Journal of Bioinformatics and Sequence Analysis Vol. 1(4), pp. 056-060, December, 2009
Available online at http://www.academicjournals.org/jbsa

Full Length Research Paper

Sequencing and identification of homologous region encoding rust resistant-gene in soybean (*Glycine max* L.)

Nguyen Thi Lang[1]*, Nguyen Cao Quan Binh[1] and Bui Chi Buu[2]

[1]Genetic and Plant Breeding Department, CuuLong Delta Rice Research Institute, Iai, CanTho City, Vietnam.
[2]Plant Breeding of Southern Agricultural Science Institute (IAS), 121, Nguyen Binh Khiem. TP Ho Chi Minh, Asia.

Accepted 29 October, 2009

Soybean (*Glycine max* L.) is one of the most important crops that is both easily cultivated and of high efficiency. Its product is commonly used as fresh-seed directly or produced into plant oil, cake, candy and milk. It can also be processed as feed for domestic animals for enhancing the protein requirement in daily meals, and it can be exported for providing valuable currency resource. On the other hand, it is also used in the process of land improvement, contributing to increase yield. Soybean is not only useful in medicine for treatment of human disease, but it is also utilized to provide materials for the processing of animal feed with rich-protein resource and also for industrial application. Based on the result of phenotypic evaluation in rust disease, 81 varieties of soybean were classified into three groups: group A consists of 37 varieties with resistance level of 0 and 1; group B has 30 varieties with resistance level of 3 and 5; and group C has 14 varieties with resistance level of level 7 and 9. We have designed two primers for targeting coding region of rust resistant-gene, represented by Langrisat1 and Langrisat2. We have shown that the soybean varieties in group A have high rust resistance capability, which can provide further application in breeding programming. A key task for the future is to determine whether rust resistant-gene can be transferred to susceptible soybean varieties. Two designed primers are needed for synthesis and testing through Polymerase Chain Reaction (PCR) upon request. To the best of our knowledge, the results presented here are the first characterization of nucleotide of rust resistant-gene in soybean. Further analysis is necessary for identifying nucleotide sequence and its functional site accurately.

Key words: Rust resistant gene, soybean, *Glycine max* L., sequencing.

INTRODUCTION

Due to its significant economic value and well characterized morphology, *Glycine max* L. is the most important crop among rice and maize. Due to the high content of protein, soybean seeds and its products are used widely to feed human and cattle worldwide. Soybean seeds contain high nutritional components: content of protein is approximately 38 - 40%, lipid ranges from 18 - 20%, with rich vitamin and mineral salt. It is the only type of plant whose value is evaluated both in terms of protein and lipid. Soybean is able to fix nitrogen through symbiotic bacteria that invade the roots of leguminous plants and form root nodules in which they fix nitrogen, supplying both the bacteria and the plants. Since each nodule is considered as “tiny nitrogen fixation-manufactory”, soybean cultivation not only reduces used nitrogen, but also contributes significantly to the improvement and reinforcement of the land.

Some studies conducted allelic tests including cross linking between a group of rust resistant genotypes from Embrapa's soybean germplasm collection and PI 230970 and PI 459025, which carry the *Rpp2* and *Rpp4* genes, respectively (Larissa et al., 2008). However, its commercial value was not appreciated correctly although it can be considered as resources for high foreign earnings and also as a potential material for human in the future.

