Arch Microbiol (2009) 191:113–122
DOI 10.1007/s00203-008-0432-0

ORIGINAL PAPER

Genetic differences between *Bradyrhizobium japonicum* variant strains contrasting in N_2-fixation efficiency revealed by representational difference analysis

Fernando Gomes Barcellos · Jesiane Stefânia da Silva Batista · Pâmela Menna · Mariangela Hungria

Received: 28 February 2008 / Revised: 1 July 2008 / Accepted: 5 September 2008 / Published online: 15 October 2008
© Springer-Verlag 2008

Abstract Two variant strains of *Bradyrhizobium japonicum*, derived from SEMIA 566, adapted to the stressful environmental conditions of the Brazilian Cerrados and characterized by contrasting capacities for N_2 fixation, were compared by representational difference analysis (RDA). Twenty-four gene sequences that are unique to the highly effective strain S 370 were identified, eight showing high similarity to known genes, nine encoding putative proteins and seven representing conserved hypothetical or hypothetical proteins; they were classified in eight functional categories. Among those genes, some were highlighted for their known or potential functions in plant–microbe interactions. The nodulation outer protein P (*nopP*), related to the type-III secretion system (TTSS) and a major determinant of nodulation of some tropical legumes, was detected in the genome of strain S 370. Three coding sequences (CDS) identified by RDA were expressed in proteomics experiments with *B. japonicum* strain USDA 110 (ChvE and NopP). The use of the sequences identified by RDA in the highly effective strain S 370 might represent an important tool to speed up strain selection programs, accelerating pre-screening procedures. Additionally, the conserved hypothetical and hypothetical proteins identified in strain S 370 might encode important but still unknown proteins related to the symbiosis that deserve further study.

Keywords *Bradyrhizobium* · Bacterial genome · Nitrogen fixation · Nodulation · Representational difference analysis

J. S. S. Batista and P. Menna contributed equally to the manuscript.

Communicated by Ursula Priefer.

F. G. Barcellos · J. S. S. Batista · P. Menna · M. Hungria (✉)
Laboratório de Biotecnologia dos Solos, Embrapa Soja,
Cx. Postal 231, Londrina 86001-970, PR, Brazil
e-mail: hungria@cnpso.embrapa.br; hungria@pq.cnpq.br

F. G. Barcellos
e-mail: fgbarcel@yahoo.com.br

F. G. Barcellos · J. S. S. Batista · P. Menna
Departamento de Microbiologia,
Universidade Estadual de Londrina,
Cx. Postal 60001, Londrina 86051-990, PR, Brazil

Introduction

Soybean (*Glycine max*) is particularly important in Brazil, where the crop occupies about half of all cropped land, with the Cerrados region—an edaphic type of tropical savannah—representing the major producing area. The legume is capable of establishing a symbiotic partnership and the process of biological N_2 fixation with strains of *Bradyrhizobium japonicum* and *B. elkanii*. However, Brazilian soils were originally void of effective soybean bradyrhizobia; therefore, exotic strains have been introduced as inoculants since the late 1950s in the south and from the early 1970s in the Cerrados. The use of selected strains in inoculants has resulted in established populations estimated at 10^3–10^6 per g of soil in the great majority of the 22 million ha cropped today (Hungria and Vargas 2000; Hungria et al. 2005, 2006a, b).

Since the late 1990s, our group has characterized several rhizobial strains obtained from nodules of field-grown soybean in areas of the Cerrados that had been massively inoculated with exotic *B. japonicum* and *B. elkanii* strains. Intriguingly, many isolates differ in morphological, serological, physiological, genetic and symbiotic properties when compared to the putative parental strains (Barcellos et al. 2007; Batista et al. 2007; Boddey and Hungria 1997;

Galli-Terasawa et al. 2003; Hungria and Vargas 2000; Hungria et al. 1996, 1998; Mendes et al. 2004; Nishi et al. 1996; Santos et al. 1999). Those differences may be caused by stress-induced mutagenesis (Tenaillon et al. 2004; Wright 2004) related to the often stressful environmental conditions of the Cerrados (Hungria and Vargas 2000); furthermore, horizontal gene transfer, leading to adaptive evolution, has been broadly observed there (Barcellos et al. 2007; Batista et al. 2007; Galli-Terasawa et al. 2003).

In addition to understanding the ecology of soybean bradyrhizobia, a main goal of our group is to search, within the naturalized population, for strains adapted to the tropics and characterized by higher efficiency of N_2 fixation and competitiveness. This approach has been very successful, such that the commercial strains available today are capable of completely fulfilling the plant's demand on N (Hungria and Vargas 2000; Hungria et al. 2005, 2006a, b). However, if the current soybean demand is for about 300 kg of N ha^{-1}, estimates are that more productive cultivars, yielding up to 8,000 kg ha^{-1} will need up to 1,000 kg of N ha^{-1} (Hungria et al. 2006b). The identification of *Bradyrhizobium* genes related to higher capacity for N_2 fixation is thus crucial to speed up strain selection programs, and the genetically stable natural variants showing a range of capacities for N_2 fixation, available in our laboratory (Boddey and Hungria 1997; Hungria et al. 1998; Santos et al. 1999) represent potentially useful material.

Subtractive DNA hybridization methodologies have been broadly used to isolate unique sequences within genomes (Ermolaeva and Sverdlov 1996; Sagerström et al. 1997). Representational difference analysis (RDA) was originally described in 1993 by Lisitsyn et al. (1993) and with the incorporation of the PCR technique with subtractive DNA hybridization allows the selective amplification of unique targets. Using this approach, successful results have been described for several bacterial species (Allen et al. 2003), including the identification of novel DNA sequences from natural strains of the diazotrophic bacterium *Sinorhizobium* (= *Ensifer*) *meliloti* (Guo et al. 2005).

