



UNIVERSIDADE
ESTADUAL DE LONDRINA

ARTHUR FERNANDES SIQUEIRA

**GENÔMICA ESTRUTURAL E PAPEL DOS
TRANSPORTADORES ABC EM *BRADYRHIZOBIUM*
JAPONICUM ESTIRPE CPAC 15 E *BRADYRHIZOBIUM*
DIAZOEFFICIENS ESTIRPE CPAC 7**

ARTHUR FERNANDES SIQUEIRA

**GENÔMICA ESTRUTURAL E PAPEL DOS
TRANSPORTADORES ABC EM *BRADYRHIZOBIUM
JAPONICUM* ESTIRPE CPAC 15 E *BRADYRHIZOBIUM
DIAZOEFFICIENS* ESTIRPE CPAC 7**

Dissertação de Mestrado apresentada ao
Departamento de Bioquímica e Biotecnologia
da Universidade Estadual de Londrina.

Orientadora: Dra. Mariangela Hungria

Londrina
2013

**Catálogo elaborado pela Divisão de Processos Técnicos da Biblioteca Central da
Universidade Estadual de Londrina**

Dados Internacionais de Catalogação-na-Publicação (CIP)

S618g	<p>Siqueira, Arthur Fernandes. Genômica estrutural e papel dos transportadores ABC em <i>Bradyrhizobium japonicum</i> estirpe CPAC 15 e <i>Bradyrhizobium diazoefficiens</i> estirpe CPAC 7 / Arthur Fernandes Siqueira. – Londrina, 2013. 101 f. : il.</p> <p>Orientador: Mariangela Hungria. Dissertação (Mestrado em Biotecnologia) – Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em Biotecnologia, 2013. Inclui bibliografia.</p> <p>1. Biotecnologia agrícola – Teses. 2. Rizóbio – Teses. 3. Microorganismos do solo – Teses. 4. Nitrogênio – Fixação – Teses. 5. Genética bacteriana – Teses. I. Hungria, Mariangela. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Biotecnologia. III. Título.</p> <p style="text-align: right;">CDU 663.1</p>
-------	---

ARTHUR FERNANDES SIQUEIRA

**GENÔMICA ESTRUTURAL E PAPEL DOS TRANSPORTADORES
ABC EM *BRADYRHIZOBIUM JAPONICUM* ESTIRPE CPAC 15 E
BRADYRHIZOBIUM DIAZOEFFICIENS ESTIRPE CPAC 7**

Dissertação de Mestrado apresentada ao
Departamento de Bioquímica e Biotecnologia
da Universidade Estadual de Londrina.

BANCA EXAMINADORA

Profa. Orientadora: Dra. Mariangela Hungria
UEL – Londrina -PR

Componente da Banca: Andre Shigeyoshi
Nakatani
Embrapa Soja – Londrina -PR

Componente da Banca: Fernando Gomes
Barcellos
UEL – Londrina -PR

Londrina, 25 de março de 2013.

Aos que apoiam a ciência para fazer o bem...

AGRADECIMENTOS

A minha gratidão pode ser dividida em várias partes. Inicialmente gostaria de agradecer a minha família, que é o meu apoio e o meu incentivo para seguir na ciência. A solidez na educação e nos bons ensinamentos estão refletindo hoje, na idade adulta. Estou conseguindo colher os frutos, acreditem. Amo vocês!

A Mariangela Hungria, minha querida orientadora, a pessoa que fez com que cada segundo em Londrina valesse a pena. Minha mãe científica que me educou e me ensinou como devo seguir em frente. Jamais será esquecida, de fato, sempre estará no meu coração. Obrigado por me aceitar como orientando, você foi tudo que eu sempre quis no meu mestrado e na minha vida!

Ao Ernesto Ormeño, que posso agora, chamar de amigo. Você é uma inspiração para mim, um grande pesquisador auto-didata. Este trabalho estaria muito incompleto sem a tua ajuda. Espero que sejamos parceiros científicos.

Ao André Nakatani (ギバさん) pelos conselhos da qualificação. E ao Willy Castagnotto, grande amigo, pela ajuda na correção ortográfica e linguística do texto, どうもありがとうともだち!

A Marisa Nicolás, minha madrinha de iniciação a bioinformática. Obrigado por me apresentar a este fantástico mundo! Você é o meu incentivo.

A todos os amigos do Laboratório de Biotecnologia de Solos – Embrapa Soja. Os momentos que passamos juntos serão, no mínimo, inesquecíveis.

Ao Alessandro Coutinho Ramos, meu orientador, pai científico da graduação e amigo. O homem que me mostrou a famosa atmosfera científica. Sou grato a você, ETERNAMENTE e a todos os integrantes do Laboratório de Microbiologia Ambiental e Biotecnologia (LMAB) que fizeram parte de minha formação.

A CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela concessão da bolsa de estudos, pois sem ela sequer teria ingressado no programa de Biotecnologia. Obrigado!

Aos invisíveis aos olhos e visíveis ao coração. Amigos cientistas, obrigado...VAMOS CONSEGUIR!!

“Biologists would rather share their toothbrush than their gene name.”

SIQUEIRA, Arthur Fernandes. **Genômica Estrutural e o Papel dos Transportadores ABC em *Bradyrhizobium japonicum* Estirpes CPAC 15 e *Bradyrhizobium diazoefficiens* estirpe CPAC 7**. 101 p. Dissertação (Mestrado em Biotecnologia) – Universidade Estadual de Londrina, Londrina, 2013.

RESUMO

A cultura de soja (*Glycine max* (L.) Merrill) possui grande importância para a economia brasileira. Tal leguminosa é utilizada, principalmente, na alimentação humana e animal. Para o bom desenvolvimento da soja, são necessárias grandes quantidades de nitrogênio (N) e, para tornar isso possível, a planta estabelece simbiose com bactérias do solo, principalmente do gênero *Bradyrhizobium*. No Brasil, *Bradyrhizobium diazoefficiens* estirpe CPAC 7 (= SEMIA 5080) e *Bradyrhizobium japonicum* estirpe CPAC 15 (= SEMIA 5079) são amplamente utilizadas como inoculantes para a cultura de soja. De forma a incrementar o conhecimento do potencial biotecnológico, o DNA genômico total de ambas estirpes foi submetido ao pirosequenciamento pela plataforma 454Flx (Roche). Em seguida, os *reads* foram montados utilizando ferramentas adequadas de bioinformática e anotados automaticamente e manualmente pelo *software* denominado SABIA. Os resultados mostraram que há certa similaridade estrutural entre ambos os genomas no percentual de conteúdo CG (aprox. 63%) e percentual da porção codificadora do genoma (aprox. 82%). Contudo, o genoma da estirpe CPAC 15 se mostrou maior que CPAC 7 em, aproximadamente, 500.000 pb; além disso, houve conformidade com outras estirpes já sequenciadas, *B. japonicum* USDA 6^T e *B. diazoefficiens* USDA 110^T. Após a anotação, as regiões codificadoras do genoma (*Open Read Frames*, *ORFs*) foram classificadas em 18 categorias pelo banco de dados KEGG. Os transportadores desempenham importantes papéis na ecologia de tais bactérias, sendo assim, foram analisados os transportadores do tipo ABC de íons orgânicos e inorgânicos, oligossacarídeos, polióis, monossacarídeos, fosfato, aminoácidos, peptídeos, níquel, cátions metálicos, ferro-sideróforos e outros tipos de transportadores incluindo os ABC-2. Após a curação manual, a estirpe CPAC 7 apresentou maior número de transportadores de carboidratos em relação à CPAC 15, já a CPAC 15 apresentou maior número de transportadores relacionados à captação de opinas, fosfonatos e poliaminas em relação à CPAC 7. Possivelmente na estirpe CPAC 7 há um transportador de manopina, talvez o primeiro relatado em rizóbios, de interesse para estudos futuros. Os demais dados abrirão novas perspectivas para pesquisas futuras com essas estirpes, principalmente de natureza bioquímica para a aplicação biotecnológica.

Palavras-chave: Genoma. *Bradyrhizobium japonicum*. *Bradyrhizobium diazoefficiens*. Transportadores.

SIQUEIRA, Arthur Fernandes. **Structural Genomics and Role of ABC Transporters in *Bradyrhizobium japonicum* strain CPAC 15 and *Bradyrhizobium diazoefficiens* strain CPAC 7.** 2012. 101 p. Dissertation (Biotechnology Master's Degree) – Universidade Estadual de Londrina, Londrina, 2013.

ABSTRACT

The soybean crop (*Glycine max* (L.) Merrill) has a major importance for the Brazilian economy. The legume is used mainly as human and animal food, and requires great amounts of nitrogen (N) for its growth. To supply this N the plant can associate symbiotically with soil bacteria, mainly belonging to the *Bradyrhizobium* genus. In Brazil, the *Bradyrhizobium diazoefficiens* strain CPAC 7 (=SEMIA 5080) and *Bradyrhizobium japonicum* strain CPAC 15 (=SEMIA 5079) are broadly used as soybean inoculants. To increase the biotechnological knowledge, total genomic DNA of both strains was pirosequenced using the 454Flx (Roche) sequencing analyzer. The reads obtained were assembled using bioinformatic tools, and manually and automatically annotated with the software SABIA. The results have shown structural similarity between both genomes in the %C/G content (about 63%) and % Genome Coding Region (about 82%); however, the genome of CPAC 15 genome is larger than CPAC 7 in approximately 500,000 bp. Furthermore, the genomes were in accordance to the genomes of other strains, *B. japonicum* USDA 6^T and *B. diazoefficiens* USDA 110^T. After annotation, Open Read Frames (ORFs) were classified in 18 classes by KEGG. Transporters have key roles in bacterial ecology; thus, ABC transporters, such as those related to the transport of minerals, organic ions, oligosaccharides, polyols, monosaccharides, phosphate, amino acids, peptides, nickel, metallic cations, ion-siderophores, ABC-2 and other transporters were analysed. After manual curation, CPAC 7 strain has shown more carbohydrate transporters and CPAC 15 more transporters related to opine, phosphonate and polyamine uptake. CPAC 7 probably has a manopine transporter, the first described in rhizobia, that deserves further analysis. The remaining data bring new perspectives for future researchs with these strains, mainly related to biochemistry and biotechnological applications.

Key words: Genome. *Bradyrhizobium japonicum*. *Bradyrhizobium diazoefficiens*.
Transporters

LISTA DE ILUSTRAÇÕES

Figura 1 -	Diagrama da planta de soja juvenil em estado vegetativo.	19
Figura 2 -	Nódulos presentes na raiz de soja.....	20
Figura 3 -	Evolução da produção de soja no Brasil.....	21
Figura 4 -	Processo de infecção durante a organogênese do nódulo. (A) Rizóbio se liga a um pelo radicular emergente em resposta à quimiotaxia por compostos liberados pela planta. (B) Em resposta aos fatores produzidos pela bactéria, o pelo radicular exibe um crescimento encurvado e as células do rizóbio proliferam-se dentro da região encurvada. (C) Ocorre degradação localizada da parede do pelo radicular e leva à infecção e formação do cordão de infecção através das vesículas secretadas do complexo-de-Golgi das células da raiz. (D) O cordão de infecção alcança o final da célula, e sua membrana se funde com a membrana plasmática da célula do pelo radicular. (E) Os rizóbios são lançados no apoplasto e penetram na lamela central até a membrana plasmática das células subepidérmicas, dando origem a um novo cordão de infecção que forma um canal aberto com o primeiro. (F) O cordão de infecção estende-se e ramifica-se até que encontre células alvo, onde vesículas compostas de membrana da planta que confinam as células bacterianas, são lançadas para o citosol.....	25
Figura 5 -	Filogenia de rizóbios e membros de alfa-Proteobacteria utilizando o rDNA 16S.....	29
Figura 6 -	Evolução dos projetos genoma de 1995 a 2012 indicando a curva de crescimento exponencial.....	37
Figura 7 -	Projetos Genoma referentes ao domínio Bacteria, indicando os seus status atuais.	38
Figura 8 -	Mapa do cromossomo de <i>Escherichia coli</i> K-12. As setas azuis marcam a metade do cromossomo atravessado por duas forquilhas de replicação indicando a origem e o término. O círculo concêntrico exterior indica a posição dos genes que são	

transcritos em sentido horário (laranja) para uma fita ou sentido anti-horário (amarelo) da fita complementar. A grande porção amarela no centro é um histograma em que o comprimento de cada faixa é proporcional à aleatoriedade da utilização do códon dentro de cada sequência codificante..... 40

Figura 9 - *Motifs* conservados no ABC. Os três *motifs* característicos encontrados em todos as ATPases de transportadores ABC são as caixas hachuradas em vermelho, que juntas formam a região de ligação de nucleotídeos da alça P da família da ATPase. Outros *motifs* característicos, as regiões *Q-loop* e *H-loop* (também chamada de região de troca), contêm somente um resíduo altamente conservado e são representados por caixas verdes hachuradas. Esses resíduos fazem contato com o γ -fosfato do ATP 50

Figura 10 - A arquitetura molecular de transportadores ABC. a) Um desenho da organização modular dos transportadores ABC, que são compostos de dois domínios transmembrana (*transmembrane domain – TMD*) e dois domínios ABC ou domínios de ligação a nucleotídeos (*nucleotide-binding domain*). A proteína de ligação que é requerida pelos importadores também é mostrada. Dois estados conformacionais do transportador ABC – voltado para o interior e voltado para o exterior com o sítio de ligação ao substrato orientado em direção às regiões periplasmáticas (extracelular) e citoplasmática (intracelular), respectivamente – são retratados para demonstrar o mecanismo de acesso alternado do transportador. b) O importador de vitamina B12 de *E. coli*, BtuCDF (Código 2QI9 no *Protein Data Bank – PDB*). c) Exportador multidroga Sav1866 de *Staphylococcus aureus* (Código 2ONJ no *Protein Data Bank – PDB*). ICL = loop intracelular (*intracellular loop*)..... 51

Figura 11 - Resultados finais da classificação das ORFs no genoma de *B. diazoefficiens* estirpe CPAC 7 (SEMIA 5080). 62

Figura 12 - Resultados finais da anotação do genoma de <i>B. japonicum</i> estirpe CPAC 15 (=SEMIA 5079).....	63
Figura 13 - Representação circular do cromossomo de <i>Bradyrhizobium japonicum</i> CPAC 15 (=SEMIA 5079). O círculo mais externo e o adjacente mostram as posições dos possíveis genes codificadores de proteínas no sentido horário e anti-horário. O terceiro círculo representa as comparações de BLASTN com <i>Bradyrhizobium diazoefficiens</i> USDA 110 ^T superior a 95% de identidade (E-value <10 ⁻¹⁰). Os dois últimos círculos mais internos representam o conteúdo GC (<i>GC content</i>) e <i>GC Skew</i> , respectivamente. As marcações no círculo mais interno representam as posições no genoma em Mbp. As setas em verde indicam a possível origem de replicação e término.	65
Figura 14 - Classificação funcional pelo KEGG das ORFs encontradas para <i>Bradyrhizobium japonicum</i> estirpes CPAC15 (= SEMIA 5079) e CPAC 7 (= SEMIA 5080).	67
Figura 15 - CDSs de <i>Bradyrhizobium diazoefficiens</i> estirpe CPAC 7 (=SEMIA 5080) indicando o percentual de todas as categorias do KEGG encontrados no genoma após a anotação.....	68
Figura 16 - CDSs de <i>Bradyrhizobium japonicum</i> estirpe CPAC 7 (=SEMIA 5080) indicando o percentual de todas as categorias encontrados no genoma após a anotação.....	69
Figura 17 - Classificação dos genomas de <i>Bradyrhizobium japonicum</i> CPAC 15 (=SEMIA 5079) e CPAC 7 (=SEMIA 5080) pelo banco de dados do RAST e gráficos gerados pelo <i>Seed Viewer</i> evidenciando a cobertura das sequências pelo RAST (<i>subsystem coverage</i>) e a distribuição das categorias (<i>subsystem category distribution</i> e <i>subsystem feature counts</i>).....	70
Figura 18 - Diagrama de Venn dos domínios comuns a transportadores do tipo ABC em <i>B. japonicum</i> estirpe CPAC 15 (amarelo) e <i>B. diazoefficiens</i> estirpe CPAC 7 (vermelho), indicando o percentual de domínios compartilhados e exclusivos em ambos os genomas.	72

Figura 19 - Imagem obtida do KEGG, indicando a presença ou ausência dos transportadores de monossacarídeos nos genomas de CPAC 15 e CPAC 7. Os transportadores pintados em cinza estão presentes em ambos os genomas.	73
Figura 20 - Via metabólica indicando a família do aspartato incluindo, todos os compostos que são seus precursores.	75
Figura 21 - Transportador de complexos férricos (sistema <i>fhuDBC</i>) presente no genoma indicando a analogia com o sistema <i>hmuTUV</i>	76
Figura 22 - Transportador de açúcares múltiplos completo presente nos genomas de <i>B.japonicum</i> estipes CPAC 7 e CPAC 15.	80
Figura 23 - Esquema do transporte de monossacarídeo realizado pela proteína GruA (AraG), onde ocorre o transporte do monossacarídeo pela proteína devido a liberação de energia contida nas ligação do fosfato ao ATP.	81
Figura 24 - Esquema do transportador de putrescina (1,4-diaminobutano) do tipo ABC, indicando as proteínas F,H,I e G presentes na membrana de <i>Bradyrhizobium diazoefficiens</i> CPAC 7 e CPAC15.	82
Figura 25 - Resultado da curação manual dos transportadores ABC em <i>Bradyrhizobium japonicum</i> CPAC 15, evidenciando as famílias e o percentual representativo no genoma.	85
Figura 26 - Resultado da curação manual dos transportadores ABC em <i>Bradyrhizobium japonicum</i> CPAC 7, evidenciando as famílias e o percentual representativo no genoma.	85

LISTA DE TABELAS

Tabela 1 -	Parâmetros fisiológicos e morfológicos avaliados <i>in vivo</i> e <i>in vitro</i> para as estirpes CPAC 7 e CPAC 15.....	36
Tabela 2 -	Classificação de genes codificadores de proteínas utilizando o banco de dados COG (<i>Clusters of Orthologous Groups</i>) para o genoma de <i>Bradyrhizobium japonicum</i> estirpe USDA6.	45
Tabela 3 -	Número de sequências codificantes do DNA (ORFs) de <i>B.japonicum</i> estirpe CPAC 15 classificada em categorias pelo COG (<i>Clusters of Orthologous Groups of Proteins</i>).	46
Tabela 4 -	Informações gerais do genoma da estirpe CPAC 7 de <i>Bradyrhizobium diazoefficiens</i> obtidas pelo software SABIA.	63
Tabela 5 -	Informações gerais do genoma da estirpe CPAC 15 de <i>Bradyrhizobium japonicum</i> obtidas pelo software Sabia.	64
Tabela 6 -	Percentual de identidade entre os genes referentes à codificação do transportador de sulfato de CPAC 7 e CPAC 15, indicando a diferença encontrada entre o tamanho das sequencias (Δ pb).	78
Tabela 7 -	Tabela final dos transportadores ABC encontrados nas estirpes de <i>B.japonicum</i> estirpes CPAC 15 e CPAC 7 após a curação manual.	84

SUMÁRIO

1	INTRODUÇÃO	15
2	OBJETIVOS	17
2.1	Objetivos Gerais	17
2.2	Objetivos Específicos	17
3	REVISÃO BIBLIOGRÁFICA	18
3.1	A Soja	18
3.2	Nutrição de Plantas: Nitrogênio	21
3.3	Fixação Biológica de Nitrogênio e Simbiontes	23
3.4	Classificação dos Rizóbios	26
3.4.1	Gênero Bradyrhizobium: Ecologia e Aplicações	30
3.4.1.1	Bradyrhizobium japonicum	31
3.5	Genomas Estruturais e Funcionais dos Rizóbios	37
3.6	Transportadores de membrana do tipo ABC	49
3.7	A Bioinformática: Conceitos e Aplicações na Genômica	52
4	MATERIAL E MÉTODOS	57
4.1	Extração de DNA Genômico	57
4.2	Sequenciamento: Pirosequenciamento	58
4.3	Anotação do Genoma	58
4.4	Análise de Bioinformática Fina dos Transportadores do Tipo ABC	60
4.5	Construção do Mapa Genômico	61
5	RESULTADOS E DISCUSSÃO	62
5.1	Panorama Geral da Anotação	62
5.2	ORFs Relacionados ao Transporte de Membrana	71
5.2.1	Transportadores do Tipo ABC	72
5.2.1.1	Transportadores de monossacarídeos e oligossacarídeos	73
5.2.1.2	Transportadores de peptídeos e aminoácidos	74

5.2.1.3	Transportadores de ions metálicos, ferro-sideróforos e níquel	76
5.2.1.4	Transportadores de sulfato e sulfonato	77
5.2.1.5	Transportadores de molibdênio.....	78
5.2.1.6	Transportadores de ions orgânicos e polióis.....	79
5.2.1.7	Transportadores do tipo ABC-2.....	83
5.2.1.8	Outros transportadores	83
5.2.2	Curaço Manual dos Transportadores do tipo ABC	84
6	CONSIDERAÇÕES FINAIS	87
	REFERÊNCIAS	88

1 INTRODUÇÃO

A soja (*Glycine max* (L.) Merrill) é uma leguminosa, inicialmente cultivada na China e que atualmente representa uma das principais culturas no mundo, responsável por movimentar bilhões de dólares anualmente. Sua principal utilização é na alimentação humana e animal, devido ao seu alto valor nutritivo e proteico. O Brasil é o segundo maior produtor de soja no mundo, perdendo apenas para os Estados Unidos. Dada a importância da cultura de soja no país, a economia tem reflexos distintos relacionados à produção e comercialização destes grãos. (ALI, 2000; QIU; CHANG, 2010).

Por se tratar de uma leguminosa, a soja, necessita de grandes quantidades de nitrogênio para se desenvolver adequadamente. O uso de fertilizantes nitrogenados implicava em custo elevado, uma vez que a síntese química requer uma grande quantidade de energia para a quebra de tripla ligação (fortemente estável) do nitrogênio gasoso de forma a transformá-lo em uma forma livre assimilável pelas plantas, além disso, o uso dos fertilizantes nitrogenados implica em poluição ambiental, percolando pelo solo e poluindo os lençóis freáticos, rios, lagos e reservatórios de água. Em contraste a isso, uma das características marcantes da planta de soja é a sua capacidade de estabelecer simbiose com bactérias específicas do solo capazes de fixar nitrogênio atmosférico (N₂). Este processo chamou a atenção dos pesquisadores em todo o mundo, pois se tratava de um processo barato, natural e que não agride o meio ambiente. O potencial biotecnológico dessas bactérias passou a ser explorado em escala crescente. Tais bactérias pertencem, principalmente, ao gênero *Bradyrhizobium*, e às espécies *Bradyrhizobium japonicum* e *Bradyrhizobium elkanii* que estão sendo amplamente estudadas devido ao benefício para o cultivo de soja. Um vasto trabalho de seleção das melhores estirpes foi iniciado, ganhando força principalmente nos Estados Unidos e no Brasil. Após vários estudos e testes de natureza fisiológica, genética e simbiótica duas estirpes de *Bradyrhizobium japonicum* foram autorizadas para a produção de inoculantes comerciais para a cultura de soja no Brasil, a estirpe CPAC 7 (=SEMIA 5080) e a CPAC 15 (=SEMIA 5079). Atualmente a estirpe CPAC 7 foi reclassificada em *Bradyrhizobium diazoefficiens* (SILVA, 1948; LOPES, GIARDINI, 1981).

Para melhor compreensão do funcionamento e do metabolismo dos organismos, a genômica desempenha importante papel. Estudos genômicos eram demorados e custosos há 30 anos, atualmente, com o advento das novas plataformas de sequenciamento em larga escala, o sequenciamento de um genoma procariótico se tornou mais rápido e barato, entretanto, o processamento dos dados requer um maior poder computacional. Com isso, a área da bioinformática teve uma grande ascensão, uma vez que o tratamento dos dados, montagem e anotação eram requeridos em maior escala para estudos de genômica, transcriptômica e metagenômica (SNUSTAD; SIMONS, 2000; PIERCE, 2004).

Contudo, apesar da importância da fixação biológica de nitrogênio com a cultura de soja, as estirpes CPAC 7 e CPAC 15 ainda não haviam sido sequenciadas.

Este trabalho vem, pioneiramente, apresentar os dados resultantes do sequenciamento pela plataforma 454 FLX - Roche para os genomas de ambas as estirpes. O principal objetivo foi o sequenciamento completo do genoma e o objetivo específico, um estudo mais detalhado sobre os transportadores do tipo ABC, que desempenham importantes funções na célula.

2 OBJETIVOS

2.1 OBJETIVOS GERAIS

Realizar o sequenciamento e a análise utilizando a plataforma 454 Roche (pirosequenciamento) dos genomas de *Bradyrhizobium japonicum* CPAC 15 (= SEMIA 5079) e *Bradyrhizobium diazoefficiens* estirpe CPAC 7 (=SEMIA 5080), autorizadas para a produção de inoculantes comerciais para cultura de soja (*Glycine max* (L.) Merrill) no Brasil.

2.2 OBJETIVOS ESPECÍFICOS

Realizar a anotação manual do genoma das estirpes com o programa SABIA.

Analisar os transportados de membrana do tipo ABC e comparação dos resultados encontrados entre as duas estirpes de *Bradyrhizobium japonicum*.

3 REVISÃO BIBLIOGRÁFICA

3.1 A SOJA

Mundialmente, a soja (*Glycine max* (L.) Merrill) representa a maior fonte de proteína, óleo e fitoquímicos promotores de saúde para a nutrição humana e para a alimentação de animais. O cultivo de soja também pode trazer benefícios à qualidade do solo, devido a sua habilidade em fixar nitrogênio atmosférico e seu sistema radicular profundo. A soja tornou-se uma importante *commodity* devido a sua grande adaptação geográfica, composição química, bom valor nutricional, benefícios funcionais à saúde e aplicações industriais. Há um grande potencial para a produção e utilização de soja e seus derivados nos alimentos, como forragem e aplicações farmacêuticas por todo o mundo (ALI, 2000).

O conteúdo de proteína na soja é de, aproximadamente, 40% e a quantidade de óleo de, aproximadamente, 20%. Essa cultura apresenta o maior ganho bruto de óleo vegetal dentre as culturas com maior produção no mundo. Em 2011, a área mundial total cultivada de soja foi de 102,99 milhões de hectares, resultando em uma produção total de 260,91 milhões de toneladas de grãos (FAO, 2013).

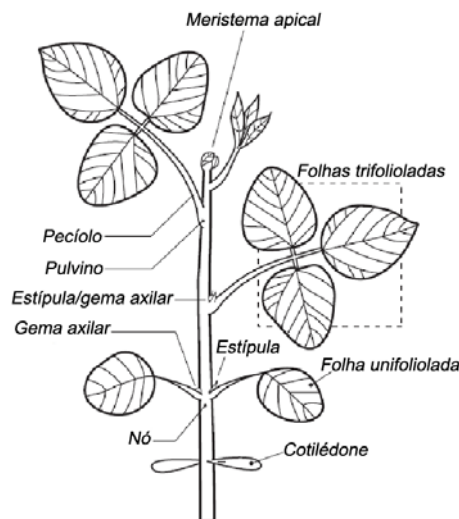
Estudiosos geralmente aceitam que a origem do cultivo da soja é a China. Inicialmente, a soja selvagem (*Glycine soja*), ancestral da soja atualmente cultivada (*G. max*), é encontrada por toda a China, sendo que a distribuição da *G. soja* limita-se à China, Japão, Coreia e oriente distante da Rússia, no Leste da Ásia, mas a sua distribuição na China é a mais rica em diversidade; a China também apresenta os registros mais antigos do cultivo dessas leguminosas. (QIU; CHANG, 2010).

Durante a primeira metade do século XX, a China foi o maior produtor e exportador de soja no mundo. Na década de 1950 a produção de soja avançou rapidamente nos EUA, agora o maior país produtor de grãos dessa leguminosa no mundo. Na década de 1970 a produção de soja avançou no Brasil, atualmente o segundo maior produtor mundial. Desde então, a produção de soja se desenvolveu rapidamente na Argentina, agora o terceiro maior país produtor. Em tais países, a taxa de *commodities* é alta, pois eles não são somente grandes produtores de soja, como também grandes exportadores. Na Índia, a produção de soja está se

desenvolvendo rapidamente com uma área de tamanho aproximado ao da China, contudo a produção por unidade de área é ainda relativamente baixa (QIU; CHANG, 2010).

De modo geral, as cultivares brasileiras têm ciclos entre 100 e 160 dias, e podem ser classificadas em grupos de maturação precoce, semi precoce, médio, semi tardio e tardio, dependendo da região. A altura da planta depende das condições ambientais e do genótipo (CISOJA, 2011). O esquema evidenciando a estrutura básica da planta de soja pode ser observado na Figura 1.

Figura 1 - Diagrama da planta de soja juvenil em estado vegetativo.



Fonte: Kumudini, 2010, (Modificado).

Nódulos são formados nas raízes de plantas de soja como consequência de uma relação mutualística benéfica entre a planta e bactérias pertencentes, principalmente, às espécies *Bradyrhizobium japonicum* e *Bradyrhizobium elkanii*. Essas são bactérias Gram-negativas presentes no solo. Outra espécie importante que se associa simbioticamente às raízes de soja é *Sinorhizobium (=Ensifer) fredii*, além de outras relatadas na literatura. No interior dos nódulos essas bactérias obtêm compostos orgânicos ricos em energia da planta. Em troca, as bactérias reduzem o nitrogênio atmosférico a amônia, a qual é subsequentemente transformada em compostos nitrogenados utilizados pela planta. A fixação biológica de nitrogênio é restrita a este e a outros procaríotos que possuem a enzima nitrogenase (SIEGBAHN et al., 1998). Na ausência da nitrogenase, a

redução do nitrogênio atmosférico não seria possível. A enzima nitrogenase é irreversivelmente inativada por oxigênio, portanto a fixação biológica de nitrogênio requer condições anóxicas ou quase anóxicas (KUMUDINI, 2010).

