Marker-assisted selection strategies for developing resistant soybean plants to cyst nematode

Fernanda Abreu Santana¹*, Martha Freire da Silva², Julierme Kellen Freitas Guimarães², Marcia Flores da Silva Ferreira³, Waldir Dias Pereira⁴, Newton Deniz Piovesan² and Everaldo Gonçalves de Barros⁵

Received 07 March 2013

Accepted 20 January 2014

Abstract – Resistant lines can be identified by marker-assisted selection (MAS), based on alleles of genetic markers linked to the resistance trait. This reduces the number of phenotypically evaluated lines, one of the limitations in the development of cultivars with resistance to soybean cyst nematode (SCN). This study evaluated the efficiency of microsatellites near quantitative trait loci (QTL) for SCN resistance, in the linkage groups (LG) G and A2 of soybean, for the selection of resistant genotypes in populations originated from crosses between the cultivars Vmax and CD201. The QTL of LG A2 was not detected in ‘Vmax’ (derived from PI 88788). In MAS, the microsatellites of LG G were efficient in selecting F₆:₇ families with resistance and moderate resistance to SCN race 3. The selection efficiency of the microsatellites Sat_168, Satt309 and Sat_141 was greater than 93%.

Key words: MAS, Glycine max, SCN, microsatellites, QTL.

INTRODUCTION

The soybean cyst nematode (SCN), *Heterodera glycines* Ichinohe, is worldwide the main pathogen of soybean (Wrather et al. 2001). The most efficient and economical control method is the use of resistant cultivars, together with rotation with non-host crops (Embrapa 2010). However, the development of resistant cultivars is limited by factors such as phenotypic analysis of segregating populations, which is time consuming, labor-intensive and requires much space in the greenhouse (Young and Mudge 2002, Cervigni et al. 2004, Concibido et al. 2004).

The development of 1,000 microsatellites (Simple Sequence Repeat) led to the construction of an integrated and saturated consensus map for soybean (Cregan et al. 1999a, Song et al. 2004). Thus, the markers near important QTL (Quantitative Trait Loci) can be used as anchors for locating regions in the linkage map in different populations (Schuster et al. 2001).


Marker-assisted selection (MAS) is an important tool to overcome difficulties of phenotypic selection, in the identification of SCN-resistant lines in segregating populations (Young and Mudge 2002, Concibido et al. 2004) and represents a useful alternative in the development of resistant cultivars.

This study evaluated the effectiveness of using microsatellite near the loci *rhg1* and *Rhg4*, for the selection of soybean lines resistant to SCN race 3.

MATERIAL AND METHODS

1. Universidade Estadual do Norte Fluminense Darcy Ribeiro, Laboratório de Melhoramento Genético Vegetal, 28.013-602, Campos dos Goytacazes, RJ, Brazil. *E-mail: fsantana.ufv@yahoo.com.br*

2. Universidade Federal de Viçosa, Instituto de Biotecnologia Aplicada à Agropecuária, 36.570-900, Viçosa, MG, Brazil

3. Universidade Federal do Espírito Santo, Alto Universitário, S/N, 29.500-000, Alegre, ES, Brazil

4. Embrapa Soja, C. P. 231, 86.001-970, Londrina, PR, Brazil

5. Universidade Católica de Brasília, 70.790-160, Brasília, DF, Brazil
Plant material

From crosses between isolines derived from the cultivars Vmax (resistant to SCN races 3 and 14) and CD 201 (SCN-susceptible), 65 F$_5$ soybean populations were obtained by the single pod descent (SPD) method. These populations were derived from the breeding program for soybean quality of the institute BIOAGRO at the Federal University of Viçosa (UFV) (Figure 1A).

Marker-assisted breeding strategy

The selected microsatellites were chosen for being in regions close to QTL for SCN resistance, i.e., in the region from 0.0 to 10.06 cM of the LG G on the consensus map, comprising the region of the SSR Satt163, Satt038, Satt275, Sat_168, Satt309, Sat_141, and Sat_163, as well as the region of LG A2, from 51.57 to 58.44 cM, with the SSR Sat_157, Sat_162, BLT 065, Satt187, GMENOD 2B (Song et al. 2004).