Bradyrhizobium japonicum strains S 370 and S 516 are putative variants of SEMIA 566 (Barcellos et al. 2007; Boddey and Hungria 1997; Hungria et al. 1998), a strain used in Brazilian commercial inoculants from 1966 to 1978 that belongs to the same serogroup as strain USDA 123. Several morpho-physiological, genetic and symbiotic characteristics of Brazilian isolates belonging to this serogroup have been described by our group (Barcellos et al. 2007; Batista et al. 2007; Boddey and Hungria 1997; Galli-Terasawa et al. 2003; Hungria et al. 1996, 1998; Mendes et al. 2004; Nishi et al. 1996; Santos et al. 1999). As S 370 and S 516 show contrasting N_2-fixation capacity (Hungria et al. 1998), we compared them using RDA, with the objective of finding genes putatively related to symbiotic effectiveness.

Methods

Rhizobial strains

Bradyrhizobium japonicum strains S 370 and S 516 have been described before (Barcellos et al. 2007; Boddey and Hungria 1997; Hungria et al. 1998) and are deposited at the Diazotrophic and Plant Growth Promoting Bacteria Culture Collection of Embrapa Soja. Information about the strains is also available at http://www.bmrc.lncc.br.

Genomic DNA extraction

For the total genomic DNA extraction, bacterial batch cultures were grown in 100 mL of yeast extract–mannitol (YM) broth until late exponential phase (10^9 cells mL^{-1}). Each bacterial culture was then centrifuged at 5,000 rpm (1,677*g*) for 10 min, resuspended in 50 mL of TES, centrifuged again at 5,000 rpm for 10 min and resuspended in 10 mL of TES. One-hundred microliters of lysozyme (200 μg mL^{-1}) and 30 μL of RNAse (1 mg mL^{-1}) were added to the resuspended cells and the mixture was incubated at 30°C for 4 h. Following that, 10 μL of proteinase K (20 μg μL^{-1}) were added and the mixture was incubated overnight at 37°C. An equal volume (10 mL) of Tris-equilibrated phenol, pH 8.0, was added, the tube gently inverted for 30 min and the mixture centrifuged at 10,000 rpm (6,708*g*) for 15 min at room temperature (22 ± 2.5°C). The supernatant was removed, an equal volume of phenol was added, and the mixture gently homogenized for 20 min. After that, 5 mL of chloroform was added and the mixture homogenized by gentle inversion for 10 min and centrifuged at 10,000 rpm for 15 min at room temperature. The upper aqueous phase was transferred to a new tube and an equal volume of isopropanol was added and homogenized by gentle inversion for 5 min, followed by centrifugation at 5,000 rpm for 15 min. The isopropanol was removed, 1 mL of 70% ethanol was added to the pellet and centrifuged at 5,000 rpm for 5 min. The ethanol was then removed and the DNA dried at room temperature for 20 min. The DNA was then resuspended in 50 μL of TE and incubated overnight at 4°C. The quality and quantity of the DNA were verified in 1% agarose gels using appropriate DNA markers.

Representational difference analysis (RDA)

The RDA method (Allen et al. 2003) was performed using two rounds before cloning; with strain S 370 used as "tester" and S 516 as "driver." The oligonucleotide adapters were based on those described by Lisitsyn et al. (1993): R*Hind*24 (5′ AGCACTCTCCAGCCTCTCACCGCA 3′) and R*Hind*12 (5′ AGCTTGCGGTGA 3′) in the first RDA

round and J*Hin*d24 (5′ ACCGACGTCGACTATCCA TGAACA 3′) and J*Hin*d12 (5′ AGCTTGTTCATG 3′) in the second RDA round. For the preparation of the "tester" genomic DNA (S 370), 3 μg of DNA was digested with *Hin*dIII in three reactions, each containing 1 μg of DNA and 20 U of *Hin*dIII (Sambrook et al. 1989). The digestion products were purified with the QIAquick PCR purification kit (Qiagen Inc., Valencia, CA). After checking the DNA in 1% agarose gel, the fragments were dephosphorylated with 300 U of bacterial alkaline phosphatase (BAP) (Invitrogen, Carlsbad, CA) for a final volume of 40 μL. The reaction was maintained at 65°C for 90 min and then purified with the QIAquick purification kit. After dephosphorylation, the DNA tester was mixed with 0.5 nmol of each adapter (R*Hin*d24 and R*Hin*d12) in a final volume of 20 μL and the mixture held at 55°C for hybridization and then maintained at 4°C for 1 h. Following that, 10 U of T4 DNA ligase were added, the mixture was incubated at 14°C for 16–20 h, and then purified with the QIAquick purification kit. For the preparation of the "driver" DNA (S 516), 50 μL of DNA (1 μg μL^{-1}) were randomly sheared by nebulization (6 s, 40 psi; Invitrogen nebulizer) using the manufacturer's protocol to generate fragments ranging from 1 to 10 kb and the fragments were verified in 1% agarose gels.

