

**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Mapeamento de QTLs para caracteres relacionados com a fixação
biológica de nitrogênio (FBN) em soja**

Maria Aparecida dos Santos

**Tese apresentada para obtenção do título de Doutor
em Ciências. Área de concentração: Genética e
Melhoramento de Plantas**

**Piracicaba
2009**

Maria Aparecida dos Santos

Bióloga

Mapeamento de QTLs para caracteres relacionados com a fixação biológica de nitrogênio (FBN) em soja

Orientador:

Prof. Dr. **ISAIAS OLÍVIO GERALDI**

Tese apresentada para obtenção do título de Doutor em Ciências. Área de concentração: Genética e Melhoramento de Plantas

**PIRACICABA
2009**

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP**

Lourenço, Marcus Venicius de Mello

Seleção de leveduras para bioconversão de D-xilose em xilitol / Marcus Venicius de Mello
Lourenço. - - Piracicaba, 2009.
76 p. : il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2009.
Bibliografia.

1. Candida 2. Fermentação aeróbica 3. Leveduras - Isolamento e purificação 4. Sequência do DNA I. Título

CDD 660.62
L892s

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

Aos meus pais Antônio e Amélia e meus irmãos: Francisco, Edneia, Eliane, Antônio, Edson e Rose. De forma especial a minha mãe, por seu amor incondicional e por seu exemplo de coragem, perseverança e humildade.

DEDICO

AGRADECIMENTOS

A Deus pela vida e pelos sonhos que colocou no meu coração;

Aos meus pais pelo carinho e apoio nos tempos difíceis;

A Embrapa soja e ao laboratório de Biotecnologia dos Solos, ambiente onde tive todo o suporte técnico e incentivo para o desenvolvimento desse trabalho;

A Dra. Mariangela Hungria pela co-orientação neste trabalho, pela amizade, paciência e por ser “mãe” muitas vezes;

A Universidade de São Paulo, que por meio da Escola Superior de Agricultura “Luiz de Queiroz” me permitiu realizar este trabalho;

Ao Prof. Dr. Isaias Olívio Geraldi pela orientação na realização do trabalho e pelos conhecimentos transmitidos;

Aos professores e funcionários do Departamento de Genética pela contribuição à minha formação. E em especial ao professor Antonio Augusto Franco Garcia pela contribuição indispensável nas análises e discussão dos dados e a secretária Candida Vanderleia de Oliveira pela prontidão e eficiência na resolução das questões práticas e cumprimento dos prazos;

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de estudos de doutorado;

Aos muitos amigos que fazem parte da história do laboratório de Biotecnologia dos Solos e da minha história - por ordem de tempo de convívio: Ligia, Rinaldo, Leni, Marisa, Pâmela, Dino, Luciana Grange, Everson, Manézino, Jesiane, Alan, Fabinho, Fábio Mostasso, Lílian, Juninho, Fernanda, Tatiana, Rose, Adriana, Letícia, Glaciela,

Odair, Fernando, Renan magrinho, Renan da Nágila, Adalgisa, Luciana, Simone, Ilmara, Renata, Elisete. De maneira especial a Ligia, Rinaldo, Taila, Michele, Nágila e Aline pelo apoio imprescindível na execução do trabalho;

Agradeço a amizade, carinho, compreensão, paciência e a generosidade das minhas amigas “wild type” Glaciela, Adalgisa e Karina que cada uma a seu modo me ajudou a superar momentos difíceis e provavelmente intransponíveis sem o poder curador da amizade.

Aos amigos do laboratório de Genética e Melhoramento de Soja: Colombari, Marco, Márcia, Agnaldo, Guilherme, Larissa e Rosa pelo convívio, dicas importantes e companheirismo;

A todos os colegas do programa de Pós-Graduação em Genética e Melhoramento de Plantas da ESALQ pela convivência e troca de experiências;

Aos três mosqueteiros: Onildo, Carlos Eduardo e Luciélío companheiros fiéis e amigos de todas as horas;

As irmãs do Cenáculo pela ótima convivência e acolhimento ao longo dos dois anos de convivência em Piracicaba;

A dona Maria e dona Cida de Arapoti que me acolheram em suas casas e me abriram a oportunidade de cursar o ensino médio e me preparar para o vestibular. E a dona Maria de Londrina pelo apoio e “comidinhas” prontas durante a redação da tese;

Ao senhor Aristides B. Franco da Luz por acreditar no meu potencial e me dar o apoio que precisava para cursar uma universidade.

Ao Luís Antônio por encher minha vida de alegria e pelo apoio na finalização da tese.

SUMÁRIO

RESUMO.....	9
ABSTRACT.....	11
LISTA DE TABELAS.....	13
LISTA DE FIGURAS.....	15
1 INTRODUÇÃO.....	17
2 DESENVOLVIMENTO.....	21
2.1 Revisão Bibliográfica.....	21
2.1.1 A fixação Biológica de Nitrogênio (FBN) nos Nódulos das leguminosas.....	21
2.1.2 O melhoramento da FBN em Soja.....	24
2.1.2.1 Avaliação da Variabilidade Genética da FBN.....	24
2.1.2.2 Herdabilidade e Correlações dos Caracteres de Nodulação e Fixação de N ₂	27
2.1.2.3 Programas de Melhoramento para FBN em Soja.....	30
2.1.2.4 A FBN na Cultura da Soja no Brasil.....	34
2.1.3 Mapa Genético e Mapeamento de QTLs em Soja.....	35
2.2 Material e Métodos.....	46
2.2.1 Material Vegetal e Análise dos Caracteres Relacionados com a FBN.....	46
2.2.2 Genotipagem e Construção do Mapa de Ligação.....	48
2.2.2.1 Análise com Marcadores Microsatélites.....	48
2.2.2.2 Análise com ESTs (<i>Expressed Sequence Tags</i>).....	51
2.2.2.3 Construção do Mapa de Ligação.....	54
2.2.3 Mapeamento de QTLs.....	55
2.2.3.1 Obtenção dos Limites Críticos (<i>Threshold</i>).....	57
2.3 Resultados e Discussão.....	59
2.3.1 Análise Fenotípica.....	59
2.3.2 Análise de Correlação.....	60
2.3.3 Mapa de Ligação.....	62
2.3.4 Mapeamento de QTLs.....	64
2.3.4.1 Regiões Coincidentes entre QTLs para FBN e outros Caracteres.....	72
2.3.5 Considerações Finais.....	78
3 CONCLUSÕES.....	81
REFERÊNCIAS.....	83
ANEXO.....	95

RESUMO

Mapeamento de QTLs para caracteres relacionados com a fixação biológica de nitrogênio (FBN) em soja

A soja, [*Glycine max* (L.) Merrill] é uma das espécies com maior teor protéico, contendo cerca de 40% de proteína nos grãos. Em consequência disso demanda alta quantidade de nitrogênio (N), o qual pode ser suprido pelo processo de fixação biológica de nitrogênio (FBN), através da simbiose com as bactérias do gênero *Bradyrhizobium*. No Brasil, a FBN é capaz de suprir toda a demanda de N da cultura da soja, dispensando a aplicação de fertilizantes nitrogenados. No entanto, os caracteres relacionados à FBN não têm sido diretamente considerados em programas de melhoramento genético, em função das dificuldades inerentes às avaliações dos mesmos, que requerem a destruição das plantas. O objetivo deste trabalho foi mapear os locos de caracteres quantitativos (*Quantitative Trait Loci*: QTLs) dos caracteres relacionados à FBN, visando identificar associações úteis para a seleção assistida por marcadores, bem como obter outras informações sobre a base genética destes caracteres em soja. Uma população composta de 157 F_{2:7} linhagens endogâmicas recombinantes (*Recombinant Inbred Lines* – RILs), derivada de um cruzamento biparental, foi genotipada com 105 marcadores microssatélites, bem como avaliada para os seguintes caracteres relacionados com a FBN: número de nódulos (NN); peso seco dos nódulos (MNS); peso médio dos nódulos secos (MNS/NN) e peso seco da parte aérea (MPAS). Utilizando o método de mapeamento por intervalo composto para múltiplas características (mCIM) foram mapeados os QTLs para os quatro caracteres. Um mapa genético foi construído com um tamanho estimado em 1.263,2 cM, correspondendo a uma cobertura de 50% do genoma. Oito regiões genômicas foram associadas com os caracteres de FBN. Quatro dessas regiões, localizadas nos grupos de ligação (GL) C1, C2, E e I, foram associadas a mais de um caráter: no GL C1 (Satt190-Satt136) foram mapeados QTLs para MNS, MNS/NN e MPAS; no GL C2 (Satt460-Satt307) foram mapeados QTLs para NN e MNS; no GL E (Satt573-Satt185) foram mapeados QTLs para MPAS e NN; e no GL I (Satt239-Satt354) foram mapeados QTLs para MNS/NN e NN. A associação dos QTLs nos GL C1, C2 e E foi atribuída à pleiotropia, enquanto que no GL I foi atribuída à ligação gênica. Os QTLs influenciando apenas um caráter foram mapeados nos GL A2, B1, G e L: no GL A2 (Sct067-Satt589) foi mapeado um QTL para MNS/NN; no GL B1 (Satt509-Satt251) foi mapeado um QTL para NN; no GL L (Satt232-Satt418) foi mapeado um QTL para MPAS; e no GL G (Satt394-Satt288) foi mapeado um QTL para MPAS. Os QTLs explicaram, individualmente, muito pouco da variação fenotípica ($R^2 = 1,2\%$ a $10,0\%$), sendo o QTL mais significativo mapeado no GL L para MPAS, e explicou $10,0\%$ da variação do caráter, com um efeito aditivo de $0,57 \text{ g planta}^{-1}$. Portanto, foram detectados QTLs em quatro regiões para MPAS (C1, E, G e L), em cinco regiões para NN (B1, C1, C2, E, G); em duas regiões para MNS (C1, C2) e em três regiões para MNS/NN (A2, C1, I) que explicaram $23,0\%$, $20,0\%$, $11,8\%$ e $16,0\%$ da variação fenotípica total, respectivamente. Estes resultados estão de acordo com as herdabilidade relativamente baixas dos caracteres (28% a 49%) e refletem a natureza complexa da FBN, que está sob influência ambiental.

Palavras-chave:

Glycine max; Marcador molecular; Mapeamento de QTL; Fixação biológica de nitrogênio (FBN)

ABSTRACT

Mapping QTLs for traits associated with biological nitrogen fixation (BNF) in soybeans

Soybeans [*Glycine max* (L.) Merrill] is a crop with high protein content (about 40%) in the seeds. As a result, the crop demands high nitrogen (N) inputs, which can be supplied by the process of biological nitrogen fixation (BNF), through the symbiosis with bacteria of the genus *Bradyrhizobium*. In Brazil the BNF allows to fulfill all the demand for N; therefore N fertilizers are not required. However, the traits associated with BNF have not been directly considered in breeding programs due to the difficulties to its evaluation, which require, generally, the plant destruction. The objective of this study was to map Quantitative Trait Loci (QTLs) of traits related to BNF and to identify useful associations for marker-assisted selection, as well as to obtain other information about the genetic basis of these traits in soybeans. A population of 157 F_{2:7} recombinant inbred lines (RILs), derived from a two-way cross, were genotyped with 105 microsatellite markers as well as evaluated for the following traits related with BNF: number of nodes (NN); nodule dry weight (NDW); mean nodule dry weight (NDW/NN); and shoot dry weight (SDW). Using the composite interval mapping for multiple traits (mCIM) method, the QTLs were mapped for all the traits. A genetic map was constructed with an estimated size of 1,263.2 cM, covering about 50% of the genome. Eight genomic regions were associated with the four traits and four of these regions, located on linkage groups (LG) C1, C2, E and I, were associated with more than one trait: in LG C1 (Satt190-Satt136), QTLs for NDW, NDW/NN and SDW were mapped; in LG C2 (Satt460-Satt307), QTLs for NN and NDW were mapped; in LG E (Satt573-Satt185), QTLs for SDW and NN were mapped; and in LG I (Satt239-Satt354) QTLs for NDW/NN and NN were mapped. The pleiotropy was attributed to QTL association in the LG C1, C2, and E, whereas genetic linkage was attributed to QTL association in LG I. QTLs affecting only one trait were mapped in LG A2, B1, G and L: in LG A2 (Sct067-Satt589) a QTL was mapped, for NDW/NN; in LG B1 (Satt509-Satt251) a QTL was mapped for NN; in LG L (Satt232-Satt418) a QTL was mapped for SDW; and in LG G (Satt394-Satt288) a QTL was mapped for SDW. The QTLs individually explained very little of the phenotypic variation ($R^2 = 1.2\%$ to 10.0%), and the most significant QTL was mapped in the LG L, explaining 10.0% of the variation for SDW, with an additive effect of $0.57 \text{ g plant}^{-1}$. Therefore, QTLs in four regions were detected for SDW (C1, E, G, and L), in five regions for NN (B1, C1, C2, E, and G), in two regions for NDW (C1, and C2) and in three regions for SDW/NN (A2, C1, and I), which explained 23.0% , 20.0% , 11.8% and 16.0% of phenotypic variation, respectively. These results are in agreement with the relatively low heritability of the traits (28% to 49%) and reflect the complex nature of BNF traits, which are influenced by the environmental effects.

Key words: *Glycine max*, Molecular marker; QTL mapping; Biological nitrogen fixation (BNF)

LISTA DE TABELAS

Tabela 1 – Marcadores microssatélites polimórficos para os genitores Bossier e Embrapa 20, utilizados para genotipar a população de linhagens.....	50
Tabela 2 – Lista das sequências ESTs e dos marcadores microssatélites dentro de nodulinas conhecidas, que foram utilizados para buscar polimorfismos entre os parentais Bossier e Embrapa 20.....	52
Tabela 3 – Coeficientes de correlação genética (acima da diagonal) e fenotípica (abaixo da diagonal) entre os caracteres relacionados com a FBN.....	61
Tabela 4 – Mapeamento de QTLs (Quantitative Trait Loci) usando o modelo mCIM para os caracteres massa da parte aérea seca (MPAS, g planta ⁻¹), número de nódulos (NN, nódulos planta ⁻¹), massa de nódulos secos (MNS, mg planta ⁻¹) e massa média de nódulos secos (MNS/NN, mg nódulos. ⁻¹).....	66

LISTA DE FIGURAS

- Figura 1 – Diagrama representando os procedimentos adotados para obtenção dos dados de genotipagem. a) testes de polimorfismo entre os genitores Bossier e Embrapa 20 (P1 e P2), para os marcadores microssatélites (M1, M2) e para as etiquetas de seqüências expressas (Expressed sequence tags, ESTs). b) Genotipagem da população de linhagens (L1,...,L7) com os marcadores polimórficos. PM - marcador de peso molecular 1 KB plus DNA..... 49
- Figura 2 – Mapa genético da soja baseado na população de 157 F_{2:7} RILS derivadas dos cultivares Bossier x Embrapa 20. Os grupos de ligação e marcadores microssatélites foram designados de acordo com o mapa consenso de Cregan et al. (1999) e Song et al. (2004). As barras fechadas e pontilhadas representam as regiões mapeadas e as regiões não mapeadas na ação..... 63
- Figura 3 – QTLs significativos para mais de um caráter, identificados pelo método mCIM. As linhas pontilhadas em preto representam os níveis de significância para a análise conjunta e a linha em vermelho o nível de significância estabelecido para a análise individual. MPAS – massa da parte aérea seca, MNS – massa de nódulos secos, NN – número de nódulos, MNS/NN – massa média de nódulos secos..... 68
- Figura 4 – QTLs significativos para um caráter, identificados pelo método mCIM. As linhas pontilhadas em preto representam os níveis de significância para a análise conjunta e a linha em vermelho o nível de significância estabelecido para a análise individual. MPAS – massa da parte aérea seca, MNS – massa de nódulos secos, NN – número de nódulos, MNS/NN – massa média de nódulos secos..... 69
- Figura 5 – Co-localização entre QTLs-FBN e QTLs descritos na literatura. MNS–massa de nódulos secos, NN–número de nódulos, MNS/NN–massa média de nódulos secos, MPAS–massa da parte aérea seca, MTP–massa total da planta, PG–produção de grãos, AP–altura de planta, ER–estádio reprodutivo, PS–peso de 100 sementes, VN–vagens por nó, VP–vagens por planta, DM–dias para maturidade, AC–acamamento, PROT–proteína, OL–óleo, GLI–gliciteína, GEN–genisteína, A_{sat}–taxa de assimilação líquida de CO₂, AF–área foliar, AFE–área foliar específica, EAP–eficiência na aquisição de P, EUP–eficiência na utilização de P, AFP–atividade da enzima fosfatase. Os QTL sublinhados foram mapeados nesse trabalho..... 73

1 INTRODUÇÃO

A fixação biológica de nitrogênio (FBN) juntamente com a fotossíntese são os processos biológicos fundamentais para a manutenção dos ecossistemas terrestres. Apesar de representar 78% da constituição gasosa da atmosfera, o nitrogênio molecular (N_2) não é absorvido pelos eucariotos, que não dispõem de um sistema enzimático para quebrar a tripla ligação entre os átomos de nitrogênio. Os microrganismos procariotos que possuem o complexo enzimático da nitrogenase são os únicos capazes de assimilar o N_2 e convertê-lo em produtos assimiláveis pelos demais organismos. Esses microrganismos também são denominados como diazotrofos ou diazotróficos e o processo como FBN.

Os microrganismos diazotrofos podem ser de vida livre ou associativa, sendo que a associação mais estudada é a relação simbiótica que se estabelece entre as bactérias denominadas coletivamente como rizóbios, e as raízes das leguminosas. Estima-se que cerca de 90×10^9 kg de N_2 são fixados simbioticamente por ano, principalmente por *Rhizobium*, *Bradyrhizobium* e *Sinorhizobium* (= *Ensifer*), nas raízes de leguminosas cultivadas (GRAHAM e VANCE, 2000). Dentre as simbioses mais estudadas e de maior relevância para a agricultura, destaca-se a associação entre a soja (*Glycine max* (L.) Merrill) e bactérias das espécies *Bradyrhizobium japonicum* e *B. elkanii*.

No Brasil, o cultivo de soja em escala comercial iniciou-se na década de 1940, no Rio Grande do Sul e, a partir da década de 1960 houve um incremento na produção, quando a soja passou a ser cultivada em rotação com o trigo. Atualmente, a soja é cultivada em quase todo o País, com destaque para a região dos Cerrados, um ecossistema que abrange 207 milhões de hectares (ha). No cenário atual, o Brasil é o segundo maior produtor de grãos desta leguminosa, e a FBN contribui, significativamente, para a sua produção, pois toda a demanda de N da cultura é suprida, o que representa uma economia estimada, atualmente, em US\$ 6,6 bilhões por safra (HUNGRIA et al., 2007).

Além da importância econômica, a FBN tem um papel fundamental para a sustentabilidade dos sistemas agrícolas, uma vez que o N proveniente desta fonte é usado diretamente pela planta, sendo, portanto, menos susceptível à volatilização,

desnitrificação e lixiviação. Os esforços dos pesquisadores para potencializar a FBN compreendem a seleção de estirpes de rizóbios mais eficientes, a adoção de manejos mais adequados à simbiose, como o sistema de plantio direto e o melhoramento da capacidade de fixar N_2 das leguminosas cultivadas (RENGEL, 2002).

Alguns programas de melhoramento específicos para a FBN em soja foram iniciados na década de 1980 na Austrália, nos Estados Unidos e na África. Esses programas, apesar de obterem êxito na seleção de materiais com melhor potencial de FBN, com exceção da soja com nodulação promíscua na África, não liberaram nenhum material no mercado. Os principais fatores que limitaram o êxito destes programas foram a menor produtividade dos materiais desenvolvidos quando comparados com materiais dos programas principais e a susceptibilidade a doenças. A experiência desses trabalhos mostra que o melhoramento para a FBN precisa estar integrado ao programa principal de melhoramento (HERRIDGE e ROSE, 2000).

A avaliação dos parâmetros relacionados com a nodulação e com a FBN é trabalhosa e envolve a destruição de plantas, o que dificulta a avaliação destes parâmetros pelos programas de melhoramento que trabalham com um grande número de genótipos e selecionam para múltiplas características. Nesse contexto, a seleção assistida por marcadores caracteriza-se como uma estratégia útil para selecionar os caracteres relacionados com a FBN.

A disponibilidade de marcadores moleculares e o desenvolvimento de métodos estatísticos de análise permitiram a construção de mapas genéticos e o mapeamento de QTL para características de interesse agrônomo nas principais espécies cultivadas. Para a soja, diversos trabalhos relatam o mapeamento de QTLs para teor de proteína nos grãos, teor de óleo, resistência a doenças, produtividade, resistência a insetos, teor de isoflavonas nas sementes, assimilação de nutrientes, entre outros (SOYBASE, 2009).

Poucas informações sobre QTLs para FBN estão disponíveis na literatura. Alguns trabalhos foram conduzidos com feijão comum (*Phaseolus vulgaris*) e soja. Em feijão comum são descritos oito QTLs para número de nódulos (NN), mapeados em duas populações derivadas do cruzamento entre os cultivares Bat-93 x Jalo EEP558. Os QTLs relatados foram individualmente de pequeno efeito e explicaram em média 35,6%

da variação fenotípica total. O QTL que explicou a maior parte da variação ($10\% < R^2 < 18\%$) foi mapeado no cromossomo 7 também associado a resistência a *Xanthomonas axopodis* pv *phaseoli* (NODARI et al., 1993; TSAI et al., 1998 e SOUZA et al., 2000).

Em soja, Tanya et al. (2005) mapearam QTLs para caracteres relacionados com a FBN: massa de nódulos secos, (MNS); massa de nódulos frescos (MNF); número de nódulos (NN); massa da parte aérea seca (MPAS) e atividade da nitrogenase em 136 linhagens endogâmicas recombinantes (*Recombinant Inbred Lines*, RILs) derivadas dos cultivares SJ2 (Tailandês) e Suwon157 (Coreano), nos grupos de ligação (GL) A1, I, K, J, D1b e O. Todas as regiões foram significativas para mais de um dos caracteres avaliados e os QTLs mapeados foram, em geral, de pequeno efeito ($R^2 < 8,3\%$), com exceção dos QTLs para MNS, MNF, NN, MPAS e atividade da nitrogenase, mapeados no GL O ($12,3\% < R^2 < 20,8\%$). Nicolás, Hungria e Arias (2006) mapearam QTLs para FBN em A2, B1, B2, C2, H, F, todos de pequeno efeito ($R^2 = 2\%$ a 9%) e quatro regiões foram associadas a mais de um caráter. Algumas das associações descritas por Nicolás, Hungria e Arias (2006) foram verificadas por Santos, Nicolás e Hungria (2006) em outra população composta por 157 $F_{2:7}$ RILs oriundas do cruzamento entre os cultivares Bossier e Embrapa 20. Nenhuma das regiões descritas por Tanya et al. (2005) foi identificada nos dois trabalhos posteriores e isso deve-se em parte à baixa cobertura do genoma, pois um número reduzido de marcadores foi utilizado nesses trabalhos. A saturação dos mapas iniciais desenvolvidos por Nicolás, Hungria e Arias (2006) e Santos, Nicolás e Hungria (2006) faz-se necessária e pode permitir a identificação de novos QTL associados com a FBN, a melhor delimitação da localização e dos efeitos dos QTLs identificados e uma melhor compreensão das causas genéticas da co-localização dos caracteres de FBN (ligação gênica ou pleiotropia). Nesse contexto, os objetivos deste trabalho, com a cultura da soja, compreenderam: *i*) saturar o mapa genético de ligação da população descrita no trabalho de Santos, Nicolás e Hungria (2006) com marcadores microssatélites; *ii*) incorporar marcadores funcionais, a partir da seleção de sequências gênicas expressas sob condições de nodulação; *iii*) mapear QTLs para caracteres relacionados com FBN.

2 DESENVOLVIMENTO

2.1 Revisão Bibliográfica

2.1.1 A fixação Biológica de Nitrogênio (FBN) nos Nódulos das Leguminosas

Uma das simbioses mutualísticas mais estudadas é a que se estabelece entre as leguminosas e bactérias do solo, genericamente denominadas como rizóbios. As bactérias reduzem o N_2 atmosférico à amônia, que é assimilada pela planta. Em troca, a planta fornece compostos carbônicos para as bactérias. O estabelecimento da simbiose é o resultado da expressão coordenada de genes da planta e da bactéria, envolvidos no surgimento e na regulação de um novo órgão nas raízes das plantas, o nódulo, dentro do qual ocorrem as trocas de C e N entre a planta e a bactéria.

A interação planta-bactéria inicia-se com a ativação dos genes de nodulação (genes *nod*, *noe* e *nol*) nos rizóbios por isoflavonóides exsudados das raízes das leguminosas. Os genes de nodulação estão envolvidos com a síntese de outra molécula sinal, os fatores de nodulação (Fator Nod). A liberação dos fatores de nodulação pelos rizóbios induz alterações morfológicas nas raízes, como o encurvamento dos pêlos radiculares, a alcalinização do meio extracelular, oscilações na concentração de Ca^+ , a formação do cordão de infecção e a multiplicação das células corticais que formam os primórdios nodulares. O cordão de infecção cresce até o primórdio nodular e a bactéria é entregue no citoplasma da célula hospedeira em um envelope derivado da membrana da planta (membrana peribacteróide – PBM). O primórdio nodular se desenvolve em nódulo maduro abrigando a bactéria diferenciada em sua forma endossimbiótica, conhecida como bacteróide. Bacteróides juntos com a PBM são chamados simbiossoma. Neste estágio, a bactéria sintetiza a enzima nitrogenase que catalisa a fixação de N_2 , o produto da fixação a amônia é exportada para os outros tecidos da planta (SCHULTZE e KONDOROSI, 1998).

Todos os passos do desenvolvimento do nódulo envolvem a expressão de genes específicos ou com expressão diferenciada. A obtenção de mutantes de *Lotus japonicus* e *Medicago truncatula* defectivos quanto aos processos iniciais da nodulação

(encurvamento do pêlo radicular, formação do cordão de infecção, formação do primórdio nodular), bem como o desenvolvimento de mapas genéticos e físicos para essas duas leguminosas, consideradas como modelos para estudos de FBN, têm permitido a clonagem de genes envolvidos com o processo de FBN. O primeiro gene de leguminosa associado com a nodulação clonado em *L. japonicus* foi o gene *nin-1* (*non-nodulating mutant*), que codifica um fator de transcrição envolvido na regulação da expressão gênica durante a formação do cordão de infecção (SCHAUSER et al., 1999). Genes que codificam os receptores da planta responsáveis pela percepção do Fator Nod foram clonados em *L. japonicus*, LjNFR1 e LjNFR5 (*nod-factor receptor*) e em *M. truncatula* MtLYK3 e MtLYK4 (*lysine motif-receptor-like kinase*) (MADSEN et al., 2003; RADUTOIU et al., 2003; LIMPENS et al., 2003). Genes ativados imediatamente após a percepção do fator Nod, denominados DMI1, DMI2 e DMI3 (*doesn't make infections*) e NSP1 e NSP2 (*nodulation signaling pathway*), também foram isolados em mutantes de *M. truncatula* (CATOIRA et al., 2000). Estes genes fazem parte da via de transdução do sinal que desencadeia a expressão de outros genes envolvidos na formação do nódulo.

Em soja, estudos com mutantes que apresentam alterações nos padrões de nodulação (número e distribuição dos nódulos nas raízes), denominados supernodulantes, levaram à clonagem do gene *nark* (*nodule autoregulation receptor kinase*), que codifica uma proteína receptora do tipo quinase. Essas proteínas receptoras possuem um domínio extracelular rico em leucina e um domínio quinase citoplasmático, sendo algumas de suas funções a recepção de hormônios, sensores de elicitores de fungos e a ativação do desenvolvimento da parte aérea da planta e de meristemas florais e apicais (SEARLE et al., 2003). Um gene similar foi clonado em *L. japonicus* denominado *har1* (*hypernodulation and altered root*) (NISHIMURA et al., 2002). *Gmnark* e *Ljhar1* definem o mesmo gene e estão envolvidos no controle da proliferação celular do meristema nodular. Um modelo proposto para o controle da nodulação considera um mecanismo de auto-regulação onde os primórdios nodulares que começam a ser formados logo após a percepção do fator Nod produzem um sinal, o qual é percebido nas folhas, que por sua vez produzem outro sinal, que atua como inibidor do desenvolvimento de novos nódulos (GRESSHOFF, 2003).

Os genes da planta cuja expressão é induzida pelo fator Nod são denominados de nodulinas, pois as proteínas que codificam são requeridas no desenvolvimento nodular e na manutenção da FBN. De acordo com o padrão de expressão, as nodulinas são classificadas em nodulinas precoces ou em nodulinas tardias. As nodulinas precoces são expressas imediatamente após o contato com o rizóbio e estão envolvidas com a formação da estrutura nodular (divisão das células do córtex, síntese do cordão de infecção). As nodulinas tardias são expressas em nódulos maduros e desempenham um papel no metabolismo, na manutenção ou senescência dos nódulos (STOUGAARD, 2000).

Muitas nodulinas iniciais são proteínas da parede celular, ressaltando o papel da constituição bioquímica da parede durante a infecção e desenvolvimento do nódulo. Exemplos incluem genes que codificam para proteínas ricas em prolina, Enod2, Enod5, Enod10, Enod11, Enod12, PRP4, bem como extensinas e proteínas ricas em glicina (GRPs) (SCHULTZE e KONDOROSI, 1998).

