

**SATURACIÓN DEL MAPA GENÉTICO DE UNA POBLACIÓN INTRA-ACERVO
DE FRÍJOL (*Phaseolus vulgaris* L.) EMPLEANDO MARCADORES
MOLECULARES TIPO SSR Y SSCP**

NATALIA FRANCO HERRERA

**UNIVERSIDAD DEL QUINDÍO
FACULTAD DE CIENCIAS BÁSICAS Y TECNOLOGÍAS
PROGRAMA DE BIOLOGÍA
ARMENIA QUINDÍO
2011**

**SATURACIÓN DEL MAPA GENÉTICO DE UNA POBLACIÓN INTRA-ACERVO
DE FRÍJOL (*Phaseolus vulgaris* L.) EMPLEANDO MARCADORES
MOLECULARES TIPO SSR Y SSCP**

NATALIA FRANCO HERRERA

**Trabajo de grado presentado como requisito parcial para optar
al título de BIÓLOGA**

**Matthew W. Blair, Ph D.
Director**

**Nohra Cecilia Rodríguez
Codirectora**

**UNIVERSIDAD DEL QUINDÍO
FACULTAD DE CIENCIAS BÁSICAS Y TECNOLOGÍAS
PROGRAMA DE BIOLOGÍA
ARMENIA QUINDÍO
2011**

NOTA DE ACEPTACIÓN

El trabajo de grado titulado “**SATURACIÓN DEL MAPA GENÉTICO DE UNA POBLACIÓN INTRA-ACERVO DE FRÍJOL (*Phaseolus vulgaris* L.) EMPLEANDO MARCADORES MOLECULARES TIPO SSR Y SSCP**”, presentado por la estudiante **NATALIA FRANCO HERRERA**, para optar al título de Bióloga, fue revisado y calificado como:

Aprobada

Matthew W. Blair Ph.D.
Especialista en Germoplasma
LCGF – CIAT

Nohra Cecilia Rodríguez Castillo Ms.C
Codirectora
Universidad Del Quindío

Firma de jurado

Firma de jurado

Firma de jurado

Armenia, Julio de 2011

DEDICATORIA

A mis padres, por todo lo que me han brindado no solo en el transcurso del camino hacia la vida profesional, sino también por la formación como persona, por mostrarme el sendero que debía seguir para salir adelante.

A mis hermanos y a mis ti@s, por su compañía, ayuda, preocupación y apoyo en todos los momentos importantes de mi vida.

AGRADECIMIENTOS

En primer lugar quiero agradecer a Dios, por haberme puesto en el camino a todas aquellas personas que de una u otra manera contribuyeron a mi formación profesional, y al desarrollo de este trabajo.

Al Dr. Matthew W. Blair, por brindarme la oportunidad de realizar esta investigación en el LCGF, por sus sugerencias, correcciones y por su amistad.

Al Dr. Stephen E. Beebe, por permitirme continuar con mi trabajo a pesar de las adversidades.

A Carlos H. Galeano, mi tutor en CIAT, por sus explicaciones, paciencia, tiempo y buenos consejos a pesar de los km que nos distanciaron durante el proceso.

A Jorge Mario Londoño C, por ser la persona que siempre estuvo pendiente de mí durante este tiempo, en los buenos y malos momentos, por todos sus conocimientos y su grata compañía.

En general a todo el equipo de trabajo del LCGF, por su colaboración, de todos aprendí un poco para la formación profesional y como persona. De manera especial quiero agradecer a Héctor Fabio Buendía, por todas las recomendaciones, ayuda y amistad brindada durante mi estadía en CIAT, a Álvaro Soler, por su paciencia, amistad y colaboración.

A la profesora Nohra C. Rodríguez por su amistad, su tiempo y correcciones realizadas a este escrito.

A mis amigos del Laboratorio de Genética y Genómica de Arroz: César Iván Murcia y Ana María Leiva, por su compañía y por tantos momentos de esparcimiento y felicidad compartidos.

A mis amigos de la Universidad, Laura Andrea Taborda, Cristiam Albeiro Fernández, Eliana Peña y Carlos Mario González, por su valiosa amistad, comprensión y ayuda durante toda mi vida universitaria.

A mi director del Programa de Biología Arlés López E y a Dianita, la secretaria, por toda la colaboración brindada durante estos 5 años, además por su cooperación para hacer posible mi estancia en CIAT.

TABLA DE CONTENIDO

	Pág.
1. INTRODUCCIÓN	14
2. MARCO TEÓRICO	16
2.1. Origen y domesticación del fríjol común.....	16
2.2. Generalidades de fríjol	16
2.3. Genoma de <i>Phaseolus vulgaris</i> L.....	17
2.4. Mapas genéticos en fríjol.....	18
2.5. Fijación de nitrógeno y nodulación	21
2.5.1. Fijación de nitrógeno	21
2.5.2. Biología de la nodulación	22
2.5.3. Genética de la nodulación.....	22
2.5.3.1 Genes involucrados en la nodulación	23
2.6. Mapeo genético.....	27
2.6. 1. Mapeo genético.....	27
2.6.2. Principios del mapeo genético	28
2.6.3. Grupos de ligamiento (GL).....	29
2.6.4. Distancia del mapa y orden de los marcadores	30
2.6.5. Prueba para la distorsión de la segregación	31
2.7. Poblaciones de mapeo y RILs.....	32
2.7.1. Poblaciones de mapeo.....	32

2.7.2. Líneas Endogámicas Recombinantes (RILs)	33
2.8. Marcadores moleculares	34
2.8.1. Marcadores microsatélites	34
2.8.2. BAC-ends.....	36
2.8.3 SSCP (<i>Single-Strand Conformational Polymorphism</i>)	36
2.8.4. Polimorfismo de un único nucleótido (<i>Single Nucleotide Polymorphisms</i> “SNP”)	37
3. OBJETIVOS	40
3.1. Objetivo general	40
3.2. Objetivos específicos	40
4. MATERIALES Y MÉTODOS	41
4.1. Material Vegetal	41
4.1.1 Población DOR364 x BAT477	41
4.1.2 Población DOR364 x G19833	42
4.2. Extracción de ADN	44
4.3. Amplificación de microsatélites y SSCPs	44
4.4. Diseño de marcadores basados en genes de nodulación	46
4.5. Mapeo y análisis de los marcadores	46
4.6 Secuenciación y SNPs	47
5. RESULTADOS Y DISCUSIÓN	49
5.1. Evaluación de marcadores en los parentales.....	49

5.2. Marcadores polimórficos y mapeo en la población.....	52
5.2.1 Segregación de los marcadores.....	53
5.2.1.1 Población DOR364 x BAT477.....	53
5.2.1.2 Población DOR364 x G19833.....	54
5.3. Mapas de ligamiento	58
5.3.1 Población DOR364 x BAT477.....	58
5.3.2 Población DOR364 x G19833.....	63
5.4. Mapeo comparativo.....	70
5.5. Secuenciación y SNPs.....	72
6. CONCLUSIONES	75
7. RECOMENDACIONES	76
8. BIBLIOGRAFÍA	77
9. ANEXOS	101

ÍNDICE DE FIGURAS

Figura 1. Semillas de las líneas parentales DOR364 y BAT477 y de las 90 RILs empleadas.....	41
Figura 2. Árbol genealógico de los genotipos DOR364 y BAT477.....	43
Figura 3. Semillas de las líneas parentales DOR364 y G19833.....	44
Figura 4. Patrones de segregación de tres marcadores mapeados en la población DB.....	52
Figura 5. Distribución de la segregación de los marcadores evaluados en la población proveniente de cruce DOR364 x BAT477.....	53
Figura 6. Distribución de la segregación de los marcadores evaluados en la población proveniente de cruce DOR364 x G19833.....	54
Figura 7. Número de nuevos marcadores ligados al mapa DOR364 x BAT477...	59
Figura 8. Mapa de ligamiento de la población derivada de la cruce intra-acervo DOR364 x BAT477 de frijol común.....	62
Figura 9. Número de nuevos marcadores ligados al mapa DOR364 x G19833...	64
Figura 10. Mapa de ligamiento de la población derivada de la cruce inter-acervo DOR364 x G19833.....	69
Figura 11. Evaluación de los productos de amplificación en gel de agarosa al 1,8%.....	72
Figura 12. Alineamiento de las secuencias los padres DOR364, BAT477 y G19833 para el marcador BSn241.....	72

ÍNDICE DE TABLAS

Tabla 1. Porcentaje de homocigosis en generaciones RIL (Törjék <i>et al.</i> , 2006) ...	34
Tabla 2. Resumen de todos los marcadores evaluados en los padres DOR364 y BAT477.	50
Tabla 3. Porcentajes de los marcadores amplificados en la prueba de padres. ...	51
Tabla 4. Resumen de los marcadores actualmente integrados en el mapa genético de la población DOR364 x BAT477.....	60
Tabla 5. Resumen de los marcadores actualmente integrados en el mapa genético de la población DOR364 x G19833.....	63
Tabla 6. Número de transversiones y transiciones encontradas en las secuencias de los tres padres (DB y DG).	74

ÍNDICE DE ANEXOS

ANEXO 1. Protocolo de extracción de ADN por CTAB	101
ANEXO 2. Procedimiento para realizar tinción de geles de acrilamida con Nitrato de Plata	102
ANEXO 3. Marcadores desarrollados a partir de genes de nodulación.	103
ANEXO 4. Protocolo para envío de productos de PCR a secuenciar en la <i>Universidad de Yale</i>	116
ANEXO 5. Segregación mendeliana de los 49 marcadores mapeados en la población proveniente de la cruce DOR364 x BAT477.	116
ANEXO 6. Segregación mendeliana de los 32 marcadores mapeados en la población proveniente de la cruce DOR364 x G19833	117
ANEXO 7. Marcadores mapeados compartidos y exclusivos en cada uno de los dos mapas empleados en este trabajo.....	118
ANEXO 8. Marcadores secuenciados con el peso obtenido luego de editar las secuencias	119

RESUMEN

Con la saturación de un mapa genético, se busca incrementar el número de marcadores en los grupos de ligamiento ya establecidos especialmente en las regiones menos saturadas. El objetivo central de este trabajo fue la saturación del mapa genético de la población proveniente de la cruce intra-acervo DOR364 x BAT477 (DB). Para tal fin se utilizó una población de mapeo obtenida por el método de descendencia de semilla única, constituida por 90 líneas endogámicas recombinantes (RILs). Se lograron anclar 49 marcadores nuevos sobre un mapa base que presentaba 245 marcadores. La longitud total del mapa resultante fue de 1771,3cM, con una distancia promedio entre marcadores de 7,2cM. El 16,3% de los marcadores estuvieron sesgados hacia uno de los dos parentales. En cuanto a la población DOR364 y G19833 (DG), esta también fue obtenida por el método de descendencia de semilla única, constituida por 87 líneas endogámicas recombinantes (RILs). Se lograron anclar 32 marcadores distribuidos a lo largo de todos los grupos de ligamiento (GL). El mapa obtenido presentó una longitud total de 2439,7cM, con una distancia promedio entre marcadores de 4,9cM. El GL más largo fue el B02 con 342,6cM, mientras que el B05 fue el más corto con 143,3cM. Para incrementar el número de marcadores, se diseñaron marcadores basados en genes de nodulación, flanqueando regiones intrónicas. En total se lograron diseñar 204 marcadores, pero solamente se lograron mapear 12 (dos en DB y 10 en DG). Entre los marcadores anclados al mapa DG, se encontraron los marcadores basados en los genes *NIN* y *LHK1*. El primero es un regulador en la etapa de infección controla la curvación de los pelos de la raíz; el otro es un receptor quinasa de citoquinina (implicado en la organogénesis de nódulos de la corteza). Respecto a la búsqueda de SNPs, se lograron reportar 343, de los cuales 150 fueron localizados entre los genotipos DOR364 y BAT477 y 193 entre DOR364 y G19833. En siete de las 44 secuencias analizadas no se encontraron SNPs. Entre los genotipos DOR364 y BAT477 las mutaciones tipo transversión fueron más abundantes (67,4%) que las transiciones (32,6%) y para DOR364 y G19833, ambos tipos de mutaciones se presentaron casi en iguales proporciones (50,3% de transiciones y 49,7% de transversiones).

PALABRAS CLAVE: *Phaseolus vulgaris* L., DOR364, BAT477, G19833, RILs, mapeo genético, SSR, SSCP.

1. INTRODUCCIÓN

Las legumbres (Leguminosae o Fabaceae) son la tercera familia de plantas más numerosa a nivel mundial (Gepts *et al.*, 2005) con alrededor de 643 géneros (18.000 especies) que son agrupados en 40 tribus (Broughton *et al.*, 2003). Las leguminosas, después de los cereales incluyen las especies más importantes de interés en la agricultura, considerando el área cultivada y la producción total (Gepts *et al.*, 2005). La familia Leguminosae se subdivide en tres subfamilias, de las cuales la subfamilia Papilionoidae contiene las legumbres y leguminosas de grano de mayor importancia económica, incluido el género *Phaseolus*. Entre las áreas de aparición natural de *P. vulgaris*, está el rango que va desde México hasta Argentina, a través de la región tropical y subtropical de América (Debouck *et al.*, 1986). La especie fue domesticada por culturas de Centro y Sur América (Grisi *et al.*, 2007) hace de 4 a 7 mil años y progresivamente se fue dispersando alrededor del mundo a otras regiones. El frijol común cultivado fue domesticado desde un *P. vulgaris* silvestre: una pequeña planta con crecimiento indeterminado, que crece en las regiones de clima medio del Neotrópico y subtrópico, y tiene una amplia distribución (Blair *et al.*, 2006a).

De las 50 especies descritas dentro de este género, cinco (*P. vulgaris*, *P. lunatus*, *P. coccineus*, *P. acutifolius*, *P. polyanthus*) son cultivadas para el consumo humano. La tribu Phaseoleae [frijol común (*P. vulgaris*), caupi (*Vigna unguiculata*), y soya (*Glycine max*)], es el grupo más importante económicamente, pues contiene cerca del 75% de las leguminosas comercializadas (Broughton *et al.*, 2003).

Grandes colecciones de germoplasma de las formas domesticadas y silvestres se encuentran en el CIAT (Cali, Colombia) y el USDA (Pullman, Washington, EE.UU) (Gepts, 2001).

El fríjol común ha sido utilizado para derivar importantes principios genéticos. Mendel lo empleó para confirmar sus resultados derivados de los guisantes. Johanssen, usó la especie para ilustrar la naturaleza cuantitativa de la herencia de ciertos caracteres como el peso de las semillas (Gepts, 2001). El fríjol común cultivado presenta dos acervos genéticos: el Mesoamericano o fríjol de semillas pequeñas y el Andino o fríjol de semillas grandes (Debouck, 1999). El acervo genético Mesoamericano es la subdivisión de fríjol común más cultivada que incluye en su mayoría variedades de semillas de colores rojo, negro o crema, y tamaños pequeños (Blair *et al.*, 2010a).

Los acervos Mesoamericano y Andino, por sus características contrastantes han sido muy útiles para desarrollar líneas endogámicas a partir de las cuales se han construido mapas de ligamiento, el primer mapa basado en caracteres morfológicos para fríjol estuvo disponible desde 1990 (Gepts, 1999), pero en la actualidad existen varios y pueden ser distinguidos por muchas características entre las que se encuentran tipos de parentales utilizados, tipo de cruce, población y su respectiva segregación, el tipo y número de marcadores utilizados y la longitud total del mapa obtenido. Gepts *et al.* (2001) realizaron el primer mapa consenso, donde se correlacionaron 12 mapas que habían sido establecidos sobre RFLP, RAPD, isoenzimas, AFLP, ISSR, marcadores microsatélites y marcadores fenotípicos. La distancia promedio total del mapa fue 1200cM, consistentes con el número promedio de quiasmas por bivalente (1,9).

El objetivo principal de este trabajo fue la saturación del mapa genético proveniente de la cruce de los genotipos DOR364 y BAT477. Para estos se evaluaron marcadores moleculares tipo microsatélites (SSR) y marcadores basados en genes. Anexo a esto, también se anclaron marcadores en el mapa genético de la población DOR364 y G19833. Por último se realizaron alineamientos de secuencias para encontrar SNPs entre los tres genotipos antes nombrados.

2. MARCO TEÓRICO

2.1. Origen y domesticación del frijol común

Fue solo hasta finales del siglo XIX cuando los científicos aceptaron que el origen del frijol común se remontaba al Nuevo Mundo. Con base en las primeras observaciones realizadas sobre los restos arqueológicos en el Perú y posteriormente en el suroeste de los Estados Unidos, Wittmack (1888b) concluyó que el frijol común se había originado en las Américas. Esto era contrario a la creencia del origen asiático, que se mantuvo durante varios siglos. Por ejemplo, Linneo en 1753, sugirió que *P. vulgaris* era de origen indio.

Después de los descubrimientos arqueológicos iniciales de Reiss y Stübel en Ancón (Perú), (Wittmack 1880; 1888a), varios restos de frijol se han descubierto, no sólo en los Andes sino también en América Central y Norte América. Estos restos consistían en semillas (Brooks *et al.*, 1962; Kaplan *et al.*, 1973), fragmentos de vaina (Kaplan, 1981), e incluso plantas completas (Kaplan & MacNeish, 1960). Varios de los hallazgos arqueológicos son fenotípicamente similares a los cultivares actuales localizados en las mismas áreas. (Gepts & Debouck, 1991).

2.2. Generalidades de frijol

El frijol (*Phaseolus* spp. L.), es uno de los cultivos más antiguos del nuevo mundo. Junto con el maíz, la ahuyama y la yuca han sido alimentos básicos importantes en las altitudes bajas y medias de América por milenios. El frijol es muy diverso en términos de métodos de cultivo, usos, rangos de ambientes a los que se han adaptado y variabilidad morfológica. Ellos se han encontrado desde 0 a 3000 msnm; son cultivados como monocultivos, en asociación o en rotación. El frijol es

consumido como grano maduro, semilla inmadura o como un vegetal (hojas y vainas) (Broughton *et al.*, 2003).

La variedad que presenta hábito de crecimiento arbustivo genera una competencia mínima y permite la siembra con otras especies, por ejemplo, en proyectos de reforestación, entre los árboles frutales o las plantaciones de café durante los primeros años hasta la cosecha. El frijol de hábito de crecimiento voluble también ofrece potencial para cultivos intercalados, pero requiere de tutorado por ejemplo con maíz o con estacas de guadua.

En los últimos 20 años, el frijol ha sido cada vez más cultivado a escala comercial y ahora se ofrece en los mercados nacionales, regionales e internacionales. Este provee proteínas a la dieta, las cuales desempeñan un papel esencial en la nutrición humana, en complemento de otros alimentos (por ejemplo, el maíz en las tierras altas de América Latina y del Este de África y el arroz en Brasil) que son principalmente cereales, fuente de carbohidratos. Las semillas de frijol contienen entre 20 y 25% de proteínas, muchas de las cuales está formada por el almacenamiento de faseolina (Ma & Bliss, 1978). La faseolina es un factor determinante de la cantidad y la calidad nutricional de las proteínas en las semillas de frijol (Broughton *et al.*, 2003).

2.3. Genoma de *Phaseolus vulgaris* L.

Phaseolus vulgaris L., es una especie diploide ($2n = 2x = 22$), anual y predominantemente autógama (Bellucci *et al.*, 2010). El tamaño del genoma de frijol es pequeño (635 Mpb / genoma haploide) y similar en su naturaleza como diploide verdadero al de arroz (340 hasta 560 Mpb / genoma haploide), que es generalmente considerada como la planta de importancia económica con el genoma más pequeño. El tamaño de los cromosomas mitóticos y meióticos es de aproximadamente 1-3 μ m de longitud, la mayoría de los cromosomas son

metacéntricos o submetacéntricos, y presentan una media de 1.9 quiasmas por cromosomas bivalentes en la meiosis por un entrecruzamiento por brazo del cromosoma (Maréchal, 1971). Un cariotipo fue desarrollado para *P. vulgaris* y *P. coccineus* basados en cromosomas politénicos de células embrionarias en suspensión (Pedrosa *et al.*, 2003).

En cuanto a trabajos realizados en citogenética, Pedrosa *et al.* (2003), reportaron la integración de los mapas genéticos y cromosomales de frijol común por medio de la hibridación *in situ* de sondas RFLP para los cromosomas mitóticos de dos cultivares. Pedrosa-Harand *et al.* (2009) establecieron mapas citogenéticos basados en la hibridación de fluorescencia *in situ* de sondas hechas de clones BAC completos y algunos otros clones genómicos para los primeros tres cromosomas de frijol común. Estos mapas permitieron la evaluación de las frecuencias de recombinación a lo largo de estos cromosomas, y generaron una visión general de la organización del genoma del frijol común.

2.4. Mapas genéticos en frijol

Existen varios mapas de ligamiento desarrollados en frijol común y se diferencian principalmente por características como tipos de parentales y de población utilizados, número de individuos, tipo de marcadores usados y la longitud total del mapa obtenido (Gepts, 1999). Freyre *et al.* (1998) construyeron un mapa de ligamiento integrado basado en una población endogámica recombinante de la cruce inter-acervo BAT93 x JaloEEP558 (BJ), con una longitud total de 1226cM, que alineaba tres mapas de RFLP desarrollados en frijol común. Este mapa comprendía 563 marcadores moleculares, incluyendo RFLP, RAPD, isoenzimas y *loci* de marcadores fenotípicos. Sin embargo, la ubicación general en el mapa como número de marcadores por GL y ubicación aproximada dentro de los grupos, se determinó por unos 1070 marcadores. La integración de microsatélites

en el mapa molecular de frijol, fue realizada por primera vez por Yu *et al.* (2000) con una asignación de 15 marcadores.

Blair *et al.* (2003) desarrollaron un mapa con un mayor número de marcadores microsatélites ligados al genoma de frijol común. Estos marcadores, fueron utilizados para unir dos mapas genéticos basados en RFLP: el de Vallejos *et al.* (1992), donde lograron un mapa con una longitud total 960cM y el de Freyre *et al.* (1998), en el cual obtuvieron un mapa cuya longitud total fue de 1226cM.

El esfuerzo por desarrollar marcadores basados en genes comenzó con Yu *et al.*, (2000) y Blair *et al.* (2003). Hugaard *et al.* (2008), también diseñaron marcadores de este tipo denominados “Leg”. Estos fueron basados en el alineamiento de los ESTs homólogos de *Lotus japonicus*, *Medicago truncatula* y soya; las regiones exónicas conservadas fueron identificadas. Para este trabajo, el objetivo era convertir el 50% de estas secuencias en marcadores de anclaje para varias leguminosas y posterior a esto, comparar posiciones correspondientes de estos marcadores en los mapas de ligamiento de *Lotus japonicus*, *Medicago truncatula*, frijol y maní (*Arachis hypogea*), mostrando los diferentes niveles de conservación del genoma. Encontraron un alto nivel de macrosintenia entre frijol y *Lotus*, incluyendo grupos de ligamiento completos. También encontraron sintenia entre frijol y *Medicago*, en regiones que cubrían segmentos sin aparente sintenia de frijol y *Lotus*.

Hanai *et al.* (2009) realizaron una investigación en la que pusieron a prueba nuevos marcadores SSR desarrollados a partir de ESTs de frijol (PvM) en un conjunto de genotipos con el fin de incrementar la densidad del mapa BJ con otros marcadores como RGA (*Resistance Gene Analog*) y AFLP (*Amplified Fragment Length Polymorphism*). Con esto, obtuvieron un mapa cuya longitud total fue de 1259cM y una distancia promedio entre marcadores de 3cM.

En cuanto a mapas de ligamiento intra-acervo, Cichy *et al.*, (2009), construyeron uno para la población endogámica recombinante derivada del cruce G19833 y AND696, que son dos genotipos Andinos de frijol común contrastantes para tolerancia a suelos con bajo fósforo. La longitud total del mapa fue de 1105cM, dentro de los cuales contenía 167 marcadores con 64 SSR, 11 RAPD, 91 AFLP, y un marcador fenotípico. Otro mapa de ligamiento generado a partir del cruce de dos genotipos Andinos fue realizado por Blair *et al.* (2010b). Para la generación de la población, se realizaron cruces a partir de los genotipos G21242 y G21078. Al mapa se lograron anclar 118 marcadores (74 SSR, 46 RAPD y 2 marcadores basados en genes), presentando una longitud total de 726,0cM.

Actualmente existen pocas publicaciones de mapeo derivados de cruces entre genotipos mesoamericanos como el de la población intra-acervo DOR364 y BAT477 (DB), la cual fue realizada por Blair *et al.* (2010a), en donde lograron anclar 186 marcadores entre los cuales habían AFLP (22) RAPD (104) y SSR (60), dando como longitud total del mapa 1087,5cM. Otro trabajo realizado con cruce entre genotipos mesoamericanos (BAT881 y G21212) fue desarrollado por Frei *et al.* (2005), quienes lograron obtener un mapa de ligamiento con 115 marcadores (26 microsatélites y 89 RAPD), una longitud total de 611,2cM y una distancia promedio entre marcadores de 4,3cM. Blair *et al.* (2010c), desarrollaron una población RIL a partir de dos genotipos mesoamericanos: G4825 y G14519, con la cual construyeron un mapa con longitud total de 915,4cM, que contenía 68 marcadores tipo SSR y 46 RAPD, dando una distancia promedio entre marcadores de 8cM.

2.5. Fijación de nitrógeno y nodulación

2.5.1. Fijación de nitrógeno

El nitrógeno es probablemente uno de los nutrientes más importantes que requieren las plantas, siendo un componente esencial de todos los aminoácidos y ácidos nucleicos. Sin embargo, la disponibilidad de nitrógeno es limitada en muchos suelos, y aunque la atmósfera terrestre se compone de 78.1% de gas nitrógeno (N_2), las plantas no pueden utilizarlo de esta forma. Para compensar esta limitación, la agricultura moderna ha sido altamente dependiente de fertilizantes industriales con nitrógeno químico para lograr la máxima productividad de los cultivos. La fijación biológica de nitrógeno es una alternativa a este tipo de fertilizantes. Esta fijación se lleva a cabo por procariontas utilizando un complejo enzimático llamado nitrogenasa y como resultado el N_2 atmosférico se reduce a formas en que la planta es capaz de utilizar como el amoníaco (Ferguson *et al.*, 2010).

Las leguminosas, han desarrollado una relación simbiótica con unas bacterias específicas del suelo, llamadas rizobios (incluyendo los géneros *Azorhizobium*, *Allorhizobium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* y *Sinorhizobium*). Una vez la simbiosis se establece, los rizobios fijan el nitrógeno atmosférico y lo proporcionan a la planta hospedera. Debido a que el nitrógeno es un factor clave que limita el crecimiento y desarrollo vegetal y la capacidad de las leguminosas para entrar en una simbiosis con fijadores de nitrógeno, los rizobios les proporcionan una clara ventaja sobre otras especies de plantas en suelos pobres de este nutriente esencial. Esta asociación simbiótica es la más importante en términos de fijación biológica de nitrógeno, produciendo alrededor de 200 millones de toneladas de nitrógeno al año (Graham & Vance 2003; Peoples *et al.*, 2009).

2.5.2. Biología de la nodulación

La biología de la interacción leguminosa-rhizobio es compleja, por tener varias etapas y requerimientos para su funcionalidad. En esta interacción se da lugar a la formación de un nuevo órgano en la raíz llamado *nódulo* en el que la bacteria convierte el gas nitrógeno atmosférico en formas de nitrógeno que son utilizables por la planta (Ferguson *et al.*, 2010). La planta controla estrictamente el número de nódulos que va a formar a través de una señalización compleja raíz-brote-raíz, ciclo denominado autorregulación de la nodulación (AON del inglés *Autoregulation Of Nodulation*). Este proceso de regulación involucra hormonas peptídicas, receptores de quinasas y pequeños metabolitos.

La AON es un sistema de retroalimentación en el que la infección en las primeras etapas restringe el desarrollo de una nueva infección. En los mutantes, la pérdida de función de AON da lugar a abundantes nódulos llamado hiper o super nodulación (Gresshoff, 2005). Usando técnicas modernas de genética y genómica, muchos de los componentes necesarios para la formación y AON de nódulos se han aislado (Ferguson *et al.*, 2010) tanto de la parte de la planta como de su ente simbiótico.

2.5.3. Genética de la nodulación

Los genomas de las bacterias *Rhizobium* albergan alrededor de 50 genes relacionados con la nodulación que son transcripcionalmente activos por los exudados de las raíces que generalmente pertenecen a la familia de las flavonas/ isoflavonas. Los genes de nodulación cooperan para sintetizar una nueva clase de reguladores del crecimiento en la planta llamados "factores de nodulación (Nod)". Estos son oligosacáridos que poseen de tres a cinco unidades ligadas b(1-4) N-acetil-D glucosamina modificado por un gran ácido graso en la posición del

C₂ del azúcar en su extremo no reductor y varios radicales menores que involucran sulfato, acetato, o fucosa, (Gresshoff, 2005).

Diferentes especies de *Rhizobium* producen diversos radicales, lo que ayuda a ser muy específico frente al hospedero. Algunos rhizobios (tal como la cepa NGR234) produce un coctel de factores que son capaces de formar múltiples nódulos en el hospedador. Los factores Nod son percibidos por un mecanismo de reconocimiento específico recientemente dilucidado mediante la clonación posicional de genes que controlan los fenotipos no nodulantes (Gresshoff, 2005).

