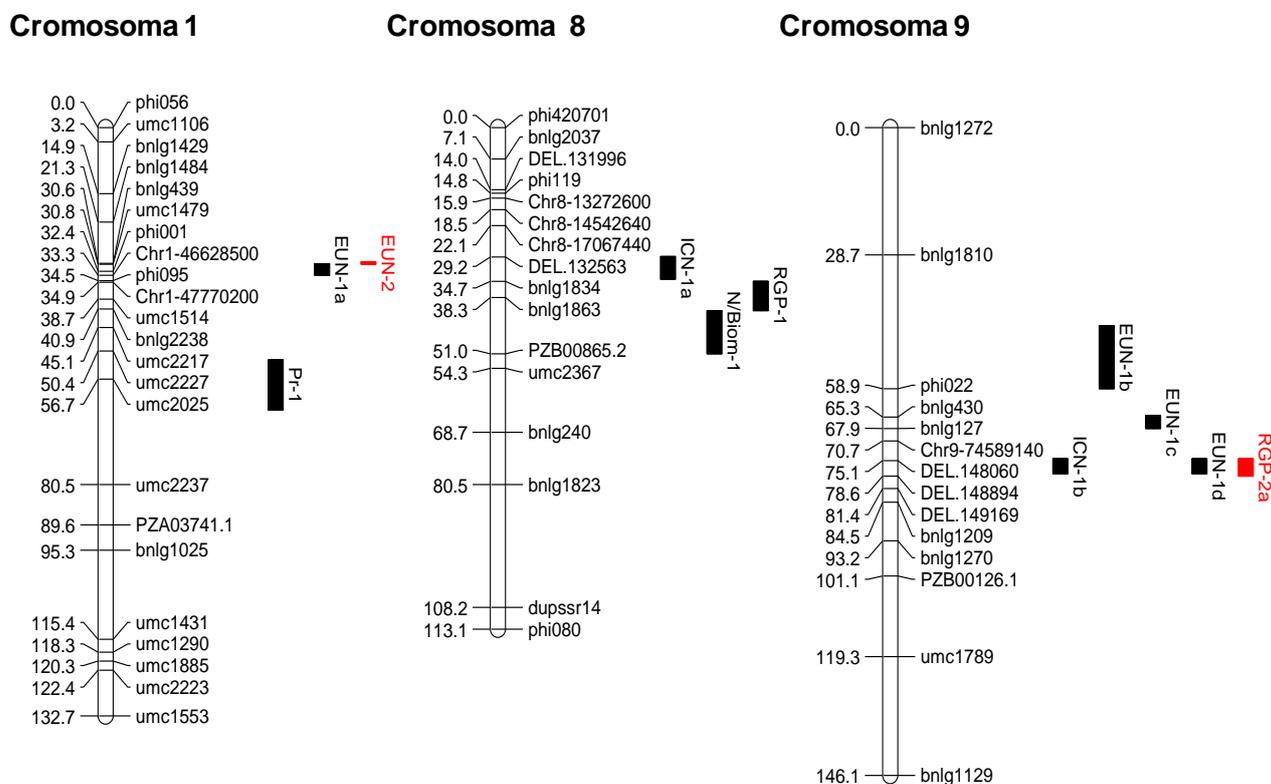
Los QTL EUN-1d, ICN-1b y RGP-2 comparten el intervalo de confianza y se encuentran dentro de la región de interés.

Tabla 4. QTLs detectados mediante el método de Mapeo por Intervalo Compuesto.

QTL	Año ^a	Posición del QTL					LODmax	Efecto Genético	
		Cr	cM ^b	IC ^c (cM)	IC (Mpb)	MM Flanqueantes		R ²	A
EUN									
EUN-1a	Exp 1	1	31.8	30.7-33.3	35.3-46.6	bnlg1484/Chr1-46628500	3.03	5.9	1.11
EUN-1b		9	58.7	44.7-58.9	9.12-23.3	bnlg1810-phi022	4.96	10.1	1.45
EUN-1c		9	67.3	64.9-67.9	23.3-51.7	phi022/bnlg127	5.56	11.4	1.55
EUN-1d	Exp 2	9	75.1	74.7-78.1	77.6-93.0	Chr9-74589140/DEL.148060	6.10	12.1	1.58
EUN-2		1	30.7	30.3-30.8	35.3-43.1	bnlg1484/umc1479	3.29	6.8	2.26
ICN									
ICN-1a	Exp 1	8	30.2	29.1-34.2	17.1-21.0	Chr8-1706740/DEL.132563	3.67	7.2	-1.90
ICN-1b		9	75.1	74.7-78.1	77.6-93.0	Chr9-74589140/DEL.148060	6.50	12.6	2.45
N/BiomMF									
N/Biom-1	Exp 1	8	47.4	41.3-51.0	91.6-125.0	bnlg1863/PZA0865.2	3.82	10.3	-0.18
Proteína									
Pr-1	Exp 1	1	56.4	52.4-63.7	67.8-91.2	umc2227/umc2025	3.73	8.4	-0.15
RGP									
RGP-1	Exp 1	8	37.7	34.7-41.3	64.0-91.6	bnlg1834/bnlg1863	4.04	8.6	-3.05
RGP-2	Exp 2	9	75.1	74.7-78.6	77.6-105.0	Chr9-74589140/DEL.148894	3.27	6.7	2.06

Cr: Cromosoma; LODmax: máximo valor de LOD para el QTL; MM: Marcadores Moleculares; ^a Sólo aquellos con dato; ^b Posición del valor de LODmax en cM; ^c Intervalo de confianza calculado ± 1 LOD de LODmax; A: Efecto aditivo con respecto al alelo del parental LP2; R² proporción de la varianza fenotípica explicada por el QTL. En los nombres de los QTL el número indica si corresponde al Exp 1 o Exp 2, la letra enumera los diferentes QTL para un determinado carácter en el Exp.

Figura 4. Localización de los QTL en los cromosomas 1, 8 y 9 de maíz detectados por MIC. El tamaño de las barras es proporcional a la longitud del intervalo de confianza de cada QTL en cM. Las barras negras corresponden a los QTL detectados con los datos del Exp. 1 y las barras rojas corresponden a los QTL detectados en el Exp. 2.



Identificación de genes candidatos

Se identificaron 3155 genes distintos dentro de los intervalos de los QTL detectados y se clasificaron según su función putativa en seis categorías principales (Tabla 5) basado en el criterio establecido por el centro de investigación MIPS (*Munich Information Center for Protein Sequences* - Ruepp *et al.*, 2004). El 67% de los genes identificados correspondieron a proteínas con función putativa, mientras que el 33% restante aún no tienen una función caracterizada (Figura 5).

Figura 4. Clasificación de los genes identificados dentro de los intervalos de los QTL detectados.

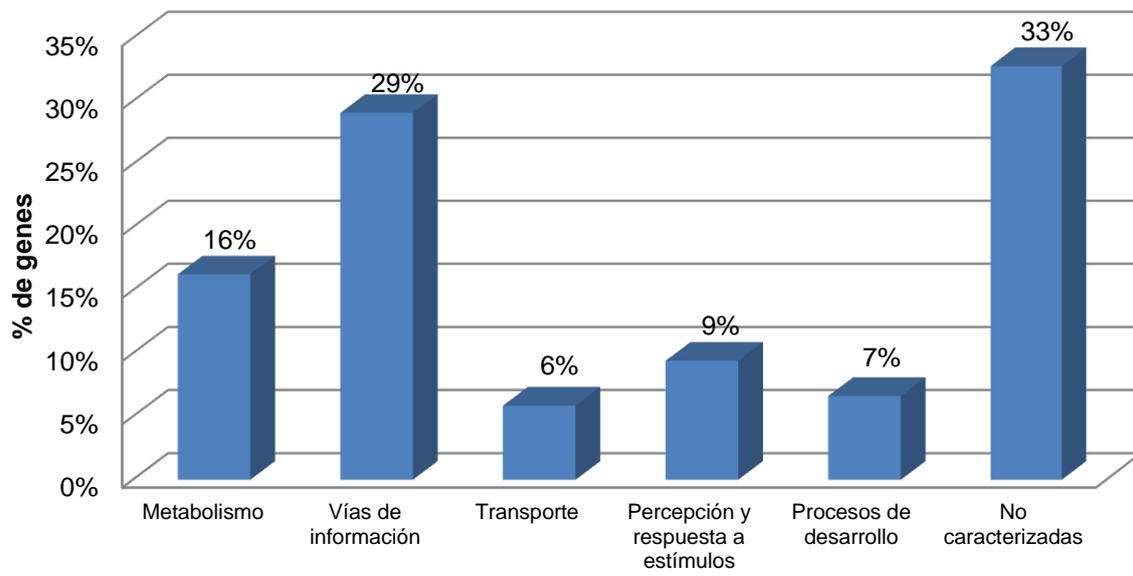


Tabla 5. Número de genes identificados dentro de los intervalos de cada uno de los QTL detectados, clasificados según su función putativa.

