

## Tesis Doctoral

# Loci de caracteres cuantitativos (QTLs) de la soja (*Glycine max* L. Merr) asociados a su capacidad de nodular y fijar nitrógeno

Salvucci, Rubén Darío

2015-06-26

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en [digital.bl.fcen.uba.ar](http://digital.bl.fcen.uba.ar). Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in [digital.bl.fcen.uba.ar](http://digital.bl.fcen.uba.ar). It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Salvucci, Rubén Darío. (2015-06-26). Loci de caracteres cuantitativos (QTLs) de la soja (*Glycine max* L. Merr) asociados a su capacidad de nodular y fijar nitrógeno. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

Cita tipo Chicago:

Salvucci, Rubén Darío. "Loci de caracteres cuantitativos (QTLs) de la soja (*Glycine max* L. Merr) asociados a su capacidad de nodular y fijar nitrógeno". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2015-06-26.

**EXACTAS** UBA

Facultad de Ciencias Exactas y Naturales



**UBA**

Universidad de Buenos Aires



**Universidad de Buenos Aires  
Facultad de Ciencias Exactas y Naturales**

**Loci de caracteres cuantitativos (QTLs) de la soja  
(*Glycine max* L. Merr) asociados a su capacidad de  
nodular y fijar nitrógeno**

**Tesis para optar por el título de Doctor de la Universidad de  
Buenos Aires en el área: CIENCIAS BIOLÓGICAS**

**Rubén Darío Salvucci**

Director de Tesis: **Dr. Pedro A. Balatti**

Directora Asistente: **Dra. Mónica B. Aulicino**

Consejero de Estudios: **Dra. Viviana Confalonieri**

Lugar de trabajo: **Instituto de Fisiología Vegetal - CONICET – UNLP, Instituto  
Fitotécnico Santa Catalina - UNLP**

Buenos Aires, 2015

Fecha de defensa: 26/06/2015







## Resumen

### Loci de caracteres cuantitativos (QTLs) de la soja (*Glycine max* L. Merr) asociados a su capacidad de nodular y fijar nitrógeno

La soja (*Glycine max* L. Merr) es la leguminosa más importante cultivada a nivel mundial y en Argentina que, junto con Estados Unidos y Brasil, son los países que lideran la producción y el comercio de esta oleaginosa. Para obtener materiales con altos rendimientos en las producciones agrícolas, el N se incorpora por medio de la fertilización química y/o biológica. Esta última tiene ventajas por sobre la fertilización química ya que presenta menores riesgos para la contaminación del medio ambiente y contribuye a un manejo sustentable. Ésta consiste en adicionar bacterias del suelo Gram (-) del género *Bradyrhizobium* que tienen la capacidad de interactuar con las raíces de la planta, como resultado de lo cual, desarrollan estructuras denominados nódulos, en los cuales la bacteria fija nitrógeno. La soja usa casi todo el nitrógeno fijado en la producción de granos. Las compañías de inoculantes se han focalizado en producir formulados con cepas de bacterias seleccionadas que fijen más eficientemente el N y que puedan ser utilizadas en una amplia variedad de leguminosas. Sin embargo, el foco del mejoramiento en leguminosas ha sido principalmente el rendimiento y la resistencia a factores abióticos y bióticos. Desde 1950 se han caracterizado varios genes involucrados en el reconocimiento de las señales disparadas por los rizobios y las distintas etapas de la infección y nodulación. Sin embargo, los caracteres capacidad de nodulación y de fijación biológica de nitrógeno (FBN) presentan una arquitectura genética compleja, y además depende de la interacción planta-bacteria y del ambiente. Debido al gran tamaño y complejidad del genoma de la soja, todavía resta identificar genes de la planta involucrados en la nodulación y fijación de nitrógeno. Por lo tanto, en el objetivo del presente trabajo de tesis fue identificar a los loci de caracteres cuantitativos (QTLs) asociados a la capacidad de nodulación de la soja, de manera de disponer de la información y de los marcadores moleculares que permitan mejorar al hospedante en lo que hace a la nodulación y por ende a la fijación de nitrógeno. Para ello se analizó la habilidad de 31 cultivares comerciales argentinos de soja para desarrollar nódulos cuando estos fueron inoculados con una mezcla de las cepas *Bradyrhizobium japonicum* SEMIA 5079 (CPAC 7) y SEMIA 5080 (CPAC 15). Los experimentos se realizaron en condiciones controladas de luz y temperatura. Se utilizaron para evaluar la capacidad de nodulación tres variables cuantitativas: número de nódulos (NN), peso seco de nódulos (PSN) y peso seco de la parte aérea de la

planta (PSA). Los resultados permitieron seleccionar cultivares con alta, media y baja capacidad de nodulación. En un segundo experimento se confirmó la respuesta diferencial para la capacidad de nodulación y en un tercer experimento se comprobó que la capacidad de nodulación no es dependiente de la cepa inoculada ya que los cultivares de alta y baja capacidad de nodulación mostraron el mismo fenotipo cuando fueron inoculados con la estirpe E109. Los cultivares con alta capacidad de nodular tuvieron el doble de nódulos que aquellos con menor capacidad, confirmando resultados de trabajos previos. Además, se demostró la existencia de una variabilidad fenotípica amplia para las variables estudiadas. Los resultados señalaron valores medianos a altos de heredabilidad  $H^2$ , siendo la variable NN la que denotó mayor variabilidad genética. La respuesta de la planta a la inoculación con una concentración determinada de bacterias estuvo condicionada al número de bacterias inoculadas. Un aumento en la concentración de células bacterianas resultó en un incremento del número de nódulos, hasta una concentración umbral, por encima de la cual la nodulación fue inhibida, lo que probablemente esté vinculado al mecanismo de “quorum sensing”. Esto último se observó tanto en los dos cultivares de alta como en los dos de baja capacidad de nodulación. La densidad celular de los rizobios no alteró la respuesta al comparar los cultivares a la inoculación. Luego se procedió a caracterizar la diversidad genética de los cultivares argentinos de alta y baja capacidad de nodulación por medio de SSRs ligados a QTLs asociados a nodulación en cultivares brasileros. Se determinó la asociación de estos marcadores con la capacidad de nodulación y otras características agronómicas de los cultivares de soja y se identificó la generación parental con capacidad de nodulación contrastante para que a partir del cruzamiento entre ellas se obtenga poblaciones segregantes aplicables al análisis de QTLs. Se encontró polimorfismo para 10 marcadores SSR. En los cultivares argentinos se encontró asociación entre la capacidad de nodulación y los marcadores SSR definidos por los cultivares brasileros. Metodologías de análisis multivariados permitieron tomar la decisión de seleccionar un cultivar de alta nodulación (NA 5485 RG) y uno de baja nodulación (A 7053 RG). Éstos fueron los parentales que se utilizaron para generar la población segregante  $F_{2:3}$  como población de mapeo. En la generación  $F_2$  se realizó el genotipado de los individuos a partir de ADN foliar con un “screening” de 221 marcadores SSR a lo largo del genoma de los cuales sólo 36 fueron polimórficos. Se construyó un mapa genético parcial conteniendo 5 grupos de ligamiento. El fenotipado se realizó sobre 94 familias  $F_3$  en un ensayo de nodulación en condiciones controladas de luz y temperatura y se evaluaron los caracteres cuantitativos que definen la capacidad de nodulación: número de nódulos (NN), peso seco de nódulos (PSN) y peso seco de nódulo por nódulo (PSNN) calculado como PSN/NN. También otros caracteres

relacionados con el crecimiento de la planta fueron considerados: la biomasa seca de hojas (PSH) y biomasa seca de tallo (PST) y la suma de ambas, peso seco de la parte aérea (PSA). Se realizó un análisis de correlaciones entre las variables y de coeficientes de paso. A partir de una matriz básica combinada con los datos genotípicos (SSR) y fenotípicos se realizó finalmente el análisis mapeo de QTLs para estos caracteres mediante tres metodologías: por locus simple (ANOVA y regresión), por intervalo (SIM) e intervalo compuesto (CIM). Los resultados encontrados para la mayoría de los QTLs y sus efectos génicos coincidieron en general por los dos primeros métodos. Se pudo confirmar que el QTL Satt414 se asoció a NN y PSN (locus simple). Se concluyó que las variables asociadas a capacidad de nodulación en la población F<sub>3</sub>, presentaron altos valores de varianza genética y heredabilidad. Se localizaron QTLs para las variables estudiadas que explicaron valores mayores al 15% de la varianza fenotípica total de la población segregante. Se observó una segregación transgresiva para la mayoría de las variables asociadas a capacidad de nodulación. Se probó asociación en acoplamiento entre variables relacionadas con la capacidad de nodular y el aumento de la biomasa. Para caracteres de capacidad de nodulación, se localizaron QTLs en los siguientes grupos de ligamiento: GL H para NN; PSN y PSNN; GL D1b para NN; GL A2 para NN; GL J para NN y PSN; GL B2 para PSN y PSNN; GL B1 para PSNN, GL C1 para PSNN. Estos QTLs explicaron entre 12.2-13.4% de la variación fenotípica total. Para caracteres de biomasa, relacionados con la FBN, se localizaron QTLs en los siguientes grupos de ligamiento: GL H para AP y PST; GL M para AP. GL D1b para AP, PST, PSA y PSH; GL B2 para AP; GL N para PSA. GL L para PSA; GL G para PSH. Estos QTLs explicaron entre 8.3-25.6% de la variación fenotípica total. Si bien la metodología aplicada permitió localizar algunos QTLs de nodulación en intervalos. Estas regiones deberían saturarse con otros marcadores de mayor eficiencia y variabilidad como los SNPs. Una mayor saturación conjuntamente con la información disponible sobre el genoma de la soja y los bancos de ESTs permitirán identificar genes candidatos relacionados con el carácter en estudio. Estos resultados aportan herramientas útiles como punto de partida para realizar mejoramiento asistido por marcadores para la capacidad de nodular y fijar nitrógeno en soja promoviendo rendimientos superiores bajo una agricultura sustentable.

**Palabras clave:** Soja (*Glycine max*), *Bradyrhizobium japonicum*, marcador molecular, análisis de mapeo de QTLs, nodulación, fijación biológica de nitrógeno.





## Abstract

### Quantitative trait loci (QTLs) associated to nodulation capacity and nitrogen fixation in soybean (*Glycine max* L. Merr)

Soybean (*Glycine max* L. Merr) is the most important legume cultivated worldwide. Argentina, Brazil and USA are the most important countries in terms of production and exports of soybean products. In order to make get the most of the production potential of soybean cultivars, N is incorporated by chemical and/or biological fertilization methods. The latter is aimed at using Biological Nitrogen Fixation (BNF) that has an advantage over the chemical fertilization since it reduces the risks of environmental pollution and contributes to a sustainable agriculture. BNF is the result of the plant rhizobia interaction, because of this a biotechnological process was developed and is known as inoculation. It consists in adding Gram-negative soil bacteria from the genus *Bradyrhizobium* that has the ability to interact with the soybean root, to the plant that as a result of this develops structures known as nodules, where bacteria fix nitrogen. In soybean almost most of the nitrogen fixed biologically is used by the plant growth therefore production is improved. Companies improved nitrogen fixation by developing formulations containing bacteria. selected based on its outstanding efficiency to fix nitrogen in an interaction with a wide array of cultivars and/or legume species. However, on the plant side breeding programs have been focused mostly in improving yield and resistance to biotic and abiotic factors. Starting in 1950, several genes that are involved in rhizobia-soybean interactions have been characterized. However, the ability of soybean cultivars to nodulate and fix nitrogen has a complex genetic architecture that depends on plant-bacterial and environment interactions. The size and complexity of the soybean genome, delayed the identification of plant genes involved in nitrogen fixation. Therefore the aim of this work was to identify quantitative trait loci (QTLs) associated to the ability of soybean to develop nodules, as well as to identify molecular markers associated to this trait. I analyzed the nodulation ability of 31 Argentinian commercial soybean cultivars which were inoculated with a mixture of *Bradyrhizobium japonicum* strain SEMIA 5079 (CPAC 7) and SEMIA 5080 (CPAC 15). The results allowed me to classify cultivars in high, medium and low nodulating cultivars. In a second experiment, a subset of the already analyzed cultivars was selected for further analysis and their response to inoculation was confirmed. In a third experiment, four cultivars two with high and two with low nodulation capacity were inoculated with another bacterial strain, *B. japonicum* E109 (recommended by INTA for the soybean crop in Argentina). Cultivars

responded like in previous experiments suggesting that the response was governed by the plant genome. Cultivars with high nodulation ability doubled the number of nodules developed by low nodulating cultivars. Heritability was estimated ( $H^2$ ) for the nodule number (NN), nodule dry weight (NDW) and shoot dry weight (SDW) and the results showed that the values of heritability was either medium or high and that NN was the most variable genetic trait. Soybean cultivars, whether with low or high nodulation capacity, increased nodulation as the concentration of bacteria in the inoculum increased though this occurred till a threshold value, decreasing thereafter. Probably nodulation was inhibited by a "quorum sensing" mechanism. Then we analyzed the genetic diversity of Argentinian soybean cultivars by means of SSR markers linked to QTLs linked to nodulation capacity. In addition to this, also the association between markers with nodulation capacity and other agronomic traits was studied and cultivars with contrasting nodulation capacity were identified, to cross them and in order to get a segregating population for QTL analysis. It was found that 10 SSR loci were polymorphic among Argentinian cultivars and that NN, NDW and SDW were associated to SSR markers previously described in Brazilian cultivars. Cultivar NA 5485 RG (high nodulating) and cultivar A 7053 RG (low nodulating) were selected as parents to generate the mapping population  $F_{2:3}$ . Genotyping was carried out in the  $F_2$  generation. Genomic DNA was screened 221 SSR loci along the whole soybean genome. Among them only 36 were polymorphic. A partial genetic map was constructed containing 5 linkage groups. Phenotyping was carried out over 94 families ( $F_3$ ) in a nodulation essay with light and temperature controlled conditions and quantitative traits that define the nodulation capacity: nodule number (NN), of nodule dry weight (NDW) and nodule dry weight per nodule (NDWN) calculated as NDW/NN. Other traits related to plant growth also were considered in the analysis: leaves dry biomass (LDB) and stem dry biomass (SDB) and the sum of both, shoot dry weight (SDW) and plant height (PH). A matrix with genotypic and phenotypic data was constructed and QTL mapping analysis was performed by three methods: single marker (ANOVA and regression), interval mapping (IM) and composite interval mapping (CIM). Satt414 locus was confirmed as QTL associated with NN and NDW (single marker). It was concluded that the variables associated with nodulation capacity in the  $F_3$  population presented high values of genetic variance and heritability. Identified QTLs for studied variables explained values higher than 15% of the phenotypic variance. A transgressive segregation was observed for most of the variables associated with nodulation capacity. In addition I also found that variables related to the nodulation capacity and biomass were associated. QTLs associated to nodulation capacity were located in linkage group GL H for NN; NDW and NDWN; GL D1b for NN; GL A2 for NN; GL J for NN and NDW; GL B2 for NDW and NDWN; GL B1 for NDWN,

GL C1 for NDWN. These QTLs explained 12.2-13.4 % of the phenotypic variation. QTLs related to growth and biomass were located in the following linkage groups: GL H for PH and SDB; GL M for PH. GL D1b for PH, SDB, SDW and LDB; GL B2 for PH; GL N for SDW; GL L for SDW and GL G for LDB. These QTLs explained 8.3-25.6% of the phenotypic variation. The applied methodology allowed me to locate a few QTLs for nodulation in intervals. However, these regions should be saturated with more and/or other molecular markers with better efficiency and variability as e.g. SNPs. A higher level of saturation as well as the information regarding the soybean genome sequence and EST's should help in the identification of genes related to nodulation and nitrogen fixation. To my knowledge this is the first work aimed at identifying nodulation and nitrogen fixation genes in soybean cultivars from Argentina and provide the starting point for marker assisted selection for nodulation capacity and nitrogen fixation in soybean in order to promote better yields under a sustainable agriculture.

**Keywords:** Soybean (*Glycine max*), *Bradyrhizobium japonicum*, molecular marker, QTL mapping analysis, nodulation, biological nitrogen fixation.



## **Agradecimientos**

A la Comisión Nacional de Investigaciones Científicas y Tecnológicas (CONICET) por la beca doctoral otorgada.

Al Instituto de Fisiología Vegetal (INFIVE-CONICET) y al Instituto Santa Catalina de la Universidad Nacional de La Plata, al EMBRAPA Soja de Londrina (Brasil) y Nidera Semillas SA. por darme sus lugares y la oportunidad de desarrollar este trabajo de investigación.

A la Facultad de Ciencias Exactas y Naturales de la UBA por aceptarme como alumno de doctorado y seguir siendo parte de esta gran Casa de Estudios en la que me he formado como biólogo.

Al Dr Pedro A. Balatti por darme la posibilidad de formarme bajo su dirección, por su constante apoyo, confianza por darme consejos y estimularme en este trabajo y su constante preocupación.

A la Dra Mónica B. Aulicino, por darme su orientación como Directora Asistente, haberme estimulado en todo este trayecto. Por su valiosa ayuda en el análisis estadístico de los resultados por formarme en genética cuantitativa y la constante preocupación en el desarrollo de este trabajo.

A la Dra. Viviana Confalonieri, como Consejera de Estudios me brindó su ayuda en la orientación en los cursos durante mi carrera de doctorado.

A la Dra. Mariangela Hungria por abrirme las puertas de su laboratorio, EMBRAPA Soja, en Londrina Brasil, y a Maria Santos por brindarme su conocimiento y poder realizar parte del trabajo molecular de microsatélites que me ayudó a resolver el capítulo 3 de la presente tesis.

Al Ing. Rodolfo Rossi por cederme gentilmente las semillas de soja de la Compañía Nidera Semillas SA y abrirme las puertas de la empresa para poder realizar allí parte de mi trabajo.

Al Ing. Carlos Sala, del Departamento de biotecnología de Nidera por ayudarme en la obtención de la población segregante como también abrirme las puertas para poder realizar parte de mi trabajo molecular dentro de la empresa. Hoy lamento mucho la pérdida de un excelente profesional y persona.

A Emiliano Altieri, del Departamento de Nidera, por darme la posibilidad de aprender técnicas nuevas en marcadores moleculares brindarme las herramientas para realizar el trabajo molecular de mapeo que generaron resultados en el corto plazo, lo cual permitió culminar mi proyecto de tesis.

Al Ing. Roberto Wright (lamento su fallecimiento) y Santiago Zujic, del Criadero Santa Rosa de Salto por haberme enseñado a realizar los cruzamientos en soja, un grupo humano excelente que me abrió las puertas del criadero para poder aprender

A la Dra. Marta Galván, por ser la primera persona que me enseñó y formó en técnicas de biología molecular.

Al Dr. Mario Saparrat, por sus consejos, por ayudarme en el laboratorio, su constante preocupación y transmitir su estímulo y pasión por la ciencia.

A la Lic. Virginia Martínez Alcántara por haberme transmitido sus conocimientos de microbiología y haberme ayudado en los ensayos de nodulación.

A mis compañeros del laboratorio: Ale, Pepe, Ernesto, Silvina, Rocío, Emiliano, Inés, Grace, por bancarme y alentarme día a día. Por los mates y anécdotas, por compartir conmigo toda esta etapa como una aventura y sobre todo haberme dado una mano en los ensayos.

A mis queridos amigos biólogos de la FCEyN: por el apoyo brindado, los consejos y bancarme en este camino y las cenas compartidas, Marian Masso, Dra. Guillermina Bilbao (DNína Protea), Dra. Patricia Demkura, Leila Haiek, Mecha, Agustina de Dios, Romina Cutrullis, Jerónimo Auzmendi, Valeria Botti (Harás magia!!), Juli Rojas *et al.*

A mis amigos de la vida que me alentaron durante toda esta etapa:

A Gustavo González por haberme ayudado en la edición de las figuras de la tesis y consejos que me ha brindado, a su hermana Sandra G. por los mates compartidos y a su hermano Matías G., a mi amigo biólogo y piscicultor Martín Sotelo, por compartir la amistad desde la facultad y aguante, a Maxi Garzón, a Felipe Whalberg, a Eliana y Mer Soria. Les agradezco a todos los momentos de desenchufe y desestresamiento.

A mi amigo de Carlos Casares Damián Regules que me seguía alentando y bancando a la distancia.

Al Dr. Pavan Akkiraju, ayer compañero de un curso de QTLs, hoy amigo desde el otro lado del mundo (India), por la corrección del abstract de la tesis.

Y Gracias sobre todo a mi familia, a mis padres: Sara y Rubén que me brindaron el apoyo incondicional durante toda esta etapa hasta el final, a mi abuela Perla, mi hermano Javier y Laura por acompañarme, aconsejarme y ayudarme a superar escollos durante este camino.

A todos aquellos, que de alguna u otra forma aportaron su grano de arena para que este trabajo fuera posible.

A todos ¡MUCHAS GRACIAS!



**Parte de los resultados de la presente Tesis Doctoral fueron publicados en:**

### **TRABAJOS PUBLICADOS**

**Salvucci Rubén Darío**, Aulicino M, Hungria M, Balatti PA. (2012 ). Nodulation capacity of Argentinean soybean (*Glycine max* L Merr) cultivars inoculated with commercial strains of *Bradyrhizobium japonicum* . *Am J Plant Sci* 3:130-140. doi: [104236/ajps201231015](https://doi.org/10.4236/ajps201231015)

### **COMUNICACIONES A CONGRESOS**

VI Reunión Nacional Científico Técnica de Biología del Suelo. VI Encuentro sobre Fijación Biológica de Nitrógeno.

**Nodulación de distintos cultivares comerciales de soja de Argentina.** Salvucci R. D., Aulicino M.B., Balatti P. 2007. Universidad Nacional de Río Cuarto, Córdoba, Argentina.

Tipo de trabajo: Resumen. Fuente: VI Reunión Nacional Científico Técnica de Biología del Suelo. VI Encuentro sobre Fijación Biológica de Nitrógeno. Pág. 85.

XXXVI Congreso Argentino de Genética.

**Variabilidad en cultivares argentinos de soja con distintas capacidades de nodulación.** 23 al 26 de septiembre de 2007. UNNOBA - Pergamino - Prov. de Buenos aires. Salvucci RD, Aulicino MB, Galván MZ y Balatti, PA. Publicado en el libro de resúmenes Basic and Applied Genetics vol 18(suplemento), pág. 79. 2007. Institución organizadora: Institución organizadora: Sociedad Argentina de Genética.

XXII Congreso Argentino de la Ciencia del Suelo.

**Respuestas diferenciales de variedades argentinas de soja a la capacidad de nodulación y fijación de nitrógeno.** 2010. Salvucci R.D., M.B. Aulicino y P.A. Balatti. Rosario. Argentina. Tipo de trabajo: Resumen expandido presentado en forma oral. Fuente: Actas XXII Congreso Argentino de la Ciencia del Suelo.

5° Congreso de la Soja del Mercosur, mercosoja 2011. Rosario.

**Potencial aplicación de marcadores ssr para el mejoramiento en la capacidad de nodulación y caracteres de ciclo en soja.** 2011. RD Salvucci, MB Aulicino, MA Santos, M Hungria, R Rossi, C Sala y PA Balatti. Institución organizadora: Asociación de la Cadena de la Soja Argentina ACSOJA.

XLII Congreso Argentino de Genética y III Reunión Regional de SAG-NOA.

**Identificación de QTLs asociados a nodulación en soja (*Glycine max* L. Merr).** 20 al 23 de octubre de 2013. Salta. Salvucci RD, Aulicino MB, Altieri E, Sala C y Balatti, PA. Publicado en el libro de resúmenes Basic and Applied Genetics vol 24(1), pág. 173. 2013. Institución organizadora: Institución organizadora: Sociedad Argentina de Genética.

XLIII Congreso Argentino de Genética y IV Reunión Regional de SAG-NOA.

**Variabilidad genética en una F<sub>3</sub> de la cruce entre cultivares de soja contrastantes en nodulación.** 19 al 22 de octubre de 2014. Bariloche. Salvucci RD, Aulicino MB, Altieri E, Sala C y Balatti, PA. Publicado en el libro de resúmenes Basic and Applied Genetics vol 25(1), pág. 204. 2014. Institución organizadora: Institución organizadora: Sociedad Argentina de Genética.

## CONGRESOS INTERNACIONALES

The 21st North American Symbiotic Nitrogen Fixation Conference.

**Nodulation and SSR markers to evaluate the BNF potential of Argentinean Cultivars of Soybean (*Glycine max* L. Merr).** 2010. Pedro Alberto Balatti, Darío Salvucci, Mónica Aulicino, Mariangela Hungria. Christopher S. Bond Life Sciences Center University of Missouri- Columbia, MO USA Tipo de trabajo: Resumen. Publicado en la versión electrónica en pdf.

## CAPÍTULOS DE LIBRO

Título: Los rizobios que nodulan la soja en sitios con ambientes nativos y cultivados de la Argentina.

Autores: Silvina M. Y. López; Graciela N. Pastorino; Virginia Martínez Alcántara; Darío Salvucci; Pedro A. Balatti; Ada S. Albanesi

Libro: Microbiología Agrícola - Un aporte de la investigación en Argentina. Segunda edición.

Editorial: Magna Publicaciones

Referencias: Lugar: Tucumán; Año: 2013; p. 237-252



# Índice

Resumen.....	i
Abstract.....	v
Agradecimientos .....	ix
Índice de Figuras.....	xix
Índice de Tablas.....	xxii
<b>1. Introducción general.....</b>	<b>1</b>
1.1. Descripción botánica de la soja ( <i>Glycine max</i> )	1
1.2. Usos	2
1.3. Origen y difusión en el mundo de la soja	4
1.4. Historia de la soja en Argentina	4
1.5. Desarrollo	7
1.5.1. Etapas de desarrollo	7
1.5.2. Factores que afectan el desarrollo	10
1.5.2.1. Fotoperíodo	10
1.5.2.2. Temperatura	13
1.5.2.3. Otros	13
1.6. Tipo de crecimiento	14
1.7. Importancia del nitrógeno en los ecosistemas. Ciclo biogeoquímico	15
1.7.1. Procesos implicados en el ciclo del nitrógeno en los sistemas terrestres	16
1.8. Fijación Biológica de Nitrógeno rizobio-leguminosas	20
1.8.1. Taxonomía de rizobios	20
1.8.2. Descripción del género <i>Bradyrhizobium</i>	21
1.8.3. Infección y nodulación	24
1.8.4. Fijación biológica de nitrógeno.	28
1.8.4.1. <i>Nitrogenasa</i>	28
1.9. Genética de la nodulación en el hospedante	30
1.10. Marcadores moleculares	33
1.10.1. Polimorfismo en la longitud de los fragmentos de restricción (RFLP)	34
1.10.2. Polimorfismos de ADN amplificados al azar (RAPD)	36
1.10.3. Polimorfismo de la longitud de fragmentos amplificados (AFLP)	36
1.10.4. Repeticiones de secuencia simple (SSR)	37
1.10.5. Polimorfismo de nucleótido simple (SNP)	38
1.11. Carácter cuantitativo y loci de carácter cuantitativo (QTL)	39
1.11.1. Pasos para el mapeo de QTLs	39
1.11.1.1. Selección de progenitores	39
1.11.1.2. Población de mapeo	40
1.11.1.3. Construcción del mapa	41
1.11.1.4. Análisis de QTLs	42
1.11.2. Aplicación del mapeo de QTLs	43
1.12. El genoma de la soja y mapa de ligamiento	45
1.12.1. Breve reseña citogenética evolutiva de <i>Glycine max</i>	45
1.12.2. Mapas genéticos en soja	45
1.12.3. Genómica	47
1.13. Objetivos	47

<b>2. Capacidad de nodulación de cultivares argentinos de soja (<i>Glycine max</i> L. Merr).....</b>	<b>51</b>
2.1. Introducción	51
2.2. Materiales y métodos	56
2.2.1. Mantenimiento de la bacteria y producción del inóculo	56
2.2.2. Material vegetal y ensayos de nodulación	57
2.2.3. Nodulación de cultivares de soja	59
2.2.3.1. <i>Primer experimento</i>	59
2.2.3.2. <i>Segundo experimento</i>	59
2.2.3.3. <i>Tercer experimento</i>	60
2.2.4. Respuesta de la nodulación frente a la concentración bacteriana	60
2.2.5. Análisis estadístico	61
2.3. Resultados	62
2.3.1. Primer Experimento	62
2.3.2. Segundo Experimento	65
2.3.3. Tercer experimento	67
2.3.4. Efecto de la concentración bacteriana en la respuesta a la nodulación	68
2.3.5. Análisis combinado y estimaciones de heredabilidades	70
2.4. Discusión	71
2.5. Conclusiones	73
<b>3. Diversidad genética de cultivares de soja de Argentina, capacidad de nodulación y asociación con caracteres agronómicos .....</b>	<b>77</b>
3.1. Introducción	77
3.2. Materiales y métodos	81
3.2.1. Amplificación y análisis de SSR	82
3.2.2. Análisis estadísticos	83
3.3. Resultados	85
3.3.1. Comparaciones de las clasificaciones moleculares y morfológicas	91
3.3.2. Diversidad genética de cultivares de soja y microsatélites	93
3.4. Discusión	98
3.5. Conclusiones	102
<b>4. Mapeo e identificación de QTLs asociados al carácter “capacidad de nodulación” .....</b>	<b>107</b>
4.1. Introducción	107
4.2. Materiales y métodos	114
4.2.1. Obtención de la población segregante F <sub>2:3</sub>	114
4.2.2. Evaluación fenotípica de la población F <sub>2:3</sub>	115
4.2.3. Análisis fenotípico de familias F <sub>3</sub>	116
4.2.3.1. <i>Análisis de asociación entre variables</i>	118
4.2.3.2. <i>Análisis de coeficientes de paso.</i>	119
4.2.4. Evaluación genotípica de la población segregante F <sub>2</sub>	121
4.2.4.1. <i>Prueba de segregación de marcadores</i>	122
4.2.4.2. <i>Construcción del mapa de ligamiento</i>	122
4.2.5. Análisis de mapeo de QTLs	122
4.2.5.1. <i>Locus simple. Método ANOVA de un solo factor</i>	123
4.2.5.2. <i>Locus simple. Método Regresión lineal</i>	125
4.2.5.3. <i>Intervalo simple. Método máxima verosimilitud</i>	126

4.2.5.4. <i>Mapeo por intervalo compuesto (CIM)</i>	129
4.3. Resultados	131
4.3.1. Análisis fenotípico y genético de la F <sub>3</sub>	131
4.3.2. Análisis de varianza de los caracteres en las familias F <sub>3</sub>	134
4.3.3. Análisis de correlación y coeficientes de paso	135
4.3.4. Análisis molecular	138
4.3.4.1. <i>Selección y segregación de marcadores</i>	138
4.3.4.2. <i>Construcción de Mapa de ligamiento</i>	138
4.3.5. Análisis de QTLs	141
4.3.5.1. <i>Mapeo de locus simple. Método ANOVA</i>	141
4.3.5.2. <i>Mapeo de locus simple. Análisis de Regresión</i>	149
4.3.5.3. <i>Mapeo por intervalo simple. Análisis de regresión por máxima verosimilitud</i>	152
4.3.5.4. <i>Mapeo por intervalo compuesto</i>	157
4.4. Discusión	158
4.4.1. Otros QTLs coincidentes relacionados con FBN	165
4.5. Conclusiones	166
Consideraciones finales.....	169
Referencias bibliográficas .....	171
ANEXO .....	203



## Índice de Figuras

<b>Figura 1.1</b>   Ilustración de <i>Glycine max</i> L. Merr (soja).....	3
<b>Figura 1.2</b>   Evolución anual del cultivo de soja en Argentina: superficie sembrada y producción entre 1970 y 2013..	6
<b>Figura 1.3</b>   Superficie sembrada, producción y exportación mundial de soja 2012/2013..	7
<b>Figura 1.4</b>   Estados de desarrollo vegetativo y reproductivo de la soja.....	10
<b>Figura 1.5</b>   Fajas latitudinales de variedades de soja según su grupo de maduración. ....	13
<b>Figura 1.6</b>   Hábitos de crecimiento de la soja.....	15
<b>Figura 1.7</b>   Ciclo del nitrógeno.....	19
<b>Figura 1.8</b>   Colonias de <i>Bradyrhizobium japonicum</i> cepa E109.....	23
<b>Figura 1.9</b>   Desarrollo del nódulo radicular.....	27
<b>Figura 1.10</b>   Esquema del mecanismo de acción del complejo enzimático nitrogenasa.....	29
<b>Figura 1.11</b>   Diagrama representando marcadores de ADN entre los genotipos A, B, C y D.....	35
<b>Figura 1.12</b>   Comparación entre un marcador (a) codominante y (b) dominante.....	36
<b>Figura 1.13</b>   Principios del mapeo de loci de caracteres cuantitativos.....	43
<b>Figura 1.14</b>   Técnicas de mapeo de QTLs en la población.....	44
<b>Figura 2.1</b>   Respuesta a la nodulación de 30 cultivares de soja.....	64
<b>Figura 2.2</b>   Respuesta de genotipos de soja a la inoculación con una mezcla de cepas <i>Bradyrhizobium japonicum</i> cepas SEMIA 5079 y SEMIA 5080 (valores medios $\pm$ desvío estándar).....	66
<b>Figura 2.3</b>   Respuesta a la nodulación de cuatro cultivares de soja frente a la inoculación con <i>Bradyrhizobium japonicum</i> cepa E109 (inoculante comercial) en (valor medio $\pm$ desvío estándar).....	67
<b>Figura 2.4</b>   Efecto de la densidad celular de <i>Bradyrhizobium japonicum</i> sobre la nodulación de <i>Glycine max</i> L. Merr cultivar NA 5485 RG (blanco) y A 6411 RG (negro), valores medios $\pm$ desvío estándar. <b>a.</b> Número de nódulos, <b>b.</b> Peso seco de nódulos y <b>c.</b> Peso seco aéreo.....	69
<b>Figura 3.1</b>   Análisis de cluster de quince cultivares comerciales argentinos de soja y tres cultivares de Brasil, Bossier (BOS), BRS133 y Empbrapa20 (E20) usando datos moleculares (10 SSR's polimórficos).....	88
<b>Figura 3.2</b>   Análisis de cluster de 15 cultivares comerciales argentinos de soja usando caracteres cuanti y cualitativos.....	90
<b>Figura 3.3</b>   GGE Biplot (Modelo I) mostrando la relación entre genotipos argentinos (cuadrados y rombos) y marcadores SSR (círculos).....	92



<b>Figura 3.4</b>   Biplot del análisis de correspondencia para categorías de capacidad de nodulación (baja, media y alta) y los diferentes alelos microsatélites.....	96
<b>Figura 3.5</b>   Análisis de cluster usando 10 SSR's polimórficos y variables cuali y cuantitativas consideradas como multiestado.....	97
<b>Figura 4.1</b>   Diagrama de coeficientes de paso.....	121
<b>Figura 4.2</b>   Histogramas de distribución de frecuencias en la población F <sub>3</sub> , para los siete caracteres cuantitativos evaluados.....	133
<b>Figura 4.3</b>   Diagrama de coeficientes de paso describiendo la relación entre el peso seco aéreo (PSA), número de nódulos (NN), peso seco de nódulos (PSN), peso seco de hojas (PSH), peso seco de tallo (PST) y altura de planta (AP) de 94 familias F <sub>3</sub> de soja.....	136
<b>Figura 4.4</b>   Mapa genético producido a partir de los marcadores microsatélites de una población F <sub>2</sub> proveniente del cruzamiento entre NA 5485 RG x A 7053 RG.....	141
<b>Figura 4.5</b>   Valores fenotípicos del NN medios condicionados a cada clase genotípica de cada marcador SSR en la población F <sub>2</sub> .....	145
<b>Figura 4.6</b>   Valores fenotípicos del PSN medios condicionados a cada clase genotípica de cada marcador SSR en la población F <sub>2</sub> .....	146
<b>Figura 4.7</b>   Valores fenotípicos de la AP medios condicionados a cada clase genotípica de cada marcador SSR en la población F <sub>2</sub> .....	146
<b>Figura 4.8</b>   Valores fenotípicos del PST medios condicionados a cada clase genotípica de cada marcador SSR en la población F <sub>2</sub> .....	147
<b>Figura 4.9</b>   Valores fenotípicos del PSH medios condicionados a cada clase genotípica de cada marcador SSR en la población F <sub>2</sub> .....	147
<b>Figura 4.10</b>   Valores fenotípicos del PSA medios condicionados a cada clase genotípica de cada marcador SSR en la población F <sub>2</sub> .....	148
<b>Figura 4.11</b>   Valores fenotípicos del PSNN medios condicionados a cada clase genotípica de cada marcador SSR en la población F <sub>2</sub> .....	149
<b>Figura 4.12</b>   Mapeo por intervalo aplicando el método máxima verosimilitud del QTL número de nódulos por planta (NN) en el grupo de ligamiento J.....	152
<b>Figura 4.13</b>   Mapeo por intervalo por el método máxima verosimilitud del QTL del peso seco de nódulos por nódulo por planta (PSNN) en el grupo de ligamiento C1.....	153
<b>Figura 4.14</b>   Mapeo por intervalo por el método máxima verosimilitud del QTL del peso seco de nódulos por nódulo por planta (PSNN) en el grupo de ligamiento M.....	154
<b>Figura 4.15</b>   Mapeo por intervalo aplicando el método máxima verosimilitud del QTL de la altura de la planta (AP) en el grupo de ligamiento M.....	155

<b>Figura 4.16</b>   Mapeo por intervalo por el método máxima verosimilitud del QTL del peso seco de hojas por planta (PSH) en el grupo de ligamiento D1b.....	156
<b>Figura 4.17</b>   Mapa parcial de ligamiento con la localización de los cinco QTLs y efecto aditivo (a) o dominante (d) usando el método de mapeo por intervalo simple (SIM)..	157
<b>Figura 4.18</b>   Mapeo de QTLs por intervalo compuesto.....	158

## Índice de Tablas

<b>Tabla 2.1</b>   Cultivares de soja incluidos en los ensayos (Nidera Semillas).....	58
<b>Tabla 2.2</b>   Análisis estadístico del número de nódulos, peso seco de nódulos y parte aérea de cultivares de soja inoculados con una mezcla de cepas de <i>Bradyrhizobium japonicum</i> .....	70
<b>Tabla 3.1</b>   Descriptores agronómicos y caracteres de nodulación para cada cultivar de soja.....	82
<b>Tabla 3.2</b>   Grupo de ligamiento, tamaño alélico, número de alelos y diversidad genética (H) de 10 loci SSR loci en cultivares argentinos.....	87
<b>Tabla 3.3</b>   Grupo de ligamiento, tamaño alélico, número de alelos y diversidad genética (H) de 10 loci SSR loci en cultivares argentinos y brasileros juntos.....	87
<b>Tabla 3.4</b>   Clasificación de los clusters por ANOVA y diferencias por el test de Tukey.....	90
<b>Tabla 4.1</b>   Esquema del análisis de la varianza para una población $F_2$ .....	124
<b>Tabla 4.2</b>   Análisis de la varianza de los valores fenotípicos de las 94 familias $F_3$ provenientes del cruzamiento entre NA 5485 RG y A7053 RG.....	134
<b>Tabla 4.3</b>   Matriz de correlaciones fenotípica (encima de la diagonal) y genética (debajo de la diagonal) entre todas las variables cuantitativas en 94 familias $F_3$ de soja proveniente del cruzamiento NA 5485 RG x A 7053 RG.....	136
<b>Tabla 4.4</b>   Coeficientes de paso que muestran los efectos directos e indirectos para la variable dependientes PSA.....	137
<b>Tabla 4.5</b>   Coeficientes de paso que muestran los efectos directos e indirectos para la variable dependientes PSH (izquierda), PST (centro) y AP (derecha).....	137
<b>Tabla 4.6</b>   Segregación de los 36 marcadores microsatélites según tres pruebas: Chi-cuadrado, criterio individual, Bonferroni y FDR (False Discovery Rate).....	140
<b>Tabla 4.7</b>   Marcadores microsatélites asociados a los caracteres cuantitativos por medio del análisis de la varianza.....	144
<b>Tabla 4.8</b>   Marcadores microsatélites asociados significativamente a los caracteres cuantitativos por medio del análisis de regresión.....	151

# **CAPÍTULO 1**

## **INTRODUCCIÓN GENERAL**



# 1. INTRODUCCIÓN GENERAL

## 1.1. Descripción botánica de la soja (*Glycine max*)

La soja pertenece a la familia de las Leguminosas, subfamilia *Papilionoideas*, tribu *Faseoleas* y al género *Glycine* Willd. Este género se divide en dos subgéneros según Hymowitz y Newell (1981) *Soja* y *Glycine*. Dentro del subgénero *Soja* se incluye la especie cultivada *Glycine max* (L.) Merr (Hymowitz y Singh, 1987) y al antecesor silvestre *Glycine soja* Sieb. & Zucc. Ambas son especies diploides ( $2n = 40$ ) y presentan homología cromosómica lo que permite la obtención de híbridos fértiles entre sí.

La soja *Glycine max* (L.) Merrill se destaca por su alto contenido de proteína y por su calidad nutritiva. El grano de soja contiene más proteínas que la mayoría de las legumbres, pero menos materia grasa que la mayor parte de las oleaginosas. Considerando que existe diversidad morfológica en función del ambiente que se considere, la planta de soja puede alcanzar una altura promedio de 83 cm (Kantolic *et al.*, 2006), con valores máximos de 123 cm y mínimos de 41 cm (Toledo *et al.*, 2008). Presenta las dos primeras hojas unifoliadas opuestas y el resto trifoliadas dispuestas en forma alterna. A partir de algunas yemas axilares pueden desarrollarse ramas con una estructura similar al tallo principal. El sistema radical está compuesto por una raíz principal pivotante donde, según el genotipo, la profundidad exploratoria estaría próxima a los 2 m (Kantolic *et al.*, 2006).

Las flores presentan características típicas de las *Papilionoideas*, forman racimos axilares con 2 a 35 flores cada uno. Presentan un cáliz tubular y cinco pétalos desiguales, cuyos colores varían entre blanco y púrpura, siendo el tamaño de la flores no mayor a 5 mm. Las flores son autógamas, predomina la autofecundación, por lo

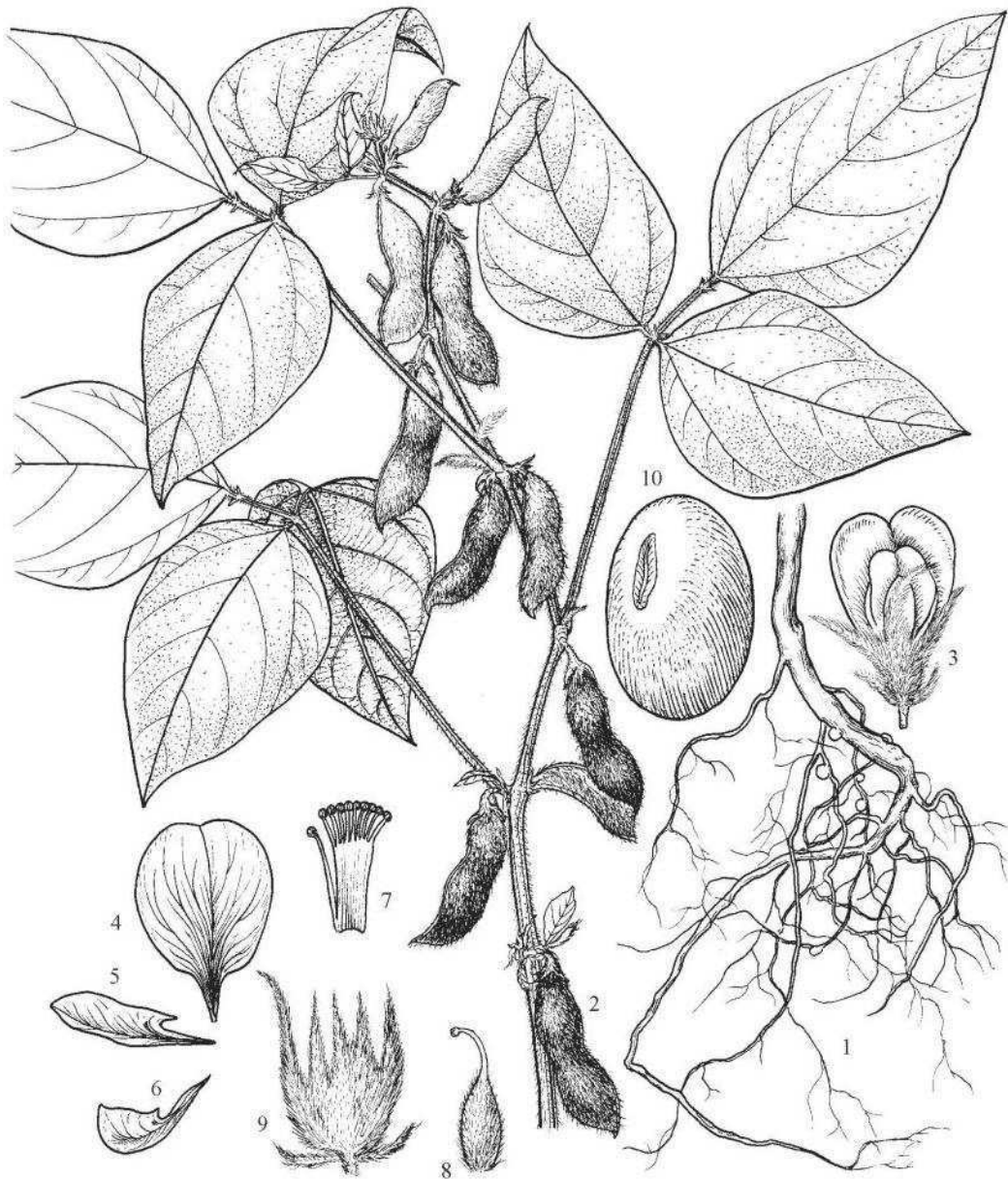
cual conservan casi intacta su información genética a lo largo de las generaciones. Este tipo de reproducción conduce a que en un cultivar se conserven las características genéticas de las plantas parentales aunque la propagación se realice por medio de semillas. Las flores (y, por tanto, las legumbres) se ubican en el punto de unión de las hojas con el tallo principal o con las ramas laterales (axila de la hoja), y se agrupan formando un racimo. Las vainas son pubescentes y de forma achatada y levemente curvada con un largo entre 2 y 7cm; cada una puede contener entre 1 y 5 granos pero generalmente presentan 2 o 3 granos. En cada racimo se pueden encontrar de 2 a 20 vainas que a la madurez presentan colores muy variados entre el amarillo claro y el marrón oscuro, incluso negro en algunas variedades (Kantolic *et al.*, 2006).

Las semillas son redondeadas con una coloración habitualmente amarilla, el peso promedio aproximado es de 130 mg, pero estos valores pueden variar en un rango de 112 mg y 165 mg (EEA INTA Marcos Juárez, 2007) de peso de cada semilla. La cicatriz de la semilla (hilo) que presenta colores diversos desde amarillo a negro pasando por diferentes tonalidades de marrón, es una característica que permite la identificación de los cultivares (Kantolic *et al.*, 2006) (**Figura 1.1**).

## 1.2. Usos

Entre los usos destacados se enumera: la producción de aceite para la elaboración de margarina y aceite de cocina. También el aceite de soja es útil para la elaboración de pinturas, barnices, cosmética, plásticos y recientemente como biodisel. Luego del procesamiento para la extracción de aceites, el residuo que se conoce como torta es utilizado en productos alimenticios como fuente de proteína para la elaboración de alimento animal (ganado y aves de corral). También la semilla de soja sirve para elaborar alimentos procesados como brote de soja, tofu, salsa y leche de

soja. Otras aplicaciones son materiales de construcción, velas, comida preventiva para osteoporosis, cáncer y diabetes.



Fuente: Flora of China. [http://www.efloras.org/object\\_page.aspx?object\\_id=116736&flora\\_id=2](http://www.efloras.org/object_page.aspx?object_id=116736&flora_id=2)

**Figura 1.1** | Ilustración de *Glycine max* L. Merr (soja). 1. Raíces con nódulos, 2. Fruto, 3. Flor, 4. Pétalo estandarte, 5. Ala, 6. Quilla, 7. Androceo, 8. Pistilo, 9. Cáliz. 10. Semilla (agrandada).



### **1.3. Origen y difusión en el mundo de la soja**

La soja es originaria de Asia, según Vavilov pertenece al centro de diversificación de especies N°3 (China y Japón). Comenzó a domesticarse en el norte y centro de China entre 1100 y 700 AC, luego se expandió a otros países de Asia como India, Nepal, Birmania, Tailandia, Indochina, Corea, Japón, Malasia, Indonesia y Filipinas antes del primer milenio DC (Smartt y Hymowitz, 1985). Recién a partir de 1880, Estados Unidos comenzó a cultivar la soja adquiriendo importancia como productor. Si bien inicialmente fue usada como forraje, luego en la década del 30 se priorizó la producción de grano. En Brasil fue introducida en 1882, pero su difusión se inició a principios del siglo XX y la producción comercial comenzó en la década del 40 (Verneti, 1983).

### **1.4. Historia de la soja en Argentina**

La historia de la soja en la Argentina, al igual que en otros países como Brasil, que hoy aparecen como principales protagonistas del mercado mundial, está fuertemente condicionada por la evolución de la demanda externa de este producto. En estas dos naciones del Mercosur, pasó de ser una especie exótica a convertirse en el cultivo de mayor expansión en cuanto a superficie sembrada, volumen de producción en toneladas y nivel de exportación de sus derivados, puede decirse que se inicia a partir de los años 70 en consonancia con los cambios operados en dicho mercado.

Las primeras plantaciones de soja en Argentina se remontan 1862, si bien en aquellos años no se encontró eco entre los productores agrícolas. En 1925, el Ministro de Agricultura Le Bretón, introdujo nuevos lotes de semillas de soja desde Europa y trató de difundir su cultivo, conocido en esa época entre los agrónomos del Ministerio como arveja peluda o soja hispida. Hacia 1956 en la Argentina aún no se conocían los

aspectos básicos del cultivo de la soja, y por ello se produjeron fracasos en la implantación que llevaron a considerarla cultivo “tabú”. La primera vez que Argentina exportó soja fue en 1962, a través del buque “Alabama”, que partió con 6.000 toneladas con destino a Hamburgo (Alemania). La campaña 61/62 mostró el primer salto importante en la producción de esta oleaginosa, ya que la superficie sembrada se extendió más de 10 veces en relación a la campaña anterior, 1.014 ha a 10.260 ha (Martínez Dougnac, 2004).

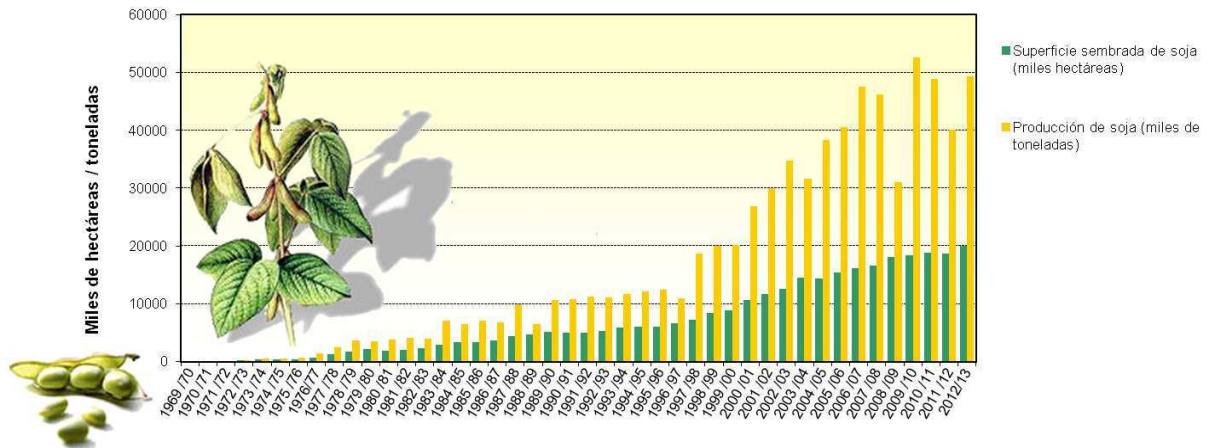
Durante la década del 60 confluyen una serie de factores y se crean condiciones favorables para el desarrollo del cultivo. En 1965 la situación planteada llevó al Ministerio de Agricultura y Ganadería, a través de la Junta Nacional de Granos, a fijar por primera vez un precio mínimo oficial para la semilla de soja. Esto ocurrió en un período de precios en relativo crecimiento, lo que generó condiciones favorables para su cultivo. Así al comenzar la década del 70, se inicia un período de expansión que irá orientando el cultivo hacia las más ricas tierras de la zona pampeana, aquellas tradicionalmente ocupadas por maíz (“donde se da bien el maíz se da bien la soja”), y modificando otra característica del período “experimental”: una mayor dispersión geográfica, con predominio de provincias extrapampeanas<sup>1</sup>.

La década del 70 fue la más relevante, ya que en ésta ocurrió el despegue del cultivo en el país (30.470 ha sembradas, se cosecharon 26.800 tn con un rendimiento promedio de 1.031 kg ha<sup>-1</sup>). La campaña 76/77, en la que se cosecharon 1.400.000 tn (más del doble que en el año anterior), puede considerarse como el inicio del verdadero incremento del cultivo en Argentina, alcanzando en la última campaña 2012/13, 20 millones de ha, con una producción de 49 millones de toneladas y un rendimiento de 2.500 kg ha<sup>-1</sup>, convirtiendo a la Argentina en el tercer productor y exportador mundial del grano (**Figura 1.2** y **Figura 1.3**). Además Argentina lidera las exportaciones de aceite de soja a nivel mundial (Argentina: 4,2 millones tn, Brasil: 1,3

---

<sup>1</sup> A mediados de los 60 Misiones producía cerca de la mitad de la soja argentina.

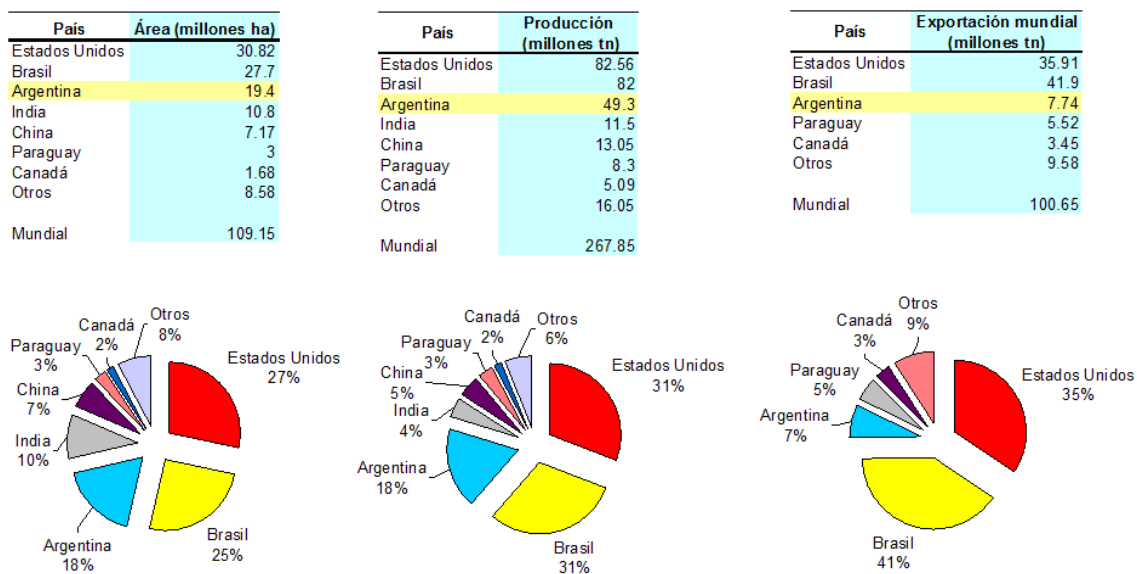
millones tn, Unión Europea: 1 millón tn, Estados Unidos: 982 mil tn y Paraguay: 558 mil tn) y también es el primer exportador de harina de soja (Argentina con 23,7 millones tn, Brasil con 13,2 millones tn, Estados Unidos con 10,1 millones tn, India con 4,3 millones tn, y Paraguay con 2,1 millones tn)<sup>2</sup>.



**Figura 1.2** | Evolución anual del cultivo de soja en Argentina: superficie sembrada y producción entre 1970 y 2013. Elaboración propia. Fuente: SAGPyA. Elaboración propia.

El principal país comprador de la producción de esta oleaginosa es China, cuya demanda se ha mantenido estable durante los últimos 30 años, en donde además el consumo interno se incrementó exponencialmente, lo que llevó a un aumento de las importaciones. La principal utilización del grano de soja en China es para engorde de cerdos y pollos.

<sup>2</sup> Fuente: USDA. Campaña 2012/2013. <http://apps.fas.usda.gov/psdonline/circulars/oilseeds.pdf>



Fuente: USDA

**Figura 1.3** | Superficie sembrada, producción y exportación mundial de soja 2012/2013. Elaboración propia. Fuente: USDA.

## 1.5. Desarrollo

### 1.5.1. Etapas de desarrollo

Existen varias clasificaciones para identificar los distintos estados de desarrollo en soja, la más difundida es la escala desarrollada por Fehr *et al.* (1971), donde se describe los estadios fenológicos externos del cultivo de soja, distinguiéndose dos etapas principales; una que describe los estados vegetativos y la otra los reproductivos (**Figura 1.4**).

#### Etapa vegetativa

Los 2 primeros estados vegetativos se los identifican con letras.

- **VE - Emergencia** - Se observa el hipocótilo, en forma de arco, empujando al epicótilo y a los cotiledones, haciéndolos emerger sobre la superficie del suelo.
- **VC - Etapa cotiledonar** - El hipocótilo se endereza, los cotiledones se despliegan totalmente y en el nudo inmediato superior los bordes de las hojas unifoliadas no se tocan.

A partir de aquí el resto de los estados vegetativos se los identifican con el número de nudos.

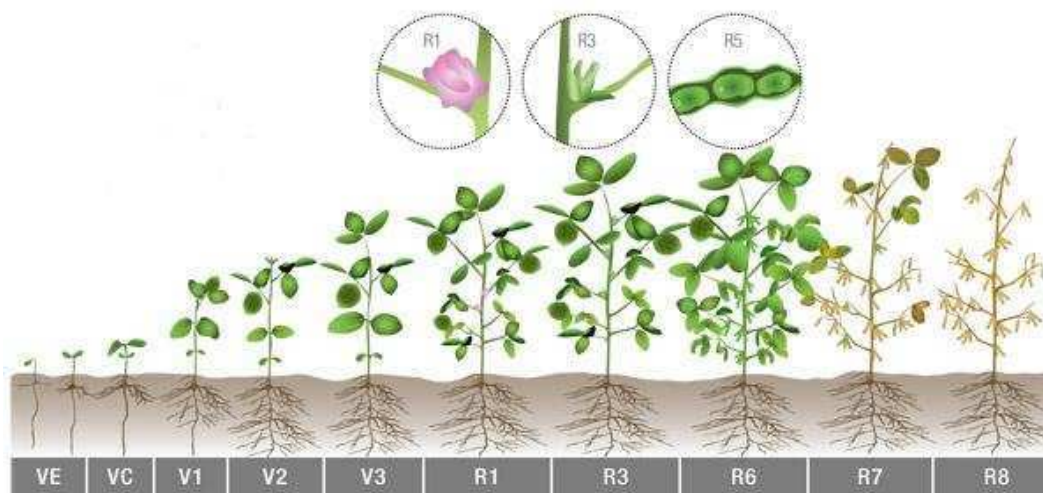
- **V1 - (1er nudo)** - El par de hojas opuestas unifoliadas están expandidas totalmente, y en el nudo inmediato superior se observa que los bordes de cada uno de los folíolos de la 1er hoja trifoliada no se tocan.
- **V2 - (2do nudo)** - La 1er hoja trifoliada está totalmente desplegada, y en el nudo inmediato superior los bordes de cada uno de los folíolos de la 2da hoja trifoliada no se están tocando.
- **Vn - (n: número de nudos)** - La hoja trifoliada del nudo (n) está expandida totalmente, y en el nudo inmediato superior los bordes de cada uno de los folíolos no se tocan.