Os nódulos das raízes de soja mostram-se como visíveis inchaços esféricos na raiz (Figura 2). Na presença da bactéria, dezenas de nódulos por planta podem ser encontrados, tanto nos primeiros centímetros abaixo da superfície, como em maiores profundidades (GRUBINGER et al., 1982).

Figura 2 - Nódulos presentes na raiz de soja.



Fonte: Marcos Javier de Luca, 2011.

Há uma atividade interativa e conectada entre a bactéria e a raiz da soja a partir do momento de identificação mútua entre o macro e o micro simbiontes, além disto, ocorrem mudanças e processos que levam ao início e ao desenvolvimento do nódulo da raiz da soja (LHUISSIER et al., 2001; CARLSON; LERSTEN, 2004; OLDROYD; TIMMERS, 2007; DOWNIE, 2008;).

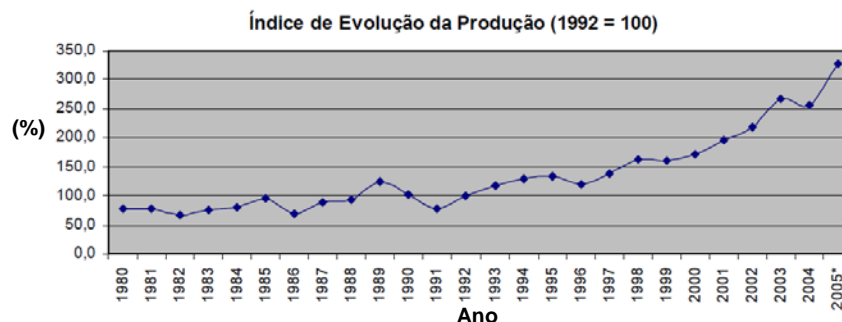
A produção de inoculantes visando culturas comerciais no Brasil iniciou-se em 1950 na Seção de Microbiologia Agrícola (SEMIA) do Rio Grande do Sul, com culturas líquidas e sobre ágar. Em 1954 foi adotado, pela primeira vez no país, o veículo turfoso e a produção alcançou 10.000 doses para soja, o suficiente para cerca de 8.000 hectares; na ocasião, era o único inoculante disponível, além de pequena quantidade de inoculante importado dos Estados Unidos. Em 1956, a

primeira indústria aprovada foi estabelecida no Rio Grande do Sul, com assistência do grupo FEPAGRO. Com a criação do Mercado Comum do Extremo Sul (MERCOSUL), os produtores passaram a considerar, seriamente, a melhoria da qualidade, em vista da competição de produtos da Argentina e do Uruguai. Em 1997, o Ministério da Agricultura tornou a fiscalização mais rigorosa. (FREIRE; VERNETTI, 1999).

A melhoria da qualidade também está sendo buscada pelas indústrias através da introdução de outros tipos de inoculantes, com base de óleo, líquidos e liofilizados com veículos de polímeros (FREIRE; VERNETTI, 1999).

O crescimento da produção de soja é, sem exagero, algo extraordinariamente notável, principalmente na região Central do Brasil. Este excelente desenvolvimento pode ser melhor avaliado a partir do índice de evolução da produção desta cultura no período 1980/2005 que foi elaborado tendo como base o ano de 1992 (Figura 3) (GARCIA; ZAPAROLLI, 2006). É importante salientar que esse incremento na produção não esteve só relacionado à expansão na área cultivada, mas também ao aumento do rendimento de grãos (CONAB, 2012).

Figura 3 - Evolução da produção de soja no Brasil



Fonte: IBGE (Elaboração por GARCIA; ZAPAROLLI, 2006), (Modificado).

3.2 NUTRIÇÃO DE PLANTAS: NITROGÊNIO

Dentre os elementos químicos presentes no solo, alguns são considerados essenciais para o crescimento de plantas. Um elemento essencial é definido como aquele cuja ausência impede a planta de completar o seu ciclo de vida, ou aquele que possui um papel fisiológico importante (TAIZ; ZEIGER, 2002).

Em termos globais, consideram-se a disponibilidade de água e dos nutrientes, nitrogênio (N) e fósforo (P), como os principais fatores limitantes à produção agrícola. Em caso de deficiência hídrica, a água precisa ser fornecida às culturas por meio de irrigação. Em relação ao P, deve ser suprido como fertilizante, mas diversos microrganismos, como os fungos micorrízicos, podem incrementar a absorção dos nutrientes, enquanto outros, como espécies de *Aspergillus*, são capazes de solubilizar fontes fosfatadas com baixa disponibilidade para as plantas. Já o N pode ser, em diversos casos, fornecido, parcial ou totalmente, pelo processo da fixação biológica do nitrogênio atmosférico (N₂) (HUNGRIA; CAMPO, 2007).

O N compõe, aproximadamente, 1,5% da massa seca das plantas e faz parte de muitos componentes celulares da planta, como os aminoácidos, amidos, proteínas, ácidos nucleicos, nucleotídeos, coenzimas, hexoaminas, etc. É um elemento mineral que as plantas requerem em grande quantidade, pertencendo ao grupo dos macronutrientes (TAIZ; ZEIGER, 2002).

As fontes para o fornecimento desse elemento são: 1) o solo, principalmente pela decomposição da matéria orgânica; 2) o processo de fixação não-biológica, resultante de descargas elétricas, combustão e vulcanismo; 3) os fertilizantes nitrogenados; e 4) o processo de fixação biológica do N₂ (POSTGATE; HILL, 1979; POSTGATE, 1982; NEWTON, 2000).

A deficiência em N rapidamente inibe o crescimento da planta. Se a deficiência persistir, muitas espécies apresentam clorose (amarelamento das folhas), especialmente nas folhas velhas próximas à base da planta. Sob severa deficiência de N, essas folhas podem se tornar completamente amarelas (ou escuras) e caírem (TAIZ; ZEIGER, 2002). Com isso, ocorre comprometimento em todo o aparato fotossintético, dificultando o seu desenvolvimento.

Como os solos brasileiros são, em geral, pobres em N e P, o aporte contínuo desses nutrientes é necessário para a manutenção da fertilidade do solo e para a nutrição das plantas. A produção dos fertilizantes nitrogenados envolve um consumo elevado de derivados de petróleo, portanto, com oscilações constantes de preço. Além disso, a maior parte dos fertilizantes nitrogenados e dos fosfatados mais solúveis utilizados no Brasil é importada, também com preços atrelados ao dólar. Como resultado, em períodos de preços baixos dos produtos agrícolas, o agricultor não consegue repor todos os nutrientes que foram retirados pelas culturas, resultando no empobrecimento gradual do solo. Nesse contexto, particularmente nas

últimas décadas, vem sendo dada cada vez mais importância aos processos biológicos envolvendo os microrganismos do solo capazes de aportarem nutrientes como o N e, desse modo, enriquecendo o solo e diminuindo as perdas de nutrientes, por lixiviação, ou na forma de gases (HUNGRIA; CAMPO, 2007).

3.3 FIXAÇÃO BIOLÓGICA DE NITROGÊNIO E SIMBIONTES

O solo é um complexo físico, químico e biológico, heterogêneo, contendo fases sólidas, líquidas e gasosas. Todas essas fases interagem com os nutrientes minerais. As partículas da fase sólida promovem reserva de potássio (K), cálcio (Ca), magnésio (Mg) e ferro (Fe). Também associado a essa fase sólida há os compostos orgânicos contendo N, P, enxofre (S), dentre outros elementos (TAIZ; ZEIGER, 2002).

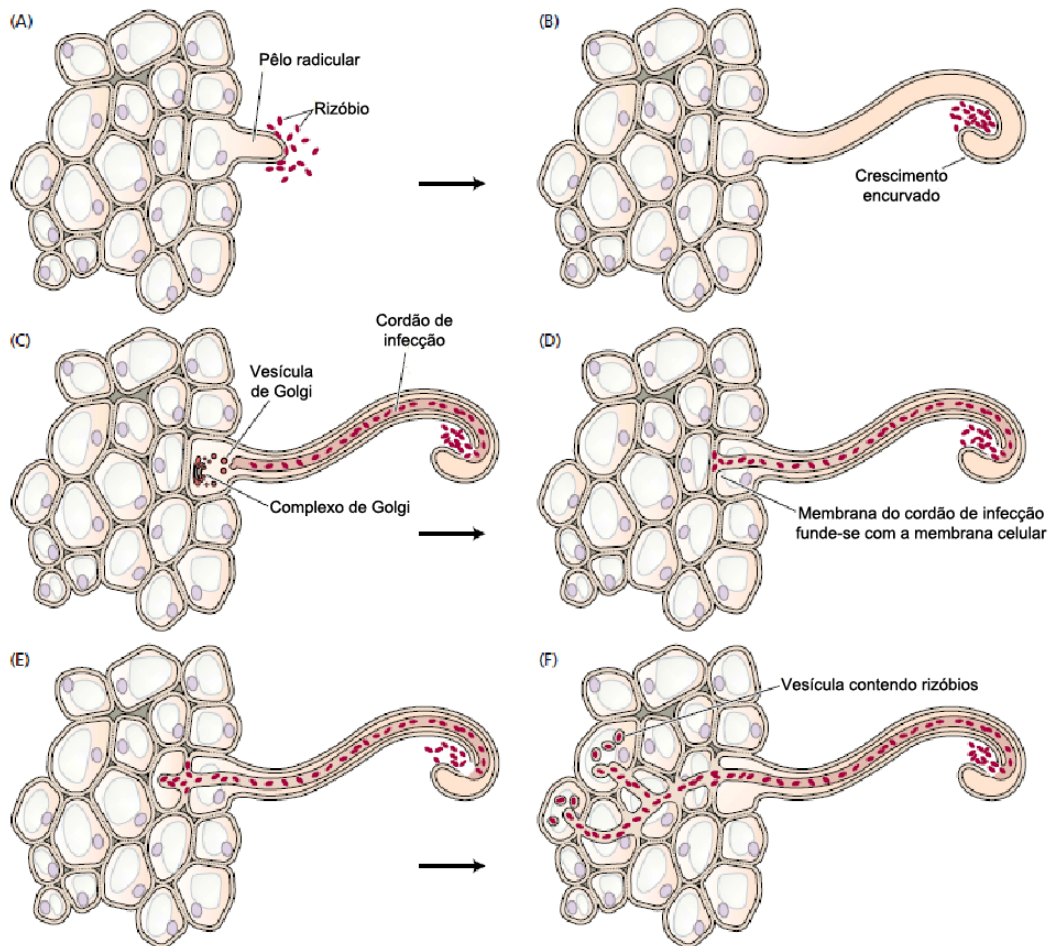
Muitos microrganismos estão presentes no solo e em maior concentração na rizosfera das plantas. Alguns destes microrganismos realizam associações simbióticas com plantas. O termo simbiose foi introduzido por Anton de Bary em 1879, este mesmo se refere a organismos “vivendo juntos” e, nesse sentido amplo, engloba situações que variam do parasitismo ao mutualismo (HARRISON, 1998). Os processos evolutivos determinam a especificidade da interação microrganismo/planta, com isso, seletos grupos de microrganismos colonizam determinados grupos de plantas e vice-versa, atuando na interface entre o solo e a raiz, funcionando como agentes facilitadores da captação de nutrientes, sendo assim, a associação simbiótica se tornou uma estratégia importante para a manutenção da homeostase de nutrição de plantas.

Na natureza, o N, presente na forma de N_2 , constitui cerca de 80% do gás atmosférico, todavia, organismos eucariontes, tais como as plantas, são incapazes de absorver N_2 e convertê-lo a uma forma assimilável devido à tripla ligação existente entre os átomos do N_2 , que é uma das mais fortes que se tem conhecimento na natureza (HUNGRIA et al., 1994). Com isso, as bactérias fixadoras de N_2 assumem um papel importante nesta etapa, pois são capazes de converter o N_2 atmosférico em amônia. (TAIZ; ZEIGER, 2002).

A infecção de uma leguminosa por um rizóbio representa uma associação íntima, que depende do reconhecimento específico para a efetiva fixação do N_2 (VINCENT, 1982). Dois processos, infecção e organogênese do nódulo,

ocorrem simultaneamente durante a formação do nódulo na raiz. Inicialmente, a planta libera sinais moleculares, na maioria flavonóides, os quais são necessários para ativar genes reguladores da nodulação. Durante o processo de infecção, o rizóbio, que está unido aos pelos radiculares, libera os fatores Nod, que induzem uma série de mudanças nas raízes, como o encurvamento do pelo radicular (Figuras 4 A e B). O rizóbio, então, fica confinado em um pequeno compartimento formado pelo encurvamento do pelo radicular. A parede celular do pelo radicular degrada-se nessas regiões, também em resposta aos fatores Nod e outras enzimas produzidas, permitindo às células bacterianas acesso direto à superfície externa da membrana plasmática da planta (LAZAROWITZ; BISSELING, 1997). A próxima etapa consta na formação do cordão de infecção (Figura 4 C), um extenso túbulo interno da membrana plasmática que é produzido pela fusão de membranas derivadas do complexo-de-Golgi no sítio de infecção. O cordão de infecção alcança o final da célula, e sua membrana se funde com a membrana plasmática da célula do pelo radicular. No córtex da raiz, próximo ao xilema, as células corticais diferenciam-se e começam a se dividir, formando uma área distinta no interior do córtex, chamado de primórdio nodular, do qual o nódulo vai se desenvolver (TAIZ; ZEIGER, 2002). O cordão de infecção cheio de rizóbios proliferados alonga-se dentro do pelo radicular e pelas camadas de células corticais, na direção do primórdio nodular. Quando o cordão de infecção alcança células especializadas dentro do nódulo, funde-se com a membrana plasmática da célula hospedeira, liberando células bacterianas que estão envolvidas em uma membrana derivada da membrana plasmática da célula hospedeira (Figura 4 D). A ramificação do cordão de infecção dentro dos nódulos permite à bactéria infectar muitas células (Figuras 4 E e F) (MYLONA; PAWLOWSKI; BISSELING, 1995).

Figura 4 - Processo de infecção durante a organogênese do nódulo. (A) Rizóbio se liga a um pelo radicular emergente em resposta à quimiotaxia por compostos liberados pela planta. (B) Em resposta aos fatores produzidos pela bactéria, o pelo radicular exibe um crescimento encurvado e as células do rizóbio proliferam-se dentro da região encurvada. (C) Ocorre degradação localizada da parede do pelo radicular e leva à infecção e formação do cordão de infecção através das vesículas secretadas do complexo-de-Golgi das células da raiz. (D) O cordão de infecção alcança o final da célula, e sua membrana se funde com a membrana plasmática da célula do pelo radicular. (E) Os rizóbios são lançados no apoplasto e penetram na lamela central até a membrana plasmática das células subepidérmicas, dando origem a um novo cordão de infecção que forma um canal aberto com o primeiro. (F) O cordão de infecção estende-se e ramifica-se até que encontre células alvo, onde vesículas compostas de membrana da planta que confinam as células bacterianas, são lançadas para o citosol.



Fonte: Taiz; Zeiger, 2002, (Modificado).

3.4 CLASSIFICAÇÃO DOS RIZÓBIOS

A classificação taxonômica dos rizóbios foi, inicialmente, baseada em sua especificidade hospedeira, a qual, embora ambígua, tinha um aspecto prático pela sua utilidade para os usuários e fabricantes de inoculantes (PINTO, 2007).

Nas duas décadas passadas, houve um aumento espetacular no número confirmado de bactérias fixadoras de N_2 . Isso foi relacionado tanto ao uso de métodos aprimorados para detecção e confirmação da capacidade de fixação de N_2 , como ao desenvolvimento de técnicas moleculares para identificação de bactérias diazotróficas. Uma lista de diazotróficos de vida livre e os tipos de habitats designados no qual a fixação de nitrogênio se estabelece foi compilado por Postgate (1982). Com exceção das bactérias fotossintéticas fixadoras de N_2 e cianobactérias, a maioria das bactérias diazotróficas de vida livre são heterotróficas, requerendo habitats capazes de prover substrato com fontes de carbono utilizáveis, que são necessárias para suprir a demanda energética para a fixação de N_2 . Uma relação mais estreita com o hospedeiro ocorre com microrganismos classificados como simbióticos, os rizóbios, que, de modo geral, revelam sua capacidade de fixar N_2 somente quando em simbiose com plantas leguminosas e algumas não-leguminosas. No entanto, os rizóbios são bactérias do solo de vida livre e sobrevivem *in situ* à custa de compostos nitrogenados do solo; possuem os genes necessários que lhes permitem fixar o N_2 quando encontrarem as plantas hospedeiras específicas e mesmo quando estão no solo, sem contato com plantas hospedeiras os genes continuam presentes de forma inativada. Estes utilizam energia dos fotossintatos fornecidos do nódulo, onde então é incorporado em esqueletos de carbono, transformado em compostos nitrogenados e exportado para planta hospedeira e excretam os íons NH_4^+ produzidos diretamente no tecido vegetal. (HUNGRIA; VARGAS; ARAUJO, 1997).

O gênero *Rhizobium* já havia sido descrito por Kirchner, em 1886 e por Frank, em 1889 (FRED et al. 1932 apud VARGAS; HUNGRIA, 1997). Esse agrupamento inicial dos rizóbios foi baseado em características do fenótipo e, principalmente, na sua capacidade em nodular leguminosas específicas, o que gerou o conceito de “grupos de inoculação cruzada”. Esse conceito resulta do princípio de que a infecção de plantas pelo rizóbio apresenta certa especificidade,

como por exemplo, *R. leguminosarum* bv. trifolii nodula os trevos (*Trifolium* spp.), *R. leguminosarum* bv. viceae nodula ervilha, ervilhaca e fava (*Pisum sativum*, *Vicia sativa*, *Vicia faba*, respectivamente) e *R. leguminosarum* bv. phaseoli o feijoeiro (*Phaseolus vulgaris*). O conceito de especificidade hospedeira, porém, há muito foi criticado, contudo, somente após meio século, Jordan (1984), considerando características morfológicas, fisiológicas, bioquímicas e genéticas, reclassificou algumas espécies de rizóbio e separou as que apresentavam crescimento lento em meio de cultura específico em um novo gênero, *Bradyrhizobium* (*bradus*, grego, lento), posteriormente, outros gêneros foram descritos, como *Azorhizobium* (DREYFUS et al., 1988) e *Sinorhizobium* (CHEN; YAN; LI., 1988).

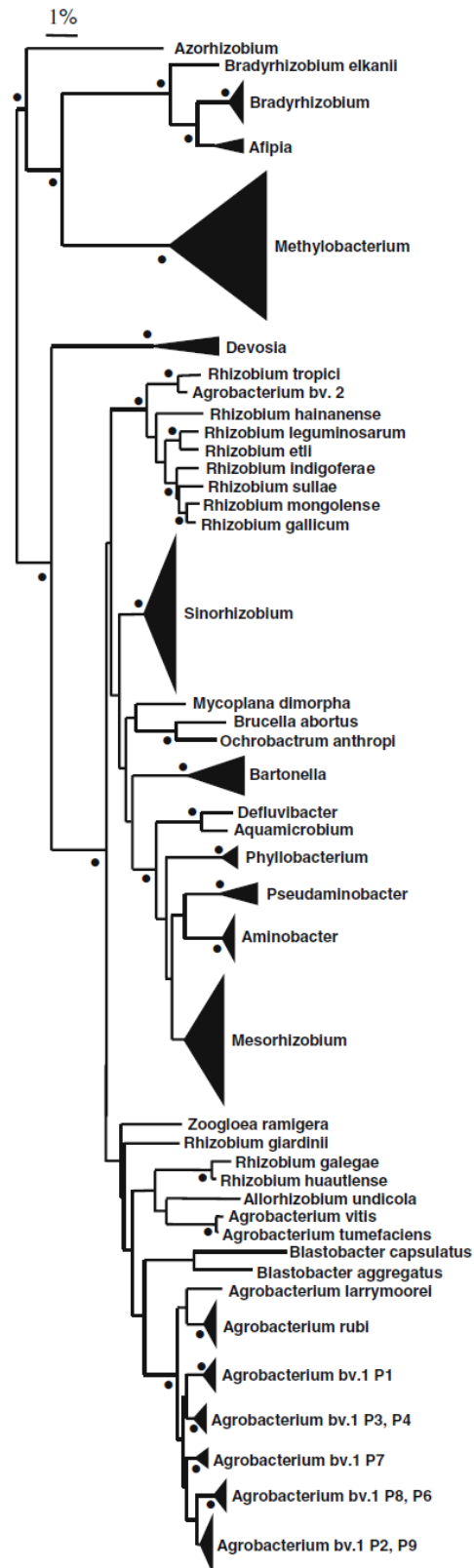
O termo “rizóbio”, no sentido exato, deveria se referir a membros do gênero *Rhizobium*. Com o passar dos anos, entretanto, o termo passou a ser usado para designar todas as bactérias que fossem capazes de nodular e fixar nitrogênio em associação com leguminosas, pertencentes a um gênero que foi uma vez parte do gênero *Rhizobium*, ou intimamente relacionado com ele (WILLEMS, 2006).

Para a classificação das diversas espécies de rizóbio são, tradicionalmente, utilizados testes morfológicos, fisiológicos, bioquímicos, enzimáticos e simbióticos, como a taxa de crescimento em meio de cultura contendo manitol como fonte de carbono, a habilidade de utilizar fontes de carbono e de nodular leguminosas hospedeiras, entre outros (VINCENT, 1970; JORDAN, 1984). A partir de 1984, novos métodos para a classificação passaram a ser adotados para a caracterização taxonômica, como a hibridização do DNA-DNA, análises de RNA, composição de exopolissacarídeos, requerimentos nutricionais, entre outros (ELKAN, 1992). Atualmente o sequenciamento do rRNA 16S é uma das técnicas utilizadas para estudar e caracterizar diversos tipos de comunidades bacterianas. O principal aspecto que torna o rRNA 16S um bom marcador para este tipo de tarefa é a onipresença entre os procariotos. Além disso, há muitas sequências de rRNA 16S depositadas nos bancos de dados online, o que torna a comparação entre procariotos mais abrangente e eficiente para a construção de árvores filogenéticas (MIZRAHI-MAN; DAVENPORT; GILAD, 2013).

Com isso, a bioinformática e a biologia molecular tornaram-se importantes ferramentas para a classificação taxonômica dos rizóbios, pois com estas ferramentas é possível fazer uma análise das sequências de nucleotídeos e, assim, realizar uma comparação de genomas, ao passo que *softwares* específicos

ajudam a propor árvores filogenéticas, explicando a relação de parentesco entre os rizóbios em questão. Willems (2006) compilou a filogenia de rizóbios e membros da classe alfa-Proteobacteria utilizando o rDNA 16S. As árvores foram calculadas pelo método "Neighbor Joining", usando correções Kimura-2. Uma análise de *bootstrap* foi realizada com 500 repetições e os agrupamentos que foram recuperados em 95% ou mais das árvores estão marcados no dendograma por um ponto na ramificação (Figura 5)

Figura 5 - Filogenia de rizóbios e membros de alfa-Proteobacteria utilizando o rDNA 16S.



Fonte: Willems, (2006).

Desde a subdivisão das estirpes classificadas em *B. japonicum* em 1992, nenhuma outra mudança foi proposta apesar de vários relatos de subgrupos dentro dessas espécies. Esses relatos tiveram início em 1980: análises de hibridização DNA:DNA definiram dois grupos homólogos para estirpes de *B. japonicum* (HOLLIS et al., 1981) – I e Ia – confirmado por Kuykendall e colaboradores (1992). Urtz & Elkan (1996) também relataram baixa homologia DNA:DNA entre a estirpe tipo USDA 6 e a estirpes do grupo Ia amplamente estudadas de USDA 110.

Nas duas últimas décadas, foram relatadas diferenças entre morfo-fisiológicas, genéticas e simbióticas dentro de *B. japonicum* (BODDEY; HUNGRIA, 1997; DELAMUTA et al., 2012; FERREIRA; HUNGRIA, 2002; GERMANO et al., 2006; MENNA; HUNGRIA, 2011; MENNA et al., 2006; SANTOS et al., 1999.) incluindo uma proposta, baseado na técnica de MLSA, de que estirpes do grupo Ia podem representar uma nova espécie (MENNA et al., 2009).

Recentemente, Delamuta et al., 2013 realizou um estudo utilizando as estirpes diversas estirpes de *Bradyrhizobium* dentre elas: SEMIA 6059, SEMIA 5060, CPAC 7 e USDA 110. Após as análises do rRNA 16S, de fragmentos dos genes *atpD*, *glnII*, *recA*, *gyrB*, *rpoB*, and *dnaK* e da região ITS (*Intergenic Transcribed Spacer*) do rRNA 16S-23S os resultados baseados em todas as evidências genotípicas e fenotípicas levaram a proposta da reclassificação do grupo Ia de *B. japonicum* em uma nova espécie denominada *Bradyrhizobium diazoefficiens*.

3.4.1 Gênero *Bradyrhizobium*: Ecologia e Aplicações

Segundo Holt et al. (1994), bactérias pertencentes ao gênero *Bradyrhizobium* são aeróbias, uma vez que o oxigênio é o aceptor final de elétrons. Também são quimio-organotróficas, pois utilizam pentoses como fonte de carbono. Como fonte de N, utilizam sais de amônia (geralmente nitratos e aminoácidos). No caso da simbiose com a cultura da soja, são duas as principais espécies que se associam à leguminosa, *B. japonicum* e *B. elkanii*.

As estirpes microssimbiontes da soja pertencentes às espécies *B. japonicum* e *B. elkanii* foram introduzidas nos solos brasileiros, principalmente a

partir das décadas de 1960 e 1970, com as primeiras inoculações na Região Sul (HUNGRIA et al., 1994). A inoculação em maior escala iniciou com a estirpe SEMIA 566, que foi isolada no Rio Grande do Sul em 1966 de um inoculante americano. A estirpe foi considerada um isolado adaptado e eficiente. Em 1970, tiveram início os programas de seleção de estirpes para inoculação de soja nos Cerrados. A SEMIA 566 foi utilizada nos primeiros inoculantes empregados na região dos Cerrados e, embora mostrasse incompatibilidade com o cultivar IAC-2, conseguiu estabelecer-se no solo devido a sua alta capacidade saprofítica (VARGAS; HUNGRIA, 1997). Atualmente, no Brasil, quatro estirpes são autorizadas para a fabricação de inoculantes comerciais: SEMIA 5019 (29w), SEMIA 587, SEMIA 5079 (= CPAC 15) e SEMIA 5080 (= CPAC 7). Essas estirpes podem ser utilizadas individualmente ou combinadas duas a duas, a critério do fabricante de inoculantes (ZILLI, et al., 2006).

A diversidade de rizóbios varia segundo diferentes tipos de solos e, também, conforme o tipo manejo. Grossman e colaboradores (2011), por exemplo, realizaram um estudo de diversidade de rizóbios em solo sob diferentes condições de manejo: sistema de agricultura orgânica e sistema convencional. Os resultados mostraram que a diversidade de rizóbios foi maior no solo sob sistema orgânico, uma vez que o manejo tradicional utiliza fertilizantes e pesticidas sintéticos, os quais são ricos em compostos como $(\text{NH}_4)_2\text{SO}_4$, NH_4Cl ou NH_4NO_3 , principais responsáveis pelo decréscimo da diversidade.

A investigação sobre a fixação de N_2 , com ênfase na identificação de estirpes *Bradyrhizobium* microssimbiontes da soja mais eficientes e competitivas tem sido fundamental para o sucesso da cultura no Brasil, promovendo altos rendimentos e baixos custos de produção (HUNGRIA et al. 2006).

3.4.1.1 *Bradyrhizobium japonicum*

As células da espécie *B.japonicum* são Gram-negativas, não formadoras de esporo, em forma de bastonete curto, com tamanho de 0,5 a 0,9 μm por 1,2 a 3,0 μm e são, comumente, pleomórficas sob diversas condições de crescimento. De um modo geral, contêm grânulos de poli- β -hidroxibutirato que são refringentes por microscopia por contraste de fase. Possuem mobilidade por um flagelo polar ou subpolar. As colônias são circulares, opacas, raramente

translúcidas, brancas e convexas e tendem a ser granular em textura (SANTOS, 1998). Apresentam um milímetro de diâmetro dentro de 5 a 7 dias de incubação em meio com extrato de levedura-manitol-ágar (SANTOS, 1998). Crescem em meio com carboidrato usualmente associado à secreção extracelular. Segundo Santos (1998), este gênero também se caracteriza por apresentar crescimento lento, com tempo de geração de 7 a 13 horas. Muitas estirpes crescem em meios com sais minerais contendo extrato de levedura, glicose, galactose, gliconato, glicerol, frutose, arabinose ou manitol. A maltose é utilizada por cerca de 10% das estirpes, mas a lactose, ramnose, rafinose, trealose, sacarose, dulcitol e dextrina são raramente utilizadas. Ácidos orgânicos como fumarato, malato, succinato, citrato e piruvato são utilizados, desde que o meio possua Ca^{2+} e Mg^{2+} suficientes para sobrepujar o efeito quelante desses ácidos. Celulose e amido não são utilizados (JORDAN, 1984).

Algumas estirpes podem utilizar sais de amônio ou nitrato como a única fonte de nitrogênio. Certos aminoácidos (glutamato, histidina, aspartato e prolina) servem como única fonte de nitrogênio, apesar de possuírem quantidade menor de vitaminas em relação à caseína hidrolisada. A caseína e o ágar não são hidrolizados e a peptona é pobremente utilizada (JORDAN, 1984).

