



Proteomic map of *Bradyrhizobium diazoefficiens* strain CPAC 7, an outstanding symbiont of soybean

Proteomic analysis of free-living *Bradyrhizobium diazoefficiens*: highlighting potential determinants of a successful symbiosis

Gomes *et al.*

RESEARCH ARTICLE

Open Access

Proteomic analysis of free-living *Bradyrhizobium diazoefficiens*: highlighting potential determinants of a successful symbiosis

Douglas Fabiano Gomes^{1,2}, Jesiane Stefânia da Silva Batista^{1,3}, Amanda Alves Paiva Rolla¹, Luciano Paulino da Silva⁴, Carlos Bloch⁴, Lygia Vitoria Galli-Terasawa² and Mariangela Hungria^{1*}

Abstract

Background: Strain CPAC 7 (=SEMIA 5080) was recently reclassified into the new species *Bradyrhizobium diazoefficiens*; due to its outstanding efficiency in fixing nitrogen, it has been used in commercial inoculants for application to crops of soybean [*Glycine max* (L.) Merr.] in Brazil and other South American countries. Although the efficiency of *B. diazoefficiens* inoculant strains is well recognized, few data on their protein expression are available.

Results: We provided a two-dimensional proteomic reference map of CPAC 7 obtained under free-living conditions, with the successful identification of 115 spots, representing 95 different proteins. The results highlighted the expression of molecular determinants potentially related to symbiosis establishment (e.g. inositol monophosphatase, IMPase), fixation of atmospheric nitrogen (N₂) (e.g. NifH) and defenses against stresses (e.g. chaperones). By using bioinformatic tools, it was possible to attribute probable functions to ten hypothetical proteins. For another ten proteins classified as “NO related COG” group, we analyzed by RT-qPCR the relative expression of their coding-genes in response to the nodulation-gene inducer genistein. Six of these genes were up-regulated, including blr0227, which may be related to polyhydroxybutyrate (PHB) biosynthesis and competitiveness for nodulation.

Conclusions: The proteomic map contributed to the identification of several proteins of *B. diazoefficiens* under free-living conditions and our approach—combining bioinformatics and gene-expression assays—resulted in new information about unknown genes that might play important roles in the establishment of the symbiosis with soybean.

Keywords: Symbiosis, Nitrogen fixation, Two-dimensional proteomics, RT-qPCR, *Bradyrhizobium*

Background

Biological N₂ fixation (BNF) is a fundamental component of the global nitrogen (N) cycle, both in natural and agricultural environments. The symbiosis of legumes with soil-borne symbiotic N₂-fixing bacteria, which are frequently referred to as rhizobia, can often provide more than 60% of the plant's N requirements [1,2]. Regarding the concept of agriculture sustainability, BNF contributes to the improvement of food production without cultivation of new lands, to lowering input costs for the farmers and to mitigating environmental degradation. Such benefits occur when BNF replaces chemical N-fertilizers, which

are expensive, and, among other harmful environment impacts, foment greenhouse-gas emissions [3,4].

Cultivation of soybean [*Glycine max* (L.) Merr.] has increased globally, mainly due to its high protein and oil contents, and plant breeding has resulted in increasing yields [5]. Certainly, efficient BNF is a major contributor to the achievement of high yields with low input costs [6]. An important example is the contribution of BNF to soybean cropping in Brazil, associated with application to the seeds at sowing of inoculants containing elite strains of *Bradyrhizobium*, including CPAC 15 (=SEMIA 5079) and CPAC 7 (=SEMIA 5080) [7,8]. The combination of these strains can fulfill much of the crop's N needs, resulting in estimated savings of about US \$15 billion in N-fertilizers per cropping season [9].

* Correspondence: mariangela.hungria@embrapa.br

¹Embrapa Soja, Embrapa Soja, C.P. 231, 86001-970 Londrina, Paraná, Brazil
Full list of author information is available at the end of the article

Bradyrhizobium diazoefficiens was recently reclassified as a novel species on the bases of morpho-physiological, genotypic and genomic differences from *Bradyrhizobium japonicum* [10]. Strain CPAC 7 (=SEMIA 5080) has outstanding efficiency in fixing N₂ with soybean and good adaptation to the often-stressful edaphoclimatic conditions of the tropics [11,12]. These features are responsible for the inclusion of this strain in inoculants applied to soybean in Brazil since 1992 [7,9].

The type strain of *B. diazoefficiens*, USDA 110^T has had its genome elucidated; however, of the 8,317 potential protein-encoding genes, 30% were assigned as hypothetical and 18% showed no similarity to any known gene [13]. Later, the expression of several predicted protein-encoding genes in USDA 110^T was reported in transcriptomic and proteomic studies [14-19]. Nevertheless, despite the economic importance of *B. diazoefficiens* as a component of soybean inoculants worldwide [9,10], few data are available on the proteins it synthesizes in the free-living state. It is well known that major attributes of successful elite strains, such as saprophytic competence, adaptation to stressful conditions and nodulation competitiveness must be expressed when free-living.

Our research group has just completed the genome sequencing of strain CPAC 7 [20] and, as occurred with USDA 110 [13], it was not possible to attribute functions to about 50% of the genes. Therefore, the establishment of a proteomic reference map for this strain in the free-living state can both add valuable protein-expression data to the genomic-annotation process [21-23] and help to attribute probable functions to hypothetical proteins [21,23].

Here we present the first two-dimensional proteomic reference map for free-living *B. diazoefficiens* strain CPAC 7, emphasizing molecular determinants of symbiosis-establishment and of tolerance of environmental stresses. Additionally, we ascribe putative functions to some hypothetical proteins detected at the proteomic level. For other hypothetical proteins without available information, we analyzed the relative expression of their coding-genes in response to the main soybean-nodulation-inducing molecule, the flavonoid genistein.

Results and discussion

Two-dimensional gel electrophoresis and protein identification

In studies with two-dimensional gels, it is necessary to optimize the resolution of the protein maps as a function of the nature and characteristics of the samples studied. With this goal, a previous experiment to obtain an overview of the protein distribution of *B. diazoefficiens* strain CPAC 7 was carried out with a broad-range IPG strip (pH 3–10) for the first-dimensional protein separation. After SDS-PAGE, the results showed that most of protein spots remained clustered in the pI range of pH 5–7 (data

not shown). To improve the separation of the proteins, we then employed in the first-dimension electrophoresis IPG strips with a narrower pH range (pH 4–7), that confirmed, in triplicated gels, better resolution than the strips with pH 3–10.

Using computer-assisted gel-image analysis software, well-defined spots were detected and the majority of their molecular weights ranged between 14 kDa and 97 kDa (Figure 1). Among these, 150 spots were selected and analyzed by MALDI-TOF MS or, when necessary, by MS/MS. Mass spectra of peptide fragments were compared with database entries, regarding the statistical requirements, and 115 spots were successfully identified, representing 95 different proteins (Tables 1 and 2). Information on the spectrometry data set is available in Additional file 1: Table S1. The presence of distinct spots for the same protein may be the result of posttranslational modifications.

Protein functional classification and cellular location

According to the functional classification in COG, proteins were distributed in 16 categories, belonging to four functional groups (Figure 2). In the metabolism-related functional group, there were eight categories, comprising the greatest number of experimentally identified proteins (40%). Next, 21% of the proteins were clustered in three functional categories related to information storage and processing, while the cellular-processes signaling group encompassed 13% of the proteins distributed in three categories. Finally, seven proteins were pooled in two other categories of a poorly characterized group, and 20 proteins did not fit any of the COG categories, being assigned as “not in COG” (Figure 2).

The high percentage of proteins with metabolic functions in CPAC 7 (Figure 2) is consistent with rhizobia's ability to adapt to varied edaphoclimatic conditions. Several of these proteins are related to energy metabolism, e.g. succinate dehydrogenase flavoprotein subunit and malate dehydrogenase, which participate in the tricarboxylic acid (TCA) cycle, the main pathway for obtaining energy and also important in the synthesis of precursors of the biosynthesis of amino acids, purines, pyrimidines and vitamins [23].

Proteins associated with amino-acid and lipid metabolism may be particularly important for free-living cells, and several of those proteins identified in CPAC 7 have been previously reported in *B. japonicum* CPAC 15 [23]. Beyond their main functions, proteins in these two categories may also play important roles at various stages of the symbiotic interaction, since auxotrophic mutants in both of them are defective in both nodulation and in N₂-fixation abilities [24].

The second largest functional group—information storage and processing (Figure 2)—encompassed several transcriptional and translational factors (Table 1). These proteins have fundamental roles in controlling metabolic pathways because they regulate and ensure the accuracy

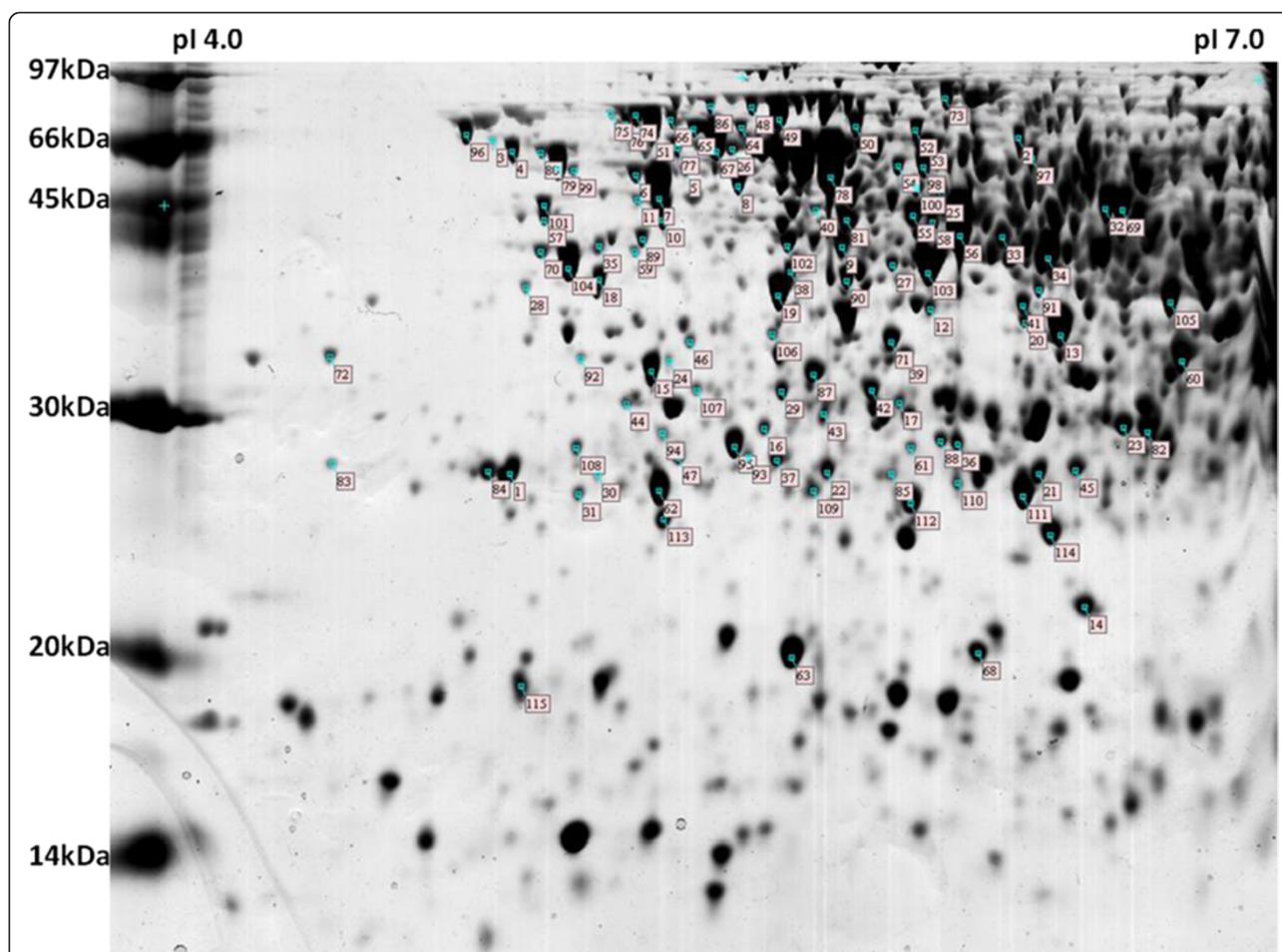


Figure 1 Two-dimensional electrophoresis protein profile of *Bradyrhizobium diazoefficiens* CPAC 7 whole cell extract at free-living state. More information about expressed proteins is available in Tables 1 and 2.

of gene expression [25]; furthermore, under stress conditions they can also perform as chaperonins, helping in *de novo* protein folding and preventing damaged proteins from forming aggregates [26,27].