For the DNA subtraction hybridization, 0.1 μg of the "tester" DNA (ligated to the R*Hin*d24 and R*Hin*d12 adapters) was mixed with 20 μg of the "driver" DNA (fragments of 1–10 kb). The mixture was precipitated with 100% ethanol and 3 M NaOAc (3/1, v/v), resuspended in 6 μL of 3X EE buffer [1XEE: 10 mM *N*-(2-hydroxyethyl)-piperazine-*N*′-(3-propanesulfonic acid)/1 mM EDTA pH 8.0], denatured at 100°C for 2 min, completed with NaCl to a final concentration of 1 M and then maintained at 67°C for 20 h. The mixture was then diluted tenfold with pre-heated 3× EE containing 1 M NaCl and immediately placed on ice. Five to ten PCR reactions were performed to obtain approximately 3 μg of the amplification product. In each reaction, 1 μL of the diluted hybridization mixture was added to 21.5 μL of the PCR reaction mix (10 mM Tris-HCL pH 8.3, 3.5 mM $MgCl_2$, 25 mM KCl, 0.2 mM of each dNTP, 2.5 U *Taq* polymerase) and incubated at 70°C for 3 min to fill the extremities of the re-annealed fragments. The PCR mix was heated to 94°C for 5 min, the R*Hin*d24 adapter (1 pmol μL^{-1} of the PCR mix) was added and submitted to 40 PCR cycles (1 min at 94°C; 1 min at 62°C and 1 min at 72°C; a final extension cycle of 10 min at 72°C). Previous tests indicated no need for the mung-bean nuclease treatment suggested by Allen et al. (2003) to avoid background amplification of the single-stranded "driver" DNA. The PCR reactions were purified (QIAquick purification kit) and the products were quantified in 1% agarose gels. The purified products of the first RDA round were digested with *Hin*dIII to detach the R*Hin*d24 and R*Hin*d12 adapters that were then removed by purification (QIAquick purification kit). The PCR products were then dephosphorylated with BAP, as described above, and the products (1 μg) were ligated to 0.5 nmol of J*Hin*d24 and of J*Hin*d12 adapters, as described above. For the second RDA round, 50 ng of the "tester" DNA (products of the first RDA round linked to the J*Hin*d24 and J*Hin*d12 adapters) were mixed with 20 μg of sheared "driver" DNA and the hybridization was performed as described for the first RDA round. After that, 1 or 2 μL of the hybridization mixture were amplified by PCR with primer J*Hin*d24 (five to ten reactions), as described for the first RDA round; the products were submitted to electrophoresis in a 3% agarose gel and the fragments obtained were excised with the help of the QIAquick gel extraction kit (Qiagen). The fragments obtained were cloned into the TOPO TA vector (Invitrogen), using the manufacturer's protocol.

DNA sequencing and BLAST analysis

We have obtained about 380 different clones in each of the five libraries built; therefore, summing 1,900 clones. All cloned RDA fragments were sequenced following the TOPO TA cloning kit (Invitrogen) protocol, using the DYEnamic ET terminator chemistry (GE Healthcare, Piscataway, NJ) and analyzed in a MegaBace 1000 DNA Analysis System (GE Healthcare) following the parameters described by Menna et al. (2006). The high-quality sequences obtained in both 3′ and 5′ directions were assembled using the programs phred (Ewing and Green 1998; Ewing et al. 1998), phrap (http://www.phrap.org) and Consed (Gordon et al. 1998), and were submitted to the GenBank (http://www.ncbi.nlm.nih.gov/blast) and Rhizobase databases (http://bacteria.kazusa.or.jp/rhizobase/) seeking significant alignments. The sequences were always verified in both GenBank and Rhizobase databases. The highest matches were almost always with the genome of *B. japonicum* strain USDA 110 genome or, in few cases, with the genomes of *Bradyrhizobium* strains ORS278 and BTAi1, but then they have also shown high similarity with homologous sequences of the genome of USDA 110. Therefore, the sequences in our study were identified and functionally classified based on the similarities with the genome of *B. japonicum* strain USDA 110.

Nucleotide sequence acquisition numbers

GenBank accession numbers EI710909–EI710932 were given to the RDA nucleotide sequences determined in this study (Table 1).

Table 1 Similarity of RDA sequences identified in *B. japonicum* strain S 370, highly effective in N_2 fixation, when compared to the genome of USDA 110

GenBank #	Length[a] (bp)	Identity[b] (%)	Length[c] (bp)	Descriptive information about matching sequences from the genome of *B. japonicum* strain USDA 110			
				ORF number	Gene name	Protein length (aa)	Putative function
Transport and binding proteins							
EI710909	278	93	132	blr4511		339	ABC transporter substrate-binding
EI710915	216	93	187	blr4449	*aapP*	245	L-amino-acid ABC transporter ATP-binding
EI710916	367	95	214	bll7921		597	ABC transporter substrate-binding
EI710917	345	99	314	blr1895	*malK*	370	Maltose/maltodextrin import ATP-binding protein MalK
EI710910	290	90	200	bll5893		262	ABC transporter ATP-binding
EI710918	323	94	256	blr3208	*chvE*	357	Multiple sugar-binding periplasmic receptor ChvE precursor
Transposon-related functions							
EI710919	285	100	176	bll2138		414	Probable site-specific integrase/recombinase
EI710920	355	100	284	bll1898		503	Putative transposase
Central intermediary metabolism							
EI710911	320	88	198	bll6723		568	Putative acid phosphatase
Fatty acid, phospholipids and sterol metabolism							
EI710921	344	96	239	bll7739	*pldB*	315	Lysophospholipase L2 (lecithinase B)
EI710922	372	89	264	bll0338	*AMACR*	388	Alpha-methylacyl-CoA racemase (AMACR)
Amino acid biosynthesis-branched chain family							
EI710923	420	95	327	bll4763	*ilvD*	574	Probable dihydroxy-acid dehydratase
Regulatory functions							
EI710912	401	91	89	bll4775		126	Two-component response regulator
EI710924	371	87	94	bll3106		1,024	Two-component sensor histidine kinase
DNA replication, recombination, and repair							
EI710913	392	82	160	bll8051	*uvrA*	847	Putative excinuclease ABC subunit A (UvrA protein)
EI710925	300	79	245	blr0145		353	Putative SNase-like nuclease precursor
Symbiosis							
EI710926	424	100	392	blr1752	*nopP*	244	Host-inducible protein A homolog-nodulation outer protein P
Conserved hypothetical							
EI710927	311	95	251	bll5459		409	
EI710928	350	85	249	blr3812		492	
EI710930	375	91	307	blr3822		495	
EI710929	333	89	130	blr1302		702	
EI710931	420	87	306	blr6271		316	

Table 1 continued

GenBank #	Length[a] (bp)	Identity[b] (%)	Length[c] (bp)	Descriptive information about matching sequences from the genome of *B. japonicum* strain USDA 110			
				ORF number	Gene name	Protein length (aa)	Putative function
Hypothetical							
EI710914	405	91	325	blr4829		313	
EI710932	438	94	380	blr2368		450	