Ya se sabía que algunos mutantes de leguminosas (como *Ljsym1* y *Ljsym5*) únicamente carecían de la capacidad para responder a la infección de *Rhizobium*, mientras que otros (tales como *Ljsym2*, *Mtdmi2* y *MsMN1007*) mostraban el mismo fenotipo no nodulante también carecían de la capacidad para formar simbiosis con hongos micorrícicos. Así, el *Rhizobium* y la simbiosis con la micorriza comparten componentes comunes, reflejando la modificación evolutiva de la simbiosis antigua con el hongo para la moderna simbiosis con el nódulo. Es aún desconocida la manera como se produce la señalización de la simbiosis de hongos, pero es probable que el hongo libere un equivalente al factor Nod, tal vez similar a los elicitores de los hongos patógenos conocidos en el campo (Gresshoff, 2005).

2.5.3.1 Genes involucrados en la nodulación

Mediante la clonación desde varias legumbres simbióticas mutantes, se han logrado reconocer genes involucrados durante varias etapas del proceso: en el reconocimiento de señales de nodulación con rhizobios, en las cascadas de señalización simbiótica temprana, en la infección, proceso de nodulación y la regulación de la fijación de nitrógeno.

*Percepción de las señales microbianas: En cuanto a la percepción de los Factores de Nodulación (en inglés NF), el modelo actual muestra a los receptores de quinasa (en inglés RLK) localizados sobre las células epidermales. Estos han sido identificados en varias especies de legumbres: en *L.japonicus* y soya (*NFR1* y *NFR5*) (Indrasumunar *et al.*, 2010), en *P.sativum* (*PsSYM2A* y *PsSYM10*) y en *M.truncatula* (*MtLYK3/ MtLYK4* y *MtNFP*) (Ferguson *et al.*, 2010).

*Señalización simbiótica temprana: Cerca de la mitad de las leguminosas mutantes no nodulantes aisladas hasta el momento son defectuosas en simbiosis de micorriza arbuscular (en inglés AM), lo que implica que los genes responsables de estos mutantes son requeridos para ambas simbiosis: raíz nodulante (en inglés RN) y AM. Las señales de transducción mediadas por estos genes es llamada “*common symbiosis pathway*” (CSP). La actividad de estos genes gira en torno a la liberación de Ca^{+2} , que es la reacción fisiológica central en el CSP. Los genes cumplen funciones antes y después de la liberación del calcio. En *L.japonicus* por ejemplo, existen cinco genes CSP (*SYMRK*, *CASTOR*, *POLLUX*, *NUP85* y *NUP133*) que son requeridos para la liberación de calcio (Stracke *et al.*, 2002; Imaizumi-Anraku *et al.*, 2005; Kanamori *et al.*, 2006; Saito *et al.*, 2007), y dos más (*CCaMK* y *CYCLOPS*.) se encuentran luego que esto sucede (Tirichine *et al.*, 2006; Yano *et al.*, 2008), cumpliendo funciones importantes para la transmisión de señales simbióticas mediadas por la liberación de calcio.

Otros factores importantes en este proceso de señalización fueron aislados de *M. truncatula* (*HMGR1*) y *L.japonicus* (*SIP*). Se cree que el *HMGR1* está involucrado en el proceso de infección por parte del rizobio y la organogénesis del nódulo (Kevei *et al.*, 2007), mientras que *SIP* es una nueva proteína de unión al ADN con dominios ricos en A-T, que se une específicamente al promotor del gen *NIN* (Zhu *et al.*, 2008). Recientemente, Lefebvre *et al.* (2010) aislaron una remorin simbiótica (*MtSYMREM1*) de *M. truncatula*, que es expresada exclusivamente en el proceso de la nodulación. Remorin son unas proteínas filamentosas que han sido implicadas en funciones de membrana celular y algunos miembros de esta familia

multigénica han sido descritos por estar involucrados con interacciones bióticas en la planta. Esta es una proteína oligomérica que se une a la membrana plasmática que rodea la bacteria y controla la infección y la liberación de los *rizhobium* en el citoplasma del hospedero. También está relacionada con el conjunto básico de receptores simbióticos que son esenciales para la percepción de las moléculas de señalización bacteriana.

La característica más destacada de *MtSYMREM1* es que este interactúa con *LYK3*, *NFP* y *DM12* que son los ortólogos de *NFR1*, *NFR5* y *SYMRK* en *L.japonicus*, respectivamente.

*Regulación sistémica de la simbiosis: Aunque la simbiosis para la fijación de nitrógeno es altamente benéfica para la legumbre hospedera, una excesiva nodulación interfiere en el crecimiento de la planta porque este proceso es energéticamente costoso. Para balancear la simbiosis, las plantas tienen mecanismos internos y externos para controlar el número de nódulos. Una señal interna importante es el sistema de reacción que involucra señales de larga distancia (desde la raíz hasta el ápice) (AON), que muestran una estrecha relación con la señalización simbiótica temprana provocada por los NF (Caetano-Anollés & Gresshoff 1991; Oka-kira & Kawaguchi, 2006). En leguminosas, los genes representados en *L.japonicus* *HAR1* son ampliamente expresados en varios órganos pero no en el ápice, sugiriendo que estos genes están involucrados únicamente en la producción de inhibidores de nodulación desde la raíz. Junto con *HAR1*, *KLAVIER (KLV)* también es indispensable para la señalización de AON en *L.japonicus* (Oka-kira *et al.*, 2005). Independientemente del AON, el papel del etileno es caracterizado como un regulador negativo de la nodulación.

Se ha demostrado que el gen *SICKLE* en *M.truncatula* codifica una proteína de señalización de etileno, que es ortólogo al gen *EIN* en *Arabidopsis*. (Penmetza *et al.*, 2008). Una de las señales externas, es el nitrógeno del suelo. Altas concentraciones de nitrógeno tales como nitrato o amonio abolen la nodulación, pero han encontrado que en *L.japonicus* el gen *CLE-RS2* regula de manera

eficiente la respuesta de las raíces ante la presencia de nitrato (Okamoto *et al.*, 2009).

*Proceso de infección y organogénesis del nódulo: La infección del rhizobio se lleva a cabo inicialmente en la epidermis de la raíz (pelos radiculares). La respuesta de los pelos radiculares a los NFs es restringida a una región axial en particular, la zona donde se desarrollan los pelos radiculares. Con el fin de incorporar eficientemente la bacteria, muchas legumbres han evolucionado a una estructura para lograr la invasión: los hilos de infección (IT del inglés *Infection Threads*) (Sprent, 2007). Hasta el momento se han clonado y reportado once genes que son importantes para la coordinación progresiva del crecimiento de los IT y para la organogénesis del nódulo (*NIP*, *LATD*, *RPG*, *ERN1*, *CERBERUS*, *LIN*, *PIR1*, *NAP1*, *CYCLOPS*, *HK1*, *CRE1*) (Kouchi *et al.*, 2010). Sin embargo hay casos como el del gen *LHK1*, el cual juega un papel importante solo durante el proceso de organogénesis del nódulo (Tirichine *et al.*, 2007).

El gen *NSP2* codifica para un GRAS, que es un tipo de regulador de transcripción que es requerido para la División Celular Cortical (CCD) y el curvamiento de los pelos radiculares. El *NSP2* interactúa con otro gen que al igual que el anterior codifica para un GRAS, el *NSP1*. Esta interacción genética es necesaria para la organogénesis y el *NSP1* se asocia con promotores de genes de nodulación temprana tales como *ENOD11*, *NIN* y *ERN1* (Hirsch *et al.*, 2009).

Los genes descritos arriba son esenciales para la iniciación de la IT o su crecimiento, pero no para la inducción de la CCD, excepto *LHK1*. Sin embargo, ellos también están involucrados en la organogénesis del nódulo de una u otra manera (Kouchi *et al.*, 2010).

*Regulación del hospedero de la fijación de nitrógeno: Casi simultáneamente con la aparición de los nódulos, los rhizobios son liberados desde los IT hacia las células del nódulo que a su vez están presentes en el citoplasma de las células del hospedero encerradas por una membrana peribacterioide (PBM del inglés *PeriBacteroid Membrane*). Se han encontrado cerca de 1000 genes que son

inducidos en mayores proporciones en los nódulos, entre estos, los genes de nodulación tardía, que son activados justo antes de iniciar la fijación de nitrógeno, estando involucrados en la regulación y fijación en el hospedero. Un ejemplo es la Leghemoglobina (Lb), que es una proteína específica del nódulo y presenta unión con el oxígeno. Su función es mantener las concentraciones bajas de O₂ en las células infectadas por el nódulo y facilitar el suministro de O₂ a las bacterias para su respiración aeróbica (Kouchi *et al.*, 2010). En *L.japonicus* fueron identificados tres genes (*SST1*, *FEN1* y *IGN1*) y en *M. truncatula* uno (*DNF1*), los cuales fueron registrados por análisis de fijación de mutantes con el fin de tratar de tener más claridad en cuanto a los mecanismos de regulación del hospedador. En el caso del gen *FEN1* este fue clonado y se demostró que codifica para una homocitrato sintasa (HCS) que es expresada específicamente en las células de los nódulos (Hacoyama *et al.*, 2009). El homocitrato es un compuesto inusual en las plantas superiores, de hecho se encontró en abundancia solamente en nódulos de raíces de leguminosas. Este es conocido por ser un componente del cofactor hierro-molibdeno en el complejo nitrogenasa, en el que se cree que ocurre la fijación de nitrógeno (Hoover *et al.*, 1987).

Los nódulos formados en una *L.japonicus* mutante para el gen *IGN1* fueron caracterizados por una senescencia prematura, rápida y severa. Esto es causado por el gen *IGN1*, que codifica una nueva proteína de membrana: repeticiones de anquirina que está presente en la membrana plasmática (Kumagai *et al.*, 2007). La expresión de este gen se da en todos los órganos de la planta en niveles bajos.

2.6. Mapeo genético

2.6. 1. Mapeo genético

Mapeo genético, (también conocido como mapeo de ligamiento o mapeo meiótico), hace referencia a la determinación de la posición relativa y distancia entre marcadores a lo largo de los cromosomas. La distancia genética entre dos

marcadores es definida por el número de eventos de recombinación que se dan en una cromátida (Semagn *et al.*, 2006). La construcción de un mapa genético requiere que el investigador desarrolle una población de mapeo apropiada, decida el tamaño de la muestra y el tipo de marcadores moleculares a emplear. La construcción de mapas genéticos detallados con altos niveles de cubrimiento del genoma es el primer paso para algunas de las aplicaciones de los marcadores moleculares en el mejoramiento de plantas (Tanksley *et al.*, 1989). La elaboración de este tipo de mapas: a) permite el análisis genético detallado de características cuantitativas y cualitativas que ayudan a la localización de genes (Mohan *et al.*, 1997; Doerge, 2002; Yim *et al.*, 2002), b) facilita la introgresión de genes deseables o QTLs a través de marcadores de selección asistida, c) permite el mapeo comparativo entre diferentes especies para evaluar la similaridad entre genes y la funcionalidad en la expresión de un fenotipo (Ahn & Tanksley, 1993; Paterson *et al.*, 2000), d) provee un marco para el anclaje del mapa físico basado en translocación de cromosomas, secuencia de ADN u otras medidas directas (Yim *et al.*, 2002), y e) constituye el primer paso hacia la clonación posicional de genes responsables para características de importancia económica (Mohan *et al.*, 1997; Vuysteke *et al.*, 1999).

2.6.2. Principios del mapeo genético

El mapeo genético está basado sobre el principio que los genes (marcadores o *loci*) segregan a través de la recombinación de los cromosomas durante la meiosis (reproducción sexual), lo cual permite su análisis en la descendencia (Paterson, 1996). Durante la meiosis, los cromosomas se ponen al azar en los gametos, de modo que la segregación de los alelos en los genes es independiente. Esto se afirma en la segunda ley de Mendel que es conocida como la ley de segregación independiente. Esta ley siempre ocurre con los genes que se encuentran en cromosomas diferentes, pero no siempre es válida para los genes que están en el mismo cromosoma. Cuando dos genes están juntos, ellos no segregan de manera independiente y se dice que están ligados. Los genes que están

estrechamente ligados pueden ser transmitidos juntos desde sus padres hacia la prole más frecuentemente que aquellos genes que se localizan distantes entre sí. En el inicio de la meiosis, un par de cromosomas homólogos se entrelazan formando un quiasma para intercambiar secciones cromosomales. Tal proceso es llamado *recombinación*, porque las moléculas de ADN cambian fragmentos continuos entre sí para generar un rearrreglo de la información genética en un organismo. La pareja de cromosomas se separan para formar gametos con nuevas combinaciones de genes que difieren de cualquiera de los padres (Semagn *et al.*, 2006), por ejemplo en la generación F₂.

La fracción de recombinación observada entre dos *loci* es un estimativo de la mitad del número de quiasmas, porque este ocurre en un estado de cuatro hebras o tétradas para un único evento de entrecruzamiento, donde son dos de las cromátidas no hermanas las que participan en el cruzamiento, mientras que en las otras dos no hay intercambio de segmentos cromosomales. Tal proceso produce dos tipos de gametos: a) si la recombinación no ocurre, el producto son gametos parentales. b) si hay recombinación los productos son gametos recombinantes. La posibilidad que un cruce produzca recombinación entre los genes está directamente relacionada con la distancia que hay entre estos. La frecuencia de recombinación es menor entre dos marcadores que se encuentran situados muy cerca uno del otro dentro de un cromosoma y en cuanto mayor sea la frecuencia de recombinación entre dos marcadores, es que se encuentran más distantes en un cromosoma o en cromosomas separados (Semagn *et al.*, 2006).

2.6.3. Grupos de ligamiento (GL)

Los marcadores son asignados a los grupos de ligamiento (GL) usando la relación entre las probabilidades que dos *loci* estén ligados con un valor de recombinación dado, sobre la probabilidad que no estén ligados. Esta relación es llamada como el logaritmo de odds (LOD) (Risch, 1992; Stam, 1993a). Los valores de LOD

usados para establecer los GL y calcular las distancias del mapa son llamadas “linklod” y “maplod”, respectivamente (Stam, 1993b; Ortiz *et al.*, 2001).

Los pares de marcadores con un valor de LOD por encima del “linklod” crítico son considerados como ligados, mientras que aquellos con un LOD menor al valor del “linklod”, se determinan como no ligados. Varios investigadores usan un valor de “linklod” de tres como umbral mínimo para decidir si un *loci* está o no ligado. Un LOD de tres entre dos marcadores indica que el ligamiento es 1.000 veces más probable que el no ligamiento (Stam, 1993a). Los valores más altos de LOD resultan en un mayor número de GL, cada uno con un número más pequeño de marcadores, mientras que los valores bajos de LOD tenderán a crear pocos GL con gran número de marcadores por grupo. Lo ideal es poder llegar a tener tantos GL como número de cromosomas haploides tiene la especie de estudio (Semagn *et al.*, 2006).

2.6.4. Distancia del mapa y orden de los marcadores

Para calcular la distancia del mapa y determinar el orden, los investigadores necesitan varios patrones específicos, incluyendo un valor del umbral de la recombinación, un “maplod” y una función de mapeo (Semagn *et al.*, 2006). Existen dos funciones de mapeo (Haldane y Kosambi), que traducen las frecuencias de recombinación dentro del mapa en distancias y viceversa. La función de mapeo de Haldane (Haldane, 1931) asume que no hay interferencia entre los cruces en la meiosis, mientras que en la función de mapeo de Kosambi (Kosambi, 1944) se asume un cierto grado de interferencia. La interferencia es tomada como el efecto que genera la aparición de un cruce en una determinada región y que reduce la probabilidad de recombinación en la región adyacente.

Tales funciones de mapeo convierten las fracciones de recombinación dentro del mapa en unidades llamadas centimorgans (cM) en honor al genetista Thomas

Hunt Morgan. Por definición, una unidad de mapeo (m.u.) es igual al 1% de fenotipos recombinantes o 1cM. Para las frecuencias de recombinación superiores al 10%, la función de mapeo de Haldane da mayores distancias en el mapa que Kosambi (Semagn *et al.*, 2006).

Vallejos *et al.* (1992), estimaron un valor promedio relacionando las distancias físicas y genéticas, dando una relación de 530 kb/cM para *P. vulgaris*, que es sólo dos veces la estimada para *Arabidopsis thaliana* (230 kb / cm).

2.6.5. Prueba para la distorsión de la segregación

Para cada marcador segregante, un análisis de chi cuadrado (X^2) es necesario para determinar la desviación de la segregación esperada para la población de mapeo (1:1 para marcadores dominantes y codominantes en: BC, RIL, DH y NIL; 1:2:1 para marcadores codominantes en F_2 y 3:1 para marcadores dominantes en F_2) (Semagn *et al.*, 2006). Una desviación de las frecuencias genotípicas observadas desde las esperadas dentro de una población segregante se llama distorsión de la segregación. Esta distorsión puede ocurrir debido a diferentes razones: sesgos estadísticos, genotipificación y errores de puntuación (Plomion *et al.*, 1995) y por razones biológicas como la pérdida de cromosomas, la competencia entre los gametos para la fertilización preferencial gametocidal, o genes asesinos de polen (aborto de gametos masculinos o femeninos), los genes de incompatibilidad o rearreglo de cromosomas no homólogos (Semagn *et al.*, 2006).

2.7. Poblaciones de mapeo y RILs

2.7.1. Poblaciones de mapeo

El primer paso para producir una población de mapeo es seleccionar dos padres genéticamente divergentes, que muestren diferencias fenotípicas y alelos diferenciales para uno o más caracteres de interés, que exhiban suficiente polimorfismo y que al mismo tiempo no sean genéticamente tan distantes con el fin de evitar esterilidad en la progenie y mostrar altos niveles de distorsión en la segregación durante el análisis de ligamiento (Semagn *et al.*, 2006).

Los estudios de simulación realizados con un tamaño de muestra que van desde 50 a 1000 individuos de poblaciones F_2 (retrocruzas), BC (Back Cross), RILs y DHs (Dobles Haploides del inglés *Doubled Haploid*), muestran que el tipo y el tamaño de la población puede ejercer una influencia sobre el mapa genético (Ferreira *et al.*, 2006):

1. Poblaciones con un número bajo de individuos provee varios grupos de ligamiento y un orden inexacto de los locus.
2. Los mapas más precisos fueron obtenidos por poblaciones F_2 y RIL con marcadores co-dominantes, de una a varias generaciones de recombinaciones respectivamente.
3. Un alto número de individuos en la población de estudio ayuda a la generación de un mapa más preciso.
4. Para todos los tipos de poblaciones, un total de 200 individuos es requerido para la construcción de un mapa de ligamiento razonablemente preciso.

En la práctica, el tamaño de las poblaciones usado en estudios preliminares de mapeo genético varió de 50 a 250 individuos (Mohan *et al.*, 1997), pero una población de gran tamaño es necesaria para obtener una alta resolución en mapeo fino.

Las BC y las F_2 son poblaciones de mapeo sencillas porque son fáciles de construir y se requiere de poco tiempo para producirlas. Sin embargo, ambas son consideradas poblaciones temporales porque son altamente heterocigotas y no pueden ser propagadas de manera indefinida a través de semillas. En cambio, las RILs, NILs (Líneas Isogénicas Cercanas del inglés *Nearly Isogenic Lines*) y DHs son poblaciones permanentes porque ellas son líneas homocigotas que pueden ser multiplicadas por semilla y reproducibles sin que ocurran cambios genéticos (Semagn *et al.*, 2006). Las poblaciones F_1 híbridas entre padres heterocigotos o poblaciones F_2 son reproducibles solamente de manera clonal.

Por lo general, sólo las plantas autocompatibles permiten la generación de líneas con un máximo grado de homocigosis. En conclusión, el material vegetal determina la elección de una población de mapeo. Otros factores son el tiempo disponible para la construcción de la población, el tamaño poblacional y la resolución de la cartografía necesaria (Schneider, 2005).

2.7.2. Líneas Endogámicas Recombinantes (RILs)

Las líneas híbridas recombinantes (RILs de las siglas en inglés *Recombinant Inbred Lines*) son la progenie homocigota resultado de la autofecundación en las generaciones de la F_2 hasta la F_6 o más. El concepto de RIL para el mapeo de genes fue desarrollado originalmente para la genética del ratón. En los animales, son necesarias, aproximadamente 20 generaciones de apareamientos entre hermanos para obtener niveles útiles de endogamia y homocigosis (Schneider, 2005).

En plantas, las líneas RI son producidas por autofecundación, a menos que la especie sea totalmente auto-incompatible. En el proceso de autofecundación una semilla de cada línea es la fuente para la próxima generación, y por eso las RILs

también se conocen como líneas descendientes por semilla única. La autopolinización permite la producción de RILs en un número relativamente corto de generaciones. De hecho, dentro de seis generaciones, se puede conseguir la homocigosis casi completa (Schneider, 2005), como se ve en la Tabla 1.

Tabla 1. Porcentaje de homocigosis en generaciones RIL (Törjék *et al.*, 2006)

Generaciones de endogamia (RIL)	% homocigosis dentro de cada locus de las líneas
F3:4	75.0
F4:5	87.5
F5:6	92.25
F6:7	96.875
F7:8	98.4375
F8:9	99.21875

En el genoma de cada RIL, los bloques de los alelos derivados de cada parental se alterna. Debido a que la recombinación ya no puede cambiar la constitución genética de las RILs, la segregación en la progenie de estas líneas es ausente. Una ventaja de los RILs es que debido a que se someten a varias rondas de la meiosis antes que la homocigosis sea alcanzada, el grado de recombinación es mayor en comparación con las poblaciones F_2 de las cuales son derivadas. En consecuencia, las poblaciones creadas a partir de las RIL muestran mejores resultados en cuanto a la alta resolución de los mapas generados versus los mapas realizados a partir de poblaciones F_2 (Schneider, 2005).

2.8. Marcadores moleculares

2.8.1. Marcadores microsatélites

Los microsatélites o SSR (del inglés *Simple Sequence Repeat*) son marcadores moleculares que consisten en repeticiones en *tándem* de uno a seis nucleótidos, siendo los más comunes los di-, tri- y tetranucleótidos (Buso *et al.*, 2006; Benchimol *et al.*, 2007). Los marcadores microsatélites son útiles para los estudios

genéticos porque ellos son codominantes, multialélicos, son transferibles entre diferentes genotipos y poseen altas tasas de polimorfismo. Esto quiere decir que distinguen múltiples alelos en una especie de planta debido a su variación en el número de unidades de repetición (motivos). Están dispersos principalmente en las regiones que se encuentran entre los genes y algunas regiones no codificantes a lo largo del genoma (Nguyen & Wu, 2005). La información generada por estos marcadores permite la comparación e intercambio de información entre diferentes estudios, especialmente para comparar mapas genéticos (Grisi *et al.*, 2007). Este sistema de marcadores está basado en PCR y es usado de manera amplia para huellas de ADN, mapeo genético, MAS, estudios de diversidad genética y genética de poblaciones. Varios grupos de investigación han tenido avances en el desarrollo de marcadores microsatélites para varias especies de la familia Leguminosae (Grisi *et al.*, 2007).

Los microsatélites son clasificados de acuerdo al tipo de repeticiones en su secuencia como: perfecto, imperfecto, interrumpidos o compuestos. En un microsatélite perfecto, la secuencia de repetición no es interrumpida por ninguna base diferente al motivo (*ej.* TATATATATATATA), mientras que en un microsatélite imperfecto hay un nucleótido entre los motivos repetidos que no coincide con la secuencia (*ej.* TATATATA**C**TATATA). En el caso de los interrumpidos hay una pequeña secuencia dentro de la secuencia repetitiva (*e.g.* TATATA**CGTGT**TATATATA), mientras que en un microsatélite compuesto la secuencia contiene dos secuencias repetitivas distintas pero adyacentes (*ej.* TATATATATA**AGTGTGTGTGTGT**) (Oliveira *et al.*, 2006).

El uso exclusivo de microsatélites y la disponibilidad de un mapa referencia puede ayudar en el desarrollo de otros mapas genéticos para frijol común, a través de la transferencia de información del orden y ligamiento de los marcadores, quienes podrían permitir análisis comparativos e integración de mapa (Grisi *et al.*, 2007; Blair *et al.*, 2008).

2.8.2. BAC-ends

Los BAC-ends corresponden a los extremos de los clones BAC y pueden relacionarse con porciones de ADN génico o genómico. El uso de BAC-ends derivados de BAC mapeados físicamente en *contigs* tiene la ventaja de facilitar la integración de los mapas físico y genético, supliendo necesidades, como la localización física de genes o regiones del genoma que codifiquen para características de interés y el desarrollo de análisis comparativos con especies cercanas (Córdoba, 2009). Muchas veces los BAC-ends proveen mucha información como marcadores físicos/genéticos para el mapa del genoma de fríjol como en los estudios realizados por Córdoba *et al.* (2010a y 2010b).

2.8.3 SSCP (*Single-Strand Conformational Polymorphism*)

La técnica de SSCP se basa en las diferencias conformacionales de fragmentos de ADN de cadena simple que pueden ser detectadas como los cambios de la movilidad en la electroforesis de gel de poliacrilamida no denaturante (Orita *et al.*, 1989). Esta técnica es fácil y económica para implementar, y se ha utilizado para el análisis de genes o EST derivados de marcadores SNP en varias especies de plantas (Galeano *et al.*, 2009a). Los marcadores tipo SSCP, tal como los microsatélites, están basados en PCR, son polimórficos y codominantes, pero, debido a que son desarrollados en SNPs, son más frecuentes y más fáciles de diseñar que los microsatélites. (Kuhn *et al.*, 2008).

Bajo esta técnica, Galeano *et al.* (2009a), desarrollaron un mapa basado en marcadores EST, logrando identificar 118 nuevos marcadores en la población DG.

2.8.4. Polimorfismo de un único nucleótido (*Single Nucleotide Polymorphisms* “SNP”)

Entre los muchos tipos de mutaciones que ocurren naturalmente en el genoma, los cambios en un único nucleótido (“mutación puntual”, *Single Nucleotide Polymorphisms* SNP), se destacan por su presencia en enormes cantidades en el genoma, por su relativa baja tasa de mutación (en comparación con los microsatélites de todo tipo), y su detección relativamente sencilla. En esencia, el término SNP describe la posición de un par de bases en el genoma de dos o más individuos en los cuales diferentes secuencias alternativas (alelos) existen en una población. Como un umbral, el alelo menos frecuente podría tener una abundancia del 1% o más. Los SNPs son las secuencias de variación más prevalentes entre los individuos de todas las especies y son una consecuencia de eventos de transición o transversión (Kahl *et al.*, 2005):

Individuo A 5'-ATGGATCGCCA-3'
Individuo B 5'-ATGGACCGCCA-3'

SNP transicional

Individuo A 5'-CGTTACAGGTT-3'
Individuo B 5'-CGTTCCAGGTT-3'

SNP transversional

Las mutaciones por transición, son aquellas en las que el cambio en un nucleótido se da de una base púrica por otra púrica (A↔G) o de una pirimidínica por otra pirimidínica (C↔T). En las transversiones, el cambio se da de una base púrica por una pirimidínica o viceversa.

Los SNPs pueden tener una variedad de propósitos, incluyendo la identificación rápida de variedades de cultivos, la construcción de mapas genéticos ultra-saturados y estudios de asociación relacionados con desórdenes genéticos (Douabin-Gicquel *et al.*, 2001).

Generalmente, los SNPs son abundantes, pero su densidad difiere sustancialmente en distintas regiones del genoma y entre genomas de cualquier

especie. Por ejemplo, el número promedio de SNPs en el genoma humano es de dos a tres por kb. Sin embargo, en regiones no codificantes, la densidad de SNPs (frecuencia de SNPs por unidad de longitud de ADN genómico, usualmente expresado como N° de SNPs por 100kb) aumenta drásticamente (Horton *et al.* 1998). El total de SNPs contenido en plantas es desconocido y sólo en unos pocos casos han sido rigurosamente buscados, pesar de un alto nivel de polimorfismo de la secuencia en muchas plantas cultivadas (e.g. maíz). En trigo por ejemplo, la densidad de SNPs es de dos por cada kb, pero podría ser mucho más alta en genes que codifican las enzimas de la biosíntesis del almidón (Bhatramakki y Rafalski 2001).

Los SNPs se dividen en varias clases, dependiendo de (1) su ubicación precisa en el genoma y (2) el impacto de su presencia dentro de la codificación o regiones reguladoras en la proteína codificada o fenotipo. Dado que la mayoría del ADN del genoma de organismos eucariotas es no-codificante y principalmente es ADN repetitivo, la mayoría de los SNPs están en esta categoría. Ellos son llamados SNPs no codificantes (ncSNPs). Entre estos ncSNPs, muchos se encuentran dentro de los intrones (SNP intrónicos). Pocos SNPs pueden ser encontrados en regiones codificantes ("SNP codificantes"), como los exones (SNP exónicos, o también SNPs basados en genes). Los SNPs exónicos también aparecen en el ADNc y son llamados SNPs cDNA, o cSNPs (Kahl *et al.*, 2005).

La presencia de un SNP en un exón de un gen que pueda tener un impacto en la función de la proteína codificada es llamado como SNP candidato. Los SNPs que se encuentran en promotores y otras regiones reguladoras del genoma son llamados SNPs de regulación y promotor respectivamente (pSNPs). Es obvio que cualquier SNP promotor puede influenciar dramáticamente la actividad del gen promotor. Si, por ejemplo, un pSNP impide la unión de un factor de transcripción a su secuencia de reconocimiento en el promotor, este se convierte en una parte disfuncional. Por el contrario, los SNPs intrónicos son considerados como inertes. Sin embargo, muchos investigadores emplean este tipo de SNPs para estudios de

asociación y mapeo de desequilibrio de ligamiento de todo el genoma (Kahl *et al.*, 2005)

La aparición de un SNP en un exón puede ser totalmente neutral (si este no cambia la composición de aminoácidos del dominio o de la proteína codificada, por lo que no tiene efecto sobre su función). En estos casos, el SNP es denominado como SNP sinónimo (synSNP). Por el contrario, un SNP no sinónimo (nsSNP) puede cambiar la cadena de aminoácidos. Los nsSNPs, puede provocar la síntesis de una proteína no funcional y por lo tanto estar involucrados en enfermedades (SNPs de diagnóstico) (Kahl *et al.*, 2005).