#### **Etapa reproductiva**

- **R1 - Inicio de Floración** - Se observa una flor abierta en cualquier nudo del tallo principal.
- **R2 - Floración completa** - Se observa una flor abierta en uno de los nudos superiores del tallo principal con hojas totalmente desplegadas.
- **R3 - Inicio de formación de vainas** - Una vaina de 5 milímetros de largo en uno de los 4 nudos superiores del tallo principal. Es común encontrar en un mismo momento vainas en desarrollo, flores marchitas, flores abiertas y yemas florales.
- **R4 - Vainas completamente desarrolladas** - Una vaina de 2 cm en uno de los 4 nudos superiores del tallo principal con hojas totalmente desplegadas. En esta etapa comienza el periodo crítico del cultivo; entre R4,5 y R5,5 es el momento más crítico, ya que ha finalizado la floración y cualquier situación de estrés: déficit hídrico, de nutrientes, defoliación por orugas, enfermedades foliares, ataque de chinches, granizo, etc., afectará el número final de vainas y de granos, provocando la reducción del rendimiento.

- **R5 - Inicio de formación de semillas** – En este estado se inicia el crecimiento rápido de la semilla, o llenado de granos y redistribución de nutrientes desde la planta a las semillas. A inicios de R5 el grado de desarrollo reproductivo varía desde flores recién abiertas a vainas conteniendo semillas de 8 mm de largo. En este estadio se producen las mayores tasas de fijación biológica de nitrógeno, que luego comienzan a caer abruptamente.
- **R6 - Semilla completamente desarrollada** - Una vaina, en cualquiera de los cuatro nudos superiores del tallo principal, contiene una semilla verde que llena la cavidad de dicha vaina, con hojas totalmente desplegadas. En esta etapa termina el período crítico del cultivo. En este período se puede observar semillas de todos los tamaños y se alcanza el máximo peso de vainas. Las tasas de crecimiento de la semilla y de la planta entera aún son altas. Poco después de R6 comienza la senescencia, caracterizada por el amarillamiento de las hojas.
- **R7 - Inicio de maduración** - Una vaina normal en cualquier nudo del tallo principal ha alcanzado su color de madurez. La semilla, en este momento, contiene el 60 % de humedad.
- **R8 - Maduración completa** - El 95 % de las vainas de la planta han alcanzado el color de madurez.

Luego de R8, se necesitan cinco a diez días de tiempo seco (baja humedad relativa ambiente), para que las semillas reduzcan su humedad por debajo del 15%.



Fuente: 1999. McWilliams, D.A., D.R. Berglund, and G.J. Endres. Soybean Growth and Management Quick Guide, A-1174. North Dakota State U., U. of Minnesota.

**Figura 1.4** | Estados de desarrollo vegetativo y reproductivo de la soja.

## 1.5.2. Factores que afectan el desarrollo

El fotoperíodo y la temperatura son los dos factores más importantes que afectan el crecimiento y desarrollo del cultivo de soja.

### 1.5.2.1. Fotoperíodo

Las hojas son órganos de la planta en los que se sensa el estímulo fotoperiódico que inicia la transformación de los meristemas vegetativos en reproductivos. En general, los cultivares comerciales de soja pueden recibir el estímulo fotoperiódico cuando las hojas unifolioladas se encuentran totalmente expandidas y están desplegando la primera hoja trifoliolada (Wilkerson, *et al.*, 1989).

La soja es una especie de días cortos con respuesta cuantitativa (Cregan y Hartwig, 1984). Significa que cada cultivar tiene un fotoperíodo crítico, por debajo del cual el período de emergencia floración no ve incrementada su duración por efecto fotoperiódico. Con fotoperíodos más largos que el crítico, la tasa o velocidad en el desarrollo de los órganos reproductivos se vuelve más lenta y la floración se retrasa

(Thomas y Raper (Jr), 1983a). El control fotoperiódico en soja ocurre hasta prácticamente la madurez.

El fotoperíodo varía con la latitud y con la época del año. En el Ecuador (Lat. 0°), el día y la noche tienen la misma duración, es decir de 12 horas cada uno. En los espacios geográficos alejados del Ecuador hacia los polos, los fotoperíodos se alargan en verano y se acortan en invierno. Los distintos genotipos de soja exhiben amplios rangos de sensibilidad fotoperiódica. Así se tiene desde cultivares insensibles al fotoperíodo, pasando por cultivares con fotoperíodos críticos altos, adaptados a latitudes altas (Inouye *et al.*, 1979), que florecen con fotoperíodos muy largos, hasta genotipos adaptados a bajas latitudes que florecen con fotoperíodos más cortos y que poseen alta sensibilidad fotoperiódica.

Entre estas respuestas fotoperiódicas externas existen situaciones intermedias, las que interactuando con las temperaturas permiten agrupar genotipos de similares exigencias fototérmicas. La clasificación más difundida es la originada en los EE.UU., que divide a los cultivares en grupos de madurez (GM), comenzando con el grupo 000, tolerante a las bajas temperaturas y prácticamente indiferente al fotoperíodo. En Argentina cada GM está adaptado a una franja latitudinal relativamente estrecha (200 Km). En consecuencia existe un rango de GM adaptados a cada región que se comportarán como ciclo corto, medio o largo (**Figura 1.5**). Si se traslada el cultivar a una latitud más alta que la de su rango de adaptación, el fotoperíodo natural que encontrará será relativamente más largo, con lo cual se favorecerá el crecimiento vegetativo, pudiendo resultar excesivo, mientras que la fructificación y maduración se retrasarán, con riesgo de exponer al cultivo a heladas antes de la maduración. Contrariamente, llevar un cultivar a latitudes más bajas que las de su rango de adaptación lo expondrá a fotoperíodos relativamente más cortos, que anticiparán la inducción floral. La antesis y la maduración, dando como resultado plantas de menor tamaño.



En síntesis podemos decir que los **cultivares de ciclo corto** requieren mayor stand de plantas, mejor distribución del stand de plantas, suelo con menores limitaciones físico-químicas, mayor control de plagas y malezas. Se caracterizan por presentar menor vuelco, mayor rendimiento en condiciones de alta fertilidad y disponibilidad hídrica, mayor posibilidad de escape a algunos problemas sanitarios, menor calidad de semilla debido a la mayor temperatura ambiente durante la madurez, y desocupan antes el lote (Giorda y Baigorri, 1997).

Los **cultivares de ciclo medio** en relación a los cultivares de ciclo corto, requieren menor stand de plantas y suelo con mayores limitaciones físico-químicas, para reducir su desarrollo vegetativo. Se caracterizan por presentar mayor plasticidad en fecha de siembra y densidad de siembra aún con mismo hábito de crecimiento, mayor estabilidad de rendimiento, al retrasar su llenado de granos hacia un período con menor probabilidad de ocurrencia de estrés hídrico, mayor tolerancia a errores en el control de plagas y malezas, mayor tendencia al vuelco, especialmente en campañas con buena disponibilidad hídrica y en especial en fecha de siembra del mes de noviembre, mayor predisposición a ser afectados por la podredumbre húmeda del tallo, mejor calidad de semilla (Giorda y Baigorri, 1997).

Los **cultivares de ciclo largo** en relación a los cultivares de ciclo corto y medio, requieren menor stand de plantas y suelo con mayores limitaciones físico-químicas. Se caracterizan por presentar mayor susceptibilidad al vuelco. Son los cultivares que manifiestan esta característica en campañas con buena disponibilidad hídrica y en especial en fechas de siembra del mes de noviembre y principios del mes de diciembre, mayor tolerancia a errores en el control de plagas y malezas, mejor comportamiento en suelos con limitaciones físicas y/o químicas, mayor competencia con las malezas por su mayor crecimiento, mejor comportamiento ante deficiencias en el manejo del cultivo (Giorda y Baigorri, 1997).

### 1.5.2.2. Temperatura

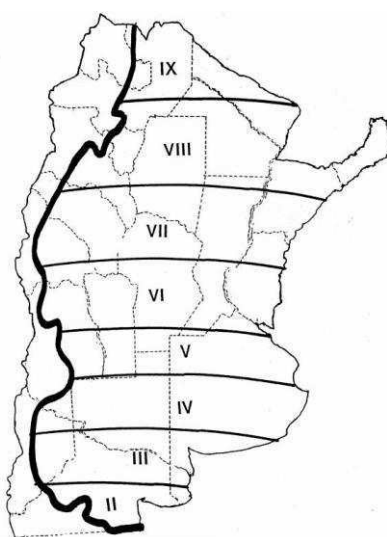
La temperatura es otro factor que incide sobre el desarrollo. Bajo períodos inductivos, los procesos se hacen más lentos, con temperaturas sub-óptimas y supra-óptimas. Las temperaturas bajo las cuales el proceso es más rápido oscilan entre 26 y 34 °C diurnos, y entre 22 y 30 °C nocturnos (Thomas y Raper (Jr), 1983b).

Las bajas temperaturas provocan una reducción en el número de primordios reproductivos y su tasa de desarrollo, estimulándose el crecimiento vegetativo (Thomas y Raper (Jr), 1983b).

### 1.5.2.3. Otros

Otro factor que afecta al cultivo es el estrés hídrico, que reduce el número de estructuras reproductivas y modifica la tasa de desarrollo hasta antesis. La magnitud de este efecto varía con el momento, la extensión y la intensidad del estrés.

La deficiencia de nutrientes, la humedad u otras condiciones de estrés en general alargan la duración de las etapas vegetativas y acortan la duración de las etapas reproductivas (Ritchie *et al.*, 1985).



Fuente: H. Baigorri, 1991 (adaptado de Remussi y Pascale, 1977). Guía práctica para el cultivo de la soja.

**Figura 1.5** | Fajas latitudinales de variedades de soja según su grupo de maduración.

## 1.6. Tipo de crecimiento

Bernard (1972) describió tres hábitos de crecimiento (HC) del tallo principal y de iniciación floral: determinado, indeterminado y semideterminado, que están controlados genéticamente (**Figura 1.6**).

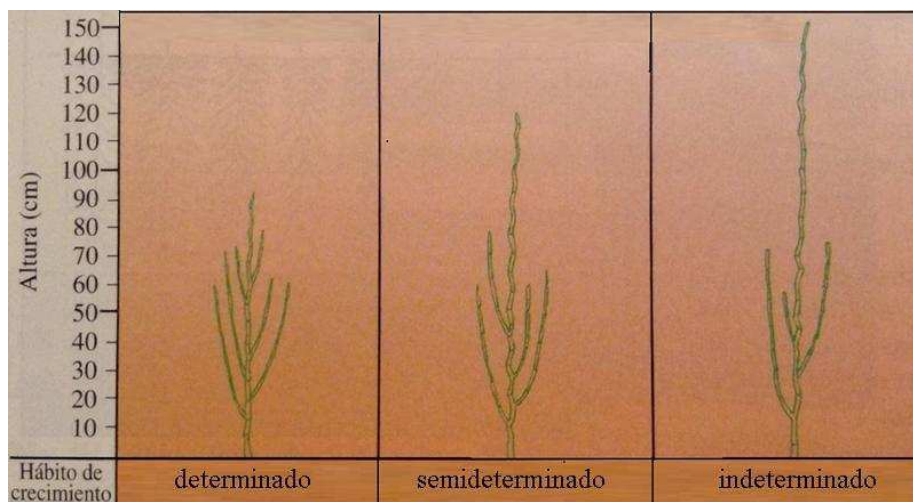
En los cultivares de **crecimiento determinado** en el tallo principal se detiene la formación de nudos poco después de iniciada la floración, en consecuencia, su crecimiento en altura es limitado. Hasta ese momento las plantas han producido la mayor parte del crecimiento vegetativo, por lo que el tiempo de superposición del crecimiento vegetativo con el reproductivo es del orden del 20% del total del ciclo de vida de la planta (Egli *et al.*, 1973; Egli *et al.*, 1985; Foley *et al.*, 1986). La mayor parte del crecimiento vegetativo entre la floración y la formación de vainas ocurre sobre las ramas. La floración comienza en la parte media del tallo principal y en menos de una semana alcanza el nudo terminal, que presenta numerosas flores.

Los cultivares de **crecimiento indeterminado**, luego del comienzo de la floración continúan desarrollando nudos sobre el tallo principal y, en consecuencia, su altura puede ser considerablemente mayor que la de los cultivares determinados de la misma longitud de ciclo y fecha de floración. El número de nudos producidos después de la floración puede ser el doble o aún más, dependiendo esta cantidad fundamentalmente del GM del cultivar, la latitud y la fecha de siembra. Hacia el ápice, el tallo reduce su diámetro y el número de vainas por nudo.

Después de la floración el crecimiento vegetativo continúa por varias semanas, existiendo una superposición con el reproductivo del 40% o más según las condiciones ambientales (Egli *et al.*, 1973; Egli *et al.*, 1985; Foley *et al.*, 1986). En estos cultivares la floración comienza en la porción inferior del tallo principal y progresa en forma ascendente a medida que se produce el alargamiento de los entrenudos superiores, pudiendo presentar vainas totalmente desarrolladas en los nudos basales y vainas recién formadas en el ápice.

Los cultivares de **crecimiento semideterminado** forman después de la floración un número de nudos intermedio entre los de hábito de crecimiento determinado e indeterminado, presentando al igual que los cultivares determinados un racimo terminal.

En el mejoramiento genético del cultivo se han seleccionado, generalmente, cultivares con hábito de crecimiento indeterminado en los GM 000 al IV y cultivares con hábito de crecimiento determinado en los GM V al X.



Fuente: Giorda y Baigorri, 1997. INTA

**Figura 1.6** | Hábitos de crecimiento de la soja.

## 1.7. Importancia del nitrógeno en los ecosistemas. Ciclo biogeoquímico

La atmósfera contiene casi un 80% de N en forma gaseosa esto es  $N_2$ , y aun cuando el N es clave para el crecimiento de las plantas, en este estado no puede ser utilizado por la mayoría de los seres vivos (a excepción de algunas bacterias), debido al fuerte triple enlace que une sus átomos.

El ciclo biogeoquímico del nitrógeno (**Figura 1.7**) incluye todos aquellos procesos bióticos y abióticos implicados en el equilibrio dinámico de este elemento en la biosfera, propiciando su biodisponibilidad en formas químicas asimilables para los seres vivos. En el ciclo del nitrógeno en el suelo, los procesos de asimilación, amonificación y nitrificación generan amonio ( $\text{NH}_4^+$ ) y nitrato ( $\text{NO}_3^-$ ), sustancias que son solubles en agua y por lo tanto son lavadas, principalmente el  $\text{NO}_3^-$ , por la escorrentía y la infiltración. Los procesos de fijación de nitrógeno atmosférico y desnitrificación completan y equilibran el ciclo, de manera que la atmósfera constituye un depósito de nitrógeno importante.

### 1.7.1. Procesos implicados en el ciclo del nitrógeno en los sistemas terrestres

#### Fijación

Las entradas de nitrógeno mineral al suelo se producen desde la atmósfera a través del agua de la lluvia, en forma de amonio o de óxidos de nitrógeno disueltos, o también a través de la fijación por parte de bacterias que en forma libre o en simbiosis como la que establecen los rizobios con plantas leguminosas y cianobacterias, que usan el  $\text{N}_2$  del aire y lo convierten en amonio (**Figura 1.7**). La fijación simbiótica del nitrógeno supone la entrada de nitrógeno más importante para los ecosistemas terrestres (Tamm, 1991). Eventos de alta energía natural (relámpagos, los fuegos forestales, flujos de lava) pueden oxidar también el nitrógeno atmosférico. Las emisiones de óxidos de nitrógeno procedentes del uso de combustibles fósiles constituyen otra fuente potencial de fijación de nitrógeno atmosférico.

#### Asimilación

El N del suelo como amonio y/o nitrato pueden ser absorbido por las plantas, incorporándose, de este modo, a la materia orgánica. La mayor parte del nitrógeno mineral del suelo se encuentra en forma de nitrato, que es la forma más oxidada de

nitrógeno y, también, la más comúnmente asimilada por los vegetales. La otra forma de N en el suelo es el  $\text{NH}_4^+$ . La asimilación de estas dos formas es diferencial, ya que el ión nitrato se encuentra disuelto en la solución del suelo, mientras que gran parte del ión amonio se adsorbe a las superficies de las arcillas (Pacheco *et al.*, 2002). El dominio en el suelo de coloides cargados negativamente que ocupan los espacios interlaminares de los minerales de la arcilla propicia la adsorción del ión amonio, reduciendo su movilidad en el suelo.

### **Amonificación**

La mineralización del N orgánico del suelo consiste en la descomposición de la materia orgánica que provoca la actividad de las bacterias y hongos del suelo que cuyo resultado es la generación de amonio, estado de oxidación en el que el N vuelve a estar disponible para la nutrición de los organismos vegetales y algunos tipos de bacterias (Keeney y Walsh, 1972). Este proceso es importante, por cuanto la mayor parte del nitrógeno presente en el suelo está formando parte de la materia orgánica (alrededor del 90%) (Pacheco *et al.*, 2002), integrado a la biomasa viva o a la materia orgánica detrítica (restos orgánicos de las plantas, excrementos animales y humus).

### **Nitrificación**

El amonio presente en el suelo sufre dos procesos sucesivos de oxidación como resultado de la actividad de ciertos grupos de bacterias que actúan en condiciones aerobias: bacterias del género *Nitrosomonas*, que oxidan el  $\text{NH}_4^+$  a nitrito ( $\text{NO}_2^-$ ), y bacterias del género *Nitrobacter*, que oxidan el  $\text{NO}_2^-$  a nitrato ( $\text{NO}_3^-$ ), en un proceso denominado nitrificación (Sharma y Ahlert, 1977; Watson *et al.*, 1981; Schmidt, 1982; Wetzel, 2001). Este proceso requiere la presencia de oxígeno y ocurre rápidamente a temperaturas entre 15 y 30°C y pH entre 6.5 y 7.5 (Keeney y Walsh, 1972). Otros factores que regulan la nitrificación son la concentración de amonio, la concentración de dióxido de carbono, la relación C:N y el potencial redox del suelo.

### **Desnitrificación**

Es el proceso que tiene lugar por la acción de bacterias reductoras (anaerobias facultativas) como algunas bacterias del género *Pseudomonas*, que bajo condiciones anaeróbicas (suelos encharcados) transforman el nitrógeno mineral del suelo en  $N_2$  (Tiedje *et al.*, 1984; Austin, 1988; Stumm y Morgan, 1996; Wetzel, 2001; Pacheco *et al.*, 2002; Paerl *et al.*, 2002) y, en menor medida, en óxidos nitrosos. La desnitrificación se produce cuando las concentraciones de carbono soluble en el suelo son mayores que  $2 \text{ mg l}^{-1}$  y la de oxígeno disuelto menores de  $2 \text{ mg l}^{-1}$  (Thomasson *et al.*, 1991). Este proceso se ve favorecido a temperaturas de  $25 \text{ }^\circ\text{C}$  o superiores, pH entre 6 y 8, bajo drenaje del suelo y existencia de altas concentraciones de compuestos orgánicos que actúen como donadores de electrones (Tiedje *et al.*, 1984; Postma *et al.*, 1991; Korom, 1992).

### **Volatilización**

Es el proceso por el cual el nitrógeno mineral en forma de ión amonio ( $NH_4^+$ ) de la superficie del suelo pasa a la atmósfera en forma de amoníaco ( $NH_3$ ). Está regulado fundamentalmente por el pH del suelo, al influir en la ecuación de equilibrio entre el amonio y el amoníaco (Russo, 1985). Las condiciones de alcalinidad elevada en el suelo conducen a que una mayor proporción de N se encuentre como  $NH_4^+$  en el suelo. Otros factores implicados son la capacidad de intercambio catiónico, el contenido en materia orgánica, la textura del suelo y la temperatura.

### **Lixiviación**

Es el proceso por el cual el nitrógeno inorgánico es arrastrado por el agua de infiltración del suelo hacia los horizontes más profundos del mismo y, finalmente, hacia los acuíferos subyacentes. Las pérdidas por este proceso ocurren, normalmente, en forma de nitrato ya que este compuesto no forma asociaciones estables con las partículas edáficas. El amonio, por el contrario, puede ser atraído por las partículas del suelo cargadas negativamente, como la arcilla y la materia orgánica particulada, y de esta forma es resistente a la percolación (Pacheco *et al.*, 2002).

El riesgo de lixiviación de nitrógeno en condiciones naturales se ve limitado por la escasez de nitrógeno mineral en los suelos y el equilibrio existente entre su biodisponibilidad y la absorción de las plantas. Sin embargo, el impacto de las actividades humanas ha modificado el ciclo del nitrógeno, con un mayor aporte de nitrógeno mineral y orgánico al suelo (uso de fertilizantes), un incremento del contenido de compuestos nitrogenados reactivos en la atmósfera, susceptibles de ser fijados y un mayor riesgo de lavado por las prácticas de riego. Entre todas las actividades de origen antrópico, la aplicación de fertilizantes agrícolas constituye la mayor alteración del ciclo del nitrógeno a escala global (Townsend *et al.*, 2003).

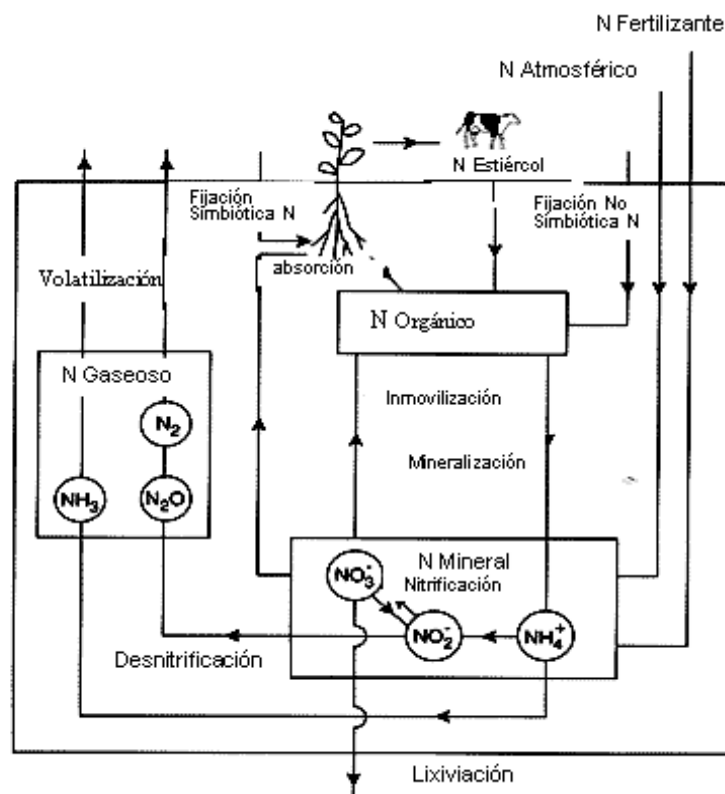


Figura 1.7 | Ciclo del nitrógeno. Fuente: Rowell (1996).



## 1.8. Fijación Biológica de Nitrógeno rizobio-leguminosas

Los rizobios son bacterias del suelo capaces de formar nódulos y establecer simbiosis con las raíces o tallos de plantas leguminosas. Durante el proceso simbiótico, los rizobios reducen el nitrógeno atmosférico en una forma directamente asimilable por las plantas (amonio).

Este tipo de interacción es una de las relaciones mutualistas más estudiada y de mayor importancia en el mundo por la cantidad de N que aporta a los ecosistemas naturales y agrícolas, aporte que puede alcanzar a  $150 \text{ kg N ha}^{-1} \text{ año}^{-1}$  en el suelo, cantidad que es superior a lo aportado por los fertilizantes, pero adicionalmente no provocan alteraciones ambientales (Labandera, 2007).

### 1.8.1. Taxonomía de rizobios

La taxonomía de rizobios está basada en aproximaciones polifásicas (Graham *et al.*, 1991), es decir, incluye características morfológicas y bioquímicas, perfiles genéticos ("fingerprint") (Versalovic *et al.*, 1994), análisis de metil éster de ácidos grasos (FAME) (Jarvis y Tighe, 1994; Tighe *et al.*, 2000) y análisis de secuencias del gen 16S rRNA (Laguerre *et al.*, 1994; Ludwig *et al.*, 1998). Estudios recientes han mostrado la existencia de una gran diversidad de bacterias fijadoras de nitrógeno aisladas de diferentes leguminosas. A la fecha, se describen más de 98 especies pertenecientes a 14 géneros de  $\alpha$ - y  $\beta$ -proteobacteria que se han descrito como rizobios. Estos son: *Rhizobium*, *Mezorhizobium*, *Ensifer* (antes *Sinorhizobium*), *Bradyrhizobium*, *Phyllobacterium*, *Microvirga*, *Azorhizobium*, *Ochrhobactrum*, *Methylobacterium*, *Devosia*, *Shinella* (Clase  $\alpha$ -proteobacteria), *Burkholderia*, *Cupriavidus* (antes *Ralstonia*) (Clase  $\beta$ -proteobacteria) y algunas  $\gamma$ -proteobacteria, forman el conjunto de bacterias conocidas como simbioses de leguminosas (Berrada y Fikri-Benbrahim, 2014).

La fijación simbiótica de nitrógeno puede suministrar todo o parte del N que requiere la leguminosa para su desarrollo (Racca y Collino, 2005; Moreno, 2010). La asociación entre rizobios y leguminosas es específica y depende de las interacciones moleculares que ocurren en las etapas iniciales de la formación de los nódulos (Debellé *et al.*, 2001); en general cada rizobio coloniza un tipo de leguminosa.

En el caso de la soja, cultivo muy importante a nivel mundial que es usado principalmente como fuente de proteína y biocombustibles, es una leguminosa que forma nódulos luego de interactuar con rizobios específicos. Los principales rizobios noduladores de soja son *Bradyrhizobium japonicum* (Jordan, 1982) *Bradyrhizobium elkani* (Kuykendall *et al.*, 1992), y *Sinorhizobium (Ensifer) fredii* (Chen *et al.*, 1988). Además *Bradyrhizobium liaoningense* (Xu *et al.*, 1995), *Sinorhizobium xinjiangense* (Chen *et al.*, 1988), y *Mesorhizobium tianshanense* (Jarvis *et al.*, 1997) han sido clasificados también como rizobios noduladores de soja.

### **1.8.2. Descripción del género *Bradyrhizobium***

La clasificación taxonómica permaneció inalterada hasta principios de 1980, cuando la taxonomía numérica consideraba patrones morfológicos, fisiológicos y genéticos posicionando algunas cepas de rizobios en un nuevo género, *Bradyrhizobium*, con una especie única definida *B. japonicum* (Jordan, 1982).

Las bacterias del género *Bradyrhizobium* incluyen bacterias de crecimiento lento (Jordan, 1982) que tienen un tiempo generacional de 10, 12 h. Son bacilos con un solo flagelo polar o subpolar. Las colonias no exceden 1 mm de diámetro en medio YMA después de 5-10 días de incubación a 28°C (Vincent, 1970). Estas bacterias simbióticas usan como fuente de carbono azúcares y ácidos orgánicos pero prefieren las pentosas y cuando se cultivan en medios ricos producen y exudan polisacáridos que le dan una apariencia mucosa brillante a las colonias.

La diversidad creciente que se detectó entre cepas de *Bradyrhizobium* hizo que se crearan nuevas especies, lo que se vio acompañado del desarrollo de técnicas moleculares que contribuyeron a identificar la alta diversidad entre cepas de *Bradyrhizobium* aisladas en un amplio rango de leguminosas (por ej.: Germano *et al.*, 2006; Menna *et al.*, 2006; Menna *et al.*, 2009).

La secuenciación del 16S rRNA se ha vuelto un método alternativo para trazar filogenias bacterianas (Woese, 1987; Weisburg, *et al.*, 1991; Garrity *et al.*, 2002), pero en varios géneros incluyendo *Bradyrhizobium*, la variabilidad en el 16S rDNA es a menudo bajo y puede no reflejar la diversidad detectada por otros parámetros morfo-fisiológicos y genéticos (por ej.: Vinuesa *et al.*, 1998; van Berkum y Fuhrmann, 2000; Willems *et al.*, 2001; Germano *et al.*, 2006; Menna *et al.*, 2009). Para detectar una mayor diversidad, otros genes ribosomales o regiones han sido elegidas, como el 23S rDNA y la región intergénica transcrita entre el 16S-23S rDNA (ITS), ya que ellos evolucionan a una tasa más rápida que el 16S rDNA, sumando información al análisis (por ej.: Tesfaye y Holl, 1998; Vinuesa *et al.*, 1998; van Berkum y Fuhrmann, 2000; Willems *et al.*, 2001; Menna *et al.*, 2009). Sin embargo la proximidad de los genes ribosomales puede no indicar adecuadamente la filogenia correcta en el caso de la transferencia horizontal de genes (van Berkum *et al.*, 2003). Así se han descrito hasta la fecha 23 especies dentro del género *Bradyrhizobium*<sup>3</sup>.

El método de análisis de secuencia multilocus (MLSA) es cada vez más usado para estudios de filogenia y taxonomía. El método consiste en el análisis de varios genes conservados (“housekeeping genes”) dispersos en al menos 100 Kb del genoma (Brett *et al.*, 1998; Maiden *et al.*, 1998; Godoy *et al.*, 2003; Cooper y Feil, 2004). De esta manera con el análisis MLSA se han definido con éxito diversos grupos filogenéticos de rizobios (Moulin *et al.*, 2004; Menna *et al.*, 2009; Ribeiro *et al.*, 2009; Delamuta *et al.*, 2012).

---

<sup>3</sup> <http://www.dsmz.de/bacterial-diversity/prokaryotic-nomenclature-up-to-date/prokaryotic-nomenclature-up-to-date.html>

Como se dijo anteriormente, dentro de este género las tres especies simbiotes más representativas que nodulan *Glycine max* son: *Bradyrhizobium japonicum* (Jordan, 1982), *B. elkanii* (Kuykendall *et al.*, 1992) y *B. liaoningense* (Xu *et al.*, 1995). *B. japonicum* y *B. elkanii* son las especies más ampliamente utilizadas como inoculantes comerciales para incrementar los rendimientos de los cultivos (**Figura 1.8**).

Otras especies nodulan otras leguminosas, algunos ejemplos son una nodulante de *Lespedeza cuneata* [*Bradyrhizobium yuanmingense* (Yao *et al.*, 2002)], otra de arbustos de las tribus *Genisteeae* y *Loteae* [*Bradyrhizobium canariense* (Vinuesa *et al.*, 2005a)], *Bradyrhizobium betae* endófito de remolacha (*Beta vulgaris* L.) (Rivas *et al.*, 2004) y dos especies que nodulan *Pachyrhizus erosus* (*Bradyrhizobium pachyrhizi* y *B. jicamae*) (Ramírez-Bahena *et al.*, 2009). Algunos bradyrizobios fotosintéticos fueron también reportados como endófitos de arroz silvestre africano (Chaintreuil *et al.*, 2000).



**Figura 1.8** | Colonias de *Bradyrhizobium japonicum* cepa E109 en medio agarizado YEM con rojo congo (Vincent, 1970). Colonias de 7 días de crecimiento. Arriba se puede observar detalle de las colonias.

### 1.8.3. Infección y nodulación

El establecimiento de la simbiosis entre el rizobio y la leguminosa es el resultado de tres eventos principales: i) infección que sufren las células del hospedante por el microsimbionte; ii) la organogénesis del nódulo; y iii) el proceso de fijación de  $N_2$ : Los primeros dos ocurren simultáneamente, mientras que la fijación de  $N_2$  ocurre luego de que se completó la organogénesis del nódulo (**Figura 1.9**).

La simbiosis es un proceso que se inicia con un diálogo químico cuando interaccionan el rizobio y la leguminosa hospedadora. Las raíces de las plantas secretan compuestos fenólicos, conocidos como flavonoides que son reconocidos por la proteína sensora bacteriana NodD, lo que activa la transcripción de los operones de nodulación (genes *nod*, *nol* o *noe*). Los genes de nodulación codifican enzimas involucradas en la síntesis de moléculas señales lipoquitoligosacáridos (LCOs) bacterianos conocidas como factores Nod. Estas consisten en un esqueleto oligomérico de quitina acetilada con varias sustituciones en los residuos terminales o no terminales no reductores. Cada especie de rizobio posee un conjunto de genes de nodulación, algunos de los cuales (*nodABCIJ*) son comunes a los rizobios en general, cumplen en diversas especies bacterianas la misma función. Las proteínas NodA, NodB y NodC cumplen roles claves en la síntesis del esqueleto del LCO, mientras que NodIJ están involucrados en su secreción. Este esqueleto es luego “decorado” por sustituciones químicas, las cuales determinan la especificidad de la interacción con la leguminosa hospedadora, estas modificaciones son llevadas a cabo por proteínas codificadas por los genes de especificidad, los que muchas veces son alélicos. Las plantas hospedantes por otro lado presentan receptores específicos de los factores Nod de sus pares simbióticos compatibles, un sistema fundamental molecular de tipo “llave-cerradura” para la especificidad por el hospedante en la simbiosis *Rhizobium*-leguminosa (Spaink, 1994; Schultze y Kondorosi, 1995; Dénarié *et al.*, 1996; Long,

1996; Mergaert *et al.*, 1997; Downie, 1998; Broughton *et al.*, 2000; Kobayashi y Broughton, 2008).

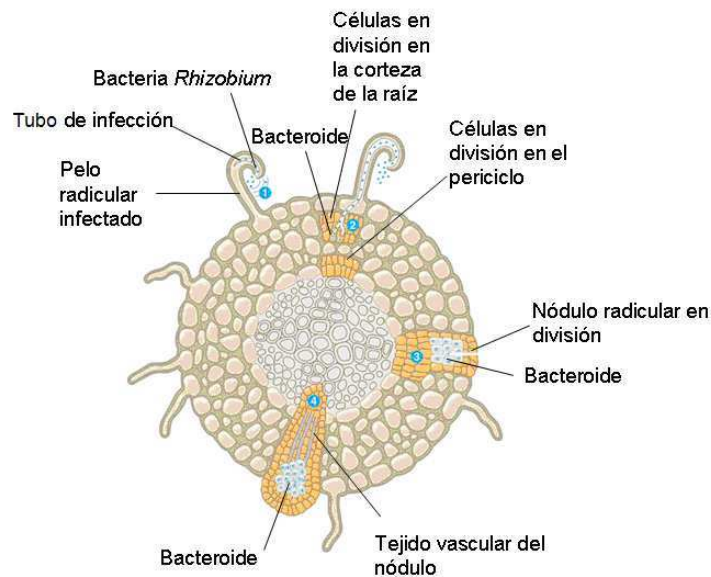
El reconocimiento de los factores Nod por los receptores de las plantas abre la puerta para el comienzo de la infección. Por ejemplo los genes cuyo producto son NFR1/NFR5 son genes situados en la cima de la cascada y mutaciones en estos genes han mostrado falta de respuesta al factor Nod, éste codifica para un complejo receptor transmembrana con actividad kinasa. Se piensa que una parte específica de las proteínas kinasa se une al esqueleto N-acetilglucosamina del factor Nod e inicia el proceso de nodulación (Indrasumunar *et al.*, 2010; 2011, Hayashi *et al.*, 2012). Luego, los reconocimientos conducen a la ocurrencia de fenómenos de transformaciones morfológicas y fisiológicas, los pelos radiculares se curvan y los niveles de calcio intracelular en la cadena de transducción de señales son alterados (Oldroyd y Downie, 2008). De esta manera una bacteria asociada a la superficie del pelo radicular puede quedar atrapada al curvarse el mismo, se genera por invaginación de la membrana de las células del hospedante un canal que es el hilo de infección. Éste guía al rizobio dentro de los tejidos de la planta y los conduce hacia lo que será el centro fisiológico de la fijación de nitrógeno. Al mismo tiempo, los factores Nod activan a las células corticales opuestas al sitio de infección, generando su desdiferenciación y división, y la formación del primordio del nódulo. Dependiendo de la planta hospedante, el primordio se origina en la corteza interna (para nódulos indeterminados) o en la corteza media/externa (para nódulos determinados) que es el caso de la soja. Cuando el hilo de infección alcanza las células del primordio, las bacterias son liberadas dentro de sus citoplasmas, sitio en el cual las bacterias modifican su morfología motivo por el cual toman el nombre de “bacteroides”. Una vez que el primordio generó un número importante de nuevas células, gran parte de éstas son infectadas por las bacterias que llegan por los hilos infectivos. Las bacterias en el interior de las células se rodean de una membrana y forman el simbiosoma que es la unidad funcional de fijación de N. El simbiosoma desplaza hacia los costados al citoplasma y organelas de las células

infectadas, estas por otro lado están acompañadas en el nódulos por células no infectadas que es donde se metaboliza el N fijado (Maunoury *et al.*, 2008) (**Figura 1.9**).

En los nódulos indeterminados, la actividad de división celular es mantenida en la región distal del primordio, el cual genera un meristema apical activo durante toda la vida del nódulo. La producción de nuevas células meristemáticas es compensada por la salida del mismo número de células del ciclo de división celular. A medida que los pelos de infección proliferan debajo del meristema y se introducen en estas células post-mitóticas, el proceso de infección y diferenciación de las células simbióticas ocurre permanentemente en la región apical del nódulo. La diferenciación de las células infectadas da lugar a una renovación continua del pool de células fijadoras de nitrógeno. Al mismo tiempo, las células fijadoras viejas pierden actividad y entran en senescencia. El crecimiento del nódulo indeterminado resulta del incremento constante del número celular y en el agrandamiento de las células infectadas. Este tipo de nódulos es mas frecuente en leguminosas de climas templados, como la alfalfa o la vicia que desarrollan nódulos en simbiosis con *Sinorhizobium meliloti* o *Rizobium leguminosarum* bvar *viciae*, respectivamente. El escenario es diferente para los nódulos determinados, en el primordio del nódulo, toda división celular cesa rápidamente y si bien también se forma un meristema, éste se mantiene activo por un período de tiempo corto por ello; el número de células en los nódulos jóvenes y viejos es el mismo. En las células infectadas, la bacteria se divide repetidamente y las células del hospedante se agrandan para hospedarlas. La fijación de N<sub>2</sub> comienza sincronizadamente en todas las células infectadas, antes que el agrandamiento celular se haya completado. El crecimiento de los nódulos determinados es principalmente dependiente de la multiplicación temprana ocurrida y del agrandamiento de las células infectadas. Cuando la diferenciación celular finaliza, el crecimiento se detiene, todo este proceso resulta en la generación de nódulos esféricos conteniendo una sola zona homogénea central de fijación de N<sub>2</sub>. La fijación de nitrógeno ocurre durante un período de tiempo debido a que no se renuevan las células en donde se fija el

nitrógeno; la vida del nódulo finaliza con la senescencia en forma radial que ocurre desde el centro hacia afuera. Las células de la planta se desintegran, y liberan los bacteroides los que en muchos casos revierten a la forma de bacilo con la que se los encuentra en los suelos. Este tipo de nódulos es más frecuente en especies de leguminosas “tropicales” como la soja (Maunoury *et al.*, 2008).

Dentro del nódulo la fisiología de la bacteria cambia notablemente. Ocurre un proceso de diferenciación en el que se expresan los genes que codifican las proteínas que conforman la maquinaria de fijación biológica de nitrógeno. Entre ellas, la más importante es la nitrogenasa, que es la enzima que transforma el nitrógeno atmosférico en amonio.



Fuente: Campbell y Reece (2007)

**Figura 1.9** | Desarrollo del nódulo radicular. **1.** La raíz exuda flavonoides que atraen a los rizobios. La bacteria emite señales que estimulan el alargamiento de los pelos radiculares que forman un hilo de infección por invaginación de la membrana plasmática. **2.** Las bacterias penetran en la corteza dentro del tubo de infección. Las células de la corteza y del periciclo comienzan a dividirse, el hilo de infección se ramifica y las vesículas que contienen bacterias forman brotes en las células corticales. Mediante este proceso se forman los bacteroides. **3.** El crecimiento continúa en la región afectada de la corteza y el periciclo y estas dos masas de células en división se fusionan y forman el nódulo. **4.** El nódulo desarrolla tejido vascular que le suministra nutrientes y transporta compuestos nitrogenados hacia el cilindro vascular para su distribución en toda la planta.



### 1.8.4. Fijación biológica de nitrógeno.

#### 1.8.4.1. *Nitrogenasa*

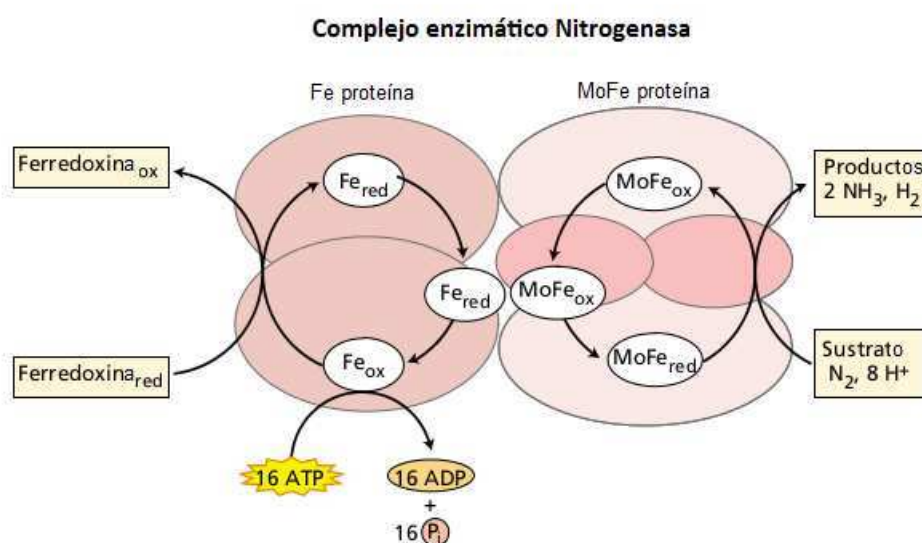
Dentro del nódulo, los bacteroides comienzan con el proceso de fijación biológica de  $N_2$  a  $NH_4^+$ , el cual es catalizado por la enzima nitrogenasa (Mylona *et al.*, 1995). Este complejo enzimático rompe la triple ligadura del  $N_2$  atmosférico que es reducido a  $NH_4^+$  en condiciones de temperatura y presión normales utilizando la energía metabólica del ATP (Becana y Bedmar, 1991). El complejo enzimático está formado por 2 metaloproteínas, es sintetizado como resultado de la expresión de los genes *nif* los cuales están altamente conservados en secuencia y estructura en los organismos fijadores de nitrógeno (Dean y Jacobsen, 1992).

La nitrogenasa reductasa o Fe-proteína (complejo II), es la más pequeña de los dos componentes y está formada por un dímero  $\alpha_2$  de 62 kDa codificado por el gen *nifH*. Tiene cuatro centros Fe-S que participan en las reacciones redox implicadas en la conversión del nitrógeno a amonio. La dinitrogenasa, o Fe-Mo-proteína (complejo I), es de mayor tamaño, es un tetramero  $\alpha_2\beta_2$  de 240 kDa codificado por los genes *nifD* y *nifK*. Cada una de estas subunidades tiene dos centros Mo-Fe-S y un número variable de centros Fe-S. El mecanismo comienza cuando los donadores de electrones como la ferredoxina o flavodoxina reduce la Fe-proteína, que a su vez reduce a la Fe-Mo-proteína, ésta utiliza estos electrones para reducir el  $N_2$  a  $NH_4^+$ . La transferencia de electrones desde el componente reductasa al componente nitrogenasa está acoplada a la hidrólisis de ATP por parte de la reductasa. Este proceso tiene una demanda energética muy alta, necesiándose, si el proceso fuese 100% eficiente, 16 moléculas de ATP por cada molécula diatómica de nitrógeno atmosférico asimilado (**Figura 1.10**). La siguiente ecuación muestra la conversión del  $N_2$  atmosférico a amonio (Taiz y Zeiger, 2010):



La nitrogenasa se encuentra en el citoplasma del bacteroide. La leghemoglobina es una de las proteínas más abundantes en los nódulos de las leguminosas fijadoras. Estas se encuentran en el citoplasma y sus funciones son proveer una adecuada demanda de oxígeno para que ocurra la respiración del bacteroide que es una fuente de energía y proteger a la nitrogenasa sensible a las tensiones de oxígeno (Lee *et al.*, 1983).

Una vez reducido, el nitrógeno puede atravesar membranas por difusión y ser asimilado por la planta o bien ser transaminado por el bacteroide a esqueletos carbonados; el metabolismo de carboxilatos es muy importante en la simbiosis y se ha observado un importante flujo de ácidos di o tri-carboxílicos hacia el bacteroide (Lodwig *et al.*, 2003).



**Figura 1.10** | Esquema del mecanismo de acción del complejo enzimático nitrogenasa. La ferredoxina reduce la Fe-proteína. Unión e hidrólisis de ATP a la Fe-proteína provocando un cambio conformacional en dicha proteína que facilita las reacciones redox. La Fe-proteína reduce la FeMo-proteína y esta reduce el N<sub>2</sub>. Fuente: Taiz y Zeiger, 2010.

## 1.9. Genética de la nodulación en el hospedante

La formación del nódulo es el resultado de la expresión de genes específicos. En la leguminosa hospedante han sido estudiados genes involucrados en el proceso de infección y la nodulación. Las plantas modelos más utilizadas para estos estudios han sido *Lotus japonicus* y *Medicago truncatula*, de las cuales se han obtenido mutantes deficientes en alguno de los procesos iniciales de la nodulación (curvamiento del pelo radicular, formación del hilo de infección, formación del primordio nodular). Además se han desarrollado mapas genéticos y físicos de estas dos leguminosas. El primer gen de leguminosa asociado a la nodulación clonado en *L. japonicus* fue el gen *nin-1* (non nodulating mutant), que codifica un factor de transcripción que regula la expresión génica durante la formación del hilo de infección (Schauser *et al.*, 1999). Los genes que codifican los receptores de la planta responsables del sensado del factor Nod fueron clonados en *L. japonicus*, *LjNFR1* y *LjNFR5* (nod-factor receptor) y en *M. truncatula* *MtLYK3* y *MtLYK4* (lysine motif-receptor-like kinase) (Madsen *et al.*, 2003; Radutoiu *et al.*, 2003; Limpens *et al.*, 2003). Los genes que son activados inmediatamente después de la interacción con el factor Nod, denominados *DM1*, *DM2* y *DM3* (doesn't make infections) y *NSP1* y *NSP2* (nodulation signal pathway), también fueron aislados en mutantes de *M. truncatula* (Catoira *et al.*, 2000). Estos genes forman parte de la vía de transducción de señales asociada a la expresión de otros genes involucrados en la formación del nódulo.

En soja, estudios con mutantes que presentaban alteraciones en los patrones de nodulación (número y distribución de nódulos en las raíces), denominados supernodulantes condujeron al clonado del gen *nark* (nodule autoregulation receptor kinase), que codifica una proteína receptora del tipo kinasa. Estas proteínas receptoras poseen un dominio extracelular rico en leucina y un dominio kinasa citoplasmático, siendo algunas de sus funciones la recepción de hormonas, sensores involucrados en la respuesta frente a hongos y activación del desarrollo de la parte

aérea de la planta y de los meristemas florales y apicales (Searle *et al.*, 2003). Un gen ortólogo fue clonado en *L. japonicus* denominado *har1* (hypernodulation aberrant root) (Nishimura *et al.*, 2002). Gmnrk y Ljhar1 definieron el mismo gen y encontraron que está involucrado en el control de la proliferación celular del meristema nodular y por lo tanto en la autorregulación de la nodulación, a través de la señalización entre las hojas y la raíz. Se propuso un modelo para el control de la nodulación que considera un mecanismo de autorregulación donde los primordios nodulares, que comienzan a formarse poco después del sensado del factor Nod, producen una señal que es percibida en las hojas, que a su vez producen otra señal, que actúa como inhibidor del desarrollo de nuevos nódulos (Gresshoff, 2003).

Los genes de la planta cuya expresión es inducida por el factor Nod, son denominadas nodulinas, porque codifican proteínas requeridas para el desarrollo de los nódulos y el mantenimiento de la fijación biológica de nitrógeno (FBN). De acuerdo al patrón de expresión, las nodulinas se agrupan en tempranas y tardías. Las nodulinas tempranas son las que se expresan inmediatamente después que el rizobio interactuó con la planta y están involucradas fundamentalmente en la formación de la estructura nodular (división de las células de la corteza, síntesis del cordón de infección). Las nodulinas tardías se expresan en los nódulos cuando ya se han formado y están involucrados en el metabolismo y en el mantenimiento o senescencia de los nódulos (Stougaard, 2000).

Muchas nodulinas tempranas son proteínas de la pared celular, es decir que forman parte de la constitución bioquímica de la pared durante la infección y el desarrollo del nódulo. Ejemplos incluyen proteínas ricas en prolina, Enod2, Enod5, Enod10, Enod11, Enod12, PRP4, así como extensinas y proteínas ricas en glicina GRPs) (Schultze y Kondorosi, 1998).

La nodulina tardía más abundante en los nódulos activos es la leghemoglobina, esta regula la tensión de O<sub>2</sub> en el nódulo. Otras nodulinas tardías, como la glutamina sintetasa (GS) y la glutamato sintetasa (GOGAT), actúan en la asimilación del

nitrógeno fijado ( $\text{NH}_3$ ), produciendo glutamina y glutamato, respectivamente. Esos dos compuestos nitrogenados son utilizados en la síntesis de amidas y ureídos, que son exportados a otros órganos de la planta. Muchas leguminosas de clima templado, que pertenecen a géneros como *Lupinus*, *Medicago*, *Pisum*, *Trifolium* y *Vicia* exportan el N fijado como asparagina y/o glutamina. Mientras que las leguminosas de clima tropical que incluyen géneros como *Glycine*, *Phaseolus* y *Vigna*, exportan el nitrógeno fijado principalmente como ureídos, incluyendo alantoína y ácido alantoico (Tajima *et al.*, 2004).

La construcción del banco de datos de ESTs (Expressed Sequence Tags) y el uso de técnicas de expresión génica como microarreglos de cDNA condujo a la identificación de patrones globales de expresión génica durante los procesos específicos del desarrollo de las plantas. ESTs específicos de nódulos fueron identificados en *M. truncatula* y *L. japonicus* (Fedorova *et al.*, 2002; Journet *et al.*, 2002). El empleo de ESTs en el estudio de la simbiosis en soja permitió identificar genes que se expresan diferencialmente durante la formación de nódulos. Lee *et al.* (2004) identificaron 3264 ESTs que se expresan exclusivamente en las células de los nódulos de soja, incluyendo genes del metabolismo de C, de N, de vitaminas y de cofactores, del metabolismo secundario, del control del ciclo celular y de la expresión de nodulinas conocidas. Jeong *et al.* (2006) realizaron un screening en los bancos de datos de EST de soja para identificar genes específicos de los nódulos a través de los perfiles de análisis de expresión *in silico*. Ellos identificaron 156 ESTs que se expresan exclusivamente en los nódulos de soja y mapearon dos nodulinas: GmPGN (poligalactouronasa específica de nódulo) y GmEKN (E-ácido glutámico, K-lisina, N-nodulina) en los grupos de ligamiento B1 y A2, respectivamente.

Estudios de expresión génica globales permiten la identificación de genes que son inducidos o reprimidos en una condición experimental dada como en diferentes estadios del desarrollo de los nódulos, permitiendo proponer posibles funciones para estos genes de acuerdo con las informaciones disponibles en los bancos de datos

públicos entre los cuales se encuentran el NCBI (National Center for Biotechnology Information). A esto hay que sumarle el banco de datos del genoma de soja (Schmutz *et al.*, 2010) recientemente secuenciado disponible en <http://www.phytozome.net/soybean>. Los trabajos de generación y caracterización de mutantes complementado con clonado basado en mapeo permite la identificación y el estudio de la función de genes específicos. Otro abordaje interesante a ser explorado es la asociación de polimorfismos naturales de esos genes con una variación cuantitativa de la capacidad de nodulación observada en las leguminosas, la cual será abordada a continuación.

## 1.10. Marcadores moleculares

Los marcadores moleculares son herramientas usadas en el mapeo de caracteres cuantitativos (que se describirá más adelante).

Un marcador molecular es un segmento de ADN con una ubicación física identificable (locus) en un cromosoma y cuya herencia genética se puede seguir en la descendencia. Un marcador puede ser un gen, o puede ser un fragmento del ADN sin función conocida. Estos marcadores suelen ocurrir repetidamente en el genoma ya que surgen como diferentes clases de mutaciones, tales como sustituciones (mutación puntual), rearreglos (inserciones y deleciones) o errores en la replicación de DNA repetido en tandem (Paterson, 1996). Estos marcadores no están sometidos a presión selectiva, o sea que son neutrales porque usualmente se encuentran localizados en regiones no codificantes del genoma. A diferencia de los marcadores morfológicos o bioquímicos, los marcadores de ADN son prácticamente ilimitados en número y no son afectados por factores ambientales y/o por diferentes estadios de desarrollo de la planta (Winter y Kahl, 1995). En base a su método de detección, los marcadores de ADN se pueden dividir en tres clases: (1) basados en hibridación (RFLP), (2) basados en amplificación por la reacción en cadena de la polimerasa (PCR) (RAPD, SSR) y (3)

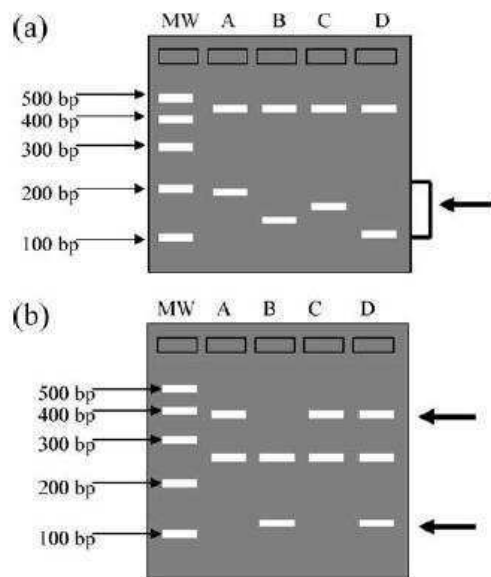
en secuenciación (SNP) (Gupta *et al.*, 1999). El marcador AFLP como se describirá más adelante está basado en una amplificación selectiva de regiones conservadas entre sitios de restricción de enzimas. Esencialmente, los marcadores de ADN revelan diferencias que se visualizan por medio de la técnica llamada electroforesis en gel y teñida con químicos (bromuro de etidio o plata) o bien pueden detectarse por medio de sondas colorimétricas o radiactivas. Los marcadores son particularmente útiles cuando revelan diferencias entre individuos de la misma o diferentes especies. Estos marcadores son llamados polimórficos, mientras que aquellos que no discriminan entre genotipos son conocidos como marcadores monomórficos (**Figura 1.11**). Un loci es polimórfico cuando la frecuencia del alelo más raro es igual o mayor al 0.05 (Nei, 1987), para tamaños muestrales inferiores a 50 individuos, tomando como criterio la frecuencia del alelo más común igual o inferior a 0.95. A su vez los marcadores polimórficos pueden ser codominantes o dominantes, lo que permite discriminar entre homocigotas y heterocigotas (**Figura 1.12**). Los marcadores codominantes indican diferencias en tamaño, mientras que los dominantes están presentes o ausentes. Estrictamente hablando, las diferentes formas de un marcador de ADN (por ejemplo bandas de diferentes tamaños en un gel) son llamados “alelos” de un marcador. Los marcadores codominantes pueden tener varios alelos mientras que los marcadores dominantes solo pueden presentar dos alelos.

A continuación se detallan los principales marcadores a ADN usados.

### **1.10.1. Polimorfismo en la longitud de los fragmentos de restricción (RFLP)**

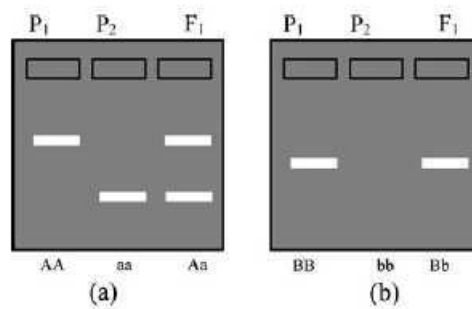
RFLP es una técnica que usa enzimas de restricción para cortar el ADN en fragmentos de diferentes tamaños. Las endonucleasas de restricción reconocen específicamente secuencias de 4, 5, 6 u 8 pares de bases (pb). Los fragmentos digeridos son transferidos a una membrana sintética o de nitrocelulosa para detectar

polimorfismos a través de una sonda específica usando análisis de Southern Blot (Southern, 1976). Son marcadores de tipo codominante, no se necesita información previa sobre la secuencia del ADN, y todos los alelos son vistos al mismo tiempo. Las desventajas de este tipo de marcador es su bajo rendimiento, es caro y laborioso y requiere grandes cantidades de ADN.



**Figura 1.11** | Diagrama representando marcadores de ADN entre los genotipos A, B, C y D. Los marcadores polimórficos se indican con flechas. Los marcadores que no discriminan entre los genotipos son llamados marcadores monomórficos. (a) Ejemplo de marcador SSR. El marcador polimórfico revela las diferencias de tamaño para los alelos del marcador para los cuatro genotipos y representa un locus simple. (b) Ejemplo de marcador generados por la técnica RAPD. Nótese que los marcadores están presentes o ausentes. En general se suelen mostrar los tamaños de estos marcadores en pares de bases (pb). Estos tamaños son estimados a partir de un marcador de peso molecular conocido (MW). Para los marcadores polimórficos, hay solamente dos alelos diferentes del marcador (Collard *et al.*, 2005).





**Figura 1.12** | Comparación entre un marcador (a) codominante y (b) dominante. A diferencia de los marcadores dominantes, los marcadores codominantes pueden claramente discriminar entre homocigotas y heterocigotas. Los genotipos son indicados en dos loci marcadores (A y B) (Collard *et al.*, 2005).

### 1.10.2. Polimorfismos de ADN amplificados al azar (RAPD)

Los marcadores RAPDs son un tipo de marcador basado en la amplificación diferencial por medio de la reacción en cadena de la polimerasa (PCR) de una muestra de ADN de secuencias cortas de oligonucleótidos (Williams *et al.*, 1990). Este marcador molecular es de tipo dominante y se basa en la amplificación de loci aleatorios del genoma. Esta técnica, utiliza como primers oligonucleótidos de 10 bases para amplificar el ADN genómico. El número de productos amplificados está directamente relacionado con el número y orientación de las secuencias complementarias al genoma. La ventaja del RAPD es que el análisis no demanda el conocimiento previo de la secuencia ADN blanco, aunque es una amplificación multilocus, es barata, y fácil de usar. La desventaja de este marcador es su reproducibilidad e interpretación.