[a] Length in bp of the sequenced DNA fragment

[b] Identity (%) of the best match found in *B. japonicum* USDA 110 genome

[c] Length of the sequence in bp that showed similarity with the sequences deposited in the database

GenBank accession number

Results

Identification and functional classification of the sequences

The second RDA round resulted in six products ranging from 230 to 430 bp, with the last fragment showing higher concentration. The six bands obtained were extracted from the gel and cloned separately. Five out of the six cloned bands gave high quality DNA sequence, and only the minor band (of about 230 pb) resulted in DNA sequences of poor quality. As stated in the Sect. "Methods", the highest DNA sequence matches were almost always with the genome of *B. japonicum* strain USDA 110 genome or, in few cases, with the genomes of *Bradyrhizobium* strains ORS278 and BTAi1, but then they have also shown high similarity with homologous sequences in the genome of USDA 110. Furthermore, all clones analyzed had significant hits to only one defined ORF, with no matches to more than one gene or boundaries between two genes. The sequences obtained for all five products were compared with those deposited in the GenBank and Rhizobase databases and classified according to the highest similarities with the genome of *B. japonicum* USDA 110 (Table 1). The nucleotide similarities ranged from 79 to 100% for fragments ranging from 89 to 392 bp and the functional categories identified were: (1) transport and binding (mainly ABC transporters—substrate and ATP-binding proteins); (2) transposon-related functions; (3) central intermediary metabolism; (4) fatty acid, phospholipid and sterol metabolism; (5) amino acid biosynthesis–branched chain family; (6) regulatory functions; (7) DNA replication, recombination, and repair; and (8) symbiosis. In addition, seven sequences showed no similarity with known proteins and were classified as conserved hypothetical or hypothetical (Table 1). In this paper, we will consider only the proteins with known or hypothetical function on plant-molecular interaction.

Transport and binding: *aapP*, *malK* and *chvE* genes

Six putative genes were classified in the transport and binding functional category (Table 1): two putative ABC transporter substrate-binding proteins, a putative ABC transporter ATP-binding protein, *aapP* (L-amino-acid ABC transporter ATP-binding protein), *malK* (maltose/maltodextrin import ATP-binding protein) and *chvE* (multiple sugar-binding periplasmic receptor ChvE precursor).

The *aapP* gene encodes the ATP-binding protein that is part of the Aap L-amino acid ABC transporter (Hosie and Poole 2001; Walshaw and Poole 1996), a general amino acid permease first described in *Rhizobium leguminosarum* that transports a wide range of L-amino acids (Day et al. 2001; Hosie et al. 2001; Walshaw and Poole 1996). The

Aap ABC transporter is codified by a cassette consisting of four genes, *aapJQMP* (Day et al. 2001): the AapJ is a periplasmic binding protein, the AapQM are integral membrane proteins and the AapP represents the ATP-binding protein (Day et al. 2001), probably responsible for energy coupling to the transport system. In *R. leguminosarum* the Aap system represents the main transporter of a wide range of amino acids including glutamate, aspartate, proline and histidine and is also a major transporter of aliphatic amino acids (Day et al. 2001). It has been suggested that the Aap ABC transporter is bidirectional, being responsible for the uptake and the efflux of solutes (Day et al. 2001; Hosie et al. 2001).

The *malK* gene encodes an ATP-binding protein of the ABC transporter complex $MalEFGK_2$, involved in the import of maltose/maltodextrin and responsible for energy coupling to the transport system (Boos and Shuman 1998; Schneider 2003); maltose/maltodextrin ABC uptake systems are widespread among Gram-negative and Gram-positive bacteria (Boss and Böhm 2000; Schneider 2003). The maltose/maltodextrin transporter of *Escherichia coli* is composed of the periplasmic receptor, the maltose binding protein (MalE), and the membrane-bound complex comprising two integral membrane proteins, MalF and MalG, as well as two copies of the ATPase subunit, MalK (Schneider 2003). Recent studies revealed that the MalK protein of *E. coli* down-regulates the activity of MalT, a transcriptional activator that controls the expression of the maltose regulon (Boss and Böhm 2000). Therefore, the MalK protein is considered as a bifunctional protein that regulates *mal* gene expression and maltose transport (Boss and Böhm 2000).

The *chvE* gene encodes a periplasmic sugar-binding protein identified in *Rhizobium* (= *Agrobacterium*) *tumefaciens* that is chromosomally encoded and located in the periplasmic space (Shimoda et al. 1993; Brencic and Winans 2005). In *R. tumefaciens*, ChvE is part of an operon encoding an ABC-type uptake system that both detects sugars (monosacharides) that occur in the plant-wound environment, thus participating in the chemotaxis, and interacts with the VirA periplasmic domain, enhancing the phenolic-induced activity of VirA (Brencic and Winans 2005; Kemmer et al. 1997; McCullen and Binns 2006; Shimoda et al. 1993).

Fatty acid, phospholipid and sterol metabolism: *pldB* and *AMACR* genes

Gene *pldB* was first characterized by Kobayashi et al. (1984) in two *E. coli* K-12 mutants for lysophospholipase L_2; it is located in the inner membrane and classified as a peripheral membrane protein (Kobayashi et al. 1985). Hsu et al. (1991) showed that in *E. coli* the *pldB* gene product has 2-acyl-GPE acyltransferase and acyl-ACP synthetase activities, participating in membrane phospholipid turnover and governing the acyl-CoA independent incorporation of exogenous fatty acids and lysophospholipids into the membrane. The *AMACR* gene encodes the α-methylacyl-CoA racemase (AMACR), a member of the family III CoA transferases, which catalyzes the racemization of α-methyl-branched CoA esters (Savolainen et al. 2005); in *Mycobacterium*, it plays a critical role in the β-oxidation of methyl-branched alkanes (Sakai et al. 2004).

Amino acid biosynthesis-branched chain family: *ilvD*

The *ilvD* gene encodes dihydroxy-acid dehydratase (EC 4.2.1.9), one of the enzymes involved in leucine and isoleucine-valine biosynthesis (Flint et al. 1993; Garault et al. 2000).

