Subtractive library of soybean roots in response to inoculation with *Bradyrhizobium japonicum*

Biblioteca subtrativa de raízes de soja em resposta à inoculação com *Bradyrhizobium japonicum*

Gesiele Almeida Barros de Carvalho[1,2], Jesiane Stefânia Silva Batista[1],
Francismar Corrêa Marcelino[1], Mariangela Hungria[11*]

ABSTRACT

Soybean [*Glycine max* (L.) Merr.] is the most important legume cropped worldwide, with an economic value attributed mainly to the high protein content of the grain, which, in turn, demands a heavy supply of nitrogen. An environmentally friendly source of nitrogen at low cost to farmers is represented by the biological fixation of atmospheric nitrogen that takes place is association with symbiotic bacteria. For the soybean, the association occurs mainly with bacteria belonging to the species *Bradyrhizobium japonicum* and *Bradyrhizobium elkanii*, but our knowledge about gene expression associated with the symbiosis is still poor. This work—part of an effort by the Soybean Genome Consortium (Genosoja)—was conducted with the aim of achieving better understanding about the functionality of the genes expressed in soybean roots in the early stages of the interaction with *B. japonicum*. The study was performed with soybean plants growing under greenhouse conditions, inoculated (or not) with strain CPAC 15 (=SEMIA 5079) of *B. japonicum*, which is broadly used in Brazilian commercial inoculants. Total RNA was extracted, the mRNA was isolated and a subtractive library was constructed. Sequencing analysis generated 4,621,072 reads, which were assembled in 3,776 sequences, defined as the differentially expressed sequences of the inoculated *versus* non-inoculated treatments. The sequences were categorized according to their molecular function and grouped with representatives of antioxidant activities, molecular transduction, transporters and enzyme regulation activities, but with an emphasis on catalytic and molecular binding/signaling.

Key words: Biological nitrogen fixation, differential gene expression, genome, *Glycine max.*

[1] *Embrapa Soja, Londrina, Paraná, Brazil.*
[2] *Department of Biochemistry and Biotechnology, Universidade Estadual de Londrina, Londrina, Paraná, Brazil.*
[*] *Author for correspondence: E-mail: hungria@cnpso.embrapa.br; hungria@pq.cnpq.br phone: (+55) 43-33716206*

RESUMO

A soja [*Glycine max* (L.) Merr.] é uma leguminosa de grande expressão mundial, com valor econômico atribuído, principalmente, ao teor proteico elevado dos grãos, que, por sua vez, demandam altas doses de nitrogênio. Uma fonte de nitrogênio ambientalmente favorável e de baixo custo para os agricultores é representada pela fixação biológica do nitrogênio atmosférico, que ocorre através da associação com bactérias simbióticas. Na soja, essa associação ocorre, principalmente, com bactérias das espécies *Bradyrhizobium japonicum* e *Bradyrhizobium elkanii*, mas o conhecimento sobre a expressão gênica associada com a simbiose ainda é escasso. Este trabalho – parte integrante do Consórcio do Genoma da Soja (Genosoja) – foi conduzido com o objetivo de obter um melhor entendimento sobre a funcionalidade dos genes expressos nas raízes de soja nos estágios iniciais da interação com *B. japonicum*. O estudo foi conduzido com plantas de soja cultivadas sob condições de casa de vegetação, inoculadas (ou não) com a estirpe CPAC 15 (=SEMIA 5079) de *B. japonicum*, utilizada em larga escala em inoculantes comerciais no Brasil. O RNA foi extraídos das raízes, o mRNA foi isolado e a biblioteca subtrativa construída. O Sequenciamento gerou 4.621.072 leituras que, após a montagem resultaram em 3.776 sequências, definidas como diferencialmente expressas dos tratamentos inoculado *versus* não inoculado. As sequências foram agrupadas de acordo com sua função molecular, que resultou nas categorias de atividade antioxidante, transdução molecular, transporte, regulação de atividades enzimáticas, mas com ênfase nas atividades catalíticas e de ligação/ sinalização molecular.

Palavras-chave: Fixação biológica de nitrogênio, expressão diferencial de genes, genoma, *Glycine max*.

INTRODUCTION

Soybean [*Glycine max* (L.) Merr.] is the most important legume cropped worldwide, with an economic value attributed mainly to the high protein content of the grain. Due to the high demand of nitrogen (N) needed to the synthesis of proteins and other biomolecules, this nutrient usually represents a limiting factor for soybean productivity (Panzieri *et al.,* 2000). An important feature of this legume is the ability to supply most of the crop´s N requirement by the establishment of a symbiotic partnership with bacteria—belonging mainly to the species *Bradyrhizobium japonicum* and *Bradyrhizobium elkanii*—that converts atmospheric N_2 within root nodules into forms usable by the plant (Graham and Vance, 2003).

