

## Mapping of QTLs associated with biological nitrogen fixation traits in soybean

MARIA APARECIDA SANTOS<sup>1,2</sup>, ISAIAS OLÍVIO GERALDI<sup>2</sup>, ANTONIO AUGUSTO FRANCO GARCIA<sup>2</sup>, NÁGILA BORTOLATTO<sup>1,3</sup>, ALINE SCHIAVON<sup>1,3</sup> and MARIANGELA HUNGRIA<sup>1</sup>

<sup>1</sup>Embrapa Soja, Londrina, Paraná, Brazil

<sup>2</sup>Escola Superior de Agricultura Luiz de Queiroz (ESALQ), Dep. Genética, Piracicaba, São Paulo, Brazil

<sup>3</sup>Centro Universitário Filadélfia (UNIFIL), Center of Biology, Londrina, Paraná, Brazil

**Santos, M. A., Geraldi, I. O., Garcia, A. A. F., Bortolatti, N., Schiavon, A. and Hungria, M. 2013.** Mapping of QTLs associated with biological nitrogen fixation traits in soybean. – *Hereditas* 150: 17–25. Lund, Sweden. eISSN 1601-5223. Received 9 July 2012. Accepted 3 May 2013.

Biological nitrogen fixation (BNF) is a key process, but despite the economic and environmental importance, few studies about quantitative trait loci (QTL) controlling BNF traits are available, even in the economically important crop soybean *Glycine max* (L.) Merr. In this study, a population of 157 F<sub>2,7</sub> RILs derived from crossing soybean cultivars Bossier (high BNF capacity) and Embrapa 20 (medium BNF capacity) was genotyped with 105 simple sequence repeat markers (SSRs). The genetic map obtained has 1231.2 cM and covers about 50% of the genome, with an average interval of 18.1 cM. Three traits, nodule number (NN), the ratio nodule dry weight (NDW)/NN and shoot dry weight (SDW) were used to evaluate BNF performance. A composite interval mapping for multiple traits method (mCIM) analysis mapped two QTLs for SDW (LGs E and L), three for NN (LGs B1, E and I), and one for NDW/NN (LG I); all QTLs were of small effect (R<sup>2</sup>-values ranging from 1.7% to 10.0%) and explained 15.4%, 13.8% and 6.5% of total variation for these three traits, respectively.

Mariangela Hungria, Embrapa Soja, Cx. Postal 231, 86001-970, Londrina, Paraná, Brazil. E-mail: mariangela.hungria@embrapa.br

Nitrogen (N), the element most limiting for crop growth, is usually supplied by application of fertilizer, but at substantial costs to farmers and with potentially adverse effects on the environment. Soybean *Glycine max* (L.) Merr., among other legumes, can obtain N through biological nitrogen fixation (BNF), and breeding for yield improvement can be performed considering BNF, inorganic-N supply, or both (COALE et al. 1985; KUMUDINI et al. 2008). The BNF results from the symbiosis between legumes with a diverse group of nitrogen fixing soil bacteria collectively known as rhizobia. This association is mutually beneficial: while bacteria provide a source of nitrogen to plant development, plants provide carbon sources for maintaining bacteria metabolism.

The establishment of the symbioses involves complex mechanisms starting with the mutual exchange of diffusible signal molecules. Plants secrete seed and root molecules – mainly flavonoids – that are sensed by specific rhizobia. In response, the rhizobia produce lipochitooligosaccharides namely Nod factors (NFs) that are recognized by LysM receptor-like kinase of the legumes. A new organ – the nodule – is then formed on the root and hosts the nitrogen fixing bacteria (SUBRAMANIAN et al. 2006; KOUCHI et al. 2010).

The elicitation and the development of an effective nodule are accompanied by the expression of specific genes in both rhizobia and their host-plants. The genetic

mechanisms of the bacteria are best understood and several rhizobial genes called *nod*, *nol* and *noe* have been described (MASSON-BOIVIN et al. 2009). On the other side, studies with the host-plants are more complex, due to biological characteristics and genome sizes. More recently, the integration of genetic and genomic approaches and the adoption of legume models with *Medicago truncatula* and *Lotus japonicum* opened a framework to the investigation of host genes essential for the BNF process. Twenty-six genes involved with Nod perception and different steps of nodule formation were characterized based on mutants, cloning and mapping strategies (KOUCHI et al. 2010). Furthermore, dozens of genes involved in nodule formation and in the BNF process called nodulin genes have been prospected by high throughput tools, such as expressed sequence tags and microarray analysis (BRECHENMACHER et al. 2008; LIBAULT et al. 2010).

In soybean, several loci controlling nodulation have been described since the 1950s. WILLIAMS and LYNCH (1954) reported the recessive gene *rl*<sub>1</sub> in a spontaneous non-nodulating mutant. Genes *Rj*<sub>2</sub> (CALDWELL 1966), *Rj*<sub>3</sub> (VEST 1970) and *Rj*<sub>4</sub> (VEST and CALDWELL 1972) have natural occurrence and control strain specific restricted nodulation. In addition, genes *rl*<sub>5</sub> and *rl*<sub>6</sub> (HARPER and NICKELL 1995) – conferring non-nodulation ability – have been identified by chemical induced mutation and

correspond to the same loci of the  $rj_i$ . Supernodulating mutants ( $rj_7$  and  $rj_8$ ) have also been generated by chemical mutagenesis (VUONG et al. 1996).

Some of these genes controlling nodulation in soybean have been cloned and their products have been described. The first gene related to nodulation control in soybean was cloned by SEARLE et al. (2003) using the  $rj_7$  and  $rj_8$  supernodulating mutants. By similar approaches the Nod factor receptor genes called NFR1 $\alpha$  and NFR1 $\beta$  (INDRASUMUNAR et al. 2011) and NFR5 $\alpha$  and NFR5 $\beta$  (INDRASUMUNAR et al. 2010) were cloned. Similar genes were first identified by mutant directed cloning in *Medicago truncatula* (AMOR et al. 2003; LIMPENS et al. 2003), *L. japonicus* (MADSEN et al. 2003; RADUTOIU et al. 2003) and pea (*Pisum sativum*) (ZHUKOV et al. 2008).