DNA replication, recombination and repair: *uvrA*

The *uvrA* gene encodes the ABC excinuclease subunit A protein (UvrA protein) and forms a complex with UvrB protein (Orren and Sancar 1989). The UvrA is an ATPase and a DNA-binding protein that together with UvrB and UvrC proteins acts in the excision repair of DNA (Bennett 1990).

Symbiosis: *nopP*

The *nopP* gene encodes a type-III secretion system (TTSS) effector protein, the nodulation outer protein P (NopP), identified in *Rhizobium* sp. (*S. fredii*) strain NGR234 (Ausmees et al. 2004). The TTSS are molecular devices that Gram-negative bacteria use to inject “effector” proteins into the cytoplasm of eukaryotic cells (Skorpil et al. 2005). The TTSS machinery usually consists of more than 20 proteins that form a complex apparatus across the inner membrane, the periplasmic space, the peptidoglycan layer, the outer membrane, the extracellular space and the host cellular membrane (Yip and Strynadka 2006). In rhizobial species, the TTSSs play an important role in symbiosis, acting as host-specificity determinants: in some hosts, e.g., *Crotalaria juncea*, *Pachyrhizus tuberosus,* the proteins are perceived as pathogenic and their absence results in increased nodulation. A second group of hosts, e.g., *Tephrosia vogelii*, *Flemingia congesta*, is benefited by the proteins, while in a third group those proteins have no effect (Skorpil et al. 2005; Viprey et al. 1998).

Discussion

Twenty-four DNA sequences identified in the second RDA round should be unique to the “tester” strain S 370

(Table 1), characterized by a higher capacity (three times more N_2 fixed) and efficiency (two times more N_2 fixed per g of nodule) of the N_2-fixation process than the "driver" strain S 516 (Hungria et al. 1998). Eight of those sequences have shown high similarity to known genes, nine encoded putative proteins and seven represented conserved hypothetical or hypothetical proteins, and we focused our study on the genes more directly related to the plant-rhizobium interaction.

In the transport and binding category, we have identified genes *aapP*, *malK* and *chvE*. Lodwig et al. (2003) studied the Aap and Bra broad-specificity ABC transporters in *R. leguminosarum* and the inoculation of a double mutant strain nodulated pea (*Pisum sativum*) but resulted in less shoot dry weight and N content. The effects were attributed to a disruption in the amino acid transport, once the normal phenotype was restored with the introduction of the *aap* operon (Lodwig et al. 2003). A model was then proposed in which the Aap/Bra system, localized in the bacteroid membrane, would transport the amino acids (as glutamate and aspartate) received from the plant, enabling cessation of their assimilation of ammonium, and in return, bacteroids would act like organelles cycling amino acids back to the plant for asparagine synthesis (Lodwig et al. 2003). The mutual dependence of this exchange would provide a selective pressure for the evolution of mutualism (Lodwig et al. 2003) and thus the identification of *aapP* gene might indicate that the less-effective strain S 516 has a non-functional Aap system, affecting the transport of amino acids between the bacteroid and the host soybean and resulting in lower N_2-fixation rates. However, as the main N compounds transported by soybean fixing N_2 are ureides and not asparagine as in peas, it would be interesting to restore the complete gene sequence of *aapP* gene in strain S 516, in order to estimate its contribution to symbiotic effectiveness.

Again, in the transport category, in *R. tumefaciens* the *chvE* gene encodes a periplasmic sugar-binding protein involved in chemotaxis and virulence-gene induction (Brencic and Winans 2005; Cangelosi et al. 1990; Kemmer et al. 1997; McCullen and Binns 2006; Shimoda et al. 1993). A ChvE-homologous protein has also been identified in the associative diazotroph *Azospirillum brasiliense* (Bastelaere et al. 1999; Dommelen et al. 1997), and was designated "SbpA" (sugar-binding protein A). In this bacterium, the protein is involved in the uptake of D-galactose and required for chemotaxis towards sugars as D-galactose, L-arabinose and D-fucose, being expressed in the presence of plant-root exudates (Bastelaere et al. 1999; Dommelen et al. 1997). In a proteomic study with *B. japonicum* USDA 110, Sarma and Emerich (2006) identified the protein open reading frame (ORF) blr3208, similar to ChvE of *R. tumefaciens*, which occurs in free-living conditions but not in bacteroids. A constitutively expressed ChvE-like protein was also identified in the proteome analysis of *R. leguminosarum* (Guerreiro et al. 1997). Finally, a putative ABC transporter substrate-binding protein encoded by ORF blr4511, expressed in USDA 110 in both the free-living and bacteroid states (Hoa et al. 2004; Sarma and Emerich 2006), was also identified in the effective strain S 370, but a more detailed role for this transporter remains to be elucidated.

The gene *ilvD*, which encodes a dihydroxy-acid dehydratase acting in amino acid synthesis pathways was also detected in strain S 370. Aguilar and Grasso (1991) obtained a mutant of *S. meliloti* that required isoleucine and valine for growth on minimal medium and was also symbiotically defective; the genetic analysis of this mutant revealed the lack of *ilvC* gene, encoding for the enzyme acetohydroxy acid isomeroreductase. Complementation of this mutant was achieved with *ilvC* gene from either *S. meliloti* or *E. coli*, and resulted in restoration of the nodulation phenotype in alfalfa (*Medicago sativa*) (Aguilar and Grasso 1991). Later, López et al. (2001) reported that the symbiotic interaction between *S. meliloti ilvC* mutants and alfalfa was blocked at a very early step, in the induction of the *nodABC* promoter, and the impact of mutation on *nod*-gene expression varied among different mutant strains; furthermore, there were indications that *ilvC* mutation might affect other molecules essential for effective nodulation. However, the impact of *ilvD* mutants of *B. japonicum* on the release of molecules affecting nodulation remains to be determined.