Most proteins belonging to the cellular-processes-and-signaling functional category were correlated with defense against stressful conditions (Table 1). Mechanisms of response to stresses are usually conserved among bacterial species, and required for rapid adaptation to environmental and metabolic changes. One of these responses comprises the expression of molecular chaperones, such as DnaK, GrpE and GroEL [28], all of which were detected in our study (Table 1). Also related to the mediation of adaptive responses to adverse conditions [29,30], two Clp proteases, ClpB and ClpX, were expressed in *B. diazoefficiens* CPAC 7. Finally, there was a cytoplasmic protein member of the two component response regulator OmpR family; proteins of this family are amongst the best characterized bacterial positive regulators, improving the transcriptional capacity of RNA polymerase, with reported effects in osmoregulation in *Escherichia coli* [31].

The cellular locations of all 115 identified proteins, predicted by PSORT-B and PSLpred, are listed in Tables 1 and 2. Although the majority of the proteins extracted in our study are located in cell cytoplasm and periplasm, two inner-membrane and one extracellular protein were obtained. Similar results have been reported in previous rhizobial proteomic studies by our group [23,32].

Symbiosis establishment and N₂-fixation-related proteins

In the establishment of the legume-rhizobia symbiosis, an exchange of molecular signals starts with the host plant's release of molecules, mainly flavonoids, that induce expression of rhizobial nodulation (*nod*) genes. The products of *nod*-genes, the Nod factors (lipo-chitin oligosaccharides, LCOs), play critical roles in root nodulation [33]. Host specificity is also determined by the Nod factors, by means of the incorporation of α,β -unsaturated acyl chains in the backbone structure [34]. In our proteomic reference map, we identified one acyl carrier protein, 3-oxoacyl-ACP synthase, and one ACP S-malonyl transferase (Table 2), both required for the synthesis of essential fatty acyl chains.

Table 1 Identified proteins of *Bradyrhizobium diazoefficiens* CPAC 7 whole cell extract and classification according to COG

Spot ID	Gene	Product	NCBI ID	Cellular location	*T/**E pI	*T/**E MW	Organism
Metabolism							
Energy production and conversion							
1	<i>nuoC</i>	NADH dehydrogenase subunit C	gi 27380028	Cytoplasmic	4.88/4.94	23201/27000	<i>B. diazoefficiens</i> USDA 110
2	<i>sdhA</i>	Succinate dehydrogenase flavoprotein subunit	gi 27375625	Periplasmic	5.91/6.34	66903/70000	<i>B. diazoefficiens</i> USDA 110
3	<i>pdhB</i>	Pyruvate dehydrogenase subunit beta	gi 27379893	Cytoplasmic	4.81/4.90	48906/65000	<i>B. diazoefficiens</i> USDA 110
4	<i>pdhB</i>	Pyruvate dehydrogenase subunit beta	gi 27379893	Cytoplasmic	4.81/4.95	48906/65000	<i>B. diazoefficiens</i> USDA 110
5	<i>atpD</i>	ATP synthase F0F1 subunit beta	gi 27375551	Cytoplasmic	5.13/5.42	50987/62000	<i>B. diazoefficiens</i> USDA 110
6	<i>atpD</i>	ATP synthase F0F1 subunit beta	gi 27375552	Cytoplasmic	5.13/5.28	50987/62000	<i>B. diazoefficiens</i> USDA 110
7	<i>atpD</i>	ATP synthase F0F1 subunit beta	gi 27375551	Cytoplasmic	5.13/5.35	50987/57000	<i>B. diazoefficiens</i> USDA 110
8		Aldehyde dehydrogenase	gi 27379895	Cytoplasmic	6.04/5.57	55297/60000	<i>B. diazoefficiens</i> USDA 110
9		Succinate-semialdehyde dehydrogenase	gi 27379109	Cytoplasmic	5.3/5.86	50087/51000	<i>B. diazoefficiens</i> USDA 110
10	<i>eno</i>	Phosphopyruvate hydratase	gi 27379905	Cytoplasmic	5.08/5.36	45314/53000	<i>B. diazoefficiens</i> USDA 110
11	<i>eno</i>	Phosphopyruvate hydratase	gi 27379905	Cytoplasmic	5.08/5.28	45314/55000	<i>B. diazoefficiens</i> USDA 110
12	<i>mdh</i>	Malate dehydrogenase	gi 384214148	Cytoplasmic	5.88/6.09	34259/41000	<i>B. japonicum</i> USDA 6
13	<i>mdh</i>	Malate dehydrogenase	gi 27375567	Cytoplasmic	5.88/6.09	34275/41000	<i>B. diazoefficiens</i> USDA 110
14		Rieske iron-sulfur protein	gi 354959910	Cytoplasmic	5.9/6.52	15192/22000	<i>B. japonicum</i> USDA 6
15	<i>etfL</i>	Electron transfer flavoprotein large subunit	gi 27376489	Cytoplasmic	5.14/5.33	32186/34000	<i>B. diazoefficiens</i> USDA 110
16		Ferredoxin NADP + reductase	gi 27375211	Cytoplasmic	5.54/5.64	31764/29000	<i>B. diazoefficiens</i> USDA 110
Carbohydrate transport and metabolism							
17		Inositol monophosphatase	gi 27382842	Cytoplasmic	5.61/6.02	28290/31000	<i>B. diazoefficiens</i> USDA 110
18		Sugar kinase	gi 27375915	Cytoplasmic	5.06/5.20	35336/38000	<i>B. diazoefficiens</i> USDA 110
19		Sugar ABC transporter substrate-binding protein	gi 27378319	Periplasmic	5.63/5.68	38378/38000	<i>B. diazoefficiens</i> USDA 110
20		6-Phosphogluconate dehydrogenase	gi 27381870	Cytoplasmic	5.88/6.35	35880/36000	<i>B. diazoefficiens</i> USDA 110
Amino acid transport and metabolism							
21	<i>trpF</i>	N-(5'-phosphoribosyl)anthranilate isomerase	gi 27375855	Cytoplasmic	5.82/6.36	23966/27000	<i>B. diazoefficiens</i> USDA 110
22	<i>leuD</i>	Isopropylmalate isomerase small subunit	gi 27375606	Cytoplasmic	5.52/5.74	22781/27000	<i>B. diazoefficiens</i> USDA 110
23		Amino acid ABC transporter substrate-binding protein	gi 27379557	Periplasmic	6.21/6.60	36860/30000	<i>B. diazoefficiens</i> USDA 110
24	<i>dapF</i>	Diaminopimelate epimerase	gi 27375588	Cytoplasmic	5.09/5.30	31803/33000	<i>B. diazoefficiens</i> USDA 110
26	<i>glnA</i>	Glutamine synthetase	gi 27380060	Cytoplasmic	5.44/5.47	52623/66000	<i>B. diazoefficiens</i> USDA 110
27		L-asparaginase	gi 27380061	Periplasmic	5.93/5.93	39549/40000	<i>B. diazoefficiens</i> USDA 110
28	<i>serB</i>	Phosphoserine phosphatase	gi 27381616	Cytoplasmic	4.83/4.93	32322/39000	<i>B. diazoefficiens</i> USDA 110
29	<i>argB</i>	Acetylglutamate kinase	gi 27383212	Cytoplasmic	5.33/5.60	31294/32000	<i>B. diazoefficiens</i> USDA 110

Table 1 Identified proteins of *Bradyrhizobium diazoefficiens* CPAC 7 whole cell extract and classification according to COG (Continued)

Coenzyme transport and metabolism							
31	<i>thiE</i>	Thiamine-phosphate pyrophosphorylase	gi 27381769	Cytoplasmic	4.99/5.06	22484/26000	<i>B. diazoefficiens</i> USDA 110
32	<i>ahcY</i>	S-Adenosyl-L-homocysteine hydrolase	gi 27381055	Cytoplasmic	6.00/6.55	52318/45000	<i>B. diazoefficiens</i> USDA 110
33	<i>metK</i>	S-Adenosylmethionine synthetase	gi 27381056	Cytoplasmic	5.88/6.25	43613/42000	<i>B. diazoefficiens</i> USDA 110
34	<i>metK</i>	S-Adenosylmethionine synthetase	gi 27381056	Cytoplasmic	5.88/6.38	43613/41000	<i>B. diazoefficiens</i> USDA 110
35	<i>hemB</i>	Delta-aminolevulinic acid dehydratase/Prophobilinogen synthase	gi 1170210	Cytoplasmic	4.99/5.11	38843/42000	<i>B. diazoefficiens</i> USDA 110
Nucleotide transport and metabolism							
38	<i>hemH</i>	Phosphoribosylaminoimidazole-succinocarboxamide synthase	gi 27375923	Cytoplasmic	5.35/5.64	33736/40000	<i>B. diazoefficiens</i> USDA 110
39	<i>hemH</i>	Phosphoribosylaminoimidazole-succinocarboxamide synthase	gi 27375923	Cytoplasmic	5.35/5.90	33736/34000	<i>B. diazoefficiens</i> USDA 110
Lipid transport and metabolism							
40		3-Oxoacyl-ACP synthase	gi 27378919	Cytoplasmic	5.56/5.71	44874/45000	<i>B. diazoefficiens</i> USDA 110
41	<i>ispH</i>	4-Hydroxy-3-methylbut-2-enyl diphosphate reductase	gi 27376425	Cytoplasmic	5.75/6.3	35239/37000	<i>B. diazoefficiens</i> USDA 110
42	<i>fabD</i>	Nitrogenase iron protein ACP S-malonyl transferase	gi 27379193	Cytoplasmic	5.57/5.87	32462/32000	<i>B. diazoefficiens</i> USDA 110
43	<i>fadB</i>	Enoyl CoA hydratase	gi 27378147	Cytoplasmic	5.44/5.73	27829/30000	<i>B. diazoefficiens</i> USDA 110
44	<i>lpk</i>	4-Diphosphocytidyl-2-C-methyl-D-erythritol kinase	gi 27377637	Cytoplasmic	5.03/5.19	31076/31000	<i>B. diazoefficiens</i> USDA 110
Inorganic ion transport and metabolism							
45	<i>modA</i>	ABC transporter molybdenum-binding protein	gi 27383271	Periplasmic	6.62/6.46	27290/27000	<i>B. diazoefficiens</i> USDA 110
46	<i>nifH</i>	Nitrogenase iron protein	gi 128264	Cytoplasmic	5.03/5.36	31902/35000	<i>B. diazoefficiens</i> USDA 110
Secondary metabolites biosynthesis, transport and catabolism							
47		Thiol oxidoreductase FrnE	gi 27380891	Cytoplasmic	5.1/5.33	24357/28000	<i>B. diazoefficiens</i> USDA 110
Information storage and processing							
Translation, ribosomal structure and biogenesis							
48	<i>fusA</i>	Elongation factor G	gi 27380514	Cytoplasmic	5.32/5.52	76067/84000	<i>B. diazoefficiens</i> USDA 110
49	<i>fusA</i>	Elongation factor G	gi 27380514	Cytoplasmic	5.32/5.60	76067/81000	<i>B. diazoefficiens</i> USDA 110
50	<i>typA</i>	GTP-binding tyrosin phosphorylated protein	gi 27375651	Cytoplasmic	5.49/5.82	67131/74000	<i>B. diazoefficiens</i> USDA 110
51	<i>typA</i>	GTP-binding tyrosin phosphorylated protein	gi 27375651	Cytoplasmic	5.49/5.26	67131/72000	<i>B. diazoefficiens</i> USDA 110
52	<i>aspS</i>	Aspartyl-tRNA synthetase	gi 27379254	Cytoplasmic	5.71/6.00	66989/73000	<i>B. diazoefficiens</i> USDA 110
53		Transporter ATP-binding protein	gi 27381796	Inner-membrane	5.58/6.00	61850/67000	<i>B. diazoefficiens</i> USDA 110
54	<i>gatA</i>	Aspartyl/glutamyl-tRNA amidotransferase subunit A	gi 13470863	Cytoplasmic	5.66/5.95	55801/59000	<i>Mesorhizobium loti</i> MAFF303099
55	<i>tuf</i>	Elongation factor Tu	gi 27380513	Cytoplasmic	5.78/5.99	43569/44000	<i>B. diazoefficiens</i> USDA 110
56	<i>tuf</i>	Elongation factor Tu	gi 27380513	Cytoplasmic	5.79/6.12	43569/43000	<i>B. diazoefficiens</i> USDA 110
57	<i>tuf</i>	Elongation factor Tu	gi 27380513	Cytoplasmic	5.78/5.00	43569/42000	<i>B. diazoefficiens</i> USDA 110
58	<i>tuf</i>	Elongation factor Tu	gi 27380513	Cytoplasmic	5.79/6.05	43569/43000	<i>B. diazoefficiens</i> USDA 110

Table 1 Identified proteins of *Bradyrhizobium diazoefficiens* CPAC 7 whole cell extract and classification according to COG (Continued)

