



UNIVERSIDADE
ESTADUAL DE LONDRINA

MARIA CLARA DE OLIVEIRA URQUIAGA

**DIVERSIDADE TAXONÔMICA DE ESTIRPES
SIMBIONTES DE *Chamaecrista fasciculata* E DESCRIÇÃO DE
UMA NOVA ESPÉCIE DO GÊNERO *Bradyrhizobium***

Londrina
2019

MARIA CLARA DE OLIVEIRA URQUIAGA

**DIVERSIDADE TAXONÔMICA DE ESTIRPES
SIMBIONTES DE *Chamaecrista fasciculata* E DESCRIÇÃO DE
UMA NOVA ESPÉCIE DO GÊNERO *Bradyrhizobium***

Dissertação apresentada ao Programa de Pós-Graduação em Microbiologia, da Universidade Estadual de Londrina, como requisito para obtenção do título de mestre.

Orientadora: Dra. Mariangela Hungria

Co-orientadora: Dra. Jakeline Renata Marçon
Delamuta

Londrina
2019

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

Urquiaga, Maria Clara de Oliveira.

Diversidade taxonômica de estirpes simbiontes de *Chamaecrista fasciculata* e descrição de uma nova espécie do gênero *Bradyrhizobium* / Maria Clara de Oliveira Urquiaga. - Londrina, 2019.

117 f. : il.

Orientador: Mariangela Hungria .

Dissertação (Mestrado em Microbiologia) - Universidade Estadual de Londrina, Centro de Ciências Biológicas, , 2019.

Inclui bibliografia.

1. MLSA - Tese. 2. Caesalpinioideae - Tese. 3. fixação biológica de nitrogênio - Tese. 4. rizóbios - Tese. I. Hungria , Mariangela . II. Universidade Estadual de Londrina. Centro de Ciências Biológicas. . III. Título.

MARIA CLARA DE OLIVEIRA URQUIAGA

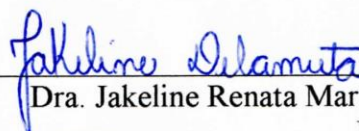
**DIVERSIDADE TAXONÔMICA DE ESTIRPES SIMBIONTES DE
Chamaecrista fasciculata E DESCRIÇÃO DE UMA NOVA ESPÉCIE
DO GÊNERO *Bradyrhizobium***

Dissertação apresentada ao Programa de Pós-Graduação em Microbiologia, da Universidade Estadual de Londrina, como requisito para obtenção do título de mestre.

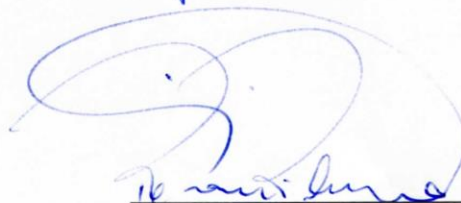
BANCA EXAMINADORA



Orientadora: Dra. Mariangela Hungria
Embrapa Soja



Dra. Jakeline Renata Marçon Delamuta
Embrapa Soja



Dr. Renan Augusto Ribeiro
Embrapa Soja

Londrina, 12 de Agosto de 2019

Dedico, com muito carinho, esse trabalho a todos aqueles que me apoiaram nessa trajetória infinita de busca por conhecimentos e autoconhecimento.

AGRADECIMENTOS

A Deus, em primeiro lugar, pelas oportunidades de crescimento moral e encontros tão especiais nessa existência, e por me fortalecer dia após dia durante toda minha caminhada.

Aos meus pais, Clarice e Segundo. Agradeço por todas as palavras e momentos de afeto, pelos conselhos tão valiosos e pela imensa generosidade. Obrigada por sempre me apoiarem nas minhas escolhas e sonhos, e por me darem todo apoio e carinho possível. Vocês são minha maior inspiração, meus eternos professores e orientadores. Todo o meu amor e esforço dedico a vocês.

Às minhas queridas irmãs, Bárbara e Ana Alice, aquelas que dividem comigo laços de amizade eternos. Obrigada por todos os momentos de descontração e pelo apoio incondicional.

Às minhas queridas avós Maria Pereira, Maria Caballero e Rosa Caivano (*in memoriam*).

À minha querida orientadora, Dra. Mariangela Hungria, o meu profundo agradecimento pela orientação e toda sua contribuição para minha formação, sendo sempre grande fonte de inspiração profissional e pessoal.

Aos queridos professores que tive durante o desenvolvimento desse trabalho, aqueles que sempre se mostraram disponíveis e que tanto contribuíram para a minha formação com seus ensinamentos e orientação: Dra. Jakeline Delamuta, Dr. Renan Ribeiro e Dr. Marco Antônio Nogueira.

Às grandes amigas Milena Klepa e Fabiane Paulitsch. Talvez não seja possível exprimir tamanha gratidão que tenho por vocês. Obrigada por dividirem comigo suas vidas, casas e tantos momentos felizes e especiais. Mas, principalmente, agradeço por vocês estarem ao meu lado nos momentos mais difíceis e por terem se tornado minha família londrinense.

Aos grandes amigos que o Laboratório de Biotecnologia de Solos me permitiu conhecer e conviver: Vanessa Fogaça, Leandro Tullio, Camila Bocatti, Débora Bueno, Ligia Chueire, Anderson Scherer, Fernanda Moura, Eduara Ferreira, Artur Rondina, Marcos Vinícios Garcia, Paula Cerezini, Mariana Sanches, Luiza Helene, Gabriel Guimarães, Thiago Fernandes, Rafaella Liviero, Alisson Sanzovo, Flávia Bender,

Amanda Pit, Rinaldo e Dona Rosa. Agradeço àqueles que ajudaram diretamente nos experimentos, mas principalmente a todos por dividirem comigo suas experiências, tempo e pela amizade.

Aos meus queridos amigos cariocas que estiveram sempre ao meu lado, apesar da distância física, torcendo pelas minhas conquistas e sendo meus grandes refúgios quando eu mais precisei, o meu profundo agradecimento: Karina Peres, Pedro Henrique Carneiro, Marina Ferreira, Carla Cândido, Natália Carla, Larissa Gomes, Larissa Campos, Gabriella Gomes e Julia Moulin.

À Patrícia Montessoro, pela amizade genuína. Agradeço por ter me apoiado do início ao fim dessa trajetória, pela paciência para me ouvir e aconselhar, e por me incentivar a seguir o meu coração.

Ao Axel Katz, pelo apoio incondicional. É difícil expressar tamanha gratidão por todos os conhecimentos e experiências que você compartilhou comigo. Agradeço por você ter dividido comigo as alegrias de cada conquista, e por ter sido refúgio nos momentos de dificuldade. Mas, acima de tudo, obrigada por sempre acreditar em mim até mesmo quando eu não acredito.

À Embrapa Soja, por ceder estrutura inigualável de trabalho e os materiais necessários à realização desse trabalho,

À CAPES pela concessão de bolsa de iniciação científica.

À UEL, em especial ao Instituto de Microbiologia, pelos grandes ensinamentos e por permitir que eu obtenha o título de mestre.

“Gentlemen, it is the microbes who
will have the last word.”

Louis Pasteur

URQUIAGA, Maria Clara de Oliveira. **Diversidade taxonômica de estirpes simbiotes de *Chamaecrista fasciculata* e descrição de uma nova espécie do gênero *Bradyrhizobium*.** 2019. 117 f. Dissertação (Mestrado em Microbiologia) – Universidade Estadual de Londrina, Londrina, 2019.

RESUMO

Os microrganismos são capazes de interagir positivamente com as plantas em sistemas agrícolas e ecossistemas sustentáveis, proporcionando ganhos nutricionais para ambos os parceiros envolvidos. A principal contribuição para a fixação biológica de nitrogênio (FBN), processo chave para a manutenção do balanço de nutrientes na agricultura, ocorre pela associação simbiótica entre leguminosas pertencentes à família Fabaceae e bactérias do solo coletivamente referidas como rizóbios. Bactérias do gênero *Bradyrhizobium* são consideradas como os microssimbiontes predominantes em leguminosas tropicais. Nos últimos anos, um número crescente de estudos filogenéticos relata alta diversidade genética dentro desse gênero, que se destaca pela taxa de crescimento lenta (5 a 7 dias) *in vitro* a 28 °C, ampla distribuição geográfica e gama de hospedeiros, incluindo leguminosas basais da subfamília Caesalpinioideae. A ocorrência de nodulação é rara entre os gêneros de Caesalpinioideae, sendo *Chamaecrista* o maior gênero nodulífero dentro dessa subfamília. Neste estudo, foi conduzida uma análise polifásica com 11 estirpes isoladas de nódulos radiculares de *Chamaecrista fasciculata*, uma herbácea anual multifuncional dos estados do centro-oeste, leste e sul dos EUA. Com base na filogenia do gene 16S RNAr, as estirpes foram agrupadas no superclado de *B. japonicum*. A análise da seqüência de multilocus (MLSA, *multilocus sequencing analysis*) com quatro genes *housekeeping* (*glnII*, *gyrB*, *recA* e *rpoB*) corroborou a classificação das 11 estirpes como um novo grupo, compartilhando menos de 95,2 % de identidade nucleotídica com outras espécies de *Bradyrhizobium*. A alta diversidade genética entre as estirpes foi confirmada nas análises do espaço intergênico entre os genes 16S RNAr e 23S RNAr (ITS), MLSA e BOX-PCR. A média de identidade de nucleotídeos (ANI, *average nucleotide identity*) e a análise da hibridação DNA-DNA digital (dDDH, *DNA-DNA hybridization*) apresentaram valores abaixo do limiar de delimitação de espécies quando comparados com as espécies descritas de *Bradyrhizobium*, com valores de 89,74 % e 40 %, respectivamente. Além disso, outras propriedades fenotípicas, genotípicas e simbióticas foram avaliadas. Cabe destacar a particularidade de crescimento rápido, três dias, de todas as 11 estirpes *in vitro* a 37 °C, sugerindo um mecanismo eficiente de tolerância a altas temperaturas. Os dados obtidos neste estudo suportam a descrição das 11 estirpes como representantes de uma nova espécie, para a qual foi sugerido o nome *Bradyrhizobium frederickii* sp. nov.

Palavras-chave: MLSA. dDDH. Caesalpinioideae. Fixação biológica de nitrogênio. Rizóbios.

URQUIAGA, Maria Clara de Oliveira. **Taxonomic diversity of *Chamaecrista fasciculata* symbionts and description of a new species of *Bradyrhizobium* genus.** 2019. 133 p. Thesis (Master in Microbiology) – Universidade Estadual de Londrina, Londrina, 2019.

ABSTRACT

Microorganisms can interact positively with plants in agricultural systems and sustainable ecosystems, providing nutritional gains for both partners involved. The major contribution to the biological nitrogen fixation (BNF), a key process for maintenance of the nutrients balance in agriculture, occurs by the symbiotic association between legumes belonging to the Fabaceae family and soil bacteria collectively referred as rhizobia. *Bradyrhizobium* is considered as the dominant symbiotic genus of tropical legumes. In the last few years, an increasing number of phylogenetic studies report high genetic diversity within the genus *Bradyrhizobium*, which stands out due to their slow growth rate (5 to 7 days) *in vitro* at 28 °C, broad geographic distribution and wide host range, including basal legumes of the subfamily Caesalpinioideae. Among the Fabaceae subfamilies, nodulation has long been known to be rare in Caesalpinioideae. *Chamaecrista* is considered the largest genus of nodulating caesalpinioids. In this study a polyphasic analysis was performed with 11 strains isolated from root nodules of *Chamaecrista fasciculata*, an annual multi-functional native legume of USA. Based on the 16S rRNA gene phylogeny, the strains were clustered in the *B. japonicum* superclade. Multilocus sequence analysis (MLSA) with four housekeeping genes (*glnII*, *gyrB*, *recA* and *rpoB*) confirmed the new group, sharing less than 95.2 % nucleotide identity with other species. Noteworthy, high genetic diversity among the strains was confirmed in the analyses of the intergenic space between the 16S rRNA and the 23S tRNA (ITS), MLSA and BOX-PCR. Average nucleotide identity (ANI) and digital DNA–DNA hybridization (dDDH) values were below the threshold of described *Bradyrhizobium* species, with values of 89.74 % and 40 %, respectively. Other phenotypic, genotypic and symbiotic properties were evaluated. It is important to highlight the unusual fast growth—three days—by all 11 strains *in vitro* at 37 °C, suggesting a very efficient mechanism of growth tolerance at high temperatures. The data obtained in this study support the description of the 11 strains as representatives of a new species, for which the nomenclature of *Bradyrhizobium frederickii* sp. nov. has been suggested.

Keywords: MLSA. dDDH. Caesalpinioideae. Biological nitrogen fixation. Rhizobia.

LISTA DE ILUSTRAÇÕES

- Figura 1.** O processo de infecção e desenvolvimento de nódulos..... 17
- Figura 2.** Modelo para evolução de vias de infecção em leguminosas.....18

LISTA DE TABELAS

- Tabela 1.** Estirpes isoladas de nódulos de *Chamaecrista fasciculata* utilizadas neste estudo, com a identificação dos locais de procedência dos nódulos..... 35
- Tabela 2.** Sequência de *primers* e as condições de amplificação do gene 16S RNAr, dos genes *housekeeping glnII*, *gyrB*, *recA* e *rpoB*, da região intergênica (ITS), dos genes de nodulação *nodC* e fixação *nifH*..... 42

SUMÁRIO

1.	INTRODUÇÃO	1
2.	OBJETIVOS	3
2.1.	OBJETIVO GERAL.....	3
2.2.	OBJETIVOS ESPECÍFICOS.....	3
3.	REFERENCIAL TEÓRICO	4
3.1.	O PAPEL DO N NA MANUTENÇÃO DA VIDA.....	4
3.1.1.	A Importância e a Distribuição de N na Atmosfera.....	4
3.1.2.	Fixação Biológica de Nitrogênio (FBN).....	6
3.2.	BACTÉRIAS FIXADORAS DE NITROGÊNIO.....	8
3.2.1.	Gênero <i>Bradyrhizobium</i>	9
3.3.	INTERAÇÃO SIMBIÓTICA RIZÓBIOS-LEGUMINOSAS.....	11
3.3.1.	A FBN em Leguminosas.....	11
3.3.2.	Gênero <i>Chamaecrista</i>	12
3.3.3.	Processos de Infecção e Nodulação.....	14
3.4.	TAXONOMIA BACTERIANA.....	19
3.4.1.	Taxonomia Polifásica.....	21
3.4.1.1.	Caracterização fenotípica.....	22
3.4.1.2.	Caracterização filogenética e genética.....	24
4.	MATERIAIS E MÉTODOS	34
4.1.	ISOLAMENTO, SELEÇÃO E CULTIVO DAS ESTIRPES UTILIZADAS.....	34
4.2.	CARACTERIZAÇÃO MORFOLÓGICA E FISIOLÓGICA.....	36
4.2.1.	Crescimento e Diâmetro das Colônias.....	36
4.2.2.	Forma, Elevação, Borda e Superfície das Colônias.....	36
4.2.3.	Produção de Muco e Consistência da Massa de Crescimento.....	37
4.2.4.	Detalhes Ópticos.....	37
4.2.5.	Alterações no pH do Meio.....	37
4.2.6.	Cromogênese das Colônias em Meio YMA com Azul de Bromotimol e Vermelho Congo.....	37
4.3.	CARACTERIZAÇÃO FISIOLÓGICA.....	38

4.3.1.	Capacidade de Utilização de Fontes de Carbono.....	38
4.3.2.	Tolerância a Antibióticos.....	39
4.3.3.	Crescimento em Meio Luria-Bertani (LB).....	39
4.3.4.	Crescimento em NaCl 1 %.....	39
4.3.5.	Crescimento a 37 °C.....	40
4.3.6.	Crescimento em pH 4 e 9.....	40
4.3.7.	Atividade da Enzima Urease.....	40
4.4.	CARACTERIZAÇÃO BIOQUÍMICA.....	40
4.4.1.	Perfil de Ácidos Graxos.....	40
4.5.	CARACTERIZAÇÃO GENÉTICA E FILOGENÉTICA.....	41
4.5.1.	Extração de DNA.....	41
4.5.2.	Análise de BOX-PCR (<i>Polimerase Chain Reaction</i>).....	41
4.5.3.	Amplificação por PCR.....	42
4.5.4.	Sequenciamento Gênico e Genômico.....	43
4.5.5.	Análise Filogenética do Gene 16S RNAr.....	44
4.5.6.	Análise Filogenética da Região Intergênica 16S-23S DNAr (ITS).....	44
4.5.7.	MLSA e Filogenia dos Genes de Nodulação (<i>nodC</i>) e de Fixação De N ₂ (<i>nifH</i>).....	45
4.5.8.	<i>Nucleotide Identity</i> (NI).....	45
4.5.9.	<i>Average Nucleotide Identity</i> (ANI).....	45
4.5.10.	Hibridação DNA-DNA Digital (dDDH).....	46
4.5.11.	Conteúdo G+C.....	46
4.6.	CAPACIDADE DE NODULAÇÃO E ATIVIDADE DOS NÓDULOS.....	46
4.6.1.	Reisolamento das Estirpes.....	48
5.	REFERÊNCIAS BIBLIOGRÁFICAS.....	50
6.	RESULTADOS E DISCUSSÃO.....	62
7.	CONSIDERAÇÕES FINAIS.....	117

1. INTRODUÇÃO

A agricultura global enfrentará diversos desafios nas próximas décadas. A Organização das Nações Unidas para a Alimentação e Agricultura (FAO, *Food and Agriculture Organization*) prevê que, em 2050, a população mundial será de 9,1 bilhões, 29% superior à atual. A necessidade de aumentar a produção de alimentos para manter uma população mundial cada vez mais próspera e crescente exigirá que a agricultura global enfrente uma maior competição por usos alternativos de terras, expandindo a produtividade agrícola de maneira sustentável.

Os microrganismos podem interagir positivamente com as plantas em sistemas agrícolas, proporcionando ganhos nutricionais para ambos os parceiros envolvidos. A fixação biológica de nitrogênio (FBN) é um processo chave para a manutenção do balanço de nutrientes e energia na agricultura, promovendo a melhoria da fertilidade do solo em sistemas produtivos e sustentáveis. A principal contribuição para o processo da FBN ocorre pela associação simbiótica entre leguminosas pertencentes à família Fabaceae e bactérias do solo coletivamente conhecidas como rizóbios. O gênero *Bradyrhizobium* é considerado predominante na família Fabaceae, atuando como foco em muitos estudos filogenéticos, que revelaram uma alta diversidade genética dentro do gênero.

Além de se caracterizarem pela lenta taxa de crescimento (em geral, 5 a 7 dias, podendo ser maior) *in vitro* a 28 °C, as espécies de *Bradyrhizobium* se destacam pela ampla distribuição geográfica e gama de hospedeiros, incluindo leguminosas basais da subfamília Caesalpinioideae. Portanto, seus representantes são considerados os simbiontes ancestrais das leguminosas. A ocorrência de nodulação é rara entre os representantes da subfamília Caesalpinioideae. O gênero *Chamaecrista* é considerado o maior gênero nodulífero dentro dessa subfamília, tendo o Brasil como centro de radiação da sua diversidade.

Estudos filogenéticos relatam que a nodulação em *Chamaecrista* parece ser um evento independente na evolução de simbioses que apresentam FBN. Até o momento, os poucos trabalhos que relatam a FBN em *Chamaecrista* são estudos amplos que pontuam apenas a ocorrência de nódulos, e incluem poucos isolados. Apesar dos poucos relatos existentes, os simbiontes isolados foram identificados como pertencentes ao gênero *Bradyrhizobium*.

Considerando as expressivas contribuições que a FBN proporciona à sustentabilidade agrícola e à manutenção de diversos ecossistemas, a expansão do conhecimento sobre a biodiversidade de rizóbios e, dessa forma, a identificação de novas espécies nodulíferas tem-se mostrado essencial. Em especial, estudos a respeito da diversidade de simbioses do gênero *Chamaecrista* poderiam contribuir para o entendimento de como a “primitiva” subfamília Caesalpinioideae evoluiu para as simbioses mais “avançadas” encontradas no clado Mimosoideae e na subfamília Papilionoideae.

2. OBJETIVOS

2.1. OBJETIVO GERAL

Realizar uma análise de taxonomia polifásica em 11 estirpes de *Bradyrhizobium* isoladas de nódulos de *Chamaecrista fasciculata*, uma leguminosa anual nativa das regiões centro-oeste, leste e sul dos Estados Unidos, visando à caracterização da diversidade taxonômica e filogenética de estirpes simbiotes dessa planta hospedeira.

2.2. OBJETIVOS ESPECÍFICOS

- Caracterizar fenotipicamente, por meio de testes morfofisiológicos e perfis de ácidos graxos, estirpes isoladas de nódulos de *Chamaecrista fasciculata* nativa dos Estados Unidos;
- Avaliar a diversidade genética entre as estirpes utilizando as técnicas de BOX-PCR, ANI (*Average Nucleotide Identity*), NI (*Nucleotide Identity*), dDDH (Hibridação DNA-DNA digital) e conteúdo G+C do genoma;
- Realizar análises filogenéticas com o gene 16S RNAr e a região intergênica 16S RNAr-23S RNAr (ITS) e aplicar a metodologia de MLSA (*Multilocus Sequence Analysis*) com genes *housekeeping* (*glnII*, *gyrB*, *recA* e *rpoB*) para inferir a posição taxonômica das estirpes;
- Inferir propriedades simbióticas das estirpes por meio da filogenia dos genes simbióticos *nifH* e *nodC*;
- Avaliar a capacidade nodulífera das estirpes em *Chamaecrista fasciculata* e em outras leguminosas;
- Avaliar, de forma integrada, todas as características analisadas.

3. REFERENCIAL TEÓRICO

3.1. O PAPEL DO N NA MANUTENÇÃO DA VIDA

3.1.1. A Importância e a Distribuição de N na Atmosfera

Dentre os nutrientes essenciais ao metabolismo vegetal, o N é aquele requerido em maior quantidade pelas plantas (ANDREWS et al., 2009). Diversos compostos bioquímicos importantes das células vegetais possuem o N em sua composição, como os aminoácidos, que fazem parte da constituição de proteínas (80 a 85 % do N total), enzimas, fitocromos, coenzimas, ácidos nucleicos, bases nitrogenadas (TAIZ; ZEIGER, 2004), fitohormônios e a molécula da clorofila. Os vegetais geralmente utilizam o amônio (NH_4^+) ou nitratos solúveis como única fonte de N.

De modo geral, pode-se dizer que o N é determinante para o crescimento e o rendimento de biomassa ou de grãos das plantas, em função da sua indispensável participação nos principais processos fisiológicos dos organismos vegetais, como o desenvolvimento de estruturas reprodutivas e a manutenção da atividade fotossintética, por participar da constituição da clorofila. Dessa forma, a disponibilidade de N pode afetar o tempo e a taxa de germinação de sementes, a expansão foliar e a fotossíntese, a partição da matéria seca entre parte aérea e raiz, a arquitetura das raízes e praticamente todos os demais processos celulares (BASI et al., 2011).

O N está presente em diversas formas na biosfera, estando principalmente concentrado na atmosfera como N_2 , compondo aproximadamente 80 % do volume total. Entretanto, na maior parte, esse grande reservatório não está diretamente disponível aos organismos vivos, já que a utilização do N_2 requer a quebra de uma ligação tripla covalente de excepcional estabilidade entre os dois átomos de N para produzir NH_3 (amônia). Essas reações bioquímicas são conhecidas como fixação do N_2 , e podem ser realizadas por um processo industrial, natural ou biológico (TAIZ; ZEIGER, 2004).

De acordo com Hungria et al. (2001), o aporte de N para as plantas pode se dar a partir de quatro fontes distintas: o solo, principalmente em função da decomposição da matéria orgânica; a fixação não-biológica, resultante de descargas elétricas, combustão e vulcanismo; os fertilizantes nitrogenados; e o processo de fixação biológica do N_2 (FBN).

O aporte de N presente na matéria orgânica do solo é limitado, podendo ser esgotado rapidamente após alguns cultivos. Além disso, diversos outros fatores, como temperatura e umidade são capazes de acelerar os processos de decomposição da matéria

orgânica, resultando em solos com baixos teores desse nutriente (HUNGRIA et al., 2013). No entanto, essa situação torna-se ainda mais séria quando se observam as características predominantes dos solos tropicais, como ocorre no Brasil. Em função de aspectos como altas temperaturas, baixa umidade e elevados níveis de acidez, além do uso inapropriado de tecnologias de campo, os trópicos apresentam solos muito desgastados por erosões, com uma estrutura frágil e, conseqüentemente, com baixa disponibilidade de matéria orgânica (GILLER; WILSON, 1991).

A fixação abiótica de N_2 , isto é, independente da ação de organismos vivos, é pequena e variável, representando cerca de 10 % das entradas de N no planeta. Dessa forma, os fertilizantes nitrogenados e o processo de FBN constituem as principais fontes de N para as plantas (HUNGRIA et al., 2001).

Os fertilizantes nitrogenados representam a forma assimilada com maior rapidez pelas plantas, considerando que esse N comercial é particularmente solúvel para fácil absorção vegetal. Além disso, outras vantagens que tornam esse aporte de N uma opção viável é a simplicidade de seu armazenamento e manuseio, fazendo com que esse nutriente seja facilmente aplicado. Os fertilizantes minerais comumente aplicados nos solos são NH_4 , CH_4N_2O , $(NH_4)_2SO_4$ e $(NH_4)(NO_3)$ (HIREL et al., 2011). Entretanto, é preciso considerar que o uso desses fertilizantes inorgânicos é bastante custoso, principalmente do ponto de vista ambiental. O processo industrial Haber-Bosh que transforma o N_2 em NH_3 requer: hidrogênio (derivado de gás de petróleo); catalisador contendo ferro; altas temperaturas (300° a $600^\circ C$) e altas pressões (200 a 800 atm) (HUNGRIA et al., 2001). Desse modo, é possível perceber que o gasto de fontes energéticas não-renováveis torna o fertilizante nitrogenado energeticamente custoso, sendo estimado um custo em seis barris de petróleo por tonelada de NH_3 sintetizada (ERISMAN et al., 2008).

No entanto, como já foi comentado anteriormente, é importante ressaltar que o uso indiscriminado de fertilizantes nitrogenados resulta em poluição ambiental, e um dos fatores que agrava essa situação é a baixa eficiência de utilização desse nutriente pelas plantas, a qual raramente ultrapassa uma taxa de 50 % (HUNGRIA et al., 2001). Dessa forma, em pouco tempo, o N dos fertilizantes pode ser perdido para o ambiente pelos processos de lixiviação, resultando no acúmulo de formas nitrogenadas nas águas dos rios e lençóis subterrâneos; bem como pela sua transformação em formas gasosas, tanto pela desnitrificação, pela ação de microrganismos (VAN DER HEIJDEN et al., (2008), como

pela volatilização. Essas emissões contribuem para a chuva ácida e representam uma fonte de emissões de gases de efeito estufa de N_2O , que contribui para a destruição da camada de ozônio, resultando em mudanças climáticas (CAMERON et al., 2013).

Com o objetivo de atender às necessidades nutricionais das culturas, ao mesmo tempo em que se espera manter os índices máximos de produtividade, a agricultura moderna tem concretizado suas metas a partir do processo biológico de fixação de N_2 , com a melhor relação custo/benefício e menor impacto ambiental. A FBN é o principal processo de incorporação de N na biosfera, contribuindo com 65 % do total de entrada de N na Terra, ou 96 % do N proveniente de processos naturais (ORMEÑO-ORRILLO et al., 2013).

3.1.2. Fixação Biológica de Nitrogênio (FBN)

As plantas absorvem o N_2 na forma de íons NO_3^- e NH_4^+ (MADIGAN; MARTINKO; BENDER, 2016). A FBN corresponde a um processo restrito a organismos procariotos composto por diversas espécies de bactérias que habitam o solo. As bactérias capazes de fixar biologicamente o N_2 são chamadas de bactérias diazotróficas, devido a sua capacidade de romper a tripla ligação do N_2 atmosférico, provocando a sua redução até NH_3 , através da atividade do complexo enzimático nitrogenase. A NH_3 sintetizada é imediatamente transformada em íons NH_4^+ , sendo exportada e assimilada pelas plantas sob formas orgânicas, como aminoácidos, amidas e ureídeos. A habilidade de fixar N_2 libera o organismo da dependência de N fixado no ambiente e confere uma vantagem ecológica significativa quando o N é limitante.

O complexo enzimático responsável por catalisar o processo de FBN, denominado nitrogenase, consiste de duas proteínas, a dinitrogenase e a dinitrogenase redutase. Ambas as proteínas contêm ferro, e a dinitrogenase contém, ainda, molibdênio, os quais atuam como cofatores na redução do N_2 . A atividade da nitrogenase pode ser avaliada pelo método de redução de acetileno, visto que esse é um substrato alternativo da enzima, ou pelo método isotópico ^{15}N (SHEARER; KOHL, 1986; ZAHRAN, 1999). A nitrogenase é irreversivelmente inativada na presença de oxigênio (O_2). Entretanto, muitas bactérias fixadoras de nitrogênio são aeróbias obrigatórias. Nesses organismos, a nitrogenase é protegida da inativação pelo O_2 por meio de uma combinação da rápida remoção de O_2 pela respiração, e pela produção de barreiras à entrada de O_2 , criando um ambiente fundamental para esses microrganismos microaerófilos. A participação da

leghemoglobina, uma proteína de ligação ao O_2 , também é crucial para manter um ambiente microaeróbico no nódulo, uma vez que ela atua como um “tampão de O_2 ”, alternando-se entre as formas oxidada e reduzida do ferro, mantendo o O_2 livre no interior do nódulo em concentrações baixas, fornecendo apenas as quantidades às bactérias (OTT et al., 2005). Devido à estabilidade da ligação tripla da molécula de N_2 , sua ativação e redução demandam muita energia. Seis elétrons são necessários para reduzir o N_2 a NH_3 , e as etapas de redução ocorrem diretamente sobre as dinitrogenases. Entretanto, durante a reação de redução, oito elétrons são efetivamente consumidos, sendo dois elétrons perdidos como H_2 para cada mol de N_2 reduzido (NELSON; COX, 2014). O N_2 é reduzido em etapas sucessivas, liberando como produto final NH_3 . Em seguida, à NH_3 são incorporados íons H^+ , em contato com o ambiente aquoso das células, sendo convertida em íons NH_4^+ que serão, então, exportados para a planta hospedeira e incorporados em diversas formas de N orgânico, como os ureídeos, aminoácidos e amidas, contribuindo para o crescimento e manutenção das células (HUNGRIA et al., 2001).

Em sistemas agrícolas produtivos e sustentáveis, a FBN aparece como uma grande contribuição, sendo importante não apenas para o aumento dos rendimentos de culturas, pastagens e plantas arbóreas, mas também porque atua substituindo a prática de aplicação de fertilizantes nitrogenados, destacando-se como uma fonte sustentável desse nutriente em sistemas de cultivo e nos solos (BOHLOOL et al., 1992). Em especial, a maioria das espécies de plantas pertencentes à família Fabaceae, conhecidas como leguminosas, pode utilizar o N_2 a partir da sua associação com bactérias diazotróficas. No Brasil, a soja (*Glycine max*) aparece em grande destaque, sendo uma das leguminosas que mais se beneficia da FBN, contribuindo de forma significativa para a lucratividade dessa cultura no país e em outros países da América do Sul (HUNGRIA et al., 2006). Dentro desse contexto, também é importante salientar o papel da FBN na sustentabilidade e manutenção de ecossistemas, os quais dependem da grande biodiversidade de microrganismos presentes no solo, capazes de estabelecer interações mutualísticas ou protooperativas com muitas espécies de plantas que compõem a flora predominante (DALL'AGNOL et al., 2017).

Diversos trabalhos já foram realizados tendo como foco principal a influência da FBN nos solos tropicais. Nessas regiões, a maioria dos solos úmidos e semiúmidos possui frações minerais compostas de argilas reticulares de capacidade relativamente baixa para reter nutrientes e água. A chave para a fertilidade em longo prazo de tais solos é manter

a matéria orgânica do solo preservada pelos resíduos das colheitas e pela seleção de rotações de culturas adequadas. Entretanto, de acordo com Boddey et al. (1997), a introdução de leguminosas forrageiras em pastagens tropicais corresponde a uma estratégia para aumentar e sustentar a produtividade. Dessa forma, conclui-se que o N derivado de leguminosas pode contribuir para as reservas do solo em benefício de outras culturas, sendo eficientemente transferido para estas. Assim, observando em um contexto de políticas gerais, as pesquisas sobre FBN demonstram que esse processo tem muito a contribuir para a sustentabilidade da agricultura (PEOPLES; CRASWELL, 1992).

3.2. BACTÉRIAS FIXADORAS DE NITROGÊNIO

A capacidade de fixar N_2 é encontrada em vários grupos de microrganismos. As bactérias diazotróficas podem viver de forma livre no ambiente, fixando N_2 para seu próprio uso, ou associam-se com vegetais, contribuindo para o crescimento das plantas. Com exceção das bactérias fotossintéticas e cianobactérias que fixam N_2 , a maioria das bactérias diazotróficas de vida livre é heterotrófica e requer uma fonte de C assimilável para a manutenção dos seus processos biológicos.

Diversas bactérias diazotróficas podem se associar a plantas com diferentes graus de especificidade, baseando-se principalmente na troca de N fixado pelos produtos da fotossíntese. Dessa forma, essas bactérias podem ser classificadas em endofíticas (ex.: *Herbaspirillum seropedicae*, *Gluconacetobacter diazotrophicus*, *Klebsiella* spp., *Azoarcus* spp.), caracterizadas por colonizar o interior de raízes, caules e folhas; associativas (ex.: *Azospirillum* spp., *Azotobacter* spp.), habitando tanto a rizosfera quanto o interior de raízes; ou simbióticas (ex.: *Rhizobium* spp., *Bradyrhizobium* spp., *Azorhizobium* spp.), coletivamente denominadas rizóbios, onde há a formação de estruturas específicas nas raízes, os nódulos, local no qual as bactérias ficarão alojadas fixando N_2 (ORMEÑO-ORRILLO et al., 2013).

As bactérias classificadas como endofíticas ou associativas são comumente encontradas em associação com diferentes espécies vegetais, incluindo alguns representantes da família Poacea, tais como arroz (*Oryza sativa*), milho (*Zea mays*) e cana-de-açúcar (*Saccharum* spp.) (BHATTACHARYYA; JHA, 2012). Entretanto, a principal contribuição para o processo da FBN ocorre pela associação simbiótica estabelecida entre rizóbios e as plantas pertencentes à família Fabaceae. Os rizóbios abrangem uma importante biodiversidade, tanto no grupo das Alfa-Proteobactérias,

bactérias Gram-negativas que apresentam representantes nos gêneros *Allorhizobium*, *Aminobacter*, *Azorhizobium*, *Bradyrhizobium*, *Devosia*, *Mesorhizobium*, *Methylobacterium*, *Microvirga*, *Neorhizobium*, *Ochrobactrum*, *Pararhizobium*, *Rhizobium*, *Ensifer* (= *Sinorhizobium*), *Shinella*; como no de Beta-Proteobactérias, nos gêneros *Paraburkholderia* (*Burkholderia*), *Cupriavidus* e *Trinickia* (CHEN et al., 2003, MOULIN et al., 2001 e SANTOS et al., 2018).