The establishment of a successful symbiosis involves an intensive molecular dialogue between the host plant and the microsymbiont. Molecular signals released by the roots, mainly flavonoids, result in the transcription of nodulation genes in compatible rhizobia, leading to the formation of lipochitooligosaccharide molecules (Nod factors) that, in turn, act as signals to the host plant, initiating nodule organogenesis (Long, 1996). Numerous changes in gene expression occur both in the host plant and in the bacterium during root infection, nodule development and onset of N_2 fixation (Vance, 2002).

Inoculation of soybean seeds with selected strains *of B. japonicum* and *B. elkanii* represents an environmental friendly technology that can fulfill most of the crop's need for N. Also, inoculation is highly economically advantageous in comparison with applying N fertilizers (Hungria *et al.,* 2006).

Advanced genetic analyses are providing improved comprehension of the complex process of the symbiosis, from the first steps of molecular signal exchange between the host plant and the microsymbiont to the full development of active nodules fixing nitrogen (Stacey *et al.*, 2006). To improve our understanding of nodule formation, we used subtractive library technology, described by Diatchenko *et al.* (1996), to identify genes differentially expressed in soybean roots in symbiosis with compatible *B. japonicum* strain CPAC 15 (=SEMIA 5079), used in commercial inoculants in Brazil, to get a better understanding of the functionality of the genes expressed in response to the bacterium.

MATERIAL AND METHODS

Plant material

Soybean seeds of cultivar MG/BR46-Conquista were carefully selected for homogeneity and surface sterilized with a solution of 70% ethanol (one minute), followed by hypochlorite solution (three minutes), and then rinsed five times with sterile distilled water (Vincent, 1970). The seeds were placed between two layers of germination paper, moistened with sterile distilled water and incubated at 22 ± 2°C (in the dark) for three days. Healthy-looking germinated seeds were selected and transferred to sterile plastic bags, each containing 200 mL of N-free nutrient solution (Broughton and Dilworth, 1970).

Treatments

Treatments consisted of: soybean inoculated with *Bradyrhizobium japonicum* strain CPAC 15 (=SEMIA 5079) and non-inoculated soybean (mock-inoculated with sterile distilled water). The experimental design was completely randomized, with three replicates, each of twenty plants.

Inoculum preparation and inoculation

The microorganism used for inoculation was *B. japonicum* strain CPAC 15, which has been used in commercial inoculants in Brazil since 1992 for its outstanding symbiotic efficiency and competitiveness (Hungria *et al*., 2006). The bacterium was cultured in yeast-mannitol broth (YMB) until the exponential phase of growth (Vincent, 1970), when the cells were centrifuged and washed with saline solution (0.85% NaCl). Subsequent counting of a washed cell suspension in YMB revealed a concentration of 2.27×10^7 cells mL^{-1}.

For the inoculation treatment, 1 mL aliquots of washed bacterial were used to inoculate the basis of the soybean radicles. Plants were grown under greenhouse conditions, with a 12-h day/night period and mean temperature of 25-38°C/15-18°C (day/night) for ten days. Subsequently, the roots were separated from shoots, immediately frozen in liquid N and stored at –80°C pending RNA extraction.

RNA extraction and isolation of mRNA

For the subtractive library construction, roots of both, inoculated (treatment receiving inoculum) and non-inoculated (treatment without nodules, control) were macerated in liquid N and frozen at -80°C until total RNA extraction by using Trizol® Reagent (Invitrogen), at a rate of 1 mL 100 mg^{-1} tissue, at room temperature

(25±2°C). Assessment of RNA integrity and quantification were verified in a Nanodrop 1000 spectrophotometer (Agilent). The mRNA was obtained from 2 µg of total RNA using the FastTrack MAG mRNA Isolation Kit (Invitrogen), according to the manufacturer's specifications.

Construction of suppressive subtractive hybridization library

The suppressive subtractive hybridization (SSH) library was obtained by using the PCR-Select cDNA Subtraction Kit (Clontech), according to the manufacturer's instructions. A cDNA library was built, containing clones from the subtraction of inoculated *versus* non-inoculated soybean plants. This subtractive hybridization was performed using a sample of cDNAs (*tester*) from inoculated plants, which was subtracted with cDNAs (*driver*) from non-inoculated plants.