As condições ótimas de temperatura e pH para o crescimento são de 20 a 30°C e pH 6 a 7. Usualmente são tolerantes à acidez, muitas estirpes crescem em pH 4,5, mais de 30% das estirpes crescem em pH 4,0 e poucas em pH inferior a 3,5. Não ocorre crescimento em pH superior a 9,0, não crescem em meio contendo 2% NaCl e não produzem H_2S . A produção de penicilinase é comum (JORDAN, 1984). Todas as características gerais descritas a espécie *B. japonicum* são também são comuns a *B. diazoefficiens* (Dealamuta et al., 2013).

A enzima chave na via da pentose fosfato é a NADP⁺ ligada a 6-fosfogliconato desidrogenase (EC 1.1.1.43). A glicose é metabolizada amplamente pela via metabólica de Entner-Doudoroff. Usualmente não há requerimento por vitaminas extracelulares, com a rara exceção da biotina. O ácido heteropolissacarídeo da secreção extracelular é heterogêneo, tanto em estrutura, quanto em composição e contém ácido d-galacturônico e, frequentemente, açúcares metilados; glicanos neutros da secreção também são heterogêneos (JORDAN, 1984).

A atividade da nitrogenase por células livres ocorre em certas estirpes, mas somente em meios contendo fontes de carbono selecionadas e sobre

reduzida tensão de oxigênio, ainda que tipicamente quimioorganotróficas, algumas estirpes possuem uma hidrogenase ativa que permite a elas crescer quimiolitotroficamente em uma atmosfera de hidrogênio, dióxido de carbono e baixos níveis de oxigênio (JORDAN, 1984). Normalmente causa a formação de nódulos efetivos nas raízes de *Glycine* e em *Macroptilium atropurpureum* (JORDAN, 1984).

Foram realizadas seleções das estirpes de *Bradyrhizobium* para utilização em inoculantes de soja. Em primeira instância, no Brasil, foram introduzidos inoculantes comerciais, entretanto, não houve sucesso na inoculação devido às condições inapropriadas de armazenamento, atrasos na liberação alfandegária e interação entre as estirpes. Com isso, partir da década de 1960, a comissão nacional de soja no Brasil determinou que a FBN era uma importante característica que precisava ser levada em consideração e para que tal objetivo fosse alcançado foi iniciado a busca por estirpes adequadas para inoculação (SILVA, 1948; LOPES, GIARDINI, 1981). A lista de estirpes recomendadas para uso em inoculantes comerciais foi primeiramente publicada em 1956 (HUNGRIA et al., 1994, 2005). No Brasil, a SEMIA 566 foi a primeira estirpe “selecionada”, isolada de um nódulo de soja cultivar Hardee, pertencente ao mesmo sorogrupo da USDA 123 e nodulava eficientemente todos os cultivares de soja da época. Foi recomendada para uso comercial de 1966 a 1978, contribuindo notavelmente para o estabelecimento da cultura no sul do Brasil. Uma segunda estirpe, SEMIA 587, foi isolada em 1967, sendo recomendada para inoculação junto com a SEMIA 566 e SEMIA 543 de 1968 a 1975 (HUNGRIA et al., 2006).

Com a expansão da cultura de soja pelo Cerrado, um novo problema surgiu. Inicialmente, alguns estudos atribuíram o insucesso da inoculação nesses solos à criação de um ambiente desfavorável ao rizóbio, causado pelo aumento na população de actinomicetos produtores de antibióticos, que ocorria pelas alterações na microbiota dos solos, provocados pelas práticas agrícolas da adubação e a calagem (COELHO; DROZDOWICZ, 1979 apud VARGAS; HUNGRIA, 1997).

Várias combinações de estirpes foram sugeridas, de 1976 a 1978, incluindo as SEMIA 566, SEMIA 527 e SEMIA 532. Contudo, o desempenho foi insatisfatório. Provavelmente, devido às condições do bioma, ou seja, altos períodos de seca, elevada temperatura do solo (superior a 40°C), baixa umidade do solo, pouca fertilidade, acidez do solo e toxicidade por alumínio (HUNGRIA et al., 2006).

Outra estirpe promissora, a 29W (= SEMIA 5019), foi isolada em

1979. Além de ser altamente competitiva (PERES, 1979), a estirpe provou ser tolerante a antibióticos produzidos pela população local de actinomicetos (SCOTTI et al., 1982), considerado um importante fator em sua performance. Em 1979 a combinação da SEMIA 587 e 29W foi recomendada para uso no Cerrado (PERES, 1979; VARGAS; SUHET, 1980; VARGAS et al., 1981).

Uma década mais tarde, novos genótipos de soja de maior produtividade surgiram, e com isso necessitaram maiores demandas de N. Então foi necessária uma nova etapa de seleção de estirpes que tivessem excelente capacidade de fixação de nitrogênio, combinada com tolerância às condições estressantes do Cerrado. Após vários ensaios, a estirpe CPAC 15 (= SEMIA 5079) foi identificada e mostrou ter a capacidade de aumentar a produtividade da soja em comparação com as estirpes SEMIA 587 e 29W (VARGAS; DENARDIN; VAN BERKUM, 1992, 1993, 1994; PERES et al. 1993; HUNGRIA; VARGAS, 1996).

Uma segunda seleção de estirpes foi iniciada no mesmo período. A estirpe CB 1809 (= SEMIA 586, =USDA 136) foi isolada nos EUA e recomendada para uso em inoculantes comerciais no Brasil em 1977. Todavia, a baixa nodulação do cultivar IAC-2 (PERES, 1979) fez com que esta estirpe fosse removida da lista no ano seguinte. Novas seleções e extensivos testes de campo levaram à identificação da estirpe CPAC 7 (= SEMIA 5080), uma variante da CB 1809, com maior capacidade competitiva. Posteriormente, em 1992, a CPAC 7 e a CPAC 15 foram incluídas na lista de estirpes autorizadas para uso em inoculantes comerciais para a soja. (VARGAS et al., 1992; PERES et al., 1993).

Baseado nas análises de sequências de rRNA 16S realizadas por Chueire et al. (2003), foi observado que a estirpe CPAC 15 difere em apenas um par de base da SEMIA 566. A CPAC 15 produz mais muco (NISHI et al., 1996; BODDEY; HUNGRIA, 1997), aumenta o número de pelos radiculares (fenótipo Hai) (HUNGRIA et al., 1998), apresenta resistência a vários antibióticos, alta síntese de ácido indolacético (AIA) *in vitro* e fenótipo negativo da hidrogenase (Hup⁻) quando comparada com a SEMIA 566.

Assim como a CPAC 15, a CPAC 7 também pertence ao gênero *Bradyrhizobium*. A estirpe CPAC 7 apresenta fenótipo Hup⁺ (NISHI et al., 1996), alta eficiência no processo de fixação de N (DÖBEREINER et al., 1970; NEVES et al., 1985), competitividade (HUNGRIA et al., 1998; HUNGRIA; VARGAS, 2000) e também é produtora de muco (NISHI et al., 1996; BODDEY; HUNGRIA, 1997).

Além das características morfológicas comuns às colônias de *Bradyrhizobium*, as estirpes CPAC 15 e CPAC 7 apresentam bordas lisas e produzem álcalis quando incubadas em meio YMA (VINCENT, 1970) a 28°C. A estirpe CPAC 7 é resistente aos antibióticos ($\mu\text{g/ml}$): tetraciclina, 100; ácido nalidíxico, 50; rifampicina, 500; streptomina, 100. Além destes antibióticos mencionados a estirpe CPAC 15 também é resistente à carbenicilina, 500; e eritromicina, 250; entretanto não é resistente a cloranfenicol, 500 (NISHI, et al., 1996).

Todos os parâmetros fisiológicos avaliados *in vivo* e *in vitro* já descritos anteriormente para ambas as estirpes podem ser somados com os resultados da síntese de ácido indolacético (AIA), fenótipo Hup, toxicidade por rhizobitoxina e nodulação do cultivar com o alelo Rj₄. É importante destacar que os estudos de toxicidade, realizados por Nishi e colaboradores (1996) tiveram como parâmetros de avaliação o alongamento da raiz de alfafa, cultivar Crioula, ou clorose nas folhas de soja do cultivar Lee e cultivar BR-16. Todas as diferenças encontradas entre as estirpes CPAC 7 e CPAC 15 descritas aqui estão sumarizadas na Tabela 1 (NISHI, et al., 1996).

Comparando a estirpe CPAC 7 com CPAC 15 podemos destacar que a CPAC 7 possui maior eficiência no processo de fixação de nitrogênio (mg de N fixado por grama de nódulo) e a CPAC 15 maior competitividade e capacidade saprofítica (SANTOS et al., 2011).

Tabela 1 - Parâmetros fisiológicos e morfológicos avaliados *in vivo* e *in vitro* para as estirpes CPAC 7 e CPAC 15.

Estirpes	Parâmetros												AIA (μM)	Hup	RT			Rj ₄
	Diâmetro (mm)	Cor	Forma	Muco	Antibióticos							Alfafa Crioula			Soja Lee	Soja BR-16		
					Tet	Chl	Car	Ery	Nal	Rif	Str							
CPAC 7	0,5-2,0	Branca	Circular	Muito	+	-	-	-	+	+	-	13,12	+	-	-	-	+	
CPAC 15	0,5-1,0	Branca	Circular	Muito	+	-	+	+	+	+	+	31,46	-	+	+	+	-	

Fonte: Nishi et al. (1996), adaptado.

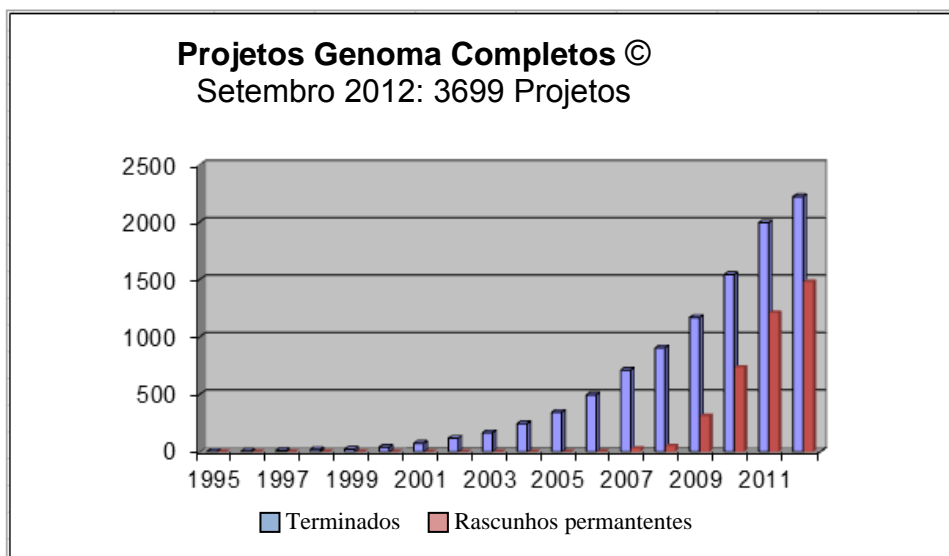
3.5 GENOMAS ESTRUTURAIS E FUNCIONAIS DOS RIZÓBIOS

Os procariotos foram estabelecidos na Terra bilhões de anos antes da primeira célula eucariótica desenvolver-se, e o longo período de adaptação a uma grande variedade de ambientes explica a sua enorme diversidade e versatilidade. O genomas completos e incompletos de 14.666 procariotos estão depositados no banco de dados do Centro Nacional para Informação Biotecnológica (*National Center for Biotechnology Information, NCBI*) (NCBI, 2012). A variabilidade em tamanho, número, densidade e organização dos genes em operons foram demonstradas e cada novo genoma fornece informações inesperadas e, às vezes, intrigantes (NCBI, 2012).

No atual cenário, a taxa de geração de dados está crescendo de forma exponencial. Em 2009, o Laboratório Europeu de Biologia Molecular registrou o crescimento da taxa de submissão à base de 200% ao ano, além disso, a produção de genomas completos avança rapidamente o nosso entendimento de biologia e evolução (SOUTHAN; CAMERON, 2009).

De acordo com a estatística apresentada no banco de dados de genomas online (*Genome Online Database, GOLD*) de 1995 a 2011 houve um crescimento exponencial nestes tipos de estudos, em 2012 houve 3699 projetos de genomas completos no mundo (GOLD, 2012). (Figura 6).

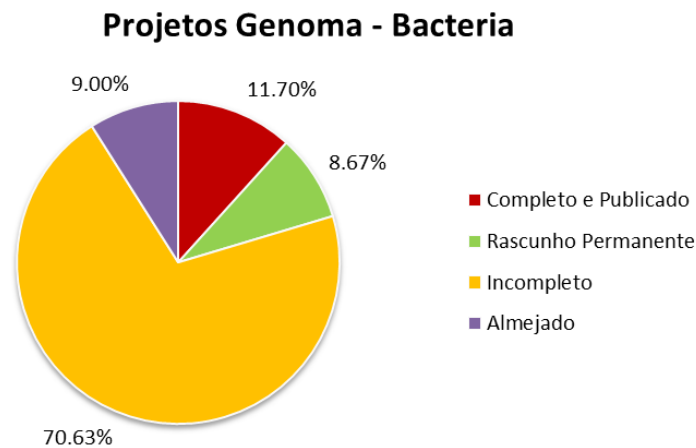
Figura 6 - Evolução dos projetos genoma de 1995 a 2012 indicando a curva de crescimento exponencial.



Fonte: GOLD, 2012, (Modificado).

Considerando os estudos com bactérias, pode-se afirmar que nesse cenário, quase toda sua totalidade (70,63%) estão incompletos e apenas 11,70% estão completos e publicados (Figura 7).

Figura 7 - Projetos Genoma referentes ao domínio Bacteria, indicando os seus status atuais.



Fonte: GOLD, 2012 (Modificado).

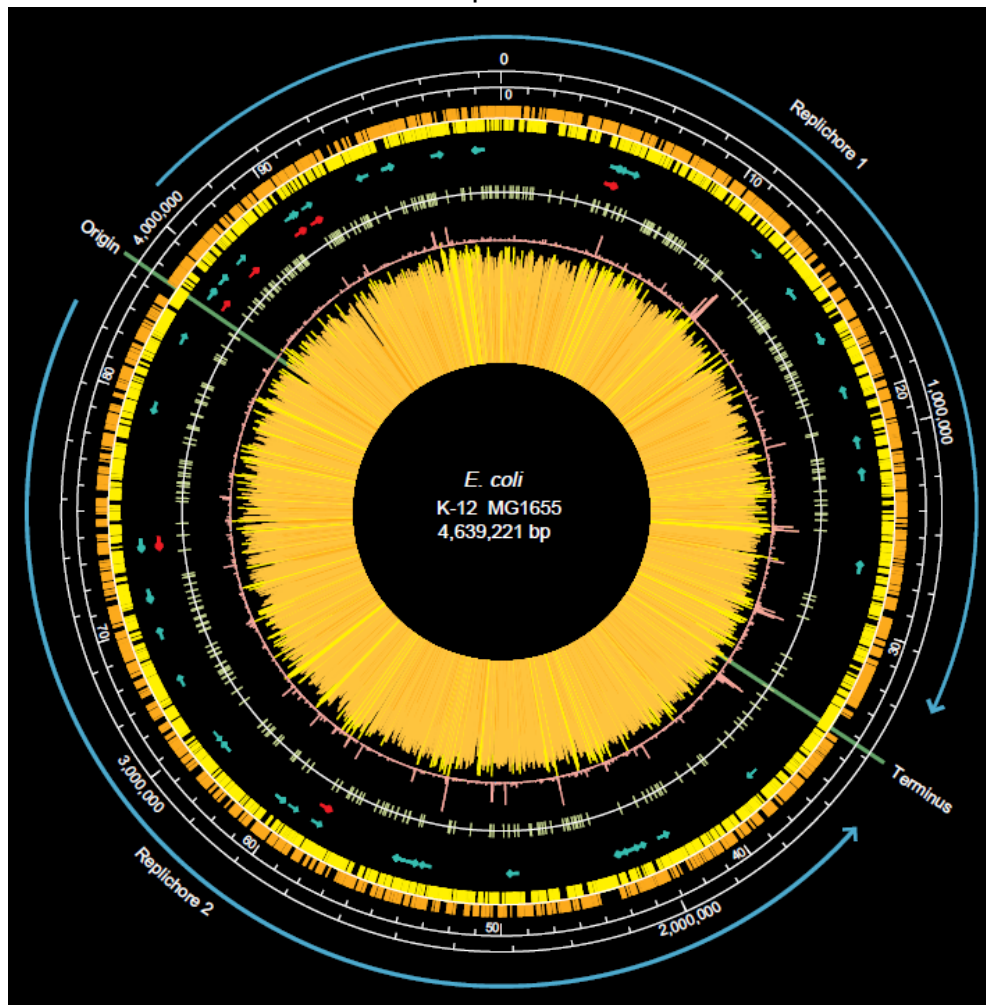
A disponibilidade de mapas detalhados e sequências nucleotídicas de genomas de muitas espécies permitem aos cientistas realizar buscas computacionais por sequências que codificam enzimas com as atividades desejadas, isolar essas sequências e introduzir no genoma de outras espécies. Já foram criadas, por exemplo, plantas resistentes a herbicidas e insetos, assim como plantas sintetizadoras de anticorpos, drogas, altos níveis de vitamina e até mesmo plástico (SNUSTAD; SIMONS, 2000).

No ano de 1995, *Haemophilus influenza* foi a primeira bactéria a ter o genoma inteiramente sequenciado (FLEISCHMANN et al., 1995). Mais tarde, em 1998, as sequências completas de 18 genomas bacterianos estavam disponíveis em bancos de dados públicos. Genomas de variados tamanhos, desde 580.070 pb para *Mycoplasma genitalium*, 4.411.529 pb para *Mycobacterium tuberculosis*, que causou mais mortes humanas do que qualquer outra bactéria infecciosa, a 4.639.221 bp para *Escherichia coli*, o microrganismo celular melhor conhecido cientificamente (SNUSTAD; SIMONS, 2000).

Dos genomas bacterianos sequenciados, a sequência de *E. coli* (Figura 8) foi, indubitavelmente, a mais excitante entre os geneticistas. *E. coli* é o organismo celular melhor estudado e compreendido do planeta. Geneticistas, bioquímicos e biólogos moleculares têm utilizado *E. coli* como organismo modelo por décadas. Muito do que é conhecido sobre genética bacteriana foi aprendido por meio de pesquisas com *E. coli*, então, a publicação do genoma completo de *E. coli* em 1997 foi um marco significativo na história da genética (SNUSTAD; SIMONS, 2000).

As sequências que podem dar origem a possíveis proteínas são chamadas de *open read frames* ou ORFs. Uma ORF é uma sequência nucleotídica que começa com um códon de iniciação (usualmente AUG), continua com os códons em triplete especificando os aminoácidos e termina com um de seus três códons de terminação (usualmente UAA, UAG ou UGA). Uma vez que a sequência do genoma completo está disponível, buscas podem ser realizadas por meio computacional com o objetivo de se comparar com outros genomas sequenciados. Tais comparações de sequências podem, frequentemente, ser usadas para a obtenção de inferências sobre a função do gene. Uma vez que muito é conhecido sobre as funções do gene de *E. coli*, comparações de outros genomas bacterianos sequenciados com o genoma de *E. coli* podem trazer várias informações. Por conseguinte, pela comparação de potenciais codificantes de genomas inteiros, os geneticistas podem inferir sobre vários aspectos do metabolismo, crescimento e interações com ambiente de várias bactérias. Tais informações são inestimáveis para o desenvolvimento de estratégias nos estudos científicos (SNUSTAD; SIMONS, 2000).

Figura 8 - Mapa do cromossomo de *Escherichia coli* K-12. As setas azuis marcam a metade do cromossomo atravessado por duas forquilhas de replicação indicando a origem e o término. O círculo concêntrico exterior indica a posição dos genes que são transcritos em sentido horário (laranja) para uma fita ou sentido anti-horário (amarelo) da fita complementar. A grande porção amarela no centro é um histograma em que o comprimento de cada faixa é proporcional à aleatoriedade da utilização do códon dentro de cada sequência codificante.



Fonte: Blattner et al. (1997).

Neste contexto, as ciências genômicas combinadas com a bioinformática permitem compreender a base genética desses organismos (MIR, 2004). Genômica é o campo da genética que tenta compreender o conteúdo, a organização, a função e a evolução da informação genética contida em genomas inteiros; consiste, também, em dois campos complementares: genômica estrutural e genômica funcional (PIERCE, 2004). O campo da genômica é de grande importância para a biologia moderna e, as informações resultantes da pesquisa nesse campo deram contribuições significativas para a saúde humana, agricultura e várias outras

áreas. Além disso, as comparações de sequências de genomas de organismos diferentes estão levando a uma melhor compreensão da evolução e da história da vida (PIERCE, 2004).

A genômica estrutural tem como objetivo entender como as proteínas funcionam, através da elucidação de suas estruturas moleculares tridimensionais e a relação com a estrutura-função biológica (MIR, 2004), ela determina a organização e a sequência da informação genética contida dentro de um genoma e relaciona-se ao sequenciamento e à compreensão do conteúdo dos genomas. Em geral, uma das primeiras etapas na caracterização de um genoma é a de preparar mapas genéticos e físicos de seus cromossomos. Esses mapas fornecem informações sobre as localizações relativas dos genes, marcadores moleculares e segmentos cromossômicos, que comumente são essenciais para o posicionamento de segmentos cromossômicos e para o alinhamento de trechos de sequenciamentos de DNA em uma sequência total do genoma (PIERCE, 2004).

A genômica funcional caracteriza a função das sequências elucidadas pela genômica estrutural; representa, em essência, a sondagem das sequências genômicas quanto ao significado, identificando genes, reconhecendo a sua organização e compreendendo a sua função. As metas da genômica funcional incluem a identificação de todas as moléculas de RNA transcritas de um genoma (o transcriptoma) e todas as proteínas codificadas pelo genoma (o proteoma). A genômica funcional explora tanto a bioinformática, quanto os enfoques experimentais baseados em experimentos, na busca para definir a função de sequências de DNA (MIR, 2004).

Com isso, os projetos de genômica estrutural aumentam as informações disponíveis para diferentes famílias de proteínas homólogas, bem como, as hipotéticas, com função biológica não determinada. Com isso, o sequenciamento representa apenas o primeiro passo nas ciências genômicas, sendo acompanhado de intensa atividade de bioinformática e de experimentação para a identificação dos genes ali codificados e suas funções (genômica funcional) (MIR, 2004).

Um fenômeno muito importante e com grande influência nos estudos de genômica e filogenia entre bactérias é a transferência gênica horizontal (*horizontal gene transfer – HGT*), relatada pela primeira vez por Victor J. Freeman (1951), que demonstrou que a transferência de um gene viral para a bactéria *Corynebacterium diphtheriae* era capaz de criar uma forma virulenta a partir de

cepas não-virulentas; todavia, devido às limitações da época, esse fenômeno permaneceu inexplicado. Atualmente, sabe-se que este evento, apesar de ser raro em eucariotos é muito comum entre bactérias. Esse é o caso, por exemplo, de *E. coli*, que parece ter adquirido pelo menos 18% do seu genoma de outras espécies nos últimos 100 milhões de anos. Essa troca de DNA fornece, à bactéria receptora, diversas vantagens adaptativas e gera novas estirpes (ALBERTS et al., 2006).

Outros elementos que contribuem para a dinâmica dos genomas das bactérias são os elementos transponíveis ou *transposons*, que conferem a habilidade de que as sequências de DNA se movimentem de um lugar para outro no genoma. Existem três tipos principais: as sequências de inserção ou elementos IS, os *transposons* compostos e os elementos Tn3. Esses três tipos de *transposons* diferem em tamanho e estrutura. As sequências de inserção são os mais simples, contendo somente genes que codificam proteínas envolvidas na transposição. Os *transposons* compostos e os elementos Tn3 são mais complexos, contendo alguns genes que codificam produtos não relacionados ao processo de transposição. Os elementos transponíveis estão presentes nos genomas de muitos organismos e desempenham um papel importante na evolução de muitas espécies (SNUSTAD; SIMONS, 2000).

As sequências de inserção, ou elementos IS, são assim chamadas porque conseguem se inserir em diferentes sítios de cromossomos bacterianos e plasmídeos; são compactamente organizadas. Tipicamente, consistem de menos de 2.500 pb e contêm apenas genes cujos produtos estão envolvidos em processar ou regular a transposição. Cada tipo de elemento IS possui uma sequência curta idêntica (ou quase idêntica) nas extremidades. Já que essas sequências são sempre invertidas em orientação de uma extremidade em relação a outra, são denominadas de repetições terminais invertidas e possuem tamanho de 9 a 40 pb. Cada IS codifica uma proteína que é necessária para a transposição, chamada transposase, que se liga no final de cada elemento, ou próximo a ele, e cliva a fita de DNA, tornando-a livre para ser inserida em uma nova posição (SNUSTAD; SIMONS, 2000). Quais as consequências que a movimentação das sequências transponíveis causam nas novas localidades do genoma? Uma possível vantagem ao organismo reside nos rearranjos genéticos, seja direta, ou indiretamente. O evento de transposição, por si só, causa deleção ou inversão, ou pode levar ao transporte de uma sequência para uma nova localização (LEWIN, 1983).

Os rizóbios são o grupo de microrganismos que possuem maior importância na agricultura, e a simbiose entre leguminosas e rizóbios promove a maior eficiência no sistema de fixação do nitrogênio dentre as formas simbióticas. Pesquisas extensivas e profundas em nível fisiológico e molecular e têm sido feitas em rizóbios uma vez que esses organismos são, atualmente, de grande interesse científico (GAO et al., 2008). Nesse contexto, a genética vem podendo auxiliar diversos fatos relacionados com a fixação biológica de nitrogênio. As ilhas simbióticas do genoma podem ser definidas por um conjunto de genes provenientes de transferência horizontal de outras linhagens, com funções relacionadas a simbiose. (GÖTTFERT et al., 2001) O operon *nif*, é constituído de genes que são responsáveis pela fixação biológica de nitrogênio estão presentes no grupo dos rizóbios (BRILL, 1980).

Em *Bradyrhizobium*, os genes envolvidos no evento de nodulação (genes *nod*, *nol* e *noe*), fixação do N₂ (genes *fix*) e síntese e atividade da nitrogenase (genes *nif* e *fix*) estão localizados no cromossomo. A sequência completa do genoma de *B. diazoefficiens* estirpe USDA 110^T confirmou que seu genoma consiste de um único cromossomo de 9.105.828 pb, sem a presença de plasmídeos (KANEKO et al., 2002a,b). O cromossomo contém 8.317 genes codificadores de proteína, além dos genes para RNAs ribossômicos e transportadores. Dos genes putativos, 52% puderam ter suas funções determinadas por comparações com bancos de dados. Um segmento de 681 kb foi identificado como uma possível ilha simbiótica, previamente identificada por Göttfert et al. (2001), e inclui a região cromossômica de 410 kb, onde estão agrupados os genes de nodulação e fixação do N₂. Também foram atribuídos 655 genes codificantes de proteínas para essa região, sendo que 301 deles estão envolvidos na transmissão de DNA e fixação de N₂. Foram anotados 167 genes para transposases, em 104 cópias de sequências de inserção identificadas no genoma, além de 100 outros genes relacionados às transposases localizados na ilha simbiótica (KANEKO et al. 2002a,b).

O gene *nifA* codifica uma proteína requerida para a expressão de todos os outros operons, com exceção do *nifRLA* (ROBERTS; BRILL, 1980). O gene *nifL* parece codificar uma proteína relacionada com a repressão de todo o agrupamento *nif* em função da presença de O₂ (DIXON et al., 1980; HILL et al., 1981). Acredita-se que o gene *nifR* seja o promotor do operon *nifRLA* (ROBERTS; BRILL, 1980); entretanto, ainda não foi possível relacioná-lo com a produção de

alguma proteína. O gene *nifH* codifica a dinitrogenase redutase (HAGEMAN; BURRIS, 1979), os *nifD* e *nifK* codificam as subunidades alfa e beta da dinitrogenase, respectivamente (HAGEMAN; BURRIS, 1979). O gene *nifM* está envolvido na maturação da dinitrogenase redutase (ROBERTS et al., 1978) e os gene *nifV* está relacionado com o processamento pós-traducional da dinitrogenase (DIXON et al., 1980), os produtos dos genes *nifS* e *nifU* estão envolvidos com a maturação desta proteína. Os genes *nifB*, *nifN*, *nifE* são requeridos para o processamento ou síntese de um cofator da dinitrogenase (ROBERTS et al., 1978). Para que haja potencial redutor, os genes *nifF* e *nifJ*, após a transcrição e tradução, proporcionam transporte de elétrons para a nitrogenase (ST JOHN et al., 1975; ROBERTS et al., 1978). O gene *nifQ* é responsável pelo transporte de molibdato, que desempenha importante papel para os genes *nif* (IMPERIAL, et al., 1984). Já o gene *nifW* faz parte do sistema que protege a proteína de ligação aos cofatores Mo-Fe dos danos causados pelo O₂ (KIM; BURGESS, 1996).

O genoma completo da estirpe USDA 6 de *B. japonicum* (a estirpe tipo dessa espécie) publicado por Kaneko e colaboradores (2011) apresentou 9.207.384 pb, com uma média de 63.67% no conteúdo GC. Foram encontradas duas cópias do *cluster* gênico do RNA ribossômico (RNAr) e 51 genes para RNA transportador (RNAt), a maioria dispersos pelo genoma e transcritos em pequenas unidades. A maioria desses RNAt apresentaram similaridade (97% de identidade ou mais) com a estirpe USDA 110. Para realizar a predição de potenciais genes codificadores de proteínas foi utilizado o programa MetaGeneAnnotator, o qual faz uma varredura no genoma buscando por códons responsáveis por aminoácidos iniciais da cadeia polipeptídica. Cada um desses genes começou com um dos três códons: ATG (7.183 genes encontrados), GTG (1.084 genes encontrados) e TTG (562 genes encontrados). Quando feita uma comparação com o banco de dados COG (*Clusters of Orthologous Groups*), foram encontrados 5.859 possíveis genes, classificados em 18 categorias (Tabela 2).