In the symbiosis category, the effective strain S 370 carried the *nopP* gene (Table 1), an effector protein of the type-three secretion system (TTSS) (Ausmees et al. 2004). In plant and animal bacterial pathogens, effector proteins act on the cytoskeleton or on intracellular signaling cascades, subverting the host's metabolism and defense reactions (Cornelis 2000; Cornelis and Gijsegem 2000). NopP and the NopL are the only two rhizobial TTSS effector proteins characterized to date (Skorpil et al. 2005). Skorpil et al. (2005) showed that NopP and NopL of *Rhizobium* sp. (*S. fredii*) strain NGR234 are phosphorylated by the plant and that the NopP is a major determinant of nodulation of some tropical legumes. In *B. japonicum*, the genes encoding the TTSS are clustered (the *tts* gene cluster) in the symbiotic island and the same transcriptional activators involved in *nod* gene regulation are also required for the expression of genes within this cluster (Krause et al. 2002). In a proteomics study, Süβ et al. (2006) identified the protein expressed by ORF blr1752 of *B. japonicum* USDA 110, which was also present in the effective strain S 370 of this study (Table 1). In USDA 110, the protein is secreted into the extracellular medium and is genistein-inducible (Süβ et al. 2006), but it remains to be determined if this protein acts as an effector in the TTSS of *B. japonicum*.

In addition to the genes identified that have known or potential functions on the symbiosis, other genes identified by RDA in this study might also affect symbiosis. In the regulatory functions, a probable two-component response regulator and a sensor histidine kinase were identified (Table 1). Two-component systems usually consist of a membrane-bound sensor protein (histidine kinase) that transfers a high-energy phosphoryl group to the response regulator, which is often a transcription factor (Bijlsma and Groisman 2003). Those systems are considered the predominant forms of signal transduction used by bacteria to respond to environmental stresses (Bijlsma and Groisman 2003); therefore, it is possible that those genes could be related to the adaptation of strain S 370 to the Cerrados. Analyzing the genome of *B. japonicum* USDA 110, Ashby (2004) found 99 potential response regulators. Examples in *B. japonicum* include the *nodV* and *nodW* genes, which are essential for the nodulation of cowpea (*Vigna unguiculata*), siratro (*Macroptilium atropurpureum*), and mung bean (*Vigna radiata*), but not soybean; NodV is the sensor protein and the NodW is the response regulator that activates the transcription of the *nodYABC* operon (Loh et al. 2002; Loh and Stacey 2003). Another two-component system identified in *B. japonicum* is represented by NwsA, a sensor protein, and NwsB, a response-regulator protein, with an action similar to NodV and NodW in *nod*-gene expression (Grob et al. 1994; Loh and Stacey 2003); the action of NwsB goes even further, and is involved in the quorum regulation of the *nod* genes (Loh and Stacey 2003). CDSs blr4775 and bll3106, identified by RDA in strain S 370 could thus encode proteins affecting *nod* gene regulation.

In conclusion, we have shown the feasibility of using the RDA technique to identify unique genes in the comparison of two *B. japonicum* variant strains contrasting in N_2-fixation properties. Some of the genes identified in the efficient strain S 370 should be involved in mechanisms related to the effectiveness of N_2 fixation. The next step in our study is thus to obtain mutants for those genes and confirm their phenotypic effects on soybean. Those genes might then represent important tools to speed up strain selection programs, accelerating pre-screening procedures. Additionally, the conserved hypothetical and hypothetical CDSs identified in the effective strain S 370 might encode important but still unknown proteins related to the symbiosis that deserve to be studied.

Acknowledgments The work was partially supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil), Universal Project (475761/2006-5), Millennium Institute and Culture Collections (552393/2005-3). The authors thank Ligia Maria Oliveira Chueire (Embrapa Soja) for help in several steps of this work and Dr. Allan R. J. Eaglesham for helpful discussion. F. G. Barcellos acknowledges a pro-doc fellowship from CAPES, P. Menna and J. S. S. Batista DTI fellowships from CNPq and M. Hungria a research fellowship from CNPq.

References

Aguilar OM, Grasso DH (1991) The product of the *Rhizobium meliloti ilvC* gene is required for isoleucine and valine synthesis and nodulation of alfalfa. J Bacteriol 173:7756–7764

Allen NL, Penn CW, Hilton AC (2003) Representational difference analysis: critical appraisal and method development for the identification of unique DNA sequences from prokaryotes. J Microbiol Meth 55:73–81

Ashby MK (2004) Survey of the number of two-component response regulator genes in the complete and annotated genome sequences of prokaryotes. FEMS Microbiol Lett 231:277–281

Ausmees N, Kobayashi H, Deakin WJ, Marie C, Krishnan HB (2004) Characterization of NopP, a type III secreted effector of *Rhizobium* sp. strain NGR234. J Bacteriol 186:4774–4780

Barcellos FG, Menna P, Batista JSS, Hungria M (2007) Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs *Sinorhizobium* (*Ensifer*) *fredii* and *Bradyrhizobium elkanii* in a Brazilian Savannah soil. Appl Environ Microbiol 73:2635–2643

Bastelaere EV, Lambrecht M, Vermeiren H, Dommelen AV, Keijers V, Proost P, Vanderleyden J (1999) Characterization of a sugar-binding protein from *Azospirillum brasilense* mediating chemotaxis to and uptake of sugars. Mol Microbiol 32:703–714

Batista JSS, Hungria M, Barcellos FG, Ferreira MC, Mendes IC (2007) Variability in *Bradyrhizobium japonicum* and *B. elkanii* seven years after introduction of both the exotic microsymbiont and the soybean host in a Cerrados soil. Microbiol Ecol 53:270–284

Bennett VH (1990) Nucleotide excision repair in *Escherichia coli*. Microbiol Rev 54:18–51

Bijlsma JJE, Groisman EA (2003) Making informed decisions: regulatory interactions between two-component systems. Trends Microbiol 11:359–366