There are over 100 diseases that destroy soybean in the world, with thirty-five of them affecting its economic value. In Vietnam, according to the statistic of 1990, there

*Corresponding author. E-mail: ntlang@hcm.vnn.vn.

were about 30 soybean rust diseases which are caused by fungi. It is considered that the rust, caused by *Phakopsora pachyrhizi* H and P Syd., was one of the most serious diseases of soybean because it has a widely dispersal area and causes considerable reduction of the crop yield. Thus, plant breeders attempt to find much new methods so that they can select promising soybean varieties that are resistant to rust disease, constantly. However, because of the natural pressure of selection, herbicide and pesticide, they have broken environmental balance, leading to unstable rust resistance. Although during the last decades, few progresses have been made on the studies of the elucidation of genes involved in rust resistant genes. Blad and Baker (1972) indicated that Komata variety (PI0492) contained a dominant allele resistant to rust strain in Australia. Singh and Thypliyal (1977) reported that Ankur variety (denoted by PI 462312) also has a dominant allele resistant to rust disease. Some soybean diversities by SSR and RAPD markers are used for rust resistant-gene (Lang et. al., 2008). The aim of this study is to identify nucleotide sequence and design primer for rust resistant gene which can serve for further research.

MATERIALS AND METHODS

Plant material

The experiment was conducted with nine soybean varieties from CuuLong Delta Rice Research Institute, Genebank, which consists of NamVang and OMDN87 as two checked varieties. Seventy-two soybean varieties were obtained from the Institute of Science and Agriculture in Vietnam in Southern Asia for screening rust disease. Crossing technique in soybean was done by CLDRRI protocol (Nguyen, 2002). Seedling leaf material was harvested after 30 days and stored in a fresh 1.5 ml tube at 4°C.

Evaluation of phenotype

P. phakyrizhi fungus from Cuulong Delta Rice Research Institute (CLRRI), Vietnam, Genebank was cultured in M1 medium with a total volume of 1 L including 10 g glucose, 1 g K_2HPO_4, 5 g peptone, 0.5 $MgSO_4.7H_2O$, 20 g agar and adjust sterile distilled water of 1 L. Fungal colonies were transferred from a master plate to two pre-prepared M1 medium plate and incubated at room temperature for 36 - 48 h. The experimental soybean varieties were designed by randomly completely block design (RCBD) with three replication; two leaves of each plant variety per replication were infected. After 8 - 10 days, evaluation of affected level was recorded following by IRRI's standard protocol. Performance of rust disease on leaves was grouped using NTSYS pc software version.

DNA isolation

Genomic DNA was isolated from seedling leaves, using CTAB method. The protocol DNA suitable for PCR analysis was prepared using a simplified procedure of IRRI (1997), which was modified by Lang (2002). A piece of young rice leaf (2 cm) was collected and placed in a labeled 1.5 ml centrifuge tube in ice. The leaf was ground using a polished glass rod in a well of a Spot Test Plate (Thomas Scientific) after adding 400 µl of extraction buffer (1 L containing 20 g CTAB, 81.82 g NaCl, 100 ml 1M Tris pH 8.0, 40 ml 0.5 M EDTA pH 8.0 and sterile distilled water). Grinding was done until the buffer turned green, an indication of cell breakage and release of chloroplasts and cell contents. Another 400 µl of extraction buffer was added into the well through pipette, and about 400 µl of the lysate was transferred to the original tube of the leaf sample. The lysate was deproteinized using 400 µl chloroform. The aqueous supernatant was transferred to a new 1.5 ml tube and DNA precipitated using absolute ethanol. DNA was air- dried and re-suspended in 50 µl of TE buffer.

DNA quality checks used 1% agarose by melting 3 g agarose in 300 ml TAE buffer (Appendix Table 1). The mixture was heated in microwave for 5 - 6 min and then cooled to around 55 - 60°C. This was then poured on prepared electrophoresis box with combs. Gels were ready and combs removed after about 45 min. Seven microliliters of DNA sample plus 3 µl loading buffer (Tris 1M pH = 8.0, glycerol, EDTA 0.5M pH = 8.0, xylene cyanol 0.2%, bromphenol blue 0.2% and distilled water) was run at 70-80v, 60mA for 45 min or until loading buffer dye moved far from the wells. Gel was then taken out and stained with ethidium bromide after which it was visualized under UV light.