A nodulina tardia mais abundante nos nódulos ativos é a leghemoglobina, cuja função é a de facilitar a difusão de O₂ na região infectada do nódulo. Outras nodulinas tardias, como a glutamina sintetase (GS) e a glutamato sintase (GOGAT), atuam na assimilação do nitrogênio fixado (NH₃), produzindo glutamina e glutamato, respectivamente. Esses dois compostos nitrogenados são utilizados na síntese de amidas e ureídeos, que são exportados para outros órgãos da planta. Muitas leguminosas de clima temperado, tais como as dos gêneros *Lupinus*, *Medicago*, *Pisum*, *Trifolium* e *Vicia*, exportam os aminoácidos asparagina e/ou glutamina. Já as leguminosas tropicais dos gêneros *Glycine*, *Phaseolus* e *Vigna* transportam principalmente os ureídeos, incluindo alantoina e ácido alantóico (TAJIMA, NOMURA e KOUCHI, 2004).

A construção de banco de dados de ESTs (*Expressed Sequence Tags*) e o uso de técnicas de expressão gênica como microarranjos de cDNA permitem a identificação de padrões globais de expressão gênica durante processos específicos do desenvolvimento das plantas. ESTs específicas de nódulos foram identificadas em *M. truncatula* e *L. japonicus* (FEDOROVA et al., 2002; JOURNET et al., 2002). O emprego de ESTs no estudo da simbiose em soja também tem possibilitado a identificação de

genes diferencialmente expressos durante a formação do nódulo. Lee et al. (2004) identificaram 3.264 ESTs expressos exclusivamente nas células dos nódulos de soja, incluindo genes envolvidos no metabolismo do C, do N, de vitaminas e de cofatores, no metabolismo secundário, no controle do ciclo celular e na expressão de nodulinas conhecidas. Jeong et al. (2006) realizaram um *screening* nos bancos de dados de EST de soja para identificar genes específicos dos nódulos (perfis de análise de expressão *in silico*). Por esta abordagem, os autores identificaram 156 ESTs exclusivamente expressas nos nódulos de soja e mapearam duas nodulinas: *GmPGN* (poligalacturonase específica de nódulo) e *GmEKN* (E-ácido glutâmico, K-lisina, N-nodulina) nos grupos de ligação B1 e A2, respectivamente.

Estudos de expressão gênica globais permitem a identificação de genes que são induzidos ou reprimidos em uma dada condição experimental como nos diferentes estágios de desenvolvimento dos nódulos, permitindo propor possíveis papéis para estes genes de acordo com as informações disponíveis nos bancos de dados públicos, como o do National Center for Biotechnology Information - NCBI. Os trabalhos de geração e caracterização de mutantes aliados à clonagem baseada em mapeamento permitem a identificação e o estudo da função de genes específicos. Outra abordagem interessante a ser explorada é a da associação de polimorfismos naturais desses genes com a variação quantitativa da capacidade de nodulação e fixação observada nas leguminosas.

2.1.2 O Melhoramento da FBN em Soja

2.1.2.1 Avaliação da Variabilidade Genética da FBN

Os caracteres fenotípicos normalmente avaliados em estudos de FBN referem-se à nodulação (massa e número de nódulos) e à capacidade de assimilação de N pela planta (N₂ fixado pelo processo biológico). Estudos sobre a genética da fixação biológica em soja foram intensificados a partir da década de 1980 e a variação genética natural quanto à FBN em soja foi avaliada em diferentes condições ambientais, como o

excesso de nitrato no solo, o estresse hídrico e interações planta-bactéria mais eficientes.

Um dos fatores ambientais que afetam negativamente a FBN é a presença de elevados níveis de nitrato no solo na fase inicial de desenvolvimento da planta, acarretando em queda da produtividade caso as duas fontes de N (N mineral e a FBN) não forem suficientes para suprir a demanda da planta. Para discriminar genótipos tolerantes ao nitrato, Betts e Herridge (1987) avaliaram uma coleção de 489 genótipos cultivados na presença de alta concentração de nitrato. Após duas etapas de avaliações em casa de vegetação em que as plantas foram cultivadas em vaso contendo areia e solução nutritiva rica em nitrato (2,5 M de nitrato) foram selecionados 66 genótipos que apresentam boa nodulação e fixação de N_2 nos tratamentos com alta dose de nitrato. Em uma terceira avaliação realizada em condições de campo, foram selecionados quatro genótipos de origem coreana (K464, K466, K468, K469). Diferenças quanto à capacidade de FBN na presença de nitrato também foram identificadas entre cultivares da época, por exemplo, Gibson e Harper (1985) identificaram os cultivares Elf e Avoyelles como os mais tolerantes ao nitrato entre doze cultivares avaliados.

Outro fator ambiental que afeta sensivelmente a FBN é o estresse hídrico. Leguminosas como a soja, nas quais o N_2 fixado é translocado para a parte aérea sob a forma de ureídos (alantoina e ácido alantoico) são particularmente sensíveis a níveis moderados de dessecação do solo (SERRAJ; SINCLAIR e PURCELL, 1999). Sall e Sinclair (1991) observaram que o cultivar Jackson, lançado em 1953, apresenta melhor desempenho da FBN sob condições de estresse hídrico. Essa observação levou Sinclair et al. (2000) a realizarem um *screening* de mais de 3000 PIs (*Plant Introduction*) para identificar linhas com FBN tolerante ao estresse hídrico; essa avaliação resultou na identificação de oito PIs (PI222547, PI2277557, PI374163, PI423886, PI429328, PI507039, PI507414 e PI578315B). Estudos sobre a base dessa tolerância relatam que em condições de estresse hídrico ocorre um aumento considerável na concentração de ureídos na parte aérea dos materiais susceptíveis, o que não ocorre com os materiais tolerantes e que esse acúmulo inibe a atividade dos nódulos. Uma maior eficiência no catabolismo de ureídos é proposta como a possível causa da tolerância. Duas enzimas estão envolvidas com o catabolismo de ácido alantoico em soja, a enzima alantoato

amidinohidrolase e a enzima alantoato amidohidrolase, essa última dependente de Mn como cofator. Resultados experimentais sobre a sensibilidade dos genótipos susceptíveis e tolerantes ao estresse hídrico mostram que nos genótipos susceptíveis os níveis de ureídos nas folhas diminuem com os tratamentos com Mn, enquanto que os genótipos tolerantes são insensíveis ao Mn. A conclusão desses estudos é que nos genótipos tolerantes a alantoato amidinohidrolase seria a principal enzima responsável pelo catabolismo de ureídos e que essa variação natural é percebida na presença do estresse hídrico. (VADEZ e SINCLAIR, 2002; SINCLAIR, VADEZ e CHENU, 2003).

Incrementos na proporção de N proveniente da FBN dependem também do genótipo das estirpes bacterianas e diferenças quantitativas devem-se, em grande parte, mais a interações planta-bactéria do que aos efeitos genéticos atribuídos individualmente aos genótipos das plantas ou dos rizóbios (MYTTON, 1984). Duas abordagens constituem exemplos de como a interação planta-bactéria pode ser explorada para potencializar a FBN: uma primeira abordagem visa identificar interações específicas e mais eficientes entre cultivares e estirpes bacterianas veiculadas por meio de inoculantes; a segunda abordagem, ao contrário, configura-se pela busca por genótipos capazes de formar nódulos com uma gama de rizóbios do solo.

A presença de uma população de *Bradyrhizobium* do sorogrupo USDA 123, altamente competitiva e estabelecida nos solos americanos constitui um limitante na resposta à inoculação com estirpes mais eficientes dos sorogrupos USDA 110 e USDA 138; mesmo com altas taxas de inoculação a proporção de nódulos formados pelos sorogrupos veiculados pelos inoculantes é, geralmente, inferior a 20%. Cregan e Keysey (1986) relatam um trabalho com 1278 acessos do banco de germoplasma americano incluindo 1213 PIs e 65 cultivares avaliados quanto à capacidade de restringir a nodulação com o sorogrupo 123. Treze acessos apresentaram restrição da nodulação com o sorogrupo 123 (número de nódulos inferior a dez) e dentre esses materiais se destacaram as PIs PI371607 e PI377578. Em ensaios de competição em que os sorogrupos USDA 123 e USDA 138 competem simultaneamente pela ocupação dos nódulos, a PI 371607 e a PI377587 nodularam preferencialmente com o sorogrupo USDA 138, que ocupou mais de 90% dos nódulos.

A utilização de genótipos capazes de formar nódulos com uma gama de rizóbios do solo é uma alternativa interessante para regiões onde não existe produção de inoculantes com qualidade e quantidade suficientes. Pesquisadores e agricultores africanos observaram que o cultivar Hernon 147 formava nódulos efetivos em solos onde não havia histórico de produção de soja. Investigações subsequentes levaram à identificação de 55 genótipos com capacidade de nodulação com rizóbios nativos. Esses materiais são denominados de nodulação promiscua e a maioria desses cultivares tem em seu *background* genético os cultivares Gilbert ou Hernon (MPEPEREKI et al. 2000).

2.1.2.2 Herdabilidade e Correlações dos Caracteres de Nodulação e Fixação de N₂

A compreensão dos fatores genéticos e ambientais que afetam os caracteres morfológicos associados com a fixação de N₂ é fundamental para a seleção visando aumento da FBN. Alguns trabalhos sobre a genética da FBN relatam estimativas de herdabilidade e correlações para os caracteres de nodulação e fixação de N₂. Ronis, Sammons e Kenwoathy (1985) estimaram a herdabilidade no sentido amplo do total de N₂ fixado nas sementes em três populações F₂, cada uma composta por 110 plantas (utilizando N marcado - N¹⁵) e os valores encontrados variaram de 53% a 60%. Herridge e Rose (1994) estimaram a herdabilidade no sentido amplo do N₂ fixado nos tecidos (determinação de ureídos no xilema das plantas) de onze populações F₂ e as estimativas variaram de 0% a 31%. Nesse mesmo estudo foram determinados os valores de herdabilidade das gerações F_{3:5} e F_{3:6} que variaram de 32% a 52%. Outros trabalhos estudaram a herdabilidade da nodulação. Greder, Orf e Lambert (1986) avaliaram três populações F_{3:5} e F_{3:6}, em três locais, quanto à capacidade de nodulação (massa de nódulos) e as estimativas de herdabilidade variaram de 0% a 66%. Em outro trabalho, visando explorar a capacidade de nodulação precoce, Pazdernik, Graham e Orf (1996) relatam valores de herdabilidade altos, de 79% e 82%, para massa de nodulos frescos em duas populações F₃ e F₄ avaliadas em casa de vegetação 17 dias após o plantio. Em um estudo sobre a genética da nodulação, Nicolás, Hungria e Arias (2002) avaliaram a herdabilidade dos caracteres de nodulação: massa de nódulos secos

(MNS), número de nódulos (NN), massa média de nódulos secos (MNS/NN) e de crescimento da planta: massa da parte aérea seca (MPAS) em quatro populações F_3 derivadas do cruzamento entre cultivares brasileiros. Os valores de herdabilidade variaram de 13% a 58% para NN, de 25% a 70% para MNS, de 20% a 47% para MNS/NN e de 53% a 68% para MPAS.

Estimativas sobre a correlação entre os caracteres de nodulação e o N_2 fixado nos tecidos das plantas, indicam que esses caracteres são correlacionados positivamente. Em seis experimentos (dois locais e três anos) realizados por Sinclair et al. (1991) foram estimados coeficientes de correlações altos entre MNS e NN ($r_F = 0,84$ a $0,91$) e correlações medianas a altas para os caracteres de nodulação e crescimento da planta (massa de raiz e da parte aérea), variando de $0,31$ a $0,67$ para massa da raiz e $0,40$ e $0,70$ para massa da parte aérea. Em um experimento com 152 cultivares avaliados em casa de vegetação em condições controladas, na ausência de N mineral e com inoculação com estirpes eficientes, Bohrer e Hungria (1998) verificaram que, entre os caracteres de nodulação, a MNS apresenta melhor correlação com o nitrogênio total da planta ($0,80$ a $0,86$), enquanto que para o NN foram inferiores ($0,34$ a $0,58$). Os autores observaram, também, correlações altas entre a MPAS e nitrogênio total da planta ($r_F = 0,87$ a $0,94$). Resultados semelhantes foram relatados por Hungria e Bohrer (2000) ao avaliarem os mesmos genótipos em solo contendo uma população naturalizada de rizóbios. Souza et al. (2008) em uma rede de ensaios a campo realizados em seis Estados brasileiros estimaram correlações altas entre MNS e NN ($r = 0,91$) e entre MPAS e o nitrogênio total da planta ($r = 0,88$); esses autores sugerem que, em solos pobres em N e na ausência de fertilização nitrogenada, os caracteres MNS e MPAS constituem boas estimativas da FBN.

Alguns estudos relacionam a FBN e a produção de grãos. Ronis et al. (1985) trabalharam com dois cruzamentos, um entre genitores contrastantes quanto à capacidade de FBN Ware (alta capacidade de FBN) x D66-5566 (baixa capacidade de FBN) e o outro entre genitores intermediários (Willians x Calland). As populações F_2 foram avaliadas quanto à massa de N_2 fixado nas sementes, massa total da planta e produção de grãos. As linhas com maior FBN também apresentaram maior massa total da planta e maior produção de grãos. A massa de N_2 fixado nas sementes apresentou

uma correlação alta com produção de grãos ($r = 0,99$). Greder, Orf e Lambert (1986) estimaram correlações altas ($r_F = 0,85$ a $0,96$) entre MNS e produção de grãos em três populações $F_{3:5}$ e $F_{3:6}$. Burias e Planchon (1992) determinaram o N_2 fixado sob condições controladas em casa de vegetação em três populações F_2 (210 plantas) e realizaram uma seleção divergente com base nos dados de FBN selecionando 11% das melhores e 11% das piores plantas F_2 quanto à capacidade de FBN, resultando na seleção de 43 plantas divididas em dois grupos. Duas sementes de cada planta F_2 selecionada foram usadas para obter as linhas F_4 que foram avaliadas em condições de campo quanto à produtividade. Foram encontrados valores de correlação entre produtividade e quantidade de N_2 fixado ($r_F = 0,39$ a $0,60$), produtividade e MNS ($0,40$ a $0,52$) e produtividade e volume dos nódulos ($0,44$ a $0,58$). A população foi avançada até a geração F_6 pelo método SSD (*Single Seed Descent*) e novas avaliações foram feitas quanto à FBN e produtividade nas linhas F_6 oriundas da seleção divergente em F_2 . As linhas F_6 resultantes da seleção positiva em F_2 apresentaram um acréscimo de 20% a 30% na FBN em relação aos genitores e as linhas F_6 oriundas da seleção negativa tiveram um decréscimo de 16% a 47% em relação aos genitores. Os coeficientes de variação fenotípica entre os caracteres de fixação de N_2 e nodulação, avaliados em F_2 e a produtividade em F_6 variaram de ($r_F = 0,39$ a $0,68$) para fixação de N_2 , para os parâmetros de nodulação, as correlações foram positivas mas não significativas.

Coale, Meisinger e Wiebold (1985) realizaram um estudo visando investigar os efeitos do melhoramento sobre a FBN. Para tanto, avaliaram a produção de matéria seca, porcentagem de N nos grãos, N total, N proveniente da fixação e porcentagem do N total oriundo da FBN de três cultivares Lincoln (Mandarin x Manchu), lançada em 1943; Shelby (Lincoln² x Richland), lançada em 1958; e Williams (((Lincoln² x Richland) x L49-4091) x L57-0034) lançada em 1972. As avaliações foram feitas em cinco estádios de desenvolvimento (V2, V7, R2, R5 e R7) e na colheita dos grãos. Para todos os caracteres, as diferenças entre os cultivares só foram evidentes a partir da maturidade fisiológica (R7), quando o cultivar Williams apresentou valores superiores e o cultivar Lincoln apresentou os menores valores para todos os parâmetros avaliados. O maior acúmulo de N no cultivar Williams em R7 deveu-se, em parte, ao maior acúmulo de N proveniente da FBN e também a maior capacidade de assimilar o N

mineral. Portanto, o melhoramento para produtividade promoveu tanto a FBN quanto a assimilação de N mineral. Segundo os autores, a melhor nutrição em N do cultivar Williams no início da maturação teria efeito direto sobre a concentração de N nos grãos e na produtividade, pelo aumento do período de enchimento de grãos e pela promoção da fotossíntese durante o enchimento de grãos.

2.1.2.3 Programas de Melhoramento para FBN em Soja

Alguns programas específicos para o melhoramento da FBN em soja foram desenvolvidos na Austrália, na África e nos EUA. O programa australiano explorou a variação natural de tolerância ao nitrato. O programa compreendeu dois ciclos de seleção entre 1986 e 1997. No primeiro ciclo foram feitos cruzamentos entre linhagens coreanas tolerantes ao nitrato (K464, K466, K468 e K469) e cultivares adaptadas (Valder, Reynolds, Forrest e Bossier). Ao todo, 849 plantas F_2 foram avaliadas para fixação de N_2 em solo rico em nitrato e as plantas F_2 que apresentaram mais de 31% do N oriundo da FBN foram selecionadas para a obtenção das gerações $F_{2:6}$ e $F_{2:7}$, quando novas avaliações foram realizadas para determinar a FBN e produção de grãos. O desempenho dessas linhagens foi comparado com o cultivar Forrest (30% de N derivado da FBN e $2,6 \text{ t ha}^{-1}$) e a linhagem coreana K478 (50% de N derivado da FBN e $0,95 \text{ t ha}^{-1}$). Com base nos resultados obtidos foram selecionadas quatro linhagens (A82-3, D22-8, K78-1 e E68-5), que apresentaram desempenhos semelhantes ao da linhagem coreana para FBN (50% do N derivado da FBN) e uma produtividade que variou de $1,7 \text{ t ha}^{-1}$ a $2,67 \text{ t ha}^{-1}$.

No segundo ciclo, as quatro linhagens selecionadas foram cruzadas com cultivares mais produtivos (Lachlan e Valiant) e o método SSD foi utilizado para obtenção de linhagens $F_{2:7}$, $F_{2:8}$ e $F_{2:9}$. Esses materiais foram avaliados em conjunto com as linhas do programa principal e o desempenho dos materiais do programa de FBN, em termos de produtividade superou o dos genitores Lachlan e Valiant, entretanto, os materiais não superaram a produtividade de linhagens do programa principal de melhoramento australiano. Nenhum dos materiais foi lançado como cultivar,

pois apresentaram susceptibilidade à *Phytophthora sojae* e o programa foi encerrado em 1998 (HERRIDGE; TURPIN e ROBERTSON, 2001; HERRIDGE e ROSE, 2000).

Outra abordagem para o desenvolvimento de materiais tolerantes ao nitrato consistiu na geração de mutantes *nts* (*nitrate-tolerant symbiotic*). São descritos mutantes dos cultivares Bragg (CARROL; McNEIL e GRESSHOFF, 1985), Williams (GREMAUD e HARPER, 1989) e Enrei (AKAO e KOUCHI, 1992). Esses mutantes *nts* apresentam cerca de dez vezes mais nódulos do que os tipos selvagens e altas taxas de FBN, tanto na ausência, quanto na presença de nitrato. Contudo, apresentam uma redução de 30% a 40% da biomassa vegetal total produzida e um crescimento restrito das raízes, comparados com os genótipos genitores (GREMAUD e HARPER, 1989).

Estudos sobre a herança do fenótipo *nts* têm verificado a natureza recessiva desse fenótipo e testes de alelismo indicam que todas as mutações ocorreram no mesmo gene, sugerindo que este gene é particularmente susceptível à mutagênese química (VUONG e HARPER, 2000). Experimentos realizados na Austrália e nos Estados Unidos envolvendo plantas mutantes dos cultivares Bragg e Williams, evidenciaram um decréscimo de 20% a 41% na produtividade em relação aos genitores (HERRIDGE, TURPIN e ROBERTSON, 2001). Um efeito positivo observado nos campos cultivados com os mutantes *nts* foi o aumento de produtividade de cereais cultivados logo após a colheita da soja. Song et al. (1995) atribuem esse efeito a dois fatores: à menor absorção de N mineral do solo pelas plantas *nts* e ao aumento de N no solo devido à decomposição de resíduos vegetais dessas plantas contendo altos teores de N.

Sinclair et al. (2007) relatam os progressos obtidos em um programa iniciado em 1993, com o objetivo de explorar a característica de fixação de N₂ tolerante à dessecação do solo da cultivar Jackson, visando o desenvolvimento de linhagens mais produtivas em ambientes secos. O programa compreendeu quatro fases: cruzamento e seleção de progênies F₃; *screening* preliminar em casa de vegetação para seleção de linhagens F_{3:4} com maior FBN quando submetidas ao estresse hídrico; avaliações de campo para identificar genótipos com maior produtividade em ambientes pouco produtivos (ambientes secos) e experimentos em casa de vegetação para comprovar a tolerância da FBN ao estresse hídrico nas linhagens mais produtivas. O cruzamento

entre o cultivar Jackson e a linhagem KS4895 (linhagem altamente produtiva e de adaptação ampla) foi identificado como o mais promissor. As progênies F_3 desse cruzamento foram submetidas a ensaios de campo e avaliadas para maturidade e vigor da planta. Com base nesses ensaios, foram selecionadas cem linhagens $F_{3:4}$ pertencentes ao mesmo grupo de maturação (grupo V). As cem linhagens $F_{3:4}$ foram submetidas a estresse hídrico em casa de vegetação, coletadas e avaliadas quanto a concentração de N nos tecidos. A seguir, as linhagens $F_{3:4}$ que apresentaram maiores níveis de FBN sob condições de seca foram submetidas à avaliações de campo. Após sete anos de avaliações em seis locais diferentes, foram selecionadas duas linhagens (R01-416F e R01-581F) mais produtivas que os cultivares (Hutcheson, Ozark e Dillon) usados como testemunhas em ambientes de produtividade média (ambientes secos). Finalmente, para comprovar que a FBN tolerante ao estresse hídrico foi o fator responsável pela produtividade dessas linhagens, foram feitas determinações da FBN em casa de vegetação em condições de estresse hídrico. As linhagens R01-416F e R01-581F apresentaram uma atividade de FBN superior (25% e 15%) comparados com o do cultivar genitor KS4895.

Incrementos na proporção de N proveniente da FBN pelas leguminosas dependem também do genótipo das estirpes bacterianas e diferenças quantitativas devem-se, em grande parte, mais às interações planta-bactéria específicas do que aos efeitos genéticos atribuídos individualmente aos genótipos das plantas ou dos rizóbios (MYTTON, 1984). Isso pode ser um fator limitante para os programas de melhoramento, mas, em geral, esse efeito é minimizado pela homogeneização dos genótipos bacterianos, que se dá pela seleção e veiculação das melhores estirpes por meio de inoculantes. Os Estados Unidos, a África e o Brasil constituem três situações diferentes que ilustram bem os desafios e as possibilidades geradas pela interação planta-bactéria.

Nos Estados Unidos, a existência de estirpes de *Bradyrhizobium* spp. bem estabelecidas nos solos, altamente competitivas, mas pouco eficientes no processo de FBN tem inviabilizado a introdução de novas estirpes mais eficientes. A busca por alternativas que possibilitem a introdução efetiva de estirpes superiores levou à identificação dos genes *rj1*, *Rj2*, *Rj3* e *Rj4*, em alguns materiais americanos, que

restringem a nodulação com certos sorogrupos de bactérias (DEVINE e BREITHAUPT, 1980). Foi sugerido que o uso destas incompatibilidades poderia ser útil para controlar a simbiose do hospedeiro, ou seja, permitir a formação de nódulos apenas com estirpes de rizóbios geneticamente superiores. Como os locos para os genes *rj1*, *Rj2*, *Rj3* e *Rj4* não estão ligados, os melhoristas poderiam desenvolver genótipos contendo uma ou todas as combinações (DEVINE e O'NEIL, 1989) e, desse modo, seria possível identificar associações altamente produtivas que viabilizassem o desenvolvimento de cultivares capazes de nodular especificamente com estirpes eficientes dos inoculantes. No entanto, a obtenção de um único genótipo vegetal que combine todos os níveis de restrição da nodulação desejados com as estirpes estabelecidas no solo é difícil e pode, também, restringir a nodulação com estirpes desejáveis. Além disso, a expressão do fenótipo de restrição da nodulação pode ser influenciada por fatores ambientais, como a temperatura (SADOWSKY et al., 1995).

Uma situação inversa ocorre na África, onde pequenos agricultores têm cultivado variedades locais de soja com baixa ou nenhuma aplicação de fertilizantes nitrogenados e sem o uso de inoculantes. Essa característica é denominada nodulação promíscua e as variedades são capazes de nodular com diversas espécies de rizóbios nativos presentes nos solos africanos, nodulando e fixando eficientemente nitrogênio (HERRIDGE e ROSE, 2000; MPEPEREKI et al., 2000). O programa de melhoramento desenvolvido desde o início dos anos 1980 pelo Instituto Internacional de Agricultura Tropical (IITA), na Nigéria, tem explorado a combinação dessa capacidade de nodulação promíscua dos materiais locais com as características melhoradas dos cultivares americanos, visando obter linhagens que combinem produtividade e nodulação promíscua (HERRIDGE e ROSE, 2000). O programa tem obtido sucesso no desenvolvimento de linhagens mais produtivas e de nodulação promíscua. No entanto, a habilidade de nodulação promíscua dessas linhagens tem se mostrado limitada e dependente da população de rizóbios do solo, apresentando desempenhos melhores em um número limitado de ambientes. Outra estratégia desenvolvida no Zimbábue e Zâmbia consiste na identificação de materiais superiores selecionados entre as variedades locais, com base no potencial de produção e na habilidade de nodulação promíscua. Esse programa tem tido sucesso em assegurar a nodulação promíscua em

diferentes ambientes, variedades locais como Magoy e Hermom 147 foram recomendadas e são cultivadas por pequenos agricultores (MPEPEREKI et al., 2000).

No Brasil, ao contrário dos EUA e da África, apesar de não haver um programa de melhoramento especificamente voltado para a FBN, a maximização da FBN foi favorecida pela condução dos programas de melhoramento em solos pobres em N, e pela inoculação das sementes com estirpes eficientes de *Bradyrhizobium*, sem a aplicação de fertilizantes nitrogenados. A FBN em soja no Brasil é apontada como um dos fatores responsáveis pelo sucesso do cultivo dessa leguminosa no País (ALVES, BODDEY e URQUIAGA, 2003; HUNGRIA et al., 2006).

2.1.2.4 A FBN na Cultura da Soja no Brasil

O cultivo em escala comercial de soja no Brasil iniciou-se na década de 1940 no Rio Grande do Sul, a partir da década de 1960 houve uma expansão da área plantada, quando a soja passou a ser cultivada em rotação com o trigo e na década de 1970, com a expansão para o cerrado, a soja tornou-se um dos principais produtos agrícolas do País (ALVES, BODDEY e URQUIAGA, 2003). A área cultivada com essa leguminosa no Brasil atualmente é superior a 21 milhões de hectares, com uma produção estimada em aproximadamente 57 milhões de toneladas para a safra 2008/2009 (CONAB, 2009).

A soja é uma cultura exigente em N, para cada 1000 kg de grãos produzidos, com um teor médio de 6,5% de N são necessário cerca de 80 kg de N ha⁻¹ (65 kg alocados nas sementes e 15 kg nas folhas, caules e raízes). Para um rendimento médio de 3000 kg de grãos ha⁻¹, a planta necessita, portanto, 240 kg N ha⁻¹. Considerando que o potencial genético da soja é estimado em, aproximadamente 8.000 kg de grãos ha⁻¹, a demanda por N pode ser superior a 1000 kg (HUNGRIA, CAMPO e MENDES, 2007). Salvagiotti et al. (2008) relatam que em média 50% a 60% da demanda de N da soja é suprida pela FBN. Uma situação privilegiada é encontrada no Brasil, onde as taxas de fixação de N₂ em condições de campo podem exceder 300 kg de N ha⁻¹. Para garantir que a FBN consiga suprir essas quantidades crescentes de N, e assegurar os benefícios da FBN o melhoramento deve considerar essa fonte de N durante os processos de seleção de plantas.

Assim como os cultivares de soja, os primeiros inoculantes utilizados foram importados dos EUA. A seleção de estirpes adaptadas às condições brasileiras a partir do re-isolamento de bactérias dos nódulos e avaliação da eficiência do processo de fixação de N₂ resultaram na identificação de estirpes mais eficientes que são utilizadas nos inoculantes comerciais. Por outro lado, por determinação de uma comissão formada em 1963, os programas de melhoramento passaram a omitir a fertilização nitrogenada e a inocular as plantas com as estirpes recomendadas. Credita-se a esses dois esforços da pesquisa, o sucesso da FBN na soja no País (HUNGRIA, FRANCHINI, CAMPO e GRAHAM, 2005; ALVES, BODDEY e URQUIAGA, 2003).