These discoveries provide important clues to understand the molecular mechanisms of host regulation of the symbiosis. However, more information with field orientation is necessary to access the potential of breeding for BNF. The occurrence of natural variability on nodulation and efficiency of BNF in soybean have been reported (NEUHAUSEN et al. 1988; SINCLAIR et al. 1991; PAZDERNIK et al. 1997; BOHRER and HUNGRIA 1998; HUNGRIA and BOHRER 2000). However, due to technical difficulties in evaluating phenotypic traits (nodulation and nitrogen fixation rates), the mechanisms and genetic control of these traits are poorly understood. Therefore, studies of genetic architecture of nodulation and nitrogen fixation are still required. An interesting and informative way to dissect the genetic architecture is to map quantitative trait loci (QTL). This approach has been broadly used in soybean to identify traits of interest, such as yield potential, disease resistance, contents of protein and oil in seeds (SOYBASE 2012).

Preliminary studies about genomic regions controlling BNF in soybean have been performed (TANYA et al. 2005; NICOLÁS et al. 2006; SANTOS et al. 2006). There is also information about other legumes, such as pea (BOURION et al. 2010) and common bean (*Phaseolus vulgaris* L.) (NODARI et al. 1993; TSAI et al. 1998). These studies have demonstrated the potential of mapping QTL for BNF traits, but additional information is still required to improve the existent knowledge.

The objective of this study was to improve the genome cover reported by SANTOS et al. (2006) in this same population; these authors used 24 simple sequence repeat markers (SSRs) to identify QTLs related to BNF and covered a small percentage (5%) of the genome. Now we have analyzed 432 SSRs, from which we selected 105 SSRs; the method of composite interval mapping for multiple traits (mCIM) was employed, altogether resulting in 50% coverage of the genome.

## MATERIAL AND METHODS

### *Plant materials and biological nitrogen fixation traits*

A population of 157 F<sub>2:7</sub> inbred lines derived from a cross between Bossier (high BNF capacity) and Embrapa 20 (medium BNF capacity), developed by NICOLÁS et al. (2002), was used and the experimental design and phenotypic evaluations were as described before (SANTOS et al. 2006). In brief, plants were grown in pots containing 4 kg of non-sterile soil and sand, with one plant per pot under greenhouse conditions, and the experiment was performed in a completely randomized design with eight replications (one plant per replication). The plants were inoculated at the V2 stage of development (FEHR et al. 1971), by adding 1 ml of inoculum containing two strains, *Bradyrhizobium japonicum* SEMIA 5079 (=CPAC 15 and belonging to the same serogroup as SEMIA 566 and USDA 123) and *Bradyrhizobium elkanii* SEMIA 587 (1:1, v/v). Six weeks after emergence plants were harvested for evaluation of nodulation (nodule number, NN; nodule dry weight, NDW, NDW/NN ratio) and plant growth (shoot dry weight, SDW). Phenotypic correlations for combinations of traits were estimated using the CORR procedure of SAS STAT 9.1, while genetic correlations were estimated according to FALCONER and MACKAY (1996).

### *Genotyping and construction of a genetic linkage map*

Leaf tissue DNA extraction and genotyping methods have been reported by SANTOS et al. (2006). A total of 432 simple sequence repeat (SSR) markers covering the 20 chromosomes of soybean genome were screened against the parents. The population was genotyped with 105 SSRs. To analyze the linkages between markers and to create a genetic map, the program Mapmaker/EXP ver. 3.0 was used (LANDER et al. 1987). Microsatellite markers were initially grouped using the default parameters that included the Kosambi's function, the linkage LOD score greater than 3.0 and a maximum genetic distance of 37.2 cM. The commands "group", "compare" and "map" were used for building the linkage groups. All non-anchored markers were added by using the "try" multipoint analysis command, and the results were checked by the "ripple" command with a window size of 5. Linkage groups were named using the designation of the consensus map (CREGAN et al. 1999; SONG et al. 2004).

### *QTL mapping of biological nitrogen fixation*

To map QTL controlling BNF, we used the multiple-trait composite interval mapping (mCIM) proposed by JIANG and ZENG (1995), that considers all traits simultaneously. The hypothesis tested was that an interval flanked by two

adjacent markers contains a pleiotropic QTL controlling at least one BNF trait. The likelihood-ratio statistical test (LR) is estimated by  $-2\ln(L_0/L_1)$ , where  $L_0$  is the maximum likelihood under the null hypothesis that  $a_1 = a_2 = \dots a_i = 0$ , where  $a_i$  is the additive effect of the QTL for trait  $i$ , and  $L_1$  is the maximum likelihood under the alternative hypothesis that at least one  $a_i \neq 0$ . In other words, this is a test for the presence of a pleiotropic QTL controlling one or more traits in one determined interval. The mapping analysis was performed using the module *JZmapqtl* of the program QTL Cartographer ver. 1.17 (BASTEN et al. 2003). Model 6 was adopted and a limited number of background markers (BG) for the mCIM analysis was identified via the forward/backward stepwise regression option using conservative probability thresholds ( $P_{in} = 0.05$ ;  $P_{out} = 0.05$ ). Genome scan was performed using a window size of 5 cM and a walk speed of 1 cM.