PCR reaction for sequencing

The PCR reaction was conducted with a final volume of 20 µl using GenAmp®PCR system 9700 that contains 4 µl of terminator ready reaction mixture (2.5X for 384-well plates), 2 µl of Bigdye ® terminator V1.1/3.1 sequencing buffer 5X, 1 µl of -21M13 forward primer, 1 µl of M13 reverse primer, 1µl DNA template and deionized water. The reaction conditions for PCR included a denaturing step of 96°C for 1 min which was followed by 25 cycles of 10 s at 96°C, 5 sec at 50°C and 4 min at 60°C, and ended with a step of sample storage at 4°C. Rapid thermal ramp rose to 4°C and was held until ready for purification. After agarose gel (1%) electrophoresis, PCR products were purified by using QIAquick gel Extraction Kit (QIAGen, USA).

Sequencing and sequence analysis

PCR products, pre-denatured with hidi-formamide at 95°C for 5 min, were sequenced by using a capillary automated Applied Biosystem 3130 DNA sequencer until the whole fragments were sequenced completely following standard sequencing process by the manufacturer. Search similarity between nucleotide and deduced amino acid sequences was done by BLAST NCBI and EMBL/SwissProt available online (www.expasy.org). A multi-sequence alignment was performed using ClustalW (www.ebi.ac.uk). Prediction of coding regions and gene structure was done by ORF finder and CLC combined workbench version 3.0.3, respectively. Signal peptide was predicted by SiganlP 3.0 server (www .cbs.dtu.dk/services). Scanprosite was used to analyze functional site of putative amino acid sequence (www.expasy.org/Prosite). Primers were designed by Primer3-BLAST available online.

PCR amplification

PCR amplification components and conditions were done based on the methods used by Lang (2002). The PCR reaction mixture contained 20-50 nanogram (ng) template DNA, 50 ng of each primers, 0.05 mM dNTP's, 1xPCR buffer (10 mM Tris pH 8.4, 50mM KCl, 1.8 mM $MgCl_2$ and 0.01 mg/ml gelatin) and 1 unit of *Taq* DNA polymerase in a total volume of 20 µl. Template DNA was initially denatured at 94°C for 5 min followed by 30 cycles of PCR amplification, using the following parameters: 30 s denaturation of 94°C, 30 s primer annealing at 55°C or 60°C and 1 min primer ex-

tension at 72°C. Completion of primer extension was allowed by a final 5 min incubation at 72°C.

An aliquot of 10 µl of the PCR product was routinely taken for gel electrophoresis to determine if amplification was successful. When the primers detected an amplicon length polymorphism, the samples were readily scored. The remaining 10 µl of PCR products was used for restriction digestion to detect PBR polymorphism in the case of other primers. The digestion reaction normally contained 3.2 µl sterile distilled water, 1.5 µl restriction buffer (10X), 0.3 µl restriction enzyme (10 U/µl) and 10.0 µl of PCR products in a total volume of 15 µl. The digestion reaction was incubated for 4 h overnight at appropriate incubation temperature for the enzyme used.

The PCR products or the DNA fragments produced by restriction digestion were resolved electrophoretically on 1% agarose gel in 1 X TAE buffer.

RESULTS

Classifying based on phenotype

With the level of genetic variability of 2.92, soybean varieties were classified into three major groups: A, B and C. Group A has 37 varieties, out which 27 varieties such as PI0830881, LS201, AGS374, GC84058, OMDN1, BR23, HL203, 9800410, SSE137559, BR24, MTD483-4, PI200429, LEIRCHART, MTD164-1, OMDN111, DT85, ALIANT, ATF15, MTD652-5, QUANGPHU, MTD652-2, AGS360, ATS16, AGS371, DT94, PI085089, PI548484 belong to level 1, while the remaining 10 varieties such as GLS2111, CPAC365-76, OMDN29, DT2000, AGS376, DH4, GC990013-12-15-10, IAC100, 96033B, AGS365 belong to level 0; group B has 30 varieties in which there were 14 varieties for level 5 such as DT200, L07515, OMDN109, ATF8, 13176, PRANA, DT93, 9907A-4, MSBR22, 9005A-7, GC90013-21-23, MTD664, MANTA, MTD517-8, and 16 varieties such as 903551CR, 9603331-1-1-1, TL57, AGS367, HL2, GC990013-1-1-39, GC90013-12-15-6, 9804512, G85-5126, OMDN64, MSBR17, 96033B, OMDN87, PI417088, HL92, OMDN130, Dau trang DT for level 3. Finally, group C has 14 varieties in which 7 varieties such as 980464, PI518759, MTD176, Dau den DT, MSPR20, OMDN110, MTD514-6 belong to level 9, while the remaining 7 varieties such as 95389, 95389-1, AGS129, 5113, Nam Vang, HQ1, PI103 belong to level 7.