Principales categorías funcionales	Subcategorías	QTL										
		EUN				ICN		N/Biom	Proteína	RGP		
		EUN-1a	EUN-1b	EUN-1c	EUN-1d	EUN-2	ICN-1a	ICN-1b	N/Biom-1	Pr-1	RGP-1	RGP-2
Metabolismo	Energía	6	4	6	0	5	1	0	11	11	8	1
	Metabolismo	59	75	54	28	36	17	28	93	60	41	66
Vías de información	Ciclo celular y procesamiento de ADN	3	17	17	1	2	4	1	20	6	10	4
	Procesamiento de proteínas (plegado, modificación y destino)	26	41	31	6	26	5	6	38	29	29	22
	Proteína con función de unión o requisito de cofactor	19	14	15	12	14	4	12	33	35	19	25
	Regulación de la actividad proteica	3	1	0	0	2	0	0	2	3	2	2
	Síntesis de proteína	7	13	12	5	4	1	5	20	7	16	11
	Transcripción	25	57	46	25	21	12	25	58	40	58	55
Transporte	Transporte	12	23	20	13	6	6	13	48	19	28	29
Percepción y respuesta a estímulos	Interacción con el ambiente	0	3	4	0	0	1	0	5	3	7	0
	Mecanismo de comunicación celular/transducción de señales	22	13	19	7	10	3	7	25	12	15	11
	Rescate celular, defensa y virulencia	18	18	19	5	11	7	5	29	34	18	11
Procesos de desarrollo	Biogénesis de componentes celulares	1	3	6	3	1	3	3	18	6	22	5
	Desarrollo	11	9	26	2	8	7	2	22	19	15	2
	Destino celular	0	1	3	0	0	0	0	6	6	1	0
	Diferenciación (célula, tejido, órgano)	0	0	0	5	0	1	5	1	5	1	9
No caracterizadas		85	128	150	81	55	29	81	208	130	139	165
Total		297	420	428	193	201	101	193	637	425	429	418

De los genes con función predicha, 513 genes fueron asignados a la categoría metabolismo. Dentro de esta categoría se agruparon proteínas implicadas en la síntesis y degradación de aminoácidos, nucleótidos, carbohidratos, lípidos y metabolitos secundarios, y proteínas involucradas en la obtención de energía como la glucólisis, ciclo de Krebs, respiración, fermentación, oxidación de ácidos grasos y fotosíntesis. En la categoría vías de información se asignaron 917 genes. Dentro de esta se encuentran las proteínas implicadas en el ciclo celular y procesamiento de ADN (síntesis, replicación, recombinación, reparación y modificación de ADN y proteínas involucradas en el ciclo celular y en el control del ciclo celular), el destino de las proteínas sintetizadas a través de procesos como el plegamiento y modificación de las proteínas sintetizadas, el ensamblaje de complejos proteicos y la degradación proteica, proteínas con función de unión o que requieren cofactor, proteínas que regulan la actividad proteica (regulación de la actividad enzimática, factores de transcripción y señalización), proteínas que participan en la síntesis de proteínas (proteínas ribosomales y otras proteínas involucradas en la traducción) y en la transcripción (síntesis, procesamiento, modificación de rRNA, mRNA y tRNA). Por otro lado, 185 genes corresponden a proteínas que participan en el transporte de iones, carbohidratos, aminoácidos, proteínas y lípidos. Asimismo, 297 genes fueron asignados a la categoría percepción y respuesta a estímulos como la interacción con el ambiente, la comunicación celular y transducción de señales, la defensa a estrés biótico y abiótico y la detoxificación celular. Por último, 209 genes corresponden a proteínas que participan en procesos de desarrollo como la biogénesis de componentes celulares, el destino celular (crecimiento celular, muerte celular) y la diferenciación de tipos celulares, tejidos y órganos. Si bien algunos QTL comparten parcial o totalmente el intervalo, el número total de genes por categoría se determinó considerando una sola vez cada gen.

DISCUSIÓN

Caracterización de la población

La caracterización genotípica de la población de RILs derivada del cruzamiento de B100xLP2 se realizó utilizando marcadores SSR, INDEL y SNP, siendo más del 60% SSRs y casi el 35% INDEL. Un alto porcentaje de los SSRs resultaron polimórficos coincidiendo con los resultados obtenidos por Olmos *et al.* (2014). Contrariamente, casi el 50% de los marcadores INDEL resultaron monomórficos entre los parentales, indicando que los SSRs serían más informativos que los INDEL para el estudio de esta población.

De los 94 marcadores moleculares usados para el análisis de segregación, 17 presentaron segregación distorsiva, por lo que se usaron 77 marcadores para construir el mapa de ligamiento. La segregación distorsiva se define como una desviación de las frecuencias genotípicas observadas respecto de sus valores esperados. Hay regiones en los cromosomas que tienden a segregar de manera distorsiva, llamadas SDR (del inglés *Segregation Distortion Regions*) (Vogl & Xu, 2000). Una gran variedad de factores fisiológicos y genéticos podrían ser la causa de la distorsión de la segregación. Los mecanismos para la segregación preferencial incluyen la competencia del tubo polínico, la fertilización preferencial y la eliminación selectiva de cigotos. En maíz, los factores genéticos más comunes asociados a la segregación distorsiva son factores gametofíticos (Lu *et al.*, 2002). El uso de marcadores distorsivos en la construcción del mapa genético puede alterar el orden de los marcadores y/o la longitud del mapa (Cheema & Dicks, 2009). Para la construcción del mapa genético en este trabajo sólo se usaron los marcadores que segregaron en la proporción 1:1 de acuerdo a lo esperado para una población de RILs.

Mapeo de QTL

Mapas de ligamiento

Con la información obtenida a partir de los marcadores que presentaron la segregación esperada (1:1) se construyeron los mapas de ligamiento de los cromosomas 1, 8 y 9 de maíz saturando las regiones de interés con marcadores moleculares. La incorporación de marcadores moleculares nuevos a los mapas de ligamiento de los cromosomas 1, 8 y 9 permitió una mejor representación de los cromosomas y mayor precisión en la localización de los QTL. Para todos los cromosomas se logró incrementar la longitud del mapa con respecto al obtenido previamente (Mandolino, 2015) y aumentar la cobertura genómica. En el cromosoma 1 se incorporaron cuatro marcadores nuevos en la región de interés: *phi001*, Chr1-46628500, Chr1-47770200 y *umc2217*, y la longitud del mapa del cromosoma 1 construido con 23 marcadores fue de 132,7 cM aumentando la longitud del mismo con respecto al realizado en el trabajo doctoral de Mandolino que fue de 126,4 cM construido con 19 marcadores. Para el cromosoma 8 se agregaron un total de siete marcadores, cinco corresponden a marcadores nuevos: DEL.131996, Chr8-13272600, Chr8-14542640, Chr8-17067440 y DEL.132563 y los otros dos son marcadores corridos en la población en la tesis doctoral de Mandolino 2015: *umc2367* y *bnlg1823*, debido a que anteriormente no coincidían con el orden del mapa esperado. La longitud del mapa del cromosoma 8 construido con 16 marcadores se amplió a 113,1 cM con respecto a la longitud del mapa citado de Mandolino, que fue de 105,7 cM y construido con 9 marcadores. Para el cromosoma 9 se adicionaron cuatro marcadores: Chr9-74589140, DEL.148060, DEL.148894 y DEL.149169 y la longitud del mapa se amplió a 146,1 cM con 14 marcadores, con respecto a la longitud de 120,6 cM previamente construido con 10 marcadores.