The polymorphic microsatellites between the parents “Vmax” and “CD201” were amplified in DNA seed bulks of each of the 65 F$_5$ populations and seven F$_5$ populations were selected on microsatellite alleles close to the resistance QTL. Of these seven, four populations were simultaneously selected on polymorphism of microsatellites of LG G and A2 and three on microsatellite polymorphism in LG A2 only. The selected populations were sown in bulks by the SPD method to obtain the F$_6$ generation. At harvest, the plants were threshed separately and 64 F$_{6:7}$ families were obtained (Figure 1B). These families were phenotyped for race 3 (HG Type 5.7) of SCN.

Phenotypic evaluation

The experiment was carried out in a greenhouse (25-30 °C and 16 hours of light) of Embrapa Soja, in Londrina, Paraná. The experiment was arranged in a completely randomized design with seven replications.

Seeds from 64 F$_{6:7}$ families were sown separately in plastic pots with sand. The same procedure was applied to the seven soybean lines (‘Peking’, PI 88788, PI 90763, PI 437654, PI 209332, PI 89772, and PI548316), to classify the HG types of the SCN populations, as proposed by Niblack et al. (2002), and to the susceptibility control Lee 74. Three days after germination, the seedlings were transplanted to 1 kg clay pots, containing a soil-sand mixture (1:3). At transplanting, each plant was inoculated with 4,000 SCN eggs of race 3 as described by Dias et al. (1998). The soybean plants were grown in a greenhouse for 28 days. Thereafter, leaves of each plant were collected for DNA extraction and recovery of the nematode females.

For female extraction, each plant was carefully removed from the pot and the root system was washed on sieves of 20 and 100 mesh under a strong water jet. After quantifying the females with a gridded acrylic plate and stereoscopic microscope, the female index (FI) was calculated for each F$_{6:7}$ family. The reaction of the F$_{6:7}$ families was classified by the criterion of Schmitt and Shannon (1992), i.e., families with FI<10% were considered resistant; 10% ≤ FI ≤ 30% moderately resistant; and FI ≥ 31% susceptible.

Genotypic analysis

For DNA extraction, leaves were collected from plants of different families prior to phenotypic analysis. The extraction followed the protocol of Doyle and Doyle (1990), modified by Abdelnoor et al. (1995). From each F$_{6:7}$ family, DNA bulks with seven plants were obtained for subsequent genotyping with the microsatellites Satt 309, Sat_141 and Sat_168 (Figure 1B).

Amplification reactions were performed in a final volume
of 15 µL containing 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2 mM MgCl₂, 100 µM of each deoxynucleotide, 0.3 µM of each primer, one unit of Taq polymerase, and 30 ng of DNA. The PCR program consisted of: 94 °C for 4 min, 30 cycles of 94 °C for 1 min, 55 °C for 1 min, 72 °C for 2 min, and subsequent 72 °C for 7 min. The amplification products were separated by electrophoresis in 10% native polyacrylamide gels using 1X TAE buffer (1 mM Tris-acetate 40 mM, EDTA) at 140 volts, subsequently stained with ethidium bromide (10 mg mL⁻¹) and photographed.

Statistical analysis

For analysis of variance and establishment of the genetic parameters, we used software Genes (Cruz 2013). The efficiency of selection (SE) of the microsatellite loci linked to SCN resistance was based on the comparison between the phenotypic and genotypic analyses and was calculated as described by Silva et al. (2007b):

\[ SE = 100 \times \frac{MFMF + mfmf}{(MM + mm)} \]

where:

- \( MFMF \) - number of families selected correctly as resistant, based on the marker and phenotypic analysis;
- \( mfmf \) - number of families selected correctly as susceptible, based on the marker and phenotypic analysis, and
- \( MM + mm \) - total of families selected as resistant and susceptible, based on markers only.