Tabela 2 - Classificação de genes codificadores de proteínas utilizando o banco de dados COG (*Clusters of Orthologous Groups*) para o genoma de *Bradyrhizobium japonicum* estirpe USDA6.

COG code	USDA110	USDA6 ^T	
	entire genome	entire genome	GIs ^a
Information storage and processing			
J: Translation, ribosomal structure, and biogenesis	197 (2.4%)	203 (2.3%)	1 (0.2%)
K: Transcription	468 (5.6%)	487 (5.5%)	10 (1.7%)
L: Replication, recombination, and repair	333 (4.0%)	290 (3.3%)	30 (5.1%)
Cellular processes and signaling			
D: Cell cycle control, cell division, and chromosome partitioning	33 (0.4%)	30 (0.3%)	0 (0.0%)
T: Signal transduction mechanisms	330 (4.0%)	346 (3.9%)	30 (5.1%)
M: Cell wall/membrane/envelope biogenesis	263 (3.2%)	245 (2.8%)	9 (1.5%)
N: Cell motility	243 (2.9%)	233 (2.6%)	9 (1.5%)
O: Posttranslational modification, protein turnover, and chaperones	213 (2.6%)	216 (2.4%)	16 (2.7%)
Metabolism			
C: Energy production and conversion	419 (5.0%)	411 (4.7%)	4 (0.7%)
G: Carbohydrate transport and metabolism	389 (4.7%)	409 (4.6%)	4 (0.7%)
E: Amino acid transport and metabolism	717 (8.6%)	710 (8.0%)	14 (2.4%)
F: Nucleotide transport and metabolism	93 (1.1%)	89 (1.0%)	0 (0.0%)
H: Coenzyme transport and metabolism	183 (2.2%)	179 (2.0%)	0 (0.0%)
I: Lipid transport and metabolism	289 (3.5%)	271 (3.1%)	2 (0.3%)
P: Inorganic ion transport and metabolism	281 (3.4%)	296 (3.4%)	5 (0.8%)
Q: Secondary metabolites biosynthesis, transport and catabolism	356 (4.3%)	356 (4.0%)	11 (1.9%)
Poorly characterized			
R: General function prediction only	583 (7.0%)	599 (6.8%)	26 (4.4%)
S: Function unknown	444 (5.3%)	489 (5.5%)	19 (3.2%)
not in COGs	2,483 (29.9%)	2,970 (33.6%)	402 (67.9%)
Total	8,317	8,829	592

The percentage of assigned genes out of the total number of genes is shown in parentheses. ^a The number of predicted genes assigned inside fifteen GIs of USDA6^T is shown.

As estirpes USDA 110¹ de *B. diazoefficiens* e USDA 6¹ de *B.*

Fonte: Kaneko et al. (2011).

Japonicum compartilham 8,5% de genes idênticos, 68,5% com significativa similaridade e 23% sem similaridade, segundo análises em nível nucleotídico com BLASTN (KANEKO et al., 2011).

Sabe-se que as bactérias fixadoras de nitrogênio possuem dois grupos distintos de genes: *nif* e *fix*. Os genes *nif* aparentemente estão envolvidos com a síntese da nitrogenase, enquanto os *fix* produzem as moléculas que regulam o processo de fixação de N₂. No processo da formação de nódulos em plantas leguminosas, induzidos por bactérias pertencentes, principalmente, à ordem Rhizobiales, também é fundamental a participação dos genes de nodulação. Os genes de nodulação são encontrados na bactéria, sendo responsáveis pela

produção da proteína que recebe o sinal químico da planta hospedeira (em sua maioria flavonoides) e pela produção das enzimas que sintetizam os fatores de nodulação (MORGANTE, 2003).

No caso do Brasil, até o presente momento foi realizado apenas um genoma de *B. japonicum*, com a análise parcial da CPAC 15, autorizada para o uso de inoculantes comerciais de soja desde 1992. Godoy et al. (2008) sequenciaram aproximadamente 16% do genoma de *B. japonicum* da estirpe CPAC 15, com o objetivo de detectar as principais classes de genes e suas similaridades com outros genomas rizobiais. Foi realizada uma classificação funcional das sequências codificantes do DNA (ORFs) pelo programa COG. Os resultados mostraram que 18% (n=797) das ORFs foram distribuídas em 20 categorias (tabela 3), percentual similar ao relatado para a estirpe USDA 110^T (Kaneko et al., 2002a).

Tabela 3 - Número de sequências codificantes do DNA (ORFs) de *B.japonicum* estirpe CPAC 15 classificada em categorias pelo COG (*Clusters of*

COG functional category	# CDSs	COG functional category	# CDSs
J Translation	31	O Post-translational modification, protein turnover, chaperones	17
A RNA processing and modification	0	C Energy production and conversion	43
K Transcription	63	G Carbohydrate transport and metabolism	57
L Replication, recombination and repair	27	E Amino acid transport and metabolism	105
B Chromatin structure and dynamics	0	F Nucleotide transport and metabolism	3
D Cell cycle control, mitosis and meiosis	2	H Coenzyme transport and metabolism	25
Y Nuclear structure	0	I Lipid transport and metabolism	47
V Defense mechanisms	12	P Inorganic ion transport and metabolism	53
T Transduction mechanisms	30	Q Secondary metabolites biosynthesis, transport and catabolism	45
M Cell-wall/membrane biogenesis	29	R General function prediction only	130
N Cell motility	5	S Function unknown	64
Z Cytoskeleton	0		
W Extracellular structures	0	Total	797
U Intracellular trafficking and secretion	9		

Fonte: Godoy *et al.* (2008).

Segundo Godoy et al. (2008), muitas características no genoma da estirpe CPAC 15 são consistentes com a sua capacidade saprofítica, como a adaptação a uma ampla faixa de condições e com a capacidade de nutrição da estirpe em solos com baixa fertilidade, oitenta e uma ORFs da CPAC 15 codificam proteínas regulatórias e 21 delas estão relacionadas com reguladores de dois componentes.

Segundo Göttfert, Grob e Hennecke (1990), o sistema regulatório de dois componentes *nodV/nodW* foi localizado na ilha simbiótica da estirpe USDA 110^T. A NodV responde ao isoflavonóide genisteína e, subsequentemente, fosforila o seu regulador cognato NodW, que por sua vez é requerido para regular

positivamente para a transcrição de um ou vários genes envolvidos na nodulação de plantas hospedeiras.

O gene *nodW* de *B. japonicum* é o ativador chave de quase todos os genes induzidos por ginesteína (LANG, et al. 2008) e é crítico para a nodulação da planta hospedeira. Sanjuan et al. (1994) confirmaram em experimentação, que a proteína NodW é essencial para a indução de genes *nod* comuns em resposta a isoflavonóides. O *NodD* liga-se aos flavonóides e regula a expressão dos genes *nodABCFE*, os quais codificam outras proteínas de nodulação, ou seja, a proteína NodD é um regulador transcricional positivo; os genes *nod* são essenciais para a função do promotor e flavonóides (usualmente exsudatos radiculares de plantas) são indutores (SCHLAMMAN; OKKER; LUGTENBERG, 1992).

A estirpe CPAC 15 apresentou 5,3% das ORFs relacionadas à produção e conversão de energia, demonstrando sua capacidade para fermentação, respiração aeróbica e anaeróbica, indicando, então, uma ampla capacidade de obter energia sobre várias condições, desde solos bem aerados a solos alagados. Oxidoredutases e desidrogenases podem ajudar a estirpe a explorar diversas fontes de energia e, na ausência de oxigênio, formato, fumarato e nitrato podem ser utilizados como acceptor final de elétrons. Muitos são os genes relacionados à degradação xenobiótica, os quais são importantes e refletem o potencial biotecnológico da estirpe e sua capacidade de sobreviver em uma ampla faixa de fontes de carbono. A maioria das ORFs estão relacionadas à degradação do benzoato via coenzima-A e a via de hidroxilação (GODOY et al., 2008). Apesar da pequena cobertura do genoma da estirpe CPAC 15, vários novos genes de interesse foram descobertos e relatados (GODOY et al., 2008), indicando que várias informações importantes podem ser obtidas pelo sequenciamento completo.

Atualmente, a maior parte dos genomas de procariontes é obtida por pirosequenciamento. Esse método permite a obtenção de 25 milhões de bases em 4 horas, entretanto, com leituras mais curtas do que no sistema de sequenciamento tradicional (SANGER), sendo necessária a elaboração de *primers* para o fechamento de *gaps*, sequenciamento de regiões de baixa qualidade e anotação do genoma.

Recentemente, foi publicado por Souza e colaboradores (2012) o *draft* do genoma de *Bradyrhizobium elkanii* estirpe semia 587, também utilizada em inoculantes comerciais no Brasil. Foram utilizados *reads* de Illumina e Sanger e o tamanho final da montagem foi de 8.8 Mpb.

Um estudo recente feito por Okubo et al. (2012) mostrou à comunidade científica que a bactéria *Bradyrhizobium* sp. estirpe S23321, apesar de ser filogeneticamente próxima à espécie *B. japonicum* não apresenta fenótipo de nodulação. Com isso, essa estirpe se tornou um bom elo para estudos de questões de como os membros do gênero *Bradyrhizobium* conseguiram obter diversas funcionalidades em nível bioquímico. Uma vez sequenciada a estirpe S23321 foram realizadas análises de genômica comparativa entre as estirpes USDA 110^T e ORS 278 e outras estirpes como a BTAi1, por estas estirpes serem filogeneticamente próximas, após à construção de árvores filogenéticas utilizando o RNAr 16S e a região ITS (entre o RNAr 16S e 23S). Os resultados mostraram que a estirpe S23321 não possuía a ilha simbiótica. A análise feita com BLASTP contra a estirpe USDA 110^T detectou genes relacionados à FBN, mas nenhum relacionado à nodulação (*nodDYABCSUIJ*). A genômica comparativa sugere que o genoma de S23321 é um tipo ancestral que precede a aquisição da ilha simbiótica, baseado nas análises com gene tRNA-Val e a ausência dos genes de nodulação. Interessantemente, o genoma de S23321 contém um conjunto completo de genes fotossintéticos similar a *Rhodopseudomonas palustres* em termos de organização e homologia. Entretanto, nas árvores filogenéticas construídas para o RNAr 16S e região ITS, a estirpe S23321 não permaneceu no mesmo cluster das fotossintéticas *R. palustres* e *Bradyrhizobium* sp. estirpe BTAi1. Com isso, os autores sugeriram duas explicações para essa discrepância. Uma é a de que bradyrhizobios não noduladores, como é o caso da estirpe S23321, possam ter carregado o protótipo fotossintético e *clusters* dos genes *nif* similares à ORS 278, BTAi1 e *R. palustres*. Após a inserção da ilha simbiótica, esse cluster gênico foi deletado do genoma. Outra possibilidade é a de que o genoma de S23321 possa ter adquirido *clusters* de genes *nif* ou fotossintéticos via transferência horizontal de uma linhagem diferente de bactéria. Com isso, a *Bradyrhizobium* sp. S23321 pode ser considerada um organismo modelo para ser utilizado em estudos da evolução simbiótica de bradyrhizóbios.

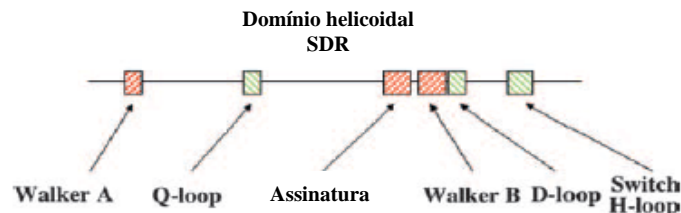
Os resultados citados, mostram, de fato, a importância dos estudos genômicos em rizóbios, a fim de que se possa elucidar aspectos da fisiologia, filogenética e ecologia destes organismos.

3.6 TRANSPORTADORES DE MEMBRANA DO TIPO ABC

Os transportadores do tipo ABC (*ATP-binding cassette (ABC) systems*) constituem uma das maiores superfamílias de sequências parálogas (DAVIDSON, et al. 2008). Esses transportadores usam o ligamento e a hidrólise do ATP para energizar a translocação de diversas variedades de substratos pela membrana, variando de íons a macromoléculas. Os transportadores ABC funcionam tanto como importadores, levando nutrientes e outras moléculas para as células, ou exportadores, que bombeiam toxinas, drogas e lipídios pelas membranas, mas não há exemplos de transportadores que possam funcionar fisiologicamente em ambas as direções. Em procariotos, importadores ABC mediam a absorção de nutrientes como aminoácidos, açúcares e metais essenciais. Os substratos dos importadores ABC variam grandemente em tamanho e natureza química, variando de oligopeptídeos e oligossacarídeos a íons pequenos. Os exportadores ABC são cruciais participantes na exportação de lipídios, ácidos graxos e colesterol e o seu mal funcionamento leva a vários problemas (REES; JOHNSON; LEWINSON, 2009). A sobrevivência celular requer a geração e a manutenção de gradientes de concentração químicos e elétricos entre as membranas celulares e os transportadores do tipo ABC são elementos chave que participam desse processo (REES; JOHNSON; LEWINSON, 2009).

Todos os sistemas ABC compartilham domínio hidrolizador de ATP conservado ou uma proteína (o ABC, também referido como o domínio de ligação ao nucleotídeo - *nucleotide binding domain [NBD]*), isso é inequivocamente formado por três sequências *motifs* curtas (padrão de sequências de nucleotídeos ou aminoácidos muito difundidos): eles são o *Walker A* e *Walker B*, indicativo da presença de um sítio de ligação do nucleotídeo e o *motif* de assinatura, única as proteínas ABC conforme esquematizado na Figura 9 (SCHNEIDER; HUNKE, 1998 apud DAVIDSON et al., 2008).

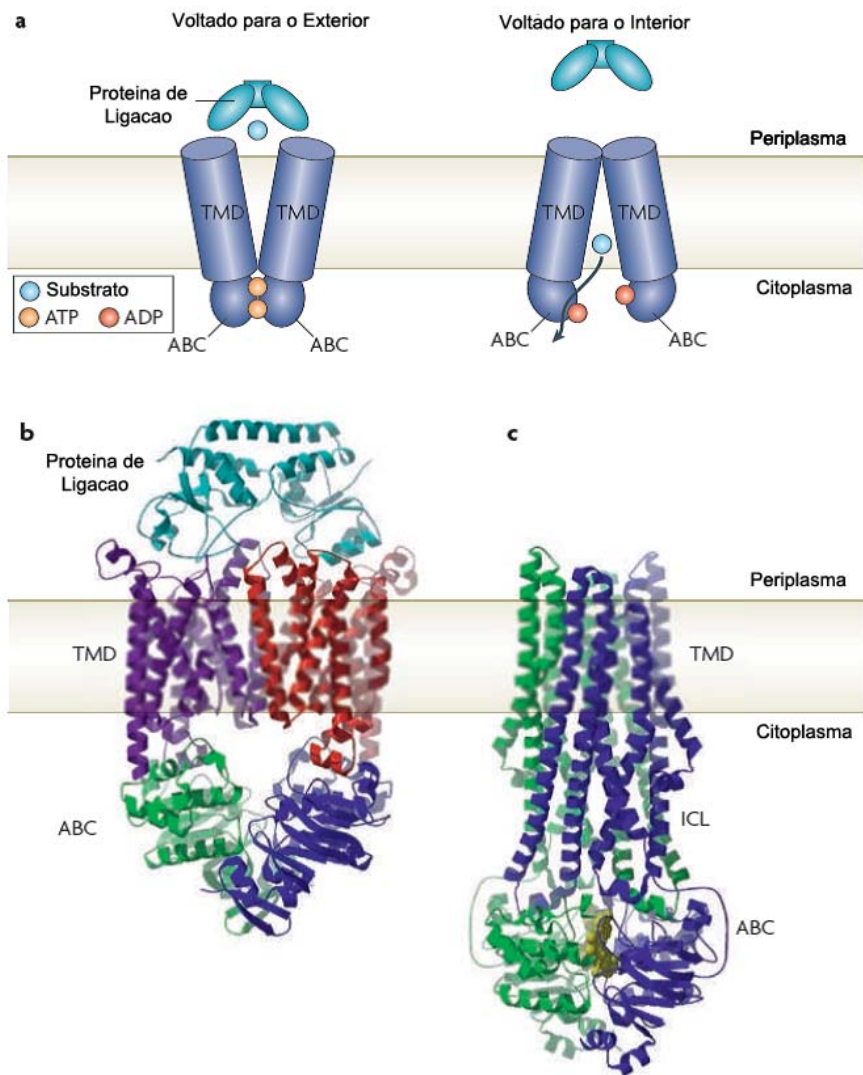
Figura 9 – *Motifs* conservados no ABC. Os três *motifs* característicos encontrados em todas as ATPases de transportadores ABC são as caixas hachuradas em vermelho, que juntas formam a região de ligação de nucleotídeos da alça P da família da ATPase. Outros *motifs* característicos, as regiões *Q-loop* e *H-loop* (também chamada de região de troca), contêm somente um resíduo altamente conservado e são representados por caixas verdes hachuradas. Esses resíduos fazem contato com o γ -fosfato do ATP



Fonte: Davidson et al., 2008, (Modificado).

Os transportadores ABC têm uma arquitetura característica que consiste minimamente de quatro domínios: dois domínios transmembrana (*transmembrane domain – TMD*), que são incorporados na bicamada lipídica e dois ABCs ou domínios de ligação à proteína (*nucleotide binding domain – NBD*), que estão localizados no citoplasma. Para transportadores ABC procarióticos que funcionam como importadores, a translocação do substrato é, também, dependente de um outro componente proteico, uma proteína de ligação de alta afinidade que se associa especificamente com ligantes no periplasma para entregar ao transportador ABC apropriado. Basicamente, ocorre a captura do substrato, que é levado ao vestíbulo de entrada das unidades translocadoras da membrana, como pode ser visualizado na Figura 10 (JONES; GEORGE, 2007; REES; JOHNSON; LEWINSON, 2009).

Figura 10 – A arquitetura molecular de transportadores ABC. a) Um desenho da organização modular dos transportadores ABC, que são compostos de dois domínios transmembrana (*transmembrane domain* – *TMD*) e dois domínios ABC ou domínios de ligação a nucleotídeos (*nucleotide-binding domain*). A proteína de ligação que é requerida pelos importadores também é mostrada. Dois estados conformacionais do transportador ABC – voltado para o interior e voltado para o exterior com o sítio de ligação ao substrato orientado em direção às regiões periplasmáticas (extracelular) e citoplasmática (intracelular), respectivamente – são retratados para demonstrar o mecanismo de acesso alternado do transportador. b) O importador de vitamina B12 de *E. coli*, BtuCDF (Código 2QI9 no *Protein Data Bank* – *PDB*). c) Exportador multidroga Sav1866 de *Staphylococcus aureus* (Código 2ONJ no *Protein Data Bank* – *PDB*). ICL = loop intracelular (*intracellular loop*).



Fonte: Rees, Johnson, Lewinson, 2009, (Modificado).

Os transportadores ABC representam um de vários tipos de transportadores diferentes, operando em bactérias e outros organismos. Transportadores são de crítica importância para organismos vivos. Interessante, a primeira propriedade distintiva de células primitivas foi a permeabilidade seletiva a nutrientes e metabólitos regulada por proteínas de membrana. Transportadores funcionalmente e estruturalmente diferentes têm sido identificados em organismos vivos. Os transportadores ABC possuem um grande impacto na fisiologia bacteriana, e sua disfunção pode ter fortes efeitos deletérios. Enquanto alguns transportadores ABC são claramente dedicados à exportação de fatores de virulência, em condições apropriadas muitos outros transportadores ABC de bactérias podem se tornar importantes para a viabilidade, virulência e patogenicidade (DAVIDSON et al., 2008).

3.7 A BIOINFORMÁTICA: CONCEITOS E APLICAÇÕES NA GENÔMICA

A bioinformática é um campo emergente que consiste na biologia molecular e ciência da computação, que tem como foco o desenvolvimento de banco de dados, algoritmos de busca em computadores, *software* de previsão genética e outras ferramentas analíticas que são usadas para se compreender sequências de dados do DNA, RNA e proteínas. A bioinformática desenvolve e aplica essas ferramentas para a “mineração de dados” (*data mining*), extraindo informações úteis de projetos de sequenciamento (PIERCE, 2004). A bioinformática e a gestão de dados científicos são críticos para apoiar as descobertas das ciências da vida (LACROIX; CRITCHLOW, 2003).

Antes de serem sequenciados, a maioria dos genomas contém poucos genes cujas localizações já foram determinadas. A enorme quantidade de DNA em um genoma somado à possíveis complexidades da estrutura gênica, torna o encontro de genes uma tarefa difícil. Foram desenvolvidos programas de computação para procurar sequências específicas no DNA que estão associadas a alguns genes. Por exemplo, genes codificantes de proteínas são caracterizados por uma matriz de leitura aberta (ORF), que inclui um códon inicial e um códon finalizador na mesma matriz de leitura, como já explicado anteriormente. Nos eucariontes sequências específicas marcam os sítios de corte no começo e no final dos íntrons. Outras sequências específicas estão presentes em promotores imediatamente antecedentes aos códons de início. Já outras estão associadas a

funções, particulares em algumas classes de proteínas. Foram desenvolvidos programas de computação que podem escanear bancos de dados do DNA para essas sequências e identificam os genes com base em sua presença e posição. Alguns desses programas são capazes de examinar bancos de dados de sequências expressas marcadas (ETSs) e sequências de proteínas para ver se há evidências de que um gene potencial é expresso (PIERCE, 2004).

A biologia dependeu da química para fazer grandes avanços, e isso levou ao desenvolvimento da bioquímica. Similarmente, para explicar fenômenos biológicos em nível atômico houve o desenvolvimento da biofísica. A enorme quantidade de dados gerados pelos biólogos—e a necessidade de interpretá-los—requer ferramentas que estão no domínio da ciência da computação, chamada bioinformática (COHEN, 2004).

Vale ressaltar que os programas que foram desenvolvidos para identificar genes com base na sequência de DNA não são perfeitos. Portanto, os números de genes relatados na maioria dos projetos de genoma são estimativas. A presença de múltiplos íntrons, recomposição alternativa, múltiplas cópias de alguns genes e muito DNA não-codificantes entre os genes dificultam a identificação e a contagem precisa dos genes (PIERCE, 2004). Uma dificuldade significativa em obter o DNA genômico inteiro é o fato de que as sequências geradas em laboratório, consistem de sequências relativamente curtas e aleatórias, que têm de ser reagrupadas usando programas de computador. Uma vez que o DNA contém muitas subsequências repetidas, a montagem pode ser complicada. Isso é devido ao fato que o fragmento pode ser colocado em duas ou mais posições do genoma que está sendo montado (COHEN, 2004).

A sequência de nucleotídeos de um gene pode ser usada pra prever a sequência de aminoácidos da proteína que ele codifica. A proteína pode, então, ser sintetizada ou isolada e suas propriedades estudadas para determinar uma função. Entretanto, esse enfoque bioquímico para compreender o funcionamento do gene é demorado e caro. Uma meta importante da genômica funcional tem sido desenvolver métodos computacionais que permitam que o funcionamento gênico seja identificado apenas a partir da sequência de DNA, contornando os processos trabalhosos de isolamento e caracterização de proteínas individuais (PIERCE, 2004).

Um método computacional (em geral o primeiro empregado) para determinar o funcionamento gênico consiste em conduzir uma pesquisa de

homologia, que é baseada em comparação das sequências de DNA e proteínas do mesmo organismo e de organismos diferentes. Diz-se que os genes que são evolutivamente correlacionados são homólogos. De forma geral, a identificação do gene (ou predição gênica) é uma anotação automática, que identifica todas as porções bioquimicamente ativas do genoma por algoritmos de processamento da sequência. Os genes homólogos encontrados em espécies diferentes que evoluem a partir desse mesmo gene em um ancestral comum são chamados de ortólogos, como é o caso das subunidades alfa hemoglobina presentes em camundongos e humanos. Os genes homólogos em geral têm a mesma função ou funções correlatas. Assim, após uma função ter sido atribuída a um determinado gene, ela pode dar um indício quanto à função de um gene homólogo (PIERCE, 2004; RAMSDEN, 2009).

Os bancos contendo genes e proteínas encontrados em uma grande gama de organismos estão disponíveis para pesquisas de homologia. Foram desenvolvidos poderosos programas de computação para escanear esses bancos de dados em busca de determinadas sequências. Um programa de busca de homologia comumente usado é o BLAST (*Basic Local Alignment Search Tool*) (ALTSCHUL et al., 2013), que encontra regiões com similaridade local entre sequências. O programa compara sequências de nucleotídeos ou proteínas com as sequências do banco de dados e calcula a significância estatística da combinação (PIERCE, 2004). O BLAST pode ser usado para inferir relações evolucionárias e funcionais entre as sequências, bem como ajuda a identificar membros das famílias do gene (ALTSCHUL et al., 2013). De um modo similar, os programas de computação podem pesquisar um único genoma quanto a parálogos, que são genes derivados de um gene ancestral via duplicação gênica. Se é encontrado um gene parólogo e sua função já foi previamente identificada, pode ser encontrada uma possível função de um gene desconhecido (PIERCE, 2004).

O arquivamento de dados de bioinformática foi originalmente realizado por grupos de pesquisa individuais motivados por um interesse em ciência associada. À medida que os requerimentos para equipamento e pessoal cresceu – e a natureza das habilidades requeridas mudou, para incluir maior ênfase em computação (LESK, 2002).

Mundialmente, o arquivamento de sequências de aminoácidos é uma parceria tripla do *National Center for Biotechnology Information – NCBI* (EUA), *EMBL Data Library* (Instituto Europeu de Bioinformática, Reino Unido) e *DNA Data*

Bank of Japan (Instituto Nacional de Genética, Japão). Estes grupos trocam informações diariamente. Como resultado, os dados brutos são idênticos, apesar do formato em que são armazenados e a natureza da anotação variar ligeiramente entre eles. Esses bancos de dados curam, arquivam e distribuem sequências de DNA e RNA coletadas de projetos genoma, publicações científicas e pedidos de patente. Para ter a certeza de que esses dados fundamentais estão disponíveis gratuitamente, revistas científicas requerem a deposição de novas sequências nucleotídicas no banco de dados como condição para publicação em um artigo. Condições similares se aplicam a sequências de aminoácidos e estruturas de proteínas (LESK, 2002).

Proteínas complexas em geral, contêm regiões que têm formas ou funções específicas chamadas de domínios. Cada domínio proteico tem um arranjo de aminoácidos comum e, provavelmente, existe um número de domínios proteicos limitado, embora grande, que se misturam e correspondem ao longo do tempo evolutivo para gerar a diversidade proteica vista nos organismos atuais. Muitos domínios proteicos já foram caracterizados, e suas funções moleculares já foram determinadas. A sequência de um gene recém identificado pode ser escaneada para comparar com um banco de dados de domínios conhecidos. Se a sequência gênica codifica um ou mais domínios cujas funções foram determinadas previamente, a função do domínio pode dar informações importantes sobre a possível função do novo gene (PIERCE, 2004).

Outro método computacional para prever o funcionamento da proteína é o perfil filogenético. Nesta metodologia, a presença ou ausência de uma proteína particular é examinada em um conjunto de organismos cujos genomas foram sequenciados. A análise dos padrões de fusão de proteínas (também chamado de método Pedra de Rosetta) é capaz de prever relações funcionais. As proteínas separadas funcionalmente relacionadas em um organismo às vezes existem como uma única proteína fusionada em outro organismo. Uma forma para se identificar função de genes, muito comum em bactérias, é a análise do gene vizinho. Os genes que codificam proteínas funcionalmente relacionadas em geral são organizados em sequência (ou *operons*) nas bactérias (PIERCE, 2004).

As funções sugeridas pelos métodos computacionais, tais como buscas de homologias, perfis filogenéticos, proteínas de fusão e análise de vizinhança não definem com 100% de precisão a função de uma proteína. Esses

métodos conferem pistas sobre possíveis funções que podem ser estudadas por meio de análises detalhadas da bioquímica e localização celular da proteína. Entretanto, esses métodos computacionais e outros são inestimáveis para a determinação das funções de genes revelados nos estudos de genômica (PIERCE, 2004).

Como modelos computacionais de proteínas, células e organismos têm se tornado cada vez mais realistas, muitos pesquisadores biólogos devem migrar da ciência laboratorial tradicional (*wet science*) para ciência computacional (*dry science*). A realização bem sucedida dessa transição, requer acesso a uma grande quantidade de informação de toda a comunidade de pesquisa (LACROIX; CRITCHLOW, 2003).

As tecnologias de sequenciamento incluem um número de métodos que são agrupados amplamente desde a preparação do *template*, sequenciamento e imagem e análise dos dados. A única combinação de protocolos específicos distingue uma tecnologia e outra e determina o tipo de dado produzido por cada plataforma. O método automatizado de Sanger é considerado uma tecnologia de 'primeira geração', e os novos métodos são referidos como sequenciamento de nova geração (*Next Generation Sequencing – NGS*). A grande vantagem oferecida pelo sequenciamento de nova geração é o enorme volume de dados de forma barata (METZKER, 2009). A plataforma 454 GS FLX da Roche, realiza o processo de pirosequenciamento e possui diversas vantagens. Cada corrida varia de 10h a 23h e produz aproximadamente 1.000.000 de *reads* com aproximadamente 700 pb cada com acurácia que varia de 9.995 a 9.997%. (ROCHE, 2013). Apesar da grande quantidade de dados produzida o sequenciamento de Sanger ainda é utilizado, principalmente para fechar gaps no genoma, auxiliando o sequenciamento de regiões do DNA com alta repetição de nucleotídeos.

