



# FERTBIO 2012

A responsabilidade socioambiental da pesquisa agrícola  
17 a 21 de Setembro - Centro de Convenções - Maceió/Alagoas

## Anotação Genômica de *Bradyrhizobium japonicum* estirpes CPAC 15 e CPAC 7: Resultados parciais e sua importância para a cultura de soja

**Arthur Fernandes Siqueira<sup>1,2</sup>; André Shigueyoshi Nakatani<sup>1</sup>; Ana Tereza Ribeiro de Vasconcelos<sup>3</sup>; Luiz Gonzaga Paula de Almeida<sup>3</sup>; Mariangela Hungria<sup>1</sup>**

<sup>1</sup>Laboratório de Biotecnologia de Solos; Embrapa Soja; Rod. Carlos João Strass - Distrito de Warta Londrina- Paraná- Brasil Caixa Postal 231 - CEP 86001-970 hungria@cnpso.embrapa.br, nakatani@cnpso.embrapa.br; <sup>2</sup>Mestrando em Biotecnologia, Departamento de Bioquímica e Biotecnologia; Universidade Estadual de Londrina, Rodovia Celso Garcia Cid - Pr 445 Km 380, Cx. Postal 6001 - CEP 86051-980, arthur\_fsiqueira@hotmail.com; <sup>3</sup>Laboratório Nacional de Ciência da Computação, Av. Getúlio Vargas, 333, Quitandinha, Petrópolis - CEP: 25651-075, atrv@lnc.br, lgonzaga@lnc.br.

**RESUMO** – O solo é um importante complexo onde encontram-se nutrientes essenciais para as plantas. A maioria das plantas necessita de grandes quantidades de nitrogênio, todavia, este em geral se encontra de forma pouco assimilável no solo. Contudo, existem bactérias que são capazes de fixar o nitrogênio atmosférico (N<sub>2</sub>), disponibilizando-o para as plantas através de uma relação simbiótica. Dentre estes, destaca-se a espécie *Bradyrhizobium japonicum*, com as estirpes CPAC 15 (=SEMIA 5079) e CPAC 7 (=SEMIA 5080), as quais são amplamente utilizadas como inoculantes para a cultura da soja (*Glycine max* (L.) Merr.) no Brasil. O presente trabalho teve como objetivo realizar a anotação manual do genoma de ambas as estirpes. Para isso, as sequências obtidas foram submetidas à anotação e à montagem utilizando o *software* “System for Automated Bacterial Integrated Annotation” (SABIA), que integra vários programas de domínio público. Foram definidas as *coding sequences* (CDSs), que foram classificadas como hipotéticas, hipotéticas conservadas, válidas e não válidas. Os resultados mostraram elevada similaridade entre ambas as estirpes, com pequena diferença na quantidade de CDSs relacionadas ao transporte transmembrana e replicação em reparo, as quais estavam em maior quantidade na estirpe CPAC 15, provavelmente devido ao seu destaque como maior competitividade. Cerca de 50% do genoma apresentou CDSs classificadas como hipotéticas, sendo então necessários mais estudos para determinar a função desses genes.

**Palavras-chave:** Bactérias diazotróficas, Genoma, Bioinformática, SABIA

**INTRODUÇÃO** - O solo é um complexo físico, químico e biológico. É um material heterogêneo, contendo fases sólidas, líquidas e gasosas. Todas essas fases interagem com os nutrientes minerais (Taiz e Zeiger, 2002). O macronutriente requerido em maior quantidade pelas plantas é o N, entretanto, as plantas são incapazes de absorver o nitrogênio que se encontra de forma abundante na atmosfera (N<sub>2</sub>) e convertê-lo a uma forma assimilável pelas plantas (Hungria et al., 1994). Com isso, as

bactérias fixadoras de N<sub>2</sub> assumem um importante papel nesta etapa, pois são capazes de converter o N<sub>2</sub> atmosférico em amônia. (Taiz e Zeiger, 2002). Dentre as bactérias do solo, a espécie *Bradyrhizobium japonicum* é particularmente importante por estabelecer simbiose com plantas de soja (*Glycine max* L. Merr.), espécie de grande importância comercial para o Brasil (Kumudini, 2010). As estirpes CPAC 15 (=SEMIA 5079) e CPAC 7 (=SEMIA 5080) são usadas e recomendadas para inoculação de soja, desde 1992 (Hungria 2006, 2007). Mundialmente, a soja é representada a maior fonte de proteína, óleo e fitoquímicos promotores de saúde para a nutrição humana e para a alimentação de animais. A soja também se tornou uma importante *commodity*, devido a sua grande adaptação geográfica, composição química, bom valor nutricional, benefícios funcionais à saúde e aplicações industriais (Ali, 2010). Com isso, o presente trabalho teve como objetivo a anotação do genoma das estirpes CPAC 15 e CPAC 7 de *B. japonicum*, a fim de potencializar o seu uso na cultura de soja.