Análisis de QTL

El mapeo de QTL se realizó con el fin de determinar las nuevas localizaciones más acotadas correspondientes a los intervalos de los QTL detectados para EUN y caracteres relacionados en la población de RILs. Se identificaron 11 QTL por MIC en los cromosomas 1, 8 y 9.

En el cromosoma 1 se detectaron tres QTL, dos para EUN, uno para cada año de ensayo (EUN-1a y EUN-2). Estos QTL comparten parcialmente sus intervalos de confianza, coincidiendo en parte con la región de interés en el cromosoma 1 (entre los marcadores *phi339017* – *phi095*). El QTL EUN-2 no fue detectado automáticamente por el Win QTL Cartographer, sin embargo fue considerado como tal debido a que superó el valor umbral de LOD 3. De este modo, el QTL EUN-2 sería un QTL putativo. Debido a los antecedentes de Mandolino (2015) y a que se detectó un QTL en el ensayo anterior con los mismos datos, se considera que puede tratarse de un QTL significativo. El otro QTL detectado en el cromosoma 1 estuvo asociado a porcentaje de proteína en el primer ensayo (Pr-1), pero localizado en otra región del cromosoma que no coincide con la región de interés estudiada. Para el carácter EUN el efecto aditivo es positivo, indicando que el alelo del parental LP2 incrementa el valor de la media cuando la disponibilidad de N en el suelo es escasa.

En el cromosoma 8 se detectaron tres QTL, todos en el ensayo 1 donde la disponibilidad de N en el suelo fue buena y ninguno de ellos comparte intervalo de confianza. Los QTL detectados fueron, uno para ICN (ICN-1a) que se localizó fuera de la región de interés, tratándose de un QTL nuevo, un QTL para N/BiomMF (N/Biom-1) localizado en la misma región que el QTL para N/BiomMF detectado en la trabajo doctoral de Mandolino, y un QTL para rendimiento (RGP-1) que se localiza fuera de la región de interés, tratándose de un QTL nuevo.

El QTL para N/BiomMF, a diferencia del QTL que se detectó en el trabajo de Mandolino como un QTL que se manifiesta en los dos ambientes, en este trabajo solo se detectó en el primer ensayo, esto podría deberse a que al aumentar el número de marcadores, el cromosoma 8 se encuentra mejor representado y por lo tanto hay una mejora en la precisión del mapeo y como consecuencia desapareció el QTL en el Exp. 2.

Para los caracteres porcentaje de proteína y N/BiomMF los efectos aditivos son negativos, indicando que el alelo del parental LP2 disminuye el valor de la media.

En el cromosoma 9 se detectaron cinco QTL, todos en un solo ensayo, tres QTL para EUN en el ensayo 1 (EUN-1b, EUN-1c y EUN-1d). El QTL EUN-1d es el único localizado dentro del intervalo de interés del cromosoma 9 (entre los marcadores *bnlg127* – *bnlg1209*) y la incorporación de los marcadores SSR e INDEL permitió acotar el intervalo del QTL detectado sobre el cromosoma 9 de 55,3 Mpb a 15,4 Mpb. Por otro lado, los QTL EUN-1b y EUN-1c se detectaron fuera de la región de interés, tratándose de QTL nuevos. Asimismo, se detectó un QTL para ICN en el ensayo 1 (ICN-1b) y un QTL para rendimiento en el segundo año de ensayo (RGP-2) dentro de la región de interés. Los QTL EUN-1d e ICN-1b se encuentran dentro de la región de interés sobre el cromosoma 9 y comparten intervalo de confianza entre los marcadores Chr9-74589140 y DEL.148060, por lo que podría tratarse de un único QTL con posible efecto pleiotrópico.

Sólo el QTL para EUN en el cromosoma 1 parecería manifestarse en los dos ambientes, el resto de los QTL se expresaron en uno de los dos ambientes, indicando la importancia de los efectos ambientales sobre la expresión de los caracteres cuantitativos.

Algunos de los QTL detectados coinciden con QTL encontrados en bibliografía. El QTL para rendimiento (RGP-1) del cromosoma 8 coincide con un QTL detectado para rendimiento en grano por año en maíz en un ensayo con alto N (Liu *et al.*, 2010). El QTL ICN-1a del cromosoma 8 coincide con un QTL detectado para la fecha de emisión de estigmas (floración femenina) (Coque *et al.*, 2008). El QTL N/Biom-1 del cromosoma 8

coincide con un QTL para contenido de clorofila a los 30 días de la emisión de estigmas (carácter relacionado con la senescencia de las hojas) en un ensayo con alto N (Coque *et al.*, 2008). Los QTL EUN-1b y EUN-1c del cromosoma 9 se corresponden con dos QTL detectados para rendimiento en un ensayo con alta disponibilidad de N (Liu *et al.*, 2010). Los QTL EUN-1d, ICN-1b y RGP-2 que colocan en el cromosoma 9 coinciden con un QTL para caracteres evaluados a la madurez: índice de cosecha (rendimiento de grano/rendimiento de la planta entera) en un ensayo con alta disponibilidad de N (Coque *et al.*, 2008). El QTL para EUN del cromosoma 1 coincide con un QTL para el carácter fisiológico “actividad de glutamina sintetasa (GS)” en la hoja encontrado por Hirel *et al.* (2001). Esta es una de las principales enzimas involucradas en la asimilación y reciclado de N y se encontró que la actividad de GS en la hoja se correlaciona positivamente con el rendimiento de grano, número de granos por planta y la eficiencia metabólica del grano cuando el aporte de N es limitado y con rendimiento de N del grano cuando hay un alto aporte de N.

La selección asistida por marcadores se puede ejecutar como una selección de un rasgo basado en el genotipo, utilizando marcadores asociados en lugar del fenotipo del rasgo (Boopathi, 2013). La detección de asociaciones significativas entre rasgos relacionados a la EUN y marcadores moleculares posibilitará el uso de métodos de selección asistida por marcadores para mejorar la EUN.

Identificación de genes candidatos

Entre los genes con función predicha detectados en los QTL, los más importantes relacionados a la EUN se detallan en la Tabla 6.

Tabla 6. Genes candidatos detectados en los intervalos correspondientes a los QTL.

QTL	ID	Descripción del gen	Nombre del gen
EUN-1a	Zm00001d028549	Proteína de la superfamilia de las cisteínas proteasas	
	Zm00001d028750	Asparagina sintetasa 3	asn3 - asparagine synthetase3
	Zm00001d028766	Asparagina sintetasa 1	
	Zm00001d028718	Proteína hidrofóbica RCI2B	pmpm4 - proteolipid membrane potential regulator4
	Zm00001d028793	Receptor de ácido abscísico PYL5	
	Zm00001d028721	Proteína con dominio LOB	lbd4 - LBD-transcription factor 4
EUN-1b	Zm00001d045226	Proteína de la superfamilia de las cisteínas proteasas	
	Zm00001d045132	Alantoinasa	
	Zm00001d044994	Transportador de aminoácidos	
	Zm00001d045392	Nodulina temprana 93	early nodulin 93
	Zm00001d045390	Proteína relacionada a nodulina temprana	
EUN-1c	Zm00001d045675	Asparagina sintetasa 1	asparagine synthetase1
	Zm00001d045740	Transportador de aminoácidos transmembrana	
EUN-1d / ICN-1b	Zm00001d046383	Proteína transportadora de nitrato NRT1	
N/Biom-1	Zm00001d010355	Asparagina sintetasa 1	
	Zm00001d010196	Asparagina sintetasa 3	
	Zm00001d010190	Glutamato-oxaloacetato transaminasa 5	got5 - glutamate-oxaloacetate transaminase5
	Zm00001d010662	SNF1- Subunidad reguladora relacionada a proteína quinasa 1	snrkb1 - SNF1-related protein kinase regulatory subunit beta1
ICN-1a	Zm00001d008744	Transaminasa/transferasa	
	Zm00001d008691	Proteína relacionada a la autofagia	atg18d - autophagy18d
Pr-1	Zm00001d029732	Glutamato sintasa 1 [NADH] cloroplástica	
	Zm00001d029510	Ureasa	
RGP-1	Zm00001d009843	Proteína homóloga de ROOT HAIR DEFECTIVE 3	
	Zm00001d009580	Proteína accesoria de Ureasa (UreG)	
	Zm00001d009823	Proteína con dominio LOB	LBD-transcription factor 35

ID: identificador (del inglés identifier).