3.2.1. Gênero *Bradyrhizobium*

A primeira classificação sistemática de rizóbios fixadores de N₂, realizada por Fred et al. (1932), baseava-se no conceito de inoculação cruzada em relação à leguminosa hospedeira, e resultou na descrição do gênero *Rhizobium*. Nessa época, as bactérias que formavam nódulos nas raízes da soja eram classificadas na espécie *Rhizobium japonicum*. No entanto, no início dos anos 1980, com o fortalecimento da taxonomia numérica, a qual inclui às técnicas de identificação microbiana análises das propriedades bioquímicas desses microrganismos, levou ao reposicionamento de algumas linhagens em um novo gênero, *Bradyrhizobium*. A etimologia do nome (*bradus*, do grego, significando lento) faz referência à baixa taxa de crescimento em meios de cultura contendo manitol como fonte de C; além disso, com essa fonte de C, o crescimento resulta em reação alcalina (JORDAN, 1982), enquanto que os membros de *Rhizobium* são produtores de ácido e apresentam crescimento rápido (JORDAN, 1984). No início, apenas uma espécie havia sido definida como pertencente ao novo gênero, *B. japonicum*, sendo as demais classificadas como *Bradyrhizobium* spp.

De acordo com Norris (1965), o gênero *Bradyrhizobium* teria se originado nos solos ácidos dos trópicos, podendo ser considerado o ancestral de todas as espécies de rizóbios fixadores de N₂, uma hipótese que foi confirmada em estudos posteriores (LLORET; MARTÍNEZ-ROMERO, 2005; PROVOROV; VOROB'EV, 2000). O gênero abrange uma ampla variedade de bactérias, mas a diversidade verificada na análise do gene 16S RNAr é, frequentemente, baixa (DELAMUTA et al., 2012). Entretanto, com os avanços nas técnicas moleculares dentro da taxonomia bacteriana foram detectadas diferenças genéticas cruciais entre as estirpes, resultando na reclassificação em duas novas espécies, *B. japonicum* e *B. elkanii* (KUYKENDALL et al., 1992).

O gênero *Bradyrhizobium* engloba bactérias diazotróficas que podem viver em associação simbiótica e endofítica com leguminosas e não leguminosas, ocupando uma

grande diversidade de ambientes, e habitando nódulos de raízes ou, mais raramente, caule de algumas espécies (EAGLESHAM et al.; 1990; ISLAM et al.; 2008). Além disso, as bactérias que pertencem a esse gênero são caracterizadas por sua versatilidade fisiológica e ampla distribuição geográfica, sendo mais abundantes em associação com raízes de leguminosas que crescem em áreas tropicais e subtropicais (MENNA et al., 2006; MENNA et al., 2009; MOREIRA et al., 1993).

A persistente ocorrência desse gênero em áreas tropicais está relacionada com a sua adaptação para se associar a uma grande variedade de hospedeiros da região. Dessa forma, novos representantes de *Bradyrhizobium* foram descritos como simbiotes de leguminosas como *Vigna* spp., de onde foram isolados os simbiotes *B. manausense* (SILVA et al., 2014), *B. subterraneum* (GRONEMEYER et al., 2015), *B. vignae* (GRONEMEYER et al., 2016) e *B. brasilense* (MARTINS et al., 2017); leguminosas genitóides que se associam a *B. canariense* (VINUESA et al., 2005); e isoladas de *Pachyrhizus erosus*, popularmente conhecido como o inhame mexicano, Ramírez-Bahena et al., (2009) descreveram duas espécies, *B. pachyrhizi* e *B. jicamae*. Da leguminosa *Cytisus villosus*, foram descritas *B. cytisi* (CHAHBOUNE et al., 2011) e *B. rifense* (CHAHBOUNE et al., 2012) e do amendoim (*Arachis hypogaeae*) foram isoladas quatro espécies na China, *B. lablabi* (CHANG et al., 2011), *B. arachidis* (WANG et al., 2013), *B. guangdongense* e *B. guangxiense* (LI et al., 2015). Em relação à soja, novas espécies foram descritas na China, como a *B. liaoningense*, que apresenta um crescimento muito lento, nodulando genótipos de soja primitivos e modernos (XU et al., 1995); *B. huanghuaihaiense* (ZHANG et al., 2012) e *B. daqingense* (WANG et al., 2013a), ambas aparecendo como nodulíferas bastante eficientes da soja; além de uma quarta espécie, *B. ottawaense*, descrita no Canadá (YU et al., 2014). Além disso, (DELAMUTA et al., 2013) reclassificaram algumas estirpes previamente classificadas como *B. japonicum* grupo Ia em uma nova espécie, *B. diazoefficiens*.

Como já havia sido comentado, o gênero *Bradyrhizobium* também inclui bactérias endofíticas. A espécie *B. betae* corresponde a endofíticos isolados de raízes de beterraba (*Beta vulgaris*) afetadas por deformações semelhantes a tumores (RIVAS et al., 2004), assim como *B. iriomotense*, espécie isolada de uma raiz com tumores da leguminosa *Entada koshunensis* (ISLAM et al., 2008).

Nos últimos anos, uma grande biodiversidade entre as estirpes desse gênero tem sido relatada, mas ainda são poucas as espécies descritas (AHNIA et al., 2018; BÜNGER et

al., 2018; DELAMUTA et al., 2016; HELENE et al., 2017; MARTINS et al., 2017; 2018; DE MATOS et al., 2017; MICHEL et al., 2017). Isso reforça a sugestão de que a diversidade de *Bradyrhizobium* nos trópicos é pouco documentada (MENNA et al., 2009; MENNA et al., 2009a), e que, provavelmente, um número significativo de espécies desse gênero ainda não foi descrita.

3.3. INTERAÇÃO SIMBIÓTICA RIZÓBIOS-LEGUMINOSAS

3.3.1. A FBN em Leguminosas

A importância das leguminosas como culturas alimentares, forrageiras e adubos verdes está principalmente associada à sua capacidade de estabelecer associações simbióticas eficientes com os rizóbios. Como já foi mencionado nos tópicos anteriores, essas bactérias têm sido objeto de muitos estudos, com ênfase em seu papel fundamental na substituição de fertilizantes químicos, melhorando a sustentabilidade agrícola, especialmente em solos tropicais, com profundos impactos ambientais.

A família Fabaceae é considerada a terceira maior família de angiospermas, englobando mais de 18.000 espécies, classificadas em cerca de 750 gêneros, ocupando quase todos os biomas terrestres (PEIX et al., 2015; POLHILL; RAVEN, 1981). Tradicionalmente, foi dividida em três subfamílias: Papilionoideae, Mimosoideae e Caesalpinioideae (LEWIS et al., 2005) mas, recentemente, foi reclassificada em seis subfamílias constituídas pelas Papilionoideae, que permaneceu praticamente inalterada, e cinco novas subfamílias criadas a partir da antiga Caesalpinioideae parafilética: Caesalpinioideae (*sensu stricto*), Duparquetioideae, Cercidoideae, Detarioideae e Dialioideae (LPWG, 2017). A antiga subfamília Mimosoideae teve o número de gêneros reduzido e é, atualmente, referida como o clado Mimosoideae dentro da subfamília Caesalpinioideae (*sensu stricto*).

A maioria das espécies de rizóbios descritas foi isolada de nódulos radiculares da subfamília Papilionoideae, na qual 97% dos seus representantes apresentam a propriedade de realizar a FBN. Entretanto, na grande maioria dos gêneros dentro das subfamílias recém-circunscritas a partir da subfamília Caesalpinioideae (*sensu lato*) o aparecimento de nódulos radiculares é raro (ALLEN; ALLEN, 1981 e GILLER; WILSON 1991). Os poucos gêneros capazes de nodular estão restritos à subfamília Caesalpinioideae (*sensu stricto*), assim como ao clado Mimosoideae que apresenta representantes amplamente nodulados. A posição basal dos membros de Caesalpinioideae capazes de nodular, e o

fato de estarem espalhados entre antigas tribos que, em grande parte, não nodulam, Cassiae e Caesalpiniae, significa que eles são fundamentais em estudos sobre as origens da simbiose de leguminosa-rizóbio (SPRENT et al., 2013; 2017).

Na última década, houve um grande avanço nos estudos taxonômicos de leguminosas e, também, uma expansão na gama geográfica de representantes estudadas. Esses estudos incluíram vários novos registros de nodulação, análises detalhadas da estrutura de nódulos e um aumento no número de gêneros e espécies de rizóbios identificados (SPRENT et al., 2017). Ainda não se chegou a um consenso claro sobre o motivo pelo qual a nodulação evoluiu. Entretanto, considerando a preponderância de leguminosas nodulíferas nos 40 principais gêneros com maior número de espécies, a nodulação foi seguramente uma estratégia bem sucedida para maximizar recursos (LEWIS et al., 2005).

As leguminosas e alguns processos de nodulação evoluíram há cerca de 55 a 60 milhões de anos. Esse processo evolutivo envolveu diferentes tribos e gêneros, e processos de nodulação distintos, mas com a característica comum de que as leguminosas que habitam latitudes mais altas são potencialmente nodulíferas, enquanto que, em regiões tropicais, onde todas as subfamílias são encontradas, muitas espécies não apresentam capacidade de nodular (SPRENT et al., 2013).

3.3.2. Gênero *Chamaecrista*

Apenas nove gêneros de Caesalpinioideae apresentam registros de nodulação, com a presença de nódulos ramificados indeterminados em todas as representantes. As poucas leguminosas capazes de nodular residem, principalmente, na tribo Caesalpiniae, incluindo os gêneros *Campsiandra* Benth., *Chidlowia* Hoyle, *Dimorphandra*, *Erythrophleum*, *Jaqueshuberia*, *Melanoxylon*, *Moldenhawera* e *Tachigali* (SPRENT et al., 2013), todas nativas dos neotrópicos, com exceção de *Erythrophleum*, que tem espécies nativas da África, Ásia e Austrália. Além desses, o único outro gênero que apresenta relatos de nodulação é *Chamaecrista* Moench. (FARIA et al., 1984 e NAISBITT et al., 1992), uma característica que o diferencia dos demais gêneros que compõem a tribo Cassieae (LEWIS et al., 2005). O gênero *Chamaecrista* aparece em grande evidência, ocupando o lugar de oitavo maior gênero de leguminosas, apresentando 330 espécies, sendo nove cultiváveis (LEWIS et al., 2005).

As representantes de *Chamaecrista* podem apresentar hábitos arbóreo, arbustivo ou herbáceo, e podem atuar como alimento para vida selvagem em seu entorno, controle

biológico de pragas, planta ornamental em jardins, e no controle de erosão e restauração de áreas perturbadas, por serem capazes de fixar N₂ (MORRIS, 2012).

Elas estão principalmente distribuídas na América tropical, mas também apresentam algumas espécies nativas da África, Ásia e Austrália (LEWIS et al., 2005). O Brasil aparece como principal centro de radiação desse gênero (SANTOS et al., 2017), apresentando 256 espécies, sendo 207 endêmicas (SOUZA; BORTOLUZZI, 2015). Além disso, *Chamaecrista* também desempenha um importante papel ecológico como uma leguminosa anual nativa do leste, centro-oeste e sul dos Estados Unidos. A espécie *Chamaecrista fasciculata*, objeto de estudo desse trabalho, é distribuída nos Estados Unidos, estando presente nos estados continentais desde o Novo México, Texas, Oklahoma, Kansas, Nebraska, Dakota do Sul, e indo para o leste, incluindo todos os estados, exceto New Hampshire, Vermont e Maine (USDA NRCS, 2012).

Ao mesmo tempo em que estudos filogenéticos utilizando marcadores moleculares têm apoiado o monofiletismo de *Chamaecrista* (CONCEIÇÃO et al., 2009), a nodulação no gênero tem sido apontada como um evento independente na evolução das simbioses envolvendo fixação de N₂ (DELAUX et al., 2015). Estudos anteriores demonstraram que *C. fasciculata*, uma planta herbácea nativa da América do Norte, é nodulada por estirpes de *Bradyrhizobium* (PARKER, 2015 e PARKER; ROUSTEAU, 2014), enquanto que em *C. ensiformis*, uma árvore nativa do Brasil, foi encontrado ocorrência de nodulação por *Mesorhizobium* (MOREIRA et al., 1998). O estudo realizado por Santos et al., 2017 foi o primeiro trabalho que teve como objetivo reportar a diversidade genética de rizóbios capazes de nodular as representantes desse gênero no Brasil. Os resultados obtidos apontam que rizóbios do gênero *Bradyrhizobium* são os simbioses que, preferencialmente, se associam a espécies herbáceas de *Chamaecrista* no Brasil, sendo verificado um alto nível de diversidade genética entre as estirpes. Além disso, foi demonstrado, pela primeira vez, em detalhes, que os nódulos de espécies herbáceas de *Chamaecrista* brasileiras são indeterminados e estruturalmente semelhantes a outros descritos nas Caesalpinioideae, com cordões de infecção nas zonas de invasão (SANTOS et al., 2017).

Em *Chamaecrista* foi reportada a existência de um espectro de estruturas de cordões de infecção. Em espécies arbóreas foi verificado que os bacteroides são retidos dentro de cordões de infecção modificados durante a fixação de N₂, conhecidos como cordões de infecção persistentes; em espécies arbustivas, as paredes desses cordões são

mais finas em relação às arbóreas; e em espécies herbáceas, como a *C. fasciculata*, os bacteroides são liberados dos cordões de infecção em simbiosomas antes que o processo de fixação de N₂ inicie, tal qual ocorre em todas as leguminosas pertencentes ao clado Mimosoideae e, predominantemente, em Papilionoideae (NAISBITT et al., 1992; SPRENT, 2007). Dessa forma, uma análise mais detalhada dessas espécies poderia fornecer informações sobre a evolução dos simbiosomas.

O crescente interesse em utilizar a espécie *C. fasciculata* como um modelo de leguminosa não Papilionoideae inclui motivos como sua posição filogenética, ocorrência de nodulação dentro de um clado de espécies nodulíferas limitadas, presença de morfologia floral não Papilionoideae, hábito de crescimento herbáceo e facilidade de cultivo em ambientes laboratoriais e de campo. Nesse sentido, o sequenciamento completo do transcriptoma da parte aérea, raízes e nódulos de *C. fasciculata*, juntamente com análises de expressão gênica atuam fornecendo recursos para estudar questões fundamentais sobre a evolução das leguminosas (SINGER et al., 2009).

3.3.3. Processos de Infecção e Nodulação

A simbiose entre leguminosas e rizóbios, em condições de restrição de N, leva ao desenvolvimento de novos órgãos na planta conhecidos como nódulos, geralmente formados em raízes, mas que também podem ser encontrados nos caules de poucas espécies (OLDROYD, 2013). A formação de um simples nódulo é resultante de um processo complexo, que ocorre em diversos estágios envolvendo mudanças fisiológicas e morfológicas tanto na célula hospedeira como na bactéria. As mudanças na bactéria visam o recebimento de fontes de C da planta hospedeira, para prover o ATP e o poder redutor necessários para o processo de fixação biológica, já as da planta hospedeira visam, principalmente, assimilar a NH₃ produzida pelas bactérias. Assim, a principal função do nódulo é produzir um ambiente propício à FBN, impondo restrições ao livre fluxo de O₂ (ARAUJO; HUNGRIA, 1994).

Diversos eventos pré-infecção precedem a formação do nódulo, como a liberação pela planta de moléculas que atraem o rizóbio, ativação dos genes da nodulação, adesão da bactéria à superfície das raízes e troca de sinais moleculares entre o hospedeiro e a bactéria. Inicialmente, diversos compostos são exsudados pela raiz da planta hospedeira, como aminoácidos, açúcares, ácidos carboxílicos, homoserina e flavonoides, que agem como substâncias quimiotáticas, atraindo o rizóbio (DHARMATILAKE; BAUER, 1992;

GAWORZEWSKA; CARLILE, 1982; JONES et al., 2007; KAPE et al., 1991). Como resultado desse estímulo, ocorre um grande aumento na população de rizóbios (DOWLING; BROUGHTON, 1986). No entanto, essas bactérias também precisam competir por esses exsudatos com outros microrganismos do solo, incluindo estirpes ineficientes mas adaptadas, e outros microrganismos que podem atuar como antagonistas por meio da produção de antibióticos (HOLLAND, 1970). Essa competição pode prejudicar a colonização das raízes e a nodulação.

Os rizóbios respondem à liberação de exsudatos com a indução da transcrição de genes essenciais à nodulação, genes *nod*, *nol* e *noe* (CASSAB, 1998; SCHULTZE; KONDOROSI, 1998). Os genes *nod* codificam aproximadamente 25 proteínas necessárias para a síntese e exportação do fator Nod, molécula de origem bacteriana essencial ao início do processo de nodulação. A estrutura básica do fator Nod é sintetizada pelos genes *nodC* (codifica uma N-acetil glucosaminiltransferase), *nodB* (codifica uma deacetilase que remove o grupo N-acetil em uma terminação não redutora), e *nodA* (codifica uma aciltransferase que liga a cadeia acil no oligossacarídeo deacetilado) (BROUGHTON et al., 2003). O fator Nod consiste em um sinal de lipo-oligossacarídeo que inicia muitas das alterações de desenvolvimento observadas na planta hospedeira no início do processo de nodulação, incluindo a deformação e o aumento dos pelos radiculares (BREWIN, 2004; VAN BRUSSEL et al., 1986; MIWA et al., 2006; ZAAT et al., 1987), a despolarização da membrana, oscilações do cálcio intracelular e o início da divisão celular no córtex da raiz, que estabelece um meristema e primórdio nodular.

No entanto, apesar da ideia da presença onipresente de genes *nod* e fatores Nod em todos os rizóbios caracterizados estar bem estabelecida, em 2007, foi demonstrado que algumas estirpes de *Bradyrhizobium* fotossintéticas eram capazes de induzir a formação de nódulos tanto em raízes quanto em caules, fixando N₂, em algumas espécies de leguminosas aquáticas do gênero *Aeschynomene*, mesmo na ausência de genes *nodABC*, necessários para a síntese de fatores Nod (GIRAUD et al., 2007). O trabalho realizado por Giraud et al. (2007) mostrou, pela primeira vez, que uma interação simbiótica alternativa, independente de fatores Nod, poderia desencadear a organogênese de nódulos em algumas leguminosas a partir de outras estratégias de sinalização, sem chegar a uma explicação definitiva. Desde então, dois outros exemplos de simbiose independente de fatores Nod foram relatados. Na planta modelo *Lotus japonicus*, normalmente envolvida em um processo simbiótico clássico dependente desses fatores,

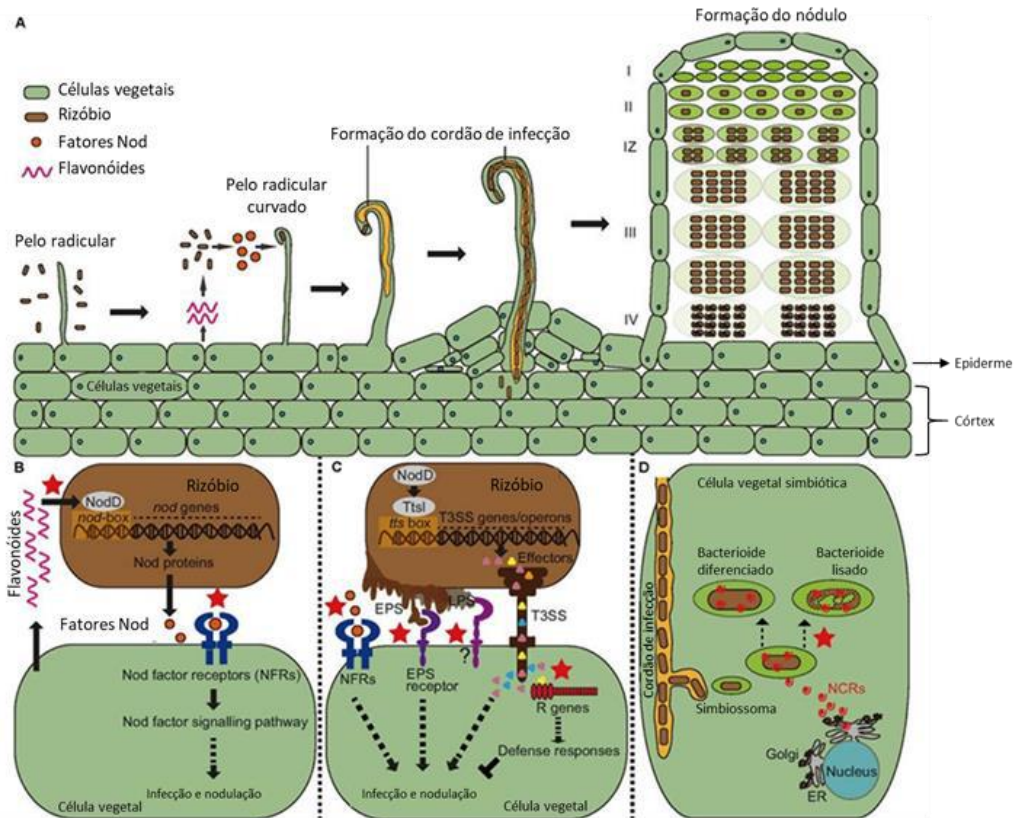
observou-se que estirpes de *Mesorhizobium loti* mutantes incapazes de produzir esses fatores eram ocasionalmente capazes de formar nódulos infectados (MADSEN et al., 2010). Mais recentemente, foi demonstrado que um mutante de *Bradyrhizobium elkanii* deficiente na produção de fatores Nod poderia induzir a organogênese de nódulos em plantas de soja selvagens e mutantes para a expressão do receptor de reconhecimento desses fatores (OKAZAKI et al., 2013). Nesse caso, a simbiose independente de fatores Nod se mostrou possível somente se a bactéria possuísse um sistema de secreção funcional do tipo III (T3SS).

O processo de adesão bacteriana ocorre de forma simultânea à troca de sinais moleculares e pode ocorrer de três maneiras, sendo o processo de aderência à superfície dos pelos radiculares o mais comum e compreendido. Primeiramente, células bacterianas isoladas aderem a sítios específicos (BHUVANESWARI et al., 1980; GULASH et al., 1984) e, depois, outras células aderem às bactérias que já estão presas aos pelos. Em seguida, inicia-se a colonização dos tecidos internos da raiz, ocorrendo a dissolução da parede celular vegetal. Um cordão de infecção cresce da extremidade do pelo radicular encurvado em direção ao córtex, conduzindo o rizóbio para as células interiores. Com a divisão das células do córtex e a indução de um meristema nodular, forma-se o primórdio nodular (Figura 1).

O cordão de infecção ramifica-se à medida que cresce através da raiz, chegando às células que constituem os primórdios nodulares. Por conseguinte, ocorre um aumento dos locais celulares que pode sofrer infecção, garantindo que um número suficiente de primórdios seja colonizado. Uma vez dentro dos primórdios, as bactérias encontram-se envolvidas em uma membrana de origem vegetal, a membrana peribacteroidal (CHEON et al., 1993; VERMA et al., 1993). Os primórdios diferenciam-se em um nódulo e a bactéria sofre um processo de diferenciação, transformando-se em bacteroide (KARR; EMERICH, 1989), conforme demonstrado na Figura 1. Em seguida, inicia-se a expressão dos genes *nif* e *fix*, responsáveis pela síntese e funcionamento do complexo da nitrogenase.

Os nódulos induzidos por rizóbios podem ser classificados em determinados e indeterminados quanto ao crescimento. Os indeterminados são alongados e têm um meristema persistente que continuamente dá origem a novos primórdios nodulares que são subsequentemente infectados. Já os determinados apresentam um meristema transitório, sendo geralmente redondos (GAGE, 2004).

Figura 1 - O processo de infecção e desenvolvimento de nódulos.

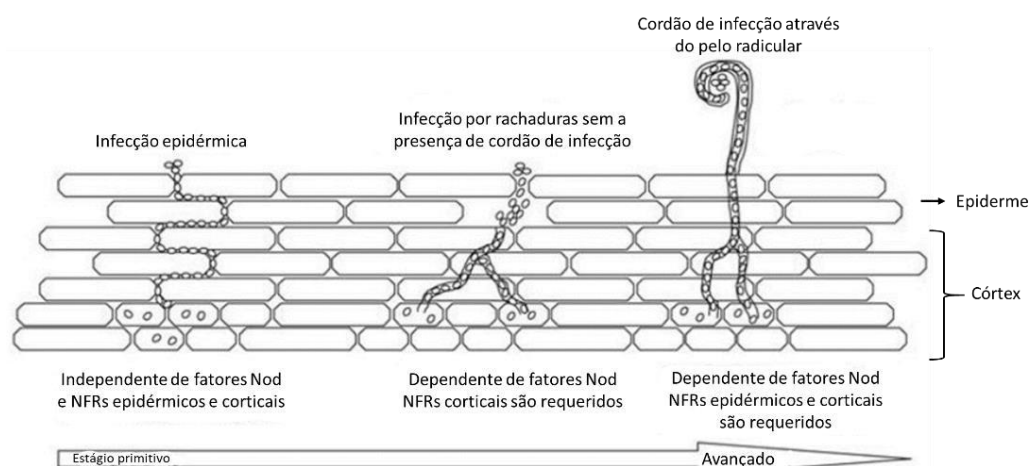


Um nódulo maduro indeterminado contém uma zona de meristema (I), uma zona de infecção (II), uma interzona (IZ), uma zona de fixação de nitrogênio (III) e uma zona senescente (IV). **(B)** O hospedeiro secreta flavonóides para induzir a expressão do gene da nodulação bacteriano (*nod*) através da ativação de proteínas NodD. As enzimas codificadas pelos genes *nod* levam a síntese dos fatores Nod (NF) que são reconhecidos pelos receptores de fatores Nod do hospedeiro (*host Nod factor receptors* - NFRs). **(C)** Além da sinalização de NF, as bactérias também produzem polissacarídeos extracelulares (*extracellular polysaccharides* - EPS) e efetores tipo III para facilitar sua infecção em interações compatíveis. **(D)** Determinadas leguminosas, como o *Medicago*, codificam peptídeos ricos em cisteína (*nodule-specific cysteine-rich* - NCR) específicos de nódulos antimicrobianos para conduzir as bactérias à diferenciação terminal necessária para a FBN.

Fonte: Modificada de Wang et al., (2018).

O segundo modelo de infecção bacteriana é encontrado em leguminosas dalbergióides (LAVIN et al., 2014). Estudos estruturais sobre a formação desses nódulos mostraram a ocorrência de infecção através de rachaduras nos locais de emergência das raízes laterais ou adventícias, não havendo a formação de cordão de infecção (Figura 2). A estrutura interna dos nódulos, embora sejam determinados, é bem diferente daqueles formados após a infecção capilar (SPRENT et al., 2013). Em terceiro lugar, algumas leguminosas, como os membros da tribo Genisteae, sofrem uma infecção epidérmica quando as bactérias aderem entre as células epidérmicas vegetais, sem a presença de cordão de infecção (SPRENT et al., 2013), conforme demonstrado na figura 2.

Figura 2 - Modelo para evolução de vias de infecção em leguminosas.



Estudos sugerem que a infecção epidérmica seja o estado fundamental da infecção por rizóbios nas raízes, seguida pelos modelos de infecção através de rachaduras e infecção de pelos radiculares. Os componentes das vias de infecção dos dois primeiros modelos são menos bem descritos.

Fonte: Modificada de Madsen et al., (2010).

Além da já mencionada capacidade inferior de nodulação de Caesalpinioideae (*sensu lato*) em relação às demais subfamílias de leguminosas, Faria et al. (1984) destacaram que as representantes dessa subfamília comumente apresentam nódulos com morfologia e estrutura primitivas em relação às outras leguminosas. Além disso, em leguminosas mais evoluídas verificou-se que os rizóbios são liberados do cordão de infecção antes de iniciar o processo de fixação de N_2 . Entretanto, em todas as representantes de Caesalpinioideae (*sensu stricto*), com exceção de algumas espécies herbáceas de *Chamaecrista*, as bactérias são retidas dentro do cordão de infecção durante o processo de fixação de N_2 , como encontrado também em outros cinco gêneros pertencentes à subfamília Papilionoideae (FARIA et al., 1987). Dessa forma, tem sido sugerido que a retenção de bactérias dentro de linhas de infecção representa um estágio primitivo (NAISBITT et al., 1992; SPRENT; RAVEN, 1985).

Os resultados obtidos por Naisbitt et al. (1992) sugerem que o gênero *Chamaecrista* pode contribuir para o entendimento de como a “primitiva” Caesalpinioideae evoluiu para as simbioses mais “avançadas” encontradas no clado Mimosoideae e na subfamília Papilionoideae. Considerações filogenéticas realizadas por Doyle (1998) sugerem que a nodulação pode ter evoluído em *Chamaecrista* de maneira independente às representantes de Papilionoideae, e juntamente ou independentemente do clado Mimosoideae. Todas as espécies de *Chamaecrista* têm nódulos indeterminados do tipo comumente encontrado em Caesalpinioideae, e também comum a todas as

espécies do clado Mimosoideae e a muitas pertencentes a Papilionoideae, embora algumas espécies de *Chamaecrista* apresentem cordões de infecção persistentes (NAISBITT et al., 1992). Embora ainda limitadas, as pesquisas sobre a nodulação em *C. fasciculata* fornecem indicações sobre a evolução dos nódulos. A nodulação em *C. fasciculata* foi descrita por Naisbitt et al. (1992), sendo caracterizada por uma infecção por meio da adesão aos pelos radiculares, com a presença de um cordão de infecção não persistente. Dentro dos nódulos, uma hemoglobina com reduzida afinidade a O₂ tem sido caracterizada como intermediária entre hemoglobinas não-simbióticas e a leghemoglobina (SINGER et al., 2009).

3.4. TAXONOMIA BACTERIANA

A taxonomia ou sistemática constitui uma disciplina científica básica que tradicionalmente engloba classificação, nomenclatura e identificação dos organismos, incluindo, sobretudo, estudos sobre mecanismos genéticos, que sustentam os processos evolutivos e a filogenia (ZHI et al., 2012). No entanto, diferentemente de animais e plantas, as bactérias possuem características morfológicas e fisiológicas simples, limitadas e insuficientes para descrições taxonômicas (WOESE, 1987). Atualmente, reconhece-se que a classificação bacteriana deve refletir as relações filogenéticas entre as bactérias. Nesse contexto, a genômica tornou-se uma metodologia promissora, pois fornece meios reprodutíveis, confiáveis e altamente informativos para inferir essas relações (CHUN et al., 2018).

A classificação envolve a geração de uma estrutura ordenada e confiável para acomodar os organismos com base nas semelhanças e diferenças de seus caracteres. A nomenclatura lida com termos usados para reconhecer postos na hierarquia taxonômica e com a prática de rotulagem das unidades definidas, que deve ser realizada de forma rigorosa e internacionalmente reconhecida em grupos taxonômicos, seguindo as regras estabelecidas no Código Internacional de Nomenclatura de Bactérias (LAPAGE, 1992). Já a identificação atua buscando alocar os organismos em níveis taxonômicos validamente publicados e, no caso dos isolados não pertencerem aos grupos conhecidos, eles serão descritos e classificados como um novo táxon. O reconhecimento dessas etapas só se torna possível se as principais características dos isolados forem determinadas por métodos padronizados (KÄMPFER; GLAESER, 2012; VANDAMME et al., 1996; ZHI et al., 2012).

A ciência taxonômica teve seu início no século XVIII, a partir dos intensos estudos de Lineu. Posteriormente, estudos envolvendo principalmente análises de microscopia levaram Ferdinand Cohn, em 1872, a classificar as bactérias em seis gêneros com base nas suas propriedades morfológicas. Inicialmente, os marcadores taxonômicos mais importantes usados nessa abordagem clássica estavam limitados à morfologia, exigências de crescimento e potencial patogênico. Em seguida, foram incluídos à classificação traços sorológicos e outros caracteres fisiológicos, utilizados para distinguir principalmente bactérias patogênicas. Na primeira metade do século passado, novos dados bioquímicos enriqueceram as análises enzimáticas e metabólicas, contribuindo, significativamente, para a resolução das classificações quando comparadas com as baseadas apenas em características morfológicas (SCHLEIFER; STACKEBRANDT, 1983). Nesse momento, surge a taxonomia numérica, a qual permitiu uma análise conjunta de todos os dados a partir de métodos estatísticos que forneceram medidas quantitativas de similaridade entre as bactérias, de uma forma menos subjetiva (ROSSELLO-MORA; AMANN, 2001; SCHLEIFER, 2009; SENTAUSA; FOURNIER, 2013).

Nas décadas seguintes, o conhecimento mais aprofundado da composição do DNA e o desenvolvimento de técnicas de biologia molecular promoveram uma revolução na taxonomia bacteriana. Nesse sentido, o avanço de maior reconhecimento consistiu no sequenciamento da subunidade menor do RNA ribossômico (gene 16S RNAr), que possibilitou, pela primeira vez, o estabelecimento de um sistema taxonômico hierárquico com base em um eficiente marcador molecular (LUDWIG; SCHLEIFER, 1994). O marco que resultou na taxonomia moderna foi conduzido por Woese et al. (1990) ao introduzirem o conhecimento de que a vida neste planeta se divide em três domínios, comumente conhecidos como eubactérias, arqueobactérias e eucariotos. A partir de uma análise comparativa das sequências do RNAr, construiu-se uma árvore filogenética universal, fornecendo uma visão global sobre a filogenia de procariotos, e tornando-a, dessa forma, a espinha dorsal da classificação microbiana (KÄMPFER; GLAESER, 2011; VANDAMME et al., 1996).

Diante do surgimento de diversas abordagens capazes de enriquecer os dados já pré-existentes, a taxonomia bacteriana passou a integrar vários níveis de informação, com o objetivo de construir uma classificação multidimensional. Neste contexto, COLWELL (1970) introduz a taxonomia polifásica, que se baseia no estudo integrado de

características fenotípicas, genotípicas e filogenéticas (ROSSELLO-MORA; AMANN, 2001; SENTAUSA; FOURNIER, 2013; VANDAMME et al., 1996).

3.4.1. Taxonomia Polifásica

Conforme comentado no tópico anterior, Colwell (1970) introduziu o termo taxonomia polifásica para se referir a uma nova era de classificações baseada na ideia de construir um consenso de classificação com um propósito geral, utilizando-se de todos os métodos disponíveis até aquele momento. Entretanto, por mais que Colwell (1970) não tenha feito uma referência a interpretações filogenéticas, Vandamme et al. (1996) definiram a taxonomia polifásica como um processo que geraria classificações baseadas em dados reunidos por diversos métodos, mas que deveriam ser consistentes com uma classificação filogenética, estabelecida na época, principalmente, a partir das sequências do gene 16S RNAr (KÄMPFER; GLAESER, 2012). Nesse sentido, a taxonomia polifásica é utilizada até os dias de hoje, integrando todas as informações genotípicas, fenotípicas e filogenéticas disponíveis.