59	<i>hisZ</i>	ATP phosphoribosyltransferase	gi 27382635	Cytoplasmic	5.1/5.22	41078/41000	<i>B. diazoefficiens</i> USDA 110
60		Elongation factor Ts	gi 27379971	Cytoplasmic	6.17/6.77	32175/33000	<i>B. diazoefficiens</i> USDA 110
61		Sigma-54 modulation protein	gi 27375835	Cytoplasmic	5.62/5.97	21727/29000	<i>B. diazoefficiens</i> USDA 110
62	<i>rplI</i>	50S ribosomal protein L9	gi 27379187	Cytoplasmic	5.08/5.22	21886/26000	<i>B. diazoefficiens</i> USDA 110
63	<i>rpsF</i>	ACP S-malonyltransferase	gi 27379190	Cytoplasmic	5.46/5.64	18616/20000	<i>B. diazoefficiens</i> USDA 110
RNA processing and modification							
64	<i>rpsA</i>	30S ribosomal protein S1	gi 384214454	Cytoplasmic	5.27/5.49	62737/71000	<i>B. japonicum</i> USDA 6
65	<i>rpsA</i>	30S ribosomal protein S1	gi 27375851	Cytoplasmic	5.27/5.38	64213/70000	<i>B. diazoefficiens</i> USDA 110
66	<i>rpsA</i>	30S ribosomal protein S1	gi 27375851	Cytoplasmic	5.27/5.31	64213/71000	<i>B. diazoefficiens</i> USDA 110
67	<i>rpsA</i>	30S ribosomal protein S1	gi 27375851	Cytoplasmic	5.27/5.42	64213/64000	<i>B. diazoefficiens</i> USDA 110
Transcription							
68	<i>greA</i>	Transcription elongation factor GreA	gi 27382489	Cytoplasmic	5.67/6.18	17169/20000	<i>B. diazoefficiens</i> USDA 110
69	<i>rho</i>	Transcription termination factor Rho	gi 27375746	Cytoplasmic	6.08/6.6	47121/45000	<i>B. diazoefficiens</i> USDA 110
70	<i>rpoA</i>	DNA-directed RNA polymerase alpha subunit	gi 354957001	Cytoplasmic	4.91/4.97	38082/41000	<i>B. diazoefficiens</i> USDA 110
71	<i>rpoA</i>	DNA-directed RNA polymerase subunit alpha	gi 27380487	Cytoplasmic	4.9/5.5	38035/35000	<i>B. diazoefficiens</i> USDA 110
Cellular processes and signaling							
Posttranslational modification, protein turnover, chaperones							
73	<i>clpB</i>	ATP-dependent protease ATP-binding subunit	gi 27376515	Cytoplasmic	5.7/6.10	96620/90000	<i>B. diazoefficiens</i> USDA 110
74	<i>dnaK</i>	Heat shock protein	gi 12642164	Cytoplasmic	5.27/5.21	65113/72000	<i>Bradyrhizobium</i> sp. WM9
75	<i>htpG</i>	Heat shock protein 90	gi 27382900	Cytoplasmic	5.08/5.16	69004/72000	<i>B. diazoefficiens</i> USDA 110
76	<i>htpG</i>	Heat shock protein 90	gi 27382900	Cytoplasmic	5.08/5.19	69004/710000	<i>B. diazoefficiens</i> USDA 110
77	<i>groEL</i>	Molecular chaperone GroEL	gi 27377170	Cytoplasmic	5.19/5.32	57749/65000	<i>B. diazoefficiens</i> USDA 110
78	<i>groEL</i>	Molecular chaperone GroEL	gi 27380737	Cytoplasmic	5.45/5.75	57716/58000	<i>B. diazoefficiens</i> USDA 110
79	<i>tig</i>	Trigger factor	gi 27380056	Cytoplasmic	4.87/5.06	50061/57000	<i>B. diazoefficiens</i> USDA 110
80	<i>tig</i>	Trigger factor	gi 27380056	Cytoplasmic	4.87/4.97	50061/64000	<i>B. diazoefficiens</i> USDA 110
81	<i>clpX</i>	ATP-dependent protease ATP-binding subunit ClpX	gi 27380054	Cytoplasmic	5.575.80	46932/44000	<i>B. diazoefficiens</i> USDA 110
82		Anti-oxidant protein	gi 27380419	Cytoplasmic	6.1/6.66	24420/29000	<i>B. diazoefficiens</i> USDA 110
84	<i>grpE</i>	Heat shock protein GrpE	gi 27375787	Cytoplasmic	4.84/4.84	21642/27000	<i>B. diazoefficiens</i> USDA 110
85	<i>Pcm</i>	Protein-L-isoaspartate O-methyltransferase	gi 27379583	Cytoplasmic	5.95/5.93	25182/27000	<i>B. diazoefficiens</i> USDA 110
Cell cycle control, cell division, chromosome partitioning							
86	<i>ftsZ</i>	Cell division protein FtsZ	gi 27381707	Periplasmic	5.21/5.40	62990/72000	<i>B. diazoefficiens</i> USDA 110
Signal transduction mechanisms							
88		Two-component response regulator OmpR	gi 27377311	Cytoplasmic	5.69/6.07	26233/29000	<i>B. diazoefficiens</i> USDA 110

Table 1 Identified proteins of *Bradyrhizobium diazoefficiens* CPAC 7 whole cell extract and classification according to COG (Continued)

Poorly characterized							
General function prediction only							
89	<i>ychF</i>	GTP-dependent nucleic acid-binding protein EngD	gi 27382550	Cytoplasmic	5.13/5.22	39493/42000	<i>B. diazoefficiens</i> USDA 110
90	<i>cobS</i>	Cobalt insertion protein	gi 383768898	Cytoplasmic	5.62/5.80	37304/39000	<i>Bradyrhizobium</i> sp. S23321
91		Dehydrogenase	gi 27378316	Cytoplasmic	5.92/6.32	33290/38000	<i>B. diazoefficiens</i> USDA 110
93	<i>cinA</i>	Competence-damage associated protein	gi 27380722	Cytoplasmic	5.29/5.51	26451/28000	<i>B. diazoefficiens</i> USDA 110
NO related COG							
96	<i>nusA</i>	Transcription elongation factor NusA	gi 27375896	Cytoplasmic	4.70/4.77	59333/67000	<i>B. diazoefficiens</i> USDA 110
97		ATP-dependent phosphoenolpyruvate carboxykinase	gi 398824719	Cytoplasmic	6.01/6.34	59235/62000	<i>Bradyrhizobium</i> sp. YR681
98		ABC transporter substrate-binding protein	gi 27380707	Periplasmic	6.93/6.00	59019/58000	<i>B. diazoefficiens</i> USDA 110
100	<i>tldD</i>	TldD protein	gi 27376279	Cytoplasmic	5.51/6.00	51101/51000	<i>B. diazoefficiens</i> USDA 110
102		ABC transporter substrate-binding protein	gi 27382938	Periplasmic	5.57/5.66	47774/42000	<i>B. diazoefficiens</i> USDA 110
104	<i>glpX</i>	Fructose 1,6-bisphosphatase II	gi 27379474	Cytoplasmic	5.68/6.06	35603/40000	<i>B. diazoefficiens</i> USDA 110
104		Substrate-binding protein	gi 27382959	Periplasmic	6.93/5.35	40839/41000	<i>B. diazoefficiens</i> USDA 110
105	<i>ilvC</i>	Ketol-acid reductoisomerase	gi 384216635	Cytoplasmic	6.09/6.70	37076/38000	<i>B. diazoefficiens</i> USDA 110
106		Dioxygenase	gi 27377910	Cytoplasmic	5.35/5.58	34110/35000	<i>B. diazoefficiens</i> USDA 110
111		ATP-dependent Clp protease proteolytic subunit	gi 338974245	Cytoplasmic	5.97/6.32	22404/26000	<i>B. diazoefficiens</i> USDA 110

*Theoretical and **Experimental.

Matched peptides masses and MS/MS combined results are available at the Additional file 1: Table S1.

Table 2 Functional prediction of hypothetical proteins identified in *Bradyrhizobium diazoefficiens* CPAC 7 whole protein extract based on protein sequences, conserved domains and motifs, protein-protein interactions and cellular locations

Spot ID	Hypothetical protein	NCBI ID	Cellular location	*T/**E pI	*T/**E MW	Predicted function
Metabolism						
Amino acid transport and metabolism						
25	Blr3064	gi 27378175	Cytoplasmic	5.64/6.00	50837/49000	Succinyl-diaminopimelate desuccinylase (COG/NCBI); Peptidase family M20 - dimerisation domain (Pfam); GO: Hydrolase activity (InterPro); Lysine biosynthesis EC:3.5.1.18 (KEGG).
30	Blr5678	gi 27380789	Periplasmic	5.84/5.11	33744/27000	L-aminopeptidase/D-esterase (COG); Peptidase family S58 (Pfam); GO:arginine biosynthetic process; DmpA/ArgJ- Like domains (InterPro).
Coenzyme transport and metabolism						
36	Blr3798	gi 27378909	Cytoplasmic	5.66/6.13	27452/29000	Demethylmenaquinone methyltransferase (COG); Methyltransf_6 (Pfam);GO: methyltransferase activity (UniProtKB); Ribonuclease E inhibitor RraA domain (InterPro).
37	Bll4565	gi 27379676	Cytoplasmic	5.3/5.59	24885/28000	Demethylmenaquinone methyltransferase (COG); Methyltransf_6 (Pfam);GO: methyltransferase activity (UniProtKB); Ribonuclease E inhibitor RraA domain (InterPro).
Information storage and processing						
Transcription						
72	Bll4752	gi 27379863	Cytoplasmic	4.41/4.43	27960/32000	Predicted transcriptional regulator containing the HTH domain (COG); Putative transcriptional regulators (Ypuh-like)(Pfam); Winged helix-turn-helix DNA-binding domain (InterPro).
Cellular processes and signaling						
Posttranslational modification, protein turnover, chaperones						
83	BJ6T_08050***	gi 354953419	Cytoplasmic	4.46/4.43	20515/28000	Thioredoxin-like proteins and domains(COG); Scaffold protein Nfu/NifU N terminal (Pfam); GO: Iron-sulfur cluster binding (InterPro).
Signal transduction mechanisms						
87	Blr2761	gi 27377872	Cytoplasmic	5.465.70	29257/33000	Universal stress protein UspA (COG); Universal stress protein family (Pfam); GO: response to stress (InterPro).
Poorly characterized						
General function prediction only						
92	Bll5663	gi 27380774	Cytoplasmic	4.92/5.08	33547/34000	MoxR-like ATPases (COG); ATPases Associated with diverse cellular Activities – AAA proteins (Pfam); GO: ATPase activity (InterProt).

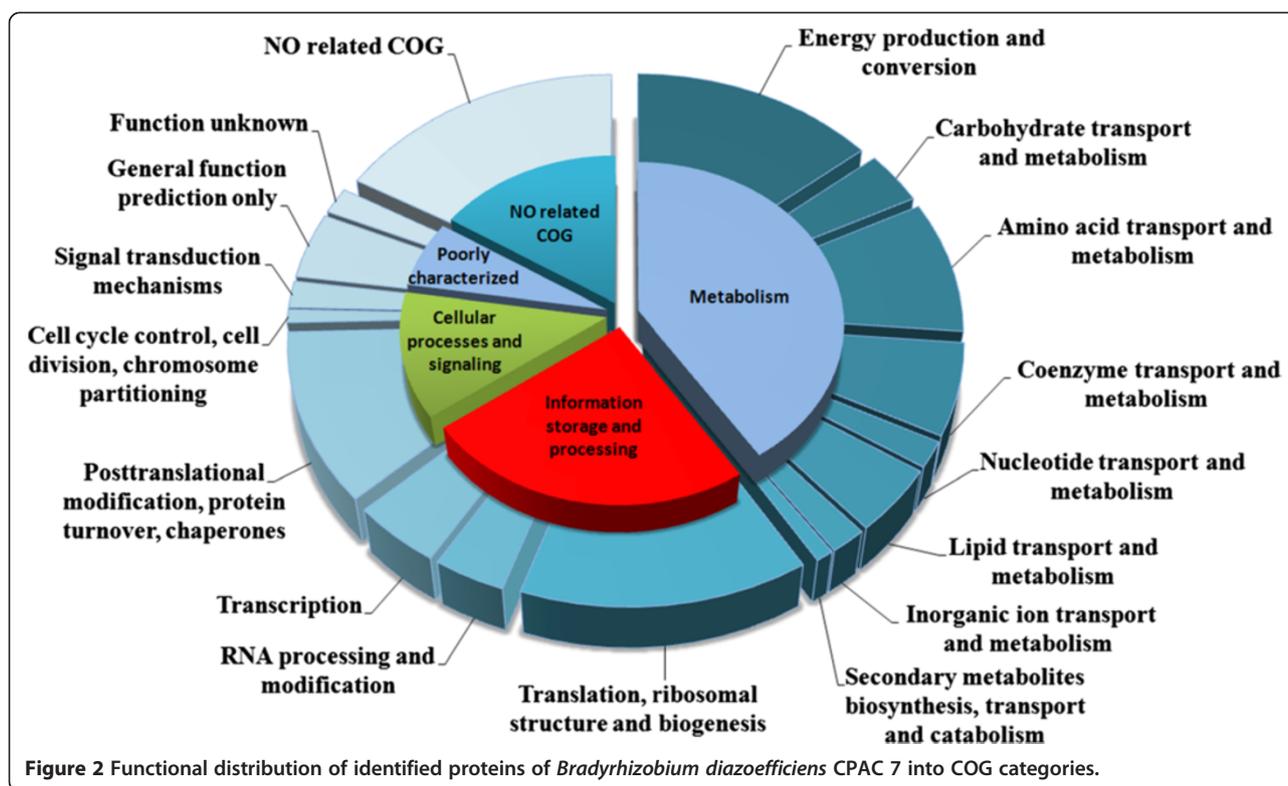
Table 2 Functional prediction of hypothetical proteins identified in *Bradyrhizobium diazoefficiens* CPAC 7 whole protein extract based on protein sequences, conserved domains and motifs, protein-protein interactions and cellular locations (Continued)