Sequence analysis of PCR products

PCR reaction successful amplified inserted DNA sequences with a band pattern of gel that is approximately 110 bp in size. This will be convenient for subsequent sequencing.

Through analysis of ORF finder, we have identified coding region in the nucleotide sequence of OMDN1 and OMDN110 with 2 exon which includes exon 1 (114 nucleotide which starts in 85 position and ends in 198 position), exon 2 (156 nucleotide from 281-436 position) and 3 exon which consists of exon 1 (216 nucleotide from 1-215 position), exon 2 (168 nucleotide from 48-215 position), exon 3 (210 nucleotide from 5-214 position), respectively. The nucleotide sequence of OMDN1 encoded a protein with pI/Mw (9.99 / 15339.14) and a protein with pI/Mw (5.23 / 6954.88) for OMDN110 nucleotide sequence. Using BLAST search similarity, the nucleotide sequence of OMDN1 and OMDN110 was identical with nucleotide sequence of rust resistant gene in wheat with accession number S79982, S79983 in Genebank database. Two nucleotide sequences of OMDN1 and OMDN110 were also submitted to Genebank Database with accession number EU077601.2 and EU077601.1, respectively. The result of multi-sequence alignment using Clustal W showed that there was a relatively low similarity between sequences OMDN1, OMDN110, S79982 and S79983 (Figure 1).

Primer design

With two sequenced nucleotide sequences, we have designed two primers, denoted by Langrisat1: forward primer5'GGCCCCGAGTGATTTAAGGA3', reverse primer 5'GGTGGTGACTCACGCTGTGG 3' and Langrisat2: forward primer 5'GTGGTGTAGCTGAATCTGGA 3', reverse primer 5' GAGTGGTGTGAGGTGATGGT 3' for amplifying coding region of rust resistant gene. The GC ratio account for 50 - 55% for forward primer and 55 - 65% for reverse primer. These results were relatively suitable for the theory of primer design following Lang's standard protocol.

Analysic of the F2 population

DNA genome from the F2 population, one of the cross between OMDN 29 and MTD 176 was assessed through PCR amplification using primer Langrisat L-R. The resultant PCR products were spliced out by double enzymes digestion with HinfI and AluI. Fragments were resolved on agarose gels and the banding patterns were scored with reference to those of the parents. The banding pattern of the F2's individuals could be classified into homozygote for the OMDN 29 type marker 250 bp fragment, homozygote for MTD 176 type marker 210 bp fragment and heterozygotes displaying both fragments OMDN 29 and MTD 176 (Figure 2).

DISCUSSION

With phenotypic evaluation of 81 soybean varieties, tolerant level of the varieties were grouped, which allowed the screening of soybean varieties with stable resistance to rust disease. These results open a new opportunities to utilize them as materials in plant breeding programming upon request. We assumed that morphology markers are very useful for screening directly based on phenotype. However, the plant phenotype is an interac-