The threshold for significance of QTL detection was adjusted for the number of independent tests in the genome scan, following what has been suggested by ZENG (1994), JIANG and ZENG (1995) and VIEIRA et al. (2000). Due to the properties of the multiple regression model used for mCIM, the effective number of independent tests depends on the size of the genome region at each side of the test interval in which marker cofactors are not fitted (window size). We used a window size of 37.5 cM (recombination fraction of 0.50 considering Kosambi's map function) as independent. The number of independent tests was then estimated by  $\sum_i [(T_i/28.11) + 1]$ , where  $T_i$  is the total estimated map length in cM of the  $i$ th linkage group, excluding gaps greater than 37.5 cM. Consequently, for our data, there are 27.67 independent tests in total. Thus, a type I error rate of  $\alpha = 0.05/27.67$  was used in the joint mapping analysis, corresponding to a  $\chi^2$ -value of 17.52 (five degrees of freedom, one for the additive effect of each trait plus one for estimating QTL position), equivalent to an LOD score of 3.80. Since for marginal analysis of each trait statistical tests are applied only in positions with putative QTL identified by the pleiotropic model, a type I error of  $\alpha = 0.05$  (LOD score of 0.83) could be used (JIANG and ZENG 1995; VIEIRA et al. 2000).

The proportion of the variance ( $R^2$ ) explained by each particular QTL and the additive (a) effects were obtained from the mCIM analysis. The total phenotypic variance explained by two or more QTLs simultaneously for a given trait was determined by adjusting a model with all QTLs using a Windows version of QTL Cartographer (WANG et al. 2005). Regarding the signals of additive effects, "+" indicates increasing allele from the Bossier cultivar; "-", indicates decreasing allele from Embrapa 20. The use of mCIM for mapping QTLs for correlated traits is also important for evaluating the cause of correlation

between traits (pleiotropy or linkage); in our study, pleiotropy was assumed if the map position for two marginal QTLs were equal or smaller than 3 cM.

## RESULTS

### *Phenotypic analysis*

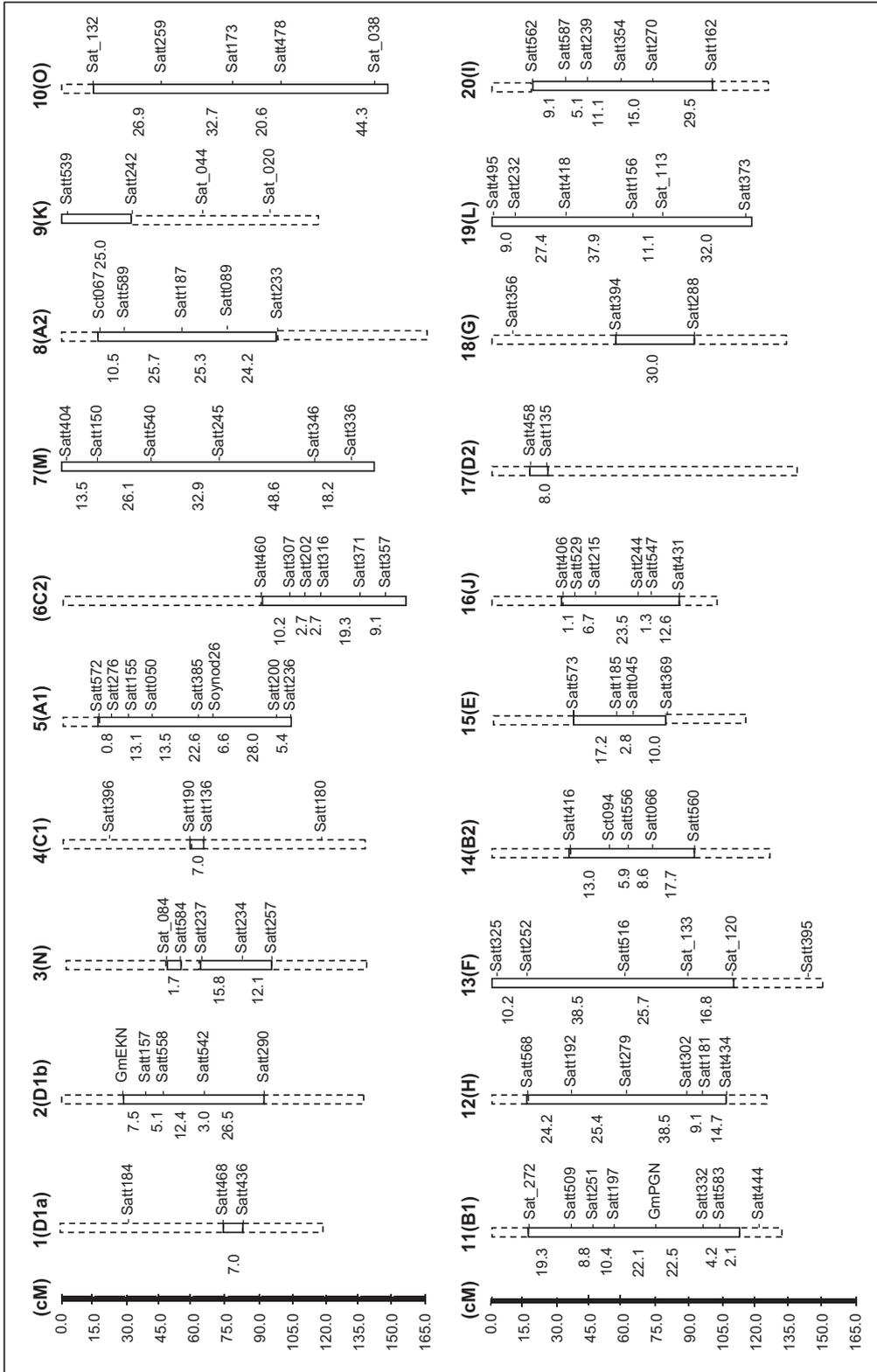
SANTOS et al. (2006) reported significant variance among nodulation traits (NDW, NN, NDW/NN) and plant growth (SDW) means, indicating genetic variability of the RILs population. As expected (SOUZA et al. 2008), high values of coefficient of variation (>25.6) and medium to low values of broad-sense heritabilities (0.49 for SDW, 0.33 for NDW, 0.33 for NN and 0.27 for NDW/NN) were observed in the dataset. We estimated positive and high genetic and phenotypic correlations for the traits NDW and SDW ( $r_G = 0.83$  and  $r_p = 0.66$ ), and for NDW and NN ( $r_G = 0.65$  and  $r_p = 0.64$ ); medium for SDW and NN ( $r_G = 0.51$  and  $r_p = 0.43$ ); and low for SDW and NDW/NN ( $r_G = 0.26$  and  $r_p = 0.24$ ), NDW and NDW/NN ( $r_G = 0.23$  and  $r_p = 0.33$ ). Negative correlations were verified between NN and NDW/NN ( $r_G = -0.60$  and  $r_p = -0.40$ ).