Los genes que codifican proteasas son sobreexpresados durante la senescencia de la hoja y la degradación de proteínas de la hoja durante la senescencia foliar es requerida para removilizar el N y exportarlo para el desarrollo y crecimiento del grano (Donnison *et al.*, 2007). Otros de los genes de gran importancia encontrados codifican para transaminasas, enzimas que participan en el flujo de N entre compuestos orgánicos y

asparagina sintetasa, una enzima que es esencial para la asimilación, distribución y removilización de N dentro de la planta a través del floema (Chevalier *et al.*, 1996).

El gen *snf-1-related protein kinase regulatory subunit beta1* participa en la asimilación de N mediante la fosforilación de la enzima nitrato reductasa, ya que el nitrato después de entrar a la célula de la planta se reduce a nitrito por la nitrato reductasa y luego a amonio por la enzima nitrito reductasa. La actividad de la nitrato reductasa está regulada postraduccionalmente por un mecanismo de fosforilación-defosforilación (Crawford and Arst, 1993).

El gen que codifica la proteína glutamato sintasa 1 está involucrado en la biosíntesis de glutamato y probablemente participa en la asimilación primaria de amonio en las raíces. El amonio derivado de nitrato o de la absorción directa de amonio por los transportadores se asimila en aminoácidos a través del ciclo de la glutamina sintetasa/glutamato sintasa (Lancien *et al.*, 2002).

También se encontraron genes relacionados al estrés. El ácido abscísico (ABA) es una hormona involucrada en respuestas de estrés abiótico y biótico y está implicada en el crecimiento de la raíz, regulada por N y en la adquisición de N (Xu *et al.*, 2012). La proteína hidrofóbica RCI2B está involucrado en la respuesta al ABA y el receptor de ácido abscísico PYL5 participa en la vía de señalización activada por ABA y es requerido para las respuestas mediadas por ABA como el cierre estomático y la inhibición de la germinación (Ma *et al.*, 2009). La alantoinasa participa en la respuesta a la inanición de N y en Arabidopsis el metabolito intermediario alantoína participa en la tolerancia al estrés abiótico a través de la activación del metabolismo de ABA (Werner *et al.*, 2013).

Los genes *lbd* (*Lateral Organ Boundaries Domain*) son factores de transcripción que además de ser reguladores del desarrollo actúan como represores de la biosíntesis de antocianinas y señales de disponibilidad de N. El nitrógeno y nitrato suprimen parte del metabolismo secundario, incluida la síntesis de antocianinas. Cuando el N es limitante la

ruta metabólica se detiene y se dirige hacia la producción de antocianina (Rubin *et al.*, 2009).

Uno de los genes más relevantes es Zm00001d046383 (Protein NRT1/ PTR FAMILY 4.6), un transportador de nitrato dependiente de protones de baja afinidad y está involucrado en la asimilación de nitrato e implicado en el transporte de ABA (Kanno *et al.*, 2013).

La nodulina temprana ENOD93 y los genes relacionados a nodulina temprana participan en la nodulación y pueden tener un efecto positivo en la EUN (Kouchi *et al.*, 1993).

Entre los genes más importantes se encuentra un gen relacionado a la autofagia (*atg18d*), la cual tiene un rol central en la removilización de N en el maíz, contribuyendo al mantenimiento y el reciclaje celular y, por lo tanto, es esencial para aliviar el estrés nutricional, donde el material citoplásmico y los orgánulos dañados se envían a la vacuola para su descomposición. Se ha demostrado que los niveles de transcripción de los genes *atg* del maíz están estrechamente regulados durante el desarrollo y en la respuesta a la disponibilidad de nutrientes y, que la removilización de N está alterada en las plantas cuando genes relacionados a la autofagia están mutados y las condiciones de campo no son las óptimas, lo que disminuye significativamente el rendimiento de semillas y el índice de cosecha de nitrógeno. Por lo tanto la autofagia, aunque no es esencial, se vuelve crítica durante el estrés de nitrógeno e impacta severamente en la productividad del maíz cuando las condiciones de campo no son óptimas. Dado que la autofagia juega un papel central en la utilización del N en maíz, el control del reciclaje de nutrientes a través de la autofagia podría ofrecer nuevas estrategias para mejorar la utilización de N (Li *et al.*, 2015).

Un gen que codifica una proteína homóloga de ROOT HAIR DEFECTIVE 3 está involucrado en la expansión de la pared celular y es necesario para la ampliación de las células de la raíz.

La ureasa está involucrada en el reciclaje de N a partir del catabolismo de ureido, purina y arginina y la proteína accesoria de ureasa (UreG) es requerida para la maduración y activación de la ureasa (Witte *et al.*, 2005).

Si bien con la incorporación de marcadores moleculares se logró acotar el intervalo de un QTL sobre el cromosoma 9 también se detectaron QTL nuevos con intervalos muy amplios, por lo tanto, será necesario un mapeo más fino de modo que el QTL de interés se encuentre próximo al marcador ligado y se eliminen genes no deseados que se encuentren cerca del QTL.

CONCLUSIÓN

El 32,2 % de los marcadores moleculares probados entre las líneas parentales fueron polimórficos. Casi el 50% de los SSR públicos y el 32% de los INDEL fueron polimórficos.

La incorporación de marcadores nuevos a los mapas de los cromosomas 1, 8 y 9 de maíz construidos previamente (Mandolino 2015), permitió aumentar la longitud de los mapas y mejorar la representación de los mismos. La mayoría de los marcadores incorporados a los mapas fueron INDEL.

Se detectaron 11 QTL por MIC en los cromosomas 1, 8 y 9. En el cromosoma 1 se detectaron dos QTL para EUN, uno para cada año de ensayo y que comparten parte del intervalo de confianza. Este podría tratarse de un QTL estable en los diferentes ambientes. La incorporación de marcadores permitió acotar el intervalo del QTL para EUN detectado sobre el cromosoma 9 de 55,3 Mpb a 15,4 Mpb. Los QTL EUN-1d e ICN-1b colocalizaron entre los mismos marcadores indicando que se podría tratar de un único QTL pleiotrópico. En los cromosomas 8 y 9 se detectaron QTL nuevos para EUN, ICN y rendimiento fuera de la región de interés.

Se lograron identificar genes relacionados al metabolismo del nitrógeno en la mayoría de los intervalos de los QTL. La identificación de genes candidatos dentro de los intervalos de los QTL resulta interesante, ya que permite la selección de regiones de interés dentro de los intervalos amplios de los QTL. Algunos de los genes de mayor importancia son los que participan en la asimilación, distribución y removilización de N, tales como asparagina sintetasa y glutamato sintasa, genes involucrados en la respuesta al estrés y en el transporte de aminoácidos.

Los intervalos de los QTL aún son muy amplios, por lo tanto, será necesario un mapeo más fino para acercar nuevos marcadores a los QTL de interés de los distintos caracteres.

BIBLIOGRAFÍA

Agrama, H.A. (2006). Application of molecular markers in breeding for nitrogen use efficiency. *Journal of crop improvement*, 15 (2): 175-211.

Arellano, J.V., Castillo, F.G., Alcántar, G.G. & Martínez, A.G. (1996). Parámetros genéticos de la eficiencia en el uso de nitrógeno en líneas de maíz de valles altos. In Edmeades, G.O; Bänziger, M; Mickelson, H.R; Peña-Valdivia, C.B. eds. *Developing drought and soil nitrogen tolerant in maize*. CIMMYT, El Batán México, pp. 320-325.