**MATERIAL E MÉTODOS** - As sequências obtidas foram submetidas à montagem e anotação. 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. Regiões codificadoras de genes presentes no genoma foram 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), COG (Clusters of Orthologous Groups of proteins), INTERPRO (uma fonte de busca de famílias, domínios e sítios funcionais de proteínas) e 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* SABIÁ. As leituras de bases de boa qualidade definirão as “CDSs” (*coding sequences*, ou seja, região codificadora de gene) válidas, hipotéticas e conservadas. Todos esses programas constam do programa SABIÁ, que integra a montagem e a anotação automática de genomas (Almeida et al., 2004).

Cada CDS, também denominada como ORF (*Open Reading Frame*), foi conferida manualmente, uniformizada quanto às informações e validada, ou não. O tamanho mínimo de cada CDS foi estabelecido em  $\approx 50$  bp e, no caso de sequências mais curtas, foi feita uma observação. Os seguintes critérios foram aplicados à anotação: CDS foram classificadas como hipotéticas quando nenhum homólogo foi detectado; conservadas hipotéticas quando várias proteínas hipotéticas apresentaram similaridade ou quando houve similaridade baixa com genes conhecidos; válidas quando alta apresentaram alta similaridade com o mesmo gene.

**RESULTADOS E DISCUSSÃO** – O resultado final da anotação foi de 8.258 CDSs para *B. japonicum* estirpe CPAC 7 (Tabela 1).

**Tabela 1** – Resultados finais da anotação do genoma de *B. japonicum* estirpe CPAC 7..

Classificação	CDSs
Conservada Hipotética	2.437
Não Válida	82
Parcial	1
Hipotética	1.635
Válida	4.103

O cromossomo de *B. japonicum* estirpe CPAC 7 consiste de uma molécula circular de 9,050,817 pb com uma média de conteúdo G+C de 63.28% indicando conformidades a outros estudos já realizados com *B. japonicum*. Do total de CDSs, 7.559, ou seja, 92% tiveram cobertura pelo KEGG, entretanto, 50% do genoma foi classificado em proteínas hipotéticas (30% de hipotéticas conservadas + 20% de hipotéticas). Além disso, a anotação de 8690 ORFs de *B. japonicum* CPAC 15 apresentou alta similaridade com a da CPAC 7. O panorama geral é dado na Tabela 2.

**Tabela 2** – Resultados finais da anotação do genoma de *B. japonicum* estirpe CPAC 15.

Classificação	CDSs
Conservada Hipotética	2.520
Não Válida	298
Parcial	4
Hipotética	1.743
Válida	4.124

O cromossomo da estirpe SEMIA 5079 consiste de uma molécula circular de 9,520,722 pb, com um percentual médio de conteúdo G+G de 62,96%..

A classificação funcional pelo KEGG de ambas as estirpes também foi semelhante, entretanto, foi encontrado um número maior de CDSs com função relacionada ao transporte de membrana e replicação e reparo na CPAC 15 conforme por ser visualizado na figura 1.

Tais dados são condizentes com a ecologia desta estirpe, uma vez que possui elevada competitividade em relação à CPAC 7. Na figura 1 também é notável o elevado número de CDSs relacionadas ao metabolismo de aminoácidos e carboidratos, entretanto, os valores foram maiores do que os encontrados por Kaneko e colaboradores (2011) em seus estudos com a estirpe tipo USDA 6. 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, para isso necessitam de diferentes isoenzimas envolvidas no metabolismo de aminoácidos e carboidratos.

**CONCLUSÕES** – Os resultados deste estudo mostraram o panorama geral dos grupos de genes de *B. japonicum* estirpes CPAC 7 e CPA 15, entretanto, há uma grande lacuna de proteínas hipotéticas, o que pode esconder um grande potencial biotecnológico que precisa ser estudado.