### *Polymorphism and genetic mapping*

A total of 432 SSR markers were used to detect polymorphisms between the two parental genotypes and 105 polymorphic markers were used to genotype the population of RILs. From these, 97 SSR markers were mapped in 20 LGs encompassing about 1231.2 cM with a mean distance of 18.1 cM between markers, representing a coverage of about 50% of the genome if we consider a consensus linkage of 2536 cM (SONG et al. 2004). Gaps of more than 30 cM between linked markers were detected in the LGs F, G, L, M and O (Fig. 1). The remaining eight markers (Satt421, Satt396, Satt180, Satt184, Satt395, Satt356, Sat\_044 and Sat\_020) were not linked to any group. In general, the order and relative distances of the SSR markers within LGs were similar to the information obtained in the public soybean genetic map (CREGAN et al. 1999; SONG et al. 2004). An exception was observed for the SSR marker GmEKN, located within a hypothetical nodulin gene and mapped in LG A2 by JEONG et al. (2006), while in our study it was mapped on LG D1b (LOD > 10). On average, five markers were mapped per LG, ranging from two on LGs C1, D1a, D2, G and K to eight on LG A1 and B1.

### *QTL mapping of BNF traits*

As explained before, three traits were used in this analysis to avoid redundancy, NN, NDW/NN and SDW. Four genomic regions were significantly associated with BNF traits, mapped on LGs B1, E, L and I. Two QTLs (*nm1-B1*



**Fig. 1.** Soybean linkage map based on a 157 F<sub>2:7</sub> RILs population derived from Bossier (high capacity of biological nitrogen fixation) × Embrapa 20 (medium capacity of biological nitrogen fixation). The linkage groups were named with designations of the consensus map of CREGAN et al. (1999) and SONG et al. (2004). The uninterrupted rectangles correspond to the covered regions of the genome and the disrupted rectangles represent not covered regions in the present study.

and *sdw2-L*) were mapped with the LOD score threshold ( $\text{LOD} \geq 3.80$ ) and two QTLs (*bnf3-E* and *bnf4-I*) were mapped with a LOD score slightly smaller than the threshold ( $\text{LOD} = 3.60$ ), considered as probable QTLs. From these, two regions located on LGs E and I were associated with more than one trait, while the regions located on LGs B1 and L were significant only for one trait (Table 1, Fig. 2). The QTLs mapped jointly were named *bnf3-E* and *bnf4-I*. The QTL *bnf3-E* located in the interval Satt573-Satt185 was significant for SDW and NN traits. The estimated additive effects of alleles were  $-0.29 \text{ g plant}^{-1}$  for SDW and  $-10.23 \text{ nodules plant}^{-1}$  for NN and explained 5.4% and 4.3% of the individual phenotypic variance ( $R^2$ ), respectively. The QTL *bnf4-I* located in the interval Satt587-Satt354, was significant for NN and NDW/NN. The estimated additive effects of alleles were 8.92 nodules  $\text{plant}^{-1}$  and  $-0.21 \text{ mg nodule}^{-1}$  and explained 1.7% and 6.5% of the individual phenotypic variance. The correlation for BNF traits observed in QTLs *bnf3-E* and *bnf4-I* in this work was due to genetic linkage (Table 1).

Two specific QTLs controlling SDW were mapped on LGs E and L, and three QTLs controlling nodulation traits (NN, NDW/NN) were mapped on LG B1, E and I (Table 1). With LOD scores of 4.20 and 3.80, respectively, the QTLs *sdw2-L* and *nn1-B1* were the only ones that

reached the empirical LOD score threshold. The additive effects of these QTLs were of  $0.57 \text{ g plant}^{-1}$  for SDW and of  $-13.8 \text{ nodules plant}^{-1}$  for NN, and explained 10.0% and 7.8% of individual phenotypic variance, respectively, corresponding to the highest values explained by the mapped QTLs in this study.

In summary, in this study we have mapped two QTLs for SDW (LGs E and L), three for NN (LGs B1, E and I), and one for NDW/NN (LG I). All QTLs were of small effect ( $R^2$ -values ranging from 1.7% to 10.0%) and explained 15.4%, 13.8% and 6.5% of total variation for these three traits, respectively, from the adjusted model with all QTLs.