Beadle, G. (1980). The ancestry of corn. *Scientific American*, 242: 112-119.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B*, 57 (1): 289-300.

Bhatramakki, D., Dolan, M., Hanafey, M., Wineland, R., Vaske, D., Register J.C., Tingey, S.V. & Rafalski, A. (2002). Insertion-deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant molecular biology*, 48 (5-6): 539-547.

Boopathi, N.M. (2013). QTL Identification. En *Genetic Mapping and Marker Assisted Selection: Basics, Practice and Benefits*. Ed Springer, Coimbatore. pp. 117-163.

Cárdenas-Navarro, R., Sánchez-Yáñez, J., Farías-Rodríguez, R. & Peña-Cabriales, J. (2004). Los aportes del nitrógeno en la agricultura. *Chapingo*, 10 (2): 173-178.

Cervigni, D.L., Ortiz, J.P. & Feingold, S.E., (2010). Construcción de mapas de ligamiento genético, localización de genes y regiones cromosómicas asociadas a caracteres de interés en plantas. En: Levitus, G., Echenique, V., Rubinstein, C., Hopp, E. & Mroginski, L. (eds). *Biotecnología y Mejoramiento Vegetal II*. Ediciones INTA, Bs. As. pp. 86-99.

Cheema, J. & Dicks, J. (2009). Computational approaches and software tools for genetic linkage map estimation in plant. *Briefing in Bioinformatics*, 10 (6): 595-608.

Chevalier, C., Bourgeois, E., Just, D. & Raymond, P. (1996). Metabolic regulation of asparagine synthetase gene expression in maize (*Zea mays* L.) root tips. *The Plant Journal*, 9 (1): 1-11.

Collard, B., Jahufer, M., Brouwer, J. & Pang, E. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica*, 142: 169-196.

Coque, M., Martin, A., Veyrieras, B., Hirel, B. & Gallais, A. (2008). Genetic variation for N-remobilization and postsilking N-uptake in a set of maize recombinant inbred lines. 3. QTL detection and coincidences. *Theoretical and Applied Genetics*, 117: 729–747.

Crawford, N.M. & Arst Jr, H.N. (1993). The molecular genetics of nitrate assimilation in fungi and plants. *Annual review of genetics*, 27 (1): 115-146.

Cruz, C.D. & Schuster, I. (2008). GQMol V2008 6.1. *Genética Quantitativa e Molecular*. Universidad Federal de Viçosa. Laboratório de Bioinformática. Consultado 22 oct. 2015. Disponible en <http://www.ufv.br/dbg/gmol.htm>.

Doebley, J., Stec, A., Wendel, J. & Edwards, M. (1990). Genetic and morphological analysis of a maize-teosinte F2 population: implications for the origin of maize. *Proceedings of the National Academy of Sciences*, 87 (24): 9888-9892.

Doerge, R.W. (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nature Reviews Genetics*, 3 (1): 43-52.

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K. & Mattick, J.S. (1991). "Touchdown PCR to circumvent spurious priming during gene amplification". *Nucleic Acids Research*, 19 (14): 4008.

Donnison, I.S., Gay, A.P., Thomas, H., Edwards, K.J., Edwards, D., James, C.L. & Ougham, H.J. (2007). Modification of nitrogen remobilization, grain fill and leaf senescence in maize (*Zea mays*) by transposon insertional mutagenesis in a protease gene. *New Phytologist*, 173 (3): 481-494.

Galinat, W.C. (1988). The origin of corn. En: 'Corn and corn improvement'. Eds: Sprague G.F., Dudley, J.W. Madison. WI. USA. American Society of Agronomy. 1-31

Harrison, P. (2002). Agricultura mundial: hacia los años 2015/2030. Dirección de Información de la FAO.

Helentjaris, T., Weber, D. & Wright, S. (1988). Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. *Genetics*, 118 (2): 353-363.

Hirel, B., Bertin, P., Quilleré, I., Bourdoncle, W., Attagnant, C., Dellay, C., Gouy, A., Cadiou, S., Retalliau, C., Falque, M. & Gallais A. (2001). Towards a better understanding of the genetic and physiological basis for nitrogen use efficiency in maize. *Plant Physiology*, 125: 1258-70.

Kanno, Y., Kamiya, Y. & Seo, M. (2013). Nitrate does not compete with abscisic acid as a substrate of AtNPF4.6/NRT1.2/AIT1 in Arabidopsis. *Plant signaling & behavior*, 8 (12): e26624.

- Kant, S., Bi, Y.M. & Rothstein, S.J. (2011). Understanding plant response to nitrogen limitation for the improvement of crop nitrogen use efficiency. *Journal of Experimental Botany*, 62 (4): 1499-1509.
- Kleinhofs, A., SaghaiMaroof, M., Biyashev, R.M., Hayes, P., Chen, F.Q., Lapitan, N., Fenwick, A. & Steffenson, B.J. (1993). A Molecular isozyme and morphological map of the barley (*Hordeum vulgare*) genome. *Theoretical and Applied Genetics*, 86: 705-712.
- Kosambi, D.D. (1944). The Estimation of Map Distances from Recombination Values. *The Annals of Human Genetics. Ann Eugenics*, 12: 172-175.
- Kouchi, H. & Hata, S. (1993). Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. *Molecular and General Genetics MGG*, 238 (1-2): 106-119.
- Lancien, M., Martin, M., Hsieh, M.H., Leustek, T., Goodman, H. & Coruzzi, G.M. (2002). *Arabidopsis* *glt1-T* mutant defines a role for NADH-GOGAT in the non-photorespiratory ammonium assimilatory pathway. *The Plant Journal*, 29 (3): 347-358.
- Levinson, G. & Gutman, G.A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.*, 4: 203-221.
- Li, F., Chung, T., Pennington, J.G., Federico, M.L., Kaeppler, H.F., Kaeppler, S.M. & Vierstra, R.D. (2015). Autophagic recycling plays a central role in maize nitrogen remobilization. *The Plant Cell*, 27 (5): 1389-1408.
- Liu, X., He, S., Zheng, Z., Huang, Y., Tan, Z., Li, Z., He, C., Wu, X. & Pu, Q. (2010). Identification of the qtls for grain yield using RIL population under different nitrogen regimes in maize. *African Journal of Agricultural Research*, 5 (15): 2002-2007.

- Liu, J., Qu, J., Yang, C., Tang, D., Li, J., Lan, H. & Rong, T. (2015). Development of genome-wide insertion and deletion markers for maize, based on next-generation sequencing data. *BMC genomics*, 16 (1): 601.
- Lu, H., Romero-Severson, J. & Bernardo, R. (2002). Chromosomal regions associated with segregation distortion in maize. *Theoretical and Applied Genetics*, 105 (4): 622-628.
- Ma, Y., Szostkiewicz, I., Korte, A., Moes, D., Yang, Y., Christmann, A. & Grill, E. (2009). Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science*, 324 (5930): 1064-1068.
- Mandolino, C.I. (2015). Mapeo de QTL para Eficiencia de Uso de Nitrógeno en Maíz mediante SSRs y SNPs. Tesis doctoral, Rosario, Argentina, Facultad de Cs. Agrarias, Universidad de Rosario.
- Mangelsdorf, P.C. (1986). The origin of corn. *Scientific American*, 255 (2): 80-86.
- Martínez, M.C., Helguera, M. & Carrera, A. (2010). Marcadores Moleculares. En: Levitus, G., Echenique, V., Rubinstein, C., Hopp, E. & Mroginski, L. (eds). *Biotecnología y Mejoramiento Vegetal II*. Ediciones INTA, Bs. As. pp. 70-85.
- Martins, W.S., Lucas, D.C.S., de Souza Neves, K.F. & Bertioli, D.J. (2009). WebSat-A web software for microsatellite marker development. *Bioinformatics*, 3 (6): 282.
- Mohan, M., Nair, S., Bhagwat, A., Krishna, T., Yano, M., Bhatia, C. & Sasaki, T. (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular breeding*, 3 (2): 87-103.
- Olmos, S.E., Delucchi, C., Ravera, M., Negri, M.E., Mandolino, C. & Eyhérbide, G.H. (2014). Genetic relatedness and population structure within the public Argentinean collection of maize inbred lines. *Maydica*, 59: 16-31.