4 MATERIAL E MÉTODOS

4.1 EXTRAÇÃO DE DNA GENÔMICO

As bactérias foram crescidas em meio YMB (modificado para conter 5 g L⁻¹ de manitol), overnight, a 28°C, centrifugadas a 10.000 rpm por 10 min, ressuspendidas em 25 mL de TES (Tris-EDTA-NaCl) (Tris-HCl 50 mM equilibrado pH 8,0 + EDTA 20 mM pH 8,0 + NaCl 200 mM), centrifugadas por 10.000 rpm por 10 min e ressuspendidas em 10 mL de TES. Foi adicionado 1 mL de SDS (dodecil sulfato de sódio 10%) e os tubos foram invertidos lentamente até que fosse notado que o material ficasse viscoso, o que indica o rompimento das células; após este processo o material foi deixado em repouso por 5 minutos. Foram acrescentados 200 µL de solução estoque 10 mg mL⁻¹ de proteinase K (para um volume final de 10 mL), homogeneizado lentamente, mantido overnight à 37°C. Foi adicionado fenol (equilibrado pH 8,0) na mesma proporção de volume do tubo (10 mL). Em seguida, os tubos foram homogeneizados por 30 min e centrifugados por 15 min a 10.000 rpm em temperatura ambiente para retirada da fase aquosa. O fenol equilibrado foi novamente adicionado à fase aquosa ao mesmo volume dos tubos (10 mL), homogeneizado por 20 minutos, centrifugado por 15 min a 10.000 rpm em temperatura ambiente. A fase aquosa das amostras foi adicionada em 5 ml de clorofórmio, homogeneizado por 10 min, centrifugado a 15 min em 10.000 rpm à temperatura ambiente e a fase aquosa foi retirada. Em seguida, foi acrescentado o mesmo volume de isopropanol gelado ao sobrenadante e homogeneizado lentamente por 5 minutos.

Por fim, os tubos foram centrifugados a 10.000 rpm para obtenção do DNA recuperado em forma de *pellet*. O DNA foi transferido com o auxílio de capilares para tubos com etanol 70% e postos para secar em temperatura ambiente ou 6 horas em estufa a 37°C. O DNA foi então ressuspendido com 100 µL de água milli-Q por 12 horas.

Para quantificação do DNA recuperado, foi realizada eletroforese em gel agarosa 0,7% com corrente de, aproximadamente, 20 volts utilizando 2 µL de DNA. Em seguida, o DNA foi quantificado em espectrofotômetro com o DNA diluído em 1:100, com leitura em 260 nm e 280 nm. Considerando que 1 densidade ótica

equivale a $50 \mu\text{g mL}^{-1}$ de DNA.

4.2 SEQUENCIAMENTO: PIROSEQUENCIAMENTO

O DNA extraído foi submetido ao sequenciamento pela técnica do pirosequenciamento (Roche, 454 flx). Nessa abordagem, as sequências fragmentadas *in vitro* são submetidas à emulsão PCR20, com amplicons capturados sob a superfície de esferas de $28 \mu\text{m}$, denominadas *beads*. Após a emulsão os *beads* são tratados com desnaturantes para remover fitas não tratadas e, então, são sujeitos à hibridização. A sequência do *primer* é hibridizada ao adaptador universal em posição e orientação apropriadas, isto é, imediatamente adjacente ao início da sequência desconhecida. Em seguida é dado início ao pirosequenciamento, onde os *beads* (que carregam as sequências) são pré-incubados com polimerase de *Bacillus stearothermophilus* (BST) e *single strand binding protein* (SSB) e, então, depositados em uma placa que possui poços fabricados na ordem de picomolares (onde as dimensões permitem somente a entrada de um *bead* por poço). *Beads* menores também são adicionados, carregando enzimas imobilizadas requeridas para o pirosequenciamento (ATP sulfúrilase e luciferase). Durante o sequenciamento, há um fluxo ordenado de introdução e remoção de reagentes associados a um dispositivo (CCD) para detecção de sinais. Durante as muitas centenas de ciclos, uma única espécie de nucleotídeo não marcado é introduzida, o que resulta em um evento de incorporação onde o pirofosfato é liberado. Através da ATP sulfúrilase e luciferase, os eventos de incorporação conduzem à geração de explosões de luz, as quais são detectadas pelo CCD correspondente às matrizes específicas dos poços e os sinais são transferidos para o computador, onde são analisados e as sequências determinadas (SHENDURE; JI, 2008). Todo este processo foi realizado no Laboratório Nacional de Computação Científica – LNCC, Pretrópolis – RJ.

4.3 ANOTAÇÃO DO GENOMA

As sequências obtidas foram submetidas à montagem e à anotação no Laboratório Nacional de Computação Científica – LNCC. Para isso foi utilizado o *software* “System for Automated Bacterial Integrated Annotation” (SABIA; ALMEIDA *et al.*, 2004), desenvolvido para integrar vários *softwares* de domínio público. A

montagem das sequências foi realizada utilizando o programa “phred” (EWING; GREEN, 1998; EWING *et al.*, 1998), que também atribui valores de qualidade a cada base e identifica as sequências repetitivas e aquelas pertencentes a vetores. A visualização e edição de montagens das sequências foi realizada usando os programas “Consed” (GORDON; ABAJIAN; GREEN, 1998) e “phrapview”.

Regiões codificadoras de genes presentes no genoma foram, a seguir, identificadas utilizando outras ferramentas de bioinformática, como os programas “Glimmer” e, por comparação com o banco de dados de sequências do GenBank, com o BLAST, do NCBI (NCBI, www.ncbi.nlm.nih.gov). Os programas KEGG (“*Kyoto Encyclopedia of Genes and Genomes*”) (KANEHISA; GOTO 2000), COG (“*Clusters of Orthologous Groups of proteins*”) (TATUSOV *et al.*, 2000), INTERPRO (uma fonte de busca de famílias, domínios e sítios funcionais de proteínas), PSORT (“*Protein Subcellular Localization Prediction*”, desenvolvido para localização das proteínas), TCDB (<http://tcdb.org>) e UniProt (*Swiss-Prot Protein knowledgebase* e TrEMBL (*Computer-annotated supplement*, Swiss-Prot, <http://expasy.org/sprot/>) foram utilizados para a identificação, predição e localização das proteínas.

Todos esses programas estão acoplados ao *software* SABIA. As leituras de bases de boa qualidade definem as “CDSs” (*coding sequences*, ou seja, região codificadora de gene) válidas ou hipotéticas. A seguir, cada CDS, também denominada como ORF (*Open Reading Frame*), foi conferida manualmente, uniformizada quanto às informações e validada, ou não. O tamanho de cada ORF foi estabelecido em maior que ≈ 70 pb e, no caso de sequências mais curtas, foi feita uma observação. As ORF com dúvidas foram manualmente curadas, usando um corte considerando E value de $\approx 10^{-10}$ e ≈ 50 pb de identidade considerando BLASTP. Os seguintes critérios foram aplicados à anotação: ORF classificadas como hipotéticas quando nenhum homólogo for detectado; ORF com alta similaridade são aquelas apresentando alta similaridade com o mesmo gene.

A partir da anotação dos genes, foi feito um estudo funcional comparativo entre as duas estirpes e, também, com outras bactérias já sequenciadas utilizando o *software* SABIA e os dados também foram depositados no RAST (“*Rapid Annotations using Subsystems Technology*”) (AZIZ *et al.* 2008) para a comparação dos genomas.

4.4 ANÁLISE DE BIOINFORMÁTICA FINA DOS TRANSPORTADORES DO TIPO ABC

Com o objetivo de fazer uma análise mais precisa dos transportadores do tipo ABC presentes em *B. japonicum* estirpes CPAC 15 e *B. diazoefficiens* CPAC 7, foi realizado um estudo mais aprofundado.

Para tal processo, foi utilizado o banco de dados *Transporter Classification Database – TCDB* (TCDB, 2012). Este banco contém informações compiladas de, aproximadamente, 10.000 referências e, aproximadamente, 3.000 proteínas distintas de vários tipos de organismos vivos organizadas em mais de 400 famílias de transportadores baseados no sistema TC (TCDB, 2012). O sistema TC consiste em um conjunto representativo de sequências de proteínas, a maioria das quais funcionalmente caracterizada. Estes transportadores são classificados por uma designação de cinco caracteres, do seguinte modo: D1.L1.D2.D3.D4. D1 (um único dígito) corresponde à classe do transportador (ex: canal, carreadora etc.). L1 (uma letra), corresponde à subclasse do transportador. D2 (um número) corresponde à família do transportador (às vezes, na realidade a superfamília). D3 (um número) corresponde à subfamília (ou à família de uma superfamília) no qual o transportador é encontrado. D4 (um número) corresponde ao transportador em sí (SAIER; TRAN; BARABOTE, 2006).

Complementamente ao TCDB foi utilizado o banco de dados de famílias de proteínas (*Protein Family Database – Pfam*). O Pfam é um banco de dados de famílias de proteínas conservadas e apresenta uma coleção de aproximadamente, 12.000 famílias e 100.000 sequências depositadas (FINN, et al., 2009)

Foram escolhidas sequências no banco de dados do TCDB; dentro das superfamílias (ABC 1, 2 e 3) foi escolhida uma sequência representante para cada 10 famílias de transportadores ABC (10 sequências para ABC1, 10 sequências para ABC2 e 8 para ABC3), totalizando um conjunto de 28 sequências. As sequências proteicas foram confrontadas com os dados no PFAM, para que fosse possível encontrar os domínios conservados e comuns às superfamílias de transportador ABC 1,2 e 3. Uma vez encontrados os domínios comuns a todas essas famílias, essa informação foi utilizada para realizar uma busca do proteoma (*in silico*) de *B. japonicum* estirpes CPAC 15 e *B. diazoefficiens* CPAC 7, com o objetivo de descobrir as proteínas transportadoras ABC no genoma de forma mais apurada.

Para isso foi utilizado o programa HMMER, juntamente com o HMMSEARCH, que utilizam os modelos ocultos de Markov (*Hidden Markov Models*) e métodos probabilísticos para procurar similaridades de sequências proteicas (FINN; CLEMENTS; EDDY, 2011). O *e-Value* adotado no programa foi $<1e-4$, para que fossem reportadas as sequências com melhor acurácia. As tabelas geradas foram trabalhadas no Linux em associação com o Excel, resultando em um conjunto de dados que pode ser confrontado com os dados encontrados pela anotação feita no SABIA.

Para a construção dos Diagramas de Venn foi utilizado o programa GeneVenn, aplicativo web para comparar lista de genes (PIROOZNIYA; NAGARAJAN; DENG, 2007).

4.5 CONSTRUÇÃO DO MAPA GENÔMICO

O mapa do genoma circular de *Bradyrhizobium japonicum* foi construído para estirpe CPAC 15, uma vez que seu genoma estava fechado em um único *contig*. Para isso foi utilizado o programa CGView Server (GRANT; STOTHARD, 2008). Os parâmetros foram ajustados para exibir ORFs acima de 300 códon combinadamente para uma melhor visualização. O BLASTN foi realizado contra o genoma completo de *B. diazoefficiens* USDA 110^T, depositado no NCBI por Kaneko e colaboradores em 2002 (BioProject PRJNA57599, PRJNA17) com *cutoff* de 10^{-15} e identidade superior a 95%. Os demais parâmetros usados foram os padrões adotados pelo programa.

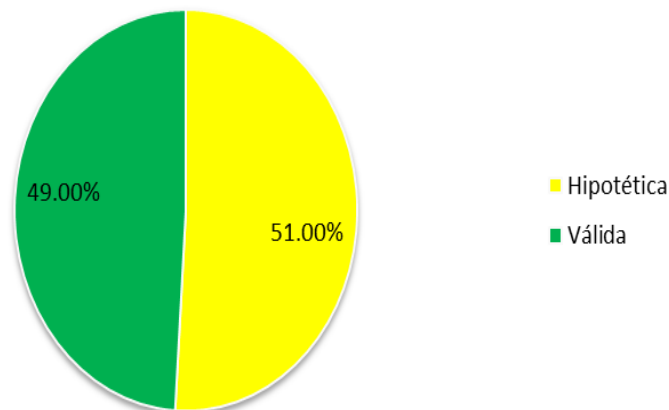
5 RESULTADOS E DISCUSSÃO

5.1 PANORAMA GERAL DA ANOTAÇÃO

O resultado final da anotação foi de 8.258 ORFs para *B. diazoefficiens* estirpe CPAC 7, onde 49% das ORFs foram validadas, uma vez que possuíam alta similaridade com sequências de proteínas já conhecidas e estudadas; todavia, a outra metade das OFSs foi classificada como hipotéticas (51%), como pode ser observado na Figura 11.

Figura 11 - Resultados finais da classificação das ORFs no genoma de *B. diazoefficiens* estirpe CPAC 7 (SEMIA 5080).

Classificação das ORFs - CPAC 7 (= SEMIA 5080)



O cromossomo de *B. japonicum* estirpe CPAC 7 consiste de uma molécula circular de, aproximadamente, 9.085.545 pb, com uma média de conteúdo G+C de 63,98% indicando conformidades com outros estudos já realizados com *B. japonicum* (KANEKO et al., 2010; KANEKO et al., 2011). Do total de ORFs, 7.737, ou seja 92%, tiveram cobertura pelo KEGG e 7.399 (89%) tiveram cobertura pelo InterPro. O genoma está dividido em 13 contigs no momento, onde 83% do genoma é representado por regiões codificantes (Tabela 4). Entretanto é necessária uma etapa complementar de sequenciamento para fechar os *gaps* entre os *contigs*.

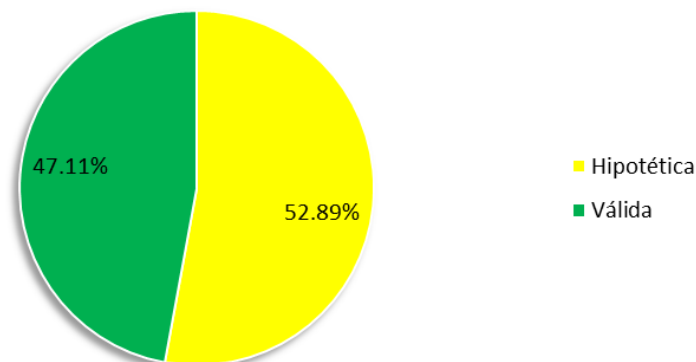
Tabela 4 - Informações gerais do genoma da estirpe CPAC 7 de *Bradyrhizobium diazoefficiens* obtidas pelo software SABIA.

Itens		
Cobertura do genoma		23x
Tamanho (pb)		9.085.545
G+G (%)		63.98%
Número de <i>contigs</i>		13
Número total de exons		52
Região codificante (% do tamanho do genoma)		83%
Tamanho médio das ORFs (pb)		919 pb
<hr/>		
Número de ORFs conhecidas		4037
Número de ORFs hipotéticas		4203
rRNA		3
- RNA ribossomal 16S		1
- RNA ribossomal 23s		1
- RNA ribossomal 5s		1
tRNA		52
<hr/>		
Número total		8.295
<hr/>		
Cobertura com o KEGG	7.737 ORFs	93%
Cobertura com o InterPro	7.399 ORFs	89%

A anotação de 8.690 ORFs de *B. japonicum* estirpe CPAC 15 apresentou alta similaridade com as da CPAC 7. Pouco mais de 47% das ORFs presentes no genoma foram classificadas como válidas (1,89% de diferença entre ambas estirpes). As ORFs restantes foram classificadas em hipotéticas (52,89%) (Figura 12).

Figura 12 - Resultados finais da anotação do genoma de *B. japonicum* estirpe CPAC 15 (=SEMIA 5079)

Classificação das ORFs - CAPC 15 (= SEMIA 5079)



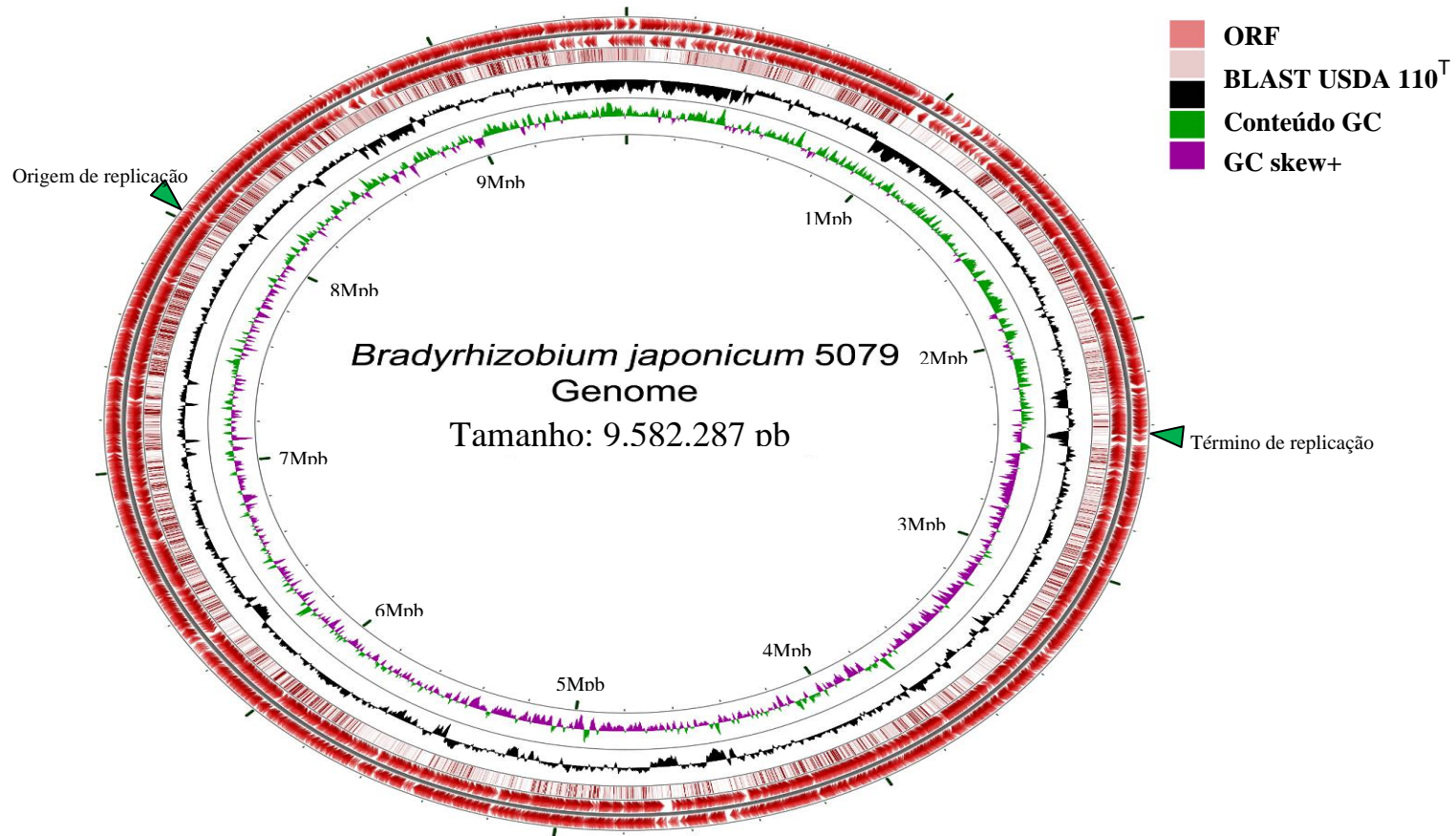
O cromossomo da estirpe CPAC 15 consiste de uma molécula circular de 9.582.287 pb, com um percentual médio de conteúdo G+C de 62,96%. As informações gerais do genoma de CPAC 15 também foram semelhantes à CPAC 7 (Tabela 5). Os dados apresentados são condizentes com a ecologia desta estirpe, uma vez que a CPAC 15 possui elevada competitividade em relação à CPAC 7 (HUNGRIA et al., 2006).

Tabela 5 - Informações gerais do genoma da estirpe CPAC 15 de *Bradyrhizobium japonicum* obtidas pelo software Sabia.

Informações Gerais		
Itens		
Cobertura do genoma		23x
Tamanho (pb)		9.582.287
G+G (%)		63.54%
Número de <i>contigs</i>		1
Número total de exons		2
Região codificante (% do tamanho do genoma)		83%
Tamanho médio das ORFs (pb)		903 pb
Número de ORFs conhecidas		4072
Número de ORFs parciais		3
Número de ORFs hipotéticas		4576
rRNA		6
- RNA ribossomal 16S		2
- RNA ribossomal 23s		2
- RNA ribossomal 5s		2
tRNA		51
Número total		8.708
Cobertura com o KEGG	8.046 ORFs	92%
Cobertura com o InterPro	7.682 ORFs	88%

O genoma da estirpe CPAC 15 encontra-se totalmente montado e anotado, com isso foi possível realizar a representação do cromossomo circular (Figura 13). O conteúdo GC (*GC content*) mostra o desvio de média do conteúdo GC de uma sequência inteira (em verde quando o conteúdo GC é superior a média e em roxo quando é menor).

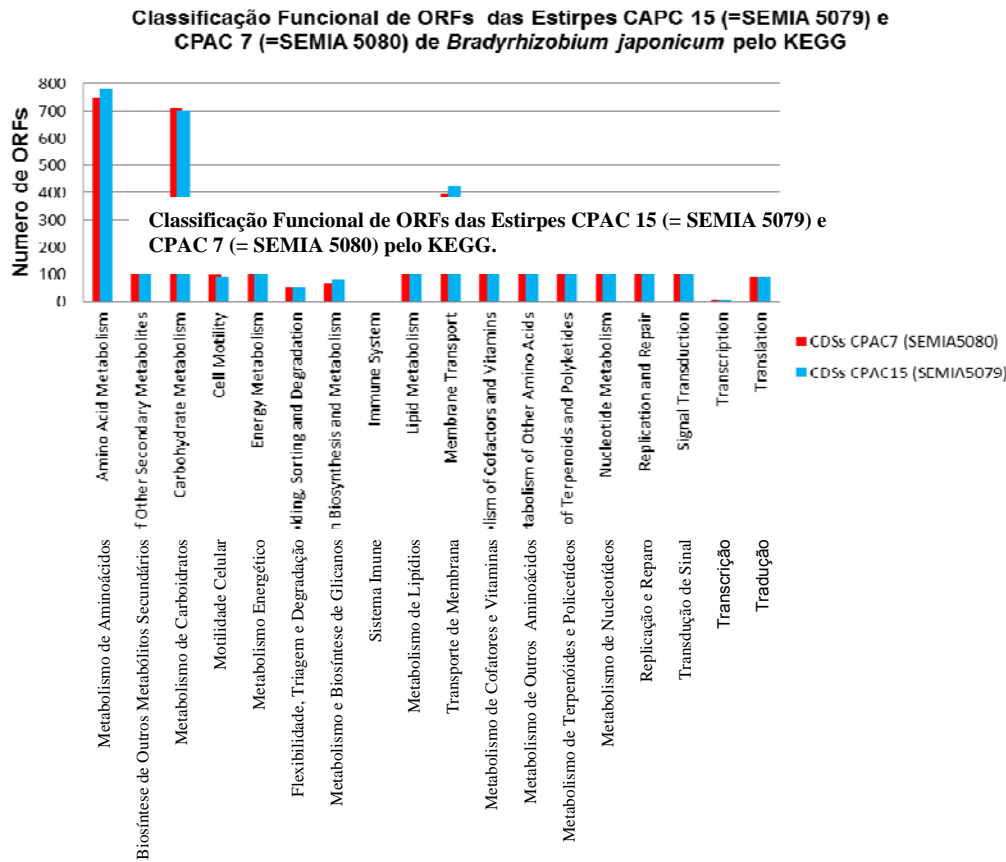
Figura 13 – Representação circular do cromossomo de *Bradyrhizobium japonicum* CPAC 15 (=SEMIA 5079). O círculo mais externo e o adjacente mostram as posições dos possíveis genes codificadores de proteínas no sentido horário e anti-horário. O terceiro círculo representa as comparações de BLASTN com *Bradyrhizobium diazoefficiens* USDA 110^T superior a 95% de identidade (E-value 10^{-10}). Os dois últimos círculos mais internos representam o conteúdo GC (GC content) e GC Skew, respectivamente. As marcações no círculo mais interno representam as posições no genoma em Mbp. As setas em verde indicam a possível origem de replicação e término.



É possível observar, no genoma, regiões com baixo conteúdo CG que podem indicar que, possivelmente, foram adquiridas por transferência horizontal. O BLASTN com USDA 110^T mostrou regiões únicas para CPAC 15 onde não houve alinhamento. Interessante, na região 800.000~1300000pb houve poucos alinhamentos com a USDA 110^T, além de um baixo conteúdo GC, indicando que essa deve ser melhor investigada, para buscar características exclusivas da estirpe CPAC 15.

Na Figura 14 também é notável o elevado número de ORFs relacionadas ao metabolismo de aminoácidos e carboidratos, com valores superiores aos encontrados por Kaneko e colaboradores (2011) em seus estudos com a estirpe tipo USDA 6. Além disso, foi encontrado um número maior de ORFs com função relacionada ao transporte de membrana e replicação e reparo na CPAC 15. Ambas as estirpes (CPAC 7 e CPAC 15) são utilizadas comercialmente como inoculantes de soja no Brasil, devido a sua alta eficiência em fixar o nitrogênio; logo, é condizente que possuam uma maior eficiência na transformação química de compostos carbônicos e nitrogenados e, para isso, necessitam de diferentes isoenzimas envolvidas no metabolismo de aminoácidos e carboidratos.

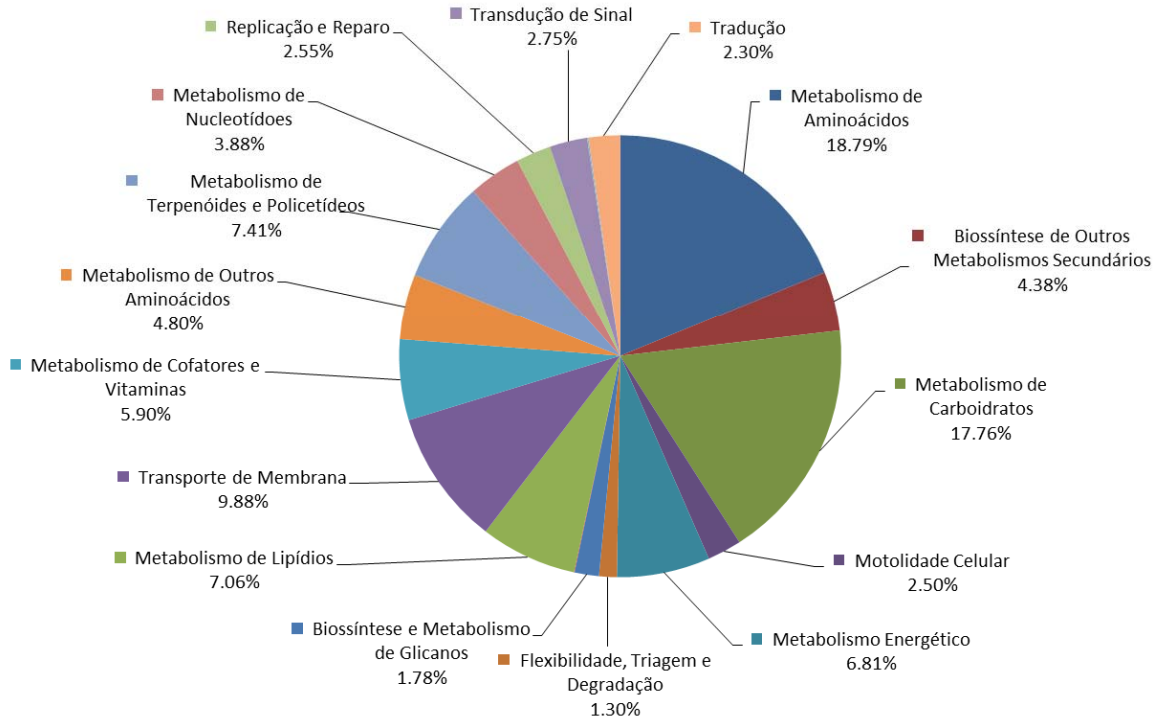
Figura 14 - Classificação funcional pelo KEGG das ORFs encontradas para estirpes CPAC15 (= SEMIA 5079) e CPAC 7 (= SEMIA 5080).



Em CPAC 7 as CDSs responsáveis pelo metabolismo de aminoácidos, metabolismo de carboidratos e transporte de membrana representam 18,79%, 17,76% e 9,88% do genoma, respectivamente. A menor porção do genoma foi representada por CDSs relacionadas ao sistema imune (0,05%) e à transcrição (0,10%) (Figura 15).

Figura 15 – CDSs de *Bradyrhizobium diazoefficiens* estirpe CPAC 7 (=SEMIA 5080) indicando o percentual de todas as categorias do KEGG encontrados no genoma após a anotação.

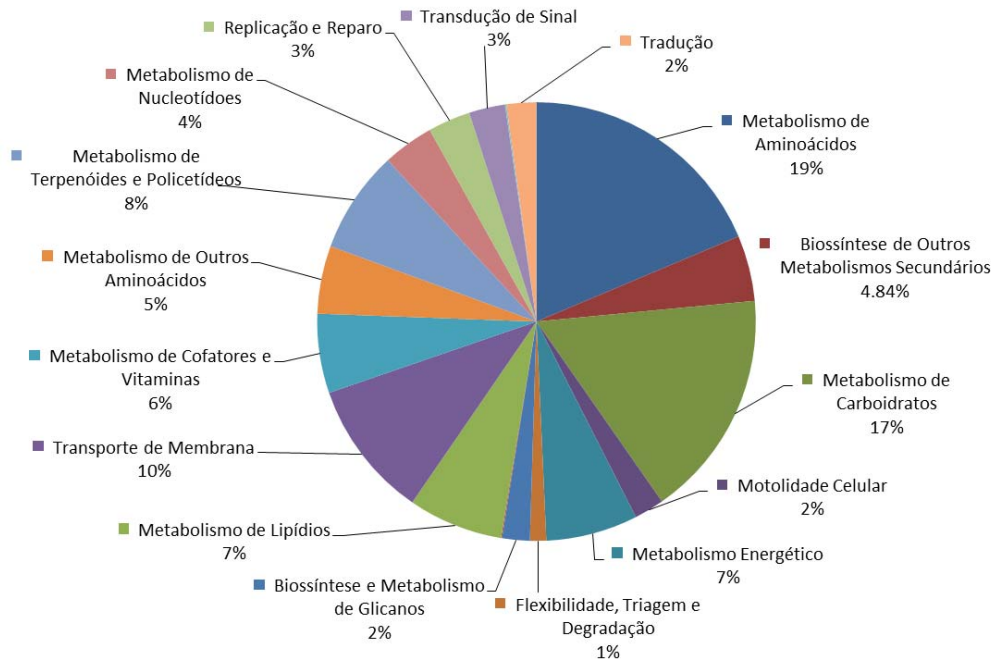
CDSs *Bradyrhizobium diazoefficiens* estirpe CPAC7 (= SEMIA 5080)



Em CPAC 15 os percentuais foram muito similares. As CDSs relacionadas ao metabolismo de aminoácidos, metabolismo de carboidratos e transporte de membrana representam, em ordem decrescente, a primeira (18,65%), a segunda (16,81%) e a terceira (10,15%) fração no genoma, respectivamente (Figura 16).