The construction of the subtractive library consists of: *(i)* synthesis from the first and second strands of cDNA; *(ii)* digestion of the cDNAs with the restriction enzyme *Rsa*I; *(iii)* binding of adapters to the cDNAs; *(iv)* first and second hybridization reactions; *(v)* amplification of the first product hybridization by PCR (*Nested* PCR); and (*vi*) the second amplification by PCR. This last step allowed an enrichment of differentially expressed cDNA sequences. PCR products resulting from the library were transferred for *in vitro* cloning and subjected to sequence analysis.

Sequencing and analysis of sequences

Sequencing analysis was performed in a Genome Analyzer II (Illumina), at the Fasteris S.A., Switzerland. Generated sequences were assembled and analyzed at the Laboratory of Genomic and Expression-UNICAMP, Campinas, Brazil. The construction of the soybean database, to serve as a reference genome, was conducted using the database of ESTs (expressed sequence tags) from the National Center of Biotechnology Information (NCBI) and the data of the soybean genome were provided by the DOE Joint Genome Institute (*Phytozome*). The SOAP2—Short Oligonucleotide Alignment Program (Li *et al*., 2009)—was used to anchor all reads in the reference genome with parameters that allowed alignment with, at most, two mismatches between the sequences. Finally, the *de novo* assembly was conducted with the EDENA software—Exact De Novo Assembler (Hernandez *et al*., 2008)—using its default parameters. Both SOAP2 and EDENA also verify the quality of the sequences. The assembled contigs were aligned to the reference genome. The library was also validated by the comparison of the differential genes in the control and in the inoculated treatment. Only the genes over expressed or absent in the control were confirmed in the subtraction library.

Bioinformatics tools used to analyze the data

Due to the large number of reads, the AutoFACT program (Koski *et al*., 2005) was used for automatic annotation. Subsequently, the resulting sequences were deposited in the database library of GenoSoja, which is available at http://bioinfo03.ibi.unicamp.br/soja. Several public databases were used to support the analysis, e.g. NCBI (http://www.ncbi.nlm.nih.gov/), Gene Ontology (http://www.geneontology.org/), KEGG (Kanehisa and Goto, 2000) and Phytozome (http://www.phytozome.net/soybean).

RESULTS AND DISCUSSION

Methodology chosen for the study

Using intact, high-quality mRNA represents a key step for the successful application of modern molecular biological methods, especially in cDNA-library constructions and quantitative RT-PCR assays (Fleige and Pfaffl, 2006). Instability and degradation of mRNA have been attributed mainly to the action of RNAses, introduced during handling (Fleige and Pfaffl, 2006). In general, the best RNA yields are obtained from tissues that have been diced into small fragments prior to being frozen by submerging in liquid N, and then homogenized by using a bead mill or a mechanical homogenizer (Bustin and Nolan, 2004). Using this procedure, we obtained adequate quantities of good quality RNA. For RNA extraction, we chose Trizol, which is a monophasic solution of phenol and guanidine isothiocyanate that allows extraction of RNA in a single step.

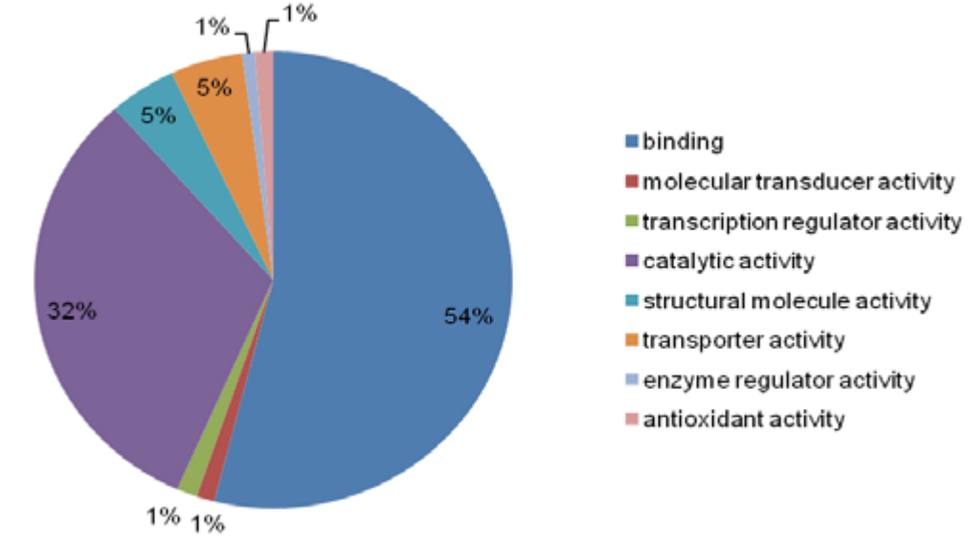
Suppression subtractive hybridization (SSH) was chosen to investigate the molecular changes in soybean roots caused by inoculation with *B. japonicum*. It's an effective approach to analyze gene expression. The technique is often used in higher eukaryotes to study molecular-regulation mechanisms, improving our insights about the molecular mechanisms underlying biological processes (Huang *et al.,* 2007).