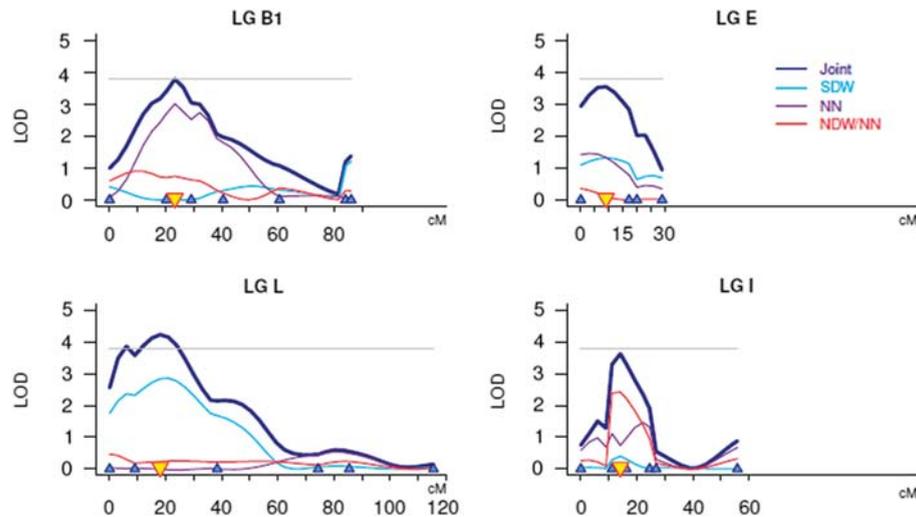
## DISCUSSION

The role of BNF in world crop production has been emphasized for several decades. In Brazil, estimates are of an economy of about US\$ 7 billion per season only with the soybean crop. Furthermore, the environment is benefited by the lower pollution of water reservoirs, and by the reduction of the emission of gases with greenhouse effects (HUNGRIA et al. 2007). However, as pointed out by GRAHAM and VANCE (2000), there are several areas where scientists need to concentrate efforts to ensure BNF benefits, including environmental conditions (soil degradation,

Table 1. A summary of QTLs controlling multiple traits for biological nitrogen fixation (BNF) in a population of 157  $F_{2:7}$  recombinant inbred lines (RILs) of soybean, derived from Bossier  $\times$  Embrapa-20 parents. The traits related to BNF are: nodule number (NN, no.  $\text{plant}^{-1}$ ), nodule dry weight ratio (NDW/NN, mg no.  $^{-1}$ ), and plant growth: shoot dry weight (SDW, g  $\text{plant}^{-1}$ ).

| LG/QTL            | Interval        | Analysis | Position (cM) | LOD   | Effect a | $R^2$ % | Linkage vs pleiotropy                 |
|-------------------|-----------------|----------|---------------|-------|----------|---------|---------------------------------------|
| B1/ <i>nn1-B1</i> | Satt509-Satt251 | jointly  | 23.28         | 3.80  |          |         |                                       |
|                   |                 | SDW      | 38.12         | 0.26  | -0.13    | 0.2     |                                       |
|                   |                 | NN       | 23.28         | 3.00* | -13.8    | 7.8     |                                       |
|                   |                 | NDW/NN   | 23.28         | 0.74  | 0.09     | 2.8     |                                       |
| E/ <i>bnf3-E</i>  | Satt573-Satt185 | jointly  | 8.01          | 3.60  |          |         |                                       |
|                   |                 | SDW      | 9.01          | 1.30* | -0.29    | 5.4     | genetic linkage between SDW and NN    |
|                   |                 | NN       | 4.01          | 1.50* | -10.23   | 4.3     |                                       |
|                   |                 | NDW/NN   | 0.01          | 0.40  | 0.07     | 0.3     |                                       |
| L/ <i>sdw2-L</i>  | Satt232-Satt418 | jointly  | 18.00         | 4.20  |          |         |                                       |
|                   |                 | SDW      | 20.00         | 2.86* | 0.57     | 10.0    |                                       |
|                   |                 | NN       | 9.00          | 0.02  | -1.20    | 0.2     |                                       |
|                   |                 | NDW/NN   | 24.00         | 0.25  | 0.09     | 0.5     |                                       |
| I/ <i>bnf4-I</i>  | Satt587-Satt354 | jointly  | 14.23         | 3.60  |          |         |                                       |
|                   |                 | SDW      | 14.23         | 0.40  | -0.16    | 2.0     | genetic linkage between NN and NDW/NN |
|                   |                 | NN       | 9.12          | 1.28* | 8.92     | 1.7     |                                       |
|                   |                 | NDW/NN   | 13.12         | 2.50* | -0.21    | 6.5     |                                       |

\*traits significantly mapped.



**Fig. 2.** QTL mapped obtained using multiple-trait composite interval mapping (mCIM) involving four biological nitrogen fixation traits (nodule dry weight, NDW; nodule number, NN; nodule dry weight ratio, NDW/NN; and shoot dry weight, SDW) in a population of 157 recombinant inbred lines (RILs) of soybean. The blue triangles indicate the position of markers on the linkage groups; putative QTL mapped under the pleiotropic model (blue lines) are indicated by yellow triangles. Marginal analysis for each trait is represented by different colors on the figure.

soil acidification and salinization, phosphorus supply and crop rotation), strain and inoculant improvement, and plant breeding.

The importance of plant breeding in increasing global BNF has been highlighted a long time ago (LARUE 1978), and since then, our understanding of the molecular mechanisms involved in the interaction between legumes and rhizobia has greatly improved (KOUCHI et al. 2010). However, few progresses have been achieved towards establishing a procedure for assessment of BNF by breeding programs, that should include simple, rapid and non-destructive field-based techniques (LARUE 1978; GRAHAM and VANCE 2000). Nowadays, the molecular assisted selection is an outstanding alternative for this proposal.

We present here an effort to identify molecular markers associated with BNF traits. In agreement with the complex nature of BNF, the mapped QTLs explained a small proportion of the total phenotypic variation effect. In soybean, phenotypic evaluations showing considerable variation under different environmental conditions were reported by CREGAN and KEYSER (1986), BETTS and HERRIDGE (1987), BOHRER and HUNGRIA (1998) and HUNGRIA and BOHRER (2000). Furthermore, the proportion of the genetic/environment contribution to the phenotype within experimental populations is highly influenced by environmental conditions (RONIS et al. 1985; PAZDERNIK et al. 1996; NICOLÁS et al. 2002; SANTOS et al. 2006). In addition, the nature of the BNF parameter is also distinct and quite variable, and

coefficients of variation larger than 35% are often reported (BOHRER and HUNGRIA 1998; HUNGRIA and BOHRER 2000; SOUZA et al. 2008).