Nos últimos quarenta anos, muitas técnicas de caracterização bacteriana foram desenvolvidas em cada nível de informação, tornando a classificação bacteriana mais estável, ao mesmo tempo em que novos softwares se tornaram disponíveis para otimizar a análise numérica de dados. As informações genotípicas advêm dos ácidos nucleicos (DNA e RNA), a partir de técnicas como a quantificação do conteúdo G+C, hibridação DNA-DNA, padrões de restrição (RFLP, AFLP) e PCR *fingerprinting*; enquanto que as informações fenotípicas abrangem análises morfológicas, fisiológicas e quimiotaxonômicas. As informações filogenéticas são obtidas, sobretudo, a partir das sequências dos genes ribossomais e, em menor escala, de genes que codificam proteínas conservadas (ROSSELLO-MORA; AMANN, 2001; SENTAUSA; FOURNIER, 2013). Classificações em nível de gênero ou espécie podem, muitas vezes, ser obtidas por apenas uma ou poucas técnicas. No entanto, se o intuito é descrever uma nova espécie, são necessárias informações fenotípicas, genotípicas e filogenéticas e, normalmente, mais de uma técnica é aplicada para cada informação.

O conceito de espécie microbiana já passou por diversas discussões na comunidade científica, que busca alcançar uma definição que seja universal. Com o progresso das metodologias representando um grande potencial sob a sistemática bacteriana, foi formado o Comitê Internacional para a Sistemática dos Procariontes

(STACKEBRANDT et al., 2002) para reavaliar a definição de espécies de procariotos. O comitê reconheceu a definição de espécies descrita por Rossello-Mora; Amann (2001), considerada pragmática, operacional e universalmente aplicável, e que define que uma espécie é uma categoria que circunscreve um grupo genomicamente coerente de isolados individuais compartilhando um alto grau de similaridade em recursos independentes (hibridação DNA-DNA, descrição fenotípica e análise filogenética do gene ribossomal 16S) (VANDAMME et al., 1996) e testados comparativamente sob condições altamente padronizadas. Em função de algumas desvantagens em relação à reprodutibilidade, os valores de hibridação DNA-DNA não permanecem como o padrão-ouro para delimitação de espécies, sendo principalmente adotadas metodologias com base na similaridade derivada de sequências genômicas pareadas (CHUN et al., 2018; GORIS et al., 2007; KONSTANTINIDIS; TIEDJE, 2005; MEIER-KOLTHOFF et al., 2013; RICHTER; ROSSELLO-MORA, 2009).

Com o advento das análises genômicas, pesquisadores encorajaram o desenvolvimento de novas técnicas, como a identidade média de nucleotídeos (*Average Nucleotide Identity* - ANI) e a hibridação DNA-DNA digital (dDDH), com o objetivo de propor novas espécies com base em outros métodos, desde que esses possam demonstrar que, dentro dos táxons estudados, exista um grau suficiente de congruência entre a técnica usada e os valores de hibridação DNA-DNA.

3.4.1.1. Caracterização fenotípica

Antes das técnicas moleculares estarem disponíveis para os microbiologistas, a classificação era baseada exclusivamente nas características morfológicas, fisiológicas e bioquímicas. Individualmente, muitas dessas características mostraram-se irrelevantes como parâmetros para indicar o parentesco genético, mas como um todo, fornecem informações descritivas que nos permitem reconhecer grupos. Essas investigações estavam diretamente ligadas ao uso de culturas puras e à capacidade laboratorial de cultivar os organismos e analisar suas propriedades (ROSSELLO-MORA; AMANN, 2001).

A morfologia de uma bactéria inclui características celulares, como forma, presença de endósporos, flagelos, corpos de inclusão, coloração de Gram e coloniais, a partir da observação da coloração, dimensões, forma e aspecto. As características fisiológicas apresentam dados sobre o crescimento bacteriano em diferentes temperaturas, valores de

pH, tolerância à salinidade, outras condições abióticas, crescimento frente a diferentes fontes de carbono e nitrogênio, tolerância a agentes antimicrobianos, atividade enzimática, entre outros (ON; HOLMES, 1992; VANDAMME et al., 1996).

Uma das desvantagens de analisar apenas o fenótipo reside em que nem todo o potencial de informação de um genoma procariótico é expresso, uma vez que a expressão gênica está diretamente relacionada às condições ambientais, como as condições de crescimento no laboratório. Dessa forma, sabe-se que o fenótipo procariótico não pode ser baseado na simples observação do organismo, uma vez que a maioria mostra ciclos de vida com diferentes estágios morfológicos. Nesse sentido, torna-se difícil realizar análises filogenéticas com base apenas nesses dados, apesar da grande diversidade de testes existentes atualmente. Isso ocorre porque não é possível distinguir os caracteres em autapomórficos (caracteres homólogos originados no mais recente ancestral comum do táxon, e exclusivo para o grupo) ou sinapomórficos (caracteres homólogos comuns a dois ou mais táxons originados no ancestral comum mais recente). Por isso, o sequenciamento completo do genoma se tornou esclarecedor e necessário à homologia das características fenotípicas. As análises morfológicas são as mais trabalhosas da classificação bacteriana, já que requerem muito tempo e habilidade, sendo necessária uma alta padronização dos procedimentos realizados, garantindo reprodutibilidade de resultados dentro e entre laboratórios (VANDAMME et al., 1996). Outro aspecto importante que se recomenda ao analisar o fenótipo de uma espécie procariótica é que as estirpes devem ser escolhidas para representar a diversidade conhecida e os nichos ambientais do grupo em estudo (ROSSELLO-MORA; AMANN, 2001).

A caracterização bioquímica da célula bacteriana é conhecida como quimiotaxonomia e utiliza vários elementos estruturais da célula, incluindo as camadas celulares externas, como peptidoglicano, ácidos teicóicos e ácidos micólicos; a membrana celular, que apresenta ácidos graxos, lipídios polares e pigmentos; ou constituintes do citoplasma, como as poliaminas. Esses compostos se apresentam em diferentes concentrações e composições nas células, mas podem aparecer como um padrão para grupos de bactérias relacionadas, podendo ser utilizados como biomarcadores taxonômicos (VANDAMME et al., 1996). Como recomendação geral para a descrição de um novo grupo taxonômico, é necessário estudar as características quimiotaxonômicas dos grupos mais intimamente relacionados para comparação (TINDALL et al., 2010).

Muitas diferenças taxonômicas significativas podem ser detectadas nos estudos referentes à composição bioquímica bacteriana, o que demonstra que a abordagem quimiotaxonômica melhorou significativamente a resolução das classificações quando comparada com a baseada apenas em características morfológicas (SCHLEIFER; STACKEBRANDT, 1983; ZHI et al., 2012). No entanto, é importante que a variação observada na composição química celular seja resultado de diferenças genéticas e não dependente da variação nas condições de cultivo. Dessa forma, é necessário cultivar as culturas sob regimes de crescimento cuidadosamente padronizados (ROSSELLO-MORA; AMANN, 2001).

Dentre os diversos compostos que podem ser analisados durante a classificação e identificação microbiana, os lipídeos são os que aparecem com maior frequência, já que uma variedade de lipídios está presente nas células bacterianas. Os ácidos graxos são os principais constituintes de lipídios e lipopolissacarídeos e têm sido amplamente utilizados para fins taxonômicos. A variabilidade nos grupos de comprimento de cadeia, posição de ligação dupla e grupos substituintes das estruturas químicas dos ácidos graxos provou ser muito útil para a caracterização de taxa bacteriana (VANDAMME et al., 1996).

Diversos protocolos de extração foram propostos na tentativa de melhorar a recuperação de compostos das amostras, com o objetivo de gerar comparações e identificações em um curto período de tempo. Um estudo realizado por Tighe et al. (2000) levou à caracterização do perfil de ácidos graxos em mais de 600 estirpes de rizóbios a partir do *Sherlock Microbial Identification System* (MIS), um sistema de cromatografia gasosa que analisa a mistura de lipídios do extrato celular. O sistema utilizado configura uma abordagem automatizada, sendo muito sensível e reprodutível, em função da sua capacidade de gerar uma biblioteca a partir dos perfis gerados, fornecendo um meio seguro para comparar e contribuir na identificação de um grande número de estirpes.

3.4.1.2. Caracterização filogenética e genética

Na década de 1980, a amplificação direcionada por PCR e o sequenciamento de genes de RNA ribossômico (RNAr), em especial a subunidade menor 16S RNAr, foram considerados uma revolução nos estudos taxonômicos procarióticos (KÄMPFER; GLAESER, 2012; SENTAUSA; FOURNIER, 2013). A introdução da análise do gene 16S RNAr forneceu, pela primeira vez, uma hipótese na qual outros elementos, como genes, poderiam ser comparados ao investigar a taxonomia e a evolução dos procariotos

(TINDALL et al., 2010; WOESE, 1987). Conforme o sequenciamento das moléculas de RNAr se tornou mais acessível, ocorreu a formação gradual de uma base de dados de sequências de 16S RNAr, favorecendo um compartilhamento seguro de dados laboratoriais (VANDAMME et al., 1996).

A utilização do gene 16S RNAr se deve ao fato dele estar presente em todas as células bacterianas, apresentar expressão constante e ser composto de domínios altamente conservados, mas que também apresentam variabilidade, o que torna esse gene um marcador genético universal para identificação taxonômica bacteriana (RIVAS et al., 2009). A proporção do conteúdo de informação para o comprimento da sequência gênica é relativamente alta, levando a construção de árvores filogenéticas de alta resolução e com bom suporte estatístico (LUDWIG; SCHLEIFER, 1994). Entretanto, as sequências de 16S RNAr frequentemente contêm informações insuficientes para mostrar as relações em posições taxonômicas mais baixas, particularmente no nível de espécies (STACKEBRANDT et al., 2002).

Dessa forma, há vários pontos que precisam ser considerados ao avaliar os dados, sendo o poder de resolução da sequência do gene 16S RNAr abaixo do nível de gênero um dos principais pontos mais críticos. Em alguns gêneros procarióticos, além da conservação da sequência, é possível que um único genoma abrigue mais de um operon ribossomal, além de existirem evidências de que ele pode sofrer transferência horizontal (KÄMPFER; GLAESER, 2012). Os organismos que compartilham sequências desse gene muito semelhantes ou idênticas podem apresentar outras regiões do genoma com maiores níveis de variação nucleotídica. Assim, à medida que mais sequências do genoma se tornarem disponíveis, uma maior seleção de genes com diferentes graus de resolução estará disponível. Várias abordagens genômicas resumidas por Cole et al. (2010), baseadas em comparações genômicas de sequências concatenadas de todos os genes conservados ou selecionados, confirmaram o sistema hierárquico do 16S RNAr pelo menos a nível de gênero.

Por isso, o 16S RNAr permanece como a abordagem de primeira linha para avaliar o status taxonômico de uma bactéria. Kim et al. (2014) realizaram um estudo para determinar o limiar ótimo da similaridade da sequência do 16S RNAr para delineamento de espécies que corresponde a 95-96 % do ANI. Foi proposto que duas estirpes sejam consideradas como membros da mesma espécie se a identidade de sequência for > 99 %, e que haverá um indicativo de nova espécie com uma identidade de <98,65 % (KIM et

al., 2014; KONSTANTINIDIS; TIEDJE, 2005). No entanto, é necessário realizar essa análise de similaridade em conjunto com as espécies-tipo do mesmo gênero mais "intimamente relacionadas" (TINDALL et al., 2010).

Com o objetivo de melhorar a informação filogenética obtida a partir de regiões ribossômicas, é possível realizar análises da região intergênica (*Intergenic Transcribed Spacer* - ITS) 16S-23S RNAr, que em estudos prévios já demonstrou ser útil em análises de filogenia e taxonomia de *Bradyrhizobium* (DELAMUTA et al., 2013; 2016; MENNA; BARCELLOS et al., 2009; WILLEMS et al., 2001). No entanto, a análise é válida para bactérias que apresentam apenas uma cópia do ITS, com uma sequência longa, apresentando maior variabilidade e, por sua vez, uma boa resolução filogenética (HUNGRIA et al., 2015). Em um estudo anterior realizado por Willems et al. (2001) com estirpes de *Bradyrhizobium* os autores observaram uma alta correlação entre o valor de similaridade de sequências de ITS e o valor obtido para hibridação DNA-DNA para essa região, sugerindo que estirpes com mais de 95,5 % de similaridade em sequências de ITS pertenciam à mesma espécie, apresentando mais de 60 % de DDH (WILLEMS et al. 2001; 2003).

Considerando as limitações comentadas anteriormente em relação ao gene 16S RNAr, surgiu a necessidade de buscar marcadores filogenéticos alternativos. Nesse contexto, a análise combinada de vários genes codificadores de proteínas ou análise conjunta de múltiplos genes (loci) (*Multilocus Sequence Analysis*, MLSA) foi usada com sucesso para elucidar relações taxonômicas entre espécies. Essa metodologia propõe a análise conjunta de genes conservados e presentes em procariotos relacionados, atuando no metabolismo basal das células, sendo conhecidos como genes *housekeeping*. Esses genes devem ser onipresentes no táxon em estudo e estar em cópia única no genoma, porém, com taxas evolutivas mais rápidas que o 16S RNAr, permitindo um poder de resolução superior para avaliação taxonômica e evolutiva entre espécies bastante próximas. Entretanto, os resultados devem apresentar dados correlacionados com 16S RNAr e com os percentuais de similaridade obtidos no ANI (GEVERS et al., 2005; SCHLEIFER, 2009; STACKEBRANDT et al., 2002; VANDAMME; PEETERS, 2014).

A abordagem mais simples consiste em concatenar as sequências de múltiplos genes e usar essas concatenações para realizar uma análise filogenética, de modo que os efeitos de recombinação genética ou transferência horizontal sejam "tamponados" (SCHLEIFER, 2009). Dessa forma, a seleção dos genes analisados em conjunto constitui

um ponto crítico no MLSA. Considerando que os genes informativos dentro de um gênero ou família podem não ser úteis ou mesmo estarem presentes em táxons mais distantes, seria insustentável propor um conjunto universal de genes para realizar a classificação taxonômica de todos os procariotos. Além disso, o alto grau de conservação de um determinado gene presente em todas as espécies pode indicar que ele não evoluiu com rapidez suficiente para distinguir táxons intimamente relacionados. A taxonomia a nível de espécie tem como foco relações filogenéticas profundas, e, por isso, visa à identificação de conjuntos de genes que possam ser usados em todas as estirpes de um grupo particular, ainda que alguns genes possam ser informativos em mais de um grupo específico (GEVERS et al., 2005).

Em estudos de diversidade e identificação de bactérias fixadoras de N₂, em especial do gênero *Bradyrhizobium* (RIVAS et al., 2009), são comumente utilizados os genes *housekeeping atpD*, *dnaK*, *glnII*, *gyrB*, *recA* e *rpoB* (DELAMUTA et al, 2012; 2013; MENNA et al., 2009; PARKER, 2004; VINUESA et al., 2005), os quais são combinados e, por sua vez, concatenados em função da resolução. Esses genes estão envolvidos em funções celulares fundamentais: o gene *atpD* codifica a subunidade beta da membrana ATP sintase; o gene *dnaK* codifica uma chaperona da classe Hsp70; o gene *glnII* codifica a enzima glutamina sintetase II; o gene *gyrB* codifica uma topoisomerase II responsável por introduzir torções negativas nos cromossomos durante a replicação; o gene *recA* codifica a recombinase A, que está envolvida na recombinação de sequências de DNA complementares e permite a troca de fragmentos, e o produto do gene *rpoB* desempenha um papel essencial na transcrição do DNA.

Um estudo prévio realizado por Durán et al. (2014) relatou que os valores de similaridade entre as espécies descritas do gênero *Bradyrhizobium* até o momento variava de 87 a 97%, sugerindo que, para uma sequência contendo cinco genes concatenados, o valor de 97% poderia ser considerado um nível de corte para delimitação de espécies desse gênero.

No caso dos rizóbios, além do MLSA, também é possível realizar inferências evolutivas com base no processo de simbiose, utilizando análises filogenéticas de genes simbióticos, como aqueles envolvidos na nodulação e FBN. Esse processo envolve a expressão de diversos genes, incluindo aqueles relacionados à síntese e atividade da nitrogenase (*nif* e *fix*) e outros relacionados à formação dos nódulos (*nod*, *nol* e *noe*). Estes genes estão agrupados em operons complexos localizados em plasmídeos ou em ilhas

simbióticas flanqueadas por sequências de inserção, como ocorre em *Bradyrhizobium* (MENNA; HUNGRIA, 2011). O gene de nodulação *nodC* representa o marcador simbiótico mais utilizado para estudos de diversidade e evolução envolvendo genes de nodulação, enquanto que o gene *niH* é geralmente o marcador molecular escolhido para monitorar a fixação biológica de N₂.

Algumas das metodologias que contribuem para a caracterização genômica se baseiam em variações características nos sítios genômicos de restrição (sítios de restrição polimórficos), como a ocorrência de inserções e deleções, repetições de sequências de DNA ou microssatélites, polimorfismos de nucleotídeo único (SNPs) ou outras variações de sequências distribuídas nos genomas bacterianos, que geram um perfil de bandas de DNA. Esses padrões de bandas, ou as "impressões digitais de DNA", são usados para comparar um isolado com outro, a fim de avaliar a diversidade intraespecífica presente. A maioria das técnicas de *fingerprinting* de DNA baseia-se na presença ou ausência dos sítios de restrição, enquanto outros são baseados nas homologias de *primers* curtos de oligonucleotídeos (FAKRUDDIN et al., 2013).

Conforme revisado por Fakruddin et al. (2013), diversas metodologias de *fingerprinting* de DNA foram desenvolvidas e aperfeiçoadas desde os anos 1970 e muitas são empregadas até os dias de hoje. Tais metodologias conferem um poder discriminatório muito maior do que aqueles obtidos por análises puramente fenotípicas. Dentre essas metodologias, aquelas que aparecem em maior frequência na identificação de rizóbios são a RFLP (*Restriction Fragment Length Polymorphism*) e a rep-PCR (*Repetitive sequence-based PCR*). A primeira técnica foca na natureza polimórfica dos locais de digestão, que pode ser visualizada a partir da atuação de enzimas de restrição dentro de uma região genética definida. O DNA total é digerido por essas enzimas formando diversos fragmentos, os quais serão imobilizados numa membrana e hibridados com sondas marcadas (GREEN; SAMBROOK, 2012). Apenas os fragmentos de DNA que apresentam especificidade para a sonda aplicada são detectados, simplificando grandemente a análise da endonuclease de restrição, a partir de uma diferenciação das estirpes pela variação da posição e comprimento das sequências alvo. Em se tratando de análises realizadas com rizóbios, diversos trabalhos demonstraram que a caracterização intraespecífica das estirpes pode ser eficientemente realizada através da análise do polimorfismo do DNA gerado pela digestão do mesmo com enzimas de restrição, que pode ser feita utilizando-se o DNA total (RFLP) ou uma parte dele (PCR- RFLP), sem a

ocorrência de imobilização e hibridação (GERMANO et al., 2006; LAGUERRE et al., 1994; 1996).

A rep-PCR é uma metodologia baseada em amplificação de DNA envolvendo elementos repetitivos e conservados do DNA, normalmente no espaço intergênico, ou seja, em regiões não codificantes do genoma bacteriano, e em ambas as orientações. Há três tipos principais de sequências repetitivas abordadas pela técnica: sequências palindrômicas extragênicas repetitivas (elementos REP) (STERN et al., 1984), sequência de consenso intergênico repetitivo enterobacteriano (ERIC) (HULTON et al., 1991) e elementos BOX (KOEUTH et al., 1995). Uma vez que esses elementos são universais em eubactérias e estão associados a elevados graus de polimorfismo, essas regiões gênicas provavelmente refletem os processos de adaptação evolutiva dos microrganismos (VAN BELKUM, 1999), sendo uma metodologia empregada com frequência nos estudos de taxonomia. As sequências BOX são alvos mais comumente utilizados em bactérias Gram-negativas, consistindo em três subunidades (boxA, boxB e boxC) que somadas representam aproximadamente 154 pb (MARTIN et al., 1992). Em 1994, Versalovic et al. (1994) realizaram estudos de hibridação e PCR usando sondas e *primers* que demonstraram que sequências oligonucleotídicas dentro da subunidade boxA (59 pb) parecem ser altamente conservadas entre diversas espécies bacterianas. Essas sequências repetitivas podem estar presentes em ambas as orientações e são amplificadas por um único *primer*, o BOX-AIR. Uma vantagem de interesse na rep-PCR é o fato dela possuir um bom poder discriminatório, que pode ser aumentado ainda mais pelo uso de conjuntos múltiplos de *primers* (VERSALOVIC et al., 1994).

Uma desvantagem das metodologias baseadas em *fingerprinting* de DNA é que os resultados são, frequentemente, de difícil comparação entre laboratórios, devido à falta de padronização. Dessa forma, os métodos de *fingerprinting* de DNA são limitados para a descrição de espécies. Essas técnicas de identificação podem, no entanto, ser muito úteis para demonstrar se os isolados de um novo táxon constituem ou não clones, sendo uma metodologia de análise da diversidade intraespecífica (TINDALL et al., 2010).

Nos últimos 50 anos, grandes conquistas foram obtidas no desenvolvimento de análises genéticas para a caracterização de microrganismos. Iniciando com a estimativa do percentual de nucleotídeos guanina e citosina presentes no DNA genômico total (conteúdo G+C), evoluindo para a metodologia de hibridação DNA-DNA (DDH), e chegando ao desenvolvimento de novas e rápidas tecnologias de sequenciamento

genômico de alto rendimento, uma técnica mais rigorosa para distinguir espécies altamente relacionadas. Nesse sentido, tornou-se viável examinar detalhadamente as sequências dos genomas com o objetivo de prever dados com potencial valor molecular filogenético. No entanto, de acordo com Klenk; Göker (2010), essa grande riqueza de informações pode apresentar dificuldades intrínsecas durante a inferência de relações filogenéticas, considerando a frequência significativa de transferência gênica lateral ou horizontal no DNA bacteriano.

Desde a década de 1970, a medida da similaridade genética geral entre isolados tem sido a base da caracterização genotípica, sendo avaliada pelo grau de hibridação entre genomas de duas espécies em condições padrão. Essa técnica, relatada por Wayne et al. (1987) ficou amplamente conhecida como hibridação DNA-DNA ou DDH. Dessa maneira, antes da disponibilidade da informação da sequência de DNA, o DDH forneceu um meio padronizado para identificar e classificar os procariotos que não possuíam características morfológicas ou fenotípicas bem definidas (GEVERS et al., 2005). As condições necessárias para empregar a técnica incluem altas temperaturas para promover a desnaturação da dupla-hélice do DNA e posterior diminuição da mesma, para permitir a reassociação das fitas. A similaridade é medida pela porcentagem de associação das fitas e pela diferença na temperatura de *melting* (T_m), necessária para dissociar o DNA. As fitas híbridas pouco similares não apresentam pareamentos fortes, e uma temperatura baixa é capaz de dissociá-las, em comparação com o pareamento da molécula original (GEVERS et al., 2005). Nesse momento, a definição filogenética pragmática de espécie se fundamentou no pressuposto de que uma espécie incluiria estirpes com uma similaridade de 70 % ou mais, baseada em experimentos de DDH, apresentando $\Delta T_m < 5$ °C (VANDAMME et al., 1996; WAYNE et al., 1987).

Atualmente, a verificação do conteúdo dos nucleotídeos guanina e citosina (G+C) no DNA ainda é um parâmetro útil, sendo uma avaliação necessária na lista de dados imprescindíveis para a descrição de um novo táxon, e um pré-requisito importante para determinar as condições usadas nos experimentos de DDH. No entanto, essa análise é apenas um indicativo de exclusão na classificação de bactérias, em que duas estirpes diferentes em mais de 3 % mol não devem ser consideradas como membros da mesma espécie e em mais de 10 % mol do mesmo gênero (MEIER-KOLTHOFF et al., 2014; MESBAH et al., 2011). Entretanto, uma composição similar de bases G+C nos DNAs não comprova que as duas estirpes estão intimamente relacionadas.

O método DDH passou a ser frequentemente criticado em função da baixa reprodutibilidade e por ser um procedimento muito laborioso e custoso (GEVERS et al., 2005; ROSSELLO-MORA; AMANN, 2001). Outra dificuldade consiste em não existir um banco de dados que permita o depósito da informação de pesquisas realizadas em diferentes laboratórios, impedindo a busca e a comparação das informações (KÄMPFER; GLAESER, 2011). Além disso, não existe uma teoria ecológica ou evolucionista sobre como as demarcações de espécies deveriam se relacionar com a extensão do compartilhamento de genes entre organismos (GEVERS et al., 2005). Nesse sentido, o valor estabelecido de 70 % de parentesco entre os DNA parece ser apenas indicativo e não absoluto (VANDAMME et al., 1996).

Dessa forma, a metodologia de DDH vem sendo substituída como "padrão ouro" na taxonomia bacteriana pelo sequenciamento do genoma e, por sua vez, pelas análises de similaridade derivadas de sequências genômicas (CHUN et al., 2018). Nesse sentido, tem havido uma série de esforços para o desenvolvimento de métodos bioinformáticos capazes de substituir o DDH. Considerando que o valor de DDH reflete basicamente o grau de parentesco ou similaridade entre dois genomas, esses esforços concentraram-se na criação de valores análogos aos valores de DDH. Esses valores aparecem como forma de similaridade ou distância evolutiva e foram referidos como índice geral relacionado ao genoma (*Overall Genome Related Index - OGRI*) (CHUN; RAINEY, 2014).

Uma das plataformas mais adequadas para o sequenciamento de genomas de procariotos dentre os sequenciadores de última geração é o MiSeq da Illumina, por permitir o sequenciamento de pequenos genomas com custo acessível e excelente qualidade (COIL et al., 2015). Primeiramente, é extraído o DNA total do microrganismo em estudo para a preparação de bibliotecas genômicas, ocorrendo a fragmentação do DNA e a ligação de adaptadores em ambas as extremidades, sendo os fragmentos fixados ao suporte por meio de seus terminais 5'. As extremidades 3' permanecem livres servindo de iniciadores nos ciclos subsequentes de amplificação por PCR. Posteriormente, o DNA fragmentado é injetado no sequenciador Illumina MiSeq e as sequências são lidas, corrigidas e "filtradas" de acordo com a qualidade. Os *contigs* gerados são verificados por meio de um mapeamento dos fragmentos (*reads*) e sobrepostos. Nesse momento, o genoma pode ser montado e analisado por softwares específicos (GLENN, 2011).

Entre as OGRIs, a identidade média de nucleotídeos (*Average Nucleotide Identity - ANI*) e a hibridação DNA-DNA digital (*Digital DNA-DNA hybridization - dDDH*) têm

sido metodologias amplamente utilizadas, havendo um fácil acesso a softwares relevantes com ferramentas prontamente disponíveis para o cálculo dos valores de similaridade (CHUN et al., 2018).

O valor de ANI consiste na comparação entre estirpes através do sequenciamento de seus genomas e alinhamento em programas específicos de bioinformática, de forma a avaliar a porcentagem de similaridade nucleotídica entre as mesmas. Os estudos realizados por Konstantinidis; Tiedje (2005) sugerem que o ANI é um método excepcionalmente robusto para medir a relação evolutiva entre estirpes bacterianas relacionadas. A primeira avaliação dos valores ANI realizada por Konstantinidis; Tiedje (2005) baseava-se em uma comparação pareada de genoma, incluindo todos os genes ortólogos codificadores de proteínas. Com base nesses estudos, Goris et al. (2007) recomendaram que os genomas comparados que apresentam mais de 95 % de ANI devem pertencer à mesma espécie, de acordo com o padrão de 70 % de similaridade da DDH. Entretanto, Richter; Rossello-Mora (2009) recomendaram o uso de um limite ainda mais estreito de 95-96 % de ANI para definir uma espécie bacteriana e mostraram que a comparação de pelo menos 20 % dos genomas gerados por sequenciamento aleatório simples seria suficiente para determinar valores confiáveis de ANI (KÄMPFER; GLAESER, 2012).

A grande vantagem da utilização do ANI se dá na possibilidade de armazenar as sequências genômicas em bancos de dados, reduzindo os custos empregados nas análises de comparação. O ANI tem sido a OGRM mais amplamente utilizada atualmente. Nesse sentido, diversas ferramentas de softwares para calcular o valor de ANI e suas versões melhoradas foram desenvolvidas para propósitos taxonômicos (RICHTER et al., 2016; RICHTER; ROSSELLO-MORA, 2009; RODRIGUEZ-R; KONSTANTINIDIS, 2016).

Além disso, a partir do advento da metodologia de sequenciamento gênico, tornou-se possível medir o valor de NI (*Nucleotide Identity*) entre as sequências de genes específicos que atuam como biomarcadores moleculares taxonômicos (genes *housekeeping*), ou genes concatenados. O valor de NI contribui para a realização de inferências taxonômicas, sem adicionar pesos evolutivos.

A metodologia de DDH digital configura uma alternativa ao ANI, sendo também amplamente utilizada para fins taxonômicos (CHUN et al., 2018; MEIER-KOLTHOFF et al., 2013). Os valores de dDDH representam uma comparação de sequências genômicas

pareadas *in silico* entre dois genomas bacterianos (AUCH et al., 2010). Os cálculos se baseiam em algoritmos que pontuam pares de segmentos, gerando uma base de dados que é usada para inferir distâncias intergenômicas, sendo calculados por uma calculadora de distância de genoma para genoma a partir do servidor GGDC (Genome-to-Genome Distance Calculator), seguindo as configurações recomendadas (AUCH et al., 2010; MEIER-KOLTHOFF et al., 2013). O limite de espécies proposto para os valores de dDDH é de 70 % (MEIER-KOLTHOFF et al., 2013).

4. MATERIAIS E MÉTODOS

Os experimentos de caracterização morfológica, fisiológica, genética e filogenética foram conduzidos no Laboratório de Biotecnologia de Solos da Empresa Brasileira de Pesquisa Agropecuária – Centro Nacional de Pesquisa de Soja (Embrapa Soja) em Londrina, Paraná. A análise de perfil de ácidos graxos para caracterização bioquímica foi realizada na Embrapa Meio Ambiente, em Jaguariúna, São Paulo.

4.1. ISOLAMENTO, SELEÇÃO E CULTIVO DAS ESTIRPES UTILIZADAS

Para esse estudo foram utilizadas 11 estirpes, selecionadas após uma avaliação prévia de 37 estirpes, isoladas de cinco locais diferentes nos estados de Missouri e Nebraska, na região centro-oeste dos Estados Unidos da América (EUA), conforme pode ser visualizado na Tabela 1. As estirpes foram isoladas da espécie *Chamaecrista fasciculata* (Michx.). Os isolamentos foram realizados na BIOZYME Inc., em Saint Joseph, Missouri (USA), pelo Dr. Padma Somasegaran e sua equipe. Nos locais de amostragem, as plantas foram cuidadosamente retiradas do solo para obter sistemas radiculares nodulados intactos, que foram separados da parte aérea e embalados em sacos plásticos rotulados. Em seguida, as embalagens com o conteúdo foram armazenadas no laboratório a 4 °C durante dois a três dias antes do isolamento das bactérias.

Primeiramente, o sistema radicular de cada amostra foi embebido em água e lavado, eliminando resquícios de solo. Cerca de 10 a 15 nódulos foram retirados de cada raiz e passaram por uma lavagem superficial com imersões de cinco a 10 minutos em uma solução de Tween 80 a 0,1 %. No passo seguinte, os nódulos foram colocados em uma placa plástica de Petri contendo papel de filtro para remover a umidade superficial. Em sequência, procedeu-se à desinfestação superficial, com imersão de três a cinco minutos em solução de H₂O₂ a 3 %, seguida por lavagem em água estéril. Após novamente remover a umidade da superfície dos nódulos, um nódulo foi selecionado e colocado em um recipiente com 0,5 mL de água destilada esterilizada. Posteriormente, cada nódulo foi gentilmente esmagado com uma haste de vidro estéril para obter uma suspensão com os bacteroides, que foi semeada em placas contendo meio *yeast mannitol agar* (YMA) modificado (manitol 5 g/L; K₂HPO₄ 0,5 g/L; MgSO₄.7H₂O 0,2 g/L; NaCl 0,1 g/L; extrato de levedura 0,4 g/L; pH 6,8; ágar 2 %) (HUNGRIA et al., 2016) com vermelho Congo a 0,25 %, e em placas com meio YMA modificado acrescido do indicador azul de bromotimol 0,5 %. Em seguida, as placas inoculadas foram colocadas em uma incubadora

a 26-28 °C, durante seis a oito dias. Nesse período, as placas foram examinadas quanto ao seu crescimento e pelas características que apresentavam em ambos os meios, e, em sequência, iniciou-se o processo de obtenção de culturas puras, como descrito por Somasegaran; Hoben (1985). As culturas purificadas das placas de YMA modificado acrescido de vermelho Congo foram semeadas em placas contendo apenas YMA, sendo mantidas sob refrigeração para armazenamento a curto prazo. Por fim, três réplicas de cada cultura, crescidas em seis a sete dias em tubos inclinados contendo YMA, foram transferidas para 1,0 mL de glicerol estéril a 50 % (v/v) em frascos criogênicos e armazenadas a -80 °C.

As estirpes estudadas foram depositadas na “Coleção de Culturas de Microrganismos Multifuncionais da Embrapa Soja: Bactérias Diazotróficas e Promotoras do Crescimento de Plantas” (WFCC Collection # 1213, WDCM Collection # 1054), recebendo as devidas identificações (CNPSO), conforme demonstrado na Tabela 1. Para iniciar os experimentos de caracterização, as estirpes foram crescidas em meio YMA modificado e incubadas por sete dias a 28°C até a formação de colônias visíveis (VINCENT, 1970).

TABELA 1. Estirpes isoladas de nódulos de *Chamaecrista fasciculata* utilizadas neste estudo, com a identificação dos locais de procedência dos nódulos.

Identificação da estirpe (CNPSO)	Local de procedência do nódulo
3421	Ao longo do rio St. Joseph, MO, EUA (89 Km ao Norte da cidade de Kansas, MO)
3411	Ao longo do rio St. Joseph, MO, EUA (89 Km ao Norte da cidade de Kansas, MO)
3413	Ao longo do rio St. Joseph, MO, EUA (89 Km ao Norte da cidade de Kansas, MO)
3417	Ao longo da Rodovia I-29, Bellavue, NE, EUA (284 Km ao Norte da cidade de Kansas, MO)
3423	Ao longo da Rodovia I-29, Bellavue, NE, EUA (284 Km ao Norte da cidade de Kansas, MO)
3424	Ao longo da Rodovia I-29, Oregon, MO, EUA (130 Km ao Norte da cidade de Kansas, MO)
3426 ^T	Ao longo da Rodovia I-29, Oregon, MO, EUA (130 Km ao Norte da cidade de Kansas, MO)
3431	Ao longo da 192 nd. Street, Omaha, NE, EUA (310 Km ao Norte da cidade de Kansas, MO)
3443	Ao longo da 192 nd. Street, Omaha, NE, EUA (310 Km ao Norte da cidade de Kansas, MO)

3446	Ao longo da 192 nd. Street, Omaha, NE, EUA (310 Km ao Norte da cidade de Kansas, MO)
3447	Ao longo da Rodovia I-29, Savannah, MO, EUA (106 Km ao Norte da cidade de Kansas, MO)

4.2. CARACTERIZAÇÃO MORFOLÓGICA E FISIOLÓGICA

A caracterização morfológica das 11 estipes (Tabela 1) foi realizada conforme Somasegaran; Hoben (1985) e Vincent (1970), acrescentando-se algumas adaptações. As análises incluíram a avaliação da taxa de crescimento e morfologia das colônias em relação à forma, elevação, bordas, superfície, produção de muco, consistência e detalhes ópticos; alterações no pH do meio; e cromogênese das colônias em meio YMA contendo azul de bromotimol e vermelho congo. Todos os testes foram realizados com três repetições.