Functional distribution of genes of the SSH library of inoculated X non-inoculated soybean roots

Sequencing of the SSH library (inoculated *versus* non-inoculated plants) in the Genome Analyzer II platform, with thirty-six cycles, generated 4,621,072 reads. The assemblage generated 3,776 differentially expressed sequences. Following, Blast2GO—with numerous functions for visualization, management, and statistical analysis (Conesa and Götz, 2008)—was chosen for functional annotation and data mining, with grouping based primarily on gene ontology (GO).

Sequences differentially expressed in soybean roots in response to inoculation were annotated according to their molecular function (level 3) and are displayed in Figure 1.

Figure 1. Functional categorization of the transcripts obtained in the analysis of a suppressive subtractive hybridization library built with soybean roots inoculated versus non-inoculated with Bradyrhizobium japonicum strain CPAC 15 (=SEMIA 5079). Percentage of sequences (%) classified according to their molecular functions.



The major category of differentially expressed genes (54%) fit into the binding function (GO:0005488). As the establishment of the symbiosis requires close coordination in the exchange of molecular signals between the host plant and rhizobia, genes that perform this function may be involved in signaling (Ferguson et al. 2010). Within this category, the interaction of molecules can occur with one or more specific sites. The symbiosis starts with the release of phenolic compounds by the host plant, particularly flavonoids, that induce the transcription of nodulation genes in the bacterium (Hungria, 1994). The rhizobium responds with the production and secretion of other signaling molecules, the Nod factors (lipochitooligosaccharides), that can partially determine host specificity (Hungria, 1994). Nod factors are recognized by the host plant by binding to specific receptors, initiating a signal transduction cascade (Stougaard, 2000; Ferguson *et al.,* 2010). This downstream response involves the interaction/binding of biomolecules, in order to transmit the required signals for starting root deformation, allowing the invasion of bacteria and the formation of infection threads, necessary for nodule establishment (Limpens and Bisseling, 2003; Oldroyd and Downie, 2004). Directly related to signaling are also the genes in the category of molecular transducer activity (GO:0060089), representing 1% of the genes differentially expressed. In addition, about 1% of the genes are in the category of regulators of transcription (GO:0001071). It is well known that a successful nodulation requires a balance between induction and repression of genes (Ferguson and Mathesius 2003); therefore, regulators include key genes for the symbiosis.

One category that showed a large number of genes (32%) is related to catalytic activity (GO:0003824). These genes may act on specific substrates in response to the bacterium, such binding of the bacterium to the plant cell (Ferguson *et al*., 2010), and modifications of the plant cell wall (Reiter, 2002). For nodule organogenesis, the signaling cascade triggers several changes in the cells, including root hair deformation, development of infection threads and nodule development (Schmidt and Panstruga, 2007). Therefore, genes with catalytic activity are directly involved with root development. In addition, 5% of the genes were related to structural molecule activity (GO:0005198), a category that includes, among others, genes acting in the organization of the cytoskeleton of the roots; due to the contact with rhizobia, the roots undergo a rearrangement of cells, critical for nodule organogenesis (Blancaflor *et al*., 2006).

This study also highlights genes that play a role in transport activity (GO:0005215)—associated with all kinds of transport of molecules in a cell—that represented about 5% of the genes; they are responsible for the movement of biomolecules within the cells or transmembranes. This functional class also participaes in the rapid response of the plant to the presence of rhizobia (Kaldenhoff and Fischer, 2006). Enzyme regulator activity (GO:0030234) is a

functional class that includes genes that modulate the activity of enzymes involved in the infection process, and accounted for 1% of the genes.