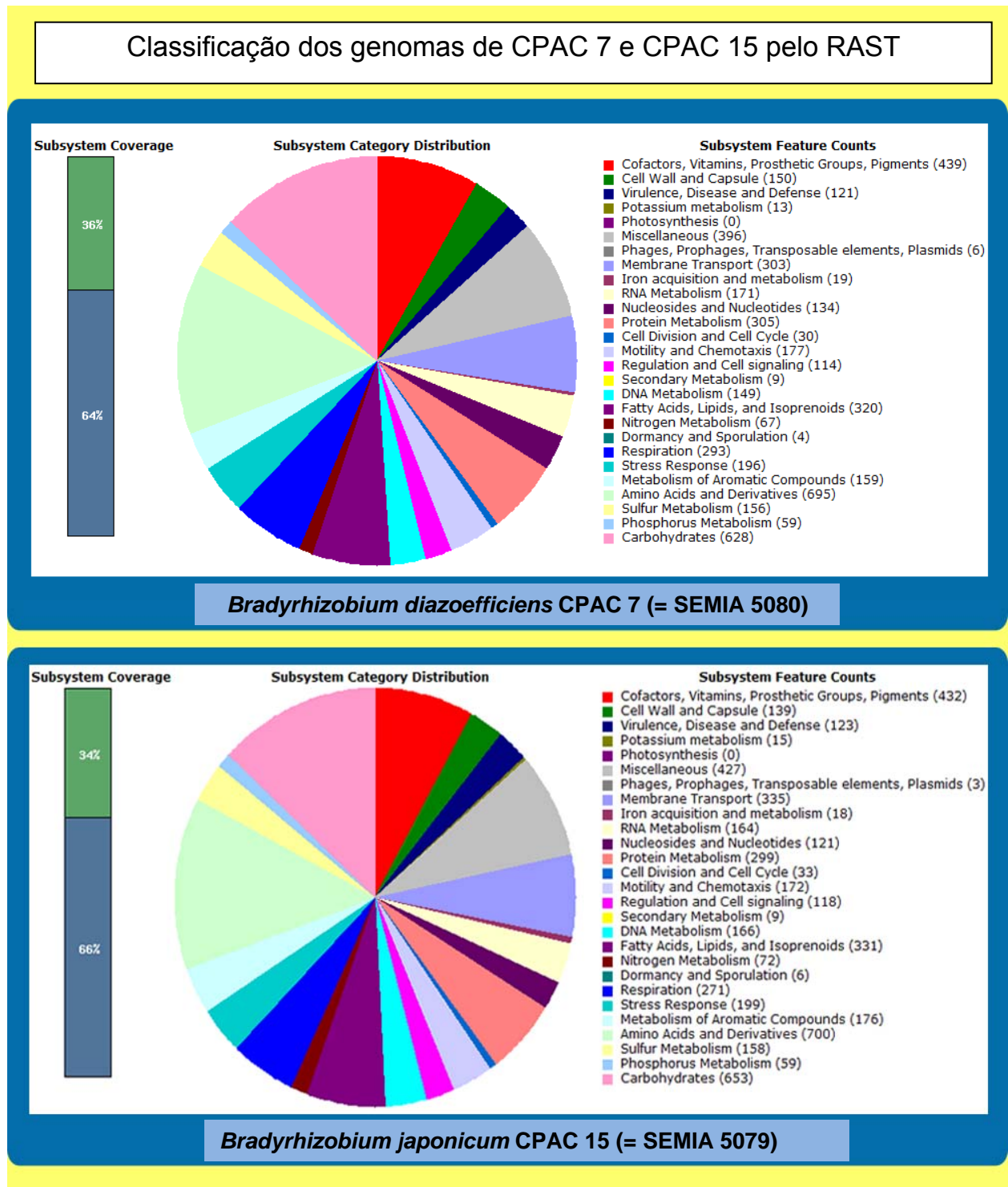
Figura 16 - CDSs de *Bradyrhizobium diazoefficiens* estirpe CPAC 7 (=SEMIA 5080) indicando o percentual de todas as categorias encontrados no genoma após a anotação.

CDSs *Bradyrhizobium japonicum* estirpe CPAC15 (= SEMIA 5079)



As sequências também foram depositadas no programa RAST: (“*Rapid Annotations using Subsystems Technology*”) (AZIZ et al. 2008), para uma melhor análise e visualização dos dados; entretanto, a predição de genes com os dados contidos no sistema foi baixa, 66% e 64% de não cobertura para as estirpes CPAC 15 e CPAC 7, respectivamente. Uma vez que o RAST trabalha com seu próprio banco de dados, provindo de sequências depositadas e anotadas por outros pesquisadores, espera-se uma menor cobertura quando comparada com o *software* SABIA, que reúne os principais bancos de dados do mundo. As sequências que tiveram cobertura pelo RAST (34% para CPAC 15 e 36% para CPAC 7) foram classificadas em diferentes categorias, conforme pode ser visualizado na Figura 17.

Figura 17 - Classificação dos genomas de *Bradyrhizobium japonicum* CPAC 15 (=SEMIA 5079) e CPAC 7 (=SEMIA 5080) pelo banco de dados do RAST e gráficos gerados pelo *Seed Viewer* evidenciando a cobertura das sequências pelo RAST (*subsystem coverage*) e a distribuição das categorias (*subsystem category distribution* e *subsystem feature counts*).



Os resultados foram similares para ambas as estirpes. A categoria mais abundante foi, conjuntamente com o grupo dos aminoácidos e derivados, relacionada a degradação e a biossíntese de aminoácidos que possuam cadeias ramificadas. Já o segundo grupo mais abundante na classificação foi o dos carboidratos, relacionados ao metabolismo central e a fermentação.

5.2 ORFs RELACIONADOS AO TRANSPORTE DE MEMBRANA

Genes relacionados a vários processos do transporte de membrana foram identificados no genoma de *B. japonicum*, em ambas as estirpes. Foram encontradas 384 ORFs para a CPAC 15: 337 (87,76%) relacionadas aos transportadores do tipo ABC, 43 (11,20%) ao sistema de secreção bacteriano e 4 (1,04%) ao transporte de membrana geral. Para a CPAC 7 foram encontradas 395 ORFs: 357 (90,38%) relacionados aos transportadores do tipo ABC, 34 (8,61%) ao sistema de secreção bacteriano e 4 (1,01%) ao transporte de membrana.

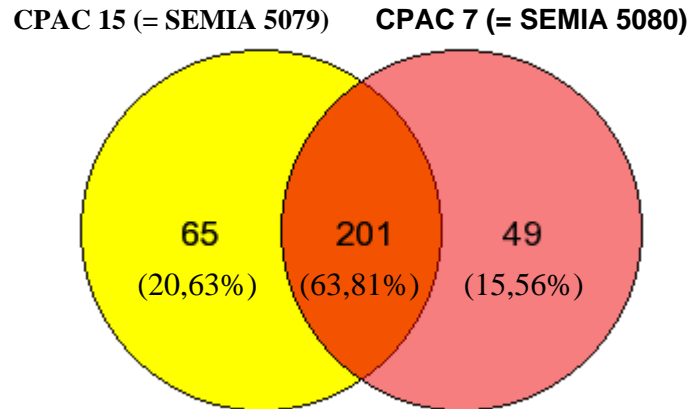
Como já dito, os transportadores do tipo ABC desempenham funções importantes na célula. Tipicamente, transportam ligantes pela bicamada lipídica, os quais são críticos para muitos aspectos da fisiologia celular, incluindo a absorção de nutrientes e a eliminação de resíduos celulares, geração de energia e sinalização celular (LINTON, 2006). Em *Rhizobium leguminosarum* Mulley, (2009), destacou a importância dos transportadores ABC de aminoácidos no genoma. Mutantes que apresentavam supressão da glutamina 2-oxoglutarato amidotransferase, apresentaram um aumento na regulação de mRNAs que codificam componentes do sistema transportador ABC incluindo os transportadores de L-aminoácidos, oligopeptídeos, dipeptídeos e aminoácidos gerais, o que permitiu que esses mutantes crescessem em meio de cultura usando glutamato (e muitos outros aminoácidos), apesar da incapacidade de assimilar o NH_4 .

Com isso fica clara a importância do estudo das ORFs relacionadas ao transporte de membrana.

5.2.1 Transportadores do Tipo ABC

Os transportadores do tipo ABC de CPAC 15 e CPAC 7 representam de 7,8% a 8,6% da capacidade codificante do genoma, respectivamente. Após uma análise mais específica dos transportadores ABC, levando em consideração as sequências de domínios comuns de todos os transportadores já descritos na literatura, foi possível observar, através do diagrama de Venn, que as estirpes compartilham 63,81% dos mesmos domínios de transportadores ABC. As diferenças de 20,63% em CPAC 15 e 15,56% em CPAC 7 provavelmente estejam atreladas a diferenças filogenéticas e adaptações evolutivas que cada estirpe desenvolveu no ambiente de onde foram isoladas, sendo necessários estudos posteriores (Figura 18)

Figura 18 – Diagrama de Venn dos domínios comuns a transportadores do tipo ABC em *B. japonicum* estirpe CPAC 15 (amarelo) e *B. diazoefficiens* estirpe CPAC 7 (vermelho), indicando o percentual de domínios compartilhados e exclusivos em ambos os genomas.



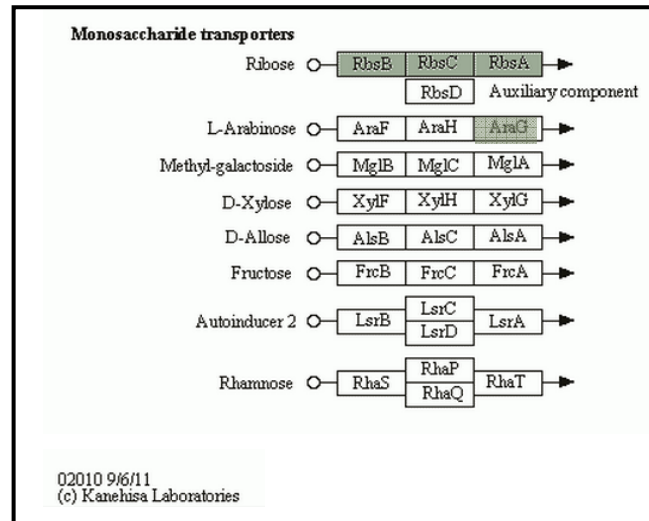
Nos tópicos subsequentes as posições dos transportadores ABC foram descritas quando encontrados em operons, ou em uma região próxima no genoma. Para a estirpe CPAC 7 de *B. diazoefficiens* as posições das ORFs podem não ser definitivas, uma vez que o genoma não se encontra inteiramente montado e, sim, dividido em 13 *contigs*.

5.2.1.1 Transportadores de monossacarídeos e oligossacarídeos

Açúcares desempenham papéis importantes em níveis metabólicos. Os monossacarídeos são açúcares mais simples em sua estrutura química, além de serem menores e mais facilmente metabolizados.

Foi encontrado o transportador completo de ribose RbsB, RbsC, RbsA, que na estirpe CPAC 15 inicia-se na posição 3.650.366 pb e termina em 3653307 pb; já em CPAC 7, o início encontra-se em 665.337 pb e o término em 664397 pb. Entretanto, os demais transportadores de açúcar (metil galactosídeo, D-xilose, D-alose, frutose e ramnose) encontram-se ausentes, como pode ser visualizado na Figura 19. Em ambas as estirpes foi verificada a presença da subunidade de transporte de L-arabinose AraG.

Figura 19 - Imagem obtida do KEGG, indicando a presença ou ausência dos transportadores de monossacarídeos nos genomas de CPAC 15 e CPAC 7. Os transportadores pintados em cinza estão presentes em ambos os genomas.



Apesar da ausência dos demais transportadores, há no genoma proteínas que realizam funções similares, como é o caso GruA, que realiza função análoga à AraG, ou seja, transporte de L-arabinose. Tal princípio pode ser aplicado em outras situações no genoma.

O papel biológico dos oligossacarídeos transpõe os espectros triviais aos que são cruciais ao desenvolvimento, crescimento, funcionamento e sobrevivência de um microrganismo. Inclui papéis estruturais que conferem

conformação e estabilidade às proteínas, fornecimento e defesa de toxinas e anticorpos, controle da meia vida de proteínas e células, modulação de funções proteicas e o fornecimento de eventos de ligações específicos no que diz respeito às interações célula-matriz e célula-célula. Eles geram muito da diversidade funcional requerida para o desenvolvimento e diferenciação de organismos complexos, além de favorecer sua interação com outros organismos no ambiente (VARKI, 1993).

Em ambos os genomas foram encontradas todas as proteínas que compõem o transportador do tipo ABC de sn-glicerol-3-fosfato (regiões 7901708~7904529 pb na estirpe CPAC 15 e 801878~804713 pb em CPAC 7). No Swiss-Prot/UniProtKB a anotação padrão está nomeada como “*sn-glycerol-3-phosphate-binding periplasmic protein*”. O complexo deste transportador ABC é composto pela proteína de ligação ao ATP (UgpC), duas proteínas transmembrana (UgpA e UgpE) e uma proteína de ligação ao soluto (UgpB).

Ao menos uma das subunidades também está presente em *Bradyrhizobium* sp. estirpes STM 3843, ORS 278, BTai1, S23321, WSM471, WSM1253, ORS 375, ORS 285, YR681, STM 3809 e em *B. japonicum* USDA 6^T e *B. diazoefficiens* USDA 110^T, segundo o banco de dados do Swiss-Prot/Uniprot-KB. Apesar de serem amplamente presentes em várias bactérias e, inclusive no grupo dos rizóbios, em *B. japonicum* somente a subunidade UgpC é considerada válida na estirpe USDA 110^T, sendo necessários estudos posteriores para validação em CPAC 7 e CPAC 15 uma vez que o glicerol-3-fosfato é um composto químico intermediário para várias vias metabólicas e pode conferir vantagens a estas estirpes.

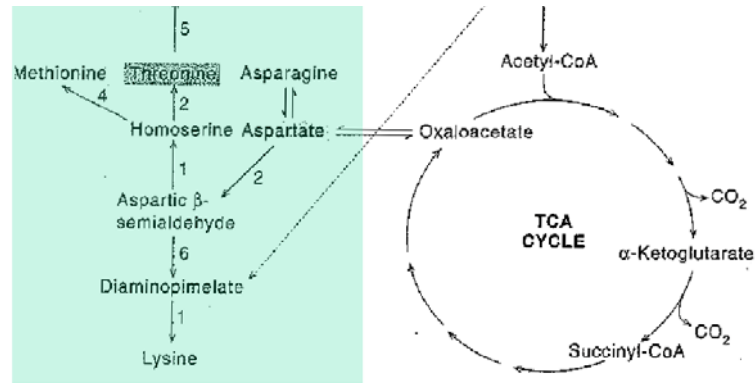
5.2.1.2 Transportadores de peptídeos e aminoácidos

Interessante, o KEGG traz a informação de que há todos os genes codificadores (*yejABEF*) do transportador de microcina C, que é um peptídeo pertencente ao grupo químico das bacteriocinas, também considerado como toxina. Entretanto, quando as sequências de nucleotídeos são confrontadas com os demais bancos de dados, como o NCBI e o Swiss-Prot/Uni-Prot, pode-se observar que as anotações presentes são superficiais, e que são necessários mais estudos.

O ácido aspártico (aspartato) possui um papel importante para os microrganismos, pois é precursor de vários aminoácidos como isoleucina, lisina,

metionina, treonina e asparagina (Figura 20).

Figura 20 - Via metabólica indicando a família do aspartato incluindo, todos os compostos que são seus precursores.



Fonte: <http://memo.cgu.edu.tw/chia-chu/p2.gif> (Modificado).

Em bactérias, o glutamato (ácido glutâmico) também é importante em vários processos metabólicos; além disso, desempenha um papel central contra estresses ácidos e outros tipos de estresses que a bactéria possa sofrer, como demonstrado nos estudos com *E. coli* recentemente descritos por Feehily e Karatzas (2012). Além disso, também funcionam como fatores de crescimento para algumas bactérias lácticas (POLLACK; LINDNER, 1942).

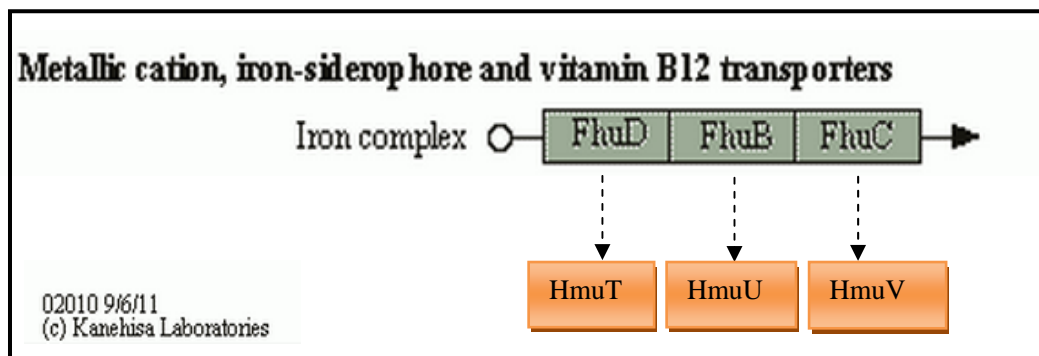
O transportador de aspartato/glutamato está presente em ambos os genomas, sendo representado pelo sistema *glfIKJL* (posição 672196 pb~670636 pb e 6932596 pb~6931027 pb em CPAC 7 e CPAC 15, respectivamente). Transportadores que são mais abrangentes também são encontrados em ambos os genomas: (1) sistema *aapJQMP* (posições 2523033~2525787 em CPAC 15 e 2680318~2684829 em CPAC 7), transportador de L-aminoácidos gerais e (2) sistema *livKHMGF*, transportador de aminoácidos com cadeia ramificada. O gene *aapP* é altamente conservado em *B. japonicum* e está relacionado com o processo de FBN o que pode estar relacionada com a função adaptativa que estes possam exercer durante o processo evolutivo das espécies de rizóbios (PERIN, 2012). Uma particularidade de ambos os genomas é a presença do transportador exclusivo para D-metionina (*operon metQIN*), que é um aminoácido essencial e códon de iniciação AUG. Esse transportador foi encontrado na posição 2379853 pb~2379169 pb em

CPAC 15; já em CPAC 7, o gene responsável pela *metQ* (BU01372) não está adjacente à região 2747810pb~2749622 pb, onde se encontram o *metI* e *metN*. Esse *operon* também foi encontrado em *Bradyrhizobium* sp. estirpes BTAi1 e ORS 278, todavia, no genoma de USDA 110^T foi encontrado somente o gene *metN*.

5.2.1.3 Transportadores de ions metálicos, ferro-sideróforos e níquel

Desta categoria de transportadores só foi encontrado um único transportador em ambos os genomas (102513 pb~105357 pb em CPAC 7 e 6413759 pb~6416597 pb em CPAC 15). O KEGG classifica os transportadores como fazendo parte do sistema *fhuDBC*; todavia, com uma melhor análise das sequências, pode-se observar que essas proteínas são análogas ao sistema *hmuTUV* (Figura 21).

Figura 21 - Transportador de complexos férricos (sistema *fhuDBC*) presente no genoma indicando a analogia com o sistema *hmuTUV*.



O transportador formado pelas proteínas HmuT, HmuU e HmuV é responsável pelo transporte de hemina, molécula de porfina quelada de ferro. Em *B. japonicum* a síntese de hemes é essencial para a expressão da holoproteína leghemoglobina em nódulos de raízes de soja (O'BRIAN; KIRSHBOM; MAIER, 1987). Sendo assim, destaca-se a importância desse transportador presente em ambas as estirpes.

O teor de níquel (Ni) na crosta terrestre é de, aproximadamente, 0,16 g kg⁻¹, como componente comum de rochas ígneas (FERNANDES, 2006). Este elemento é o 24º mais abundante na crosta terrestre, e compreende, aproximadamente, 3% da composição da terra e também é um metal traço essencial

para muitas espécies de microrganismos e plantas, pois faz parte de metaloproteínas (CEMPEL; NIKEL, 2005). O transportador de níquel *oppABCDF* está presente em ambos os genomas, entretanto, não estão dispostos em *cluster*. Esse transportador encontra-se completo no genoma *Bradyrhizobium* sp. BTAi1 e, em *Bradyrhizobium* sp. ORS 278, os genes *oppA* e *oppB* estão ausentes. Este transportador também é comum a alguns peptídeos ainda não descritos na literatura.

5.2.1.4 Transportadores de sulfato e sulfonato

O enxofre (S) é um átomo muito reativo e as reações químicas envolvendo esse elemento consomem grande quantidade de energia. O S está presente essencialmente na forma de aminoácidos, cisteína (e seus produtos de oxidação: homocisteína e homocistina) e metionina, bem como em moléculas de reserva. Também é encontrado em derivados do metabolismo secundário (SEKOWSKA; KUNG; DAHCHIM, 2000).

O sulfato, em primeira instância, entra na célula, onde o potencial elétrico intracelular é fortemente negativo (-70mV), mudando do estado altamente oxidado para o estado reduzido. Isso requer significativo consumo de energia. (SEKOWSKA; KUNG; DANCHIN, 2000). Transportadores de sulfato estão envolvidos na translocação de sulfato ou compostos que contenham sulfato, como tiosulfato, através de membranas biológicas. Em ambos os genomas foram encontrados ORFs referentes a todas as subunidades do transportador de sulfato. Os genes *cysP*, *cysT*, *cysW* e *cysA* tiveram uma alta similaridade entre CPAC 7 e CPAC 15, conforme pode ser visualizado na tabela 6. Na estirpe CPAC 15 estes genes encontram-se na região 9256844~9260676 pb e, em CPAC 7, na região 1671267~1675785 pb.

Tabela 6 - Percentual de identidade entre os genes referentes à codificação do transportador de sulfato de CPAC 7 e CPAC 15, indicando a diferença encontrada entre o tamanho das sequências (Δ pb).

Gene	Identidade	Δ pb
<i>cysP</i>	92,3%	0
<i>cysT</i>	93,5%	21
<i>cysW</i>	89,8%	0
<i>cysA</i>	90,0%	0

Em *Escherichia coli* K-12 essa região encontra-se em um *cluster* gênico, organizado em operons (SIRKO et al., 1995). Em *B. diazoefficiens* estirpe CPAC 7 e *B. japonicum* estirpe CPAC 15 foi encontrado o mesmo padrão. O primeiro gene no operon, *cysP*, codifica uma proteína de ligação ao tiosulfato [*tiosulfate-binding protein* (TSBP)] e os outros (*cysT*, *cysW*, *cysA*) codificam componentes associados à membrana, referentes ao sistema transportador sulfato-tiosulfato.

Sequências de todas as proteínas de transportadores de sulfonato (também comuns a nitrato e taurina), *ssuA*, *ssuC* e *ssuB*, também foram encontradas em ambos os genomas e são responsáveis pelo transporte de compostos sulfurados.

5.2.1.5 Transportadores de molibdênio

O molibdênio (Mo) é um elemento essencial para as bactérias, uma vez que serve como cofator para várias enzimas envolvidas no metabolismo do C, N e S. Em *B. japonicum*, a primeira etapa do processo de desnitrificação (redução do NO_3^- ou nitrito (NO_2^-) via óxido nítrico (NO) e óxido nitroso (N_2O para N_2) é realizada pela nitrato redutase, a qual utiliza Mo como cofator (DELGADO et al. 2006). Em *B. japonicum* CPAC 15 e *B. diazoefficiens* estirpe CPAC 7, a incorporação de Mo pode ser realizada pelo sistema transportador do tipo ABC de alta afinidade, codificado pelos genes *modABC*. Não foram encontrados os genes *modEF*. O proteína ModA (5219795~5220484 pb em CPAC 15 e 53071~53835 pb em CPAC 7) se liga ao molibdato no periplasma, ModB (8814845~8815540 pb em CPAC 15 e 52379~53065 pb em CPAC 7) é a proteína transmembrana da permease e o ModC

(8814180~8814842 pb em CPAC 15 e 50571~52382 pb em CPAC 7) promove energia e encontra-se no lado citoplasmático da membrana. Sirko et al. (1990), em seus estudos com *E. coli*, relataram que o *cluster* gênico *cysPTWA*, posteriormente traduzido em proteínas que compõem transportadores de sulfato, também pode realizar o transporte de molibdato, apesar de apresentar menor afinidade. Neste estudo, foram encontrados os mesmos *clusters*, com isso, é provável que em ambas as estirpes este processo também ocorra da mesma maneira.

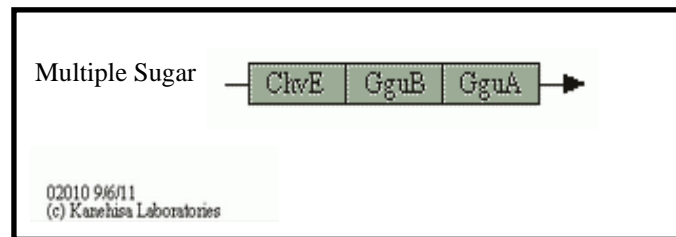
5.2.1.6 Transportadores de ions orgânicos e polióis

O transportador completo de manitol/sorbitol (*smoE*, *smoF*, *smoG*, *smoK*) está presente em ambos os genomas nas coordenadas 3.767.333 – 3.765.427 e 1.301.142 – 1.305.278, para CPAC 15 e CPAC 7, respectivamente. Entretanto, não foi encontrado em outras estirpes de *B. japonicum* e *B. diazoefficiens* com o genoma sequenciado, sendo exclusivo em CPAC 15 e CPAC 7. O poliól de 6 carbonos, d-manitol, é o poliól mais abundante na natureza (SONG; VIELLE, 2009), sendo produzido por bactérias, leveduras, fungos, algas, líquens e muitas plantas (WISSELINK et al. 2002). Este composto possui grande importância na relação simbiótica entre *B. japonicum* e leguminosas. Nas plantas, o manitol é formado no citosol de células fotossintéticas de folhas pela redução do NADPH da manose-6-fosfato para manitol-1-fosfato. Em microrganismos, o manitol é formado pela redução do carbono-2 em frutose ou frutose-6-fosfato, para formar manitol ou manitol-1-fosfato, respectivamente (PHARR et al., 1995). A alta capacidade de capturar energia em forma química reduzida de fonte fotossintética pode representar uma das vantagens de se utilizar o manitol como fotoassimilado translocado, visto que este é quimicamente mais reduzido que a sacarose e a oxidação completa do manitol rende mais energia por mol do que aldoses e cetoses análogas. Além disso, o manitol pode ser utilizado como fonte de carbono quando utilizado em meio de cultura (PHARR et al., 1995).

Por fim, em ambos os genomas foram encontrados transportadores de açúcares múltiplos, representados pelas três proteínas ChvE, GguB, GguA que trabalham em conjunto, uma vez que suas sequências estão em *cluster* (região 1289129~1291947 pb em CPAC 7 e 3781165~3779489 pb em CPAC 15) e

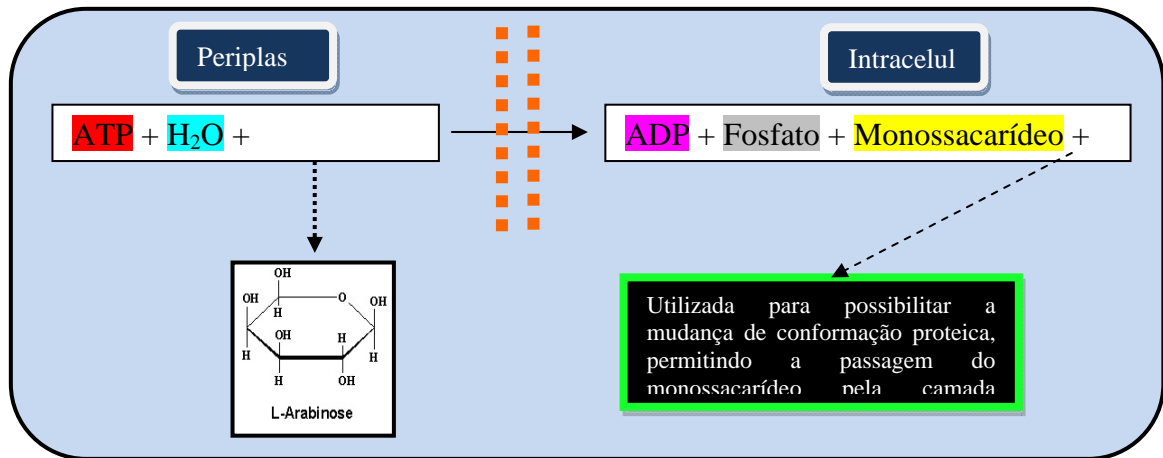
as informações presentes no mapa metabólico do KEGG validam essa informação, conforme pode ser visualizado na Figura 22.