2.1.3 Mapa Genético e Mapeamento de QTL em Soja

Os primeiros mapas genéticos da soja foram baseados em marcadores RFLP (*restriction fragment length polymorphism*), RAPD (*random amplified polymorphic DNA*), isoenzimas e marcadores morfológicos (SHOEMAKER e SPECHT, 1995; LARK et al., 1993; KEIM et al., 1990). Cregan et al. (1999) desenvolveram um conjunto de 606 marcadores SSR (*simple sequence repeat*) e utilizaram esses marcadores para construir mapas genéticos em duas populações F₂, uma oriunda do cruzamento interespecífico entre a PI 468.916 (*G. soja*) e a linhagem A81-356022 (*G. max*) e a outra do cruzamento entre os cultivares Clark x Harosoy. E também uma população de RILs (Minsoy x Noir 1). Pelo alinhamento dos mapas gerados para essas três populações construíram um mapa consenso com 20 grupos de ligação (GL), correspondendo os 20 pares de cromossomos homólogos da soja. Posteriormente 420 novos marcadores SSR, dentre os quais 24 derivados de sequências ESTs, foram incorporados aos desenvolvidos por Cregan et al. (1999), resultando em um novo mapa com 1.026 marcadores SSR (SONG et al. 2004). Choi et al. (2007) desenvolveram marcadores baseados em polimorfismos de base única (*single-nucleotide polymorphisms* – SNPs), a partir de sequências ESTs, possibilitando o mapeamento de 1.141 SNPs. Atualmente o mapa genético da soja tem um tamanho estimado em aproximadamente 2.523,6 cM (distância de mapa de Kosambi), com 20 GL contendo 1.015 marcadores microssatélites, 709 RFLPs, 73 RAPDs, 6 AFLPs, 10 isoenzimas, 24 marcadores

morfológicos e 1.141 SNPs (SONG et al., 2004; CHOI et al., 2007), proporcionando uma ferramenta de trabalho para estudos de mapeamento de QTL, mapeamento físico e sequenciamento do genoma da soja.

Os marcadores moleculares, principalmente os SSR, têm sido extensivamente empregados em estudos de mapeamento de QTL para características de interesse para o melhoramento da soja. A base de dados SOYBASE (2009) contém 95 categorias de características que incluem: qualidade de semente, resistência a doenças, absorção de nutrientes, ciclo reprodutivo, produção, etc. A seguir, são revisados alguns trabalhos que relatam mapeamento de QTLs para produção de grãos, atributos de qualidade dos grãos como concentração de isoflavonóides e proteína, resistência ao nematóide de cisto e absorção de N e P. Optou-se por essa ênfase em função da importância desses caracteres para o melhoramento, o que leva a um esforço no sentido de mapear QTLs, e procurando também relações entre essas características e a FBN.

A produção é um dos caracteres de maior interesse para o melhoramento. Mas, também, um dos caracteres agrônômicos mais complexos. Palomeque et al., (2009a) e Palomeque et al., (2009b) mapearam QTLs para produção e para componentes de produção (altura de planta, maturidade, acamamento, peso de cem sementes, estágio reprodutivo) em uma população de RILs, visando identificar QTLs estáveis em múltiplos ambientes. Para tanto, foram realizados experimentos em três locais no Canadá e em dois locais na China, em três anos de avaliações e que resultaram na identificação de sete QTLs para produção de sementes. Os marcadores Satt139 (C1), Satt100 (GL C2), Satt277 (C2), Satt162 (GL I) e Sat_126 (GL K) foram associados à QTLs para produção de sementes em pelo menos um ambiente no Canadá e um ambiente na China, sendo denominados QTLs ambiente-universal e explicaram de 9% a 37% da variação fenotípica. Dois QTLs ligados aos marcadores Satt194 (GL C1) e Satt259 (GL O) foram identificados em ambientes específicos. Dos sete QTLs mapeados para produtividade, cinco foram associados também aos componentes de produção. Na região delimitada pelos marcadores Satt277-Satt100 no GL C2, por exemplo, foram mapeados QTLs para peso de cem sementes, vagens por planta, vagens por nó, altura de planta, estágio reprodutivo (R1, R5 e R8), conteúdo de proteína e óleo.

Outros trabalhos sobre mapeamento de QTLs para produção têm utilizado populações de retrocruzamento onde o genitor doador são PIs, e o genitor recorrente

são cultivares derivados dos programas de mapeamento. Com essa abordagem, tem-se buscado identificar QTLs para produtividade em materiais não adaptados (SMALLEY et al., 2004; GUZMAN et al., 2007; LI, PFEIFFER e CORNELIUS, 2008). Guzman et al. (2007) mapearam treze QTLs em três populações de retrocruzamento e em oito desses QTLs o alelo favorável foi derivado do genitor doador (PIs), sendo cinco localizados nos GL A1, B2, C1, C2 e K, em regiões onde outros trabalhos também relatam QTLs para produtividade em pelo menos duas populações independentes. O QTL localizado no grupo C2 foi significativamente associado ao SSR Satt557 em três ambientes diferentes e explicou, em média, 20% da variação. O marcador Satt557 está localizado entre os marcadores Satt277-Satt100. Nessa região, outros trabalhos também descrevem QTLs para produção e componentes de produção (ORF et al., 1999; SPECHT et al., 2001; SMALLEY et al., 2004; PALOMEQUE et al., 2009a e PALOMEQUE et al., 2009b). Indicando que os efeitos dessa região sobre caracteres relacionados com produção são independentes de ambientes e backgrounds genéticos. Próximo a essa região Nicolás, Hungria e Arias (2006) e Santos, Nicolás e Hungria (2006) identificaram QTLs para caracteres relacionados com a FBN (NN e MNS).

O nematóide de cisto da soja (NCS) (*Heterodera glycines* Ichinohe) causa consideráveis quedas de produtividade na cultura da soja no mundo. Entre as causas dos danos causados está a inibição da nodulação por *B. japonicum*, causando deficiência de N para crescimento da planta (KO, BARKER e HUANG, 1984; WU et al., 2009). Concibido et al. (2004) sumarizaram os resultados de dezesseis trabalhos sobre mapeamento de QTLs para resistência ao NCS. No total são descritos 60 QTLs, distribuídos em dezessete grupos de ligação (A1, A2, B1, B2, C1, C2, D1a, D2, E, F, G, H, I, J, L, M e N), compreendendo 31 regiões. Duas regiões consistentes correspondem a QTL de efeito maior, uma no GL A2 (*Rhg4*) e outra no GL G (*rhg1*). Outras regiões onde foram mapeados QTL em mais de uma população e diferentes fontes de resistência estão distribuídas nos grupos B1, E, F e J.

Para avaliar a confiabilidade dos dados sobre mapeamento de QTLs disponíveis, Guo et al. (2006) utilizaram as informações descritas acima em um estudo de meta-análise, baseado no intervalo de confiança. QTLs mapeados por diferentes trabalhos e QTLs significativos para diferentes raças de NCS foram agrupados quando seus

intervalos de confiança foram sobrepostos e considerados como um mesmo locus. Os QTLs foram classificados de acordo com a evidência estatística em três classes: sugestivos ($LOD \leq 3$), significativos ($LOD \leq 4$) e confirmados ($LOD \geq 4$ mapeado em mais de um estudo ou populações com fonte de resistência diferente). Com esses critérios foram considerados como confirmados os QTL localizados nos GL A2, G, E, B1, e J e, como sugestivos ou significativos, os QTL descritos nos demais grupos de ligação. Os QTLs mapeados para diferentes raças (1, 2, 3, 4, 5 e 14) foram agrupados nas mesmas regiões nos grupos A2, G, E, B1 e J. Por exemplo, o QTL do grupo de ligação G (*rhg1*) foi frequentemente, associado com resistência às raças 1, 2, 3 e 5 e o QTL no grupo A2 (*Rhg4*) é associado à resistência à raça 3 e, em alguns casos, com a resistência às raças 2, 5 e 14. Wu et al. (2009), utilizando uma população grande de RILs (275 F_{7:9} Essex x PI 437654), com boa cobertura do genoma (2.406 cM), intervalo médio de 8,7 cM e empregando o método de mapeamento por intervalo composto, identificaram os QTLs de efeito maior e já bem descritos GL A2, B1 e G e o QTL de pequeno efeito para resistência às raças 5 e 14, no grupo de ligação I (Sat_299-Sct_189) classificado como QTL sugestivo por Guo et al. (2006). Além de confirmar essas regiões, quatro novos QTLs de pequeno efeito foram mapeados nos grupos H e K para a raça 1, no grupo C1 para a raça 3 e no grupo D1a para as raças 2 e 3.

As isoflavonas daidzeína, genisteína e gliciteína encontradas nas sementes de soja, possuem efeitos benéficos sobre vários aspectos da saúde humana, osteoporose, menopausa, doenças cardiovasculares e alguns tipos de câncer (MUNRO et al., 2003). Na planta, as isoflavonas estão relacionadas com várias interações planta-microrganismos como na resposta a patógenos e na interação com as bactérias fixadoras de N₂. As isoflavonas exsudadas na rizosfera são importantes no processo inicial da nodulação, ativando a transcrição de genes bacterianos (genes de nodulação) que codificam as enzimas envolvidas na síntese da molécula sinal da nodulação (lipoquitooligossacarídeo) (HUNGRIA e STACEY, 1997; MEKSEM et al., 2001).

Trabalhos sobre mapeamento de QTLs para isoflavonóides nas sementes de soja têm identificado QTLs para daidzeína, genisteína, gliciteína e isoflavonas totais em quatro populações de RILs diferentes. Os primeiros trabalhos foram realizados por Njiti et al. (1999), Meksem et al. (2001) e Kassem et al. (2004), com uma população

constituída por 100 RILs derivadas do cruzamento entre os cultivares Essex x Forrest. Esses trabalhos empregaram os métodos de análise de QTL de marca simples e mapeamento por intervalo simples (IM), usando o mesmo conjunto de dados fenotípicos e aumentando a cobertura do genoma pela inclusão de novos marcadores e relataram o mapeamento de sete QTLs em cinco grupos de ligação (A1, B1, K, H e N). Os QTLs para gliciteína (K, B1, H e N), para daidzeína (A1 e K) e genisteína (K) explicaram 70,6%, 31,9% e 21% da variação, respectivamente. No entanto, quando esse mesmo conjunto de dados foi analisado por Kassem et al. (2006) com o emprego do método de mapeamento por intervalo composto (CIM), apenas o QTL para gliciteína no GL B1 e o QTL para daidzeína no GL N permaneceram significativos e onze novos QTLs foram identificados, dois para gliciteína (A2, D1a e N), dois para daidzeína (A2 e M), cinco para genisteína (A2, M, G) e dois para isoflavonas totais (A2 e M).

Primomo et al. (2005) mapearam QTLs para isoflavonas em outra população composta por 207 RILs (Ac756 x RCAT Angora), em dois locais. Considerando os QTLs identificados pelo método de mapeamento CIM, no primeiro ambiente, dezessete QTLs foram mapeados em onze grupos de ligação (A1, A2, C2, D1a, F, G, H, J, K, M e N); individualmente, esses QTL explicaram de 3,5% a 10,5% da variação. No segundo ambiente, apenas os QTL para genisteína e isoflavonas totais no grupo M foram mapeados.

Em um estudo recente, Zeng et al. (2009) avaliaram uma população de 130 RILs (Zhongdou 27 x Jiunong 20) em seis ambientes (três locais e dois anos) e identificaram 15 QTLs para isoflavonas nas sementes de soja, três para daidzeína (F, K e I), quatro para genisteína (A2, C2, e F), três para gliciteína (I, M e O) e cinco para isoflavonas totais (F, G, O e M). O marcador SSR Satt144 no grupo F foi associado a três características (daidzeína, genisteína e isoflavona total) nos seis ambientes. No grupo M o marcador Satt245 foi associado com genisteína e isoflavonas totais em cinco ambientes; esse marcador também foi associado a essas características por Primomo et al. (2005) em dois ambientes e em outra constituição genética. Em média os QTLs foram significativos em mais de quatro ambientes e explicaram de 3% a 18% da variação fenotípica, sendo que a maioria da variação individual foi inferior a 10%.

Gutierrez-Gonzales et al. (2009) também mapearam QTLs para isoflavonas em outra população formada por 197 RIL F₇ (Essex x PI437654), em dois ambientes. Dos 26 QTL mapeados, dezenove foram significativos para mais de uma característica, sendo quatro para genisteína, daidzeína e isoflavonas totais (A1, D1b, M e L), um para genisteína, daidzeína e gliciteína (E), dois para daidzeína e isoflavonas totais (A2, H). A maioria dos QTLs foram de efeitos menores ($R^2 < 10\%$). Os QTLs mapeados no grupo A1 para genisteína, daidzeína e isoflavonas totais explicaram a maior parte da variação (22,1%, 18,8% e 30%) e esses QTL coincidem com os QTLs mapeados por Primomo et al. (2005).

Considerando os resultados dos estudos de mapeamento de QTLs nas quatro populações, observa-se que dois QTLs de efeitos maiores mapeados na população derivada dos cultivares Essex e Forrest para genisteína no grupo B1 ($R^2 > 40\%$) e para daidzeína no grupo N ($R^2 > 50\%$) não foram identificados nas outras três populações reportadas. Possivelmente essas populações não segregaram para esses QTLs, em contrapartida, vários QTLs de efeitos menores foram identificados. Duas regiões, uma no GL M (Satt323-Satt463) e outra no A1 (Satt336-Sat_271) foram associadas às três características, genisteína, daidzeína e isoflavonas totais em duas populações e em diferentes locais, caracterizando-se como regiões estáveis em diferentes ambientes e *backgrounds* genéticos.

Os grãos de soja contêm em torno de 40% de proteína e 20% de óleo, razão pela qual a cultura é uma das principais fontes de proteína para alimentação animal e óleo para consumo humano. O melhoramento para aumento da concentração de proteína resulta, geralmente, em redução da concentração de óleo e vice-versa, devido à correlação negativa entre esses dois caracteres. O mapeamento de QTLs pode fornecer informações sobre as razões genéticas dessa correlação e a possibilidade de melhorar essas características de forma independente. Mais de 150 QTL para proteína são descritos na base de dados SOYBASE (2009). Considerando os intervalos onde esses QTL foram mapeados, observa-se que correspondem a 64 regiões genômicas distribuídas ao longo dos 20 grupos de ligação da soja. Desses QTLs, 42 foram mapeados em apenas uma população, 12 QTLs foram mapeados em duas populações, oito QTL em três populações e apenas dois QTL foram mapeados em quatro

populações, de um total de dezesseis populações de mapeamento onde os estudos foram realizados. Apesar de um grande número de QTL mapeados, a maioria tem sido detectada em um número pequeno de populações e em poucos ambientes e muitos QTL foram mapeados pelo emprego de métodos de análise de marca simples, ou mapeamento por intervalo simples.

Uma das regiões onde foram mapeados QTL para proteína em quatro populações diferentes, corresponde à região no grupo I onde Diers et al. (1992) identificaram um QTL de efeito maior controlando a concentração de proteína em uma população de 60 linhagens $F_{2:3}$ do cruzamento entre o acesso com alto teor de proteína PI 468916 (*G. soja*) e a linhagem produtiva A81-356022 (*G. max*). Essa região tem sido extensivamente caracterizada. Sebolt et al. (2000) desenvolveram uma população de retrocruzamento RC_3 utilizando o genitor recorrente A81-356022 e uma linhagem $F_{2:3}$ homocigota para os alelos da PI 468916 oriunda da população desenvolvida por Diers et al. (1992) e observaram que, as linhagens homocigotas para os alelos dos marcadores ligados ao QTL no GL I apresentavam um aumento na concentração de proteína quando comparadas com as linhagens homocigotas para os alelos do genitor recorrente. As linhagens homocigotas para os alelos da PI 468916 apresentaram, também, uma redução no teor de óleo, produção e peso de semente, além de maturidade precoce. A mesma região também foi mapeada por Chung et al. (2003) em uma população de 76 RIL F_5 derivadas do cruzamento entre a PI 437088A (acesso de *G. max* com alta concentração de proteína) e o cultivar Asgrow A3733 (altamente produtiva e baixo teor de proteína). O QTL foi delimitado pelos marcadores SSR Satt496 e Satt239, compreendendo uma região de, 6,6 cM. Os alelos da PI 437088A aumentaram a concentração de proteína nos grãos em 18 g kg^{-1} , diminuíram o conteúdo de óleo em $11,1 \text{ g kg}^{-1}$ e reduziram a produtividade em 154 kg ha^{-1} , resultados semelhantes aos relatados por Sebolt et al. (2000).

Os esforços para o mapeamento fino dessa região resultaram na delimitação do QTL em um intervalo de 3 cM entre o marcador SSR Satt239 e o marcador AFLP (*amplified fragment length polymorphisms*) ACG9b (NICHOLS et al., 2006). Para o mapeamento fino foram utilizados dois conjuntos de populações de retrocruzamento desenvolvidos a partir da população de retrocruzamento RC_3 desenvolvida por Sebolt et

al., (2000) descrita anteriormente. O primeiro conjunto incluiu três populações de RC₄ e o segundo conjunto incluiu quatro populações RC₅. Essas populações segregam para diferentes segmentos da região genômica que delimitam o QTL. O mapeamento fino é um passo importante para a clonagem baseada em mapeamento dessa região. A identificação do gene ou genes que estão relacionados com o controle genético da concentração de proteína e óleo é importante para a compreensão dos mecanismos que regulam esses caracteres.

A soja é uma das culturas com maior teor de proteína nos grãos. Estima-se que para produzir 1000 kg de grãos são necessários 80 kg de N, que pode ser obtido, principalmente de duas fontes, o N mineral disponível no solo e veiculado por meio de fertilizantes, ou a FBN por meio da simbiose com bactérias fixadoras de N₂. (HUNGRIA, CAMPO e MENDES, 2007). Regiões genômicas governando o acúmulo de N nas sementes foram identificadas por Panthee et al. (2004), em uma população de 101 RILs F_{6:8} do cruzamento entre as linhagens N87-984-16 x TN93-99. Foram identificadas 17 associações significativas nos grupos A2, B2, D1b, E, G e M, todas de pequeno efeito, (R²= 5,9 a 11,6). O QTL detectado no grupo B2 (Satt556) coincide com a região onde Nicolás, Hungria e Arias (2006) mapearam QTLs para nodulação (NN e MNS). Os QTLs associados aos marcadores Satt274 no GL D1b e Satt570 no GL M também foram relacionados com QTLs que explicaram 20,2% e 24,9% do teor de proteína nos grãos (PANTHEE et al., 2005).

O fósforo (P) é depois do N, o elemento mineral mais limitante para o crescimento da planta e produção de grãos. Plantas que assimilam N por meio da FBN geralmente requerem mais P do que as plantas que utilizam N mineral. A atividade da enzima nitrogenase requer altos níveis de ATP, além disso, o P é necessário para a transdução do sinal, biossíntese de membrana, desenvolvimento e funcionamento do nódulo. A fertilização fosfatada promove incrementos na atividade da nitrogenase, no número e na massa de nódulos e, no acúmulo de N nos tecidos das plantas dependentes de FBN como fonte de N (ISRAEL, 1987). Em um estudo de mapeamento de QTLs para assimilação de P, Zhang et al. (2009) avaliaram os seguintes caracteres fenotípicos: aquisição eficiente de P (AEP), calculado como o total de P na planta (mg P planta⁻¹), uso eficiente de P (UEP), que corresponde ao peso da planta seca produzido

por unidade de P (g matéria seca mg de P⁻¹) e atividade da fosfatase ácida em uma população de 152 RILs e identificaram QTLs nos grupos de ligação B2, D1b, D2, C2, H, E e L e, em geral, os QTLs explicaram individualmente 12% a 19% da variação. Li et al. (2005) mapearam QTLs para tolerância ao estresse por falta de P em outra população de RILs, e detectaram dois QTLs os para conteúdo de P nas folhas e conteúdo de P na raiz, localizados no GL F. Esses QTLs explicaram 25% e 12% da variação, respectivamente.

Poucos estudos foram realizados sobre o mapeamento de QTL para FBN. Nicolás, Hungria e Arias (2006), identificaram cinco regiões associadas com nodulação nos grupos de ligação B1, B2, C2, F, D1b e L, explicando de 3% a 9% da variação fenotípica dos caracteres de nodulação (NN e MNS), em uma população de 160 F_{2:3} linhagens do cruzamento entre BRS 133 x Embrapa 20. Todos os marcadores SSR associados com nodulação nesse trabalho, são localizados em regiões onde estão descritos QTL para teor de proteína nos grãos (SOYBASE, 2009). O QTL para MNS no GL B1 também está localizado próximo da região onde Kassem et al. (2006) mapearam um QTL de efeito maior explicando >50% da variação fenotípica para o isoflavonóide gliciteína. Os marcadores associados com caracteres de FBN identificados por Nicolás, Hungria e Arias (2006) foram utilizados por Santos, Nicolás e Hungria (2006) para verificar essas associações em outra população de mapeamento formada por 159 RILs F_{2:7} do cruzamento entre os cultivares Bossier x Embrapa 20. As regiões nos GL B1, H, C2 e D1b, foram significativamente associadas com os caracteres de nodulação, indicando que essas regiões estão associadas com a FBN em soja. Regiões genômicas associadas com caracteres de FBN também foram identificadas nos GL J, I, A1, D1b e O por Tanya et al. (2005). Os QTL identificados também são de pequeno efeito $R^2 < 6,84$, exceto para o QTL localizado no GL O, que foi significativo para todas as características de FBN avaliadas (NN, MNS, MPAS, N₂ fixado) e explicou de 12,91% a 20,83% da variação fenotípica. Como observado nos trabalhos anteriores, os QTL para FBN localizados nos GL J, I, A1 e D1b também coincidem com regiões onde foram mapeados QTL para teor de proteína nos grãos.

Apesar de promissores, esses trabalhos são iniciais e apresentam uma baixa cobertura do genoma, razão pela qual a identificação de QTLs foi feita com métodos de

análise baseados em marcas simples, portanto, esforços no sentido de obter mapas mais saturados, que possibilitem o mapeamento de QTL por métodos mais robustos como o CIM podem contribuir para o mapeamento mais preciso dos QTL relacionados com FBN na soja.

Outra abordagem empregada em estudos de mapeamento é o desenvolvimento de marcadores baseados em sequências de genes com função conhecida (genes candidatos), como marcadores no estudo de QTLs. Assume-se, como hipótese de trabalho, que o polimorfismo dentro de genes candidatos pode estar relacionado com a variação fenotípica. Esses genes candidatos podem ser genes estruturais ou genes envolvidos com a regulação de vias metabólicas (PFLIEGER, LEFEBVRE e CAUSSE, 2001). Genes candidatos são empregados, principalmente, para o mapeamento de QTLs relacionados à resistência a doenças, pois muitos genes envolvidos na interação planta-patógeno foram clonados e caracterizados. Byrne et al. (1996) demonstraram a ligação entre genes envolvidos na via de síntese de isoflavona e QTLs para resistência à lagarta do cartucho do milho. No caso, o loco p1 que codifica um ativador transcricional e outros três genes candidatos foram responsáveis por 75,9% da variação fenotípica de resistência. Faris et al. (1999) mapearam QTLs de pequeno efeito para resistência à mancha amarela e à ferrugem foliar em linhagens recombinantes de trigo, utilizando genes que codificam proteínas de resposta de defesa da planta (genes DR). Esses genes são, geralmente, superexpressos no processo de resposta da planta ao ataque por patógenos.

Em soja, estudos sobre expressão gênica a partir de nódulos também têm disponibilizado um grande número de sequências ESTs que correspondem a genes envolvidos com o desenvolvimento do nódulo (LEE et al., 2004; JEONG et al., 2006; LESTARI et al., 2006). No trabalho desenvolvido por Lee et al. (2004), foi utilizada uma abordagem para identificar genes expressos especificamente nos nódulos; para tanto, construíram um banco de dados de ESTs de nódulos e compararam esse banco de ESTs com mais de 330.000 ESTs derivados de bibliotecas construídas para vários órgãos da soja disponíveis nos bancos de dados públicos, visando identificar ESTs presentes especificamente em bibliotecas construídas a partir de raízes e nódulos. Com essa abordagem, identificaram 3.267 ESTs encontradas apenas em bibliotecas construídas a

partir de nódulos de soja e 72 ESTs presentes em ambos (raízes e nódulos). Posteriormente, analisaram a expressão gênica em três fases de desenvolvimento dos nódulos, utilizando um arranjo com 382 cDNA selecionados a partir dessas ESTs e verificaram que 58 genes apresentaram um aumento no padrão de expressão gênica, dentre os quais 20 são genes que codificam para nodulinas conhecidas, como a nodulina 20, a nodulina 93, e a nodulina 96 e a leghemoglobina. Treze são genes envolvidos com metabolismo de C e N, e genes envolvidos com metabolismo secundário.

Lestari et al. (2006) avaliaram o padrão de expressão gênica em nódulos do cultivar Sinpaldalkong-2 e o tipo mutante SS2-2, caracterizado por apresentar um grande número de nódulos, redução do tamanho das raízes e da parte aérea, maior conteúdo de proteína nos grãos e menor conteúdo de carboidratos no início do desenvolvimento. Nesse estudo foram identificados 140 transcritos que são diferencialmente expressos, 65 expressos preferencialmente no cultivar selvagem e 40 no tipo mutante. A maioria dos transcritos identificados corresponde a genes que codificam enzimas envolvidas com o metabolismo do C e do N, e com os processos de energia e transporte, provavelmente para suprir a demanda dos nódulos, como as enzimas piruvato carboxilase e acetil coenzima A carboxilase, envolvidas com o metabolismo de oxalacetato e glutamato, respectivamente. Transcritos potencialmente envolvidos com a simbiose também foram identificados, por exemplo, um transcrito específico identificado no cultivar Sinpaldalkong-2 apresenta homologias com proteínas receptoras tipo-kinase e pode estar relacionado com o controle da nodulação que distingue o selvagem do supernodulante.

Borevitz e Chory (2004) consideram que, no processo de escolha dos genes candidatos devem ser consideradas as informações disponíveis obtidas por meio da genética tradicional e por abordagens moleculares. Os estudos de expressão gênica que visam identificar genes sob condições específicas podem ser boas fontes para a seleção de sequências candidatas para o mapeamento de características de interesse. Assim, os trabalhos sobre expressão gênica em raízes de soja noduladas com *B. japonicum* foram usados como base para seleção de genes candidatos avaliados nesse trabalho.

2.2 Material e Métodos

2.2.1 Material Vegetal e Análise dos Caracteres Relacionados com a FBN

A população utilizada neste estudo foi desenvolvida por Nicolás, Hungria e Arias (2002), e é constituída por 157 linhagens endogâmicas recombinantes (Recombinant Inbred Lines, RILs) $F_{2:7}$ oriundas do cruzamento entre os cultivares Bossier (oriunda de seleção em Lee) com alta capacidade de FBN e Embrapa 20 (oriunda do cruzamento Doko RC4 x IAC-7R) com capacidade intermediária de FBN (HUNGRIA e BOHRER, 2000). A população de RILs e os cultivares genitores foram avaliados no verão de 2003 por Santos, Nicolás e Hungria (2006), conforme descrição a seguir. As avaliações foram conduzidas em casa de vegetação, em vasos plásticos contendo quatro kg de uma mistura de solo e areia. O solo (oxissolo) utilizado foi coletado em Londrina (Paraná, Brasil), e tem as seguintes propriedades químicas: pH em CaCl_2 : 5,1; N: $0,15 \text{ g dm}^{-3}$; Al: $0,0 \text{ cmol}_c \text{ dm}^{-3}$; K: $0,25 \text{ cmol}_c \text{ dm}^{-3}$; Ca: $5,2 \text{ cmol}_c \text{ dm}^{-3}$; Mg: $1,5 \text{ cmol}_c \text{ dm}^{-3}$; H + Al: $4,16 \text{ cmol}_c \text{ dm}^{-3}$; C: $2,5 \text{ g dm}^{-3}$ e P: $5,1 \text{ mg dm}^{-3}$. Embora o solo apresentasse uma população elevada de bactérias capazes de nodular a soja, para assegurar uma boa nodulação as plantas foram inoculadas com uma mistura das estirpes de *Bradyrhizobium japonicum* SEMIA 5079 e *B. elkanii* SEMIA 587, duas semanas após a semeadura. O delineamento experimental foi inteiramente casualizado, com oito repetições, contendo cada parcela uma planta. Aos 44 dias após o plantio as plantas foram coletadas individualmente.

Na coleta, o sistema radicular e a parte aérea foram separados e mantidos em uma estufa com circulação forçada de ar, a 65°C , por quatro dias, quando atingiram massa constante. Foram avaliadas as características relacionadas com a FBN: massa de nódulos secos em mg planta^{-1} (MNS); número de nódulos (NN) e massa média de nódulos secos (MNS/NN); e crescimento da planta: massa da parte aérea seca em g planta^{-1} (MPAS).