4.2.1. Crescimento e Diâmetro das Colônias

As estirpes foram crescidas a 28 °C em placas contendo meio YMA modificado e o indicador de contaminação vermelho Congo pela técnica de semeadura por esgotamento para avaliação quanto à taxa de crescimento. Foram necessárias pelo menos três colônias isoladas para o registro da manifestação de crescimento, seguindo a classificação da taxa em rápida (até três dias), intermediária (quatro a cinco dias), lenta (seis a nove dias) ou muito lenta (superior a 10 dias). Em seguida, o diâmetro das colônias foi medido após a manifestação de crescimento das colônias isoladas, com um paquímetro, sendo realizados registros da média do diâmetro das três colônias, analisadas em milímetros (mm).

4.2.2. Forma, Elevação, Borda e Superfície das Colônias

Da mesma forma que na análise anterior, foram selecionadas três colônias isoladas para avaliar todos os parâmetros indicados. Essas observações foram realizadas em um estereomicroscópio, com luz incidente superior. A classificação do formato das colônias pode se dar como circular, oval (elíptica) ou irregular. A elevação foi analisada observando-se a placa lateralmente, sendo classificadas em colônias planas ou convexas. Quanto à borda, as colônias foram agrupadas em lisas, onduladas ou filamentosas, enquanto que a superfície foi classificada em lisa ou rugosa.

4.2.3. Produção de Muco e Consistência da Massa de Crescimento

A verificação da produção de muco foi realizada nas estirpes após a manifestação de crescimento em YMA modificado com vermelho Congo. As estirpes foram agrupadas em função de pouca produção (tamanho das colônias $\leq 1,9$ mm), produção moderada (tamanho das colônias entre 2 a 3,9 mm) e produção abundante de muco (tamanho das colônias ≥ 4 mm). Em relação à análise da consistência da massa de crescimento, a metodologia baseou-se na utilização de uma alça de platina, previamente flambada e resfriada, que entrou em contato com uma colônia isolada para observar o seu aspecto. A observação foi realizada sob estereomicroscópio, com luz incidente superior e as colônias foram enquadradas nas características de aquosas (desmanchando-se totalmente quando tocadas pela alça), secas (apresentando pouca quantidade de muco), gomosas (com um aspecto cremoso, deixando rastros), e viscosas (com uma consistência elástica).

4.2.4. Detalhes Ópticos

Nesse teste, utilizou-se o fundo branco de um estereomicroscópio e uma caneta preta (retroprojeter) para marcar um risco no fundo da placa contendo YMA modificado e vermelho congo, a fim de facilitar a visualização dos detalhes ópticos. A massa de crescimento foi classificada em opaca ou translúcida, dependendo de como o risco foi visualizado na luz, estando totalmente coberto pela massa celular ou sobrepondo a mesma.

4.2.5. Alterações no pH do Meio

Para realização desse teste, as estirpes foram inoculadas em placas contendo meio YMA modificado, acrescido de 2 mL/L de uma solução de azul de bromotimol a 0,5 %, um indicador ácido-base orgânico sintético. Na presença do indicador, quando o meio se torna ácido a coloração é alterada para amarelo, enquanto que em reação básica ele assume a cor azul; e meios neutros permanecem com a coloração original esverdeada obtida na preparação do meio. Dessa forma, as estirpes foram classificadas como apresentando reações de acidificação, alcalinização ou neutralidade no meio de cultura.

4.2.6. Cromogênese das Colônias em Meio YMA com Azul de Bromotimol e Vermelho Congo

As culturas foram crescidas em YMA modificado com dois indicadores distintos, azul de bromotimol e vermelho Congo. A classificação da coloração da massa em crescimento

na presença de azul de bromotimol baseou-se nas colorações creme, amarelo, rosa e azul; já no meio acrescido de vermelho congo as colônias foram classificadas em brancas, rosadas, avermelhadas (centro) ou vermelhas.

4.3. CARACTERIZAÇÃO FISIOLÓGICA

4.3.1. Capacidade de Utilização de Fontes de Carbono

Para caracterizar a utilização de fontes de carbono da nova espécie descrita, foi realizado um experimento em duplicata com as estirpes CNPSO 3426^T, estirpe de referência ou tipo, e CNPSO 3424, assim como para as estirpes *B. liaoningense* LMG 18230^T e *Bradyrhizobium daqingense* CCBAU 15774^T, as espécies filogeneticamente mais relacionadas. Os experimentos foram realizados utilizando-se o kit API 50CH (BioMérieux), o qual é amplamente empregado para identificação taxonômica microbiana (BROLAZO et al., 2011). Nesse kit foram testadas 49 fontes de carbono, as quais podem ser utilizadas dependendo da capacidade de metabolização das estirpes. Inicialmente, as estirpes foram crescidas em duplicata em 10 mL de meio YM e incubadas sob agitação a 100 rpm, durante sete dias. Após o período de incubação e manifestação de crescimento, as culturas foram centrifugadas a 6.000 rpm por 10 minutos e o sobrenadante descartado. No passo seguinte, o pélete foi ressuspensionado em 5 mL de solução salina a 0,85 %, e centrifugado sob as mesmas condições anteriores e, novamente, o sobrenadante foi descartado. Em seguida, foram adicionados 2 mL de água destilada autoclavada para ressuspender o pélete. A solução foi transferida com uma pipeta Pasteur para um tubo de ensaio estéril contendo 5 mL de água destilada, onde a concentração de células foi ajustada para 10⁹ células/mL, utilizando-se a escala de McFarland como padrão. Uma vez que o ajuste foi realizado, o dobro do volume de inóculo necessário para atingir a escala foi transferido com a pipeta Pasteur para um frasco de vidro com 10 mL de meio YM modificado (sem manitol, com extrato de levedura a 0,3 g/L e acrescido de azul de bromotimol). Utilizando-se o kit API 50CH, foram colocados 230 µL do inóculo preparado em cada uma das 49 fontes de C, que, em seguida, foram recobertas com 3-5 gotas de vaselina autoclavada. As placas do API foram incubadas a 28°C, e os resultados observados em cinco e sete dias após a inoculação. Os resultados foram interpretados de acordo com a intensidade de alteração da coloração, sendo positivo, e confirmando que a estirpe metaboliza a fonte de C, quando a alteração de cor é intensa; negativo, e indicando

a não metabolização da fonte de C, quando não há alteração de cor; ou fraco, com a estirpe metabolizando muito pouco desta fonte de C, quando há leve alteração na coloração.

4.3.2. Tolerância a Antibióticos

As estirpes foram testadas quanto à resistência a 10 antibióticos: bacitracina (0,04 U/disco), ampicilina (10 µg/mL), neomicina (30 µg/mL), estreptomicina (10 µg/mL), tetraciclina (30 µg/mL), eritromicina (15 µg/mL), cefuroxima (30 µg/mL), cloranfenicol (30 µg/mL), penicilina (10 U/disco) e ácido nalidíxico (30 µg/mL). O método utilizado para verificação de tolerância foi o disco-difusão, o qual avalia a resposta da bactéria ao antimicrobiano pela medida direta do diâmetro do halo de inibição que se forma ao redor do disco contendo antibiótico (BAUER et al., 1966). Primeiramente, foi feito um pré-inóculo, no qual as bactérias foram crescidas em meio YM durante sete dias, sob agitação constante de 100 rpm. Em seguida, 100 µL do pré-inóculo foram transferidos para uma placa contendo YMA e espalhados com uma alça de Drigalski por toda superfície do meio, e, por fim, dois discos de um mesmo antibiótico foram adicionados. Cada disco foi depositado a uma distância segura do outro, permitindo uma medição correta do halo após o período de incubação. Após sete dias de crescimento, o diâmetro dos halos de inibição presentes foi medido com um paquímetro. Posteriormente, as bactérias tiveram suas resistências classificadas em sensível, moderadamente resistente ou resistente, de acordo com o padrão determinado para cada antibiótico. O diâmetro do halo de inibição final foi calculado em função média dos diâmetros dos halos formados nos dois discos da placa.

4.3.3. Crescimento em Meio Luria-Bertani (LB)

A caracterização fisiológica inclui a realização de testes de crescimento das bactérias em outros meios de cultura, ou meios inespecíficos. Nesse estudo, foi realizado um experimento de crescimento das estirpes em meio LB (triptona 10g/L, NaCl 5g/L, extrato de levedura 5g/L, ágar 15g/L, pH 7,5).

4.3.4. Crescimento em NaCl 1 %

A capacidade de crescimento das estirpes em condição de alta salinidade também foi testada. Nesse experimento as bactérias foram semeadas em meio YMA modificado com 1 % de NaCl. Após um período de incubação de sete dias, as placas foram observadas e os resultados registrados quanto à presença (+) ou ausência (-) de crescimento.

4.3.5. Crescimento a 37 °C

Além de observar o crescimento das estirpes em suas condições ideais de temperatura (28 °C), também foram realizados testes para avaliar a performance de crescimento das mesmas em placas de YMA na temperatura de 37 °C. Os resultados foram analisados quanto à presença (+) ou ausência (-) de crescimento.

4.3.6. Crescimento em pH 4 e 9

Outra propriedade a ser avaliada na caracterização fisiológica é o crescimento das bactérias em condições de pH diferentes. A capacidade de crescimento bacteriano foi avaliada em condições de pH 4 e pH 9. As estirpes foram inoculadas em meio YMA com pH ajustados, seguido de um período de incubação a 28 °C por sete dias, quando os resultados foram interpretados em relação à presença (+) ou ausência (-) de crescimento.

4.3.7. Atividade da Enzima Urease

A atividade da enzima urease foi outro aspecto fisiológico testado entre as estirpes. O experimento baseou-se em inocular as bactérias em placas de meio YM contendo 2 % de ureia e o corante vermelho de fenol (0,012 g/L). A coloração inicial do meio é amarelada e, se após o período de incubação a cor do meio permanecer inalterada, interpreta-se que a estirpe não possui a enzima urease e, portanto, não degrada ureia. Se a cor do meio for alterada para rosa, constata-se a presença da urease e, conseqüente degradação da ureia. A fraca alteração de coloração indica baixa atividade da urease. As bactérias foram classificadas em urease positiva (+), negativa (-) ou baixa (+/-).

4.4. CARACTERIZAÇÃO BIOQUÍMICA

4.4.1. Perfil de Ácidos Graxos

A análise do perfil de ácidos graxos foi realizada em parceria com a Embrapa Meio-Ambiente. Esse experimento de caracterização foi realizado apenas com a estirpe de referência CNPSo 3426^T e com uma das estirpes mais próxima filogeneticamente, *B. liaoningense* LMG 18230^T, as quais foram crescidas em uma placa contendo YMA modificado por sete dias a 28 °C. Posteriormente, foram enviadas para a realização do perfil de ácidos graxos, conforme descrito por Delamuta et al. (2013). Os perfis foram determinados usando o MIDI, Microbial Identification System (SASSER, 1990).

4.5. CARACTERIZAÇÃO GENÉTICA E FILOGENÉTICA

4.5.1. Extração de DNA

Inicialmente, as estirpes foram semeadas em 10 mL de YM sob agitação de 100 rpm, a 28 °C, durante sete dias. A extração de DNA de 1 mL do inóculo das 11 estirpes foi realizada utilizando o kit DNeasy Blood & Tissue (Quiagen), conforme as instruções do fabricante. Posteriormente, as amostras foram submetidas à eletroforese em gel de agarose 1 %, a fim de realizar uma verificação prévia da qualidade de concentração do DNA extraído. Os estoques foram armazenados a -20 °C.

4.5.2. Análise de BOX-PCR (*Polimerase Chain Reaction*)

Foi realizada uma análise do perfil genético das estirpes pela metodologia de BOX-PCR, que compara os perfis de banda de cada estirpe em um gel de agarose, a partir da realização de uma eletroforese nas amostras de DNA total. A reação de amplificação apresenta volume final de 25 µL, contendo 12,8 µL de água Milli-Q estéril; 5,0 µL de dNTPs, (estoque com 1,5 mmol/L de cada base nitrogenada); 2,5 µL de tampão 10X (500 mM KCl; 200 mM Tris-HCl, pH 8,4); 1,5 µL de MgCl₂, (50 mmol/L); 2,0 µL de DNA (50 ng/µL); 0,2 µL de Taq (5 U/µL); e 1,0 µL do BOX-AIR (5'-CTACGGCAAGGCGACGCTGACG-3') (50 pmol/µL). As reações ocorreram em um termociclador (ProFlex PCR System (*Applied Biosystems*)) com uma etapa inicial de desnaturação a 95 °C por 7 min; 30 ciclos de 1 min a 94 °C (desnaturação), 1 min a 53 °C (hibridação) e 8 min a 65 °C (extensão); etapa final de extensão a 65 °C por 16 min; manutenção a 4 °C. O perfil de bandas formado foi visualizado em gel de agarose 1,5 % (20,5 X 25,5 cm), com tempo de corrida de cinco horas, sendo corado com brometo de etídio a 0,00005 %. A partir do programa Bionumerics (*Applied Mathematics, Kortrijk, Bélgica, v.7.6*) foi construído um dendrograma com os perfis de *fingerprinting* ou bandas obtidos. Inicialmente, os tamanhos dos fragmentos em todos os produtos das reações foram normalizados de acordo com o tamanho dos fragmentos do marcador de DNA (1 Kb plus, Invitrogen®). As análises de agrupamento dos perfis de *fingerprinting* foram realizadas utilizando o algoritmo UPGMA (Unweighted Pair-Group Method with Arithmetic mean) (SNEATH; SOKAL, 1973) e o coeficiente de Jaccard (JACCARD, 1912), com 1 % de tolerância.

4.5.3. Amplificação por PCR

Para cada estirpe foi realizada uma reação de amplificação por PCR do gene 16S RNAr, dos genes *housekeeping*: *glnII*, *gyrB*, *recA* e *rpoB*, da região ITS e dos genes de nodulação *nodC* e fixação *nifH*. Para cada reação houve a preparação de um mix contendo tampão 10X (Tris- base a 200 mM, pH 8,4, KCl a 500 mM), dNTPs (0,3 mM de cada base), MgCl₂ (1,5 mM), enzima Taq DNA polimerase (1U), iniciadores referentes a cada gene a ser amplificado (10 pmol/reação) (Tabela 2), DNA (de 10 a 20 ng/μL) e água ultrapura esterilizada até completar o volume final, específico para cada gene. Os *primers* e as condições de amplificação de cada gene são apresentados na Tabela 2. Após a amplificação, foi realizada uma eletroforese para verificar a qualidade dos produtos amplificados nas reações. Em seguida, os produtos de PCR passaram por uma purificação, utilizando-se o kit PureLink (Invitrogen®), de acordo com as instruções do fabricante. Após a purificação, as amostras foram novamente submetidas à eletroforese em gel de agarose a 1 % para verificação da concentração, tal como feito após extração do DNA. As amostras que apresentaram bandas fracas, em função de uma perda de parte do material durante a etapa de purificação, foram submetidas a um speedvac para concentrar as amostras (Concentrador Plus, Eppendorf).

TABELA 2. Sequência de *primers* e as condições de amplificação do gene 16S RNAr, dos genes *housekeeping* *glnII*, *gyrB*, *recA* e *rpoB*, da região intergênica (ITS), dos genes de nodulação *nodC* e fixação *nifH*.

Primer	Sequência	Gene alvo	Condições do ciclo	Referência
fD1 rD1	AGAGTTTGATCC TGGCTCAG CTTAAGGAGGTG ATCCAGCC	16S RNAr	2 min 95°C, 30 X (15s 94°C, 45s 93°C, 45s 55°C), 2 min 72°C e 5 min 72°C)	WEISBURG et al. (1991)
TSglnII _f TSglnII _r	AAGCTCGAGTAC ATCTGGCTCGAC GG SGAGCCGTTCCA GTCGGTGTCG	<i>glnII</i>	2 min 95°C, 35 X (45s 95°C, 30s 58°C), 1.5 min 72°C e 7 min 72°C	STEPKOWSKI et al. (2005)
gyrB343F gyrB1043R	TTCGACCAGAAAY TCCTAYAAGG AGCTTGTCCTTS GTCTGCG	<i>gyrB</i>	5 min 95°C, 5X (2 min 94°C, 2 min 58°C, 1 min 72°C) 28 X (30 s 94°C, 1 min 58°C, 1 min 72°C e 5 min 72°C)	MARTENS et al. (2008)
TSrecA _f TSrecA _r	CAACTGCMYTG CGTATCGTCGAAGG CGGATCTGGTTG ATGAAGATCACCATG	<i>recA</i>	2 min 95°C, 35 X (45s 95°C, 30s 58°C), 1.5 min 72°C e 7 min 72°C	STEPKOWSKI et al. (2005)

rpoB83F rpoB1061R	CCTSATCGAGGT TCACAGAAGGC AGCGTGTGCGG ATATAGGCG	<i>rpoB</i>	5 min 95°C, 3X (2 min 94°C, 2 min 58°C, 1 min 72°C) 30 X (30s 94°C, 1 min 58°C), 1 min 72°C e 5 min 72°C	MARTENS et al. (2008)
FGPS1490 FGPS130	TGCGGCTGGATCACCTCCTT CCGGGTTTCCCCATTCGG	ITS	3 min 94°C, 35 X (1 min 94°C, 1 min 55°C, 2 min 72°C e 6 min 72°C)	LAGUERRE et al. (1996)
nodCF nodCI	AYGTHGTYGAYGACGGITC CGYGACAGCCANTCKCTATTG	<i>nodC</i>	30s 94°C, 40x (30s 96°C, 1 min 53°C, 30s 72°C), 5min 72°C	LAGUERRE et al. (2001)
nifHF nifHR	TACGGNAARGGSGGNATCGGCAA AGCATGTCTCSAGYTCNTCCA	<i>nifH</i>	2 min 95°C, 35x(1 min 94°C, 45s 60°C, 2 min 72°C), 7 min 72°C	LAGUERRE et al. (2001)

4.5.4. Sequenciamento Gênico e Genômico

Os genes 16S RNAr, *glnII*, *gyrB*, *recA*, *rpoB*, *nodC* e *nifH* e a região ITS foram sequenciados para análises genômicas e filogenéticas. Durante o sequenciamento, os produtos de PCR obtidos após purificação foram submetidos a uma nova amplificação, a qual consiste na adição de 3 µL de DNA do produto amplificado (50 ng/µL), 1,3 µL de BigDye (compatível com o sequenciador ABI3500xL sequence analyser (Applied Biosystems), 1 µL de cada *primer* (3,2 pmol/µL) conforme demonstrado em Delamuta et al. (2017), 2,5 µL de tampão 10x (500 mM KCl; 200 mM Tris-HCl, pH 8,4) e 12,2 µL de H₂O milli-Q estéril, totalizando um volume final de 20 µL por reação, que foram distribuídos em placas de 96 poços. As condições da PCR foram as mesmas para todos os genes, com um ciclo inicial de 96 °C por 1 min, seguido de 35 ciclos de 96 °C por 15 s, 50 °C por 15 s e 60 °C por 4 min, em um termociclador (Veriti 96-Well Thermal Cycler (Applied Biosystems)).

Após amplificação, as amostras foram purificadas. Em todos os poços da placa, foram adicionados 2 µL de acetato de amônio estéril (7,5 M) e 65 µL de etanol a 99,5 % em temperatura ambiente. Em seguida, a placa foi selada e homogeneizada, sendo submetida a uma centrifugação de 4000 rpm por 45 min. O sobrenadante foi descartado e, posteriormente, foram adicionados 150 µL de etanol 70 % em cada poço da placa, repetindo-se o passo anterior de centrifugação por 20 min. Após o descarte do

sobrenadante, uma nova centrifugação foi realizada em baixa rotação (300 rpm) durante 25 s. Por fim, a placa foi colocada em uma estufa a 37 °C por cerca de 30 min para secar.

Após secagem completa, os péletes foram ressuspensos em 10 µL de formamida, aquecidos à 95 °C durante 5 min e resfriados por mais 2 min em gelo, sendo submetidos ao sequenciador ABI 3500XL (*Applied Biosystems*). Uma vez que os resultados foram obtidos pelo sequenciador, as sequências foram analisadas com o programa Bionumerics versão 7.6 e corrigidas manualmente com base em sequências depositadas no banco de dados do NCBI (*National Center for Biotechnology Information*).

O sequenciamento genômico do DNA total das estirpes CNPSo 3424 e CNPSo 3426^T foi realizado no sequenciador Illumina MiSeq utilizando-se o kit Nextera XT (*Nextera DNA Library Preparation Kits*) para contruir uma biblioteca, de acordo com as instruções do fabricante. Os esboços de ambos os genomas foram montados utilizando o A5-miseq pipeline (*de novo*) v.20140604. Utilizando-se os programas QUAST v.2.0 (GUREVICH et al., 2013) e RAST v.2.0 (RODRIGUEZ-R; KONSTANTINIDIS, 2016), diversos parâmetros do genoma como cobertura, número de contigs, valor de N50, tamanho e conteúdo G+C foram avaliados.

4.5.5. Análise Filogenética do Gene 16S RNAr

Após a etapa de correção das sequências do gene 16S RNAr, elas foram alinhadas e cortadas por meio do software MEGA 6 (*Molecular Evolutionary Genetics Analysis*) (TAMURA et al., 2013), utilizando-se o algoritmo MUSCLE (EDGAR, 2004). A seleção do modelo evolutivo de melhor ajuste foi realizada pelo software JModelTest 2 (DARRIBA et al., 2012). As análises de máxima verossimilhança (*Maximum Likelihood*, ML) (FELSENSTEIN, 1981) para construção da árvore filogenética foram realizadas utilizando o software GARLI 2.0 (ZWICKL, 2006). Para testar o suporte estatístico na análise de ML, foram realizadas 1.000 réplicas de *bootstrap* (FELSENSTEIN, 1985) no software GARLI 2.0. A estirpe *Xanthobacter autotrophicus* Py2 foi utilizada como *outgroup* (DELAMUTA et al., 2013).

4.5.6. Análise Filogenética da Região Intergênica 16S-23S DNAr (ITS)

Em sequência à etapa de correção, as sequências obtidas no sequenciamento foram alinhadas e editadas utilizando-se o software MEGA 6 (TAMURA et al., 2013) a partir do algoritmo MUSCLE (EDGAR, 2004). A construção da árvore filogenética a partir dos dados de distância evolutiva foi realizada com base em análises de *Neighbor-Joining* (NJ)

(SAITOU; NEI, 1987), geradas pelo software MEGA 6. Para testar o suporte estatístico dos nós na análise de NJ, 1.000 réplicas de *bootstrap* foram realizadas.

4.5.7. MLSA e Filogenia dos Genes de Nodulação (*nodC*) e de Fixação De N₂ (*nifH*)

Os genes *housekeeping glnII*, *gyrB*, *recA* e *rpoB* analisados foram selecionados com base em estudos anteriores (DELAMUTA et al. 2012, 2013, HELENE et al. 2015, 2017). Dessa forma, após a obtenção das sequências e da etapa de correção manual, elas foram alinhadas e editadas conforme realizado para o gene 16S RNAr. Posteriormente, as quatro sequências foram concatenadas, utilizando-se o software *SeaView* (GOUY et al., 2010). As análises de máxima verossimilhança para construção da árvore filogenética foram realizadas utilizando o software GARLI 2.0 (ZWICKL, 2006), enquanto que a seleção do modelo evolutivo foi realizada pelo software JModelTest 2 (DARRIBA et al., 2012). Foram realizadas 1.000 réplicas de *bootstrap* (FELSENSTEIN, 1985) pelo software GARLI 2.0. A estirpe *Xanthobacter autotrophicus* Py2 foi utilizada como *outgroup* (DELAMUTA et al., 2013).

As análises filogenéticas dos genes *nodC* e *nifH* foram realizadas tal qual aquelas feitas para o gene 16S RNAr e para o MLSA, sendo construídas uma árvore filogenética para cada gene.

4.5.8. Nucleotide Identity (NI)

Após a realização do sequenciamento e edição das sequências no Bionumerics versão 7.6 dos genes 16S RNAr, ITS, *glnII*, *gyrB*, *recA* e *rpoB*, foi calculado o valor percentual de identidade nucleotídica (% NI) para cada gene e para os genes *housekeeping* concatenados. Os percentuais foram calculados comparando as sequências entre as 11 estirpes estudadas e em comparação com as estirpes tipo do gênero *Bradyrhizobium* pertencentes ao superclado de *B. japonicum*. Essas análises foram realizadas pelo programa BioEdit (HALL, 1999).

4.5.9. Average Nucleotide Identity (ANI)

Os valores de ANI foram estimados para os genomas das estirpes CNPSo 3426^T e CNPSo 3424 e cinco espécies de *Bradyrhizobium* estreitamente relacionadas com genomas disponíveis: *B. daqingense* CCBAU 15774^T (JGI Project Id: 1041378 - JGI), *B. liaoningense* LMG 18230^T (Project ID: 1052895 - JGI), *B. ottawaense* OO99^T (CP029425.1), *B. shewense* ERR11^T (NZ_FMAI01000001.1), *B. stylosanthis* BR 446^T

(LVEM00000000.1), *B. neotropica* BR 10247^T (NZ_LFJC01000003.1), *B. guangxiense* CCBAU 53363^T (NZ_CP022220.1) e *B. huanghuaihaiense* CCBAU 23303^T (SAMN02927902). Dessa forma, foi utilizado um programa específico desenvolvido por Rodriguez-R; Konstantinidis (2014) capaz de calcular a % média de similaridade nas sequências de DNA dos genomas totais entre duas estirpes. A delimitação de espécies considerada para os valores de ANI foi de 96 % (GORIS et al., 2007 e KONSTANTINIDIS et al., 2006).

4.5.10. Hibridação DNA-DNA Digital (dDDH)

Todos os valores obtidos por meio de pareamentos foram determinados pelo programa GGDC v2.1, usando a "Fórmula 2" recomendada (<http://ggdc.dsmz.de/distcalc2.php>) (RODRIGUEZ-R; KONSTANTINIDIS, 2016). Os valores de dDDH foram obtidos para uma comparação entre o genoma das estirpes CNPSo 3426^T e CNPSo 3424; e em relação ao genoma das espécies mais estreitamente relacionadas: *B. daqingense* CCBAU 15774^T (JGI Project Id: 1041378 - JGI), *B. liaoningense* LMG 18230^T (Project ID: 1052895 - JGI), *B. ottawaense* OO99^T (CP029425.1), *B. shewense* ERR11^T (NZ_FMAI01000001.1), *B. stylosanthis* BR 446^T (LVEM00000000.1), *B. neotropica* BR 10247^T (NZ_LFJC01000003.1), *B. guangxiense* CCBAU 53363^T (NZ_CP022220.1) e *B. huanghuaihaiense* CCBAU 23303^T (SAMN02927902).

4.5.11. Conteúdo G+C

A partir dos resultados obtidos com o sequenciamento do genoma das estirpes CNPSo 3426^T e CNPSo 3424 foi estimado, em percentual, o conteúdo dos nucleotídeos G+C presente no DNA (% molar). Inicialmente, os *contigs* gerados como resultado do sequenciamento foram montados pelo programa A5-miseq pipeline (COIL et al., 2015) e, posteriormente, as análises dos genomas foram realizadas nos programas QUAST v.2.0 (GUREVICH et al., 2013) e RAST v.2.0 (RODRIGUEZ-R; KONSTANTINIDIS, 2016).

4.6. CAPACIDADE DE NODULAÇÃO E ATIVIDADE DOS NÓDULOS

Para avaliar a capacidade de nodulação das estirpes em estudo foram realizados dois experimentos, um em casa de vegetação e o outro em sacos de polipropileno, utilizando-se quatro estirpes pré-selecionadas: CNPSo 3411, CNPSo 3424, CNPSo 3426^T e CNPSo 3443. Ambos os experimentos testaram a capacidade de nodulação em associação com a

planta hospedeira, *Chamaecrista fasciculata*, e outras duas leguminosas, soja (*Glycine max*) e siratro (*Macroptilium antropurpureum*). Todos os tratamentos apresentaram três réplicas biológicas.

Inicialmente, foram crescidos os inóculos de cada estirpe em meio YM modificado. O crescimento dos inóculos foi acompanhado por leituras em espectrofotometro, até que a concentração de 10^9 células/mL foi atingida. O tratamento relativo ao controle negativo não recebeu inóculo, enquanto que no tratamento do controle positivo as plantas receberam inóculos da estirpe *B. diazoefficiens* SEMIA 5080.

A germinação das sementes de *C. fasciculata* e siratro dependeu da execução de protocolos de quebras de dormência mecânico e químico, respectivamente. A quebra de dormência das sementes de *C. fasciculata* se baseou na realização de escarificações mecânicas utilizando lixas, ao passo que as sementes de siratro foram colocadas em uma solução de ácido sulfúrico concentrado por três minutos, sendo, em seguida, realizadas cinco lavagens com água destilada. Posteriormente, todas as sementes passaram por um protocolo de esterilização superficial a partir de uma lavagem com etanol 70 % por 1 min, seguida de lavagem com água destilada esterilizada; uma lavagem com uma solução de NaClO 10 % por cinco minutos, seguida por cinco lavagens com água destilada esterilizada.

O experimento em casa de vegetação foi realizado em vasos de Leonard (VINCENT, 1970) contendo um substrato inerte e autoclavado (areia e carvão 1:1 v/v). Em cada vaso foram adicionadas três sementes, as quais receberam 1 mL de inóculo cada. Cerca de seis dias após a emergência, duas plantas foram desbastadas, restando apenas uma planta por vaso. Ao longo do experimento as plantas receberam solução nutritiva sem N (0,00036 % Fe-EDTA, 0,0061 % $MgSO_4 \cdot 7H_2O$, 0,0043 % K_2SO_4 , 0,00016 % $MnSO_4 \cdot H_2O$, 0,014 % $CaCl_2 \cdot 2H_2O$, 0,0068 % KH_2PO_4). A coleta e observação da capacidade de nodulação foram realizadas 30 dias após a inoculação.

Em relação ao experimento em sacos de polipropileno, os inóculos foram preparados seguindo o mesmo protocolo descrito para o experimento em casa de vegetação. As sementes de *C. fasciculata* e siratro passaram pelo mesmo protocolo anteriormente relatado de quebra de dormência. Todas as sementes foram esterilizadas seguindo o protocolo anteriormente descrito. As sementes de soja foram pré-germinadas em papel germiteste por três dias, enquanto que as sementes das demais planta foram pré-

germinadas em meio Ágar água 1 % (ágar 10 g/L). Posteriormente, as plântulas foram colocadas em sacos de polipropileno contendo solução nutritiva sem N, tal qual a descrita para o experimento anterior, quando foram inoculadas com 1 mL de inóculo cada. A coleta e observação da capacidade da nodulação foram realizadas 30 dias após a inoculação.

Além disso, foi realizada uma avaliação quanto à atividade dos nódulos em ambos os experimentos. A partir de cada raiz nodulada, foram retirados três nódulos, que em seguida foram cortados ao meio com auxílio de uma lâmina. A atividade dos nódulos foi interpretada em relação à coloração interna dessas estruturas, manifestando atividade positiva ao apresentar uma coloração rosada, e negativa com uma coloração branca.

4.6.1. Reisolamento das Estirpes

Para confirmar a presença das estirpes CNPSo em estudo nos nódulos verificados em soja e siratro, foram coletados três nódulos de cada raiz a fim de reisolar as bactérias ali presentes. Dos nódulos coletados, foi selecionado um nódulo por raiz.

Utilizando uma placa de 96 poços para realizar a desinfestação dos nódulos, os poços foram preenchidos com soluções de Tween 80 (0,1 %), H₂O₂ (3 %) e água estéril, conforme o número de lavagens indicadas. Primeiramente, os nódulos foram lavados em solução de Tween 80 por 5-10 min. Em seguida, os nódulos foram colocados por alguns instantes sob uma folha de papel filtro a fim de retirar o excesso de solução. Posteriormente, os nódulos foram lavados em uma solução de H₂O₂ durante 3-5 min, seguido por 10 lavagens em água estéril. Utilizando-se um microtubo, os nódulos foram macerados em 500 µL de água estéril com auxílio de uma ponteira. Em seguida, 100 µL do inóculo gerado foram transferidos para uma placa contendo YMA e espalhados com uma alça de Drigalski por toda superfície do meio. Entretanto, com o objetivo de gerar placas com fácil reisolamento de colônias, 200 µL do inóculo inicial foram diluídos em 4 mL de água estéril, sendo semeados 100 µL em uma placa contendo YMA. As placas foram incubadas a 28 °C por sete dias.

Após o período de incubação, as colônias características de *Bradyrhizobium* foram reisoladas e semeadas em novas placas contendo YMA. As placas foram incubadas a 28 °C por sete dias, confirmando-se a ausência de contaminantes. Em sequência, os isolados foram crescidos em meio YM, passando por uma etapa de extração de DNA, chegando

até a realização de uma reação de PCR para amplificar o gene *gyrB*, tendo sido realizado o posterior sequenciamento dessas sequências.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- AHNIA, H.; BOUREBABA, Y.; DURÁN, D.; et al. *Bradyrhizobium algeriense* sp. nov., a novel species isolated from effective nodules of *Retama sphaerocarpa* from Northeastern Algeria. **Systematic and Applied Microbiology**, v. 41, n. 4, p. 333–339, 2018.
- ALLEN, O. N.; ALLEN, E. **The Leguminosae: a source book of characteristics, uses and nodulation**. Wisconsin-USA: University of Wisconsin Press, 1981. p. 812.
- ANDREWS, M.; LEA, P. J.; RAVEN, J. A.; AZEVEDO, R. A. Nitrogen use efficiency. 3. Nitrogen fixation: Genes and costs. **Annals of Applied Biology**, v. 155, p. 1–13, 2009.
- ARAUJO, R. S.; HUNGRIA, M. **Microrganismos de importância agrícola**. Brasília; EMBRAPA-SPI, 1994. p. 236.
- AUCH, A. F.; JAN, M. VON; KLENK, H.; GÖKER, M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. **Standards in Genomic Sciences**, v. 2, p. 117–134, 2010.
- BASI, S.; NEUMANN, M.; MARAFON, F.; UENO, R. K.; SANDINI, I. E. Influência da adubação nitrogenada sobre a qualidade da silagem de milho. **Revista Brasileira de Tecnologia Aplicada nas Ciências Agrárias**, v. 4, n. 3, p. 219–234, 2011.
- BAUER, A.; KIRBY, M.; SHERRIS, J.; TURCK, M. Antibiotic susceptibility testing by a standardized single disk method. **The American Journal of Clinical Pathology**, v. 45, n. 4, p. 493–496, 1966.
- BHATTACHARYYA, P. N.; JHA, D. K. Plant growth-promoting rhizobacteria (PGPR): Emergence in agriculture. **World Journal of Microbiology and Biotechnology**, v. 28, n. 4, p. 1327–1350, 2012.
- BHUVANESWARI, T. V.; TURGEON, B. G.; BAUER, W. D. Early events in the infection of soybean (*Glycine max* L. Merr) by *Rhizobium japonicum*. **Plant Physiology**, v. 66, n. 709, p. 1027–1031, 1980.
- BODDEY, R. M.; SÁ, J. C. M.; ALVES, B. J. R.; URQUIAGA, S. The contribution of biological nitrogen fixation for sustainable agricultural systems in the tropics. **Soil Biology and Biochemistry**, v. 29, n. 516, p. 787–799, 1997.
- BOHLOOL, B. B.; LADHA, J. K.; GARRITY, D. P.; GEORGE, T. Biological nitrogen fixation for sustainable agriculture: a perspective. **Plant and Soil**, v. 141, p. 1–11, 1992.
- BREWIN, N. J. Plant cell wall remodelling in the rhizobium-legume symbiosis. **Critical Reviews in Plant Sciences**, v. 23, n. 4, p. 293–316, 2004.
- BROLAZO, E. M.; LEITE, D. S.; TIBA, M. R.; et al. Correlation between API 50 ch and multiplex polymerase chain reaction for the identification of vaginal lactobacilli in isolates. **Brazilian Journal of Microbiology**, v. 42, p. 225–232, 2011.
- BROUGHTON, W. J.; ZHANG, F.; PERRET, X.; STAEHELIN, C. Signals exchanged between legumes and *Rhizobium*: agricultural uses and perspectives. **Plant and Soil**, v. 252, n. 1, p. 129–137, 2003.
- BÜNGER, W.; GRÖNEMEYER, J.; SARKAR, A.; REINHOLD-HUREK, B. *Bradyrhizobium ripae* sp. nov., a nitrogen-fixing symbiont isolated from nodules of wild legumes in Namibia. **International Journal of Systematic and Evolutionary Microbiology**, v. 68, p. 3688–3695, 2018.
- CAMERON, K. C.; DI, H. J.; MOIR, J. L. Nitrogen losses from the soil/plant system: a review.