Todo lo anterior nos proporciona una idea más global y completa a cerca de las implicaciones e importancia del mapeo genético para la ubicación de regiones de interés.

3. OBJETIVOS

3.1. Objetivo general

Saturar el mapa genético de la población Mesoamericana intra-acervo de frijol DOR364 x BAT477 con marcadores moleculares tipo SSR y SSCP.

3.2. Objetivos específicos

- Evaluar el nivel de polimorfismo de los marcadores moleculares (SSR y SSCP) en la prueba de padres (catastro), para generar un primer acercamiento en la obtención de marcadores aptos para ser empleados en la saturación del mapa genético DOR364 x BAT477 (DB).
- Diseñar y evaluar marcadores basados en genes de nodulación.
- Ubicar marcadores basados en genes en el mapa de la población inter-acervo DOR364 x G19833 (DG), mediante la técnica SSCP.
- Comparar entre los mapas de ligamiento la ubicación de los marcadores basados en genes.
- Detectar SNPs entre los parentales de ambas poblaciones para posteriormente ser utilizados en técnicas de SNP alelo específicas.

4. MATERIALES Y MÉTODOS

4.1. Material Vegetal

4.1.1 Población DOR364 x BAT477

Se tomaron los dos genotipos DOR364 y BAT477, y su progenie de 90 líneas F_{5:7} endogámicas recombinantes (RILs) (la población será denominada en este documento como DB) (Figura 1). Las RILs fueron desarrolladas por descendencia de semilla única (SSD), para la generación de F₅, seguido por selección masal e incremento para pruebas de campo e invernadero hasta la generación F₇.(Remans *et al.*, 2007).

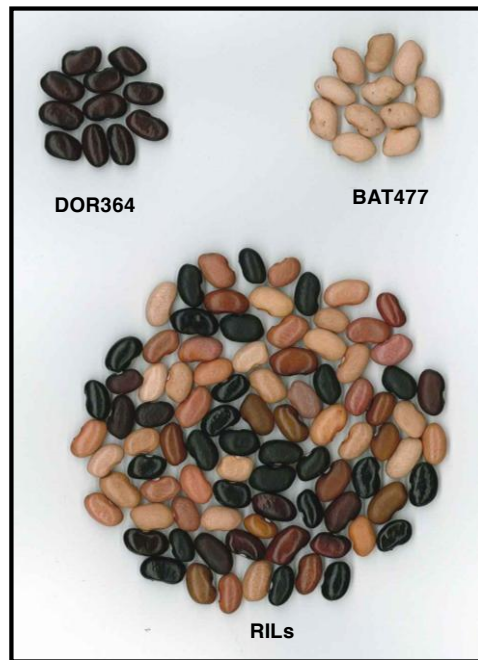


Figura 1. Semillas de las líneas parentales DOR364 y BAT477 y de las 90 RILs empleadas (una semilla por cada RIL).

Ambos parentales pertenecen a la raza Mesoamericana (M) del acervo genético Mesoamericano (Singh *et al.*, 1991). Al revisar la genealogía, estos genotipos comparten un padre a través de su historia, lo cual los hace estar más cerca a nivel genético (Figura 2). BAT477 es una línea avanzada del CIAT derivada de la cruce (G3834 x G4493) x (G4792 x G5694), que tiene hábito de crecimiento tipo III y pequeñas semillas de color crema. Además de su adaptación a la sequía, tiene una alta capacidad de fijación de nitrógeno (Lynch y White 1992). DOR364 fue seleccionado como el otro parental porque es un tipo comercial, presenta semillas rojas y pequeñas, además es una variedad mejorada original de CIAT que se cultiva ampliamente en América Central y en algunas partes del Caribe. DOR364 es resistente al virus del mosaico dorado (Beebe *et al.*, 1995), pero no es productivo en suelos pobres o infértiles (Blair *et al.*, 2006b). El árbol genealógico de DOR364 se basa en la cruce de BAT1215 x (RAB166 x DOR125), este genotipo tiene un hábito de crecimiento tipo II, posee buen crecimiento y rendimiento cuando el fósforo (P) no es limitante, pero es susceptible a la sequía. Ambos parentales no sólo difieren en la tolerancia a la sequía, sino también en la arquitectura de la planta y color de la semilla, esta última característica se utiliza como marcador para el éxito de la hibridación (Blair *et al.*, 2010a).

4.1.2 Población DOR364 x G19833

Para la generación de esta población, se emplearon como genotipos parentales DOR364 y G19833 (la población será denominada en este documento como DG) (Figura 3). La descripción del parental DOR364 se encuentra en la población anterior. Por otro lado, G19833 es una variedad Peruana, llamada Chaucha Chuga, perteneciente a la raza Perú (Blair *et al.*, 2006b) del pool genético Andino. Tiene un hábito de crecimiento tipo II y presenta una semilla rojo-amarilla moteada, tiene una alta calidad nutricional (Blair *et al.*, 2009a) y gran capacidad de adaptación a condiciones de estrés por deficiencia de fósforo. Comercialmente, este genotipo presenta gran tamaño (peso de 100 semillas: 46 g).

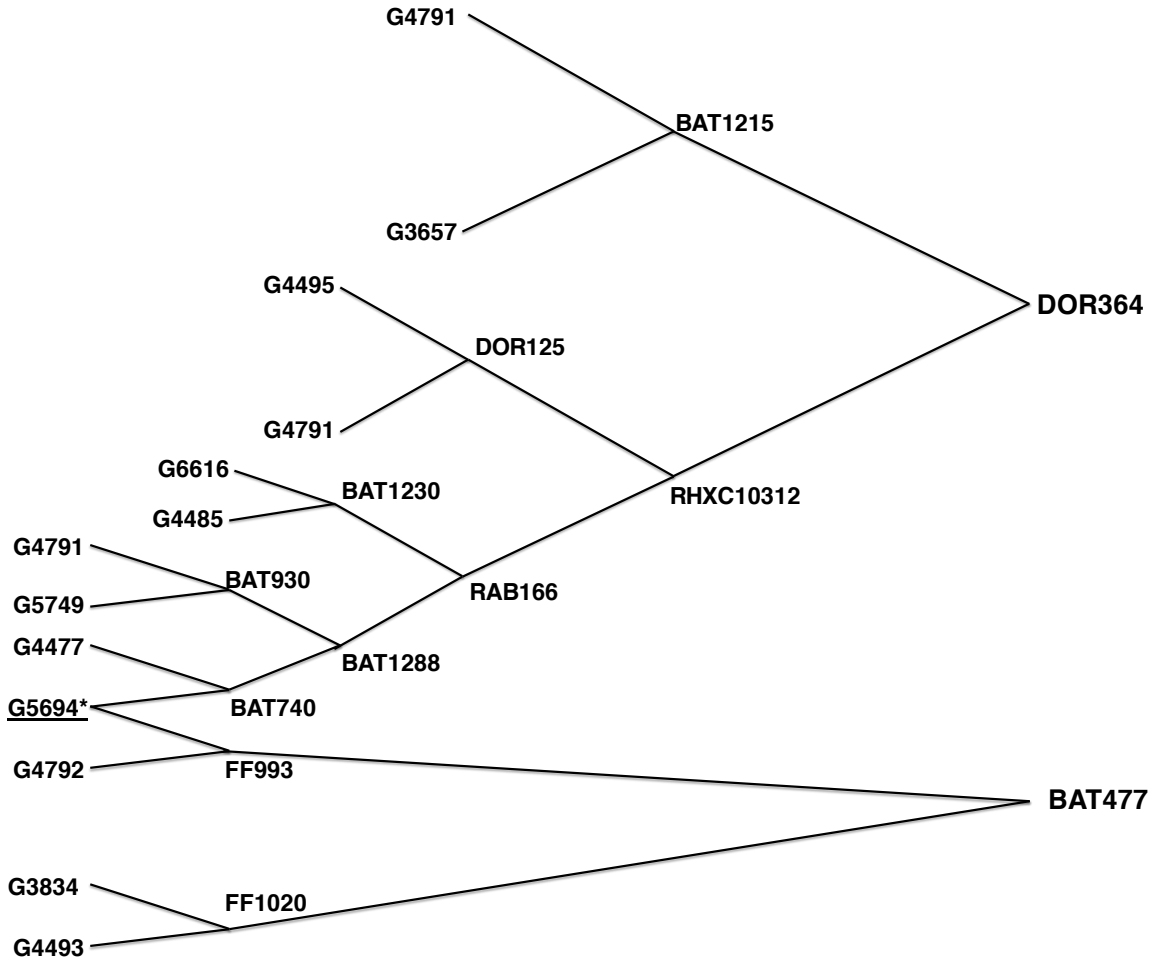


Figura 2. Árbol genealógico de los genotipos DOR364 y BAT477. (El genotipo subrayado hace referencia al padre que tienen en común)

La población consta de 87 individuos $F_{7:11}$ y fue desarrollada a partir del método SSD en el Centro Internacional de Agricultura Tropical (CIAT). Esta población, ha sido tomada como consenso para diferentes proyectos de investigación en frijol y por ende ha sido ampliamente estudiada en CIAT.



Figura 3. Semillas de las líneas parentales DOR364 y G19833.

4.2. Extracción de ADN

El ADN de los parentales y de la descendencia fue extraído del tejido de hojas jóvenes usando el método de Dellaporta *et al.*, (1983) y Vallejos (1992) (Anexo 1). Posterior a la extracción, la calidad del ADN de cada una de las muestras se estimó por medio de una electroforesis en gel de agarosa al 1%, el cual fue teñido con el colorante Gel red (Biotium) y visualizado con luz ultravioleta. La cuantificación de la concentración de ADN se llevó a cabo con el software Quantity One® v 4.0.3. Este programa calcula la cantidad de ADN en cada una de las bandas obtenidas en gel de agarosa de acuerdo a las extrapolaciones que realiza con los patrones de concentración de ADN lambda de 100, 200, 300 y 400ng/ μ l. Se realizó dilución de ADN a 5ng/ μ l para la amplificación en PCR.

4.3. Amplificación de microsatélites y SSCPs

Se realizó una evaluación preliminar de 970 marcadores en los dos parentales (DB), incluyendo 555 microsatélites genómicos (BMb) Córdoba *et al.*, (2010a y 2010b), 100 microsatélites génicos (PvM) (Hanai *et al.*, 2009), 111 marcadores basados en genes (Leg) (Hougaard *et al.*, 2008) y otros 204 de este mismo tipo diseñados en este estudio.

Todas las reacciones de PCR se realizaron en un volumen final de 15µL con 20ng/µL de ADN genómico, 0.2µM de los primers sentido (*forward*) y antisentido (*reverse*), 2.0mM MgCl₂, 0.2mM de dNTP's totales y 0.2µL de *Taq* polimerasa por cada reacción. Las condiciones de amplificación iniciaron con un "touch-down" durante los primeros 10 ciclos. La temperatura de inicio fue de 94 °C durante 3 minutos (min); seguido por 30 segundos (seg) de denaturación a 94 °C durante 40 seg. La temperatura de hibridación dependió del T_m (temperatura de melting) de los primers, con una disminución de 1 °C por ciclo y la extensión a 72 °C durante 50 seg. por 10 ciclos. A continuación, se denaturó nuevamente a 94 °C durante 30 seg, seguido de 40 seg de alineamiento a una temperatura de hibridación dependiendo del T_m de los primers. Esta segunda parte tuvo una duración de 27 ciclos, para un total de 35 ciclos. Finalmente, se realizó una extensión a 72 °C durante 5 min.

Luego de realizar la amplificación a cada una de las muestras se le adicionó 5µl de formamida compuesta por 0.4% de azul de bromofenol y 0.25% p/v de xylene cyanol y se procedió a denaturarlas a 96 °C durante 5 min. Para los marcadores tipo microsatélite, las muestras fueron cargadas en gel de acrilamida al 4% (29:1 acrilamida:bisacrilamida) que contenía 5M úrea y TBE 0.5x. Los geles se corrieron en un cámara T-Rex™ Aluminium Backed Sequencer (OWL separation systems, Portsmouth, NH, UK) aproximadamente a 110 Watts y una temperatura de 50 °C. Para la visualización del gel y la detección de los productos de PCR se empleó la metodología de tinción con nitrato de plata, según lo reportado por Blair *et al.*, (2003) (Anexo 2). Para estimar el tamaño de los alelos, se sembró junto a las muestras en cada gel, 3µl de marcador de peso de 10 y 25 pb. La caracterización molecular de los parentales permitió establecer cuáles marcadores eran polimórficos, además de medir el tamaño de los alelos observados para cada SSR. En cuanto a las condiciones de amplificación, electroforesis y tinción para los SSCP, se siguió la metodología empleada por Galeano *et al.*, (2009a).

En los padres de la población DG únicamente fueron evaluados los marcadores Leg y los diseñados a partir de genes de nodulación (BSn).

4.4. Diseño de marcadores basados en genes de nodulación

A partir de la revisión de literatura y conociendo cuales son algunos de los genes involucrados durante el proceso de nodulación, se lograron diseñar 204 primers con la ayuda del Programa *Geneious Pro* v5.3.4 (Drummond *et al.*, 2010), flanqueando intrones o regiones variables para evaluar SSCP en los parentales. Posterior a esto, se realizó la evaluación de los marcadores que fueron determinados como polimórficos sobre la población de estudio. Los marcadores aquí diseñados recibieron el nombre de BSn (*Bean SNP nodulation*).

4.5. Mapeo y análisis de los marcadores

Los marcadores que se determinaron como polimórficos entre los parentales de ambas poblaciones (DB y DG) fueron evaluados en todas las RILs de las respectivas poblaciones de mapeo. A partir de los datos obtenidos se analizó la segregación presentada por cada uno de los marcadores, utilizando la prueba estadística de chi-cuadrado (χ^2), según las proporciones esperadas (1:1) para determinar distorsiones en la segregación o predominio de alguno de los alelos de los padres.

La información de la segregación de los marcadores fue empleada para establecer su ubicación relativa en el mapa genético ya establecido para cada una de las poblaciones: El mapa de DB que fue descrito por Blair *et al.* (2010a), el cual incluye marcadores AFLP, RAPD y SSR, obteniendo una longitud total de 1087,5cM en 11 grupos de ligamiento, cubriendo el genoma entero con una saturación de un marcador cada 5,9cM, y el mapa de DG, descrito por Galeano *et*

al. (2009a), que posee una longitud total de 1,900cM y una distancia promedio entre marcadores de 6,8cM.

Para el mapeo se empleó el software MapDisto v. 1.7 beta (Lorieux, 2008). Las distancias genéticas se reportaron en centiMorgans (cM) estimadas con la función de Kosambi. El umbral mínimo para considerar un marcador como ligado fue de un LOD de 3.0. Los marcadores con un $LOD < 3.0$ o asignados a grupos de ligamiento ambiguos se excluyeron del estudio. LOD es el $\text{Log}_{10}(L_1/L_0)$, donde L_0 representa la probabilidad de no ligamiento y L_1 representa la hipótesis alternativa de ligamiento.

Los comandos “order sequence”, “check inversions” y “ripple order” del programa MapDisto v1.7 beta fueron empleados para identificar el orden más probable de los marcadores dentro de los grupos de ligamiento. Estas opciones fueron empleadas luego de adicionar cada marcador al mapa.

4.6 Secuenciación y SNPs

Se amplificaron 204 primers (BSn) diseñados para nodulación (Anexo 3), en los genotipos DOR364, BAT477 y G19833. Las reacciones de PCR fueron llevadas a cabo en un volumen final de 30 μ L para cada uno de los parentales, manteniendo las condiciones y concentraciones para la realización de PCR mencionados anteriormente. Posterior a esto, se alicuotaron 5 μ L de cada producto de amplificación, el cual fue mezclado con igual cantidad de BlueJuice y GelRed (Biotium) para determinar la calidad del producto de amplificado. Las muestras fueron evaluadas en gel de agarosa a una concentración de 1.8%. En los extremos del gel se sembraron 7 μ L de marcador de peso de 1kb Plus (BRL Markers & Standards Ladder). Las condiciones de electroforesis fueron 100 voltios durante 30 min. Luego, las bandas se visualizaron con luz U.V. Una vez corroborada la presencia y el peso del producto de PCR, las muestras fueron

preparadas según las exigencias de la *Universidad de Yale*, para llevar a cabo el proceso de secuenciación (Anexo 4).

La secuenciación de las muestras en la *Universidad de Yale*, fue realizada por medio del equipo *Applied Biosystems (ABI) 3730*. Para las reacciones de secuenciación, se utilizaron dideoxinucleótidos (ddNTP) (se diferencian de los deoxinucleótidos (dNTP) en que carecen del grupo OH en el carbono 3' de la ribosa) marcados con fluorescencia (Big Dye Terminators) y Taq polimerasa FS (fluorescente sequencing), en un protocolo de ciclos térmicos. Empleando una buena cantidad de ADN plantilla y de primer, además de Taq/ dye-terminator, el ciclo de secuenciación puede generar entre 500 y 600 bases de secuencia con una precisión del 98-99%. (Tomado de <http://medicine.yale.edu/keck/dna/protocols/index.aspx>).

Después de obtener los resultados, estos fueron editados, eliminando los extremos de las secuencias que no presentaban buena calidad respecto a la diferenciación de los picos mostrados en el electroferograma. El paso a seguir fue alinear las tres secuencias por medio del programa *Geneious Pro v5.3.4* (Drummond *et al.*, 2010), para así poder encontrar los SNP presentes entre los genotipos. Luego de detectar los SNPs, estos fueron discriminados en transversiones y transiciones para cada combinación de padres (DB y DG).

5. RESULTADOS Y DISCUSIÓN

5.1. Evaluación de marcadores en los parentales

De los 970 marcadores evaluados en los padres DOR364 y BAT477 (Tabla 2), 152 resultaron polimórficos, 628 monomórficos y 190 se descartaron debido a que presentaron altas inespecificidades o no mostraron producto de amplificación a pesar de cambiar condiciones de PCR tales como concentración de $MgCl_2$ y temperaturas de alineamiento.

En el caso de los padres DOR364 y G19833, se evaluaron 315 marcadores, dentro de los cuales hubo 130 monomórficos, 58 polimórficos, y 127 presentaron problemas de amplificación.

La falta de amplificación de un alelo en las líneas de fríjol puede ser el resultado de mutaciones en el sitio de unión al primer que impiden que este reconozca la secuencia y por lo tanto se una a ella, creando un alelo nulo. Sin embargo, esto podría también resultar de la producción de una cantidad indetectable de producto de PCR, por condiciones de amplificación inadecuadas para el par de primers (Lavi *et al.*, 1994).

Debido a que un alelo nulo es tomado como un dato perdido durante la evaluación de los marcadores en la población, su presencia puede interferir en la interpretación de la segregación debido a que ni los alelos nulos ni los genotipos heterocigotos son considerados en el mapeo pudiendo causar una desviación en la segregación mendeliana esperada (Gupta & Varshney, 2000; Varshney *et al.*, 2005).

Tabla 2. Resumen de todos los marcadores evaluados en los padres DOR364 y BAT477. (No se muestran los AFLP ni RAPD).

FUENTE	NOMBRE DE PRIMER	TIPO DE MARCADOR	GÉNICO/ GENÓMICO	N° PRIMERS EVALUADOS
Córdoba <i>et al.</i> , 2010a y b	BMb/ BMbw	SSR-BES	Genómico	555*
Hanai <i>et al.</i> , 2009	PvM	SSR-EST	Génico	100*
Hougaard <i>et al.</i> , 2008	Leg	SSR-EST	Génico	111*
Franco <i>et al</i> (sin publicar)	BSn	Basados en genes	Génico	204*
Hurtado (sin publicar)	BMe	SSR-EST	Génico	106
Blair (sin publicar)	Bmd	SSR	Génico	128
Guerra-Sanz 2004	Genebank number	SSR	Génico	20
de Campos <i>et al.</i> , 2007	FJ	SSR	Genómico	20
Caixeta <i>et al.</i> , 2005	PV	SSR	Genómico	21
Yaish and Perez, 2003	Genebank number	SSR	Genómico	21
Yu <i>et al.</i> , 2000	PV	SSR	Génico	37
Blair <i>et al.</i> , 2009b	BM	SSR	Genómico	85
Blair <i>et al.</i> , 2009c	BMc	SSR	Génico	258
Blair <i>et al.</i> , 2003	BMd	SSR	Génico	57
Hanai <i>et al.</i> , 2007	PVM, FJ	SSR	Génico/Genómico	80
Benchimol <i>et al.</i> , 2007	IAC	SSR	Genómico	123
Grisi <i>et al.</i> , 2007	PVBR	SSR	Genómico	98
Gaitán-Solís <i>et al.</i> , 2002	BM, GATS, AG	SSR	Genómico	68
Buso <i>et al.</i> , 2006	PVBR	SSR	Genómico	20
Galeano <i>et al.</i> 2009a	BSNP, g	Basados en genes	Génico	393
Blair <i>et al.</i> , 2008	BMa, HRG, ME, PG2, GA, CAC	SSR	Genómico	85
TOTAL				2590

* Marcadores evaluados en este trabajo

En la evaluación parental realizada entre los genotipos DOR364 y BAT477, el porcentaje de polimorfismo más alto fue exhibido por los marcadores genómicos (22,5%), Córdoba *et al.* (2010b) quienes emplearon los mismos marcadores genómicos (BMb), encontraron que el 65% de estos fueron polimórficos para los padres DOR364 y G19833, debido a las diferencias genéticas existentes entre los genotipos empleados. Los marcadores génicos mostraron los porcentajes más

bajos de polimorfismo (entre 3% y 9,3%), de igual forma, en este tipo de marcadores se obtuvo el porcentaje más alto de fragmentos no amplificados (45,1%) y monomórficos (86%). En los parentales DOR364 y G19833, solo se emplearon marcadores génicos obteniendo bajos niveles de polimorfismo (18%). Galeano *et al.* (2009a), también emplearon marcadores basados en genes y encontraron tasas de polimorfismo del 25,6 y 22% entre los padres DOR364 y G19833, y BAT93 x JaloEEP558 respectivamente, datos que no son muy distantes a lo encontrado en esta investigación. Blair *et al.* (2003), utilizando los mismos padres (DOR364 y G19833), reportaron un 64.3% de polimorfismo en los microsatélites genómicos y un 46.3% para los marcadores génicos. Medina, (2006), también halló el mismo comportamiento de los marcadores en las diferentes poblaciones empleadas, obteniendo de manera general para los marcadores de origen génico un 35% de polimorfismo, mientras que para los genómicos el polimorfismo fue 57%. Galeano *et al.* (2009b), encontró que el 59% de los marcadores EST evaluados en los padres (DOR364 y G19833) fueron polimórficos (21/37), siendo este el valor más alto encontrado para marcadores de origen génico. Sin embargo, el porcentaje de marcadores génicos polimórficos encontrados en este trabajo entre los padres DOR364 y G19833 es 2 veces mayor (37%) respecto al mostrado por los parentales DOR364 y BAT477 (16,8%) (Tabla 3).

Tabla 3. Porcentajes de los marcadores amplificados en la prueba de padres.

Padres evaluados	Tipo de Marcador	Nombre del Marcador	Tipo de marcador	POLIMÓRFICO (%)	MONOMÓRFICO (%)	NO AMPLIFICÓ (%)
DOR364 y BAT477	Genómicos	BMb	SSR-BES	22,5	68,1	9,4
		PvM	SSR-EST	3	86	11
	Génicos	Leg	SSR-EST	4,5	63,9	31,6
		BSn	Basados en genes	9,3	45,6	45,1
DOR364 y G19833	Génicos	Leg	SSR-EST	18,9	49,6	31,5
		BSn	Basados en genes	18,1	36,8	45,1

El porcentaje de polimorfismo es mayor en los marcadores genómicos, debido a que provienen de regiones no codificantes del genoma (como intrones o espacios intergénicos, regiones teloméricas y centroméricas), mientras que los génicos provienen de regiones codificantes (Blair *et al.*, 2003), que en muchas especies son secuencias altamente conservadas. (Varshney *et al.*, 2005). Los SSR génicos y genómicos tienden a ser complementarios para el mapeo genético (Varshney *et al.*, 2005).

Los resultados obtenidos en las dos combinaciones de padres, indican claramente que el número de marcadores monomórficos es mucho más elevado para los padres DOR364 y BAT477, lo cual está indicando que existe una cercanía genética muy estrecha que hace que no se presente mucha variación en el genoma y que por el contrario se encuentren muchas regiones conservadas.

5.2. Marcadores polimórficos y mapeo en la población

De los 152 marcadores tanto SSR como SSCP determinados como polimórficos en los parentales DOR364 y BAT477, solo 49 (44 genómicos y cinco génicos) pudieron ser anclados al mapa, 89 marcadores, al ser evaluados en la población fueron monomórficos y 14 al ser mapeados obtuvieron un LOD menor a tres y mapeaban en varios grupos de ligamiento distintos. (En la Figura 4, se muestran ejemplos de marcadores BMb y PvM mapeados en la población DB).

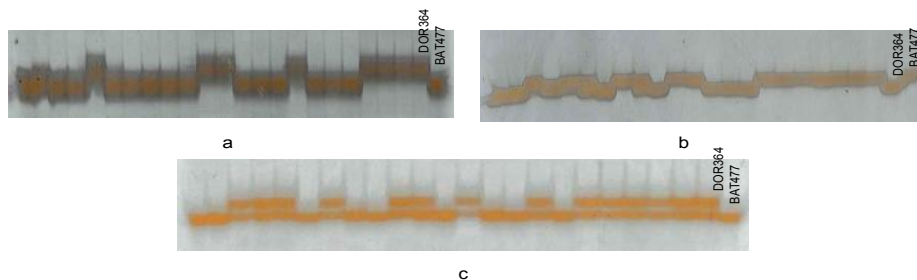


Figura 4. Patrones de segregación de tres marcadores mapeados en la población DB. a) BMb737, b) PvM127, c) BMb1939. Las primeras dos últimas líneas muestran los alelos de los padres.

Al mapa obtenido de la cruce de los padres DOR364 y G19833, se lograron anclar 32 marcadores (19 Leg y 13 BSn), y los 26 restantes al ser corridos en la población se presentaron como monomórficos. Para esta cruce no hubo presencia de marcadores con LOD bajo (menor a 3) ni con ambigüedades al momento de ser mapeados.

5.2.1 Segregación de los marcadores

5.2.1.1 Población DOR364 x BAT477

Luego de realizar la prueba χ^2 ($p < 0,05$), para los marcadores de la población proveniente de la cruce DB, se encontró que el 84% (41/49) de los marcadores presentaba una segregación 1:1; mientras el 12% (6/49) estuvieron sesgados hacia el parental DOR364 y el sesgo hacia el parental BAT477 fue de 4% (2/49), es decir que el 16% (8/49) de los marcadores no presentó la segregación esperada (Figura 5).

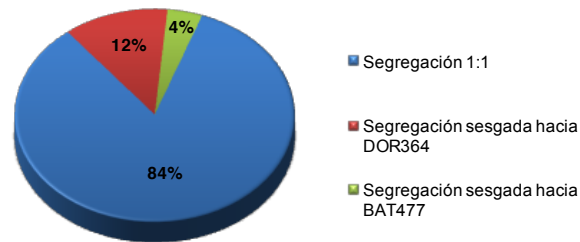


Figura 5. Distribución de la segregación de los marcadores evaluados en la población proveniente de cruce DOR364 x BAT477.

Los GL B04 y B05 fueron los que presentaron más marcadores distorsionados (2), mientras que 5 grupos (B03, B06, B07, B08 y B10) mostraron únicamente marcadores con la segregación esperada. De acuerdo con la prueba χ^2 , el locus más distorsionado (BMb643 p : 0,001) se ubicó en el GL B05, mientras que el menos distorsionado (BMb653 p : 0,036) fue encontrado en el GL B11 (Anexo 5).

5.2.1.2 Población DOR364 x G19833

La prueba χ^2 para los marcadores anclados al mapa de la población DG, muestra que el 60% de los marcadores (19/32) presenta la segregación esperada para este tipo de población (1:1), mientras que el 40% (13/32) restante exhibe una distorsión en la segregación (Figura 6). Se encontraron tres marcadores sesgados hacia el parental DOR364 (9,3%) y diez hacia el padre G19833 (31,2%). Los alelos Andinos fueron predominantes en los cromosomas B02 y B08 (con tres cada uno), lo cual es concordante con los datos encontrados en varios estudios (Blair *et al.*, 2003., Hurtado 2009; Galeano *et al.*, 2009a; Blair *et al.*, 2010c; Córdoba *et al.*, 2010a) donde reportan al GL B02 como el que presenta mayor número de marcadores distorsionados hacia ese mismo padre.

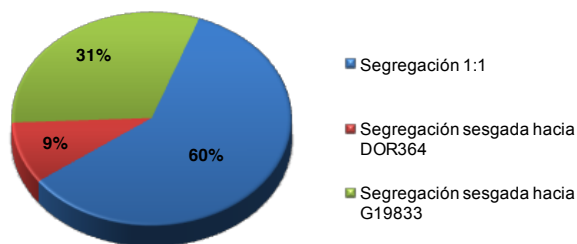


Figura 6. Distribución de la segregación de los marcadores evaluados en la población proveniente de cruce DOR364 x G19833.