Os dados fenotípicos obtidos por Santos, Nicolás e Hungria (2006) foram utilizados para determinar os coeficientes de correlações genéticas e fenotípicas entre os caracteres MNS, NN, MNS/NN e MPAS segundo procedimentos propostos por

Vencovsky e BARRIGA (1992). Para isso foram realizadas as análises de variância para as características individuais e de covariância para os pares de características, segundo o delineamento inteiramente casualizado, utilizando o software estatístico SAS (Statistical Analysis System –SAS INSTITUTE, 2003) versão 9.1, módulo “proc GLM”. A partir das análises de variância foram estimadas as variâncias genéticas e fenotípicas para cada caráter e, a partir das análises de covariância, as covariâncias genéticas e fenotípicas entre os pares de caracteres. Desse modo, as correlações genéticas e fenotípicas foram determinadas pelas expressões:

Correlação Genética:

$$r_{G(x,y)} = \frac{COV_{G(x,y)}}{\sqrt{\sigma_{G(x)}^2 \sigma_{G(y)}^2}} \quad (1)$$

em que:

$r_{G(x,y)}$: é o coeficiente de correlação genética entre as características x e y;

$COV_{G(x,y)}$: é a covariância genética entre as características x e y;

$\sigma_{G(x)}^2$ e $\sigma_{G(y)}^2$: são as variâncias genéticas das características x e y;

Correlação Fenotípica:

$$r_{F(x,y)} = \frac{COV_{F(x,y)}}{\sqrt{\sigma_{F(x)}^2 \sigma_{F(y)}^2}} \quad (2)$$

em que:

$r_{F(x,y)}$: é o coeficiente de correlação fenotípica entre as características x e y;

$COV_{F(x,y)}$: é a covariância fenotípica entre as características x e y;

$\sigma_{F(x)}^2$ e $\sigma_{F(y)}^2$: são as variâncias fenotípicas das características x e y;

2.2.2 Genotipagem e Construção do Mapa de Ligação

2.2.2.1 Análise com Marcadores Microssatélites

Para as análises moleculares foi extraído o DNA genômico de folhas jovens (primeiras folhas trifoliadas) pelo método descrito por Keim, Olson e Shoemaker (1988). Inicialmente foram feitos testes de polimorfismos entre os genitores para identificar os marcadores úteis para construção do mapa de ligação (Figura 1). Os marcadores microssatélites para o teste de polimorfismo entre os genótipos genitores foram selecionados ao longo dos 20 grupos de ligação (GL) da soja (CREGAN et al., 1999; SONG et al., 2004), de forma a permitir uma boa cobertura do genoma. No total 432 marcadores microssatélites foram utilizados para amplificar o DNA dos genitores, com uma média de vinte e um marcadores por GL. Com exceção de três regiões, uma no GL B2 e duas no GL C1, onde existem distâncias superiores a 28 cM, a distância em cM entre os marcadores selecionados foi em geral menor que 20 cM (Tabela 1). Dos marcadores testados, 153 foram polimórficos (35,83%), mas apenas 107 foram utilizados para a genotipagem da população de linhagens, pois alguns marcadores estão localizados muito próximos, como, por exemplo, no GL E, onde a distância entre os marcadores Satt598 e Satt573 é de 1,59 cM. Numa segunda etapa, estes marcadores poderão ser utilizados para saturar as regiões onde foram mapeados QTLs.

As reações de amplificação foram realizadas em um volume de 10 µL contendo: tampão de PCR 1X (2,5 mM Tris-HCL pH 8,3, 62,5 mM de KCl); 2,5 mM de MgCl₂; 125 µM de cada dNTP; 0,2 µM de cada oligonucleotídeo (*primer*) (*forward* e *reverse*), uma unidade de *Taq* DNA polimerase e 20 ng de DNA genômico. Para a amplificação foi utilizado um ciclo iniciado com 7 min a 94°C, seguidos de 30 ciclos de 1 min a 94°C, 1 min a 50°C e 2 min a 72 °C e finalizado com uma etapa de 7 min a 72°C.

Após a amplificação, os fragmentos foram separados por eletroforese vertical em gel de poliacrilamida (29:1, acrilamida:bis-acrilamida) a 10%, as amostras foram submetidas a um campo elétrico de 120 volts por três horas, corados em uma solução de brometo de etídio (0,5 µg ml⁻¹) e fotografados sob luz UV com câmera digital Kodak

DC 120 e o software para PC Kodak Digital Science 1D (Eastman Kodak Company). Para estimação do tamanho dos fragmentos amplificados foi utilizado o marcador de peso molecular 1KB plus DNA Ladder (Invitrogen) aplicado nas duas extremidades dos géis.

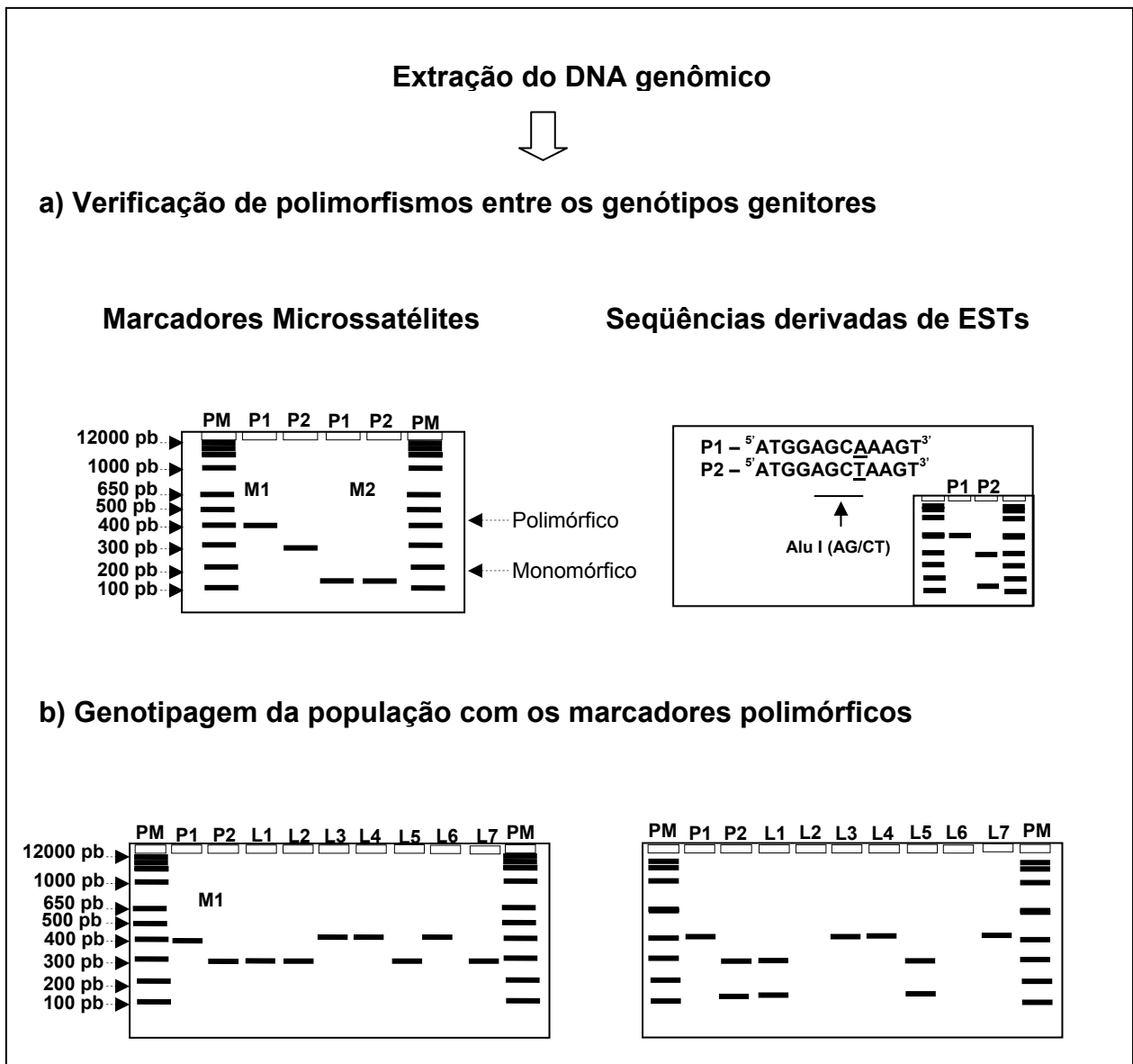


Figura 1 - Diagrama representando os procedimentos adotados para obtenção dos dados de genotipagem. a) testes de polimorfismo entre os genitores Bossier e Embrapa 20 (P1 e P2), para os marcadores microssatélites (M1, M2) e para as etiquetas de seqüências expressas (Expressed sequence tags, ESTs). b) Genotipagem da população de linhagens (L1,...,L7) com os marcadores polimórficos. PM - marcador de peso molecular 1 KB plus DNA Ladder

Tabela 1 - Marcadores microssatélites polimórficos para os genitores Bossier e Embrapa 20, utilizados para genotipar a população de linhagens

Microssatélite	GL ^a	Posição (cM) ^b	Microssatélite	GL ^a	Posição (cM) ^b
Satt572	A1 (5)	14,65	Satt516	F (13)	44,42
Satt276	A1 (5)	17,16	Sat_133	F (13)	50,78
Satt155	A1 (5)	32,68	Sat_120	F (13)	75,97
Satt050	A1 (5)	46,45	Satt395	F (13)	146,42
Satt385	A1 (5)	64,74	Satt356	G (18)	12,18
Soynod26	A1 (5)	66,80	Satt394	G (18)	43,38
Satt200	A1 (5)	92,89	Satt288	G (18)	76,77
Satt236	A1 (5)	93,23	Satt568	H (12)	27,64
Sct_067	A2 (8)	14,99	Satt192	H (12)	44,04
Satt589	A2 (8)	33,96	Satt279	H (12)	68,50
Satt233	A2 (8)	100,09	Satt302	H (12)	81,04
Satt421	A2 (8)	115,93	Satt181	H (12)	91,12
Sat_272	B1 (11)	14,32	Satt434	H (12)	105,74
Satt509	B1 (11)	32,51	Satt562	I (20)	22,84
Satt251	B1 (11)	36,48	Satt587	I (20)	31,49
Satt197	B1 (11)	46,39	Satt239	I (20)	36,94
GmPGN	B1 (11)	-	Satt354	I (20)	46,22
Satt332	B1 (11)	80,89	Satt270	I (20)	50,11
Satt583	B1 (11)	84,19	Satt162	I (20)	86,74
Satt444	B1 (11)	85,92	Satt406	J (16)	38,19
Satt416	B2 (14)	56,96	Satt529	J (16)	41,90
Sct_094	B2 (14)	70,56	Satt215	J(16)	44,08
Satt556	B2 (14)	73,21	Satt244	J(16)	65,04
Satt066	B2 (14)	78,84	Satt547	J(16)	67,79
Satt560	B2 (14)	97,92	Satt431	J(16)	78,57
Satt396	C1 (4)	24,11	Satt539	K (9)	1,80
Satt190	C1 (4)	73,32	Satt242	K (9)	14,35
Satt136	C1 (4)	75,11	Sat_044	K (9)	58,01
Satt180	C1 (4)	127,77	Sat_020	K (9)	103,60
Satt460	C2 (6)	117,77	Satt495	L (19)	0,00
Satt307	C2 (6)	121,27	Satt232	L (19)	10,35
Satt202	C2 (6)	126,24	Satt418	L (19)	30,93
Satt316	C2 (6)	127,67	Satt156	L (19)	56,14
Satt371	C2 (6)	145,48	Sat_113	L (19)	68,26
Satt357	C2 (6)	151,91	Satt373	L (19)	107,24
Satt184	D1a (1)	17,52	Satt404	M (7)	0,84
Satt468	D1a (1)	69,91	Satt150	M (7)	18,58
Satt436	D1a (1)	70,69	Satt540	M (7)	35,85
GmEKN	D1b (2)	-	Satt245	M (7)	53,54
Satt157	D1b (2)	37,07	Satt346	M (7)	112,79
Satt558	D1b (2)	43,91	Satt336	M (7)	133,83
Satt296	D1b (2)	52,61	Sat_084	N (3)	36,86
Satt542	D1b (2)	53,02	Satt584	N (3)	37,98
Satt290	D1b (2)	73,35	Satt237	N (3)	74,99
Satt458	D2 (17)	24,52	Satt234	N (3)	84,60
Satt135	D2 (17)	26,05	Satt257	N (3)	92,56
Satt573	E (15)	35,79	Sat_132	O (10)	8,75
Satt185	E (15)	44,76	Satt259	O (10)	39,82
Satt045	E (15)	46,65	Satt173	O (10)	58,40
Satt369	E (15)	56,27	Satt478	O (10)	71,10
Satt325	F (13)	2,23	Sat_038	O (10)	112,17
Satt252	F (13)	16,08			

^aGrupo de ligação; ^bPosição no mapa consenso (Cregan et al., 1999; Song et al., 2004).

2.2.2.2 Análise com ESTs (Expressed Sequence Tags)

Com base em dados de literatura (LEE et al., 2004; JEONG et al., 2006; LESTARI et al., 2006), foram selecionados 29 pares de oligonucleotídeos (*primers*) que amplificam ESTs de genes com provável função no processo de FBN. O procedimento para buscar polimorfismos nos fragmentos gerados por esses *primers* compreendeu a amplificação e o sequenciamento dos produtos de PCR e posterior clivagem com enzima de restrição e visualização do polimorfismo em gel de agarose (Figura 1).

Inicialmente, todos os pares de *primers* foram utilizados para amplificar um único genótipo (cultivar Bossier), visando a padronização das condições de amplificação e seleção dos pares de *primers* que produzissem um único produto de amplificação (fragmento). As reações de amplificação foram preparadas com 30 ng de DNA genômico, 1,5 mM de MgCl₂, 0,15 µM de *primer* (*forward* - F ou *reverse* - R), 100 µM de cada nucleotídeo, 1 X tampão (10 mM de Tris-HCl pH 8,3, 50 mM KCl), 2 U de *Taq polimerase*, em um volume final de 50 µL. O ciclo de amplificação foi o mesmo utilizado nas reações de sequenciamento e consistiu de uma etapa de desnaturação de 92°C por 45 segundos, anelamento de 50°C por 45 segundos e uma temperatura de extensão da fita de 68°C por 45 segundos por 32 ciclos. Para os *primers* que produziram bandas inespecíficas foram ajustadas as temperaturas de anelamento, de forma a obter fragmentos únicos (Tabela 2).

Os pares de *primers* que produziram banda única foram sequenciados. Para tanto, os produtos de PCR foram purificados (PureLink™ PCR Purification Kit Invitrogen). Para cada EST foram feitas duas reações de sequenciamento, utilizando um dos *primers* e uma mistura de 3 µL de *dye* (DYEnamic ET terminator reagent premix for the MegaBACE, Amersham Biosciences), para os genitores Bossier e Embrapa 20. Após a amplificação, as amostras foram purificadas, ressuspendidas em água e sequenciadas no sequenciador MegaBACE 1000 DNA Analysis System (Amersham Biosciences). Os parâmetros para a eletroforese foram: voltagem de injeção 2 kV e tempo de injeção 60 segundos; voltagem de corrida 5 kV e tempo de corrida de 300 minutos. As sequências de alta qualidade foram agrupadas em *contigs* usando os

Tabela 2 - Lista das sequências ESTs e dos marcadores microssatélites dentro de nodulinas conhecidas, que foram utilizados para buscar polimorfismos entre os parentais Bossier e Embrapa 20

(continua)

Nº de acesso no NCBI	Gene - características	Primers (5' – 3') ^e	pb ^f	T °C ^g	Função
AI973769 ^a	SSR dentro da enzima poligalacturonase	F-CAACCTCCTAGGCTCCTATT R-CAAGACTCTTCCAGCTTCAC	132	53	Participa de processos de modificações da parede celular (maturação de frutos, deiscência, patogênese e expansão celular).
AI973715 ^a	SSR-gene rico em lisina e ác. glutâmico	F-TGTAACCGTGTACACCAAAA R-CACCATCTGGGAACAAGTAT	480	63	-
AI973690 ^a	Citocinina sintase	F-AGAAGATTCACCGCCTTTAT RGTAGTAGTGTCATCGGAGGC	300	55	Enzima da via de síntese de citocininas, hormônios vegetais com atividade sobre a divisão celular.
AI973624 ^a	WD-repeat protein	F-AATGGATAACGATTATGAAA R-TGATAAGTGTTCTTCAGGTC	300	53	-
AI973769 ^a	Poligalacturonase	F-TAGGCACATAAGTGGTCTT R-ATTCTTTATGACAATGTTGG	500	53	Participa de processos de modificações da parede celular (maturação de frutos, deiscência, patogênese e expansão celular).
M16772 ^d	Urease	F-GGCAAACCTGCTGATAAGATGAAGTC R-CGCCTGCAAACCTGGTCTAAATGA	500	65	Enzima cuja função principal em plantas é reciclar o N oriundo da uréia, possível papel na degradação de ureídeos em soja com FBN
M10594 ^b	Uricase II (ENOD35) (nodulina tardia)	F-TCCTTGCTTTGTTTGTGTTAG R-ATTGCCATTTTATTTGTTTGATT	700	55	Enzima relacionada com o metabolismo de ureídeos nas células nodulares não infectadas.
*M10595 ^b	ENOD 24H (nodulina tardia)	F-AGTCCTAATAACAAGAGTAGTA R-GCAGCAGCCATGCCATGTAG	800	50	Proteína da membrana peribacteróide, provável função no transporte de íons.
L20310 ^b	Nodulina 20 (nodulina inicial)	F-CATGCCATCTCTCAATCAAT R-ATGTGAAGCCTGTTATAT	500	57	Provável papel na reorganização da parede celular durante a formação do cordão de infecção e diferenciação celular na formação nodular.
SOYCHS7A ^b	Chalcona sintase	F-CTGCGAATCATTCTTGAAATAAATAA R-TTTGGTGAGTTGGTAATCAG	530	59	Enzima da via central de síntese de isoflavonas.
*SOYAKI70 ^b	Nodulina ENOD 70 (nodulina tardia)	F-CCTGGGGTTTTGATAGTCTCTT R-GAAAGGCAATTGTTGTTTGTTA	750	60	Proteína da membrana do simbiossoma, possível papel no transporte de (NO ₃).
AB003908 ^b	Fosfoenolpiruvato carboxilase	F-GATGGTCGGAAATGAAAGAAG R-TACCCCAAAACCTGCCAACATC	630	59	Enzima da via de síntese de malato e succinato utilizados na respiração pelos bacteróides.
AF195819 ^b	Isoflavona sintase 2	F-GCGGGGTTTTCTTGACACTTT R-CTTGCCATACATTGAAGAGAATC	600	63	Enzima que catalisa a síntese das isoflavonas (Daidzeína e Genisteína).

Tabela 2 - Lista das sequências ESTs e dos marcadores microssatélites dentro de nodulinas conhecidas, que foram utilizadas para buscar polimorfismos entre os parentais Bossier e Embrapa 20

(conclusão)

N de acesso no NCBI	Gene - características	Primers (5' – 3') ^e	pb ^f	T °C ^g	Função
AB018378 ^b	Nodulina inicial ENOD93 (nodulina tardia)	F-CGCTCCAATTCCTTTATTGTTTATT R-CGGGACTAACAACACTATTCACAGAA	600	56	Possível papel na regulação da adaptação dos nódulos a estresse osmótico.
X16875 ^b	Nodulina 2A (ENOD75) (nodulina inicial)	F-ATCCACCATCCAAGAAAACTAA R-AGACCAAGCAACTTTACACATTATCA	400	60	Nodulina inicial rica em prolina, com provável função na morfogênese do nódulo
DY762978 ^c	Nodulina GmN36b (nodulina inicial)	F-ATGGTGTGAGAGGACTAATGCTGG R-ATCTCTGTTGGGTGTGTTACGAGG	430	60	Localizada nas células do periciclo e nas células do sistema vascular, com provável função na morfogênese e no sistema vascular do nódulo.
TC213078	SSR dentro de nodulina	F-ACCTAACTTGCGTGGGAGTG R-TTTGTCTTGTTCCCTCCGAGC	300	60	-
M94764 ^e	SSR dentro da nodulina 26 (nodulina tardia)	F-GAACAAAATGCCTGAGTT R-CCATTCTCAATTCTTCTAGTT	150	60	Proteína de canal (aquaporina) membrana peribacteróide, possível papel na osmoregulação
V00452 ^e	SSR dentro da leghemoglobina	F-GTGGTTTCAGTGAGTGATC R-AAGGTGGAGATGAACTCA	150	50	Nodulina tardia abundante nos nódulos funcionais com função na regulação do transporte de O ₂ .
AF180335 ^b	Malate desidrogenase (Mdh1)	F-TTGCTCTGAACACACTTCTAACA R-TTTTGCCTAACAACATCAAG	400	58	Catalisa a reação oxalacetado - malato, fornecendo o C para necessário para a geração de energia durante a FBN

^aJeong et al. (2006); ^bZhu et al. (2003); ^cLestari et al. (2006); ^dHisano et al. (2007); ^eCregan et al. (1999). ^fF – Forward, R – Reverse. ^gPb- tamanho do fragmento em pares de bae, ^gT°C- temperatura de anelamento do *primer* para amplificação ; * indicam quais ESTs foram polimórficas para os genitores parentais Bossier e Embrapa 20.

programas Phred (Ewing e Green, 1998), Phrap (Ewing et al., 1998) e Consed (Gordon, Abajian e Green 1998). As seqüências obtidas foram submetidas ao banco de dados do NCBI (National Center Biotechnology Information – <http://www.ncbi.nlm.nih.gov/>) para verificação do alinhamento com o gene correspondente.

As seqüências obtidas com a amplificação dos dois genótipos genitores foram analisadas visualmente para verificação da qualidade das seqüências e identificação de possíveis sítios polimórficos com o programa Consed (Gordon, Abajian e Green 1998). Dezessete ESTs foram sequenciadas e destas, apenas na EST da nodulina 70 (ENOD 70) foi verificado polimorfismo de base entre os dois genitores; essa alteração resulta na geração de um sítio de restrição para a enzima *Alu I* (AG/CT). A clivagem do produto de PCR com a enzima *Alu I* gerou o polimorfismo de tamanho de fragmentos entre os genitores, revelado em gel de agarose 1,5%. Posteriormente essa abordagem foi utilizada para genotipar a população de linhagens. As reações de clivagem com a enzima de restrição continham 40 ng de DNA, 0,4 U da enzima *Alu I*, um µL de tampão (específico para enzima) em um volume de 10 µL e foram incubadas por uma hora a 37°C.

Outro polimorfismo observado foi no tamanho dos fragmentos gerados para a EST da nodulina 24H (ENOD 24H). Quando os genótipos genitores (Bossier e Embrapa 20) foram amplificados com o par de *primers* desta nodulina e visualizados em gel de agarose, observou-se que foram produzidos fragmentos de tamanhos diferentes. Para essa nodulina a genotipagem foi feita pela amplificação e visualização direta dos produtos gerados em gel de agarose 1,5%, sem a necessidade da etapa de sequenciamento e clivagem com enzimas de restrição.

2.2.2.3 Construção do Mapa de Ligação

Para a obtenção do mapa genético foram utilizados os dados de segregação de 105 marcadores microssatélites e dos dois marcadores baseados em ESTs (ENOD 70 e ENOD 24H), amplificados numa população de 157 $F_{2:7}$ RILs. Os dados foram analisados por meio do programa MAPMAKER/EXP versão 3.0 (LANDER et al., 1987).

Os critérios adotados para a formação dos grupos de ligação consistiram em um LOD de 3,0, uma distância entre marcas adjacentes de 50 cM e a função de mapeamento de Kosambi (1944) foi empregada para a conversão das unidades de recombinação em distâncias genéticas.

A ligação entre duas marcas foi calculada usando a razão de chances (*odds ratio*), ou seja, a razão entre ligação vs não ligação, que é usualmente expressa como logaritmo da razão (*logarithm off odds*) ou LOD score. Um LOD de 3,0 significa que a ligação é mil vezes mais provável (1000:1) do que a hipótese nula (não ligação) sendo um critério geralmente adotado para a obtenção dos mapas genéticos (RISCH, 1992). Os comandos “group”, “compare” e “map”, foram empregados para formar os grupos de ligação, estabelecer a ordem dos marcadores dentro dos grupos e inferir as distâncias, respectivamente. Os marcadores que pertenciam a grupos de ligação conhecidos, de acordo com o mapa genético (CREGAN et. al., 1999), foram inseridos por meio dos comandos “try” e “riple”.

2.2.3 Mapeamento de QTLs

Os dados de literatura, bem como os resultados desse trabalho, demonstraram que os caracteres relacionados com a fixação biológica de nitrogênio (MNS, NN, MNS/NN e MPAS) são correlacionados e em função disso, o método de mapeado empregado foi o método de mapeamento por intervalo composto para múltiplos caracteres (mCIM) proposto por Jiang e Zeng (1995). Esse método é uma extensão do modelo de mapeamento por intervalo composto (CIM) desenvolvido por Zeng (1993) para caracteres individuais e permite realizar análises de múltiplos caracteres simultaneamente, o que resulta em maior precisão dos testes estatísticos, e em melhor entendimento do controle genético dos caracteres correlacionados. Para testar se há um QTL num intervalo entre dois marcadores controlando pelo menos um dos m caracteres avaliados, foi usado o modelo proposto por Jiang e Zeng (1995) e adaptado para RILs como segue:

$$Y_{jm} = b_{0m} + b_m^* x_j^* + \sum_k^t (b_{km} x_{jk}) + e_{jm}$$

em que:

Y_{jm} é o valor fenotípico da j -ésima linhagem para o m -ésimo carácter ($j= 1, 2, \dots, 157, m= 1, 2 \dots 4$);

b_{0m} é o efeito médio do modelo para o carácter m ;

b_m^* é o efeito aditivo (a^*) do provável QTL sobre o carácter m ;

x_j^* é a probabilidade condicional da variável indicadora θ_m , considerando os valores genotípicos observados para os locos do marcador m e $(m+1)$ que flanqueiam o provável QTL (θ_m assume valores de 1 ou 2, dependendo do genótipo da linhagem para o QTL, se homocigoto para o genitor 1 ou genitor 2, respectivamente);

b_{km} é o coeficiente de regressão parcial entre os valores fenotípicos e os marcadores (k), assumindo t marcadores incluídos no modelo como cofatores;

x_{jk} é a variável indicadora que assume valores 1 e 2, dependendo do genótipo do marcador k , assumindo t marcadores selecionados como cofatores;

e_{jm} é o efeito residual do modelo.

As análises de mapeamento de QTL foram realizadas utilizando o modelo 6 do módulo *JZmapqtl* do programa QTLCartographer versão 1.17 (BASTEN; WEIR; ZENG, 2003). O número de cofatores utilizados foi obtido pela soma dos cofatores que foram selecionados para cada carácter individualmente pelo método *stepwise (forward/backward)*, com probabilidade de entrada e saída de 5%. Para determinar a presença de QTL em cada posição e a estimativa dos efeitos genéticos aditivos (b), foram consideradas as hipóteses:

$$H_0 : b_1 = b_2 \dots = b_m = 0$$

$$H_1 : \text{pelo menos um } b \neq 0$$

As hipóteses foram testadas com base na estatística razão de verossimilhança (*Likelihood-Ratio*, LR) dada por:

$$LR = \left(\frac{L_0}{L_1} \right) \quad (4)$$

em que:

L_0 é o valor da função de verossimilhança sob a hipótese de nulidade H_0 ;

L_1 é o valor da função de verossimilhança sob a hipótese alternativa H_1 .

O teste da razão de verossimilhança (LR) foi realizado a cada 1 cM de ligação (*walking speed*), com uma janela (*windows size*) de 5 cM. Os valores obtidos para os testes LR em cada posição foram plotados em um gráfico onde os pontos no eixo das abscissas corresponderam à posição em que os testes foram realizados, em cM, e os pontos no eixo das ordenadas corresponderam aos valores da LR em cada posição, sendo obtida a curva LR x posição para a análise conjunta para múltiplas características. A partir da curva LR x posição foi realizado o mapeamento de QTL, como segue: para cada intervalo entre marcadores foi observado o maior valor da curva LR x posição (pico) e, caso esse valor ultrapassasse o limite crítico, ou *threshold*, definido para a análise conjunta, o QTL era declarado presente naquele ponto.

Para a posição onde o QTL era detectado na análise conjunta, foi observado o valor de LR correspondente à análise individual de cada caráter. Para cada caráter o efeito do QTL foi considerado significativo se o valor de LR do caráter excedesse o limite crítico estabelecido para a análise individual.

2.2.3.1 Obtenção dos Limites Críticos (Threshold)

O valor do limite crítico (*threshold*) apropriado para declarar a presença de um QTL para múltiplas características foi ajustado pelo número de testes independentes realizados no genoma, seguindo procedimento descrito por Vieira et al. (2000). O número de testes independentes para o mapeamento de QTL foi estimado por:

$$N_t = \sum_i [(T_i / 26) + 1] \quad (5)$$

Em que:

N_t é o número de testes independentes para presença do QTL;

T_i é o comprimento do i -ésimo grupo de ligação formado na elaboração do mapa.

Os grupos de ligação correspondem aos grupos formados no mapeamento e o valor 26 é resultante do valor do intervalo médio que foi de 16,8 cM, mais o valor do “*window size*”, utilizado na análise, que foi de 5 cM, de cada lado do intervalo. Assim, foram considerados independentes os testes realizados a cada 26,8 cM (16,8+5+5). E o número de testes independentes para cada grupo de ligação foram 3,35; 3,20; 3,34; 1,69; 0,26; 1,61; 0,26; 2,03; 0,30; 1,12; 3,04; 1,12; 3,44; 2,10; 1,66; 0,93; 4,38; 5,20; 3,10 e 4,65 que correspondem a aproximadamente 47 testes independentes no total. Assim, a probabilidade de erro tipo I $\alpha=0,001$ (0,05/47) foi usada na análise conjunta. O teste LR possui distribuição aproximada de χ^2 com m graus de liberdade (m =número de caracteres avaliados); assim, considerando o valor de $\alpha= 0,001$ e quatro graus de liberdade tem-se o valor de LR de 18,43 que corresponde ao valor de LOD de 4,01 (18,43 x 0,2172). Para as análises individuais foi considerado também um $\alpha= 0,05$ com um grau de liberdade (número de classes genotípicas – 1); assim, foi obtido o valor de LR de 3,8 que corresponde ao LOD de 0,83. Adicionalmente para a análise conjunta foram considerados dois níveis de significância; considerando $\alpha=0,01$ e $\alpha=0,05$ que corresponderam aos valores de LOD de 2,88 e 2,06 respectivamente. Assim, os QTLs detectados pela análise conjunta foram considerados como mapeados ($LOD \geq 4,01$), prováveis ($2,88 < LOD < 4,01$) e como possíveis ($2,06 < LOD < 2,88$).