Boddey LH, Hungria M (1997) Phenotypic grouping of Brazilian *Bradyrhizobium* strains which nodulate soybean. Biol Fertil Soils 25:407–415

Boos W, Shuman H (1998) Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. Microbiol Mol Biol Rev 62:204–229

Boss W, Böhm A (2000) Learning new tricks from an old dog: MalT of the *Escherichia coli* maltose system is part of a complex regulatory network. Trends Genet 16:404–409

Brencic A, Winans SC (2005) Detection of and response to signals involved in host–microbe interactions by plant-associated bacteria. Microbiol Mol Biol Rev 69:155–194

Cangelosi GA, Ankenbauer RG, Nester EW (1990) Sugars induce the *Agrobacterium* virulence genes through a periplasmic binding protein and a transmembrane signal protein. Proc Natl Acad Sci USA 87:6708–6712

Cornelis GR (2000) Type III secretion: a bacterial device for close combat with cells of their eukaryotic host. Philos Trans R Soc London Ser B 355:681–693

Cornelis GR, Gijsegem FV (2000) Assembly and function of type III secretory systems. Annu Rev Microbiol 54:735–774

Day DA, Poole PS, Tyerman SD, Rosendahl L (2001) Ammonia and amino acid transport across symbiotic membranes in nitrogen-fixing legume nodules. Cell Mol Life Sci 58:61–71

Dommelen AV, Bastelaere EV, Keijers V, Vanderleyden J (1997) Genetics of *Azospirillum brasilense* with respect to ammonium transport, sugar uptake, and chemotaxis. Plant Soil 194:155–160

Ermolaeva OD, Sverdlov ED (1996) Subtractive hybridization, a technique for extraction of DNA sequences distinguishing two closely related genomes: critical analysis. Genet Anal Biomol E 13:49–58

Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8:186–194

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8:175–185

Flint DH, Emptage MH, Finnegan MG, Fu W, Johnson MK (1993) The role and properties of the iron-sulfur cluster in *Escherichia coli* dihidroxy-acid dehydratase. J Biol Chem 268:14732–14742

Galli-Terasawa LV, Glienke-Blanco C, Hungria M (2003) Diversity of soybean rhizobial population adapted to a Cerrados soil. World J Microb Biotechnol 19:933–939

Garault P, Letort C, Juillard V, Monnet V (2000) Branched-chain amino acid biosynthesis is essential for optimal growth of *Streptococcus thermophilus* in milk. Appl Environ Microbiol 66:5128–5133

Gordon D, Abajian C, Green P (1998) Consed: a graphical tool for sequence finishing. Genome Res 8:195–202

Grob P, Hennecke H, Göttfert M (1994) Cross-talk between the two-component regulatory systems NodVW and NwsAB of *Bradyrhizobium japonicum*. FEMS Microbiol Lett 120:349–354

Guerreiro N, Redmond JW, Rolfe BG, Djordjevic MA (1997) New *Rhizobium leguminosarum* flavonoid-induced proteins revealed by proteome analysis of differentially displayed proteins. Mol Plant Microbe Interact 10:506–516

Guo H, Sun S, Finan TM, Xu J (2005) Novel DNA sequences from natural strains of the nitrogen-fixing symbiotic bacterium *Sinorhizobium meliloti*. Appl Environ Microbiol 71:7130–7138

Hoa LTP, Nomura M, Tajima S (2004) Characterization of bacteroid proteins in soybean nodules formed with *Bradyrhizobium japonicum* USDA110. Microbes Environ 19:71–75

Hosie AHF, Poole PS (2001) Bacterial ABC transporters of amino acids. Res Microbiol 152:259–270

Hosie AHF, Allaway D, Jones MA, Walshaw DL, Johnston AWB, Poole PS (2001) Solute-binding protein-dependent ABC transporters are responsible for solute efflux in addition to solute uptake. Mol Microbiol 40:1449–1459

Hsu L, Jackowski S, Rock CO (1991) Isolation and characterization of *Escherichia coli* K-12 mutants lacking both 2-acylglycerophosphoethanolamine acyltransferase and acyl-acyl carrier protein synthetase activity. J Biol Chem 266:13783–13788

Hungria M, Vargas MAT (2000) Environmental factors affecting N_2 fixation in grain legumes in the tropics, with an emphasis on Brazil. Field Crop Res 65:151–164

Hungria M, Nishi CYM, Cohn J, Stacey G (1996) Comparison between parental and variant soybean *Bradyrhizobium* strains with regard to the production of lipo-chitin nodulation signals, early stages of root infection, nodule occupancy, and N_2 fixation rates. Plant Soil 186:331–341

Hungria M, Boddey LH, Santos MA, Vargas MAT (1998) Nitrogen fixation capacity and nodule occupancy by *Bradyrhizobium japonicum* and *B. elkanii* strains. Biol Fertil Soils 27:393–399

Hungria M, Franchini JC, Campo RJ, Graham PH (2005) The importance of nitrogen fixation to soybean cropping in South America. In: Werner D, Newton WE (eds) Nitrogen fixation in agriculture, forestry, ecology and environment. Kluwer Academic Publishers, Dordrecht, pp 25–42

Hungria M, Franchini JC, Campo RJ, Crispino CC, Moraes JZ, Sibaldelli RNR, Mendes IC, Arihara J (2006a) Nitrogen nutrition of soybean in Brazil: contributions of biological N_2 fixation and of N fertilizer to grain yield. Can J Plant Sci 86:927–939

Hungria M, Campo RJ, Mendes IC, Graham PH (2006b) Contribution of biological nitrogen fixation to the nitrogen nutrition of grain crops in the tropics: the success of soybean (*Glycine max* L. Merr.) in South America. In: Singh RP, Shankar N, Jaiwal PK (eds) Nitrogen nutrition in plant productivity. Studium Press, Houston, pp 43–93

Kemmer JM, Liang X, Nester EW (1997) The *Agrobacterium tumefaciens* virulence gene *chvE* is part of a putative ABC-type sugar transport operon. J Bacteriol 179:2452–2458