From the four regions controlling BNF identified in this study, two (LGs B1 and L) have been previously reported by NICOLÁS et al. (2006), confirming the association of these regions with BNF traits. Furthermore, two new QTLs were mapped (LGs E and I) as a result of a greater coverage of the genome. On the other hand, some QTLs mapped on D1b, H and B2 (NICOLÁS et al. 2006; SANTOS et al. 2006), and on A1, D1b, J, I and O (TANYA et al. 2005) were not identified in our study. We attributed this to the fact that these studies used a single marker approach (regression analysis), that could result in false positives, since cofactors were not included in the model. Additionally, it is likely that QTLs are highly dependent on specific environments, or that they do not always segregate in the mapping population.

The method of QTL mapping using a composite interval mapping for multiple traits method (mCIM) takes into account correlations among traits. This joint analysis improves the statistical power of the tests, the precision of parameter estimation, and allows inference of the genetic causes (genetic linkage or pleiotropy) of the occurrence of coincident QTLs (JIANG and ZENG 1995). We identified two coincident regions controlling more than one trait of BNF (LGs E and I). Genetic linkage was the cause of coincident QTLs, SDW and NN on LG E, and NN and NDW/NN on LG I (Table 1). In the case of linked QTLs, a selection based on molecular markers could

promote increases in NN and NDW/NN despite the negative correlation observed for these two traits.

The QTL *nn1-B1* was located in a region previously associated with nodulation traits by NICOLÁS et al. (2006). Coincidentally, this region has been associated with isoflavones (TODA et al. 2002; KASSEM et al. 2006) and, as is well known, these metabolites are involved in molecular signaling between the host plant and the rhizobium, starting the process of nodule formation (HUNGRIA and STACEY 1997; SUBRAMANIAN et al. 2006). This QTL may result from a pleiotropic effect, or might be closely linked to genes involved in the metabolism of isoflavones.

Other interesting observation is that in the region of the QTLs *nn1-B1*, INDRASUMUNAR et al. (2010) cloned a gene (Nod Factor Reception 5 $\alpha$  - NFR5 $\alpha$ ) involved with the recognition of Nod factors secreted by *Bradyrhizobium*. This result corroborates with the evidence that this region host genes that control nodulation in soybean.

For decades we have performed extensive work showing that SDW can be considered an excellent parameter to estimate the contribution of BNF on soybean. These studies included experiments with 152 cultivars inoculated with *Bradyrhizobium* and non-inoculated controls with and without chemical N-fertilizers under greenhouse controlled conditions in sterile substrate (BOHRER and HUNGRIA 1998) and pots containing non-sterile soils (HUNGRIA and BOHRER 2000). The feasibility of using SDW to evaluate BNF in soybean was also confirmed under field conditions, in several trials performed in soils poor on N (SOUZA et al. 2008). In this study, for the SDW trait, we identified two QTLs (LGs E and L) that together explained 15.4% of the phenotypic variation (Table 1). The QTL *sdw2-L* showed the highest R<sup>2</sup>, explaining 10% of the phenotypic variation for SDW and was also reported by NICOLÁS et al. (2006). Moreover, QTLs for plant height and lodging are described in this region (CHASE et al. 1996), so we consider this QTL as highly significant for SDW.

The QTL *bnf4* coincides with a major QTL for seed protein content in LG I (NICHOLS et al. 2006). Soybean is one of the highest protein content crop species and thus shows a high demand on N. There are studies showing that the N derived from BNF is more easily moved to pods during the grain filling than is N-nitrate, both in soybean (ISRAEL et al. 1985) and in common bean (HUNGRIA et al. 1985; HUNGRIA and NEVES 1987).

The QTLs identified in this study may be very useful in future marker-assisted selections to improve BNF in soybean. Furthermore, the soybean genome recently published reports the identification of fifty-two genes related to BNF (28 nodulins and 24 regulatory genes) (SCHMUTZ et al. 2010), and the crossing of information from the genome with QTL mapping can provide evidence about the function of these genes in the BNF process.

## Conclusions

Nitrogen fixation is, undoubtedly, an important component of sustainable agricultural systems and is specifically important for soybean, which accumulates high amounts of N. However, BNF is dependent on genetics and environmental factors, and the evaluation of traits related to it is laborious; therefore, selection assisted by molecular markers may be an approach worthy of consideration for improvement of BNF in breeding programs. We reported here four genomic regions associated with BNF; two of them, B1 and L, have been described in other mapping populations and can be useful in marker-assisted selection. Subsequent efforts to saturate the genetic map may allow the identification of other QTLs controlling BNF.

**Acknowledgements** – This research was partially supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CNPq-MAPA (577933/2008) and CNPq-Repensa (562008/2010-1). M. A. Santos had a PhD fellowship from CNPq. I. O. Geraldi, A. A. F. Garcia and M. Hungria have research fellowships from CNPq. Authors thank Rinaldo B. Conceição and Ligia M. O. Chueire for technical help and Dr. Allan R. J. Eaglesham for scientific and English suggestions.