A proporção da variância fenotípica total (R^2) explicada pelos QTLs mapeados para cada caráter foi obtida a partir do ajuste de um modelo com todos os QTLs mapeados para o caráter simultaneamente, usando a versão Windows do QTL Cartographer (WANG; BASTEN; ZENG, 2005). Os sinais dos efeitos aditivos indicam a origem do alelo que aumenta a expressão de cada caráter: se positivo, os alelos são provenientes do cultivar Bossier; se negativo, do cultivar Embrapa 20.

2.3 Resultados e Discussão

2.3.1 Análise Fenotípica

As estimativas descritas a seguir foram obtidas por Santos, Nicolás e Hungria (2006), e correspondem aos valores médios para os caracteres massa de nódulos secos (MNS), número de nódulos (NN), massa média de nódulos (MNS/NN) e massa da parte aérea seca (MPAS), avaliados na população de 157 RILs $F_{2:7}$ e nos cultivares genitores Bossier e Embrapa 20. O cultivar genitor Bossier apresentou médias superiores, embora não diferindo estatisticamente para todas as características: MNS ($699,93 \pm 331,77$ mg planta⁻¹), NN ($206,40 \pm 86,96$ nódulos planta⁻¹), MNS/NN ($4,04 \pm 2,28$ mg nódulo⁻¹) e MPAS ($10,19 \pm 2,35$ g planta⁻¹), em comparação ao genitor Embrapa 20: MNS ($529,96 \pm 347,30$ mg planta⁻¹), NN ($183,57 \pm 95,81$ nódulos planta⁻¹), MNS/NN ($2,75 \pm 1,28$ mg nódulo⁻¹) e MPAS ($9,11 \pm 4,02$ g planta⁻¹). Para as linhagens, as médias foram: MNS ($754,12 \pm 165,34$ mg planta⁻¹), NN ($216,00 \pm 46,78$ nódulos planta⁻¹), MNS/NN ($3,78 \pm 0,74$ mg nódulo⁻¹) e MPAS ($11,47 \pm 1,51$ g planta⁻¹). Os quatro caracteres apresentaram uma distribuição normal e nas análises de variância os tratamentos foram significativos para todas as características, caracterizando o conjunto de dados como satisfatório para o mapeamento de QTLs. O desempenho dos genitores confirma as informações prévias descritas por Hungria e Bohrer (2000) e Nicolás, Arias e Hungria (2002), que classificam o cultivar Bossier como um cultivar com capacidade de FBN alta e o cultivar Embrapa 20 como intermediário quanto à FBN.

Santos, Nicolás e Hungria (2006), também estimaram herdabilidades no sentido amplo (h^2) de 49% para MPAS; 33% para MNS; 30% para NN e 27% para MNS/NN. Na literatura, são descritas estimativas de herdabilidade para FBN que variam de 0% a 82% (GREDER, ORF E LAMBERT, 1986; HERRIDGE; DANSO, 1995; PAZDERNIK, GRAHAM E ORF, 1996). Nicolás, Hungria e Arias (2002) avaliaram a herdabilidade dos caracteres de FBN em quatro populações F_3 derivadas do cruzamento entre cultivares brasileiros, em condições experimentais semelhantes às dos dados descritos nesse trabalho e obtiveram estimativas de herdabilidade que variaram de 13% a 58% para NN, 25% a 70% para MNS, 20% a 47% para MNS/NN e de 53% a 68% para MPAS. A

herdabilidade é um parâmetro que expressa a proporção da variação fenotípica que é herdável, quantificando quanto da variação fenotípica é atribuída a variação genética (FALCONER e MACKAY, 1996). Valores baixos de herdabilidade para um dado caráter indicam uma base genética complexa, ou uma baixa precisão experimental, que também compromete as estimativas da variação genética, resultando em valores baixos de herdabilidade. Esses dois fatores são reais para a FBN, onde a avaliação dos caracteres é trabalhosa e para os quais são, frequentemente, relatados coeficientes de variação experimental superiores a 30% (HUNGRIA e BORHER, 2000; SANTOS, NICOLÁS e HUNGRIA, 2006; SOUZA et al., 2008). Na revisão sobre estudos de genética da FBN realizada por Herridge e Danso (1995), foi ressaltado que os caracteres avaliados como componentes do desempenho simbiótico estão sob o controle genético e que a variabilidade genética existe, mas que, a variação ambiental é um fator que mascara o componente genético.

2.3.2 Análise de Correlação

As informações sobre a correlação entre caracteres são importantes nos programas de melhoramento, permitindo a previsão de mudanças no desempenho de um determinado caráter quando a seleção é praticada em outro caráter. Segundo Falconer e Mackay (1996), a associação entre caracteres pode ter causas genéticas e ambientais. São causas genéticas, a ligação gênica e a pleiotropia (um gene controlando mais de um caráter simultaneamente), sendo que a ligação gênica pode ser quebrada pela recombinação, enquanto que a correlação devido à pleiotropia é permanente.

Para os quatro caracteres avaliados (MPAS, MNS, NN, MNS/NN) os coeficientes de correlações genéticas e fenotípicas foram similares, com valores superiores para as correlações genéticas (Tabela 3). Correlações altas e positivas foram encontradas entre os caracteres MPAS e MNS ($r_G = 0,83$ e $r_F = 0,66$), MNS e NN ($r_G = 0,65$ e $r_F = 0,64$), medianas entre MPAS e NN ($r_G = 0,51$ e $r_F = 0,43$) e baixas entre MPAS e MNS/NN ($r_G = 0,26$ e $r_F = 0,24$) e MNS e MNS/NN ($r_G = 0,23$ e $r_F = 0,33$). Para NN e MNS/NN, as correlações foram negativas ($r_G = -0,60$ e $r_F = -0,40$). A seleção de caracteres é baseada

nos dados fenotípicos, mas valores de correlações genéticas superiores às correlações fenotípicas indicam que o fenótipo é um indicador seguro do genótipo

Tabela 3 – Coeficientes de correlação genética (acima da diagonal) e fenotípica (abaixo da diagonal) entre os caracteres relacionados com a FBN

Caráter^a	MPAS	NN	MSN	MNS/NN
MPAS	-	0,51	0,83	0,26
NN	0,43	-	0,65	-0,60
MNS	0,66	0,64	-	0,23
MNS/NN	0,24	-0,40	0,33	-

^a MPAS, NN, MNS e MNS/NN referem-se à massa da parte aérea seca (g planta^{-1}), número de nódulos ($\text{nódulos planta}^{-1}$), massa de nódulos secos (mg planta^{-1}), massa média de nódulos (mg nódulo^{-1}).

Estimativas sobre as correlações fenotípicas entre os caracteres de nodulação e o N_2 fixado nos tecidos das plantas realizados por outros estudos indicam que esses caracteres são correlacionados positivamente. Sinclair et al. (1991) estimaram coeficientes de correlação altos entre os caracteres MNS e NN ($r_F = 0,84$ a $0,91$) e correlações medianas a altas para os caracteres de nodulação e características da planta (massa de raiz e da parte aérea), variando de $0,31$ a $0,67$ para massa da raiz e de $0,40$ e $0,70$ para massa da parte aérea. Dados similares foram relatados por Hungria e Borher (2000) em um experimento com 152 cultivares avaliados em casa de vegetação, onde foram verificados coeficientes de correlação de $0,58$ (MNS e NN), $0,66$ (MNS e MPAS) e $0,31$ (NN e MPAS). Hungria e Borher (2000) também relatam que os caracteres MNS e MPAS foram altamente correlacionados com o N total da parte aérea, NTPA ($0,70$ e $0,91$) respectivamente. Assim esses dois caracteres, MNS e MPAS, constituem bons indicadores da nodulação e do N_2 fixado. Além disso, Souza et al. (2008) estimaram correlações altas entre MNS e NN ($r = 0,91$) e entre MPAS e o NTPA ($r = 0,88$), em experimentos conduzidos em condições de campo em seis Estados brasileiros. Esses autores sugerem que em solos pobres em N e na ausência de fertilização nitrogenada, os caracteres MNS e MPAS são bons indicadores da FBN.

Neste trabalho, a verificação de que os caracteres componentes da FBN são correlacionados foi importante para a escolha do método de análise para o mapeamento de QTL, conforme descrito no item Materiais e Métodos. Para características correlacionadas, recomenda-se o método de mapeamento mCIM, que considera todas os caracteres simultaneamente, aumentando a capacidade de detecção de QTLs e, também, possibilitando inferir as causas genéticas (ligação gênica ou pleiotropia) da ocorrência de QTL coincidentes.

2.3.3 Mapa de Ligação

Os genitores foram avaliados com 432 marcadores microssatélites (anexo 1) e 17 marcadores ESTs (Tabela 2). Destes, 105 marcadores microssatélites (Tabela 1) e 2 ESTs (Enod 70 e Enod 24H) foram polimórficos, resultando em um total de 107 marcadores utilizados para genotipar a população de 157 linhagens. Desses, 97 marcadores SSR foram agrupados nos 20 grupos de ligação da soja (Figura 2), enquanto que oito marcadores microssatélites (Satt421, Satt396, Satt180, Satt184, Satt395, Satt356, Sat_044, Sat_020) e os dois marcadores baseados nas ESTs permaneceram não ligados. O mapa genético obtido tem um comprimento de 1.263,2 cM (Kosambi), o que corresponde a uma cobertura de aproximadamente 50% do genoma, estimado em 2.523,6 cM (SONG et al. 2004). A distância média entre os marcadores foi de 16,4 cM e oito intervalos apresentam distâncias superiores a 30 cM, localizados nos GL F, G, L, M e O. Os grupos formados, bem como a ordem dos marcadores dentro dos grupos foram semelhantes aos grupos de ligação descritos para a soja (CREGAN et al., 1999; SONG et al., 2004). A única exceção foi o marcador GmEKN que, segundo Jeong et al. (2006), pertence ao grupo A2 no intervalo delimitado pelos marcadores (Satt589-Satt187) mas que neste estudo foi agrupado no grupo D1b (LOD >10); esse marcador microssatélite foi desenvolvido dentro de um gene hipotético identificado em estudos de expressão gênica em nódulos de soja. O número médio de marcadores por grupo de ligação foi de cinco. Os grupos com maior número de

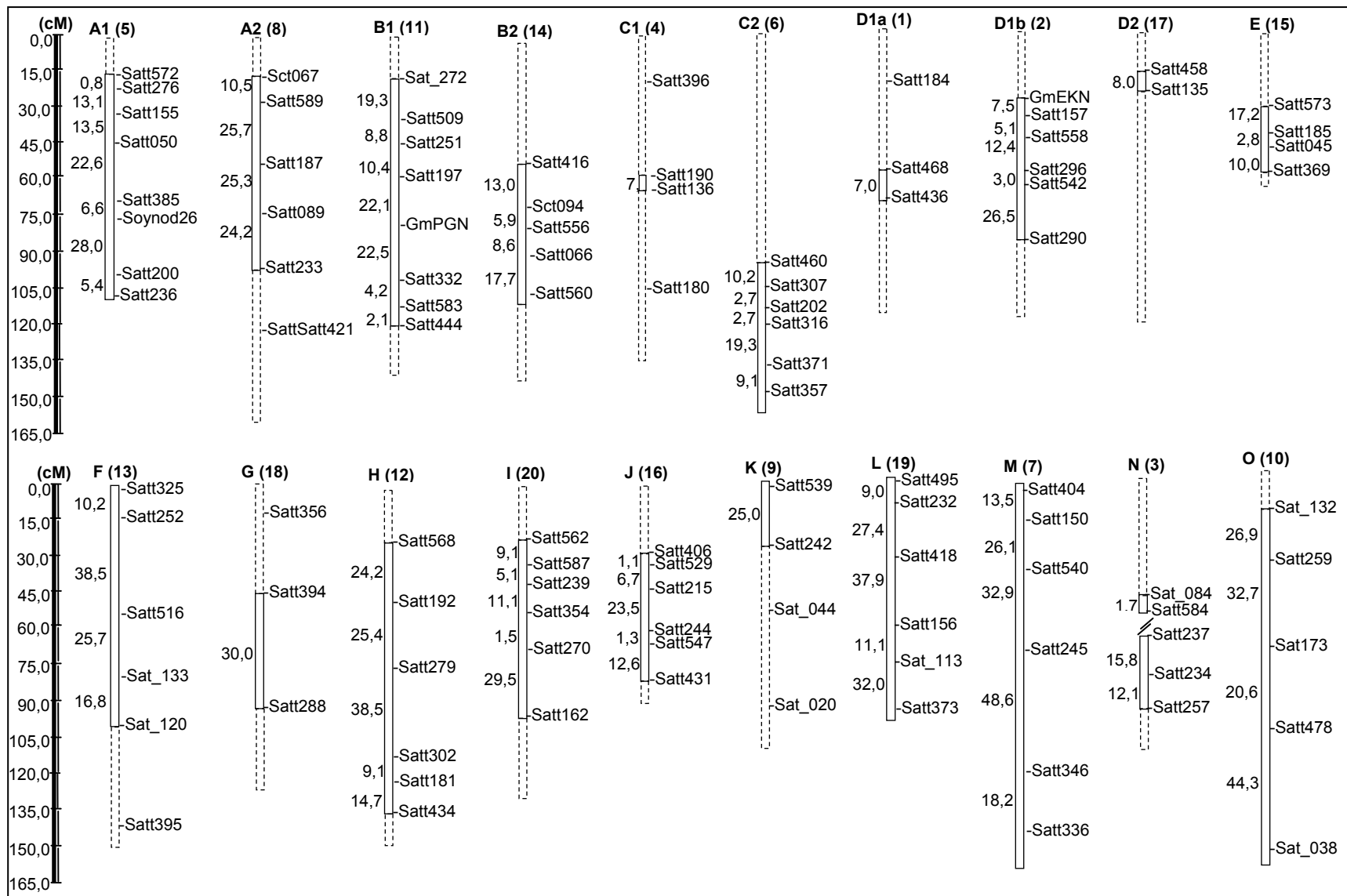


Figura 2 – Mapa genético da soja baseado na população de 157 F_{2:7} RILs derivadas dos cultivares Bossier x Embrapa 20. Os grupos de ligação e marcadores microssatélites foram designados de acordo com o mapa consenso de Cregan et al. (1999) e Song et al. (2004). As barras fechadas e pontilhadas representam as regiões mapeadas e as regiões não mapeadas na população, respectivamente

marcadores foram os A1 e B1 com oito marcadores cada e os grupos com menor número foram os C1, C2, D1a, D2, G e K com dois marcadores cada. Na Figura 2 são destacadas as regiões que foram mapeadas nesse trabalho, comparando com os tamanhos estimados para os GLs, de acordo com o mapa consenso (CREGAN et al., 1999 e SONG et al., 2004), evidenciando as regiões cobertas pelo trabalho e as regiões que ainda precisam ser mapeadas. Coberturas parciais do genoma são relatadas em trabalhos de mapeamento de QTL em soja, como no trabalho de Zeng et al. (2009), que usaram um mapa composto por 99 marcadores e um intervalo médio de 20 cM para mapear QTLs para isoflavonas. Em outros exemplos, Palomeque et al. (2009a) e Palomeque et al. (2009b) utilizaram um mapa com 105 marcadores agrupados em vinte e seis GL para mapear QTLs para produção de grãos de soja, enquanto que Panthee et al. (2004) mapearam QTLs para acúmulo de N em soja com um mapa construído com 86 marcadores. Uma cobertura maior do genoma é necessária para permitir a identificação de novas regiões associadas com as características de interesse. Além disso, a saturação de regiões já mapeadas permite melhorar as estimativas de posição e efeito dos QTLs identificados.

2.3.4 Mapeamento de QTL

Com o mapa genético construído e o método de mapeamento por intervalo composto para múltiplas características (mCIM), que utiliza os dados fenotípicos de todas as características conjuntamente, foram mapeados oito QTLs significativos para pelo menos um dos caracteres nos grupos de ligação (GL) A2, B1, C1, C2, E, G, I e L (Tabela 4). Destes, os QTLs localizados nos GL C1 (FB1), C2 (FBN2), E (FBN3) e I (FBN4) influenciaram mais de um caráter (Tabela 4 e Figura 3) e quatro QTLs foram responsáveis pela expressão de apenas um caráter (Tabela 4 e Figura 4): dois QTLs para MPAS mapeados nos GL G (MPAS1) e L (MPAS2), um QTL para NN no grupo B1 (NN1) e um QTL para MNS/NN no grupo A2 (MNS/NN1). De acordo com os critérios estabelecidos para o nível de significância neste estudo, foi considerado como mapeado o QTL no GL L ($LOD \geq 4,01$), como prováveis os QTL dos GL B1, C1, E, G e I ($2,88 < LOD < 4,01$) e como possíveis os QTL mapeados nos GL A2 e C2 ($2,06 < LOD < 2,88$). Os trabalhos que mapeiam QTLs em soja normalmente adotam

valores de LOD entre 2 e 3 como critérios para determinar uma associação significativa (KASSEM et al., 2006; GUZMAN et al., 2007; GUTIERREZ-GONZALEZ et al. 2009; ZHANG et al., 2009).

Os caracteres de nodulação e de fixação de N_2 , avaliados como medidas do desempenho da FBN são correlacionados e o mapeamento conjunto de QTLs identificou quatro regiões significativas para mais de um caráter, evidenciando que ambos estão sob um controle genético comum (Tabela 4, Figura 3). O QTL FBN1 foi significativo para os caracteres MPAS, MNS e MNS/NN, e os alelos favoráveis para estes caracteres vieram do cultivar Embrapa 20, correspondendo a $-0,39 \text{ g planta}^{-1}$ para MPAS, $-36,66 \text{ mg planta}^{-1}$ para MNS e $-0,14 \text{ mg nódulo}^{-1}$ para MNS/NN e explica individualmente, 3,3%, 2,5% e 2,5%, respectivamente, da variação fenotípica. O QTL FBN2 localizado no GL C2 foi significativo para NN e MNS e explica 1,7% e 3,6% da variação desses caracteres; os alelos favoráveis derivaram do cultivar Bossier e os efeitos aditivos foram de $8,8 \text{ nod planta}^{-1}$ e $28,97 \text{ mg planta}^{-1}$. Para o QTL FBN3 que explica 5,4% e 4,3% da variação para MPAS e NN, os alelos favoráveis vieram do cultivar Embrapa 20 correspondendo a $-0,29 \text{ mg planta}^{-1}$ para MPAS e $-10,23 \text{ nod planta}^{-1}$ para NN. Localizado no GL I, o QTL FBN4 compreende o intervalo entre os marcadores Satt587-Satt354, com efeitos significativos para MNS/NN e NN. Na região delimitada por esse intervalo, foram mapeados dois QTLs para NN, explicando 1,7% e 1,2% da variação e um QTL para MNS/NN, explicando 6,5% da variação fenotípica desse caráter. O alelo favorável para MNS/NN pertence ao cultivar Embrapa 20, com um efeito de $6,5 \text{ mg nódulo}^{-1}$ e os alelos favoráveis para NN são do cultivar Bossier ($8,92$ e $10,76 \text{ nódulos planta}^{-1}$). Para os QTL FBN1 (MPAS, NN, MNS/NN), FBN2 (MNS, NN) e FBN3 (MPAS e NN), as causas genéticas do mapeamento conjunto dos caracteres foram os efeitos pleiotrópicos enquanto que para o QTL FBN4 (MNS/NN e NN) a causa genética da correlação foi ligação gênica, como é evidenciado pela posição dos QTLs individuais (Figura 3) e pela origem dos alelos favoráveis (Tabela 4).

Dos quatro QTLs que influenciam apenas um caráter, dois são para MPAS (MPAS1 e MPAS2) e foram localizados nos grupos G e L. Os alelos do cultivar Bossier aumentam o caráter para ambos QTLs e os valores dos efeitos aditivos foram de $0,29$ e $0,57 \text{ g planta}^{-1}$ (Tabela 4 e Figura 4). O QTL localizado no grupo L (MPAS2), delimitado

Tabela 4 - Mapeamento de QTLs (*Quantitative Trait Loci*) usando o modelo mCIM para os caracteres massa da parte aérea seca (MPAS, g planta⁻¹), número de nódulos (NN, nódulos planta⁻¹), massa de nódulos secos (MNS, mg planta⁻¹) e massa média de nódulos secos (MNS/NN, mg nódulos⁻¹)

Grupo de ligação/QTL	Intervalo	Análise	Posição (cM)	LOD	Efeito	R ²	Ligação versus Pleiotropia
					a	%	
C1/FBN1	Satt190-Satt136	Conjunta		3,2			Evidência de um QTL com efeitos pleiotrópicos para MPAS, MNS e MNS/NN.
		MPAS	7,01	2,7*	-0,39	3,3	
		NN	7,01	0,2	-2,90	1,1	
		MNS	7,01	1,7*	-36,66	2,5	
		MNS/NN	4,01	1,2*	-0,14	2,5	
C2/FBN2	Satt460-Satt307	Conjunta	0,01	2,7			Possível QTL com efeitos pleiotrópicos para NN e MNS.
		MPAS	0,01	0,3	-0,12	0,0	
		NN	9,01	1,6*	8,80	1,7	
		MNS	10,25	1,0*	28,97	3,6	
		MNS/NN	10,25	0,0	-0,03	0,0	
E/FBN3	Satt573-Satt185	Conjunta	8,01	3,6			Provável QTL com efeitos pleiotrópicos para MPAS e NN.
		MPAS	9,01	1,3*	-0,29	5,4	
		NN	4,01	1,5*	-10,23	4,3	
		MNS	13,01	0,2	-9,89	1,1	
		MNS/NN	0,01	0,4	0,07	0,3	
I/FBN4	Satt587-Satt354	Conjunta	14,23	3,6			Evidência de ligação entre os QTLs para NN e MNS/NN
		MPAS	14,23	0,4	-0,16	2,0	
		NN	9,12	1,28*	8,92	1,7	
		NN	23,23	1,5*	10,76	1,2	
		MNS	14,23	0,6	-20,81	2,1	
		MNS/NN	13,12	2,5*	-0,21	6,5	
G/MPAS1	Satt394-Satt288	Conjunta	30,01	3,3			QTL significativo somente para MPAS
		MPAS	30,00	1,3*	0,29	1,2	
		NN	30,00	0,2	-3,47	1,1	
		MNS	0,01	0,3	14,80	0,8	
		MNS/NN	0,01	0,3	-0,07	0,6	
L/MPAS2	Satt232-Satt418	Conjunta	18,00	4,2			QTL significativo somente para MPAS
		MPAS	20,00	2,9*	0,57	10,0	
		NN	9,00	0,0	-1,20	0,2	
		MNS	36,00	0,5	19,52	0,6	
		MNS/NN	24,00	0,3	0,09	0,5	
B1/NN1	Satt509-Satt251	Conjunta	23,28	3,8			QTL significativo somente para NN
		MPAS	38,12	0,3	-0,13	0,2	
		NN	23,28	3,0*	-13,80	7,8	
		MNS	23,28	0,5	-22,21	0,9	
		MNS/NN	23,28	0,7	0,09	2,8	
A2/MSN/NN	Sct067-Satt589	Conjunta	0,01	2,7			QTL significativo somente para MNS/NN
		MPAS	0,01	0,7	0,20	0,8	
		NN	0,01	0,1	1,60	0,2	
		MNS	0,01	0,7	-21,31	0,2	
		MNS/NN	0,01	1,7*	-0,15	3,0	

Para a análise conjunta, foram estabelecidos três níveis de significância: LOD 4,01, 2,88 e 2,06 e de acordo com esses critérios os QTLs foram considerados como mapeados, prováveis e possíveis (ver material e métodos). * indica o caráter que é influenciado pelo QTL mapeado pela análise conjunta; a – corresponde ao efeito aditivo e % - a proporção da variação fenotípica que é explicada pelo caráter individualmente.

pelos marcadores Satt232-Satt418 foi o único que atingiu o limite crítico determinado com base nos dados experimentais ($LOD > 4,01$), sendo considerado como QTL mapeado. Esse QTL também foi o que apresentou variação individual maior, explicando 10,0% da variância fenotípica. O alelo favorável foi o do cultivar Bossier, com um efeito aditivo de $0,57 \text{ g planta}^{-1}$. Associados a essa região, Nicolás Arias e Hungria (2006) mapearam um QTL para MPAS e Orf et al. (1999) mapearam um QTL para altura de planta e acamamento. Dos outros dois QTLs, um foi significativo para NN e o outro para MNS/NN. O QTL para NN foi mapeado no GL B1 e explica 7,8% do caráter, com o alelo favorável derivado do cultivar Embrapa 20 e o valor do efeito aditivo foi de $-13,8 \text{ nódulo planta}^{-1}$. Para o caráter MNS/NN, um possível QTL mapeado no GL A2 respondendo por 3% da variação com alelos favoráveis do cultivar Embrapa 20 ($-0,15 \text{ mg planta}^{-1}$).

Foram identificados QTLs em oito regiões genômicas sendo que quatro regiões foram associadas a mais de um caráter. No total foram mapeados quatro QTLs para o caráter MPAS (C1, E, G e L) cinco para NN (B1, C1, C2, E, G), dois para MNS (C1, C2) e três para MNS/NN (A2, C1, I) que explicaram 23,0%, 20,0%, 11,8% e 16,0% respectivamente, da variação fenotípica total dos caracteres. Resultados similares são relatados em outros estudos de mapeamento para FBN em soja (TANYA et al., 2005; NICOLÁS, HUNGRIA e ARIAS, 2006), como exemplo, no trabalho de Tanya et al. (2005) foram identificadas cinco regiões genômicas nos GL A1, I, J, D1b e O associadas a mais de um caráter e explicaram 28,1% da variação fenotípica total para MPAS, 28,7% para MNS e 33,5% para NN. Os resultados deste trabalho bem como os resultados relatados por outros autores, evidenciam a natureza complexa da FBN, pois os fenótipos resultantes são fortemente influenciados pelo ambiente. No entanto, considerando os valores de herdabilidade dos caracteres MPAS (49%), NN (30%), MNS (33%), e MNS/NN (27%) (SANTOS, NICOLÁS e HUNGRIA, 2006) e a cobertura genômica de 50%, constata-se que é possível identificar novas regiões controlando os caracteres de FBN, principalmente para MPAS que possui herdabilidade mais alta. QTLs para MPAS, MNS e NN foram mapeados por Tanya et al. (2005) nos grupos A1, I, J e O, isto é, em regiões não cobertas pelo mapa genético do presente trabalho e constituem regiões que deverão ser priorizadas na seqüência do trabalho.