Lo anterior indica que existe una prevalencia de distorsión de la segregación en algunos grupos de ligamiento para la cruce DG, posiblemente debido a que los factores genéticos responsables de ella, se ubican en regiones cromosómicas que se mantienen en la población (Lu *et al.*, 2002) y que son responsables de la eliminación selectiva de alelos en respuesta a diferentes condiciones ambientales en las que se cultiva la planta (Xu *et al.*, 1997).

Rodrigues & Santos (2006), también obtuvieron una segregación sesgada hacia uno de los padres, y señalan que todos los *loci* fueron afectados por la selección

natural, que favoreció a uno de los dos parentales, pues encontraron que 29 de los 30 *loci* evaluados eran favorecidos hacia el genotipo *Carioca MG*.

La mayoría de los marcadores distorsionados se ubican en los GL B02, B08 y B03 con tres y cuatro marcadores respectivamente, mientras que los grupos B01, B07 y B10 solo mostraron un marcador sesgado cada uno. De manera particular, en el GL B03, se encuentran ubicados los *loci* para el marcador más distorsionado que fue el Leg213 (p: 0,002), y el que presentó menor distorsión: Leg237 (p: 0,04) (Anexo 6). El agrupamiento de los marcadores con segregación distorsionada, posiblemente refleja la localización de genes que fueron favorecidos por la selección meiótica o selección involuntaria de la población (Hanai *et al.*, 2009).

En plantas, el porcentaje, grado, origen y efectos genéticos de la distorsión de la segregación varía significativamente entre especies, tipos de poblaciones, cruces y tipo de marcadores. Poblaciones como DH y RILs, generalmente presentan una alta distorsión en la segregación, contrario a lo que sucede con poblaciones F₂ (Semagn *et al.*, 2006). La distorsión observada en las poblaciones RILs derivadas por descendencia de semilla única representan un efecto de los factores genéticos y ambientales acumulados por múltiples generaciones, incrementándose esta interacción con cada generación (Torjék *et al.*, 2006).

Sin embargo, el porcentaje de distorsión hallado para la población DB, fue intermedio (16%) respecto a lo encontrado por Blair *et al.* (2003), quienes mostraron que el 23 y el 9% de los marcadores estaban sesgados en las poblaciones RILs DG y BJ respectivamente, pero al comparar estos resultados, con los obtenidos para la población DG, se observa que el porcentaje de marcadores distorsionados encontrado en este estudio es mayor (40%).

Los patrones de distorsión en la segregación, probablemente reflejan la ubicación de genes de incompatibilidad o adaptación que influyen la selección en la población (Beebe *et al.*, 1998). Estos patrones también han sido asociados a

través de múltiples cruces con genes de esterilidad (Xu *et al.*, 1997). La segregación distorsionada de los marcadores no los excluye de ser potencialmente informativos para la detección de recombinantes. Por esta razón, se incorporaron al análisis. Según Shizhong (2008), el uso de estos marcadores puede incrementar la saturación de mapeo y el poder estadístico para mapear QTLs.

El fenómeno de marcadores con segregación distorsionada en frijol común ha sido descrito por varios autores (Freyre *et al.*, 1998; Blair *et al.*, 2003; Ochoa *et al.*, 2006; Grisi *et al.*, 2007; Hougaard *et al.*, 2008; Galeano *et al.*, 2009a; Hanai *et al.*, 2009; Blair *et al.*, 2010c; García *et al.*, 2011). Esto fue observado no solo para marcadores microsatélites, sino también para marcadores basados en genes y RFLP en leguminosas (Hougaard *et al.*, 2008; Freyre *et al.*, 1998). La razón para esta distorsión puede estar asociada a regiones genómicas relacionadas con las ventajas adaptativas que fueron seleccionadas durante cada ciclo de generación en el desarrollo de la población de mapeo (de Campos *et al.*, 2011).

A nivel molecular, este fenómeno es pobremente entendido, pero se sabe que la distorsión puede ser causada por una variedad de factores ambientales, fisiológicos y genéticos (Torjék *et al.*, 2006). Entre ellos está la incompatibilidad entre polen-pistilo, la competencia gamética (Lu *et al.*, 2002), las interacciones epistáticas negativas (Li *et al.*, 1997) y el aborto de gametos (Sano, 1990).

El porcentaje de polimorfismo obtenido a partir de los nuevos marcadores para la población procedente de la cruce DB fue del 19,5%, sin embargo, Blair *et al.* (2010c) realizaron un trabajo con una población derivada de un cruce intra-acervo de dos genotipos mesoamericanos, obteniendo un porcentaje de polimorfismo aún más bajo de 11,9%. Chavarro (2006), encontró un porcentaje de polimorfismo del 23% en su población de mapeo intra-acervo Andino derivado de la cruce G122 x Montcalm. Otro trabajo realizado con una población del acervo Andino fue el desarrollado por Cichy *et al.*, (2009), en el cual, el porcentaje de polimorfismo fue

mayor respecto a los datos anteriores (31%). Estos resultados se ajustan a lo indicado por Blair *et al.* (2006b), quienes proponen que la diversidad genética intra-acervo es más alta en cruces Andinos que en los Mesoamericanos con ambos tipos de marcadores: microsatélites genómicos y basados en genes. No obstante existen algunas excepciones, pues Blair *et al.* (2010b) emplearon una población RIL a partir de un cruce intra-acervo (Andino x Andino), pero tan solo encontraron una tasa de polimorfismo del 10,6% empleando marcadores SSR. En contraste, Frei *et al.* (2005) desarrollaron un trabajo con un cruce intra-acervo entre dos genotipos Mesoamericanos, donde hubo una tasa de polimorfismo del 30,5% utilizando marcadores microsatélites.

La baja tasa de polimorfismo registrada en este trabajo, puede ser explicada porque los genotipos parentales de la cruce pertenecen al mismo pool genético y además a la misma raza (Mesoamericana), lo cual indica que están estrechamente relacionados, y se sabe que los cruces inter-acervos presentan mayor porcentaje de polimorfismo que cruces intra-acervo, ya que el polimorfismo está representando la variabilidad existente entre estos donde definitivamente hay una base genética más amplia (Blair *et al.*, 2003). Por lo tanto, se espera que el porcentaje de polimorfismo sea mayor cuando se comparan cruces entre genotipos de acervos o razas distintas, dado que comparten menos alelos, como en el cruce DG (Mesoamericano x Andino), donde se encontró que el porcentaje de polimorfismo fue del 30,8%.

En cuanto a los marcadores basados en genes, Hougaard *et al.* (2008), lograron anclar al mapa genético de frijol (entre los mapas de las poblaciones DG y BJ) 104 de los 111 marcadores basados en genes (Leg), mientras que en este trabajo tan solo pudieron ser mapeados 19 en DG y dos en DB empleando estos mismos marcadores. Galeano *et al.* (2009a), anclaron al mapa el 28% (118/418) de los marcadores basados en genes probados en la población DG y en esta investigación se lograron anclar únicamente el 10 y el 1,2% de los marcadores basados en genes en los mapas de las poblaciones DG y DB respectivamente.

En síntesis, los porcentajes de polimorfismo hallados en este estudio son correspondientes con lo planteado por Blair *et al.* (2006b), donde proponen que las combinaciones de padres inter-acervo (Mesoamericano x Andino) tienen una tasa de polimorfismo más alta que las combinaciones intra-acervo (Mesoamericano x Mesoamericano y Andino x Andino).

5.3. Mapas de ligamiento

5.3.1 Población DOR364 x BAT477

Posterior al análisis de segregación, se obtuvo el mapa para la población DB, el cual presentó una longitud total de 1771,3cM. La longitud de este mapa es 683,8cM (38,6%) más largo que el publicado por Blair *et al.* (2010a), debido a que los nuevos marcadores fueron posicionados en los extremos de los GL, quedando aislados, generando así *gaps*, como en el caso del B11, donde la inclusión de los nuevos marcadores expandió el GL de 29,3cM (Blair *et al.*, 2010a) a 108,07cM. Sin embargo, es consistente con otros estudios de mapeo genético (Blair *et al.*, 2003; Grisi *et al.*, 2007; Blair *et al.*, 2010a; de Campos *et al.*, 2011).

La presencia de datos perdidos también influye en el cálculo de las distancias, lo cual se ve reflejado en los *gaps* que se generan a través de los grupos de ligamiento.

Dado que las distancias genéticas entre los marcadores fueron amplias respecto a lo que se esperaría de un mapa genético saturado, es por esto, que se hace necesaria la inclusión de más marcadores que saturen aquellas zonas que presentan una baja densidad.

El número total de marcadores nuevos anclados al mapa por GL varió de dos (grupos B01 y B06) a diez (grupo B02) (Figura 7). Los marcadores nuevos se

distribuyeron a lo largo de los once grupos de ligamiento, los cuales presentaron una longitud promedio de 161cM.

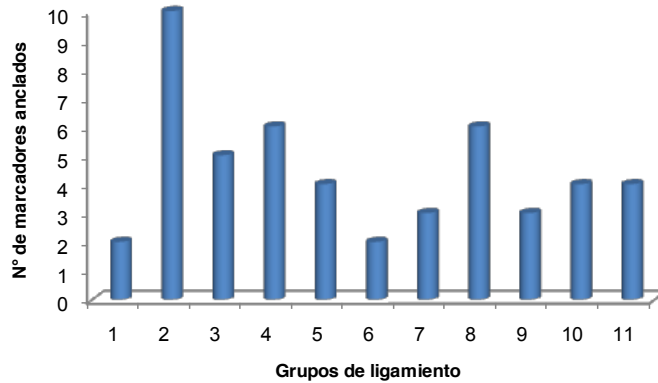


Figura 7. Número de nuevos marcadores ligados al mapa DOR364 x BAT477.

La longitud de los GL varió entre 75,9cM (para el B09) y 267,9cM (para el B04) (Figura 8). La distancia promedio entre marcadores reportada por Blair *et al.*, (2010a) es de 5,9cM, mientras que para este estudio se amplió a 7,2cM. Algo similar obtuvieron en el trabajo realizado por de Campos *et al.* (2011), donde lograron un mapa de 1865,9cM y una distancia promedio entre marcadores de 9,4cM. También presentaron grupos de ligamiento con distancias genéticas mayores a 37,5cM.

En promedio, se lograron ubicar 4,4 marcadores por GL. La mayor cobertura fue presentada en el grupo B02 con ocho marcadores nuevos, mientras que en los grupos B01 y B06 solo se anclaron dos marcadores nuevos en cada uno. En términos generales, fueron los marcadores genómicos SSR-BES (BMb) quienes estuvieron representados en todos los grupos de ligamiento, mientras que los basados en genes (BSn) y Leg solo se restringieron a dos grupos (Tabla 4). Blair *et al.* (2003), en su investigación encontraron que los microsatélites génicos presentaron una mejor distribución respecto a los microsatélites genómicos, caso contrario a lo observado en este trabajo, lo que muy posiblemente esté relacionado con el diseño aleatorio (en cuanto a su distribución en el genoma) de los marcadores genómicos aquí empleados.

El GL con el mayor número de marcadores (54) fue el B02, mientras que el grupo con el número menor (9) fue el B11 (Tabla 4). Estos resultados son parcialmente coherentes con los encontrados por Blair *et al.*, (2010a), quienes reportan también el grupo B02 con el mayor número de marcadores, pero en ese momento, este tan solo poseía 12 de ellos. Sin embargo, en esa misma publicación, ellos señalan el B08 como el grupo que exhibe el menor número de marcadores (8).

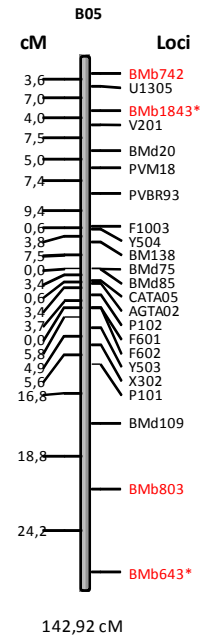
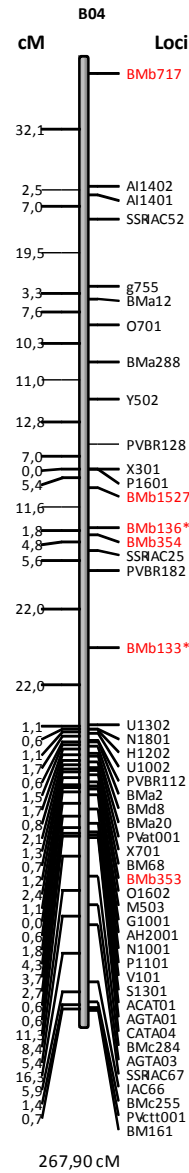
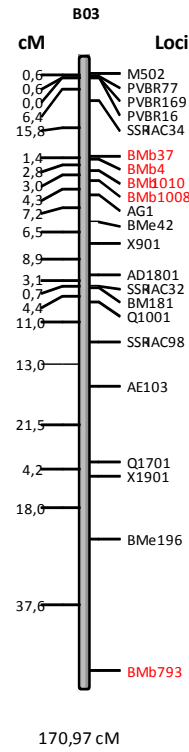
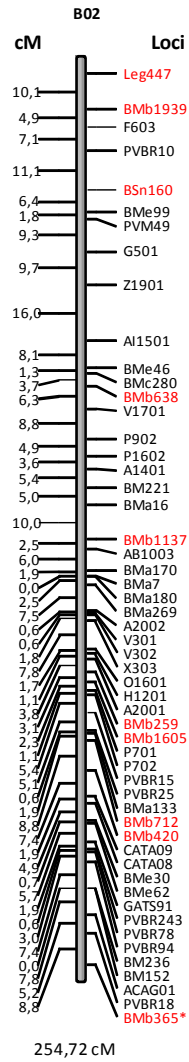
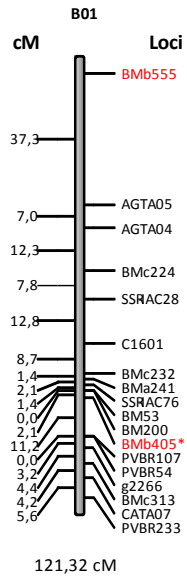
En términos generales, no se observó una relación directa entre el número de marcadores y la longitud de los GL. El GL B09 solo posee 9 marcadores y presenta una longitud de 108cM, mientras que el GL B06 tiene anclados 23 marcadores y registra una longitud menor (105,8cM). Esto puede ser explicado por los patrones de agrupamiento en algunas regiones del genoma de frijol (Hanai *et al.*, 2009).

Tabla 4. Resumen de los marcadores actualmente integrados en el mapa genético de la población DOR364 x BAT477.

Grupo de ligamiento (GL)	AFLP	RAPD	SSR_BES	SSR	SSR_EST	Basados en genes	Leg	TOTAL	Distancia promedio entre marcadores (cM)	Distancia total de cada GL(cM)
B01	3	1	2	11		1		18	6,0	121,3
B02	3	18	8	19	4	1*	1	54	3,3	254,7
B03		7	5	8	2			22	3,5	171
B04	4	19	6	18		1		48	3,6	267,9
B05	2	10	4	6				23	7,5	142,9
B06	1	7	2	13				23	4,6	105,8
B07	3	5	2	6	1	3	1	21	6,8	173,6
B08	3	10	5	9		1*		28	9,6	171,0
B09		3	3	6		1		13	13,2	75,9
B10	3	16	3	11	2			35	7,3	179,0
B11		2	4	3				9	13,5	108,0
TOTAL	22	98	44	111	9	8	2	294	7,2	1771,3

*Marcadores de nodulación

Al considerar todos los marcadores, algunos *gaps* mayores a 25 cM fueron encontrados en los grupos de ligamiento B01, B03, B04, B09 y B11. La distancia genética mayor, se presentó en los cromosomas B01, B03 y B11, entre los marcadores BMb555 y AGTA05; BMe196 y BMb793 al igual que entre el BMb654 y BM239 (37,3; 37,6 y 40,9 cM, respectivamente) (Figura 8).



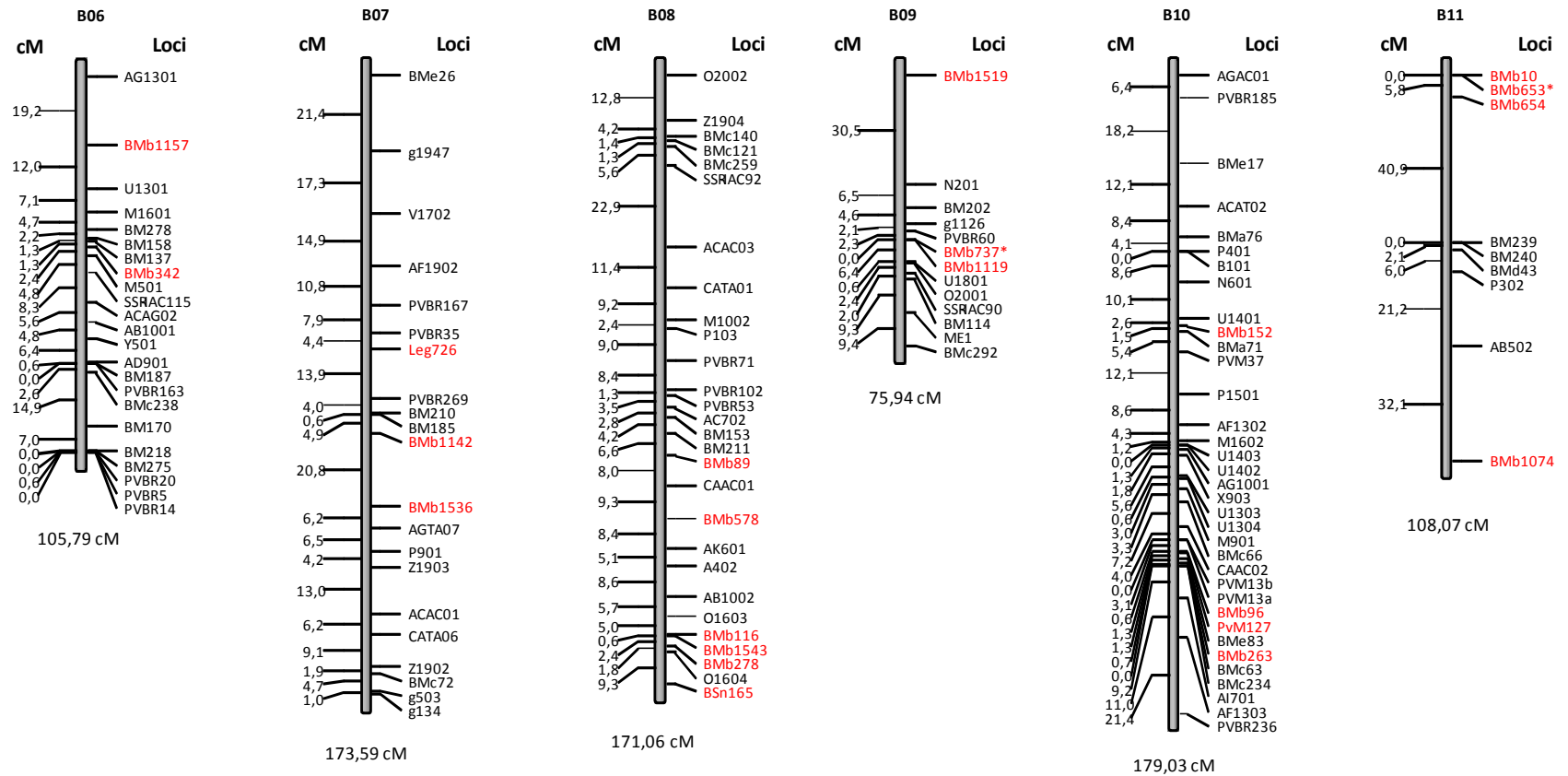


Figura 8. Mapa de ligamiento de la población derivada de la cruda intra-acervo DOR364 x BAT477 de frijol común. Cada barra vertical representa un GL. Los nombres de los marcadores se muestran al lado derecho de cada GL, las líneas conectoras indican la posición de cada *locus* sobre el mapa, los valores de las distancias dadas en centiMorgan están al lado izquierdo y las distancias de cada grupo se muestran debajo de cada GL. Los marcadores presentados en letras rojas, son los que se lograron anclar en este trabajo. Los asteriscos indican los marcadores con segregación distorsionada. (Longitud total del mapa: 1771,3cM)

5.3.2 Población DOR364 x G19833

Para la población derivada de la crucea DG, se obtuvo un mapa que presentó una longitud de 2439,7cM. Este mapa es 539,7cM más largo que el obtenido por Galeano *et al.* (2009a), el cual solo presentaba marcadores tipo SSR, STS, RFLP y SNP (288 marcadores), mientras que el mapa actual posee 526, combinando marcadores tipo RFLP, SNP SSR, basados en genes, STS, y Leg, como lo muestra la siguiente tabla:

Tabla 5. Resumen de los marcadores actualmente integrados en el mapa genético de la población DOR364 x G19833.

<i>Grupos de ligamiento(GL)</i>	<i>RFLP</i>	<i>SNP</i>	<i>SSR</i>	<i>Basados en genes*</i>	<i>STS</i>	<i>LEG</i>	<i>TOTAL</i>	<i>Distancia promedio entre marcadores (cM)</i>	<i>Distancia total de cada GL (cM)</i>
B01	3	15	46	1			65	4,3	272,7
B02	3	16	56	3		2	80	4,3	342,6
B03	5	12	33	2	1	5	58	4,7	265,1
B04	2	10	30	3	1		46	4	179,1
B05	1	5	19	1			26	5,7	143,3
B06	7	8	22		1	3	41	4,3	173
B07	6	6	29			3	44	5,2	225,5
B08	1	17	34	3	1	2	58	4,5	256,8
B09	1	8	21			1	31	7,1	212,1
B10	1	6	28		1	2	38	4,7	173,2
B11	3	9	26			1	39	5,2	196,6
TOTAL	33	112	344	13	5	19	526	4,97	2439,7

*Genes de nodulación

En este trabajo se lograron mapear marcadores en todos los grupos de ligamiento (en promedio, tres por cada grupo), siendo el B03, el grupo al cual se lograron anclar más marcadores (7), mientras que en grupos como el B01, B05, B09 y B11, únicamente se ubicó un marcador respectivamente (Figura 9).

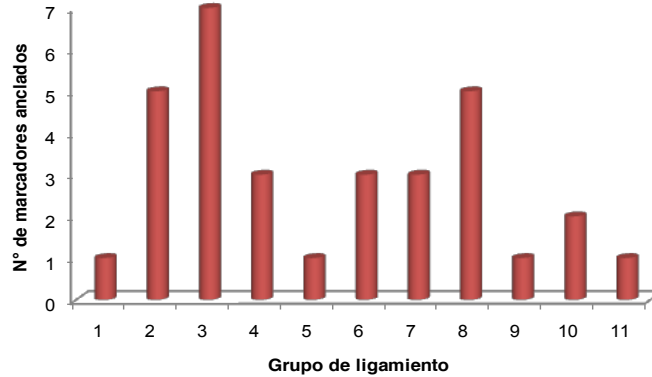


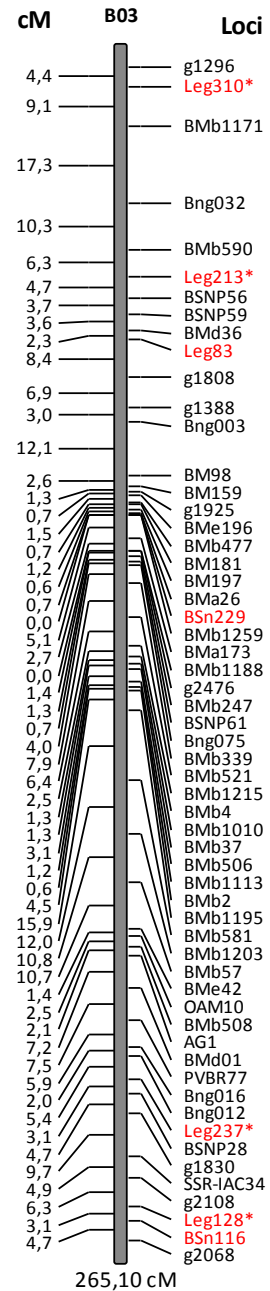
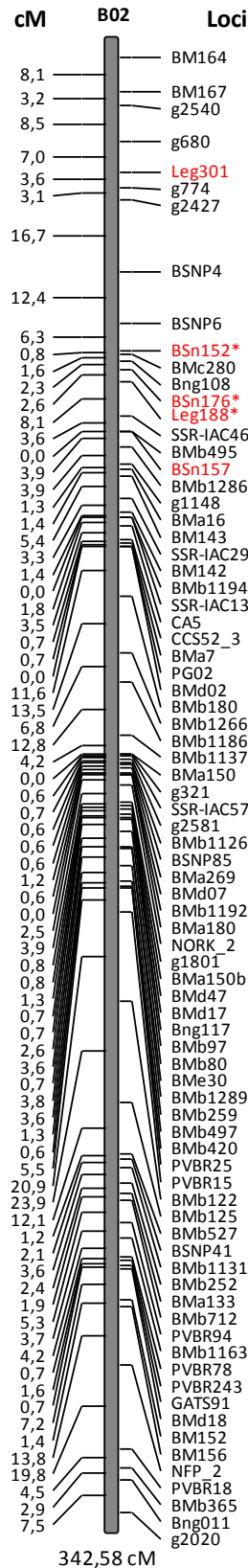
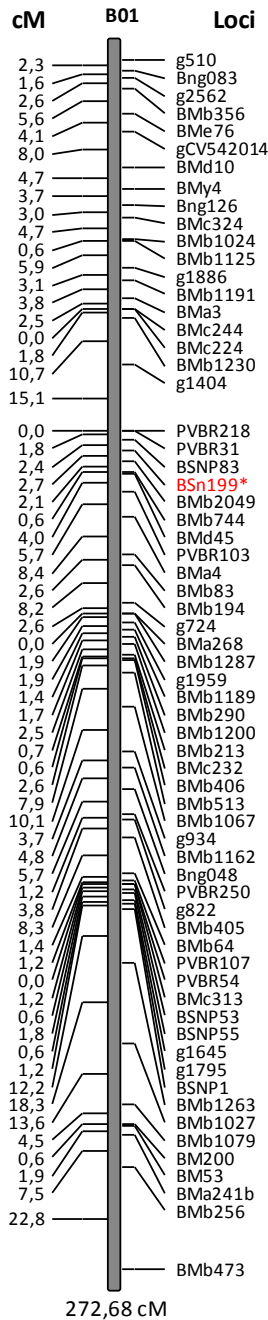
Figura 9. Número de nuevos marcadores ligados al mapa DOR364 x G19833.

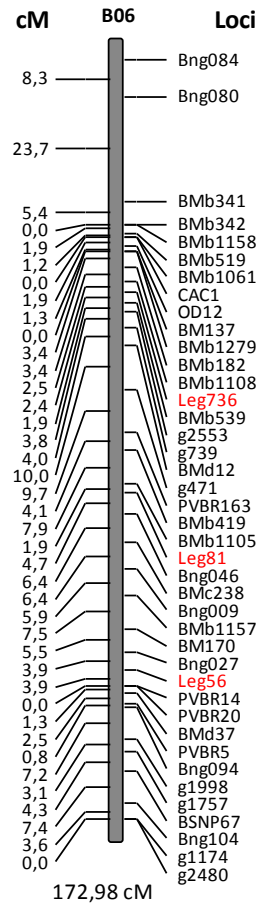
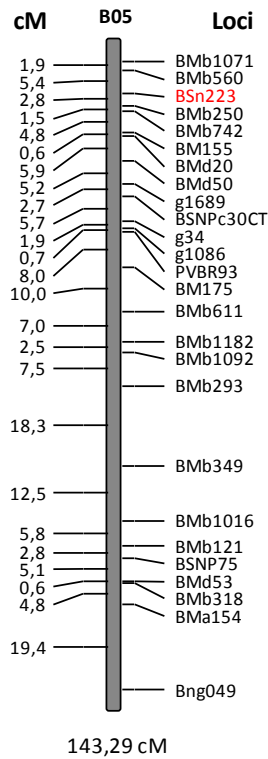
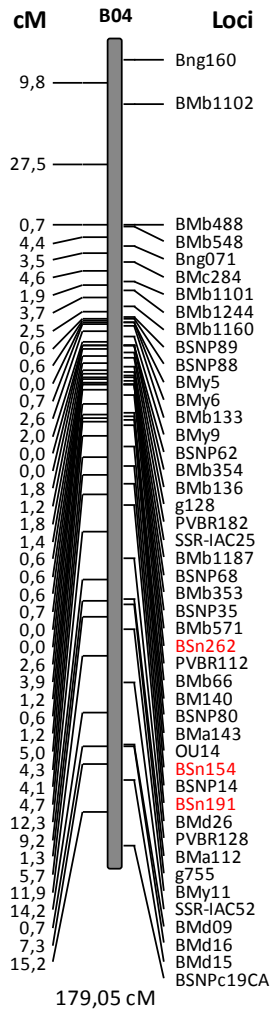
Contrario a lo sucedido en el mapa genético de la población anterior, para este caso, el número de marcadores está relacionado con la longitud del GL, pues el grupo B02 presenta anclados 80 marcadores y es el más largo (342,6cM), mientras que el B05 fue el más corto (143,3cM) con únicamente 26 marcadores. La longitud promedio de los GL fue de 221,81cM (Figura 10). Estos resultados son parcialmente coherentes con lo encontrado por Galeano *et al.* (2009a), quienes presentan también al GL B02 y B05 como los grupos con mayor y menor número de marcadores, sin embargo, para ellos estos no fueron los grupos con las longitudes contrastantes. El GL B02, ha sido reportado en varios estudios como el cromosoma más largo empleando diversos tipos de marcadores (Blair *et al.*, 2003; Córdoba *et al.*, 2010a; García *et al.*, 2011).