Paliwal, R. L., Granados, G., Lafitte, H. & Violic, A. (2001). El maíz en los trópicos: mejoramiento y producción. Organización de las Naciones Unidas para la Agricultura y la Alimentación, FAO, Roma.

Paterson, A.H. (1996). Making genetic maps. In Paterson, AH. ed. Genome mapping in plants. Landes Company. San Diego. California, Academic Press. Austin. Texas. 23–39.

Prasanna, B.M. (2012). Diversity in global maize germplasm: characterization and utilization. *Journal of biosciences*, 37 (5): 843-855.

Promega. (2007). Technical manual: Silver sequence DNA sequencing system. Instructions for use of products Q4130. Disponible en <http://www.promega.com>.

Raun, W.R. & Johnson, G.V. (1999). Improving nitrogen use efficiency for cereal production. *Agronomy journal*, 91 (3): 357-363.

Rubin, G., Tohge, T., Matsuda, F., Saito, K. & Scheible, W.R. (2009). Members of the LBD family of transcription factors repress anthocyanin synthesis and affect additional nitrogen responses in *Arabidopsis*. *The Plant Cell*, 21 (11): 3567-3584.

Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mpkrejs, M., Tetko, I., Güldener, U., Mannhaupt, G., Münsterkötter, M. & Mawes, W. (2004). The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Research*, 32 (18): 5539-5545.

Schnable, P.S., Ware, D., Fulton, R.S., *et al.* (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science*, 326 (5956): 1112-1115.

Semagn, K., Bjornstad, A. & Xu, Y. (2010). The genetic dissection of quantitative trait in crops. *Electronic Journal of Biotechnology*, 13 (5): 1-45.

Settles, A.M., Bagadion, A.M., Bai, F., Zhang, J., Barron, B., Leach, K., Mudunkothge, J., Hoffner, C., Bihmidine, S., Finefield, E., Hibbard, J., Dieter, E., Malidelis, A., Gustin, J.L., Karoblyte, V., Tseung, C. & Braun, D.M. (2014). Efficient molecular marker design using the MaizeGDB Mo17 SNPs and Indels track. *G3: Genes| Genomes| Genetics*, 4 (6): 1143-1145.

Shiferaw, B., Prasanna, B.M., Hellin, J. & Bänziger, M. (2011). Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. *Food Security*, 3 (3): 307-327.

Tanksley, S.D. (1993). Mapping polygenes. *Annual review of genetics*, 27 (1): 205-233.

Van Ooijen, J.W. (1999). LOD significance thresholds for QTL analysis in experimental populations of diploid species. *Heredity*, 83 (5): 613-624.

Vogl, C. & Xu, S. 2000. Multipoint mapping of viability and segregation distorting loci using molecular markers. *Genetics*, 155: 1439–1447.

Voorrips, R.E. (2006). MapChart 2.2. Copyright Plant Research International. Licensed to Mapchart used.

Wang, S., Basten, J. & Zeng, Z. (2013). Windows QTL Cartographer 2.5. Raleigh, NC: Department of Statistics, North Carolina State University.

Werner, A.K., Medina-Escobar, N., Zulawski, M., Sparkes, I.A., Cao, F.Q. & Witte, C.P. (2013). The ureide-degrading reactions of purine ring catabolism employ three amidohydrolases and one aminohydrolase in Arabidopsis, soybean, and rice. *Plant physiology*, 163 (2): 672-681.

Winter, P. & Kahl, G. (1995). Molecular marker technologies for plant improvement. *World Journal of Microbiology & Biotechnology*, 11: 438–448.

Witte, C.P., Rosso, M.G. & Romeis, T. (2005). Identification of three urease accessory proteins that are required for urease activation in Arabidopsis. *Plant physiology*, 139 (3): 1155-1162.

Xu, Y. (2010). *Molecular plant breeding*. Ed. CAB International. 11 p.

Xu, G., Fan, X. & Miller, A.J. (2012). Plant nitrogen assimilation and use efficiency. *Annual Review of Plant Biology*, 63: 153-182.

ANEXO 1

1.1. Marcadores SSR diseñados.

Bin	SSR	PM		Cebadores	Posición	Repetición
1.03	N001	173	N001f	ATTGCGTTGGCTGTCTGTAG	46.869.721	(GA) ₉
			N001r	AGCATCGCATCTCGTGTCC	46.869.894	
1.03	N002	175	N002f	CTGTTAAGAGAAAGGGGTACACGA	46.854.897	(TA) ₇
			N002r	AATCGCAATTCCCCCTTACA	46.855.072	
1.03	N003	200	N003f	GCCATGTGTCCTTGGTCAAAC	47.473.699	(TA) ₆
			N003r	TTAGTTTGCCGGTCTGTCTCT	47.473.899	
1.03	N004	125	N004f	GCGTGTGTTCTGTCTGTCACT	47.295.116	(GCC) ₆
			N004r	AGAGACCGATGAAAGCCTCA	47.295.241	
1.03	N005	178	N005f	GACGCCCTAAAGGACCGAAC	47.463.878	(TA) ₆
			N005r	GGCTCTCACAGTAGCCAAACTA	47.464.056	
9.03	N006	250	N006f	ACGACGTAAAAACGACGCAC	65.019.923	(CT) ₈
			N006r	CCTTTCCCGGTCCGTGTG	65.020.173	
9.03	N007	231	N007f	TCCCTAGCATTGCGGGTTTT	51.769.833	(TA) ₃₆
			N007r	TCGTACAGGACAGTTGCATCC	51.770.064	
9.03	N008	155	N008f	ACCGTGTCCGCGTCAATG	75.079.158	(GTCCAT) ₆
			N008r	CTGACTCGTGGGATGTGGAC	75.079.313	
9.03	N009	306	N009f	GACACGCCCAACGATGAC	80.081.721	(GAT) ₉
			N009r	AACAACCTCCCGCACGACTA	80.082.027	
9.03	N010	219	N010f	CGTAGGTTCCGGTGTGGATG	95.336.552	(TA) ₃₅
			N010r	TTCTACCGTGTGCGAACTGGC	95.336.771	
9.03	N011	240	N011f	GCCCTGGTATTGCTGGTG	100.040.055	(TA) ₃₃
			N011r	AGAGCGGCAGAAAGAGATGG	100.040.295	
9.04	N012	167	N012f	CATACACCCACATCCTCCAGAT	105.348.403	(TA) ₇
			N012r	GCCAGTTCACGTACATCCAGT	105.348.570	

Posición basada en la anotación del B73 RefGen_v3 genome sequence.

Bin: intervalo cromosómico determinado por marcadores moleculares; *PM*: peso molecular.

ANEXO 2

2.1. Marcadores moleculares utilizados para el análisis.

Tabla 1. Marcadores SSR públicos utilizados para el análisis.

Bin	Marcador	Secuencia de oligonucleótidos	Repetición
1.03	phi001	F: 5' TGACGGACGTGGATCGCTTCAC 3' R: 5' AGCAGGCAGCAGGTGAGCAGCG 3'	AG
1.03	umc2217	F: 5' AAGACCTGTTCGTCAACCATGTTC 3' R: 5' ATGTCCCATTACCAAGGTCTGTG 3'	TG
1.04	umc2229	F: 5' GAGAAGGGCGGGAGGAATAAC 3' R: 5' CGAAGAGCACGATGTTGACG 3'	CGC
8.03	umc1735	F: 5' TTTGCCAAAAAGACATGACACCTA 3' R: 5' GCGCACAATATAATGAACCAACAG 3'	AGC
9.03	umc2337	F: 5' GAGTACCTCCGCCCACTCATC 3' R: 5' CTAACGTACACACATCTTGGCTGG 3'	AGC
9.03	umc2370	F: 5' TAAAACCTTTGCGTGCATGTATGT 3' R: 5' TCAGTGGATTGCTGGATCATAGAA 3'	TCTC
9.03	umc1191	F: 5' ACTCATCACCCCTCCAGAGTGTC 3' R: 5' AAGTCATTGCCCAAAGTGTTC 3'	CT
9.03	umc1743	F: 5' TGGACTTCGAAAATTCTTTCAGC 3' R: 5' GAGAGGAGGAGCTTCACGAGC 3'	GGC
9.03	umc2338	F: 5' ATCTCACACTCACGACCTGTTGTC 3' R: 5' AGCATTGGGACTGTTCAAGAGAAG 3'	GCC

Bin: intervalo cromosómico determinado por marcadores moleculares; PM: peso molecular.