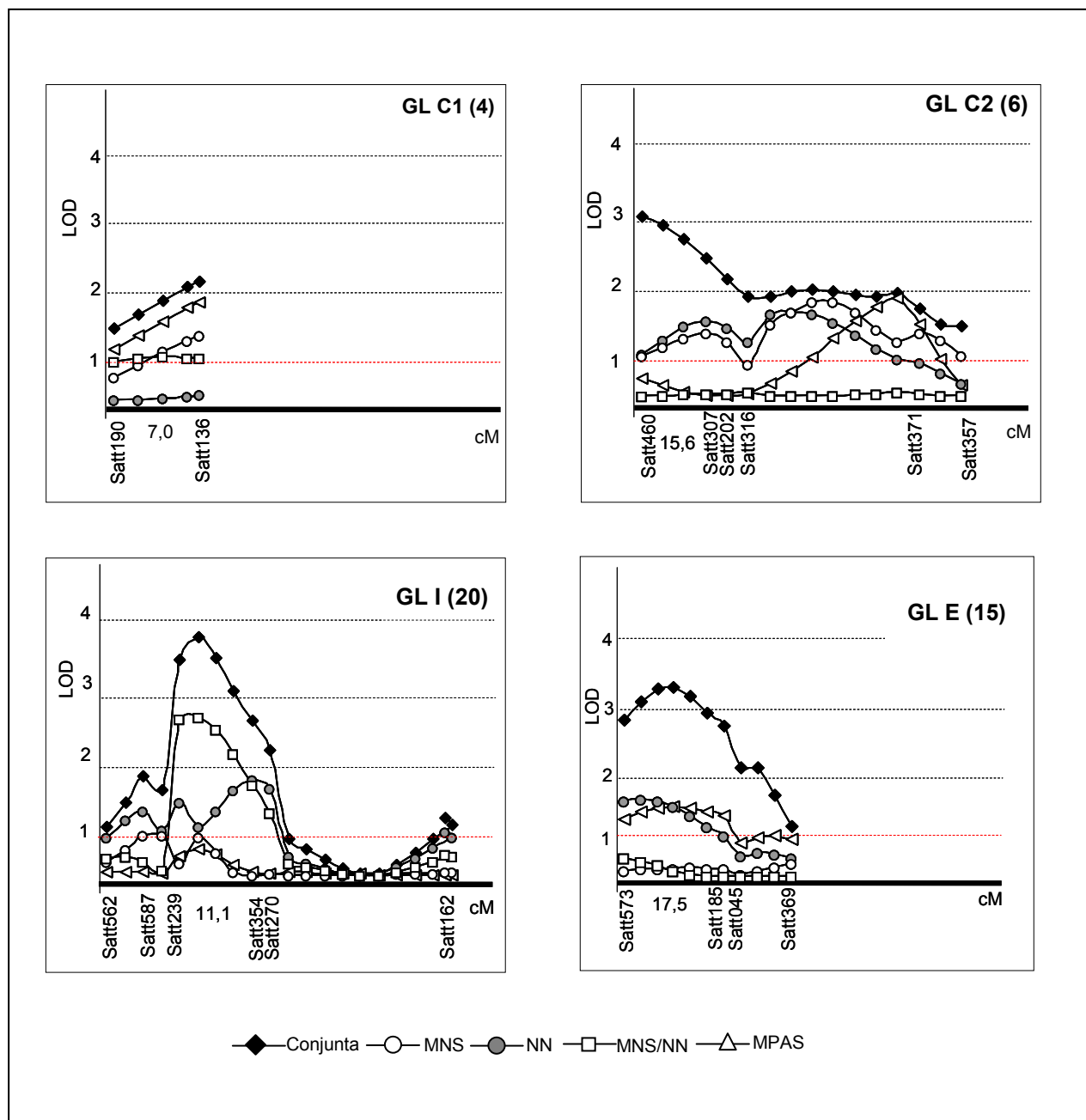


Figura 3 – QTLs significativos para mais de um caráter, identificados pelo método mCIM. As linhas pontilhadas em preto representam os níveis de significância para a análise conjunta e a linha em vermelho o nível de significância estabelecido para a análise individual. MPAS – massa da parte aérea seca, MNS – massa de nódulos secos, NN – número de nódulos, MNS/NN – massa média de nódulos secos

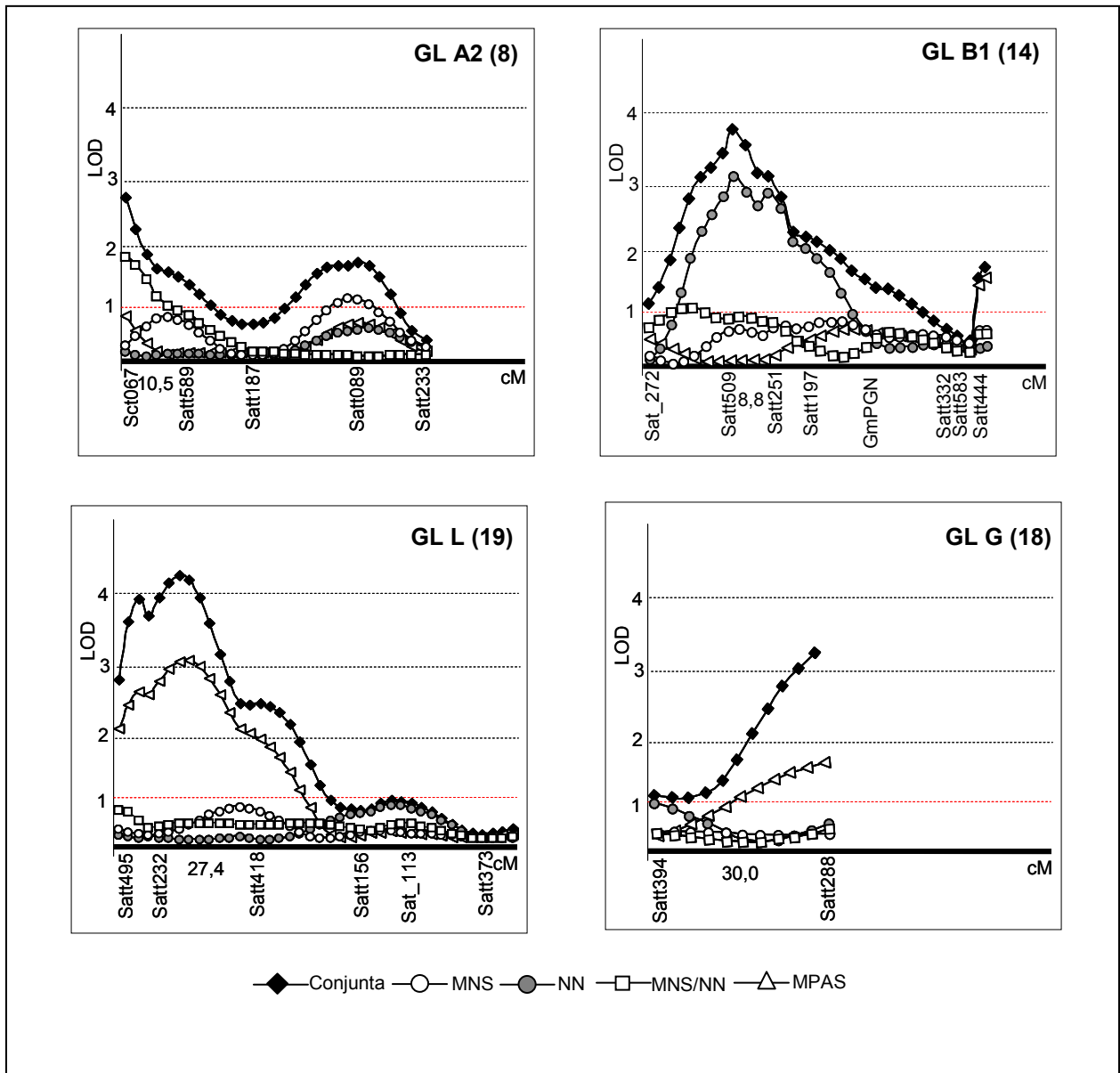


Figura 4 – QTLs significativos para um caráter, identificados pelo método mCIM. As linhas pontilhadas em preto representam os níveis de significância para a análise conjunta e a linha em vermelho o nível de significância estabelecido para a análise individual. MPAS – massa da parte aérea seca, MNS – massa de nódulos secos, NN – número de nódulos, MNS/NN – massa média de nódulos secos

Poucos trabalhos sobre mapeamento de QTLs para FBN são relatados na literatura, podendo-se citar estudos com soja (TANYA et al., 2005; NICOLÁS, ARIAS e HUNGRIA, 2006; SANTOS et al., 2006) e feijão comum (NODARI et al., 1993; TSAI et al., 1998 e SOUZA et al., 2000). Além disso, nos trabalhos de mapeamento para FBN em soja citados acima, foram empregados métodos de mapeamento baseados em análise de marcas simples (análise de regressão e análise de variância). Desse modo, este trabalho constitui um passo importante para a compreensão da base genética da FBN em soja, identificando oito regiões genômicas relacionadas com a expressão dos caracteres componentes da FBN, sendo estimados os efeitos aditivos, a proporção da variação fenotípica explicada pelos QTLs, a origem dos alelos envolvidos, bem como as causas das correlações entre os caracteres. Todos os 14 QTLs, sendo quatro para MPAS, cinco para NN, dois para MNS e três para MNS/NN foram de pequeno efeito (a variação do R^2 individual foi de 1,2% a 10,0%). Valores de R^2 individuais pequenos (<8%) foram relatados por Tanya et al. (2005), Nicolás, Arias e Hungria (2006) e Santos et al. (2006) para os caracteres de FBN na soja. E em feijão comum, Souza et al. (2000), Tsai et al. (1998) e Nodari et al. (1993) identificaram QTLs para NN que apresentaram valores de R^2 individuais inferiores a 10% e explicaram, conjuntamente, em média, 35,5% da variação fenotípica.

Ambos os genitores contribuíram para o aumento dos caracteres (Tabela 4), sendo que em quatro regiões os alelos favoráveis vieram do cultivar Embrapa 20 (A2, B1, C1, E) e em três regiões os alelos favoráveis vieram do cultivar Bossier (C2, G e L). A identificação de alelos favoráveis em genitores com menor desempenho para caracteres mapeados é comum nos dados de mapeamento em soja, como pode ser observado, por exemplo, nos trabalhos de Chapman et al., (2003); Li et al., (2005) e Palomeque et al., (2009a) e, explica a segregação transgressiva observada nas linhagens segregantes (SANTOS, NICOLÁS e HUNGRIA, 2006), pela combinação de alelos favoráveis oriundos dos dois genitores.

Dos QTL mapeados, dois foram considerados como possíveis (A2 e C2), cinco como prováveis (B1, C1, E, G e I) e um como mapeado (L). A identificação desses QTLs em outras populações de mapeamento e em outras condições experimentais é necessária para a confirmação dessas associações e estudos das interações com o

ambiente. Os QTLs FBN2 (Figura 3), MPAS2 e NN1 (Figura 4) foram mapeados em regiões onde Nicolás, Arias e Hungria (2006) também detectaram QTLs significativos para NN e MNS em uma população de linhas $F_{2:3}$ do cruzamento entre os cultivares BRS 133 e Embrapa 20, corroborando com os dados deste estudo, de que essas regiões estão envolvidas com o controle genético da FBN. Os QTLs FBN1, MNS/NN e MPAS1 foram mapeados nas extremidades dos grupos de ligação C1, A2 e G (Figuras 3 e 4) e a inserção de novos marcadores nessas regiões é necessária para uma melhor estimativa dos efeitos e da localização desses QTLs.

Nos trabalhos de Nicolás, Arias e Hungria (2006) e Tanya et al. (2005) um mesmo marcador foi geralmente associado a mais de um caráter de FBN e o mesmo padrão foi verificado neste estudo, sendo que das oito regiões mapeadas, quatro regiões (C1, C2, E e I) foram associadas a mais de um caráter (Tabela 4 e Figura 3) e o mapeamento conjunto revelou que, para os caracteres componentes da FBN, as causas dessas coincidências são devidas, principalmente, a efeitos pleiotrópicos dos locos que afetam simultaneamente mais de um caráter. Isto explica as correlações altas e positivas relatadas para os caracteres MNS e NN e para MNS e MPAS. Para o QTL FBN4 (MNS/NN e NN) no grupo I a causa da associação foi a ligação gênica; esses dois caracteres foram correlacionados negativamente, indicando que o aumento no número de nódulos implica na formação de nódulos menores; contudo a evidência de que para o QTL FBN4 a causa da correlação é a ligação gênica, indica ser possível o melhoramento para ambos os caracteres pela quebra da ligação gênica.

A FBN é um fator importante para a produção de grãos das leguminosas cultivadas, resultando em benefícios econômicos e ambientais, pela economia no uso de fertilizantes nitrogenados e pela redução da contaminação de cursos de água com nitrato. A natureza complexa da FBN e as dificuldades encontradas nas avaliações fenotípicas (que incluem a destruição de plantas) são fatores que dificultam a avaliação dessa característica em programas de melhoramento (HERRIDGE e ROSE, 2000). O mapeamento de regiões genômicas controlando a FBN em soja permite a identificação de associações marcador-QTL potencialmente úteis para a seleção assistida por marcadores para os caracteres de FBN.

2.3.4.1 Regiões Coincidentes entre QTLs para FBN e outros Caracteres

Para as regiões onde foram identificados QTLs para FBN, foi feita uma busca no banco de dados SOYBASE (2009) e na literatura, com o objetivo de verificar QTLs descritos para outras características nesses intervalos (Figura 5). A ênfase nessa busca foi para QTLs relacionados com a produção de grãos e o teor de proteína nas sementes, por serem caracteres que demandam grandes quantidades de N. Também foram feitas buscas para os caracteres de produção de isoflavonóides e absorção de P, devido à importância desses dois fatores na comunicação entre planta e bactéria e na geração de energia necessária para o processo de fixação de N₂.

As oito regiões onde foram mapeados QTLs para FBN (A2, B1, C1, C2, G, E, I e L), foram coincidentes com QTL para outros caracteres. O QTL FBN1, localizado no grupo de ligação C1 (Satt190-Satt136) foi coincidente com QTLs para teor de proteína (ORF et al., 1999), número de vagens e área foliar específica (VIEIRA et al., 2006). No grupo C2, na região onde foi mapeado o QTL FBN2 (Satt307-Satt406), foram descritas associações com produção de grãos e com componentes de produção em diferentes populações de mapeamento (ORF et al., 1999; SPECHT et al., 2001; SMALLEY et al., 2004; PALOMEQUE et al., 2009a; PALOMEQUE et al., 2009b) e QTLs para outros caracteres, como massa de raiz seca e eficiência na absorção de P (ZHANG et al., 2009), teor de proteína (HYTEN et al., 2004), isoflavona genisteína (ZENG et al., 2009) e para NN (NICOLÁS, HUNGRIA e ARIAS, 2006). Nessa região também está localizado o locus do marcador morfológico T, que controla a cor da pubescência; segundo Toda et al. (2002), o marcador T codifica a enzima F3'H (flavonóide 3'-hidroxilase) responsável pela hidroxilação da posição 3' dos flavonóides e os autores denominaram este loco como Sf3'h1 (soybean flavonoide 3'-hidroxilase). Takahashi e Asanuma, (1996) associaram o marcador T a uma maior capacidade de FBN (determinação de N₂ fixado) sob condições de baixa temperatura. Na região delimitada pelo QTL FBN3 (Satt573-Satt185) no GL E, Zhang et al. (2009) identificaram QTLs para caracteres relacionados à tolerância ao estresse por deficiência de P (massa da parte aérea seca, eficiência na absorção de P, eficiência na utilização de P e atividade da enzima fosfatase ácida). Também foram descritos QTLs para teor de proteína (LEE et al., 1996), e para absorção

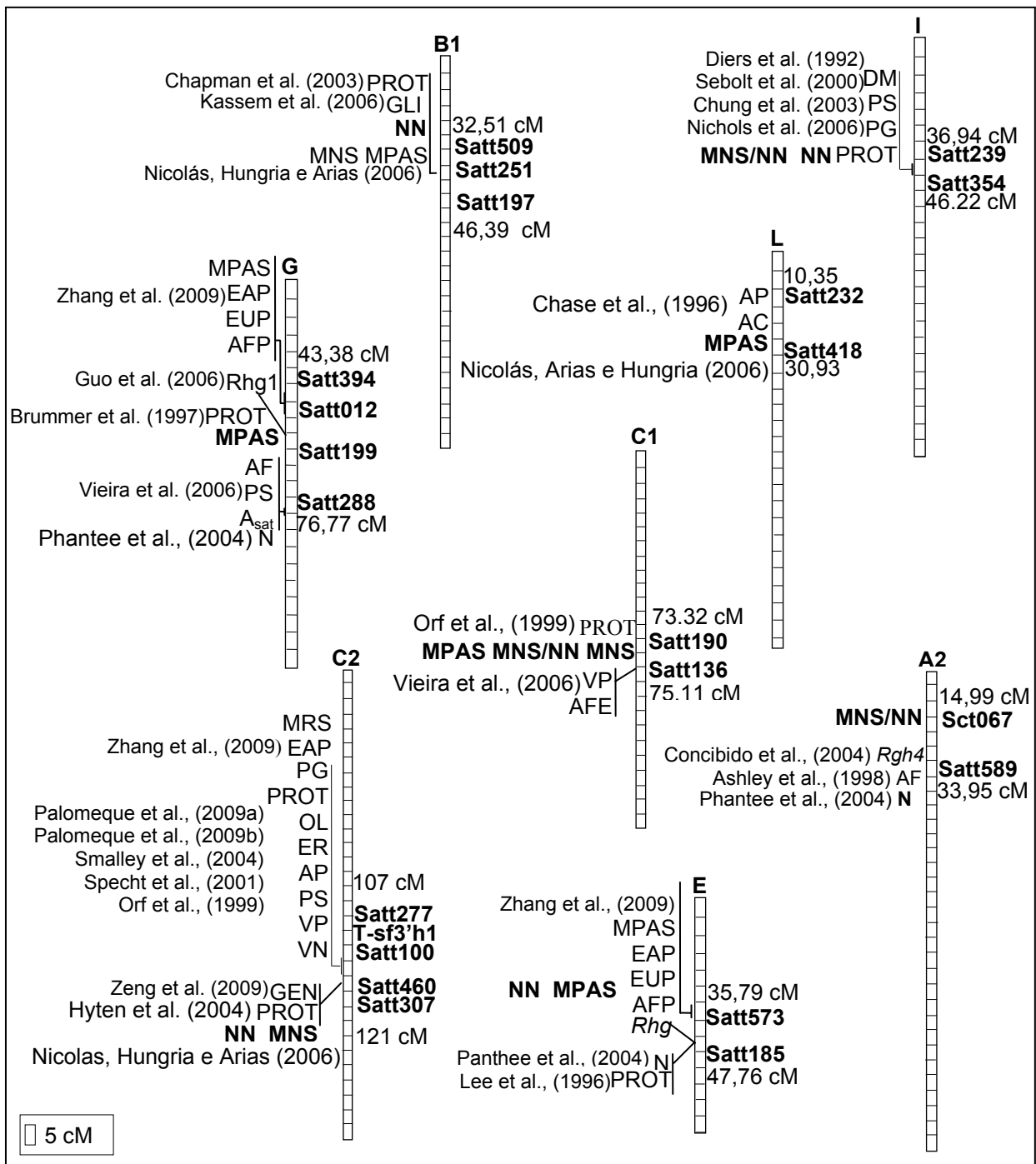


Figura 5 – Co-localização entre QTLs-FBN e QTLs descritos na literatura. MNS–massa de nódulos secos, NN–número de nódulos, MNS/NN–massa média de nódulos secos, MPAS–massa da parte aérea seca, MTP–massa total da planta, PG–produção de grãos, AP–altura de planta, ER–estádio reprodutivo, PS–peso de 100 sementes, VN–vagens por nó, VP–vagens por planta, DM–dias para maturidade, AC–acamamento, PROT–proteína, OL–óleo, GLI–gliciteína, GEN–genisteína, A_{sat}–taxa de assimilação líquida de CO₂, AF–área foliar, AFE–área foliar específica, EAP–eficiência na aquisição de P, EUP–eficiência na utilização de P, AFP–atividade da enzima fosfatase. Os QTL em negrito foram mapeados nesse trabalho

de N (PANTHEE et al., 2004). O QTL FBN4 foi mapeado em uma região do grupo I (Satt239-Satt354) que tem sido extensivamente caracterizada pela localização de um QTL de efeito maior para teor de proteína (DIERS et al., 1992; SEBOLT, et al., 2000; CHUNG et al., 2003 e NICHOLS et al., 2006).

No grupo B1 (NN: Satt509-Satt251) foram descritos QTLs para teor de proteína (CHAPMAN et al., 2003) e para a isoflavona gliciteína (NJITI et al., 1999; MEKSEM et al., 2001; KASSEM et al., 2004 e KASSEM et al., 2006). No grupo G (MPAS: Satt394-Satt288) Zhang et al. (2009) mapearam QTLs relacionados à tolerância ao estresse por deficiência de P (eficiência da absorção de P, eficiência na utilização de P e atividade da fosfatase ácida), bem como um QTL para o caráter MPAS. Também nesta região do grupo G, Brummer et al. (1997) mapearam um QTL para teor de proteína e Vieira et al. (2006) mapearam QTLs para peso de sementes e área foliar. Finalmente, no grupo L, na região onde neste trabalho foi mapeado o QTL para MPAS, Orf et al. (1999) mapearam QTLs para altura de planta e acamamento e Nicolás, Hungria e Arias (2006) identificaram um QTL para MPAS.

O nitrogênio é o nutriente mineral requerido em maior quantidade pelas plantas, participando da constituição de diversas moléculas como proteínas, DNA, RNA, clorofila, bem como na constituição das moléculas de transferência de energia ATP (adenosina trifosfato). Tem, portanto, um papel fundamental na constituição estrutural dos tecidos da planta, bem como, do metabolismo de DNA e na geração de energia para os processos de desenvolvimento da planta. As leguminosas dispõem de duas fontes de nitrogênio: o N disponível no solo, na forma de nitrato e amônia e o N proveniente da FBN. A soja consegue suprir todas as necessidades de N pelo processo da FBN.

Ao verificar se regiões genômicas que controlam o processo de FBN são coincidentes com regiões que controlam outros fenótipos da planta, algumas constatações interessantes foram feitas. Cinco das oito regiões onde foram mapeados QTLs para FBN coincidem com regiões onde foram mapeados QTLs para teor de proteína nos grãos. Dados de literatura indicam que o N oriundo da FBN é deslocado preferencialmente para os grãos. Como exemplo, Israel, Burton e Wilson (1985) verificaram que a re-mobilização de N dos tecidos vegetativos para os tecidos

reprodutivos foi superior (40%) em plantas noduladas do que em plantas tratadas com nitrato (24%); essa diferença pode estar associada com a deposição e estocagem de diferentes formas de N nos tecidos vegetativos em função das duas fontes de N (N-nitrato e N-FBN). Nas plantas que utilizam o N-nitrato ocorre um estímulo maior do crescimento da parte vegetativa e, nas plantas dependentes da FBN, ocorre um acúmulo de ureídos nos caules e folhas antes do início do enchimento de grãos. Esses ureídos são deslocados para as vagens durante o enchimento de grãos e utilizados na síntese de proteínas. Para esses autores, essa seria uma explicação para as observações de que a re-mobilização do N dos tecidos vegetativos é iniciada mais precocemente nas plantas de soja noduladas e ao fato de que essas plantas retêm menos N (39%) nos tecidos vegetativos do que aquelas cultivadas sob o sistema N-nitrato. Constatações semelhantes foram relatadas em trabalhos realizados com feijão comum (HUNGRIA, NEVES e VICTORIA, 1985; HUNGRIA e NEVES, 1987). Uma das regiões identificadas neste estudo está localizada no GL I e é delimitada pelos marcadores Satt239-Satt354, sendo uma região onde tem sido concentrados esforços de mapeamento fino, visando a clonagem de um QTL de efeito maior para aumento no teor de proteína nos grãos. Essa mesma região também tem sido relacionada com a redução no teor de óleo, na produção de sementes e com a maturidade e peso de sementes. O QTL denominado FBN4 (NN e MNS/NN), localizado nesse intervalo, foi o único que apresentou evidência de ligação gênica controlando os dois caracteres de FBN e a clonagem e a caracterização dessa região pode contribuir para a identificação dos fatores genéticos que determinam os dois fenótipos de FBN, bem como para elucidar se os fatores que controlam a FBN são os mesmos que controlam o teor de proteína nos grãos nessa região.

Alguns trabalhos afirmam que a FBN deve ser considerada como um dos componentes de produção e relatam correlações altas entre os caracteres MNS e PG ($r_F = 0,85$ a $0,96$) e N_2 fixado e PG ($r_F = 0,99$) em soja (RONIS et al., 1985; GREDEK, ORF e LAMBERT, 1986). O QTL FBN2 (NN e MNS) foi mapeado no grupo C2, em uma região relacionada frequentemente com a produção de grãos (PG) e com caracteres componentes da produção de grãos, sendo considerado um QTL para produção, também denominado universal (PALOMEQUE et al., 2009a e PALOMEQUE et al.,

2009b), por ter sido mapeado em ambientes muito divergentes. Essa região também tem sido relacionada com a fixação biológica por Nicolás, Hungria e Arias (2006), que detectaram QTL para NN em outra população de mapeamento e por Takahashi e Asanuma (1996) que observaram uma ligação entre o marcador morfológico T (que determina a cor da pubescência) com FBN em duas linhagens isogênicas avaliadas em condições de baixa temperatura. Mais recentemente, Toda et al. (2002) demonstraram que os alelos T (cor de pubescência marrom) e t (pubescência cinza) correspondem ao gene que codifica uma enzima responsável pela hidroxilação da posição 3' dos flavonóides, com uma única deleção de uma citosina (C) observada no alelo t. Nesta região, Zeng et al. (2009) também mapearam um QTL para genisteína, um isoflavonóide envolvido com a sinalização inicial entre a soja e *Bradyrhizobium*. Todos os QTLs mapeados nesta região do grupo C2 estão localizados num intervalo de 8,11 cM, delimitado pelos marcadores Satt277-Satt307 (Figura 5). Os dados gerados pelos estudos de mapeamento de QTLs, associados com as informações oriundas do sequenciamento do genoma da soja constituem ferramentas importantes para elucidar as causas desse *cluster* de QTLs.

Outra coincidência interessante foi observada entre os QTLs para tolerância ao estresse de deficiência de P mapeados por Zhang et al. (2009) e os QTL para FBN nos grupos de ligação G, E e C2. Nos grupos G (MPAS) e E (NN e MPAS) os QTL para FBN foram mapeados nos mesmos intervalos em que Zhang et al. (2009) identificaram QTLs para os caracteres de tolerância ao estresse por P (MPAS, eficiência da utilização de P, eficiência da absorção de P e atividade da fosfatase ácida), enquanto que no grupo C2 (FBN2) os autores mapearam QTLs para massa seca de raiz e eficiência no uso de P. O P é um fator limitante para a FBN, pois a atividade da enzima nitrogenase demanda muita energia, ou seja, requer muito ATP. Como consequência, a concentração de P nos nódulos de soja é elevada, em geral superior a 6 mg g⁻¹ de massa seca de nódulo, enquanto que na parte aérea esse valor é de 2-3 mg g⁻¹ de massa seca. Em ambientes deficientes em P, a fertilização fosfatada das plantas dependentes do N fornecido pela FBN promove incrementos na atividade da nitrogenase, no número e massa de nódulos e no acúmulo de N na planta (ISRAEL, 1987).

Dado à grande demanda da FBN por P, uma abordagem para aumentar a capacidade de aquisição de P pela planta deve ser desenvolvida para potencializar a FBN. Possíveis estratégias incluem a acidificação da rizosfera pela atividade da fosfatase ácida e mudanças na arquitetura das raízes para aumentar a área de absorção de P, entre outros. Se os dois processos (FBN e absorção de P) estiverem sob o mesmo controle genético, ambos podem ser considerados conjuntamente.

Práticas agrícolas sustentáveis requerem que os pesquisadores identifiquem os mecanismos das plantas responsáveis pela absorção de nutrientes minerais, visando explorar esses mecanismos no sentido de potencializar a capacidade da planta de absorver nutrientes do solo. Nesse contexto, a identificação de QTLs representa um ponto inicial importante para a caracterização molecular desses processos e para a compreensão da base genética que controla esses mecanismos, ao mesmo tempo que pode representar uma ferramenta extremamente útil aos programas de melhoramento, auxiliando no processo de seleção assistida. Este trabalho teve como objetivo contribuir com o conhecimento sobre a base genética da FBN em soja e identificar associações QTL-marcador úteis para a seleção assistida. Todos os QTLs identificados foram de pequeno efeito, o que está de acordo com a natureza quantitativa dos caracteres de FBN. Observou-se, também, que as causas das correlações entre os caracteres de FBN devem-se predominantemente a efeitos pleiotrópicos. Os QTLs FBN1 (MPAS, MNS e MNS/NN), FBN2 (NN e MNS), e FBN3 (MPAS e NN) promovem incrementos a todos os caracteres relacionados com a FBN e a ligação verificada no grupo de ligação I (FBN4: NN e MNS/NN) pode ser quebrada com a utilização dos marcadores na seleção assistida.

Os QTLs MPAS2, NN1 e FBN2 são relatados em outra população de mapeamento (NICOLÁS, HUNGRIA e ARIAS, 2006) e constituem bons candidatos para serem utilizados na seleção de genótipos visando incrementos na FBN, mas experimentos adicionais devem ser conduzidos para avaliar a estabilidade desses QTLs em ambientes diferentes. Os demais QTLs também são promissores e a confirmação desses QTLs em outra população de mapeamento é necessária; nesse contexto, a população descrita por Nicolás, Hungria e Arias (2006) poderá ser utilizada.

2.3.5 Considerações Finais

Com uma cobertura parcial do genoma, cerca de 50%, foi possível identificar oito regiões genômicas associadas com a FBN em uma população de linhagens oriundas dos genitores Bossier (alta capacidade de FBN) e Embrapa 20 (capacidade intermediária de FBN). Com exceção do QTL MPAS1 ($R^2 = 10,0\%$) os QTLs explicaram individualmente menos de 7,8% da variação fenotípica o que é coerente com os coeficientes de herdabilidade estimados e também com informações de literatura que indicam que a FBN está sob uma forte influência ambiental.

O emprego do método de análise de QTL mCIM permitiu inferir que a pleiotropia é a causa principal das correlações observadas entre os caracteres componentes da FBN, sendo observada em três das quatro regiões que foram significativas para mais de um caráter. No grupo de ligação C1, por exemplo, foi mapeado um provável QTL denominado FBN1, que foi significativo para MPAS, MNS e MNS/NN. Os caracteres MPAS e MNS são os que, normalmente, apresentam as melhores correlações com N total da planta em trabalhos que avaliam a FBN; desse modo a identificação dessa região, bem como de outras regiões em outras condições experimentais e em outras populações de mapeamento é importante para a confirmação desses QTLs. Nesse sentido, o QTL considerado como QTL mapeado, localizado no grupo L, também foi descrito em outra população onde o cultivar Embrapa 20 era um dos genitores. Isso também ocorreu com o possível QTL FBN2 no grupo C2 e com o QTL NN1 no grupo B1 (NICOLÁS, HUNGRIA e ARIAS, 2006).

Para este trabalho foram considerados três níveis de significância: um baseado nos dados experimentais que consistiu de um $LOD=4,01$, e dois níveis arbitrários para os valores de LOD 2,88 e 2,06. De acordo com esses valores os QTLs foram considerados como mapeados, prováveis e possíveis, respectivamente. Nesse trabalho apenas o QTL mapeado no grupo L atingiu o nível de significância de 4.01, mas a maioria foi mapeada com um LOD superior a 3,0, com exceção dos QTL localizados nos grupos C2 e A2, para os quais os valores de LOD foram iguais a 2,7. Contudo, com base nas informações disponíveis na literatura sobre mapeamento de QTL em soja, são em geral considerados como significativos valores de $LOD \geq 2,0$. São raros os trabalhos que empregaram o método de análise mCIM, que considera todos caracteres

simultaneamente, aumentando o poder de detecção de QTL. Portanto, mesmo os QTL classificados como possíveis são fortes candidatos a regiões controlando caracteres de FBN em soja.