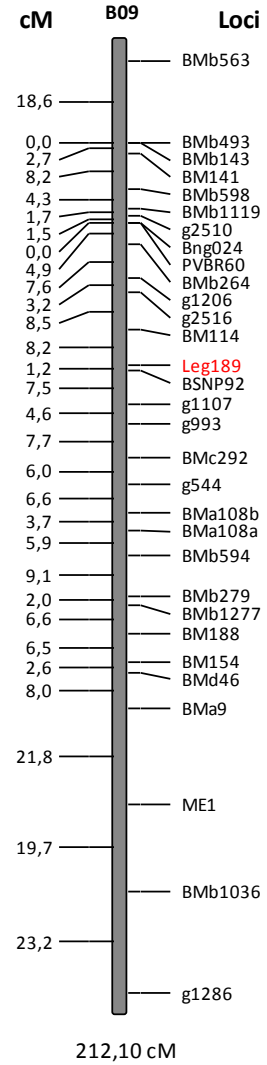
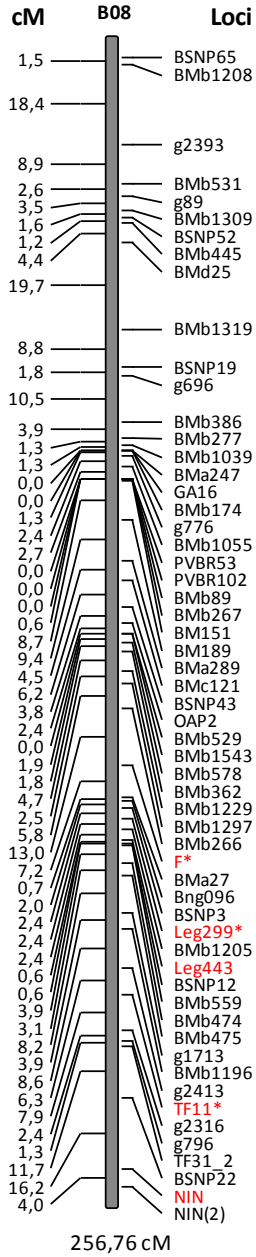
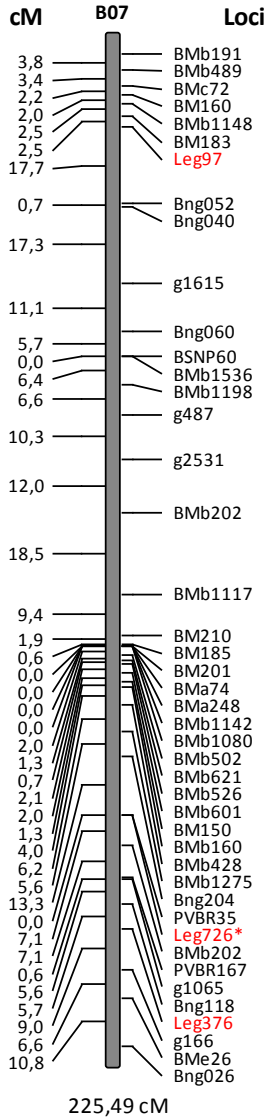
La distancia promedio entre marcadores fue de 4,97cM, que es menor a las reportadas en otros trabajos (Beebe *et al.*, 2006; Grisi *et al.*, 2007; Galeano *et al.*, 2009a; Blair *et al.*, 2010c; de Campos *et al.*, 2011; Córdoba *et al.*, 2011b, García *et al.*, 2011)

Los marcadores Leg presentaron una mayor distribución a lo largo de los cromosomas, pues estuvieron representados en ocho de los grupos (B02, B03, B06, B07, B08, B09 B10 y B11), mientras que los BSn estuvieron restringidos en los cinco primeros GL. Cada uno de los tipos de marcadores se presentó de

manera exclusiva en tres de los GL: los BSn estuvieron en los grupos B01, B04 y B05 y los Leg en los GL B06, B09 y B10.







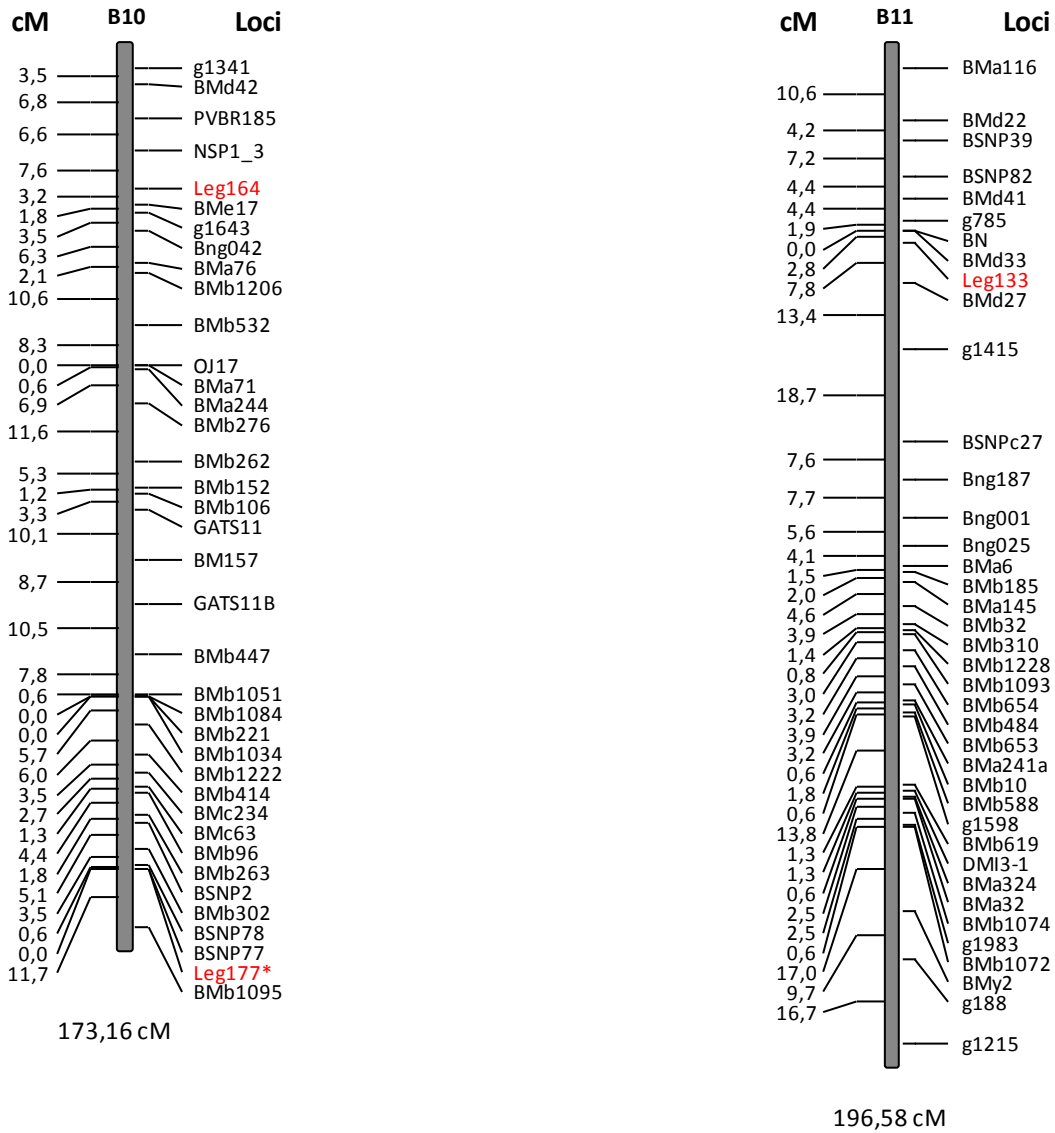


Figura 10. Mapa de ligamiento de la población derivada de la cruce inter-acervo DOR364 x G19833 de frijón común. Cada barra vertical representa un GL. Los nombres de los marcadores se muestran al lado derecho de cada GL, las líneas conectoras indican la posición de cada *locus* sobre el mapa, los valores de las distancias dadas en centiMorgan están al lado izquierdo y las distancias de cada grupo se muestran debajo de cada GL. Los marcadores presentados en letras rojas, son los que se lograron anclar en este trabajo. Los asteriscos indican los marcadores con segregación distorsionada (Longitud total del mapa: 2439,7cM).

5.4. Mapeo comparativo

Los marcadores que se anclaron a cada uno de los mapas en este estudio fueron específicos de un solo lugar en el genoma, lo cual permitió que pudiesen ser usados al momento de confrontarlos en los dos mapas. Un total de 30 marcadores (29 genómicos y 1 génico) de los que fueron anclados al mapa de DB, presentan reciprocidad a través del mapa de DG (Anexo 7). Blair *et al.*, (2003), realizaron un mapeo comparativo entre las poblaciones DG y BJ, encontrando únicamente 14 microsatélites recíprocos para ambas poblaciones de mapeo. Este mapeo comparativo muestra la consistencia de la localización de los marcadores sobre ambas poblaciones: mapeando sobre el mismo cromosoma y la ubicación equivalente en el mapa de cada población (Blair *et al.*, 2003).

El 38,8% (19/49) de los marcadores mapeados en DB no aparecen en el mapa de DG, lo cual indica que esta última población no presenta polimorfismo para las regiones amplificadas por estos marcadores o esto no han sido evaluados, como es el caso del PvM127, que se ubica en el GL B10 en el mapa de la población BJ (Hanai *et al.*, 2009) y se mantiene en esa misma posición al ser mapeado en DB. Para este marcador, Hanai *et al.* (2009) encontraron que tiene homología con una proteína hipotética de *V. vinífera*.

Mapear en dos poblaciones es útil para ubicar marcadores que son monomórficos en una u otra población (Blair *et al.*, 2003). Para este caso por ejemplo, 18 Leg y 10 BSn, fueron monomórficos para la población DB, pero se lograron anclar al mapa de DG. Al comparar la ubicación de los marcadores Leg anclados en los mapas DB y DG con el mapa de BJ (Hougaard *et al.*, 2008), se encontró una concordancia total, pues todos los marcadores mantienen la misma posición en los grupos de ligamiento GL.

Respecto a los marcadores de nodulación (BSn), no se tiene un mapa de referencia para determinar y corroborar su ubicación, debido a que fueron

diseñados en esta investigación y solamente se han anclado algunos de ellos a los dos mapas aquí empleados. En total se lograron mapear 10 marcadores de este tipo (dos en DB y 10 en DG), pero no hubo presencia de marcadores polimórficos para ambas poblaciones, es decir que no hubo marcadores en común.

Entre los marcadores basados en genes de nodulación que lograron ser anclados al mapa de la población DG estuvieron *NIN* y *LHK1*. *NIN* es un regulador putativo de la transcripción que es requerido durante la etapa de infección (Libault *et al.*, 2010), División Celular Cortical (CCD) y control de la curvación de los pelos de la raíz (Schauser *et al.*, 1999). La expresión de *NIN* es drásticamente inducida en la epidermis poco después de la percepción de los factores de nodulación (NF) (Radutoiu *et al.*, 2003). Los resultados obtenidos por Marsh *et al.* (2007) en su investigación sugieren que *NIN* en *M.truncatula* juega un papel en la limitación de los patrones espaciales de la expresión del gen nodulin y que este papel puede estar relacionado con el control del número de nódulos.

Se ha demostrado genéticamente que hormonas como la citoquinina son esenciales para la organogénesis de los nódulos. En *L.japonicus*, *LHK1*, codifica un receptor quinasa de citoquinina, que sólo está implicado en la organogénesis de nódulos en la corteza y no en la formación de hilos de infección.

Las funciones antes nombradas, han sido encontradas en especies que tienen una estrecha relación con fríjol, entre las cuales se han realizado trabajos de sintenia, obteniendo resultados de alta fidelidad entre las secuencias y marcadores estudiados (Galeano *et al.*, 2009a; Hougaard *et al.*, 2008), lo cual puede estar indicando para este caso la presencia de secuencias homólogas con fríjol que podrían tener una funcionalidad similar a la presentada en las especies relacionadas.

5.5. Secuenciación y SNP

Aunque la calidad del amplificado fue chequeado antes de ser enviado (Figura 11), 18 marcadores (29%) no pudieron ser analizados debido a la baja calidad encontrada en los electroferogramas.

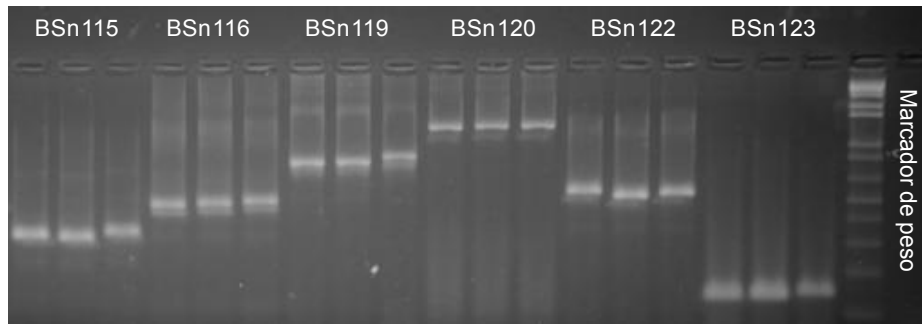


Figura 11. Evaluación de los productos de amplificación en gel de agarosa al 1,8% de los marcadores diseñados en este estudio. Se comparan los genotipos BAT477, DOR364 y G19833 respectivamente (marcador de peso de 1kb).

Una vez editada y verificada la calidad de las secuencias, se realizaron los alineamientos de los tres padres para cada marcador y registrar así los SNPs correspondientes (Figura 12).

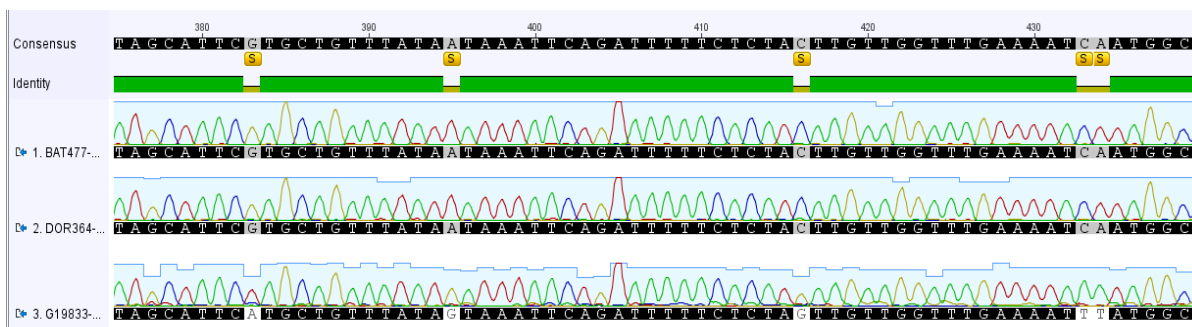


Figura 12. Alineamiento de las secuencias los padres DOR364, BAT477 y G19833 para el marcador BSn241. Se muestran cinco SNPs encontrados (S) en esta región.

En frijol común, Ramírez *et al.* (2005), analizaron EST y lograron identificar 529 SNPs en 214kb de contigs, dando aproximadamente un SNP cada 387pb. En este estudio se analizaron 26,4kb de secuencias, (sumatoria de cada una de las secuencias comparadas) encontrándose un SNP cada 92pb. Dadas estas frecuencias, los SNP pueden ser útiles para realizar un mapa de alta resolución de caracteres en frijol común y otras legumbres, también para estudios de asociación basados en genes candidatos o análisis de genoma completo (Galeano *et al.*, 2009b).

La alta frecuencia de SNPs encontradas en este trabajo es debido a que los marcadores están flanqueando intrones, los cuales presentan muchas más variaciones respecto a las regiones exónicas, como en el caso de los EST evaluados por Ramírez *et al.* (2005), donde la cantidad de SNPs reportados es menor, debido a que son regiones codificantes altamente conservadas.

En total, se analizaron 44 secuencias cuyo tamaño osciló entre 148pb y 979pb luego de ser editadas (Anexo 8). Siete de estas secuencias fueron muy conservadas, pues no se hallaron SNPs a lo largo de las mismas (BSn115, BSn131, BSn165, BSn195, BSn221, BSn229 y BSn 233). En el 48,6% de los alineamientos se encontraron SNPs para los tres genotipos; en 16 marcadores (43,3%) se reportaron SNPs únicamente para los genotipos inter-acervo (DOR364 y G19833) y tan solo el 8,1% de las secuencias exhibió cambios de bases de manera exclusiva para los genotipos Mesoamericanos DOR364 y BAT477, esta baja tasa de SNPs es explicada a partir de la estrecha diferencia genética encontrada entre estos dos padres. En la secuencia del marcador BSn158 se encontró el mayor número de SNPs (50) para los genotipos DOR364 y G19833 y fue en el marcador Bsn211 donde se registraron 30 SNPs, el número más alto para los padres DOR364 y BAT477.

Se identificaron 193 SNPs entre los genotipos DOR364 y G19833, dentro de los cuales las transversiones y las transiciones se presentaron casi en iguales proporciones (50,3% y 49,7% respectivamente), mientras que los 150 SNPs encontrados al comparar las secuencias de DOR364 y BAT477, señalaron que las transiciones (32,6%) fueron menos comunes que las transversiones (67,4%) y entre estas, las mutaciones A↔T (28,8%) y G↔C (26,8%) se presentaron en mayores proporciones, aunque no hubo una diferencia muy marcada respecto a las demás (Tabla 6). Los resultados aquí encontrados son contrarios a lo obtenido por Ramírez *et al.* (2005), quienes reportaron que el porcentaje mayor (65,6%) correspondía a mutaciones de transición, mientras que el 34,5% pertenecía a transversiones, estos resultados sugieren un patón de sustitución de nucleótido no aleatoria, con un fuerte sesgo hacia las mutaciones en transversión. Por el contrario, Zhu *et al.* (2003), encontraron que la mayoría de SNPs fueron mutaciones tipo transversión (52%), mientras que las transiciones se presentaron en un porcentaje menor (48%), lo que se hace concordante con los resultados aquí obtenidos entre los genotipos DOR364 y BAT477.

Tabla 6. Número de transversiones y transiciones encontradas en las secuencias de los tres padres (DB y DG).

Tipo de mutación	N° de cambios en DB	(%)	Cambio de bases	N° de cambios en DG	(%)
TRANSVERSIÓN	22	21,7	G↔T	28	29,2
	29	28,8	A↔T	20	20,7
	27	26,8	G↔C	27	28,1
	23	22,7	A↔C	21	22
TOTAL	101	100		96	100
TRANSICIÓN	30	61,2	G↔A	54	55,7
	19	38,8	C↔T	43	44,3
TOTAL	49	100		97	100

El análisis de SNPs podría ser útil para la detección de polimorfismo en marcadores COS (del inglés *Conserved Ortholog Set*) (Choi *et al.*, 2006), que son a menudo monomórficos en tamaño, pero variables en secuencia, y muchos de los cuales aún no se han mapeado en frijol común (Galeano *et al.*, 2009b).

6. CONCLUSIONES

- En la prueba de padres, el porcentaje de polimorfismo fue mucho mayor entre los parentales inter-acervo DG (30,5%), respecto a lo encontrado entre los genotipos Mesoamericanos DB (19,5%), lo cual guarda concordancia con otros estudios.
- Para la saturación del mapa DB fueron los marcadores genómicos los más útiles, puesto que presentaron mayor polimorfismo y distribución a lo largo de los 11 grupos de ligamiento.
- En el mapa de DG, los marcadores basados en genes estuvieron representados a través de los 11 grupos de ligamiento, mostrando así una distribución al azar.
- El mapeo comparativo, permite compartir información y determinar de una u otra forma regiones conservadas entre poblaciones.
- Como era de esperarse por su estrecha cercanía genética, se encontró un menor número de SNPS entre los genotipos Mesoamericanos en contraste con lo hallado entre los padres DOR364 y G19833.
- La abundancia de SNPS encontrados en las regiones intrónicas propician una alternativa para la generación de nuevos marcadores.
- Las transversiones fueron más prevalentes que las transiciones entre los genotipos DOR364 y BAT477.

7. RECOMENDACIONES

- Al momento de saturar un mapa genético es mejor emplear marcadores genómicos, pues presentan altas tasas de polimorfismo, lo cual ayuda a generar grupos de ligamiento más saturados.
- Generar un mapa consenso para las poblaciones de base genética estrecha, donde la saturación del mapa se hace más difícil.
- Emplear metodologías de genotipificación de SNP que permitan evaluar más muestras a bajo costo.
- Incrementar esfuerzos en el fenotipaje "*Phenotyping*" para poder ser utilizados y correlacionar de forma coherente la información de los mapas genéticos.

8. BIBLIOGRAFÍA

- AHN, S., TANKSLEY, S.D. 1993. Comparative linkage maps of the rice and maize genomes. *Proc. Natl. Acad. Sci.* 90 (17): 7980-7984.
- BHATTRAMAKKI, D. & RAFALSKI, A. 2001. Discovery and application of single nucleotide polymorphism markers in plants. In “Plant Genotyping: the DNA Fingerprinting of Plants”. CAB International. Capítulo 12: 179.
- BEEBE, S.E., OCHOA, I., SKROCH, P., NIENHHUIS, J., TIVANG, J. 1995. Genetic diversity among common bean breeding lines developed for Central America. *Crop Sci.* 35 (4): 1178-1183.
- BEEBE, S.E., PEDRAZA, F., ROJAS, M., GUTIERREZ, J., THOME, J. 1998. A genetic map combining RFLP, RAPD, SCAR y AFLP markers. *Bean Improv Coop.* 41: 95-96.
- BEEBE, S.E., ROJAS-PIERCE, M., YAN, X., BLAIR, M.W., PEDRAZA, F., MUÑOZ, F., THOME, J., LYNCH, J.P. 2006. Quantitative Trait *Loc*i for root architecture traits correlated with phosphorus acquisition in Common Bean. *Crop Science.* 46: 413-423.
- BELLUCCI, E., GORETTI, D., BITTOCCHI, E., ROSSI, M., NANNI, L., ATTENE, G., PAPA, R. 2010. Nucleotide diversity analysis in wild and domesticated *Phaseolus vulgaris* L. from Mesoamerica.). Vth International Congress on Legume Genetics and Genomic.
- BENCHIMOL, L. L., DE CAMPOS, T., CARBONELL, S.A., COLOMBO, C. A., CHIORATTO, A. F., FORMIGHIERI, E., GOUVÊA, L., SOUZA, A. 2007.

Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers. *Genetic Resources and Crop Evolution*. 54 (8): 1747-1762.

- BLAIR, M.W., PEDRAZA, F., BUENDÍA, H.F., GAITÁN-SOLÍS, E., BEEBE, S.E., GEPTS, P., & TOHME, J. 2003. Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics*. 107: 1362-1374.
- BLAIR, M.W., IRIARTE, G., BEEBE, S. 2006a. QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean · wild common bean (*Phaseolus vulgaris* L.) cross. *Theor Appl Genet*. 112: 1149-1169.
- BLAIR, M. W., GIRALDO, M. C., BUENDÍA, H. F., TOVAR, E., DUQUE, M. C., BEEBE, S. E. 2006b. Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.). *TAG. Theoretical and applied genetics*. 113: 100-109.
- BLAIR, M. W., BUENDÍA, H. F., GIRALDO, M. C., MÉTAIS, I., PELTIER, D. 2008. Characterization of AT-rich microsatellites in common bean (*Phaseolus vulgaris* L.). *TAG. Theoretical and applied genetics*. 118: 91-103.
- BLAIR, M. W., ASTUDILLO, C., GRUSAK, M.A., GRAHAM, R., BEEBE, S.E. 2009a. Inheritance of seed iron and zinc concentrations on common bean (*Phaseolus vulgaris* L.). *Molecular Breeding*. 23: 197-207.
- BLAIR, M. W., MUÑOZ, M., PEDRAZA, F., GIRALDO, M. C., BUENDÍA, H. F., HURTADO, N. 2009b. Development of microsatellite markers for

common bean (*Phaseolus vulgaris* L.) based on screening of non-enriched, small-insert genomic libraries. *Genome*. 52: 772-782.

- BLAIR, M. W., TORRES, M. M., GIRALDO, M. C., PEDRAZA, F. 2009c. Development and diversity of Andean-derived, gene-based microsatellites for common bean (*Phaseolus vulgaris* L.). *BMC Plant Biology*. 9 (100): 1-14.
- BLAIR, M.W., GALEANO, C.H., TOVAR, E., MUÑOZ TORRES, M.C., VELASCO CASTRILÓN, A., BEEBE, S.E., RAO, I. M. 2010a. Development of a Mesoamerican intra-genepool genetic map for quantitative trait *loci* detection in a drought tolerant susceptible common bean (*Phaseolus vulgaris* L.) cross. *Mol Breeding*.1-18.
- BLAIR, M.W., ASTUDILLO, C., RENGIFO, J., BEEBE, S.E., GRAHAM, R. 2010b. QTL analyses for seed iron and zinc concentrations in an intra-genepool population of Andean common beans (*Phaseolus vulgaris* L.). *TAG. Theoretical and applied genetics*. 122 (3): 511-21.
- BLAIR, M. W., MEDINA, J. I., ASTUDILLO, C., RENGIFO, J., BEEBE, S. E., MACHADO, G., GRAHAM, R. 2010c. QTL for seed iron and zinc concentration and content in a Mesoamerican common bean (*Phaseolus vulgaris* L.) population. *TAG. Theoretical and applied genetics*. 121 (6): 1059-1070.
- BROOKS, R. H., KAPLAN, L., CUTLER, H. C., WHITAKER, T. W. 1962. Plant material from a cave on the Río Zape, Durango, México. 27 (3): 356-369.

- BROUGHTON, W. J., HERNANDEZ, G., BLAIR, M., BEEBE, S., GEPTS, P., VANDERLEYDEN, J. 2003. Beans (*Phaseolus* spp.) – model food legumes. *Plant and Soil*. 252: 55-128.
- BUSO, G.S.C., AMARAL, Z.P.S., BRONDANI, R.P.V., FERREIRA, M.E. 2006. Microsatellite markers for the common bean *Phaseolus vulgaris*. *Molecular Ecology Notes*. 6: 252-254.
- CAETANO-ANOLLES, G. & GRESSHOFF, P.M. 1991. Plant genetic control of nodulation. *A nnu. Rev. Microbiol.* 45: 345-82.
- CAIXETA, E., BORÉM, A., KELLY, J. 2005. Development of microsatellite markers based on BAC common bean clones. *Crop Breeding and Applied Biotechnology*. 5: 125-133.
- CHAVARRO, M.C. 2006. Contribución a la construcción de mapas moleculares e identificación de QTL en frijol común *Phaseolus vulgaris* L. para tolerancia a calor y sequía. Trabajo de grado. Pontificia Universidad Javeriana. Facultad de Ciencias. Carrera de Biología. Bogotá D.C.
- CHOI, H.-K., M. LUCKOW, J. DOYLE, AND D. COOK. 2006. Development of nuclear gene-derived molecular markers linked to legume genetic maps. *Mol. Genet. Genomics*. 276 (1): 56-70.
- CICHY, K. A., BLAIR, M.W., GALEANO, C.H., SNAPP, S.S., KELLY, J.D. 2009. QTL Analysis of Root Architecture Traits and Low Phosphorus Tolerance in an Andean Bean Population. *Crop Science*. 49 (1): 59-68.
- CÓRDOBA, J.M. 2009. Desarrollo de marcadores microsatélites a partir de las secuencias BAC-END de frijol común (*Phaseolus vulgaris* L.) y

ubicación en los mapas genético y físico. Trabajo de grado. Universidad Nacional de Colombia, Facultad De Agronomía, Escuela de Posgrados. Bogotá D.C.

- CÓRDOBA, J.M., CHAVARRO, C., ROJAS, F., MUÑOZ, C., BLAIR, M.W. 2010a. Identification and Mapping of Simple Sequence Repeat Markers from Common Bean (*Phaseolus vulgaris* L.) Bacterial Artificial Chromosome End Sequences for Genome Characterization and Genetic–Physical Map Integration. *The Plant Genome*. 10: 436.
- CÓRDOBA, J.M., CHAVARRO, C., SCHLUETER, J.A., JACKSON, S.A., BLAIR, M.W. 2010b. Integration of physical and genetic maps of common bean through BAC-derived microsatellite markers. *BMC Genomics*. 11 (436): 436.
- DEBOUCK, D.G. 1986. Primary diversification of *Phaseolus* in the Americas: Three centres. *Plant Genet. Resour. Newsl.* 67.
- DEBOUCK, D.G. 1999. Diversity in *Phaseolus* species in relation on the common bean. En: Frei, A., Blair, M.W., Cardona, C., Beebe, S.E., Hu, H., Dorn, S. 2005. QTL Mapping of Resistance to Thrips palmi Karny in Common Bean (Genomics, Molecular Genetics & Biotechnology). *Crop Science*. 45: 379-387.
- DE CAMPOS, T., BENCHIMOL, L., CARBONELL, S.A., CHIORATTO, A.F., FORMIGHIERI, E, F., SOUZA, A. P. 2007. Notas Científicas. Microsatellites for genetic studies and breeding programs in common bean. *Pesq. agropec. Bras.* 42 (4): 589-592.