Finally, antioxidant activity (GO:0016209) included 1% of the genes, possibly related to differentiation of tissues and/or to the production of signal molecules involved in auto-regulation of nodulation by the host plant (Bueno *et al.*, 2001). The presence of microorganisms results in large amounts of ROS (reactive oxygen species), which are described as secondary messengers in many signaling processes in plants (Cárdenas *et al.,* 2008; Marino *et al.*, 2009); therefore, genes with antioxidant activity help in the protection of cells against oxidative damage by regulating the levels of ROS within the cells (Marino *et al.*, 2009).

CONCLUSION

High throughput analysis of differential gene expression represents a powerful tool helping to obtain additional information about genomic responses in a given condition (Holtorf *et al.,* 2002). In this preliminary analysis, we may say that the approach of sequencing and analysis of a suppressive subtractive hybridization library of inoculated versus non-inoculated soybean roots has allowed the identification of groups of genes differentially expressed in response to *B. japonicum*. Key genes were identified that will be now studied with more details.

ACKNOWLEDGMENTS

Authors thank Aline Luiza Schiavon and Dr. Elisete Pains Rodrigues for technical and scientific advice. The work was partially supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil) Project Genosoja (577933/2008) and CNPq-Repensa (562008/2010-1). Fellowships from CNPq are acknowledge by G.A.B. Carvalho (M.Sc.), J.S.S. Batista (pos-doc) and M. Hungria (research).

REFERENCES

BLANCAFLOR, Elison B.; WANG, Yuh-Shuh; MOTES, Christy M. **Organization and function of the actin cytoskeleton in developing root cells.** International Review of Cytology, v. 252, p. 219-264, 2006.

BROUGHTON, William J.; DILWORTH, Michael J. Methods in legume-*rhizobium* technology: plant nutrient solutions. In: SOMASEGARAN, Padmanabhan; HOBEN, Heinz J. (ed). **Handbook for Rhizobia**. Hawaii: NifTAL Project and University of Hawaii, 1970. p. 245–249.

BUENO, Pablo; SOTO, María José; RODRIGUEZ-ROSALES, María Pilar; SANJUAN, Juan; OLIVARES, José; DONAIRE, Juan Pedro. Time-course of lipoxygenase, antioxidant enzyme activities and H_2O_2 accumulation during the early stages of *Rhizobium*-legume symbiosis. **New Phytologist**, v. 152, p. 91–96, 2001.

BUSTIN, Stephen A.; NOLAN, Tania. Template handling, preparation, and quantification. In: BUSTIN, Stephen A. (ed) **The Real-Time PCR Encyclopedia A–Z of Quantitative PCR.** International University Line, La Jolla, CA, 2004. p. 87–120.

CÁRDENAS, Luis; MARTÍNEZ, Adán; SÁNCHEZ, Federico; QUINTO, Carmen. Fast, transient and specific intracellular ROS changes in living root hair cells responding to Nod factors (NFs). **The Plant Journal**, v. 56, p. 802-813, 2008.

CONESA, Ana; GÖTZ, Stefan. Blast2GO: a comprehensive suite for functional analysis in plant genomics. **International Journal of Plant Genomics**, 2008:619832, 2008.

DIATCHENKO, Luda; LAU, Yun-Fai Chris; CAMPBELL, Aaron P.; CHENCHIK, Alex; MOQADAM, Fauzia; HUANG, Betty; LUKYANOV, Sergey; LUKYANOV, Konstantin; GURSKAYA, Nadya; SVERDLOV, Eugene D.; SIEBERT, Paul D. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. **Proceedings of the National Academy of Sciences of the United States of America**, v. 93, p. 6025-6030, 1996.

FERGUSON, Brett James; INDRASUMUNAR, Arief; HAYASHI, Satomi; LIN Meng-Han; LIN, Yu-Hsiang; REID, Dugald E.; GRESSHOFF, Peter M. Molecular analysis of legume nodule development and autoregulation. **Journal of Integrative Plant Biology**, v. 52, p. 61–76, 2010.

FERGUSON, Brett James; MATHESIUS, Ulrike. Signaling interactions during nodule development. **Journal of Plant Growth Regulation**, v. 22, p. 47–72, 2003.

FLEIGE, Simone; PFAFFL, Michael W. RNA integrity and the effect on the real-time qRT-PCR performance. **Molecular Aspects of Medicine**, v. 27, p. 126–139, 2006.

GRAHAM, Peter H.; VANCE, Carroll P. Legumes: importance and constraints to greater use. **Plant Physiology**, v. 131, p. 872–877, 2003.