Annals of Applied Biology, v. 162, p. 145–173, 2013.

CASSAB, G. I. Plant cell wall proteins. **Annual Review of Plant Physiology and Plant Molecular Biology**, v. 49, n. 1, p. 281–309, 1998.

CHAHBOUNE, R.; CARRO, L.; PEIX, A.; et al. *Bradyrhizobium cytisi* sp. nov., isolated from effective nodules of *Cytisus villosus*. **International Journal of Systematic and Evolutionary Microbiology**, v. 61, n. 12, p. 2922–2927, 2011.

CHAHBOUNE, R.; CARRO, L.; PEIX, A.; et al. *Bradyrhizobium rifense* sp. nov. isolated from effective nodules of *Cytisus villosus* grown in the Moroccan Rif. **Systematic and Applied Microbiology**, v. 35, n. 5, p. 302–305, 2012.

CHANG, Y. L.; YU WANG, J.; TAO WANG, E.; et al. *Bradyrhizobium lablabi* sp. nov., isolated from effective nodules of *Lablab purpureus* and *Arachis hypogaea*. **International Journal of Systematic and Evolutionary Microbiology**, v. 61, n. 10, p. 2496–2502, 2011.

CHEN, W.; MOULIN, L.; BONTEMPS, C.; VANDAMME, P.; et al. Legume symbiotic nitrogen fixation by beta-Proteobacteria is widespread in nature. **Journal of Bacteriology**, v. 185, n. 24, p. 7266–7272, 2003.

CHEON, C. I.; LEE, N. G.; SIDDIQUE, A. B.; BAL, A. K.; VERMA, D. P. Roles of plant homologs of Rab1p and Rab7p in the biogenesis of the peribacteroid membrane, a subcellular compartment formed *de novo* during root nodule symbiosis. **The EMBO Journal**, v. 12, n. 11, p. 4125–35, 1993.

CHUN, J.; OREN, A.; VENTOSA, A.; et al. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. **International Journal of Systematic and Evolutionary Microbiology**, v. 68, p. 461–466, 2018.

CHUN, J.; RAINEY, F. A. Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*. **International Journal of Systematic and Evolutionary Microbiology**, v. 64, p. 316–324, 2014.

COIL, D.; JOSPIN, G.; DARLING, A. E. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. **Bioinformatics**, v. 31, n. 4, p. 587–589, 2015.

COLE, J. R.; KONSTANTINIDIS, K.; FARRIS, R. J.; TIEDJE, J. M. Microbial diversity and phylogeny: extending from rRNAs to genomes. In: LIU, W. T.; JANSSON, J. K. (eds.) **Environmental molecular microbiology**, Norfolk-UK: Caister Academic Press, 2010. p.1-19.

COLWELL, R. R. Polyphasic taxonomy of the genus *Vibrio*: numerical taxonomy of *Vibrio cholerae*, *Vibrio parahaemolyticus*, and related *Vibrio* species. **Journal of Bacteriology**, v. 104, n. 1, p. 410–433, 1970.

CONCEIÇÃO, A.; QUEIROZ, L. P. DE; LEWIS, G. P.; et al. Phylogeny of *Chamaecrista* Moench (Leguminosae-Caesalpinioideae) based on nuclear and chloroplast DNA regions. **Taxon**, v. 58, n. 4, p. 1168–1180, 2009.

DALL'AGNOL, R.; BOURNAUD, C.; MOULIN, L.; et al. Genetic diversity of symbiotic *Paraburkholderia* species isolated from nodules of *Mimosa pudica* (L.) and *Phaseolus vulgaris* (L.) grown in soils of the Brazilian Atlantic Forest (Mata Atlântica). **FEMS Microbiology Ecology**, v. 93, p. 1–15, 2017.

DARRIBA, D.; TABOADA, G. L.; DOALLO, R.; POSADA, D. JModelTest 2: More models, new heuristics and parallel computing. **Nature Methods**, v. 9, n. 8, p. 772, 2012.

DELAMUTA, J.; RIBEIRO, R. A.; ORMEÑO-ORRILLO, E.; et al. Polyphasic evidence supporting the reclassification of *Bradyrhizobium japonicum* group Ia strains as *Bradyrhizobium diazoefficiens* sp. nov. **International Journal of Systematic and Evolutionary Microbiology**,

v. 63, p. 3342–3351, 2013.

DELAMUTA, J.; RIBEIRO, R.; ARAÚJO, J.; et al. *Bradyrhizobium stylosanthis* sp. nov., comprising nitrogen-fixing symbionts isolated from nodules of the tropical forage legume *Stylosanthes* spp. **International Journal of Systematic and Evolutionary Microbiology**, v. 66, n. 8, p. 3078–3087, 2016.

DELAMUTA, J.; RIBEIRO, R.; MENNA, P.; BANGEL, E.; HUNGRIA, M. Multilocus sequence analysis (MLSA) of *Bradyrhizobium* strains: revealing high diversity of tropical diazotrophic symbiotic bacteria. **Brazilian Journal of Microbiology**, v. 62, p. 698–710, 2012.

DELAMUTA, J.; MENNA, P.; RIBEIRO, R.; HUNGRIA, M. Phylogenies of symbiotic genes of *Bradyrhizobium* symbionts of legumes of economic and environmental importance in Brazil support the definition of the new symbiovars pachyrhizi and sojae. **Systematic and Applied Microbiology**, v. 40, n. 5, p. 254–265, 2017.

DELAUX, P.; RADHAKRISHNAN, G.; OLDROYD, G. Tracing the evolutionary path to nitrogen-fixing crops. **Current Opinion in Plant Biology**, v. 26, p. 95–99, 2015.

DHARMATILAKE, A. J.; BAUER, W. D. Chemotaxis of *Rhizobium meliloti* towards nodulation gene-inducing compounds from alfalfa roots. **Applied and Environmental Microbiology**, v. 58, n. 4, p. 1153–1158, 1992.

DOWLING, D. N.; BROUGHTON, W. J. Competition for nodulation of legumes. **Annual Review of Microbiology**, v. 40, n. 1, p. 131–157, 1986.

DOYLE, J. J. Phylogenetic perspectives on nodulation: evolving views of plants. **Trends in Plant Sciences**, v. 3, n. 12, p. 473–478, 1998.

DURÁN, D.; REY, L.; MAYO, J.; et al. *Bradyrhizobium paxllaeri* sp. nov. and *Bradyrhizobium icense* sp. nov., nitrogen-fixing rhizobial symbionts of Lima bean (*Phaseolus lunatus* L.) in Peru. **International Journal of Systematic and Evolutionary Microbiology**, v. 64, p. 2072–2078, 2014.

EAGLESHAM, A.; ELLIS, J.; EVANS, W.; et al. The first photosynthetic N₂-fixing rhizobium: characteristics. IN: GRESSHOFF, P.M; ROTH, L.E.; STACEY, G.; NEWTON, W.E. (Eds.). **Nitrogen fixation: achievements and objectives**. New York: Chapman and Hall, 1990. p.805-811.

EDGAR, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, v. 32, n. 5, p. 1792–1797, 2004.

ERISMAN, J. W.; SUTTON, M. A.; GALLOWAY, J.; KLIMONT, Z.; WINIWARTER, W. How a century of ammonia synthesis changed the world. **Nature Geoscience**, v. 1, n. 10, p. 636–639, 2008.

FAKRUDDIN, M.; BIN MANNAN, K. S.; MOHAMMAD MAZUMDAR, R. M.; et al. Identification and characterization of microorganisms: DNA-fingerprinting methods. **Songklanakarin Journal of Science and Technology**, v. 35, n. 4, p. 397–404, 2013.

FARIA, S. M.; FRANCO, A. A.; JESUS, R.; et al. New nodulating legume trees from south-east Brazil. **New Phytologist**, v. 98, p. 317–328, 1984.

FARIA, S. M.; MCINROY, S. G.; SPRENT, J. I. The occurrence of infected cells, with persistent infection threads, in legume root nodules. **Canadian Journal of Botany**, v. 65, p. 553–558, 1987.

FELSENSTEIN, J. Evolutionary trees from DNA sequences: a Maximum Likelihood approach. **Journal of Molecular Evolution**, v. 17, p. 368–376, 1981.

FELSENSTEIN, J. Confidence limits on phylogenies: an approach using the bootstrap.

Evolution, v. 39, n. 4, p. 783–7791, 1985.

FRED, E.; BALDWIN, I.; MCCOY, E. **Root nodule bacteria of leguminous plants**. Madison: The University of Wisconsin, 1932. p. 167.

GAGE, D. J. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. **Microbiology and Molecular Biology Reviews**, v. 68, n. 2, p. 280–300, 2004.

GAWORZEWSKA, E. T.; CARLILE, M. J. Positive chemotaxis of *Rhizobium leguminosarum* and other bacteria towards root exudates from legumes and other plants. **Microbiology**, v. 128, n. 6, p. 1179–1188, 1982.

GERMANO, M.; 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, n. 1, p. 217–229, 2006.

GEVERS, D.; COHAN, F. M.; LAWRENCE, J. G.; et al. Re-evaluating prokaryotic species. **Nature Reviews Microbiology**, v. 3, n. 9, p. 733–739, 2005.

GILLER, K.; WILSON, K. J. **Nitrogen fixation in tropical cropping systems**. Wallingford: CAB International, 1991. p. 313.

GIRAUD, E.; AVARRE, J.; JAUBERT, M.; et al. Legumes symbioses: absence of *nod* genes in photosynthetic bradyrhizobia. **Science**, v. 316, p. 1307–1312, 2007.

GLENN, T. C. Field guide to next-generation DNA sequencers. **Molecular Ecology Resources**, v. 11, n. 5, p. 759–769, 2011.

GORIS, J.; KONSTANTINIDIS, K. T.; KLAPPENBACH, J. A.; et al. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. **International Journal of Systematic and Evolutionary Microbiology**, v. 57, n. 1, p. 81–91, 2007.

GOUY, M.; GUINDON, S.; GASCUEL, O. Sea view version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. **Molecular Biology and Evolution**, v. 27, n. 2, p. 221–224, 2010.

GREEN, M. R.; SAMBROOK, J. **Molecular cloning: a laboratory manual**. 4th edition. New York: Cold Spring Harbor Laboratory Press, 2012. p. 75-78.

GRONEMEYER, J.; CHIMWAMUROMBE, P.; REINHOLD-HUREK, B. *Bradyrhizobium subterraneum* sp. nov., a symbiotic nitrogen-fixing bacterium from root nodules of groundnuts. **International Journal of Systematic and Evolutionary Microbiology**, v. 65, p. 3241–3247, 2015.

GRONEMEYER, J.; HUREK, T.; BUNGER, W.; REINHOLD-HUREK, B. *Bradyrhizobium vignae* sp. nov., a nitrogen-fixing symbiont isolated from effective nodules of *Vigna* and *Arachis*. **International Journal of Systematic and Evolutionary Microbiology**, v. 66, p. 62–69, 2016.

GULASH, M.; AMES, P.; LAROSILIERE, R. C.; BERGMAN, K. Rhizobia are attracted to localized sites on legume roots. **Applied and Environmental Microbiology**, v. 48, n. 1, p. 149–152, 1984.

GUREVICH, A.; SAVELIEV, V.; VYAHHI, N.; TESLER, G. QUASt: Quality assessment tool for genome assemblies. **Bioinformatics**, v. 29, n. 8, p. 1072–1075, 2013.

HALL, T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium Series**, v. 41, p. 95–98, 1999.

HELENE, L. C. F.; DELAMUTA, J. R. M.; RIBEIRO, R. A.; et al. *Bradyrhizobium viridifuturi* sp. nov., encompassing nitrogen-fixing symbionts of legumes used for green manure and

- environmental services. **International Journal of Systematic and Evolutionary Microbiology**, v. 65, n. 12, p. 4441–4448, 2015.
- HELENE, L. C. F.; DELAMUTA, J. R. M.; RIBEIRO, R. A.; HUNGRIA, M. *Bradyrhizobium mercantei* sp. nov., a nitrogen-fixing symbiont isolated from nodules of *Deguelia costata* (syn. *Lonchocarpus costatus*). **International Journal of Systematic and Evolutionary Microbiology**, v. 67, n. 6, p. 1827–1834, 2017.
- HIREL, B.; TÉTU, T.; LEA, P. J.; DUBOIS, F. Improving nitrogen use efficiency in crops for sustainable agriculture. **Sustainability**, v. 3, n. 9, p. 1452–1485, 2011.
- HOLLAND, A. A. Competition between soil and seed borne *Rhizobium trifolii* in nodulation of introduced *Trifolium subterraneum*. **Plant and Soil**, v. 1217, n. 32, p. 293–302, 1970.
- HULTON, C. S. J.; HIGGINS, C. F.; SHARP, P. M. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria. **Molecular Microbiology**, v. 5, n. 4, p. 825–834, 1991.
- HUNGRIA, M.; MENDES, I. C.; MERCANTE, F.M. **A fixação biológica do nitrogênio como tecnologia de baixa emissão de carbono para as culturas do feijoeiro e da soja**. Londrina: Embrapa Soja, 2013. p. 14 (Embrapa Soja. Documentos, 337).
- HUNGRIA, M.; FRANCHINI, J. C.; CAMPO, R. J.; et al. Nitrogen nutrition of soybean in Brazil: contributions of biological N₂ fixation and N fertilizer to grain yield. **Canadian Journal of Plant Science**, v. 86, p. 927–939, 2006.
- HUNGRIA, M.; CAMPO, R. J.; MENDES, I. C. **Fixação biológica do nitrogênio na cultura da soja**. Londrina: Embrapa Soja, 2001. p. 48. (Embrapa Soja. Circular Técnica, 35).
- HUNGRIA, M.; MENNA, P.; DELAMUTA, J. *Bradyrhizobium*, the ancestor of all rhizobia: phylogeny of housekeeping and nitrogen-fixation genes. In: DE BRUIJN, F. (Ed.). **Biological nitrogen fixation**. Chapter 18. New Jersey: John Wiley & Sons, Inc., 2015. vol.1. p.191-202.
- HUNGRIA, M.; O'HARA, G. W.; ZILLI, J. E.; et al. Isolation and growth of rhizobia. In: HOWIESON, J.G.; DILWORTH, M.J. (Eds.) **Working with rhizobia**. Chapter 3. Canberra, Australia: Australian Center for International Agricultural Research (ACIAR), 2016. p. 39-60.
- ISLAM, M. S.; KAWASAKI, H.; MURAMATSU, Y.; NAKAGAWA, Y.; SEKI, T. *Bradyrhizobium iriomotense* sp. nov., isolated from a tumor-like root of the legume *Entada koshunensis* from Iriomote island in Japan. **Bioscience, Biotechnology, and Biochemistry**, v. 72, n. 6, p. 1416–1429, 2008.
- JACCARD, P. The distribution of flora in the alpine zone. **New Phytologist**, v. 11, n. 2, p. 37, 1912.
- JONES, K. M.; KOBAYASHI, H.; DAVIES, B. W.; et al. How rhizobial symbionts invade plants: the *Sinorhizobium - Medicago* model. **Nature Reviews Microbiology**, v. 5, n. 8, p. 619–633, 2007.
- JORDAN, D. C. Transfer of *Rhizobium japonicum* Buchananm 1980 to *Bradyrhizobium japonicum* gen. nov., a genus of slow growing, root nodule bacteria from leguminous plants. **International Journal of Systematic Bacteriology**, v. 32, n. 1, p. 378–380, 1982.
- JORDAN, D. C.; ALLEN, O. N. Family III. Rhizobiaceae Conn 1938. In KRIEG N. R.; HOLT J. G. (Eds.) **Bergey's Manual of Systematic Bacteriology**, Baltimore, MD: Williams & Wilkins, 1984. p.234–235.
- KÄMPFER, P.; GLAESER, S. P. Under the microscope: prokaryotic taxonomy in the sequencing era and the role of MLSA in classification. **Microbiology Australia**, v. 32, n. 2008, p. 66–70, 2011.

- KÄMPFER, P.; GLAESER, S. P. Prokaryotic taxonomy in the sequencing era - the polyphasic approach revisited. **Environmental Microbiology**, v. 14, n. 2, p. 291–317, 2012.
- KAPE, R.; PARNISKE, M.; WERNER, D.; et al. Chemotaxis and *nod* gene activity of *Bradyrhizobium japonicum* in response to hydroxycinnamic acids and isoflavonoids. **Applied and Environmental Microbiology**, v. 57, n. 1, p. 316–319, 1991.
- KARR, D. B.; EMERICH, D. W. Protein phosphorylation in *Bradyrhizobium japonicum* bacteroids and cultures. **Journal of Bacteriology**, v. 171, n. 6, p. 3420–3426, 1989.
- KIM, M.; OH, H. S.; PARK, S. C.; CHUN, J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. **International Journal of Systematic and Evolutionary Microbiology**, v. 64, p. 346–351, 2014.
- KLENK, H. P.; GÖKER, M. En route to a genome-based classification of *Archaea* and *Bacteria*? **Systematic and Applied Microbiology**, v. 33, n. 4, p. 175–182, 2010.
- KOEUTH, T.; VERSALOVIC, J.; LUPSKI, J. R. Differential subsequence conservation of interspersed repetitive *Streptococcus pneumoniae* BOX elements in diverse bacteria. **Genome Research**, v. 5, n. 4, p. 408–18, 1995.
- KONSTANTINIDIS, K. T.; RAMETTE, A.; TIEDJE, J. M. Toward a more robust assessment of intraspecies diversity, using fewer genetic markers. **Applied and Environmental Microbiology**, v. 72, n. 11, p. 7286–7293, 2006.
- KONSTANTINIDIS, K. T.; TIEDJE, J. M. Genomic insights that advance the species definition for prokaryotes. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 7, p. 2567–2572, 2005.
- 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, n. 6, p. 501–505, 1992.
- LAGUERRE, G.; ALLARD, M. R.; REVOY, F.; AMARGER, N. Rapid identification of rhizobia by restriction fragment length polymorphism analysis of PCR-Amplified 16S rRNA genes. **Applied and Environmental Microbiology**, v. 60, p. 56–63, 1994.
- LAGUERRE, G.; MAVINGUI, P.; ALLARD, M. R.; et al. Typing of rhizobia by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions: application to *Rhizobium leguminosarum* and its different biovars. **Applied and Environmental Microbiology**, v. 62, n. 6, p. 2029–2036, 1996.
- LAGUERRE, G.; NOUR, S. M.; SANJUAN, J.; DROUIN, P.; et al. Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. **Microbiology**, v. 147, p. 981–993, 2001.
- LAPAGE, S.P.; SNEATH P.H.A.; LESSEL E.F.; SKERMAN V.B.D.; SEELIGER H.P.R.; CLARK W.A. (Eds). **International code of nomenclature of Bacteria: bacteriological code** (1990 Revision). Washington, DC: American Society for Microbiology, 1992. p. 232.
- LAVIN, M.; PENNINGTON, R. T.; KLITGAARD, B. B.; et al. The Dalbergioid legumes (Fabaceae): delimitation of a pantropical monophyletic clade. **American Journal of Botany**, v. 88, n. 3, p. 503–533, 2014.
- LEWIS, G.; SCHRIRE, B.; MACKINDER, B.; LOCK, M. **Legumes of the world**. Kew, UK: Royal Botanic Gardens, 2005.
- LI, Y. H.; WANG, R.; ZHANG, X. X.; et al. *Bradyrhizobium guangdongense* sp. nov. and *Bradyrhizobium guangxiense* sp. nov., isolated from effective nodules of peanut. **International**

Journal of Systematic and Evolutionary Microbiology, v. 65, p. 4655–4661, 2015.

LLORET, L.; MARTÍNEZ-ROMERO, E. Evolución y filogenia de *Rhizobium*. **Revista Latinoamericana de Microbiología**, v. 47, n. 1–2, p. 43–60, 2005.

LPWG, L. P. W. G. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. **Taxon**, v. 66, p. 44–77, 2017.

LUDWIG, W.; SCHLEIFER, K. H. Bacterial phylogeny based on 16S and sequence analysis 23S rRNA. **FEMS Microbiology Reviews**, v. 15, p. 155–173, 1994.

MADIGAN, M.; MARTINKO, J.; BENDER, K. **Microbiologia de Brock**. 14th edição. Porto Alegre: Artmed Editora LTDA, 2016. p. 987.

MADSEN, L. H.; JURKIEWICZ, A.; SULLIVAN, J. T.; et al. The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. **Nature Communications**, v.1, n. 10, p. 1–12, 2010.

MARTENS, M.; DAWYNDT, P.; COOPMAN, R.; et al. Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). **International Journal of Systematic and Evolutionary Microbiology**, v. 58, p. 200–214, 2008.

MARTIN, B.; HUMBERT, O.; CAMARA, M.; et al. A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. **Nucleic Acids Research**, v. 20, p. 3479–3483, 1992.

MARTINS, E.; AZARIAS, A.; RAYSSA, G.; et al. *Bradyrhizobium brasilense* sp. nov., a symbiotic nitrogen fixing bacterium isolated from Brazilian tropical soils. **Archives of Microbiology**, v. 17, p. 1–11, 2017.

MARTINS, E.; AZARIAS GUIMARÃES, A.; SOARES DE CARVALHO, T.; et al. *Bradyrhizobium forestalis* sp. nov., an efficient nitrogen-fixing bacterium isolated from nodules of forest legume species in the Amazon. **Archives of Microbiology**, v. 200, n. 5, p. 743–752, 2018.

DE MATOS, G. F.; ZILLI, J. E.; DE ARAÚJO, J. L. S.; et al. *Bradyrhizobium sacchari* sp. nov., a legume nodulating bacterium isolated from sugarcane roots. **Archives of Microbiology**, v. 199, n. 9, p. 1251–1258, 2017.

MEIER-KOLTHOFF, J.; AUCH, A.; KLENK, H.; GOKER, M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. **BMC Bioinformatics**, v. 14, p. 1–14, 2013.

MEIER-KOLTHOFF, J. P.; KLENK, H.; GO, M. Taxonomic use of DNA G+C content and DNA–DNA hybridization in the genomic age. **International Journal of Systematic and Evolutionary Microbiology**, v. 64, p. 352–356, 2014.

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, n. 12, p. 2934–2950, 2009.

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, n. 12, p. 3052–3067, 2011.

MENNA, P.; HUNGRIA, M.; BARCELLOS, F. G.; et al. Molecular phylogeny based on the 16S rRNA gene of elite rhizobial strains used in Brazilian commercial inoculants. **Systematic and**

Applied Microbiology, v. 29, n. 4, p. 315–332, 2006.

MENNA, P.; PEREIRA, A. A.; BANGEL, E. V.; HUNGRIA, M. Rep-PCR of tropical rhizobia for strain fingerprinting, biodiversity appraisal and as a taxonomic and phylogenetic tool. **Symbiosis**, v. 48, n. 1–3, p. 120–130, 2009a.

MESBAH, N. M.; WHITMAN, W. B.; MESBAH, M. Determination of the G+C Content of Prokaryotes. In: RAINEY, F.; OREN, A. (eds.) **Taxonomy of Prokaryotes**, London: Elsevier Ltd, 2011.

MICHEL, D. C.; PASSOS, S. R.; SIMÕES-ARAUJO, J. L.; et al. *Bradyrhizobium centrolobii* and *Bradyrhizobium macuxiense* sp. nov. isolated from *Centrolobium paraense* grown in soil of Amazonia, Brazil. **Archives of Microbiology**, v. 199, n. 5, p. 657–664, 2017.

MIWA, H.; SUN, J.; OLDROYD, G. E. D.; DOWNIE, J. A. Analysis of Nod-factor-induced calcium signaling in root hairs of symbiotically defective mutants of *Lotus japonicus*. **Molecular Plant-Microbe Interactions**, v. 19, n. 8, p. 914–923, 2006.

MOREIRA, F.; HAUKKA, K.; YOUNG, J. P. W. Biodiversity of rhizobia isolated from a wide range of forest legumes in Brazil. **Molecular Ecology**, v. 7, p. 889–895, 1998.

MOREIRA, F. M. S.; GILLIS, M.; POT, B.; KERSTERS, K.; FRANCO, A. A. Characterization of rhizobia isolated from different divergence groups of tropical leguminosae by comparative polyacrylamide gel electrophoresis of their total proteins. **Systematic and Applied Microbiology**, v. 16, n. 1, p. 135–146, 1993.

MORRIS, J. B. Showy partridge pea [*Chamaecrista fasciculata* (Michx.) Greene] with potential for cultivation as a multi-functional species in the United States. **Genetic Resources and Crop Evolution**, v. 59, p. 1577–1581, 2012.

MOULIN, L.; MUNIUE, A.; DREYFUS, B.; BOIVIN-MASSON, C. Nodulation of legumes by members of the beta-subclass of Proteobacteria. **Nature**, v. 411, p. 948–950, 2001.

NAISBITT, T.; JAMES, E. K.; SPRENT, J. I. The evolutionary significance of the legume genus *Chamaecrista*, as determined by nodule structure. **New Phytologist**, v. 122, n. 3, p. 487–492, 1992.

NELSON, D. L.; COX, M. M. Princípios de bioquímica de Lehninger. 6ª edição. Porto Alegre: Artmed, 2014. p. 1298.

NORRIS, D. O. Acid production by *Rhizobium* a unifying concept. **Plant and Soil**, v. 22, n. 2, p. 143–166, 1965.

OKAZAKI, S.; KANEKO, T.; SATO, S.; SAEKI, K. Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. **Proceedings of the National Academy of Sciences of the United States of America**, v. 110, n. 42, p. 17131–17136, 2013.

OLDROYD, G. E. D. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. **Nature Reviews Microbiology**, v. 11, n. 4, p. 252–263, 2013.

ON, S. L. W.; HOLMES, B. Assessment of enzyme detection tests useful in identification of Campylobacteria. **Journal of Clinical Microbiology**, v. 30, n. 3, p. 746–749, 1992.

ORMEÑO-ORRILLO, E.; HUNGRIA, M.; MARTÍNEZ-ROMERO, E.. Dinitrogen-Fixing Prokaryotes. In: ROSENBERG, E.; DELONG, E. F.; STACKEBRANDT, E.; LORY, S.; THOMPSON, F. **The Prokaryotes: Prokaryotic physiology and biochemistry**, Berlin: Springer-Verlag, 2013. p. 1–662.

OTT, T.; VAN DONGEN, J. T.; GÜNTHER, C.; et al. Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. **Current Biology**, v. 15, n. 6, p. 531–535, 2005.

- PARKER, M. A. rRNA and *dnaK* relationships of *Bradyrhizobium* sp. nodule bacteria from four Papilionoid legume trees in Costa Rica. **Systematic and Applied Microbiology**, v. 27, n. 3, p. 334–342, 2004.
- PARKER, M. A. The spread of *Bradyrhizobium* lineages across host legume clades: from *Abarema* to *Zygia*. **Microbial Ecology**, v. 69, p. 630–640, 2015.
- PARKER, M. A.; ROUSTEAU, A. Molecular phylogenetics and evolution mosaic origins of *Bradyrhizobium* legume symbionts on the Caribbean island of Guadeloupe. **Molecular Phylogenetics and Evolution**, v. 77, p. 110–115, 2014.
- PEIX, A.; RAMÍREZ-BAHENA, M. H.; VELÁZQUEZ, E.; BEDMAR, E. J. Bacterial associations with legumes. **Critical Reviews in Plant Sciences**, v. 34, p. 17–42, 2015.
- PEOPLES, M. B.; CRASWELL, E. T. Biological nitrogen fixation: Investments, expectations and actual contributions to agriculture. **Plant and Soil**, v. 141, p. 13–39, 1992.
- POLHILL, R.; RAVEN, P. **Advances in legume systematics**. Kew-UK: Royal Botanic Gardens, 1981.
- PROVOROV, N. A.; VOROB'EV, N. I. Evolutionary genetics of nodule bacteria: molecular and population aspects. **Russian Journal of Genetics**, v. 36, n. 12, p. 1573–1587, 2000.
- RAMÍREZ-BAHENA, M. H.; PEIX, A.; RIVAS, R.; et al. *Bradyrhizobium pachyrhizi* sp. nov. and *Bradyrhizobium jicamae* sp. nov., isolated from effective nodules of *Pachyrhizus erosus*. **International Journal of Systematic and Evolutionary Microbiology**, v. 59, n. 8, p. 1929–1934, 2009.
- RICHTER, M.; ROSSELLO-MORA, R. Shifting the genomic gold standard for the prokaryotic species definition. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 45, p. 19126–19131, 2009.
- RICHTER, M.; ROSSELLO, R.; GLOCKNER, F.; PEPLIES, J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. **Bioinformatics**, v. 32, n. 6, p. 929–931, 2016.
- RIVAS, R.; GARCÍA-FRAILE, P.; VELÁZQUEZ, E. Taxonomy of bacteria nodulating Legumes. **Microbiology Insights**, v. 2, p. 51–69, 2009.
- RIVAS, R.; MARTENS, M.; DE LAJUDIE, P.; WILLEMS, A. Multilocus sequence analysis of the genus *Bradyrhizobium*. **Systematic and Applied Microbiology**, v. 32, n. 2, p. 101–110, 2009.
- RIVAS, R.; WILLEMS, A.; PALOMO, J. L.; et al. *Bradyrhizobium betae* sp. nov., isolated from roots of *Beta vulgaris* affected by tumour-like deformations. **International Journal of Systematic and Evolutionary Microbiology**, v. 54, n. 4, p. 1271–1275, 2004.
- RODRIGUEZ-R, L. M.; KONSTANTINIDIS, K. T. Bypassing cultivation to identify bacterial species. **Microbe**, v. 9, n. 3, p. 111–118, 2014.
- RODRIGUEZ-R, L. M.; KONSTANTINIDIS, K. T. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes microbial genomes and metagenomes. **PeerJ Preprints**, v. 4, p. 1–15, 2016.
- ROSSELLO-MORA, R.; AMANN, R. The species concept for prokaryotes. **FEMS Microbiology Review**, v. 25, n. 1, p. 39–67, 2001.
- SAITOU, N.; NEI, M. The Neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v. 4, p. 406–425, 1987.
- SANTOS, J. M. F. DOS; CASAES ALVES, P. A.; SILVA, V. C.; et al. Diverse genotypes of *Bradyrhizobium* nodulate herbaceous *Chamaecrista* (Moench) (Fabaceae, Caesalpinioideae)

- species in Brazil. **Systematic and Applied Microbiology**, v. 40, n. 2, p. 69–79, 2017.
- SANTOS, P. E. L.; PALMER, M.; BEUKES, C.; et al. Whole genome analyses suggests that *Burkholderia* sensu lato contains two additional novel genera (*Mycetohabitans* gen. nov., and *Trinickia* gen. nov.): implications for the evolution of diazotrophy and nodulation in the *Burkholderiaceae*. **Genes**, v. 9, n. 389, p. 1–23, 2018.
- SASSER, M. **Identification of bacteria by gas chromatography of cellular fatty acids**. MIDI - Technical Note #101, 1990.
- SCHLEIFER, K. H. Classification of *Bacteria* and *Archaea*: Past, present and future. **Systematic and Applied Microbiology**, v. 32, n. 8, p. 533–542, 2009.
- SCHLEIFER, K. H.; STACKEBRANDT, E. Molecular systematics of prokaryotes. **Annual Review of Microbiology**, v. 37, p. 143–87, 1983.
- SCHULTZE, M.; KONDOROSI, A. Regulation of symbiotic root nodule development. **Annual Review of Genetics**, , n. 32, p. 33–57, 1998.
- SENTAUSA, E.; FOURNIER, P. E. Advantages and limitations of genomics in prokaryotic taxonomy. **Clinical Microbiology and Infection**, v. 19, n. 9, p. 790–795, 2013.
- SHEARER, G.; KOHL, D. H. N₂-fixation in field settings: estimations based on natural ¹⁵N abundance. **Australian Journal of Plant Physiology**, v. 13, p. 699–756, 1986.
- SILVA, F. V.; DE MEYER, S. E.; SIMÕES-ARAÚJO, J. L.; et al. *Bradyrhizobium manausense* sp. nov., isolated from effective nodules of *Vigna unguiculata* grown in Brazilian Amazonian rainforest soils. **International Journal of Systematic and Evolutionary Microbiology**, v. 64, p. 2358–2363, 2014.
- SINGER, S. R.; MAKI, S. L.; FARMER, A. D.; et al. Venturing beyond beans and peas: what can we learn from *Chamaecrista*? **Plant Physiology**, v. 151, p. 1041–1047, 2009.
- SNEATH, P. H. A.; SOKAL, R. R. **Numerical taxonomy**: the principles and practice of numerical classification. San Francisco: Freeman, 1973. p. 573.
- SOMASEGARAN, P.; HOBEN, H. Methods in Legume-*Rhizobium* technology. In: SOMASEGARAN, P.; HOBEN, H. (eds.) **Handbook for rhizobia**, New York: Springer, 1985. p. 450.
- SOUZA, V. C.; BORTOLUZZI, R. L. C. *Chamaecrista* In: **Lista de espécies da flora do Brasil**, Rio de Janeiro: Jardim Botânico do Rio de Janeiro. Disponível em: <<http://www.floradobrasil.jbrj.gov.br/jabot/floradobrasil/FB22876>>. Acesso em: 16 fev, 2019.
- SPRENT, J. I. Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. **New Phytologist**, v. 174, p. 11–25, 2007.
- SPRENT, J. I.; ARDLEY, J.; JAMES, E. K. Biogeography of nodulated legumes and their nitrogen-fixing symbionts. **New Phytologist**, v. 215, p. 40–56, 2017.
- SPRENT, J. I.; ARDLEY, J. K.; JAMES, E. K. From North to South: a latitudinal look at legume nodulation processes. **South African Journal of Botany**, v. 89, p. 31–41, 2013.
- SPRENT, J. I.; RAVEN, J. A. Evolution of nitrogen-fixing symbioses. **Proceedings of the Royal Society of Edinburgh**, v. 85, p. 215–237, 1985.
- STACKEBRANDT, E.; FREDERIKSEN, W.; GARRITY, G. M.; et al. Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. **International Journal of Systematic and Evolutionary Microbiology**, v. 52, p. 1043–1047, 2002.
- STEPKOWSKI, T.; MOULIN, L.; KRZYZANSKA, A.; et al. European origin of

- Bradyrhizobium* populations infecting lupins and serradella in soils of Western Australia and South Africa. **Applied and Environmental Microbiology**, v. 71, n. 11, p. 7041–7052, 2005.
- STERN, M. J.; AMES, G. F. L.; SMITH, N. A.; ROBINSON, E. C.; et al. Repetitive extrapalindromic sequences: a major component of the bacterial genome. **Cell**, v. 37, p. 1015–1026, 1984.
- TAIZ, L.; ZEIGER, E. Nutrição mineral. In: TAIZ, L.; ZEIGER, E. (eds.) **Fisiologia vegetal**, Porto Alegre: Artmed, 2004. p.722.
- TAMURA, K.; STECHER, G.; PETERSON, D.; FILIPSKI, A.; KUMAR, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. **Molecular Biology and Evolution**, v. 30, n. 12, p. 2725–2729, 2013.
- TIGHE, S. W.; LAJUDIE, P. DE; DIPIETRO, K.; LINDSTRO, K. Analysis of cellular fatty acids and phenotypic relationships of *Agrobacterium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium* species using the Sherlock Microbial Identification System. **International Journal of Systematic and Evolutionary Microbiology**, v. 50, p. 787–801, 2000.
- TINDALL, B. J.; ROSSELLÓ-MÓRA, R.; BUSSE, H. J.; LUDWIG, et al. Notes on the characterization of prokaryote strains for taxonomic purposes. **International Journal of Systematic and Evolutionary Microbiology**, v. 60, n. 1, p. 249–266, 2010.
- USDA, NRCS. **The plants database** (<http://plants.usda.gov>). National Plant Data Team, Greensboro, 2012.
- VAN BELKUM, A. Short sequence repeats in microbial pathogenesis and evolution. **CMLS Cellular and Molecular Life Sciences**, v. 56, n. 9–10, p. 729–34, 1999.
- VAN DER HEIJDEN, M. G. A.; BARDGETT, R. D.; VAN STRAALLEN, N. M. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. **Ecology Letters**, v. 11, n. 3, p. 296–310, 2008.
- VANDAMME, P.; PEETERS, C. Time to revisit polyphasic taxonomy. **Antonie van Leeuwenhoek**, v. 106, n. 1, p. 57–65, 2014.
- VANDAMME, P.; POT, B.; GILLIS, M.; et al. Polyphasic taxonomy, a consensus approach to bacterial systematics. **Microbiological Reviews**, v. 47, n. 2, p. 98–108, 1996.
- VERMA, D. P. S.; CHEON, C.; LEE, N.; HONG, Z.; MIAO, G. Biogenesis of peribacteroid membrane (PBM) forming a subcellular compartment essential for symbiotic nitrogen fixation. In: PALAIOS, R.; MORA, J.; NEWTON, W. **New horizons in nitrogen fixation**, Cancún: Kluwer Academic Publishers, 1993. p. 269–274, 1993.
- VERSALOVIC, J.; SCNEIDER, M.; BRUIJIN, F.; LUPSKI, J. R. Genomic fingerprinting of bacteria using repetitive sequence- based polymerase chain reaction. **Methods in Molecular and Cell Biology**, v. 5, n. 1, p. 25–40, 1994.
- VINCENT, J. **A Manual for the practical study of the root-nodule Bacteria**. Oxford-Edinburgh: Blackwell Scientific Publ., 1970. p. 164.
- VINUESA, P.; LEÓN-BARRIOS, M.; SILVA, C.; et al. *Bradyrhizobium canariense* sp. nov., an acid-tolerant endosymbiont that nodulates endemic genistoid legumes (Papilionoideae: Genisteeae) from the Canary islands, along with *Bradyrhizobium japonicum* bv. genistearum, *Bradyrhizobium* genospecies alpha and *Bradyrhizobium* genospecies beta. **International Journal of Systematic and Evolutionary Microbiology**, v. 55, n. 2, p. 569–575, 2005.
- WANG, J. Y.; WANG, R.; ZHANG, Y. M.; et al. *Bradyrhizobium daqingense* sp. nov., isolated from soybean nodules. **International Journal of Systematic and Evolutionary Microbiology**, v. 63, p. 616–624, 2013a.