### 1.10.3. Polimorfismo de la longitud de fragmentos amplificados (AFLP)

Este tipo de marcador molecular es generado por la combinación de una digestión con enzimas de restricción y amplificación (Vos *et al.*, 1995). La particularidad de esta metodología consiste en la adición a fragmentos de ADN

generados por digestión completa del ADN genómico de adaptadores de secuencia conocida. Primers pre-selectivos y selectivos son usados para detectar bandas y separarlas en un gel de electroforesis. El poder de la técnica es que genera un gran número de fragmentos lo que permite analizar gran parte del genoma con una simple amplificación. Esto facilita la saturación de una región del genoma (Keim *et al.*, 1997). La eficiencia de esta técnica es su rápida generación, alta reproductibilidad y polimorfismo alto. Sin embargo este marcador es costoso en cuanto a su tinción con plata, por fluorescencia o compuestos radiactivos.

#### **1.10.4. Repeticiones de secuencia simple (SSR)**

Los SSR o microsatélites o repeticiones en tandem cortas (STR) o polimorfismos de longitud de secuencia simple (SSLPs) es una repetición de secuencia nucleotídica de 2-7 pb. Las secuencias repetitivas son el resultado del desplazamiento en la replicación (Shlötterer y Tautz, 1992), y crossing-over desigual (Smith, 1976).

Consiste en la amplificación de regiones de ADN con estas secuencias repetidas y adyacentes. Para esto se utilizan primers de secuencia específica de 20-30 pb, que son complementarios a las secuencias únicas que flanquean al microsatélite (Tautz, 1989). Es decir que en este caso es indispensable tener un conocimiento previo de las secuencias del genoma de la especie a analizar. Este método tiene muchas ventajas, es rápido, seguro (Diwan y Cregan, 1997), son marcadores frecuentes, ampliamente distribuidos (Lagercrantz *et al.*, 1993), codominantes (Akkaya *et al.*, 1992), de alta heterocigocidad (Powell *et al.*, 1996), y presentan un alto nivel de polimorfismo (Akkaya *et al.*, 1995).

### 1.10.5. Polimorfismo de nucleótido simple (SNP)

Con el desarrollo de la secuenciación del ADN y de su automatización, surgieron técnicas más complejas para generar marcadores, entre los que se encuentran los SNP. Los SNP identifican diferencias entre individuos a través de variaciones en apenas un nucleótido, las que se detectan en secuencias de ADN codificante o no codificante. Existe un gran número de métodos para registrar polimorfismos en SNP. A diferencia de los marcadores mencionados anteriormente, los SNP no necesitan de la separación de fragmentos de ADN por tamaño, sino que pueden ser visualizados en placas para 96 o 384 muestras (Brookes, 1999). Aunque los SNP pueden ser detectados por secuenciación, frecuentemente los productos de amplificación se caracterizan por PCR mediante el uso de primers específicos.

Los marcadores SNP pueden ser utilizados para la construcción de mapas genéticos de alta resolución, mapeo de caracteres, diagnóstico genético, análisis de la estructura genética de poblaciones y análisis filogenético, entre otros. Las principales ventajas de los SNP son: la alta frecuencia con que ocurren en el genoma, el elevado grado de polimorfismo (identificación de variaciones de ADN hasta un nucleótido) y la posibilidad de detectar diferentes alelos para genes de interés.

La tecnología para la identificación de SNP es más sofisticada y costosa que la de otros marcadores, y su desarrollo depende de secuenciar ADN a gran escala. Una vez que se dispone de los SNPs, estos tienen innumerables aplicaciones. Con el avance de los proyectos genómicos en diversas especies será posible desarrollar marcadores de tipo SNP y esto con seguridad conducirá a que sean utilizados con más frecuencia en los programas de mejoramiento.

Además de los marcadores moleculares mencionados, basados en la amplificación del ADN mediante PCR, existen otras variantes de dichas técnicas con modificaciones menores.

## 1.11. Carácter cuantitativo y loci de carácter cuantitativo (QTL)

Como ya bien se sabe, los caracteres cualitativos presentan patrones de herencia simples, que se basan en proporciones de clases fenotípicas, evaluadas en las descendencias de los cruzamientos. Los caracteres cuantitativos presentan una herencia compleja, que en contraposición con los cualitativos, son en su mayoría, condicionados por muchos genes con efectos individuales pequeños, que además se encuentran bajo la gran influencia del ambiente.

Por su naturaleza compleja y por la importancia que presentan, tanto desde el punto de vista evolutivo en cuanto al mejoramiento genético, los caracteres cuantitativos son muy estudiados, pero muy poco comprendidos (Coehlo, 2000). Los modelos propuestos para elucidar su control genético son extremadamente simplificados, si se considera las complejas vías biosintéticas que llevan a la expresión, que dependen, adicionalmente del efecto del ambiente.

Entre los caracteres cuantitativos más estudiados se encuentran características como la producción de granos, altura de la planta, contenido de proteínas, resistencia a enfermedades, que resultan de la acción acumulativa de un conjunto de genes (poligen). Uno de los objetivos más importantes para la construcción de mapas genéticos es la localización de genes que controlan estos caracteres o sus loci. Estos son llamados “loci de carácter cuantitativo” o QTLs (Tanksley, 1993).

### 1.11.1. Pasos para el mapeo de QTLs

#### 1.11.1.1. *Selección de progenitores*

Los progenitores seleccionados que deben generar la población de mapeo, deben ser contrastantes en cuanto a la expresión fenotípica para caracteres de interés agronómico, y deben presentar polimorfismo genético en cuanto a los loci marcadores (Young, 1994).

### 1.11.1.2. *Población de mapeo*

La construcción del mapa de ligamiento requiere una población segregante (por ejemplo poblaciones derivadas de reproducciones sexuales). Los padres seleccionados para generar la población deben tener como requisito diferir en uno o más caracteres de interés. Los tamaños poblacionales usados en estudios de mapeo preliminares generalmente son de entre 50 a 250 individuos (Mohan *et al.*, 1997), sin embargo para realizar mapeos con alta resolución se requieren tamaños poblacionales mayores. Si el mapa será usado para estudios de QTLs, es importante notar que la población de mapeo debe ser evaluada fenotípicamente (por ejemplo se debe registrar los datos del carácter).

Generalmente en especies autóгамas, las poblaciones de mapeo se originan de padres que presentan un alto grado de homocigocis (endocríados). En especies alógamas, la situación es mas compleja ya que la mayoría de estas especies no toleran la endocría (sucesivas autofecundaciones). Muchas plantas alógamas son también poliploides (contienen varios juegos de pares de cromosomas). Las poblaciones de mapeo usadas para especies alógamas pueden generarse también a partir de un cruzamiento entre un padre heterocigota y un padre haploide u homocigota (Wu *et al.*, 1992).

Se pueden usar diferentes poblaciones para el mapeo de diversas especies de plantas, teniendo ventajas y desventajas (McCouch y Doerge, 1995; Paterson, 1996). Las poblaciones  $F_2$ , derivadas de híbridos  $F_1$ , y poblaciones de retrocruzas (RC), derivadas de cruzar el híbrido  $F_1$  con uno de los padres, son los tipos más simples de mapeo desarrollados para especies autóгамas. Sus principales ventajas son que estas poblaciones son fáciles de construir y requieren un tiempo corto para producirlas. La endocría de plantas individuales  $F_2$  permite la construcción de líneas recombinantes endocríadas (RILs), que consisten en una serie de líneas homocigotas, cada una conteniendo una combinación de segmentos cromosómicos únicos, que

proviene de los padres que les dieron origen. El principal inconveniente que presentan las poblaciones RILs es que para generarlas se necesitan 6 a 8 generaciones. Las poblaciones doble haploide (DH) pueden ser producidas regenerando plantas al inducir la duplicación de granos de polen, sin embargo, la producción de poblaciones DH son sólo posibles en especies que toleran el cultivo de tejidos (por ejemplo en cereales como arroz, cebada y trigo). Las principales ventajas de las poblaciones RILs y DH son que ellas producen líneas homocigotas o “verdaderas” (true-breeding lines) que pueden ser multiplicadas y reproducidas sin la ocurrencia de cambio genético. Esto permite tener material para realizar ensayos con réplicas en distintas localidades y años. Así ambas poblaciones, representan recursos “eternos” para el mapeo de QTLs (Lynch y Walsh, 1998). Además, la semilla de una RIL o DH puede ser transferida de un laboratorio a otro para realizar análisis de ligamiento y la adición de marcadores a los mapas existentes, asegurando de esta manera que todos los colaboradores examinen materiales idénticos (Paterson, 1996; Young, 1994).

### 1.11.1.3. *Construcción del mapa*

Los datos que se requieren para la construcción de un mapa incluyen marcadores morfológicos, marcadores bioquímicos y varios tipos de marcadores de ADN<sup>4</sup>. Los datos genotípicos pueden ser usados para construir un mapa en diversas etapas. En un análisis de locus simple, cada población de mapeo dará segregaciones específicas en cada locus. La información que dan estas relaciones determina si la población está expresando un desvío en la segregación de cualquier locus. En una población F<sub>2</sub>, una relación de marcadores codominantes es 1:2:1, mientras que para un marcador dominante es 3:1. En una población de RC y RIL la relación en ambos es 1:1, sea el marcador dominante o codominante. Esta relación puede ser puesta a

---

<sup>4</sup> Ver sección 1.10 Marcadores moleculares.

prueba con el test de bondad de ajuste *Chi-cuadrado*. En el análisis de dos loci, el orden y la distancia entre ambos pueden ser medidas. El análisis de tres o más loci es más preciso ya que permite identificar los dobles entrecruzamientos mejorando la estimación del porcentaje de recombinación y la distancia correspondiente de mapa. El porcentaje de recombinación es una medida directa de las distancia entre genes. La distancia de mapa puede ser calculada con una función de mapa Morgan, Haldane o Kosambi. Estos mapas se construyen con la ayuda de softwares, MAPMAKER (Lincoln *et al.*, 1993), QMOL, entre otros (Schuster y Cruz, 2004).

#### 1.11.1.4. **Análisis de QTLs**

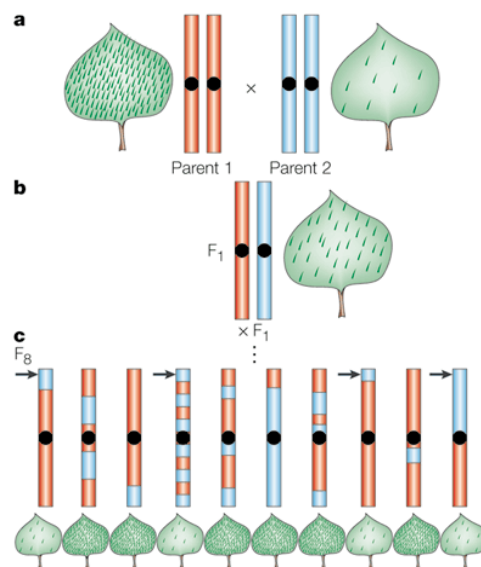
El mapeo de QTLs permite identificar el número de loci involucrados en mecanismos o sistemas de herencia compleja, así como sus localizaciones cromosómicas, modo de acción génica (aditividad, dominancia, heterosis y epistasia), además de posibilitar la descomposición de la interacción genotipos por ambiente a nivel de cada QTL (Ferreira y Grattapaglia, 1996; Austin y Lee, 1998). La capacidad de detección de un QTL es función de la magnitud de su efecto sobre la característica, del tamaño de la población segregante evaluada, de la frecuencia de recombinación entre el marcador y el QTL, así como la heredabilidad de la característica (Paterson *et al.*, 1991; Tanksley, 1993).

Dentro de los métodos propuestos para localización de QTLs se encuentran los modelos lineales (ANOVA y regresión lineal simple), mapeo por intervalo simple ("simple interval mapping" o SIM), mapeo por intervalo compuesto ("composite interval mapping" o CIM) (Zeng, 1993; 1994; Jansen, 1993) y mapeo multipunto ("multipoint mapping") (Kearsey y Hyne, 1994).

Se muestra a continuación la **Figura 1.13** y **Figura 1.14** un esquema resumido del análisis de QTLs usando RILs como ejemplo de población segregante de mapeo y donde el carácter estudiado es la densidad de tricomas en hojas (Mauricio, 2001).

### 1.11.2. Aplicación del mapeo de QTLs

Una vez detectados los marcadores ligados a la característica de interés, es posible seleccionar individuos en base al fenotipo del marcador, sin necesidad de evaluar el fenotipo de la característica, de este modo se elimina el efecto ambiental asociado al carácter cuantitativo. Ese abordaje es conocido como mejoramiento asistido por marcadores moleculares (MAS). Es importante, principalmente, en los casos que la característica de interés de especies perennes de ciclo largo es difícil de evaluar o es costosa, y para características fenotípicas de heredabilidad baja. Esto puede conducir a reducir los tiempos y costos de los programas de mejoramiento (Souza, 2001).

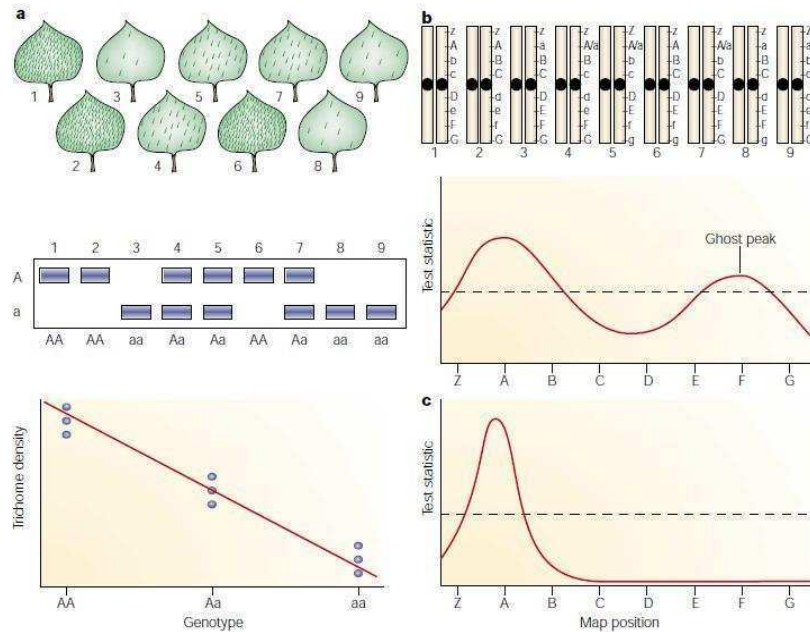


Nature Reviews | Genetics

Fuente: Mauricio, R. Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. Nature Rev.Genet. 2001; 2: 370-381.

**Figura 1.13** | Principios del mapeo de loci de caracteres cuantitativos. La estrategia básica detrás del mapeo de QTL está ilustrado para a) líneas genéticamente puras plantas "muy pilosas" y "poco pilosas" que son la generación parental. b) Híbrido F1 obtenido es autofecundado para obtener individuos F<sub>2</sub>. c) Cada individuo F<sub>2</sub> es autofecundado hasta 6-8 generaciones hasta formar varias líneas endocriadas recombinantes (RILs). Cada RIL es homocigota para una sección de de un cromosoma parental. A cada RIL se realiza el genotipado con varios marcadores moleculares así como también su fenotipado. En c) se marcan con flechas una sección del cromosoma que deriva del padre "poco piloso". Las hojas de todos los individuos que han heredado esta sección del cromosoma del padre "poco piloso" también son poco pilosos, indicando que esta región cromosómica probablemente contiene un QTL para este carácter.





Fuente: Mauricio, R. Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nature Rev. Genet.* 2001; 2: 370-381.

**Figura 1.14** | Técnicas de mapeo de QTLs en la población. Siguiendo con el mismo ejemplo de la densidad de tricomas en hoja, se ilustran los principales métodos. **a**) En la técnica de regresión, el fenotipo es correlacionado con cada genotipo marcador (el panel del medio representa la migración diferencial del ADN en el gel). En este caso un marcador A es usado para el genotipado. Los individuos que son homocigotas para el alelo "A" tienen alta densidad de tricomas, los individuos que son homocigotas para el alelo "a" tienen baja densidad de tricomas, y los heterocigotas tienen densidad intermedia de tricomas. La regresión lineal de la densidad de tricomas con el número de alelos "A" muestra una relación significativa entre el marcador y el fenotipo, lo cual indica que un QTL para la densidad de tricomas esté probablemente ligado al marcador. El método de regresión simple descrito presenta un uso limitado en localizar el segmento cromosómico que contiene un QTL. El método subestima el efecto del QTL. Cuanto más alejado esté el QTL del marcador, más débil será su efecto. **b**) El mapeo por intervalo simple usa uno o dos pares de marcadores flanqueantes a la vez (Lander y Botstein, 1989; Thoday, 1961). En este método, el QTL es localizado dentro de un intervalo cromosómico, definido por los marcadores flanqueantes. La técnica demanda usar un alto número de marcadores, como se indica en la parte superior, y luego se evalúa la probabilidad de que un intervalo entre dos marcadores esté asociado con un QTL que afecta el carácter de interés. Los resultados del análisis son graficados usando el test estadístico de máxima verosimilitud (Likelihood-ratio test statistic) versus la posición de mapa, medido en unidades de recombinación (Morgans). La línea punteada representa el punto de corte significativo, tal que por encima del cual el test de máxima verosimilitud se ajusta significativamente al modelo de los datos. La mejor estimación de la localización del QTL está dado por la localización cromosómica que corresponde con el valor significativamente más alto de máxima verosimilitud. Si bien el mapeo por intervalo ha sido un importante avance, también tiene un cierto sesgo. En particular, los QTLs fuera del intervalo bajo consideración pueden afectar la capacidad de encontrar un QTL dentro de éste (Zeng, 1993). Además, puede surgir la identificación falsa de un QTL si otro QTL está en el intervalo de interés (el falso "pico fantasma" a la derecha) (Zeng,

1993). **c)** Un tercer método conocido como mapeo por intervalo compuesto, combina la técnica de mapeo por intervalo simple con el análisis de regresión múltiple (Jansen, 1993; Zeng, 1993; Zeng, 1994). En el mapeo por intervalo compuesto se evalúa la probabilidad de que un intervalo entre dos marcadores esté asociado con un QTL que afecta el carácter de interés, y controla también los efectos de otros marcadores (como cofactores) sobre el carácter (esto provee resultados más precisos). Los resultados del análisis son graficados como en **b)**. El test estadístico es independiente del QTL en otras regiones de los cromosomas (aunque todavía puede presentar sesgo si el QTL está en el intervalo inmediatamente adyacente al intervalo de interés) (Zeng, 1994).

Zhao-Bang Zeng y sus colegas han extendido este método a una técnica de mapeo por intervalo múltiple. Éste combina análisis de mapeo de múltiples QTLs con el análisis de la arquitectura genética usando un algoritmo para buscar el número, las posiciones, los efectos y las interacciones de los QTLs significativos (Kao *et al.*, 1999; Zeng *et al.*, 1999).

## 1.12. El genoma de la soja y mapa de ligamiento

### 1.12.1. Breve reseña citogenética evolutiva de *Glycine max*

Investigadores evolucionistas han informado que *Glycine max* es una especie autotetraploide (Gurley *et al.*, 1979; Lee y Verma, 1984; Skorupska *et al.*, 1989) con una constitución cromosómica  $2n = 4x = 40$ . Sin embargo evidencias citológicas (Hymowitz, 1970; Xu *et al.*, 2000a,b) y moleculares (Doyle *et al.*, 2003a,b; Rauscher *et al.*, 2004; Zhu, *et al.*, 1994; Straub *et al.*, 2006) sugieren que esta especie es un tetraploide que se ha diploidizado generado de un organismo aloploide, por lo que se puede referirse como un organismo paleopoliploide en cuanto a su constitución genómica. Adicionalmente, estos trabajos revelaron que *G. soja* es un antecesor de *G. max* (Doyle *et al.*, 2003b, 2004).

### 1.12.2. Mapas genéticos en soja

El primer mapa de soja fue publicado en 1990 y este se construyó sobre la base de 150 marcadores RFLP (Keim *et al.*, 1990). En 1990 Williams *et al.* (1990) usó marcadores RAPDs a lo largo del genoma y en 1992, Shoemaker *et al.* (1992) sobre la base de un mapa de RFLPs, identificó polimorfismo entre distintos genotipos de soja.

Todos estos marcadores fueron herramientas útiles para analizar la estructura del genoma y para el análisis de pedigree. Sin embargo, los marcadores SSR son altamente polimórficos (Akkaya *et al.*, 1992), y con ellos Morgante (1994) identificó a un mayor número de alelos por locus que los RFLPs. En 1995, Shoemaker y Specht (1995) usaron 7 marcadores para pigmentación, 6 morfológicos, 7 isoenzimas, 8 RAPDs y 110 RFLPs para mapear 60 familias  $F_{2:3}$  derivadas del cruzamiento Clark x Harosoy. Además, estas mismas familias fueron caracterizadas con 40 marcadores SSR (Akkaya *et al.*, 1995). Mansur *et al.* (1996) mapearon caracteres agronómicos usando líneas recombinantes endocriadas (RILs) y encontraron que las RILs constituyen una población permanente que puede ser evaluada en el espacio y en el tiempo, reduciendo el error experimental y el error en el ligamiento, en comparación con una población  $F_2$  y de retrocruza. Keim *et al.* (1997) realizaron un mapa de la soja con 840 marcadores: 165 RFLPs, 25 RAPDs y 650 AFLPs. Éstos se distribuyeron a lo largo de 28 grupos de ligamiento que cubrieron una distancia de 3441 cM. Cregan *et al.* (1999) utilizaron 606 SSRs y con ellos construyeron el mapa genético partiendo de tres poblaciones:  $F_2$  proveniente del cruzamiento interespecífico PI 468.916 (*G. soja*) x A81-356022 (*G. max*), en otra se usó también una población  $F_2$  del cruzamiento Clark x Harosoy y también una población de RILs (Minsoy x Noir). Los mapas de esas tres poblaciones fueron alineados y se construyó un mapa consenso con 20 grupos de ligamiento (GL), correspondiente a los 20 cromosomas homólogos de soja. Posteriormente, Song *et al.* (2004) incorporaron 420 marcadores nuevos SSR y obtuvieron un mapa integrado a partir de 5 mapas resultando en un total de 1849 marcadores: 1015 SSRs, 709 RFLPs, 73 RAPDs, 24 caracteres morfológicos, 6 AFLPs y 10 isoenzimas. Este mapa genético de soja presentó una cobertura de 2523.6 cM en unidades de distancia Kosambi que representó 20 GL asociados a los 20 cromosomas. A este último mapa se le adicionaron 1141 SNPs (Choi *et al.*, 2007), de esta manera se dispuso de un mapa genético más preciso, que ha sido utilizado como

base para otros estudios de mapeo de QTL, mapeo físico y para asistir en la secuenciación del genoma de la soja.

### 1.12.3. Genómica

El avance de nuevas tecnologías aplicadas al estudio del ADN ha permitido un rápido y eficiente desarrollo de los recursos genómicos en una amplia variedad de leguminosas de importancia agronómica. En este sentido, la técnica de “shotgun”, ha contribuido a obtener la secuenciación completa del genoma de la soja, que tiene un tamaño de 1.1 Gb (Schmutz *et al.*, 2010). Los análisis bioinformáticos del genoma predijeron que la soja contiene 46.430 genes de los cuales 1.127 corresponden al metabolismo de los lípidos. En cuanto a genes de nodulación, se identificaron un total de 52 genes relacionados, de los cuales 28 corresponden a nodulinas y 24 son genes regulatorios claves, que probablemente representen verdaderos ortólogos de genes de nodulación conocidos en otras especies de leguminosas (Schmutz *et al.*, 2010). La integración del mapa físico (secuencia completa del genoma) con el mapa genético o de ligamiento saturado que hoy existe en soja (<http://www.Soybase.org>) ha permitido asociar los fenotipos ya mapeados con la secuencia de ADN. Esto a su vez ha ayudado a avances relevantes como por ejemplo el clonado de un gen de resistencia a roya asiática en soja (Meyer *et al.*, 2009).

### 1.13. Objetivos

En base a lo expuesto anteriormente se propone en la presente tesis como:

#### **Objetivo general**

Identificar QTLs asociados a la capacidad de nodulación de la soja, de manera de disponer de la información y de los marcadores moleculares que permitan mejorar al hospedante en lo que hace a la nodulación y por ende a la fijación de nitrógeno.

### **Objetivos particulares**

1. Identificar entre los cultivares comerciales de soja argentina, materiales que difieran en su capacidad de nodulación (Cap. 2).
2. Estimar la variabilidad fenotípica y los mecanismos de la herencia de los caracteres asociados a la capacidad de nodular y fijar nitrógeno (Cap. 2).
3. Identificar la diversidad genética de los cultivares argentinos por medio de marcadores microsatélites ligados a QTLs de nodulación en cultivares brasileros (Cap. 3).
4. Construir un mapa genético que permita establecer las distancias entre los marcadores moleculares y los QTLs o genes de interés y determinar su modo de herencia (Cap 4).

## **CAPÍTULO 2**

# **CAPACIDAD DE NODULACIÓN DE CULTIVARES COMERCIALES ARGENTINOS DE SOJA (*GLYCINE MAX* L. MERR)**



## 2. CAPACIDAD DE NODULACIÓN DE CULTIVARES ARGENTINOS DE SOJA (*GLYCINE MAX L MERR*)

### 2.1. Introducción

El nitrógeno es uno de los nutrientes que más frecuentemente limita el crecimiento de las plantas, por lo que la baja disponibilidad del mismo afecta el rendimiento. Usualmente en las producciones agrícolas el N se incorpora por medio de la fertilización química y/o biológica. Esta última tiene alguna ventaja por sobre la fertilización química ya que presenta menores riesgos para la contaminación del medio ambiente y contribuye a un manejo sustentable.

Las leguminosas pertenecen a la tercera familia en importancia de las plantas con flor (*Fabaceae*), que incluye aproximadamente 20.000 especies (Polhill y Raven, 1981). La subfamilias de las *Papilionoideas* contiene la mayor cantidad de plantas que además tienen la capacidad de establecer una relación simbiótica con bacterias del suelo Gram (-) conocidas como rizobios. Como resultado de la asociación, en la raíz se desarrollan estructuras llamadas nódulos los cuales proveen el ambiente en donde se desarrollan las bacterias. La planta proporciona a la bacteria carbohidratos que son el sustrato respiratorio que genera la energía demandada por la bacteria para fijar el nitrógeno gaseoso a amonio, por medio de la enzima, nitrogenasa que es sensible a las altas tensiones de oxígeno.

La soja (*Glycine max* L. Merr) es la leguminosa cultivada más importante a nivel mundial. Hasta no hace tanto tiempo se consideraba que nodulaba solo con *Bradyrhizobium japonicum* (Jordan, 1982), una bacteria de crecimiento lento. Sin embargo, Trinick (1980) en Papua Nueva Guinea, Keyser *et al.* (1982) y Dowdle y Bohloul (1985) en suelos de China describieron la existencia de estirpes de rizobios de



crecimiento rápido que también nodulan la soja. A esto le siguieron investigaciones realizadas por Cleyet Marel (1987) en Vietnam en donde también aisló bacterias de crecimiento rápido que nodulan soja. Más tarde investigadores chinos y Europeos aislaron rizobios e hicieron un detallado estudio de los suelos y los rizobios contenidos por estos (Camacho *et al.*, 2002; Thomas Oates *et al.*, 2003; Chen *et al.*, 2004; Li *et al.*, 2011; Zhang *et al.*, 2011;). Yang (2006) y Barcellos *et al.* (2007) encontraron que en las regiones tropicales y subtropicales de China con climas húmedos y suelos ácidos, la soja establece simbiosis con *B. japonicum* (Jordan, 1982), *B. elkanii* (Kuykendall *et al.*, 1992) y difícilmente con *B. liaoningense* (Xu *et al.*, 1995), *B. yuanmingense* (Yao *et al.*, 2002) y *B. canariense* (Vinuesa *et al.*, 2005), excepto en Hubei sitio de donde se aislaron diversas cepas de *Ensifer fredii* (Camacho *et al.*, 2002). En Heilongjiang, una provincia al Noreste de la China que tiene clima húmedo y suelos neutros o ligeramente ácidos, la soja es nodulada primordialmente por *B. japonicum*, muy poco por *B. elkanii* y nunca por *E. fredii* (Wang *et al.*, 2009). Tanto la provincia de Hubei como de Heilongjiang fueron posiblemente los centros de origen de la soja. En la provincia de Xinjiang, que tiene clima seco y suelos alcalinos, las especies de rizobios que nodulan soja son *Ensifer fredii* y *B. liaoningense*. Es importante destacar que en los sitios descritos se cultivan distintos cultivares de soja. Todos estos resultados sugieren, entre otras cosas, que la diversidad de los rizobios que interactúan con la soja no es completamente clara, más aún algunos de estos rizobios nodulan preferencialmente con ciertos materiales genéticos, por lo tanto la diversidad de rizobios se que se identifican usando plantas trampa es una función del genotipo de los cultivares utilizados. Por otro lado, también se concluyó que el pH de los suelos y la disponibilidad de un conjunto de nutrientes claves definen la distribución geográfica de los rizobios (López *et al.*, 2013).

Los materiales de soja responden de diversas maneras a los rizobios. Mientras que algunos cultivares son incompatibles con algunos rizobios, otros excluyen o restringen el proceso de nodulación a ciertos serogrupos de *Bradyrhizobium japonicum*

(Van *et al.*, 2007). Alternativamente hay genotipos, como es el caso de los cultivares “Tropical Glycine cross” (TGx), que nodulan con varias poblaciones de *Bradyrhizobium* sp. presentes en los suelos indígenas africanos (Abaidoo *et al.*, 2000; Osunde *et al.*, 2003). Estos cultivares fueron evaluados en diversos países de África (Javaheri, 1996; Mpeperekki *et al.*, 2000) y se demostró que los cultivares desarrollaron nódulos eficientes lo que sugiere que los cultivares son promiscuos (Osunde *et al.*, 2003).

El análisis de la respuesta de los diversos cultivares de soja condujo a identificar genes que controlan la nodulación en soja. Se demostró la formación de nódulos en la soja esta controlado por genes ubicados en varios loci, llamados *rj(s)* o *Rj(s)* (Hayashi *et al.*, 2012). Los alelos recesivos, *rj1* y *rj5* fueron identificados como receptores putativos, *nfr1* y *nfr5*, respectivamente (Indrasumunar *et al.*, 2010; 2011; Hayashi *et al.*, 2012), y *rj7* fue identificado como *nts1*, el cual está involucrado en la regulación sistémica de la nodulación (Nishimura *et al.*, 2002, Searle *et al.*, 2003). En contraste, los alelos dominantes, *Rj2*, *Rj3*, *Rj4* y *Rfg1* han tenido características únicas, ya que restringen la nodulación con cepas específicas (o serogrupos) de *Bradyrhizobium* o *Ensifer/Sinorhizobium*: ejemplos incluyen *Rj2* con *B. japonicum* USDA122, *Rj3* con *B. elkanii* USDA33, *Rj4* con *B. elkanii* USDA61 y *Rfg1* con *E./S. fredii* USDA257 (Caldwell, 1966; Caldwell *et al.*, 1966; Trese, 1995; Vest, 1970; Vest y Caldwell, 1972; Weiser *et al.*, 1990).

Debido al gran tamaño del genoma de la soja, los caracteres relacionados con la capacidad de nodulación y fijación de nitrógeno son complejos. Es de esperar que varios genes estén involucrados en el proceso. Tanya *et al.* (2005), Nicolás *et al.* (2006), Santos *et al.* (2006, 2013), identificaron a partir de cruzamientos con cultivares de soja con capacidades de nodulación contrastantes, varios loci de caracteres cuantitativos (QTLs) que están involucrados en la fijación biológica de nitrógeno (FBN). Esta información podría ser aplicada para mejorar la capacidad simbiótica de los cultivares de soja.

También hay factores bacterianos que contribuyen a la nodulación en soja y que pueden ser influidos por la densidad de inóculo, mecanismo conocido como el “quorum sensing”. Esta respuesta es probable que sea atenuada en el suelo (Jitacksorn y Sadowsky, 2008). El “quorum sensing” está relacionado con la síntesis y sentido de moléculas, conocidas como autoinductores, éstas son producidas en condiciones de una alta densidad poblacional y alteran la expresión de genes bacterianos (Waters y Bassler, 2005). Takats (1986) encontró que la nodulación en plantas era inhibida cuando la inoculación se realizaba con una alta densidad inóculo. El compuesto que se identificó se denomina bradyoxetina, que es un autoinductor producido por *Bradyrhizobium japonicum* cuando este se desarrolla en un medio y se alcanza una alta densidad poblacional. Este autoinductor presenta una estructura similar al sideróforo ácido muginéico (Loh *et al.*, 2002a) y su producción está regulada por las concentraciones de hierro en el medio, siendo máxima en condiciones de baja concentración de este metal (Loh *et al.*, 2002a; Loh y Stacey, 2003). Cuando aumenta la densidad poblacional, los genes *nod* son reprimidos por un mecanismo de señalización en cascada iniciado por el regulador NswB, en respuesta a altas concentraciones de bradyoxetina, NswB activa la proteína NolA, la cual a su vez activa a NodD2, esta última reprime los genes *nod* (Loh y Stacey, 2001, Loh *et al.*, 2002b). Esta represión ha sido demostrada *in vitro* e *in planta* (Loh *et al.* 2002b). Pero además Pongsilp *et al.* (2005) encontraron que varias cepas de *Bradyrhizobium* también producen moléculas tipo N-acyl homoserina lactona. Vale la pena destacar que no se ha descrito un rol de las mismas en el proceso de la nodulación. Estos hallazgos sugieren que la expresión de los genes *nod* en *B. japonicum* está regulado por autoinductores de “quorum sensing” que es probable que alteren además, la nodulación de la soja.

Los programas de mejoramiento en general han priorizado la identificación y selección de materiales en base a caracteres relacionados con el rendimiento y la resistencia a enfermedades. Sin embargo se pueden citar unos pocos planes de

mejora que tuvieron como objetivo a la FBN. En 1980 en Australia se seleccionaron cultivares en base a su capacidad de fijación de nitrógeno, aunque el plan no fue exitoso debido principalmente a problemas como el “screening fenotípico” y con la imposibilidad de asociar la fijación de nitrógeno con caracteres relacionados principalmente con rendimiento y resistencia a enfermedades (Herridge *et al.*, 2001). Sólo en África se han liberado cultivares que *ex profeso* se seleccionaron por su promiscuidad. Sin embargo, esta tecnología ha funcionado solamente en dicho continente porque no existía disponibilidad cepas selectas ni tecnologías para su producción y además porque los suelos contienen naturalmente rizobios simbioses de soja (Mperekki *et al.*, 2000; Herridge *et al.*, 2001).

La capacidad de fijación de nitrógeno de la soja depende principalmente de la biomasa de nódulos y de la actividad específica de la nitrogenasa (Döbereiner, 1966; Bohrer y Hungria, 1998; Hungria y Bohrer, 2000). En Brasil, se han realizado estudios pioneros relacionados con la evaluación del potencial de nodulación y otros parámetros de la FBN en cultivares de soja, incluyendo materiales de origen Brasileño derivados de cultivares norteamericanos (Döbereiner y Arruda, 1967; Brose *et al.*, 1979; Vargas *et al.*, 1982). Aún así, son pocos los trabajos realizados en las dos últimas décadas en Argentina y en Brasil, que tuvieron como objetivo estudiar la variación entre genotipos de soja en cuanto a eficiencia del proceso de nodular y fijar nitrógeno. La falta de información de las características relacionadas a la FBN y sus factores genéticos condicionantes, ha contribuido a aumentar en un 30% las pérdidas en rendimiento por deficiencias en la FBN. Esto fue observado en cultivares brasileños y en sus parentales norteamericanos que les dieron origen (Bohrer y Hungria, 1998; Hungria y Bohrer, 2000).

Tanto en la Argentina, como en Brasil, así como en otras regiones agrícolas del mundo en donde se cultiva soja como una especie exótica, y donde los suelos carecen de los rizobios que interactúan con ella. Por ello la soja debe ser inoculada. Si bien hay un gran desarrollo en la industria de inoculantes en Argentina, tanto en las compañías

privadas como públicas, los investigadores han estado buscando nuevas cepas con una alta eficiencia de fijación de nitrógeno. Sin embargo pocos son los trabajos que se han realizado seleccionando plantas en base a su capacidad para nodular y fijar nitrógeno.

### **Hipótesis:**

- Existe variabilidad fenotípica para los caracteres de capacidad de nodulación entre cultivares argentinos de soja.

### **Objetivos**

- Evaluar la capacidad de nodulación de 31 cultivares comerciales argentinos de soja e identificar aquellos con capacidad contrastante en cuanto al carácter, los que podrían ser potencialmente seleccionados como material de partida en un análisis de QTLs.

## **2.2. Materiales y métodos**

### **2.2.1. Mantenimiento de la bacteria y producción del inóculo**

Las cepas de *B. japonicum* que se utilizaron alternativamente para inocular las plantas fueron SEMIA 5079, SEMIA 5080 y E109 (INTA Argentina). Los cultivos “stock” fueron mantenidos a -70°C en glicerol al 10%. Los cultivos se iniciaron a partir de cultivos crecidos en tubo de pico de flauta que fueron mantenidos en heladera a 4°C. Las bacterias se cultivaron en medio manitol-agar- extracto de levadura (YEM) (Vincent, 1970) hasta fase log en un agitador a 150 rev min<sup>-1</sup> a 28°C. La densidad celular se ajustó midiendo la D.O. a 625 nm de una dilución seriada que se comparó con el recuento en placa. La densidad celular fue ajustada con buffer PBS (7.2 g NaCl, 2.79 g, 0.43 g Na<sub>2</sub>PO<sub>4</sub>H por litro, pH=7.2).

### 2.2.2. Material vegetal y ensayos de nodulación

Las semillas de los cultivares comerciales de soja pertenecientes a grupos de madurez II a VIII fueron provistas por la compañía Nidera Semillas S.A. (**Tabla 2.1**). Éstas se esterilizaron superficialmente sumergiéndolas 5 minutos en 20.83 M de etanol y 0.4 M de NaHPO<sub>4</sub>, luego de lo cual se lavaron con agua corriente 10 veces. Luego, se colocaron en placas de petri conteniendo 15 ml de agar-agua 1% m/v y se incubaron por 48 hs a 30°C en oscuridad. En este período de tiempo la radícula alcanzó un tamaño de aproximadamente 1 cm. En el momento en que las plántulas se trasplantaron a las jarras de Leonard, se inocularon con 1 ml de una suspensión de bacterias conteniendo 10<sup>8</sup> UFC ml<sup>-1</sup> de una mezcla de cepas de *B. japonicum* SEMIA 5079 (CPAC 15) y SEMIA 5080 (CPAC 7), o *B. japonicum* E109 (cepa comercial que se sabe deriva de la estirpe USDA 138). Se usó vermiculita como sustrato de las jarras, éstas fueron regadas con solución nutritiva de Jensen libre de nitrógeno (Vincent, 1970). Al cabo de 3 días se realizó un raleo para dejar una planta por jarra, las que se mantuvieron a 25°C con un fotoperíodo de 14 hs de luz. Periódicamente, se adicionó a las jarras agua destilada para mantener la humedad. Así se evaluaron 30 cultivares de soja en los que se analizó la capacidad de nodulación en plantas inoculadas con una mezcla de cepas de *B. japonicum* SEMIA 5079 (CPAC 15) y SEMIA 5080 (CPAC 7). En base a los resultados de este experimento se seleccionaron 14 cultivares para un segundo experimento, la mitad de ellos con alta nodulación y la otra mitad con baja nodulación.

**Tabla 2.1** | Cultivares de soja incluidos en los ensayos (Nidera Semillas). Grupos de maduración (100 a 150 días de maduración), medio (116 a 150 días), tardíos (mas de 150 días de maduración). Hábito de crecimiento: I: crecimiento indeterminado, D: crecimiento determinado, S: crecimiento semideterminado. Los días de floración indica el número de días desde que germina hasta R8. RG: poseen gen de resistencia a glifosato.

Cultivar	Hábito de crecimiento	Días a floración	Días a maduración
NA 2018 RG <sup>1,2</sup>	I	28	112
A 3302 RG <sup>1,2</sup>	I	32	120
A 3550 RG <sup>1,2</sup>	I	39	126
NA 3628 RG <sup>1</sup>	I	37	128
NA 3731 RG <sup>1</sup>	I	37	128
NA 3933 RG <sup>1</sup>	I	39	129
NA 4209 RG <sup>1,2,3</sup>	I	41	132
NA 4241 RG <sup>1</sup>	I	41	132
A 4303 RG <sup>1</sup>	I	43	135
NA 4553 RG <sup>1</sup>	I	43	136
NA 4613 RG <sup>1</sup>	I	43	140
A 4725 RG <sup>1,2</sup>	I	43	140
A 4910 RG <sup>1</sup>	I	47	146
NA 4990 RG <sup>1</sup>	I	47	147
NA 5009 RG <sup>1,2</sup>	I	48	151
NA 5485 RG <sup>1,2,3</sup>	D	62	152
NA 5579 RG <sup>1</sup>	I	64	158
NA 5920 RG <sup>1,2</sup>	I	68	162
NA 6010 RG <sup>1</sup>	D	68	164
NA 6126 RG <sup>1,2</sup>	D	69	166
NA 6355 RG <sup>1</sup>	D	70	167
A 6411 RG <sup>2,3</sup>	D	71	167
NA 6509 RG <sup>1</sup>	I	72	168
NA 7000 RG <sup>1</sup>	D	72	169
A 7053 RG <sup>1,2,3</sup>	D	75	171
NA 7708 RG <sup>1,2</sup>	D	78	177
A 8000 RG <sup>1,2</sup>	D	82	180
NA 8010 RG <sup>1,2</sup>	D	83	181
NA 8164 RG <sup>1</sup>	D	83	183
NA 8413 RG <sup>1</sup>	D	86	187
NA 8499 RG <sup>1,2</sup>	S	86	187

Cultivares ensayados en el experimento 1, <sup>2</sup> Cultivares ensayados en el experimento 2, <sup>3</sup> Cultivares ensayados en el experimento 3. RG: resistente a glifosato.

Las plantas se cultivaron en las condiciones descritas por un período de 35 días. Al final del cual se cosecharon las plantas, separando el tallo, raíz y los nódulos. Se determinó el número de nódulos (NN), el peso seco de la parte aérea (PSA) y de nódulos (PSN); estos dos últimos fueron evaluados después de secarlos en estufa a 65°C hasta peso seco constante (aproximadamente 48 h).

A continuación se detallan la metodología y diseños de cada ensayo. En los experimentos con bajo número de tratamientos (cultivares) se aumentó el número de réplicas. Esta es la razón por la cuál se explican los diferentes números de repeticiones usadas en los experimentos.

### **2.2.3. Nodulación de cultivares de soja**

#### **2.2.3.1. Primer experimento**

Los cultivares de soja se inocularon con una mezcla de cepas SEMIA 5079 y SEMIA 5080 las que fueron adicionadas en una concentración de  $10^8$  UFC ml<sup>-1</sup>. Las plantas se cultivaron bajo condiciones de luz y temperatura controladas por un período de 5 semanas. Se evaluó la capacidad de nodulación de 30 cultivares comerciales de soja de ciclo de maduración temprano, medio y tardío (**Tabla 2.1**). Para mostrar que no existió contaminación bacteriana de las plantas, se incluyeron como controles, plantas inoculadas con agua destilada estéril, pertenecientes a cada grupo de maduración. El experimento se realizó en un diseño completamente al azar con 4 réplicas por tratamiento.

#### **2.2.3.2. Segundo experimento**

A partir de los resultados del primer experimento descrito se seleccionaron en base al NN y al PSN, 14 cultivares que representaron a tres grupos alta, media y baja capacidad de nodulación (**Tabla 2.1**). El procedimiento fue idéntico al experimento 1. A



los cultivares seleccionados se le adicionó un nuevo cultivar, A 6411 RG (ya que en el primer ensayo se perdieron estas plantas y este tratamiento se excluyó del experimento 1). Se realizó un diseño completamente aleatorizado con 12 réplicas por tratamiento.

### **2.2.3.3. Tercer experimento**

A partir de los resultados de los experimentos previos, se seleccionaron cuatro cultivares: dos de los cuales se caracterizaron por su baja capacidad de nodulación A 6411 RG y A 7053 RG y otros dos por su alta capacidad de nodulación (NA 5485 RG y NA 4209 RG). Éstos se inocularon con la cepa *B. japonicum* E109. El diseño fue completamente al azar con 10 réplicas por cultivar. Se inocularon plantas control pertenecientes a los cultivares NA 4209 RG y A 7053 RG con agua destilada estéril.

### **2.2.4. Respuesta de la nodulación frente a la concentración bacteriana**

El objetivo de este ensayo fue evaluar el efecto que la concentración de bacterias del inoculante sobre la capacidad de nodulación de 2 cultivares con capacidad de nodulación extremas: NA5485 RG de alta capacidad y NA 6411 RG de baja. Para ello, se cultivó *B. japonicum* E109 en medio YEM hasta fase log y a partir de este caldo se realizó en medio YEM estéril una dilución en serie de  $10^2$ ,  $10^4$ ,  $10^6$  y  $10^8$  UFC ml<sup>-1</sup>. Se tomó 1 ml de cultivo con el que se inocularon las semillas de *Glycine max* (L.) Merr pregerminadas, que previamente fueron esterilizadas superficialmente. Los ensayos se hicieron en jarras de Leonard conteniendo vermiculita. Después de la inoculación, las semillas del cultivar NA 5485 RG y NA 6411 RG fueron cubiertas con 1 cm de vermiculita, luego de la germinación se raleó a 2 plantas por jarra. Las plantas fueron regadas con una solución libre de N (Vincent, 1970) o agua destilada en

cámara de cultivo con un fotoperíodo de 14 h luz a 25°C. A los 35 días post-inoculación se determinó el NN, PSN y PSA. Los nódulos y las partes aéreas fueron secados en estufa a 65°C hasta peso constante. El diseño experimental fue completamente aleatorizado con 10 réplicas por tratamiento.

### 2.2.5. Análisis estadístico

Se realizó un ANOVA para probar las diferencias entre cultivares y se hicieron comparaciones múltiples con el test HSD-Tukey ( $\alpha = 5\%$ ). Las diferencias con los controles se realizaron con el test de Dunnet ( $\alpha = 5\%$ ).

A partir de los datos obtenidos de los 14 cultivares que fueron incluidos en los dos primeros experimentos se realizó un análisis factorial combinado de la varianza. Ésto nos permitió determinar los efectos del cultivar y ambiental (ensayos) y el efecto de la interacción cultivar x ambiente en la capacidad de la nodulación (McIntosh, 1983). Las esperanzas de los cuadrados medios fueron estimados según:

$$\text{Ambiente (E): } \hat{\sigma}_e^2 + g\hat{\sigma}_{R \times E}^2 + r\hat{\sigma}_{GE}^2 + rg\hat{\sigma}_E^2 \quad (1)$$

$$\text{Bloques en ambientes (Rep/E): } \hat{\sigma}_e^2 + g\hat{\sigma}_{R \times E}^2 \quad (2)$$

$$\text{Cultivares (G): } \hat{\sigma}_e^2 + r\hat{\sigma}_{RE}^2 + re\hat{\sigma}_G^2 \quad (3)$$

$$\text{Cultivar x Ambiente (GxE): } \hat{\sigma}_e^2 + r\hat{\sigma}_{GE}^2 \quad (4)$$

$$\text{Error: } \hat{\sigma}_e^2 \quad (5)$$

Donde  $\hat{\sigma}_G^2$  : representa la varianza entre cultivares;  $\hat{\sigma}_E^2$ : varianza entre ambientes (ensayos);  $\hat{\sigma}_{GE}^2$  : varianza debido a interacciones cultivares x ambiente;  $\hat{\sigma}_e^2$ : varianza debido al error;  $g$  = número de cultivares;  $e$  = número de ambientes y  $r$  = el número de bloques o repetición por tratamiento.

Las varianzas calculadas (Ec. (1) a Ec. (5)) permitieron estimar la correlación intraclase, que es una medida de la heredabilidad en sentido amplio ( $He$ ) de los caracteres (Hanson *et al.*, 1956) como:

$$He = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{GE}^2}{e} + \frac{\sigma_e^2}{er}} \quad (6)$$

Para medir la importancia relativa de cada factor estudiado (“Cultivar”, “Ambiente” y la interacción “Cultivar x Ambiente”), se determinó el porcentaje de variación de cada efecto como el cociente entre las sumas de cuadrado de la fuente correspondiente a la variación sobre la suma de cuadrados de los tratamientos (Sokal y Rohlf, 1995).

También se analizó el grado de asociación calculando el coeficiente de correlación de Pearson (Sokal y Rohlf, 1995) con las variables NN, PSN y PSA.

## 2.3. Resultados

Uno de los problemas a la hora de evaluar un número considerable de cultivares es el número de repeticiones. Por ello primero se evaluó la capacidad de nodulación de 30 cultivares y en este caso el número réplicas fue reducido. Con el fin de aumentar el número de réplicas, el número de cultivares se redujo, seleccionando materiales dentro de los grupos de alta, media y baja capacidad de nodular. En el tercer experimento, se redujo aún más, el número de cultivares, sólo se evaluaron 2 de alta y 2 de baja capacidad de nodulación y se aumentó el número de réplicas.

### 2.3.1. Primer Experimento

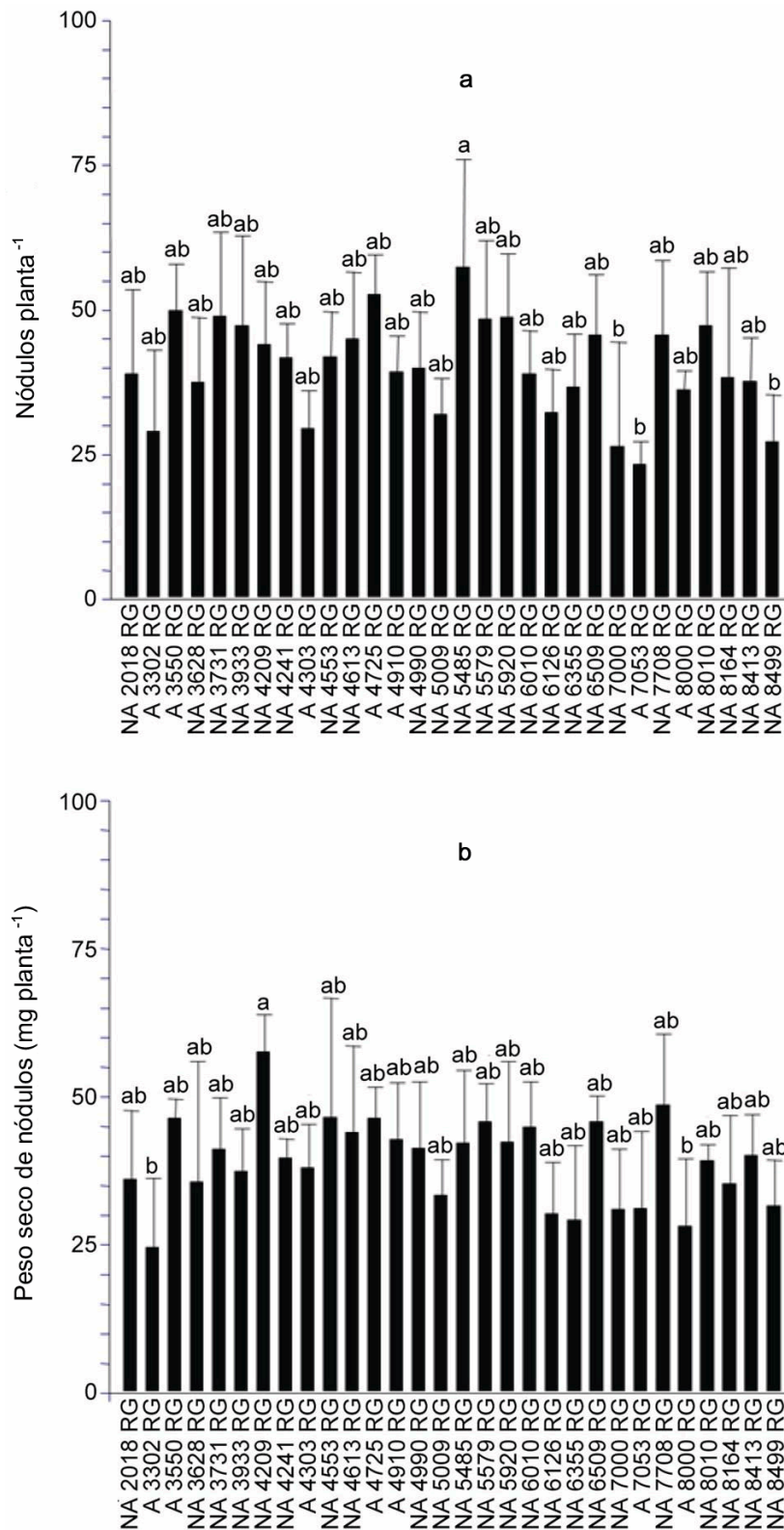
Los cultivares de soja respondieron diferencialmente a la inoculación con *Bradyrhizobium japonicum*. Se encontraron diferencias significativas para el NN ( $F = 2.33, P < 0.01$ ) y PSN ( $F = 1.96, P < 0.05$ ), sin embargo no existieron diferencias entre

cultivares para el PSA ( $F = 1.84$ ,  $P > 0.05$ ). Los coeficientes de variación (CV%) fueron: 26.5 para NN, 27.3 para el PSN y 28.9 para el PSA.

Para el NN y el PSN, el test de Tukey agrupó a los cultivares en 2 grupos, uno con cultivares con alta capacidad de nodulación y otro con cultivares de baja capacidad de nodulación. El cultivar NA 5485 RG fue el que presentó el mayor NN (**Figura 2.1a**) y los cultivares NA 7000 RG, A 7053 RG y NA 8499 RG fueron los que presentaron el menor NN. Mientras que NA 4209 RG presentó la mayor biomasa de nódulos (PSN) y los cultivares A 3302 RG y NA 8000 RG mostraron los valores más bajos (**Figura 2.1b**). El resto de los cultivares desarrollaron un número y peso seco de nódulos entre estos dos extremos. El cultivar A 7053 RG que presentó el NN más bajo desarrolló 23 nódulos planta<sup>-1</sup> que es aproximadamente un 60% más bajo que el NN del cultivar de alta nodulación NA 5485 RG (57.5 nódulos planta<sup>-1</sup>). En base a estos resultados los cultivares de soja se agruparon en tres grupos, de alta, media y baja capacidad de nodulación (**Figura 2.1a**).

La masa de nódulos más baja observada fue 24.5 mg planta<sup>-1</sup> para el cultivar A 3302 RG y la más alta fue 57.6 mg planta<sup>-1</sup> que correspondió al cultivar NA 4209 RG (**Figura 2.1b**).

El análisis de correlación se efectuó sobre el primer experimento por ser éste el ensayo donde participó el mayor número de cultivares. Tal cual era de esperar, se encontró una correlación altamente significativa entre el NN y el PSN ( $r = 0.69$ ,  $P < 0.01$ ). Además el PSA estuvo altamente correlacionado con ambas variables, NN y PSN ( $r = 0.53$ ,  $P < 0.01$  y  $r = 0.67$ ,  $P < 0.01$ ) respectivamente.



**Figura 2.1** | Respuesta a la nodulación de 30 cultivares de soja. Las inoculaciones fueron realizadas con una mezcla de cepas *Bradyrhizobium japonicum* cepas SEMIA 5079 y SEMIA 5080 (valores medios  $\pm$  desvío estándar), la concentración de los inóculos fue de  $10^8$  rizobios  $ml^{-1}$ . **a.** Número de nódulos, **b.** Peso seco de nódulos. Mismas letras indican diferencias no significativas (Tukey 5%).

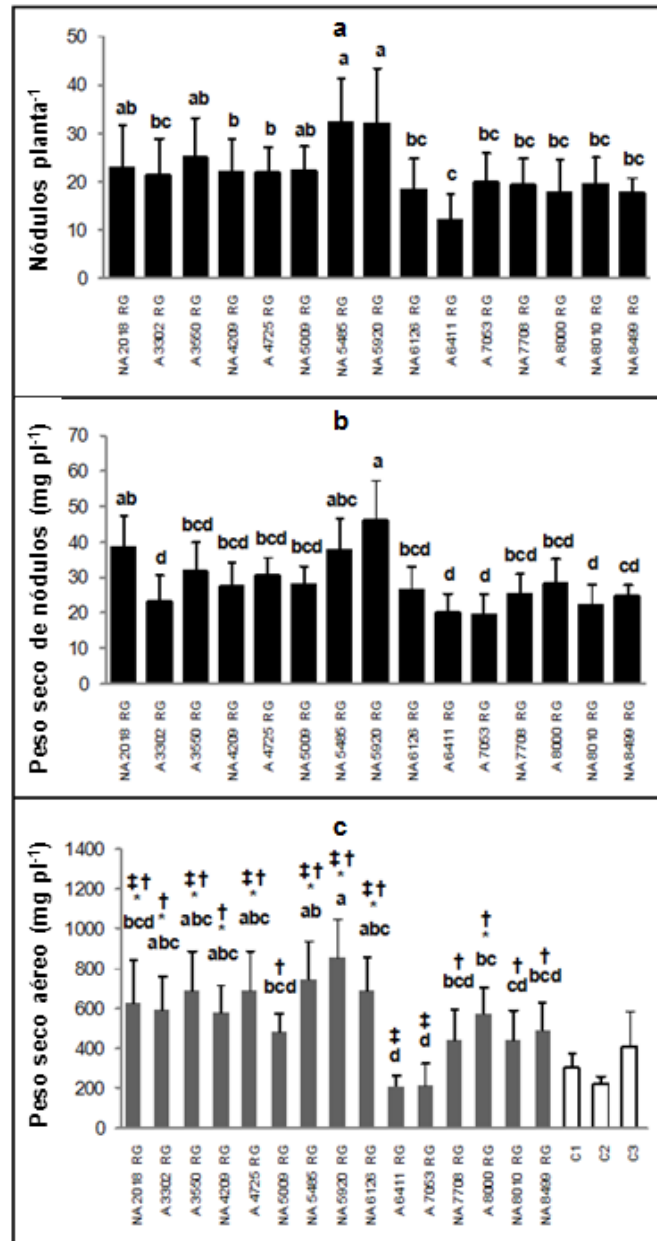
### 2.3.2. Segundo Experimento

En el segundo experimento se seleccionaron 14 cultivares (**Tabla 2.1**), entre ellos, genotipos representativos de los grupos de alta, media y baja capacidad de nodulación, y se incluyó un nuevo cultivar introducido en este ensayo, A 6411 RG, que no sobrevivió en el primer ensayo.

Los cultivares se diferenciaron significativamente para el NN ( $F = 8.00$ ,  $P < 0.01$ ), y PSN ( $F = 6.65$ ,  $P < 0.01$ ) y PSA ( $F = 13.00$ ,  $P < 0.01$ ). Los CV% fueron: 31.3 para el NN, 33.9 para el PSN y 24.9 para el PSA.

Los cultivares NA 5920 RG y NA 5485 RG, fueron las de mayor capacidad de nodulación ya que presentaron los valores más altos en las 3 variables (**Figura 2.2a,b,c**), el cultivar NA 2018 RG sólo presentó una biomasa de nódulos significativamente más alta. Por el contrario, los cultivares A 6411 RG, A 7053 RG, NA 8010 RG y A 3302 RG desarrollaron la biomasa más baja de nódulos (**Figura 2.2b**). A 6411 RG y NA 7053 RG no difirieron significativamente de los controles no inoculados en cuanto a PSA (Dunnet,  $\alpha = 5\%$ ) (**Figura 2.2c**). Estos dos mostraron una menor capacidad de nodulación, ya que presentaron un menor número de nódulos lo que se reflejó en una menor biomasa aérea.