- DE CAMPOS, T., RODRIGUES OBLESSUC, P., SFORÇA, D. A., KUPPER CARDOSO, J.M., MORO BARONI, R., BARBOSA DE SOUSA, A.C., MORAIS CARBONELL, S. A., CHIORATTO, A.F., FRANCO GARCIA, A.A., BENCHIMOL RUBIANO, L., PEREIRA DE SOUZA, A. 2011. Inheritance of growth habit detected by genetic linkage analysis using microsatellites in the common bean (*Phaseolus vulgaris* L.). *Mol Breeding*. 27 (4): 549-560.
- DELLAPORTA, S., WOOD, L., HICKS, J. 1983. A plant DNA miniprep: Version ii. *Plant Molecular Biology Reporter*. 1 (14): 19-21.
- DOERGE, R. 2002. Mapping and analysis of quantitative trait *loci* in experimental populations. En: Semagn, K., Ndjiondjop, A., Bjørnstad, A. 2006. Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.
- DOUABIN-GICQUEL, V., SORIANO, N., FERRAN, H., WOJCIK, F., PALIERNE, E., TAMIN, S., JOVELIN, T., MCKEI, A.T., LE GALL, J.Y., DAVID, V., MOSSER, J. (2001). Identification of 96 single nucleotide polymorphisms in eight genes involved in iron metabolism: efficiency of bioinformatic extraction compared with a systematic sequencing approach. En: Soleimani, V. D., Baum, B. R., Johnson, D. A. (2003). Efficient Validation of Single Nucleotide Polymorphisms in Plants by Allele-Specific PCR, With an Example From Barley. *Plant Molecular Biology*. 21: 281-288.
- DRUMMOND, A.J., ASHTON, B., BUXTON, S., CHEUNG, M., COOPER, A., HELED, J., KEARSE, M., MOIR, R., STONES-HAVAS, S., STURROCK, S., THIERER, T., WILSON, A. 2010. *Geneious* v5.3.4, Available from <http://www.geneious.com>
- FERGUSON, B.J., INDRASUMUNAR, A., HAYASHI, S., LIN, M.H., LIN, Y.H., REID, D.E., P.M., GRESSHOFF. 2010. Molecular Analysis of Legume

Nodule Development and Autoregulation. *Journal of Integrative Plant Biology*. 52 (1): 61-76.

- FERREIRA, A., DA SILVA, M.F., DA COSTA E SILVA L., CRUZ, C.D., 2006. Estimating the effects of population size and type on the accuracy of genetic maps. *Genet. Mol. Biol.* 29 (1): 187-192.
- FREI, A., BLAIR, M.W., CARDONA, C., BEEBE, S.E., HU, H., DORN, S. 2005. QTL Mapping of Resistance to *Thrips palmi* Karny in Common Bean. *Crop Science*. 45: 379-387.
- FREYRE, R., SKROCH, P. W., GEFFROY, V., ADAM-BLONDON, A-F., SHIRMOHAMADALI, A, JOHNSON, W. C., LLACA, V., NODARI, R. O., PEREIRA, P. A., TSAI, S.-M., TOHME, J., DRON, M. NIENHUIS, J., VALLEJOS, C. E., GEPTS, P. 1998. Towards an integrated linkage map of common bean. 4. Development of a core linkage map and alignment of RFLP maps. *TAG Theoretical and Applied Genetics*. 97: 847-856.
- GAITÁN-SOLÍS, E., DUQUE, M. C., EDWARDS, K. J., TOHME, J. (2002). Microsatellite repeats in common bean (*Phaseolus vulgaris*): Isolation, characterization, and cross-species amplification in *Phaseolus* ssp. *Crop Science*. 42: 2128-2136.
- GALEANO, C. H., FERNÁNDEZ, A. C., GÓMEZ, M., BLAIR, M. W. 2009a. Single strand conformation polymorphism based SNP and Indel markers for genetic mapping and synteny analysis of common bean (*Phaseolus vulgaris* L.). *BMC genomics*. 10: 629-643.
- GALEANO, C. H., GÓMEZ, M., RODRÍGUEZ, L. M., BLAIR, M. W. 2009b. CEL I nuclease digestion for SNP discovery and marker development in common bean. (*Phaseolus vulgaris* L.). *Crop science*. 49 (2): 381-394.

- GARCIA, R.A.V., RANGEL, P.M., BRONDANI, C., MARTINS, W.S., MELO, L.C., CARNEIRO, M. S., BORBA, T,C.O., BRONDANI, R.P.V. 2011. The characterization of a new set of EST-derived simple sequence repeat (SSR) markers as a resource for the genetic analysis of *Phaseolus vulgaris*. *BMC Genetics*. 12 (41).
- GEPTS, P. 1999. Development of an Integrated Linkage Map. In: Common Bean Improvement in the Twenty-First Century. S .P. Singh. Kluwer Academic Publishers. En: CHAVARRO, M.C. 2006. Contribución a la construcción de mapas moleculares e identificación de QTL en frijol común *Phaseolus vulgaris* L. para tolerancia a calor y sequía. Trabajo de grado. Pontificia Universidad Javeriana. Facultad de Ciencias. Carrera de Biología. Bogotá D.C.
- GEPTS, P. & DEBOUCK, D., 1991. Origen, domestication, and evolution of the common bean (*Phaseolus vulgaris*). Research for crop Improvement. van Shoonhoven, A., Voysest, O. CIAT. CAB International. 7-53.
- GEPTS, P. 2001. *Phaseolus vulgaris* (Beans). Encyclopedia of Genetics. Department of Agronomy and Range Science, University of California, Davis, CA 95616-8515, USA. 1444-1445.
- GEPTS, P., BEAVIS, W.D., BRUMMER, E.C., SHOEMAKER, R.C., STALKER, H.T., WEEDEN, N.F., YOUNG, N.D. 2005. Legumes as a model plant family. Genomics for food and feed. Report of the Cross-legume Advances through Genomics conference. *Plant Physiology*.137: 1228-1235.
- GRAHAM, P.H., VANCE, C.P. 2003. Legumes: Importance and constraints to greater use. *Plant Physiol*. 131 (3): 872-877.

- GRESSHOFF, P.M. 2005. Positional Cloning of Plant Developmental Genes. En: *The Handbook of Plant Genome Mapping Genetic and Physical Mapping*. Meksem, K., Kahl, G. Wiley-VCH. Capítulo 10.
- GRISI, M. C. M., BLAIR, M. W., GEPTS, P., BRONDANI, C., PEREIRA, P.A.A., BRONDANI, P.V. 2007. Genetic mapping of a new set of microsatellite markers in a reference common bean (*Phaseolus vulgaris*) population BAT93 x Jalo EEP558. *Genetics and Molecular Research*. 6 (3): 391-706.
- GUERRA-SANZ, J., 2004. New SSR markers of *Phaseolus vulgaris* from sequence database. *Plant Breeding*. 123: 87-89.
- GUPTA, P.K., & VARSHNEY, R.K. 2000. The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica*. 113: 163-185.
- HANAI, L. R., DE CAMPOS, T., ARANHA, L.E., BENCHIMOL, L.L., PEREIRA, A., MELOTTO, M., MORAES, S.A., CHIORATTO, A.F., CONSOLI, L., FORMIGHIERI, E.F., VINÍCIUS, M., MUI, S., CARNEIRO, M.L. 2007. Development, characterization, and comparative analysis of polymorphism at common bean SSR *loci* isolated from genic and genomic sources. *Biosystems*, 50 (3): 266-77.
- HANAI, L. R., SANTINI, L., CAMARGO, L. E. A., FUNGARO, M. H. P., GEPTS, P., TSAI, S. M., CARNEIRO, M.L. 2009. Extension of the core map of common bean with EST-SSR, RGA, AFLP, and putative functional markers. *Molecular*. 25: 25-45.

- HALDANE, J.B.S. 1931. The cytological basis of genetical interference. *Cytologia*. Vol. 3. En: Semagn, K., Ndjioudjop, A., Bjørnstad, A. 2006. Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.
- HIRSCH, S., KIM, J., MUÑOZ, A., HECKMANN, A.B., DOWNIE, J.A., OLDROYD, G.E. 2009. GRAS proteins form a DNA binding complex to induce gene expression during nodulation signaling in *Medicago truncatula*. *Plant Cell*. 21: 545-557.
- HOOVER, T.R., ROBERTSON, A.D., CERNY, R.L., HAYES, R.N., IMPERIAL, J., SHAH, V.K., LUDDEN, P.W. 1987. Identification of the V factor needed for synthesis of the iron–molybdenum cofactor of nitrogenase as homocitrate. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51(9): 1381-1397.
- HOUGAARD, B. K., MADSEN, L. H., SANDAL, N., CARVALHO MORETZSOHN, M. DE, FREDSLUND, J., SCHAUSER, L., NIELSEN, A.M., ROHDE, T., SATO, S., TABATA, S., BERTIOLI, D.J., STOUGAARD, J. 2008. Legume anchor markers link syntenic regions between *Phaseolus vulgaris*, *Lotus japonicus*, *Medicago truncatula* and *Arachis*. *Genetics*. 179 (4): 2299-312.
- HORTON, R., NIBLETT D., MILNE S., PALMER S., TUBBY B., TROWSDALE J., BECK S. 1998. Large-scale sequence comparisons reveal unusually high levels of variation in the HLA-DQB1 locus in the class II region of the human MHC. *J. Mol. Biol.* 282 (1): 71-97.

- HURTADO, N. 2009. Ubicación de microsatélites (SSR) diseñados a partir de secuencias codificantes ESTs en el mapa genético de frijol común (*Phaseolus vulgaris*). Trabajo de Grado. Pontificia Universidad Javeriana. Facultad de Ciencias. Carrera de Biología. Bogotá D.C.
- IMAIZUMI-ANRAKU, H., TAKEDA, N., CHARPENTIER, M., PERRY, J., MIWA, H., UMEHARA, Y. KOUCHI, H., MURAKAMI, Y., MULDER, L., VICKERS, K., PIKE, J., DOWNIE, J.A., WANG, T., SATO, S., ASAMIZU, E., TABATA, S., YOSHIKAWA, M., MUROOKA, Y., WU, G.J., KAWAGUCHI, M., KAWASAKI, S., PARNISKE, M., HAYASHI, M. 2005. Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.
- INDRASUMUNAR, A., KERESZT, A., SEARLE, I., MIYAGI, M., LI, D., NGUYEN, C.D., MEN, A., CARROLL, B.J., GRESSHOFF, P.M. 2010. Inactivation of duplicated nod factor receptor 5 (NFR5) genes in recessive loss-of-function non-nodulation mutants of allotetraploid soybean (*Glycine max* L. Merr.). *Plant Cell Physiol*. 51 (2): 201-214.
- KAHL, G., MAST, A., TOOKE, N., SHEN, R., VAN DEN BOOM, D. 2005. Single nucleotide Polymorphisms: Detection Techniques and Their Potential for Genotyping and Genome Mapping. En: *The Handbook of Plant Genome Mapping. Genetic and Physical Mapping*. Meksem, K., Kahl, G. Wiley-VCH. Capítulo 4.
- KANAMORI, N., MADSEN, L.H., RADUTOIU, S., FRANDESCU, M., QUISTGAARD, E.M., MIWA, H. DOWNIE, J.A., JAMES, E.K., FELLE, H.H.,

HAANING, L.L., JENSEN, T.H., SATO, S., NAKAMURA, Y., TABATA, S., SANDAL, N., STOUGAARD, S. 2006. A nucleoporin is required for induction of Ca²⁺ spiking in legume nodule development and essential for rhizobial and fungal symbiosis. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.

- KAPLAN, L. & MACNEISH, R. S. 1960. Prehistoric bean remains from caves in the Ocampo region of Tamaulipas, Mexico. En: Gepts, P., Debouck, D. 1991. Origen, domestication, and evolution of the common bean (*Phaseolus vulgaris*). Research for crop Improvement. van Shoonhoven, A. & Voysest, O. CIAT. CAB International. 7-53.
- KAPLAN, L., LYNCH, T. F., SMITH, C. E., Jr. 1973. Early cultivated beans (*Phaseolus vulgaris*) from an intertropical Peruvian valley. En: Gepts, P., Debouck, D. 1991. Origen, domestication, and evolution of the common bean (*Phaseolus vulgaris*). Research for crop Improvement. van Shoonhoven, A. & Voysest, O. CIAT. CAB International. 7-53.
- KAPLAN, L. 1981. What is the origin of the common bean?. En: Gepts, P., Debouck, D. 1991. Origen, domestication, and evolution of the common bean (*Phaseolus vulgaris*). Research for crop Improvement. van Shoonhoven, A. & Voysest, O. CIAT. CAB International. 7-53.
- KEVEI, Z., LOUGNON, G., MERGAERT, P., HORVATH, G.V., KERESZT, A., JAYARAMAN, D. 2007. 3-Hydroxy-3-methylglutaryl coenzyme A reductase 1 interacts with NORK and is crucial for nodulation in *Medicago truncatula*. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010.

How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.

- KOSAMBI, D.D. 1944. The estimation of map distances from recombination values. *Ann. Eugen.* Vol.12. En: Semagn, K., Ndjiondjop, A., Bjørnstad, A. 2006. Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.
- KOUCHI, H., IMAIZUMI-ANRAKU, H., HAYASHI, M., HAKOYAMA, T., NAKAGAWA, T., UMEHARA, Y., SUGANUMA, NORIO, S., MASAYOSHI, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue–Mini Review. *Plant Cell*. 51 (9): 1381-1397.
- KUMAGAI, H., HAKOYAMA, T., UMEHARA, Y., SATO, S., KANEKO, T., TABATA, S. 2007. A novel ankyrin-repeat membrane protein IGN1 is required for persistence of nitrogen-fixing symbiosis in root nodules of *Lotus japonicus*. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.
- KUHN, D. N., MOTAMAYOR, J. C., MEEROW, A. W., BORRONE, J. W., SCHNELL, R. J. 2008. SSCP markers provide a useful alternative to microsatellites in genotyping and estimating genetic diversity in populations and germplasm collections of plant specialty crops. *Electrophoresis*. 29 (19): 4096-108.

- LAVI, U., AKKAYA, M.S., BHAGWAT, A.A., LAHAV, E., CREGAN, P.B.1994. Generation of simple sequence repeat DNA markers in avocado (*Persea americana* M.). En: Yu, K., Park, J., Poysa, V., Gepts, P. 2000. Integration of Simple Sequence Repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). *The Journal of Heredity*. 91 (6): 429-434.
- LEFEBVRE, B., TIMMERS, T., MBENGUE, M., MOREAU, S., HERVE, C., TOTH, K. 2010. A remorin protein interacts with symbiotic receptors and regulates bacterial infection. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.
- LI, Z., PINSON, S.R.M., PATERSON, A.H., PARK, W.D., STANSEL, J.W., 1997. Genetics of hybrid sterility and hybrid breakdown in an intersubspecific rice (*Oryza sativa* L.) population. *Genetics*. 145 (4): 1139-48.
- LIBAULT, M., FARMER, A., BRECHENMACHER, L., DRNEVICH, J., LANGLEY, R. J., BILGIN, D. D., RADWAN, O., NEECE, D.J., CLOUGH., S.J., MAY, G.D., STACEY, G. 2010. Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection. *Plant physiology*. 152 (2): 541-52.
- LORIEUX, M. 2008. Mapdisto. [en línea]. <http://mapdisto.free.fr>.

- LU, H., ROMERO-SEVERSON, J., & BERNARDO, R. 2002. Chromosomal regions associated with segregation distortion in maize. *TAG Theoretical and Applied Genetics*. 105: 622-628.
- LYNCH, J. & WHITE, J.W. 1992. Shoot nitrogen dynamics in tropical common bean. *Crop Science*. 32 (2): 392-397.
- MA, Y. & BLISS, F. A. 1978. Seed proteins of common bean. *Crop Sci*. 17: 431-437.
- MARSH, J. F., RAKOCEVIC, A., MITRA, R. M., BROCARD, L., SUN, J., ESCHSTRUTH, A., LONG, S.R., SCHULTZE, M., RATET, P., OLDROYD, G.D.E. 2007. *Medicago truncatula* NIN is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. *Plant physiology*. 144 (1): 324-335.
- MARÉCHAL, R. 1971. Observations sur quelques hybrides dans le genre *Phaseolus*. Les phénomènes méiotiques. *Bulletin Research Agronomical of Gembloux*. 6.
- MEDINA, J.I. 2006. Saturación del mapa de ligamiento de la población G14519xG4825 de frijol común (*Phaseolus vulgaris* L.) e identificación de QTL asociados a la acumulación de hierro y cinc en la semilla. Trabajo de grado. Universidad de Antioquia. Facultad de Ciencias Exactas y Naturales Instituto de Biología. Medellín.
- MOHAN, M., NAIR, S., BHAGWAT, A., KRISHNA, T.G., YANO, M. 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. En: Semagn, K., Ndjiondjop, A., Bjørnstad, A. 2006. Principles,

requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.

- NGUYEN, H. & WU, X. 2005. Molecular Marker Systems for Genetic Mapping. En: *Handbook of Plant Genome Mapping Genetic and Physical Mapping*. Meksem, K., Kahl, G. Wiley-VCH. Capítulo 2.
- OCHOA, I. E., BLAIR, M. W., LYNCH, J. P. 2006. QTL Analysis of Adventitious Root Formation in Common Bean under Contrasting Phosphorus Availability. *Crop Science*. 46 (4): 1609-1621.
- OKA-KIRA, E., TATENO, K., MIURA, K., HAGA, T., HAYASHI, M., HARADA, K. 2005. *klv*, a novel hypernodulation mutant of *Lotus japonicus* affected in vascular tissue organization and floral induction. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.
- OKA-KIRA, E. & KAWAGUCHI, M. 2006. Long-distance signaling to control root nodule number. *Plant Biol*. 9 (5): 496-502.
- OKAMOTO, S., OHNISHI, E., SATO, S., TAKAHASHI, H., NAKAZONO, M., TABATA, S., 2009. Nod factor/nitrate-induced CLE genes that drive HAR1-mediated systemic regulation of nodulation. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.

- OLIVEIRA, J.E., PÁDUA, J.G., ZUCCHI, M.I., VENCOVSKY, R., CARNEIRO VIEIRA, M.L. 2006. Review: Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology*. 29 (2): 294-307.
- ORITA, M., IWAHANA, H., KANAZAWA, H., HAYASHI, K., SEKIYA, T. 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. En: Galeano, C. H., Fernández, A. C., Gómez, M., Blair, M. W. 2009a. Single strand conformation polymorphism based SNP and Indel markers for genetic mapping and synteny analysis of common bean (*Phaseolus vulgaris* L.). *BMC genomics*. 10: 629-643.
- ORTIZ, J.P.A., PESSINO, S.C., BHAT, V., HAYWARD, M.D., QUARIN, C.L. 2001. A genetic linkage map of diploid *Paspalum notatum*. En: Semagn, K., Ndjiondjop, A., Bjørnstad, A. 2006. Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.
- PATERSON, A.H. 1996. Making genetic maps. *Genome mapping in plants*, San Diego, California: Academic Press, Austin, Texas. 23-39.
- PATERSON, A.H., BOWERS, J.E., BUROW, M.D., DRAYE, X., ELSIK, C.G., JIANG, C.X., KATSAR, C.S., LAN, T.H., LIN, Y.R., MING. R., WRIGHT, R.J. 2000. Comparative genomics of plant chromosomes. *Plant Cell*. 12: 1523-1540.
- PEDROSA, A., VALLEJOS, C. E., BACHMAIR, A., SCHWEIZER, D. 2003. Integration of common bean (*Phaseolus vulgaris* L.) linkage and chromosomal maps. *Theoretical and applied genetics*. 106 (2): 205-212.

- PEDROSA-HARAND, A., KAMI, J., GEPTS, P., GEFFROY, V., SCHWEIZER, D. 2009. Cytogenetic mapping of common bean chromosomes reveals a less compartmentalized small-genome plant species. *Chromosome*. 17 (3): 405-417.
- PENMETSIA, R.V., URIBE, P., ANDERSON, J., LICHTENZVEIG, J., GISH, J.C., NAM, Y.W., ENGSTROM, E., XU, K., SCKISEL, G., PEREIRA, M., BAEK, J.M., LOPEZ- MEYER, M., LONG, S.R., HARRISON, M.J., SINGH, K.B., KISS, G.B., COOK, D.R. 2008) The Medicago truncatula of the Arabidopsis EIN2 gene, sickle, is a negative regulator of symbiotic and pathogenic microbial interactions. *Plant J*. 55 (4): 580-595.
- PEOPLES, M.B., BROCKWELL, J., HERRIDGE, D.F., ROCHESTER, I.J., ALVES, B.J.R., URQUIAGA, S., BODDEY, R.M., DAKORA, F.D., BHATTARAI, S., MASKEY, S.L., SAMPET, C., RERKASEM, B., KHAN, D.F., HAUGGAARD-NIELSEN, H., JENSEN, E.S. 2009. The contributions of nitrogen-fixing crop legumes to the productivity of agricultural systems. *Symbiosis*. 48 (1): 1-17.
- PLOMION, C., O'MALLEY, D.M., DUREL, C.E. 1995. Genomic analysis in maritime pine (*Pinus pinaster*): comparison of two RAPD maps using selfed and open-pollinated seeds of the same individual. En: Semagn, K., Ndjiondjop, A., Bjørnstad, A. 2006. Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.
- RADUTOIU, S., MADSEN, L.H., MADSEN, E.B., FELLE, H.H., UMEHARA, Y., GRONLUND, M. 2003. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature*. 425: 585-592.

- RAMÍREZ, M., GRAHAM, M. A., BLANCO-LÓPEZ, L., SILVENTE, S., MEDRANO-SOTO, A., BLAIR, M. W., HERNÁNDEZ, G., VANCE, C. P., LARA, M. 2005. Sequencing and Analysis of Common Bean ESTs. *Plant Physiology*. 137: 1211-1227.
- REMANS, R., BEEBE, S., BLAIR, M., MANRIQUE, G., TOVAR, E., RAO, I., CROONENBORGH, A., TORRES-GUTIERREZ, R., EL-HOWEITY, M., MICHIELS, J., VANDERLEYDEN, J. 2007. Physiological and genetic analysis of root responsiveness to auxin-producing plant growth-promoting bacteria in common bean (*Phaseolus vulgaris* L.). *Plant and Soil*. 302 (1): 149-161.
- RISCH, N. 1992. Genetic linkage: Interpreting LOD scores. En: Semagn, K., Ndjiondjop, A., Bjørnstad, A. 2006. Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.
- RODRIGUES, T.B & SANTOS, J.B. 2006. Effect of natural selection on common bean (*Phaseolus vulgaris*) microsatellite alleles. *Genetics and Molecular Biology*. 29 (2): 345-352.
- SANO, Y. 1990. The genic nature of gamete eliminator in rice. *Genetics*. 125: 183-191.
- SAITO, K., YOSHIKAWA, M., YANO, K., MIWA, H., UCHIDA, H., ASAMIZU, E. 2007. NUCLEOPORIN85 is required for calcium spiking, fungal and bacterial symbioses, and seed production in *Lotus japonicus*. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic

Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.

- SCHAUSER, L., ROUSSIS, A., STILLER, J., STOUGAARD, J. 1999. A plant regulator controlling development of symbiotic root nodules. *Nature*. 402: 191-195.
- SCHNEIDER, K. 2005. Mapping Populations and Principles of Genetic Mapping. En: *The Handbook of Plant Genome Mapping. Genetic and Physical Mapping*. Meksem, K., Kahl, G. Wiley-VCH. Capítulo 1.
- SEMAGN, K., NDJIONDJOP, A., BJØRNSTAD, A. 2006. Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.
- SHIZHONG, X .2008. Quantitative trait locus mapping can benefit from segregation distortion. *Genetics*. 180 (4): 2201-2208.
- SINGH, S.P., GEPTS, P., DEBOUCK, D.G. 1991. Races of common bean (*Phaseolus vulgaris*, Fabaceae). *Econ Bot*. 45 (3): 379-396.
- SPRENT, J.I. 2007. Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.
- STAM, P. 1993a. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. En: Semagn, K., Ndjiondjop,

A., Bjørnstad, A. 2006. Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.

- STAM, P. 1993b. JoinMap Version 1.4: A computer program to generate genetic linkage maps. En: Semagn, K., Ndjiondjop, A., Bjørnstad, A. 2006. Principles, requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.
- STRACKE, S., KISTNER, C., YOSHIDA, S., MULDER, L., SATO, S., KANEKO, T. 2002. A plant receptor-like kinase required for both bacterial and fungal symbiosis. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1318-1397.
- TANKSLEY, S.D., YOUNG, N.D., PATTERSON, A.H., BONIERBALE, M.W. 1989. RFLP mapping in plant breeding: New tools for an old science. *Bio/Technology*. 7: 257-264.
- TIRICHINE, L., IMAIZUMI-ANRAKU, H., YOSHIDA, S., MURAKAMI, Y., MADSEN, L.H., MIWA, H. 2006. Deregulation of a Ca²⁺/calmodulin-dependent kinase leads to spontaneous nodule development. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.
- TIRICHINE, L., SANDAL, N., MADSEN, L.H., RADUTOIU, S., ALBREKTSSEN, A.S., SATO, S. 2007. A gain-of-function mutation in a

cytokinin receptor triggers spontaneous root nodule organogenesis. *Science*. 315 (5808): 104-107.

- TÖRJÉK, O., WITUCKA-WALL, H., MEYER, R.C., VON KORV, M., KUSTERER, B., RAUTENGARTEN, C., ALTMANN, T. 2006. Segregation distortion in *Arabidopsis* C24/Col-0 and Col-0/C24 recombinant inbred line populations is due to reduced fertility caused by epistatic interaction of two *loci*. *Theor Appl Genet*. 113 (8): 1551-1561.
- VALLEJOS, C.E., SAKIYAMA, N.S., CHASE, C.D. 1992. A molecular marker-based linkage map of *Phaseolus vulgaris* L. *Genetics*. 131: 733-740.
- VARSHNEY, R.K., GRANER, A., & SORRELLS, M.E. 2005. Genic microsatellite markers in plants: Features and applications. *TRENDS in Biotechnology*. 23 (1): 48-55.
- VUYSTEKE, M., MANK, R., ANTONISE, R., BASTIAANS, E., SENIOR, M.L., STUBER, C.W. 1999. Two high density AFLP linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. *Theor. Appl. Genet*. 99 (6). 921-935.
- WITTRNACK, L.1880. (Untitled.) Sitzungsber. En: Gepts, P., Debouck, D. 1991. Origen, domestication, and evolution of the common bean (*Phaseolus vulgaris*). Research for crop Improvement. van Shoonhoven, A. & Voysest, O. CIAT. CAB International. 7-13.
- WITTRNACK, L. 1888a. Die Heimat der Bohnen und der Kürbisse. En: Gepts, P., Debouck, D. 1991. Origen, domestication, and evolution of the common bean (*Phaseolus vulgaris*). Research for crop Improvement. van Shoonhoven, A. & Voysest, O. CIAT. CAB International. 7-13.

- WITTRNACK, L. 1888b. Die Nutzpflanzen der alten Peruaner. Gepts, P., Debouck, D. 1991. Origen, domestication, and evolution of the common bean (*Phaseolus vulgaris*). Research for crop Improvement. van Shoonhoven, A. & Voysest, O. CIAT. CAB International. 7-13.
- XU, Y., ZHU, L., XIAO, J., HUANG, N., MCCOUCH, S.R. 1997. Chromosomal regions associated with segregation distortion of molecular markers in F₂, backcross, doubled-haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). *Molecular Genome and Genetics*. 253: 535-545.
- YAISH, M.W.F & PÉREZ DE LA VEGA, M. 2003. Isolation of (GA)_n Microsatellite Sequences and Description of a Predicted MADS-box Sequence Isolated from Common Bean (*Phaseolus vulgaris* L.). *Genetics and Molecular Biology*. 26 (3): 337-342.
- YANO, K., YOSHIDA, S., MULLER, J., SINGH, S., BANBA, M., VICKERS, K. 2008. CYCLOPS, a mediator of symbiotic intracellular accommodation. En: Kouchi, H., Imaizumi-anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, Norio, S., Masayoshi, K. 2010. How Many Peas in a Pod? Legume Genes Responsible for Mutualistic Symbioses Underground Special Issue – Mini Review. *Plant Cell*. 51 (9): 1381-1397.
- YIM, Y.S., DAVIS, G.L., DURU, N.A., MUSKET, T.A., LINTON, E.W., MESSING, J.W. 2002. Characterization of three maize bacterial artificial chromosome libraries toward anchoring of the physical map to the genetic map using high-density bacterial artificial chromosome filter hybridization. En: Semagn, K., Ndjiondjop, A., Bjørnstad, A. 2006. Principles,

requirements and prospects of genetic mapping in plants. *African Journal of Biotechnology*. 5 (25): 2569-2587.

- YU, K., PARK, J., POYSA, V., GEPTS, P. 2000. Integration of Simple Sequence Repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). *The Journal of Heredity*. 91: 429-434.
- ZHU, Y. L., SONG, Q. J., HYTEN, D. L., VAN TASSELL, C. P., MATUKUMALLI, L. K., GRIMM, D. R., HYATT, S. M., FICKUS, E. W., YOUNG, N. D., CREGAN, P. B. 2003. Single-nucleotide polymorphisms in soybean. *Genetics*. 163: 1123-1134.
- ZHU, H., CHEN, T., ZHU, M., FANG, Q., KANG, H., HONG, Z. 2008. A novel ARID DNA-binding protein interacts with SymRK and is expressed during early nodule development in *Lotus japonicus*. *Plant Physiol*. 148: 337-347.