Figura 22 - Transportador de açúcares múltiplos completo presente nos genomas de *B.japonicum* estirpes CPAC 7 e CPAC 15.



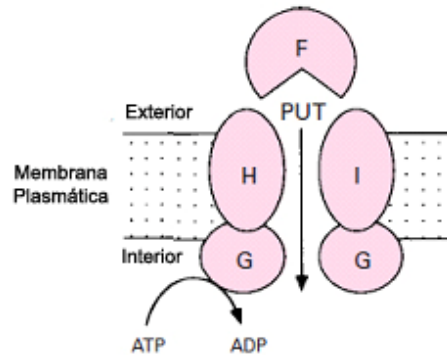
Estudos realizados por Shimoda et al. (1993) concluíram que a proteína ChvE possui estreita relação com a proteína VirA em *Agrobacterium*. O gene *virA*, juntamente com o gene *virG* são responsáveis pela codificação positiva dos demais genes *vir* presentes em *Agrobacterium*, os quais são requeridos para a tumogenicidade da bactéria. Foram encontrados, em ambos os genomas, proteínas semelhantes denominadas *virA-like proteins*, locus BU00870 e BS06042 para CPAC 7 e CPAC 15, respectivamente. Entretanto, são necessários estudos moleculares para determinar qual o tipo de interação existente entre ChvE e proteínas *virA-like*. Mas sabe-se que, primariamente, a ChvE é responsável pela quimiotaxia de certos açúcares. Análises de diferença representacional entre o genoma das estirpes CPAC 15 e S 370 mostrou que o gene *chvE* está possivelmente relacionado com a eficiência simbiótica na estirpe S 370 (BARCELLOS et al, 2009). A proteína GguB não foi caracterizada em *B. japonicum*. Levando em consideração a similaridade da sequência neste estudo com as que estão presentes em *Shigella flexneri*, pode-se inferir que, provavelmente, a proteína GguB seja responsável pelo transporte de D-xilose, pela membrana segundo o Swiss Prot/Uniprot-KB. Por sua vez, a GguA é análoga à proteína AraG, que realiza o transporte de L-arabinose acoplada ao sistema de obtenção de energia do transportador ABC, conforme o Swiss Prot/Uniprot-KB. A atividade catalítica ocorre de forma a transportar o monossacarídeo para o interior da célula com utilização da energia contida no fosfato da adenosina trifosfato (ATP), conforme pode ser visualizado na Figura 23.

Figura 23 - Esquema do transporte de monossacarídeo realizado pela proteína GguA (AraG), onde ocorre o transporte do monossacarídeo pela proteína devido a liberação de energia contida nas ligação do fosfato



Poliaminas são encontradas em uma grande quantidade de animais, bactérias, leveduras e plantas e possuem uma grande importância fisiológica, pois são requeridas durante o crescimento celular de muitos microrganismos. De maneira geral são encontradas em maior quantidade em procariotos (TABOR; TABOR, 1976). O conteúdo de poliaminas na célula é regulado pela biossíntese, degradação e transporte (IGARASHI; KASHIWAGI, 1999). Dentre os sistemas de transporte existentes nos genomas estudados, um é específico para a putrescina (1,4-diaminobutano), e consiste de uma proteína de ligação ao periplasma, PotF (3213346~3214473 pb em CPAC 15 e 1979221~1980345 pb em CPAC 7), duas proteínas transmembrana que formam um canal, PotI (6435123~6435935 pb em CPAC 15 e ausente em CPAC 7) e PotH (6435940~6436854 pb em CPAC 15 e 129933~130847 pb em CPAC 7), e uma ATPase responsável para o fornecimento da energia para este complexo, PotG (6436851~6438023 pb em CPAC 15 e 130844~132016 pb em CPAC 7) (Figura 24).

Figura 24 - Esquema do transportador de putrescina (1,4-diaminobutano) do tipo ABC, indicando as proteínas F,H,I e G presentes na membrana de *Bradyrhizobium diazoefficiens* CPAC 7 e CPAC15.



Fonte: Igarashi; Kashiwagi, (1999), (Modificado).

As opinas são carboidratos ou aminoácidos modificados (BRASILEIRO; LACORTE, 2000). No genoma de algumas espécies de *Agrobacterium* há um plasmídeo indutor de tumor (CHILTON et al. 1977) e este possui uma região responsável pela síntese de opinas (MOORE; CHILTON; CANFIELD et al. 1997), que desempenham um importante papel na ecologia do *Agrobacterium* (BOUZAR et al. 1995). Além disso, as estirpes de *Agrobacterium* que são indutoras de opinas também são capazes de utilizá-las como fonte de C e N (BRASILEIRO; LACORTE, 2000). No genoma de *B. japonicum* CPAC 7 foram encontradas todas as subunidades da proteína transportadora de manopina (AttC, AttB, AttA2 e AttA1), todavia, nenhuma dessas subunidades foi encontrada no genoma das estirpes CPAC 15, USDA 110^T e USDA 6. No genoma da CPAC 7, as posições são 1654204~1655271 pb (*attC*), 1653405~1654181 pb (*attB*), 1652497~1653378 pb (*attA2*), 1651370~1652500 pb (*attA1*). Em *Bradyrhizobium* sp. STM 3843 há anotações putativas no Swiss-Prot/Uniprot indicando a presença das subunidades AttC, AttB e AttA2. Não há relatos concretos de que este gênero seja capaz de produzir opina ou utilizá-la como fonte de carbono ou nitrogênio, assim como ocorre em algumas espécies de *Agrobacterium*, sendo necessários estudos mais profundos e de caráter fenotípico. Se tratando da família *Rhizobiaceae*, atualmente no banco de dados do NCBI existem informações sobre estas proteínas somente em *Sinorhizobium fredii* NGR234, *Sinorhizobium meliloti* 1021, *Agrobacterium vitis* S4, *Agrobacterium fabrum* str. C58, *Agrobacterium fabrum* str. C58, *Sinorhizobium fredii* USDA 257, *Sinorhizobium fredii* NGR234, *Sinorhizobium*

fredii HH103, sendo este então o primeiro relato da possível existência deste transportador em *B. japonicum*, que será melhor avaliado após o fechamento do genoma.

5.2.1.7 Transportadores do tipo ABC-2

Nesta categoria enquadram-se os transportadores formados por duas a três unidades proteicas, mas que também desempenham importantes papéis para a célula bacteriana. Foram encontrados, em ambas as estirpes, transportadores de (1) antibióticos (*yadH, yadG*), (2) lipopolissacarídeos (*rfaA, rfaB*), lipooligosacarídeos (*nodJ, nodI*), lipoproteínas (*lolC, lolE, lolD*), e de divisão celular (*ftsX, ftsE*).

5.2.1.8 Outros transportadores

Dentre os transportadores de íons orgânicos do tipo ABC foi encontrado um transportador incompleto em ambos os genomas. Trata-se de um transportador denominado osmoprotetor, formado pelas proteínas OpuBC, OpuBB e OpuBA (região 864742~867080 pb em CPAC 7 e região 5781064~5783387 pb em CPAC 15), onde o gene responsável pela codificação da OpuBB não foi encontrado. Esse transportador é responsável pelo controle da osmoregularidade através da obtenção de compostos osmoprotetores, ou seja, compostos que podem servir ou serem convertidos em solutos compatíveis, resultando em uma acumulação suficiente de soluto para prover homeostase osmótica para as células (CHEN; BEATTIE, 2007). Todavia, foi encontrada a proteína sensora de osmolaridade, EnvZ [ORF BS01149 em CPAC 15 e BU03682 em CPAC 7]. Essa é membro do sistema regulador EnvZ/OmpR, envolvido na regulação da osmoregulação. A EnvZ funciona como proteína associada à membrana que fosforila a OmpR em resposta a sinais ambientais (FORST; DELGADO; INOUE, 1989). Com isso, é possível que a ausência da proteína OpuBB possa ser suprida por outro sistema de regulação osmótica presente em ambas as estirpes.

5.2.2 Curação Manual dos Transportadores do Tipo ABC

Os resultados mostraram similaridade nos padrões de transportadores ABC para ambas as cepas (Tabela 7)

Tabela 7 – Tabela final dos transportadores ABC encontrados nas estirpes de *B.japonicum* estirpes CPAC 15 e CPAC 7 após a curação manual.

Superfamília	Família de Transportadores ABC	Descrição (Pfam)	Números de Transportadores ABC	
			CPAC 15	CPAC 7
ABC1	3.A.1.106	The Lipid Exporter (LipidE) Family	0	1
ABC2	3.A.1.102	The Lipooligosaccharide Exporter (LOSE) Family	1	0
ABC2	3.A.1.1	The Carbohydrate Uptake Transporter-1 (CUT1) Family	15	1
ABC2	3.A.1.2	The Carbohydrate Uptake Transporter-2 (CUT2) Family	5	7
ABC2	3.A.1.3	The Polar Amino Acid Uptake Transporter (PAAT) Family	8	7
ABC2	3.A.1.4	The Hydrophobic Amino Acid Uptake Transporter (HAAT) Family	29	24
ABC2	3.A.1.5	The Peptide/Opine/Nickel Uptake Transporter (PepT) Family	18	17
ABC2	3.A.1.6	The Sulfate/Tungstate Uptake Transporter (SulT) Family	1	1
ABC2	3.A.1.7	The Phosphate Uptake Transporter (PhoT) Family	1	1
ABC2	3.A.1.8	The Molybdate Uptake Transporter (MolT) Family	0	1
ABC2	3.A.1.9	The Phosphonate Uptake Transporter (PhnT) Family	1	1
ABC2	3.A.1.11	The Polyamine/Opine/Phosphonate Uptake Transporter (POPT) Family	2	8
ABC2	3.A.1.17	The Taurine Uptake Transporter (TauT) Family (Similar to 3.A.1.12 and 3.A.1.16)	6	4

O banco de dados do TCDB classifica os transportadores em três superfamílias: ABC1, ABC2 e ABC3. Entretanto, neste estudo não foram encontrados transportadores da superfamília ABC3 no genoma. Os transportadores ABC1 surgiram pela triplicação intragênica de elementos genéticos codificadores de dois segmentos transmembrana primordiais, dando origem a seis domínios transmembrana. Os transportadores ABC2 surgiram da duplicação de um elemento genético codificador de três segmentos transmembrana diferentes, não homólogos as proteínas ABC1 (WANG et al., 2009).

As maiores diferenças foram nas famílias dos transportadores ABC relacionados à captação de carboidratos (15 em CPAC 15 x 1 em CPAC 7) e nas

relacionadas ao transporte e à captação de poliaminas, opinas e fosfonatos (2 em CPAC 15 e 8 em CPAC 7). As diferenças percentuais são melhores observadas nas Figuras 25 e 26.

Figura 25 – Resultado da curação manual dos transportadores ABC em *Bradyrhizobium japonicum* CPAC 15, evidenciando as famílias e o percentual representativo no genoma.

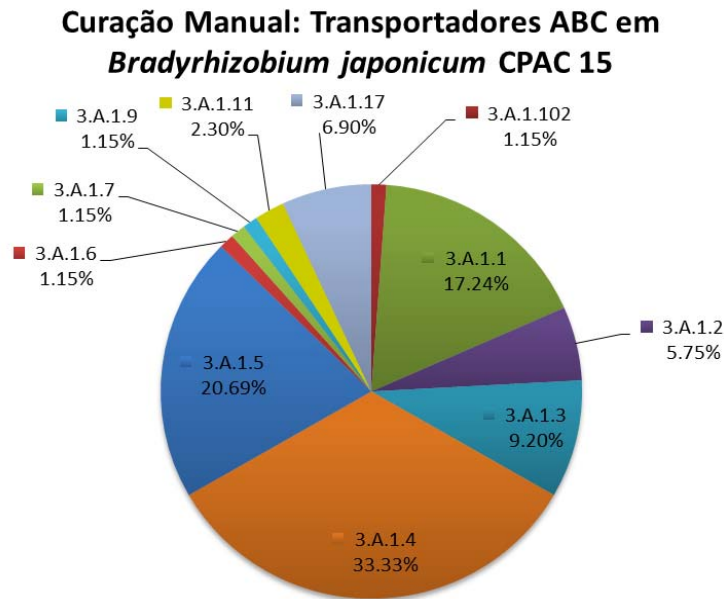
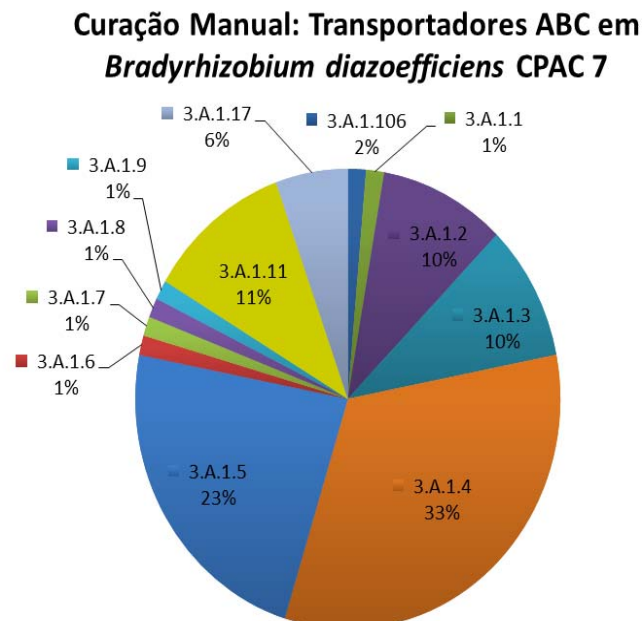


Figura 26 - Resultado da curação manual dos transportadores ABC em *Bradyrhizobium diazoefficiens* CPAC 7, evidenciando as famílias e o percentual representativo no genoma.



A CPAC 15 possui 16,24% a mais de transportadores ABC relacionados à captação de carboidratos do que a estirpe CPAC 7. Tal fato possivelmente esteja relacionado à sua ecologia, uma vez que já foi constatado que esta estirpe realiza maior produção de muco que CPAC 7 e, também, maior produção de ácido indolacético (item 3.4.1 deste trabalho), os quais são compostos ricos em carboidratos. Com isso, faz-se necessária maior captação de carboidratos para o anabolismo destes componentes químicos.

Em contrapartida, a CPAC 7 possui 8,7% a mais de transportadores relacionados ao transporte e captação de poliaminas, opinas e fosfonatos. Apesar de não ter suas funções plenamente elucidadas, as poliaminas apresentam íntima relação com o DNA, podendo inclusive agir como promotores durante a tradução (RATO et al., 2011). Em células que buscam a rápida internalização poliamina exógena, transportadores de poliamina são encontrados altamente ativos, o que é muito comum em células que necessitam se multiplicar rapidamente (WANG et al., 2003). Como já dito, as opinas são produzidas por *Agrobacterium*, encontrados em tumores de pelos radiculares, sendo responsável pelo processo de formação do tumor, conseqüentemente, multiplicação celular rápida (BOUZAR et al. 1995) O fosfonato é um importante agente quelante que impede a solubilidade de compostos metálicos di e trivalentes (NOWACK, 2003). Todas essas características possuem grande importância para a CPAC 7, uma vez que possui taxas de fixação de nitrogênio bastante elevadas, quando comparada com outras estirpes de *Bradyrhizobium*.

6 CONSIDERAÇÕES FINAIS

Foram encontradas várias semelhanças entre os genomas estruturais das estirpes de *Bradyrhizobium diazoefficiens* CPAC 7 e *Bradyrhizobium japonicum* CPAC 15, mas também foram constatadas diferenças importantes entre os dois genomas. O genoma da estirpe CPAC 15 é, aproximadamente, 500.000 pb maior do que o da estirpe CPAC 7 e acredita-se que não ocorra muita variação após o fechamento do genoma. A capacidade codificante da estirpe CPAC 7 é 1% superior à da CPAC 15.

Os transportadores do tipo ABC exercem papéis importantes em nível celular. O sequenciamento das estirpes CPAC 15 e CPAC 7 indicou pequenas diferenças destes transportadores; entretanto, mesmo com a ausência de algumas proteínas, não é garantido que o transportador terá uma perda total de função, visto que suas atividades podem ser substituídas por outras proteínas.

Este estudo, de natureza descritiva, abre caminho para a realização de outros estudos bioquímicos e biotecnológicos, como por exemplo, para a validação de enzimas/proteínas, ou a descrição em nível metabólico de determinadas moléculas assimiladas ou excretadas pelas estirpes.

Levando em consideração a importância das estirpes CPAC 7 e da CPAC 15, que são autorizadas para a inoculação da cultura soja no Brasil, estudos futuros devem ser realizados para incrementar o conhecimento do potencial biotecnológico das mesmas, já que, aproximadamente, 50% do genoma destas estirpes ainda permanece com anotações hipotéticas.

REFERÊNCIAS

- ALBERTS, B.; BRAY, D.; HOPKIN, K.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Fundamentos da biologia celular**. 2.ed. Porto Alegre: Artmed, 2006
- ALI, N. Soybean processing technology and promotional efforts for its food uses. In: Proceedings of the international soybean processing and utilization conference (ISPUC-III), 2000, Tokyo. **Annals of Proceedings of The Internacional Soybean Processing and Utilization Conference(ISPUC-III)**. Tokyo: Korin Pulishing, 2000, p. 327–328.
- ALMEIDA, L.G.P; PAIXÃO, R.; SOUZA, R.C; COSTA, G.C.; BARRIENTOS, F.J.A.; SANTOS, M.T.S.; ALMEIDA, D.F.; VASCONCELOS, A.T. A system for automated bacterial (genome) integrated annotation – SABIA. **Bioinformatics Applications Note**, v.20, p. 2832-3, 2004.
- ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215,p. 403-410, 1990.
- AZIZ, R.K.; BARTELS, D.; BEST, A.A.; DEJONGH, M.; DISZ, T.; EDWARDS, R.A.; FORMSMA, K.; GERDES, S.; GLASS, E.M.;KUBAL, M.; MEYER, F.; OLSEN, G.J.; OLSON, R.; OSTERMAN, A.L.; OVERBEEK, R.A.; MCNEIL, L.K.; PAARMANN, D.; PACZIAN, T.; PARRELLO, B.; PUSCH, G.D.; REICH C.; STEVENS, R.; VASSIEVA, O.; VONSTEIN, V.; WILKE, A.; ZAGNITKO, O. The RAST server: rapid annotations using subsystems technology. **BMC Genomics**, v.9, p. 75, 2008.
- BARBER, S.A. **Soil nutrient bioavailability**: A mechanistic approach. New York: Willey-Interscience, 1984. 389 p.
- BARCELLOS, F.G.; BATISTA, J.S.S.; MENNA, P.; HUNGRIA, M. Genetic differences between *Bradyrhizobium japonicum* variant strains contrasting in N₂-fixation efficiency revealed by representational difference analysis. **Archives of Microbiology**, v. 191, p. 113-122, 2009.
- BARDIN, S. **Phosphate Uptake in *Rhizobium meliloti***. 280. Tese (PhD em Filosofia) – School of Graduate Studies, McMaster University. 1997.
- BLATTNER, F.R.; PLUNKETT, G.; BLOCH, C.A.; PERNA, N.T.; BURLAND, V.; RILEY, M.; COLLADO-VIDES, J.; GLASNER, J.D.; RODE, C.K.; MAYHEW, G.F.; GREGOR, J.; DAVIS, N.W.; KIRKPATRICK, H.A.; GOEDEN, M.A.; ROSE, D.J.; MAU, B.; SHAO, Y. The complete genome sequence of Escherichia coli K-12. **Science**, v. 277, p.1453-1462, 1997.
- BODDEY, L.H.; HUNGRIA, M. Phenotypic grouping of Brazilian *Bradyrhizobium* strains which nodulate soybean. **Biology and Fertility of Soils**, v.25, p.407-415, 1997.
- BOUZAR, H.; CHILTON, W.S.; NESME, X.; DESSAUX, Y.; VAUDEQUIN, V.; PETIT, A.; JONES, J.B.; HODGE, N.C. A new *Agrobacterium* strain isolated from aerial

tumors on *Ficus benjamina* L. **Applied and Environmental Microbiology**, v. 61, p. 65-73, 1995.

BRASILEIRO, A.C.M.; LACORTE, C. *Agrobacterium*: um sistema natural de transferência de genes para a planta. **Biotecnologia Ciência e Desenvolvimento**, p. 12-15, 2000.

BRILL, W. Biochemical genetics of nitrogen fixation. **Microbiological Reviews**, v. 44, p. 449-467, 1980.

CARLSON, J.B.; LERSERN, N.R. Reproductive morphology. In: BOERMA, H.R.; SPECHT, J.E. **Soybeans**: improvement, production, and uses. 3. ed. Madison: American Society of Agronomy, 2004. p. 59–95.

CEMPEL, M.; NIKEL, G. Nickel: A review of its sources and environmental toxicology. **Polish Journal of Environmental Studies**, v. 15, p. 375-382, 2006.

CHEN, C.; BEATTIE, G.A. Characterization of the osmoprotectant transporter OpuC from *Pseudomonas syringae* and demonstration that cystathionine-b-synthase domains are required for its osmoregulatory function. **Journal of Bacteriology**, v. 189, p. 6901-6912, 2007.

CHEN, W.X.; YAN, G.H.; LI, J.L. Numerical taxonomic study of fast-growing soybean rhizobia and a proposal that *Rhizobium fredii* be assigned to *Sinorhizobium* gen. nov. **International Journal of Systematic Bacteriology**, v. 38, n. 4, p. 392-397, 1988.

CHILTON, M.; DRUMMOND, M.H.; MERLO, D.J.; SCIAKY, D.; MONTOYA, A.L.; GORDON, M.P.; NESTER, E.W. Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. **Cell**, v.11, p. 263-271, 1977.

CHUEIRE, L.M.O.; BANGEL, E. V.; MOSTASSO, F. L.; CAMPO, R. J.; PEDROSA, F. O.; HUNGRIA, M. Classificação taxonômica das estirpes de rizóbio recomendadas para as culturas da soja e do feijoeiro baseada no sequenciamento do gene 16S rRNA. **Revista Brasileira de Ciência do Solo**, v. 27, p. 833-840, 2003.

CISOJA. Centro de inteligência de soja, 2011. Disponível em: <<http://www.cisoja.com.br>> Acesso em: 12 de novembro 2011.

COHEN, J. Bioinformatics – an introduction for computer scientists. **AMC Computing Surveys**, v. 36, p.122-158, 2004.

CONAB. Companhia Nacional do Abastecimento. Acompanhamento de safra brasileira: grãos, primeiro levantamento, outubro 2012. Disponível em <<http://www.conab.gov.br>> Acesso em: 11 de abril 2012.

DAVIDSON, A.L.; DASSA, E.; ORELLE, C.; CHEN, J. Structure, function, and evolution of bacterial ATP-binding cassette systems. **Microbiology and Molecular Biology Reviews**, v. 72, p.317-364, 2008.

DELAMUTA, J.R.; RIBEIRO, R.A.; MENNA, P.; BANGEL, E.V.; HUNGRIA, M. Multilocus sequence analysis (MLSA) of *Bradyrhizobium* strains: revealing high diversity of tropical diazotrophic symbiotic bacteria. **Brazilian Journal of Microbiology**, v. 43, p. 698-710, 2012.

DELAMUTA, J.R.; RIBEIRO, R.A.; ORMEÑO-ORRILLO, E.; MELO, I.S.; MARTÍNEZ-ROMERO, E.; HUNGRIA M. Polyphasic evidence supporting the reclassification of *Bradyrhizobium japonicum* Group Ia strains as *Bradyrhizobium diazoefficiens* sp. nov. **International Journal of Systematic and Evolutionary Microbiology**, 2013.

DELGADO, M.J.; TRESIERRA-AYALA, A.; TALBI, C.; BEDMAR, E. J. Functional characterization of the *Bradyrhizobium japonicum modA* and *modB* genes involved in molybdenum transport. **Microbiology**, v. 152, p. 199-207, 2006.

DIXON, R.; EADY, R.R.; ESPIN, G.; HILL, S.; LACCARINO, M.; KAHN, D.; MERRICK, M. Analysis of regulation of *Klebsiella pneumonia* nitrogen fixation (*nif*) gene cluster with gene fusion. **Nature**, v. 286, p. 128-132, 1980.

DÖBEREINER, J.; FRANCO, A.A.; GUZMÁN, I. Estirpes de *Rhizobium japonicum* de excepcional eficiência. **Pesquisa Agropecuária Brasileira, Série Agronomia**, v.5, p.155-161, 1970.

DREYFUS, B.; GARCIA, J.L.; GILLIS, M. Characterization of *Azorhizobium caulinodans* gen. nov., sp. nov., a stem-nodulating nitrogen-fixing bacterium isolated from *Sesbania rostrata*. **International Journal of Systematic Bacteriology**, v. 38, n. 1, p. 89-98, 1988.

ELKAN, G. H. Taxonomy of the rhizobia. **Canadian Journal of Microbiology**, v. 38, p. 446-450, 1992.

EWING B.; GREEN, P. Base-calling of automated sequencer traces using phred. II. error probabilities. **Genome Research**, v. 8, p. 186-194, 1998.

EWING, B.; HILLIER, L.; WENDL, M.C.; GREEN, P. Base-calling of automated sequencer traces using phred. I. accuracy assessment. **Genome Research**, v. 8, p. 175-185, 1998.

FAO - FAOSTAT. **Food and Agriculture Organization of the United Nations**, 2013. Disponível em: <<http://faostat.fao.org/site/567/default.aspx#ancor>> Acesso em: 27 de fevereiro de 2013.

FEEHILY, C.; KARATZAS, K.A.G. Role of glutamate metabolism in bacterial responses towards acid and other stresses. **Journal of Applied Microbiology**, 2012. doi:10.1111/j.1365-2672.2012.05434.x (Online)

FERNANDES, M.S. **Nutrição mineral de plantas**. Viçosa: Sociedade Brasileira de Ciência do Solo, 2006. 432 p.

FERNANDES, M.F.; FERNANDES, R.P.M.; HUNGRIA, M. Caracterização genética de rizóbios nativos dos tabuleiros costeiros eficientes em culturas do guandu e caupi. **Pesquisa Agropecuária Brasileira**, v. 38, p. 911-920, 2003.

FERREIRA, M. C. & HUNGRIA, M. Recovery of soybean inoculant strains from uncropped soils in Brazil. **Field Crops Research**, v. 79, p. 139-152, 2002.

FINN, R.D.; MISTRY, J.; TATE, J.; COGGILL, P.; HEGER, A.; POLLINGTON, J.E.; GAVIN, O.L.; GUNASEKARAN, P.; CERIC, G.; FORSLUND, K.; HOLM, L.; SONNHAMMER, E.L.L.; EDDY, S.R.; BATEMAN, A. The Pfam protein families database. **Nucleic Acids Research**, v. 38, p. D211-D222, 2010.

FINN, R.D.; CLEMENTS, J.; EDDY, S.R. HMMER web server: interactive sequence similarity searching. **Nucleic Acids Research**, v. 39, p. W29-W37, 2011.

FLEISCHMANN, R.D.; ADAMS, M.D.; WHITE, O.; CLAYTON, R.A.; KIRKNESS, E.F.; et al. Wholegenome random sequencing and assembly of *Haemophilus influenzae* Rd. **Science**, v. 269, p. 496–512, 1995.

FORST, S.; DELGADO, J.; INOUE, M. Phosphorylation of OmpR by the osmosensor EnvZ modulates expression of the ompF and ompC genes in *Escherichia coli*. **Proceedings of the National Academy of Sciences USA**, v. 86, p. 6052-6056, 1989.

FREIRE, J. R.; VERNETTI, F. J. A pesquisa com soja, a seleção de rizóbio e produção de inoculantes no Brasil. **Pesquisa Agropecuária Gaúcha**, v.5, p. 117-126, 1999.

FREENIAN, V.J. Studies on the virulence of bacteriophage-infected strains of *Corynebacterium diphtheria*. **Journal of Bacteriology**, v. 61, p. 675–688, 1951.

GARCIA, M.F.; ZAPAROLLI, M.J.S. A participação brasileira no comércio mundial de soja frente ao dilema soja transgênica x soja convencional. In: IX ENCONTRO DE ECONOMIA DA REGIÃO SUL - ANPECSUL 2006, Florianópolis. **Anais do IX encontro de economia da região sul - ANPECSUL 2006**. Florianópolis: ANPECSUL – UFSC, 2006.

GAO, Y.; HAN, Y.; TANG, H.; SUN, D.; WANG, Y.; WANG, W. Analysis of simple sequence repeats in genomes of Rhizobia. **Agricultural Sciences in China**, v. 7, p. 1189-1195, 2008.

GERMANO, M.G.; MENNA, P.; MOSTASSO, F.L.; HUNGRIA, M. RFLP analysis of the rRNA operon of a Brazilian collection of bradyrhizobial strains from 33 legume species. **International Journal of Systematic and Evolutionary Microbiology**, v. 56, p. 217-229, 2006.

GODOY, L.P.; VASCONCELOS, A.T.R.; CHUEIRE, L.M.O.; SOUZA, R.C.; NICOLÁS, M.F.; BARCELLOS, F.G.; HUNGRIA, M. Genomic panorama of *Bradyrhizobium japonicum* CPAC 15, a commercial inoculant strain largely established in Brazilian soils and belonging to the same serogroup as USDA 123. **Soil Biology & Biochemistry**, v. 40, p. 2742-2753, 2008.

GOLD. Genome Online Database. 2012. Disponível em: <<http://www.genomesonline.org/cgi-bin/GOLD/index.cgi>> Acesso em: 5 de Novembro de 2012.

GORDON, D.; ABAJIAN, C.; GREEN, P. Consed: a graphical tool for sequence finishing. **Genome Research**, v. 8, p. 195-202, 1998.