En el segundo experimento, en el que se incluyeron solo 15 genotipos en base a su respuesta diferencial, éstos se agruparon en un gran número de grupos. Los rangos de valores observados para el NN y el PSN fueron 34.2 nódulos planta<sup>-1</sup> y 33.1 mg planta<sup>-1</sup> versus 15.1 nódulos planta<sup>-1</sup> y 27.3 mg planta<sup>-1</sup>, para el primer y segundo experimento, respectivamente. Para el PSA los rangos fueron de 594.1 mg planta<sup>-1</sup> y 688.9 mg planta<sup>-1</sup> para el primer y segundo experimento respectivamente.

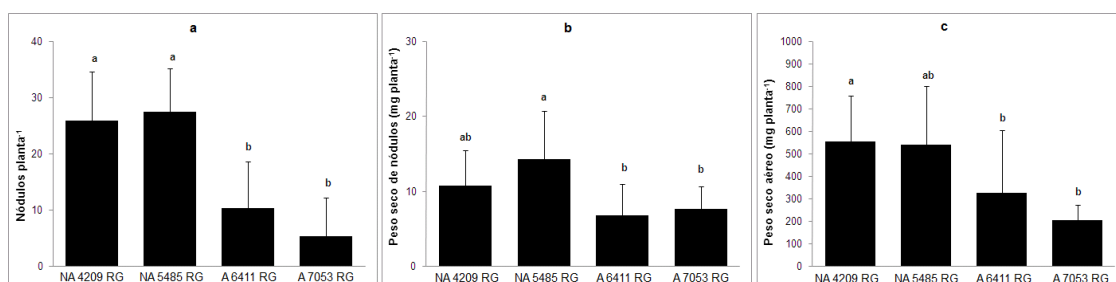


**Figura 2.2** | Respuesta de genotipos de soja a la inoculación con una mezcla de cepas *Bradyrhizobium japonicum* cepas SEMIA 5079 y SEMIA 5080 (valores medios  $\pm$  desvío estándar). La concentración de los inóculos fueron  $10^8$  rizobios  $\text{ml}^{-1}$  **a.** Número de nódulos, **b.** Peso seco de nódulos, **c.** Peso seco aéreo. Mismas letras indican diferencia no significativa (Tukey,  $\alpha=5\%$ ). Los controles fueron inoculados con agua destilada estéril, Control 1 (C1): corresponde al control perteneciente a GM temprano (cultivar NA 3731 RG), control 2 (C2) corresponde al control perteneciente al GM medio (cultivar NA 5579 RG) y control 3 (C3) corresponde al control perteneciente al GM tardío (cultivar A 8000 RG). (\*) diferencia significativa entre cultivar y C1 por el test de Dunnet ( $\alpha=5\%$ ). (†) diferencia significativa entre cultivar y C2 por el test de Dunnet ( $\alpha=5\%$ ) (‡) diferencia significativa entre cultivar y C3 por el test de Dunnet ( $\alpha=5\%$ ).

### 2.3.3. Tercer experimento

Este ensayo tuvo como objetivo confirmar la capacidad de nodulación observada en un número aún menor de genotipos pero además se evaluó si esa capacidad de nodulación del genotipo depende del genotipo de la bacteria o de la interacción bacteria planta. Por eso en este experimento se incluyeron dos cultivares, con alta capacidad de nodulación (NA 5485 R y NA 4209 RG) y otros dos (A 6411 RG y A 7053 RG) con baja capacidad. Estos fueron seleccionados considerando los resultados de los dos primeros experimentos. Todos estos cultivares fueron inoculados en este caso con la estirpe de *B. japonicum* E109.

El ANOVA mostró que los cultivares difirieron en el PSN ( $F = 9.64, P < 0.01$ ), NN ( $F = 4.94, P < 0.01$ ) y PSA ( $F = 5.00, P < 0.01$ ). Los cultivares seleccionados en base a la alta capacidad de nodulación con cepas brasileras SEMIA también desarrollaron alto NN y PSN cuando se inocularon con la cepa E109 (**Figura 2.3**). Estos resultados fueron confirmados por la biomasa de la planta que fue significativamente más alta en los cultivares con los número y peso seco de nódulos más altos.

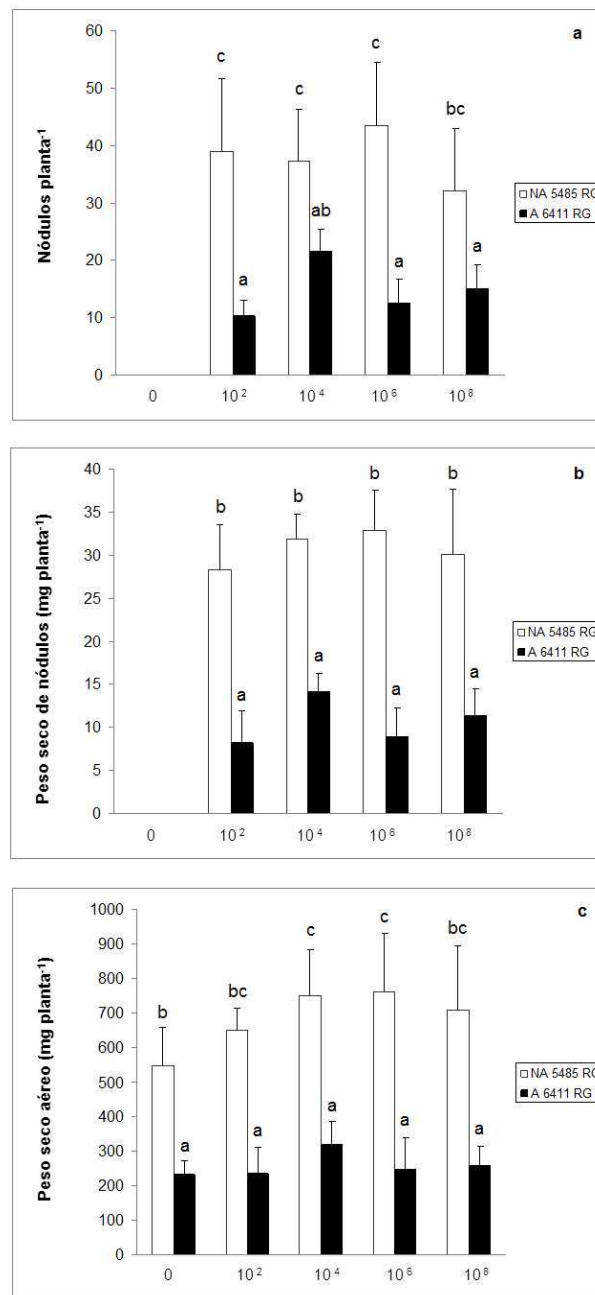


**Figura 2.3** | Respuesta a la nodulación de cuatro cultivares de soja frente a la inoculación con *Bradyrhizobium japonicum* cepa E109 (inoculante comercial) en (valor medio  $\pm$  desvío estándar). La concentración de los inóculos fueron  $10^8$  rizobios  $ml^{-1}$ . **a.** Número de nódulos, **b.** Peso seco de nódulos, **c.** Peso seco aéreo. Mismas letras indican diferencia no significativa Mismas letras indican diferencias no significativas por el test de Tukey ( $\alpha=5\%$ ).



### 2.3.4. Efecto de la concentración bacteriana en la respuesta a la nodulación

La nodulación que se produce como resultado de la interacción *Bradyrhizobium japonicum*-soja es dependiente de la densidad poblacional bacteriana, y en esto interviene un mecanismo que es conocido como “quorum sensing”. Con el fin de determinar el efecto de la población bacteriana, los cultivares de soja NA 4585 RG y A 6411 RG se inocularon con concentraciones crecientes de bradirizobios  $10^2$ ,  $10^4$ ,  $10^6$  y  $10^8$  células  $\text{ml}^{-1}$  de *B. japonicum* E 109. Los resultados se presentan en la **Figura 2.4**.



**Figura 2.4** | Efecto de la densidad celular de *Bradyrhizobium japonicum* sobre la nodulación de *Glycine max* L. Merr cultivar NA 5485 RG (blanco) y A 6411 RG (negro), valores medios  $\pm$  desvío estándar. **a.** Número de nódulos, **b.** Peso seco de nódulos y **c.** Peso seco aéreo. Valores sobre el eje X indican la densidad de *B. japonicum* (células ml<sup>-1</sup>). La concentración de los inóculos fue ajustado respuspendiendo rizobios en PBS estéril. Las plantas control fueron inoculadas con buffer PBS estéril.

Se observó que la concentración de rizobios no alteró la respuesta de los cultivares a la inoculación. Es decir, el cultivar NA 5485 RG desarrolló más nódulos

que A 6411 RG independientemente de la concentración del inoculo. El ensayo por otro lado demostró que el “quorum sensing” es un proceso que impacta en el proceso de la nodulación ya que independientemente del genotipo de soja analizado, la inoculación con concentraciones bacterianas altas provocaron una , reducción de la nodulación.

### 2.3.5. Análisis combinado y estimaciones de heredabilidades

Con los genotipos que se incluyeron en el primer experimento y el segundo se realizó un análisis y este mostró que hubo un efecto significativo de “ambiente” y “cultivar” en el NN, PSN y PSA. El efecto de la interacción “Cultivar x Ambiente” fue significativo excepto para el PSA. Esto indica que aunque la diferencia entre los experimentos fue significativa, la relación entre las respuestas de los cultivares fue similar. Ambos caracteres asociados a la biomasa, PSN y PSA, presentaron errores experimentales mayores al 40% de la variación total (**Figura 2.2**).

El NN y PSA fueron caracteres con mayores valores de  $H_e$ , señalando una variabilidad genética mayor para cada carácter.

**Tabla 2.2** | Análisis estadístico del número de nódulos, peso seco de nódulos y parte aérea de cultivares de soja inoculados con una mezcla de cepas de *Bradyrhizobium japonicum*. Cuadrados medios y porcentaje (%) de la variación total explicada por cada fuente de variación.  $H_e$  indica los valores de heredabilidad en sentido amplio.

Fuente de variación	g.l. <sup>a</sup>	Número de nódulos	Peso seco de nódulos (mg)	Peso seco aéreo (mg)
<b>Cultivares</b>	13	300 ** (28)	312 ** (30.5)	144 ** (26)
<b>Ambiente</b>	1	5842 ** (41.9)	1799 ** (13.5)	1703 ** (23.6)
<b>Cultivares x Ambiente</b>	13	115 ** (10.7)	147 ** (14.6)	56 ns (10.1)
<b>Error</b>	79	34 (19.4)	70 (41.7)	36 (40.3)
<b><math>H_e</math></b>		0.57	0.39	0.48

\*:  $P < 0.05$ , \*\*:  $P < 0.01$ , ns: no significativo, <sup>a</sup>: grados de libertad.

## 2.4. Discusión

La interacción entre la soja y *B. japonicum*, como otras interacciones planta-microorganismo tales como enfermedades causadas por patógenos, puede ser representada por un triángulo formado por el genoma de la planta y el genoma del microorganismo bajo la influencia del ambiente (Van *et al.*, 2007). Esto significa que la interacción está bajo el efecto de estos tres componentes sumado sus interacciones (Yusuf *et al.*, 2008). La simbiosis *Bradyrhizobium*-soja ha sido ampliamente estudiada. Si bien hay algunos trabajos referidos al rol del genotipo de la planta (Vorhees, 1915; Nutman, 1969; Sinclair *et al.*, 1991), éstos muy recientemente han ganado importancia (Nicolás *et al.*, 2002, Nicolás *et al.*, 2006). Esto es particularmente relevante considerando que el desarrollo de nódulos fijadores de nitrógeno es un proceso regulado y complejo que requiere la síntesis e intercambio de señales moleculares entre ambas partes, en este caso *Bradyrhizobium* - soja.

Los genotipos de soja de Argentina presentaron una respuesta diferencial a la inoculación mas probablemente debido a diferencias a nivel genético. Resultados similares se obtuvieron en cultivares brasileros (Bohrer y Hungria, 1998; Hungria y Bohrer, 2000). La continuidad en la distribución de medias sugiere que la herencia es de tipo poligénica y esto coincide con lo encontrado y descrito por Nicolás *et al.* (2002) y Sinclair *et al.* (1991). Además los cultivares con mayor nodulación presentaron hasta el doble de nódulos que los cultivares de baja nodulación. Sinclair *et al.* (1991) y Hungria y Bohrer (2000) también encontraron que los cultivares con alta capacidad de nodulación presentaron un número dos veces mayor de nódulos que los cultivares con menor número de nódulos. Por lo tanto parecería que la diferencia en el número de nódulos entre los cultivares de alta y baja capacidad de nodulación estaría mostrando la diversidad contenida en los cultivares de soja en la Argentina.

El ambiente tiene un impacto profundo en la nodulación de los genotipos (Lie, 1974; Lee y Lee, 1998). En el segundo experimento se observó que la capacidad de

nodulación se redujo a la mitad en relación a lo observado en el primer experimento. Aún así las diferencias relativas entre los cultivares se mantuvieron. Esto significa que, si bien las condiciones ambientales condujeron a un crecimiento de la planta distinto, éstas no alteraron el fenotipo de nodulación determinado por el genotipo. Más aún, el carácter de alta o baja capacidad de nodulación estuvo en función del genotipo de la planta ya que los cultivares que nodularon más respondieron de la misma manera, aun cuando la cepa de rizobio inoculada fue distinta.

El análisis de correlación mostró que el PSN fue la variable que más condicionó a la biomasa aérea. Además, su heredabilidad fue de 0.39, lo que sugiere que puede ser objetivo de los mejoradores para la selección de genotipos que tengan mayor potencial de fijación de nitrógeno. Nicolás *et al.* (2002) citaron heredabilidades de 13% a 58% y 25% a 70% para NN y PSN, respectivamente, cuando evaluaron cuatro poblaciones  $F_3$  derivadas de cruzamiento entre cultivares brasileros. Greder *et al.* (1986) encontraron una heredabilidad entre 54% y 67% para la masa de nódulos en soja inoculada con *B. japonicum*. Los resultados de este trabajo confirman lo citado por otros autores para soja, aunque cabe destacar, que la estimación de la heredabilidad en sentido amplio fue medida en un ambiente controlado.

Para asegurarnos que la respuesta diferencial de los cultivares encontrada no se debió al uso de niveles subóptimos de inoculante que fueran ineficientes en inducir nódulos en la planta, se realizó un experimento usando concentraciones crecientes de densidad bacteriana. La densidad celular de los rizobios no alteró la respuesta de los cultivares a la inoculación sugiriendo que la nodulación diferencial de los cultivares no se debió a la concentración de bacterias en el inoculante para los cultivares en estudio. Por otro lado la reducción de la nodulación observada tanto en cultivares de alta como de baja capacidad de nodulación, probablemente se deba a una alta densidad celular de inoculante. Este fenómeno puede ser una respuesta general de los cultivares de soja, tal como lo ha sugerido Jitackorn y Sadowsky (2008) debido al

mecanismo de “quorum sensing” y probablemente la producción de bradyoxetina estaría involucrada en esta respuesta de reducción de la nodulación.

Los resultados de estos estudios sugieren que los materiales cultivados en Argentina son materiales altamente diversos en términos de su capacidad para nodular. El uso de cultivares de soja con alta capacidad de nodulación asociados a bacterias que fijan N con eficiencia conducirían a aumentar la disponibilidad de N simbiótico beneficiando a los suelos, sin embargo podría impactar en forma negativa en el crecimiento y el rendimiento. El costo energético que implica una elevada masa de nódulos puede ser deletéreo para el índice de cosecha de la soja, como ya se ha observado en leguminosas supernodulantes (Novák *et al.*, 2009).

Considerando la importancia económica de la soja para Argentina y la gran demanda de N por parte de la planta y que la inoculación es una práctica ampliamente aceptada por los productores, la habilidad de los genotipos para nodular y fijar nitrógeno debería ser considerado por los mejoradores como un objetivo importante. Posiblemente la capacidad de nodulación debería ser un carácter a evaluar en genotipos que transiten la última etapa de evaluación para ser liberados al mercado.

Se confirmó en cuanto a la heredabilidad que el PSN es el carácter que más variación genética posee. Por otro lado nuestros resultados sugieren que el grupo de madurez no condiciona la producción de nódulos y, por lo tanto, materiales de diferentes ciclos podrían ser usados como progenitores en planes de mapeo de genes (QTLs) que controlan la nodulación.

## 2.5. Conclusiones

- Se identificaron en los materiales evaluados de soja diversidad fenotípica y genética para la nodulación.
- La número de nódulos fue el que presentó la mayor variabilidad genética ( $H_e = 57\%$ ).

- Se identificaron materiales con alta capacidad de nodulación y biomasa aérea:  
NA 2018 RG, NA 4209 RG, NA 5485 RG y NA 5920 RG.
- Se identificaron materiales con baja capacidad de nodulación y biomasa aérea:  
A 3302 RG, A 6411 RG, A 7053 RG y NA 8499 RG.

## **CAPÍTULO 3**

# **DIVERSIDAD GENÉTICA DE CULTIVARES DE SOJA DE ARGENTINA, CAPACIDAD DE NODULACIÓN Y ASOCIACIÓN CON CARACTERES AGRONÓMICOS**





### 3. DIVERSIDAD GENÉTICA DE CULTIVARES DE SOJA DE ARGENTINA, CAPACIDAD DE NODULACIÓN Y ASOCIACIÓN CON CARACTERES AGRONÓMICOS

#### 3.1. Introducción

El objetivo fundamental de la agricultura actual es obtener materiales con altos rendimientos. Sin embargo, para que muchos de estos materiales alcancen su rendimiento potencial necesitan condiciones óptimas de crecimiento como la provisión de nutrientes, en particular el N y P, que son los que más frecuentemente limitan el rendimiento de los cultivos.

Aún cuando es posible proceder a la fertilización con agroquímicos, la fertilización biológica es la mejor alternativa para fertilizar los suelos, ya que contribuye a realizar un manejo sustentable de la producción.

Las leguminosas, entre ellas, la soja, son plantas que tiene la habilidad de interaccionar con bacterias del suelo Gram (-), como resultado de lo cual desarrollan estructuras denominados nódulos, en los que la bacteria fija nitrógeno (Burriss y Roberts, 1993).

El origen de la soja es Asia, más precisamente China (Hymowitz, 1970) y desde allí, esta planta se distribuyó por el mundo y a los países que en la actualidad son los mayores productores como Estados Unidos, Brasil y Argentina, en donde los suelos carecen de rizobios compatibles. A causa de esto, es una práctica usual inocular el grano de soja, que consiste en adicionar en el momento de la siembra, un alto número de bacterias fijadoras de nitrógeno. Esto da como resultado el desarrollo en la raíz de nódulos fijadores de nitrógeno (Bhuvaneshwari *et al.*, 1981). Los nódulos proveen un ambiente hipóxico, que es crucial para la actividad de la nitrogenasa, enzima

bacteriana que fija nitrógeno (Oke y Long, 1999; Loh y Stacey, 2003). Por lo tanto, el potencial de fijación de la simbiosis depende, dentro de ciertos límites, del número de nódulos desarrollados (Döbereiner, 1966; Bohrer y Hungria, 1998; Hungria y Bohrer, 2000) y de la actividad específica de la nitrogenasa de la cepa de rizobio (Stougaard, 2000).

Es decir que la fijación de nitrógeno en soja es el resultado de determinantes genéticos que se encuentran en ambos simbioses y de la interacción de los mismos. Se han descrito una gran cantidad de genes bacterianos involucrados en la nodulación y fijación de nitrógeno como son los genes comunes de nodulación, los genes de especificidad (*nod/nol/noe*) y los genes de fijación *nif/nix* (Kondorosi *et al.*, 1984; Brewin, 1991; Long, 1996, Downie, 1998). Los genes de especificidad determinan el rango de especies hospedantes, es decir, la compatibilidad de la bacteria con diversas especies y/o cultivares de plantas (Cregan *et al.*, 1989; Balatti, 2004).

Entre los genes de la soja vinculados a la nodulación, se encuentran los que codifican proteínas conocidas como nodulinas, que están involucradas en el desarrollo y funcionamiento del nódulo (Brewin, 1991; Long, 1996). Además se han identificado otros genes que controlan la nodulación así como también la restricción por ciertas cepas de rizobios. Estos loci son llamados *rj(s)* o *Rj(s)* (Hayashi *et al.*, 2012). Los alelos recesivos, *rj1* y *rj5* fueron identificados como receptores putativos, *nfr1* y *nfr5*, respectivamente (Indrasumunar *et al.*, 2010; 2011; Hayashi *et al.*, 2012), y *rj7* fue identificado como *nts1*, el cual está involucrado en la regulación sistémica de la nodulación (Nishimura *et al.*, 2002, Searle *et al.*, 2003). Por otra parte, los alelos dominantes *Rj2*, *Rj3*, *Rj4* y *Rfg1* restringen la nodulación con cepas específicas (o serogrupos) de *Bradyrhizobium* o *Ensifer/Sinorhizobium*. Algunos ejemplos incluyen *Rj2* con *B. japonicum* USDA122, *Rj3* con *B. elkanii* USDA33, *Rj4* con *B. elkanii* USDA61 y *Rfg1* con *E./S. fredii* USDA257 (Caldwell, 1966; Caldwell *et al.*, 1966; Trese, 1995; Vest, 1970; Vest y Caldwell, 1972; Weiser *et al.*, 1990). Recientemente, el gen

de las variantes alélicas *Rj2/Rfg1* se identificó y codifica para una clase de proteínas de resistencia en plantas (R), las cuales confieren resistencia contra patógenos microbianos. Este descubrimiento revela un mecanismo de reconocimiento común tanto en interacciones simbióticas y en hospedador patógeno-bacteria sugiriendo que el establecimiento de la simbiosis requiere de la evasión o supresión de las respuestas inmunes disparadas por los efectores de los rizobios (Yang *et al.*, 2010). Además se identificó el gen *Rj4*, que codifica para una proteína relacionada a patogénesis 5 (PR5) (van Loon *et al.*, 2006), pertenece a una de las 17 familias de proteínas involucradas en la resistencia del hospedador a los patógenos, que además participan en varios procesos del desarrollo (Hayashi *et al.*, 2014). Además de estos genes, se ha reportado que los genotipos de soja presentan respuestas diferenciales a la inoculación, lo que estaría vinculado a genes que codifican proteínas que regulan finamente los procesos (Greder *et al.*, 1986; Nicolás *et al.*, 2002; Salvucci *et al.*, 2012). En conclusión, es evidente que tanto la bacteria como la planta contribuyen a una simbiosis exitosa y, si bien se ha trabajado mucho en la selección de cepas de bacterias, la nodulación no ha sido un carácter importante en los programas de mejoramiento de la soja, excepto en países como Brasil (Hungria y Bohrer, 2000).

En la actualidad se dispone de una considerable cantidad de información del genoma de la soja. Inicialmente, la diversidad del germoplasma de soja se caracterizó en base al análisis de los caracteres morfológicos, más recientemente por características bioquímicas y en la actualidad por marcadores moleculares (Gorman, 1984; Abe y Ohara, 1992; Xu, 1995; Zhou, 1998).

Los mapas genéticos son herramientas claves para los programas de mejoramiento vegetal (Dudley, 1993), ya que permiten asociar marcadores a fenotipos cuando ambos co-segregan y esta asociación permite realizar selección indirecta del carácter a través de los marcadores. Cregan *et al.* (1999) publicó un mapa de ligamiento genético para soja, que junto con herramientas como los SNPs (polimorfismos de nucleótidos simples), y BACs (cromosomas artificiales de bacterias)

proveyeron información clave sobre el genoma de la soja (Shoemaker *et al.*, 2003). Más recientemente, la secuencia completa del genoma de la soja fue descrita por Schmutz *et al.* (2010).

Las técnicas moleculares son particularmente útiles para caracterizar accesiones (individuos, cultivares o poblaciones) y pueden ser de gran ayuda para analizar la diversidad de potenciales genotipos parentales en programas de mejoramiento (Abu Assar *et al.*, 2005).

Los marcadores SSR (repeticiones de secuencia corta), también conocidos como microsatélites, son secuencias de ADN repetidas en tándem de 2-7 pb que se encuentran ampliamente distribuidas en el genoma, generalmente son codominantes y polimórficas (Morgante y Olivieri, 1993). Estos marcadores moleculares han sido particularmente útiles para la identificación, localización y caracterización de los principales QTLs (Tanksley, 1993; Xu *et al.*, 1999). Li *et al.* (2008a) demostraron que 53 de 100 marcadores RAPDs y 35 SSRs fueron polimórficos en una muestra de 101 cultivares de soja pertenecientes a distintas compañías semilleras y el análisis de cluster agrupó en general entre sí los cultivares de una misma compañía. Además encontraron que algunos de estos marcadores estuvieron asociados a los ciclos de maduración. Recientemente, Nicolás *et al.* (2002, 2006) y Santos *et al.* (2006) analizando progenies  $F_2$  y  $F_{2:7}$ , encontraron marcadores SSR ligados a QTLs relacionados con número y peso seco de nódulos, así como también a peso seco aéreo de la planta. Estos autores identificaron varios SSR ligados a QTLs dentro de una población de 157 RILs  $F_{2:7}$ , provenientes del cruzamiento cv Bossier x cv Embrapa 20, cultivares que diferían en la capacidad de fijar nitrógeno (alta x media). Además utilizando la población  $F_{2:3}$  provenientes del cruzamiento Embrapa 20 x BRS133 (media x baja nodulación, respectivamente), se identificaron microsatélites ligados a QTLs de nodulación en los siguientes grupos de ligamiento (GL): A2 (Satt187, Satt233), B1 (Satt251, satt191, Satt509), B2 (Satt181, Satt192), L (Satt232) (Santos *et al.*, 2006).

Hasta el momento en Argentina no existen evidencias de programas de mejoramientos para caracteres asociados a capacidad de nodular y de fijar nitrógeno.

#### **Hipótesis:**

- Existe variabilidad genética entre los cultivares de soja argentinos evaluados a través de marcadores SSR asociados a caracteres de capacidad de nodular.
- Hay una asociación significativa entre los caracteres asociados de capacidad de nodular *versus* caracteres relacionados con producción de materia seca y ciclo de la planta.

#### **Objetivos**

- a) Identificar la diversidad genética de los cultivares argentinos por medio de SSRs ligados a QTLs asociados a nodulación en cultivares brasileños.
- b) Determinar la asociación de estos marcadores con la capacidad de nodulación y otras características agronómicas de los cultivares de soja.
- c) Identificar la generación parental con capacidad de nodulación contrastante para que a partir del cruzamiento entre ellas se obtengan poblaciones segregantes que puedan utilizarse en un estudio de análisis de QTLs para nodulación.

### **3.2. Materiales y métodos**

A partir de los resultados del ensayo de nodulación de los 15 cultivares de soja (2do experimento del capítulo anterior) se tomaron los datos de las variables (NN), peso de nódulos (PSN) y peso seco de la parte aérea de la planta (PSA). También se realizaron determinaciones de ciertos descriptores cuali y cuantitativos de los cultivares de soja como altura de planta (cm), días de maduración, hábito de

crecimiento (determinado: 1/indeterminado: 0), color de flor (blanco: 0/púrpura: 1) y color de pubescencia (gris: 0/ marrón: 1) (**Tabla 3.1**).

**Tabla 3.1** | Descriptores agronómicos y caracteres de nodulación para cada cultivar de soja. NN: número de nódulos por planta, PSN: peso seco de nódulos (mg/planta), PSA: peso seco aéreo (mg/planta), AP: Altura de plata (cm), DM: Días a madurez, HC: Hábito de crecimiento (1: determinado, 0: indeterminado), CF: Color de flor (0: blanco, 1: púrpura) y PUB: color de pubescencia (0: gris, 1: marrón). Los valores son medias de doce repeticiones. HSD: Valor crítico para comparación (Tukey,  $P < 0.05$ ).

Cultivar	PSN	NN	PSA	HC*	CF*	PUB*	DM*	AP*
NA2018	38.52	22.86	629.67	0	1	1	112	75
A3302	23.24	21.42	591.25	0	1	0	120	82
A3550	31.8	25	692.75	0	1	1	126	85
NA4209	27.5	22	582.11	0	1	1	132	93
A4725	30.45	22	690.75	0	1	1	140	108
NA5009	28.06	22.27	480.14	0	0	1	151	100
NA5485	37.75	32.25	747.17	1	1	0	152	95
NA5920	46.04	32.17	857.58	0	1	0	162	98
NA6126	26.48	18.36	690.5	1	0	0	166	106
A6411	19.96	12	209.08	1	1	0	167	105
A7053	19.42	20	217.27	1	1	0	171	119
NA7708	25.39	19.27	444.25	1	1	0	177	119
A8000	28.34	17.8	575.6	1	0	0	180	118
NA8010	22.3	19.64	441.42	1	1	0	181	125
NA8499	24.62	17.75	492.5	1	1	0	187	88
HSD	16.24	10.82	314.59	NE	NE	NE	NE	NE

\*Valores medios fueron asignados por Nidera Semillas S.A. NE: No estimado.

### 3.2.1. Amplificación y análisis de SSR

El ADN genómico se extrajo a partir de 200 mg de tejido fresco de hojas utilizando el protocolo modificado CTAB (Bornet y Branchard, 2001). La cantidad y calidad del ADN extraído se determinó realizando electroforesis en geles de agarosa al 0.8% usando marcador lambda (*Inbio*).

Para la reacción de amplificación de microsatélites (SSR) se usó ADN genómico de cada cultivar (Salvucci *et al.*, 2012) y además se incluyeron como controles internos el ADN genómico de tres cultivares brasileños: Bossier (BOS),

Embrapa20 (E20) y BRS133, con alta media y baja capacidad de nodulación respectivamente (Nicolás *et al.*, 2006; Santos *et al.*, 2006). Con el fin de realizar la PCR (*Polymerase Chain Reaction*) se seleccionaron once pares de primers de microsátelites polimórficos del genoma de la soja (Cregan *et al.*, 1999) identificados previamente como ligados a caracteres asociados a nodulación y FBN, Satt066, Sat\_123, Satt181, Satt187, Satt192, Satt197, Satt216, Satt232, Satt233, Satt251 y Satt509 (Nicolás *et al.*, 2006; Santos *et al.*, 2006). Las secuencias de los primers reverse y forward están disponibles en USDA-ARS Soybean Genome Database (<http://129.186.26.94/SSR.html>).

El volumen total de las reacciones de PCR fue de 10  $\mu$ L. La misma estuvo compuesta por buffer de PCR 1X (2.5 mM Tris-HCl pH 8.3 y 62.5 mM de KCl), 2.5 mM de  $MgCl_2$ , 125  $\mu$ M de cada dNTP, 0,2  $\mu$ M de cada oligonucleótido iniciador (*forward* y *reverse*), una unidad de *Taq* ADN polimerasa y 20 ng de ADN genómico. Las amplificaciones se realizaron en un termociclador (MJ Research modelo PT-200), programado de la siguiente manera: un ciclo inicial de 7 minutos a 94°C, seguido de 30 ciclos de 1 minuto a 94°C, 1 minuto a 50°C y 2 minutos a 72°C, y seguido con una etapa de extensión final de 7 minutos a 72°C. Los productos de reacción se resolvieron en geles de poliacrilamida al 10%, los que se trataron con bromuro de etidio para observar el ADN bajo luz UV. Los tamaños de los fragmentos de ADN fueron estimados comparando las bandas de ADN con marcador de 1 Kb (GIBCO).

### 3.2.2. Análisis estadísticos

Los datos morfológicos de los 15 cultivares fueron estandarizados y se construyó un dendrograma por medio del método de agrupamiento denominado ligamiento promedio no ponderado (UPGMA). Además se calculó el coeficiente de distancia taxonómica (Sneath y Sokal, 1973). Los clusters obtenidos permitieron agrupar a los cultivares en función de su capacidad de nodulación y otras variables de



importancia agronómica. La magnitud de la disimilaridad entre clusters se determinó en las sumas de cuadrados entre clusters para cada variable. Se aplicó un ANOVA y las diferencias significativas entre grupos de cultivares del dendrograma construido en base a las variables cuali y cuantitativas se determinaron por el test de Tukey ( $\alpha=5\%$ ).

El índice de diversidad genética ( $H$ ) fue calculado como:  $1 - \sum P_{ij}^2$ , donde  $P_{ij}$  es la frecuencia de  $j^{\text{mo}}$  alelo en el locus  $i$ , sumado para todos los alelos en un locus.

Para el análisis de los SSR, las bandas se codificaron como presentes (1) o ausentes (0) por alelo de cada SSR. De esta manera se construyó una matriz con los datos (cultivares x alelos de SSR). Comparaciones de a pares fueron calculadas usando el coeficiente de Simple Matching (SM) (Sneath y Sokal, 1973). Con la matriz de similitud se construyó un dendrograma incluyendo los cultivares de Argentina y Brasil, utilizando el método de UPGMA. Se calculó el coeficiente de correlación cofenético como medida de distorsión producida por el método. También se armó otro análisis de cluster con todas las variables, moleculares, cualitativas y cuantitativas. Cada locus SSR fue tratado como variable multiestado, esto significa que cada alelo SSR fue codificado usando diferentes números (1 a 4, desde banda liviana a banda pesada). El dendrograma fue hecho con el coeficiente de correlación de Pearson y para el agrupamiento se usó el método UPGMA (Sneath y Sokal, 1973).

Se realizó un biplot GGE (Yan, 2001) para analizar la interacción entre los cultivares argentinos y los marcadores moleculares (alelos SSR). Se aplicó el modelo SREG1:

$$(y_{ij} - \bar{y}_j) = \lambda_1 \xi_{i1} \eta_{j1} + \lambda_2 \xi_{i2} \eta_{j2} + \varepsilon_{ij}$$

donde:

$y_{ij}$ : es el valor medio del genotipo  $i$  para el marcador  $j$ ;  $\bar{y}_j$ : el valor medio del marcador  $j$  sobre todos los genotipos;  $S_j$ : el desvío estándar del marcador  $j$  entre la media del genotipo;  $\lambda_1$  y  $\lambda_2$ : los autovalores para el CP1 (primer componente principal)

y CP2 (segundo componente principal), respectivamente.  $\xi_{i1}$  y  $\xi_{i2}$ : los *scores* para CP1 y CP2, respectivamente, para el genotipo  $i$ .  $\eta_{j1}$  y  $\eta_{j2}$ : los *scores* para PC1 y PC2, respectivamente, para el marcador  $j$ .  $\varepsilon_{ij}$ : el residual del modelo asociado con el genotipo  $i$  en el marcador  $j$ . El biplot fue construido con la macro de Excel dado por Lipkovich y Smith (2001).

Estas clases fueron usadas para construir una tabla de contingencia de dos vías (clases fenotípicas x SSR). Se realizó un análisis de correspondencia (AC) para investigar la estructura de los datos. La distancia de *Chi-cuadrado* fue usada para juzgar la proximidad de las clases fenotípicas (variable fila) y microsatélites (variable columna). Los valores de las coordenadas de las clases (fila) y alelos SSR (columna) fueron superpuestos en el mismo gráfico (biplot), por lo que las relaciones entre cada set de variables y la relación entre las variables filas y las variables columnas pueden ser apreciadas (Lebart *et al.*, 1984). Una matriz de correlaciones al cuadrado para los factores fila (clases de nodulación) fue calculado en orden de mostrar la proporción de varianza de la variable explicada por cada factor. El análisis numérico fue hecho usando el programa NTSYS-p 2.0 (Rohlf, 1998).

### 3.3. Resultados

La amplificación de los fragmentos SSRs a partir de ADN genómico de los cultivares argentinos de soja permitió identificar polimorfismo de marcadores. Salvo excepción del marcador Satt192, que fue monomórfico, los restantes 10 microsatélites fueron polimórficos presentando: Satt509, Sat\_123, Satt187 y Satt232, 2 alelos, Satt066, Satt251, Satt216, Satt197 y Satt233, 3 alelos y solo Satt181, 4 alelos (**Tabla 3.2**). A pesar que el número de cultivares brasileros estudiados fueron 3, éstos fueron más diversos que los materiales argentinos ya que sus microsatélites presentaron un mayor número de alelos (**Tabla 3.3**). Todos los materiales analizados presentaron un

único patrón de amplificación para cada marcador microsatélite, con la excepción del cultivar A 6411 RG que presentó 2 alelos (A1 y A2) para el Satt251.

El índice de diversidad genética medio para los cultivares de soja fue de  $0.50 \pm 0.18$  y se incrementó levemente cuando se incluyeron en el análisis a los cultivares brasileros,  $0.52 \pm 0.11$  (**Tabla 3.2** y **Tabla 3.3**). El alto nivel de diversidad se debe principalmente a los marcadores microsatélites Satt232, Satt233, Satt251, Satt187, Sat\_123, Satt066 (**Tabla 3.3**).

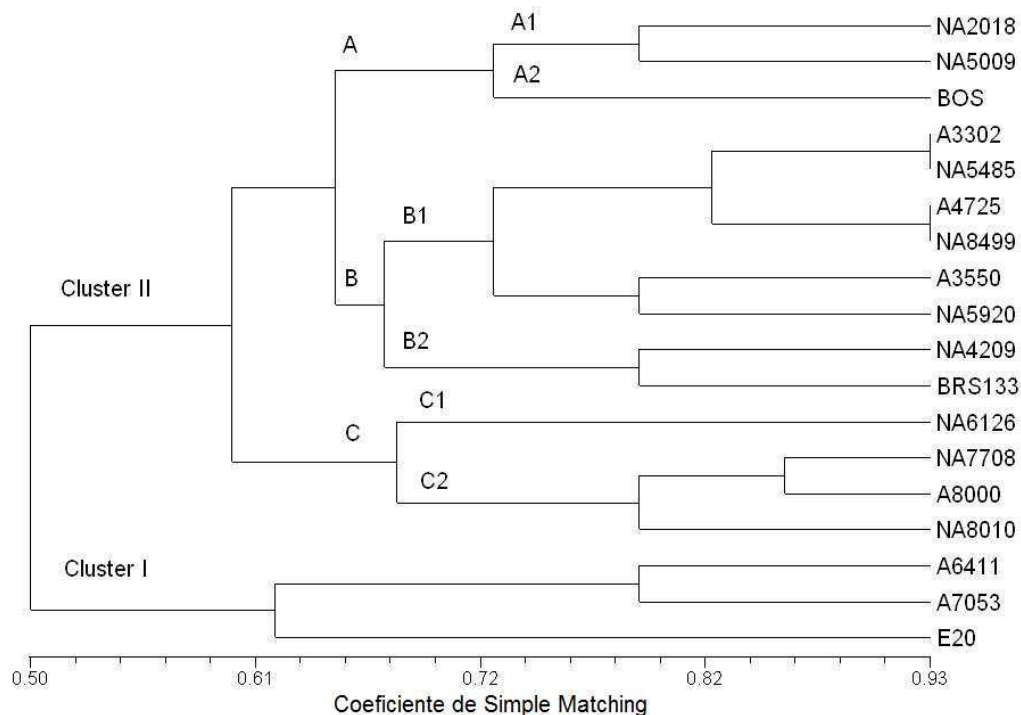
Los microsatélites polimórficos se usaron para construir un dendrograma basado en el coeficiente de SM (**Figura 3.1**). Se pueden observar claramente 2 clusters, I y II que comparten el 50% de los microsatélites. Mientras que el cluster I incluyó los cultivares A 6411 RG, A 7053 RG y E-20, el cluster II incluyó los restantes que se subagruparon en 3 clusters A, B y C. El subcluster A incluyó 3 cultivares que compartieron más del 72% de los SSR, este mismo se subdividió en A1 que incluyó 2 cultivares de Argentina, NA 2018 RG y NA 5009 RG y A2 que incluyó el cultivar brasileros BOS. El grupo B también estuvo compuesto por el cluster B1 que incluyó 6 cultivares de Argentina de ciclo de maduración corto y tardío, A 3302 RG, NA 5485 RG, A 4725 RG, NA 8499 RG, A 3550 RG y NA 5920 RG, y el subcluster B2 que incluyó un genotipo de Argentina NA 4209 RG y uno de Brasil BRS133. Los cultivares del cluster B compartieron el 66% de los microsatélites. El tercer cluster, C, fue dividido, a un nivel del 66% de similitud en 2 subcluster, C1 que incluyó el cultivar NA 6126 RG y el C2, que incluyó tres cultivares genéticamente similares, de los grupos de maduración VII y VIII, NA 7708 RG, A 8000 RG y NA 8010 RG. El coeficiente de correlación cofenética (CCC) fue 0.75 sugiriendo un buen ajuste del dendrograma y la matriz de datos (Sneath y Sokal, 1973).

**Tabla 3.2** | Grupo de ligamiento, tamaño alélico, número de alelos y diversidad genética (*H*) de 10 loci SSR loci en cultivares argentinos

Locus	Grupo de ligamiento	Rango del tamaño de alelos (pb)	Número de alelos	Frecuencia alélica	<i>H</i>
Satt509	B1	290 – 310	2	0.466, 0.533	0.5
Satt066	B2	215 – 240	3	0.866, 0.066, 0.066	0.24
Sat_123	B1	380 – 420	2	0.333, 0.666	0.45
Satt181	H	170 – 210	4	0.266, 0.333, 0.100, 0.300	0.72
Satt187	A2	255 – 300	2	0.800, 0.200	0.32
Satt251	B1	175 – 190	3	0.600, 0.166, 0.233	0.56
Satt216	D1b	200 – 315	3	0.066, 0.666, 0.266	0.48
Satt232	L	380 – 390	2	0.733, 0.266	0.39
Satt197	B1	240 – 290	3	0.266, 0.266, 0.466	0.86
Satt233	A2	250 – 260	3	0.133, 0.200, 0.666	0.5

**Tabla 3.3** | Grupo de ligamiento, tamaño alélico, número de alelos y diversidad genética (*H*) de 10 loci SSR loci en cultivares argentinos y brasileros juntos.

Locus	Grupo de ligamiento	Rango del tamaño de alelos (pb)	Número de alelos	Frecuencia alélica	<i>H</i>
Satt509	B1	290 – 310	2	0.500, 0.500	0.5
Satt066	B2	200 – 270	4	0.722, 0.056, 0.167, 0.056	0.44
Sat_123	B1	380 – 420	2	0.389, 0.611	0.48
Satt181	H	170 – 210	4	0.222, 0.389, 0.083, 0.306	0.7
Satt187	A2	255 – 300	2	0.778, 0.222	0.35
Satt251	B1	175 – 190	3	0.194, 0.250, 0.556	0.59
Satt216	D1b	200 – 315	3	0.667, 0.278, 0.056	0.48
Satt232	L	380 – 390	2	0.667, 0.333	0.44
Satt197	B1	240 – 290	4	0.056, 0.333, 0.222, 0.389	0.68
Satt233	A2	250 – 260	3	0.167, 0.222, 0.611	0.55



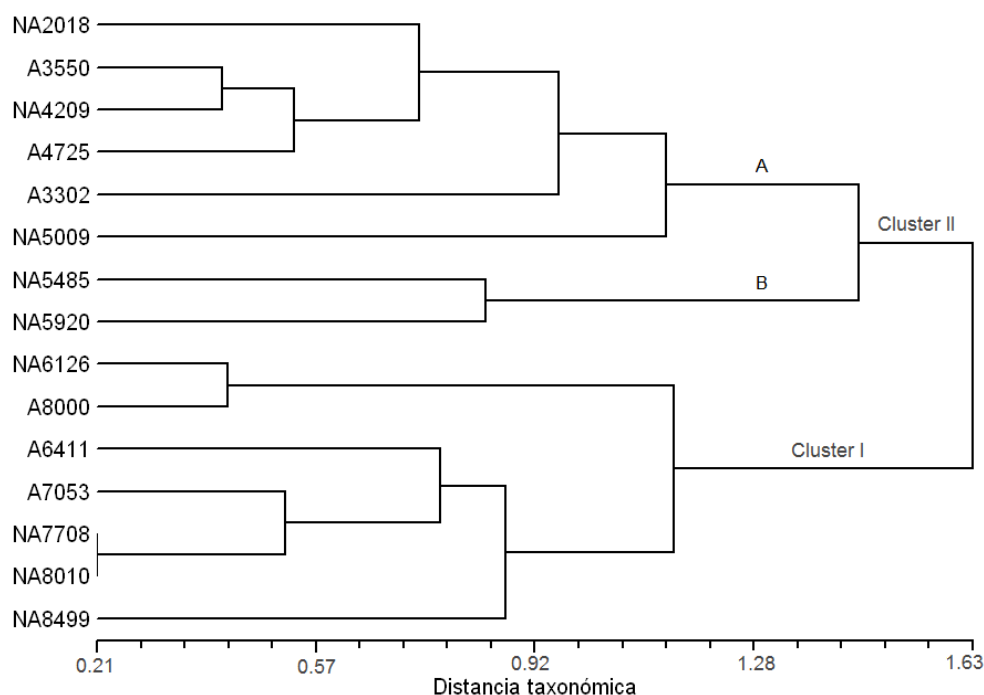
**Figura 3.1** | Análisis de cluster de quince cultivares comerciales argentinos de soja y tres cultivares de Brasil, Bossier (BOS), BRS133 y Empbrapa20 (E20) usando datos moleculares (10 SSR's polimórficos). El dendrograma se construyó con el método UPGMA y se usó el coeficiente de Simple Matching. La distorsión según se indicó fue CCC=0.75.

Dos pares de cultivares, A 3302 RG y NA 5485 RG; A 4725 RG y NA 8499 RG presentaron alto nivel de similitud (93%). Entre los cultivares brasileros, 2 se agruparon con diferentes cultivares argentinos y el tercero E20 fue genéticamente muy diferente del resto de los cultivares separandose en el dendrograma (**Figura 3.1**).

Para el estudio fenotípico de los cultivares argentinos fueron consideradas variables asociadas a capacidad de nodulación (materia seca de la parte aérea, número y peso seco de nódulos) como también otras de importancia agrónomica como altura de la planta, hábito de crecimiento, días de maduración, color de flor y presencia de pubescencia. A partir de dichas variables se construyó un dendrograma (**Figura 3.2**).

Los cultivares se agruparon en dos grupos. El cluster I incluyó plantas con hábito de crecimiento determinado, ciclo de vida largo y mayor altura de planta que desarrollaron un bajo número y peso seco de nódulos ( $< 20$  nódulos planta<sup>-1</sup>,  $< 28$  mg planta<sup>-1</sup>) y el más bajo peso seco aéreo (438.7 mg planta<sup>-1</sup>). El cluster II incluyó dos subgrupos. El subgrupo II-A agrupó cultivares con hábito de crecimiento indeterminado que desarrollaron plantas más bajas (altura  $< 108$  cm) con un ciclo de vida más corto y biomasa seca aérea intermedia (611.1 mg planta<sup>-1</sup>) y número y biomasa de nódulos intermedios ( $> 22$  nodulos planta<sup>-1</sup> y  $> 28$  mg planta<sup>-1</sup>). El subgrupo II-B incluyó dos cultivares NA 5485 RG y NA 5920 RG, que se caracterizaron por ser plantas con una altura intermedia, un ciclo de vida intermedio, que desarrollaron valores altos para número y biomasa de nódulos ( $> 38$  nodulos planta<sup>-1</sup> y  $> 32$  mg planta<sup>-1</sup>) y la biomasa aérea ( $> 802.4$  mg planta<sup>-1</sup>). Se realizó un ANOVA y contrastes entre medias de los grupos obtenidos en el dendrograma fenotípico. El contraste de las diferencias entre clusters para cada variable se realizó por medio del test de Tukey ( $\alpha=5\%$ ). Los ANOVAS demostraron que los clusters se diferenciaron significativamente al nivel de  $P < 0.01$  para NN, PSN, DM, PUB y HC y al nivel de  $P < 0.05$  para el PSA y AP (**Tabla 3.4**).

Para variables relacionadas con la capacidad de nodulación: NN, PSN y PSA se observó que el grupo IIB, presentó medias más altas y se diferenció al menos significativamente del grupo I. (**Tabla 3.4**).



**Figura 3.2** | Análisis de cluster de 15 cultivares comerciales argentinos de soja usando caracteres cuanti y cualitativos. NN: número de nódulos por planta, PSN: peso seco de nódulos, PSA: peso seco aéreo, AP: Altura de la planta, DM: Días a madurez, HC: Hábito de crecimiento, CF: Color de flor y PUB: color de pubescencia.

**Tabla 3.4** | Clasificación de los clusters por ANOVA y diferencias por el test de Tukey. Letras iguales corresponden a diferencias no significativas entre pares.

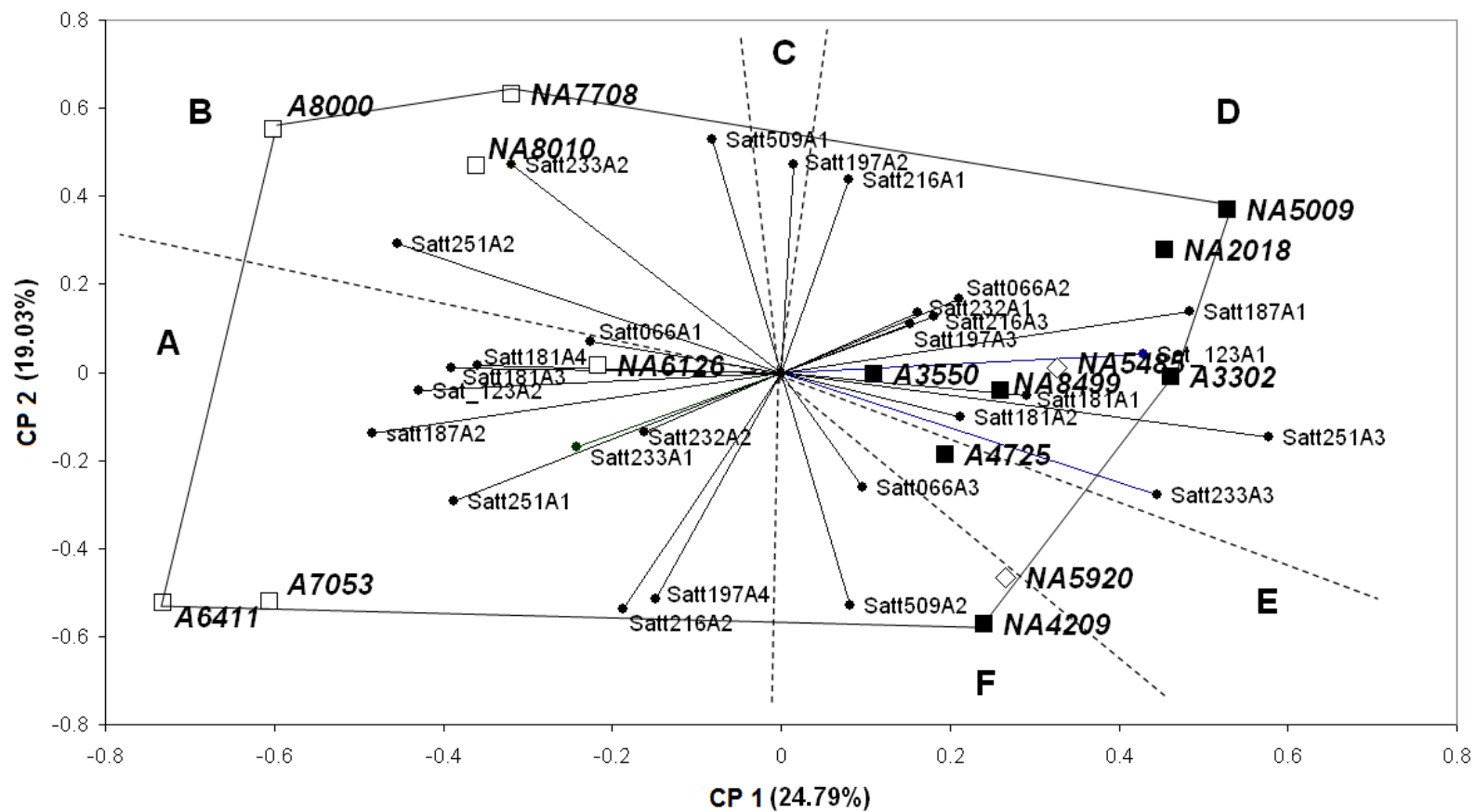
	Cluster I		Cluster IIA		Cluster IIB		CV (%)	F	P-valor
	media	Tukey	Media	Tukey	media	Tukey			
<b>NN</b>	17.83	c	22.59	b	32.21	a	9.65	37.80	<0.001
<b>PSN</b>	23.78	b	29.93	b	41.89	a	15.38	13.50	0.0008
<b>PSA</b>	438.66	b	611.11	ab	802.38	a	24.60	6.30	0.0135
<b>AP</b>	111.43	a	97.00	ab	90.50	b	11.82	5.09	0.0251
<b>PUB</b>	0	b	0.83	a	0.00	b	79.06	18.00	0.0002
<b>DM</b>	175.57	a	130.17	b	157.00	a	6.98	28.50	<0.001
<b>CF</b>	0.71	a	0.83	a	1.00	a	54.27	0.37	0.7008
<b>HC</b>	1	a	0.00	c	0.50	b	38.27	38.80	<0.001

### 3.3.1. Comparaciones de las clasificaciones moleculares y morfológicas

Se construyeron dos matrices de distancia, una basada en caracteres fenotípicos cuantitativos y cualitativos, y la otra basada en la presencia/ausencia de marcadores microsatélites. Se observó que ambas estuvieron correlacionadas significativamente ( $r = -0.39$ ,  $P < 0.01$ ), confirmando que hay cierta congruencia entre los datos morfológicos y moleculares.

La relación de los genotipos de Argentina y los alelos marcadores SSR fueron analizados por medio de un GGE biplot. Éste fue construido con el primer (CP1) y segundo (CP2) componente principal y explicó el 43.8% de la variabilidad genética (**Figura 3.3**). Por razones prácticas, los cultivares que diferían en su capacidad de nodulación que pertenecieron a diferentes clusters dentro del dendrograma (**Figura 3.2**) se identificaron con distintos símbolos. El biplot confirmó los resultados del dendrograma. Los cultivares del cluster I fueron agrupados en el lado negativo del CP1 y los cultivares incluidos en el cluster IIA y IIB se agruparon sobre el lado positivo del eje 1, (**Figura 3.3**). Coincidentemente, los cultivares a lo largo del primer componente se distribuyeron en base a su ciclo de vida, los de ciclo de vida largo sobre el lado negativo y los de ciclo corto de maduración, del lado positivo. Sólo el cultivar NA 8499 RG no se ajustó a este patrón, ya que se ubicó junto con los cultivares de ciclo intermedio a corto.





**Figura 3.3** | GGE Biplot (Modelo I) mostrando la relación entre genotipos argentinos (cuadrados y rombos) y marcadores SSR (círculos). El biplot se construyó con el primer componente principal (CP1) contra el segundo componente principal (CP2), ambos explicaron un 43.82% de la variación total. La identificación de cultivares de acuerdo a los clusters y subclusters de la **Figura 3.2**. Cuadrados blancos: baja capacidad de nodulación (cluster I), cuadrados negros: cultivares de capacidad media de nodulación (cluster II-A), rombos: cultivares de alta capacidad de nodulación (cluster II-B).

### 3.3.2. Diversidad genética de cultivares de soja y microsatélites

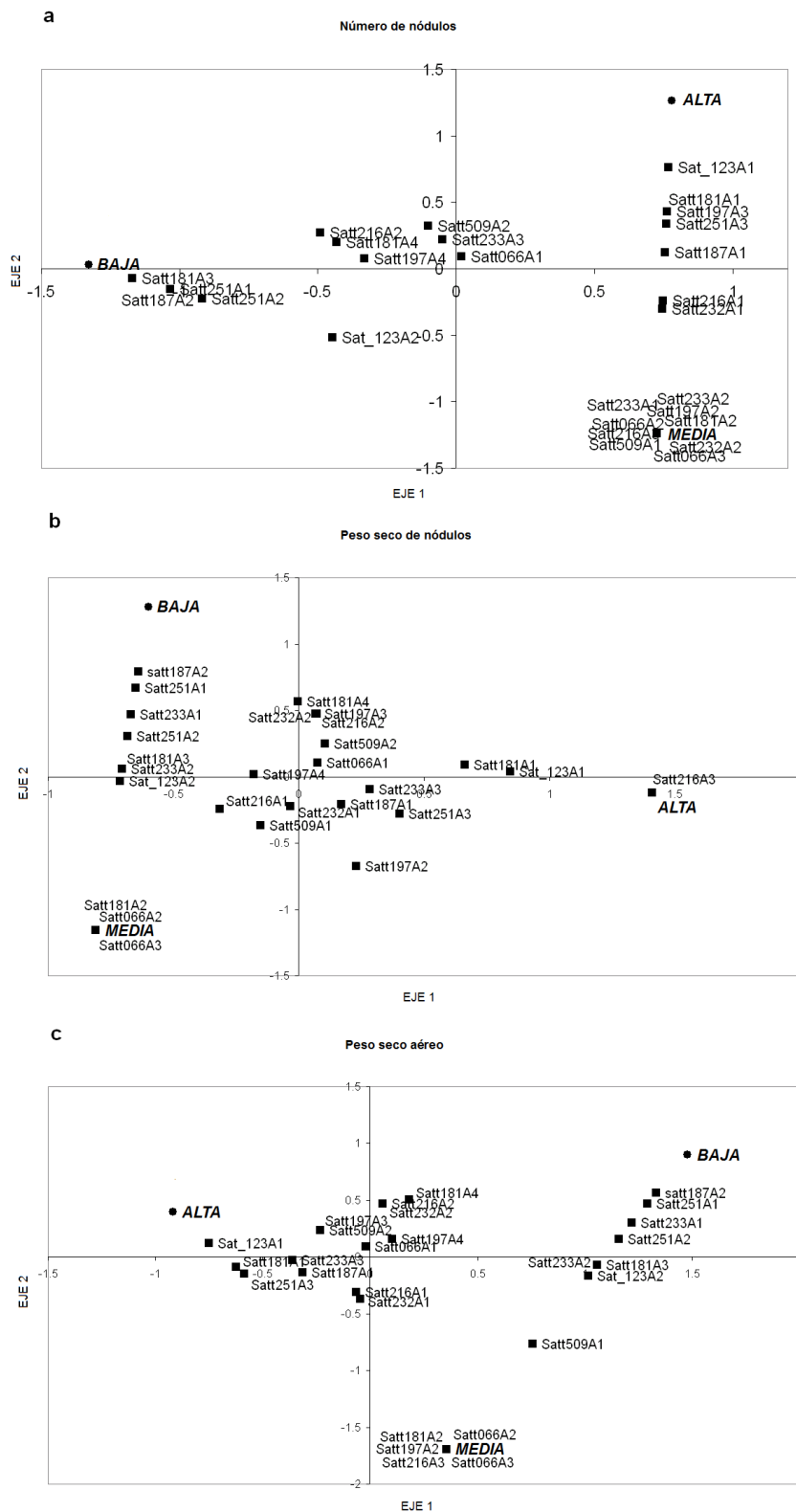
La **capacidad discriminatoria** de un microsatélite en el biplot está en relación directa con la longitud del vector, el cual está representado por la distancia entre los marcadores (círculos) y el punto de origen (0,0) (**Figura 3.3**). Los alelos de los microsatélites Satt251 A1, A2 y A3, Satt187 A1 y A2, Satt233 A3, Sat\_123 A1 y A2 y Satt181 A3 y A4, presentaron altos scores en el primer componente, sugiriendo que poseen una alta capacidad para discriminar genotipos. Los alelos de los microsatélites con mayor discriminación a lo largo del segundo componente fueron Satt216 A1 y A2, Satt509 A1 y A2, satt197 A4 y A2 y Satt233 A2. Estos marcadores fueron los que determinaron la distribución de los cultivares a lo largo del segundo componente que no mostró un patrón particular. La proyección de los microsatélites sobre el segundo componente (eje y) es un **índice de la representatividad**. Aquellos alelos cercanos al origen, tienen baja capacidad de discriminar debido a que son compartidos por genotipos agrupados en diferentes clusters (Yan, 2001). Se construyó un polígono trazando una línea llena uniendo los cultivares localizados en los extremos. Se trazó una línea punteada perpendicular en cada lado del polígono, que dividió el biplot en varias secciones que incluyeron cultivares asociados a microsatélites (Yan, 2001), (**Figura 3.3**). Se identificaron 6 secciones, que fueron definidas por 6 cultivares vértices: A 6411 RG (A), A 8000 RG (B), NA 7708 RG (C), NA 5009 RG (D), A 3302 RG (E) y NA 4209 RG (F). Mientras que los cultivares en secciones opuestas tienen patrones moleculares divergentes, aquellos en la misma sección tenían un patrón molecular similar al del cultivar vértice, compartiendo con éste los alelos SSR incluidos en la misma sección (Yan, 2001).