Function unknown						
94	Blr5067	gi 27380178	Cytoplasmic	5.25/5.29	24247/29000	Uncharacterized ACR (COG); putative metal binding site - region_name = "LabA (NCBI); NYN domain (Pfam).
95	Blr5067	gi 27380178	Cytoplasmic	5.93/5.96	16852/28000	Uncharacterized ACR (COG); putative metal binding site - region_name = "LabA (NCBI); NYN domain (Pfam).
NO related COG						
99	Bll0565	gi 27375676	Periplasmic	4.98/5.04	41554/57000	No related data
101	Blr7534	gi 27382645	Periplasmic	5.86/4.99	49518/45000	No related data
107	Bll5131	gi 27380242	Extracellular	7.68/5.80	34214/32000	Protein of unknown function DUF (Pfam/InterPro).
108	Blr2961	gi 27378072	Cytoplasmic	4.96/5.06	25510/29000	GO: catalytic activity/Metabolic process (InterPro); Protein of unknown function (Pfam).
109	Bll5307	gi 27380418	Periplasmic	5.55/5.70	14177/25000	No related data
110	Blr2191	gi 27377302	Cytoplasmic	6.58/6.12	25491/27000	Uncharacterized protein conserved in bacteria (DUF2328) (Pfam). KO: chpT histidine phosphotransferase (KEGG); Two-component system/His Kinase A (phospho-acceptor) domain (IMG).
112	Bll7551	gi 27382662	Cytoplasmic	5.95/5.92	27565/26000	No related data
113	Blr0227	gi 27375338	Periplasmic	5.17/5.29	22619/25000	PHB accumulation regulatory domain (Pfam);GO: Regulation of transcription/Transcription repressor activity (NCBI/InterPro).
114	Blr0227	gi 27375338	Periplasmic	5.17/6.40	22619/24000	PHB accumulation regulatory domain (Pfam);GO: Regulation of transcription/Transcription repressor activity (NCBI/InterPro).
115	Blr7436	gi 27382547	Cytoplasmic	4.82/4.90	15290/18000	No related data

[~]Theoretical and ^{**}Experimental.

***Best match with *Bradyrhizobium japonicum* USDA 6.

Matched peptides masses and MS/MS combined results are available at the Additional file 1: Table S1.



Also important to the production of an effective N_2 -fixing symbiosis, the exopolysaccharides (EPSs) play an essential role in the symbiotic interaction with compatible host plants [35,36]. We found one inositol monophosphatase (IMPase) (Table 1); this protein has been related with the regulation of EPS production, which, when mutated in *Rhizobium leguminosarum* bv. trifolii, resulted in defective-EPS production and a non- N_2 -fixing phenotype [37-39].

Several microbial factors, classified either as general or host-specific elicitors, are related to the induction of immune responses in plants [40,41]. General elicitors include flagellins, cold-shock proteins (CSPs), LCOs and LPSs. We identified the elongation factor Tu (EFTu), which also acts as elicitor and, in general, is conserved across multiple groups of bacteria, allowing plants to perceive and respond to an epitope common to many bacteria [42,43].

In response to general bacterial elicitors, plants have basal defense mechanisms that include increases in extracellular pH, ethylene production, and synthesis of reactive oxygen species (ROSs) [40]. Upon an initial “unfriendly” reception from the host plant, rhizobia must avoid host defenses, and elicit a successful environment to establish an effective N_2 -fixing symbiosis [44-46]. Among the several features presented by the bacteria to overcome plant defenses, GTP-binding protein TypA, which confers resistance to certain antimicrobial peptides and survival under stress conditions, has been recognized as the main contributor to

a successful interaction between *Sinorhizobium* (= *Ensifer*) *meliloti* and some *Medicago truncatula* lines [44]. A probable symbiotic function of this protein was also observed in *B. japonicum* CPAC 15 in response to the host flavonoid genistein [47], and now its constitutive expression has been detected in *B. diazoefficiens* CPAC 7.

The reduction of N_2 to ammonia by the nitrogenase complex can take place either by the rhizobial bacteroids inside the nodules, or in free-living rhizobia, including some *Bradyrhizobium* strains. In both cases, a finely balanced regulation of oxygen availability is required, since rhizobia are aerobic and need oxygen, whereas the element can denature nitrogenase. Inside the nodule, the ideal oxygen environment is reached by the participation of multiple factors, including the synthesis of heme compounds by the rhizobia [48,49]. Proteins HemB and HemH, identified in CPAC 7 (Table 1), catalyze two important steps in heme synthesis and are essential for the *Bradyrhizobium*/soybean symbiosis, since mutants defective in these genes generate a microaerobic condition with poorly developed nodules that are inefficient in fixing N_2 [48,50].

A protein related to amino acid metabolism and also key in N_2 fixation is the glutamine synthetase I (GS I), which was identified in our proteomic map (Table 1). The role of GS I in the regulation of nitrogenase has been highlighted by studies with *Rhizobium* sp. mutants, resulting in defective ability to derepress the enzyme, both *in vitro* [51] and in symbiotic conditions [52].

In our study, we also detected the constitutive expression of a sigma-54 modulation protein, which fitted in the translation, ribosomal structure and biogenesis functional category (Table 1). Similarly to NifA, this protein participates in controlling the expression of sigma-54 (RpoN, NtrA), which, in turn, helps initiate the transcription of genes encoding proteins for diverse cell functions [53,54]. Among several roles, RpoN is directly involved in the N₂-fixation process, being required for control of major N₂-fixation genes, as *nifHDK*, the products of which constitute the nitrogenase complex and accessory proteins [55]. RpoN is also related to free-living metabolic pathways, as demonstrated when rhizobia *rpoN*-mutants showed, beyond defects in symbiotic N₂-fixation [56], alterations in free-living nitrate assimilation [57]. We may suppose that the constitutive sigma-54 modulation protein may be important in mediating adaptations to changing environmental conditions, both in free-living and in symbiotic conditions.

The expression of the key protein NifH (nitrogenase iron protein, or component II) in our study may be related to the presence of sigma-54 modulation protein, which was shown to be a regulator of the sigma-54 expression. To support this, results show that the knockout of sigma-54 transcriptional factor in *B. japonicum* leads to strong pleiotropic effects, including the absence of NifH [57] and the abolishment of symbiotic N₂-fixation ability [55]. The detection of NifH in our proteomic study might correlate with reports of expression of nitrogenase in *Bradyrhizobium* strains under free-living conditions; however, measurements of nitrogenase activity under these conditions are difficult and require a fine adjustment of the oxygen concentration [58-61].

Stress-tolerance proteins

In tropical regions, crops and soil microorganisms are frequently exposed to stressful conditions, in particular high temperatures, salinity and soil acidity [6,62,63]. Therefore, in addition to high efficiency of N₂ fixation, commercial rhizobial strains must be tolerant of such adverse factors. Indeed, several soybean bradyrhizobia have been extensively studied to characterize their tolerances of salt [64], desiccation [65], antibiotics [66], acidity [67], among other stresses. Now, in *B. diazoefficiens* CPAC 7 we produced evidence of several molecular determinants related to the ability to overcome adverse conditions, including chaperonins and other proteins, such as Clp proteases and transcription-elongation factors, with roles in cell defense.

ATP-dependent Clp proteases participate in diverse cell processes, including rapid adaptive responses of bacteria to environmental changes [29] and to stressful conditions [30], and ClpB and ClpX were detected in CPAC 7. These two proteins have the properties associated with molecular chaperones, such as preventing the aggregation of denatured proteins and, in some cases, refolding them [68].

These properties are particularly important under stress conditions that exacerbate the occurrence of protein denaturation, and ClpB and ClpX help to ensure a fast return to the pre-stressed state, maintaining cell homeostasis [69,70].

Bradyrhizobium is acid-tolerant, it grows at pH 4.5, over 30% of the strains are capable of growing at pH 4.0 and a few are tolerant of pH 3.5 [71]. The proteome of *B. diazoefficiens* USDA 110, when studied at pH 4.7 [72], revealed differential expression of several proteins, eight of which—spots 5, 8, 15, 49, 70, 81, 82, and 91—were constitutively expressed in the proteome of CPAC 7 (Table 1).

Another limiting factor for rhizobia, and also for the symbiosis, is high soil temperature, which can often exceed 40°C in the tropics and limit the success of inoculants [6,8,62]. Rhizobia, similar to most organisms [73], make use of molecular chaperones to tolerate high temperatures, including heat-shock proteins (HSPs) DnaK, GrpE, GroEL and HtpG, which were identified in our proteomic reference map (Table 1). DnaK and GrpE comprise a versatile chaperone system [74] that, together with GroEL and HtpG, play a critical role in thermotolerance, routinely rescuing the majority of the proteins denatured [69,75].

Still associated with thermotolerance, it is worth mentioning the identification of three elongation factors (Ef-Tu, Ef-Ts and Ef-G) and two ribosomal proteins (30S and 50S) expressed in CPAC 7. Besides their main function of ensuring gene expression, elongation factors can also act as chaperones [26,27]. This secondary role has been demonstrated recently at the proteomic level in *Rhizobium tropici* strain PRF 81 (now reclassified as *Rhizobium freirei*) under high-temperature stress [75]. In *B. japonicum*, 30S and 50S ribosomal proteins may be involved in heat-stress defense, once they were up-regulated at 43°C [76]. Considering this finding, those authors hypothesized that ribosomes may act as sensors of heat shock in *B. japonicum*. A similar mechanism has been suggested in *E. coli*, in which ribosomes seem to be the primary sensor of conditions that evoke heat-shock responses [77].

Oxidative stress is frequently caused by cell exposure to reactive oxygen species (ROs), such as superoxide anion (O₂⁻) and hydrogen peroxide (H₂O₂). ROs are byproducts of normal metabolic processes and, if not properly detoxified, they become toxic. Oxidative stress also occurs by cell exposure to external ROs, which in bacteria such as rhizobia may take place during interactions with other microorganisms or eukaryotic hosts. Therefore, tolerating and overcoming oxidative stress is critical to bacteria viability as well as for the establishment of a successful symbiotic infection [32,78].

Of the proteins identified in CPAC 7, at least five have already been reported as showing antioxidant activity (spots 16, 50, 68, 82 and 83, Table 1). Among them, ferredoxin-NADP⁺ reductase (FRN) has been reported to

overcome the harmful effects of ROSs on DNA replication [79]. Those authors emphasized the importance of FRN to cell protection against oxidative damage by comparing its role with those of superoxide dismutases, a group of proteins well known for mitigating damage caused by ROSs.

Salinity leads to loss of intracellular water, resulting in osmotic disturbances that can influence a range of metabolic activities [80]. Indeed, several negative effects in rhizobia-plant symbioses have been attributed to salinity; e.g. in growth and survival of rhizobia in soil, in root colonization and in nodule development [81].

Mutational studies with *S. meliloti* allowed the identification of multiple genes involved in salt tolerance, including trigger factor (*tig*) [82]. When this gene was absent, *Sinorhizobium* showed reduced ability to grow in LB with high salt concentrations, as well inability to compete against the wild-type for nodule occupancy [82]. Given these results, we suggest that *tig* may also contribute to competitiveness and to saprophytic competence under environmentally stressful conditions, as reported for CPAC 7 [7,8,10,11,22].

Classified in the information storage and processing COG category, the transcription elongation factor GreA has been recognized as a general stress protein (Gsp) induced in response to various environmental conditions. Additionally to the transcription elongation activity, its role in acid-, salt-, and cold-stress responses in *Streptococcus mutans* [83], *S. meliloti* [84] and *R. tropici* [85] has been reported. Constitutive expression of GreA was previously reported in a *B. japonicum* CPAC 15 proteomic assay [23]. Mutation of *greA* in *R. tropici* impaired the establishment of an effective symbiosis as a result of the altered ability to adapt to hyperosmosis and salt stress [85], highlighting the importance of this protein in overcoming adverse conditions during symbiosis establishment.