**AGRADECIMENTOS** – À CAPES pela bolsa de mestrado. Projeto parcialmente financiado pelo CNPq-Repensa (562008/2010-1).

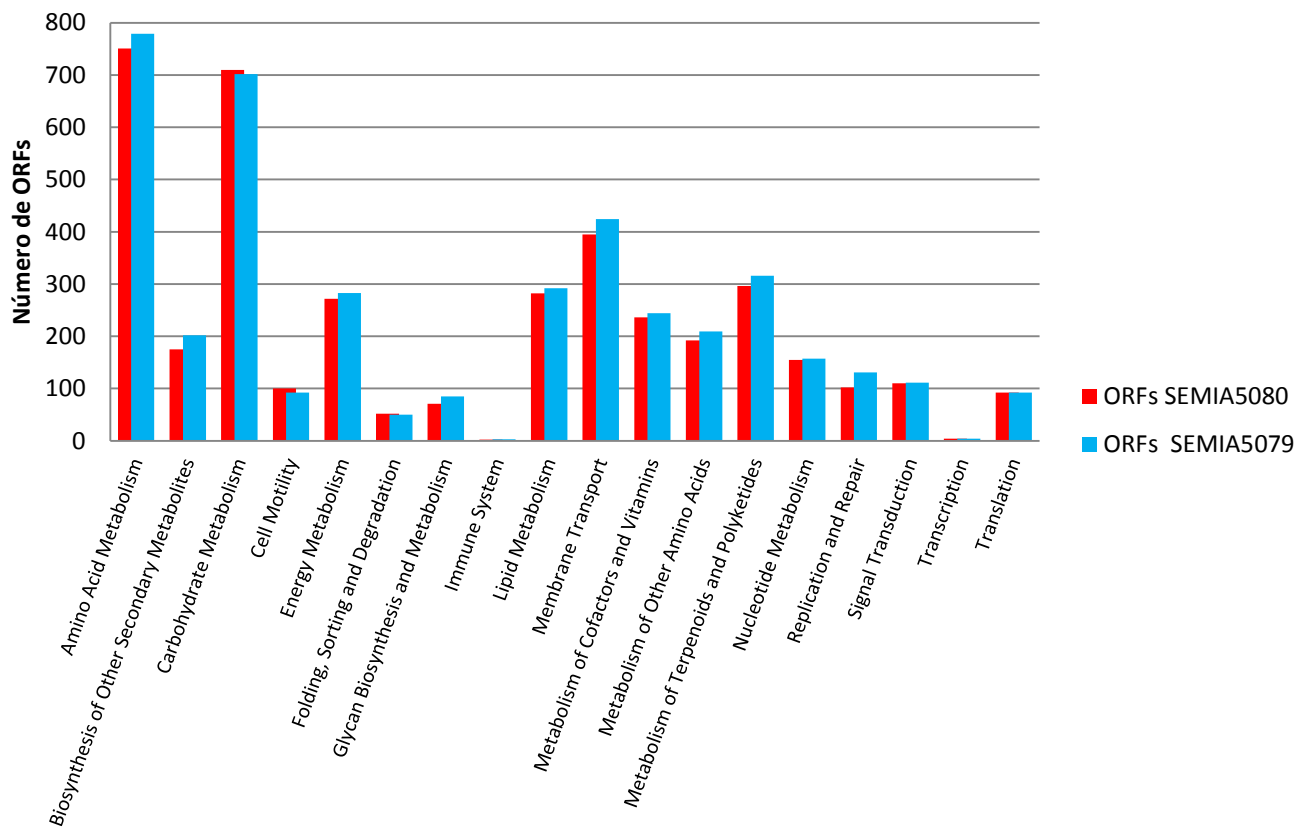
## REFERÊNCIAS

- 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,
- TAIZ, L.; ZEIGER, E. **Plant physiology**. Sinauer, 2002. 690p.
- KUMUDINI, S. Soybean growth and development. In: SINGH, G. **The Soybean: botany, productions and uses**. 1. ed. India, CABI, 2010. p. 48-73.
- HUNGRIA, M.; CAMPO, R.J.; MENDES, I.C. A importância do processo de fixação biológica do nitrogênio para a cultura de soja: componente essencial para a competitividade do produto brasileiro. Londrina: Embrapa Soja, 2007. 80p. (Embrapa Soja. Documentos, 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., (Ed.). **Nitrogen nutrition in plant productivity**. Houston: Studium, 2006. p.43-93
- ALI, N. Soybean Processing and Utilization. In: SINGH, G. **The Soybean: botany, productions and uses**. 1. ed. India: CABI, 2010. p. 345-374.

ALMEIDA, L.G.; PAIXÃO, R.; SOUZA, R.C.; COSTA, G.C.;  
BARRIENTOS, J.C.; SANTOS, M.T.; ALMEIDA, D.F.;  
VASCONCELOS, A.T.R. A system for automated bacterial

(genome) integrated annotation – SABIA. **Bioinformatics**, v.20,  
p.2832-2833, 2004.

**Classificação Funcional de ORFs das Estirpes SEMIA5079 e 5080 de  
*Bradyrhizobium japonicum* pelo KEGG**



**Figura 1** – Tabela da classificação funcional de ORFs das estirpes SEMIA 5079 e 5080 de *Bradyrhizobium japonicum* pelo KEGG, onde fica evidenciada a grande similaridade entre ambas.