Los cultivares con baja capacidad de nodulación se ubicaron en tres secciones diferentes, A, B y C. La sección A incluyó a los cultivares A 6411 RG, A 7053 RG y NA 6126 RG que compartieron los mismos alelos: Satt251 A1, Satt066 A1, Sat\_123 A2 y Satt197 A4. Los cultivares A 6411 RG y A 7053 RG son genéticamente similares, y

compartieron alelos que tuvieron alto poder discriminante en ambos ejes, tales como: Satt181 A4, Satt509 A2, Satt187 A2 y Satt216 A2. De manera similar, las secciones B y C agruparon cultivares muy relacionados como NA 8010 RG, NA 7708 RG y NA 8000 RG. Estos cultivares compartieron los alelos Satt251 A2, Satt216 A1, Satt509 A1 y Satt233 A2. Sin embargo, el cultivar NA 7708 RG vértice de la sección C fue el único cultivar de esta sección que portó el alelo Satt197 A2, que también estuvo presente en los cultivares de la sección D, NA 5009 RG y NA 2018 RG. Aquellos cultivares localizados hacia la derecha del gráfico presentaron los alelos Satt251 A3, Satt187 A1, Satt233 A3 que fueron los que más discriminaron (mayor longitud del vector). La sección D agrupó a los cultivares NA 5009 RG, NA 2018 RG, NA 5485 RG, A 3550 RG y NA 8499 RG. El cultivar A 3302 RG fue el vértice de la sección E, dentro de la cual incluyó al cultivar A 4725 RG. Ambos cultivares compartieron alelos con los cultivares de la sección D y F. Sat\_123 A1 se encontró en los cultivares NA 5009 RG, NA 2018 RG, NA 5485 RG de la sección D y en el cultivar A 3302 RG de la sección E. El cultivar NA 4209 RG definió la sección F y fue el único que presentó el alelo Satt066 A3. Además el alelo Satt509 A2 fue compartido por los cultivares de la sección F, aunque también estuvo presente en el genoma de A 7053 RG y A 6411 RG de la sección A, y los cultivares A 4725 RG y A 3302 RG de la sección E. Es importante mencionar que los cultivares con alta capacidad de nodulación, NA 5485 RG y NA 5920 RG se agruparon en dos secciones diferentes D y F, sugiriendo que hay divergencia genética entre dichos materiales. Si bien NA 5920 RG se agrupó dentro de la sección F, éste compartió los alelos Satt197 A4, Satt216 A2 y Satt509 A2 con los cultivares A 6411 RG y A 7053 RG de baja capacidad de nodulación (cluster I). Se puede observar que algunos cultivares con alta nodulación de la sección D presentaron un patrón altamente divergente de aquellos cultivares con baja capacidad de nodulación incluidos en la sección A y B (**Figura 3.3**).

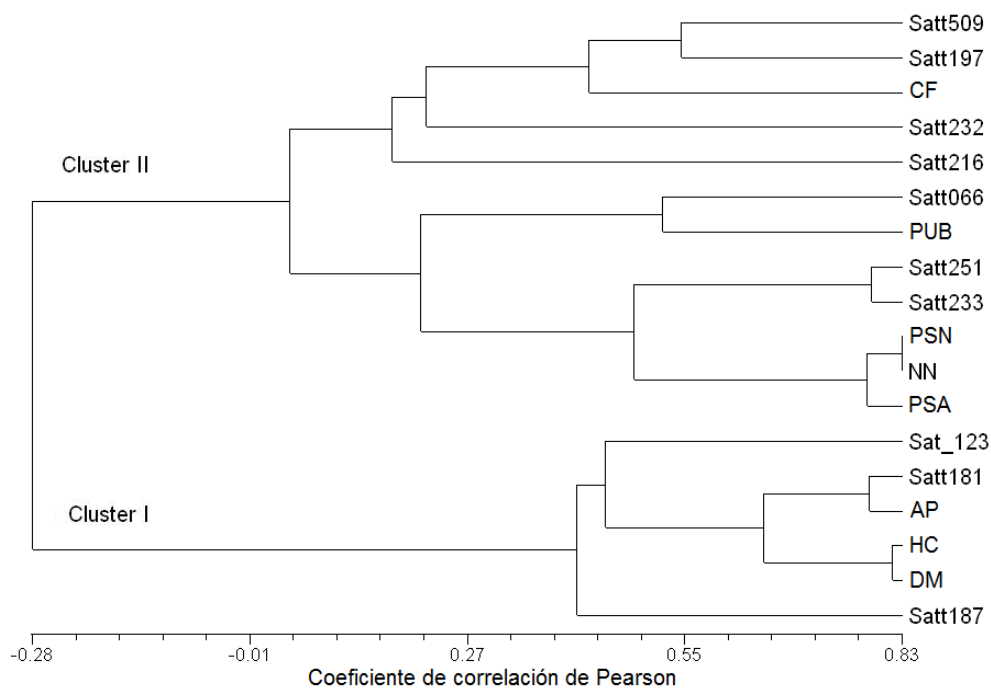
En estudios anteriores se correlacionó la capacidad de nodular con las variables en estudio: NN, PSN y PSA (capítulo 2 de la presente tesis; Hungría y

Bohrer, 2000; Sinclair *et al.*, 1991; Souza *et al.*, 2008), por lo tanto, con el objeto de asociar grupos con diferente capacidad de nodular y microsátélites se realizó un AC para cada una de estas variables (**Figura 3.4**). Los genotipos que desarrollaron un NN alto y medio se agruparon sobre el lado positivo del eje 1, aquellos que desarrollaron un bajo NN estuvieron del lado negativo. Cultivares con NN intermedios fueron discriminados por el eje 2 (cuadrante inferior derecho) (**Figura 3.4a**). Aunque el patrón no fue idéntico, los cultivares que difirieron en su masa de nódulos fueron discriminados de una manera similar por el AC. Aquellos cultivares que produjeron más biomasa de nódulos se agruparon en el lado positivo del eje 1 y aquellos que produjeron una biomasa intermedia o menor lo hicieron dentro del cuadrante negativo (**Figura 3.4b**). De cualquier manera, los materiales con biomasa de nódulos intermedia se agruparon en base al segundo eje sobre el cuadrante inferior izquierdo (**Figura 3.4b**). Con respecto a la biomasa aérea (**Figura 3.4c**) los cultivares que produjeron la mayor biomasa aérea se agruparon en el lado negativo del primer eje, y los cultivares que produjeron poca biomasa aérea se agruparon sobre el lado positivo del eje.



**Figura 3.4** | Biplot del análisis de correspondencia para categorías de capacidad de nodulación (baja, media y alta) y los diferentes alelos microsatélites. **a** Número de nódulos: eje 1 y eje 2 tuvieron 68 % y 31 % de inercia, respectivamente. **b** Peso seco de nódulos: eje 1 y eje 2 tuvieron 58 % y 42 % de inercia, respectivamente. **c** Peso seco aéreo: eje 1 y eje 2 tuvieron 61 % y 29 % de inercia, respectivamente.

Finalmente se realizó un análisis de cluster considerando la totalidad de caracteres: cuantitativos, cualitativos y moleculares (**Figura 3.5**). El dendrograma de caracteres obtenido mostró dos clusters principales, I y II que estuvieron negativamente correlacionados ( $r = -0.28$ ). El cluster I asoció el microsatélite Sat\_123, Satt181 y Satt187 con variables de ciclo (AP, HC Y DM), mientras que el cluster II agrupó a los microsatélites Satt251 y Satt233 con variables asociadas con la capacidad de nodular (**Figura 3.5**). En este mismo cluster, Satt066 se correlacionó con el color de pubescencia. Y Satt509 y Satt197 se correlacionaron con CF.



**Figura 3.5** | Análisis de cluster usando 10 SSR's polimórficos y variables cuali y cuantitativas consideradas como multiestado. Marcadores: Satt509, Satt197, Satt232, Satt216, Satt066, Satt251, Satt233, Sat\_123, Satt181 and Sat187. NN: número de nódulos por planta, PSN: peso seco de nódulos (mg/planta), PSA: peso seco aéreo (mg planta<sup>-1</sup>), AP: Altura de planta (cm), DM: Días a madurez, HC: Hábito de crecimiento, CF: Color de flor y PUB: color de pubescencia. El dendrograma se construyó con el metodo de agrupamiento UPGMA y se usó el coeficiente de correlación de Pearson.

### 3.4. Discusión

Los marcadores microsatélites se encuentran ampliamente distribuidos en el genoma, por lo que brindan información reproducible y amplia y de gran utilidad en los programas de mejora. En este trabajo se utilizaron microsatélites asociados a caracteres de nodulación en cultivares de soja de origen brasileiros (Nicolás *et al.*, 2002; 2006; Santos *et al.*, 2006), como herramientas para hacer un “screening” de este carácter en los cultivares argentinos de soja. Los 11 microsatélites identificaron 27 alelos diferentes, entre ellos el 66.67% tuvieron frecuencias superiores a 0.25 (**Tabla 3.2**). Cuando el análisis incluyó tres cultivares brasileiros se identificaron otros 2 alelos adicionales. En la población alélica, el 55.17% tuvo frecuencias superiores a 0.25, mientras que no hubo alelos con frecuencias menores a 0.05 (**Tabla 3.3**). Estos resultados confirmaron que los alelos estuvieron altamente representados en los cultivares analizados, que fueron 15 de Argentina y 3 de Brasil. El análisis mostró un considerable nivel de diversidad para la mayoría de los microsatélites utilizados y éstos, además fueron altamente eficientes para discriminar genotipos de soja. Los microsatélites Satt509, Satt251, Satt197, Satt233 presentaron valores de  $H$  superiores al 50% y fueron los que mostraron una mayor capacidad para discriminar, tal cual se observa en el GGE-biplot (**Figura 3.3**) (Yan, 2001).

El dendrograma generado con los patrones moleculares, agrupó a los cultivares en 5 grupos, con valores de similitud genética que oscilaron entre el 50 y el 93% (**Figura 3.1**). Li *et al.* (2008a) evaluó la diversidad de 101 genotipos de soja con marcadores RADP y SSR y encontró valores de similitud genética entre 45-100% y 36-100 %, respectivamente. Los resultados obtenidos de este trabajo sugieren que el nivel de diversidad de los materiales analizados es menor (0.5) que el reportado por Li *et al.* (2008b). Esto podría ser explicado por una reducción de la base genética producida durante el proceso de mejoramiento con el objetivo de generar cultivares de soja con caracteres de interés agronómico, aunque también el número de genotipos

muestreados fue considerablemente bajo. Alternativamente estos resultados podrían ser consecuencia debido a un bajo número de marcadores SSR, aunque hay que considerar que entre los once microsatélites, diez consistieron en repeticiones en tándem de trinucleótidos, y como se ha reportado este tipo de SSR muestran con mayor eficiencia la diversidad de los genomas (Song *et al.*, 1999). Estos microsatélites se seleccionaron en base a su asociación con la nodulación (Nicolás *et al.*, 2002; 2006; Santos *et al.*, 2006), y muy probablemente como el carácter no estuvo bajo presión de selección, estos determinantes genéticos son altamente variables (Greder *et al.*, 1986; Nicolás *et al.*, 2002). Sólo algunos de estos marcadores estuvieron asociados a nodulación en los genotipos de soja estudiados, lo que sugiere que es necesario identificar más microsatélites para realizar un “screening” de QTLs de nodulación en el germoplasma argentino.

Cuando en el análisis se incluyeron los 3 cultivares brasileños, el índice de diversidad poblacional fue más alto, lo que sugiere que los cultivares brasileños en general podrían contribuir a ampliar la base genética de los cultivares de soja argentinos. Es importante mencionar que los cultivares brasileños incluidos en este estudio, se eligieron por su capacidad de nodulación divergente y porque presentaron microsatélites con patrones electroforéticos distintos (Nicolás *et al.*, 2002; 2006 ; Santos *et al.*, 2006). El cultivar E-20 compartió, con el resto de los cultivares, solo el 50% de los alelos, lo cual sugiere que es el cultivar genéticamente más distante y probablemente el responsable del incremento en la diversidad. Además, estos resultados sugieren que los niveles de diversidad de los genotipos de Argentina y Brasil son diferentes, aunque la dimensión de las diferencias encontraría demandaría utilizar un mayor número de marcadores. Además, se encontró que el cultivar A 6411 RG era un material aún en segregación, ya que presentó un patrón heterocigota asociado al microsatélite Satt251, marcador que estuvo asociado a nodulación, detectándose además, segregación en las plantas del cultivar, es decir, que podrían estar segregando para el carácter de nodulación.



El análisis de cluster de las variables fenotípicas generó dos grupos de cultivares, uno de ellos presentó dos subclusters A y B que incluyeron materiales con capacidad de nodulación media y alta (**Figura 3.2**). Los cultivares con baja capacidad de nodulación tuvieron un hábito de crecimiento determinado y ciclos de vida largo, mientras que aquellos cultivares con alta capacidad de nodulación presentaron ciclo de vida intermedio. Resultados similares fueron observados por Balatti y Pueppke (1992) mientras estudiaban la eficiencia de nodulación de ciertos genotipos de soja con la bacteria de crecimiento rápido *Ensifer (Sinorhizobium fredii)*. Ellos encontraron que la longitud del ciclo de vida del genotipo estuvo correlacionada con la habilidad de los cultivares de desarrollar nódulos fijadores de nitrógeno o nódulos ineficientes. Aparentemente los caminos involucrados en la respuesta de la soja a la infección y nodulación con *Ensifer (Sinorhizobium fredii)* sería similar que con *Bradyrhizobium japonicum*. Cabe destacar que, este estudio que incluyó solamente 15 cultivares mostró una asociación entre los caracteres asociados a capacidad de nodulación y el ciclo de vida del cultivar. Esto sugeriría una asociación entre los microsatélites ligados a capacidad de nodulación y el ciclo de vida del cultivar. La distribución genotípica que se observó en el GGE biplot muestra la alta asociación entre la clasificación genotípica y fenotípica, lo que también fue confirmado por el AC. Estos resultados en conjunto sugieren la existencia de un ligamiento entre los microsatélites y los descriptores analizados en este trabajo.

Los caracteres que definen el ciclo de vida como son los días de maduración y la fecha de floración se agruparon juntos con el carácter altura de planta, el cual tiene efecto en la morfología. Reyna y Sneller (2001), Tasma *et al.* (2001) y Tasma y Shoemaker (2003) también encontraron microsatélites asociados a QTLs ligados a días de maduración. Los resultados de este trabajo confirman la asociación altamente significativa entre las variables relacionadas con ciclo de vida y altura de planta, aún cuando los cultivares desarrollaron en condiciones controladas. Li *et al.* (2008a) también encontró que los cultivares con valores similares de unidades de temperatura

se agruparon en los mismos clusters aún cuando la clasificación en el dendrograma se construyó con marcadores moleculares. Los genes determinantes del ciclo de vida podrían estar asociados a un origen geográfico particular, por consiguiente, ciclos de madurez diferentes estarían asociados a divergencia genética. Frecuentemente todas son características consideradas en los programas de mejoramiento y la selección de padres divergentes.

Es importante destacar que Satt187, Satt181 y Sat\_123 fueron los tres SSRs con mayor capacidad de discriminación y estos estuvieron principalmente asociados con ciclo de vida y con la altura de la planta. Los restantes, Satt251 y Satt233, estuvieron relacionados a capacidad de nodulación (**Figura 3.5**). La correlación negativa entre ambos grupos de microsatélites podría estar indicando un ligamiento en fase de repulsión entre variables de ciclo y nodulación y como consecuencia de esto, entre los microsatélites ligados a estos caracteres.

Aunque el biplot explicó un bajo porcentaje de la variabilidad genética (43.8%), ésta no se refiere a la variabilidad total, debido a que el biplot sólo refleja la variación debida a los cultivares (genotipo) y la interacción genotipo x marcadores (Yan *et al.*, 2001). Sin embargo, el biplot fue un análisis útil tanto para asociar el patrón fenotípico y genotípico, como para seleccionar padres putativos de futuras poblaciones segregantes. En base a lo expuesto anteriormente, los potenciales progenitores con alta capacidad de nodulación son: NA 5485 RG o NA 5920 RG, mientras que los de baja capacidad de nodulación son A 7053 RG, A 6411 RG, NA 7708 RG, A 8000 RG o NA 8010 RG. Como uno de los objetivos del mejoramiento debe ser generar una población con base genética amplia en lo que hace a la capacidad de nodulación, los mejores cruzamientos serían aquellos producidos entre cultivares ubicados en sectores opuestos del gráfico, ya que reflejan la mayor divergencia genética, por ejemplo, NA 5485 RG debería ser cruzada con A 7053 RG o A 6411 RG, mientras que el cultivar NA 5920 RG con NA 7708 RG, A 8000 RG o NA 8010 RG.

En este estudio se mostró la asociación entre los microsatélites, Satt251 y Satt233 y las variables que definen el carácter nodulación en cultivares comerciales de soja de Argentina. Como la nodulación presenta una heredabilidad poligénica (Nicolás *et al.*, 2002; 2006; Santos *et al.*, 2006; Tanya *et al.*, 2005), más genes con mayor o menor efecto en la nodulación deberían estar dispersos en el genoma de soja, es muy probable que se encuentre un número adicional de genes cuantitativos de nodulación cuya identificación requiere trabajar con un mayor número de marcadores moleculares y/o con otra población de estudio.

Los estudios descriptos condujeron a la identificación de los padres putativos para la obtención de poblaciones segregantes para el análisis de QTLs de nodulación.

### 3.5. Conclusiones

- Se confirmó una alta representatividad de los alelos en los cultivares argentinos y brasileros analizados. El análisis mostró un considerable nivel de diversidad para la mayoría de los microsatélites utilizados, y estos además fueron altamente eficientes para discriminar cultivares de soja.
- El dendrograma generado con los patrones moleculares, agrupó a los cultivares en 5 grupos, con valores de similitud genética que oscilaron entre el 50 y el 93 %.
- En los cultivares argentinos se confirmó no sólo la existencia de una asociación entre los caracteres de capacidad de nodulación, el ciclo de vida y con morfología de la planta, sino también, una asociación entre dichos caracteres con algunos microsatélites.
- La distribución genotípica que se observó en el biplot muestra la alta asociación entre la clasificación genotípica y fenotípica, lo que también fue confirmado por el AC.

- En los cultivares argentinos se confirmó una asociación entre los microsatélites, Satt251 y Satt233 y las variables que definen el carácter nodulación. Sat\_123, Satt181 y Satt187 se asociaron con variables de ciclo de vida de la planta (AP, HC Y DM). Satt066 se correlacionó con el color de pubescencia.
- Los potenciales progenitores con alta capacidad de nodulación son: NA 5485 RG o NA 5920 RG, mientras que los de baja capacidad de nodulación son A 7053 RG, A 6411 RG, NA 7708 RG, A 8000 RG o NA 8010 RG. Ambos grupos pueden ser cruzados entre si para la obtención de poblaciones segregantes aplicables al estudio de QTLs de nodulación.



## **CAPÍTULO 4**

### **MAPEO E IDENTIFICACIÓN DE QTLS ASOCIADOS AL CARÁCTER “CAPACIDAD DE NODULACIÓN”**



## 4. MAPEO E IDENTIFICACIÓN DE QTLS ASOCIADOS AL CARÁCTER “CAPACIDAD DE NODULACIÓN”

### 4.1. Introducción

La interacción simbiótica entre leguminosas y rizobios contribuye entre 91 a 163 millones de toneladas de nitrógeno por año, el 65% de éste es usado en la agricultura (Burris y Roberts, 1993). La eficiencia de la fijación biológica de nitrógeno depende del tipo de cepa de rizobio, de las variedades cultivadas así como también de los factores ambientales y prácticas de manejo del cultivo. Después de la fotosíntesis, la fijación biológica de nitrógeno (FBN) en las leguminosas puede ser considerada como el proceso biológico más importante a nivel mundial. Sin embargo, muchos países, tanto desarrollados como en desarrollo, no han adoptado o tomado ventajas del potencial de la FBN y son prácticamente dependientes de fertilizantes nitrogenados para manejar la productividad de los cultivos. Las leguminosas fijan aproximadamente 20 a 100 kg N ha<sup>-1</sup> año<sup>-1</sup> y usan casi todo el nitrógeno fijado en la producción de granos (Evans *et al.*, 1989). Sin embargo las leguminosas han sido mejoradas principalmente para rendimiento y resistencia a factores abióticos y bióticos, pero son escasos los programas de mejora para aumentar la eficiencia en la FBN. Por otro lado, las compañías de inoculantes se han dirigido sus esfuerzos a mejorar y utilizar en los formulados cepas de bacterias que fijen mas eficientemente el N y que puedan ser utilizadas en una amplia variedad de leguminosas (Abi-Ghanem *et al.*, 2013).

La interacción simbiótica entre la leguminosa y el rizobio es altamente específica e involucra mecanismos complejos que comienzan con un intercambio mutuo de moléculas señales difusibles entre la planta y el rizobio (Ferguson *et al.*, 2010). Este proceso de señalización desencadena la formación de un nuevo órgano



en la raíz, el nódulo, el cual, hospeda a la bacteria fijadora de nitrógeno (Subramanian *et al.*, 2006; Kouchi *et al.*, 2010). Es decir que la simbiosis no depende sólo del genoma de la planta y del genoma de la bacteria sino también de la interacción que se establece entre ambos.

Mediante estudios del genoma de las bacterias fijadoras de nitrógeno, se han descrito varios genes de la nodulación que se conocen como *nod*, *nol* y *noe* (Masson-Boivin *et al.*, 2009). Sin embargo, los estudios de las leguminosas hospedantes son más complejos, debido a sus características biológicas y al tamaño de sus genomas. Más recientemente, la integración de enfoques genéticos y genómicos y la adopción de modelos de leguminosas como *Medicago truncatula* y *Lotus japonicum* abrieron un marco a la investigación de genes del hospedante involucrados en el proceso de FBN (Kouchi *et al.*, 2010; Murray *et al.*, 2011; Oldroyd, 2013). Recientemente, en leguminosas por medio de experimentos con mutantes, clonado de genes y el uso de estrategias de mapeo genético, se caracterizaron veintiséis genes, involucrados en el reconocimiento de las señales disparadas por el rizobio y en los diferentes pasos necesarios para la formación del nódulo (Kouchi *et al.*, 2010). Además se han encontrado decenas de genes que codifican proteínas conocidas como nodulinas, utilizando herramientas de alto rendimiento (*high-throughput*), tales como bancos de marcadores EST (*Expressed Sequence Tags*) y análisis de microarreglos (*Microarrays*) (Brechenmacher *et al.*, 2008; Libault *et al.*, 2010).

Desde 1950, en la soja, se han descrito varios loci que controlan la nodulación y en la última década la mayoría fueron clonados y caracterizados. Estos loci son llamados *rj(s)* o *Rj(s)* y se los puede dividir en tres categorías según su función (Hayashi *et al.*, 2012). La primera categoría corresponde a genes recesivos caracterizados en plantas de soja mutantes no nodulantes. Estos genes alélicos son *rj1* y *rj5*, sus productos codifican proteínas llamadas NFR1 y NFR5 (NFR: Nodule Factor Receptor), las cuales formarían un complejo receptor e interaccionarían con el factor Nod en su región extracelular a través de su dominio de lisinas (LysM). La

respuesta consistiría en la transducción de la señal a través de su dominio con actividad kinasa intracelular, generando finalmente la infección y organogénesis del nódulo (Indrasumunar *et al.*, 2010; 2011; Hayashi *et al.*, 2012). En la segunda categoría se conoce el gen *rj7*, que fue identificado como *nts1* (*nitrate tolerant symbiosis 1*) a partir de mutantes hipernodulantes y tolerantes a altas concentraciones de nitrato en el suelo inducidos con mutagénicos químicos. Este gen está involucrado en la regulación sistémica de la nodulación (Nishimura *et al.*, 2002, Searle *et al.*, 2003). Su producto génico Rj7/NTS1/GmNARK (*G. max Nodule autoregulation receptor kinase*) formaría un complejo receptor con una proteína tipo CLV2 (CLAVATA2) ubicada en el tallo de la planta. Este complejo receptor reconoce señales derivadas que son producidas en respuesta al sensado de los factores Nod en la raíz, resultando en la producción de la señal/es de AON (autorregulación) en su dominio intracelular con actividad kinasa, inhibiendo la formación posterior de nuevos nódulos. Dentro de la tercer categoría se encuentran los alelos dominantes *Rj2*, *Rj3*, *Rj4* y *Rfg1*. Éstos restringen la nodulación con cepas específicas (o serogrupos) de *Bradyrhizobium* o *Ensifer/Sinorhizobium*. Algunos ejemplos incluyen *Rj2* con *B. japonicum* USDA122, *Rj3* con *B. elkanii* USDA33, *Rj4* con *B. elkanii* USDA61 y *Rfg1* con *E./S. fredii* USDA257 (Caldwell, 1966; Caldwell *et al.*, 1966; Vest, 1970; Vest y Caldwell, 1972; Weiser *et al.*, 1990; Trese, 1995). Los genes *Rj2/Rfg1* y *Rj4* recientemente fueron identificados y forman parte de proteínas involucradas en resistencia de las plantas a patógenos microbianos. Estos descubrimientos revelan un mecanismo de reconocimiento común tanto en interacciones simbióticas y como hospedante-patógeno (bacteria), sugiriendo que el establecimiento de la simbiosis requiere la evasión o supresión de las respuestas inmunes disparadas por los rizobios (Yang *et al.*, 2010). El gen *Rj3* aún no ha sido identificado.

Genes ortólogos a los mencionados fueron identificados usando herramientas de mutagénesis dirigida, primeramente en *Medicago truncatula* (Amor *et al.*, 2003;

Limpens *et al.*, 2003) y *Lotus japonicus* (Madsen *et al.*, 2003; Radutoiu *et al.*, 2003) y posteriormente en arveja (*Pisum sativum*) (Zhukov *et al.*, 2008).

Estos descubrimientos sientan las bases para comprender los mecanismos moleculares de la regulación de la simbiosis en el hospedante. Sin embargo la capacidad de nodulación y FBN presentan una arquitectura genética compleja, la cual, depende de la interacción planta-bacteria. No sólo está regulada por genes cualitativos de herencia mendeliana simple como los descritos anteriormente, sino también intervienen numerosos loci génicos cada uno con un efecto pequeño. Estos últimos bajo la influencia del ambiente, dan como resultado la expresión de caracteres cuantitativos. En soja se ha observado una variabilidad considerable en cuanto a la capacidad de nodulación y la eficiencia de FBN entre cultivares (Neuhausen *et al.*, 1988; Sinclair *et al.*, 1991; Pazdernik *et al.*, 1997; Bohrer y Hungria, 1998; Hungria y Bohrer 2000; Salvucci *et al.*, 2012). Esta variabilidad sería una fuente genética importante para los programas de mejoramiento con el fin de aumentar el potencial de rendimiento. Algunos de estos proyectos de mejoramiento específicos para potenciar la FBN en soja fueron iniciados en la década de 1980 en Australia, en los Estados Unidos y en África. El programa australiano exploró una variación natural de tolerancia al nitrato (altos niveles de nitratos en el suelo afectan negativamente la nodulación). El programa comprendió dos ciclos de selección entre 1986 y 1997. En un primer ciclo se realizaron cruzamientos entre líneas coreanas tolerantes a nitrato y cultivares adaptados. Finalmente se obtuvieron cuatro líneas avanzadas F<sub>7</sub>, que fueron comparadas por FBN y producción de granos con sus progenitores putativos. Las líneas presentaron desempeños semejantes a una de las líneas coreanas para FBN (50% del N proveniente de FBN) y para productividad (que varió entre 1.7 tn ha<sup>-1</sup> y 2.67 tn ha<sup>-1</sup>). Sin embargo los nuevos cultivares no superaron la productividad de las líneas del programa principal de mejoramiento australiano. Además, todos los cultivares fueron susceptibles a *Phytophthora sojae*. Este no es más que un ejemplo de las dificultades que pueden surgir durante el proceso de mejora para la obtención de

cultivares superiores con caracteres combinados: alta fijación de nitrógeno, resistencia a enfermedades y mayor rendimiento. Como consecuencia el programa fue finalizado en 1998, y ningún cultivar con mayor capacidad de nodulación pudo liberarse al mercado (Herridge y Rose, 2000; Herridge *et al.*, 2001).

Otro abordaje para el desenvolvimiento de materiales tolerantes a nitrato consistió en la generación de mutantes *nts* (gen que luego fue caracterizado, ver más arriba). Los cultivares mutantes que se describieron fueron Bragg (Carrol *et al.*, 1985), Williams (Gremaud y Harper, 1989) y Enrei (Akao y Kouchi, 1992). Estos mutantes presentan casi diez veces más nódulos que sus pares salvajes y altas tasas de FBN, tanto en ausencia como en presencia de nitrato. Estudios sobre la herencia del fenotipo de *nts* ha comprobado la naturaleza recesiva del gen y pruebas de alelismo indicaron que toda las mutaciones ocurren en el mismo gen, sugiriendo que éste es particularmente susceptible a mutagénesis química (Vuong y Harper, 2000). Experimentos realizados en Australia y Estados Unidos con mutantes de los cultivares Bragg y Williams, evidenciaron una reducción del 20% al 41% de la productividad en relación a los progenitores (Herridge *et al.*, 2001). Sin embargo, un efecto positivo que se observó en los campos cultivados con mutantes *nts*, fue un aumento en la productividad de los cereales cultivados luego de la cosecha de soja. Song *et al.* (1995) atribuyen ese efecto a dos factores: menor absorción de N mineral del suelo por las plantas *nts* y un aumento de N en el suelo debido a la descomposición de residuos vegetales de esas plantas conteniendo altas concentraciones de N.

Sinclair *et al.* (2007) describieron progresos obtenidos en un programa iniciado en 1993, con el objetivo de desarrollar líneas productivas en ambientes secos, explorando el carácter de fijación de N<sub>2</sub>. Se utilizó como uno de los progenitores el cultivar Jackson, tolerante a sequía. Como resultado, dos líneas presentaron tasas de FBN superiores a los de su progenitor y fueron tolerantes a sequía.

En África, ocurre una situación inversa, donde pequeños agricultores han cultivado variedades locales de soja con baja o ninguna aplicación de fertilizantes

nitrogenados y sin o con uso de inoculantes. Esta característica es denominada nodulación promiscua, porque las variedades son capaces de nodular con diversas especies de rizobios nativos de los suelos africanos, nodulando y fijando eficientemente nitrógeno (Herridge y Rose, 2000; Mpereki *et al.*, 2000). El programa de mejoramiento desarrollado desde inicio del año 1980 por el Instituto de Agricultura Tropical (IITA), en Nigeria, ha explorado la combinación de esa capacidad de nodulación promiscua de los materiales locales con las características mejoradas de los cultivares americanos, para obtener líneas que combinen la productividad y la nodulación promiscua (Herridge y Rose, 2000). Ese programa ha tenido éxito en asegurar una nodulación promiscua en diferentes ambientes, variedades locales como Magoy y Hermon 147 fueron recomendadas y son cultivadas por pequeños agricultores (Mpereki *et al.*, 2000).

En el caso australiano, posiblemente, la falta de comprensión de los mecanismos de control genético de estos caracteres hayan colaborado para el fracaso de los programas de mejora, especialmente si dichas variables presentan una herencia compleja o poligénica. En estos casos es indispensable estimar parámetros genéticos tales como la heredabilidad para calcular el avance genético en la selección del carácter (Falconer y Mackay, 1996). También se debe considerar las asociaciones entre caracteres, que pueden tener causas genéticas o ambientales. Entre las causas genéticas se encuentran el ligamiento génico o bien la pleiotropía, este concepto último se aplica cuando un gen controla a más de un carácter simultáneamente. Si bien el ligamiento génico puede ser quebrado por recombinación, la pleiotropía es permanente (Falconer y Mackay, 1996). Una metodología para abordar esta problemática es el análisis de coeficientes paso que ha sido frecuentemente utilizada por los mejoradores de plantas para asistir en la identificación de caracteres que son útiles durante el proceso de selección (Li, 1975). El método mide la influencia directa de una variable predictora sobre la variable respuesta. Utiliza una regresión parcial estandarizada para particionar los coeficientes de correlación en efectos directos e

indirectos, aplicados a un diagrama causal construido de acuerdo a una lógica básica. Este análisis permite un examen crítico de cuales son los factores específicos que producen una correlación dada y que pueden ser empleados en formular una estrategia de selección efectiva.

Todavía se requieren estudios de la arquitectura genética de la nodulación y fijación de nitrógeno por el lado del hospedante. Esta arquitectura genética puede construirse mapeando poligenes involucrados en la expresión del carácter cuantitativo. Para ello se utiliza como herramienta los marcadores moleculares a lo largo de los cromosomas, para elaborar mapas genéticos saturados, esto es con innumerables marcadores a lo largo de los cromosomas, permitiendo ubicar físicamente en dichos mapas regiones de actividad cuantitativa (QTLs: Quantitative Trait Loci) (Cubero, 2003). Los marcadores microsatélites (SSR) son útiles para identificar, localizar y caracterizar a los principales factores genéticos de un QTL. Éstos se caracterizan por ser secuencias cortas repetidas de 1 a 6 pares de bases de ADN distribuidas ampliamente en el genoma, son codominantes, altamente polimórficos exhibiendo un gran número de alelos por locus (Tanksley, 1993; Xu *et al.*, 1999).

No sólo es importante localizar los QTLs en el genoma, sino también estimar sus efectos génicos (Kearsey y Pooni, 1996). Esto permite estimar la heredabilidad del QTL, definida como la proporción de la varianza genética debida a la varianza fenotípica del carácter. Esta disección genética ayuda a poder asistir en los proyectos de mejoramiento de caracteres de interés. En soja se han identificado QTLs asociados a: potencial de rendimiento, resistencia a enfermedades, contenido de proteínas y aceite en semillas (Soybase 2012).

Se han realizado estudios preliminares sobre regiones genómicas que controlan la capacidad de nodulación y la FBN en soja (Tanya *et al.*, 2005; Nicolas *et al.*, 2006; Santos *et al.*, 2006). También hay información sobre otras leguminosas, como arveja (Bourion *et al.*, 2010) y poroto (*Phaseolus vulgaris* L.) (Nodari *et al.*, 1993; Tsai *et al.*, 1998). Estos estudios han demostrado el potencial del mapeo de QTLs

para dichos caracteres, sin embargo todavía se requiere mayor información en relación a los genes de la planta.

Entre 1995 y 2009 diversas publicaciones dan crédito de la aplicación de la selección asistida por marcadores, que ha provocado grandes avances en el mejoramiento de caracteres cuantitativos en cuanto a resistencia a enfermedades y rendimiento en soja (Concibido *et al.*, 2003; Zhu *et al.*, 2007; Saghai Maroof *et al.*, 2008; Shi *et al.*, 2009). Es de esperar que estos avances ocurran también para los caracteres asociados a la capacidad de nodular.

### **Hipótesis**

- El genoma de la soja contiene genes que controlan la capacidad de nodular y fijar nitrógeno y que presentan una herencia de tipo cuantitativa.

### **Objetivos**

- a) Estudiar la asociación entre variables relacionadas con la capacidad de nodular y de fijar nitrógeno y estimar la variabilidad genética existente al aplicar un modelo genético.
- b) Localizar y determinar el modo de herencia de los QTLs asociados a la capacidad de nodulación mediante la aplicación de métodos de análisis de QTLs utilizando marcadores SSR.

## **4.2. Materiales y métodos**

### **4.2.1. Obtención de la población segregante $F_{2:3}$**

Sobre la base de los resultados obtenidos en los estudios de capacidad de nodulación de los cultivares de soja y en lo que hace a la asociación de SSR con la

capacidad de nodulación, se seleccionaron los progenitores NA 5485 RG<sup>5</sup> (alta capacidad de nodulación) y A 7053 RG<sup>6</sup> (baja capacidad de nodulación), que fueron cruzados durante la campaña 2009-2010. A partir de este cruzamiento se generó la población segregante F<sub>2:3</sub>, con la que se realizó un análisis de mapeo de QTLs. Las semillas F<sub>1</sub> obtenidas a partir de los cruzamientos, fueron sembradas en Octubre de 2010 en el campo experimental del Instituto Fitotécnico de Santa Catalina. El carácter híbrido de las plantas F<sub>1</sub>, se confirmó con los marcadores moleculares tipo microsatélites (SSR), utilizando como molde ADN foliar. Los productos de amplificación de SSR-PCR se resolvieron en geles de agarosa haciéndose hincapié en aquellos que fueron contrastantes entre los progenitores. En otoño de 2011 se cosecharon 210 semillas F<sub>2</sub>, que fueron sembradas en invernáculo en condiciones controladas de luz y temperatura en la localidad de Venado Tuerto (Nidera Semillas S.A.), durante la campaña 2011-2012. Sobre cada individuo de la población F<sub>2</sub> se tomaron muestras de tejido foliar, para la evaluación genotípica. Se cosecharon las semillas de las plantas F<sub>2</sub> individualmente, disponiéndose de la generación de semillas F<sub>3</sub>. Se eliminaron aquellas familias que produjeron menos de 10 granos por planta, quedando un total de 94 familias, las cuales fueron incluidas en un ensayo con un diseño completamente al azar (DCA) con repeticiones, en condiciones controladas de luz y temperatura, durante la campaña 2012-2013.

#### 4.2.2. Evaluación fenotípica de la población F<sub>2:3</sub>

Con las 94 familias F<sub>2:3</sub> se realizó un ensayo de nodulación en cámaras de crecimiento bajo condiciones controladas de luz y temperatura (fotoperíodo de 14 hs de luz y 25 °C, respectivamente). El diseño del ensayo fue completamente al azar con

---

<sup>5</sup> y <sup>6</sup> Dado que la soja es una especie autógena, se consideró a los cultivares progenitores homocigotas.



12 repeticiones por familia  $F_3$ . La inoculación y las condiciones fueron las mismas ya explicadas en el punto 2.2.2. Se usó como inoculante *B. japonicum* cepa E109 y las plantas fueron mantenidas en speedlings.

A los 30 días del inicio del ensayo se cosecharon las plantas  $F_3$  y se determinó el número (NN) y peso seco de nódulos (PSN). Las partes aéreas de las mismas se separaron en tallo y hojas, los que se secaron en estufa a 70°C hasta peso constante. Se determinó el peso seco de hoja (PSH, en mg) y el peso seco del tallo (PST, en mg), la sumatoria de las cuales generó el peso seco de la parte aérea (PSA, en mg). Además, se consideró el peso seco de nódulo por nódulo, el cual fue calculado como la razón entre el peso seco de nódulo y el número de nódulos por planta ( $PSN/NN = PSNN$ , en  $mg\ nódulo^{-1}$ ). La altura del tallo principal (AP, en cm) fue considerada como variable adicional.

Con los datos fenotípicos, se armó una matriz con las medias de cada carácter promediando las plantas dentro de cada familia  $F_{2:3}$ .

### 4.2.3. Análisis fenotípico de familias $F_3$

Con las familias  $F_3$  se construyeron histogramas de frecuencias, para todas las variables estudiadas (R-Cran, 2008). Se aplicó un diseño completamente aleatorizado desbalanceado, para todas las variables asociadas a la capacidad de nodulación (Sokal y Rohlf, 1995).

El modelo fue:

$$y_{ij} = \mu + g_i + \varepsilon_{ij}$$

$y_{ij}$  : Observación o variable respuesta

$\mu$  : Media general

$g_i$  : Efecto del tratamiento o familia  $i$

$\varepsilon_{ij}$ : Error o variación no controlada que existe entre los individuos de cada familia y varía de una observación a otra.

Si se considera un modelo de efectos aleatorios, las varianzas se construyen por descomposición de las esperanzas matemáticas de los cuadrados medios  $E(CM)$  como se indica en la siguiente tabla de ANOVA:

FV	GL	CM	E(CM)	F
Entre familias	$g-1$	$CM_T$	$\hat{\sigma}^2 + k \hat{\sigma}_G^2$	$CM_T/CM_R$
Error (entre individuos dentro de familias)	$N-g$	$CM_R$	$\hat{\sigma}^2$	
Total	$N-1$			

$CM_T$ : Cuadrado medio entre familias

$CM_R$ : Cuadrado medio dentro de familias

Por ser un diseño desbalanceado se calculó el número de observaciones  $k$  como:

$$k = \frac{N - \frac{1}{N} \sum_{i=1}^g r_i^2}{g - 1} \text{ donde } N = \sum_{i=1}^g r_i$$

$g$ : número de familias: 94

$r_i$ : número de individuos dentro de cada familia.

Se estimaron los siguientes parámetros genéticos (Cruz, 2001):

Varianza fenotípica media:  $\hat{\sigma}_F^2 = \frac{CM_T}{k}$

Varianza ambiental media:  $\hat{\sigma}^2 = \frac{CM_R}{k}$

Varianza genotípica media:  $\hat{\sigma}_G^2 = \frac{CM_T - CM_R}{k}$

Correlación intraclase:  $\hat{\rho} = \frac{\hat{\sigma}_G^2}{\hat{\sigma}_G^2 + \hat{\sigma}^2}$

El CV% fue calculado como la razón entre la raíz de la varianza entre plantas dentro de familias ( $\sigma^2$ ) y la media general en porcentaje.

La correlación intraclase es tomada como la heredabilidad en sentido amplio para cada carácter ( $H^2$ ) (Falconer y Mackay, 1996).

En una  $F_3$ , la varianza ambiental, cuyo estimador es el  $CM_R$  está confundida con el componente genético de la varianza (Mariotti, 1986; Kearsey y Pooni, 1996).

#### 4.2.3.1. Análisis de asociación entre variables

A partir de los datos fenotípicos de las siete variables cuantitativas en las familias  $F_3$ , se armó una matriz con los índices de correlación genética y fenotípica entre todos los pares de caracteres. Varianzas para cada carácter individual y covarianzas para los pares de caracteres se estimaron usando el programa Genes (Cruz, 2001). Las correlaciones genéticas y fenotípicas entre los pares de caracteres fueron determinadas por las siguientes expresiones:

Correlación genética:

$$r_{G(x,y)} = \frac{\text{COV}_{G(x,y)}}{\sqrt{\sigma_{G(x)}^2 \sigma_{G(y)}^2}}$$

en que:

$r_{G(x,y)}$ : es el coeficiente de correlación genética entre las características x e y;

$\text{COV}_{G(x,y)}$ : es una covarianza genética entre las características x e y;

$\sigma_{G(x)}^2$  y  $\sigma_{G(y)}^2$ : son las varianzas genéticas de las características x e y;

Correlación fenotípica:

$$r_{F(x,y)} = \frac{\text{COV}_{F(x,y)}}{\sqrt{\sigma_{F(x)}^2 \sigma_{F(y)}^2}}$$

en que:

$r_{F(x,y)}$ : es el coeficiente de correlación fenotípica entre las características x e y;

$\text{cov}_{F(x,y)}$ : es una covarianza fenotípica entre las características x e y;

$\sigma_{F(x)}^2$  y  $\sigma_{F(y)}^2$ : son las varianzas fenotípicas de las características x e y.

Posteriormente se confirmó la significancia estadística de cada coeficiente de correlación ( $r$ ), planteando la hipótesis nula  $H_0: r = 0$  versus la hipótesis alternativa

$H_a: r \neq 0$ , mediante una prueba de  $T$ , dada por  $T_c = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ . La  $T$  calculada ( $T_c$ ) se

comparó con una  $T$  tabular ( $T_t$ ) al nivel de significancia de  $\alpha=5\%$  y  $\alpha=1\%$  y con

$(n-2)$  grados de libertad. La regla de decisión fue: si  $T_c \geq T_t$ , entonces el valor de  $r$

es estadísticamente diferente de cero (Espitia *et.al.*, 2008).

#### **4.2.3.2. Análisis de coeficientes de paso.**

El análisis de coeficiente de paso fue utilizado para particionar los coeficientes de correlación fenotípicos en efectos directos e indirectos (Li, 1975). Se construyó un diagrama lógico de causa-efecto con los caracteres considerados, teniendo en cuenta la significancia de los coeficientes de correlación obtenidos entre variables ( $P < 0.05$ ).

El PSA fue adoptado como variable dependiente. Las variables PSH, PST, AP fueron consideradas como variables de primer grado y NN y PSN como de segundo grado (**Figura 4.1**). Para evaluar la existencia de efectos de colinealidad se siguió el procedimiento de Montgomery y Peck (1981). Estos autores proponen el índice CN o número de condición matricial que se estima como el cociente entre el más alto y el más bajo autovalor. Si  $CN < 100$ , la multicolinealidad no es problema, entre  $100 < CN < 1000$  la colinealidad es moderada a fuerte, si presenta  $CN > 1000$  es severa. Los cálculos se realizaron con el programa Genes (Cruz, 2001). Posteriormente al estudio

de colinealidad se eliminó la variable PSNN, que no fue incluida en el diagrama. El índice de colinealidad CN del modelo usando 5 variables explicativas y como variable dependiente PSA fue de 81.07 (CN < 100 colinealidad baja).

La correlación fenotípica fue particionada dentro de efectos directos e indirectos de acuerdo a las siguientes ecuaciones lineales:

### Primer grado

Peso seco aéreo (PSA)

$$r_{21} = P_{21} + r_{23} P_{31} + r_{24} P_{41}$$

$$r_{31} = r_{23} P_{21} + P_{31} + r_{34} P_{41}$$

$$r_{41} = r_{42} P_{21} + r_{43} P_{31} + P_{41}$$

### Segundo grado

Peso seco de hoja (PSH)

$$r_{52} = P_{52} + r_{56} P_{62}$$

$$r_{62} = r_{56} P_{52} + P_{62}$$

Peso seco de tallo (PST)

$$r_{53} = P_{53} + r_{56} P_{63}$$

$$r_{63} = r_{56} P_{53} + P_{63}$$

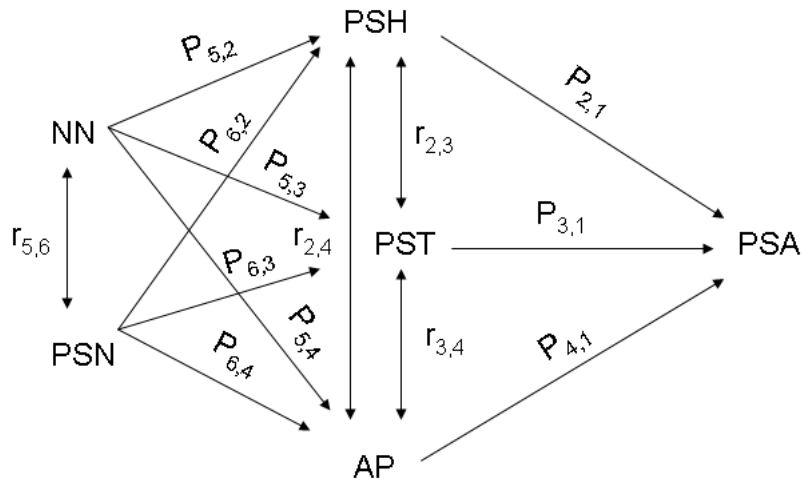
Altura de planta (AP)

$$r_{54} = P_{54} + r_{56} P_{64}$$

$$r_{64} = r_{56} P_{54} + P_{64}$$

Donde: 1-PSA, 2-PSH, 3-PST, 4-AP, 5-PSN y 6-NN (**Figura 4.1**).

Donde  $r$  es el coeficiente de correlación fenotípica lineal entre dos variables,  $P$  es el efecto directo del coeficiente de paso; el producto  $r P$  es la medida del efecto indirecto de un variable sobre otra (Li, 1975).



**Figura 4.1** | Diagrama de coeficientes de paso. P son los efectos directos y r los coeficientes de correlación de Pearson entre las variables.

#### 4.2.4. Evaluación genotípica de la población segregante $F_2$

De cada individuo de la  $F_2$  se extrajo con papel nucleico ADN genómico para proceder al “screening” usando 221 loci SSR ubicados a intervalos de entre 15-20 cM en cada cromosoma (mapa Soybase <http://soybase.org/>). Evaluando en la población parental el polimorfismo de dichos microsatélites, se seleccionaron los SSR a amplificar sobre la población  $F_2$  (**Tabla A 1**). Los SSRs polimórficos se utilizaron para determinar el genotipo (homo o heterocigota) de los individuos de la población  $F_2$  (94 individuos).

Con los datos moleculares se construyó una matriz genotípica “loci SSR x genotipo” considerando cada loci polimórfico SSR para los 94 individuos sobrevivientes de la  $F_2$ . La codificación para cada individuo según el programa GQMOL (Schuster y Cruz, 2004) y considerando el SSR como marcador codominante fue: 0 para el genotipo idéntico al progenitor de alta capacidad de nodulación (NA 5485 RG), 2 para el genotipo idéntico al progenitor de baja capacidad de nodulación (A 7053 RG) y 1 para el genotipo heterocigota, que presenta los alelos de ambos progenitores.

#### **4.2.4.1. Prueba de segregación de marcadores**

A partir de la matriz genotípica se puso a prueba la segregación esperada para la codominancia típica de los SSRs 1:2:1, para cada locus marcador polimórfico. Para esto se utilizaron tres pruebas, análisis de locus individual: Chi-cuadrado, análisis múltiple FDR (False Discovery Rate) y protección de Bonferroni (estos últimos controlan la tasa de error tipo I global). Dichas pruebas de segregación se realizaron con el programa GQMOL (Schuster y Cruz, 2004).

#### **4.2.4.2. Construcción del mapa de ligamiento**

Aquellos loci marcadores SSR que no se apartaron significativamente de la segregación esperada 1:2:1, fueron seleccionados para construir una matriz genotípica. Los marcadores que no ajustaron a la segregación no fueron considerados. La matriz resultante “loci SSR x genotipo” se utilizó para construir un mapa de ligamiento con el programa GQMOL (Schuster y Cruz, 2004) con el método de análisis multilocus, usando un LOD mínimo = 3 y una distancia máxima entre marcas adyacentes de 30 cM. Las distancias entre marcadores se estimaron con el método de mínimos cuadrados y se expresaron en unidades de mapa de Kosambi (Kosambi, 1944).

#### **4.2.5. Análisis de mapeo de QTLs**

A la matriz genotípica se le adicionó las medias fenotípicas de cada variable cuantitativa para cada familia  $F_3$ , obteniéndose una matriz mixta de datos genotípicos y fenotípicos. Con dicha matriz se realizó el análisis de mapeo de QTLs.

Los métodos de mapeo que se usaron en el presente capítulo fueron:

- i) **Mapeo por locus simple**
- ii) **Mapeo por intervalo simple**

### iii) Mapeo por intervalo compuesto

i) y ii) se realizaron con el programa QMOL (Schuster y Cruz, 2004). iii) se analizó con el programa WINQTL –Cartographer, versión 2.5\_011 (Basten *et al.*, 2003)

### i) Locus simple

El mapeo de locus simple analiza la asociación entre el marcador y el QTL utilizando ANOVA y regresión lineal. Se aplicó el test *t-Student* para comparar las medias condicionadas a las clases marcadoras, para cada variable estudiada.

#### 4.2.5.1. Locus simple. Método ANOVA de un solo factor

El ANOVA, permitió testear si existen diferencias entre los valores medios del carácter condicionado por el genotipo de la clase marcadora, o sea:

$$H_0: \mu_{MM} = \mu_{Mm} = \mu_{mm}$$

La existencia de diferencias significativas entre las medias es evidencia de ligamiento factorial entre el marcador y la característica. El modelo lineal fue propuesto por Edwards *et al.* (1987).

El cuadrado medio entre clases marcadoras, a su vez, fue particionado ortogonalmente en efecto aditivo (A) y dominante (D), de esta manera se puso a prueba usando contraste para efecto aditivo ( $C_1$ ) y para el efecto dominante ( $C_2$ ) del QTL asociado al marcador (Liu, 1998). Las hipótesis de ambos contrastes fueron:

$$C_1: \mu_{MM} = \mu_{mm}$$

$$C_2: \mu_{MM} - \frac{1}{2} (\mu_{Mm} + \mu_{mm})$$

La tabla de ANOVA y el cálculo de la F's respectivas se muestran en la **Tabla 4.1** (Liu, 1998).



**Tabla 4.1** | Esquema del análisis de la varianza para una población  $F_2$ .

FV	GL	CM	E(CM)	F
Entre clases marcadoras	2	$CM_T$	$\sigma_{gd}^2 + k\sigma_{SSR}^2$	$CM_T/CM_R$
A	1	$CM_A$		$CM_A/CM_R$
D	1	$CM_D$		$CM_D/CM_R$
Dentro de clases marcadoras	$N-3$	$CM_R$	$\sigma_{gd}^2$	
Total	$N-1$			

Donde:

$CM_T$ : Cuadrado medio entre clases marcadoras

$CM_R$ : Cuadrado medio residual

$CM_A$ : Cuadrado medio debido al efecto de aditividad

$CM_D$ : Cuadrado medio debido al efecto de dominancia

$\sigma_{gd}^2$ : Componente de la varianza residual

$\sigma_{SSR}^2$ : Componente de la varianza entre clases genotípicas condicionadas al marcador

Por ser un diseño desbalanceado se calculó el número de observaciones  $k$  como:

$$k = \frac{N - \frac{1}{N} \sum_{i=1}^t n_i^2}{t-1} \text{ donde } N = \sum_{i=1}^t n_i$$

$t$ : número de clases genotípicas para un marcador ( $t=3$ )

$n_i$ : Número de individuos de cada clase genotípica para un marcador

$N$ : Número de individuos totales

A partir de la descomposición de las  $E(CM)$  se estimó la varianza genética

( $\hat{\sigma}_{SSR}^2$ ) como:

$$\hat{\sigma}_{SSR}^2 = \frac{CM_T - CM_R}{k}$$

Se calculó la heredabilidad ( $H^2$ ) del QTL asociado al marcador como,

$$H^2 = \frac{\hat{\sigma}_{SSR}^2}{\hat{\sigma}_T^2}$$

Donde:

$$\hat{\sigma}_T^2 = \frac{CM_T}{k}$$

También se calculó la proporción de varianza genotípica entre familias  $F_3$  ( $\hat{\sigma}_G^2$ , ver 4.2.3) atribuida a la varianza del marcador ( $\hat{\sigma}_{SSR}^2$ ) como  $\frac{\hat{\sigma}_{SSR}^2}{\hat{\sigma}_G^2}$

En aquellos loci marcadores donde el ANOVA fue significativo ( $P < 0.05$ ) se compararon las medias de las tres clases genotípicas con una prueba de *t-Student* ( $P < 0.05$ ). Se analizó el estado de ligamiento entre el marcador y el QTL o sea si se encuentran ligados, en acoplamiento o repulsión (Kearsey y Pooni, 1996).

#### 4.2.5.2. Locus simple. Método Regresión lineal

La asociación entre el marcador y el QTL puede ser valorada a través de una regresión entre la característica cuantitativa en función de los scores relativos al genotipo o gen marcador. El modelo para una  $F_2$  (Schuster y Cruz, 2004) usado fue el siguiente:

$$Y_j = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \varepsilon_j$$

En que:

$Y_j$  es el valor de la característica cuantitativa tomada en el  $j$ -ésimo individuo de la población  $F_2$

$X_{1j}$  es el código del marcador para el efecto aditivo (MM = 1, Mm = 0 y mm = -1);

$X_{2j}$  es el código del marcador para el efecto dominante (MM = 0, Mm = 1 y mm = 0);

$\beta_0$  es el intercepto de la regresión (media de la característica);

$\beta_1$  es la inclinación de la recta para el efecto aditivo;

$\beta_2$  es la inclinación de la recta para el efecto dominante;

$\varepsilon_j$  es el error aleatorio manifestado en la característica del  $j$ -ésimo individuo.

Cabe aclarar que los estimadores de los coeficientes de regresión:  $\beta_1$  y  $\beta_2$ , son el efecto aditivo y dominante para un marcador.

Para evaluar y confirmar el modelo completo se aplicó el test  $F$  (modelos aditivo + dominante). Se analizó la importancia relativa de cada uno de los modelos por medio de una regresión múltiple de eliminación por paso (*stepwise*). También se estimó el coeficiente de determinación  $R^2$  que representa la proporción de la varianza fenotípica de la característica explicada por el marcador. La significación de los parámetros  $\beta_1$  y  $\beta_2$  fueron realizados mediante un test de *t-Student*.

#### **4.2.5.3. Intervalo simple. Método máxima verosimilitud**

Se realizó un mapeo por intervalo simple (SIM) por el método de máxima verosimilitud (Liu, 1998). Haley y Knott (1992), presentaron un modelo de mapeo por intervalos basados en regresión. El modelo adoptado para poblaciones  $F_2$  es:

$$y_j = \mu + ax^* + dz^* + \varepsilon_j$$

donde:  $y_j$  es el valor de la característica  $y$  en el individuo  $j$ ;

$\mu$  es la media de la característica en la población;

$a$  es el efecto aditivo del locus que está siendo estudiado, sobre la característica  $y$ ;

$d$  es el efecto de dominancia del locus que está siendo estudiado, sobre la característica  $y$ ;

$x^*$  y  $z^*$  son las variables indicadoras (utilizadas para estimar los efectos  $a$  y  $d$  del QTL) y dependen del genotipo de los marcadores flanqueantes al QTL, en el individuo  $j$ ;

$\varepsilon_j$  es el error aleatorio para el  $j$ -ésimo individuo.

Este tipo de mapeo permite localizar la posición del QTL en el intervalo determinado por dos marcadores. La búsqueda se hizo cada 2 cM y se calculó el LR o LOD, se estimaron las medias y los efectos de aditividad y dominancia de los QTLs.

La hipótesis de presencia de un QTL en el intervalo es testada por la razón de verosimilitud (LR) entre el modelo reducido y el modelo completo. En los análisis de máxima verosimilitud, el test de razón de máxima verosimilitud es,

$$LR = 2 \ln \left( \frac{L_1}{L_0} \right)$$

En que  $L_0$  es el valor de la función de verosimilitud sobre una hipótesis nula  $H_0$ ;

$L_1$  es el valor de la función de verosimilitud sobre una hipótesis alternativa  $H_1$ .

En el caso de las regresiones, cuando los errores son independientes y siguen una distribución normal, la razón de verosimilitud puede expresarse como una función de los valores de las sumas de cuadrados de los desvíos de los modelos completos y reducidos. Así, según Haley Knott (1992), el test de LR puede ser expresado como:

$$LR = n \ln \left( \frac{SCD_r}{SCD_c} \right)$$

Donde:

$n$  es el número de individuos.