Altogether, these bacterial defense mechanisms are crucial to survival in the soil and to symbiosis establishment in the tropics, where rhizobia are commonly exposed to high soil temperatures, acid pH and saline conditions [3,8].

Hypothetical proteins: function prediction with bioinformatics tools

In several genome projects, portions of the annotated sequences have been classified in “hypothetical”, “conserved hypothetical” or “of unknown function” categories [86,87]. These denominations are used when the existence of a gene is supported only by prediction of gene-finding software, and they do not show significant homology to any characterized gene [23]. With *B. diazoefficiens* USDA 110, these proteins were abundant, and, of the 8,317 protein-coding genes predicted, 30% showed similarity to hypothetical genes, whereas 18% showed no similarity to any registered gene [13]; in the genome of CPAC 7, both categories comprised 50% of all predicted genes [20]. A still modest improvement in annotation of hypothetical

proteins has been achieved with proteomic studies [23,87]. For example, the reference map of *B. japonicum* CPAC 15 contributed to the assignment of 26 hypothetical proteins by using bioinformatics tools [23].

In our study, 20 proteins were classified as hypothetical/conserved hypothetical or unknown, and by using bioinformatics tools, we were able to attribute probable functions to half of them (Table 2). Two proteins (Blr3064 and Blr5678) were assigned to the amino acid transport and metabolism COG category. The first one shows hydrolase activity and, according to the KEGG database, participates in lysine biosynthesis. Blr5678 is probably related to the arginine biosynthetic process, since it has an ArgJ-like domain.

Classified according to the COGnitor in the coenzyme transport and metabolism functional category, proteins Blr3798 and Bll4565, represented by spots 36 and 37 (Figure 1, Table 2) presented similar information in the databanks searched in our study. Both were annotated as demethylmenaquinone methyltransferase by COG and exhibited a ribonuclease E inhibitor RraA domain. The inhibitory activity of this domain was recently described in *E. coli* as having a regulatory function in gene expression, since the interactions of RraA with RNase E affect the composition of the RNA-degradosome, modulating its activity [88,89].

The cytoplasmic protein Bll4752, which belongs to the transcriptional COG group is predicted as a transcriptional regulator protein that contains a Ypuh-like helix-turn-helix domain (Table 2). This protein probably plays a role in chromosomal partitioning during cell division.

BJ6T_08050, classified in the posttranslational modification, protein turnover, chaperones functional category, shows 97% of similarity with Bll0800 of *B. diazoefficiens* USDA 110. After comparison of databases, this hypothetical protein was assigned as a thioredoxin-like protein involved in iron transport (Table 2); however, it may be related to reactivation of proteins damaged by oxidative stress [90]. Also potentially related to cell defense, Blr2761 is closely similar to universal stress protein UspA, expression of which is known to be enhanced when cells are exposed to adverse conditions [91].

Two other proteins constitutively expressed in strain CPAC 7—Bll5663 and Blr5067—remained annotated as hypothetical in the NCBI database (Table 2). Bll5663 is a member of the MoxR family of AAA + ATPases, widespread among bacteria. Associated with diverse cellular activities, MoxR ATPases display a chaperone-like function and have been found to be important modulators of multiple-stress-response pathways in various organisms, including *R. leguminosarum* [92,93]. The other protein, Blr5067, was also expressed in bacteroids of *B. diazoefficiens* USDA 110, and is still poorly characterized [14]; it could be a LabA-like protein with a putative

metal-binding site. This family of proteins has been studied in cyanobacteria and reported to affect both gene expression and cellular metabolic state [94,95]. Nevertheless, the biological role of Blr5067 in *Bradyrhizobium* remains to be determined.

Genistein effect on the expression of hypothetical protein-coding genes

Ten out of the 20 hypothetical proteins identified in our proteomic study did not fit into any of the functional categories of COG, being assigned as “NO related COG” (Table 2). The lack of homology of these sequences with those of known proteins, combined with their detection at the proteomic level, suggest that they could be interesting subjects of study, possibly providing new information and deeper understanding of the organism. We analyzed the relative expression patterns of nine hypothetical protein-coding genes (two spots—113 and 114—showed similarity with the same protein Blr0227) in response to genistein. The localization of the genes encoding these proteins in the genome of *B. diazoefficiens* CPAC 7 is shown in Additional file 1: Table S1. Identified as one of the main components present in soybean root extract [96], genistein induces the expression of *nod* genes in *Bradyrhizobium* [16,97]. In addition, it has been shown that flavonoids, such as genistein, can induce the expression of other genes besides *nod* genes [15,47].

The product of *nodC* directs the synthesis of the backbone of lipochitin oligosaccharides (LCOs), also called Nod factors, which are essential for the nodulation process [97-100]. In our study, the *nodC* gene of CPAC 7 was used as a positive control in the RT-qPCR analysis, and up-regulation was confirmed; in addition, six of the nine hypothetical protein-encoding genes were significantly up-regulated (Figure 3). Of these, bl10565 and blr2961

protein-coding genes showed the highest genistein-induction effect, and up-regulation was observed also for blr7534, bl15307, blr2191 and blr0227.

Blr2191 is a ChpT histidine phosphotransferase, well characterized in *Caulobacter crescentus*, a model organism in cell-cycle studies [101]; in this bacterium, the protein controls, via phosphorylation, the activity of the master cell-cycle regulator CtrA [102,103]. The role of histidine phosphotransferases in the cell cycle of *B. diazoefficiens* is still to be elucidated; however, with the results from our study—showing its induction by genistein—we may suppose that it affects the growth rate of *B. diazoefficiens* in a genistein-enriched environment, such as the soybean rhizosphere [104].

Protein Blr0227 has also been identified in *B. japonicum* under neutral [23] and acidic conditions [72]. A potential implication of this protein in rhizobial competitiveness might exist, since it is up-regulated in response to the flavonoid daidzein in *B. japonicum* strain 4534, a strain highly competitive for nodulation, but not in the poorly competitive *B. japonicum* strain 4222 [105]; it is noteworthy that CPAC 7 is also competitive. Blr0227 has a polyhydroxybutyrate (PHB)-accumulation regulatory domain and it is known that in *B. japonicum* large amounts of carbon are directed to the synthesis of storage compounds, especially PHB [106]. Interestingly, the PHB biosynthesis also seems to be associated with rhizosphere competitiveness, since *B. japonicum* mutants defective in PHB synthesis show reduced competitiveness [107].

Three out of the nine proteins of CPAC 7 analyzed were not significantly induced by genistein, bl15131, bl17551 and blr7436 (Figure 3, Additional file 2: Figure S1), confirming results reported with *B. japonicum* strain CPAC 15 [47]. They may play other roles, e.g. bl17551 was one of the up-regulated proteins in bacteroids of *B. japonicum*

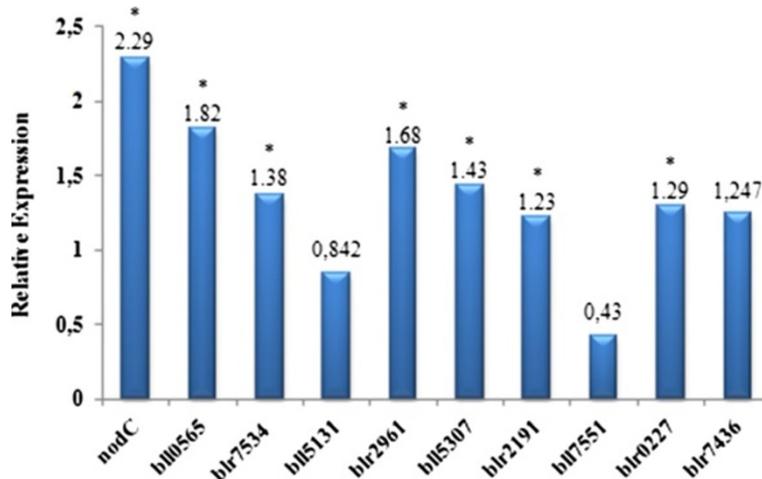


Figure 3 Expression levels of hypothetical proteins coding genes of *Bradyrhizobium diazoefficiens* strain CPAC 7 after growth to exponential phase in the presence of genistein. Relative expression was determined by REST2009 software.

USDA 110 [108], suggesting a role in latter steps in the development of the symbiosis.

Conclusions

CPAC 7 is an agronomically important strain used in commercial inoculants for application to soybean crops in Brazil and in other South American countries; it presents high N₂-fixation efficiency, and adaptability to tropical conditions [7,8,11,109]. Here we provide the first proteomic map for this bacterium, revealing molecular determinants of distinct steps in the establishment and functioning of the symbiotic biological N₂-fixation process. We also report the constitutive expression of proteins such as DnaK, ferredoxin-NADP⁺ reductase (FRN) and trigger factor (*tig*) related to cell protection against heat-, oxidative- and salt stress conditions, that should contribute to bacterial survival and symbiotic functioning under adverse environmental conditions common in the tropics. In general, no function can be attributed to more than one third of the putative genes in bacterial genomes. With the approach taken in our study—including proteomics, use of bioinformatics tools and transcriptomic assays—it was possible to obtain information about several hypothetical genes/proteins of *B. diazoefficiens*, revealing interesting information, with an emphasis on genistein-induced genes, that deserve further study to confirm their roles in the soybean root-nodule symbiosis.

Methods

Strain and culture conditions

Bradyrhizobium diazoefficiens strain CPAC 7 (=SEMIA 5080, =CNPSo 6), a natural variant of CB 1809 (=USDA 136, a subculture of USDA 122) [10,110] is used in commercial inoculants in Brazil since 1992 [7]. Information about morpho-physiologic, genetic and symbiotic properties is available elsewhere [10,11,111-113]. CPAC 7 is deposited at the “Diazotrophic and Plant Growth Promoting Bacteria Culture Collection” of Embrapa Soja (WFCC Collection # 1213, WDCC Collection # 1054).

The strain was pre-cultured in 10-mL aliquots of arabinose-gluconate (AG) medium [114], at 80 rpm and 28°C, in the dark. For the proteomic experiment, pre-cultures were transferred to Erlenmeyer flasks containing 200 mL of the same medium and were grown under the same conditions as for the pre-cultures until the exponential phase (O.D. of 0.6 at 600 nm). Low agitation (80 rpm) was employed to minimize the production of extracellular polysaccharides, which can interfere with the 2-D gel electrophoresis.

For the reverse transcription quantitative PCR (RT-qPCR), pre-cultures were transferred to Erlenmeyer flasks containing 100 mL of AG medium. Bacteria were grown to the exponential phase under two treatment conditions:

induced or not with genistein (5 μM, final concentration) dissolved in methanol, added as 50 μL per 100 mL of culture [15]; to the non-induced cultures, the same amount of methanol was added. In both proteomics and qPCR experiments, bacteria were grown in triplicates for each treatment.

Whole-cell protein extraction

Cultures were centrifuged at 5,000 g, at 4°C and cells were carefully washed with a solution containing 3 mM KCl; 1.5 mM KH₂PO₄; 68 mM NaCl; and 9 mM NaH₂PO₄. Washed cells were resuspended in 600 μL of a buffer containing 10 mM Tris-HCl pH 8.0; 1.5 mM MgCl₂; 10 mM KCl; 0.5 mM DTT; and 0.5 mM PMSF. Aliquots of 150 μL were stored in ultrafreezer (-80°C) until the analyses.

For total protein extraction, aliquots were resuspended in lysis buffer (9.5 M urea; 2% CHAPS; 0.8% v/v Pharmalyte 4-7; and 1% DTT), and submitted to forty cycles of freezing in liquid N₂ and thawing at 37°C, as described before [115]. The lysates were separated from particulate material by centrifugation at 14,000 g for 90 min, at 4°C. Protein extract was washed with phenol and the concentration was determined by NanoDrop 1000 Spectrophotometer V3.7 (Thermo Scientific).

Two-dimensional gel electrophoresis and visualization

For isoelectric focusing (IEF), 300 μg of protein extract were dissolved with DeStreak buffer (GE Healthcare) and 2% v/v IPGphor to a final volume of 250 μL. IPG-strips (pH 4-7, 13 cm, GE Healthcare) were rehydrated overnight with the protein solution and covered with Cover Fluid (GE Healthcare). Loaded strips were submitted to isoelectric focalization in an Ettan IPGphor IEF system (GE Healthcare) for 1 h at 200 V, 1 h at 500 V, a gradient step to 1,000 V for 1 h, a gradient step to 8,000 V for 2 h 30 min, and fixed at 8,000 V for 1 h 30 min. The final Vh was fixed at 24,800. Prior to second dimension, strips were equilibrated first for 20 min in 5 mL of an equilibration buffer (50 mM Tris-HCl pH 8.8; 6 M urea; 30% v/v glycerol; 2% w/v SDS; and 0.2% v/v of a 1% solution of bromophenol blue) supplemented with 50 mg DTT and then in TE buffer with 175 mg of iodoacetamine, also for 20 min.