The SE was calculated using both the criterion of resistance, considering an index of parasitism (IP) of < 10 as well as moderate resistance, with IP < 30 (Schmitt and Shannon 1992).

RESULTS AND DISCUSSION

Three microsatellites of LG G (Satt309, Sat_168 and Sat_141) and one from LG A2 (Satt187) were polymorphic in the parents. These were used in the assisted selection performed in DNA seed bulks of each of the 65 F₅ populations. Based on the resulting polymorphism, seven segregating F₅ populations were selected by microsatellites, four of which of both LG and three of LG A2 only. The other populations were not polymorphic for the tested microsatellites.

Phenotypic evaluation of the selected populations

The 64 F₆:7 families derived from the seven selected F₅ populations were phenotyped for resistance to SCN race 3 (HG type 5.7). The lowest and highest mean numbers of females were found in population 1; transgressive segregation was observed in this population only, for both reduction and increase of the number of females, compared to the means of the parents. In the other populations, transgressive segregation occurred only for increase in the number of females (Table 1). This type of segregation was also observed for resistance to SCN race 14, which was attributed to possible effects of gene interaction in the control of resistance (Silva et al. 2007a).

All populations except 3 and 4 had a higher mean number of females than the susceptible parent (Table 1). The different resistance level of the families in each population indicated that the parents (isolines) derived from Vmax may have failed to recover all resistance genes, particularly minor-effect genes. Another possibility is that the population size did not allow the detection of a combination with all alleles. In this case, it would be necessary to study strategies that would allow the evaluation of a greater number of genotypes.

By the classification based on the index of parasitism of F₆:7 families of each population (Table 2) and resistant or moderately resistant and susceptible families were identified in the four F₅ populations, selected on SSR polymorphism of LG G and A2. However, the F₆:7 families derived from the three selected F₅ populations with LG A2 only were susceptible.

The results of the analysis of variance and estimates of genetic parameters of each F₅ population are shown in Table 3. Significant genetic variance of 1% was found among F₆:7 families, originating from the F₅ populations selected on the basis of SSR of LG G and A2 (Table 3). This allowed the

<table>
<thead>
<tr>
<th>Table 1. Mean, minimum and maximum numbers of females detected by the phenotypic analysis of F₆:7 families of the selected F₅ populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Populations selected on the SSR of LG G and A2</td>
</tr>
<tr>
<td>Population</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>
Marker-assisted selection strategies for developing resistant soybean plants to cyst nematode

selection for the best families in these populations. In the F_{6:7} families of the populations selected by the microsatellite of LG A2 (Satt 187), no genetic variability for resistance to race 3 was found (Table 3), indicating the absence of segregation of the resistance gene Rhg4 of LG A2 in these populations. Thus, LG A2 markers should not be used in marker-assisted selection in Vmax-derived populations (descendant from PI 88788). For this reason, the discussion below focuses only on populations with significant genetic variability.

The heritability estimates ranged from 75.9 to 95.7%. The high heritability detected may be a result of the high level of homozygosis of the studied families, the small number of major genes involved and of the environmental control in the experiment. According to Falconer and Mackay (1996), heritability estimates depend on the plant material, the estimation method and the experiment.

Webb et al. (1995) reported a broad-sense heritability of 97% for resistance to race 3 in crosses with PI 437654.

Therefore, the efficiency of phenotypic selection is high for this SCN race.

Genotypic evaluation of selected F_{6:7} families

The second genotypic evaluation assessed DNA bulks of seven plants from each F_{6:7} F_{3} families derived from the selected populations. These families were genotyped using only the LG G microsatellites, since the F_{6:7} families of the populations selected by marker Satt 187 of LG A2 were highly susceptible, indicating that this marker was not linked to SCN resistance in the studied populations.