$SCD_r$  es la suma de cuadrados de desvío (o residual) del modelo reducido

$$y_j = \mu + \varepsilon_j.$$

$SCD_c$  es la suma de cuadrados de desvío (o residual) del modelo completo

$$y_j = \mu + ax^* + dz^* + \varepsilon_j$$

Las hipótesis que se testearon fueron las siguientes:

$$H_0: a = 0, d = 0$$

$$H_1: a \neq 0, d = 0$$

$$H_2: a = 0, d \neq 0$$

$$H_3: a \neq 0, d \neq 0$$

Las siguientes pruebas de hipótesis se realizaron con el programa GQMol (Schuster y Cruz, 2004):

- a)  $H_3:H_0$  ( $\mu_{MM} = \mu_{Mm} = \mu_{mm}$  o  $a = 0, d = 0$ ). Su efecto significativo indica presencia de aditividad y/o dominancia, indica presencia del QTL.
- b)  $H_3:H_1$  ( $\mu_{MM} + \mu_{mm} - 2\mu_{Mm}$  o  $d = 0$ ). Test de detección de efecto no aditivo del QTL.
- c)  $H_3:H_2$  ( $\mu_{MM} = \mu_{mm}$  o  $a = 0$ ). Test de detección de efecto aditivo del QTL.

Sí el valor de LR sobrepasa el límite crítico o “threshold” definido (LRT), el QTL en esa posición se considera significativo, cuando el  $LR > LRT$ . Este LRT se estimó para cada carácter permutando 1000 veces los datos fenotípicos y se eligió como punto de corte, una tasa de error experimental tipo I del 5% (Churchill y Doerge, 1994). En todos los caracteres arrojaron un  $LRT \approx 11.41$ , por lo tanto se decidió tomar como punto de corte este valor o su equivalente  $LOD = 2.5$  ( $LOD = LR / 4.61$ ).

#### 4.2.5.4. Mapeo por intervalo compuesto (CIM)

Este tipo de mapeo es derivado del mapeo por intervalo simple e incluye marcadores adicionales como co-factores. La inclusión de estos co-factores en el análisis resulta en un aumento sustancial en el poder de detección de un QTL y en la precisión de la estimación de la posición del QTL, ya que reducen la varianza residual provocadas por otros QTLs fuera del intervalo (Liu, 1998).

El modelo es el siguiente:

$$y_j = [\mu + ax_j + dz_j] + \sum_{k \neq i, i+1}^c b_k x_{kj} + \varepsilon_j$$

$y_j$  es el valor fenotípico de la  $j$ -ésimo individuo;

$\mu$  es el efecto medio del modelo (intercepto del modelo);

$a$  es el efecto aditivo del probable QTL;

$x_j$  y  $z_j$  son variables condicionadoras cuyos valores dependen del genotipo del marcador que flanquean el QTL en el individuo  $j$

$d$  es el efecto de dominancia del probable QTL;

$b_k$  es el coeficiente de regresión parcial entre los valores fenotípicos y los marcadores  $k$ , asumiendo  $c$  marcadores incluido en el modelo de co-factores.

$x_{kj}$  variable condicionadora que depende del genotipo del marcador  $k$  para los  $c$  cofactores seleccionados en el modelo.

$\varepsilon_j$  es el efecto residual del modelo.

El término entre corchetes es el mismo presentado por la regresión de Haley Knott para mapeo por locus simples. Para probar y confirmar la existencia de los QTLs encontrados con locus simple, se empleó siempre que fue posible el mapeo por intervalo compuesto. Para realizar este mapeo se utilizó el programa QTL

Cartographer v2.5 (Basten *et al.*, 2003). Se usó el modelo 6 (modelo estándar) por el método *stepwise* (forward/backward), con probabilidad de entrada y salida del 5%.

El test de razón de verosimilitud (LR) fue realizado cada 1 cM de ligamiento (*walking speed*), con un tamaño de ventana (*windows size*) de 10 cM. Los valores obtenidos para los tests LR en cada posición fueron representados en un gráfico donde los puntos en el eje de las abscisas corresponden a las posiciones en que los tests fueron realizados, en cM, y los puntos en el eje de las ordenadas corresponden a los valores de LR en cada posición, obteniéndose la curva de LR vs posición para cada característica. La curva permitió realizar el mapeo de QTL.

Al igual que en SIM, se consideró significativo un QTL cuando el  $LR > 11.41$  ( $LOD > 2.5$ ) obtenido por el test permutacional (Churchill y Doerge, 1994).

Las siguientes hipótesis fueron testeadas:

i. Significancia del efecto del QTL en el intervalo ( $H_0: a = 0, d = 0$ )

En este caso el modelo completo presenta aditividad y dominancia

ii. Significancia del efecto aditivo del QTL ( $H_0: a = 0$ )

En este caso el modelo reducido a ser utilizado en el test de razón de verosimilitud será:

$$y_j = [\mu + dz_j] + \sum_{k \neq i, i+1}^c b_k x_{kj} + e_j$$

iii. Significancia del efecto atribuido al desvío de dominancia ( $H_0: d = 0$ )

Para testear el modelo con efecto dominante del QTL, el modelo reducido será:

$$y_j = [\mu + ax_j] + \sum_{k \neq i, i+1}^c b_k x_{kj} + e_j$$

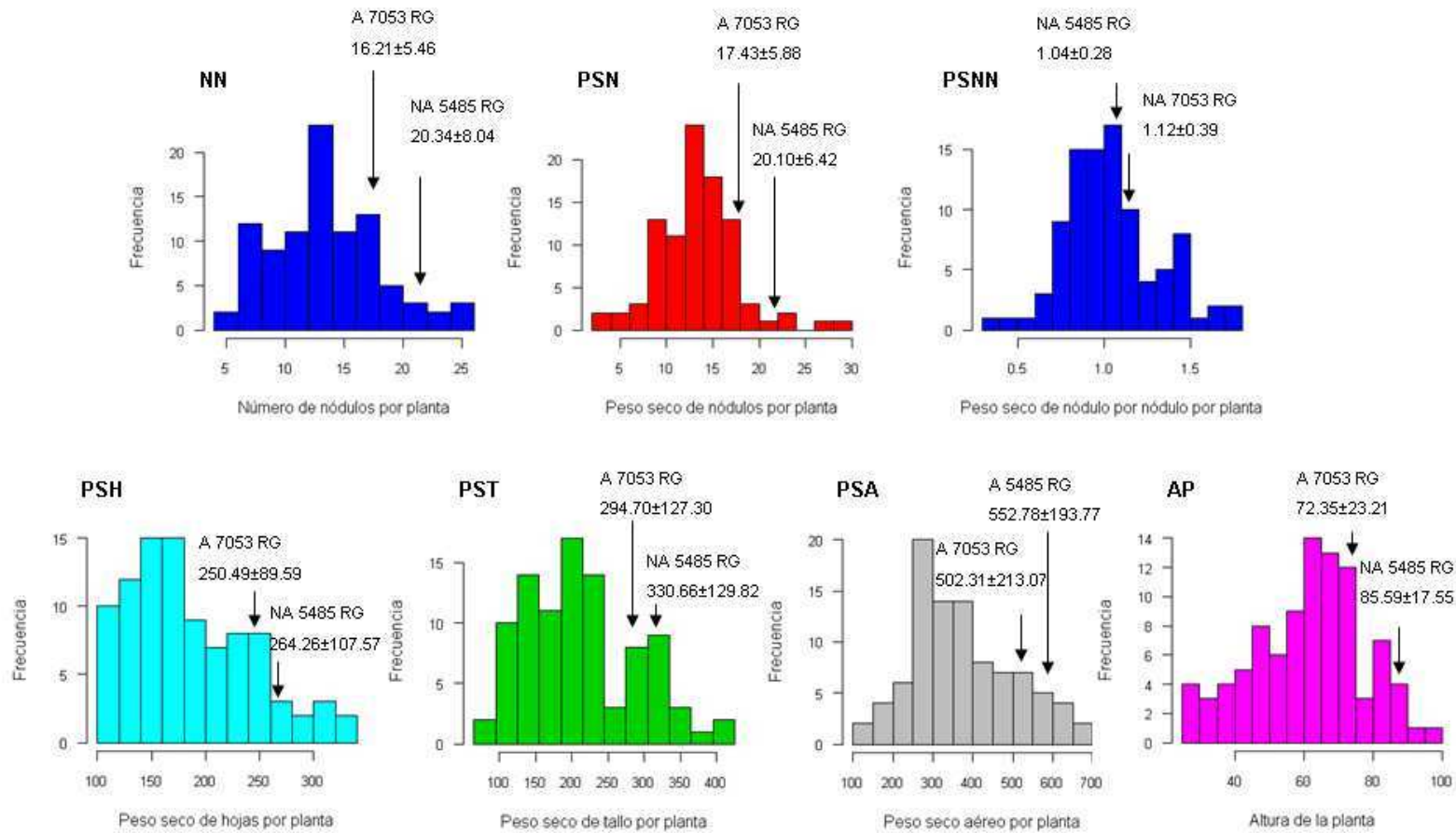
## 4.3. Resultados

### 4.3.1. Análisis fenotípico y genético de la F<sub>3</sub>

La distribución de frecuencias de las características relacionadas con el crecimiento de la planta y la nodulación no presentaron desvíos significativamente diferentes de la normalidad, confirmando que las características fenotípicas tienen distribución continua (**Figura 4.2**). El cultivar parental NA5485 RG presentó valores medios de PST ( $330.66 \pm 129.82 \text{ mg pl}^{-1}$ ), PSH ( $264.26 \pm 107.57 \text{ mg pl}^{-1}$ ), PSA ( $552.78 \pm 193.77 \text{ mg pl}^{-1}$ ), AP ( $85.59 \pm 17.55 \text{ cm pl}^{-1}$ ), NN ( $20.33 \pm 8.05 \text{ nódulos pl}^{-1}$ ), PSN ( $20.10 \pm 6.42 \text{ mg pl}^{-1}$ ); mayores que el cultivar A7053 RG con valores medios de PST ( $294.70 \pm 127.30 \text{ mg pl}^{-1}$ ), PSH ( $250.49 \pm 89.59 \text{ mg pl}^{-1}$ ), PSA ( $502.31 \pm 213.07 \text{ mg pl}^{-1}$ ), AP ( $72.35 \pm 23.21 \text{ cm pl}^{-1}$ ), NN ( $16.21 \pm 5.46 \text{ nódulos pl}^{-1}$ ), PSN ( $17.43 \pm 5.88 \text{ mg pl}^{-1}$ ). En cuanto al PSNN se observó mayor valor en el parental A 7053 RG ( $1.12 \pm 0.39 \text{ mg nódulo}^{-1}$ ) mientras que NA 5485 RG presentó el menor valor ( $1.04 \pm 0.39 \text{ mg nódulo}^{-1}$ ).







**Figura 4.2** | Histogramas de distribución de frecuencias en la población  $F_3$ , para los siete caracteres cuantitativos evaluados. NN: número de nódulos planta<sup>-1</sup>, PSN: peso seco de nódulos/planta (mg planta<sup>-1</sup>), AP: altura de planta (cm), PST: peso seco de tallo (mg planta<sup>-1</sup>), PSA: peso seco aéreo (mg planta<sup>-1</sup>), PSH: peso seco de hoja (mg planta<sup>-1</sup>), PSNN: peso seco de nódulo/nódulo (mg nódulo<sup>-1</sup>). Se indican los valores fenotípicos medios y desvíos estándar ( $\pm$  ds) de la generación parental (NA 5485 RG y A 7053 RG). Medias de la  $F_3$ : NN:13.83, PSN: 13.37, PSNN: 1.03, PSH: 189.07, PST: 220.53, PSA: 384.33, AP: 63.26.

**Tabla 4.2** | Análisis de la varianza de los valores fenotípicos de las 94 familias  $F_3$  provenientes del cruzamiento entre NA 5485 RG y A7053 RG. NN: número de nódulos planta<sup>-1</sup>, PSN: peso seco de nódulos (mg planta<sup>-1</sup>), AP: altura de planta (cm), PST: peso seco de tallo (mg planta<sup>-1</sup>), PSA: peso seco aéreo (mg planta<sup>-1</sup>), PSH: peso seco de hoja (mg planta<sup>-1</sup>), PSNN: peso seco de nódulo/nódulo (mg nódulo<sup>-1</sup>).  $\sigma_G^2$  : Varianza genética,  $\sigma_F^2$  : Varianza fenotípica,  $H^2$ : heredabilidad en sentido amplio.

Carácter	Fte. de Variación	g.l.	CM	$\sigma_G^2$	$\sigma_F^2$	$H^2$
NN	Familia	93	112.283**	13.47	18.26	0.74
	Error	495	29.479			
	Total	588				
PSN	Familia	93	78.6576**	7.8	12.86	0.61
	Error	482	30.9239			
	Total	575				
AP	Familia	93	1629.65**	216.3	250.73	0.86
	Error	518	223.78			
	Total	611				
PST	Familia	93	39654.8**	5038.91	6151.7	0.82
	Error	513	7173.2			
	Total	606				
PSA	Familia	93	120909**	15693.33	19266.74	0.81
	Error	497	22425			
	Total	590				
PSH	Familia	93	19704.8**	2214.03	3081.57	0.72
	Error	508	5547.4			
	Total	601				
PSNN	Familia	93	0.36225**	0.04	0.059	0.67
	Error	481	0.11984			
	Total	574				

g.l: grados de libertad, CM: cuadrado medio.

(\*)  $p < 0.05$ , (\*\*)  $p < 0.01$

### 4.3.2. Análisis de varianza de los caracteres en las familias $F_3$

Se probaron diferencias significativas entre familias  $F_3$  para todas las características (**Tabla 4.2**). Los coeficientes de variación (CV%) para los parámetros de crecimiento PSA, PST, PSH, AP fueron 38.96%, 38.41%, 39.39% y 23.65%, respectivamente; y para los parámetros asociados a nodulación PSN, PSNN y NN fueron 41.58%, 33.76% y 32.27%, respectivamente. Los valores de  $H^2$  estimados a partir del ANOVA, aplicados a los datos fenotípicos de la  $F_3$ , tuvieron valores que

fluctuaron entre 0.61 y 0.81 con un promedio de 0.71. La  $H^2$  de las variables PSA, PST, AP, NN y PSH alcanzaron valores superiores a 0.71, mientras que las variables PSN y PSNN presentaron valores menores (**Tabla 4.2**).

### 4.3.3. Análisis de correlación y coeficientes de paso

Se puede observar que las correlaciones genéticas y fenotípicas fueron similares. Estos resultados no sorprenden ya que los ensayos se realizaron en cámaras de crecimiento bajo condiciones controladas, esto hace que las varianzas ambientales sean bajas a nulas. Para todos los caracteres fenotípicos las correlaciones más altas se observaron entre las variables asociadas a biomasa (**Tabla 4.3**). Se destaca que las tres variables PSH, PST y PSA junto con la AP estuvieron correlacionadas significativamente entre sí y con las variables relacionadas con la capacidad de nodular, PSN y NN. Entre el NN y el PSN se encontró un  $r$  alto ( $P < 0.01$ ). También el NN y PSNN tuvieron correlaciones fenotípicas y genotípicas negativas y significativa ( $P < 0.01$ ) aunque sólo se demostró relación directa significativa al 5% entre PSN con el PSNN para la correlación fenotípica. No se demostró asociación entre PSNN con las variables de biomasa y crecimiento (PSH, PST, PSA y AP) (**Tabla 4.3**).

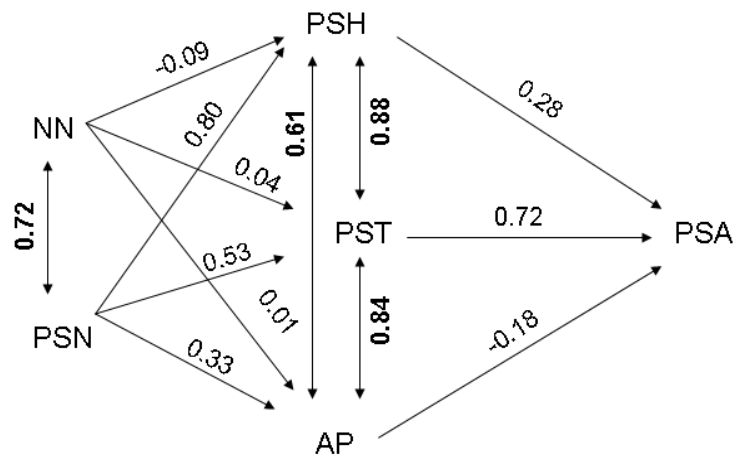
En cuanto a los distintos modelos planteados (**Figura 4.1**) para explicar el problema utilizando las variables de segundo orden, se encontró el mayor coeficiente de determinación ( $R^2$ ) para PSH (0.54) (**Tabla 4.5**). El PSN tuvo un marcado efecto directo sobre el PSH (0.8), mientras que el efecto directo del NN fue bajo (-0.09). Como consecuencia, concluimos que el PSH está influido mayormente por el PSN, mientras que el NN tendría una influencia indirecta a través de su efecto modificador del PSN. Los modelos propuestos para PST y AP resultaron tener un bajo ajuste ( $R^2 = 0.31$  y  $R^2 = 0.12$ , respectivamente) y valores residuales altos (83% y 94%). Esto

indicaría que otras variables no consideradas en el modelo estarían influyendo en la determinación del PST y la AP (Tabla 4.4 y Tabla 4.5).

**Tabla 4.3** | Matriz de correlaciones fenotípica (encima de la diagonal) y genética (debajo de la diagonal) entre todas las variables cuantitativas en 94 familias  $F_3$  de soja proveniente del cruzamiento NA 5485 RG x A 7053 RG. NN: número de nódulos planta<sup>-1</sup>, PSN: peso seco de nódulos (mg planta<sup>-1</sup>), AP: altura de planta (cm), PST: peso seco de tallo (mg planta<sup>-1</sup>), PSA: peso seco aéreo (mg planta<sup>-1</sup>), PSH: peso seco de hoja (mg planta<sup>-1</sup>), PSNN: peso seco de nódulo por nódulo (mg nódulo<sup>-1</sup>).

carácter	NN	PSN	PSNN	AP	PSH	PST	PSA
<b>NN</b>	1	0.72**	-0.49**	0.25**	0.48**	0.42**	0.40**
<b>PSN</b>	0.74**	1	0.17*	0.34**	0.73**	0.56**	0.46**
<b>PSNN</b>	-0.60**	0.03	1	-0.09	0.11	0.02	-0.02
<b>AP</b>	0.22*	0.28**	-0.08	1	0.61**	0.84**	0.59**
<b>PSH</b>	0.45**	0.74**	0.04	0.61**	1	0.88**	0.80**
<b>PST</b>	0.39**	0.51**	-0.05	0.85**	0.88**	1	0.81**
<b>PSA</b>	0.37**	0.37**	-0.12	0.57**	0.77**	0.77**	1

(\*)  $p < 0.05$ , (\*\*)  $p < 0.01$



**Figura 4.3** | Diagrama de coeficientes de paso describiendo la relación entre el peso seco aéreo (PSA), número de nódulos (NN), peso seco de nódulos (PSN), peso seco de hojas (PSH), peso seco de tallo (PST) y altura de planta (AP) de 94 familias  $F_3$  de soja. Los coeficientes de paso se muestran en flechas de una dirección, y los coeficientes de correlación se muestran en flechas con ambas direcciones.

**Tabla 4.4** | Coeficientes de paso que muestran los efectos directos e indirectos para la variable dependientes PSA.

Tipo de efecto	Valores estimados
<b>Efecto del PSH</b>	
Efecto directo	0.28
Efecto indirecto vía PST	0.63
Efecto indirecto vía AP	-0.11
Correlación total	0.80
<b>Efecto del PST</b>	
Efecto directo	0.72
Efecto indirecto vía PSH	0.25
Efecto indirecto vía AP	-0.15
Correlación total	0.82
<b>Efecto del AP</b>	
Efecto directo	-0.18
Efecto indirecto vía PSH	0.17
Efecto indirecto vía PST	0.60
Correlación total	0.59
R <sup>2</sup>	0.70
Residual	0.55

**Tabla 4.5** | Coeficientes de paso que muestran los efectos directos e indirectos para la variable dependientes PSH (izquierda), PST (centro) y AP (derecha).

Tipo de efecto	valores estimados PSH	valores estimados PST	valores estimados AP
<b>Efecto del NN</b>			
Efecto directo	-0.09	0.04	0.01
Efecto indirecto vía PSN	0.57	0.38	0.24
Correlación total	0.48	0.42	0.25
<b>Efecto del PSN</b>			
Efecto directo	0.8	0.53	0.33
Efecto indirecto vía NN	-0.07	0.03	0.007
Correlación total	0.73	0.56	0.34
R <sup>2</sup>	0.54	0.31	0.12
Residual	0.68	0.83	0.94

#### 4.3.4. Análisis molecular

##### 4.3.4.1. Selección y segregación de marcadores

Se evaluaron un total de 221 marcadores SSR, de los cuales un 62.9% no fueron polimórficos, un 20.8% no generaron amplicones y sólo un 16.3% (36 marcadores) del total amplificado fueron polimórficos para los progenitores y las plantas F<sub>2</sub>. Luego de aplicar los métodos estadísticos para comprobar la segregación esperada, finalmente se optó por el criterio FDR por ser la metodología más conservadora entre las pruebas que controlan el error global de tipo I (Schuster y Cruz, 2004). Al utilizar este método, se seleccionaron 35 marcadores que ajustaron a la segregación esperada para marcadores codominantes (1:2:1) y sólo un marcador ajustó a una segregación dominante 3:1 (**Tabla 4.6**).

##### 4.3.4.2. Construcción de Mapa de ligamiento

Sólo un tercio de los marcadores polimórficos estuvieron ligados, formando 5 grupos de ligamiento, 3 contruidos con dos marcas y 2 con tres marcas. No se detectó ligamiento entre los restantes marcadores. Las distancias de ligamiento en unidades de Kosambi fueron desde 9.2 a 17.2 cM. Los 5 grupos de ligamientos formados se correspondieron parcialmente con 5 cromosomas, del total de 20 que posee el genoma de soja (Cregan *et al.*, 1999). El grupo de ligamiento 1 (GL-1) constituido con los marcadores Sct\_065 y Satt414 fue coincidente con el cromosoma J. El Grupo de ligamiento 2 (GL-2) constituido por tres marcadores Satt646, Satt294 y Satt195 coincidió con el cromosoma C1. El grupo de ligamiento 3 (GL-3) asoció los marcadores Satt245 y Satt435 y coincidió con el cromosoma M. El grupo de ligamiento 4 (GL-4) formado por el Satt005 y Sat141, fue coincidente con el cromosoma D1b y el grupo de ligamiento 5 (GL-5) formado por 3 marcadores: Satt152, Satt584 y Satt009 fue coincidente con el cromosoma N aunque, en este último, el orden de los

marcadores fue inconsistente con el orden citado en el mapa de soja (Cregan et. al., 1999) (**Figura 4.4**).



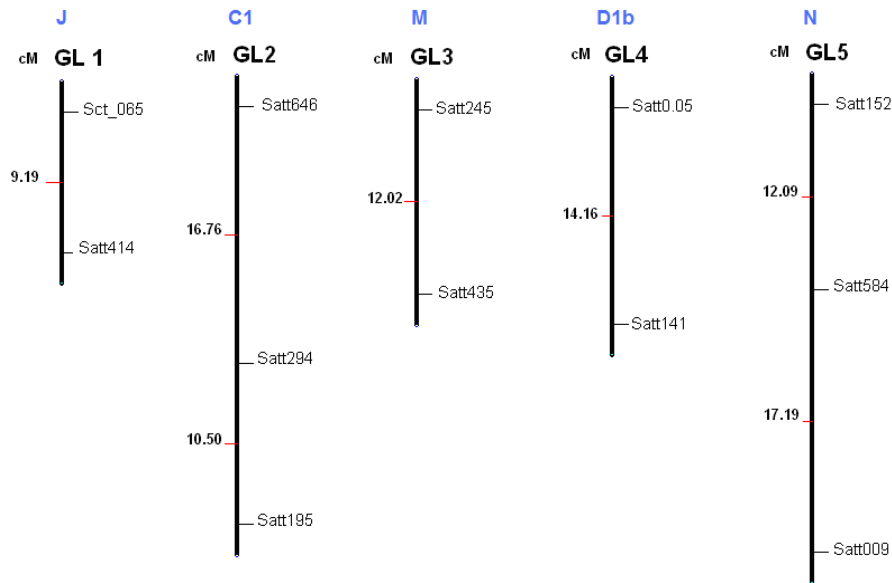
**Tabla 4.6** | Segregación de los 36 marcadores microsatélites según tres pruebas: Chi-cuadrado, criterio individual, Bonferroni y FDR (False Discovery Rate).

<i>Marcador</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>9</i>	<i>Hipótesis</i>	<i>Chi-Cuadrado</i> <sup>1</sup>	<i>Chi-Cuadrado</i> <sup>2</sup>	<i>Chi-Cuadrado</i> <sup>3</sup>
Satt197	20	46	24			4	1:2:1	0.40	0.40	0.40
Satt560	19	33	4			38	1:2:1	9.82*	9.82	9.82
Satt434	19	47	23			5	1:2:1	0.64	0.64	0.64
Satt245	16	39	29			10	1:2:1	4.45	4.45	4.45
Satt152	28	35	25			6	1:2:1	3.89	3.89	3.89
Satt005	17	34	20			23	1:2:1	0.38	0.38	0.38
Sat_114	19	44	17			14	1:2:1	0.90	0.90	0.90
Sct_065	17	48	18			11	1:2:1	2.06	2.06	2.06
Satt187	24	44	19			7	1:2:1	0.59	0.59	0.59
Satt216	33	43	14			4	1:2:1	8.20*	8.20	8.20
Satt646	22	48	22			2	1:2:1	0.17	0.17	0.17
Satt294	12	56	25			1	1:2:1	7.52*	7.52	7.52
Satt164	20	52	21			1	1:2:1	1.32	1.32	1.32
Satt195	19	39	35			1	1:2:1	7.93*	7.93	7.93
Satt458	20	42	29			3	1:2:1	2.32	2.32	2.32
Satt309	19	40	20			15	1:2:1	0.04	0.04	0.04
Satt311	14	40	25			15	1:2:1	3.08	3.08	3.08
Satt353	23	39	30			2	1:2:1	3.20	3.20	3.20
Satt296	23	46	23			2	1:2:1	0.00	0.00	0.00
Satt271	22	45	24			3	1:2:1	0.10	0.10	0.10
Sat_232	21	40	31			2	1:2:1	3.74	3.74	3.74
Satt431	32	38	23			1	1:2:1	4.85	4.85	4.85
Satt174	23	46	23			2	1:2:1	0.00	0.00	0.00
Sat_294	20	58	15			1	1:2:1	6.23*	6.23	6.23
Satt168	24	51	19			-	1:2:1	1.21	1.21	1.21
BE021153	23	47	24			-	1:2:1	0.02	0.02	0.02
Satt141	18	52	24			-	1:2:1	1.83	1.83	1.83
Satt703	19	45	28			2	1:2:1	1.80	1.80	1.80
Satt414	22	51	20			1	1:2:1	0.96	0.96	0.96
Satt435	22	41	28			3	1:2:1	1.68	1.68	1.68
Satt009	28	37	22			7	1:2:1	2.77	2.77	2.77
Satt584	29	44	21			-	1:2:1	1.75	1.75	1.75
Satt687	20	53	21			-	1:2:1	1.55	1.55	1.55
Satt462	18	10	48			18	1:2:1	64.95*	64.947*	64.95
Sat_283	16	10	58			10	1:2:1	90.76*	90.762*	90.76*
Sat_244			28		57	9	3:1	2.86	2.86	2.86

<sup>1</sup> criterio individual al 5%<sup>2</sup> Bonferroni Nivel crítico: 0.14%<sup>3</sup> FDR, nivel crítico 0.0 %

(\*) p&lt;0.05

**Figura 4.4** | Mapa genético producido a partir de los marcadores microsatélites de una población  $F_2$  proveniente del cruzamiento entre NA 5485 RG x A 7053 RG. Se indican los marcadores microsatélites y las distancias en unidades de Kosambi.



### 4.3.5. Análisis de QTLs

#### 4.3.5.1. Mapeo de locus simple. Método ANOVA

El análisis de las diferencias entre familias  $F_3$  dentro de las clases marcadoras, mediante el análisis de la varianza, demostró la existencia de diferencias significativas entre medias de los genotipos pertenecientes a cada clase marcadora. En la **Tabla 4.7** se muestran los resultados del análisis de la varianza que confirmó la existencia de 4 QTLs de NN ligados a los marcadores Satt353, Satt271, Sat\_294 y Satt414. 3 QTLs de PSN ligados a Satt560, Satt414 y Satt434. 5 QTLs de AP ligados a los marcadores Satt434, Satt245, Satt216, Satt168 y Satt435. 2 QTLs de PST ligados a Satt434 y Satt216. 5 QTLs de PSA ligados a los marcadores Satt152, Satt271, BE021153, Satt009 y Sat462. 3 QTLs de PSH ligados a los marcadores Satt005, Satt216 y

Satt309. 4 QTLs de PSNN ligados a los marcadores Satt197, Satt560, Satt195 y Satt353.

Este modelo además permitió estimar las varianzas genéticas explicadas por cada QTL. Debido a que cada clase marcadora estuvo constituida por diferente número de familias, se estimó el tamaño medio de la muestra, “ $k$ ” (calculada como media armónica) y se estimó la heredabilidad de cada QTL. La razón entre la varianza genética estimada para cada QTL condicionado a cada marcador y la varianza genética del carácter es una medida de la importancia y la magnitud de la variación que explica cada QTL. Algunos microsatélites explicaron más del 25% de la variación genética total del carácter. Por ejemplo Satt353, explicó el 25.8% de la variación genética total del NN, El Satt560, explicó el 45.4% de la variabilidad del PSN y el 32.5% en el caso del PSNN. Los marcadores Satt245 y Satt435 explicaron el 42% y 35.1%, respectivamente de la variación para la AP.

En las **Figura 4.5** a **Figura 4.11** se muestran las medias de los caracteres cuantitativos condicionadas a las 3 clases marcadoras y su comparación utilizando un test *t-Student*. La observación de las comparaciones permite entender la fase de ligamiento y el modo de herencia de cada QTL ligado a las clases marcadoras.

Para el NN, existió un efecto aditivo significativo, los QTLs asociados fueron Satt353 y Satt414, y presentaron desvíos de dominancia significativos, los marcadores Satt271 y Sat\_294. El QTL para NN podría encontrarse ligado en repulsión con el marcador Satt353 (**Figura 4.5**).

El QTL para el PSN asociado a Satt414 tuvo un efecto aditivo significativo, mientras que Satt560 y Satt434 tuvieron un efecto dominante significativo. Se puede observar una sobredominancia a favor del progenitor con mayor peso seco de nódulo. El marcador Satt434 se encuentra en fase de repulsión con el QTL para PSN (**Figura 4.6**).

Tres QTLs de AP tuvieron efecto aditivo significativo, Satt245, Satt216 y Satt435 y dos, Satt434 y Satt168, efecto dominante significativo. Como se observa en

**Figura 4.7**, estos dos últimos marcadores presentaron sobredominancia. El QTL de AP estaría ligado en repulsión al marcador Satt434.

El QTL de PSH asociado a Satt216 tuvo un efecto aditivo significativo, mientras que los QTLs asociados a Satt005 y Satt309 solo presentaron efecto dominante significativo. Únicamente se encontró fase de repulsión entre el QTL de PSH con el Satt309 (**Figura 4.9**).

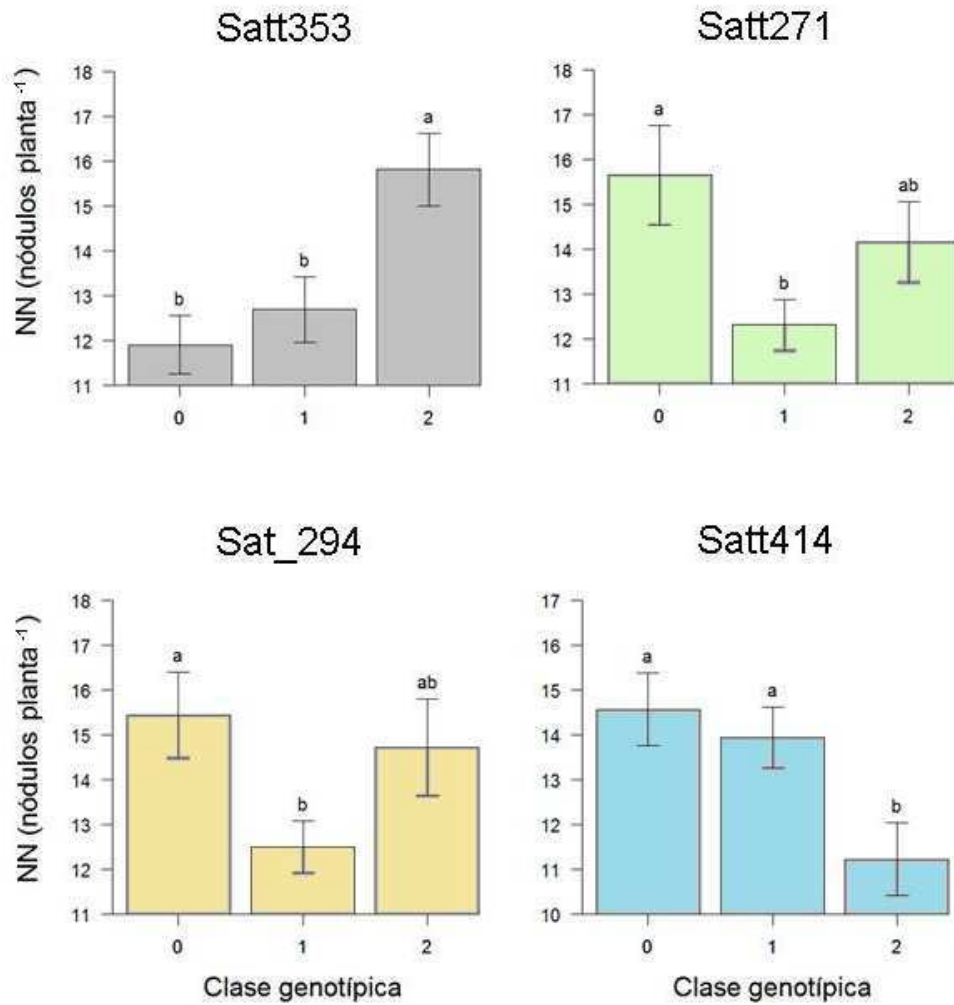
Los QTLs de PSA asociados a los marcadores Satt152, Satt271 y Satt009 presentaron efectos aditivo significativos, mientras que el Satt462 tuvo un efecto dominante significativo. Únicamente el QTL asociado al marcador BE021153 presentó ambos efectos, aditivo y dominante significativos. El locus Satt462 tuvo un grado de dominancia alto, indicando sobredominancia. Los marcadores Satt152, Satt009 y Satt462 indicarían ligamiento en repulsión con el QTL respectivo (**Figura 4.10**).

Para PSNN los QTLs asociados con los marcadores Satt560 y Satt353 presentaron el efecto aditivo significativo y Satt 197 y Satt195 efecto dominante significativo (**Figura 4.11**).

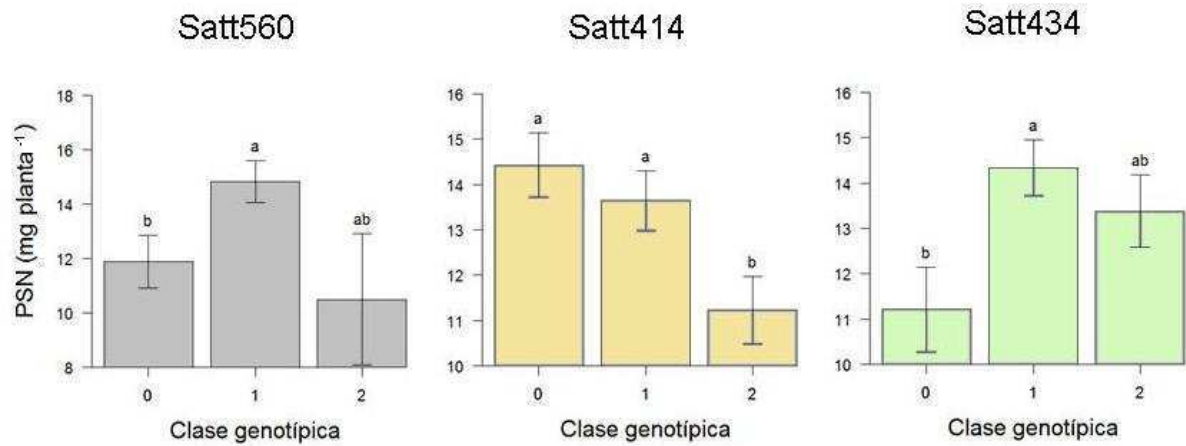
**Tabla 4.7** | Marcadores microsatélites asociados a los caracteres cuantitativos por medio del análisis de la varianza. Se indican los valores fenotípicos medios de cada carácter condicionados cada marcador. NN: número de nódulos/planta, PSN: peso seco de nódulos planta<sup>-1</sup> (mg), AP: altura de planta (cm), PST: peso seco de tallo (mg), PSA: peso seco aéreo (mg), PSH: peso seco de hoja (mg), PSNN: peso seco de nódulo nódulo<sup>-1</sup> (mg). Se estimó la varianza genética entre clases genotípicas ( $\sigma_{(SSR)}^2$ ), la heredabilidad del QTL ( $H^2$ ), proporción de la varianza genética entre familias  $F_3$  explicada por el marcador ( $\sigma_{(SSR)}^2/\sigma_G^2$ ). GL: grupo de ligamiento,  $a$  = efecto aditivo,  $d$  = efecto de dominancia.

	Marcador	GL	F	Media(0) A1A1	Media(1) A1A2	Media(2) A2A2	$\sigma_{(SSR)}^2$	$H^2$	$\frac{\sigma_{(SSR)}^2}{\sigma_G^2}$	Efecto
NN	Satt353	H	6.89**	11.91	12.7	15.81	3.47	0.85	25.76	a***
	Satt271	D1b	4.67*	15.65	12.31	14.16	2.41	0.76	17.89	d**
	Sat_294	A2	4.06*	15.44	12.5	14.72	2.33	0.75	17.3	d**
	Satt414	J	3.57*	14.56	13.95	11.23	1.79	0.72	13.28	a*
PSN	Satt560	B2	3.80*	11.89	14.83	10.5	3.54	0.74	45.38	d*
	Satt414	J	3.48*	14.43	13.65	11.24	1.53	0.71	19.62	a*
	Satt434	H	3.97*	11.21	14.34	13.38	1.82	0.75	23.33	d*
AP	Satt434	H	3.45*	54.44	65.33	59.09	23.17	0.71	9.24	d*
	Satt245	M	14.10**	71.3	65.6	50.06	105.25	0.93	41.98	a***
	Satt216	D1b	4.27*	66.77	58.72	53.2	30.1	0.77	12	a**
	Satt168	B2	3.47*	67.44	57.43	63.37	22.05	0.71	8.79	d*
	Satt435	M	13.99**	70.12	62.83	49.65	88.04	0.93	35.11	a***
PST	Satt434	H	4.45*	172.15	231.95	205.42	713.54	0.78	14.16	d**
	Satt216	D1b	6.10**	242.63	200.48	166.78	986.74	0.84	19.58	a***
PSA	Satt152	N	3.66*	321.91	392.36	411.94	1545.8 3	0.73	9.85	a*
	Satt271	D1b	3.73*	439.99	358.65	345.75	1607.8 4	0.73	10.25	a*
	BE021153	D1b	5.29**	385.68	409.47	305.5	2402.3 2	0.81	15.31	a* y d*
	Satt009	N	3.60*	323.48	386.15	422.36	1634.0 7	0.72	10.41	a**
	Satt462	L	4.32*	383.52	275.2	406.13	2716.5	0.77	17.31	d**
	PSH	Satt005	D1b	3.28*	173.22	193.6	156.34	278.69	0.7	12.59
Satt216		D1b	3.62*	201.86	179.85	157.57	275.89	0.72	12.46	a**
Satt309		G	3.42*	159.98	201.14	190.64	316.09	0.71	14.28	d*
PSNN	Satt197	B1	4.35*	0.92	1.12	1	0.01	0.77	21.25	d**
	Satt560	B2	3.69*	0.92	1.13	1.14	0.01	0.73	32.5	a*
	Satt195	C1	3.19*	0.99	1.13	0.98	0.01	0.69	13	d*
	Satt353	H	4.34*	1.1	1.1	0.93	0.01	0.77	19.5	a*

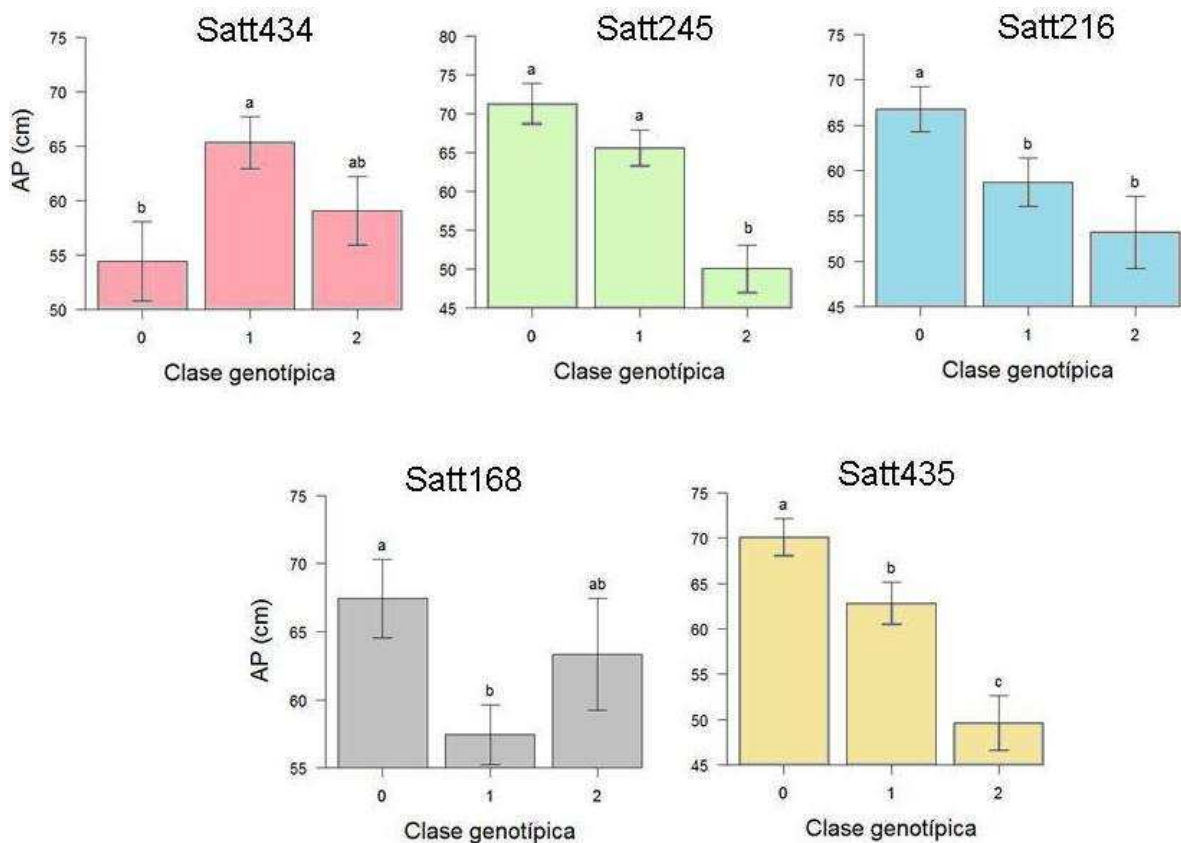
(\*)  $p < 0.5$ , (\*\*)  $p < 0.01$ , (\*\*\*)  $p < 0.001$



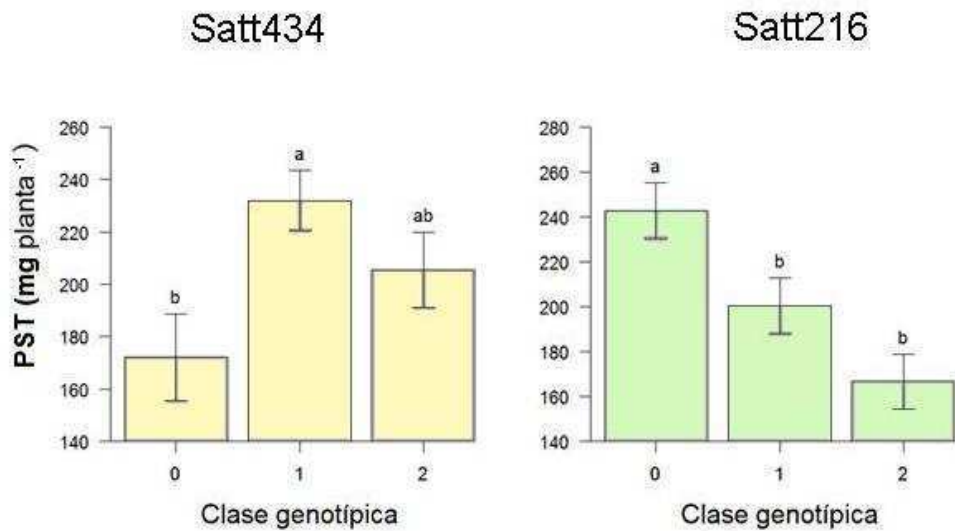
**Figura 4.5** | Valores fenotípicos del NN medios condicionados a cada clase genotípica de cada marcador SSR en la población F<sub>2</sub>. 0: homocigota para el padre mayor NN (NA 5485 RG), 2: homocigota para el padre de menor NN (A 7053 RG), 1: heterocigota (comparte alelos de ambos padres para el marcador). Mismas letras indican diferencias no significativas al 5% por el test *t-Student* entre pares de medias.



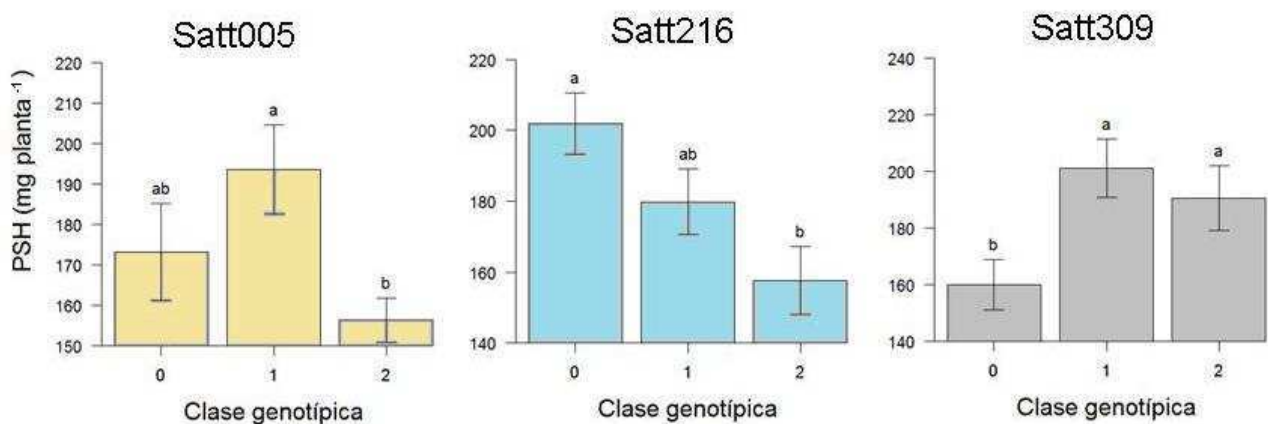
**Figura 4.6** | Valores fenotípicos del PSN medios condicionados a cada clase genotípica de cada marcador SSR en la población  $F_2$ . 0: homocigota para el padre mayor PSN (NA 5485 RG), 2: homocigota para el padre de menor PSN (A 7053 RG), 1: heterocigota (comparte alelos de ambos padres para el marcador). Mismas letras indican diferencias no significativas al 5% por el test *t-Student* entre pares de medias.



**Figura 4.7** | Valores fenotípicos de la AP medios condicionados a cada clase genotípica de cada marcador SSR en la población  $F_2$ . 0: homocigota para el padre mayor AP (NA 5485 RG), 2: homocigota para el padre de menor AP (A 7053 RG), 1: heterocigota (comparte alelos de ambos padres para el marcador). Mismas letras indican diferencias no significativas al 5% por el test *t-Student* entre pares de medias.

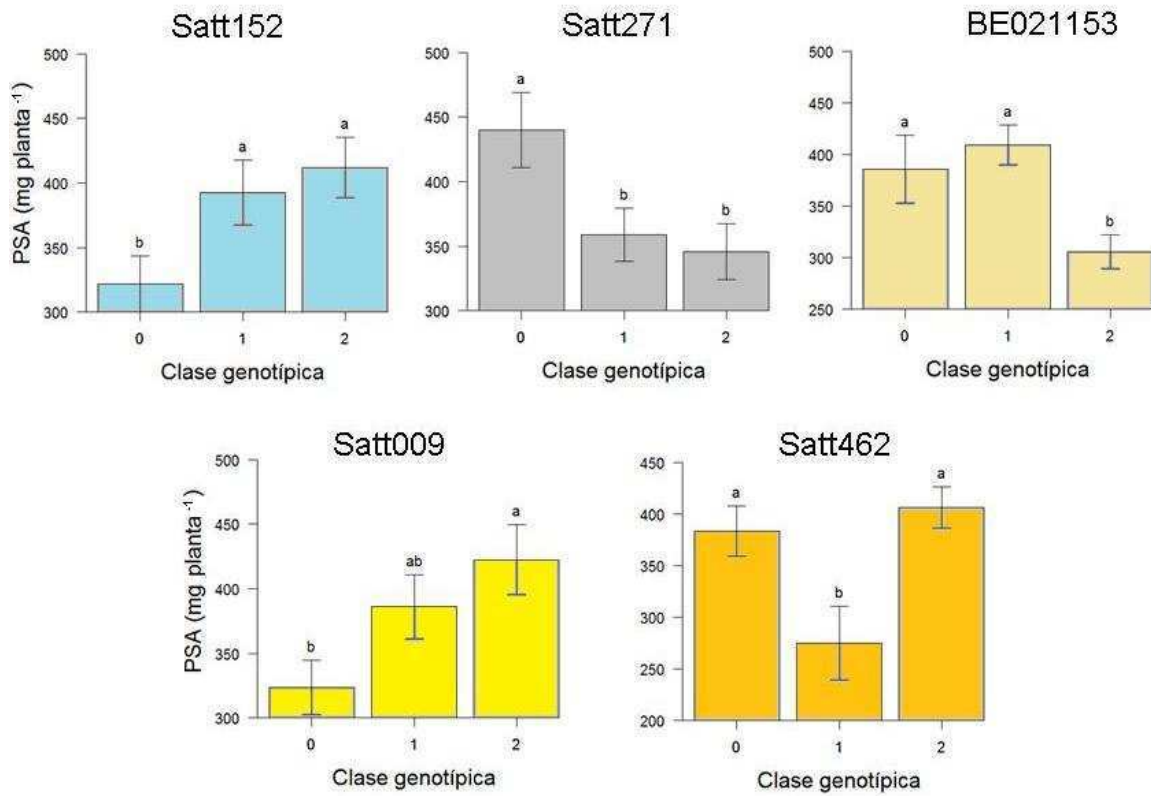


**Figura 4.8** | Valores fenotípicos del PST medios condicionados a cada clase genotípica de cada marcador SSR en la población F<sub>2</sub>. 0: homocigota para el padre mayor PST (NA 5485 RG), 2: homocigota para el padre de menor PST (A 7053 RG), 1: heterocigota (comparte alelos de ambos padres para el marcador). Mismas letras indican diferencias no significativas al 5% por el test *t-Student* entre pares de medias.

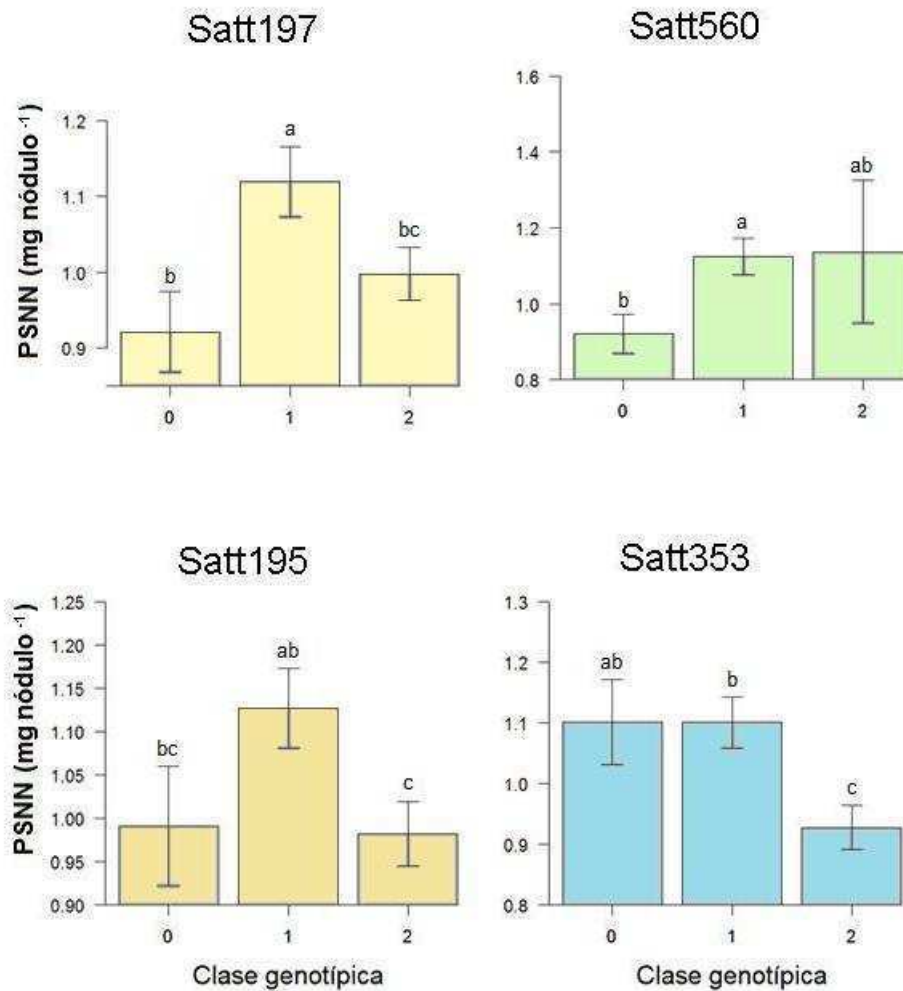


**Figura 4.9** | Valores fenotípicos del PSH medios condicionados a cada clase genotípica de cada marcador SSR en la población F<sub>2</sub>. 0: homocigota para el padre mayor PSH (NA 5485 RG), 2: homocigota para el padre de menor PSH (A 7053 RG), 1: heterocigota (comparte alelos de ambos padres para el marcador). Mismas letras indican diferencias no significativas al 5% por el test *t-Student* entre pares de medias.





**Figura 4.10** | Valores fenotípicos del PSA medios condicionados a cada clase genotípica de cada marcador SSR en la población  $F_2$ . 0: homocigota para el padre mayor PSA (NA 5485 RG), 2: homocigota para el padre de menor PSA (A 7053 RG), 1: heterocigota (comparte alelos de ambos padres para el marcador). Mismas letras indican diferencias no significativas al 5% por el test *t-Student* entre pares de medias.



**Figura 4.11** | Valores fenotípicos del PSNN medios condicionados a cada clase genotípica de cada marcador SSR en la población  $F_2$ . 0: homocigota para el padre menor PSNN (NA 5485 RG), 2: homocigota para el padre de mayor PSNN (A 7053 RG), 1: heterocigota (comparte alelos de ambos padres para el marcador). Mismas letras indican diferencias no significativas al 5% por el *test t-Student* entre pares de medias.

#### 4.3.5.2. Mapeo de locus simple. Análisis de Regresión

Los resultados del análisis de regresión aplicado al análisis de locus simple fueron coincidentes con los obtenidos por el análisis de la varianza presentado anteriormente, y confirmó el ligamiento entre los QTLs y los marcadores. La **Tabla 4.8**

muestra los valores  $\beta_0$ ,  $\beta_1$  y  $\beta_2$  y el cálculo del coeficiente de determinación  $R^2$  para cada QTL confirmado.

Los efectos aditivos ( $\beta_1$ ) fueron significativos para: los marcadores Satt353 y Satt414 para NN; por el marcador Satt414 para PSN; por los marcadores Satt245, Satt216 y Satt435 para AP; por Satt216 para PST; por Satt152, Satt271, BE021153 y Satt009 para PSA; por Satt216 para PSH y por Satt560 y Satt353 para PSNN.

Efectos de dominancia ( $\beta_2$ ) significativos y negativos fueron encontrados para la variable NN cuando se cuantificó con el marcador Satt271 y Sat\_294; para la variable AP cuantificada por el marcador Satt168; PSA cuantificada por el marcador Satt462. Se encontraron efectos de dominancia ( $\beta_2$ ) significativos y positivos para la variables PSN cuantificada por el marcador Satt560 y Satt434; para la variable AP cuantificado para el marcador Satt434; para la variable PSA cuantificada por el marcador BE021153; Para la variable PSH cuantificada para el marcador Satt005 y Satt309; para la variables PSNN cuantificada por el Satt197 y Satt195.

Solo el carácter PSA, el locus BE021153 presentó ambos efectos ( $\beta_1$  y  $\beta_2$ ) significativos (igual que en ANOVA) y con magnitudes semejantes. Comparativamente la prueba de *t-Student* no detectó diferencias significativas al comparar el heterocigota con el progenitor superior, por consiguiente, se podría considerar una herencia de dominancia completa (**Figura 4.10**).

El  $R^2$  es una medida de la magnitud de la varianza fenotípica del carácter que explica el marcador. El Satt245 explicó el 26% y el Satt435 el 24%, por lo cual en conjunto explicaron el 50 % de la variación fenotípica de AP. Satt353 explicó el 13.4% de la variación del NN; el Satt414 el 12.6% del PSN; Satt216 el 12.3% del PST; Satt560 el 12.2% del PSNN; Satt462 y BE021153 el 10.6% y 10.4, respectivamente del PSA y Satt005 y Satt309 el 8.8 y 8.3, respectivamente del PSH.