The second dimension electrophoresis was performed in a 12% polyacrylamide gel in a Ruby SE 600 vertical electrophoresis system (GE Healthcare). The run was carried out for 30 min at 15 mA/gel and 240 min at 30 mA/gel, using the Low Molecular Weight Calibration Kit for SDS Electrophoresis (Amersham Biosciences) as standard. Similarly to the protein extraction step, both dimensions of gel electrophoresis were run in triplicate. Gels were fixed overnight with an ethanol-acetic acid solution before being stained with Coomassie Blue PhastGel™ R-350 (GE Healthcare) and were then scanned (ImageScanner LabScan v5.0).

Gel image analysis and sample preparation to mass spectrometry

Protein spots were automatically detected in the high-resolution digitized gel images and analyzed by Image Master 2D Platinum v 5.0 software (GE Healthcare). Well defined spots were manually selected, excised and processed as previously described [116]. Digestion was achieved with trypsin (Gold Mass Spectrometry Grade, Promega, Madison, WI) at 37°C, overnight.

Peptides from digested proteins were mixed with saturated solution of α -cyano- 4-hydroxy-cinnamic acid (HCCA) in 50% acetonitrile, 0.1% trifluoroacetic acid (TFA). The mixture was spotted onto a MALDI (Matrix Assisted Laser Desorption Ionization) target plate and allowed to crystallize at room temperature. The same procedure was used for the standard peptide calibration mixture I (Bruker Daltonics). For mass spectra acquisition, a MALDI-TOF-TOF (MALDI-time-of-flight in tandem) UltraFlex III mass spectrometer (Bruker Daltonics) was operated in the reflector for MALDI-TOF MS peptide mass fingerprint (PMF) and in the "LIFT"[™] mode for MALDI-TOF-TOF MS/MS fragmentation experiments, on fully manual mode using FlexControl software v. 2.2. To process the data obtained, Flex Analysis v.3.0 software (Bruker Daltonics, Billerica, MA) was employed.

Protein identification

PMFs and MS/MS ion spectra generated were searched against the public database NCBI nr (National Center for Biotechnology Information non-redundant), using Mascot software search engine v. 2.3 (<http://www.matrixscience.com/>). For protein searches, performed in the Proteobacteria taxonomic group, monoisotopic masses were used, considering a peptide tolerance of 150 ppm and allowance of one missed cleavage. When MS/MS was carried out, a tolerance of 0.3 Da was acceptable. Carbamidomethylation of cysteine and oxidation of methionine were considered fixed and variable modifications, respectively.

Identifications were validated only when the Mowse (molecular weight search) score was significant. Searches on the Decoy database (Mascot) were done and both decoy score and false discovery rates were considered for the identification. The spectrometry datasets are available at the Additional file 1: Table S1.

In silico protein characterization

A set of bioinformatics tools was used to improve the characterization of the proteins. The proteins were fitted into COG (Clusters of Orthologous Groups) categories according to their functional inference, using the COGNITOR program (<http://www.ncbi.nlm.nih.gov/COG>) [117]. Software packages PSORT-B [118] and PSLpred [119] were used for the prediction of subcellular localization.

To search for putative roles of the hypothetical proteins, a package of bioinformatics tools was applied [87]. SignalP [120] was employed for the prediction of signal peptides. To determine the protein family and domains we used Pfam [121] and InterPro [122]. MicrobesOnline (<http://www.microbesonline.org>) [123], a suite of web-based comparative tools, and the Integrated Microbial Genomes system (<http://img.jgi.doe.gov>) [124] were also searched. Finally, the prediction of physical and functional protein interactions was carried out with STRING 9.1 (<http://string-db.org/>) [125].

RNA extraction and primers design

Cells from 35 mL of the control and the genistein induced bacterial cultures (item 2.1) were centrifuged (8,000 *g* for 10 min at 4°C) and the pellet was resuspended in 280 μ L of lysis buffer, consisted of 250 μ L of TE (10 mM Tris, adjusted to pH 8.0 with HCl; 1 mM EDTA), 10 μ L of lysozyme (5 mg/mL) and 20 μ L of 10% SDS solution (w/v). After resuspended, the mixture was incubated at 37°C for 5 min to achieve an efficient cell disruption. Lysates were then centrifuged (8,000 *g* for 10 min at 4°C) and the supernatants were transferred to new 2-mL tubes and homogenized with 1 mL of TRIzol[®] reagent (Life Technologies). This new mixture was centrifuged and the superior phase was transferred to another tube. After a wash step with chloroform, RNA was precipitated by adding 500 μ L of cold isopropanol, purified with RNeasy Mini Kit (Qiagen) and quantified by NanoDrop ND-1000 (NanoDrop-Technologies, Inc.). The RNA was assessed in a 1% (w/v) agarose gel.

Primers were designed using PrimerExpress 3.0 (Applied Biosystems/Life Technologies, Grand Island, NY, USA) targeting an amplicon size of 50–200 bp. The primer sequences were searched against the *B. diazoefficiens* strain USDA 110 genome (<http://www.ncbi.nlm.nih.gov/genome/18384>) to verify their specificity. Primer sequences and amplification efficiency rates are shown in Additional file 3: Table S2.

Extracted RNAs were submitted to DNase treatment (Invitrogen/Life Technologies, Grand Island, NY, USA), and high quality total RNA was used to synthesize cDNA strands (Superscript II First Strand Synthesis, Invitrogen/Life Technologies, Grand Island, NY, USA).

Relative gene expression analysis by RT-qPCR

After carrying out the amplification to determine the primers efficiency rate, nine hypothetical genes were amplified by RT-qPCR using a 7500 RT-qPCR Thermocycler (Applied Biosystems/Life Technologies, Grand Island, NY, USA) with the following manufacturer's instructions: 50°C for 2 min, 95°C for 10 min, 45 cycles at 95°C for 2 min, 60°C for 30 s and 72°C for 30 s, in 45 cycles. The 16S rRNA gene was used as endogenous control (Additional file 3: Table S2).

Rest2009 software package [126] was used to evaluate the data by providing a robust statistical analysis (Additional file 2: Figure S1). The normalization of cycle threshold (Ct) of RT-qPCR amplifications was performed based on the selected endogenous gene (16S rRNA). The genistein responsive gene *nodC* [127] was used as positive control.

Additional file

Additional file 1: Table S1. Complementary information about protein identifications. All searches were performed with Mascot software v. 2.3 (<http://www.matrixscience.com/>) against the public database NCBI nr (National Center for Biotechnology Information non-redundant). *Identified by MS; **Identified by MS/MS.

Additional file 2: Figure S1. Localization in the genome of *B. diazoefficiens* CPAC 15 of the genes coding hypothetical proteins used in our study.

Additional file 3: Table S2. Statistical results of RT-qPCR data provided by Rest2009 software package.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Conceived and designed the experiments: MH, LVGT. Performed the experiments: DFG, JSSB, AAPR, LPS, CB. Analyzed the data: DFG, JSSB, AAPR, LPS, CB, LVGT, MH. Contributed reagents/materials/analysis tools: MH, LPS, CB. Wrote the paper: DFG, JSSB, LPS, MH. All authors read and approved the final manuscript.

Acknowledgements

The study was partially supported by CNPq (National Council for Scientific and Technological Development), Project Repensa (562008/2010-1) and Science without Borders (400205/2012-5). Authors acknowledge Dr. Allan R.J. Eaglesham for suggestions on the manuscript and English review. D.F.G. acknowledges a PhD fellowship from Project Repensa and J.S.S.B. was and A.A.P.R. postdoc fellowships from CNPq. M.H. is also a research fellow from CNPq. Approved for publication by the Editorial Board of Embrapa Soja as manuscript 12/2014.

Author details

¹Embrapa Soja, Embrapa Soja, C.P. 231, 86001-970 Londrina, Paraná, Brazil. ²Departamento de Genética, Universidade Federal do Paraná, C.P. 19031, 81531-900 Curitiba, PR, Brazil. ³Present address: Departamento de Biologia Estrutural, Molecular e Genética, Universidade Estadual de Ponta Grossa, C.P. 6001, 84030-900 Ponta Grossa, PR, Brazil. ⁴Embrapa Recursos Genéticos e Biotecnologia, PqEB s/n°, 70770-901 Brasília, DF, Brazil.

Received: 5 May 2014 Accepted: 25 July 2014

Published: 3 August 2014

References

1. Saravanan VS, Madhaiyan M, Osborne J, Thangaraju M, Sa TM: Ecological occurrence of *Gluconacetobacter diazotrophicus* and nitrogen-fixing Acetobacteraceae members: their possible role in plant growth promotion. *Microb Ecol* 2008, **55**:130–140.
2. Graham PH, Vance CP: Legumes: importance and constraints to greater use. *Plant Physiol* 2003, **13**:872–877.
3. Ormeño-Orrillo E, Menna P, Almeida LGP, Ollero FJ, Nicolás MF, Rodrigues EP, Nakatani AS, Batista JSS, Chueire LMO, Souza RC, Vasconcelos ATR, Megias M, Hungria M, Martínez-Romero E: Genomic basis of broad host range and environmental adaptability of *Rhizobium tropici* CIAT 899 and *Rhizobium* sp. PRF 81 which are used in inoculants for common bean (*Phaseolus vulgaris* L.). *BMC Genomics* 2012, **13**:735.
4. Vance CP: Update on the state of nitrogen and phosphorus nutrition symbiotic nitrogen fixation and phosphorus acquisition: plant nutrition in a world of declining renewable resources. *Plant Physiol* 2001, **127**:390–397.
5. Subramanian S, Smith DL: A proteomics approach to study *Soybean* and its symbiont *Bradyrhizobium japonicum* – A Review. In *Agricultural and Biological Sciences, A Comprehensive Survey of International Soybean Research - Genetics, Physiology, Agronomy and Nitrogen Relationships*. CC BY 3.0 license. Edited by Board JE. 2013:3–30.
6. Hungria M, Andrade DDS, Chueire LMO, Probanza A, Gutierrez FJM, Megias M: Isolation and characterization of new efficient and competitive bean (*Phaseolus vulgaris* L.) rhizobia from Brazil. *Soil Biol Bioche* 2000, **32**:1515–1528.
7. Hungria M, Franchini JC, Campo RJ, Crispino CC, Moraes JZ, Sibaldelli RNR, Mendes IC, Arihara J: Nitrogen nutrition of soybean in Brazil: contributions of biological N₂ fixation and N fertilizer to grain yield. *Can J Plant Sci* 2006, **86**:927–939.
8. Hungria M, Vargas MAT: Environmental factors affecting N₂ fixation in grain legumes in the tropics, with an emphasis on Brazil. *Field Crops Res* 2000, **65**:151–164.
9. Hungria M, Mendes IC: Nitrogen fixation with soybean: the perfect symbiosis? In *Biol nitrogen Fixat*. Edited by de Bruijn F. Hoboken: Wiley-Blackwell. New Jersey, Wiley Publisher; 2014.
10. Delamuta JRM, Ribeiro RA, Ormeño-Orrillo E, Melo IS, Martínez-Romero E, Hungria M: Polyphasic evidence supporting the reclassification of *Bradyrhizobium japonicum* group Ia strains as *Bradyrhizobium diazoefficiens* sp. nov. *Int J Syst Evol Microbiol* 2013, **63**:3342–3351.
11. Hungria M, Boddey LH, Santos MA, Vargas MAT: Nitrogen fixation capacity and nodule occupancy by *Bradyrhizobium japonicum* and *B. elkanii* strains. *Biol Fertil Soils* 1998, **27**:393–399.
12. Ferreira MC, Hungria M: Recovery of soybean inoculant strains from uncropped soils in Brazil. *Field Crops Res* 2002, **79**:139–152.
13. Kaneko T, Nakamura Y, Sato S, Minamisawa K, Uchiyama T, Sasamoto S, Watanabe A, Idesawa K, Iriguchi M, Kawashima K, Kohara M, Matsumoto M, Shimpo S, Tsuruoka H, Wada T, Yamada M, Tabata S: Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res* 2002, **9**:225–256.
14. Sarma AD, Emerich DW: Global protein expression pattern of *Bradyrhizobium japonicum* bacteroids: a prelude to functional proteomics. *Proteomics* 2005, **5**:4170–4184.
15. Süß C, Hempel J, Zehner S, Krause A, Patschkowski T, Göttfert M: Identification of genistein-inducible and type III-secreted proteins of *Bradyrhizobium japonicum*. *J Biotechnol* 2006, **126**:69–77.
16. Lang K, Lindemann A, Hauser F, Göttfert M: The genistein stimulon of *Bradyrhizobium japonicum*. *Mol Genet Genomics* 2008, **279**:203–211.
17. Wei M, Yokoyama T, Minamisawa K, Mitsui H, Itakura M, Kaneko T, Tabata S, Saeki K, Omori H, Tajima S, Uchiyama T, Abe M, Ohwada T: Soybean seed extracts preferentially express genomic loci of *Bradyrhizobium japonicum* in the initial interaction with soybean, *Glycine max* (L.) Merr. *DNA Res* 2008, **15**:201–214.
18. Hempel J, Zehner S, Göttfert M, Patschkowski T: Analysis of the secretome of the soybean symbiont *Bradyrhizobium japonicum*. *J Biotechnol* 2009, **140**:51–58.
19. Delmotte N, Ahrens CH, Knief C, Qeli E, Koch M, Fischer H-M, Vorholt JA, Hennecke H, Pessi G: An integrated proteomics and transcriptomics reference data set provides new insights into the *Bradyrhizobium japonicum* bacteroid metabolism in soybean root nodules. *Proteomics* 2010, **10**:1391–1400.
20. Siqueira AF, Ormeño-Orrillo E, Souza RC, Rodrigues EP, Almeida LGP, Barcellos FG, Batista JSS, Nakatani AS, Martínez-Romero E, Vasconcelos ATR, Hungria M: Comparative genomics of *Bradyrhizobium japonicum* CPAC 15 and *Bradyrhizobium diazoefficiens* CPAC 7: elite model strains for understanding symbiotic performance with soybean. *BMC Genomics* 2014, **15**:420.
21. Gupta N, Tanner S, Jaitly N, Adkins JN, Lipton M, Edwards R, Romine M, Osterman A, Bafna V, Smith RD, Pevzner PA: Whole proteome analysis of post-translational modifications: Applications of mass-spectrometry for proteogenomic annotation. *Gen Res* 2007, **17**:1362–1377.
22. Ansong C, Purvine SO, Adkins JN, Lipton MS, Smith RD: Proteogenomics: needs and roles to be filled by proteomics in genome annotation. *Brief Funct Genomic Proteomic* 2008, **7**:50–62.
23. Batista JSS, Torres AR, Hungria M: Towards a two-dimensional proteomic reference map of *Bradyrhizobium japonicum* CPAC 15: spotlighting "hypothetical proteins". *Proteomics* 2010, **10**:3176–3189.
24. Yadav AS: Auxotrophy in rhizobia revisited. *Indian J Microbiol* 2007, **47**:279–288.