- WANG, R.; CHANG, Y. L.; ZHENG, W. T.; et al. *Bradyrhizobium arachidis* sp. nov., isolated from effective nodules of *Arachis hypogaea* grown in China. **Systematic and Applied Microbiology**, v. 36, n. 2, p. 101–105, 2013.
- WANG, Q.; LIU, J.; ZHU, H. Genetic and molecular mechanisms underlying symbiotic specificity in legume-*Rhizobium* interactions. **Frontiers in Plant Science**, v. 9, n. 313, p. 1–8, 2018.
- WAYNE, L. G.; MOORE, W. E. C.; STACKEBRANDT, E.; et al. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. **International Journal of Systematic and Evolutionary Microbiology**, v. 37, n. 4, p. 463–464, 1987.
- WEISBURG, W. G.; BARNS, S. M.; PELLETIER, D. A.; LANE, D. J. 16S Ribosomal DNA amplification for phylogenetic study. **Journal of Bacteriology**, v. 173, n. 2, p. 697–703, 1991.
- WILLEMS, A.; COOPMAN, R.; GILLIS, M. Comparison of sequence analysis of 16S – 23S rDNA spacer regions, AFLP analysis and DNA – DNA hybridizations in *Bradyrhizobium*. **International Journal of Systematic and Evolutionary Microbiology**, v. 51, p. 623–632, 2001.
- WILLEMS, A.; MUNIVE, A.; LAJUDIE, P. DE; GILLIS, M. In most *Bradyrhizobium* groups sequence comparison of 16S-23S rDNA internal transcribed spacer regions corroborates DNA-DNA hybridizations. **Systematic and Applied Microbiology**, v. 210, p. 203–210, 2003.
- WOESE, C. Bacterial Evolution. **Microbiology Reviews**, v. 51, n. 2, p. 221–271, 1987.
- WOESE, C. R.; KANDLER, O.; WHEELIS, M. L. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. **Proceedings of the National Academy of Sciences of the United States of America**, v. 87, n. 12, p. 4576–4579, 1990.
- XU, L. M.; GE, C.; CUI, Z.; LI, J.; FAN, H. *Bradyrhizobium liaoningense* sp. nov., isolated from the root nodules of soybeans. **International Journal of Systematic Bacteriology**, v. 45, n. 4, p. 706–711, 1995.
- YU, X.; CLOUTIER, S.; TAMBONG, J. T.; BROMFIELD, E. S. P. *Bradyrhizobium ottawaense* sp. nov., a symbiotic nitrogen fixing bacterium from root nodules of soybeans in Canada. **International Journal of Systematic and Evolutionary Microbiology**, v. 64, p. 3202–3207, 2014.
- ZAAT, S. A.; VAN BRUSSEL, A. A.; TAK, T.; PEES, E.; LUGTENBERG, B. J. Flavonoids induce *Rhizobium leguminosarum* to produce *nodDABC* gene-related factors that cause thick, short roots and root hair responses on common vetch. **Journal of Bacteriology**, v. 169, n. 7, p. 3388–3391, 1987.
- ZAHARAN, H. H. *Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. **Microbiology and Molecular Biology Reviews**, v. 63, n. 4, p. 968–989, 1999.
- ZHANG, Y. M.; LI, Y.; CHEN, W. F.; et al. *Bradyrhizobium huanghuaihaiense* sp. nov., an effective symbiotic bacterium isolated from soybean (*Glycine max* L.) nodules. **International Journal of Systematic and Evolutionary Microbiology**, v. 62, n. 8, p. 1951–1957, 2012.
- ZHI, X. Y.; ZHAO, W.; LI, W. J.; ZHAO, G. P. Prokaryotic systematics in the genomics era. **Antonie van Leeuwenhoek**, v. 101, n. 1, p. 21–34, 2012.
- ZWICKL, D. J. **Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion**, Tese não publicada: The University of Texas at Austin, 2006. p. 115

6. RESULTADOS E DISCUSSÃO

Bradyrhizobium frederickii* sp. nov., linhagem fixadora de nitrogênio isolada de nódulos da espécie *Chamaecrista fasciculata* e caracterizada pela tolerância a altas temperaturas *in vitro

Resumo

As simbioses estabelecidas entre leguminosas e bactérias fixadoras de nitrogênio, coletivamente referidas como rizóbios, representam a maior contribuição para a entrada de nitrogênio global através do processo de fixação biológica de nitrogênio (FBN). As espécies do gênero *Bradyrhizobium* se destacam pela lenta taxa de crescimento *in vitro* a 28 °C, vasta distribuição geográfica e ampla faixa de hospedeiros, incluindo leguminosas basais da subfamília Caesalpinioideae. No entanto, ainda existem poucos estudos sobre os simbioses dessa subfamília, incluindo o gênero *Chamaecrista*. Foi conduzido um estudo polifásico com 11 estirpes isoladas de nódulos radiculares de *Chamaecrista fasciculata*, uma herbácea anual multi-funcional dos estados do centro-oeste, leste e sul dos EUA. Com base na filogenia do gene 16S RNAr, as linhagens foram agrupadas no superclado de *B. japonicum*, apresentando uma similaridade de 95,3-99,8 % com outras espécies descritas de *Bradyrhizobium*. A análise da região intergênica 16S-23S rDNA (ITS) indicou 93,7-100 % de similaridade entre as 11 estirpes e menor que 89,9 % com outras espécies de *Bradyrhizobium*. A análise da sequência de multilocus (MLSA) com quatro genes *housekeeping* (*glnII*, *gyrB*, *recA* e *rpoB*) corroborou a classificação das 11 estirpes como um novo grupo, compartilhando menos de 95,2 % de identidade nucleotídica com outras espécies de *Bradyrhizobium*. A análise de MLSA com dez genes de *housekeeping* (*atpD*, *dnaK*, *gap*, *glnII*, *gltA*, *gyrB*, *pnp*, *recA*, *rpoB* e *thrC*) indicou *B. daqingense* como a espécie filogeneticamente relacionada. A alta diversidade genética entre as estirpes foi confirmada nas análises de ITS, MLSA e BOX-PCR. A média de identidade de nucleotídeos (ANI) e a análise da hibridação DNA-DNA digital (dDDH) apresentaram valores abaixo do limiar de delimitação de espécies quando comparados com as espécies descritas de *Bradyrhizobium*, de 89,7 % e 40 %, respectivamente. Nas filogenias dos genes *nifH* e *nodC*, as estirpes foram agrupadas juntas, mas com indicação de transferência horizontal de genes, mostrando maior similaridade com *B. arachidis* e *B. forestalis*. Além disso, outras propriedades fenotípicas, genotípicas e simbióticas foram avaliadas. Os dados obtidos neste estudo suportam a descrição das 11 estirpes como

representantes da nova espécie *Bradyrhizobium frederickii* sp. nov.. A estirpe tipo é a CNPSo 3426^T (=USDA 10052^T = U686^T =CL 20^T).

Palavras-chave: ANI; MLSA; Caesalpinioideae; fixação biológica de nitrogênio; nodulação

Bradyrhizobium frederickii* sp. nov., a nitrogen-fixing lineage isolated from nodules of the caesalpinoid species *Chamaecrista fasciculata* and characterized by tolerance to high temperature *in vitro

Maria Clara de Oliveira Urquiaga,^{1,2,3#} Milena Serenato Klepa,^{1,3,4#} Padma Somasegaran,⁵ Renan Augusto Ribeiro,⁴ Jakeline Renata Marçon Delamuta,^{1,4} and Mariangela Hungria^{1,2,4*}

¹ Embrapa Soja, C.P. 231, 86001-970 Londrina, Paraná, Brazil;

² Department of Microbiology, Universidade Estadual de Londrina, C.P. 10.011, 86057-970, Londrina, Paraná, Brazil;

³ Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, SBN, Quadra 2, Bloco L, Lote 06, Edifício Capes, 70.040-020 Brasília, Distrito Federal, Brazil;

⁴ Conselho Nacional de Desenvolvimento Científico e Tecnológico, SHIS QI 1 Conjunto B, Blocos A, B, C e D, Lago Sul, 71605-001 Brasília, Distrito Federal, Brazil.

⁵ Formerly with BIOZYME Inc., Saint Joseph, Missouri, USA

Both authors contributed equally to this study

* Corresponding author

Running title: *Bradyrhizobium frederickii* sp. nov.

Number of text pages	24
Number of Tables	04
Number of Figures	05
Number of Supplementary Tables	03
Number of Supplementary Figures	09

*Corresponding author:

Mariangela Hungria

Embrapa Soja

Cx. Postal 231

86001-970, Londrina, Paraná, Brazil

Fax: (+55)4333716100

Telephone: (+55)4333716206

E-mail: mariangela.hungria@embrapa.br; biotecnologia.solo@hotmail.com;

E-mail addresses: M.C.O. Urquiaga (mariaclara.urquiaga@gmail.com); M.S. Klepa (milenaserenato@gmail.com); P. Somasegaran (padma.somasegaran@gmail.com); R.A. Ribeiro (renanribeiro83@hotmail.com); J.R.M. Delamuta (jake_renata@hotmail.com); M. Hungria (mariangela.hungria@embrapa.br).

Bradyrhizobium frederickii* sp. nov., a nitrogen-fixing lineage isolated from nodules of caesalpinoid species *Chamaecrista fasciculata* and characterized by tolerance to high temperature *in vitro

Abstract

The symbioses between legumes and nitrogen-fixing rhizobia give the greatest contribution to the global nitrogen input via the process of biological nitrogen fixation (BNF). *Bradyrhizobium* stands out as the main genus nodulating basal Caesalpinioideae. We performed a polyphasic study with eleven strains isolated from root nodules of *Chamaecrista fasciculata*, an annual multi-functional native legume of the USA. In the 16S rRNA gene phylogeny, the strains were clustered in the *B. japonicum* superclade. The analysis of the intergenic transcribed spacer (ITS) indicated less than 89.9 % similarity with other *Bradyrhizobium* species. Multilocus sequence analysis (MLSA) with four housekeeping genes (*glnII*, *gyrB*, *recA* and *rpoB*) confirmed the new group, sharing less than 95.2 % nucleotide identity with other species. The MLSA with ten housekeeping genes (*atpD*, *dnaK*, *gap*, *glnII*, *gltA*, *gyrB*, *pnp*, *recA*, *rpoB* and *thrC*) indicated *B. daqingense* as the closest species. Noteworthy, high genetic diversity among the strains was confirmed in the analyses of ITS, MLSA and BOX-PCR. Average nucleotide identity (ANI) and digital DNA–DNA hybridization (dDDH) values were below the threshold of described *Bradyrhizobium* species, of 89.7 % and 40 %, respectively. In the *nifH* and *nodC* phylogenies, strains were grouped together, but with indication of horizontal gene transfer, showing higher similarity with *B. arachidis* and *B. forestalis*. Other phenotypic, genotypic and symbiotic properties were evaluated, altogether supporting the description of the CNPSo strains as representatives of the new species *Bradyrhizobium frederickii* sp. nov., with CNPSo 3426^T (=USDA 10052^T = U686^T =CL 20^T) as the type strain.

Keywords: ANI; MLSA; Caesalpinioideae; biological nitrogen fixation; nodulation

Microorganisms can interact positively with plants in agricultural systems, providing nutritional gains for both partners involved. Biological nitrogen fixation (BNF) is a key process for maintenance of the nutrients and energy balance in agriculture and for the improvement of soil fertility in productive and sustainable systems. The major contribution to the BNF process occurs by the symbiotic association between legumes belonging to the Fabaceae (= Leguminosae) family and soil bacteria collectively referred as rhizobia [1].

Bradyrhizobium is considered as the dominant symbiotic genus across the Fabaceae family, which includes the Papilionoideae, Caesalpinioideae, Duparquetioideae, Cercidoideae, Detarioideae and Dialioideae subfamilies. In the last few years, an increasing number of phylogenetic studies report high genetic diversity within the genus *Bradyrhizobium* [2, 3, 4, 5, 6, 7, 8, 9]. Furthermore, *Bradyrhizobium* species stands out due to their slow growth rate *in vitro* at 28 °C, broad geographic distribution and wide host range, including basal legumes of the subfamily Caesalpinioideae. Therefore, their representatives are considered the ancestral symbionts of legumes [10, 11].

Among the Fabaceae subfamilies, nodulation has long been known to be rare in Caesalpinioideae [12, 13]. The few legumes in this subfamily able to nodulate reside mainly in the tribe Caesalpiniae, with the exception of *Chamaecrista* Moench (tribe Cassieae) [14, 15]. *Chamaecrista* is considered the largest genus of nodulating caesalpinoids and has Brazil as its centre of diversity. Growth habit in the genus

Chamaecrista includes trees, shrubs and herbs and is one of the few caesalpinoid genus that has expanded into temperate areas [16].

Phylogenetic studies reported that nodulation in *Chamaecrista* appears to be an independent event in the evolution of nitrogen-fixing symbioses with legumes [16]. Until now, the number of studies about nitrogen fixation performed with *Chamaecrista* is low and has focused either on the occurrence of nodules or included very few isolates in broader studies [18, 19, 20, 21, 22]. There are relatively few reports on the rhizobial microsymbionts associated with this legume, and the symbionts have been identified as belonging to the genus *Bradyrhizobium* [23, 24, 25].

By using a polyphasic approach to elucidate the diversity of the rhizobial symbionts nodulating *Chamaecrista*, this study investigated eleven strains isolated from root nodules of *Chamaecrista fasciculata*, a caesalpinoid species native to the Midwest states of Missouri and Nebraska in the United States of America (USA).

Isolation and Ecology

Following microbiological procedures and media formulations described for rhizobia isolation [34], eleven strains were isolated by Dr. Padma Somasegaran from root nodules of *C. fasciculata*. Commonly known as showy partridge pea in the USA, *C. fasciculata* is an annual nodulating species in the subfamily Caesalpnoideae native to the Midwest, eastern and southern states of the USA, with potential for cultivation as a multi-functional species [35]. Root nodules for rhizobia isolation were collected in different sites between the states of Missouri and Nebraska (Table 1). Purified cultures were maintained on modified Yeast Mannitol Agar (YMA) medium containing Congo red (0.00125 %) [36] at 4 °C as source cultures. For long-term storage, stocks were prepared on modified YM with 30 % glycerol (v/v) in sterile cryogenic vials and stored at -80 °C and at -150 °C, and

were also lyophilized. The CNPSo (Centro Nacional de Pesquisa de Soja) strains are deposited at the “Diazotrophic and Plant Growth Promoting Bacteria Culture Collection of Embrapa Soja” (WFCC Collection #1213, WDCM Collection #1054), in Londrina, State of Paraná, Brazil, and in others culture collections of the USA (United States Department of Agriculture, USDA collection) and Uruguay (Instituto Nacional de Investigación Agropecuária, U collection), all associated with the World Federation Culture Collection (WFCC).

Phylogeny

Genomic DNA extractions were performed using the DNeasy blood and tissue kit (Quiagen). The 16S rRNA gene sequences, the partial sequences of the *glnII*, *recA*, *gyrB*, *rpoB*, *nodC* and *nifH* genes, and of the 16S–23S rRNA intergenic transcribed spacer (ITS) of the CNPSo strains were amplified using the primers and conditions indicated in previous studies [37, 38] and shown in Table S1 (available in the online Supplementary Material), and were sequenced using an ABI3500xL sequence analyser (Applied Biosystems). Gene sequences of other *Bradyrhizobium* strains presented in this study were retrieved from the GenBank database, and the accession numbers of all strains used in this study are shown in the phylograms and in Table S2. Sequences were aligned and edited with the MEGA 6 software [39] using the MUSCLE algorithm [40]. The best-fit evolutionary model selection was performed with the software JMODELTEST 2 [41]. Maximum likelihood (ML) analyses were conducted using the software GARLI 2.0 [42] and Neighbour-Joining (NJ) analyses were generated using the MEGA 6 software [39]. To test the support of the nodes in the ML and NJ analysis, 1,000 bootstrap [43] pseudoreplicates were performed.

The phylogenetic analysis of partial 16S rRNA gene sequences was performed with the Tamura Nei [44] phylogenetic model with gamma-distributed invariant sites (G+I). CNPSo strains were clustered in the *Bradyrhizobium japonicum* superclade, and their closest neighbours were *Bradyrhizobium liaoningense* LMG 18230^T, *Bradyrhizobium guangxiense* CCBAU 53363^T, *Bradyrhizobium americanum* CMVU44^T, *Bradyrhizobium centrosemae* A9^T, *Bradyrhizobium daqingense* CCBAU 15774^T and *Bradyrhizobium nanningense* CCBAU 53390^T (Fig. S1). Moreover, the 16S rRNA phylogeny was also performed considering the complete aligned sequences retrieved from *Bradyrhizobium* type strains and CNPSo 3426^T and CNPSo 3424 genomes, confirming the clustering within the *B. japonicum* superclade (Fig. 1).

Nucleotide identities (NI) were calculated with BioEdit [45] for the partial and complete 16S rRNA gene sequences, ITS region and with four partial and ten complete single and concatenated housekeeping genes (Table 2 and Table S3). The NI of partial 16S rRNA gene sequences of the CNPSo strains were above 99.8 % of identity to each other and showed 95.3-99.8 % similarity with other described *Bradyrhizobium* species (Table 2). However, despite the 16S rRNA gene being considered as reference in phylogenetic analyses, it has been shown that the proposed threshold of 98.65 % to species delineation is not adequate to the highly conservative *Bradyrhizobium* genus [27, 37, 46]. Therefore, other genomic regions were amplified with the aim to improve our phylogenetic information.

In *Bradyrhizobium*, the higher variability of the 16S–23S rDNA intergenic spacer region (ITS) allows a better definition of species than the 16S rRNA [27, 37, 46, 47]; therefore, we proceeded with this analysis. The evolutionary distances were calculated according to NJ method and based on a matrix of uncorrected distances [27, 37, 47]. The resulting NJ tree and the accession numbers are shown in Fig. 2. Sequence analysis of the ITS region

indicated 93.7-100 % similarity between the CNPSo strains, with high bootstrap support (99 %), and with similarity lower than 89.9 % in relation to the other species of the genus *Bradyrhizobium* (Table 2). Interestingly, the results highlighted high genetic diversity among the CNPSo strains, in some cases showing lower similarity than suggested to species delimitation (95.5 %) [47].

In addition to the phylogenetics approaches of ribosomal genomic regions, several authors have confirmed the multilocus sequence analysis (MLSA) as an efficient and reliable tool for providing information on the phylogenetic relationships, once it contributes to mitigate effects of gene recombination and horizontal gene transfer [5, 6, 7, 27, 46]. Partial sequences of the *glnII* (505 bp), *gyrB* (562 bp), *recA* (360 bp) and *rpoB* (369 bp) genes were obtained and analysed individually (Figs. S2 to S5) and after concatenation of the four genes (Fig. 3). The accession numbers of all genes are shown in Table S2. The consensus sequence consisted of 1,796 bp. The best-fit evolutionary model found was the General Time Reversible [48] with invariant sites and a gamma rate distribution (I+G). Similar phylogeny relationships were observed in the MLSA when compared to the 16S rRNA phylogeny (Fig. 1 and Figs. S1 to S5), and higher bootstrap values were obtained with the concatenated sequences (Fig. 3). A robust clade (bootstrap of 97 %) comprising the eleven CNPSo strains was distinctly separated from other *Bradyrhizobium* species. In addition, it was possible to endorse with robust bootstrap support high phylogenetic diversity between the CNPSo strains. Therefore, our results confirm the reliability of using the MLSA approach to identify rhizobial strains potentially representative of new species, surpassing the limitations of the 16S rRNA gene.

In order to obtain more information about the CNPSo strains, we performed a phylogenetic analysis with the complete sequences of the *aptD* (1,373 bp), *dnaK* (1,795

bp), *gap* (992 bp), *glnII* (985 bp), *gltA* (1,283 bp), *gyrB* (2,405 bp), *pnp* (2,120 bp), *recA* (1,076 bp), *rpoB* (4,125 bp) and *thrC* (1,422 bp) genes, which were retrieved from type species of *Bradyrhizobium* with genomes available in GenBank database (Table S2). For the analysis, CNPSo 3426^T and CNPSo 3424 were included as two representatives of the proposed new species. The sequences were analyzed after concatenation of the ten genes, resulting in 17,576 bp. The results obtained suggest that *Bradyrhizobium daqingense* CCBAU 15774^T represents the closest taxon (Fig. 4). Also, with the aim of extending the phylogenetic comparisons between the new species and the described *Bradyrhizobium* species, we obtained the complete sequences of the *glnII* and *recA* genes, available for all *Bradyrhizobium* representatives, and performed a concatenated analysis. Again, the CNPSo strains were grouped in a different cluster from all other type strains (Fig. S6).

Considering the four concatenated genes, the CNPSo strains shared more than 97.6 % NI among each other and ranged from 87.5 % to 95.2 % of NI with other *Bradyrhizobium* species (Table 2). Based on the concatenated analysis of the ten genes, the CNPSo strains shared 97.8 % among each other and the closest species were *B. japonicum* USDA 6^T (96.9%), *B. diazoefficiens* USDA 110^T (96.9%), *B. huanghuaihaiense* CCBAU 23303^T (96.7%) and *B. daqingense* CCBAU 15774^T (95.6%) (Table S3.). Considering the suggested species circumscription threshold of 97 % recently proposed for *Bradyrhizobium* species delineation in the analysis of five housekeeping genes sequences [49], the NI values obtained corroborate with the classification of the CNPSo strains as a new species. However, it is worth noting that the similarity values shared with the described *Bradyrhizobium* species varied in the analyses of four (Table 2) and ten (Table S3) concatenated genes. Thus, considering the great increase of information provided by the addition of six genes, the value of species delimitation should be revised, becoming more rigorous.

To obtain information about the evolutionary relationships of nitrogen-fixation and nodulation genes, sequences of the *nifH* and *nodC* genes were obtained using the primers and amplification conditions previously described [50] and shown in Table S1, or were retrieved from the GenBank database. Accession numbers are shown in the phylograms and in Table S2. In the *nifH* phylogeny, CNPSo strains formed a cluster with *B. arachidis*, *B. kavangense*, *B. forestalis* and *B. vignae* (Fig. S7). In the *nodC* phylogeny, CNPSo strains clustered in a separate group from other species of the genus *Bradyrhizobium*, with 87 % bootstrap support (Fig. 5), and again, *B. arachidis* and *B. forestalis* were the most closely related species. These results suggest that the novel CNPSo strains might have obtained their nodulation and nitrogen-fixation genes from other *Bradyrhizobium* species by horizontal gene transfer, as reported for other species of *Bradyrhizobium* [50]. Indeed, previous studies by our group have highlighted high rates of horizontal transfer of the symbiotic islands occurring in this genus, where *nifH* and *nod* genes are flanked by insertion sequence elements [51, 52], that may be intra- and interspecies transferred [51, 52, 53, 54]. Therefore, our results from this study points out distinct evolutionary histories between the nodulation and nitrogen fixation genes and the ribosomal and housekeeping genes.

Genome Features

For genome analysis, total DNA was used to build a library using the Nextera XT kit, according to the manufacturer's instructions. The library processing was realized on the MiSeq platform (Illumina) at Embrapa Soja. Drafts of the whole genomes of strain CNPSo 3426^T (SPQS00000000, Biosample SAMN11254746) and CNPSo 3424 (SPQU00000000, Biosample SAMN11254751) were assembled with the A5-miseq pipeline (*de novo*) v.20140604. The sequencing resulted in genome coverage of 147-fold

and 81-fold assembled in 144 and 145 contigs, respectively. N_{50} values were calculated as 304,001 bp for CNPSo 3426^T and 328,664 bp for CNPSo 3424, based on both RAST v.2.0 [55] and QUAST v.2.0 [56]. The genomes of CNPSo 3426^T and CNPSo 3424 were estimated at 8,294,086 bp and 8,325,748 bp, respectively, confirmed with both RAST v.2.0 [55] and QUAST v.2.0 [56], using default parameters. *B. daqingense* CCBAU 15774^T (JGI Project Id: 1041378 - JGI), *B. liaoningense* LMG 18230^T (Project ID: 1052895 - JGI), *B. stylosanthis* BR 446^T (LVEM00000000.1), *B. neotropicae* BR 10247^T (NZ_LFJC01000003.1), *B. ottawaense* OO99^T (CP029425.1) and *B. shewense* ERR11^T (NZ_FMAI01000001.1) were the type strains chosen for comparison, because they were the closest species based on the MLSA (Figs. 3 and 4) analyses; based on the NI values (Table 2 and Table S3) the genomes of *B. guangxiense* CCBAU 53363^T (NZ_CP022220.1) and *B. huanghuaihaiense* CCBAU 23303^T (SAMN02927902) were also included.

The replacement of DNA–DNA hybridization (DDH) as “the gold standard” in prokaryotic taxonomy for genome-sequence data based on similarity has been increasingly used and proposed by several authors [6, 27, 57, 58, 59, 60]. Average nucleotide identity (ANI) and digital DDH (dDDH) have been used, and currently many tools are easily available in online platforms [61]. The suggested and generally accepted species delimitation for ANI and dDDH values are 95-96 % and 70 %, respectively, corresponding to the 70 % DNA-DNA hybridization (DDH) threshold [59, 60]. ANI and dDDH were estimated for the genomes of CNPSo 3426^T and CNPSo 3424 and eight closely related *Bradyrhizobium* species with available genomes (Table S2). ANI values were calculated using ANI calculator [55]. Despite showing diversity in the ITS (Fig. 2) and MLSA (Fig. 3 and 4) phylogenetic analyses, strains CNPSo 3426^T and CNPSo 3424 showed ANI values above the species boundary when compared to each other (97.38 %)

and below the threshold when compared to all other type strains (lower than 89.7 %), supporting the indicative that the CNPSo strains represent a novel species (Table 3). The dDDH values represent a pairwise genome-sequence comparison *in silico* between two genome sequences [59, 62]. All pairwise values were determinate by GGDC v2.1 using the recommended 'Formula 2' (<http://ggdc.dsmz.de/distcalc2.php>) [55]. dDDH values were obtained for the CNPSo 3426^T, CNPSo 3424 and closest *Bradyrhizobium* type species and among them, and the values obtained with the closest species were lower than 40 % (Table 3), giving support to the finding that the CNPSo strains represent a new species. However, it is interesting to highlight the low value of 76.8 % found in the comparison of CNPSo 3426^T and CNPSo 3424, suggesting high genome diversity of the strains belonging to the novel species (Table 3).

The G+C content (mol%) of DNA was determined using the genome contigs concatenated and the proportions of G+C were calculated with RAST v.2.0 [55] and QUAST v.2.0 [56]. The DNA G+C content for strains CNPSo 3426^T and CNPSo 3424 were 63.85 mol% and 63.93 mol%, respectively, as expected for species of the genus *Bradyrhizobium* [63].

Fingerprinting analysis was performed by BOX-PCR, using BOX-A1R primer (Table S1), as previously described [64]. The profiles were compared with four *Bradyrhizobium* type strains more closely related to the novel species based on the MLSA (Fig. 4) and NI values (Table 2). The dendrogram was built with the Bionumerics software (Applied Mathematics, Kortrijk, Belgium, v.7.6) using the UPGMA (Unweighted Pair Group Method with Arithmetic mean) algorithm [65] and the Jaccard coefficient [66] with 1 % of tolerance. The analysis revealed similarity of 40 % among CNPSo strains, confirming high genetic diversity among the strains, as well as with the most related *Bradyrhizobium* species (Fig. S8).

It is important to comment that despite the high genetic diversity observed between the CNPSo strains in the analyses of the ITS (Fig. 2), of the housekeeping genes (Figs. S2 to S5), and in the BOX-PCR profiles (Fig. S8), the values of ANI (97.38 %) and of dDDH (76.8 %) were within the threshold of species definition. This might indicate that within the conserved genus *Bradyrhizobium* [10] the threshold values established to define species deserves further investigation and probably needs to be revised.

Physiology and Chemotaxonomy

The nodulation and nitrogen fixation abilities of the eleven CNPSo strains were confirmed in the original host, but also in two other legumes, siratro (*Macroptilium atropurpureum*) and soybean (*Glycine max*) (Fig. S9). The greenhouse experiments were performed in modified Leonard jars containing sand and vermiculite, receiving sterile N-free plant nutrient solution. The presence of effective nodules was confirmed by nodules with pink color inside and healthy plants with green leaves. Abundant effective nodules were confirmed in *C. fasciculata* and siratro inoculated with strains CNPSo 3411, CNPSo 3424, CNPSo 3426^T and CNPSo 3443, with outstanding performance of CNPSo 3411, CNPSo 3426^T and CNPSo 3443. In addition, we confirmed efficient nodulation and nitrogen fixation in soybean inoculated with strains CNPSo 3426^T and CNPSo 3424. The identity of the strains re-isolated from the nodules was confirmed by sequencing analysis (data not shown).

Several phenotypic features were determined for the CNPSo strains and compared to those of *B. daqingense*, *B. liaoningense*, *B. huanghuaihaiense* and *B. stylosanthis*. All tests were performed in triplicate. Morphological tests were performed in modified-YMA with Congo red, at 28 °C, and the results related to colony morphology were evaluated after seven days of growth. The acid or alkaline reaction in modified-YMA with

bromothymol blue (0.5 %) was investigated under the same conditions. We also evaluated the salinity tolerance in modified-YMA supplemented with 1 % (w/v) NaCl and the ability to grow in agar Luria-Bertani (LB) medium [36]. Enzymatic degradation of urea was determined in modified-YMA supplemented with 2 % (w/v) urea and phenol red as indicator. Growth in pHs 4.0 and 9.0, and temperatures of 28 °C and 37 °C was verified in modified-YMA. The most relevant results are presented in Table 4.

The ability to grow with different carbon sources was determined in duplicate using the API 50CH kit (BioMérieux), according to the manufacturer's instructions and the basal medium used was YM-minus-mannitol, with bromothymol blue as indicator of the pH reaction. As shown in Table 4, the results are scored as no utilization (-), weak utilization (W) and good utilization (+). Antibiotic tolerance was evaluated using the disc diffusion method on modified-YMA plates. All tests were performed in duplicate using the following antibiotics (per disc): ampicillin (10 µg), bacitracin (0.04 U), cefuroxime (30 µg), chloramphenicol (30 µg), erythromycin (15 µg), nalidixic acid (30 µg), neomycin (30 µg), penicillin G (10U), streptomycin (10 µg) and tetracycline (30 µg). Table 4 shows the most significant results. For the antibiotic tolerance, the results are included only in the species description.

In general, the CNPSo strains analysed in this study showed similar properties in most tests. We should highlight the unusual fast growth by all 11 strains in three days at 37 °C *in vitro*, suggesting a very efficient mechanism of tolerance to high temperature. Representative strains CNPSo 3426^T and CNPSo 3424 showing high growth rate at 37 °C are included in Table 4. The strains were also able to grow under acid and basic conditions, and showed positive urease activity. Tolerance of antibiotics and acid or alkaline reaction were more variable. In the comparisons made with the related type strains, phenotypic results are in agreement with those commonly found in

Bradyrhizobium species in other studies [27]. The properties that characterized the CNPSo strains are specified in the species description. However, it is important to mention that the phenotypic data can provide only limited information, as these properties are usually encoded on the accessory genome and can easily be lost or transferred.