## REFERENCES

- Amor, B. B., Shaw, S. L., Oldroyd, G. E. D. et al. 2003. The NFP locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. – *Plant J.* 34: 495–506.
- Basten, C. J., Weir, B. S. and Zeng, Z. B. 2003. QTL cartographer: ver. 1.17. – Dept of Statistics, North Carolina State Univ., Raleigh, NC.
- Betts, J. H. and Herridge, D. F. 1987. Isolation of soybean lines capable of nodulation and nitrogen fixation under high levels of nitrate supply. – *Crop Sci.* 27: 1156–1161.
- Bohrer, T. R. J. and Hungria, M. 1998. Avaliação de cultivares de soja quanto à fixação biológica do nitrogênio. – *Pesq. Agropec. Bras.* 33: 937–952.
- Bourion, V., Rizvi, S. M. H., Fournier, S. et al. 2010. Genetic dissection of nitrogen in pea through a QTL approach of root, nodule and shoot variability. – *Theor. Appl. Genet.* 121: 71–86.
- Brechenmacher, L., Kim, M. Y., Benitez, M. et al. 2008. Transcription profiling of soybean nodulation by *Bradyrhizobium japonicum*. – *Mol. Plant–Microbe Interact.* 21: 631–645.
- Caldwell, B. E. 1966. Inheritance of a strain-specific ineffective nodulation in soybean. – *Crop Sci.* 6: 427–428.
- Chase, K., Cregan, P. B., Jarvik, T. et al. 1996. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. – *Crop Sci.* 36: 1327–1336.
- Coale, F. J., Meisinger, J. J. and Wiebold, W. J. 1985. Effects of plant breeding and selection on yields and nitrogen fixation in soybeans under two soil nitrogen regimes. – *Plant Soil* 86: 357–367.
- Cregan, P. B. and Keyser, H. H. 1986. Host restriction of nodulation by *Bradyrhizobium japonicum* strain USDA 123 in soybean. – *Crop Sci.* 26: 911–916.

- Cregan, P. B., Jarvik, T., Bush, A. L. et al. 1999. An integrated genetic linkage map of the soybean genome. – *Crop Sci.* 39: 1464–1491.
- Falconer, D. S. and Mackay, T. F. C. 1996. Introduction to quantitative genetics. – Longman.
- Fehr, W. R., Caviness, C. E., Burmood, D. T. et al. 1971. Stages of development descriptions for soybeans, *Glycine max* (L.) Merrill. – *Crop Sci.* 11: 929–931.
- Graham, P. H. and Vance, C. P. 2000. Nitrogen fixation in perspective: an overview of research and extension needs. – *Field Crops Res.* 65: 93–106.
- Harper, J. E. and Nickell, C. D. 1995. Genetic analysis of non-nodulating soybean mutants in a hypernodulating background. – *Soybean Genet. Newslett.* 22: 185–190.
- Hungria, M. and Neves, M. C. P. 1987. Partitioning of nitrogen from biological fixation and fertilizer in *Phaseolus vulgaris*. – *Physiol. Plant.* 69: 55–63.
- Hungria, M. and Stacey, G. 1997. Molecular signals exchanged between host plants and rhizobia: basic aspects and potential application in agriculture. – *Soil Biol. Biochem.* 29: 819–830.
- Hungria, M. and Bohrer, T. R. J. 2000. Variability of nodulation and dinitrogen fixation capacity among soybean cultivars. – *Biol. Fertil. Soils* 31: 45–52.
- Hungria, M., Neves, M. C. P. and Victoria, R. L. 1985. Assimilação do nitrogênio pelo feijoeiro II. Absorção e translocação do N mineral e do N<sub>2</sub> fixado. – *Rev. Bras. Cienc. Solo* 9: 201–209.
- Hungria, M., Campo, R. J. and Mendes, I. C. 2007. A importância do processo de fixação biológica do nitrogênio para a cultura da soja: componente essencial para a competitividade do produto brasileiro. Documentos 283. – Embrapa-Soja, Brazil.
- Indrasumunar, A., Kereszt, A., Searle, I. et al. 2010. Inactivation of duplicate Nod Factor Receptor 5 (NFR5) genes in recessive loss-of-function non-nodulation mutants of allotetraploid soybean (*Glycine max* L. Merr.). – *Plant Cell Physiol.* 51: 201–214.
- Indrasumunar, A., Searle, L., Lin, M. H. et al. 2011. Nodulation factor receptor kinase 1 $\alpha$  controls nodule organ number in soybean (*Glycine max* L. Merr.). – *Plant J.* 65: 39–50.
- Israel, D., Burton, J. W. and Wilson, R. F. 1985. Studies on genetic male-sterile soybeans IV. Effect of male sterility and source of nitrogen nutrition on accumulation, partitioning and transport of nitrogen. – *Plant Physiol.* 78: 762–767.
- Jeong, S. C., Yang, K., Park, J. Y. et al. 2006. Structure, expression and mapping of two nodule-specific genes identified by mining public soybean EST databases. – *Gene* 393: 71–80.
- Jiang, C. and Zeng, Z. B. 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. – *Genetics* 140: 1111–1127.
- Kassem, M. A., Shultz, J., Meksem, K. et al. 2006. An updated ‘Esses’ by ‘Forrest’ linkage map and first composite interval map of QTL underlying soybean traits. – *Theor. Appl. Genet.* 113: 1015–1026.
- Kouchi, H., Imaizumi-Anraku, H., Hayashi, M. et al. 2010. How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. – *Plant Cell Physiol.* 51: 1381–1397.
- Kumudini, S., Omielan, J. and Hume, D. J. 2008. Soybean genetic improvement in yield and the effect of late-season shading and nitrogen source and supply. – *Agron. J.* 100: 400–405.
- Lander, E. S., Green, P., Abrahamson, J. et al. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkages maps or experimental and natural populations. – *Genomics* 1: 174–181.
- LaRue, T. A. 1978. Selecting and breeding legumes for enhanced nitrogen fixation. – In: Proceedings of a Workshop, Boyce Thompson Inst., Cornell, 1987.
- Libault, M., Farmer, A., Brechenmacher, L. et al. 2010. Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection. – *Plant Physiol.* 152: 541–552.
- Limpens, E., Franken, C., Smit, P. et al. 2003. LysM Domain receptor kinases regulating rhizobial nod factor-induced infection. – *Science* 302: 630–633.
- Madsen, E. B., Madsen, L. H., Radutoiu, S. et al. 2003. A receptor-like kinase gene involved in legume perception of rhizobial signal molecules. – *Nature* 425: 637–640.
- Masson-Boivin, C., Giraud, E., Perret, X. et al. 2009. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? – *Trends Microbiol.* 17: 458–466.
- Neuhausen, S. L., Graham, P. H. and Orf, J. H. 1988. Genetic variation for dinitrogen fixation in soybean of maturity group 00 and 0. – *Crop Sci.* 28: 769–772.
- Nichols, D. M., Glovers, K. D., Carlson, S. R. et al. 2006. Fine mapping of a seed protein QTL on soybean linkage group I and its correlated effects on agronomic traits. – *Crop Sci.* 46: 834–839.
- Nicolás, M. F., Arias, C. A. A. and Hungria, M. 2002. Genetics of nodulation and nitrogen fixation in Brazilian soybean cultivars. – *Biol. Fertil. Soils* 36: 109–117.
- Nicolás, M. F., Hungria, M. and Arias, C. A. A. 2006. Identification of quantitative trait loci controlling nodulation and shoot mass in progenies from two Brazilian soybean cultivars. – *Field Crops Res.* 95: 355–366.
- Nodari, R. O., Tsai, S. M., Busman, P. et al. 1993. Towards an integrated linkage map of common bean III. Mapping genetic factors controlling host bacteria interactions. – *Genetics* 134: 341–350.
- Pazdernik, D. L., Graham, P. H., Vance, C. P. et al. 1996. Host genetic variation in the early nodulation and dinitrogen fixation of soybean. – *Crop Sci.* 36: 1102–1107.
- Pazdernik, D. L., Graham, P. H. and Orf, J. H. 1997. Heritability in the early nodulation of F3 and F4 soybean lines. – *Can. J. Plant Sci.* 77: 201–205.
- Radutoiu, S., Madsen, L. H., Madsen, E. B. et al. 2003. Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. – *Nature* 425: 585–592.
- Ronis, D. H., Sammons, D. J., Kenworthy, W. J. et al. 1985. Heritability of total and fixed N content of the seed in two soybean populations. – *Crop Sci.* 23: 1–4.
- Santos, M. A., Nicolás, M. F. and Hungria, M. 2006. Identificação de QTL associados à simbiose entre *Bradyrhizobium japonicum*, *B. elkanii* e soja. – *Pesq. Agropec. Bras.* 41: 67–75.
- Schmutz, J., Cannon, S. B., Schlueter, J. et al. 2010. Genome sequence of the palaeopolyploid soybean. – *Nature* 463: 178–183.
- Searle, I. R., Men, A. E., Laniya, T. S. et al. 2003. Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. – *Science* 299: 109–112.
- Sinclair, T. R., Soffes, A. R., Hinson, K. et al. 1991. Genotypic variation in soybean nodule number and weight. – *Crop Sci.* 31: 301–304.
- Song, Q. J., Marek, L. F., Shoemaker, R. C. et al. 2004. A new integrated genetic linkage map of the soybean. – *Theor. Appl. Genet.* 109: 122–129.