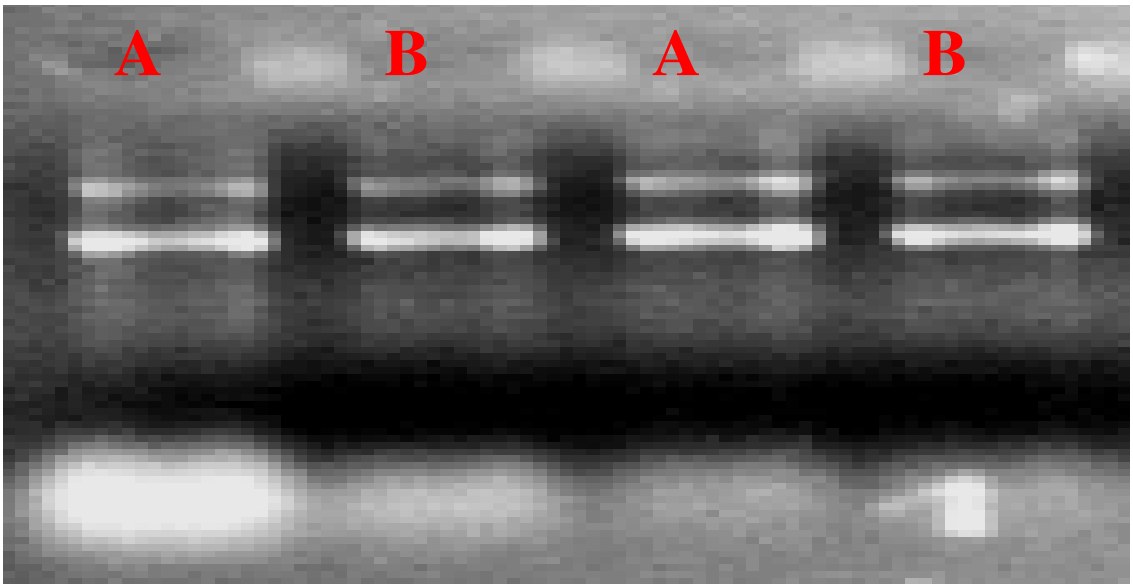

Figure 1. PCR products of OMDN1 and OMDN110 with a band pattern around 1100bp. A: OMDN1, B: OMDN110.

```
OMDN110         AGGTTAAAGGGATAGTCGAGTTGGTTTCTGAT--AGATGATAGTAAGATTAAT-----AG  54
OMDN1           AGATAAAAAAAGTGTGTGAGTTATTTGGAGGTTTATTTATTCATAAGGTAAAT-----CA 143
S79982          TGGCAAAGAGGCCTCATCCATCGCCCTGGACGGCTCGTTTTGGTCAAATCAGTGATTGCG 599
S79983          TGGCAAAGAGGCCTCATCCATCGCCCTGGACGGCTCGTTTTGGTCAAATCAGTGATTGCG 567
                 *   **              *                 *  *  * *  * * *

OMDN110         AGTGA---TGTGAATGATGG---GGTGAAAGATGTG---GTGTAGCTGAATCTGGAAAT- 104
OMDN1           GGTGA---CTTGCACAGTTGATGGGAGGAGAAGAAA---GTGGATCTGGTTGGGGGAAT- 196
S79982          GCTAAACCCATCCATCATTTCATGGTGACACATGCTCCCGTGTGGGTATTTGAAGAGATC 659
S79983          GCTAAACCCATCCATCATTTCATGGTGACACATGCTCCCGTGTGGGTATTTGAAGAGATC 627
                  * *     *  *   *     ** *    *       ***    *   *   *  **

OMDN110         -----------GAA-TAGGGATGATGGATAGAAAGTGGATAAGTTGAGTGGTGTGAGGTG 152
OMDN1           -----------AAAATTTGGTCGGTGGGGGCAAAACAAAAAGAAAAAAAAATT---GTTA 242
S79982          ----------TGCGATCCTTCTTTTGGGCTGGCAAGGAAAAAATCCAACGGTGGCCAGTG 709
S79983          GAGCAGTGGATGGGATCCTTCTTTTGGGCTGGCAAGGAAAAG-TCCAACGGTGGCCAGTG 686
                          *        ***      *    * *     *    *      *

OMDN110         ATGGTTGAAATTGAGTTTGGTTGT--GCGTTGGTCGGGGTAAAG--TGGTGGGGTGAGAG 208
OMDN1           GGGGTTAAAATCCGTTCTCGCCCCCGGCCCCGAGTGATTTAAGGATAGGTTTGTTGGGGG 302
S79982          CTTGGTAACTTGGAAATCAGTCTGCAAACCCACTTCGCTTGGTGGTCTGGGAATCCGCAA 769
S79983          CTTGGTAACTTGGAAATCAGTCTGCAAACCCACTTCGCTTGGTGGTCTGGGAATCCGCAA 746
```