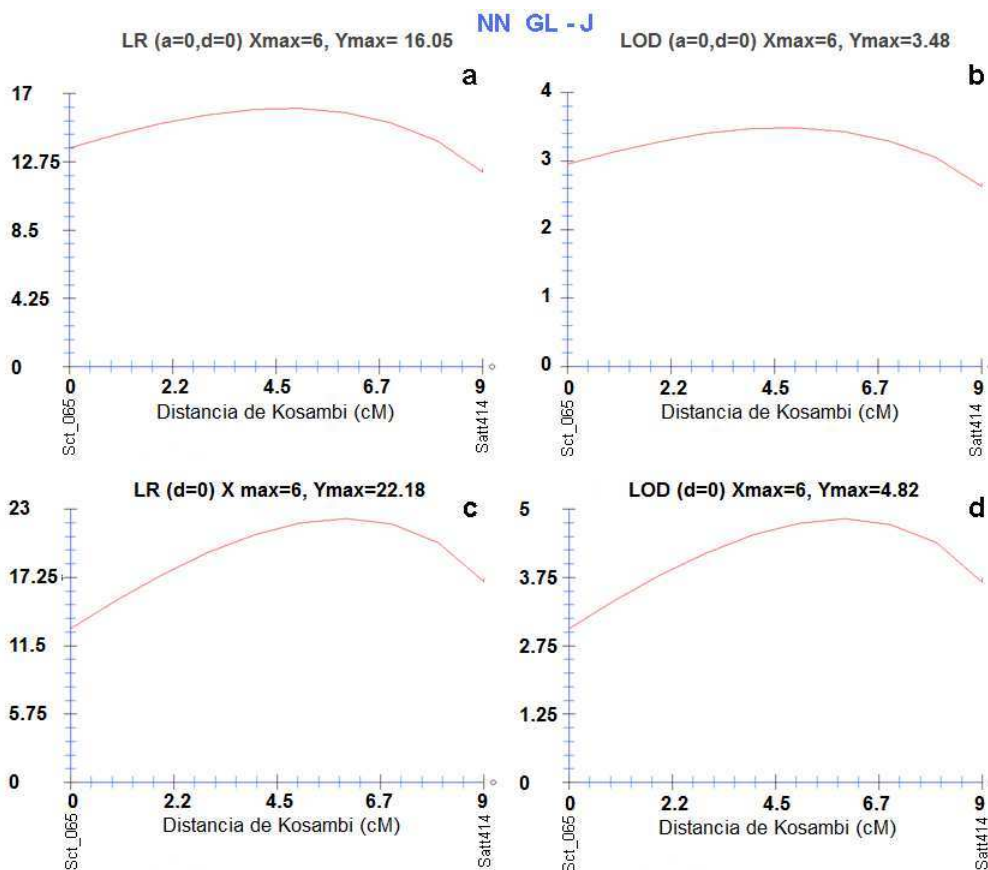
**Tabla 4.8** | Marcadores microsatélites asociados significativamente a los caracteres cuantitativos por medio del análisis de regresión. NN: número de nódulos planta<sup>-1</sup>, PSN: peso seco de nódulos planta<sup>-1</sup> (mg), AP: altura de planta (cm), PST: peso seco de tallo (mg), PSA: peso seco aéreo (mg), PSH: peso seco de hoja (mg), PSNN: peso seco de nódulo nódulo<sup>-1</sup> (mg). Se muestran la media general de los caracteres estimada como  $\beta_0$ , los valores de las pendientes  $\beta_1$  (efecto aditivo) y  $\beta_2$  (efecto dominante) y sus significaciones mediante la prueba t. GL: grupo de ligamiento.

Carácter	Marca	GL	Prob (F)	$\beta_0$	$\beta_1$	$\beta_2$	R <sup>2</sup> (%)
NN	Satt353	H	0.002**	13.859	1.95**	-1.16	13.4
	Satt271	D1b	0.012*	14.907	1.95	-2.59**	9.6
	Sat_294	A2	0.021*	15.079	1.95	-2.58**	8.27
	Satt414	J	0.032*	12.897	1.95*	1.05	7.36
PSN	Satt560	B2	.029*	11.198	-0.70	3.63*	12.55
	Satt414	J	.035*	12.832	-1.60*	0.81	7.19
	Satt434	H	.022*	12.298	1.09	2.04*	8.46
AP	Satt434	H	.036*	56.766	2.32	8.56*	7.43
	Satt245	M	.0**	60.681	-10.62**	4.92	25.82
	Satt216	D1b	.017*	59.982	-6.78**	-1.26	8.94
	Satt168	B2	.035*	65.402	-2.03	-7.97*	7.08
	Satt435	M	.0**	59.882	-10.24**	2.95	24.12
PST	Satt434	H	.015*	188.782	16.63	43.17**	9.38
	Satt216	D1b	.003**	204.703	-37.92**	-4.22	12.29
PSA	Satt152	N	.03*	366.923	45.02*	25.44	7.92
	Satt271	D1b	.028*	392.869	-47.12*	-34.21	7.81
	BE021153	D1b	.007**	345.591	-40.09*	63.88*	10.42
	Satt009	N	.032*	372.917	49.44*	13.24	7.9
	Satt462	L	.017*	394.825	11.31	-119.62**	10.58
PSH	Satt005	D1b	.044*	164.782	-8.44	28.82*	8.8
	Satt216	D1b	.031*	179.713	-22.15*	0.14	7.69
	Satt309	G	.038*	175.306	15.33	25.83*	8.26
PSNN	Satt197	B1	.016*	0.96	0.04	0.16**	9.08
	Satt560	B2	.032*	1.029	0.11*	0.10	12.21
	Satt195	C1	.046*	0.987	-0.004	0.14*	6.62
	Satt353	H	.016*	1.015	-0.09*	0.09	8.9

(\*) p<0.5, (\*\*) p<0.01

### 4.3.5.3. Mapeo por intervalo simple. Análisis de regresión por máxima verosimilitud

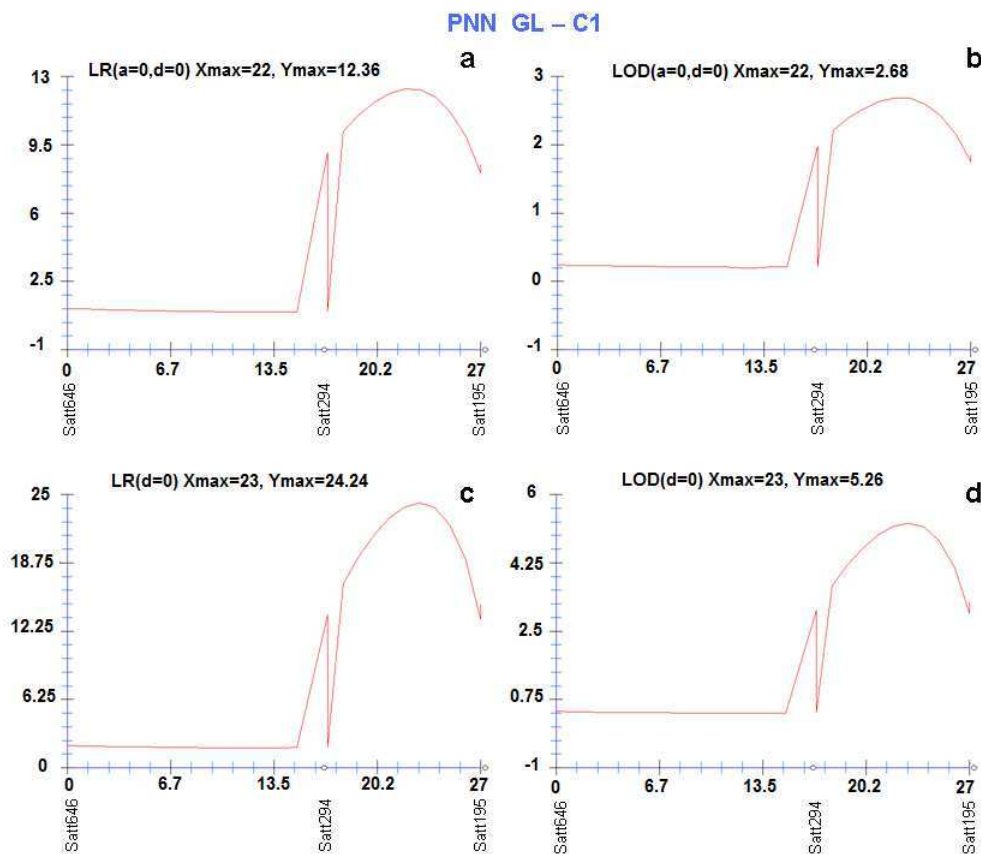
Este análisis de mapeo reveló asociación entre los 4 caracteres NN, PSNN, AP y PSH con 5 regiones genómicas. Se incluyó, entre los marcadores Sct\_065 y Satt414 en el GL J, un intervalo que tuvo un efecto significativo en la expresión fenotípica del NN (LOD = 3.43) a 3.2 cM del marcador Satt414, con un efecto dominante significativo (LOD = 4.82) (Figura 4.12).



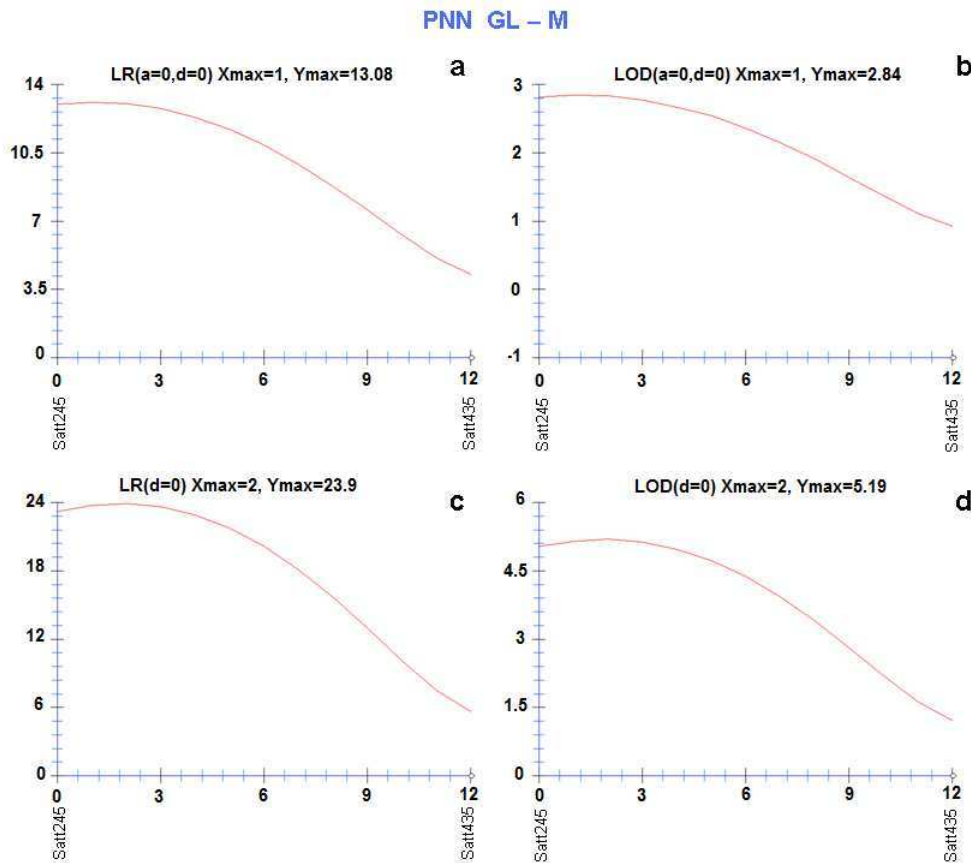
**Figura 4.12** | Mapeo por intervalo aplicando el método máxima verosimilitud del QTL número de nódulos por planta (NN) en el grupo de ligamiento J. a) y b) prueba de posición del QTL (LR y LOD, respectivamente vs distancia de Kosambi); c) y d) pruebas de efecto de dominancia del QTL (LR y LOD, respectivamente vs distancia de Kosambi).

La expresión del PSNN se mapeo en 2 regiones, una ubicada en GL-C1 dentro del intervalo formado por los marcadores Satt195 y Satt294, a 4.5 cM de Satt195 (LOD

= 2.7) y de efecto dominante (LOD = 5.27); la otra región se encuentra en el GL-M dentro del intervalo formado por los marcadores Satt245 y Satt435, a 1 cM de Satt245 (LOD = 2.84) y también con efecto dominante (LOD = 5.15) (**Figura 4.13** y **Figura 4.14**). Esta última asociación indica la presencia de un nuevo QTL en el GL-M, que no había sido confirmado previamente con la metodología de locus simple, por ninguno de los dos métodos estadísticos utilizados (ni ANOVA, ni regresión).



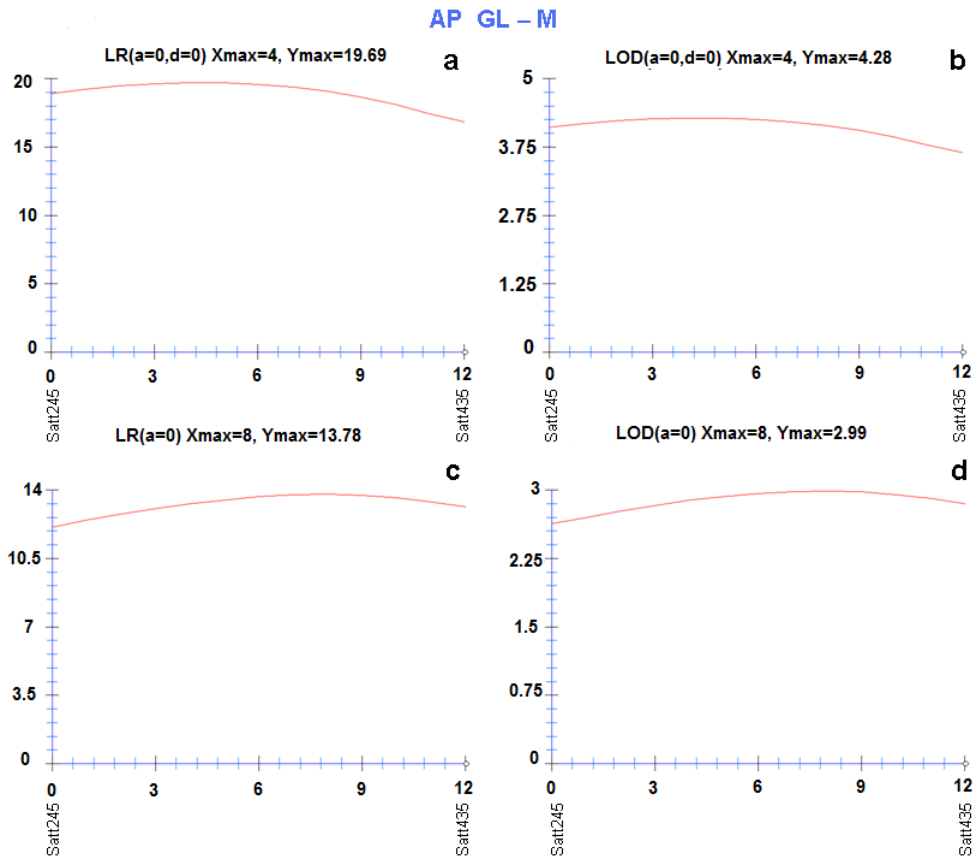
**Figura 4.13** | Mapeo por intervalo por el método máxima verosimilitud del QTL del peso seco de nódulos por nódulo por planta (PSNN) en el grupo de ligamiento C1. a) y b) prueba de posición del QTL (LR y LOD vs distancia de Kosambi); c) y d) pruebas de efecto de dominancia del QTL (LR y LOD vs distancia de Kosambi).



**Figura 4.14** | Mapeo por intervalo por el método máxima verosimilitud del QTL del peso seco de nódulos por nódulo por planta (PSNN) en el grupo de ligamiento M. a) y b) prueba de posición del QTL (LR y LOD, respectivamente vs distancia de Kosambi); c) y d) pruebas de efecto de dominancia del QTL (LR y LOD, respectivamente vs distancia de Kosambi).

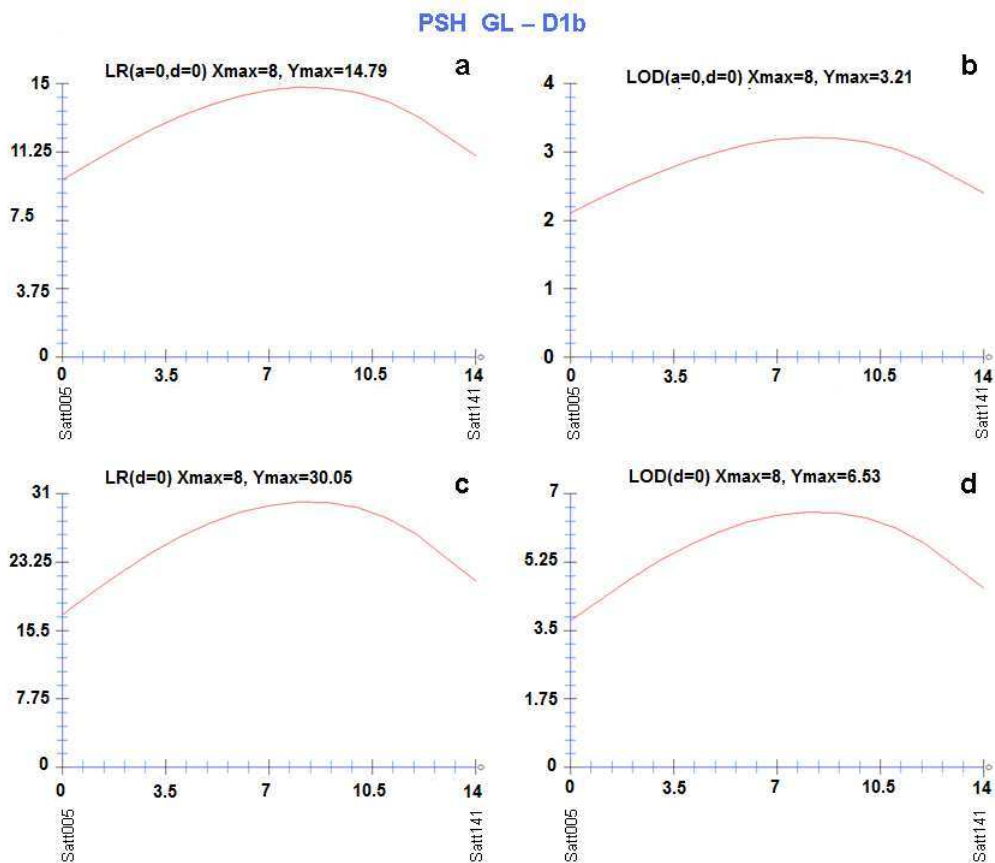
Para el carácter AP, al igual que en regresión por SIM, se ubicó un QTL en el GL M entre Satt245 y Satt435, a 4 cM de Satt245 (LOD = 4.28) y de efecto aditivo (LOD = 2.89) (**Figura 4.15**). Como señalamos anteriormente, ambos marcadores fueron los que presentaron  $R^2$  más altos, explicando el 50 % de la variación del carácter.

El PSH se localizó en el GL D1b dentro del intervalo formado por Satt005 y Satt141 a 6.2 cM de Satt141 (LOD = 3.21) y tuvo un significativo efecto dominante (LOD = 6.53) (**Figura 4.16**).



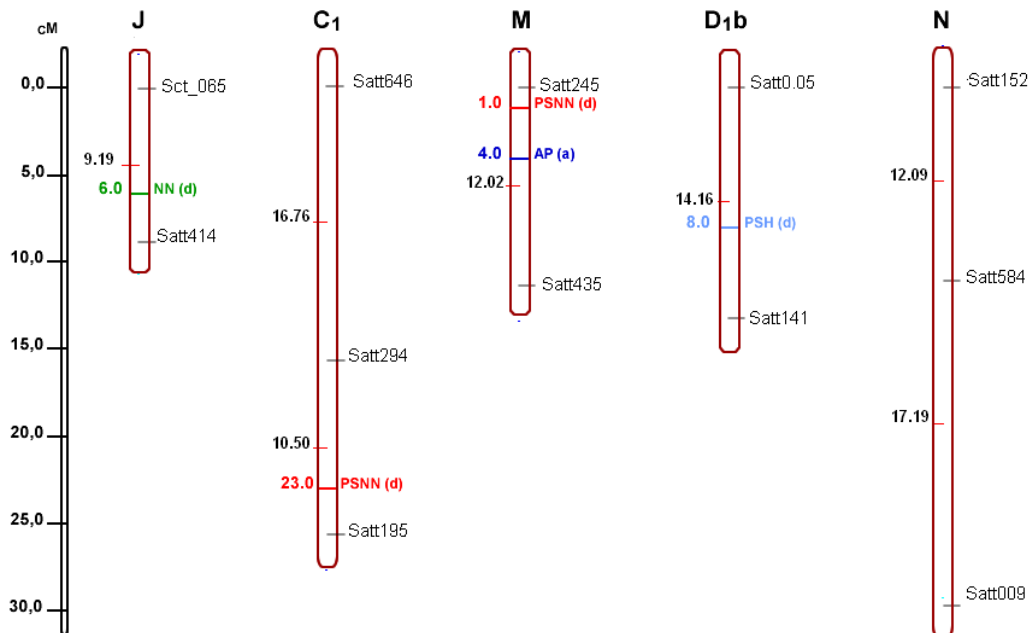
**Figura 4.15** | Mapeo por intervalo aplicando el método máxima verosimilitud del QTL de la altura de la planta (AP) en el grupo de ligamiento M. a) y b) prueba de posición del QTL (LR y LOD, respectivamente vs distancia de Kosambi); c) y d) pruebas de efecto de aditividad del QTL (LR y LOD, respectivamente vs distancia de Kosambi).





**Figura 4.16** | Mapeo por intervalo por el método máxima verosimilitud del QTL del peso seco de hojas por planta (PSH) en el grupo de ligamiento D1b. a) y b) prueba de posición del QTL (LR y LOD, respectivamente vs distancia de Kosambi); c) y d) pruebas de efecto de dominancia del QTL (LR y LOD, respectivamente vs distancia de Kosambi).

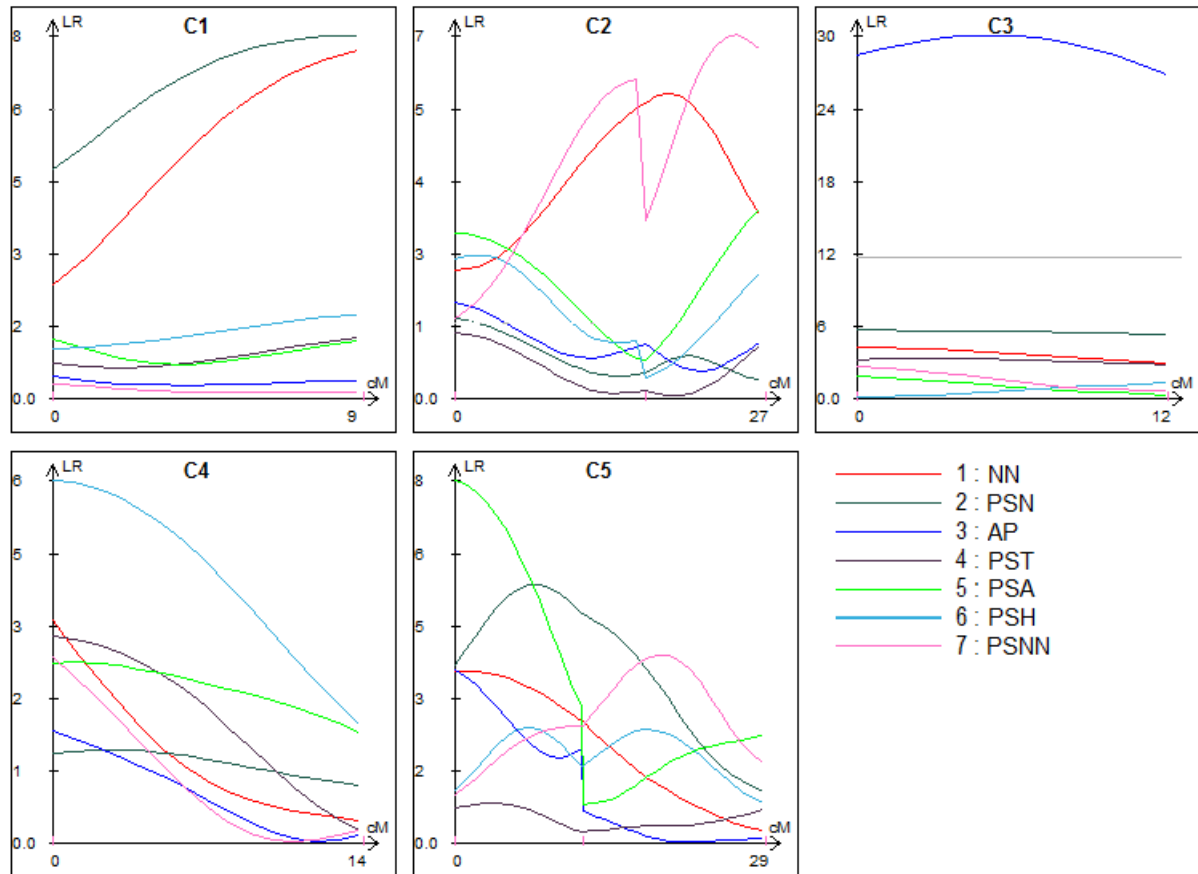
En la **Figura 4.17** se puede observar el mapa parcial con la localización de los QTLs hallados por intervalo simple.



**Figura 4.17** | Mapa parcial de ligamiento con la localización de los cinco QTLs y efecto aditivo (a) o dominante (d) usando el método de mapeo por intervalo simple (SIM). NN: número de nódulos/planta, AP: altura de planta (cm), PSH: peso seco de hoja (mg planta<sup>-1</sup>), PSNN: peso seco de nódulo/nódulo (mg nódulo<sup>-1</sup>). Se indican las posiciones de cada QTLs y los intervalos entre los marcadores en unidades cM de Kosambi.

#### 4.3.5.4. Mapeo por intervalo compuesto

Cuando se realizó el mapeo por intervalo compuesto, utilizando el programa QTL cartographer v.2.5, sólo se confirmó la localización del QTL para altura de planta (AP) en el GL-3 (cromosoma M) a 5 cM del marcador Satt245 con un LR de 29.7. Este valor fue superior al límite crítico estimado a partir de 1000 permutaciones y una significación al 5% (LRT = 11.42), con un efecto aditivo significativo de 11.46 que explicó un 29.7% de la varianza fenotípica total ( $R^2 = 0.30$ ) (**Figura 4.18**).



**Figura 4.18** | Mapeo de QTLs por intervalo compuesto. C1: cromosoma J, C2: cromosoma C1, C3: cromosoma M, C4: cromosoma D1b, C5: cromosoma N. NN: número de nódulos/planta, PSN: peso seco de nódulos/planta ( $\text{mg planta}^{-1}$ ), AP: altura de planta (cm), PST: peso seco de tallo ( $\text{mg planta}^{-1}$ ), PSA: peso seco aéreo ( $\text{mg planta}^{-1}$ ), PSH: peso seco de hoja ( $\text{mg planta}^{-1}$ ), PSNN: peso seco de nódulo/nódulo ( $\text{mg nódulo}^{-1}$ ). El eje de las x representa distancia de Kosambi (cM).

#### 4.4. Discusión

La distribución de frecuencia en la  $F_3$  de los caracteres cuantitativos estudiados, mostró una amplia norma de reacción, tal como es esperado para una población segregante cuya varianza fenotípica es explicada por la varianza genética y ambiental. Todas las variables de la población  $F_3$  presentaron una segregación transgresiva, ya que las clases fenotípicas de las colas de la distribución superaron a la medias de los progenitores. Nicolás *et al.* (2002) también reportó segregación transgresiva para la población segregante  $F_3$  producida a partir del cruzamiento entre BRS133 (baja

nodulación) x E-20 (media nodulación). La segregación transgresiva se produce cuando son distintos los genes que controlan el carácter en una y otra variedad progenitora, o bien cuando hay un número de genes menores y distintos en ellas. Esta segregación es típica cuando los parentales no están emparentados entre sí (Cubero, 2003). Esta herencia se cumplió no solo para caracteres asociados a la capacidad de nodulación sino también con las demás variables consideradas. Sin embargo, los resultados moleculares no demostraron una gran divergencia genética entre progenitores, ya que sólo un 20.6% (36 marcadores) fueron polimórficos para los progenitores y las plantas  $F_2$ . Una posible explicación sería que otras fuerzas estén actuando, tales como efectos maternos o efectos epistáticos no estimados.

Los valores fenotípicos de los caracteres de la  $F_3$  mostraron valores de CV%, generalmente superiores a 30%. Estos altos valores concuerdan con los reportados por otros autores también para variables asociadas a nodulación (Bohrer y Hungria, 1998; Hungria y Bohrer, 2000; Nicolás et al., 2002). En presencia de una población homogénea el CV% sólo puede atribuirse a causas de tipo aleatorio o más genéricamente a lo que se considera variabilidad de tipo ambiental, relacionados con errores experimentales o efectos de variabilidad microambiental que influyen en la expresión individual. En una población  $F_3$  la varianza residual está determinada por componentes genéticos y ambientales (Mariotti, 1986). El efecto ambiental produce “ruido”, pudiendo provocar un valor de varianza debida a la dominancia superior a su valor auténtico. Un buen diseño experimental con una completa distribución al azar del material en estudio disminuirá la covariación del genotipo con el ambiente (Cubero, 2003). Dado que las plantas crecieron en condiciones controladas de luz y temperatura, y que se aplicó un diseño completamente aleatorizado sería esperable una reducción mínima del componente ambiental entre y dentro de familias. Por ende, los altos valores de CV% encontrados estarían principalmente explicados por importantes efectos genéticos más que por elevados efectos ambientales. Un parámetro importante que permite medir la importancia relativa del componente

genético (G) con respecto al valor fenotípico (F) total es la heredabilidad en sentido amplio ( $H^2$ ) (Falconer y Mackay, 1996). Valores bajos de  $H^2$  para un carácter dado indican una base genética estrecha o una baja precisión experimental. En relación a lo discutido anteriormente, y en consistencia con los valores de  $H^2$ , se confirmó una alta varianza genética y una baja varianza ambiental. Las  $H^2$  para las variables PSA, PST, AP, NN y PSH alcanzaron valores superiores a 0.71, mientras que para las variables PSN y PSNN la  $H^2$  fue menor (**Tabla 4.2**). Valores de  $H^2$  moderados a altos coinciden con lo encontrado por otros autores, Tanya *et al.* (2005), estudiando una población de 136 RILs (SJ2 x Suwon 157), encontraron valores de heredabilidad para PSA, PSN y para NN de 0.66, 0.55 y 0.78, respectivamente. Nicolás *et al.* (2002) y Santos *et al.* (2006) estudiando cuatro poblaciones  $F_3$  derivadas de cruzamientos entre cultivares brasileños, citaron valores de heredabilidad en sentido estricto de moderados a bajos para estas 3 variables. Es importante destacar que las variables asociadas a capacidad de nodular presentaron valores más bajos de  $H^2$ . Esto podría indicar efectos génicos epistáticos no controlados o bien la existencia de un componente intrínseco a dichas variables que no influye sobre las demás variables. Una situación diferente podría ser la producida por la interacción compleja entre huésped (bacteria)-hospedador (planta) necesaria para la producción de nódulos.

En línea con lo arriba expuesto, fueron detectados valores de correlaciones genéticas y fenotípicas similares. Cheverud (1988) señaló que una alta heredabilidad incrementa la contribución proporcional de la correlación genética a la fenotípica en detrimento de la ambiental.

A partir de los resultados de correlaciones entre las variables, se encontraron evidencias similares a las descritas por otros autores. Sinclair *et al.* (1991), al comparar los caracteres de nodulación *versus* la masa de la raíz y de la parte aérea encontraron correlaciones medias a altas. También encontraron una alta correlación al comparar PSN *versus* el NN (entre  $r = 0.84$  y  $r = 0.91$ ). Hungria y Bohrer (2000) en un experimentos con 152 cultivares evaluados en invernáculo, citaron coeficientes de

correlación fenotípicas de 0.58 entre PSN y NN, de 0.66 entre PSN y PSA y de 0.31 entre NN y PSA. Hungria y Bohrer (2000) también citaron altas correlaciones entre el N total de la parte aérea (NTPA) *versus* PSN y PSA ( $r = 0.70$  y  $r = 0.91$ , respectivamente). Estos resultados confirman a los caracteres PSN y PSA, como buenos indicadores de nodulación y de N fijado por la planta. Por otra parte, Souza *et al.* (2008) estimaron correlaciones altas entre PSN y NN ( $r = 0.91$ ) y entre PSA y el NTPA ( $r = 0.88$ ), cuando utilizaron experimentos a campo con repeticiones en seis estados brasileros. Estos autores sugieren que en suelos pobres en N y en ausencia de fertilización nitrogenada, los caracteres PSN y PSA serían buenos indicadores de la FBN.

En cuanto a los resultados del análisis molecular se encontró bajo polimorfismo entre los progenitores y sólo un bajo porcentaje de microsatélites pudo ser utilizado (16.3%) en el análisis de QTLs. Este patrón estaría señalando una baja divergencia genética entre las variedades progenitoras que podría ser explicado por la existencia de cierto grado de parentezco entre ellas. Como consecuencia se encontraron 5 grupos de ligamiento que se correspondieron con regiones parciales de 5 cromosomas sobre un total de 20. A pesar de la reducida región genómica estudiada, las distancias encontradas entre marcadores fueron menores a 20 cM, que es la distancia recomendada para mapeo (Kearsey y Pooni, 1996).

Al realizar una comparación entre los métodos de mapeo de QTLs, se pudieron observar diferentes situaciones:

Los resultados encontrados para la mayoría de los QTLs se confirmaron por ambos métodos: el de locus simple y por SIM. Un QTL para NN se localizó cerca del Satt414 cuando se aplicó el SIM y también se probó asociación significativa al usar mapeo por locus simple (ANOVA y regresión). Para la AP dos de los marcadores (Satt245 y Satt435) que explicaron mayormente la variación fenotípica al aplicar locus simple resultaron ser los límites del intervalo donde se localizó el QTL por SIM y también fue el único QTL que se detectó por CIM.

La proporción de varianzas fenotípicas explicadas por los QTLs encontrados fueron altas para la mayoría de los QTLs localizados (> 20%). La variable que alcanzó una mayor proporción fue la AP con dos QTLs que sumaron más del 50%. Coincidentemente dichos QTLs explicaron una proporción de varianza genética mayor al 35 % de la varianza genética total.

En cuanto a la significancia de los efectos génicos de aditividad-dominancia, en general, los resultados encontrados al aplicar la metodología SIM y marca simple (ANOVA y regresión) fueron similares. Resultados dispares fueron encontrados para NN que presentó un efecto aditivo significativo al usar el locus simple (ANOVA y regresión) y dominante en SIM. El bajo porcentaje de la variación fenotípica ( $R^2 = 7.4\%$ ) explicado por el QTL conjuntamente con una baja probabilidad (5%) encontrada para el efecto aditivo explicarían dicha divergencia. Para el caso de la variable PSNN, un QTL fue mapeado a 1 cM de Satt245 en GL M. Sin embargo, cuando se aplicó el método de locus simple no se observó asociación significativa entre el QTL y dicho marcador. Dado que el método SIM ubica con mayor eficiencia QTLs podríamos confirmar la existencia de dicho QTL. Lander y Botstein (1989) y Nodari *et al.* (1993) encontraron que el mapeo de QTLs utilizando el análisis de regresión se aproxima a los resultados obtenidos por intervalo simple cuando el QTL está localizado cerca del locus marcador. La falta de saturación no permitió aplicar métodos de mapeo por intervalo, que presentan mayor precisión y eficiencia en la localización de un QTL.

Es interesante destacar que en el trabajo realizado por Nicolás *et al.* (2006), al utilizar el método de mapeo de locus simple ANOVA, detectó asociación significativa entre el locus marcador Satt414 con dos QTLs para NN y PSN, ambos con efectos dominantes significativos. En este estudio también se asociaron dichas variables con el locus marcador Satt414, sin embargo solo se probaron efectos aditivos significativos, para ambas variables al aplicar regresión. Por otro lado, el locus Satt434 estuvo asociado a varios QTLs: PSN, AP y PST, por ANOVA-regresión pero no se

encontró asociación a ningún GL. Santos *et al.* (2006) demostraron asociación entre dicho marcador y el PSA en el GL H.

En este trabajo se confirmó la asociación entre el locus marcador Satt197 y un QTL para PSNN con efecto dominante significativo, que también fue encontrado por Nicolás *et al.* (2006), utilizando una población de mapeo  $F_{2:3}$  proveniente del cruzamiento BRS133 x E-20. Estos mismos autores, aplicando el método de regresión y SIM, encontraron asociación entre este marcador y variables relacionadas a la nodulación (PSN), y la producción de biomasa vegetal (PSA), ambos con efectos aditivos significativos. Este mismo marcador estuvo también asociado al NN cuando se realizó un mapeo en una población  $F_{2:7}$  proveniente del cruzamiento entre Bossier x E-20 (Santos *et al.*, 2006). Es importante resaltar que en **capítulo 3** de la presente tesis se demostró asociación entre el alelo A3 del marcador Satt197 con un aumento del NN y del PSA cuando se realizó el biplot del AC (**Figura 3.4**). Otros alelos además estuvieron asociados con las clases fenotípicas de poblaciones con diferentes capacidades de nodulación. Entre ellos, los alelos A1 y A3 del Satt216 que discriminaron grupos de alta y medio NN, respectivamente (**Figura 3.4**).

La utilización de poblaciones derivadas de diferentes cruzamientos podría concluir en mapas distintos. Por ejemplo, la utilización de poblaciones derivadas de los cruzamientos BRS133 x E-20 (Nicolás *et al.*, 2005) y Bossier x E-20 (Santos *et al.*, 2006) generaron resultados diferentes. Esto sería causado principalmente por la disposición de los alelos en los progenitores (Mian *et al.*, 1998). Analizando la ocurrencia de QTLs en las características relacionadas a la producción de la soja en tres poblaciones diferentes de RILs, Orf *et al.* (1999) identificaron algunos QTLs en las tres poblaciones, otros en dos, y otros en apenas una población. En ese contexto, la base genética también contribuiría para la identificación. La ausencia de un QTL no puede ser simplemente atribuida a la falta de ligamiento entre un locus marcador y el QTL. En muchos casos, una asociación entre un marcador y un QTL puede ser falso positiva y en este caso se estaría cometiendo error de tipo I (Dudley, 1993), esto



significaría no aceptar una hipótesis nula cuando ésta es verdadera. Fasoula *et al.*, 2004, señalaron que los estudios sobre detección de QTLs en soja son limitados y contradictorios, por lo que sería necesario repetirlos utilizando diferentes poblaciones y ambientes para poder aplicar los conocimientos encontrados con mayor precisión en programas de mejoramiento.

Algunos marcadores se asociaron significativamente a más de una característica fenotípica o QTL. Satt434 se asoció con PSN, AP y PST; Satt271 con NN y PSA; Satt560 con PSN y PSNN y Satt216 estuvo asociado con PST, PSH y AP. Otros autores, también encontraron que algunos microsatélites se asociaron a más de un carácter, al evaluar caracteres relacionados con la FBN (Tanya, *et al.*, 2005; Nicolas *et al.*, 2006; Santos, *et al.*, 2006). Las variables estuvieron correlacionadas y esto fue mostrado por los coeficientes medios a altos y positivos, especialmente para los pares PSN vs. NN ( $r = 0.72$ ). El PSH estuvo mejor correlacionado que el PST que con el PSN ( $r = 0.73$  y  $r = 0.56$  respectivamente), siendo estas las variables que mejor explicarían la acumulación de N en el tejido foliar teniendo en cuenta la biomasa total (PSA). El análisis de sendero demostró que el PSN tiene un efecto directo importante, principalmente sobre la biomasa foliar (PSH). Además ésta también determina medianamente a la biomasa de tallo (PST). Por otro lado, el número de nódulos no influiría directamente sobre las variables de biomasa sino únicamente lo haría a través del PSN. Por otro lado el PSA estaría determinado principalmente por el PST, el que a su vez estaría influenciado por el PSH y la AP. Dado que las variables, PST y PSH constituyen al PSA se puede concluir que es el peso de los nódulos y no el número de nódulos es el carácter que contribuye mayormente al aumento de la biomasa aérea. Souza *et al.* (2008) alcanzaron resultados similares a los encontrados en este trabajo y deberían tenerse en cuenta en caso de aplicarse métodos de selección indirecta en programas de mejoramiento.

La asociación de varios QTLs con un mismo marcador podría indicar una herencia pleiotrópica o bien la existencia de varios QTLs ligados en el mismo sector de

un cromosoma. Esta última causa podría ser revertida por recombinación génica, lo que permitiría romper un ligamiento no deseado. En este estudio, se localizó en el cromosoma H, al marcador Satt434 asociado a tres QTLs, el PSN, el PST y la AP, pudiendo ser consecuencia de ligamiento en fase de acoplamiento. Estos resultados estarían confirmados además por la influencia directa que el PSN tuvo sobre las características de biomasa demostrada por el análisis de coeficientes de paso.

#### 4.4.1. Otros QTLs coincidentes relacionados con FBN

En las regiones donde se identificaron QTLs para FBN, se realizó una búsqueda en la base de datos Soybase (2009) con el objetivo de verificar QTLs descritos en esos intervalos (SIM) que estuvieran relacionados con alta producción y calidad de grano, ambas características que se asocian a una alta demanda de N.

Se encontró que en este intervalo se ha citado un QTL para peso de semilla asociado al Satt195 (Chen *et al.*, 2007). En el mapa generado en este trabajo se localizó un QTL para PSNN asociado al GL C1, en el intervalo Satt294-Satt195.

En el GL J se registró ligamiento entre el Satt414 y QTLs asociados con rendimiento (Guzman *et al.*, 2007), con NN (Nicolás *et al.*, 2006) y con PSN (Tanya *et al.*, 2005). Nuestro mapa incluyó un QTL para NN en el intervalo Sat\_065-Satt414.

En el GL M, donde se mapeó el QTL para PSNN y para AP, en el intervalo Satt245-Satt435, otros autores encontraron un QTL para contenido total de daidzeina en semillas (Satt245) (Primomo *et al.*, 2005). La daidzeina es un isoflavonoide, y pertenece a un grupo de compuestos que son secretados por la planta y que atraen químicamente a los rizobios hacia la región apical de los pelos radicales y son consideradas una señal primaria durante el proceso de infección para la producción de nódulos y fijación de nitrógeno.

Cabe destacar que el marcador Satt271 asociado con el NN, se encuentra vecino al locus nod49. Su producto génico es la proteína GmNFR1 $\alpha$ , que forma parte

del receptor de membrana y que interviene en el sensado del factor Nod en el inicio de la interacción temprana de las raíces con los rizobios. Esta primer señalización desencadena en la organogénesis del nódulo. Además, se probó en soja un aumento de la nodulación utilizando ensayos de sobreexpresión del gen (Indrasumunar *et al.*, 2011).

## **4.5. Conclusiones**

En base a los estudios genéticos realizados sobre la población  $F_{2:3}$  se presentaron las siguientes conclusiones:

- Localización de QTLs para las variables estudiadas que explican valores mayores al 15% de la varianza fenotípica total de la población segregante de cultivares argentinos de soja.
- Herencia transgresiva para la mayoría de las variables asociadas a capacidad de nodulación, típica de los caracteres cuantitativos, que presentan medios a altos valores de heredabilidad, asegurando un exitoso avance genético en programas de mejora para nodulación.
- Asociación en acoplamiento entre variables relacionadas con la capacidad de nodular y el aumento de la biomasa y la FBN.
- Para caracteres de capacidad de nodulación, se localizaron QTLs en los siguientes grupos de ligamiento: GL H para NN; PSN y PSNN; GL D1b para NN; GL A2 para NN; GL J para NN y PSN; GL B2 para PSN y PSNN; GL B1 para PSNN, GL C1 para PSNN. Estos QTLs explicaron entre 12.2-13.4% de la variación fenotípica total.
- Para caracteres de biomasa, relacionados con la FBN, se localizaron QTLs en los siguientes grupos de ligamiento: GL H para AP y PST; GL M para AP. GL D1b para AP, PST, PSA y PSH; GL B2 para AP; GL N para PSA. GL L para

PSA; GL G para PSH. Estos QTLs explicaron entre 8.3-25.6% de la variación fenotípica total.



## Consideraciones finales

Para manejar un QTL como si fuera un gen mendeliano, es necesario que tenga un efecto importante sobre la variación total, y que además los genes estén ubicados en una corta región del mapa y no dispersos a lo largo de todo el genoma. Esto es condición necesaria para ser eficientes en la selección asistida. Para esto es necesario aplicar y contar con marcadores flanqueantes y de algún marcador adicional en el interior del QTL para reducir al máximo el riesgo de recombinación entre los extremos. La existencia de un bajo nivel de divergencia genética entre los progenitores hizo que las metodologías no identificarán con mayor eficiencia QTLs.

La utilización de mapas más saturados permitirá discernir acerca de si los genes que determinan la capacidad de nodular son pocos o muchos, concentrados o esparcidos, importantes o insignificantes. Pocos y muy localizados es la condición ideal. Si son varios pero extendidos en una gran longitud y explicando una pequeña parte de la variación total, el efecto práctico sería nulo, y además un elevado número de marcadores haría más costoso el proceso de selección tradicional. Si bien la metodología aplicada permitió localizar algunos QTLs de nodulación en intervalos, estas regiones deberían saturarse con otros marcadores de mayor eficiencia y variabilidad como los SNPs. Una mayor saturación conjuntamente con la información disponible sobre el genoma de la soja y los bancos de ESTs permitirían identificar genes candidatos relacionados con el carácter en estudio.

La herencia transgresiva observada en la población de fenotipado ( $F_{2:3}$ ) permite distinguir claramente las colas de la distribución de frecuencia. Estas contienen los individuos más extremos (menos y más nodulantes) y, en teoría, presentarían los genes que controlan la capacidad de nodular en estado homocigota. La selección de individuos extremos permitirá obtener nuevos progenitores con características contrastantes extremas en homocigosis. Los cruzamientos entre ellos redundarán en

poblaciones segregantes con mayor variabilidad genética de tipo aditiva, que es la que mejor explica el parecido entre parientes y es la única que puede ser fijada por selección.

Nuestro ensayo se realizó en un único ambiente, inoculando con una cepa seleccionada de alta eficiencia, en condiciones controladas de luz y temperatura y en un medio nutritivo sin nitrógeno. Como se explicó en la introducción, la capacidad de fijar nitrógeno por parte de las leguminosas puede ser condicionada no sólo por el genotipo de la planta sino por prácticas culturales, por condiciones edafológicas particulares, y por factores climáticos (temperatura, humedad, etc). Por lo tanto, sería imprescindible en los futuros programas de mejora, incorporar ensayos experimentales bajo diferentes condiciones ambientales.

## Referencias bibliográficas

- Abaidoo, RC, Keyser, HH, Singleton, PW y Borthakur, D. 2000. *Bradyrhizobium* spp. (TGx) isolates nodulating the new soybean cultivars in Africa are diverse and distinct from *Bradyrhizobia* that nodulate North American soybeans. *Int J Syst Evol Micr* 50:225-234.
- Abe J y Ohara M. 1992. New electrophoretic mobility variations observed in wild soybean (*G. soja*) distributed in Japan and Korea. *Soyb Genet Newsl* 9:63-72.
- Abi-Ghanem R, Bodah ET, Wood M, Braunwart K. 2013. Potential breeding for high nitrogen fixation in *Pisum sativum* L.: germplasm phenotypic characterization and genetic investigation. *Am J Plant Sci* 4:1597-1600.
- Abu Assar AH, Uptmoor R, Abdelmula AA, Salih M, Ordon F y Friedt W. 2005. Genetic variation in sorghum germplasm from Sudan ICRISAT and USA assessed by simple sequence repeats (SSRs). *Crop Sci* 45:1636-1644.
- Akao S y Kouchi H. 1992. A supernodulating mutant isolated from soybean cultivar Enrei. *Soil Sci Plant Nutr* 38:183-187.
- Akkaya MS, Bhagwat AA y Cregan PB. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132:1131-1139.
- Akkaya MS, Shoemaker RC, Specht JE, Bhagwat AA y Cregan PB. 1995. Integration of simple sequence repeat DNA markers into a soybean linkage map. *Crop Sci* 35:1439-1445.
- Amor BB, Shaw SL, Oldroyd GE, Maillet F, Penmetsa RV, Cook D, Long SR, Dénarié J y Gough C. 2003. The NFP locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. *Plant J* 34:495-506.
- Austin DF y Lee M. 1998. Detection of quantitative trait loci for grain yield and yield components in maize across generations in stress and nonstress environments. *Crop Sci* 38:1296-1308.



- Austin G. 1988. Manual de Procesos Químicos en la Industria. McGraw-Hill. pp 365-366.
- Balatti PA. 2004. Competition an old unsolved problem of rhizobia nodulating soybean. En: Proceedings VII world soybean research conference, IV International soybean processing and utilization conference, III Congresso Brasileiro de soja. Foz do Iguazu. Brazil. pp. 499-508.
- Balatti PA y Pueppke SG. 1992. Identification of North American soybean lines that form nitrogen-fixing nodules with *Rhizobium fredii* USDA257. *Can J Plant Sci* 72:49-55.
- Barcellos FG, Menna P, Da Silva Batista JS y Hungria M. 2007. Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous *Sinorhizobium (Ensifer) fredii* and *Bradyrhizobium elkanii* in a Brazilian savannah soil. *Appl Environ Microb* 73:2635-2643.
- Basten CJ, Weir BS y Zeng ZB. 2003. QTL cartographer: ver. 1.17. Dept of Statistics, North Carolina State University. Raleigh, North Carolina.
- Becana M, Bedmar EJ. 1991. Metabolismo del nitrógeno y oxígeno en nódulos de leguminosas. En: Olivares J, Barea JM (eds.). Fijación y Movilización Biológica de Nutrientes. Vol 2. CSIC. Madrid. pp 33-45.
- Bernard RL. 1972. Two genes affecting stem termination in soybeans. *Crop Sci* 12:235-239.
- Berrada H y Fikri-Benbrahim K. 2014. Taxonomy of the Rhizobia: Current Perspectives. *Brit Microbiol Res J* 4:616-639.
- Bhuvaneshwari TV, Bhagwat AA y Bauer WD. 1981. Transient susceptibility of root cells in four common legumes to nodulation by rhizobia. *Plant Physiol* 68:1144-1149.
- Bohrer TRJ y Hungria M. 1998. Avaliação de cultivares de soja quanto à fixação biológica do nitrogênio. *Pesqui Agropecu Bras* 33:937-953.

- Bornet B y Branchard M. 2001. Nonanchored Inter Simple Sequence Repeat (ISSR) markers: Reproducible and specific tools for genome fingerprinting. *Plant Mol Biol Rep* 19:209-215.
- Bourion V, Rizvi SM, Fournier S, de Larambergue H, Galmiche F, Marget P, Duc G y Burstin J. 2010. Genetic dissection of nitrogen in pea through a QTL approach of root, nodule and shoot variability. *Theor Appl Genet* 121:71-86.
- Brechenmacher L, Kim MY, Benitez M, Li M, Joshi T, Calla B, Lee MP, Libault M, Vodkin LO, Xu D, Lee SH, Clough SJ y Stacey G. 2008. Transcription profiling of soybean nodulation by *Bradyrhizobium japonicum*. *Mol Plant Microbe Interact* 21:631-645.
- Brett PJ, Deshazer D y Woods DE. 1998. *Burkholderia thailandensis* sp. nov., a *Burkholderia pseudomallei*-like species. *Int J Syst Bacteriol* 48:317-320.
- Brewin NJ. 1991. Development of the legume root nodule. *Annu Rev Cell Biol* 7:191-226.
- Brookes AJ. 1999. The essence of SNPs. *Gene* 234:177-186.
- Brose E, Freire JRJ y Müller L. 1979. Relações entre genótipos de soja (*Glycine max* [L.] Merrill), fixação simbiótica do nitrogênio e rendimento de grãos. *Agron Sulriog* 15:179-198.
- Broughton WJ, Jabbouri S y Perret X. 2000. Keys to symbiotic harmony. *J Bacteriol* 182:5641-5652.
- Burris RH y Roberts GP. 1993. Biological nitrogen fixation. *Annu Rev Nutr* 13:317-335.
- Caldwell BE, Hinson K y Johnson HW. 1966. A strain-specific ineffective nodulation reaction in the soybean *Glycine max* L. Merrill. *Crop Sci* 6:495-496.
- Caldwell BE. 1966. Inheritance of a strain-specific ineffective nodulation in soybean. *Crop Sci* 6:427-428.
- Camacho M, Santamaría C, Temprano F, Rodríguez-Navarro DN, Daza A, Espuny R, Bellogín R, Ollero FJ, Lyra de MC, Buendía-Clavería A, Zhou J, Li FD, Mateos C, Velázquez E, Vinardell JM y Ruiz-Sainz JE. 2002. Soils of the Chinese

- Hubei province show a very high diversity of *Sinorhizobium fredii* strains. *Syst Appl Microbiol* 25:592-602.
- Campbell N y Reece J. 2007. Biología, 7ma ed. Médica Panamericana. Buenos Aires, Madrid.
- Carroll BJ, McNeil DL y Gresshoff PM. 1985. A supernodulation and nitrate-tolerant symbiotic (nts) soybean mutant. *Plant Physiol* 78:34-40.
- Catoira R, Galera C, de Billy F, Penmetsa RV, Journet EP, Maillet F, Rosenberg C, Cook D, Gough C y Dénarié J. 2000. Four genes of *Medicago truncatula* controlling components of a nod factor transduction pathway. *The Plant Cell* 12:1647-1665.
- Chaintreuil C, Giraud E, Prin Y, Lorquin J, Bâ A, Gillis M, De Lajudie P y Dreyfus B. 2000. Photosynthetic Bradyrhizobia are natural endophytes of the Africa wild rice *Oryza breviligulata*. *Appl Environ Microbiol* 66:5437-5447.
- Chen Q, Zhang Z, Liu C, Xin D, Qiu H, Shan D, Shan C, Hu G. 2007. QTL analysis of major agronomic traits in soybean. *Agr Sci China* 6:399-405.
- Chen WL, Huang QY y Xiong XJ. 2004. Distribution and biodiversity of soybean rhizobia in the soils of Shennongjia forest reserve, China. *Biol Fert Soils* 40:306-312.
- Chen WX, Yan GH y Li JL. 1988. Numerical taxonomic study of fast-growing soybean rhizobia and a proposal that *Rhizobium fredii* be assigned to *Sinorhizobium* gen. nov. *Int J Syst Bacteriol* 38:392-397.
- Cheverud JM. 1988. A comparison of genetic and phenotypic correlations. *Evolution* 42:958-968.
- Choi IY, Hyten DL, Matukumalli LK, Song Q, Chaky JM, Quigley CV, Chase K, Lark KG, Reiter RS, Yoon MS, Hwang EY, Yi SI, Young ND, Shoemaker RC, van Tassell CP, Specht JE y Cregan PB. 2007. A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics* 176:685-696.

- Churchill GA y Doerge RW. 1994. Empirical threshold values for quantitative trait mapping. *Genetics* 138:963-971.
- Cleyet-Marel JC. 1987. Dynamique des populations de *Rhizobium* et de *Bradyrhizobium* dans le sol et la rhizosphere. These d'Etat, Université Claude Bernard, Lyon, France. 240 p.
- Coelho ASG. 2000. Considerações gerais sobre a análise de QTL's. En: Pinheiro JB, Carneiro IF. (eds.). Análise de QTL no melhoramento de plantas. Fundação de Apoio à Pesquisa. Goiânia. pp 1-36.
- Collard BCY, Jahufer MZZ, Brouwer JB y Pang ECK. 2005 An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* 142:169-196.
- Concibido VC, La Vallee B, McLaird P, Pineda N, Meyer J, Hummel L, Yang J, Wu K, Delannay X. 2003. Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. *Theor Appl Genet* 106:575-582.
- Cooper JE y Feil EJ. 2004. Multilocus sequence typing - what is resolved?. *Trends Microbiol* 12:373-377.
- Cregan PB y Hartwig EE. 1984. Characterization of flowering response to photoperiod in diverse soybean genotypes. *Crop Sci* 24:659-662
- Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J y Specht JE. 1999. An integrated genetic linkage map of the soybean genome. *Crop Sci* 39:1464-1490.
- Cregan PB, Keyser HH y Sadowsky MJ. 1989. Host plant effects on nodulation and competitiveness of the *Bradyrhizobium japonicum* serotype strains constituting serocluster 123. *Appl Environ Microb* 55:2532-2536.
- Cruz CD. 2001. Programa Genes (Versão Windows): aplicativo computacional em genética e estatística. Universidade Federal de Viçosa. Viçosa, Brasil.

- Cubero JI. 2003. Introducción a la mejora genética vegetal. Ed. Mundi-Prensa. Madrid. 567 p.
- Dean DR y Jacobsen MR. 1992. Biochemical genetics of nitrogenase. En: Stacey G, Burris RH, Evans HJ (eds.). Biological Nitrogen Fixation. Chapman and Hall. New York. pp 763–834.
- Debellé F, Moulin L, Mangin B, Dénarié J y Boivin C. 2001. Nod genes and Nod signals and the evolution of the *Rhizobium* legume symbiosis. *Acta Biochim Pol* 48:359-65.
- Delamuta JRM, Ribeiro RA, Menna P, Bangel EV y Hungria M. 2012. Multilocus sequence analysis (MLSA) of *Bradyrhizobium* strains: revealing high diversity of tropical diazotrophic symbiotic bacteria. *Braz J Microbiol* 43:698-710.
- Dénarié J, Debellé F y Promé JC. 1996. *Rhizobium* lipo-chitooligosaccharide nodulation factors: Signaling molecules mediating recognition and morphogenesis. *Annu Rev Biochem* 65:503-35.
- Diwan N y Cregan PB. 1997. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theor Appl Genet* 95:723-733.
- Döbereiner J y Arruda NB. 1967. Interrelações entre variedades e nutrição na nodulação e simbiose da soja. *Pesq Agropec Bras* 2:475-487.
- Döbereiner J. 1966. Evaluation of nitrogen fixation in legumes by the regression of total plant nitrogen with nodule weight. *Nature* 210:850-852.
- Dowdle SF, Bohlool BB. 1985. Predominance of fast-growing *Rhizobium japonicum* in a soybean field in the People's Republic of China. *Appl Environ Microbiol* 50:1171-1176.
- Downie JA. 1998. Functions of rhizobial nodulation genes. En: Spaink HP, Kondorosi A y Hooykaas PJJ (eds.). The Rhizobiaceae. Kluwer Academic Publishers. Dordrecht, The Netherlands. pp. 387-402.

- Doyle, JJ, Doyle JL, Rauscher JT, Brown AHD. 2004. Evolution of the perennial soybean polyploid complex (*Glycine* subgenus *Glycine*): a study of contrasts. *Biol J Linn Soc* 82:583-897.
- Doyle JJ, Doyle JL y Harbison C. 2003a. Chloroplast-expressed glutamine synthetase in *Glycine* and related Leguminosae: phylogeny, gene duplication, and ancient polyploidy. *Syst Bot* 28:567-577.
- Doyle JJ, Doyle JL, Rauscher JT y Brown AHD. 2003b. Diploid and polyploidy reticulate evolution throughout the history of the perennial soybeans (*Glycine* subgenus *Glycine*). *New Phytol* 161:121-132.
- Dudley JW. 1993. Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. *Crop Sci* 33:660-668.
- Edwards MD, Stuber CW and Wendel JF. 1987. Molecular marker-facilitated investigations of quantitative trait loci in maize. I: numbers, genomic distribution, and types of gene action. *Genetics* 116:113-125.
- Egli DB y Leggett JE. 1973. Dry matter accumulation patterns in determinate and indeterminate soybeans. *Crop Sci* 13:220-222.
- Egli DB, Guffy RD y Leggett JE. 1985. Partitioning of assimilate between vegetative and reproductive growth in soybean. *Agron J* 77:917-922.
- Espitia M, Aramendiz H y Cadena J. 2008. Correlaciones y análisis de sendero en algodón *Gossypium hirsutum* L. en el Caribe colombiano. *Rev Fac Nac Agron* 61:4325-4335.
- Evans J, O'Connor GE, Turner GL, Coventry DR, Fettell N, Mahoney J, Armstrong EL y Walsgott DN. 1989. N<sub>2</sub> fixation and its value to soil N increase in lupin, field pea and other legumes in south-eastern Australia. *Aust J Agr Res* 40:791-805.
- Falconer DS y Mackay TFC. 1996. Introduction to Quantitative Genetics. 4th ed. Longman Group Ltd. Essex, UK.