Tabla 2. INDEL utilizados para el análisis.

Bin	Marcador	PM	Secuencia de oligonucleótidos	Coordenadas
1.03	chr1-46628500	153	F: 5' AGCCCCAAAGTGTGCTCAT 3' R: 5' TGATGTGATGCGATCATGTG 3'	46.628.500
1.03	chr1-46629720	230	F: 5' CGATGCACCAAACCGATAAT 3' R: 5' AAGGTCAAACCTCTGGGGCT 3'	46.629.720
1.03	chr1-47770200	246	F: 5' TGGAACAAATATTGTATTTTGACG 3' R: 5' ACCAACTTTGCCTCCCTCTT 3'	47.770.200
8.02	chr8-12684300		F: 5' TCAGTTGGCTTTGTTTGCAT 3' R: 5' TCAGCAGCACAAGAATTTGG 3'	12.684.300
8.02	chr8-13272600		F: 5' CCGGACTTCAACGTGTCTG 3' R: 5' TAAACGACCTGGCTCCATTT 3'	13.272.600
8.02	DEL.131996	19	F: 5' TTCTCTGTACTGTACGCGGC 3' R: 5' GAAGCGGCACTTTCCCTTTCC 3'	13.368.932
8.02	chr8-14542640		F: 5' AGTTACGCTCCTGCCCCC 3' R: 5' GTGGCGGACTGATTCAATTT 3'	14.542.640
8.02	chr8-17067440		F: 5' GAGCAGGCTGATGTCGAACT 3' R: 5' CACAGCTGAATGTTGCATCC 3'	17.067.440
8.02	DEL.132563		F: 5' GCTTTATCCGACGGTCTTGG 3' R: 5' TGCCAGTGAGAAAACCTGACA 3'	20.972.027
8.03	DEL.132932		F: 5' ACTGACGACAGAGCAGTTGG 3' R: 5' TATGGGCTGAAGTTGGCCTG 3'	25.947.383
8.03	DEL.133392		F: 5' CTACAGCACGCAGGTATGGA 3' R: 5' CACACACATACAACGGCTTCA 3'	34.633.815
8.03	DEL.133721		F: 5' ACAGTGATGCTTGATGCTTCTTTT 3' R: 5' AGGATGAAACGTGCATGGC 3'	39.124.283

Tabla 2. Continuación.

Bin	Marcador	PM	Secuencia de oligonucleótidos	Coordenadas
9.03	chr9-27968120	227	F: 5' TCAACAGGCTGTGAAGAAGAAA 3' R: 5' CAATGCATGCCTATGGAAGA 3'	27.968.120
9.03	DEL.144557		F: 5' TCAGACCTAGGTGAAACAACAGA 3' R: 5' TGTGTGGTACCCTGAATGGC 3'	28.431.672
9.03	INS.116020	134	F: 5' GGAAAGTTCTCGGTGCTAAA 3' R: 5' GGTTGGTGAACCTCAAGTCTG 3'	33.672.808
9.03	INS.116034	144	F: 5' TGGAAGTGAGGTCTGAATCT 3' R: 5' AGTCAACGACAACGAACAC 3'	33.730.012
9.03	DEL.145171	135	F: 5' GGTAACCGCAGTATATAGCA 3' R: 5' GCGACTGACAAGGGAAA 3'	39.336.418
9.03	chr9-45015160	228	F: 5' ATCGGCTCTCCGACACAGT 3' R: 5' ACCTTTTCCAGCCCTATGCT 3'	45.015.160
9.03	DEL.145752		F: 5' GCCGTGTGGAAATTGTTTCGT 3' R: 5' GCTTCCTCCACTGGCTTCTT 3'	47.043.881
9.03	chr9-56424320	166	F: 5' CCGAAGCTCAACACCAAGTT 3' R: 5' CAGGCAGCGTTTTTCTTTTC 3'	56.424.320
9.03	DEL.146397		F: 5' CAGGTGTTTCGTGGCAGTAGT 3' R: 5' CCGTGTGGTATGTAGCAGGG 3'	59.702.927
9.03	DEL.147207		F: 5' AAGCCACTCTGGGTAATTCTG 3' R: 5' CTGGGAGTTCACACGGTTCC 3'	77.316.243
9.03	chr9-74589140	215	F: 5' CCCCAGTATCTGCAGCAAAG 3' R: 5' GAGGGATGTACGTTTGGTGG 3'	74.589.140
9.03	chr9-74592980		F: 5' CGAAGCCCTTTTGGCTTT 3' R: 5' CCATCTGGTTCTGATCTGCC 3'	74.592.980
9.03	chr9-85234820		F: 5' CGAAAGTTTTGACCAAATCTCATA 3' R: 5' CGACGTTGGATGGCTAGAAT 3'	85.234.820
9.03	DEL.148060		F: 5' TTGCTTAGATGGACGGCTCT 3' R: 5' TGCAAGGTCCAGTCTGATGT 3'	89.605.276
9.03	DEL.148101		F: 5' TGAGCACAACCTCAGTATTGG 3' R: 5' TCAAACTACACATATCCAGCCCA 3'	90.205.338
9.03	DEL.148523		F: 5' TGCGAGCACTCGTCGATTTA 3' R: 5' CGACTTTGAGGTAGGGTGGG 3'	98.489.288
9.03	chr9-98676780	207	F: 5' AAGGTGCTGCAGCGTGTG 3' R: 5' AGCAGAGCAGCATGTCAGAA 3'	98.676.780
9.03	DEL.148894		F: 5' GACGACCTCGCCATTTTCT 3' R: 5' GGGACGGCGCCATAGTT 3'	101.686.335
9.04	DEL.149169		F: 5' GTTAGAACCTATGCGCGGGA 3' R: 5' CAGCAGTATACTCGTTGTCCCA 3'	104.915.540

Coordenadas basadas en la anotación del B73 RefGen_v2 genome sequence

Bin: intervalo cromosómico determinado por marcadores moleculares; *PM:* peso molecular.

Tabla 3. SSR de estudios anteriores (Mandolino, 2015) utilizados para el análisis.