GÖTTFERT, M.; RÖTHLISBERGER, S.; KUNDIG, C.; BECK, C.; MARTY, R.; HENNECKE, H. Potential symbiotic-specific genes uncovered by sequencing a 410-kilobase DNA region of the *Bradyrhizobium japonicum* chromosome. **Journal of Bacteriology**, v.183, p. 1405-1412, 2001.

GÖTTFERT, M.; GROB, P.; HENNECKE, H. Proposed regulatory pathway encoded by the *nodV* and *nodW* genes, determinants of host specificity in *Bradyrhizobium japonicum*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 87, p. 2680-2684, 1990.

GRANT, J.R.; STOTHARD, P. The CGView Server: a comparative genomics tool for circular genomes. **Nucleic Acids Research**, v. 36, p.181-184, 2008.

GROSSMAN, J.M.; SCHIPANSKI, M.E.; SOOKSANGUAN, T.; SEEHAVER, S.; DRINKWATER L.E. Diversity of rhizobia in soybean [*Glycine max* (Vinton)] nodules varies under organic and conventional management. **Applied Soil Ecology**, v. 50, p. 14-20, 2011.

GRUBINGER, V.; ZOBEL, R.; VENDELAND, J.; CORTES, P. Nodule distribution on roots of field-grown soybeans in subsurface soil horizons. **Crop Science**, v. 22, p. 153-155, 1982

HAGEMAN, R.V.; BURRIS, R.H. Changes in the EPR signal of dinitrogenase from *Azotobacter vinelandii* during the lag period before hydrogen evolution begins. **Journal of Biological Chemistry**, v. 254, p.11189-11192, 1979.

HARRISON M.J. Development of the arbuscular mycorrhizal symbiosis. **Current Opinion in Plant Biology**, v. 1, p. 291-503, 1998.

HILL, S.; KENNEDY, C.; KAVANAGH, E.; GOLDBERG, R.B.; HANAU, R. Nitrogen fixation gene (*nifL*) involved in oxygen regulation of nitrogenase synthesis in *K. pneumoniae*. **Nature**, v. 290, p. 424-426, 1981.

HOLLIS, A.B.; KLOOS, W.E.; ELKAN, G.H. DNA-DNA hybridization studies of *Rhizobium japonicum* and related Rhizobiaceae. **Journal of general microbiology**, v. 123, p. 215-222, 1981.

HOLT, J.G.; KRIEG, N.R.; SNEATH, P.H.A.; STALEY, J.T.; WILLIAMS, S.T. **Bergey's manual of determinative bacteriology**. Baltimore: Williams & Wilkins, Maryland-USA, 1994. p. 787.

HUNGRIA, M.; CAMPO, R.J. Inoculantes microbianos: situação no Brasil. In: IZAGUIRRE-MAYORAL, M.L.; LABANDERA, C.; SANJUAN, J., eds. **Biofertilizantes en Iberoamérica**: visión técnica, científica y empresarial. Montevideo: Cyted/Biofag, 2007. p. 22-31.

HUNGRIA, M.; CAMPO, R.J.; MENDES, I.C. **A importância do processo de fixação biológica do nitrogênio para a cultura da soja**: componente essencial

para a competitividade do produto brasileiro. Londrina: Embrapa Soja, 2007. 80 p.. (Embrapa Soja. Comunicado Técnico, 283).

HUNGRIA, M.; CAMPO, R.J.; MENDES, I.C.; GRAHAM, P.H. Contribution of biological nitrogen fixation to the N nutrition of grain crops in the tropics: the success of soybean (*Glycine max* L. Merr.) in South America. In: SINGH, R.P.; SHANKAR, N.; JAIWAL, P.K., eds. **Nitrogen nutrition and sustainable plant productivity**. Texas: Studium Press, 2006. p. 43–93.

HUNGRIA, M.; VARGAS, M.A.T. Exploring the microbial diversity and soil management practices to optimize the contribution of soil microorganisms to plant nutrition. In: STACEY, G.; MULLIN, B.; GRESSHOFF, P. eds. **Biology of plant-microbe interactions**. St. Paul: ISPMI, 1996. p.493-496.

HUNGRIA, M.; VARGAS, M. A. T. Environmental factors affecting N₂ fixation in grain legumes in the tropics, with an emphasis on Brazil. **Field Crops Research**, v. 65, p. 151-164, 2000.

HUNGRIA, M.; VARGAS, M.A.T.; ARAUJO, R.S. Fixação biológica do nitrogênio em feijoeiro: In: VARGAS, M.A.T.; HUNGRIA, M., eds. **Biologia dos solos dos Cerrados**. Planaltina, Empresa Brasileira de Pesquisa Agropecuária, Centro de Pesquisa Agropecuária dos Cerrados, 1997. p.189-295.

HUNGRIA, M.; VARGAS, M.A.T.; SUHET, A.R.; PERES, J.R.R. Fixação biológica do nitrogênio em soja. In: ARAUJO, R.S.; HUNGRIA, M., eds. **Microrganismos de importância agrícola**. Brasília: EMBRAPA-SPI, 1994. p.9-89.

HUNGRIA, M. ; BANGEL, E. V. ; CAMPO, R. J. ; MENDES, I. C. ; REIS-JUNIOR, F. B. . Proposta de certificação das estirpes recomendadas para o uso em inoculantes comerciais para a cultura da soja. In: BIOFAG - **Taller de inoculantes**, Salvador. Granada: CYTED, 2005. p. 7.

HUNGRIA, M.; BODDEY, L.H.; SANTOS, M.A.; VARGAS, M.A.T. Nitrogen fixation capacity and nodule occupancy by *Bradyrhizobium japonicum* and *B. elkanii* strains. **Biology and Fertility of Soils**, v.27, n.4, p.393-399, 1998.

IGARASHI, K.; KASHIWAGI, K. Polyamine transport in bacteria and yeast. **Biochemical Journal**, v. 344, p. 633-642, 1999.

IMPERIAL J.; UGALDE R. A.; SHAH V. K.; BRILL W. J. Role of the nifQ gene product in the incorporation of molybdenum into nitrogenase in *Klebsiella pneumoniae*. **Journal of Bacteriology**, v. 158, p. 187–194, 1984.

JONES, P.M.; GEORGE, A.M. The ABC transporter structure and mechanism: perspectives on recent research. **CMLS Cellular and Molecular Life Sciences**, v. 61, p. 682-699, 2004.

JORDAN, D.C. *Rhizobiaceae* Conn 1938. In: KRIEG, N. R.; HOLT, J.G., eds. **Bergey's manual of systematic bacteriology**. Baltimore, London: Williams & Wilkins, 1984. p. 235-244.

KANEHISA, M.; GOTO, S. KEGG: Kyoto encyclopedia of genes and genomes. **Nucleic Acids Research**, v. 28, p. 27-30, 2000.

KANEKO, T.; MAITA, H.; HIRAKAWA, H.; UCHIIKE, N.; MINAMISAWA, K.; WATANABE, A.; SATO, S. Complete genome sequence of the soybean symbiont *Bradyrhizobium japonicum* strain USDA6. **Genes**, v. 2, p. 763-787, 2011.

KANEKO, T.; NAKAMURA, Y., SATO, S.; MINAMISAWA, K.; UCHIUMI, T.; SASAMOTO, S.; WATANABE, A.; IDESAWA, K.; IRIGUCHI, M.; KAWASHIM, K.; KOHARA, M.; MATSUMOTO, M.; SHIMPO, S.; TSURUOKA, H.; WADA, T.; YAMADA, M.; TABATA, S. Complete genome sequence of the nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA 110. **DNA Research**, v. 9, p.189-197, 2002a.

KANEKO, T.; NAKAMURA, Y., SATO, S.; MINAMISAWA, K.; UCHIUMI, T.; SASAMOTO, S.; WATANABE, A.; IDESAWA, K.; IRIGUCHI, M.; KAWASHIM, 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 Research**, v. 9, supplement, p.225-256, 2002b.

KASCHUCK, G.; HUNGRIA, M.; ANDRADE, D.S.; CAMPO, R. J. Genetic diversity of rhizobia associated with common bean (*Phaseolus vulgaris* L.) grown under no-tillage and conventional systems in Southern Brazil. **Applied Soil Ecology**, v.32, p. 210-220, 2006.

KIM S.; BURGESS, B.K. Evidence for the direct interaction of the *nifW* gene product with the MoFe protein. **Journal of Biological Chemistry**, v. 271, p. 9764-9770, 1996.

KUMUDINI, S. Soybean growth and development. In: SINGH, G, 1. ed. **The soybean**: botany, productions and uses. India: CABI, 2010. p. 48-73.

KUYKENDALL, L.D.; SAXENA, B.; DEVINE, T.E.; UDELL, S.E. Genetic diversity in *Bradyrhizobium japonicum* Jordan 1982 and a proposal for *Bradyrhizobium elkanii* sp. nov. **Canadian Journal of Microbiology**, v. 38, p. 501-505, 1992.

LACROIX, Z.; CRITCHLOW, T. **Bioinformatics**: managing scientific data. United States of America: Elsevier Science (USA), 2003. 441p.

LANG, K.; LINDEMANN, A.; HAUSER, F.; GÖTTFERT, M. The genistein stimulon of *Bradyrhizobium japonicum*. **Molecular Genetics and Genomics**, v.279, p.203-211, 2008.

LAZAROWITZ, S.G.; BISSELING, T. Plant development from the cellular perspective: Integrating the signals. **Plant Cell**, v. 9, p. 1884–1900, 1997.

LESK, A.M. **Introduction to Bioinformatics**. New York: Oxford University Press, 2002. 255 p.

LEWIN, B. **Genes**. United States of America: John Wiley & Sons, 1983. 715p.

- LHUISSIER, F.G.P.; DE RUIJTER, N.C.A.; SIEBERER, B.J., ESSELING, J.J.; EMONS, A.M.C. Time course of cell biological events evoked in legume root hairs by *Rhizobium* Nod factors: state of the art. **Annals of Botany**, v. 87, p. 289–302, 2001.
- LINTON, K.J. Structure and function of ABC transporters. **Physiology**, v. 22, p. 122-130, 2006.
- LOPES, E.S.R.G.; GIARDINI, A.R. Estudos no Estado de São Paulo. In: MYIASAKA, S., MEDINA, J.C. **A soja no Brasil**. São Paulo: Instituto de Tecnologia de Alimentos, 1981. p.139-144.
- MENNA, P.; HUNGRIA, M. Phylogeny of nodulation and nitrogen-fixation genes in *Bradyrhizobium*: supporting evidence for the theory of monophyletic origin, and spread and maintenance by both horizontal and vertical transfer. **International Journal of Systematic and Evolutionary Microbiology**, v. 61, p. 3052-3067, 2011.
- MENNA, P.; BARCELLOS, F.G.; 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. **International Journal of Systematic and Evolutionary Microbiology**, v. 59, p. 2934-2950, 2009.
- MENNA, P.; HUNGRIA, M.; BARCELLOS, F.G.; BANGEL, E.V.; HESS, P.N.; MARTINEZ-ROMERO, E. Molecular phylogeny based on the 16S rRNA gene of elite rhizobial strains used in Brazilian commercial inoculants. **Systematic and Applied Microbiology**, v. 29, p. 315-332, 2006.
- METZKER, M.L. Sequencing Technologies – the next generation. **Nature Reviews Genetics**, v. 11, p. 31-46, 2010.
- MIR, L. **Genômica**. São Paulo: Atheneu, 2004. 1114 p.
- MIZRAHI-MAN, O.; DAVENPORT, E.R.; GILAD, Y. Taxonomic classification of bacterial 16S rRNA genes using short sequencing reads: evaluation of effective study designs. **Plos One**, v. 8, 2013.
- MOORE, L.W.; CHILTON, W.S.; CANFIELD, M.L. Diversity of opines and opine-catabolizing bacteria isolated from naturally occurring crown gall tumors. **Applied and Environmental Microbiology**, v. 63, p. 201-207, 1997.
- MORGANTE, P.G. Fixação biológica e assimilação de nitrogênio. 2003. Disponível em: <<http://pt.scribd.com/doc/3319817/Fixacao-Biologica-e-Assimilacao-de-Nitrogenio>>. Acesso: 15 nov. 2011.
- MULLEY, G. **Nitrogen regulation of ABC transport systems in *Rhizobium leguminosarum***. Thesis (Doctor of Philosophy) The University of Reading – School of Biological Sciences, Berkshire. 2009.
- MYLONA, P.; PAWLOWSKI, K.; BISSELING, T. Symbiotic nitrogen fixation. **Plant Cell**, v. 7, p. 869–885, 1995.

NCBI. **National Center for Biotechnology Information**. 2012. Disponível em: <<http://www.ncbi.nlm.nih.gov>>. Acesso em: 10 out. 2012.

NEWTON, W.E. Nitrogen fixation in perspective. In: PEDROSA, F.O.; HUNGRIA, M.; YATES, M.G.; NEWTON, W.E., eds. **Nitrogen fixation: from molecules to crop productivity**. Dordrecht: Kluwer, 2000. p.3-8.

NEVES, M.C.P.; DIDONET, A.D.; DUQUE, F.F.; DÖBEREINER, J. *Rhizobium* strain effects on nitrogen transport and distribution in soybeans. **Journal of Experimental Botany**, v.36, p.1179-1192, 1985.

NISHI, C.Y.M.; BODDEY, L.H.; VARGAS, M.A.T.; HUNGRIA, M. Morphological, physiological and genetic characterization of two new *Bradyrhizobium* strains recently recommended as Brazilian commercial inoculants for soybean. **Symbiosis**, v. 20, p. 147-162, 1996.

NOWACK, B. Environmental chemistry of phosphonates. **Water Research**, v. 37, p. 2533-2546, 2003.

O'BRIAN, M.; KIRSHBOM, P.M.; MAIER, R.J. Bacterial heme synthesis is required for expression of the leghemoglobin holoprotein but not the apoprotein in soybean root nodules. **Proceedings of the National Academy of Sciences**, v. 84, p. 8390-8393, 1987.

OLDROYD, G.E.D.; DOWNIE, J.A. Coordinating nodule morphogenesis with rhizobial infection in legumes. **Annual Review of Plant Biology**, v. 59, p. 519–546, 2008.

OKUBO, T.; TSUKUI, T.; MAITA, H.; OKAMOTO, S.; OSHIMA, K.; FUJISAWA, T.; SAITO, A.; FUTAMATA, H.; HATTORI, R.; SHIMOMURA, Y.; HARUTA, S.; MORIMOTO, S.; WANG, Y.; SAKAI, Y.; HATTORI, M.; AIZAWA, S.; NAGASHIMA, K.V.P.; MASUDA, S.; HATTORI, T.; YAMASHITA, A.; BAO, Z.; HAYATSU, M.; KAJIYA-KANEGAE, H.; YOSHINAGA, I.; SAKAMOTO, K.; TOYOTA, K.; NAKAO, M.; KOHARA, M.; ANDA, M.; NIWA, R.; JUNG-HWAN, P.; SAMESHIMA-SAITO, R.; TCDB. **Transporter Classification Database**. Disponível em: <<http://www.tcdb.org>>. Acesso em: 15 jan. 2013.

TOKUDA, S.; YAMAMOTO, S.; YAMAMOTO, S.; YOKOYAMA, T.; AKUTSU, T.; NAKAMURA, Y.; NAKAHIRA-YANAKA, Y.; TAKADA HOSHINO, Y.; HIRAKAWA, H.; MITSUI, H.; TERASAWA, K.; ITAKURA, M.; SATO, S.; IKEDA-OHTSUBO, W.; SAKAKURA, N.; KAMINUMA, E.; MINAMISAWA, K. Complete genome sequence of *Bradyrhizobium* sp. s23321: insights into symbiosis evolution in soil oligotrophs. **Microbes and Environments**, v. 27, p. 306–315, 2012.

PARNISKE, M. Arbuscular mycorrhiza: the mother of plant root endosymbioses. **Nature Reviews Microbiology**, v. 6, p. 763-775, 2008.

PERES, J.R.R. **Seleção de estirpes de *Rhizobium japonicum* e competitividade por sítios de infecção nodular em cultivares de soja (*Glycine max* (L.) Merrill)**. Tese (Mestrado em Agronomia - Solos) – Departamento de Solos, Universidade Federal do Rio Grande do Sul, Porto Alegre. 1979. 81p.

PERES, J.R.R.; MENDES, I.C.; SUHET, A.R.; & VARGAS, M.A.T. Eficiência e competitividade de estirpes de rizóbio para soja em solos de cerrado. **Revista Brasileira de Ciência do Solo**, v. 17, p. 357-363, 1993.

PERIN, F.G. **Caracterização dos genes *aapP* e *nopP* relacionados à eficiência da fixação biológica do nitrogênio com a soja em estirpes de *Bradyrhizobium japonicum* contrastantes para esta característica.** Dissertação (Mestrado em Biotecnologia aplicada a Agricultura) – Universidade Paranaense, Umuarama. 2012.

PHARR, D.M.; STOOP, J.M.H.; WILLIAMSON, J.D.; FEUSI, S.; MASSEL, M.O.; CONKLING, M.A. Seedling morphological and physiological adaptation to abiotic stress. **HortScience**, v. 3, p. 1182-1188, 1995.

PIERCE, B.A. **Genética: Um enfoque conceitual.** Guanabara Koogan S.A., 2004. 758p.

PINTO, F.G.S. **Genômica e diversidade de *Rhizobium tropici*, uma espécie tropical microssimbionte do feijoeiro (*Phaseolus vulgaris* L.)** Tese (Doutorado em Microbiologia) - Departamento de Microbiologia – CCB, Universidade Estadual de Londrina, Londrina. 2007. 116p.

PIROOZANIA, M.; NAGARAJAN, V.; DENG, Y. GeneVenn - A web application for comparing gene lists using Venn diagrams. **Bioinformatics**, v. 1, p. 420-422, 2007.

POLLACK, M.A.; LINDNER, M. Glutamine and glutamic acid as growth factors for lactic acid bacteria. **The Journal of Biological Chemistry**, v. 143, p. 655-661, 1942.

POSTGATE, J.R. **The fundamentals of nitrogen fixation.** Cambridge: Cambridge University Press, 1982. 252 p.

POSTGATE, J.R.; HILL, S. Nitrogen fixation. In: LYNCH, J.M.; POOLE, N.J., eds. **Microbial ecology, a conceptual approach.** Oxford: Blackwell, 1979. p.191-213.

QIU, L.-J.; CHANG, R.-Z.. The origin and history of soybean. In: SINGH, G., 1. ed. **The soybean: botany, productions and uses.** India: CABI, 2010. p. 1-23.

RAMSDEN, J.J. **Bionformatics – an introduction.** 2 ed. London: Springer-Verlag, 2009. 271p.

RATO, C.; AMIROVA, S.R.; BATES, D.G.; STANSFIELD, I. WALLACE, H.M. Translational recoding as a feedback controller: systems approaches reveal polyamine-specific effects on the antizyme ribosomal frameshift. **Nucleic Acids Research**, v. 39, p. 4587–4597, 2011.

REES, D.C.; JOHNSON, E.; LEWINSON, O. ABC transporters: the power to change. **Nature Reviews – Molecular Cell Biology**, v. 10, p. 218-227, 2009.

ROBERTS, G.P.; BRILL, W.J. Gene-product relationships of the *nif* regulon of *Klebsiella pneumoniae*. **Journal of Bacteriology**, v.144, p. 210–216, 1980.

ROBERTS, G.P.; MacNEIL, T.; MacNEIL, D.; BRILL, W.J. Regulation and characterization of protein products coded by the *nif* (nitrogen fixation) genes of *Klebsiella pneumoniae*. **Journal of Bacteriology**, v.136, p. 267-279, 1978.

ROCHE – GS FLX+ System: 454 Life Sciences. Disponível em <<http://454.com/products/gs-flx-system/index.asp>> Acesso em: 1 de abril 2013 .

SA, T.M.; ISRAEL, D.W. Energy status and functioning of phosphorus-deficient soybean nodules. **Plant Physiology**, v. 94, p. 328-333, 1990.

SAIER, M.H.; TRAN, C.V.; BARABOTE, R.D. TCDB: the transporter classification database for membrane transport protein analyses and information. **Nucleic Acids Research**, v. 34, p. D181-D186, 2006.

SANTOS, M.A. **Caracterização genotípica e fenotípica de *Bradyrhizobium noduladoras de soja (Glycine Max (L.) Merrill)* isolado da região dos cerrados**. Dissertação (Mestrado em Microbiologia) - Departamento de Microbiologia – CCB, Universidade Estadual de Londrina, Londrina. 1998.

SANTOS, M.A.; VARGAS, M.A.T.; HUNGRIA, M. Characterization of soybean *Bradyrhizobium* strains adapted to the Brazilian savannas. **FEMS Microbiology Ecology**, v. 30, p. 261-272, 1999.

SANTOS, M.H.R.; BITTENCOURT, J.V.M.; SEVERO, B.B.; FRANGO, V. Avaliação do desempenho de inoculante à base de *Bradyrhizobium japonicum* em cultura de soja na cidade de Palmeira – Paraná. **Anais do VIII EPEGE - Encontro Paranaense de Empreendedorismo e Gestão Empresarial**, Paraná, 2011.

SANJUAN, J.; GROB, P.; GÖTTTFERT, M.; HENNECKE, H.; STACEY, G. NodW is essential for full expression of the common nodulation genes in *Bradyrhizobium japonicum*. **Molecular Plant-Microbe Interactions**, v. 7, p. 71–77, 1994.

SCHLAMMAN, H.R.M.; OKKER, R.J.H.; LUGTENBERG, B.J.J. Regulation of nodulation gene expression *nodD* in rhizobia. **Journal of Bacteriology**, v. 174, p. 5177-5182, 1992.

SCHNEIDER E.; HUNKE, S. 1998 apud DAVIDSON, A.L.; DASSA, E.; ORELLE, C.; CHEN, J. Structure, function, and evolution of bacterial ATP-binding cassette systems. **Microbiology and Molecular Biology Reviews**, v. 72, p.317-364, 2008.

SCOTTI, M.R.M.M.L.; SÁ, N.M.H.; VARGAS, M.A.T.; DÖBEREINER, J. Streptomycin resistance of *Rhizobium* isolates from Brazilian cerrados. **Anais da Academia Brasileira de Ciências**, Rio de Janeiro, v.54, p.733-738, 1982.

SEKOWSKA, A.; KUNG, H.; DANCHIN, A. Sulfur metabolism in *Escherichia coli* and related bacteria: facts and fiction. **Journal of Molecular Microbiology and Biotechnology**, v. 2, p. 145-177, 2000.

SHENDURE, J.; JI, H. Next-generation DNA sequencing. **Nature Biotechnology**, v. 26, p. 1135-1145, 2008.

- SHIMODA, N.; TOYODA-YAMAMOTO, A.; AOKI, S.; MACHIDA, Y. Genetic evidence for an interaction between the VirA sensor protein and ChvE sugar-binding protein of *Agrobacterium*. **The Journal of Biological Chemistry**, v. 268, p. 26552-26558, 1993.
- SIEGBAHN, P.E.M.; WESTERBERG, J.; SVENSSON, M.; CRABTREE, R.H. Nitrogen Fixation by Nitrogenase: A Quantum Chemical Study. **The Journal of Physical Chemistry B**, v. 102, p. 1615-1623, 1998.
- SILVA, J.G. Estudos sôbre a inoculação da soja. **Revista Agrícola**, v. 23, p. 365-378, 1948.
- SIRKO, A.; HRYNIEWICZ, M.; HULANICKA, D.; BÖCK, A. Sulfate and thiosulfate transport in *Escherichia coli* K-12: nucleotide sequence and expression of the *cysTWAM* gene cluster. **Journal of Bacteriology**, v. 172, p. 3351-3357, 1990.
- SIRKO, A.; ZATYKA, M.; SADOWY, E.; HULANICKA, D. Sulfate and thiosulfate transport in *Escherichia coli* K-12: evidence for a functional overlapping of sulfate- and thiosulfate-binding proteins. **Journal of Bacteriology**, v. 177, p. 4134-4136, 1995.
- SMIL, V. Phosphorus in the environment: natural flows and human interferences. **Annual Review of Energy and the Environment**, v. 25, p. 53-88, 2000.
- SNUSTAD, D.P.; SIMONS, M.J. **Principles of genetics**. 2.ed. John Wiley & Sons, INC. 896 p.
- SONG, S. H.; VIEILLE, C. Recent advances in the biological production of mannitol. **Applied Microbiology and Biotechnology**, v. 84, p. 55-62, 2009.
- SOUTHAN, C.; CAMERON, G. Para além do tsunami: infraestrutura para lidar com dados das Ciências da Vida. In: HEY, T.; TANSLEY, S.; TOLLE, K. **O quarto paradigma**. São Paulo: Oficina de textos, 2011. p. 137-142.
- SOUZA, J.A.M; TIEPPO, E.; MAGNANI, G. S.; ALVES, L.M.C; CARDOSO, R.L.; CRUZ, L.M.; OLIVEIRA, L.F.; RAITTZ, R.T.; SOUZA, E.M.; PEDROSA, F.O.; LEMOS E.G.M. Draft genome sequence of the nitrogen-fixing symbiotic bacterium *Bradyrhizobium elkanii* 587. **Journal of Bacteriology**, v.194, p. 3547-3548, 2012
- STEVENSON, F.J.; COLE, M.A. **Cycles of soil: carbon, nitrogen, phosphorus, sulfur, micronutrients**. 2. ed. New York: Wiley & Sons, 1999. 427 p.
- ST JOHN, R.T.; JOHNSTON, M.H.; SEIDMAN, C.; GARFINKEL, D.; GORDON, J.K.; SHAH, V.K.; BRILL, W.J. Biochemistry and genetics of *Klebsiella Pneumoniae* mutant strains unable to fix N₂. **Journal of Bacteriology**, v.121, p.759-765, 1975.
- TABOR, C.W.; TABOR, H. 1,4-diaminobutane (putrescine), spermidine, and spermine. **Annual Review of Biochemistry**, v. 45, p. 285-306, 1976.
- TAIZ, L.; ZEIGER, E. **Plant physiology**. 3.ed. Sunderland:Sinauer, 2002. 690 p.

- TATUSOV, R.; GALPERIN M.Y.; NATALE, D.A.; KOONIN, E.V. The COG database: a tool for genome-scale analysis of protein functions and evolution. **Nucleic Acids Research**, v. 28, p. 33-36, 2000.
- TIMMERS, A.C.J. The role of the plant cytoskeleton in the interaction between legumes and rhizobia. **Journal of Microscopy**, v. 231, p. 247–256, 2007.
- VARGAS, M.A.T.; PERES, J.R.R.; SUHET, A.R. Reinoculação da soja em função dos sorogrupos de *Rhizobium japonicum* predominantes em solos de Cerrados. In: SEMINÁRIO NACIONAL DE PESQUISA DE SOJA, 2., 1980, Brasília: **Anais do Seminário Nacional de Pesquisa de Soja**, Embrapa-CPAC, p.715-723, 1981.
- VARGAS, M.A.T.; HUNGRIA, M. **Biologia dos solos dos Cerrados**. Planaltina: EMBRAPA – CPAC, 1977. 524 p.
- VARGAS, M.A.T.; DENARDIN, N.D.; VAN BERKUM, P. Tolerância de rizóbios nativos do feijoeiro a antibióticos e sua relação com fatores de acidez do solo. **Revista Brasileira de Ciência do Solo**, v.16, p.331-336, 1992.
- VARGAS, M.A.T.; MENDES, I.C.; SUHET, A.R. & PERES, J.R.R. Resposta do feijoeiro à inoculação com *Rhizobium leguminosarum* biovar phaseoli, em condições de cerrado. In: REUNIÃO NACIONAL DE PESQUISA DE FEIJÃO, 4, 1993, Londrina. 1993. **Resumos**. Londrina: IAPAR, 1993. p.126.
- VARGAS, M.A.T.; SUHET, A.R. Efeito de tipos e níveis de inoculante na soja cultivada em um solo de Cerrado. *Pesquisa Agropecuária Brasileira*, v. 15, p. 343-347, 1980.
- VARGAS, M.A.T.; SUHET, A.R.; MENDES, I.C.; PERES, J.R.R. **Fixação biológica de nitrogênio em solos de cerrados**. Brasília: Embrapa-CPAC, 1994. p.83.
- VARKEI, A. Biological roles of oligosaccharides: all of the theories are correct. **Glycobiology**, v. 3, p. 97-130, 1993.
- VINCENT, J.M. **Manual for the practical study of root nodule bacteria**. Oxford: Blackwell, 1970. 164p.
- VINCENT, J.M. ed. **Nitrogen fixation in legumes**. Sidney: Academic Press, 1982. p. 111-122
- WANG, B.; DUKAREVICH M.; SUN, E.I.; YEN, M.R.; SAIER M.H. Membrane porters of ATP-binding cassette transport systems are polyphyletic. **Journal of Membrane Biology**, v. 231, p. 1-10, 2009.
- WANG C.; DELCROS, J.G.; CANNON, L.; KONATE, F.; CARIAS, H.; BIGGERSTAFF, J.; GARDNER, A.G.; PHANSTIEL, O. Defining the molecular requirements for the selective delivery of polyamine conjugates into cells containing active polyamine transporters. **Journal of Medicinal Chemistry**, v. 46, p. 5129–5138, 2003.
- WILLEMS, A. The taxonomy of rhizobia: an overview. **Plant and Soil**, v. 287, p. 3-14, 2006.

WISSELINK, H. W.; WEUSTHUIS, R. A.; EGGINK, G.; HUGENHOLTZ, J. GROBBEN, G.J. Mannitol production by lactic acid bacteria: a review. **International Dairy Journal**, v. 12, p. 151-161, 2002.

ZILLI, J.E.; ; MARSON, L.C.; CAMPO, R.J.; GIANLUPPI, V.; HUNGRIA, M. **Avaliação da fixação biológica de nitrogênio na soja em áreas de primeiro cultivo no cerrado de Roraima**. Roraima: Embrapa, 2006. 9 p. (Comunicado Técnico 20).