25. Santos MF, Pádua VLM, Nogueira EM, Hemery AS, Domont GB: Proteome of *Gluconacetobacter diazotrophicus* co-cultivated with sugarcane plantlets. *J Proteomics* 2010, **73**:917–931.
26. Caldas TD, El Yaagoubi A, Richarme G: Chaperone properties of bacterial elongation factor EF-Tu. *J Biol Chem* 1998, **273**:11478–11482.
27. Caldas T, Laalami S, Richarme G: Chaperone properties of bacterial elongation factor EF-G and initiation factor IF2. *J Biol Chem* 2000, **275**:855–860.
28. Wagner MA, Zahrl D, Rieser G, Koraimann G: Growth phase- and cell division-dependent activation and inactivation of the σ 32 regulon in *Escherichia coli*. *J Bacteriol* 2009, **191**:1695–1702.
29. Gaillot O, Bregenholt S, Jaubert F, Santo JPD, Berche P: Stress-induced *clpP* serine protease of *Listeria monocytogenes* is essential for induction of listeriolysin α -dependent protective immunity. *Infect Immun* 2001, **69**:4938–4943.
30. Gerth U, Krüger E, Derré I, Msadek T, Hecker M: Stress induction of the *Bacillus subtilis clpP* gene encoding a homologue of the proteolytic component of the Clp protease and the involvement of ClpP and ClpX in stress tolerance. *Mol Microbiol* 1998, **28**:787–802.
31. Cai SJ, Inouye M: EnvZ-OmpR interaction and osmoregulation in *Escherichia coli*. *J Biol Chem* 2002, **277**:24155–24161.
32. Gomes DF, Batista JSDS, Torres AR, Andrade DS, Galli-Terasawa LV, Hungria M: Two-dimensional proteome reference map of *Rhizobium tropici* PRF 81 reveals several symbiotic determinants and strong resemblance with agrobacteria. *Proteomics* 2012, **12**:859–863.
33. van Rhijn P, Vanderleyden J: The *Rhizobium*-plant symbiosis. *Microbiol Rev* 1995, **59**:124–142.
34. Geiger O, López-Lara IM: Rhizobial acyl carrier proteins and their roles in the formation of bacterial cell-surface components that are required for the development of nitrogen-fixing root nodules on legume hosts. *FEMS Microbiol Lett* 2002, **208**:153–162.
35. Parniske M: ExoB Mutants of *Bradyrhizobium japonicum* with reduced competitiveness for nodulation of *Glycine max*. *Mol Plant-Microbe Interact* 1993, **6**:99.
36. Becker BU, Kosch K, Parniske M, Müller P: Exopolysaccharide (EPS) synthesis in *Bradyrhizobium japonicum*: sequence, operon structure and mutational analysis of an *exo* gene cluster. *Mol Gen Genet* 1998, **259**:161–171.
37. Janczarek M, Skorupska A: Exopolysaccharide synthesis in *Rhizobium leguminosarum* bv. trifolii is related to various metabolic pathways. *Res Microbiol* 2003, **154**:433–442.
38. Janczarek M, Skorupska A: Regulation of *pssA* and *pssB* gene expression in *Rhizobium leguminosarum* bv. trifolii in response to environmental factors. *Antonie Van Leeuwenhoek* 2004, **85**:217–227.
39. Janczarek M, Król J, Skorupska A: The *pssB* gene product of *Rhizobium leguminosarum* bv. trifolii is homologous to a family of inositol monophosphatases. *FEMS Microbiol Lett* 1999, **173**:319–325.
40. Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC: How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. *Nat Rev Microbiol* 2007, **5**:619–633.
41. Abramovitch RB, Anderson JC, Martin GB: Bacterial elicitation and evasion of plant innate immunity. *Nat Rev Mol Cell Biol* 2006, **7**:601–611.
42. Boller T: Peptide signalling in plant development and self/non-self perception. *Curr Opin Cell Biol* 2005, **17**:116–122.
43. Zeidler D, Zähringer U, Gerber I, Dubery I, Hartung T, Bors W, Hutzler P, Durner J: Innate immunity in *Arabidopsis thaliana*: lipopolysaccharides activate nitric oxide synthase (NOS) and induce defense genes. *Proc Natl Acad Sci U S A* 2004, **101**:15811–15816.
44. Kiss E, Huguet T, Poinsoy V, Batut J: The *typA* gene is required for stress adaptation as well as for symbiosis of *Sinorhizobium meliloti* 1021 with certain *Medicago truncatula* lines. *Mol Plant Microbe Interact* 2004, **17**:235–244.
45. Goebel W, Gross R: Intracellular survival strategies of mutualistic and parasitic prokaryotes. *Trends Microbiol* 2001, **9**:267–273.
46. van Sluys MA, Monteiro-Vitorello CB, Camargo LE, Menck CFM, Da Silva AC, Ferro JA, Oliveira MC, Setubal JC, Kitajima JP, Simpson AJ: Comparative genomic analysis of plant-associated bacteria. *Annu Rev Phytopathol* 2002, **40**:169–189.
47. Batista JSS, Hungria M: Proteomics reveals differential expression of proteins related to a variety of metabolic pathways by genistein-induced *Bradyrhizobium japonicum* strains. *J Proteomics* 2012, **75**:1211–1219.
48. Chauhan S, O'Brian MR: Transcriptional regulation of delta-aminolevulinic acid dehydratase synthesis by oxygen in *Bradyrhizobium japonicum* and evidence for developmental control of the hemB gene. *J Bacteriol* 1997, **179**:3706–3710.
49. Page KM, Guerinot ML: Oxygen control of the *Bradyrhizobium japonicum hemA* gene. *J Bacteriol* 1995, **177**:3979–3984.
50. Frustaci JM, Brian MRO: Characterization of a *Bradyrhizobium japonicum* ferredoxin mutant and isolation of the hemH gene. *J Bacteriol* 1992, **174**:4223–4229.
51. Ludwig RA, Signer ER: Glutamine synthetase and control of nitrogen fixation in *Rhizobium*. *Nature* 1977, **267**:245–248.
52. Kondorosi A, Svfib Z, Kiss GB, Dixon RA: Ammonia assimilation and nitrogen fixation in *Rhizobium meliloti*. *Mol Gen Genet* 1977, **226**:221–226.
53. Dombrecht B, Marchal K, Vanderleyden J, Michiels J: Prediction and overview of the RpoN-regulon in closely related species of the Rhizobiales. *Genome Biol* 2002, **3**. research0076.1–0076.11.
54. Shingler V: Signal sensing by sigma 54-dependent regulators: derepression as a control mechanism. *Mol Microbiol* 1996, **19**:409–416.
55. Fischer HM: Genetic regulation of nitrogen fixation in rhizobia. *Microbiol Rev* 1994, **58**:352–386.
56. Clark SR, Oresnik IJ, Hynes MF: RpoN of *Rhizobium leguminosarum* bv. viciae strain VF39SM plays a central role in FnrN-dependent microaerobic regulation of genes involved in nitrogen fixation. *Mol Gen Genet* 2001, **264**:623–633.
57. Kullik I, Fritsche S, Knobel H, Sanjuan J, Hennecke H, Fischer HM: *Bradyrhizobium japonicum* has two differentially regulated, functional homologs of the sigma 54 gene (*rpoN*). *J Bacteriol* 1991, **173**:1125–1138.
58. Alazard D: Nitrogen fixation in pure culture by rhizobia isolated from stem nodules of tropical *Aeschynomene* species. *FEMS Microbiol Lett* 1990, **68**:177–182.
59. Kurz WGW, Larue TA: Nitrogenase activity in rhizobia in absence of plant host. *Nature* 1975, **256**:407–409.
60. Pankhurst CE, Scott DB, Ronson CW: Correlation between rifampicin-resistance of slow-growing *Rhizobium* strains and their ability to express nitrogenase activity in culture. *FEMS Microbiol Lett* 1982, **15**:137–139.
61. Agarwal AK, Keister DL: Physiology of ex planta nitrogenase activity in *Rhizobium japonicum*. *Appl Environ Microbiol* 1983, **45**:1592–1601.
62. Martínez-Romero E, Segovia L, Mercante FM, Franco AA, Graham P, Pardo MA: *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* sp. trees. *Int J Syst Bacteriol* 1991, **41**:417–426.
63. Graham PH, Draeger KJ, Ferrey ML, Conroy MJ, Hammer BE, Martinez E, Aarons SR, Quinto C: Acid pH tolerance in strains of *Rhizobium* and *Bradyrhizobium*, and initial studies on the basis for acid tolerance of *Rhizobium tropici* UMR1899. *Can J Microbiol* 1994, **40**:198–207.
64. Barboza F, Correa NS, Rosas SB: Metabolic and physiological characteristics of salt-tolerant strains of *Bradyrhizobium* spp. *Biol Fertil Soils* 2000, **32**:368–373.
65. Mary P, Dupuy N, Dolhembiremon C, Defives C, Tailliez R: Differences among *Rhizobium meliloti* and *Bradyrhizobium japonicum* strains in tolerance to desiccation and storage at different relative humidities. *Soil Biol Biochem* 1994, **26**:1125–1132.
66. Raza S, Jørgensen B, Abou-Taleb H, Christiansen JL: Tolerance of *Bradyrhizobium* sp. (Lupini) strains to salinity, pH, CaCO₃ and antibiotics. *Lett Appl Microbiol* 2001, **32**:379–383.
67. Asanuma S, Ayanaba A: Variation in acid-al tolerance of *Bradyrhizobium japonicum* strains from African soils. *Soil Sci Plant Nutr* 1990, **36**:309–317.
68. Wawrzynow A, Wojtkowiak D, Marszałek J, Banecki B, Jonsen M, Graves B, Georgopoulos C, Zyllicz M: The ClpX heat-shock protein of *Escherichia coli*, the ATP-dependent substrate specificity component of the ClpP-ClpX protease, is a novel molecular chaperone. *EMBO J* 1995, **14**:1867–1877.
69. Thomas JG, Baneyx F: ClpB and HtpG facilitate de novo protein folding in stressed *Escherichia coli* cells. *Mol Microbiol* 2000, **36**:1360–1370.
70. Yu AYH: Investigation of ClpXP protease mechanism of function and its interaction with the folding chaperone trigger factor by. PhD thesis. University of Toronto; 2013.
71. Jordan DC, Bacteriology D: Transfer of *Rhizobium japonicum* Buchanan 1980 to *Bradyrhizobium* gen. nov., a genus of slow-growing, root nodule bacteria from leguminous plants. *Int J Syst Bacteriol* 1982, **32**:136–139.
72. Puranamaneewiwat N, Niamsup H: Proteomic analysis of *Bradyrhizobium japonicum* USDA110 in acidic condition. *Chiang Mai J Sci* 2006, **33**:335–345.