Ressalta-se, também, que os QTL para FBN foram mapeados em regiões para as quais são relatados QTL para características importantes para o melhoramento, como produção de grãos, teor de proteína e de isoflavonas nas sementes, absorção de N e P e resistência ao nematóide de cisto da soja. O mapeamento de QTL pode fornecer informações sobre as razões da co-localização de QTL e, assim, potencializar o uso de marcadores na seleção assistida para mais de uma característica, caso um mesmo marcador esteja associado a mais de uma característica de interesse.

3 CONCLUSÕES

- a) Foram mapeados QTLs controlando quatro caracteres relacionados com a fixação biológica de nitrogênio (FBN) em oito grupos de ligação, a maioria de efeito pequeno ($1,2\% < R^2 < 10,0\%$);
- b) Dos QTLs mapeados, quatro estão envolvidos com o controle de mais de um caráter simultaneamente, e em três deles a ligação foi atribuída à pleiotropia;
- c) Dos QTLs mapeados, três estão localizados nos grupos de ligação L, B1 e C2, que foram associados com a FBN em outra população de mapeamento em soja;
- d) Foram observadas coincidências entre os QTLs mapeados para FBN e QTLs previamente relatados para os caracteres teor de proteína, produção de grãos, nitrato (N), absorção de P, fotossíntese, isoflavonas e resistência ao nematóide de cisto;
- e) Como neste trabalho cobriu-se aproximadamente 50% do genoma da soja, sugere-se em trabalhos futuros uma maior cobertura, visando a localização de outras regiões associadas com a FBN.

REFERÊNCIAS

- AKAO, S.; KOUCHI, H. A supernodulating mutant isolated from soybean cultivar Enrei. **Soil Science and Plant Nutrition**, Tokyo, v. 38, n. 1, p. 183-187, 1992.
- ALVES, B. J. R.; BODDEY, R. M.; URQUIAGA, S. The success of BNF in soybean in Brasil. **Plant and Soil**, Dordrecht, v. 252, n. 1, p. 1-9, 2003.
- BASTEN, C. J.; WEIR, B. S.; ZENG, Z-B. **QTL cartographer**: version 1.17. Department of Statistics, North Carolina State University, Raleigh, 2003.
- BETTS, J. H.; HERRIDGE, D. F. Isolation of soybean lines capable of nodulation and nitrogen fixation under high levels of nitrate supply. **Crops Science**, Madison v. 27, n. 6, p. 1156-1161, 1987.
- BOHRER, T. R. J.; HUNGRIA, M. Avaliação de cultivares de soja quanto à fixação biológica de nitrogênio. **Pesquisa Agropecuária Brasileira**, Brasília, v. 33, n. 6, p. 937-953, 1998.
- BOREVITZ, J. O.; CHORY, J. Genomics tools for QTL analysis and gene discovery. **Current Opinion in Plant Biology**, Amsterdam, v. 7, n. 2, p. 132-136, 2004.
- BURIAS, N.; PLANCHON, C. Divergent selection for dinitrogen fixation and yield in soybean. **Theoretical Applied Genetics**, Berlin, v. 83, n. 5, p. 543-548, 1992.
- BRUMMER, E. C.; GRAEF, G. L.; ORF, J.; WILCOX, J. R.; SHOEMAKER, R. C. Mapping QTL for seed protein and oil content in eight soybean populations. **Crop Science**, Madison, v. 37, n. 2, p. 370-378, 1997.
- BYRNE, P. F.; McMULLEN, M. D.; SNOOK, M. E. MUSKET, T. A.; THEURI, J. M.; WIDSTROM, N. W.; WISEMAN, B.R.; COE, E. H. Quantitative trait loci and metabolic pathways: genetic control of the concentration of maysin, a corn earworm resistance factor, in maize silks. **Proceedings of National Academy of Science of the United States of America**, Washington, v. 93, n. 17, p.8820-8825, 1996.
- CARROLL, B. J.; McNEIL, D. L.; GRESSHOFF, P. M. A supernodulation and nitrate-tolerant symbiotic (*nts*) soybean mutant. **Plant Physiology**, Minneapolis, v. 78, n.1, p. 34-40, 1985.
- CATOIRA, R.; GALERA, C.; BILLY, F.; PENMETSA, V.; JOURNET, E.; MAILLET, F.; ROSENBERG, C.; COOK, D.; GOUGH, C.; DÉNARIÉ, J. Four genes of *Medicago truncatula* controlling components of a nod factor transduction pathway. **The Plant Cell**, Rockville, v. 12, n. 9, p. 1647-1665, 2000.

CHAPMAN, A.; PANTALONE, V. R.; USTUN, A.; ALLEN, F. L.; LANDAU- ELLIS, D.; TRIGIANO, R. N.; GRESSHOFF, P. M. Quantitative trait loci for agronomic and seed quality traits in an F₂ and F_{4:6} soybean population. **Euphytica**, Wageningen v.129, n. 3, p. 387-393, 2003.

CHASE, K.; CREGAN, P. B.; LARK, K. G.; AMSUR, L. M.; ORF, J. H. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. **Crop Science**, Madison, v.36, p. 1327-1336, 1996.

CHOI, I-Y.; HYTEN, D. L.; MATUKUMALLI, L. K.; SONG, Q.; CHAKY, J. M.; QUIGLEY, C. V.; CHASE, K.; LARK, K. G.; REITER, R. S.; YOON, M-S.; HWANG, E-Y.; YI, S-I.; YOUNG, N. D.; SHOEMAKER, R. C.; VAN TASSELL, C. P.; SPECHT, J. E.; CREGAN, P. B. A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. **Genetics**, Austin, v. 176, n. 1, p. 685-696, 2007.

CHUNG, J.; BABKA, H. L.; GREAF, G. L.; STASWICK, P. E.; LEE, D. J.; CREGAN, P. B. SHOEMAKER, R. C.; SPEACHT, J. E. The seed protein, oil, and yield QTL on soybean linkage group I. **Crop Science**, Madison, v. 43, n. 3, p. 1053-1067, 2003.

COALE, F. J.; MEISINGER, J. J.; WIEBOLD, W. J. Effects of plant breeding and selection on yields and nitrogen fixation in soybeans under two soil nitrogen regimes. **Plant and Soil**, Dordrecht, v. 86, n. 3, p. 357-367, 1985.

CONAB. Companhia Nacional de Abastecimento. Acompanhamento da safra brasileira: grãos – décimo segundo levantamento, setembro 2009. Brasília: CONAB, 43p. Disponível em: <http://www.conab.gov.br/conabweb/download/safra/12graos_08-09.pdf>. Acesso em: 15 set. 2009.

CONCIBIDO, V. C.; DIERS.; B. W. PRAKASH, R. A. A decade of QTL mapping for cyst nematode resistance in soybean. **Crop Science**, Madison, v. 44, n. 4, p. 1121-1131, 2004.

CREGAN, P. B.; JARVIK, T.; BUSH, A. L.; SHOEMAKER, R. C.; LARK, K. G.; KAHLER, A. L.; KAYA, N.; VAN TOAL, T. T.; LOHNES, D. G.; CHUNG, J.; SPECHT, J. E. An integrated genetic linkage map of the soybean genome. **Crop Science**, Madison, v. 39, n. 5, p. 1464 - 1491, 1999.

CREGAN, P. B.; KEYSER, H. H. Host restriction of nodulation by *Bradyrhizobium japonicum* strain USDA 123 in soybean. **Crop Science**, Madison, v. 26, n. 5, p. 911-916, 1986.

DIERS, B. W.; KEIM, P.; FEHR, W. R.; SHOEMAKER, R.C. RFLP analysis of soybean seed protein and oil content. **Theoretical and Applied Genetics**, Berlin, v. 83, n. 5, p. 608-612, 1992.

- DEVINE, T. E.; O'NEIL, J. J. Genetic allelism of nodulation response genes *rj1*, *Rj2* and *Rj4* in soybean. **Crop Science**, Madison, v. 29, n. 6, p. 1345-1350, 1989.
- DEVINE, T. E.; BREITHAUPT, B. H. Significance of incompatibility reactions of *Rhizobium japonicum* strains with soybean host genotypes. **Crop Science**, Madison, v. 20, n. 2, p. 269-271, 1980.
- EWING, B.; GREEN, P. Base-calling of automated sequencer traces using phred II Error probabilities, **Genome Research**, New York, v. 8, n. 3, p. 186-194, 1998.
- EWING, B.; HILLIER, L.; WENDL, M. C.; GREEN, P. Base-calling of automated sequencer traces using phred I. Accuracy assessment. **Genome Research**, Nova lorque, v. 8, n. 3, p. 175-185, 1998.
- FALCONER, D. S.; MACHAY, T. F. **C Introduction to quantitative genetics**. 4th ed. Esses: Longman, 1996. 464 p.
- FARIS, J. D.; LI, W. L.; LIU, D. J.; CHEN, P. D.; GILL, B. S. Candidate gene analysis of quantitative disease resistance in wheat. **Theoretical and Applied Genetics**, Berlin, v. 98, n. 2, p. 219-225, 1999.
- FEDOROVA, M.; VAN de MORTEL, J.; MATSUMOTO, P. A.; CHO, J.; ROWN, C. D.; VANDENBOSCH, K. A.; GRANTT, J. S.; VANCE, C. P. Genome-wide identification of nodule-specific transcripts in the model legume *Medicago truncatula*. **Plant Physiology**, Minneapolis, v. 130, n. 2, p. 519-537, 2002.
- GIBSON, A. H.; HARPER, J. E. Nitrate effect on nodulation of soybean by *Bradyrhizobium japonicum*. **Crop Science**, Madison, v. 25, n. 3, p. 497-501, 1985.
- GORDON, D.; ABAJIAN, C.; GREEN, P. Consed: a graphical tool sequence finishing. **Genome Research**, New York, v. 8, n. 3, p. 195-2002, 1998.
- GRAHAM, P. H.; VANCE, C. P. Nitrogen fixation in perspective: an overview of research and extension needs. **Field Crops Research**, Amsterdam, v. 65, n. 1-2, p. 93-106, 2000.
- GREDER, R. R.; ORF, J. H.; LAMBERT, J. W. Heritabilities and associations of nodule mass and recovery of *Bradyrhizobium japonicum* serogroup USDA 110 in soybean. **Crop science**, Madison, v. 26, n. 1, p. 33-36, 1986.
- GREMAUD, M. G.; HARPER, J. E. Selection and initial characterization of partially nitrate tolerant nodulation mutants of soybean. **Plant Physiology**, Rockville, v. 89, n. 1, p. 169-173, 1989.
- GRESSHOFF, P. Post-genomic insights into plant nodulation symbioses. **Genome Biology**, London, v.4, n. 1, p. 201-206, 2003.

GUTIERREZ-GONZALES, J.; WU, X.; ZHANG, J.; LEE, J.; ELLERSIECK, M.; SHANNON, J. G.; YU, O.; NGUYEN, H. T.; SLEPER, D. A. Genetic control of soybean seed isoflavone content: importance of statistical model and epistasis in complex traits. **Theoretical and Applied Genetics**, Berlin, v. 119, n. 6, p. 1069-1083, 2009.

GUO, B.; SLEPER, D. A.; LU, P.; SHANNON, J. G.; NGUYEN, H. T.; ARELLI, P. R. QTLs associated with resistance to soybean cyst nematode in soybean: meta-analysis of QTL locations. **Crop Science**, Madison, v.46, n. 3, p. 595-602, 2006.

GUO, B.; SLEPER, D. A.; ARELI, P. R.; SHANNON, J. G.; NGUYEN, H. T. Identification of QTL associated with resistance to soybean cyst nematode races 2, 3 and 5 in soybean PI 90763. **Theoretical and Applied Genetics**, Berlin, v. 111, n. 5, p. 965-971, 2005.

GUZMAN, P. S.; DIERS, B. W.; NEECE, D. J. ; St MARTIN, S. K. ; LeROY, A. R.; GRAU, R. C.; HUGHES, T. J. NELSON, R. L. QTL associated with yield in three backcross-derived populations of soybean. **Crop Science**, Madison, v. 47, n. 1, p. 111-122, 2007.

HERRIDGE, D.; ROSE, I. Breeding for enhanced nitrogen fixation in crop legumes. **Field Crops Research**, Amsterdam, v. 65, n. 2-3, p. 229-248, 2000.

HERRIDGE, D. F.; ROSE, I. A. Heritability of enhanced N₂ fixation in early and late inbreeding generations of soybean. **Crop Science**, Madison, v. 34, n. 2, p. 360-366, 1994.

HERRIDGE, D. F.; TURPIN, J. E.; ROBERTSON, M. J. Improving nitrogen fixation of crop legumes through breeding and agronomic management analysis with simulation modelling. **Australian Journal of Experimental Agriculture**, Melbourne, v. 41, p. 391-401, 2001.

HERRIDGE, D. F.; DANSO, S. K. A. Enhancing crop legume N₂ fixation through selection and breeding, **Plant and Soil**, Dordrecht, v. 174, n. 1-2, p. 51-82, 1995.

HISANO, H.; SATO, S.; ISOBE, S.; SASAMOTO, S.; WADA T.; MATSUNO, A.; FUJISHIRO, T.; YAMADA, M.; NAKAYAMA, S.; NAKAMURA, Y, WATANABE, S.; HARADA, K.; TABATA, S. Characterization of the soybean genome using EST-derived microsatellite markers. **DNA Research**, Amsterdam, v. 14, n. 6, p. 1-11, 2007.

HUNGRIA, M.; BOHRER, T. R. J. Variability of nodulation and dinitrogen fixation capacity among soybean cultivars, **Biology and Fertility of Soils**, Berlin, v. 31, n. 1, p. 45-52, 2000.

HUNGRIA, M.; CAMPO, R. J.; MENDES, I. C; **A importância do processo de fixação biológica do nitrogênio para a cultura da soja**: componente essencial para a competitividade do produto brasileiro. Londrina: Embrapa soja, 2007. 80 P. (Documentos 283).

- HUNGRIA, M.; CAMPO, R. J.; MENDES, I. C.; GRAHAM, P. H. Contribution of biological nitrogen fixation to the N nutrition of grain crops in the tropics: the success of soybean (*Glycine max* L. Merr.) in South America. In: SINGH, R. P.; SHANKAR, N.; JAIWAL, P. K. (Eds.). **Nitrogen nutrition in plant productivity**. Houston: Studium Press/LLC, 2006, p. 43-93.
- HUNGRIA, M.; FRANCHINI, J. C.; CMPO, R. J.; GRAHAM, P. H. The importance of nitrogen fixation to soybean cropping in South América. In: WERNER, D.; NEWTON, W. E. (Ed.). **Nitrogen fixation in agriculture, forestry, ecology and environment**. Dordrecht: Academic Publishers, 2005. p. 25-42.
- HUNGRIA, M.; NEVES, M.; C.; P.; Partitioning of nitrogen from biological fixation and fertilizer in *Phaseolus vulvaris*. **Physiologia Plantarum**, Copenhagen, v. 69, n. 1, p. 55-63, 1987.
- HUNGRIA, M.; NEVES, M.; C.; P.; VICTORIA, R. L. Assimilação do nitrogênio pelo feijoeiro II. Absorção e translocação do N mineral e do N₂ fixado. **Revista Brasileira de Ciência do Solo**. Campinas, v. 9, p. 201-209, 1985.
- HUNGRIA, M.; STACEY, G. Molecular signals exchanged between host plants and rhizobia: basic aspects and potential application in agriculture. **Soil Biology and Biochemistry**, Oxford, v. 29, n. 5-6, p. 819-830, 1997.
- HYTEN, D. L.; PANTALONE, V. R.; SAMS, C. E.; SAXTON, A. M.; LANDAU-ELLIS, D.; STEFANIAK, T. R.; SCHIMIDT, M. E. Seed quality QTL in a prominent soybean population. **Theoretical and Applied Genetics**, Berlin, v. 109, n. 3, p. 552-561, 2004.
- ISRAEL, D. W. Investigation of the role of phosphorus in symbiotic dinitrogen fixation. **Plant Physiology**, Minneapolis, v.84, n. 3, p. 835-840, 1987.
- ISRAEL, D.; BURTON, J. W.; WILSON, R. F. Studies on genetic male-sterile soybeans IV. Effect of male sterility and source of nitrogen nutrition on accumulation, partitioning, and transport of nitrogen. **Plant Physiology**, Minneapolis v.78, n. 4, p. 762-767, 1985).
- JEONG, S-C.; YANG, K.; PARK, J-Y.; HAN, K. S.; YU, S.; HWANG, T-Y.; HUR, C. G.; KIM, S. H.; PARK, P. B.; KIM, H. M.; PARK, Y-I.; LIU, J. R. Structure, expression, and mapping of two nodule-specific genes identified by mining public soybean EST databases. **Gene**, Amsterdam, v. 393, p. 71-80, 2006.
- JIANG, C.; ZENG, Z-B. Multiple trait analysis of genetic mapping for quantitative trait loci. **Genetics**, Austin, v.140, n. 3, p. 1111-1127, 1995.
- JOURNET, E. P.; VAN TUINEN, GOUZY, J.; CRESPEAU, H.; CARREAU, V.; FARMER, M. J.; NIEBEL, A.; SCHIEX, T.; JAILLON, O.; CHATAGNIER, O.; GODIARD, L.; MICHELI, F.; KAHN, D.; GIANINAZZI-PEARSON, V.; GAMAS, P. Exploring root symbiotic programs in the model *Medicago truncatula* using EST analysis. **Nucleic Acids Research**. Oxford, v. 30, n. 24, p. 5579-5592, 2002.

KASSEM, M. A.; SHULTZ, J.; MEKSEM, K.; CHO, Y.; WOOD, A. J.; IQBAL, M. J.; LIGHTFOOT, D. A. An updated 'Esses' by 'Forrest' linkage map and first composite interval map of QTL underlying soybean traits. **Theoretical and Applied Genetics**, Berlin, v. 113, n. 6, p. 1015-1026, 2006.

KASSEM, M. A.; MEKSEM, K.; IQBAL, M. J.; NJITI, V. N.; BANZ, W. J.; WINTERS, T. A.; WOOD, A.; LIGHTFOOT, D. A. Definition of soybean genomic regions that control seed phytoestrogen amounts. **Journal of Biomedicine and Biotechnology**, Cairo, v. 1, p. 52-60, 2004.

KEIM, P.; DIERS, B. W.; OLSON, T. C.; SHOEMAKER, R. C. RFLP mapping in soybean: association between marker loci and variation in quantitative traits. **Genetics**, Austin, v. 126, n. 3, p. 735-742, 1990.

KEIM, P.; OLSON, T. C.; SHOEMAKER, R. C. A rapid protocol for isolating soybean DNA. **Soybean Genetics Newsletter**, Ames, v. 15, p. 150-152, 1988.

KOSAMBI, D. D. The estimation of map distances from recombination values. **Annual of Eugenics**, Cambridge, v. 12, p. 172-175, 1944.

KO, M. P.; BARKER, K. R.; HUANG, J.-S. Nodulation of Soybeans as Affected by Half-root Infection with *Heterodera glycines*. **Journal of Nematology**, Madison, v. 16, n. 1, p. 97-105, 1984.

LANDER, E. S.; GREEN, P.; ABRAHAMSON, J.; BARLON, A.; DALY, M. J.; LINCOLN, S.; NEWBURG, L. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. **Genomics**, San Diego, v. 1, n. 2, p. 174-181, 1987.

LARK, K. G.; WEISEMANN, J. M.; MATHEWS, J. M.; PALMER, R.; CHASE, K.; MACALMA, T. A genetic map of soybean (*Glycine max L.*) using an intraspecific cross of two cultivars: 'Minsoy' and 'Noir 1'. **Theoretical and Applied Genetics**, Berlin, v. 86, n. 8, p. 901-906, 1993.

LEE, H.; HURI, G.; OH, C. J.; KIM, H. B.; PARKI, S.; AN, C. S. Analysis of the root nodule-enhanced transcriptome in soybean. **Molecules and Cells**, Seoul, v. 18, n. 1, p. 53-62, 2004.

LEE, S. H.; BAILEY, M. A.; MIAN, M. A. R.; CARTER, T. E.; SHIPE, E. R.; ASHLEY, D. A.; PARROTT, W. A.; HUSSEY, R. S.; BOERMA, H. R. RFLP loci associated with soybean seed protein and oil content across populations and locations. **Theoretical and Applied Genetics**, Berlin, v. 93, n. 5-6, p. 649-657, 1996.

LESTARI, P.; VAN K.; KIM, M. Y.; HWANG, C. H.; LEE, B.; LEE, S. Differentially expressed genes related to symbiotic association in a supernodulating soybean mutant and its wild-type. **Molecular Plant Pathology**, Hoboken, v. 7, n. 4, p. 235-247, 2006.

LI, D.; PFEIFFER, T. W.; CORNELIUS, P. L. Soybean QTL for yield and yield components associated with *Glycine soja* alleles. **Crop Science**, Madison, v. 48, n. 2, p.571-581, 2008.

LI, YI-D.; WANG, Y-J.; TONG, YI-J.; GAO, JI-G.; ZHANG, J-S.; CHEN, S-YI. QTL mapping of phosphorus deficiency tolerance in soybean (*Glycine max* L. Merr.). **Euphytica**, Wageningen, v. 142, n. 1-2, p. 137-142, 2005.

LIMPENS, E.; FRANKEN, C.; SMIT, P.; WILLEMSE, J.; BISSELING, T.; GEURTS, R.; LysM domain receptor kinase regulating rhizobial Nod factor-induced infection. **Science**, Washington, v. 302, n. 5645, p. 630-633, 2003.

MADSEN, E. B.; MADSEN, L. H.; RADUTOIU, S.; OLBRYT, M.; RAKWALSKA, M.; SZCZYGLOWSKI, K.; SATO, S.; HANEKO, T.; TABATA, S. SANDAL, N.; STOUGAARD, J. A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals, **Nature**, Washington v. 425, p. 637-640, 2003.

MEKSEM, K.; NJITI, V. N.; BANZ, W. J.; IQBAL, M. J.; KASSE, M. M.; HYTEN, D. L.; YUANG. J.; WINTERS, T. A.; LIGHTFOOT, D. A. Genomic regions that underlie soybean seed isoflavone content. **Journal of Biomedicine and Biotechnology**. Cairo, v. 1, n. 1, p. 38-45, 2001.

MPEPEREKI., S.; JAHAVERI, F.; DAVUS, P.; GILLER, K. E. Soybeans and sustainable agriculture. Promiscuous soybeans in Southern Africa. **Field Crops Research**, Amsterdam, v. 65, n. 2, p. 137-150, 2000.

MUNRO, I. C.; HARWOOD, M.; HLYWKA, J. J.; STEPHEN, A. M.; DOULL, J.; FLAMM, W. G.; ADLERCREUTZ, H. Sou isoflavones: A safety review. **Nutrition Reviews**, Nova lorque, v. 61, n. 1, p. 1-33, 2003.

MYTTON, L. R. Developing a breeding strategy to exploit quantitative variation in symbiotic nitrogen fixation, **Plant and Soil**, Dordrecht, v. 82, n. 3, p. 329-335, 1984.

NCBI National Center for Biotechnology Information. Disponível em: <<http://www.ncbi.nlm.nih.gov/sites/entrez?db>>. Acesso em: 20 nov. 2008.

NICHOLS, D. M.; GLVER, K. D.; CARLSON, S. R.; SPECHT, J. E.; DIERS, B. W. Fine mapping of seed protein QTL on soybean linkage Group I and its correlated effects on agronomic traits. **Crop Science**, Madison, v.46, n. 2, p. 834-839, 2006.

NICOLÁS, M. F.; ARIAS, C. A. A.; HUNGRIA, M. Genetics of nodulation and nitrogen fixation in Brazillian soybean cultivars. **Biology and Fertility of Soils**, Berlin, v. 36, n. 2, p. 109-117, 2002.

NICOLÁS, M. F.; HUNGRIA, M.; ARIAS, C. A. A. Identification of quantitative trait loci controlling nodulation and shoot mass in progenies from two Brazilian soybean cultivars. **Field Crops Research**, Amsterdam, v. 95, n. 2-3, p. 355-366, 2006.

NISHIMURA, R.; HAYASHI, M.; WU, G.; KOUCHI, H.; IMIZUMI-ANRAKU, H.; MURAKAMI, Y.; KAWASAKI, S.; AKAO, S.; OHMORI, M.; NAGASAWA, M.; HARADA, K.; KAWAGUCHI, M. HAR1 mediates regulation of symbiotic organ development. **Nature**, Washington, v. 420, p. 426-429, 2002.

NJITI, V. K.; MEKSEM, D. A.; LIGHTFOOT, W. J. B.; WINTERS, T. A. Molecular markers of phytoestrogen content in soybeans. **Journal of Medicine and Food**, New Rochelle, v. 2, p. 165–167, 1999.

NODARI, R. O.; TSAI, S. M.; GUZMÁN, P.; GILBERTSON, R. L.; GEPTS, P. Toward an integrated linkage map of common bean. III. Mapping genetic factors controlling host-bacteria interactions. **Genetics**, Austin, v. 134, n. 1, p. 341-350, 1993.

ORF, J. H.; CHASE, K.; JARVIK, T.; MANSUR, L. M.; CREGAN, P. B.; ADLER, F. R.; LARK, K. G. Genetics of soybean agronomic traits: I comparison of three related recombinant inbred populations. **Crop Science**, Madison, v. 39, n. 4, p. 1642-1651, 1999.

PALOMEQUE, P.; LI-JUN, L.; LI, W.; HEDGES, B.; COBER, E. R.; RAJCAN, I. QTL in mega-environments: I. universal and specific seed yield QTL detected in a population derived from a cross of high-yielding adapted x high-yielding exotic soybean lines. **Theoretical and Applied Genetics**, Berlin, v. 119, n. 3, p. 417-427, 2009a.

PALOMEQUE, L.; LI-JUN, L.; LI, W.; BRADLEY, H.; COBER, E. R.; RAJCAN, I. QTL in mega-environments: II. Agronomic trait QTL co-localized with seed yield QTL detected in a population derived from a cross of high-yielding adapted x high-yielding exotic soybean lines. **Theoretical and Applied Genetics**, Berlin, v. 119, n. 3, p. 429-436, 2009b.

PANTHEE, D. R.; PANTALONE, V. R.; VEST, D. R.; SAXTON, A. M.; SAMS, C. E. Quantitative trait loci for seed protein and oil concentration, and seed size in soybean. **Crop Science**, Madison, v. 45, n. 5, p. 2015-2022, 2005.

PANTHEE, D. R.; PANTALONE, V. R.; SAMS, C. E.; SAXTON, A. M.; WEST, D. R.; RAYFORD, W. E. Genomic regions governing soybean seed nitrogen accumulation. **Journal of the American Oil Chemists' Society**. Chicago, v. 81, n. 1, p. 77-81, 2004.

PAZDERNIK, D. L.; GRAHAM, P. H.; ORF, J. H. Heritability in the early nodulation of F3 and F4 soybean lines. **Canadian Journal of Plant Science**, Ottawa, v. 77, p. 201-205, 1996.

PFLIEGER, S.; LEFEBVRE, V.; CAUSSE, M. The candidate gene approach in the plant genetics: a review. **Molecular Breeding**, Dordrecht, v. 7, n. 4, p. 275-291, 2001.

PRIMOMO, V. S.; POYSA, V.; ABLETT, G. R.; JACKSON, C.; GIJZEN, M.; RAJCAN, I. Mapping QTL for individual and total isoflavone content in soybean seeds. **Crop Science**, Madison, v. 45, n. 6, p. 2454-2464, 2005.

RADUTOIU, S.; MADSEN, L. H.; MADSEN, E. M.; FELLE, H. H.; UMEHARA, Y.; GRONLUND, M.; SATO, S.; NAKAMURA, Y.; TABATA, S.; SANDAL, N.; STOUGAARD, L. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. **Nature**, Washington, v. 425, p. 585-592, 2003.

RENGEL, Z. Breeding for better symbiosis. **Plant and Soil**, Dordrecht, v. 245, n. 1, p. 147-162, 2002.

RONIS, D. H.; SAMMONS, D. J.; KENWORTHY, W. J.; MEISINGER, J.J. Heritability of total and fixed N content of the seed in two soybean populations. **Crop Science**, Madison, v. 23, n. 1, p. 1-4, 1985.

RISCH, N. Genetic linkage: Interpreting LOD scores. **Science**, Washington, v.255, n. 5046, p. 803-804, 1992.

SADOWSKY, M. J.; KOSSLAK, R. M.; CEZARY, J. MADRZAK, C. J. M.; GOLINSKA, B.; CREGAN, P. B. Restriction of nodulation by *Bradyrhizobium japonicum* is mediated by factors present in the roots of *Glycine max*. **Applied and Environmental Microbiology**, Washington, v. 61, n. 2, p. 832–836, 1995.

SALL, K.; SINCLAIR, T. R. Soybean genotypic differences in sensitivity of symbiotic nitrogen fixation to soil dehydration. **Plant and Soil**, Dordrecht, v. 133, n. 1, p. 31-37, 1991.