Bin	Marcador	PM	Secuencia de oligonucleótidos	Repetición
1.00	umc1106	121	F: 5' CATCCAGTCATCGGTAATACGACA 3' R: 5' AACTGTTGCCTTCTTTTTCCCTTC 3'	GAGA
1.01	phi056	84	F: 5' ACTTGCTTGCCTGCCGTTAC 3' R: 5' CGCACACCACTTCCCAGAA 3'	CCG
1.02	bnlg1429	150	F: 5' CTCCTCGCAAGGATCTTCAC 3' R: 5' AGCACCGTTTCTCGTGAGAT 3'	AG
1.02	bnlg1953	133	F: 5' CCTCGGAGCTCGATTTACAC 3' R: 5' AACATTTAACCGCCGTCATC 3'	AG
1.03	phi339017	146	F: 5' ACTGCTGTTGGGGTAGGG 3' R: 5' GCAGCTTGAGCAGGAAGC 3'	AGG
1.03	bnlg1484	118	F: 5' GTAAAAGACGACGACATTCCG 3' R: 5' GACGTGCACTCCGTTTAACA 3'	AG
1.03	umc1479	199	F: 5' CTGGCTCTTCAAGTGTAAGGAGG 3' R: 5' GGCCTTTTTCTTAGCTTCCTCATC 3'	AGA
1.03	bnlg439	228	F: 5' TTGACATCGCCATCTTGGTGACCA 3' R: 5' TCTTAATGCGATCGTACGAAGTTGTGGAA 3'	
1.03	phi095	180	F: 5' CCGATCGGCTTTTACTACTGTTTAGC 3' R: 5' ATGCACCATTCTAGCACTATAGCAACACT3'	AG
1.03	umc1514	114	F: 5' TGCTGTCTGACGATCTCTCAAATC 3' R: 5' TGCCACCAAATAGACCATCTATCA 3'	AGT
1.03	umc2096		F: 5' AACGTACCATCCTTGTGCCTGTAT 3' R: 5' CATTATTAGTACGGAGCGCTGGTC 3'	ACAT
1.04	bnlg2238		F: 5' TGCCACTCAAGCCTTCTTTT 3' R: 5' TTCTGATTGCAGTGCAGACC 3'	
1.04	umc2227	84	F: 5' ACCTTGAGCGTGGAGTCCGGT 3' R: 5' AGCTGAGCCTTCTCTTCTTGGCT 3'	GCGCTG
1.05	umc2025	129	F: 5' CGCCGTAGTATTTGGTAGCAGAAG 3' R: 5' TCTACCGCTCCTTCGTCCAGTA 3'	AGCT
1.05	umc1734	116	F: 5' TTGTGCATTTTGCAGAACTAGGA 3' R: 5' AGTACTTGTCCGGTGGAGACTGGAG 3'	TAC
1.05	umc1395	100	F: 5' TGAATGAGTGGCATTCAAATCTG 3' R: 5' CAGATTGCATGTGTGAGTGTGTGT 3'	AC
1.06	umc1335	125	F: 5' ATGGCATGCATGTGTTTGTTTTAC 3' R: 5' ACAGACGTGCTAATTCCTGAAAG 3'	AG
1.07	umc2237	156	F: 5' CTCAGCTACAGGAGCGAAGAGG 3' R: 5' GTCACTGCACGATCCATCACAT 3'	GCGACT
1.07	bnlg1025	166	F: 5' TGGTGAAGGGGAAGATGAAG 3' R: 5' CCGAGACGTGACTCCTAAGC 3'	AG
1.09	umc1290	157	F: 5' AGAGATTCATCAGAGTGGCGATG 3' R: 5' CTGCTCAGCTCATCCTCCT 3'	GCG
1.09	umc1431	135	F: 5' GTTGCTGTCCCCGACGTAGTAG 3' R: 5' GTGAGAGTACCGAGATGGCTGAG 3'	GCA
1.10	umc1885	138	F: 5' TATACCAGCATCAGGTCTCGTCG 3' R: 5' GTAGAGTGACCGTGTGTAGCAGA 3'	AAC
1.10	umc2223	151	F: 5' TTTTGGGACTGAAGAAGAAGATCG 3' R: 5' ACTTCTGCAGAGCGAGCAGG 3'	GCG
1.11	umc1553	157	F: 5' TGAATGGAAGAGAAGGGAAATCTG 3' R: 5' GCTCTGTACATCCTTAGCGACACA 3'	AG
1.11	bnlg504	180	F: 5' CGGCAGCTCCAGCACCGGCAT 3' R: 5' AGTGTCCACATACCGCCACACACGTTT 3'	
8.00	phi420701	300	F: 5' GATGTTTTCAAACCACCCAGA 3' R: 5' ATGGCACGAATAGCAACAGG 3'	CCG
8.01	bnlg2037		F: 5' GGGTGCTCGTAGTAGGGTT 3' R: 5' CTAAGGCACACGGAGAGAGG 3'	AG
8.02	phi119	164	F: 5' GGGCTCCAGTTTTAGTCATTGG 3' R: 5' ATCTTTCGTGCGGAGGAATGGTCA 3'	AG

Tabla 3. Continuación.

Bin	Marcador	PM	Secuencia de oligonucleótidos	Repetición
8.03	bnlg1834	210	F: 5' AAGGTTGGGTGTTGCTATGC 3' R: 5' TAGCTCTGCCACTGGACATG 3'	AG
8.03	bnlg1863	132	F: 5' GCGTTCGTTTTGCACTAAT 3' R: 5' CGACACAGTTGACATCAGGG 3'	AG
8.04	phi014	160	F: 5' AGATGACCAGGGCCGTCAACGAC 3' R: 5' CCAGCTTCACCAGCTTGCTCTTCGTG 3'	GGC
8.05	umc2367	147	F: 5' AACCGTTGTCGTTGTCTCTGTTC 3' R: 5' ATGAACTTGGTCACGCTCATCTC 3'	CCG
8.06	bnlg240	137	F: 5' AAGAACAGAAGGCATTGATACATAA 3' R: 5' TGCAGGTGTATGGGCAGCTA 3'	TC
8.07	bnlg1823	181	F: 5' TGTGACTCCATACCGCACAT 3' R: 5' CTCATCATGTTGTACATGGCG 3'	AG
8.08	phi080	157	F: 5' CACCCGATGCAACTTGCGTAGA 3' R: 5' TCGTCACGTTCCACGACATCAC 3'	AGGAG
8.09	dupsr14	105	F: 5' AGCAGGTACCACAATGGAG 3' R: 5' GTGTACATCAAGGTCCAGATTT 3'	(CT) T (CT) (CA)
9.00	bnlg1272	222	F: 5' ACCGAAGATGAGGTGTGACA 3' R: 5' TCAGTGCAAGGGCAATTTAG 3'	AG
9.01	bnlg1810	107	F: 5' ATGCTCCTCCTCTCCTCCAT 3' R: 5' GCGATGATGAGCTGCAAGTA 3'	AG
9.02	umc1170	151	F: 5' TGGGTGCTAAAACGTAACAACAAA 3' R: 5' GAGGACGAAGCAGAAATCCTACC 3'	TC
9.03	phi022	137	F: 5' TGCGCACCAGCGACTGACC 3' R: 5' GCGGGCGACGCTTCCAAAC 3'	GTGC
9.03	bnlg127		F: 5' CATGTATACGAGAAGCACCCCTAT 3' R: 5' ATCGTAACTCAGCGGTTTGTG 3'	
9.03	bnlg430	101	F: 5' CTTACTGAGCATCTTCCTTCTCTCC 3' R: 5' TCCGGTGATGCTCCAGCGAC 3'	AG
9.04	umc2120		F: 5' ACGTCAAGTAGTGAATCCTGCTCC 3' R: 5' ATTCTCCGCGGAAGAGACAAG 3'	CACAG
9.04	bnlg1209		F: 5' GTCCCGGGCAGAATAATACC 3' R: 5' TTCCTCCTTGAAGTGCTCGT 3'	AG
9.05	umc1654	104	F: 5' CACATTGATCAACTACCAGGTGGA 3' R: 5' GCCTCTTTTCTTCTCCATTTTCT 3'	GCA
9.05	umc1078	129	F: 5' AGGCACTAGCAGGCGAGAGG 3' R: 5' GCGTAGTAACATCCATCCAACCAA 3'	GT
9.05-9.06	bnlg1270	148	F: 5' TAGTTAACATGAGCAAATTAACAAGA 3' R: 5' TAGAAATGCAGAACCAGGGC 3'	AG
9.06	umc1789	98	F: 5' ACCTCTCCTTTTTCTCGCCTT 3' R: 5' GTCAGAGAAGAGGCCGGGTC 3'	CCG
9.07	bnlg619	255	F: 5' ACCCATCCCCTTTCCACCTCCTCCT 3' R: 5' GCTTTCAGCGAATACTGAATAACGCGGA 3'	CT
9.08	bnlg1129	197	F: 5' GAGAGTATGCTACTCGCCGC 3' R: 5' GACGAGTTTGGAGTGCCATT 3'	AG

Bin: intervalo cromosómico determinado por marcadores moleculares; PM: peso molecular.

Tabla 4. SNPs de estudios anteriores (Mandolino, 2015) utilizados para el análisis

SNPS	Cromosoma	Mutación	Proteína o dominio de proteína en la región del SNPs
PZA03741.1	1	[A/G]	Factor de unión DRE 1
PZB01385.3	8	[T/G]	Trehalosa-6-fosfato Sintasa
PZB00865.2	8	[A/G]	Almidón Sintasa IV (SSIV)
PZB00126.1	9	[A/C]	Asparagina Sintetasa