Although the Satt187 marker of LG A2 is close to the resistance locus Rhg4, there is no evidence that this resistance allele is present in the resistant parent Vmax, which was derived from the resistance source PI 88788. Glover et al. (2004) studied PI 88788 and identified QTL in LG G and J by composite interval mapping at the 5% level of significance, but mentioned no QTL for SCN resistance in the region of LG A2. Conicibo et al. (1997) found no resistance locus

**Table 2. Number of resistant, moderately resistant, moderately susceptible and susceptible F_{6:7} families, according to the criterion of the index of parasitism (IP)**

<table>
<thead>
<tr>
<th>Phenotype of F_{6:7} families</th>
<th>F_{2} populations selected on the SSRs of LG G and A2</th>
<th>F_{2} populations selected on the SSR of LG A2 only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>R</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>MR</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>S</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Total F_{6:7} families*</td>
<td>12</td>
<td>10</td>
</tr>
</tbody>
</table>

* Number of plants evaluated in each F_{6:7} family, varying from 5 to 7; R - resistant (IP<10) MR - moderately resistant (10≤IP<30) S - susceptible-S (IP≥60).

**Table 3. Estimates of genetic parameters obtained from a phenotypic evaluation for race 3 (HG Type 5.7) in F_{6:7} families, derived from seven selected F_{2} populations, based on microsatellite markers of LG G and A2**

<table>
<thead>
<tr>
<th>F_{2} populations selected on the LG G and A2</th>
<th>F_{2} populations selected on LG A2 only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>GMS</td>
<td>50004**</td>
</tr>
<tr>
<td>RMS</td>
<td>8196.30</td>
</tr>
<tr>
<td>(\hat{\sigma}^2_g)</td>
<td>6198.15</td>
</tr>
<tr>
<td>(\hat{\sigma}^2_e)</td>
<td>7413.27</td>
</tr>
<tr>
<td>(\hat{\sigma}^2_{g:e})</td>
<td>1215.13</td>
</tr>
<tr>
<td>(h^2) (%)</td>
<td>83.61</td>
</tr>
<tr>
<td>(CV^g) (%)</td>
<td>32.83</td>
</tr>
<tr>
<td>(CV^r) (%)</td>
<td>37.74</td>
</tr>
<tr>
<td>(CV^g/CV^r)</td>
<td>0.87</td>
</tr>
<tr>
<td>MNR</td>
<td>6.75</td>
</tr>
</tbody>
</table>

**No. of families | 12 | 10 | 11 | 7 | 9 | 8**

**significant at 1% probability ns: non-significant GMS and RMS; GMS: genotype mean squares and residual mean squares, respectively; \(\hat{\sigma}^2_g\): estimate of genetic variance; \(\hat{\sigma}^2_e\): estimate of phenotypic variance; \(\hat{\sigma}^2_{g:e}\): estimate of environmental variance; \(h^2\) (%): broad-sense heritability, in percentage; \(CV^g\): coefficient of genetic variation, in percentage; \(CV^r\): coefficient of experimental variation, in percentage; MNR: mean number of replications (plants); \(CV^g/CV^r\): variation index.
for race 3 in LG A2 either in the same resistance source. Therefore, the locus for SCN race-3 resistance in LG A2 is probably absent in PI 88788.

However, LG A2 is important to control resistance to SCN race 3 in different resistance sources, such as PI 209332 (Concibido et al. 1994), Peking (Mahalingam and Skorupska 1995, Chang et al. 1997, Meksem et al. 2001), PI 437654 (Webb et al. 1995, Prabhu et al. 1999), PI 90763 (Guo et al. 2005) and Hartwig (Silva et al. 2007b).

**Genotype - phenotype analysis of F₆:7 families**

The four phenotypically resistant F₆:7 families carried resistance alleles of the three microsatellite markers LG G (Satt309, Sat_168 and Sat_141), indicating that these markers are extremely efficient in the selection of resistant families.

Of the four moderately resistant F₆:7 families, three carried resistance alleles of the LG G microsatellites, indicating that these markers did not distinguish resistant from moderately resistant families. Within these families, plants with resistance, moderate resistance and moderate susceptibility were observed. This demonstrates that in addition to the resistance QTL of LG G, other resistance (minor effect) QTL are needed for complete resistance to SCN race 3 in Vmax-derived populations. The QTL of LG G is known to explain much of the resistance to SCN race 3 in different resistance sources. Complete resistance however requires minor-effect genes that are not always identified (Concibido et al. 2004).