- Souza, R. A., Hungria, M., Franchini, J. C. et al. 2008. Conjunto mínimo de parâmetros para avaliação da microbiota do solo e da fixação biológica do nitrogênio pela soja. – *Pesq. Agropec. Bras.* 43: 83–91.
- Soybase 2012. Soybase and the soybean breeder's toolbox. Knowledge and tools for soybean breeders and researchers. <[www.soybase.org/](http://www.soybase.org/)> Accessed 3 June 2012.
- Subramanian, S., Stacey, G. and Yu, O. 2006. Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. – *Plant J.* 48: 261–273.
- Tanya, P., Srinives, P., Toojinda, T. et al. 2005. Identification of SSR markers associated with N<sub>2</sub> fixation components in soybean [*Glycine max* (L.) Merr.]. – *Korean J. Genet.* 27: 351–359.
- Toda, K., Yang, D., Yamanaka, N. et al. 2002. A single-base deletion in soybean flavonoid 3'-hydroxylase gene is associated with gray pubescence color. – *Plant Mol. Biol.* 50: 187–196.
- Tsai, S. M., Nodari, R. O., Moon, D. J. et al. 1998. QTL mapping for nodule number and common bacterial blight in *Phaseolus vulgaris* L. – *Plant Soil* 204: 135–145.
- Vest, G. 1970. R<sub>3</sub> a gene controlling ineffective nodulation in soybean. – *Crop Sci.* 10: 34–35.
- Vest, G. and Caldwell, B. E. 1972. R<sub>4</sub> a gene controlling ineffective nodulation in soybean. – *Crop Sci.* 12: 692–693.
- Vieira, C., Pasyukova, E. G., Zeng, Z. B. et al. 2000. Genotype-environment interaction for quantitative trait loci affecting life span, in *Drosophila melanogaster*. – *Genetics* 154: 213–222.
- Vuong, T. D., Nickell, C. D. and Harper, J. E. 1996. Genetic and allelism analyses of hypernodulation soybean mutants from two genetic backgrounds. – *Crop Sci.* 36: 1153–1158.
- Wang, S., Basten, C. J. and Zeng, Z. B. 2005. Windows QTL cartographer: ver. 2.5. – Dept of Statistics, N. Carolina State Univ., Raleigh, NC.
- Williams, L. F. and Lynch, D. L. 1954. Inheritance of a non-nodulating character in the soybean. – *Agron. J.* 46: 28–29.
- Zeng, Z. B. 1994. Precision mapping of quantitative trait loci. – *Genetics* 136: 1457–1468.
- Zhukov, V., Radutoiu, S., Madsen, L. H. et al. 2008. The Pea Sym37 receptor kinase gene controls infection-thread initiation and nodule development. – *Mol. Plant-Microbe Interact.* 21: 1600–1608.