The fatty acid profile of strain CNPSo 3426^T was determined using the MIDI Sherlock Microbial Identification System with the TSBA6 database after growth on modified-YMA for seven days [67]. The analysis showed that the fatty acids of CNPSo 3426^T strain are composed primarily of summed feature 8 (C18:1 ω 6c/C18:1 ω 7c, 84.61 %) and, in a smaller proportion, the C16:0 (15.39 %) was also found; for strain CNPSo 3424 these values were of 82.56 % and 14.67 %, respectively, in addition to 2.77 % of C16:0. These are typical proportions described for *Bradyrhizobium* members [68]. Similar profiles were verified, but with different proportions, in the others type strains analysed, which also exhibited different varieties of fatty acids in smaller concentrations (data not shown).

Based on a complete polyphasic approach comprising phylogenetic, genotypic and phenotypic studies, the results indicate that the 11 CNPSo strains correspond to a novel species within the genus *Bradyrhizobium*, for which the name *Bradyrhizobium frederickii* sp. nov. is proposed, with CNPSo 3426^T nominated as the type strain.

Description of *Bradyrhizobium frederickii* sp. nov.

Bradyrhizobium frederickii [fre.de.rick'i.i. N.L. gen. n. *frederickii* of Frederick. Named after the late microbiologist Professor Lloyd R. Frederick (Iowa State University) whose initiatives during his tenure with USAID led to significant international development of Biological Nitrogen-Fixation technology via the University of Hawaii NifTAL Project & MIRCEN].

Cells are Gram-stain-negative rods, aerobic and do not form spores. Slightly pink colonies when grown on modified-YMA medium with Congo Red are circular, convex, opaque, translucent, showing low production of mucus with 0.57–1.25 mm in diameter, after seven days of incubation at 28 °C. CNPSo 3426^T strain presents alkaline reaction in modified-YMA containing bromothymol blue in seven days of growth. All strains grow at pH from 4.0 to 9.0, with optimum growth at pH 6.8. Abundant growth was observed after seven days of growth at 28 °C, and at 37 °C abundant growth was observed after only 3 days of incubation. Strains do not grow in LB medium or in the presence of 1 % NaCl. They are positive to urease activity. In general, CNPSo strains assimilate glycerol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-fructose, esculin iron citrate, starch, glycogen, D-lyxose, L-fucose, D-fucose, D-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium 5-ketogluconate. They grow weakly with erythritol, methyl-βD-xylopyranoside, D-galactose, D-glucose, D-mannose, L-sorbose, D-mannitol, D-melibiose, D-sucrose, D-trehalose, inulin, D-melezitose, D-raffinose, gentiobiose, D-turanose and D-tagatose, and they do not grow with D-adonitol, L-rhamnose, dulcitol, inositol, D-sorbitol, methyl-αD-mannopyranoside, N-acetylglucosamine, methyl-αD-glucopyranoside, amygdalin, arbutin, salicin, D-cellobiose, D-maltose, D-lactose, xylitol and L-arabitol. The strains show tolerance of bacitracin, penicillin G, chloramphenicol and nalidixic acid; are moderately tolerant to neomycin and are sensitive to ampicillin, cefuroxime, erythromycin, streptomycin and tetracycline.

The type strain is CNPSo 3426^T (=USDA 10052^T =U686^T =CL 20^T), isolated from effective nodules of *Chamaecrista fasciculata* in the USA; the strain is also able to nodulate other legumes such as *Macroptilium atropurpureum* and *Glycine max*. The DNA G+C content of strain CNPSo 3426^T is 63.85 mol%.

The following new sequences have been deposited in GenBank/EMBL/DDBJ: 16S rRNA of CNPSo 3426^T (MK672937), CNPSo 3421 (MK672931), CNPSo 3411 (MK672932), CNPSo 3413 (MK672933), CNPSo 3417 (MK672934), CNPSo 3423 (MK672935), CNPSo 3424 (MK672936), CNPSo 3431 (MK672938), CNPSo 3443 (MK672939), CNPSo 3446 (MK672940), CNPSo 3447 (MK672941); 16S–23S rDNA of CNPSo 3426^T (MK680950), CNPSo 3421 (MK680944), CNPSo 3411 (MK680945), CNPSo 3413 (MK680946), CNPSo 3417 (MK680947), CNPSo 3423 (MK680948), CNPSo 3424 (MK680949), CNPSo 3431 (MK680951), CNPSo 3443 (MK680952), CNPSo 3446 (MK680953), CNPSo 3447 (MK680954); *glnII* of CNPSo 3426^T (MK682688), CNPSo 3421 (MK682682), CNPSo 3411 (MK682683), CNPSo 3413 (MK682684), CNPSo 3417 (MK682685), CNPSo 3423 (MK682686), CNPSo 3424 (MK682687), CNPSo 3431 (MK682689), CNPSo 3443 (MK682690), CNPSo 3446 (MK682691), CNPSo 3447 (MK682692); *gyrB* of CNPSo 3426^T (MK682721), CNPSo 3421 (MK682715), CNPSo 3411 (MK682716), CNPSo 3413 (MK682717), CNPSo 3417 (MK682718), CNPSo 3423 (MK682719), CNPSo 3424 (MK682720), CNPSo 3431 (MK682722), CNPSo 3443 (MK682723), CNPSo 3446 (MK682724), CNPSo 3447 (MK682725); *recA* of CNPSo 3426^T (MK682710), CNPSo 3421 (MK682704), CNPSo 3411 (MK682705), CNPSo 3413 (MK682706), CNPSo 3417 (MK682707), CNPSo 3423 (MK682708), CNPSo 3424 (MK682709), CNPSo 3431 (MK682711), CNPSo 3443 (MK682712), CNPSo 3446 (MK682713), CNPSo 3447 (MK682714); *rpoB* of CNPSo 3426^T (MK682699), CNPSo 3421 (MK682693), CNPSo 3411 (MK682694), CNPSo 3413 (MK682695), CNPSo 3417 (MK682696), CNPSo 3423 (MK682697), CNPSo 3424 (MK682698), CNPSo 3431 (MK682700), CNPSo 3443 (MK682701), CNPSo 3446 (MK682702), CNPSo 3447 (MK682703); *nifH* of CNPSo 3426^T (MK682732), CNPSo 3421 (MK682726), CNPSo 3411 (MK682727), CNPSo 3413 (MK682728), CNPSo 3417 (MK682729), CNPSo 3423 (MK682730), CNPSo 3424 (MK682731), CNPSo 3431 (MK682733), CNPSo 3443 (MK682734), CNPSo 3446 (MK682735), CNPSo 3447 (MK682736); *nodC* of CNPSo 3426^T (MK682743), CNPSo 3421 (MK682737), CNPSo 3411 (MK682738), CNPSo 3413 (MK682739), CNPSo 3417 (MK682740), CNPSo 3423 (MK682741), CNPSo 3424 (MK682742), CNPSo 3431 (MK682744), CNPSo 3443 (MK682745), CNPSo 3446 (MK682746), CNPSo 3447 (MK682747).

Genome sequences of *Bradyrhizobium frederickii* CNPSo 3426^T (SPQS00000000) and of CNPSo 3424 (SPQU00000000).

Funding

Funded by INCT-Plant-Growth Promoting Microorganisms for Agricultural Sustainability and Environmental Responsibility (CNPq 465133/2014-2, Fundação Araucária-STI, CAPES), Embrapa (02.13.08.001.00.00), CNPq-Universal (400468/2016-6)

Acknowledgements

M.C. Urquiaga acknowledges an MSc fellowships from CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil - Finance Code 001) and M.S. Klepa, J.R.M. Delamuta and R.A. Ribeiro fellowships from CNPq (Brazilian National Research Council). Authors acknowledge Dr. Maria do Carmo Catanho Pereira Lyra (IPA, Brazil) and Ligia Maria Oliveira Chueire (Embrapa Soja) for help in several steps of this study, Tyson Vorderstrasse and Kelsey Goforth (BIOZYME Inc.) for their enthusiastic support in the research, Dr. Itamar S. Melo and Juliane Fontana (Embrapa Meio Ambiente, Brazil) for help with the fatty acids analyses, and Dr. Sergio Miana de Faria (Embrapa Agrobiologia, Brazil) for sending *Chamaecrista* seeds. Authors thank also Dr. Aharon Oren and Dr. Bernhard Schink for their help with the epithet designation of this species.

Ethics approval and consent to participate

The authors declare no ethical conflicts; authors declare that they have consented to participate in the manuscript and publish it.

Competing interests

The authors declare no competing interests

Abbreviations: NI, nucleotide identity; ANI, average nucleotide identity; DDH, DNA-DNA hybridization; dDDH, digital DNA-DNA hybridization; MLSA, multilocus sequence analysis.

References

1. **Ormeño-Orrillo E, Hungria M, Martínez-Romero E.** Dinitrogen-fixing prokaryotes. In: Rosenberg E (editor). *The Prokaryotes – Prokaryotic Physiology and Biochemistry*. Berlin Heidelberg: Springer-Verlag, 2013. pp. 427–451.
2. **Ahnia H, Bourebaba Y, Durán D et al.** *Bradyrhizobium algeriense* sp. nov., a novel species isolated from effective nodules of *Retama sphaerocarpa* from Northeastern Algeria. *Syst Appl Microbiol* 2018;41:333–339.
3. **Araújo J, Flores-Félix JD, Igual JM et al.** *Bradyrhizobium cajani* sp. Nov. isolated from nodules of *Cajanus cajan*. *Int J Syst Evol Microbiol* 2017;67:2236–2241.
4. **Aserse A, Woyke T, Kyrpides N, Whitman WB, Lindstrom K.** Draft genome sequences of *Bradyrhizobium shewense* sp. nov. ERR11 T and. *Stand Genomic Sci* 2017;12:1–14.
5. **Bünger W, Grönemeyer J, Sarkar A, Reinhold-Hurek. B.** *Bradyrhizobium ripae* sp. nov., a nitrogen-fixing symbiont isolated from nodules of wild legumes in Namibia. *Int J Syst Evol Microbiol* 2018;18.
6. **Helene LCF, Delamuta JRM, Ribeiro RA, Hungria M.** *Bradyrhizobium mercantei* sp. Nov., a nitrogen-fixing symbiont isolated from nodules of *Deguelia costata* (syn. *Lonchocarpus costatus*). *Int J Syst Evol Microbiol* 2017;67:1827–1834.
7. **Martins Da Costa E, Guimarães AA, Soares De Carvalho T et al.** *Bradyrhizobium forestalis* sp. nov., an efficient nitrogen-fixing bacterium isolated from nodules of forest legume species in the Amazon. *Arch Microbiol* 2018;200:743–752.
8. **De Matos GF, Zilli JE, De Araújo, JLS et al.** *Bradyrhizobium sacchari* sp. nov., a legume nodulating bacterium isolated from sugarcane roots. *Arch Microbiol* 2017; 199:1251–1258.
9. **Michel DC, Passos SR, Simões-Araujo JL et al.** *Bradyrhizobium centrolobii* and *Bradyrhizobium macuxiense* sp. nov. isolated from *Centrolobium paraense* grown in soil of Amazonia, Brazil. *Arch Microbiol* 2017;199:657–664.
10. **Hungria M, Menna P, Delamuta JRM.** *Bradyrhizobium*, the ancestor of all rhizobia: phylogeny of housekeeping and nitrogen-fixation genes. In: de Bruijn FJ (editor). *Biological Nitrogen Fixation*. John Wiley, Sons, Inc., New Jersey. 2015.pp. 191-202.

11. **Parker MA.** The spread of *Bradyrhizobium* lineages across host legume clades: from *Abarema* to *Zygia*. *Microb Ecol* 2015;69:630–640.
12. **Sprent JI, James EK.** Legume Evolution: Where Do Nodules and Mycorrhizas Fit In? *Plant Physiol* 2007;144:575–581.
13. **Allen ON, Allen EK.** The *Leguminosae*: A sourcebook of Characteristics, Uses, and Nodulation. University of Wisconsin Press/MacMillan Publishers Ltd, Madison/London, U.S.A. /U.K 1981
14. **Faria SM, Franco AA, Jesus R et al.** New nodulating legume trees from south-east Brazil. *New Phytol* 1984;98:317–328.
15. **Naisbitt T, James EK, Sprent JI.** The evolutionary significance of the legume genus *Chamaecrista*, as determined by nodule structure. *New Phytol* 1992;122:487–492.
16. **Sprent JI, Ardley J, James EK.** Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytol* 2017;215:40–56.
17. **Delaux P, Radhakrishnan G, Oldroyd G.** Tracing the evolutionary path to nitrogen-fixing crops. *Curr Opin Plant Biol* 2015;26:95–99.
18. **Diabate M, Munive A, De Faria SM, Ba A, Dreyfus B, Galiana A.** Occurrence of nodulation in unexplored leguminous trees native to the West African tropical rainforest and inoculation response of native species useful in reforestation. *New Phytol* 2005;166:231-239.
19. **de Faria SM, Diedhiou AG, de Lima HC et al.** Evaluating the nodulation status of leguminous species from the Amazonian forest of Brazil. *J Exp Bot* 2010;61:3119- 3127.
20. **Gehlot HS, Panwar D, Tak N, Tak A, Sankhla IS et al.** Nodulation of legumes from the Thar desert of India and molecular characterization of their rhizobia. *Plant Soil* 2012;357:227–243.
21. **Parker MA, Kennedy DA.** Diversity and relationships of bradyrhizobia from legumes native to eastern North America. *Can J Microbiol* 2006;52:1148-1157.
22. **Moreira F, Haukka K, Young JPW.** Biodiversity of rhizobia isolated from a wide range of forest legumes in Brazil. *Mol Ecol* 1998;7:889–895.
23. **Parker MA, Rousteau A.** Molecular Phylogenetics and Evolution Mosaic origins of *Bradyrhizobium* legume symbionts on the Caribbean island of Guadeloupe. *Mol Phylogenet Evol* 2014;77:110–115.

24. Santos JMF, Casaes Alves PA, Silva VC et al. Diverse genotypes of *Bradyrhizobium* nodulate herbaceous *Chamaecrista* (Moench) (Fabaceae, Caesalpinioideae) species in Brazil. *Syst Appl Microbiol* 2017;40:69–79.
25. Rathi S, Tak N, Bissa G, Chouhan B et al. Selection of *Bradyrhizobium* or Ensifer symbionts by the native Indian caesalpinoid legume *Chamaecrista pumila* depends on soil pH and other edaphic and climatic factors. *FEMS Microbiol Ecology* 2018;94:1-17.
26. Xu LM, Ge C, Cui Z et al. *Bradyrhizobium liaoningense* sp. nov., isolated from the root nodules of soybeans. *Int J Syst Bacteriol* 1995;45:706–711.
27. Delamuta J, Ribeiro R, Araújo J et al. *Bradyrhizobium stylosanthis* sp. nov., comprising nitrogen-fixing symbionts isolated from nodules of the tropical forage legume *Stylosanthes* spp. *Int J Syst Evol Microbiol* 2016;66:3078–3087.
28. Zhang YM, Li Y, Chen WF et al. *Bradyrhizobium huanghuaihaiense* sp. nov., an effective symbiotic bacterium isolated from soybean (*Glycine max* L.) nodules. *Int J Syst Evol Microbiol* 2012;62:1951–1957.
29. Yu X, Cloutier S, Tambong JT, Bromfield ESP. *Bradyrhizobium ottawaense* sp. nov., a symbiotic nitrogen fixing bacterium from root nodules of soybeans in Canada. *Int J Syst Evol Microbiol* 2014;64:3202–3207.
30. Aserse A, Woyke T, Kyrpides N et al. Draft genome sequences of *Bradyrhizobium shewense* sp. nov. ERR11^T and *Bradyrhizobium yuanmingense* CCBAU 10071^T. *Stand Genomic Sci* 2017;12:1–14.
31. Wang JY, Wang R, Zhang Y et al. *Bradyrhizobium daqingense* sp. nov., isolated from soybean nodules. *Int J Syst Evol Microbiol* 2013;63:616–624.
32. Zilli JE, Baraúna A, Silva K et al. *Bradyrhizobium neotropiale* sp. nov., isolated from effective nodules of *Centrolobium paraense*. *Int J Syst Evol Microbiol* 2014;64:3950–3957.
33. Li YH, Wang R, Zhang X et al. *Bradyrhizobium guangdongense* sp. nov. and *Bradyrhizobium guangxiense* sp. nov., isolated from effective nodules of peanut. *Int J Syst Evol Microbiol* 2015;65:4655–4661.
34. Somasegaran P, Hoben H. Methods in Legume-Rhizobium Technology. In: *Handbook for Rhizobia*. Springer Verlag, 1985;450.
35. Morris JB. Showy partridge pea (*Chamaecrista fasciculata* (Michx.) Greene) with potential for cultivation as a multi-functional species in the United States. *Genet Resour Crop Evol* 2012;59:1577-1581.

36. **Hungria M, O'hara G. W, Zilli JE et al.** Isolation and growth of rhizobia. *Working with rhizobia*. 2016;315, 2016.
37. **Delamuta J, Ribeiro RA, Ormeño-orriillo E et al.** Polyphasic evidence supporting the reclassification of *Bradyrhizobium japonicum* group Ia strains as *Bradyrhizobium diazoefficiens* sp. nov. *Int J Syst Evol Microbiol* 2013;63:3342–3351.
38. **Laguerre G, Nour SM, Sanjuan J, Drouin, P.** Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. *Microbiology* 2001;147:981–993.
39. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S.** MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 2013;30:2725–2729.
40. **Edgar RC.** MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 2004;32:1792–1797.
41. **Darriba D, Taboada GL, Doallo R, Posada D.** JModelTest 2: More models, new heuristics and parallel computing. *Nature Methods* 2012;9:772, 2012.
42. **Zwickl DJ.** Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. University of Texas Electronic Theses and Dissertations 2006.
43. **Felsenstein J.** Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 1985;39:783, 1985.
44. **Tamura K, Nei M.** Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 1993;10.
45. **Hall T.** BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 1999;41:95–98.
46. **Menna P, Barcellos FG, Hungria M.** Phylogeny and taxonomy of a diverse collection of *Bradyrhizobium* strains based on multilocus sequence analysis of the 16S rRNA gene, ITS region and *glnII*, *recA*, *atpD* and *dnaK* genes. *Int J Syst Evol Microbiol* 2009;59:2934–2950.
47. **Willems A, Coopman R, Gillis M.** Comparison of sequence analysis of 16S – 23S rDNA spacer regions, AFLP analysis and DNA – DNA hybridizations in *Bradyrhizobium*. *Int J Syst Evol Microbiol* 2001;51: 623-632.
48. **Tavaré S.** Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures Math Life Sci* 1986;17:57–86.

- 49. Durán D, Rey L, Mayo J et al.** *Bradyrhizobium paxllaeri* sp. nov. and *Bradyrhizobium icense* sp. nov., nitrogen-fixing rhizobial symbionts of Lima bean (*Phaseolus lunatus* L.) in Peru. *Int J Syst Evol Microbiol* 2014;64:2072–2078.
- 50. 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. *Int J Syst Evol Microbiol* 2011;61:3052–3067.
- 51. Barcellos FG, Stefa J, Batista S, Hungria M.** Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs *Sinorhizobium* (Ensifer) *fredii* and *Bradyrhizobium elkanii* in a Brazilian savannah soil. *Appl Environ Microbiol* 2007;73: 2635–2643.
- 52. Batista J, Hungria M, Barcellos FG.** Variability in *Bradyrhizobium japonicum* and *B. elkanii* seven years after introduction of both the exotic microsymbiont and the soybean host in a cerrados soil. *Microbial Ecol* 2007;53:270–284.
- 53. Vinuesa P, Silva C, Lorite MJ et al.** Molecular systematics of rhizobia based on maximum likelihood and Bayesian phylogenies inferred from *rrs*, *atpD*, *recA* and *nifH* sequences, and their use in the classification of *Sesbania* microsymbionts from Venezuelan wetlands. *Syst Appl Microbiol* 2005;28:702–716.
- 54. Remigi P, Zhu J, Young JPW et al.** Symbiosis within symbiosis: Evolving nitrogen-fixing legume symbionts. *Trends Microbiol* 2016;24:63–75.
- 55. Rodriguez-R LM, Konstantinidis KT.** The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes microbial genomes and metagenomes. *PeerJ Prepr* 2016;4:1-15.
- 56. Gurevich A, Saveliev V, Vyahhi N, Tesler G.** QUASt: Quality assessment tool for genome assemblies. *Bioinformatics* 2013;29:1072–1075.
- 57. Goris J, Konstantinidis KT, Klappenbach JA et al.** DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 2007;57:81–91.
- 58. Konstantinidis KT, Tiedje JM.** Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci USA* 2005;102:2567–2572.
- 59. Meier-Kolthoff J, Auch A, Klenk H, Goker M.** Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:1–14.chu
- 60. Richter M, Rossello-Mora R.** Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 2009;106:19126–19131.

61. **Chun J, Oren A, Ventosa A et al.** Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2018;68:461–466.
62. **Meier-Kolthoff JP, Klenk H, Go M.** Taxonomic use of DNA G+C content and DNA–DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 2014;64:352–356.
63. **Xu LM, Ge C, Cui Z, Li J, Fan H.** *Bradyrhizobium liaoningense* sp. nov., isolated from the root nodules of soybeans. *Int J Syst Bacteriol* 1995;45:706–711.
64. **Kaschuk G, Hungria M, Andrade DS, Campo RJ.** Genetic diversity of rhizobia associated with common bean (*Phaseolus vulgaris* L.) grown under no-tillage and conventional systems in Southern Brazil. *Appl Soil Ecol* 2006;32:210–220.
65. **Sneath PHA, Sokal RR.** *Numerical taxonomy: the principles and practice of numerical classification*. San Francisco: Freeman 1973;573.
66. **Jaccard P.** The distribution of flora in the alpine zone. *New Phytol* 1912;11:37.
67. **MIDI.** *Sherlock Microbial Identification System Operating Manual*, version 4.0. Newark: MIDI, Inc;2001.m
68. **Tighe SW, Lajudie P, Dipietro K, Lindstro K.** Analysis of cellular fatty acids and phenotypic relationships of *Agrobacterium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium* species using the Sherlock Microbial Identification System. *Int J Syst Evol Microbiol* 2000;50:787–801.

Figures Legends

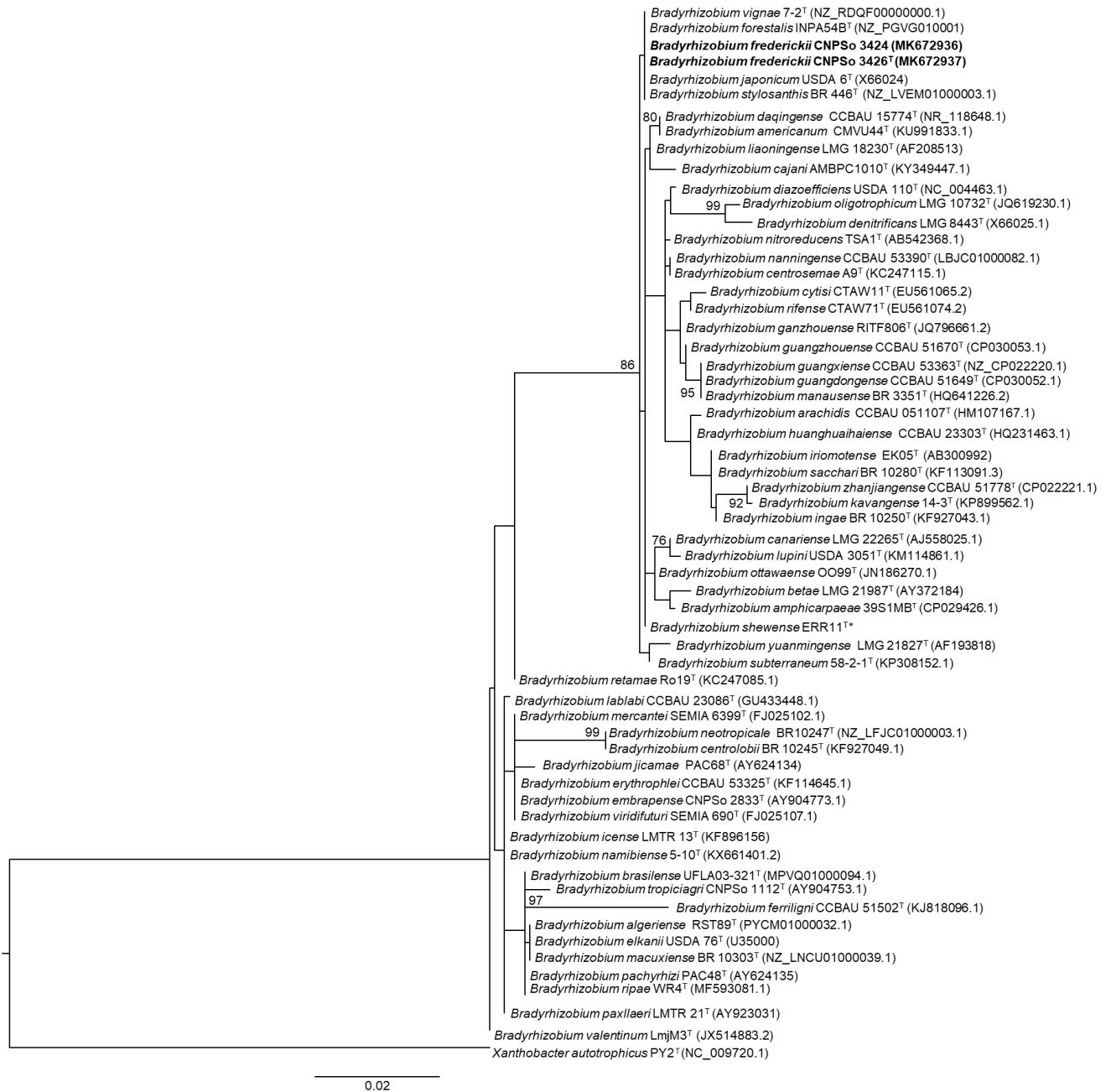


Fig. 1. Maximum Likelihood (ML) phylogeny based on the 16S rRNA aligned sequences (1,312 bp), showing in bold the phylogenetic position of the two representative strains of the novel species *Bradyrhizobium frederickii*. GenBank accession numbers are indicated in parentheses. Bootstrap values greater than 70 % are indicated at the nodes. *Xanthobacter autotrophicus* Py2 was used as outgroup. The scale bar indicates 2 substitutions per 100 nucleotide positions.

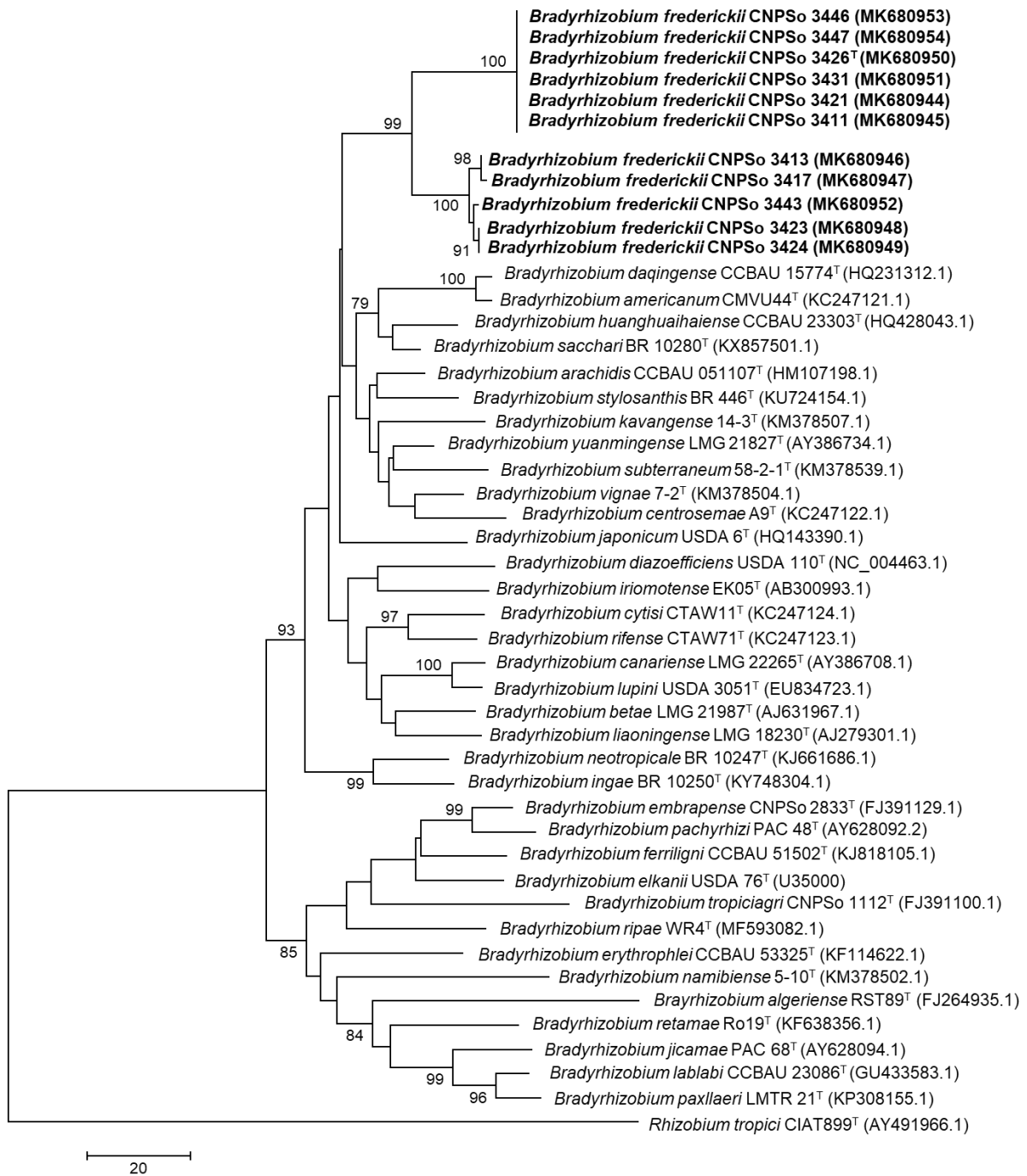


Fig. 2. Neighbor Joining (NJ) phylogeny based on the 16S-23S rRNA intergenic spacer (ITS) aligned sequences (921 bp), showing the relationships of the novel species *Bradyrhizobium frederickii* (in bold) with related species of the genus. GenBank accession numbers are indicated in parentheses. Bootstrap values greater than 70 % are indicated at the nodes. *Rhizobium tropici* CIAT 899 was used as outgroup. The scale bar indicates 20 substitutions per 100 nucleotide positions.

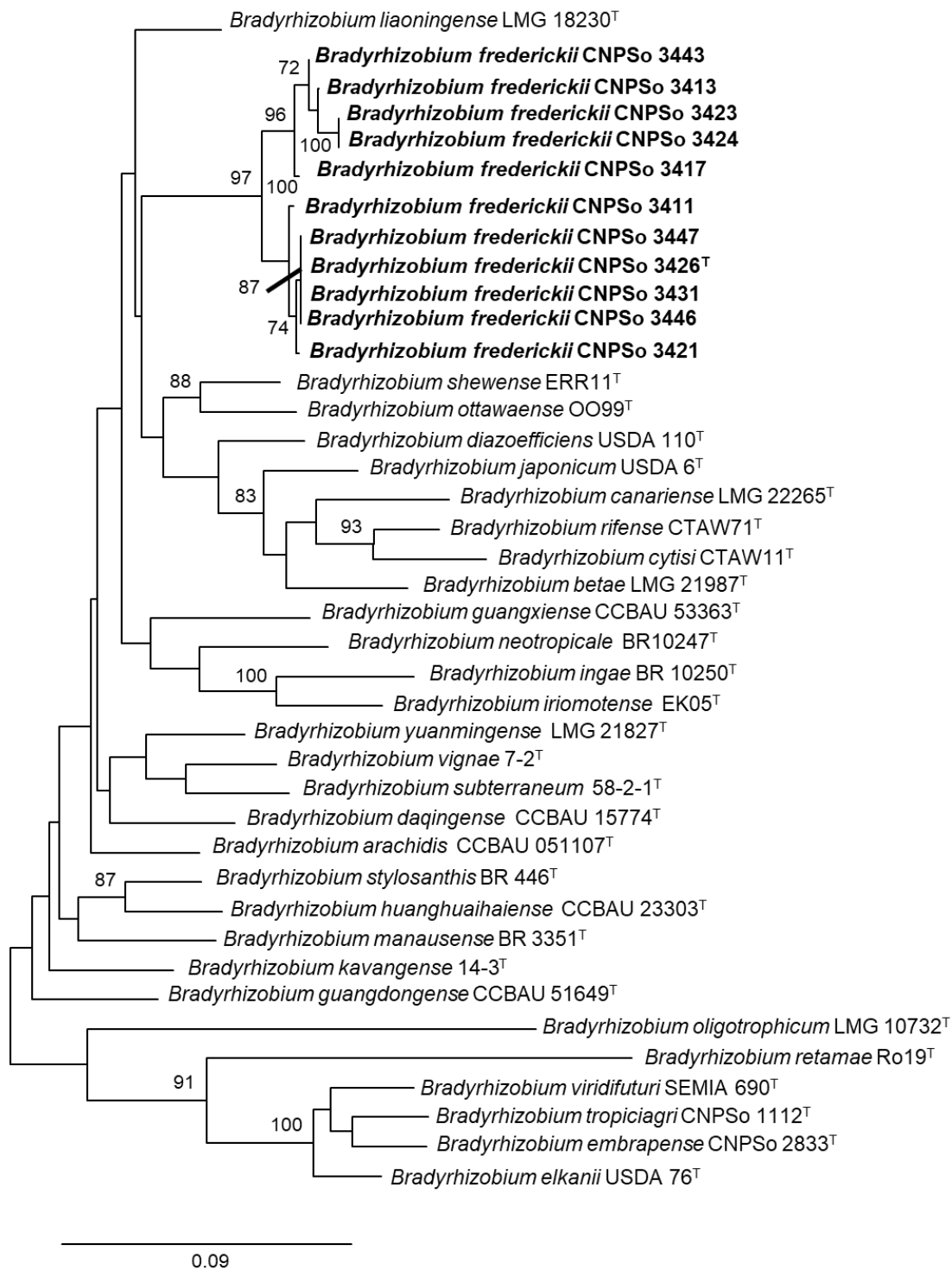


Fig. 3. Maximum Likelihood (ML) phylogeny based on the concatenated aligned sequences of the *glnII* (505 bp), *gyrB* (562 bp), *recA* (360 bp) and *rpoB* (369 bp) genes showing the relationships of the novel species *Bradyrhizobium frederickii* (in bold) with related *Bradyrhizobium* species. GenBank accession numbers are indicated in Table S2. Bootstrap values greater than 70 % are indicated at the nodes. The scale bar indicates 9 substitutions per 100 nucleotide positions.