73. Gupta RS: Evolution of the chaperonin families (Hsp60, Hsp10 and Tcp-1) of proteins and the origin of eukaryotic cells. *Mol Microbiol* 1995, **15**:1–11.
74. Hartl FU, Hayer-Hartl M: Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol* 2009, **16**:574–581.
75. Gomes DF, Batista JSS, Schiavon AL, Andrade DS, Hungria M: Proteomic profiling of *Rhizobium tropici* PRF 81: identification of conserved and specific responses to heat stress. *BMC Microbiol* 2012, **12**:84.
76. Münchbach M, Dainese P, Staudenmann W, Narberhaus F, James P: Proteome analysis of heat shock protein expression in *Bradyrhizobium japonicum*. *Eur J Biochem* 1999, **264**:39–48.
77. VanBogelen RA, Neidhardt FC: Ribosomes as sensors of heat and cold shock in *Escherichia coli*. *Proc Natl Acad Sci U S A* 1990, **87**:5589–5593.
78. Seaver LC, Imlay JA, Loewen P: Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. *J Bacteriol* 2001, **183**:7173–7181.
79. Krapp R, Division MB, Molecular B: The role of ferredoxin-NADP⁺ reductase in the concerted cell defense against oxidative damage - studies using *Escherichia coli* mutants and cloned plant genes. *Eur J Biochem* 1997, **563**:556–563.
80. Challougui Fatnassi I, Harzalli Jebara S, Jebara M: Selection of symbiotic efficient and high salt-tolerant rhizobia strains by gamma irradiation. *Ann Microbiol* 2010, **61**:291–297.
81. Kulkarni S, Surange S, Nautiyal CS: Crossing the limits of *Rhizobium* existence in extreme conditions. *Curr Microbiol* 2000, **41**:402–409.
82. Miller-Williams M, Loewen PC, Oresnik IJ: Isolation of salt-sensitive mutants of *Sinorhizobium meliloti* strain Rm1021. *Microbiology* 2006, **152**(Pt 7):2049–2059.
83. Lena CL: Stress-responsive proteins are upregulated in *Streptococcus mutans* during acid tolerance. *Microbiology* 2004, **150**:1339–1351.
84. Wei W, Jiang J, Li X, Wang L, Yang SS: Isolation of salt-sensitive mutants from *Sinorhizobium meliloti* and characterization of genes involved in salt tolerance. *Lett Appl Microbiol* 2004, **39**:278–283.
85. Nogales J, Campos R, BenAbdelkhalik H, Olivares J, Lluch C, Sanjuan J: *Rhizobium tropici* genes involved in free-living salt tolerance are required for the establishment of efficient nitrogen-fixing symbiosis with *Phaseolus vulgaris*. *Mol Plant Microbe Interact* 2002, **15**:225–232.
86. Minion FC, Lefkowitz EJ, Madsen ML, Cleary BJ, Swartzell SM, Mahairas GG: The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of wvine mycoplasmosis. *J Bacteriol* 2004, **186**:7123–7133.
87. Lubec G, Afjehi-Sadat L, Yang J-W, John JPP: Searching for hypothetical proteins: theory and practice based upon original data and literature. *Prog Neurobiol* 2005, **77**:90–127.
88. Pietras Z, Tsai Y, Callaghan AJ, Herna H: The regulatory protein RraA modulates RNA-binding and helicase activities of the *E. coli* RNA degradosome. *RNA* 2010, **16**:553–562.
89. Pandey SP, Minesinger BK, Kumar J, Walker GC: A highly conserved protein of unknown function in *Sinorhizobium meliloti* affects sRNA regulation similar to Hfq. *Nucleic Acids Res* 2011, **39**:4691–4708.
90. Fernando MR, Nanri H, Yoshitake S, Nagata-Kuno K, Minakami S: Thioredoxin regenerates proteins inactivated by oxidative stress in endothelial cells. *Eur J Biochem* 1992, **209**:917–922.
91. Freestone P, Nyström T, Trine M, Norris V: The universal stress protein, UspA, of *Escherichia coli* is phosphorylated in response to stasis. *J Mol Biol* 1997, **274**:318–324.
92. Vanderlinde EM, Magnus SA, Tambalo DD, Koval SF, Yost CK: Mutation of a broadly conserved operon (RL3499-RL3502) from *Rhizobium leguminosarum* biovar viciae causes defects in cell morphology and envelope integrity. *J Bacteriol* 2011, **193**:2684–2694.
93. Wong KS, Houry WA: Novel structural and functional insights into the MoxR family of AAA + ATPases. *J Struct Biol* 2012, **179**:211–221.
94. Taniguchi Y, Katayama M, Ito R, Takai N, Kondo T, Oyama T: *labA*: a novel gene required for negative feedback regulation of the cyanobacterial circadian clock protein KaiC. *Genes Dev* 2007, **21**:60–70.
95. Taniguchi Y, Nishikawa T, Kondo T, Oyama T: Overexpression of *lalA*, a paralog of *labA*, is capable of affecting both circadian gene expression and cell growth in the cyanobacterium *Synechococcus elongatus* PCC 7942. *FEBS Lett* 2012, **586**:753–759.
96. Porter PM, Banwart WL, Hassett JJ: HPLC isolation and GC-MS identification of genistein, daidzein, and coumestrol from unhydrolyzed soybean root extracts. *Environ Exp Bot* 1985, **25**:229–232.
97. Kosslak RM, Bookland R, Barkei J, Paaren HE, Appelbaum ER: Induction of *Bradyrhizobium japonicum* common nod genes by isoflavones isolated from *Glycine max*. *Proc Natl Acad Sci U S A* 1987, **84**:7428–7432.
98. Hungria M, Stacey G: Molecular signals exchanged between host plants and rhizobia: basic aspects and potential application in agriculture. *Soil Biol Biochem* 1997, **29**:819–830.
99. Debelle F, Mou L, Mangin B, Dénarié J, Boivin C: Nod genes and Nod signals and the evolution of the *Rhizobium* legume symbiosis. *Acta Biochim Pol* 2001, **48**:359–365.
100. Brencic A, Winans SC, Colonization P: Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. *Microbiol Mol Biol Rev* 2005, **69**:155–194.
101. Purcell EB, Boutte CC, Crosson S: Two-component signaling systems and cell cycle control in *Caulobacter crescentus*. *Adv Exp Med Biol* 2008, **631**:122–130.
102. Fioravanti A, Clantin B, Dewitte F, Lens Z, Verger A, Biondi EG, Villeret V: Structural insights into ChpT, an essential dimeric histidine phosphotransferase regulating the cell cycle in *Caulobacter crescentus*. *Acta Crystallogr Sect F: Struct Biol Cryst Commun* 2012, **68**(Pt 9):1025–1029.
103. Biondi EG, Reisinger SJ, Skerker JM, Arif M, Perchuk BS, Ryan KR, Laub MT: Regulation of the bacterial cell cycle by an integrated genetic circuit. *Nature* 2006, **444**:899–904.
104. Dudeja SS, Khurana AL: Persistence of *Bradyrhizobium* sp. (*Cajanus*) in a sandy loam. *Soil Biol Biochem* 1989, **21**:709–713.
105. Li J, Xiao W, Ma M, Guan D, Jiang X, Cao F, Shen D, Chen H, Li L: Proteomic study on two *Bradyrhizobium japonicum* strains with different competitiveness for nodulation. *Agric Sci China* 2011, **10**:1072–1079.
106. Ludwig EM, Leonard M, Marroqui S, Wheeler TR, Findlay K, Downie JA, Poole PS: Role of polyhydroxybutyrate and glycogen as carbon storage compounds in pea and bean bacteroids. *Mol Plant Microbe Interact* 2005, **18**:67–74.
107. Aneja P, Zachertowska A, Charles TC: Comparison of the symbiotic and competition phenotypes of *Sinorhizobium meliloti* PHB synthesis and degradation pathway mutants. *Can J Microbiol* 2005, **51**:599–604.
108. Nomura M, Arunothayanan H, Van Dao T, Le HT-P, Kaneko T, Sato S, Tabata S, Tajima S: Differential protein profiles of *Bradyrhizobium japonicum* USDA110 bacteroid during soybean nodule development. *Soil Sci Plant Nutr* 2010, **56**:579–590.
109. Marks BB, Nogueira MA, Hungria M, Megias M: Biotechnological potential of rhizobial metabolites to enhance the performance of *Bradyrhizobium* spp. and *Azospirillum brasilense* inoculants with soybean and maize. *AMB Express* 2013, **3**:21.
110. Santos MA, Vargas MA, Hungria M: Characterization of soybean *Bradyrhizobium* strains adapted to the Brazilian savannas. *FEMS Microbiol Ecol* 1999, **30**:261–272.
111. Menna P, Barcellos FG, Hungria M: Phylogeny and taxonomy of a diverse collection of *Bradyrhizobium* strains based on multilocus sequence analysis of the 16S rRNA gene, ITS region and *glnII*, *recA*, *atpD* and *dnaK* genes. *Int J Syst Evol Microbiol* 2009, **59**(Pt 12):2934–2950.
112. Barcellos FG, Menna P, Batista JSS, Hungria M: Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs *Sinorhizobium* (*Ensifer*) *freddiei* and *Bradyrhizobium elkanii* in a Brazilian Savannah soil. *Appl Environ Microbiol* 2007, **73**:2635–2643.
113. Barcellos FG, Batista JSS, Menna P, Hungria M: Genetic differences between *Bradyrhizobium japonicum* variant strains contrasting in N₂-fixation efficiency revealed by representational difference analysis. *Arch Microbiol* 2009, **191**:113–122.
114. Sadowsky MJ, Tully RE, Cregan PB, Keyser HH: Genetic diversity in *Bradyrhizobium japonicum* Sserogroup 123 and its relation to genotype-specific nodulation of soybean. *Appl Environ Microbiol* 1987, **53**:2624–2630.
115. Lery LMS, Coelho A, von Kruger WMA, Gonçalves MSM, Santos MF, Valente RH, Santos EO, Rocha SLG, Perales J, Domont GB, Teixeira KRS, Bertalan M, Ferreira PCG, Bisch PM: Protein expression profile of *Gluconacetobacter diazotrophicus* PAL5, a sugarcane endophytic plant growth-promoting bacterium. *Proteomics* 2008, **8**:1631–1644.
116. Chaves DFS, Souza EM, Monteiro RA, Pedrosa FO: A two-dimensional electrophoretic profile of the proteins secreted by *Herbaspirillum seropedicae* strain Z78. *J Proteomics* 2009, **73**:50–56.
117. Tatusov RL, Galperin MY, Natale DA, Koonin EV: The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 2000, **28**:33–36.

118. Gardy JL, Laird MR, Chen F, Rey S, Walsh CJ, Ester M, Brinkman FSL: **PSORTb v. 2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis.** *Bioinformatics* 2005, **21**:617–623.
119. Bhasin M, Garg A, Raghava GPS: **PSLpred: prediction of subcellular localization of bacterial proteins.** *Bioinformatics* 2005, **21**:2522–2524.
120. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340**:783–795.
121. Bateman A, Birney E, Cerruti L, Durbin R, Ewlinger L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer ELL: **The Pfam protein families database.** *Nucleic Acids Res* 2002, **30**:276–280.
122. Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Biswas M, Bradley P, Bork P, Bucher P, Copley R, Courcelle E, Durbin R, Falquet L, Fleischmann W, Gouzy J, Griffith-Jones S, Haft D, Hermjakob H, Hulo N, Kahn D, Kanapin A, Krestyaninova M, Lopez R, Letunic I, Orchard S, Pagni M, Peyruc D, Ponting CP, Servant F, *et al*: **InterPro: an integrated documentation resource for protein families, domains and functional sites.** *Brief Bioinform* 2002, **3**:225–235.
123. Alm EJ, Huang KH, Price MN, Koche RP, Keller K, Dubchak IL, Arkin AP: **The microbesOnline web site for comparative genomics.** *Genome Res* 2005, **15**:1015–1022.
124. Markowitz VM, Korzeniewski F, Palaniappan K, Szeto E, Werner G, Padki A, Zhao X, Dubchak I, Hugenholtz P, Anderson I, Lykidis A, Mavromatis K, Ivanova N, Kyrpides NC: **The integrated microbial genomes (IMG) system.** *Nucleic Acids Res* 2006, **34**:344–348.
125. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A, Simonovic M, Bork P, von Mering C: **STRING 8—a global view on proteins and their functional interactions in 630 organisms.** *Nucleic Acids Res* 2009, **37**:412–416.
126. Pfaffl MW, Horgan GW, Dempfle L: **Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR.** *Nucleic Acids Res* 2002, **30**:e36.
127. Bortolan S, Barcellos FG, Marcelino FC, Hungria M: **Expressão dos genes *nodC*, *nodW* e *nopP* em *Bradyrhizobium japonicum* estirpe CPAC 15 avaliada por RT-qPCR.** *Pesq Agropec Bras* 2009, **44**:1491–1498.

doi:10.1186/1471-2164-15-643

Cite this article as: Gomes *et al*: Proteomic analysis of free-living *Bradyrhizobium diazoefficiens*: highlighting potential determinants of a successful symbiosis. *BMC Genomics* 2014 **15**:643.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