9. ANEXOS

ANEXO 1. Protocolo de extracción de ADN por CTAB

Este método fue modificado por Dellporta *et al.* (1983), y Vallejos *et al.* (1992)

1. Colectar 3 g de tejido foliar joven (trifolios abiertos) → empacar en bolsas de papel aluminio → Colocar en Nitrógeno líquido → Freezer (-80 °C).
2. Macerar Con Nitrógeno líquido hasta obtener un polvo fino → vaciar a tubo Falcon de 50 ml → pesar 3 g de tejido a un nuevo tubo Falcon para iniciar extracción.
3. *Nota:* Cuando el tejido foliar se va a liofilizar → colectar tejido en tubos Falcon de 50 ml → colocar en Hielo → directamente a Liofilizador a - 50°C y 0.4 mbar por aprox. 20 horas → Macerar → pesar 1.5 g de tejido a un nuevo tubo Falcon para iniciar extracción.
4. Agregar 15 ml de Buffer de Extracción precalentado por 10 minutos en un baño a 65°C (1.5% CTAB; 1M NaCl; 0.015M EDTA pH=8; 0.15M T/HCl PH=8; 1 % PVP ; 1.2% B-ME).
5. Vortex por 1 minuto.
6. Incubar a 65°C por 1hora, agitar por inversión cada 15 minutos.
7. Adicionar 15 ml de Fenol: Cloroformo: Octanol (25: 24:1) → Mezclar por inversión (20 veces) hasta que se forme un emulsión → Agitar en un shaker por 10 minutos.
8. Centrifugar a 4000 r.p.m por 20 minutos a 4°C.
9. Transferir fase acuosa (15 ml) a un nuevo tubo, utilizando micropipeta (sin tocar la interfase = precipitado proteico).
10. Adicionar 15 ml de Cloroformo: Octanol (24:1) → Mezclar por inversión (20 veces)→ Agitar en un shaker por 10 minutos.
11. Centrifugar a 4000 r.p.m por 20 minutos a 4°C.
12. Transferir fase acuosa (Aprox. 10 ml) a un nuevo tubo, utilizando micropipeta.
13. Añadir 1/10V de 3M Acetato de Sodio PH = 5.2 (1 ml) y 6/10 V de Isopropanol 100% frio (6 ml).
14. Mezclar por inversión muy suavemente.
15. Incubar a -20 °C mínimo durante 1h, máximo toda la noche.
15. Centrifugar a 4.000 rpm por 20 minutos a 4 °C.
16. Eliminar el sobrenadante.
17. Secar pellet a temperatura ambiente (colocar tubo invertido sobre una toalla de papel).
18. Lavar precipitado de ADN con 1ml de etanol 70% frio.

19. Eliminar sobrenadante y colocar el tubo invertido sobre una toalla de papel hasta que seque por completo (aprox. 1h).
20. Resuspender pellet en 500µL de TE pH =8 con RNAasa A (6 µL RNAasa A / 1ml TE PH =8).
Nota: El volumen de TE utilizado depende del tamaño del pellet. Pellet grande = 500 µL ; mediano = 300 µL
21. Incubar a 37⁰C por 1 hora
22. Transferir a un nuevo eppendorff de 1.5 ml.
Nota : Si el ADN no es claro → purificar ADN : Incubar a 65°C por 10 minutos → centrifugar a 14000 rpm por 5 minutos → transferir sobrenadante a un nuevo eppendorff.
23. Observar el ADN en un gel de agarosa al 1%
Agua = 10 µL ; ADN = 2µL ; Buffer de carga = 3µL
24. Cuantificar TDNA con Lambdas: λ 100, 200 y 300ng /µL

ANEXO 2. Procedimiento para realizar tinción de geles de acrilamida con Nitrato de Plata

Para visualizar los fragmentos de ADN se empleó la tinción con plata (Promega Inc., Madison, Wisconsin., USA).

1. Introducir el gel en la solución fijadora (ácido acético 10%) durante 20 minutos
2. Realizar dos lavados con agua, cada uno de 5 minutos
3. Sumergir el gel en la solución de tinción (AgNO₃ 0.1%) por 30 min.
4. Lavar el gel por 5 segundos con agua
5. Introducir en la solución reveladora (Na₂CO₃ + formaldehído 37% + tiosulfato de sodio 10mg/ml) de 5 a 10 min.
6. Sumergir el gel en solución de parada (Glicina 2% p/v) durante 5 min
7. Por último, se realiza un lavado final con agua durante 5 minutos y se deja secar.

ANEXO 3. Marcadores desarrollados a partir de genes de nodulación.

Contig	Gen	Nombre	Secuencia
Contig8755	NIN	BSn109_F BSn109_R	GGATGCTGGGTGGAACAAGGGT TGTTGGATGCAGGGAGCCCA
Contig2226	MtHB8	BSn110_F BSn110_R	ACGTGTTCCGCCACCAAGGTT GTGTCCACAGGCCCTGGCAA
Contig4450	MtHAP2.1	BSn111_F BSn111_R	TCCATGCAATGCGGCGACCA TGAGCCAAGGATGGATTGCCAAAGT
CV535404	MtERF	BSn112_F BSn112_R	ACAAAGGCCTTCAGGGAGATGGGT CCTCCCTGCAATTGCTCAAATCTGGT
CV538571	MtCNA1	BSn113_F BSn113_R	TGCAGTGAATGGATGAGAATGCTGT TGAGCAAGGGTTTGTGCTTCAGGA
Contig4142	MtCDPK1	BSn114_F BSn114_R	GGGCAAGGCTAACTGCTGCTCA CACCGCAAAGGGGATACCGCA
CV543862	LYK3	BSn115_F BSn115_R	TTCGGCTAAGGAAGCAATTGTCAAGA TCAATTGGGTAGTTATCCCCAAGCCT
Contig1237	LHK1	BSn116_F BSn116_R	GCTGAAAGTGGCAAAGCCGCA TGCCTTGCTCTCCATCATCCGA
Contig 9301	FEN1	BSn117_F BSn117_R	AGTTCGCTCGCAGTTTGGGCT GCCACTTTCATGGGCGAAGGC
Contig7916	IGN1	BSn118_F BSn118_R	GCAGCTCGTCAGGGTCATGTGG TCCCTTGTTGTTGAGCTCGT
Contig3823	Glyma01g00530	BSn119_1F BSn119_1R	TCCTTCTCGTCTGGCCTCCGA TCAGCTAGAAGGGCCACAACACA
Contig5979	Glyma03g03990	BSn120_1F BSn120_1R	CGCTCCAACCCCAACTCCCC TGCAGGCCAAAAGAAACCTTCCACA
CV536782	Glyma01g01210	BSn121_1F BSn121_1R	ACCGCGGTGGAGGCTCTCAT GGTCCCTTGAACCTTGAGCTCCCA
Contig8860	Glyma01g03470	BSn122_1F BSn122_1R	TGGTGAGCATGAAGTTTTGTTGGCA TGGGAGGGGACAGAGATGGAGTGA
FG228747	Glyma01g04370	BSn123_1F	TCAGCTCTGGCGATGGTGAGTGT

Contig10235	Glyma01g31670	BSn123_1R BSn124_1F BSn124_1R	ACCAAACGTGGGAGTGCGACA CCGCAAATGTTGGGGTTGGACG ACAATGGCATCCCTAGTGGAGGGT
Contig8835	Glyma01g32270	BSn125_1F BSn125_1R	TGGAGGTGATGGCCGTCTACTGT CCCTGCCTCCCAGTAAGAGGCTT
CV537701	Glyma01g38040	BSn126_1F BSn126_1R	ACGCGTAGGTGAGGCCAGGA GGCTCCAATGCCAACTGTGGCT
Contig8246	Glyma01g42800	BSn127_1F BSn127_1R	TGGCCAGAAGCACGAGATTTTGT TGCCATTCCCACATCCGCTGC
Contig1302	Glyma02g02430	BSn128_1F BSn128_1R	CCCCACCCTCCCACGACCAA CAGCCGACGAGGGTTGCTGG
Contig9030	Glyma02g04760	BSn129_1F BSn129_1R	CGCCACCCAACCTCCACACC ACACCACTCTGGGCAACAGGG
CV536972	Glyma02g06730	BSn130_1F BSn130_1R	TGCCCTGGTGGGTCTGTCC GCGAGAGCAGTGGGCAGTGG
Contig8022	Glyma02g08020	BSn131_1F BSn131_1R	GCTGCATGCCTCCTTCGTGGA TCTGGAAGCTCCAAGCCAGCTA
FD790466	Glyma02g09410	BSn132_1F BSn132_1R	AGCCACCAGGGTCACCGTTCA ACGCGGTTTCCCGGTGCTTT
EC911382	Glyma02g09540	BSn133_1F BSn133_1R	TGTGTGTCTCTTATTTGTCTTGGTGGG GCCATGGCTGCAAACGTAGGACT
Contig8347	Glyma02g15390	BSn134_1F BSn134_1R	CCACCACTCTATGCTCTGGACTCTCAT CCATGGGGTGCCCCAGAGTC
Contig6416	Glyma02g26160	BSn135_1F BSn135_1R	GGTGCTTATCAGTGCAAATGGGGT AGGCCATGTGGTGCATTTGGAT
Contig1929	Glyma02g42250	BSn136_1F BSn136_1R	ACGCCCTCGAAATCACCTTCCG TCCATCATCCCGGGCCCACC
Contig1887	Glyma02g42730	BSn137_1F BSn137_1R	GTGGAAGCACTTTGCCCCGGA AGGCCTTCCACAAAGCCTCCT
Contig3105	Glyma02g43280	BSn138_1F BSn138_1R	TGGTGCAACCAGCTCTCGCC TCCCATGCGCTTACCACCG
Contig4240	Glyma03g05050	BSn139_1F BSn139_1R	GGAAGCCAGTTTCCAATTAACCCAGA AGTCCGCCACTATCACATCTCCA
Contig7898	Glyma03g22960	BSn140_1F	ACGAAGTGGCTGAAAAACATGCAGC

		BSn140_1R	CGGTCAGCTTTGGGGCAGA
Contig10444	Glyma03g26660	BSn141_1F	ATGATAAAGAGCTTGCGGCATACGA
		BSn141_1R	GGGAGTTGCCATGGCTCGGG
Contig5485	Glyma03g27250	BSn142_1F	TCCTTCTCGGAATTCTCTGCGGC
		BSn142_1R	TGCCGCATCTCGAGCACTTTACG
Contig8888	Glyma03g28080	BSn143_1F	AGAAGACCCTGCTTCCAAGCCAT
		BSn143_1R	ACCAGGACCCTGATGCCAGACT
Contig1827	Glyma03g31980	BSn144_1F	GATGGGATTGAAGAAGGGTCCTTGG
		BSn144_1R	TGGCATCACTGTTTCCTCTGACCA
Contig7396	Glyma03g34440	BSn145_1F	CGTGCTTCCGACTCGGGTCTC
		BSn145_1R	GCCACGTAAGCATTGGCACGC
Contig7179	Glyma04g02750	BSn146_1F	ACACCACCAATCCCCACCA
		BSn146_1R	CTTAGCTTCGGGGCCACGCC
Contig8086	Glyma04g05230	BSn147_1F	AGGACTCAGCTGCAAACAACCTTCA
		BSn147_1R	ACCACAGTGCAGTCATCAGGCT
Contig444	Glyma04g05980	BSn148_1F	TCCTAGCTCCCCTCAGGCCA
		BSn148_1R	ACCATGTCCAAACACAATCCCAACCA
Contig8706	Glyma04g07070	BSn149_1F	GCACTGCAAAGGTTGGGAATTCAAGG
		BSn149_1R	TTCTGCACAAGTTCGTTGGGTGA
CB541451	Glyma04g11000	BSn150_1F	ACGACGGACACTTGGGAGAGAGT
		BSn150_1R	TCACCTGAATGGCCTCCCCCT
Contig7194	Glyma04g14360	BSn151_1F	AGTGTGGCTCATACGCACAAGCA
		BSn151_1R	TCCCATACCTGGGAACAAGCGTCA
Contig9532	Glyma04g42120	BSn152_1F	TGGGAAGAGGCAGTGCAGTGGT
		BSn152_1R	TCAAACCTCAAGCAGCATTGATGGCA
Contig245	Glyma04g42290	BSn153_1F	TGCCGGGACAGAATCATTGCCA
		BSn153_1R	TGGGTATGGAAGCTGCCGAGTGT
FE705457	Glyma04g42300	BSn154_1F	GGACCTCACTGTGACCCAACATCT
		BSn154_1R	AGTGCCTTTTTGCTGCTGCCCT
Contig3124	Glyma05g22660	BSn155_1F	ACCCACCACACCACCACCCA
		BSn155_1R	TGCTGCTGCTGCTGGCATTG
Contig9199	Glyma05g25010	BSn156_1F	TCTTGGTCTCCCTCCTCGGTTG
		BSn156_1R	CCCAGTTGAGCCACCACCCC
FE710214	Glyma05g25460	BSn157_1F	CCCATAACACCACCTCAATCTCCA

		Bsn157_1R	AAGCTTGCTCCACCGCCACC
Contig7336	Glyma05g28110	Bsn158_1F	TCTGCTTTCAGCAACACCCAAATGAA
		Bsn158_1R	TGGCTGATTGGGGTCCTGTGGT
Contig9436	Glyma05g33830	Bsn159_1F	AACCGGGCACCATGGACAGC
		Bsn159_1R	GGAGCAAATCCCACCATGAAAAAGTGG
Contig9232	Glyma05g36290	Bsn160_1F	ACTTCGCCTCTTTGGCGTGC
		Bsn160_1R	AGCAGGAACGAAGAGTAAGGTCCA
Contig8706	Glyma06g07160	Bsn161_1F	GGGGAGCGGCTTTTTGAGGGG
		Bsn161_1R	TGTCCAAGGCTGAAAAGCATCAATCC
FE698744	Glyma06g07780	Bsn162_1F	TGCAAAGAGGCCTAAGGTGGTGC
		Bsn162_1R	TTTCTAGTGCTTTTAGGCAGCCATGT
Contig6676	Glyma06g08610	Bsn163_1F	TGGGCAAGGGGAGCGTGAGT
		Bsn163_1R	TCCCTTGTGGTTTGGCGGCT
FE705457	Glyma06g12510	Bsn164_1F	CACCAGGACCAAGTGGGAGGC
		Bsn164_1R	AGTGCCTTTTTGCTGCTGCCCT
Contig9195	Glyma06g36590	Bsn165_1F	GGGATCGGACTCGGTTTTGGTGG
		Bsn165_1R	AGTAGCCGGAGCGAAGGGGG
Contig3232	Glyma06g42040	Bsn166_1F	TGCTCGCTGGAGAGGCTGTGA
		Bsn166_1R	TGCTTGCTGCACCACACGCT
CB540213	Glyma06g43560	Bsn167_1F	TGTGAAACTACCCACCCTTTTGACCA
		Bsn167_1R	AGCTTCTGCTGCATTCTTTTACAGC
Contig8343	Glyma06g47690	Bsn168_1F	AGCTGCTGTTGATGCTGCTCCA
		Bsn168_1R	GGCACCAGTGACCTTGGAGTTGT
FE899738	Glyma07g00840	Bsn169_1F	AGTGAACACTTCGACAACCACCA
		Bsn169_1R	CGGGACTTTTGCCTTTGCTGGGA
Contig4799	Glyma07g08050	Bsn170_1F	GGGTCAGTGTGGTGCTCACAGC
		Bsn170_1R	AGGCCAGAAGGGCTTCCCTCA
Contig9558	Glyma07g09710	Bsn171_1F	TGGCTGCCTTCTTTGGAGTTGT
		Bsn171_1R	TGGCTTGCCATCACTGCATCA
FE707853	Glyma07g09970	Bsn172_1F	ACTTCTGCAACTGTGGTGGAGTG
		Bsn172_1R	TCCAGGGCATCCCCTTCGACC
Contig8397	Glyma07g13780	Bsn173_1F	GGGAAGTTCCTAGCTTTGATGCGA
		Bsn173_1R	GGCCTGCAAGTGCCGGTGAA
Contig3305	Glyma07g15780	Bsn174_1F	TAGTCGCGCTGTGCTGGCTC

Contig1038	Glyma07g30390	BSn174_1R	TGCTGAGCCGGGTCTCCCTC
		BSn175_1F	TTCCCACACGCATCCCGCAC
		BSn175_1R	CCTCTGGGGCCTCTGGACTGG
Contig8557	Glyma07g32150	BSn176_1F	GCCAGTTCTTCAGCAAAGTCACCCC
		BSn176_1R	TCCACCCTCCACGAACCTCG
CV543443	Glyma07g38620	BSn177_1F	GGCTGGAACCGAGGACGAGAC
		BSn177_1R	TGAGACGAAAGCCCGGCCCA
Contig9232	Glyma08g03330	BSn178_1F	TGCTCACATTGTGGAGTGATAGGGCA
		BSn178_1R	ACTCCTATGCTGGGGCTGTTGCT
Contig1419	Glyma08g05850	BSn179_1F	CAGGGAGGCCAATGCGGCAA
		BSn179_1R	CGGCCATGTCGAGGATCAGCA
Contig1495	Glyma08g09450	BSn180_1F	GGGGCCATTCTCCTTCCCCA
		BSn180_1R	AGCCGCAGGATGCAAACGAA
Contig8751	Glyma08g17270	BSn181_1F	CCTCACCCCATCACAGCGCA
		BSn181_1R	TGCTCTTCTTGCAAAGAAACCATACGC
Contig6872	Glyma08g19250	BSn182_1F	TGTGTGCGTACCCCGTGCAA
		BSn182_1R	GCTGCCGAGACGCAGCACTA
Contig5836	Glyma08g23310	BSn183_1F	GCCAGCGGTGAATGGAGCGA
		BSn183_1R	TGACCTAGCCCATATTGCCAAAAGC
Contig2969	Glyma08g41230	BSn184_1F	TCTGCTCTGCTCGTGGGCAA
		BSn184_1R	ACTCCTGAGTTCCAAAAGGCATGGA
Contig7805	Glyma08g43330	BSn185_1F	CCATGGCCATTCAAATGCTTGGAG
		BSn185_1R	GGCGTAGGCAGTGGGAGGGA
CV535867	Glyma08g48030	BSn186_1F	AGGGTGGCCAAACTGAGGTGGA
		BSn186_1R	AGAAGGAATTGCTCCATTGCAGACAC
FG229201	Glyma09g04510	BSn187_1F	GGTGGCCCCGGAACCATCAA
		BSn187_1R	ACGAGCTCTCACGATTATGCAGC
Contig7859	Glyma09g17170	BSn188_1F	GGGCTGGTGGCCGAACCTTT
		BSn188_1R	TGAGAAAACAGAACACGCGAATGAGT
Contig9031	Glyma09g21820	BSn189_1F	CTCCGGTGACGCTTCCGGT
		BSn189_1R	CCCGTCCCGTGCAGGTTTCC
Contig9198	Glyma09g27600	BSn190_1F	ACGAGGCTAGCAAACAATGGCA
		BSn190_1R	CCCTTGTGGTAAGATGGGTCACACC
CB280700	Glyma09g28750	BSn191_1F	ACCAGTAGCTTATCTGGTAAACGTTGC

Contig3677	Glyma09g30030	BSn191_1R	GGGGTCACCAAGGACGTGGC
		BSn192_1F	AGGCACGGGAGCAGCTGAGA
EG563052	Glyma09g30380	BSn192_1R	AACCATACACACTTGCATCCACCAA
		BSn193_1F	ACACCCTTTGAGTTCACACTCACAAC
		BSn193_1R	ACAGAACAGGAACAGGGGAGGCA
CV537082	Glyma09g41680	BSn194_1F	CACGCGCCGGGAAGTAGTGG
		BSn194_1R	TGCTGCAGCTGAAGAGTCCGT
FG232298	Glyma10g04150	BSn195_1F	AGACCAGGAGACCTCAAATCCCC
		BSn195_1R	GGCCAGAGGTGTTGCCACTGC
CB540018	Glyma10g04350	BSn196_1F	ACCCACACATGCCTCAAAGG
		BSn196_1R	GCGTCCAAGTTTGAAGCCACGC
CV543597	Glyma10g06610	BSn197_1F	TGACCGGGTTGCTGCAAGGG
		BSn197_1R	TCCGCGCCCTGGTTTGAAGC
Contig8962	Glyma10g06810	BSn198_1F	TGGTGGATGGTCACTGCCACA
		BSn198_1R	CCTGTGTGTATCTGCATTGGCAAGTC
CV540842	Glyma10g07500	BSn199_1F	GGCTTCGAGCATCGGCGGAG
		BSn199_1R	CGGACGAGCTGCGGAACCAG
Contig9160	Glyma10g23790	BSn200_1F	TCCGAAAAGATGGCGCAGGAAG
		BSn200_1R	TGAGCCATGAGGCTCGTCCGT
FE705457	Glyma10g38600	BSn201_1F	GCACAGGACCTCACTGTGACCC
		BSn201_1R	AGTGCCTTTTTGCTGCTGCCCT
Contig1617	Glyma10g39450	BSn202_1F	CCACCACCGGAATCGACCGC
		BSn202_1R	TGGGTAGGTGGGCCATGTAAGT
Contig655	Glyma11g07750	BSn203_1F	ACACCGTCCCCAATGATTTAAGACTGA
		BSn203_1R	GCCGGCGCATCAGAAGAGGG
Contig4291	Glyma11g13390	BSn204_1F	TGGGTGATGTTGGTGAATGATGCCG
		BSn204_1R	GGGTGCGATCACAGCTTTGCAG
Contig10485	Glyma11g14300	BSn205_1F	GCCTCTTGCTCTTCCCTCTCTCCC
		BSn205_1R	GCCCAGACGTGGCAGCCATT
Contig2910	Glyma11g26470	BSn206_1F	ACGCAACTCCTCTGTACCCGT
		BSn206_1R	TCCCACGCTGCAAAGCTTCCT
Contig5829	Glyma11g34380	BSn207_1F	GGCACACGTGGAATCGGGCA
		BSn207_1R	AGCTGCTGGCAAGCAAAGGA
CV540542	Glyma11g37620	BSn208_1F	TGCACATTGCTGGTGGCCCT

CV541683	Glyma12g01970	BSn208_1R	TCCCGTTGACCTCCAATTCCAGT
		BSn209_1F	GAGCTGCACGTCCCAGCCAG
FE710511	Glyma12g33070	BSn209_1R	TGGCCTAAGTTGCTTCTTTTGAAGGT
		BSn210_1F	TTCTCGGGCTCCGAGCTTGC
		BSn210_1R	GCAGCCACTCACTTGACCCTGC
Contig2079	Glyma12g34510	BSn211_1F	TCCCTCTTCCCCAAAACACCAACT
		BSn211_1R	CTAACCCAGCATGCCCCGC
		BSn212_2F	TACCTCGACCCCAACCGCCA
		BSn212_2R	ACAGGACCCACCCCTCTGCC
Contig5752	Glyma13g01180	BSn213_1F	ACCGAAGGTGGGCAGAGTGC
		BSn213_1R	TCGCCAGGTGCATTGAGGC
		BSn214_2F	CTGGCAGCTGTGCAGGGCAT
		BSn214_2R	TCCTCACCTTACGTTGGATGCT
Contig7984	Glyma13g17420	BSn215_1F	GGGCACGTAACGGGGAGCTG
		BSn215_1R	AGCATGTAGTCTTACACCCAACAGTGA
EX304735	Glyma13g17570	BSn216_1F	CTTCTGCAATCGAGTGGAGTGAGTC
		BSn216_1R	GCAGTGGCGTCGCCGAAGAT
Contig9258	Glyma13g23770	BSn217_1F	CACACATTCACCATGACCACAACCA
		BSn217_1R	TGCAAAGGCCAGAAACGCTACGA
CV535133	Glyma13g27300	BSn218_1F	TCACTCTTGAATCCCAAATCACTCCT
		BSn218_1R	TGAAGCAAATGATCCTCTGGCACAA
Contig4325	Glyma13g32310	BSn219_1F	CATTCACCAACACCTTCTGGTTTCCA
		BSn219_1R	GGCTGCCTTGAGCCATCCCC
Contig8880	Glyma13g38710	BSn220_1F	GCTCCCCCATCCGTTTTCGCT
		BSn220_1R	TGCACACACAAATCATTGCGCGTC
FE709707	Glyma13g40240	BSn221_1F	TGTTGGGGACGTAACTAAACGCA
		BSn221_1R	ACGCAATGCGACAACCTCCCCT
Contig8703	Glyma13g40400	BSn222_1F	ACCATTGATCCTGCAAAAGAAGCAGG
		BSn222_1R	GCTGCAAAACGCCTGCACCA
Contig8921	Glyma13g44060	BSn223_1F	TGCCTTCTCATCACATGCAACACCA
		BSn223_1R	TCAAACCATAGAAACCACCCTCAGGA
Contig10203	Glyma13g44870	BSn224_1F	TCCTGCTGGAAGCCAGATTGC
		BSn224_1R	ACCACCAGACATACTCTGCTGGCTA
Contig6500	Glyma14g02650	BSn225_1F	AGCAGAGTGTGCTACTGTACCCCA

		BSn225_1R	CCGGCAACCCCGGATCATCT
Contig9251	Glyma14g05690	BSn226_1F	TCAGAGGTGGCAGCCAGGAA
		BSn226_1R	AGTGTGCAGCTTCAGTTTTGTCTGA
EX304023	Glyma01g42560	BSn227_1F	GGCCTCAAACACTGGAGGCCTAT
		BSn227_1R	GCACCCAAACCAGGCCTAGCA
Contig897	Glyma01g45280	BSn228_1F	TCCTTCCCTTCTGCTCTGGCA
		BSn228_1R	ACCCATCACATCATACTCTCAGGA
FE686593	Glyma02g04850	BSn229_1F	TGGGTCCCACATGTCTTCGT
		BSn229_1R	GGACGGCCACAAGAGGGTCCA
CV542419	Glyma02g13800	BSn230_1F	GCATGAGAGCGTGGGGTCACA
		BSn230_1R	TGTTGCCGATAGGACTGGTCTGGA
Contig9251	Glyma2g43320	BSn231_1F	TCAGAGGTGGCAGCCAGGAA
		BSn231_1R	AGTGTGCAGCTTCAGTTTTGTCTGA
Contig5055	Glyma03g00580	BSn232_1F	CAAGGCCGGTGCCGTCAGAG
		BSn232_1R	TGCTGGCAAAAACCCACCCAGA
Contig8101	Glyma03g30930	BSn233_1F	GCAGCAACATTCTCCAGATTCATCCC
		BSn233_1R	GCACAACACTACGCATTGATGCTCCA
Contig9294	Glyma04g43090	BSn234_1F	TCACTTCCAACCTCCTCACCCCC
		BSn234_1R	TTGGACGGCCTCAGCGCAAAG
		BSn235_2F	CCGAATATGACCTCAGAGCCATTCC
		BSn235_2R	CGCGGAATTCGAACTGGACGC
Contig3468	Glyma05g03750	BSn236_1F	TGGATTCGAACCTGGTGAACGGT
		BSn236_1R	CCGAAAACACCAGCCGGCGA
Contig8367	Glyma05g12100	BSn237_1F	TGCAACAGAGTATTGTCCAGTGAGCC
		BSn237_1R	CTTCAATAGAAGCCAAGCCCCCTTT
CB530136	Glyma05g15700	BSn238_1F	GCCACCTCGCAGGTTGTCCC
		BSn238_1R	CTGTCCCGAGCCATCACCGC
CB542797	Glyma05g25920	BSn239_1F	TGAACCAGCTAAGGAAACAGGGAAACT
		BSn239_1R	TCCCAGTTTTGGGGTGGTTCACT
Contig485	Glyma05g26810	BSn240_1F	TGCGATGTTGGTAGCGGACGA
		BSn240_1R	TGCCGAATTGACAAGAATTAGCCCCA
CB539909	Glyma05g26830	BSn241_1F	TCTCTTGATTGCCCCAAAGGGACT
		BSn241_1R	CCTGATCTCTGCGCCCATCGC
CB540361	Glyma05g29440	BSn242_1F	GCACCTGCAGCAGCCCCAAA