Kobayashi T, Homma H, Natori Y, Kudo I, Inoue K, Nojima S (1984) Isolation of two kinds of *E. coli* K-12 mutants for lysophospholipase L_2: one with an elevated level of the enzyme and the other defective in it. J Biochem 96:137–145

Kobayashi T, Kudo I, Homma H, Karasawa K, Inoue K, Ikeda H, Nojima S (1985) Gene organization of *pldA* and *pldB*, the structural genes for detergent-resistant phospholipase A and lysophospholipase L_2 of *Escherichia coli*. J Biochem 98:1007–1016

Krause A, Doerfel A, Göttfert M (2002) Mutational and transcriptional analysis of the type III secretion system of *Bradyrhizobium japonicum*. Mol Plant Microbe Interact 15:1228–1235

Lisitsyn N, Lisitsyn N, Wigler M (1993) Cloning the differences between two complex genomes. Science 259:946–951

Lodwig EM, Hosie AHF, Bourdès A, Findlay K, Allaway D, Karunakaran R, Downie JA, Poole PS (2003) Amino-acid cycling drives nitrogen fixation in the legume–rhizobium symbiosis. Nature 422:722–726

Loh J, Lohar DP, Andersen B, Stacey G (2002) A two-component regulator mediates population-density-dependent expression of the *Bradyrhizobium japonicum* nodulation genes. J Bacteriol 184:1759–1766

Loh J, Stacey G (2003) Nodulation gene regulation in *Bradyrhizobium japonicum*: a unique integration of global regulatory circuits. Appl Environ Microbiol 69:10–17

López JC, Grasso DH, Frugier F, Crespi MD, Aguilar OM (2001) Early symbiotic responses induced by *Sinorhizobium meliloti ilvC* mutants in alfalfa. Mol Plant Microbe Interact 14:55–62

McCullen CA, Binns AN (2006) *Agrobacterium tumefaciens* and plant cell interactions and activities required for interkingdom macromolecular transfer. Annu Rev Cell Dev Biol 22:101–127

Mendes IC, Hungria M, Vargas MAT (2004) Establishment of *Bradyrhizobium japonicum* and *B. elkanii* strains in a Brazilian Cerrado oxisol. Biol Fertil Soils 40:28–35

Menna P, Hungria M, Barcellos FG, Bangel EV, Hess PN, Martínez-Romero E (2006) Molecular phylogeny based on the 16S rRNA gene of elite rhizobial strains used in Brazilian commercial inoculants. Syst Appl Microbiol 29:315–332

Nishi CYM, Boddey LH, Vargas MAT, Hungria M (1996) Morphological, physiological and genetic characterization of two new *Bradyrhizobium* strains recently recommended in Brazilian commercial inoculants for soybean. Symbiosis 20:147–162

Orren DK, Sancar A (1989) The (A)BC excinuclease of *Escherichia coli* has only the UvrB and UvrC subunits in the incision complex. Proc Natl Acad Sci USA 86:5237–5241

Sagerström CG, Sun BI, Sive HL (1997) Subtractive cloning: past, present, and future. Annu Rev Biochem 66:751–783

Sakai Y, Takahashi H, Wakasa Y, Kotani T, Yurimoto H, Miyachi N, Veldhoven PPV, Kato N (2004) Role of α-methylacyl coenzyme A racemase in the degradation of methyl-branched alkanes by *Mycobacterium* sp. strain P101. J Bacteriol 186:7214–7220

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Plainview, New York

Santos MA, Vagas MAT, Hungria M (1999) Characterization of soybean *Bradyrhizobium* strains adapted to the Brazilian savannas. FEMS Microbiol Ecol 30:261–272

Sarma AD, Emerich DW (2006) A comparative proteomic evaluation of culture grown vs nodule isolated *Bradyrhizobium japonicum*. Proteomics 6:3008–3028

Savolainen K, Bhaumik P, Schmitz W, Kotti TJ, Conzelmann E, Wierenga RK, Hiltunen JK (2005) α-Methylacyl-CoA racemase from *Mycobacterium tuberculosis*. J Biol Chem 280:12611–12620

Schneider E (2003) Import of solutes by ABC transporters: the maltose system. In: Holland IB, Cole SPC, Kuchler K, Higgins CF (eds) ABC protein: from bacteria to man. Elsevier Science, London, pp 157–185

Shimoda N, Yamamoto AT, Aoki S, Machida Y (1993) Genetic evidence for an interaction between the VirA sensor protein and the ChvE sugar-binding protein of *Agrobacterium*. J Biol Chem 268:26552–26558

Skorpil P, Saad MM, Boukli NM, Kobayashi H, Ares-Orpel F, Broughton WJ, Deakin WJ (2005) NopP, a phosphorylated effector of *Rhizobium* sp. strain NGR234, is a major determinant of nodulation of the tropical legumes *Flemingia congesta* and *Tephrosia vogelii*. Mol Microbiol 57:1304–1317

Süβ C, Hempel J, Zehner S, Krause A, Patschkowski T, Göttfert M (2006) Identification of genistein-inducible and type III-secreted proteins of *Bradyrhizobium japonicum*. J Biotechnol 126:69–77

Tenaillon O, Denamur E, Matic I (2004) Evolutionary significance of stress-induced mutagenesis in bacteria. Trends Microbiol 12:264–270

Viprey V, Greco AD, Golinowski W, Broughton WJ, Perret X (1998) Symbiotic implications of type III protein secretion machinery in *Rhizobium*. Mol Microbiol 28:1381–1389

Walshaw DL, Poole PS (1996) The general L-amino acid permease of *Rhizobium leguminosarum* is an ABC uptake system that also influences efflux of solutes. Mol Microbiol 21:1239–1252

Wright BE (2004) Stress-directed adaptive mutations and evolution. Mol Microbiol 52:643–650

Yip CK, Strynadka NCJ (2006) New structural insights into the bacterial type III secretion system. Trends Biochem Sci 31:223–230