It was found that five F₆:7 families derived from the four genotyped populations were classified as susceptible in spite of carrying resistance alleles of the evaluated microsatellites, indicating the existence of segregating resistance genes in these families, since plants with moderate resistance were found in the families.

**Selection efficiency**

The selection efficiency was evaluated considering all F₆:7 families of the four populations together. With the SSR Sat_168, a selection efficiency of 97.05% was reached and 96.55 and 93.55%, with Satt309 and Sat_141, respectively. The pairwise combinations of the microsatellites raised the SE to 100%, in all cases (Table 4).

The selection efficiency increased little when the criterion of moderate resistance (IP < 30) was used for calculations (Table 4). In all cases, the discrimination of resistant families did not raise the efficiency, unlike the correct classification of the susceptible ones.

These data confirmed that these markers are very close to locus rhg1, which has been extensively studied for explaining a large phenotypic variation of the SCN resistance to race 3 and other races (Concibido et al. 2004), and proved their usefulness in the selection of the assessed populations and that they can be used in other Vmax-derived populations.

Cregan et al. (1999b) reported that the SSR Satt309 and Sat_168 are at a distance of 0.4 cM from locus rhg1 and the use of one of these markers would ensure success in genotypic selection. High selection efficiency was also reported by Mudge et al. (1997) in the identification of race-3 resistant lines with SSR flanking only the region of gene rhg1, with an accuracy of 98%. Silva et al. (2007b) also reported high selection efficiency (94%) for race-3 resistance in F₂:3 families with only one SSR of LG G: Satt309.

However, the high SE values reported in the literature were obtained in mapping populations, where the QTL was detected in the study population. This particular study deals with a breeding population in advanced generations; nevertheless, it was possible to obtain high selection efficiency with the SSR of LG G, near locus rhg1.

**ACKNOWLEDGEMENTS**

The authors wish to thank the National Council for Scientific and Technological Development (CNPq) and the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES) for funding this study.

### Table 4.

<table>
<thead>
<tr>
<th>Marker</th>
<th>SE% (IP &lt; 10)</th>
<th>SE% (IP &lt; 30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sat_168</td>
<td>97.05</td>
<td>97.29</td>
</tr>
<tr>
<td>Satt309</td>
<td>96.55</td>
<td>96.87</td>
</tr>
<tr>
<td>Sat_141</td>
<td>93.55</td>
<td>94.12</td>
</tr>
<tr>
<td>Sat_168 + Sat_141</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Satt309 + Sat_141</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Sat_168 + Satt309</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
Estratégias de seleção assistida por marcadores para desenvolvimento de plantas de soja resistentes ao nematoide de cisto

Resumo – A seleção assistida por marcadores (SAM) permite identificar linhagens resistentes com base em alelos de marcadores genéticos ligados ao caráter, o que reduz o número de linhagens avaliadas fenotipicamente, uma das limitações ao desenvolvimento de cultivares resistentes ao nematoide de cisto da soja (NCS). Neste trabalho objetivou-se avaliar a eficiência de microsatélites próximos a QTLs de resistência ao NCS, nos grupos de ligação (GL) G e A2 da soja, na seleção de genótipos resistentes em populações originadas do cruzamento entre as cultivares Vmax e CD201. O QTL do GL A2 não foi detectado em ‘Vmax’ (derivada da PI 88788). A SAM por microsatélites do GL G foi eficiente na seleção de famílias F₅₇ resistentes e moderadamente resistentes à raça 3 do NCS. Os microsatélites Sat_168, Satt309 e Sat_141 apresentaram eficiência de seleção maior que 93%.

Palavras-chave: SAM, Glycine max, NCS, microsatélites, QTLs.

REFERENCES


Schuster I, Abdelnoor RV, Marin SRR, Carvalho VP, Kiihl RAS, Silva JFV,