Figure 2. The comparison of nucleotide sequence of OMDN1 and OMDN110 with others in Genebank database. Accession number of nucleotide sequence of rust resistant-gene in wheat is followed as S79982, S79983. The column conservation represents the level of homologous between sequences. Bold lines in background show homologous of two nucleotide sequence for rust resistant gene in wheat.

interaction between genotype and environment. Thus, a practical comprehension is to provide distinct approaches which evaluate on genotype to isolate stable resistance with soybean rust disease correctly.

The nucleotide sequence analysis of two soybean varieties (OMDN1 and OMDN110) indicated that they were relatively low compared to others submitted by authors in Genebank database. In addition, nucleotide data for soybean rust disease are still very restricted. This leads to difficulties in unraveling and identifying rust resistant nucleotide sequence. To the best of our knowledge, the results presented here are the first characterization of nucleotide of rust resistant-gene in soybean. For accuracy, further analysis is necessary for identifying nucleo-

tide sequence and its functional site. Two primers represented by Langrisat1 and Langrisat2 were designed to amplify coding region of target gene. However, only Langrisat1 was detected with F2 population from OMDN 1/ MTD 176.

The accuracy of marker selection for rust resistance was verified through F2, progeny tests. The accuracy of predicting homozygous resistant genotypes based on flanking marker data was 85% for rust using a single marker. A single marker could thus be as accurate as two flanking markers provided.

The successful PCR based marker for rust has made it possible to use these markers in future MAS program for the transfer of rust into elite breeding lines of soybean.

Conclusion

We have shown that the soybean varieties in group A have high rust resistance capability, which can provide further application in breeding programming. But at present, very little studies have been carried out on soybean rust. We proposed that further analysis and characterization of rust resistant gene will help to identify nucleotide sequence and its functional sites. A key task for future research is to determine whether rust resistant-gene can be transferred to susceptible soybean varieties. An additional task is to determine the efficiency of breeding methods that is very important for improving rust resistance in soybean varieties.

ACKNOWLEDGEMENTS

The authors gratefully acknowledged all colleagues for their help. Acknowledgement is also made to Genetic and Plant breeding Department Laboratory (Southern of Agriculture Institute, VietNam and CuuLong Delta Rice Research Institute) for their support to this research and all who had read and approved this final manuscript.

REFERENCES

Blad, Baker (1972) Soybean. Narsa publishing House: Replika Press Pvt. Ltd

Larissa Di-CL, Carlos AAA, Aliny SR, Breno FR, Pedro HBP, José FF; Antonio EP, Geraldo EC (2008) New genes conferring resistance to Asian soybean rust: allelic testing for the Rpp2 and Rpp4 loci . Pesquisa Agropecuária Brasileira. Print version ISSN 0100-204X

Lang NT (2002). Basic methods in studying biotechnology. Agricultural publisher in Ho Chi Minh City.

Singh BB, Thapliyal PN (1977). Breeding for resistance to soybean rust in India. p. 62-65. *In* R.E. Ford and J.B. Sinclair (ed.) Rust of soybeans, the problem and research needs. Int. Agric. Publ. INTSOY Ser. 12.Univ. of Illinois, Urbana, IL, USA.