Contig2330	Glyma06g08950	BSn242_1R	CCACTCTTGGCTGTTATGAGCTGTGA
		BSn243_1F	CGGCACCGCCGACATGGAAA
		BSn243_1R	GGCAGGTTCTTGAGGTGATGGCA
Contig1460	Glyma06g17640	BSn244_1F	GGAAAATTTGTTTCAGGATGGCCTTGTC
		BSn244_1R	GGCCTCCACTCAGTCCTACCCC
Contig9362	Glyma06g19620	BSn245_1F	GGCTCCTCGGACAGTCACATGC
		BSn245_1R	TGGGAGGTCCACCCCTTCGT
Contig3619	Glyma07g08290	BSn246_1F	TGCCTCATCTCTCTCATCACTTGAGAC
		BSn246_1R	GCATCTCTGCCGATGTTCCCCG
Contig2170	Glyma07g33090	BSn247_1F	GCACTTGGCAAAGAAGTGAATGATAGC
		BSn247_1R	TGACACTACTTCAGGTTGACGTGTTGA
Contig5523	Glyma07g35770	BSn248_1F	CGCCGAACAGGACATTTTTCAACTGG
		BSn248_1R	ACACGAGGCCGCCTTTCTGC
		BSn249_2F	AACGGCTGGCAGAAAGGCGG
		BSn249_2R	TCCAGAGGCTTTGTTTGCTCTCCA
Contig9307	Glyma08g08170	BSn250_1F	CCCTCAACCGTGGGATGAGCC
		BSn250_1R	TGGCACCAATCACACCTCCAAGA
Contig6727	Glyma08g12650	BSn251_1F	TGGCATTGCCTTCGGACGCT
		BSn251_1R	GGCCAGCCAAACCACCACCA
Contig54	Glyma08g18030	BSn252_1F	CGTGAAAGCACTCACTGAATCACCA
		BSn252_1R	GCCGGTTCACCACAGTGTCC
Contig5420	Glyma08g18080	BSn253_1F	TCCCACAGATGGCTTCACCCCT
		BSn253_1R	ACACTGCTTCTCCCAGTCCAAGT
Contig9569	Glyma08g24680	BSn254_1F	TCTGCCATGGCTTCACCAGA
		BSn254_1R	TGCGGCACCAGGTAACAAGC
Contig7488	Glyma08g48140	BSn255_1F	ACGGAGAGGGATTCTTTGAATTGGGA
		BSn255_1R	AGCATTGTGAGTCCACTTCCACCA
CB537278	Glyma09g00800	BSn256_1F	GGTGTGGTGTGGCGTGGCTT
		BSn256_1R	ACGAACTTGTTCGAGCCCCTTGC
FE899204	Glyma10g05450	BSn257_1F	TGGGGCCTTAACATTCTTGCGT
		BSn257_1R	CGGAATGCGCAACCGGACCA
Contig8894	Glyma10g11060	BSn258_1F	TTCTCGCGTCGCTCTGACGG
		BSn258_1R	ACTGCAACAGTGATTACTGGCGATGA
CB530305	Glyma10g34060	BSn259_1F	ACGCGTCCGGAGGACGTTTG

FE689976	Glyma10g4284	BSn259_1R	TCTGACCCTGCAACCTCCATTGT
		BSn260_1F	TCGGTTGGTTCCCTGTCTTCGGA
		BSn260_1R	AGTACCAGGGAATGCAGCCCAC
CB511230	Glyma11g03940	BSn261_1F	ACTTGGGAGGCCCGCTTGAA
		BSn261_1R	CGCTTCAAAGCACGCAGCGG
Contig4875	Glyma11g04620	BSn262_1F	CATGGCTCAGCCGCCTACGG
		BSn262_1R	CGGCGTTGCCTGATTCTCGC
Contig4686	Glyma11g22090	BSn263_1F	AGGCCACTGTGGTCGCCATT
		BSn263_1R	TCTCGGCTCGGCCTGCTCTT
		BSn264_2F	GCTGCGAGGGTCTGGCTTT
		BSn264_2R	TGGACGCACAACATCACGCT
Contig9624	Glyma11g34500	BSn265_1F	AGGGTTTCAGCTCTCACCAACTGT
		BSn265_1R	TTGGAGGCCCCAGAGGCAGCA
Contig5358	Glyma11g34510	BSn266_1F	TCGTTCCCTTCGATTTCGTTCCATTTCTT
		BSn266_1R	CCTCAGACAGCTCTTCTCCTGTCA
FE709391	Glyma12g02590	BSn267_1F	GCAACAACAAATCATGTTGCCCTCT
		BSn267_1R	AGGGACTGAGGAGGGGATGGC
CB539777	Glyma12g30050	BSn268_1F	CGCATGTGGAATGACCTTTGGGA
		BSn268_1R	GGAGGCATCAAACATGTGGGGGT
CV542814	Glyma12g36360	BSn269_1F	ACCCACCATGCCCTAAGCCTGA
		BSn269_1R	AGAATTGCGTCATGGATGTTGCCT
Contig4275	Glyma13g22350	BSn270_1F	CACCAGCCCAAGACTACAAGCAA
		BSn270_1R	TGACCAACGACCAACCAACCTTT
Contig8433	Glyma13g24560	BSn271_1F	TAGGGCCAGTGCCGTGAGCA
		BSn271_1R	GTGCAATTGGGCGGAGGCCGA
Contig6337	Glyma13g27640	BSn272_1F	TGGTTCTGGAAGTGGTTTTTCCTTCC
		BSn272_1R	TCACAGGGGAAAGCTATGCAGGT
Contig9456	Glyma13g44100	BSn273_1F	ACCATTGATCCTGCAAAAGAAGCAGG
		BSn273_1R	TGTGCACTGACTTGAGAGCAGGT
Contig2911	Glyma14g01160	BSn274_1F	ACCCAGCACATTTCTGTCCCG
		BSn274_1R	TGGGGTCAAAGCAAACCTACCCA
Contig1778	Glyma14g06650	BSn275_1F	GCGAACCGAGCGCCGAATGA
		BSn275_1R	GAAAGACGCCGTTTCGCCGC
Contig2131	Glyma14g06530	BSn276_1F	ACCCTCGGTCTCCCCTTCGC

Contig9423	Glyma14g23780	BSn276_1R	AACCAGGGCACAACCGTGGC
		BSn277_1F	TCCTTTGCCCTTTGCCGACTCA
		BSn277_1R	TGGCTTTGCACAACCATAGGCA
Contig6889	Glyma15g00620	BSn278_1F	CCAAGTCCAAGCACCCCGC
		BSn278_1R	ACATGGCTCATCACAACCCTCAAAGT
CV531682	Glyma15g01500	BSn279_1F	GCACAGAGCTGCTCGTTCCCC
		BSn279_1R	GGTGGCCACATAAATAAGCAACCG
Contig8750	Glyma15g05010	BSn280_1F	ACCTCCACCACCTCCACCACC
		BSn280_1R	TGGGTCGGGGTCACGCAAAC
Contig8730	Glyma15g06140	BSn281_1F	CCCCTGCGTTGCACCACCTT
		BSn281_1R	AGGAGCAAGAGCCTCCAAGTGT
Contig504	Glyma15g15610	BSn282_1F	ACTCCTCCACATTGGCTCCAGCA
		BSn282_1R	TCGACATGAACACCAACAAGAGACG
Contig2966	Glyma16g01020	BSn283_1F	AGGCAGAGATACCGTGAAGCCT
		BSn283_1R	TCTTTGCACACAACACATTGAGAATGC
Contig3292	Glyma16g03120	BSn284_1F	CCTTGCAATGGCCCGGCACT
		BSn284_1R	TGGCAACAGGCTTGGGACAGC
Contig6453	Glyma16g04260	BSn285_1F	TCGGATCCTCACCCGCCTCG
		BSn285_1R	CCCTCACTTCTCTCGGATCACA
Contig550	Glyma16g06500	BSn286_1F	AGCTGCGTACTTGCTTGTTGCT
		BSn286_1R	ACCGACACCACAAGTTCTTTCTCCA
CV530976	Glyma16g06740	BSn287_1F	CCACGCGTCCGGTCTCCCTT
		BSn287_1R	TGTGAGCAGAAGCTGCCAGG
Contig8733	Glyma16g21620	BSn288_1F	CACCGGTGGCGCGACAAGAT
		BSn288_1R	TGGTCCAGAGGGGCTCACCA
Contig3228	Glyma16g26290	BSn289_1F	GCGCCGCAGGAAGAAGCTCA
		BSn289_1R	TGGCCTAAACAGGAAGGCCTTTACA
Contig3428	Glyma17g03340	BSn290_1F	GCAGTGGCACCTGCAAGACT
		BSn290_1R	AGGCAACAACACAGACAACTCGT
CV532819	Glyma17g06220	BSn291_1F	GCCTCAAAGCTTAGTCGTGACCCT
		BSn291_1R	TGCAGGTCTAGAAAGTATTCTTGCTCT
DN153793	Glyma17g07330	BSn292_1F	GCTCCTTGTTGTGACAAAGCCAACG
		BSn292_1R	TGCTTCCAATACTGACATAAAGGCTGC
Contig1740	Glyma17g07830	BSn293_1F	ACGTTTCTCTCAACGCCTCTCCC

Contig8828	Glyma17g08110	BSn293_1R	TACCCTGTCAACCAGCGCGT
		BSn294_1F	TGTGGTTGGAGGCACTGCAAAT
		BSn294_1R	ACCAGGCATTGAAGCTAAAGGGC
Contig2308	Glyma17g14260	BSn295_1F	AACCACAGTGGGCGCTGGGA
		BSn295_1R	CAGCCGCGTTTGCAACAGGT
Contig4485	Glyma17g27490	BSn296_1F	TCAGACCTTGAACCAATGCACAAC
		BSn296_1R	GCATCCCAAGGATTTGTCTCCATCA
Contig1296	Glyma17g32050	BSn297_1F	TGATGCTCCTTCAGCTGCTCCT
		BSn297_1R	TCGGACCATGTGCGCCACCT
Contig9899	Glyma17g35230	BSn298_1F	CCCACGCGTCCGTAGAACC
		BSn298_1R	AGTCTTCCGCCGTGAAGGAGC
FD793890	Glyma18g00350	BSn299_1F	GCCATTTAAGGACATGGTTGAAGGGA
		BSn299_1R	AGCTGCAGGCAGTGACAAGAGC
Contig9001	Glyma18g02230	BSn300_1F	GCACTTCTCTGAGAGCACTAGGTTCC
		BSn300_1R	AGTCCAATGCCCACAGCAGC
CV544224	Glyma18g15530	BSn301_1F	AGTGTGGCACCAAAGAGGCCA
		BSn301_1R	CTTCAAACCTCCCACAACCCATCACA
Contig1668	Glyma18g44870	BSn302_1F	AGTGGGGTGCAAGGGGCTGA
		BSn302_1R	TCCCTCTCAAACACGGTTGGCA
Contig3701	Glyma18g49360	BSn303_1F	AGGCAAGATGGGAAGGCCACCT
		BSn303_1R	GGGTTGAAGGCCAAAGTGAGATTCAAG
FE691655	Glyma18g50180	BSn304_1F	CGTTGGAAGGGCCGCAAGGT
		BSn304_1R	AGTCTGCACCGGCCACCAT
Contig4021	Glyma18g52730	BSn305_1F	TGCTGTGCTTCTATTTCTGCAACA
		BSn305_1R	AAGCACTGAAGCAAACCTCAGACACA
Contig10472	Glyma19g22460	BSn306_1F	CCACTCCCTCCGCCGTCTCA
		BSn306_1R	GGCCAGTGTTTTGCGGGGGT
Contig6842	Glyma19g34380	BSn307_1F	CGTAACAACGTGCCTGCCCT
		BSn307_1R	GGGCACCACTAACGATGCTCGAA
Contig6670	Glyma19g35270	BSn308_1F	AGGCCTTGGGGGACTGATTGGA
		BSn308_1R	ACTCTGAATTTTTGTACACCTCGGCA
Contig5253	Glyma19g42790	BSn309_1F	AGCTTCGAATGGAAATCCACGCCT
		BSn309_1R	TCATCTAGGAAAACCCAAAAGGACACG
Contig8065	Glyma19g43370	BSn310_1F	ATGGCCGACCAGCTCACCGA

Contig2158	Glyma19g44060	BSn310_1R	ACAGCATTAGCCACCCACATGTAAGA
		BSn311_1F	TGGCAGTGCTGCTGAGAGTTCC
		BSn311_1R	GGTCCCACTACAATGGGGCTCCT
Contig2728	Glyma19g45310	BSn312_1F	GCTGACTCTAGATCTCTTGCCACTCT
		BSn312_1R	GTCCAAAACCACCTGCATCAATATCCT
CV537589	Glyma20g21860	BSn313_1F	TCTCCGCGAGTCCTGAGACCT
		BSn313_1R	GCGATCACACGGTGAGTCTCAAACCA

ANEXO 4. Protocolo para envío de productos de PCR a secuenciar en la *Universidad de Yale*

PCR Product:

- **10-20** ng template for every 200 bases of PCR fragment length
- 2 μ l **4 μ M** primer
- x μ l sterile water
- **18 μ l total volume**

e.g. 500bp PCR fragment, mix 25-50 ng PCR product with 2 μ l 4 μ M primer and sufficient water to bring the final volume to 18 μ l.

Tomado de: <http://medicine.yale.edu/keck/dna/protocols/tube/index.aspx>

ANEXO 5. Segregación mendeliana de los 49 marcadores mapeados en la población proveniente de la cruce DOR364 x BAT477.

NOMBRE DEL MARCADOR	Prueba χ^2 ($p < 0.05$)	SEGREGACIÓN
BMb1010	0,346	Segregación 1:1
BMb10	0,128	Segregación 1:1
BMb1008	0,244	Segregación 1:1
BMb1074	0,833	Segregación 1:1
BMb1119	0,169	Segregación 1:1
BMb1137	0,244	Segregación 1:1
BMb1142	0,278	Segregación 1:1
BMb1157	0,157	Segregación 1:1
BMb116	0,256	Segregación 1:1
BMb133	0,008	Sesgado hacia DOR364
BMb136	0,007	Sesgado hacia DOR364
BMb1519	0,833	Segregación 1:1
BMb152	0,233	Segregación 1:1
BMb1527	0,385	Segregación 1:1
BMb1536	0,317	Segregación 1:1
BMb1543	0,881	Segregación 1:1
BMb1605	0,673	Segregación 1:1
BMb1843	0,014	Sesgado hacia BAT477
BMb1939	0,058	Segregación 1:1
BMb259	0,447	Segregación 1:1
BMb263	0,180	Segregación 1:1
BMb278	0,637	Segregación 1:1
BMb342	0,655	Segregación 1:1
BMb353	0,233	Segregación 1:1
BMb354	0,089	Segregación 1:1

BMb365	0,023	Sesgado hacia DOR364
BMb37	0,206	Segregación 1:1
BMb4	0,108	Segregación 1:1
BMb405	0,003	Sesgado hacia DOR364
BMb420	0,082	Segregación 1:1
BMb555	0,739	Segregación 1:1
BMb578	0,447	Segregación 1:1
BMb638	0,297	Segregación 1:1
BMb643	0,002	Sesgado hacia DOR364
BMb653	0,037	Sesgado hacia DOR364
BMb654	0,346	Segregación 1:1
BMb712	0,365	Segregación 1:1
BMb717	0,092	Segregación 1:1
BMb737	0,027	Sesgado hacia BAT477
BMb742	0,881	Segregación 1:1
BMb793	0,539	Segregación 1:1
BMb803	0,256	Segregación 1:1
BMb89	0,739	Segregación 1:1
BMb96	0,059	Segregación 1:1
BSn160	0,833	Segregación 1:1
BSn165	0,244	Segregación 1:1
Leg447	0,739	Segregación 1:1
Leg726	0,136	Segregación 1:1
PvM127	0,065	Segregación 1:1

ANEXO 6. Segregación mendeliana de los 32 marcadores mapeados en la población proveniente de la cruz DOR364 x G19833

NOMBRE DEL MARCADOR	Prueba X^2 ($p < 0.05$)	SEGREGACIÓN
Leg 376	0,633	Segregación 1:1
Leg 736	0,391	Segregación 1:1
Leg 81	0,359	Segregación 1:1
Leg 443	0,084	Segregación 1:1
Leg 310	0,032	Sesgado hacia G19833
Leg 133	0,250	Segregación 1:1
Leg 83	0,056	Segregación 1:1
Leg 177	0,035	Sesgado hacia DOR364
Leg 188	0,029	Sesgado hacia G19833
Leg 299	0,018	Sesgado hacia G19833
Leg 301	0,117	Segregación 1:1
Leg 726	0,028	Sesgado hacia G19833
Leg 213	0,002	Sesgado hacia G19833
Leg 97	0,284	Segregación 1:1
Leg 56	0,133	Segregación 1:1
Leg 128	0,010	Sesgado hacia DOR364
Leg 164	0,760	Segregación 1:1

Leg 189	0,520	Segregación 1:1
Leg 237	0,041	Sesgado hacia DOR364
TF11 (60)	0,018	Sesgado hacia G19833
F (34)	0,013	Sesgado hacia G19833
NIN (45)	0,133	Segregación 1:1
BSn157	0,086	Segregación 1:1
BSn116	0,100	Segregación 1:1
BSn154	1	Segregación 1:1
BSn152	0,033	Sesgado hacia G19833
BSn223	0,633	Segregación 1:1
BSn229	0,542	Segregación 1:1
BSn176	0,018	Sesgado hacia G19833
BSn191	0,172	Segregación 1:1
BSn262	0,152	Segregación 1:1
BSn199	0,005	Sesgado hacia G19833

ANEXO 7. Marcadores mapeados compartidos y exclusivos en cada uno de los dos mapas empleados en este trabajo

MARCADOR	GL	PRESENTE EN AMBOS MAPAS	PRESENTE EN DB	PRESENTE EN DG
BMb405	1	x		
BMb712	2	x		
BMb420	2	x		
BMb365	2	x		
BMb259	2	x		
BMb1137	2	x		
BMb4	3	x		
BMb37	3	x		
BMb1010	3	x		
BMb353	4	x		
BMb136	4	x		
BMb133	4	x		
BMb354	4	x		
BMb742	5	x		
BMb342	6	x		
BMb1157	6	x		
BMb1536	7	x		
BMb1142	7	x		
Leg726	7	x		
BMb89	8	x		
BMb1543	8	x		
BMb578	8	x		
BMb1119	9	x		
BMb263	10	x		
BMb152	10	x		
BMb96	10	x		
BMb10	11	x		
BMb654	11	x		
BMb653	11	x		
BMb1074	11	x		
BMb555	1		x	
BMb1939	2		x	
BMb638	2		x	
BMb1605	2		x	
Leg447	2		x	

BSn160	2	x	
BMb793	3	x	
BMb1008	3	x	
BMb717	4	x	
BMb1527	4	x	
BMb1843	5	x	
BMb803	5	x	
BMb643	5	x	
BMb278	8	x	
BMb116	8	x	
BSn165	8	x	
BMb1519	9	x	
BMb737	9	x	
PvM127	10	x	
BSn199	1		x
BSn152	2		x
BSn157	2		x
BSn176	2		x
Leg188	2		x
Leg301	2		x
BSn116	3		x
BSn229	3		x
Leg128	3		x
Leg213	3		x
Leg237	3		x
Leg310	3		x
Leg83	3		x
BSn154	4		x
BSn191	4		x
BSn262	4		x
BSn223	5		x
Leg56	6		x
Leg736	6		x
Leg81	6		x
Leg376	7		x
Leg97	7		x
F	8		x
Leg299	8		x
Leg443	8		x
NIN	8		x
TF11	8		x
Leg189	9		x
Leg164	10		x
Leg177	10		x
Leg133	11		x

ANEXO 8. Marcadores secuenciados con el peso obtenido luego de editar las secuencias

MARCADOR	PESO DE LA SECUENCIA (PB)
BSn112	495
*BSn115	239
BSn120	936
BSn122	428
BSn124	674
*BSn131	542
BSn132	148
BSn136	656
BSn138	595
BSn139	609
BSn140	809
BSn145	873
BSn150	530

BSn152	730
BSn154	540
BSn157	618
BSn158	457
BSn161	450
*BSn165	566
BSn176	674
BSn177	369
BSn187	549
BSn191	688
BSn192	542
*BSn195	437
BSn204	419
BSn209	669
BSn211	211
*BSn221	538
*BSn229	434
*BSn233	680
BSn234	756
BSn235	759
BSn238	427
BSn241	824
BSn242	213
BSn247	946
BSn248	734
BSn249	696
BSn252	979
BSn255	639
BSn258	858
BSn261	972
BSn269	492

*No hubo presencia de SNP

ANEXO 9. Socialización de los resultados obtenidos en esta investigación.

Evento:

COMPARATIVE & REGULATORY GENOMICS IN PLANTS CONFERENCE

Period [11/04/2011 - 15/04/2011]
Event location Karel Lodewijk Ledeganckstraat VIB - Department of Plant Systems Biology, Ghent University
Host VIB University of Gent, Laboratory of Tropical Crop Improvement, K.U.Leuven
See the agenda http://bioinformatics.psb.ugent.be/crg_event/conference/programme
Event URL http://bioinformatics.psb.ugent.be/crg_event/
Release date 10/02/2011

The two-day conference (11-12 April 2011) will discuss recent advances in plant genomics. It will be followed by a workshop combining lectures with practical sessions covering data integration, comparative genomics and the biological interpretation of large-scale omics datasets (13-15 April 2011).

Certificado:



Department of Plant Systems Biology

Ghent, 8 March 2011

Subject: conference 'Comparative & Regulatory Genomics in Plants' in Ghent (Belgium) April 2011



Dear Carlos H. Galeano M.,

We are pleased to acknowledge the receipt of your registration for the **Comparative and Regulatory Genomics in Plants** event (April 2011, Ghent, Belgium) and confirm that we have received your payment for this event.



Registration type: Conference & Workshop (Academic)
Code: ABAW3C7L
Amount: 370euro



We also confirm that Carlos H. Galeano M. has contributed to the Poster session 'A new intra-gene pool linkage map: toward unified consensus linkage map in common bean.'

Best regards,

Prof. Dr. Klaas Vandepoele
conference chair

A handwritten signature in black ink, appearing to read 'Klaas Vandepoele'.

VIB

Department of Plant
Systems Biology

UGent

UGent-VIB Research
Building FSVM
Technologiepark 927
9052 GENT
BELGIUM
Tel.: +32 9 331 38 00
Fax: +32 9 331 38 09
contact@psb.ugent.be

www.vib.be
www.psb.ugent.be

A new intra-gene pool linkage map: toward unified consensus linkage map in common bean.



Carlos H Galeano¹, Andrea Fernández², Natalia Franco², Carolina Chavarro², Amina Sánchez-Rodríguez¹, Matthew W Blair², Steve Beebe², and Jos Vanderleyden¹.

¹Centre of Microbial and Plant Genetics, Department of Microbial and Molecular Systems, K.U. Leuven, Kasteelpark Arenberg 20, 3001 Heverlee, Belgium, ²CIAT-International Center for Tropical Agriculture, Biotechnology Unit and Bean Project, AA6713, Cali, Valle, Colombia



Introduction

Despite efforts to build linkage maps in common bean, the interest to develop an integrated consensus map between different mapping populations has been limited. The construction of a consensus map combining the information of multiple segregating populations, of different genetic backgrounds offers the opportunity to map a larger number of loci than in most single crosses, thus increasing the number of potentially useful markers across divergent genetic backgrounds and providing greater genome coverage, in addition to providing opportunities to validate marker order¹. This approach allows that markers, genes or QTL that could not be mapped in a particular study may be located on the basis of their relative position in other studies². The goal of our research was to increase the marker saturation of the Mesoamerican population derived of the cross BAT477×DOR364 (DB) by using the consensus strategy with inter-gene pool linkage map of DOR364×G19833 (DG) and BAT93×JALO EEP558 (BJ) populations. Additionally, the synteny analysis was developed against the model legumes and allowed confirming the consensus map order. This synteny platform is a useful tool for *in-silico* gene mapping, as well as a source of new markers for fine mapping projects and positional cloning.

Results and Discussion

A total of 2,706 markers were evaluated between the parents DOR364 and BAT477, including 1,136 genomic SSR, 866 genic SSR and 704 gene based markers. On average, the polymorphism rate was 7.9 %. On average, the polymorphism rate for DB population was 3.7 % for genic and 12.3 % for genomic SSR. The new set of gene based markers presents the lowest polymorphism rate of 1.6 %. The low polymorphism reported here can be explained by the fact that the genotypes DOR364 and BAT477 both belong to race Mesoamerica within the Mesoamerican gene pool.

The segregation data was integrated with our previous map version³. A total of 294 markers were located in the linkage map with total full map length of 1,804 cM. (Fig. 1) Linkage group size ranged from 81 cM (b09k) to 272 cM (b04b) with an average of 164 cM per linkage group. In general, the marker loci were well distributed within the linkage groups with an average 26 markers per linkage group.

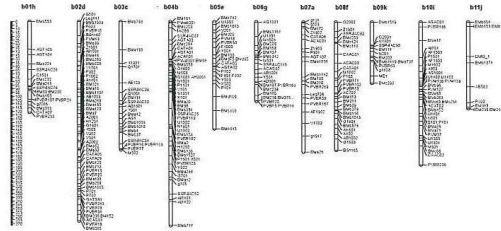


Figure 1. Linkage map of population BAT477×DOR364

The consensus map between the populations DB, BJ and DG was developed following the graph theoretic model, where the maps were modeled as a directed acyclic graph (DAG)⁴. A diagram of the shared markers and their order in consensus map are showed in fig. 2. A total of 100 and 51 common anchor markers were shared between DG-DB and DG-BJ, respectively. A total of 917 markers were placed in the consensus map, including 284 SNP, 416 SSR and 43 RFLP. The average number of markers per linkage group are 83, with a maximum of 142 on linkage group b02d and a minimum of 52 on the linkage group b09k. The total full map length was 3000 cM and ranged from 199 cM (b06g) to 440 cM (b01h) with an average of 290 cM per linkage group. Moreover, even though marker order among the four maps (consensus, DB, DG and BJ) was reliable, some small differences were observed between the consensus and complementary maps.

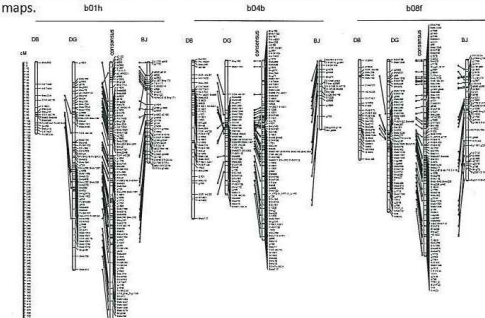


Figure 2. Comparison of marker order (consensus and single linkage map) in three linkage groups.

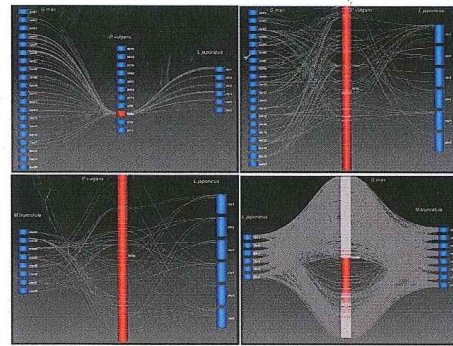


Figure 3. Synteny relationship between model legumes and common bean consensus map.

Our synteny analysis allowed comparing homologies in genomes (genomic and gene level) and the consensus map (Fig 3). Relationships between soybean with *Lotus* and *Medicago* appear to be rather complex suggesting that considerable genome shuffling occurred after separation of the species (54 mya). The relationship of soybean with common bean is closer (19 mya). However, duplicate chromosome regions in soybean are frequent because soybean has undergone two or more large scale duplications and is probably an ancient polyploid⁵.

The current common bean consensus map provides the means to anchor maps across different pedigrees, to establish linkage with genes for agronomic traits and to compare QTL of interest in breeding (QTL meta-analysis). Additionally, the inter-species synteny relationship is fundamental for transferring information between common bean and model legumes.

Material and methods

Plant material
The population DOR364×BAT477 consists of 113 F5:7 recombinant inbred lines (RILs). The generation of this population is described by Blair et al. (2010). For mapping saturation the first 92 lines were selected.

Gene based markers
A set of 200 nodulation genes or expressed sequences reported in model legumes were compared with common bean EST assembly. The contigs selected were aligned with soybean genomes to find the intron sequences. The conserved fragments flanking introns were selected for primer design.

Map saturation
The parental survey was evaluated with 1617 markers published or in preparation for publication in common bean, including genic and genomic SSR, SSR based in EST libraries and BAC-end sequences; and gene based markers. The SSR and gene based marker PCR conditions, MgCl₂ concentration, annealing temperature and electrophoresis were as described [6 and 7, respectively]. The new gene based markers were also evaluated in the core mapping population DOR364 × G19833.

Linkage analysis
Segregation data was used to place the new markers on the previous genetic map for the DOR364 × BAT477 population⁴. Linkage analysis was conducted with the Kosambi mapping function using the software MapDisto 1.7.5⁶.

Consensus map
The core mapping populations in common bean derived of inter-gene pool crosses DOR364 × G19833 (hereafter named DG) and BAT93 × JALO EEP58 (hereafter named BJ)⁷ were used to build a consensus map with the less saturated DOR364 × BAT477 population (hereafter named DB). The DG linkage map has been developed by CIAT Bean Project. The consensus map was constructed with MergeMap⁸. Comparative mapping and verification of consistency of marker order and distances between maps were done using MapChart¹¹.

Synteny analysis
Marker sequences from the consensus map were compared with the genome of soybean, *Medicago* and *Lotus* [7]. Additional comparison was done using the cDNA information of the model legumes, and drawing with Strudel 1.11.02.11 software¹².

Acknowledgements

We thank the Bean Genetics program (CIAT), especially the field workers, students and research assistants for their contribution in the construction of the DG and BJ linkage maps. Bean photos - Neil Palmer (CIAT).
CHG is supported by a doctoral research fellowship from IRO (Interfaculty Council for Development Co-operation of the Catholic University of Leuven)

References

1. E. Mace et al., BMC Plant Biology 9, 13 (2009).
2. L. V. Yap et al., Genetics 165, 2315 (2003).
3. M. W. Blair et al., Molecular Breeding, (2010)
4. Y. Wu, T. J. Close, S. Lonardi, Comput Syst Bioinformatics Conf (2008).
5. R. C. Shoemaker et al., Genetics 164, 329 (1996)
6. M.W Blair et al., Theor Appl Genet 107, 1362 (2003)
7. C. H. Galeano et al., BMC Genomics 10, (2009).
8. Lorraine M P&G XV conference (2007)
9. M. McConnell et al., Theor Appl Genet (2010)
10. R. Freyre et al., Theor Appl Genet 97, 847 (1998).
11. R. E. Voorrips, Journal of Heredity 95, 77 (2004)
12. Bayer et al., Bioinformatics (2011) .