SALVAGIOTTI, F.; CASSMAN, K. G.; SPECHT, J. E.; WALTERS, D. T.; WEISS, A.; DOBERMAN, A. Nitrogen uptake, fixation and response to fertilizer N in soybeans: A review. **Field Crops Research**, Amsterdam, v. 108, n. 6, p. 1-13, 2008.

SANTOS, M. A.; NICOLÁS, M. F.; HUNGRIA, M. Identificação de QTL associados à simbiose entre *Bradyrhizobium japonicum*, *B. elkanii* e soja. **Pesquisa Agropecuária Brasileira**, Brasília, v. 41, n. 1, p. 67-75, 2006.

SAS INSTITUTE, **Statistical analysis software**: SAS 9.1. 2003. Cary-NC, EUA.

SEARLE, I. R.; MEN, A. M.; LANIYA, T. S.; BUZAS, D. M.; ITURBE-ORMAETXE, I.; CAOOLL, B. J.; GRESSHOFF, P. M. Long-distance signalling for nodulation control in legumes requires a CLAVATAI-like receptor kinase. **Science**, Washington, v. 299, p. 109-112, 2003.

SEBOLT, A. M.; SHOEMAKER, R.C.; DIERS, B.W. Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. **Crop Science**, Madison, v. 40, n. 5, p. 1438-1444, 2000.

SCHAUSER, L.; ROUSSIS, A.; STILLER, A.; STOUGAARD, J. A plant regulator controlling development of symbiotic root nodules. **Nature**, Washington, v. 402, p. 192-195, 1999.

SCHULTZE, M.; KONDOROSI, A. Regulation of symbiotic root nodule development. **Annual Review of Genetics**, Palo Alto, v.32, p. 33-57, 1998.

SERRAJ, R.; SINCLAIR, T. R.; PUERCELL, L. C. Symbiotic N₂ fixation response to drought. **Journal of Experimental Botany**, Oxford, v. 50, n. 331, p. 143-155, 1999.

SHOEMAKER, R. C.; SPECHT, J. E. Integration of the soybean molecular and classical genetic linkage groups. **Crop Science**, Madison, v. 35, n. 2, p. 436-446, 1995.

SINCLAIR, T. R.; PURCELL, L. C.; KING, C. A.; SNELLER, C. H.; CHEN, P.; VADEZ, V. Drought tolerance and yield increase of soybean resulting from improved symbiotic N₂ fixation. **Field Crops Research**, Amsterdam, v. 101, n. 1, p. 68-71, 2007.

SINCLAIR, T. R.; VADEZ, V.; CHENU, K. Ureide accumulation in response to Mn nutrition by eight soybean genotypes with N₂ fixation tolerance to soil drying. **Crop Science**, Madison, v. 43, n. 2, p. 592-597, 2003.

SINCLAIR, T. R.; PURCELL, L. C.; VADEZ, V.; SERRAJ, R.; KING, C. A.; NELSON, R. Identification of soybean genotypes with N₂ fixation tolerance to water deficits. **Crop Science**, Madison, v. 40, n. 6, p. 1803-1809, 2000.

SINCLAIR, T. R.; SOFFES, A. R.; HINSON, K.; ALBRECHT, S. L.; PFAHLER, P. L. Genotypic variation in soybean nodule number and weight. **Crop Science**, Madison, v. 31, n. 2, p. 301-304, 1991.

SMALLEY, M. D.; FEHR, W. R.; CIANZIO, S. R.; HAN, F. H.; SEBASTIAN, S. A.; STREIT, L. Quantitative trait loci for soybean seed yield in elite and plant introduction germplasm. **Crop Science**, Madison, v. 44, n. 2, p. 436-442, 2004.

SONG, Q. J.; MAREK, L. F.; SHOEMAKER, R. C.; LARK, K. G.; CONCIBIDO, V. C.; DELANNAY, X.; SPECHT, J. E.; CREGAN, P. B. A new integrated genetic linkage map of the soybean. **Theoretical and Applied Genetics**, Berlin, v.109, n. 1, p. 122-128, 2004.

SONG, L.; CARROLL, B. J.; GRESSHOFF, P. M.; HERRIDGE, D. F. Field assessment of supernodulating genotypes of soybean for yield, N₂ fixation and benefit to subsequent crops. **Soil Biology Biochemistry**, Oxford, v. 27, n. 4-5, p. 563-569, 1995.

SOYBASE 2009. Soybase and the soybean breeder's toolbox. Knowledge and tools for soybean breeders and researchers. USDA-Iowa State University, IA, USA. Disponível em: < <http://www.soybase.agron.iastate.edu/>>. Acesso em: 10 jul. 2009.

SOUZA, A. A.; BOSCARIOL, R. L.; MOON, D. H.; CAMARGO, L. E. A.; TSAI, S. M. Effects of *Phaseolus vulgaris* QTL in controlling host-bacteria interactions under two levels of nitrogen fertilization. **Genetics and Molecular Biology**, Ribeirão Preto, v. 23, n. 1, p. 155-161, 2000.

- SOUZA, R. A.; HUNGRIA, M.; FRANCHINI, J. C.; MACIEL, C. D.; CAMPO, R. J.; ZAIA, D. A. M. Conjunto mínimo de parâmetros para avaliação da microbiota do solo e da fixação biológica do nitrogênio pela soja. **Pesquisa Agropecuária Brasileira**, Brasília, v. 43, n. 1, p. 83-91, 2008.
- SPECHT, J. E.; CHASE, K.; MACRANDER, M.; GRAEF, G. L.; CHUNG, J.; MARKWELL, J. P.; GERMANN, M.; ORF, J. H.; LARK, G. Soybean response to water: a QTL analysis of drought tolerance. **Crop Science**, Madison, v. 41, n. 2, p. 493-509, 2001.
- STOUGAARD, J. Regulators and regulation of legumes root nodule development. **Plant Physiology**, Minneapolis, v. 124, n. 2, p. 531-540, 2000.
- TAKAHASHI, R.; ASANUMA, S.; Association of T gene chilling tolerance in soybean. **Crop Science**, Madison, v. 36, n. 3, p. 559-562. 1996.
- TAJIMA, S.; NOMURA, M.; KOUCHI, H. Ureide biosynthesis in legume nodules. **Frontiers in Bioscience**, Nova Iorque, v. 9, p. 1374-1381, 2004.
- TANYA, P. SRINIVES, P.; TOOJINDA, T.; NAKHON, P.; VANAVICHIT, A.; LEE, S. H.; Identification of SSR markers associated with N₂ fixation components in soybean [*Glycine max* (L.) Merr.]. **The Korean Journal of Genetics**, Seoul, v. 27, n. 4, p. 351-359, 2005.
- TSAI, S. M.; NODARI, R. O.; MOON, D. J.; CAMARGO, L. E. A.; VENCOVSKY, R. GEPTS, P. QTL mapping for nodule number and common bacterial blight in *Phaseolus vulgaris* L. **Plant and Soil**, Dordrecht, v. 204, p. 135-145, 1998.
- TODA, K.; YANG, D.; YAMANAKA, N.; WATANABE, S.; HARADA, K.; TAKAHASHI, R. A single-base deletion in soybean flavonoid 3'-hydroxylase gene is associated with gray pubescence color. **Plant Molecular Biology**, the Netherlands, v. 50, n. 2, p. 187-196, 2002.
- VADEZ, V.; SINCLAIR, T. R. Sensitivity of N₂ fixation traits in soybean cultivar Jackson to manganese. **Crop Science**, Madison, v. 42, n. 3, p. 791-796, 2002.
- VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: sociedade Brasileira de Genética, 1992. 486 p.
- VIEIRA, J. D. V.; OLIVEIRA, D. A.; SOARES, T. C. B.; SCHUSTER, I.; PIOVESAN, D.; MARTÍNEZ, C. A.; BARROS, E. G.; MOREIRA, M. A. Use of the QTL approach to the study of soybean trait relationships in two populations of recombinant inbred lines at the F and F₈ generations. **Brazilian Journal of Plant Physiology**, Piracicaba, v. 18, n. 2, p. 281-290, 2006.

VIEIRA, C.; PASYUKOVA, E. G.; ZENG, Z. B.; HACKETTE, J. B.; LYMAN, R. F.; MACKAY, T. F. C. Genotype-environment interaction for quantitative trait loci affecting life span, in *Drosophila melanogaster*. **Genetics**, Austin, v. 154, n. 1, p. 213-222, 2000.

VUONG, T. D.; HARPER, J. E. Inheritance and allelism analysis of hypernodulating genes in the NOD3-7 and NOD2-4 soybean mutants. **Crop Science**, Madison, v. 40, n. 3, p. 700-703, 2000.

WU, X.; BLAKE, S.; SLEPER, D. A.; SHANNON, J. G.; CREGAN, P.; NGUYEN, H. T. QTL. Additive and epistatic effect for SCN resistance in PI 437654. **Theoretical and Applied Genetics**, Berlin, v. 118, n. 6, p. 1093-1105, 2009.

WANG, S.; BASTEN, C. J.; ZENG, Z-B. **Windows QTL cartographer**: version 2.5. Raleigh: North Carolina State University, Department of Statistics, 2005.

ZENG, Z. B. Theoretical basis for separation of multiple linked gene effects in mapping of quantitative trait loci. **Proceedings of the National Academy of Science of the United States of America**, Washington, v.90, p.10972-10976, 1993.

ZENG, G.; LI, D.; HAN, Y.; TENG, W.; WANG, J.; QIU, L.; LI, W. Identification of QTL underlying isoflavone contents in soybean seeds among multiple environments. **Theoretical and Applied Genetics**, Berlin, v. 118, n. 8, p. 1455-1463, 2009.

ZHANG, D.; CHENG, H.; GENG, L.; KAN, G.; CUI, S.; MENG, Q. GAI, J.; YU, D. Detection of quantitative trait loci for phosphorus deficiency tolerance at soybean seedling stage. **Euphytica**, Wageningen, v. 167, n. 3, p. 313-322, 2009.

ZHANG, W. K.; WANG, Y. J.; LUO, G. Z.; ZHANG, J. S. HE, C.Y.; WU, Y. L.; GAI, J. Y.; CHEN S. Y. QTL mapping of ten agronomic traits on the soybean (*Glycine max* L. Merr.) genetic map and their association with EST markers. **Theoretical and Applied Genetics**, Berlin, v. 108, n. 6, p. 1131-1139, 2004.

ZHU, Y. L.; SONG, Q. J. ; HYTEN, D. L.; VAN TASSELL, C. P.; MATUKUMALI, L. K.; GRIMM, D. R.; HYATT, S. M.; FICKUS, E. W.; YOUNG, N. D.; CREGAN, P. B. Single-nucleotide polymorphisms in soybean. **Genetics**, Austin, v. 163, n.11, p. 1123-1134, 2003.

ANEXO

Tabela 1 - Marcadores microssatélites selecionados para os testes de polimorfismos entre os cultivares Bossier e Embrapa 20

(continua)

Microssatélites	GL ^a	cM ^b	polimorfismo	Microssatélites	GL ^a	cM ^b	polimorfismo
Satt572	A1 (5)	14,65	P	Satt197	B1 (11)	46,39	P
Satt276	A1 (5)	17,16	P	Sat_247	B1 (11)	49,73	NP
Satt165	A1 (5)	23,00	NP	Sat_128	B1 (11)	53,41	NP
Satt042	A1 (5)	27,66	P	Sat_149	B1 (11)	54,01	P
Satt471	A1 (5)	27,96	NP	Satt519	B1 (11)	57,91	NP
Satt155	A1 (5)	32,68	P	satt298	B1 (11)	64,91	NP
Satt050	A1 (5)	46,45	P	sat_348	B1 (11)	71,97	NP
Satt385	A1 (5)	64,74	P	Satt597	B1 (11)	73,77	NP
Satt545	A1 (5)	71,39	NP	Satt332	B1 (11)	80,89	P
Satt599	A1 (5)	85,58	NP	satt415	B1 (11)	82,29	NP
Satt174	A1 (5)	88,58	NP	Satt583	B1 (11)	84,19	P
Satt200	A1 (5)	92,89	P	Satt444	B1 (11)	85,92	P
Satt236	A1 (5)	93,23	P	Sat_123	B1 (11)	100,88	NP
satt225	A1 (5)	95,16	P	Satt359	B1 (11)	102,56	NP
Satt211	A1 (5)	95,96	NP	Satt484	B1 (11)	118,53	NP
satt390	A2 (8)	9,14	NP	Satt453	B1 (11)	123,96	P
Sct_067	A2 (8)	14,99	P	Sat_331	B1 (11)	125,74	NP
Satt480	A2 (8)	28,44	NP	satt577	B1 (14)	6,05	NP
Satt589	A2 (8)	33,96	P	Satt467	B1 (14)	17,77	NP
Satt493	A2 (8)	35,02	NP	Satt126	B1 (14)	27,63	P
Satt177	A2 (8)	36,77	P	satt168	B1 (14)	55,2	NP
Satt315	A2 (8)	45,29	NP	Satt416	B1 (14)	56,96	P
Satt187	A2 (8)	54,92	P	satt304	B1 (14)	65,56	NP
satt424	A2 (8)	60,59	NP	satt020	B1 (14)	72,13	NP
Satt341	A2 (8)	77,70	NP	Satt556	B1 (14)	73,21	P
Sat_115	A2 (8)	81,02	P	sat_009	B1 (14)	78,66	NP
Sat_129	A2 (8)	84,08	NP	Satt066	B1 (14)	78,84	P
Satt089	A2 (8)	87,57	P	satt534	B1 (14)	87,59	NP
Satt377	A2 (8)	90,84	NP	Sct_064	B1 (14)	89,31	NP
Satt525	A2 (8)	96,97	NP	Satt063	B1 (14)	93,49	NP
Satt233	A2 (8)	100,09	P	Satt560	B1 (14)	97,92	P
Satt437	A2 (8)	107,05	NP	satt565	C1 (4)	0	NP
Satt508	A2 (8)	108,78	NP	Sct_186	C1 (4)	9,02	NP
satt327	A2 (8)	109,83	NP	SoyGpatr	C1 (4)	10,34	NP
Satt329	A2 (8)	110,94	P	Satt396	C1 (4)	24,11	P
Satt158	A2 (8)	115,25	NP	satt578	C1 (4)	65,08	NP
Satt421	A2 (8)	115,93	P	Satt190	C1 (4)	73,32	P
satt470	A2 (8)	116,73	NP	Satt136	C1 (4)	75,11	P
Sat_097	A2 (8)	122,05	NP	Satt361	C1 (4)	75,52	NP
satt455	A2 (8)	129,86	NP	Satt399	C1 (4)	76,23	P
Satt409	A2 (8)	145,57	P	satt294	C1 (4)	78,65	NP
Satt228	A2 (8)	154,11	NP	Sat_042	C1 (4)	82,51	P
Satt429	A2 (8)	162,03	NP	Satt195	C1 (4)	84,81	NP
satt378	A2 (8)	165,73	NP	satt524	C1 (4)	120,12	NP
Sat_272	B1 (11)	14,32	P	Satt338	C1 (4)	123,79	NP
Sat_270	B1 (11)	21,99	NP	Satt180	C1 (4)	127,77	P
Satt426	B1 (11)	28,33	NP	satt164	C1 (4)	132,46	NP
Sat_411	B1 (11)	30,87	NP	Sat_130	C2 (6)	11,39	NP
Satt509	B1 (11)	32,51	P	Satt227	C2 (6)	26,65	NP
Sat_156	B1 (11)	35,00	NP	Sat_062	C2 (6)	30,8	NP
Satt251	B1 (11)	36,48	P	Satt432	C2 (6)	38,05	NP
Satt638	B1 (11)	37,80	NP	Satt281	C2 (6)	40,30	NP

Tabela 1 - Marcadores microssatélites selecionados para os testes de polimorfismos entre os cultivares Bossier e Embrapa 20

				(continuação)			
Microssatélites	GL ^a	cM ^b	polimorfismo	Microssatélites	GL ^a	cM ^b	polimorfismo
Satt291	C2 (6)	45,76	NP	Satt216	D1b (2)	9,80	NP
Satt457	C2 (6)	56,51	NP	Satt095	D1b (2)	25,60	NP
Satt305	C2 (6)	69,67	NP	Satt157	D1b (2)	37,07	NP
Satt170	C2 (6)	70,56	NP	Satt558	D1b (2)	43,91	P
Satt322	C2 (6)	82,23	NP	Satt296	D1b (2)	52,61	P
Satt450	C2 (6)	89,31	NP	Satt542	D1b (2)	53,02	P
Satt643	C2 (6)	94,65	NP	Sat_135	D1b (2)	70,65	NP
Satt376	C2 (6)	97,83	NP	Satt141	D1b (2)	72,89	P
Satt363	C2 (6)	98,07	NP	Satt290	D1b (2)	73,35	P
Satt286	C2 (6)	101,75	NP	Satt005	D1b (2)	75,29	NP
Satt277	C2 (6)	107,59	NP	Satt428	D1b (2)	77,35	NP
Satt557	C2 (6)	112,19	NP	Satt041	D1b (2)	84,04	NP
Satt289	C2 (6)	112,35	NP	Satt546	D1b (2)	87,2	NP
Satt489	C2 (6)	113,39	NP	Satt172	D1b (2)	100,89	NP
Satt319	C2 (6)	113,42	NP	Sat_069	D1b (2)	102,60	NP
Satt100	C2 (6)	113,96	NP	Satt274	D1b (2)	116,35	NP
Satt708	C2 (6)	115,49	NP	Satt459	D1b (2)	118,62	NP
Satt460	C2 (6)	117,77	P	Satt271	D1b (2)	137,97	NP
Satt307	C2 (6)	121,27	P	Satt328	D2 (17)	16,76	NP
Satt202	C2 (6)	126,24	P	Satt458	D2 (17)	24,52	P
Satt316	C2 (6)	127,67	P	Satt135	D2 (17)	26,05	P
Satt371	C2 (6)	145,48	P	Satt498	D2 (17)	32,14	NP
Satt357	C2 (6)	151,91	P	Satt486	D2 (17)	34,09	NP
Satt184	D1a (1)	17,52	P	Satt372	D2 (17)	39,35	NP
Satt531	D1a (1)	40,87	NP	Satt002	D2 (17)	47,73	NP
Satt368	D1a (1)	43,84	NP	Satt443	D2 (17)	51,41	P
Satt482	D1a (1)	45,75	NP	Satt154	D2 (17)	57,07	NP
Satt320	D1a (1)	46,80	NP	Sat_092	D2 (17)	57,51	NP
Satt221	D1a (1)	47,04	NP	Satt447	D2 (17)	66,27	NP
Satt605	D1a (1)	47,73	NP	Satt397	D2 (17)	69,3	NP
Satt342	D1a (1)	48,14	NP	Satt389	D2 (17)	79,23	NP
Satt032	D1a (1)	48,63	NP	Satt461	D2 (17)	80,19	NP
Satt532	D1a (1)	49,07	NP	Satt311	D2 (17)	84,62	NP
Satt321	D1a (1)	50,16	NP	Satt226	D2 (17)	85,15	NP
satt548	D1a (1)	51,07	NP	Satt514	D2 (17)	85,69	NP
Satt295	D1a (1)	55,22	NP	Satt528	D2 (17)	86,34	NP
Satt179	D1a (1)	56,20	P	Satt082	D2 (17)	87,25	NP
Satt383	D1a (1)	56,57	NP	Satt574	D2 (17)	87,67	NP
Satt402	D1a (1)	57,75	NP	Satt543	D2 (17)	88,02	NP
Satt203	D1a (1)	59,00	NP	Satt488	D2 (17)	89,20	P
Satt370	D1a (1)	60,99	NP	Satt301	D2 (17)	93,71	NP
Satt580	D1a (1)	62,37	NP	Satt186	D2 (17)	105,45	NP
Satt507	D1a (1)	64,52	P	Satt310	D2 (17)	107,49	NP
Satt198	D1a (1)	68,62	P	Satt413	D2 (17)	113,61	NP
Satt468	D1a (1)	69,91	P	Satt031	D2 (17)	115,70	NP
Satt436	D1a (1)	70,69	P	Sat_086	D2 (17)	118,66	NP
Satt439	D1a (1)	72,27	NP	Sat_022	D2 (17)	120,30	P
Satt407	D1a (1)	99,59	NP	Satt256	D2 (17)	124,31	NP
Satt071	D1a (1)	100,39	NP	Satt386	D2 (17)	125,00	NP
Satt408	D1a (1)	106,69	NP	Satt575	E (15)	3,30	NP
Satt147	D1a (1)	108,89	NP	Satt213	E (15)	3,72	NP

Tabela 1 - Marcadores microssatélites selecionados para os testes de polimorfismos entre os cultivares Bossier e Embrapa 20

				(continuação)			
Microssatélites	GL ^a	cM ^b	polimorfismo	Microssatélites	GL ^a	cM ^b	polimorfismo
Satt384	E (15)	19,30	NP	Satt199	G (18)	62,16	NP
Satt212	E (15)	32,27	NP	Satt012	G (18)	66,55	NP
Satt598	E (15)	34,20	P	Satt517	G (18)	69,87	NP
Satt573	E (15)	35,79	P	Satt288	G (18)	76,77	P
Sat_136	E (15)	39,16	NP	Satt472	G (18)	94,84	NP
Sat_107	E (15)	43,05	NP	Satt191	G (18)	96,57	NP
Satt185	E (15)	44,76	P	Sat_117	G (18)	100,00	NP
Satt045	E (15)	46,65	P	SCt_187	G (18)	107,11	NP
Satt369	E (15)	56,27	P	Sat_064	G (18)	108,70	P
Satt553	E (15)	67,92	NP	Satt353	H (12)	8,48	NP
Satt231	E (15)	70,23	NP	Satt568	H (12)	27,64	P
Satt230	E (15)	71,31	NP	Sat_127	H (12)	28,80	P
Satt325	F (13)	2,23	P	Satt192	H (12)	44,04	P
Satt269	F (13)	11,37	NP	Satt442	H (12)	46,95	P
Satt252	F (13)	16,08	P	Sat_122	H (12)	61,33	NP
Sat_039	F (13)	27,87	NP	Sat_118	H (12)	61,86	NP
Satt160	F (13)	33,19	NP	Satt222	H (12)	68,08	P
Satt425	F (13)	43,44	NP	Satt279	H (12)	68,50	P
Satt516	F (13)	44,42	P	Satt314	H (12)	69,12	NP
Sat_133	F (13)	50,78	P	Satt302	H (12)	81,04	P
Satt114	F (13)	63,69	NP	Satt142	H (12)	86,49	NP
Sat_120	F (13)	75,97	P	Satt293	H (12)	89,09	NP
Satt334	F (13)	78,06	NP	Satt317	H (12)	89,52	NP
Sat_188	F (13)	85,33	NP	Satt181	H (12)	91,12	P
Satt072	F (13)	87,01	NP	Satt434	H (12)	105,74	P
Satt490	F (13)	97,97	NP	Satt571	I (20)	18,50	NP
Satt144	F (13)	102,08	NP	Satt451	I (20)	20,34	P
Satt554	F (13)	111,89	NP	Satt562	I (20)	22,84	P
Satt218	F (13)	117,65	NP	Satt367	I (20)	27,98	NP
Satt522	F (13)	119,19	NP	Satt587	I (20)	31,49	P
Sat_090	F (13)	130,64	NP	Satt127	I (20)	35,35	P
Sat_074	F (13)	142,35	NP	Satt239	I (20)	36,94	P
Satt395	F (13)	146,42	P	Satt354	I (20)	46,22	P
Satt038	G (18)	1,84	NP	Satt270	I (20)	50,11	P
Satt275	G (18)	2,20	NP	Satt049	I (20)	58,82	NP
Satt309	G (18)	4,53	NP	Sat_104	I (20)	65,62	NP
Satt356	G (18)	12,18	P	Satt330	I (20)	77,84	NP
Satt570	G (18)	12,74	NP	Satt162	I (20)	86,74	P
Satt235	G (18)	21,89	NP	Satt148	I (20)	100,78	P
Satt130	G (18)	23,10	NP	Satt440	I (20)	112,70	NP
Sat_131	G (18)	31,33	NP	SCt_189	I (20)	113,77	NP
Satt324	G (18)	33,26	NP	Satt249	J (16)	11,74	NP
Satt394	G (18)	43,38	P	Satt405	J (16)	12,41	NP
sat_358	G (18)	45,49	NP	SCt_046	J (16)	24,09	NP
Satt501	G (18)	47,27	NP	Satt285	J (16)	25,51	P
Satt340	G (18)	48,54	NP	SCt_065	J (16)	32,10	NP
Satt427	G (18)	51,69	P	Satt414	J (16)	37,04	NP
Satt594	G (18)	52,94	NP	Satt406	J (16)	38,19	P
Satt303	G (18)	53,42	NP	Satt596	J (16)	39,64	NP
Sat_094	G (18)	56,18	NP	SCaa003	J (16)	39,96	NP
Satt564	G (18)	57,32	NP	Satt456	J (16)	40,83	P

Tabela 1 - Marcadores microssatélites selecionados para os testes de polimorfismos entre os cultivares Bossier e Embrapa 20

				(conclusão)			
Microssatélites	GL ^a	cM ^b	polimorfismo	Microssatélites	GL ^a	cM ^b	polimorfismo
Satt380	J(16)	43,01	NP	Satt201	M (7)	13,56	NP
Satt215	J(16)	44,08	P	Satt150	M (7)	18,58	P
SCt_001	J(16)	44,68	P	Satt567	M (7)	33,47	NP
Satt244	J(16)	65,04	P	Satt540	M (7)	35,85	P
Satt547	J(16)	67,79	P	Satt463	M (7)	50,1	NP
Satt431	J(16)	78,57	P	Satt245	M (7)	53,54	P
Satt539	K (9)	1,80	P	Satt323	M (7)	60,05	NP
Sat_087	K (9)	4,85	NP	Satt536	M (7)	62,14	P
Satt242	K (9)	14,35	P	Sat_003	M (7)	62,31	P
Satt178	K (9)	40,86	P	Satt175	M (7)	66,99	P
Satt349	K (9)	42,39	NP	Satt494	M (7)	71,71	NP
Satt555	K (9)	42,71	NP	Satt306	M (7)	80,02	NP
Satt544	K (9)	43,35	P	Satt551	M (7)	95,45	NP
Satt247	K (9)	43,96	NP	Satt250	M (7)	107,70	NP
Satt381	K (9)	44,99	NP	Satt210	M (7)	112,08	P
Satt167	K (9)	45,74	NP	Satt346	M (7)	112,79	P
Satt375	K (9)	45,81	NP	Satt336	M (7)	133,83	P
Satt326	K (9)	49,53	NP	Satt152	N (3)	22,67	P
Satt001	K (9)	50,56	NP	Satt159	N (3)	27,13	NP
Sat_111	K (9)	55,70	NP	Satt530	N (3)	32,85	P
Satt273	K (9)	56,62	NP	Sat_084	N (3)	36,86	P
Sat_044	K (9)	58,01	P	Satt584	N (3)	37,98	P
Sat_043	K (9)	61,67	NP	Satt387	N (3)	53,25	NP
Satt499	K (9)	71,01	NP	Satt521	N (3)	65,46	NP
Satt475	K (9)	78,68	NP	Satt549	N (3)	70,60	P
Satt260	K (9)	80,12	NP	Satt237	N (3)	74,99	P
Sat_020	K (9)	103,60	P	Satt339	N (3)	75,91	P
Sat_126	K (9)	108,20	NP	Satt255	N (3)	76,49	P
Satt495	L (19)	0,00	P	Satt234	N (3)	84,60	P
Satt232	L (19)	10,35	P	Satt257	N (3)	92,56	P
Satt446	L (19)	11,47	NP	Satt358	O (10)	5,44	NP
Satt182	L (19)	14,03	NP	Sat_132	O (10)	8,75	P
Satt238	L (19)	19,93	NP	Satt445	O (10)	20,43	NP
Sat_071	L (19)	20,78	NP	Satt259	O (10)	39,82	P
Satt388	L (19)	23,55	NP	Satt347	O (10)	42,29	NP
Satt523	L (19)	27,92	NP	Satt420	O (10)	49,71	NP
Sat_134	L (19)	28,27	NP	Satt173	O (10)	58,40	P
Satt143	L (19)	30,19	NP	Satt563	O (10)	68,39	NP
Satt398	L (19)	30,58	P	Satt478	O (10)	71,10	P
Sat_195	L (19)	30,83	NP	Satt477	O (10)	82,09	NP
Satt418	L (19)	30,93	P	Satt123	O (10)	86,86	NP
Satt462	L (19)	41,00	P	SOyLBC	O (10)	95,00	P
Satt481	L (19)	54,57	P	Satt592	O (10)	100,38	P
Satt156	L (19)	56,14	P	Satt581	O (10)	106,03	P
Satt527	L (19)	70,36	NP	Sat_038	O (10)	112,17	P
Sat_099	L (19)	78,23	NP	Satt243	O (10)	119,5	NP
Satt229	L (19)	93,89	NP	Sat_109	O (10)	127,50	NP
Satt373	L (19)	107,24	P	Sat_108	O (10)	129,30	NP

^a GL – Grupos de ligação, os números entre parênteses correspondem ao número dos cromossomos, uma nomenclatura que tem sido utilizada recentemente em soja, ^b cM- centimorgan – os valores correspondem a posição dos marcadores no mapa consenso da soja (Cregan et al., 1999), P – marcadores polimórficos e NP- marcadores não polimórficos.