- Fasoula VA, Harris DK y Boerma HR. 2004. Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. *Crop Sci* 44:1218-1225.
- Fedorova M, van de Mortel J, Matsumoto PA, Cho J, Town CD, VandenBosch KA, Gantt JS y Vance CP. 2002. Genome-wide identification of nodule-specific transcripts in the model legume *Medicago truncatula*. *Plant Physiol* 130:519-537.
- Fehr W, Caviness C, Burmood D y Pennington J. 1971. Stage of development descriptions for soybeans, *Glycine max* (L.) Merrill. *Crop Sci* 11:929-931.
- Ferguson BJ, Indrasumunar A, Hayashi S, Lin MH, Lin YH, Reid DE y Gresshoff PM. 2010. Molecular analysis of legume nodule development and autoregulation. *J Integr Plant Biol* 52:61-76.
- Ferreira ME y Grattapaglia D. 1996. Introdução ao uso de marcadores moleculares em análise genética. 2° ed. EMBRAPA-CENARGEN, Brasília.
- Foley TC, Orf JH y Lambert JW. 1986. Performance of related determinate and indeterminate soybean lines. *Crop Sci* 26:5-8.
- Garrity GM, Johnson KL, Bell J y Searles DB. 2002. Bergey's Manual of Systematic Bacteriology. Second Edition. Springer-Verlag. New York.
- Germano MG, Menna P, Mostasso FL y Hungria M. 2006. RFLP analysis of the rRNA operon of a Brazilian collection of bradyrhizobial strains from 33 legume species. *Int J Syst Evol Microbiol* 56:217-229.
- Giorda LM y Baigorri H. 1997. El cultivo de la soja en argentina. Agro 4 de Córdoba. INTA. C.R. Córdoba.
- Godoy D, Randle G, Simpson AJ, Aanensen DM, Pitt TL, Kinoshita R y Spratt BG. 2003. Multilocus sequence typing and evolutionary relationships among the causative agents of melioidosis and glanders, *Burkholderia pseudomallei* and *Burkholderia mallei*. *J Clin Microbiol* 41:2068-2079.

- Gorman MB. 1984. An electrophoretic analysis of the genetic variation in the wild and cultivated soybean germplasm. *Diss Abstr Int B* 44:11.
- Graham PH, Sadowsky MJ, Keyser HH, Barnet M, Bradley RS, Cooper JE, De Ley DJ, Jarvis, BDW, Roslycky EB, Strijdom BW y Young JPW. 1991. Proposed minimal standards for the description of new genera and species of root- and stem- nodulating bacteria. *Int J Syst Bacteriol* 41:582-587.
- Greder RR, Orf JH y Lambert JW. 1986. Heritabilities and associations of nodule mass and recovery of *Bradyrhizobium japonicum* serogroup USDA 110 in soybean. *Crop Sci* 26:33-37.
- Gremaud MF y Harper JE. 1989. Selection and initial characterization of partially nitrate tolerant nodulation mutants of soybean. *Plant Physiol* 89:169-173.
- Gresshoff PM. 2003. Post-genomic insights into plant nodulation symbioses. *Genome Biol* 4:201-206.
- Gupta P, Varshney R, Sharma P y Ramesh B. 1999. Molecular markers and their applications in wheat breeding. *Plant Breed* 118:369-390.
- Gurley WP, Hepburn AG y Key JL. 1979. Sequence organization of the soybean genome. *Biochim Biophys Acta* 561:167-183.
- Guzman P, Diers B, Neece D, St. Martin S, LeRoy A, Grau C, Hughes T y Nelson R. 2007. QTL associated with yield in three backcross-derived populations of soybean. *Crop Sci* 47:111-122.
- Haley CS y Knott SA. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69:315-324.
- Hanson CH, Robinson HF y Comstock RE. 1956. Biometrical studies of yield in segregating populations of Korean *Lespedeza*. *Agron J* 48:268-272.
- Harper JE. 1999. Nitrogen fixation – Limitations and potential. World soybean research conference VI, Chicago, Illinois.



- Hayashi M, Saeki Y, Haga M, Harada K, Kouchi H y Umehara Y. 2012. Rj (rj) genes involved in nitrogen-fixing root nodule formation in soybean. *Breed Sci* 61:544-553.
- Hayashi M, Shiro S, Kanamori H, Mori-Hosokawa S, Sasaki-Yamagata H, Sayama T, Nishioka M, Takahashi M, Ishimoto M, Katayose Y, Kaga A, Harada K, Kouchi H, Saeki Y, Umehara Y. 2014. A thaumatin-like protein, Rj4, controls nodule symbiotic specificity in soybean. *Plant Cell Physiol* 55:1679-1689.
- Herridge D y Rose I. 2000. Breeding for enhanced nitrogen fixation in crop legumes. *Field Crop Res* 65:229-248.
- Herridge DF, Turpin JE y Robertson MJ. 2001. Improving nitrogen fixation of crop legumes through breeding and agronomic management analysis with simulation modelling. *Aust J Exp Agr* 41:391-401.
- Hungria M y Bohrer TRJ. 2000. Variability of nodulation and dinitrogen fixation capacity among soybean cultivars. *Biol Fert Soils* 31:45-52.
- Hymowitz T y Newell CA. 1981 Taxonomy of the genus *Glycine*, domestication and uses of soybeans. *Econ Bot* 37:371-379.
- Hymowitz T. 1970. On the domestication of the soybean. *Econ Bot* 24:408-421.
- Hymowitz, T y Singh RJ. 1987. Taxonomy and Speciation. En: Soybeans: improvement, production, and uses. 2nd ed. Agronomy Monograph no.16. pp.23-48.
- Indrasumunar A, Kereszt A, Searle I, Miyagi M, Li D, Nguyen CD, Men A, Carroll BJ, Gresshoff PM. 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* 51:201-214.
- Indrasumunar A, Searle I, Lin MH, Kereszt A, Men A, Carroll BJ y Gresshoff PM. 2011. Nodulation factor receptor kinase 1 $\alpha$  controls nodule organ number in soybean (*Glycine max* L. Merr). *Plant J* 65:39-50.

- Inouye J, Shanmagasundaram S y Masuyuma F. 1979. Effects of temperature and day length on the flowering of some photoinsensitive soybean varieties. *Japan. J. Tropical Agric* 22:167-172.
- Jansen RC y Stam P. 1994. High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* 136:1447:1455.
- Jarvis BDW y Tighe SW. 1994. Rapid identification of *Rhizobium* species based on cellular fatty acid analysis. *Plant Soil* 161:31-41.
- Jarvis BDW, van Berkum P, Chen WX, Nour SM, Fernandez MP, Cleyet-Marel JC y Gillis M. 1997. Transfer of *Rhizobium loti*, *Rhizobium huakuii*, *Rhizobium ciceri*, *Rhizobium mediterraneum*, and *Rhizobium tianshanense* to *Mesorhizobium* gen. nov. *Int J Syst Bacteriol* 47:895-898.
- Javaheri F. 1996. Naturally nodulating (promiscuous) soybean varieties for smallholder farmers: the Zambian experience. En: Mpeperekwi S, Makonese F, Giller KE (eds.). Soyabeans in smallholder cropping systems of Zimbabwe. SoilFertNet/CIMMYT, Harare, Zimbabwe, pp. 63-73.
- Jeong SC, Yang K, Park JY, Han KS, Yu S, Hwang TY, Hur CG, Kim SH, Park PB, Kim HM, Park YI, Liu JR. 2006. Structure, expression, and mapping of two nodule-specific genes identified by mining public soybean EST databases. *Gene* 393:71-80.
- Jitackson S y Sadowsky MJ. 2008. Nodulation gene regulation and quorum sensing control density-dependent suppression and restriction of nodulation in the *Bradyrhizobium japonicum*-soybean symbiosis. *Appl Environ Microbiol* 74:3749-3756.
- Jordan DC. 1982. Transfer of *Rhizobium japonicum* Buchanan 1980 to *Bradyrhizobium* gen. nov., a genus of slow-growing, root nodule bacteria from leguminous plants. *Int J Syst Bacteriol* 32:136-139.
- Journet EP, van Tuinen D, Gouzy J, Crespeau H, Carreau V, Farmer MJ, Niebel A, Schiex T, Jaillon O, Chatagnier O, Godiard L, Micheli F, Kahn D, Gianinazzi-

- Pearson V y Gamas P. 2002. Exploring root symbiotic programs in the model *Medicago truncatula* using EST analysis. *Nucleic Acids Res* 30:5579-5592.
- Kantolic A, Giménez P, de la Fuente E y Giménez P. 2006. Capítulo 2.2: Soja. En: de la Fuente EB, Gil A, Giménez PI, Kantolic AG, López Pereira M, Ploschuk E, Sorlino DM, Vilariño MP, Wassner DF y Windauer LB (eds.). Cultivos Industriales. 1ra edición. Facultad de Agronomía, Buenos Aires. pp 95-141.
- Kao CH, Zeng ZB y Teasdale RD. 1999. Multiple interval mapping for quantitative trait loci. *Genetics* 152:1203-1216.
- Kearsey MJ y Pooni HS. 1996. The genetical analysis of quantitative traits. Chapman and Hall, London.
- Kearsey MJ, Hyne V. 1994. QTL analysis: a simple 'marker regression' approach. *Theor Appl Genet* 89:698-702.
- Keeney DR y Walsh LW. 1972. Available nitrogen in rural ecosystems. Sources and fate. *Hortscience* 7:219-223.
- Keim P, Diers BW, Olson TC y Shoemaker RC. 1990. RFLP mapping in soybean: association between marker loci and variation in quantitative traits. *Genetics* 126:735-742.
- Keim P, Schupp JM, Travis SE, Clayton K, Zhu T, Shi L, Ferreira A y Webb DM. 1997. A high-density soybean genetic map based on AFLP markers. *Crop Sci* 37:537-543.
- Keyser HH, Bohlool BB, Hu TS y Weber DF. 1982. Fast-growing rhizobia isolated from root nodules of soybean. *Science* 215:1631-1632.
- Kobayashi H y Broughton WJ. 2008. Fine-tuning of symbiotic genes in rhizobia: flavonoid signal transduction cascade. En: Dilworth MJ, James EK, Sprent JI y Newton WE (eds.). Nitrogen-fixing leguminous symbioses. Springer. Dordrecht, The Netherlands, pp 117-152.

- Kondorosi E, Banfalvi Z y Kondorosi A. 1984. Physical and genetic analysis of a symbiotic region of *Rhizobium meliloti* identification of nodulation genes. *Mol Gen Genet* 193:445-452.
- Korom S. 1992. Natural denitrification in the saturated zone: A review. *Water Resour Res* 28: 1657-1668.
- Kosambi DD. 1944. The estimation of map distance from recombination values. *Ann Eugen* 12:172-175.
- Kouchi H, Imaizumi-Anraku H, Hayashi M, Hakoyama T, Nakagawa T, Umehara Y, Suganuma N y Kawaguchi M. 2010. How many peas in a pod?. Legume genes responsible for mutualistic symbioses underground. *Plant Cell Physiol* 51:1381-1397.
- Kuykendall LD, Saxena B, Devine TE y Udell SE. 1992. Genetic diversity in *Bradyrhizobium japonicum* Jordan 1982 and a proposal for *Bradyrhizobium elkanii* sp. nov. *Can J Microbiol* 38:501-505.
- Labandera C. 2007. Actividades en fijación biológica de nitrógeno: situación actual y perspectivas en Uruguay. En: Izaguirre- Mayoral ML, Labandera C, Sanjuán, J (eds.). Biofertilizantes en Iberoamérica: Una visión técnica científica y empresarial. Imprenta Denad Internacional. Montevideo. pp 75-78.
- Lagercrantz U, Ellegran H y Anderson L. 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Res* 21:1111-1115.
- Laguerre G, Allard MR, Revoy F y Amarger N. 1994. Rapid identification of rhizobia by restriction fragment length polymorphism analysis of PCR-amplified 16S rRNA genes. *Appl Environ Microbiol* 60:56-63.
- Lander ES y Botstein D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199.
- Lebart L, Morineau A y Warwick K. 1984. Multivariate Descriptive Statistical Analysis. Wiley. New York.

- Lee H, Hur CG, Oh CJ, Kim HB, Pakr SY y An CS. 2004. Analysis of the root nodule-enhanced transcriptome in soybean. *Mol Cells* 18:53-62.
- Lee HS y Lee SH. 1998. Introduction, development and characterization of supernodulating soybean mutant-nitrate inhibition of nodulation and nitrogen fixation in supernodulating soybean mutant. *Korean J Crop Sci* 43:23-27.
- Lee JS y Verma DPS. 1984. Chromosomal arrangement of leghemoglobin genes in soybean. *Nuc Acids Res* 11:5541-5553.
- Lee JS, Brown GG y Verma DP. 1983. Chromosomal arrangement of leghemoglobin genes in soybean. *Nuc Acids Res* 11:5541-5553.
- Li QQ, Wang ET, Zhang YZ, Zhang YM, Tian CF, Sui XH, Chen WF y Chen WX. 2011. Diversity and biogeography of rhizobia isolated from root nodules of *Glycine max* grown in Hebei Province, China. *Microb Ecol* 61:917-931.
- Li CC. 1975. Path analysis - A primer. Boxwood Press, Pacific Grove, California.
- Li W, Han Y, Zhang D, Yang M, Teng W, Jiang Z, Qiu L y Sun G. 2008a. Genetic diversity in soybean genotypes from north-eastern China and identification of candidate markers associated with maturity rating. *Plant Breeding* 127:494-500.
- Li Y, Guan R, Liu Z, Ma Y, Wang L, Li L, Lin F, Luan W, Chen P, Yan Z, Guan Y, Zhu L, Ning X, Smulders MJ, Li W, Piao R, Cui Y, Yu Z, Guan M, Chang R, Hou A, Shi A, Zhang B, Zhu S y Qiu L. 2008b. Genetic structure and diversity of cultivated soybean (*Glycine max* (L) Merr) landraces in China. *Theor Appl Genet* 117:857-871.
- Libault M, Farmer A, Brechenmacher L, Drnevich J, Langley RJ, Bilgin DD, Radwan O, Neece DJ, Clough SJ, May GD y 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 Physiol* 152:541-552.

- Lie TA. 1974. Environmental Effects on Nodulation and Symbiotic Nitrogen Fixation. En: Quispel A (ed.). The Biology of nitrogen fixation. North Holland Publishing Company. Amsterdam. pp. 555-582.
- Limpens E, Franken C, Smit P, Willemsse J, Bisseling T y Geurts R. 2003. LysM Domain receptor kinases regulating rhizobial nod factor-induced infection. *Science* 302:630-633.
- Lincoln SE, Daly MJ y Lander ES. 1993. Mapping genes controlling quantitative traits using MAPMAKER/QTL. Whitehead Institute for Biomedical Research. Cambridge, Massachusetts.
- Lipkovich I y Smith EP. 2001. Biplot and singular value decomposition macros for Excel. *J Stat Softw* 7:1-13. <http://www.jstatsoft.org/v07/i05/paper>
- Liu BH. 1998. Genomics, linkage mapping and QTL analysis. CRC Press, Boca Raton, Florida.
- Lodwig EM, Hosie AH, Bourdès A, Findlay K, Allaway D, Karunakaran R, Downie JA y Poole PS. 2003. Amino-acid cycling drives nitrogen fixation in the legume-Rhizobium symbiosis. *Nature* 422:722-726.
- Loh J, Carlson RW, York WS y Stacey G. 2002a. Bradyoxetin, a unique chemical signal involved in symbiotic gene regulation. *Proc Natl Acad Sci USA* 99:14446-14451.
- Loh J, Pierson EA, Pierson LS III, Stacey G y Chatterjee A. 2002b. Quorum sensing in plant-associated bacteria. *Curr Opin Plant Biol* 5:285-290.
- Loh J y Stacey G. 2003. Nodulation gene regulation in *Bradyrhizobium japonicum*: a unique integration of global regulatory circuits. *Appl Environ Microbiol* 69:10-17.
- Loh JT y Stacey G. 2001. Feedback regulation of the *Bradyrhizobium japonicum* nodulation genes. *Mol Microbiol* 41:1357-1364.
- Long SR. 1996. *Rhizobium* symbiosis: nod factors in perspective. *Plant Cell* 8:1885-1898.

- López SMY, Pastorino GN, Alcántara VM, Salvucci RD, Balatti PA y Albanesi AS. 2013. Los rizobios que nodulan la soja en sitios con ambientes nativos y cultivados de la Argentina. En *Microbiología agrícola: un aporte de la investigación en Argentina*. Magna Publicaciones, Tucumán. pp 237-252.
- Ludwig W, Amann R, Martinez-Romero E, Schonhuber W, Bauer S, Neef A y Schliefer, K-H. 1998. rRNA based identification and detection systems for rhizobia and other bacteria. *Plant Soil* 204:1-19.
- Lynch M, Walsh B. 1998. *Genetics and analysis of quantitative traits*. Sinauer Associates, Sunderland, Massachusetts.
- Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M, Szczyglowski K, Sato S, Kaneko T, Tabata S, Sandal N y Stougaard J. 2003. A receptor-like kinase gene involved in legume perception of rhizobial signal molecules. *Nature* 425:637-640.
- Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M y Spratt BG. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci USA* 95:3140-3145.
- Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB y Lark KG. 1996. Genetic mapping of agronomic traits recombinant inbred lines of soybean. *Crop Sci* 36:1327-1336.
- Mariotti JA. 1986. *Fundamentos de genética biométrica. Aplicaciones al mejoramiento genético vegetal*. OEA, Monografía N° 32. 152 p.
- Martínez Dougnac G. 2004. *Apuntes acerca de la historia de la soja en la Argentina. Elementos para delinear experiencias comparadas*. Documentos del CIEA. [http://www.econ.uba.ar/www/servicios/Biblioteca/bibliotecadigital/Indicealfabetico/archivos/Apuntes acerca de la historia de la soja en la Argentina. Elementos para delinear experiencias comparadas.pdf](http://www.econ.uba.ar/www/servicios/Biblioteca/bibliotecadigital/Indicealfabetico/archivos/Apuntes%20cerca%20de%20la%20historia%20de%20la%20soja%20en%20la%20Argentina.Elementos%20para%20delinear%20experiencias%20comparadas.pdf)

- Masson-Boivin C, Giraud E, Perret X y Batut J. 2009. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes?. *Trends Microbiol* 17: 458-466.
- Maunoury N, Kondorosi A, Kondorosi E y Mergaert P. 2008. Cell biology of nodule infection and development. En: Dilworth MJ, James EK y Sprent JI y Newton WE (eds.). *Nitrogen-fixing Leguminous Symbioses*. Springer. Dordrecht, The Netherlands. pp. 153-189.
- Mauricio R. 2001. Mapping quantitative trait loci in plants: Uses and caveats for evolutionary biology. *Nat Rev Genet* 2:370-381.
- McCouch SR y Doerge RW. 1995. QTL mapping in rice. *Trends Genet* 11:482-487.
- McIntosh MS. 1983. Analysis of combined experiments. *Agronomy J* 75:153-155.
- Menna P, Hungria M, Barcellos FG, Bangel EV, Hess PN y Martinez-Romero E. 2006. Molecular phylogeny based on the 16S rRNA gene of elite rhizobial strains used in Brazilian commercial inoculants. *Syst Appl Microbiol* 29:315-332.
- Menna P, Pereira AA, Bangel EV y Hungria M. 2009. rep-PCR of tropical rhizobia for strain fingerprinting, biodiversity appraisal and as a taxonomic and phylogenetic tool. *Symbiosis* 48:120-130.
- Mergaert P, Van Montagu M y Holsters M. 1997. Molecular mechanisms of Nod factor diversity. *Mol Microbiol* 25:811-817.
- Meyer JD, Silva DC, Yang C, Pedley KF, Zhang C, van de Mortel M, Hill JH, Shoemaker RC, Abdelnoor RV, Whitham SA y Graham MA. 2009. Identification and analyses of candidate genes for Rpp4-mediated resistance to Asian soybean rust in soybean (*Glycine max* (L.) Merr.). *Plant Physiol* 150:295-307.
- Mian MAR, Ashley DA, Vencill WK y Boerma R. 1998. QTLs conditioning early growth in a soybean population segregating for growth habit. *Theor Appl Genet* 97:1210-1216.



- Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR y Sasaki T. 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Mol Breed* 3:87-103.
- Montgomery DC y Peck EA. 1981. Introduction to linear regression analysis. John Wiley & Sons. New York. 504 pp.
- Moreno CLM. 2010. Caracterización de las cepas ICA L9 e ICA J96 bacterias simbióticas fijadoras de nitrógeno y pruebas de estabilidad de inoculantes elaborados para cultivos de alverja y soya. Tesis de Maestría. Pontificia Universidad Javeriana. Bogotá, Colombia.
- Morgante M y Olivieri A. 1993. PCR-amplified microsatellites as markers in plant genetics. *Plant J* 3:175-182.
- Morgante M. 1994. Genetic mapping and variability of seven soybean simple sequence repeat loci. *Genome* 37:763-769.
- Moulin L, Béna G, Boivin-Masson C y Stepkowski T. 2004. Phylogenetic analyses of symbiotic nodulation genes support vertical and lateral gene co-transfer within the *Bradyrhizobium* genus. *Mol Phylogenet Evol* 30:720-732.
- Mperekki S, Javaheri F, Davis P y Giller KE. 2000. Soyabeans and sustainable agriculture: Promiscuous soybean in Southern Africa. *Field Crop Res* 65:137-149.
- Murray JD, Muni RR, Torres-Jerez I, Tang Y, Allen S, Andriankaja M, Li G, Laxmi A, Cheng X, Wen J, Vaughan D, Schultze M, Sun J, Charpentier M, Oldroyd G, Tadege M, Ratet P, Mysore KS, Chen R y Udvardi MK. 2011. Vapyrin, a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of *Medicago truncatula*. *Plant J* 65:244-252.
- Mylona P, Pawlowski K y Bisseling T. 1995. Symbiotic Nitrogen Fixation. *Plant Cell* 7:869-885.
- Nei M. 1987. Molecular evolutionary genetics. Columbia University Press. New York.

- Neuhausen SL, Graham PH y Orf JH. 1988. Genetic variation for dinitrogen fixation in soybean of maturity group 00 and 0. *Crop Sci* 28:769-772.
- Nicolás MF, Arrabal Arias CA, Hungria M. 2002. Genetics of nodulation and nitrogen fixation in Brazilian soybean cultivars. *Biol Fert Soils* 36:109-117.
- Nicolás MF, Hungria M y Arias CAA. 2006. Identification of quantitative trait loci controlling nodulation and shoot mass in progenies from two Brazilian soybean cultivars. *Field Crops Res* 95: 355-366.
- Nishimura R, Hayashi M, Wu GJ, Kouchi H, Imaizumi-Anraku H, Murakami Y, Kawasaki S, Akao S, Ohmori M, Nagasawa M, Harada K y Kawaguchi M. 2002. HAR1 mediates systemic regulation of symbiotic organ development. *Nature* 420:426-429.
- Nodari RO, Tsai SM, Guzmán P, Gilbertson RL y Gepts P. 1993. Towards an integrated linkage map of common bean III. Mapping genetic factors controlling host bacteria interactions. *Genetics* 134:341-350.
- Novák K, Biedermannová E y Vondryš J. 2009. Symbiotic and growth performance of supernodulating forage pea lines. *Crop Sci* 49:1227-1234.
- Nutman PS. 1969. Genetics of symbiosis and nitrogen fixation in legumes. *P Roy Soc Lond B Bio* 172:417-437
- Oke V y Long SR. 1999. Bacteroid formation in the Rhizobium-legume symbiosis. *Curr Opin Microbiol* 2:641-646. doi:10.1016/S1369-5274(99)00035-1
- Oldroyd GED y Downie JA. 2008. Coordination of nodule morphogenesis with rhizobial infection in legumes. *Annu Rev Plant Biol* 59:519-546.
- Oldroyd GED. 2013. Speak, friend, and enter: Signalling systems that promote beneficial symbiotic associations in plants. *Nat Rev Microbiol* 11:252-263.
- Orf JH, Chase K, Jarvik T, Mansur LM, Cregan PB, Adler FR y Lark KG. 1999. Genetics of soybean agronomic traits: I. Comparison of three related recombinant inbred populations. *Crop Sci* 39:1642-1651.

- Osunde AO, Gwam S, Bala A, Sanginga N y Okogun JA. 2003. Responses to rhizobial inoculation by two promiscuous soybean cultivars in soils of the Southern Guinea savanna zone of Nigeria. *Biol Fertil Soils* 37:274-279.
- Pacheco J, Pat-Canul R, Cabrera A. 2002. Análisis del ciclo del nitrógeno en el medio ambiente con relación al agua subterránea y su efecto en los seres vivos. *Ingeniería* 6:73-81.
- Paerl HW, Dennis RL, Whitall DR. 2002. Atmospheric deposition of nitrogen: implications for nutrient over-enrichment of coastal waters. *Estuaries* 25:677-693.
- Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES y Tanksley SD. 1991b. Mendelian factors underlying quantitative traits in tomato comparison across species generations and environments. *Genetics* 127:181-198.
- Paterson AH. 1996a. Making genetic maps. En: Paterson AH (ed.). *Genome mapping in plants*. R. G. Landes Company, Academic Press. San Diego, California. pp. 23-39.
- Pazdernik DL, Graham PH y Orf JH. 1997. Heritability in the early nodulation of F<sub>3</sub> and F<sub>4</sub> soybean lines. *Can J Plant Sci* 77:201-205.
- Polhill RM y Raven PH. 1981. *Advances in legume systematics*. Royal Botanic Gardens, Kew, UK.
- Pongsilp N, Triplett EW y Sadowsky MJ. 2005. Detection of homoserine lactone-like quorum sensing molecules in *Bradyrhizobium* strains. *Curr Microbiol* 51:250-254.
- Postma D, Boesen C, Kristiansen H, Larsen F. 1991. Nitrate reduction in an unconfined sandy aquifer: Water chemistry, reduction processes, and geochemical modeling. *Water Resour Res* 27:2027-2045.

- Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S y Rafalski A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breed* 2:225-238.
- Primomo V, Poysa V, Ablett G, Jackson C, Gijzen M y Rajcan I. 2005. Mapping QTL for individual and total isoflavone content in soybean seeds. *Crop Sci* 45:2545-2464.
- R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Racca RW, Collino DJ. 2005. Bases fisiológicas para el manejo de la fijación biológica del nitrógeno en soja. Congreso Mundo Soja. Buenos Aires. pp 111-120.
- Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Grønlund M, Sato S, Nakamura Y, Tabata S, Sandal N y Stougaard J. 2003. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* 425:585-592.
- Ramírez-Bahena MH, Peix A, Rivas P, Camacho M, Rodríguez-Navarro DN, Mateos PF, Martínez-Molina E, Willems A y Velázquez E. 2009. *Bradyrhizobium pachyrhizi* sp. nov. and *Bradyrhizobium jicamae* sp. nov., isolated from effective nodules of *Pachyrhizus erosus*. *Int J Syst Evol Microbiol* 59:1929-1934.
- Rauscher JT, Doyle JJ y Brown AHD. 2004. Multiple origins and nrDNA internal transcribed spacer homeologue evolution in the *Glycine tomentella* (Leguminosae) allopolyploid complex. *Genetics* 166:987-998.
- Reyna N, Sneller CH. 2001. Evaluation of marker-assisted introgression of yield QTL alleles into adapted soybean. *Crop Sci* 41:1317-1321.
- Ribeiro RA, Barcellos FG, Thompson FL y Hungria M. 2009. Multilocus sequence analysis of Brazilian Rhizobium microsymbionts of common bean (*Phaseolus vulgaris* L.) reveals unexpected taxonomic diversity. *Res Microbiol* 160:297-306.

- Ritchie SW, Hanway JJ, Thomson HE y Benson GO. 1985. How soybean plant develops. Iowa State University of Science and Technology. Cooperative Extension Service. Ames. Iowa. Special Report 53. 20 p.
- Rivas R, Willems A, Palomo JL, García-Benavides P, Mateos PF, Martínez-Molina E, Gillis M y Velázquez E. 2004. *Bradyrhizobium betae* sp. nov., isolated from roots of *Beta vulgaris* affected by tumour-like deformations. *Int J Syst Evol Microbiol* 54:1271-1275.
- Rohlf FJ. 1998. NTSYSpc: Numerical Taxonomy and Multivariate Analysis System Version 2.02 Exeter Software Setauket. New York.
- Rowell DL. 1996. Soil Science: Methods and applications. Longman. Harlow, UK
- Russo RC. 1985. Ammonia, nitrite and nitrate. En: Rand GM, Petrocelli SR (eds.). Fundamentals of aquatic toxicology, Hemisphere Publishing Corporation. Washington DC. pp 455-471.
- Saghai Maroof MA, Jeong SC, Gunduz I, Tucker DM, Buss GR y Tolin SA. 2008. Pyramiding of soybean mosaic virus resistance genes by marker-assisted selection. *Crop Sci* 48:517-526.
- Salvucci RD, Aulicino M, Hungria M, Balatti PA. 2012. Nodulation capacity of Argentinean soybean (*Glycine max* L Merr) cultivars inoculated with commercial strains of *Bradyrhizobium japonicum*. *Am J Plant Sci* 3:130-140.
- Santos MA, Geraldi IO, Gracia AAF, Bortolatto N, Schiavon A y Hungria M. 2013. Mapping of QTLs associated with biological nitrogen fixation traits in soybean. *Hereditas* 150:17-25.
- Santos MA, Nicolás MF y Hungria M. 2006. Identificação de QTL associados à simbiose entre *Bradyrhizobium japonicum*, *B. elkanii* e soja. *Pesq Agropec Bras* 41:67-75.
- Schauser L, Roussis A, Stiller J y Stougaard J. 1999. A plant regulator controlling development of symbiotic root nodules. *Nature* 402:191-195.

- Schlötterer C y Tautz D. 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Res* 20:211-215.
- Schmidt EL. 1982. Nitrification in soil. En: Stvenson FJ (ed.). Nitrogen in agricultural soils. Am Soc Agron Inc. Madison, USA, pp 253-288.
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC y Jackson SA. 2010. Genome sequence of the palaeopolyploid soybean. *Nature* 463:178-83.
- Schultze M y Kondorosi A. 1995. What makes nodulation signals host-plant specific?. *Trends Microbiol* 3:370-372.
- Schultze M y Kondorosi A. 1998. Regulation of symbiotic root nodule development. *Annu Rev Genet* 32:33-57.
- Schuster I y Cruz CD. 2004. Estatística genômica aplicada a populações derivadas de cruzamentos controlados. Universidade Federal de Viçosa. Viçosa, Brasil. 568 p.
- Searle IR, Men AE, Laniya TS, Buzas DM, Iturbe-Ormaetxe I, Carroll BJ y Gresshoff PM. 2003. Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. *Science* 299:109-112.
- Sharma B y Ahlert RC. 1977. Nitrification and nitrogen removal. *Wat Res* 11:897-925.
- Shi A, Chen P, Li D, Zheng C, Zhang B y Hou A. 2009. Pyramiding multiple genes for resistance to soybean mosaic virus in soybean using molecular markers. *Mol Breeding* 23:113-124.
- Shoemaker RC, Guffy RD, Lorenzen LL y Specht JE. 1992. Molecular genetic mapping of soybean: map utilization. *Crop Sci* 32:1091-1098.

- Shoemaker RC, Schlueter J, Cregan PB y Vodkin L. 2003. The status of soybean genomics and its role in the development of soybean biotechnologies. *AgBioForum* 6:1-10. <http://www.agbioforum.org/v6n12a02-shoemaker.htm>
- Shoemaker RC y Specht JE. 1995. Integration of the soybean molecular and classical genetic linkage groups. *Crop Sci* 35:436-446.
- Sinclair TR, Purcell LC, King CA, Sneller CH, Chen P y Vadez, V. 2007. Drought tolerance and yield increase of soybean resulting from improved symbiotic N<sub>2</sub> fixation. *Field Crop Res* 101:68-71.
- Sinclair TR, Soffes AR, Hinson K, Albrecht SL y Pfahler PL. 1991. Genotypic variation in soybean nodule number and weight. *Crop Sci* 31:301-304.
- Skorupska H, Albertsen MC, Langholz KD y Palmer RG. 1989. Detection of ribosomal RNA genes in soybean by *in situ* hybridization. *Genome* 32:1091-1095.
- Smartt J y Hymowitz T. 1985. Domestication and evolution of grain legumes. En: Summerfield RJ y Roberts EH (eds.). Grain legume crops. Collins, London. pp. 37-72.
- Smith GP. 1976. Evolution of repeated DNA sequences by unequal crossover. *Science* 191:528-535.
- Sneath PHA y Sokal RR. 1973. Numerical taxonomy: the principles and practice of numerical classification. Freeman. San Francisco, California.
- Sokal RR y Rohlf FJ. 1995. Biometry: The principals and practice of statistics in biological research. W.H. Freeman and Company. San Francisco, California.
- Song L, Carrol BJ, Gresshoff PM y Herridge DF. 1995. Field assessment of supernodulating genotypes of soybean for yield, N<sub>2</sub> fixation and benefit to subsequent crops. *Soil Biol Biochem* 27:563-569.
- Song Q, Quigley CV, Carter TE, Nelson RL, Boerma HR, Strachan J y Cregan PB. 1999. A selected set of trinucleotide simple sequence repeat markers for soybean variety identification. *Plant Var Seeds* 12:207-220.

- Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE y Cregan PB. 2004. A new integrated genetic linkage map of the soybean. *Theo Appl Genet* 109:122-128.
- Southern EM. 1976. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98:503-517.
- Souza AP. 2001. Biología molecular aplicada ao melhoramento. En: Nass LL, Valois ACC, Melo IS y Valadares-Inglis MC (eds.). Recursos genéticos e melhoramento de plantas. Fundação MT. Rondonópolis, Mato Grosso. Brasil. pp 939-966.
- Souza RA, Hungria M, Franchini JC, Maciel CD, Campos RJ y Zaia DAM. 2008. Conjunto mínimo de parâmetros para avaliação da microbiota do solo e da fixação biológica do nitrogênio pela soja. *Pesq Agropec Bras* 43:83-91.
- Soybase 2012. Soybase and the soybean breeder's toolbox. Knowledge and tools for soybean breeders and researchers. [www.soybase.org](http://www.soybase.org)
- Spaink HP. 1994. The molecular basis of the host specificity of the Rhizobium bacteria. *Anton Van Lee J Microb* 65:81-98.
- Stougaard J. 2000. Regulators and regulation of legume root nodule development. *Plant Physiol* 124:531-540.
- Straub SCK, Pfeil BE y Doyle JJ. 2006. Testing the polyploidy past of soybean using a low-copy nuclear gene - Is Glycine (Fabaceae: Papilionoideae) an auto- or allopolyploid?. *Molec Phylogenet Evol* 39:580-584.
- Stumm W, Morgan JJ. 1996. Aquatic chemistry, chemical equilibria and rates in natural waters. 3rd edition. John Wiley and Sons. New York. 1040 p.
- Subramanian S, Stacey G y Yu O. 2006. Endogenous isoflavones are essential for the establishment of symbiosis between soybean and Bradyrhizobium japonicum. *Plant J* 48:261-273.
- Taiz L y Zeiger E. 2010. Plant Physiology. 5ed. Sinauer Associates. Sunderland, Massachusetts.



- Tajima S, Nomura M y Kouchi H. 2004. Ureide biosynthesis in legume nodules. *Front Biosci* 9:1374-1381.
- Takats ST. 1986. Suppression of nodulation in soybeans by superoptimal inoculation with *Bradyrhizobium japonicum*. *Physiol Plantarum* 66:669-673.
- Tamm CO. 1991. Nitrogen in terrestrial ecosystems. Springer-Verlag, Berlin. 115 pp.
- Tanksley SD. 1993. Mapping polygenes. *Annu Rev Genet* 27:205-233.
- Tanya P, Srinives P, Toojinda T, Nakhon P, Vanavichit A y Lee SH. 2005. Identification of SSR markers associated with N<sub>2</sub> fixation components in soybean [*Glycine max* (L.) Merr.]. *Korean J. Genet* 27: 351-359.
- Tasma IM y Shoemaker RC. 2003. Mapping flowering time gene homologs in soybean and their association with maturity (E) loci. *Crop Sci* 43:319-328.
- Tasma IM, Lorenzen, LL, Green DE y Shoemaker RC. 2001. Mapping genetic loci for flowering time, maturity, and photoperiod insensitivity in soybean. *Mol Breeding* 8:25-35.
- Tautz D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* 17:6463-6471.
- Tesfaye M y Holl FB. 1998. Group-specific differentiation of Rhizobium from clover species by PCR amplification of 23S rDNA sequences. *Can J Microbiol* 44:1102-1105.
- Thoday JM. 1961. Location of polygenes. *Nature* 191:368-370.
- Thomas JF y Raper CD (Jr). 1983a. Effects on soybean growth during the onset of reproductive development under various temperature regimes. *Bot Gaz* 144:471-476.
- Thomas JF y Raper CD (Jr). 1983b. Photoperiod and temperature regulation of floral initiation and anthesis in soybean. *Ann Bot* 51:481-489.
- Thomas-Oates J, Bereszczak J, Edwards E, Gill A, Noreen S, Zhou JC, Chen MZ, Miao LH, Xie FL, Yang JK, Zhou Q, Yang SS, Li XH, Wang L, Spaink HP, Schlaman HR, Harteveld M, Díaz CL, van Brussel AA, Camacho M,

- Rodríguez-Navarro DN, Santamaría C, Temprano F, Acebes JM, Bellogín RA, Buendía-Clavería AM, Cubo MT, Espuny MR, Gil AM, Gutiérrez R, Hidalgo A, López-Baena FJ, Madinabeitia N, Medina C, Ollero FJ, Vinardell JM y Ruiz-Sainz JE. 2003. A catalogue of molecular, physiological and symbiotic properties of soybean-nodulating rhizobial strains from different soybean cropping areas of China. *System Appl Microbiol* 26:453-465.
- Thomasson AJ, Bouma J y Leith H (eds.). 1991. Soil and Groundwater Research Report. II. Nitrate in Soils. EUR13501 Office for Official Publications of the European Communities. Luxembourg. 544 pp.
- Tiedje JM, Sextone AJ, Parkin TB y Revsbech NB. 1984. Anaerobic processes in soil. *Plant Soil* 70:197-212.
- Tighe SW, de Lajudie P, Dipietro K, Lindstrom K, Nick G y Jarvis BDW. 2000. Analysis of cellular fatty acids and phenotypic relationships of *Agrobacterium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium* species using the Sherlock Microbial Identification System. *Int J Syst Evol Microbiol* 50:787-801.
- Toledo R y Moya G. 2008. Respuesta diferenciada de grupos de madurez de soja según fecha de siembra. Informe de Actualización Técnica n° 10. EEA INTA Marcos Juárez. pp 32-34.
- Townsend AR, Howarth RW, Bazzaz FA, Booth MS, Cleveland CC, Collinge SK, Dobson AP, Epstein PR, Holland EA, Keeney DR, Mallin MA, Rogers CA, Wayne P y Wolfe AH. 2003. Human health effects of a changing global nitrogen cycle. *Front Ecol Environ* 1:240-246.
- Trese AT. 1995. A single dominant gene in McCall soybean prevents effective nodulation with *Rhizobium fredii* USDA257. *Euphytica* 81:279-282.
- Trinick MJ. 1980. Relationships amongst the fast-growing rhizobia of *Lablab purpureus*, *Leucaena leucocephala*, *Mimosa* spp., *Acacia farnesiana* and *Sesbania*

- grandiflora* and their affinities with other rhizobial groups. *J Appl Microbiol* 49:39-53.
- Tsai SM, Nodari RO, Moon DH, Camargo L.EA, Vencovsky R y Gepts P. 1998. QTL mapping for nodule number and common bacterial blight in *Phaseolus vulgaris* L. *Plant Soil* 204:135-145.
- van Berkum P y Fuhrmann JJ. 2000. Evolutionary relationships among the soybean bradyrhizobia reconstructed from 16S rRNA gene and internally transcribed spacer region sequence divergence. *Int J Syst Bacteriol* 50:2165-2172.
- van Berkum P, Terefework Z, Paulin L, Suomalainen S, Lindström K y Eardly BD. 2003. Discordant phylogenies within the *rrn* loci of rhizobia. *J Bacteriol* 185:2988-2998.
- Van K, Kim MY y Lee SH. Genomic of root nodulation. 2007. En: Varshney RK, Tuberosa R (eds.). *Genomics Assisted Crop Improvement, Vol 2: Genomics Applications in Crops*. Springer. Dordrecht, The Netherlands. pp435-452.
- Van Loon LC, Rep M y Pieterse CM. 2006. Significance of inducible defense-related proteins in infected plants. *Annu Rev Phytopathol* 44:135-162.
- Vargas MAT, Peres JRR y Suhet AR. 1982. Fixação de nitrogênio atmosférico pela soja em solos de cerrado. *Inf Agrop* 8:20-23.
- Verneti FJ. 1983. Soja: Genética e Melhoramento. Fundação Cargill. v.2. Campinas. Brasil.
- Versalovic J, Schneider M, de Bruijn FJ y Lupski JR. 1994. Genomic fingerprinting of bacteria using repetitive sequence based polymerase chain reaction. *Meth Mol Cell Biol* 5:25-40.
- Vest G y Caldwell BE. 1972. Rj4-A gene conditioning ineffective nodulation in soybean. *Crop Sci* 12:692-693.
- Vest G. 1970. Rj3-A gene conditioning ineffective nodulation in soybean. *Crop Sci* 10:34-35.

- Vincent J. 1970. A manual for the practical study of the root-nodule bacteria. International Biological Program Handbook n .15. Blackwell Scientific Publications. Oxford, England.
- Vinuesa P, León-Barrios M, Silva C, Willems A, Jarabo-Lorenzo A, Pérez-Galdona R, Werner D y Martínez-Romero E. 2005. *Bradyrhizobium canariense* sp. nov., an acid-tolerant endosymbiont that nodulates endemic genistoid legumes (Papilionoideae: Genisteae) from the Canary Islands, along with *Bradyrhizobium japonicum* bv. *genistearum*, *Bradyrhizobium* genospecies alpha and *Bradyrhizobium* genospecies beta. *Int J Syst Evol Microbiol* 55:569-575.
- Vinuesa P, Rademaker JLW, de Bruijn FJ y Werner D. 1998. Genotypic characterization of *Bradyrhizobium* strains nodulating endemic woody legumes of the Canary Islands by PCR-Restriction Fragment Length Polymorphism analysis of genes encoding 16S rRNA (16S rDNA) and 16S-23S rRNA intergenic spacers, repetitive extragenic palindromic PCR genomic fingerprinting, and partial 16S rRNA sequencing. *Appl Environ Microbiol* 64 2096-2104.
- Vinuesa P, Silva C, Werner D y Martínez-Romero E. 2005. Population genetics and phylogenetic inference in bacterial molecular systematics: the roles of migration and recombination in *Bradyrhizobium* species cohesion and delineation. *Mol Phylogenet Evol* 34:29-54.
- Voorhees JH. 1915. Variations in Soybean Inoculation. *Agron J* 7:139-140.
- Vos P, Hogers R, Bleeker M, Reijans M, Lee TVD, Hornes M, Frijters A, Pot J, Pelemen J, Kuiper M y Zabeau M. 1995. AFLP: a new technique for DNA fingerprint. *Nucleic Acids Res* 23:4407-4414.
- Vuong TD y Harper JE. 2000. Inheritance and allelism analysis of hypernodulating genes in the NOD3-7 and NOD2-4 soybean mutants. *Crop Sci* 40:700-703.

- Wang H, Man CX, Wang ET y Chen WX. 2009. Diversity of rhizobia and interactions among the host legumes and rhizobial genotypes in an agricultural-forestry ecosystem. *Plant Soil* 314:169-182.
- Waters CM y Bassler BL. 2005. Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* 21:319-346.
- Watson SW, Valos FW y Waterbury JB. 1981. The Family Nitrobacteraceae. En: Starr MP, Stolp H y Trüper H (eds.). *The Prokaryotes*. Springer-Verlag. Berlin, Germany. pp 89-101.
- Weisburg WG, Barns SM, Pelletie DA y Lane DJ. 1991. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 173:697-703.
- Weiser GV, Skipper HD y Wollum AG. 1990. Exclusion of inefficient *Bradyrhizobium japonicum* serogroups by soybean genotypes. *Plant Soil* 121:99-105.
- Wetzel RG. 2001. *Limnology, lake and river ecosystems*. 3rd edition. Academic Press, San Diego. USA. 1006 pp.
- Wilkerson GG, Jones JW, Boote KJ y Buol GS. 1989. Photoperiodically sensitive interval in time to flower of soybean. *Crop Sci* 29:721-726.
- Willems A, Coopman R y Gillis M. 2001. Comparison of sequence analysis of 16S-23S rDNA spacer regions, AFLP analysis and DNA-DNA hybridizations in *Bradyrhizobium*. *Int J Syst Evol Microbiol* 51:623-632.
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA y Tingey SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18:6531-6535.
- Winter, P y Kahl G. 1995. Molecular marker technologies for plant improvement. *World J Microbiol Biotechnol* 11:438-448.
- Woese CR. 1987. Bacterial evolution. *Microbiol Rev* 51:221-271.
- Wu K, Burnquist W, Sorrells ME, Tew T, Moore P y Tanksley SD. 1992. The detection and estimation of linkage in polyploids using single-dose restriction fragments. *Theor Appl Genet* 83:294-300.

- Xu B. 1995. Geographical distribution and genetic diversity of seed traits of wild soybean in China. *Acta Agron Sin* 21:733-739.
- Xu LM, Ge C, Cui Z, Li J y Fan H. 1995. *Bradyrhizobium liaoningense* sp. nov., isolated from the root nodules of soybean. *Int J Syst Bacteriol* 45:706-711.
- Xu ML, Melchinger AE, Xia XC y Lübberstedt T. 1999. High-resolution mapping of loci conferring resistance to sugarcane mosaic virus in maize using RFLP SSR and AFLP markers. *Mol Gen Genet* 261:574-581.
- Xu SJ, Singh RJ y Hymowitz T. 2000a. Monosomics in soybean: origin, identification, cytology, and breeding behavior. *Crop Sci* 40:985-989.
- Xu SJ, Singh RJ, Kollipara KP y Hymowitz T. 2000b. Hypertriploid in soybean: origin, identification, cytology, and breeding behavior. *Crop Sci* 40:72-77.
- Yan W. 2001. GGEbiplot-A Windows application for graphical analysis of multienvironment trial data and other types of two-way data. *Agronomy J* 93:1111-1118.
- Yan W, Hunt LA, Sheng Q y Szlavnic Z. 2000. Cultivar evaluation and mega-environment investigation based on GGE biplot. *Crop Sci* 40:597-605.
- Yang JK, Zhang WT, Yuan TY y Zhou JC. 2006. Genotypic characteristics of the *rrn* operon and genome of indigenous soybean bradyrhizobia in cropping zones of China. *Can J Microbiol* 52:968-976.
- Yang S, Tang F, Gao M, Krishnan HB y Zhu H. 2010. R gene-controlled host specificity in the legume-rhizobia symbiosis. *Proc Natl Acad Sci USA* 107:18735-18740.
- Yao ZY, Kan FL, Wang ET, Wei GH y Chen WX. 2002. Characterization of rhizobia that nodulate legume species of the genus *Lespedeza* and description of *Bradyrhizobium yuanmingense* sp. nov. *Int J Syst Evol Microbiol* 52:2219-2230.
- Young, ND. 1994. Constructing a plant genetic linkage map with DNA markers. En: Ronald IKV y Phillips L (eds.). DNA-based markers in plants. Kluwer, Dordrecht/Boston/London. pp. 39-57.

- Yusuf AA, Abaidoo RC, Iwafor ENO y Olufajo OO. 2008. Genotype Effects of Cowpea and Soybean on Nodulation, N<sub>2</sub>-Fixation and N Balance in the Northern Guinea Savanna of Nigeria. *J Agron* 7:258-264.
- Zeng ZB. 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. *Proc Natl Acad Sci USA* 90:10972-10976.
- Zeng ZB. 1994. Precision mapping of quantitative trait loci. *Genetics* 136:1457-1468.
- Zeng ZB, Kao CH y Basten CJ. 1999. Estimating the genetic architecture of quantitative traits. *Genet Res* 74:279-289.
- Zhang RJ, Hou BC, Wang ET, Li Y, Zhang XX y Chen WX. 2011. *Rhizobium tubonense* sp. nov., isolated from root nodules of *Oxytropis glabra*. *Int J Syst Evol Microbiol* 61:512-517.
- Zhou XA. 1998. The first discussion about center of genetic diversity of cultivated soybean in China. *China Agric Sci* 31:37-48.
- Zhu S, Walker DR, Warrington CV, Parrott WA, All JN, Wood ED y Boerma HR. 2007. Registration of four soybean germplasm lines containing defoliating insect resistance QTLs from PI 229358 introgressed into 'Benning'. *J Plant Reg* 1:162-163.
- Zhu T, Schupp JM, Oliphant A y Keim P. 1994. Hypomethylated sequences: characterization of the duplicate soybean genome. *Mol Gen Genet* 244:638-645.
- Zhukov V, Radutoiu S, Madsen LH, Rychagova T, Ovchinnikova E, Borisov A, Tikhonovich I y Stougaard J. 2008. The Pea Sym37 receptor kinase gene controls infection-thread initiation and nodule development. *Mol. Plant Microbe Interact* 21:1600-1608.

## ANEXO

**Tabla A 1 | Marcadores microsatélites seleccionados para el screening de polimorfismo entre los cultivares NA 5485 RG x A 7053 RG.**

Microsatélite	LG <sup>a</sup>	cM <sup>b</sup>	Polimorfismo	Microsatélite	LG <sup>a</sup>	cM <sup>b</sup>	Polimorfismo
Satt382	A1(5)	26.42	NP	Sat_367	C1(4)	28.04	NP
Satt042	A1(5)	27.66	NP	Sat_337	C1(4)	32.10	NP
Satt385	A1(5)	64.74	NP	Sat_140	C1(4)	41.43	NP
Satt174	A1(5)	88.58	P	Satt578	C1(4)	65.08	NP
Satt236	A1(5)	93.23	NP	Satt646	C1(4)	70.52	P
Satt511	A1(5)	94.20	NP	Satt294	C1(4)	78.65	P
Satt390	A2(8)	9.14	NP	Sat_042	C1(4)	82.51	NP
Sct_067	A2(8)	14.99	NP	Satt195	C1(4)	84.81	P
Satt315	A2(8)	45.29	NP	Sct_191	C1(4)	92.99	NP
Satt632	A2(8)	51.51	NP	Sat_235	C1(4)	94.62	NP
Satt187	A2(8)	54.92	P	Satt524	C1(4)	120.12	NP
GMENOD2B	A2(8)	58.44	NP	Satt164	C1(4)	132.46	P
Satt424	A2(8)	60.59	NP	Satt640	C2(6)	30.47	NP
Satt377	A2(8)	90.84	NP	Satt281	C2(6)	40.30	NP
Satt233	A2(8)	100.09	NP	Satt520	C2(6)	42.37	NA
Satt329	A2(8)	110.94	NP	Satt291	C2(6)	45.76	NP
Sat_232	A2(8)	112.88	P	Satt363	C2(6)	98.07	NP
Satt421	A2(8)	115.93	NP	Sat_076	C2(6)	99.18	NA
Satt209	A2(8)	128.44	NP	Satt202	C2(6)	126.24	NP
Satt455	A2(8)	129.86	NP	Sat_252	C2(6)	127.00	NP
Sat_294	A2(8)	131.97	P	Satt371	C2(6)	145.48	NP
Satt409	A2(8)	145.57	NP	Satt357	C2(6)	151.91	NA
Satt429	A2(8)	162.03	NP	Sat_353	D1a(1)	36.23	NP
Sat_272	B1(11)	14.32	NP	Satt531	D1a(1)	40.87	NA
Satt509	B1(11)	32.51	NP	Satt368	D1a(1)	43.84	NP
Satt251	B1(11)	36.48	NP	Satt321	D1a(1)	50.16	NP
Satt197	B1(11)	46.39	P	Satt169	D1a(1)	50.55	NP
Sct_026	B1(11)	78.13	NP	Satt254	D1a(1)	56.43	NA
Satt415	B1(11)	82.89	NP	Satt580	D1a(1)	62.37	NP
Sat_123	B1(11)	100.88	NP	Sat_106	D1a(1)	66.39	NP
Satt453	B1(11)	123.96	NP	Sat_036	D1a(1)	75.25	NP
Satt577	B2(14)	6.05	NP	Satt408	D1a(1)	106.69	NP
Satt168	B2(14)	55.20	P	Satt147	D1a(1)	108.89	NP
Satt020	B2(14)	72.13	NP	Satt129	D1a(1)	109.67	NP
Satt066	B2(14)	78.84	NP	Sat_096	D1b(2)	0.00	NP
Satt534	B2(14)	87.59	NP	Sat_279	D1b(2)	3.79	NP
Satt560	B2(14)	97.92	P	Satt216	D1b(2)	9.80	P
Satt687	B2(14)	113.61	P	Sat_351	D1b(2)	20.61	NP
Satt565	C1(4)	0.00	NA	BE021153	D1b(2)	30.23	P



Microsatélite	LG <sup>a</sup>	cM <sup>b</sup>	Polimorfismo	Microsatélite	LG	cM	Polimorfismo
Satt157	D1b(2)	37.07	NP	Satt334	F(13)	78.06	NA
Sat_211	D1b(2)	38.04	NP	Satt072	F(13)	87.01	NP
Satt698	D1b(2)	38.04	NA	Satt490	F(13)	97.97	NP
Satt634	D1b(2)	46.62	NA	Sat_074	F(13)	142.35	NP
Satt296	D1b(2)	52.61	P	Satt163	G(18)	0.00	NA
Satt266	D1b(2)	59.61	NP	Satt038	G(18)	1.84	NP
Sat_135	D1b(2)	70.65	NP	Satt309	G(18)	4.53	P
Satt141	D1b(2)	72.89	P	Satt610	G(18)	10.92	NP
Satt005	D1b(2)	75.29	P	Satt217	G(18)	18.25	NP
Satt644	D1b(2)	79.42	NP	Satt235	G(18)	21.89	NP
Satt041	D1b(2)	84.04	NA	Satt130	G(18)	23.10	NP
Sat_139	D1b(2)	93.35	NA	Sat_131	G(18)	31.33	NP
Satt703	D1b(2)	98.75	P	Satt594	G(18)	52.94	NP
Satt172	D1b(2)	100.89	NP	Satt564	G(18)	57.32	NP
Sat_069	D1b(2)	102.60	NA	Sat_088	G(18)	58.00	NA
Staga002	D1b(2)	126.45	NP	Satt199	G(18)	62.16	NA
Sat_289	D1b(2)	131.92	NP	Satt400	G(18)	63.28	NP
Sat_283	D1b(2)	136.66	P	Satt517	G(18)	69.87	NP
Satt271	D1b(2)	137.06	P	Sat_143	G(18)	73.42	NA
Satt458	D2(17)	24.52	P	Satt612	G(18)	80.38	NA
Satt002	D2(17)	47.73	NP	Satt472	G(18)	94.84	NA
Sat_092	D2(17)	57.51	NA	Sat_117	G(18)	100.00	NP
Satt447	D2(17)	66.27	NP	Sct_187	G(18)	107.11	NP
Satt389	D2(17)	79.23	NA	Sat_064	G(18)	108.70	NP
Sat_114	D2(17)	84.18	P	Satt353	H(12)	8.48	P
Satt311	D2(17)	84.62	P	Sat_122	H(12)	61.33	NA
Sat_001	D2(17)	92.12	NA	Satt253	H(12)	67.17	NP
Satt186	D2(17)	105.45	NP	Sat_205	H(12)	68.18	NA
Sat_326	D2(17)	112.85	NA	Satt181	H(12)	91.12	NP
Satt672	D2(17)	114.97	NA	Satt434	H(12)	105.74	P
Satt256	D2(17)	124.31	NP	Satt571	I(20)	18.50	NP
Satt386	D2(17)	125.00	NA	Satt127	I(20)	35.35	NP
Sat_112	E(15)	8.67	NP	Satt239	I(20)	36.94	NP
Sat_107	E(15)	43.05	NA	Satt354	I(20)	46.22	NP
Satt045	E(15)	46.65	NP	Sat_104	I(20)	65.62	NP
Satt553	E(15)	67.92	NP	Satt330	I(20)	77.84	NP
Satt231	E(15)	70.23	NP	Satt148	I(20)	100.78	NP
Satt030	F(13)	3.95	NP	Satt440	I(20)	112.70	NP
Satt269	F(13)	11.37	NP	Satt249	J(16)	11.74	NP
Sat_039	F(13)	27.87	NA	Satt285	J(16)	25.51	NP
Sat_133	F(13)	50.78	NA	Sct_065	J(16)	32.10	P
Satt114	F(13)	63.69	NA	Satt414	J(16)	37.04	P
Sct_033	F(13)	74.13	NA	Satt215	J(16)	44.08	NP

Microsatélite	LG <sup>a</sup>	cM <sup>b</sup>	Polimorfismo	Microsatélite	LG <sup>a</sup>	cM <sup>b</sup>	Polimorfismo
Sct_001	J(16)	44.68	NA	Satt494	M(7)	71.71	NA
Satt620	J(16)	53.71	NP	Satt551	M(7)	95.45	NP
Sat_396	J(16)	69.30	NP	Sat_121	M(7)	103.98	NA
Satt431	J(16)	78.57	P	Satt308	M(7)	130.76	NP
Sat_087	K(9)	4.85	NP	Satt152	N(3)	22.67	P
Satt242	K(9)	14.35	NP	Satt009	N(3)	28.52	P
Satt102	K(9)	30.28	NP	Satt683	N(3)	34.52	NP
Satt046	K(9)	45.59	NA	Sat_084	N(3)	36.86	NA
Satt417	K(9)	46.20	NP	Satt584	N(3)	37.98	P
Satt001	K(9)	50.56	NA	Satt125	N(3)	40.63	NA
Sat_044	K(9)	58.01	NP	Satt080	N(3)	45.14	NP
Sat_043	K(9)	61.67	NP	Satt387	N(3)	53.25	NP
Satt475	K(9)	78.68	NP	Satt660	N(3)	72.60	NA
Satt196	K(9)	104.79	NP	Sat_304	N(3)	77.10	NA
Satt588	K(9)	117.02	NP	Satt312	N(3)	79.86	NP
Satt232	L(19)	10.35	NP	Satt022	N(3)	102.06	NP
Sat_071	L(19)	20.78	NA	Sat_125	N(3)	103.34	NP
Satt462	L(19)	41.00	P	Satt358	O(10)	5.44	NA
Satt156	L(19)	56.14	NP	Sat_132	O(10)	8.75	NA
Satt076	L(19)	61.35	NA	Satt487	O(10)	9.53	NP
Satt006	L(19)	92.00	NP	Satt445	O(10)	20.43	NP
Satt373	L(19)	107.24	NP	Satt259	O(10)	39.82	NP
Satt590	M(7)	7.84	NP	Sat_221	O(10)	51.01	NP
Satt150	M(7)	18.58	NP	Satt094	O(10)	56.58	NA
Sat_316	M(7)	21.00	NP	Satt173	O(10)	58.40	NP
Satt435	M(7)	38.94	P	Satt241	O(10)	59.49	NP
Sat_244	M(7)	48.86	P	Sat_038	O(10)	112.17	NP
Satt245	M(7)	53.54	P	Sat_108	O(10)	129.30	NP
Sat_003	M(7)	62.31	NA				

<sup>a</sup>GL: Grupo de ligamiento, los números entre paréntesis corresponden a los números de cromosomas, <sup>b</sup>cM: centimorgan. Los valores corresponden a la posición de los marcadores en el mapa consenso de soja (Cregan *et al.*, 1999). P: polimórfico, NP: no polimórfico, NA: no presentó amplificación.