HERNANDEZ, David; FRANÇOIS, Patrice; FARINELLI, Laurent; ØSTERÅS, Magne; SCHRENZEL, Jacques. De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. **Genome Research**, v. 18, p. 802-809, 2008.

HOLTORF, Hauke; GUITTON, Marie-Christine; RESKI, Ralf. Plant functional genomics. **Naturwissenschaften**, v. 89, p. 235–249, 2002.

HUANG, Xiaowei; LI, Yunxia; NIU, Qiuhong; ZHANG, Keqin. Suppression subtractive hybridization (SSH) and its modifications in microbiological research. **Applied Microbiologyand Biotechnology**, v. 76, p. 753–760, 2007.

HUNGRIA, Mariangela. Sinais moleculares envolvidos na nodulação das leguminosas por rizóbio. **Revista Brasileira de Ciência do Solo**, v. 18, p. 339-364, 1994.

HUNGRIA, Mariangela; CAMPO, Rubens José; MENDES, Iêda Carvalho; GRAHAM, Peter 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, Rana P.; SHANKAR, Niharika; JAIWAL, Pawan K. (ed) **Nitrogen nutrition in plant productivity.** Houston: Studium Press, 2006. p. 43-93.

KALDENHOFF, Ralf; FISCHER, Matthias. Functional aquaporin diversity in plants. **Biochimica et Biophysica Acta**, v. 1758, p. 1134–1141, 2006.

KANEHISA, Minoru; GOTO, Susumu. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Research**, v. 28, 27-30, 2000.

KOSKI, Liisa B.; GRAY, Michael W.; LANG, B. Franz; BURGER, Gertraud. AutoFACT: an automatic functional annotation and classification tool. **BMC Bioinformatics**, v. 6, p. 151-162, 2005.

LI, Ruiqiang; YU, Chang; LI, Yingrui; LAM, Tak-Wah; YIU, Siu-Ming; KRISTIANSEN, Karsten; WANG, Jun. SOAP2: an improved ultrafast tool for short read alignment. **Bioinformatics**, v. 25, p. 1966-1967, 2009.

LIMPENS, Erik; BISSELING, Ton. Signaling in symbiosis. **Current Opinion in Plant Biology**, v. 6, p. 343–350, 2003.

LONG, Sharon R. *Rhizobium* symbiosis: nod factors in perspective. **The Plant Cell**, v. 8, p. 1885–1898, 1996.

MARINO, Daniel; PUCCIARIELLO, Chiara; PUPPO, Alain; FRENDO, Pierre. The redox state, a referee of the legume–rhizobia symbiotic game. **Advances in Botanical Research**, v. 52, p. 116-139, 2009.

NCBI (National Center for Biotechnology Information). Disponível em: < http://www.ncbi.nlm.nih.gov/> Acesso em: 20 out. 2010.

OLDROYD, Giles E. D.; DOWNIE, J. Allan. Calcium, kinases and nodulation signalling in legumes. **Nature Reviews Molecular Cell Biology**, v. 5, p. 566–576, 2004.

PANZIERI, Margherita, MARCHETTINI, Nadia; HALLAM, Thomas G. Importance of the *Bradyrhizobium japonicum* symbiosis for the sustainability of a soybean cultivation. **Ecological Modelling**, v. 135, p. 301–310, 2000.

REITER, Wolf-Dieter. Biosynthesis and properties of the plant cell wall. **Current Opinion in Plant Biology**, v. 5, p. 536–542, 2002.

SCHMIDT, Sarah Maria; PANSTRUGA, Ralph. Cytoskeleton functions in plant–microbe interactions. **Physiological and Molecular Plant Pathology,** v. 71, p. 135–148, 2007.

STACEY, Gary; LIBAULT, Marc; BRECHENMACHER, Laurent; WAN, Jinrong, MAY, Gregory D. Genetics and functional genomics of legume nodulation. **Current Opinion in Plant Biology,** v. 9, p. 110–121, 2006.

STOUGAARD, Jens. Regulators and regulation of legumes root nodule development. **Plant Physiology**, v. 124, p. 531-540, 2000.

VANCE, Carroll P. Root-bacteria interactions: symbiotic nitrogen fixation. In: WAISEL, Yoav; ESHEL, Amram; KAFKATI, Uzi. (ed) **Plant Roots.** The Hidden Half. 3 rd edition. New York: Marcel Dekker Publishers, 2002. p. 839–867.

VINCENT, James Matthew. ***A manual for the practical study of root nodule bacteria***. Oxford: Blackwell, 1970. 164 p.