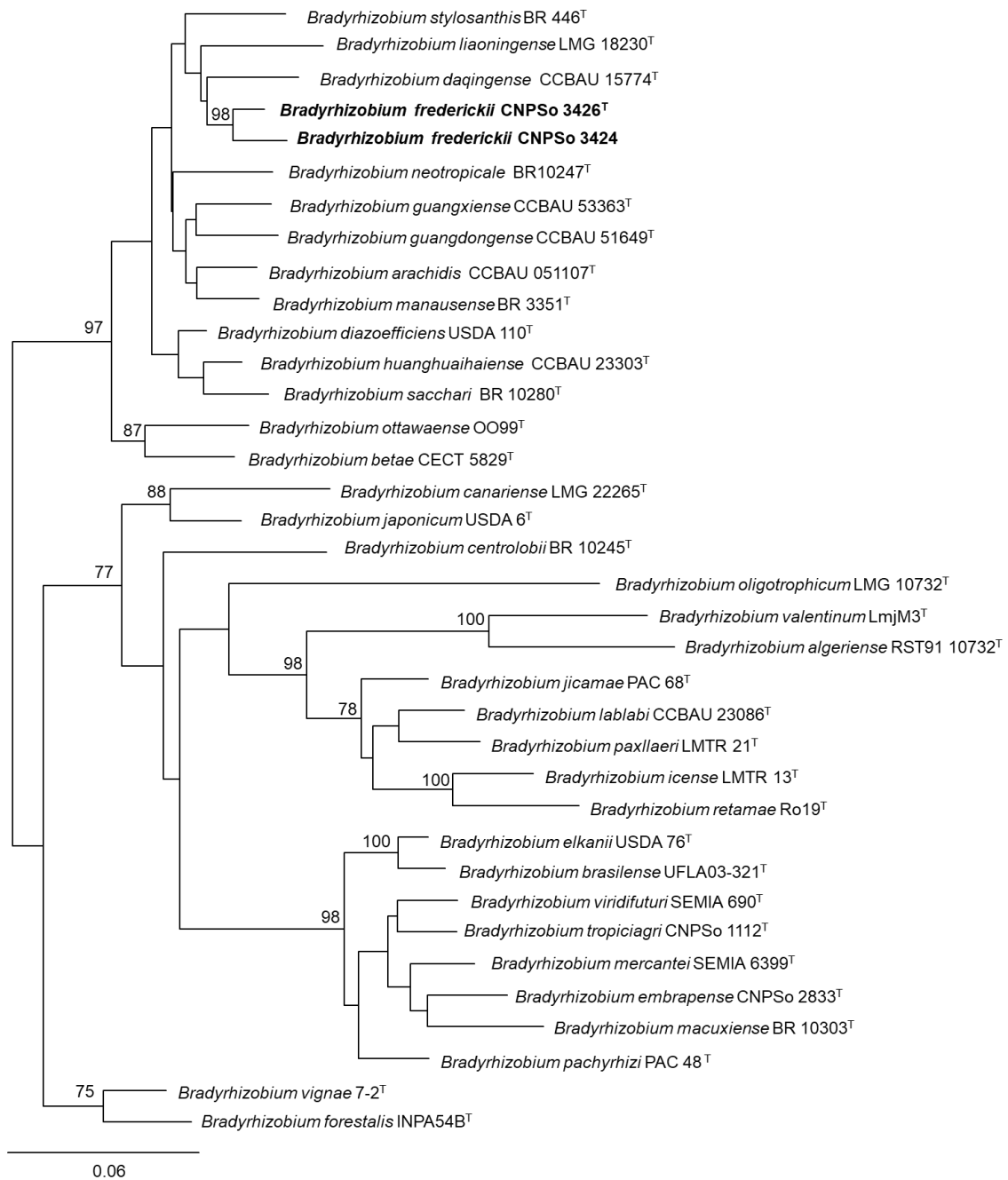


Fig. 4. Maximum Likelihood (ML) phylogeny based on the concatenated aligned sequences of the *aptD* (1,373 pb), *dnaK* (1,795 pb), *gap* (992 pb), *glnII* (985 pb), *gltA* (1,283 pb), *gyrB* (2,405 pb), *pnp* (2,120 pb), *recA* (1,076 pb), *rpoB* (4,125 pb) and *thrC* (1,422) genes showing the relationships of two representative strains of the novel species *Bradyrhizobium frederickii* (in bold) with related *Bradyrhizobium* species. GenBank accession numbers are indicated in Table S2. Bootstrap values greater than 70 % are indicated at the nodes. The scale bar indicates 6 substitutions per 100 nucleotide positions.

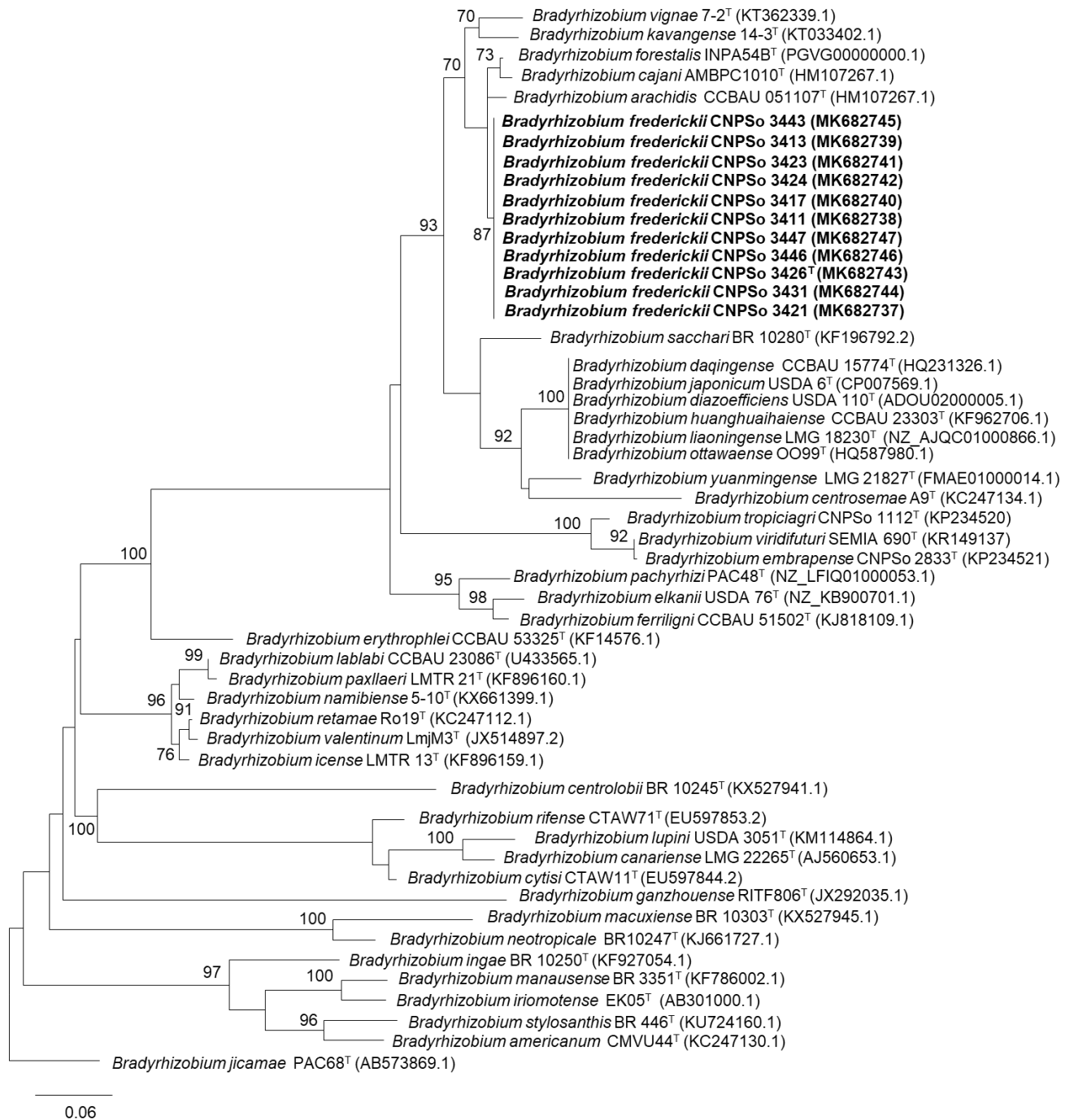


Fig. 5. Maximum Likelihood (ML) phylogenetic tree based on the *nodC* aligned sequences (426 bp) showing the relationships of the novel species *Bradyrhizobium frederickii* (in bold) with other *Bradyrhizobium* species. GenBank accession numbers are indicated in parentheses. Bootstrap values greater than 70 % are indicated at the nodes. The scale bar indicates 6 substitutions per 100 nucleotide positions.

Table 1. Strains used in this study.

Strain	Other designations	Original host legume	Nodule sampling site for isolation	Reference
<i>B. frederickii</i> CNPSo 3426 ^T	USDA 10052 ^T , U686 ^T , CL 20 ^T	<i>Chamaecrista fasciculata</i>	Along Hwy. I-29, Oregon, MO, USA (130 Km North of Kansas City, MO) 39.98715 degrees N, -95.143 degrees E	This study
<i>B. frederickii</i> CNPSo 3424	USDA 10053, U685, CL 17	<i>Chamaecrista fasciculata</i>	Along Hwy. I-29, Oregon, MO, USA (130 Km North of Kansas City, MO) 39.98715 degrees N, -95.143 degrees E	This study
<i>B. frederickii</i> CNPSo 3421	CL 4	<i>Chamaecrista fasciculata</i>	Along Riverside Rd. St. Joseph, MO, USA (89 Km North of Kansas City, MO) 39.7666 degrees N, -94.85607 degrees E	This study
<i>B. frederickii</i> CNPSo 3411	CL 6	<i>Chamaecrista fasciculata</i>	Along Riverside Rd. St. Joseph, MO, USA (89 Km North of Kansas City, MO) 39.7666 degrees N, -94.85607 degrees E	This study
<i>B. frederickii</i> CNPSo 3413	CL 8	<i>Chamaecrista fasciculata</i>	Along Riverside Rd. St. Joseph, MO, USA (89 Km North of Kansas City, MO) 39.7666 degrees N, -94.85607 degrees E	This study
<i>B. frederickii</i> CNPSo 3417	CL 13	<i>Chamaecrista fasciculata</i>	Along Hwy. I-29, Bellavue, NE, USA (284 Km North of Kansas City, MO) 41.13647 degrees N, -95.89249 degrees E	This study
<i>B. frederickii</i> CNPSo 3423	CL 15	<i>Chamaecrista fasciculata</i>	Along Hwy. I-29, Bellavue, NE, USA (284 Km North of Kansas City, MO) 41.13647 degrees N, -95.89249 degrees E	This study
<i>B. frederickii</i> CNPSo 3431	CL 32	<i>Chamaecrista fasciculata</i>	Along 192 nd. Street Omaha, NE, USA (310 Km North of Kansas City, MO) 41.26067 degrees N, -95.94056 degrees E	This study
<i>B. frederickii</i> CNPSo 3443	CL 35	<i>Chamaecrista fasciculata</i>	Along 192 nd. Street Omaha, NE, USA (310 Km North of Kansas City, MO) 41.26067 degrees N, -95.94056 degrees E	This study
<i>B. frederickii</i> CNPSo 3446	CL 38	<i>Chamaecrista fasciculata</i>	Along 192 nd. Street Omaha, NE, USA (310 Km North of Kansas City, MO) 41.26067 degrees N, -95.94056 degrees E	This study
<i>B. frederickii</i> CNPSo 3447	CL39	<i>Chamaecrista fasciculata</i>	Along Hwy. I-29, Savannah, MO, USA (106 Km North of Kansas City, MO) 39.94668 degrees N, -94.83081 degrees E	This study

<i>B. liaoningense</i> LMG 18230 ^T	DSM 24092 ^T , CECT 4845 ^T , USDA 3622 ^T	<i>Glycine max</i>	China	Xu <i>et al.</i> [26]
<i>B. stylosanthis</i> BR 446 ^T	H-8 ^T , SEMIA 6154 ^T , CNPSo 2823 ^T	<i>Stylosanthes guyanensis</i>	Brazil	Delamuta <i>et al.</i> [27]
<i>B. huanghuaihaiense</i> CCBAU 23303 ^T	LMG 26136 ^T , CGMCC 1.10948 ^T , HAMBI 3180 ^T , CNPSo 2790 ^T	<i>Glycine max</i>	China	Zhang <i>et al.</i> [28]
<i>B. ottawaense</i> OO99 ^T	LMG 26739 ^T , HAMBI 3284 ^T	<i>Glycine max</i>	Canada	Yu <i>et al.</i> [29]
<i>B. shewense</i> ERR11 ^T	HAMBI 3532 ^T , LMG 30162 ^T	<i>Erythrina brucei</i>	Ethiopia	Aserse <i>et al.</i> [30]
<i>B. daqingense</i> CCBAU 15774 ^T	LMG 26137 ^T , HAMBI 3184 ^T , CGMCC 1.10947 ^T	<i>Glycine max</i>	China	Wang <i>et al.</i> [31]
<i>B. neotropale</i> BR 10247 ^T	HAMBI 3599 ^T	<i>Centrolobium paraense</i>	Brazil	Zilli <i>et al.</i> [32]
<i>B. guangxiense</i> CCBAU 53363 ^T	CGMCC 1.15035 ^T , LMG 28621 ^T	<i>Arachis hypogaea</i>	China	Li <i>et al.</i> [33]

Table 2. Nucleotide sequence identity percentage (% NI) between *Bradyrhizobium frederickii* CNPSo strains and other *Bradyrhizobium* type strains. Similarities based on the sequence identity of the 16S rRNA, ITS, single and concatenated housekeeping genes (*glnII*, *gyrB*, *recA* and *rpoB*).

Species	Gene							
	16S rRNA (949 bp)	16S rRNA (1,312 bp)	ITS (921 bp)	<i>glnII</i> (505 bp)	<i>gyrB</i> (562 bp)	<i>recA</i> (408 bp)	<i>rpoB</i> (369 bp)	MLSA (1,796 bp)
Within CNPSo strains*	99.8 – 100	100	93.7 – 100	96 – 100	98.3 – 100	97.7 – 100	98 – 100	97.6 – 100
Between CNPSo strains and:								
<i>B. diazoefficiens</i>	99.6 – 99.7	99.3	74.1 – 76.7	94 – 95	93.1 – 93.6	92.2 – 93.8	92.6 – 94.2	93.4 – 93.8
<i>B. stylosanthis</i>	99.1 – 99.2	100	79.3 – 80.8	95.2 – 95.6	91.8 – 92.2	95.8 – 96.9	94.5 – 95.9	94.2 – 94.7
<i>B. japonicum</i>	99.6 – 99.7	100	77.5 – 78.9	93.8 – 95	92.2 – 92.7	92.2 – 93.6	91.5 – 92.9	92.9 – 93.2
<i>B. liaoningense</i>	99.7 – 99.8	99.7	78.5 – 80.1	93.8 – 95	93.4 – 93.8	95.5 – 97.2	96.4 – 97	95 – 95.2
<i>B. canariense</i>	99.3 – 99.4	99.6	80.1 – 82	93.4 – 94.8	92 – 92.9	92.5 – 94.4	88.2 – 89.6	92.2 – 92.7
<i>B. yuanmingense</i>	99.3 – 99.4	99.4	78.4 – 80.4	93.4 – 94.2	91.1 – 91.8	94.7 – 96.1	95.3 – 96.1	93.4 – 94.1
<i>B. betae</i>	99.6 – 99.7	99.1	77.8 – 79.4	93.4 – 94.2	91.3 – 91.8	93.6 – 94.7	89.9 – 91.2	92.3 – 92.7
<i>B. iriomotense</i>	98.6 – 98.7	98.7	79.5 – 86.4	93.2 – 93.4	92 – 92.2	92.7 – 94.4	91.5 – 92.3	92.5 – 92.9
<i>B. cytisi</i>	99.2 – 99.3	99	79.1 – 80.7	92 – 93.2	91.3 – 92	91.9 – 93.6	92.9 – 94	92.4 – 92.5
<i>B. daqingense</i>	99.7 – 99.8	99.6	75.1 – 77.2	92.8 – 94.8	91.6 – 92.3	94.1 – 95.2	95.6 – 96.1	93.5 – 94.2

<i>B. huanghuaihaiense</i>	99.1 – 99.2	99.1	79.7 – 81.2	94.4 – 95.8	92.3 – 92.7	93.8 – 95.2	94 – 95.6	93.9 – 94.5
<i>B. rifense</i>	99.3 – 99.4	99.1	76.5 – 78.2	93.2 – 94.6	92 – 92.2	92.5 – 94.1	92.6 – 94	92.9 – 93.3
<i>B. arachidis</i>	98.9 – 99	98.9	78 – 79.5	93.4 – 93.8	91.4 – 91.8	95.5 – 96.3	96.1 – 97.8	94 – 94.3
<i>B. ottawaense</i>	99.6 – 99.7	99.6	NA	92.8 – 93	91.1 – 91.4	95 – 95.8	95.3 – 97	93.4 – 93.7
<i>B. ingae</i>	98.7 – 98.8	98.7	71.1 – 72.8	93.6 – 95.2	90.3 – 91.1	91.6 – 93	89.9 – 90.7	91.8 – 92.4
<i>B. manausense</i>	99.3 – 99.4	98.9	NA	93 – 93.6	91.1 – 91.6	94.7 – 95.8	93.1 – 94.8	93.1 – 93.4
<i>B. subterraneum</i>	99.5 – 99.6	99.7	75.1 – 76.8	93 – 93.6	91.3 – 91.8	92.7 – 94.4	96.1 – 97	93.3 – 93.7
<i>B. kavangense</i>	98.6 – 98.7	98.9	83.6 – 89.3	92.8 – 93.8	92 – 92.9	94.7 – 95.8	93.1 – 94.8	93.3 – 93.8
<i>B. guangdongense</i>	99.4 – 99.5	98.9	NA	92 – 93.6	92.7 – 93.2	93.3 – 94.4	94.5 – 95.6	93.3 – 93.8
<i>B. guangxiense</i>	99.7 – 99.8	98.9	NA	93.2 – 94.6	91.3 – 92.2	94.7 – 96.1	94.2 – 95	93.4 – 94.1
<i>B. shewense</i>	99.6 – 99.7	99.8	NA	93.8 – 94.2	86.2 – 86.8	94.7 – 96.1	94.8 – 96.4	93.8 – 94.1
<i>B. forestalis</i>	97 – 99.1	100	NA	93 – 93.6	91.6 – 92.2	95.5 – 96.9	NA	NA
<i>B. sacchari</i>	97 – 98.8	98.8	78.8 – 80.1	93.8 – 94.2	NA	94.1 – 95.2	NA	NA
<i>B. lupini</i>	99.3 – 99.4	99.4	78.1 – 79.8	92.8 – 93.8	86.8 – 87.4	93.3 – 95	NA	NA
<i>B. centrosemae</i>	99.7 – 99.8	99.4	83 – 89.7	92.4 – 93.4	NA	95 – 96.3	NA	NA
<i>B. americanum</i>	99.7 – 99.8	99.6	73.1 – 75.2	93 – 95	NA	94.7 – 96.1	NA	NA
<i>B. cajani</i>	99.2 – 99.3	99.3	NA	92.4 – 93.2	NA	94.7 – 96.1	NA	NA
<i>B. vignae</i>	99.6 – 99.7	100	76.7 – 78.5	93.8 – 95.4	91.6 – 92.2	94.7 – 96.1	95.9 – 97	94.3 – 94.9
<i>B. ganzhouense</i>	99.3 – 99.4	99.3	NA	94.4 – 95.4	88 – 88.6	92.5 – 93.6	NA	NA
<i>B. neotropicale</i>	96.7 – 95.8	96.6	71.5 – 73.3	91.2 – 02.6	96.1 – 92	94.7 – 95.5	96.6 – 94	NA
<i>B. algeriense</i>	97 – 97.2	97.4	NA	87.7 – 87.9	NA	90.8 – 91.9	NA	NA
<i>B. macuxiense</i>	97	97.4	NA	89.3 – 89.7	86.2 – 87.4	92.5 – 93.3	88.2 – 88.8	NA
<i>B. tropiciagri</i>	96.8 – 96.9	96.8	69.1 – 74	88.3 – 88.7	91.4 – 92	93 – 94.1	86.9 – 87.4	90.1 – 90.4
<i>B. embrapense</i>	97 – 97.1	97.8	69.6 – 71.3	89.3 – 90	90.9 – 91.4	93.3 – 93.8	85 – 86.1	89.9 – 90.3
<i>B. viridifuturi</i>	97 – 97.1	97.8	NA	88.7 – 89.1	90.9 – 91.3	93.6 – 94.4	86.9 – 87.4	89.9 – 90
<i>B. elkani</i>	96.9 – 97	97.4	70.1 – 71.9	89.1 – 89.5	91.3 – 92	93.6 – 93.8	87.1 – 87.7	90.3 – 90.7
<i>B. jicamae</i>	96.6 – 96.7	97.5	66.5 – 68.5	87.1 – 88.3	89.1 – 89.6	89.7 – 90.8	85.5 – 86.9	NA
<i>B. pachyrhizi</i>	97 – 97.1	97.5	66.7 – 68.7	87.7 – 88.5	91.4 – 92	93.3 – 94.1	86.9 – 87.4	NA
<i>B. lablabi</i>	97 – 97.1	97.6	68.4 – 69.2	87.5 – 87.9	89.1 – 90	91.3 – 92.7	85.5 – 86.9	NA
<i>B. denitrificans</i>	98.1 – 98.2	98.1	NA	88.1 – 88.5	88.5	88 – 88.6	90.7 – 91	NA
<i>B. oligotrophicum</i>	98.6 – 98.7	98.7	NA	85.3 – 85.9	88 – 88.4	90.2 – 91.1	91.5 – 91.8	88.5 – 88.9
<i>B. retamae</i>	97.4 – 97.5	97.9	65.3 – 66.5	88.1 – 89.7	86 – 86.5	91.1 – 93.1	83.6 – 85	87.5 – 87.8
<i>B. icense</i>	97 – 97.1	97.6	88.7 – 89.9	88.7 – 89.9	82 – 82.6	90.8 – 91.6	NA	NA
<i>B. paxllaeri</i>	97 – 97.1	97.6	NA	87.7 – 88.5	88.5 – 89.1	90 – 90.5	84.7 – 81.6	NA
<i>B. valentinum</i>	96.9 – 97	97.6	NA	84.7 – 85.9	86.2 – 86.8	91.3 – 91.9	NA	NA

<i>B. erythrophlei</i>	97 – 97.1	97.8	70.3 – 71.9	88.7 – 89.1	89.9 – 90.3	92.5 – 93.8	89.3 – 89.9	NA
<i>B. ferriligni</i>	95.7 – 95.8	95.6	69.6 – 70.9	87.9 – 88.7	87.6 – 88.4	93 – 93.8	86.6 – 87.1	NA
<i>B. mercantei</i>	97 – 97,1	97.8	NA	88.5 – 88.9	88 – 88.6	93.8 – 94.1	NA	NA
<i>B. centrolobii</i>	95,3 – 97	96.6	83 – 89.7	93 – 93.8	89.8 – 90.4	92.7 – 94.7	92.6 – 94	NA
<i>B. namibiense</i>	97 – 97,1	97.6	74.1 -75.1	87.1 – 88.3	87.6 – 88.5	91.3 – 93.6	88 – 88.8	88.8 – 89.2
<i>B. ripae</i>	NA	97.5	70.4 – 72.5	89.1 – 89.5	90.7 – 91.3	93.3 – 94.1	86.1 – 86.6	NA
<i>B. brasilense</i>	96 – 96,1	97.5	NA	88.5 – 89.3	86.8 – 87.4	93 – 93.8	86.9 – 87.4	NA
<i>B. nitroreducens</i>	NA	99.6	NA	94 – 95	NA	95.5 – 96.8	NA	NA
<i>B. amphicarphae</i>	99.7 – 99.8	99.5	NA	93.6 – 94.2	NA	93.1 – 94.3	NA	NA
<i>B. guangzhouense</i>	99.4 – 99.5	99.2	NA	92 – 93	NA	95 – 96	NA	NA
<i>B. zhanjiangense</i>	98.5 – 98.6	98.9	NA	92.6 – 93.8	NA	93.8 – 95	NA	NA
<i>B. nanningense</i>	99.7 – 99.8	99.4	NA	94 – 94.8	NA	93.3 – 95	NA	NA

NA, Not available

*Eleven strains: CNPSO 3421, CNPSO 3411, CNPSO 3413, CNPSO 3417, CNPSO 3423, CNPSO 3424, CNPSO 3426^T, CNPSO 3431, CNPSO 3443, CNPSO 3446 and CNPSO 3447.

Table 3. Average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) values (%) obtained based on genome comparisons between *Bradyrhizobium frederickii* CNPSo 3426^T and CNPSo 3424 and the most related *Bradyrhizobium* type strains.

Strain used as query	ANI % [*]		DDH% [#]	
	CNPSo 3426 ^T	CNPSo 3424	CNPSo 3426 ^T	CNPSo 3424
<i>B. liaoningense</i> LMG 18230 ^T	88.81	88.84	37.6	37.7
<i>B. stylosanthis</i> BR 446 ^T	87.79	87.71	34.6	34.6
<i>B. huanghuaihaiense</i> CCBAU 23303 ^T	87.96	87.93	35.2	35.2
<i>B. ottawaense</i> OO99 ^T	89.39	89.31	38.7	38.5
<i>B. shewense</i> ERR11 ^T	88.9	88.72	38	36.8
<i>B. frederickii</i> CNPSo 3426 ^T	100	97.38	100	76.8
<i>B. frederickii</i> CNPSo 3424 ^T	97.38	100	76.8	100
<i>B. daqingense</i> CGMCC 1.10947 ^T	87.83	87.7	35.1	34.9
<i>B. neotropicae</i> BR 10247 ^T	85.81	85.68	31.1	31.1
<i>B. guangxiense</i> CCBAU 53363 ^T	89.7	89.44	40	39.6

^{*}ANI values were obtained using ANI calculator [55].

[#]Digital DDH analysis were estimated after [59], determined by GGDC v2.1 using the recommended 'Formula 2' (<http://ggdc.dsmz.de/distcalc2.php>).

Table 4. Distinctive phenotypic features of *Bradyrhizobium frederickii* CNPSO strains and phylogenetically related *Bradyrhizobium* species. Strains: 1, *B. frederickii* CNPSO 3426^T; 2, *B. frederickii* CNPSO 3424; 3, *B. daqingense* CCBAU 15774^T, 4, *B. liaoningense* LMG 18230^T; 5, *B. huanghuaihaiense* CCBAU 23303^T; 6, *B. stylosanthis* BR 446^T.

Characteristic*	1	2	3	4	5 [#]	6 [#]
Carbon source utilization:						
Glycerol	+	+	W	W	W	W
Erythritol	W	+	-	-	-	-
D-Arabinose	+	+	+	+	+	W
L-Arabinose	+	+	W	+	W	W
D-Ribose	+	+	+	+	+	W
D-Xylose	+	+	W	+	W	W
D-Adonitol	-	+	-	-	-	-
Methyl-βD-Xylopyranoside	W	+	-	-	W	W
D-Galactose	W	+	W	W	W	W
D-Glucose	W	+	W	W	W	W
D-Fructose	+	+	W	-	W	W
D-Mannose	W	+	W	W	W	W
L-Sorbose	W	+	-	-	-	-
L-Rhamnose	-	+	W	W	W	W
Dulcitol	-	W	-	-	-	-
D-Mannitol	W	+	W	-	W	W
Methyl-αD-glucopyranoside	-	+	-	-	-	-
N-Acetylglucosamine	-	+	-	-	-	-
Esculin Iron Citrate	+	+	+	+	W	+
D-Melibiose	W	W	-	-	-	-
D-Sucrose	W	W	-	-	-	-
D-Trehalose	W	W	-	W	-	-
Inulin	W	W	-	-	-	-
D-Melezitose	W	W	-	-	-	-
D-Raffinose	W	W	-	-	-	-
Glycogen	+	W	+	+	+	+
Xylitol	-	W	-	-	-	-
Gentiobiose	W	W	-	-	-	-
D-Turanose	W	+	-	-	-	-
D-Tagatose	W	W	-	-	-	-
D-Fucose	+	+	W	+	+	W
L-Arabitol	-	W	-	-	-	-
D-Arabitol	+	+	W	-	W	W
Potassium 2-ketogluconate	+	+	+	+	-	-
Tolerance (µg per disc):						
Chloramphenicol (30)	+	+	+	+	-	-
Erythromycin (15)	-	-	W	+	+	+
Nalidixic Acid (30)	+	+	+	+	+	W
Neomycin (30)	W	W	+	+	W	-
Penicillin G (10U)	+	+	+	+	-	-
Tetracycline (30)	-	W	+	+	W	-
Growth at:						
37 °C	+	+	-	-	-	-
pH 9.0	+	+	W	-	+	+

Alkaline reaction	+	-	+	+	+	+
Urease activity	+	+	+	W	+	+

*Data were evaluated as +, growth; W, weakly positive; -, no growth, after seven days of growth.

#Data obtained in the same laboratory conditions, by Delamuta *et al.* [30].

SUPPLEMENTARY MATERIAL

Bradyrhizobium frederickii* sp. nov., a nitrogen-fixing lineage isolated from nodules of the caesalpinoid species *Chamaecrista fasciculata* and characterized by tolerance to high temperature *in vitro

Maria Clara de Oliveira Urquiaga,^{1,2,3#} Milena Serenato Klepa,^{1,3,4#} Padma Somasegaran,⁵ Renan Augusto Ribeiro,⁴ Jakeline Renata Marçon Delamuta,^{1,4} and Mariangela Hungria^{1,2,4*}

¹ Embrapa Soja, C.P. 231, 86001-970 Londrina, Paraná, Brazil;

² Department of Microbiology, Universidade Estadual de Londrina, C.P. 10.011, 86057-970, Londrina, Paraná, Brazil;

³ Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, SBN, Quadra 2, Bloco L, Lote 06, Edifício Capes, 70.040-020 Brasília, Distrito Federal, Brazil;

⁴ Conselho Nacional de Desenvolvimento Científico e Tecnológico, SHIS QI 1 Conjunto B, Blocos A, B, C e D, Lago Sul, 71605-001 Brasília, Distrito Federal, Brazil.

⁵ Formerly with BIOZYME Inc., Saint Joseph, Missouri, USA

Both authors contributed equally to this study

Corresponding author: Mariangela Hungria, Embrapa Soja, Cx. Postal 231, 86001-970, Londrina-PR, Brazil. E-mail: mariangela.hungria@embrapa.br; biotecnologia.solo@hotmail.com

Table S1. Primers and DNA amplification conditions used in this study.

Primer	Sequence (5' - 3')	Target gene	*Length	PCR cycling	Reference
fD1 rD1	AGAGTTTGATCC TGGCTCAG CTTAAGGAGGTG ATCCAGCC	16S rRNA	1312 bp	2 min 95°C, 30 X (15s 94°C, 45s 93°C, 45s 55°C), 2 min 72°C e 5 min 72°C)	Weisburg et al. [1]
TSglnII _f TSglnII _r	AAGCTCGAGTAC ATCTGGCTCGAC GG SGAGCCGTTCCA GTCGGTGTCG	<i>glnII</i>	505 bp	2 min 95°C, 35 X (45s 95°C, 30s 58°C), 1.5 min 72°C e 7 min 72°C	Stepkowski et al. [2]
gyrB343F gyrB1043R	TTCGACCAGAAY TCCTAYAAGG AGCTTGTCCTTS GTCTGCG	<i>gyrB</i>	562 bp	5 min 95°C, 5X (2 min 94°C, 2 min 58°C, 1 min 72°C) 28 X (30 s 94°C, 1 min 58°C, 1 min 72°C e 5 min 72°C)	Martens et al. [3]
TSecAf TSecAr	CAACTGCMYTG CGTATCGTCGAAGG CGGATCTGGTTG ATGAAGATCACCATG	<i>recA</i>	360 bp	2 min 95°C, 35 X (45s 95°C, 30s 58°C), 1.5 min 72°C e 7 min 72°C	Stepkowski et al. [2]
rpoB83F rpoB1061R	CCTSATCGAGGT TCACAGAAGGC AGCGTGTTGCGG ATATAGGCG	<i>rpoB</i>	371 bp	5 min 95°C, 3X (2 min 94°C, 2 min 58°C, 1 min 72°C) 30 X (30s 94°C, 1 min 58°C), 1 min 72°C e 5 min 72°C	Martens et al. [3]
FGPS1490 FGPS130	TGCGGCTGGATCACCTCCTT CCGGGTTTCCCCATTCCG	ITS	921 bp	3 min 94°C, 35 X (1 min 94°C, 1 min 55°C, 2 min 72°C e 6 min 72°C	Laguerre et al. [4]
nodCF nodCI	AYGTHGTYGAYGACGGITC CGYGACAGCCANTCKCTATTG	<i>nodC</i>	462 bp	30s 94°C, 40x (30s 96°C, 1 min 53°C, 30s 72°C), 5min 72°C	Laguerre et al. [5]

nifHF	TACGGNAARGGSGGNATCGGCAA					
nifHR	AGCATGTCYTC SAGYTCNTCCA	<i>nifH</i>	286 bp	2 min 95°C, 35x(1 min 94°C, 45s 60°C, 2 min 72°C), 7 min 72°C		Laguerre et al. [5]
A1R	ATGTAAGCTCCTGGGGATTAC	BOX	-	7 min 94°C, 35 x (1min 94°C, 1min 52°C, 8 min 72°C) 16min 65°C		Versalovic et al. [6]

*Sequence alignment length.

- 69. Weisburg WG, Barns SM, Pelletier DA, Lane DJ.** 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 1991;173:697-703.
- 70. Stepkowski T, Moulin L, Krzyzanska A, McInnes A et al.** European origin of *Bradyrhizobium* populations infecting lupins and serradella in soils of western Australia and South Africa. *Appl Environ Microbiol* 2005;71:7041-7052.
- 71. Martens M, Dawyndt P, Coopman R, Gillis M et al.** Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *Int J Syst Evol Microbiol* 2008;58:200-214.
- 72. Laguerre G, Mavingui P, Allard MR, Charnay MP et al.** Typing of rhizobia by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions: application to *Rhizobium leguminosarum* and its different biovars. *Appl Environ Microbiol* 1996;62:2029-2036.
- 73. Laguerre G, Nour SM, Sanjuan J, Drouin P.** Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. *Microbiology* 2001;147:981-993.
- 74. Versalovic J, Schneider M, de Bruijn FJ, Lupski JR.** Genomic fingerprinting of bacteria using repetitive sequence based PCR (rep-PCR). *Methods in Molecular and Cellular Biology* 1994;5:25-40.

Table S2. Accession numbers of the housekeeping genes sequences used in this study. Sequences obtained in this study are shown in bold.

Strain	Genome	<i>glnII</i>	<i>gyrB</i>	<i>recA</i>	<i>rpoB</i>	<i>nifH</i>	<i>nodC</i>
<i>B. frederickii</i> CNPSO 3421		MK682682	MK682715	MK682704	MK682693	MK682726	MK682737
<i>B. frederickii</i> CNPSO 3411		MK682683	MK682716	MK682705	MK682694	MK682727	MK682738
<i>B. frederickii</i> CNPSO 3413		MK682684	MK682717	MK682706	MK682695	MK682728	MK682739
<i>B. frederickii</i> CNPSO 3417		MK682685	MK682718	MK682707	MK682696	MK682729	MK682740
<i>B. frederickii</i> CNPSO 3423		MK682686	MK682719	MK682708	MK682697	MK682730	MK682741
<i>B. frederickii</i> CNPSO 3424	SPQU0000000	MK682687	MK682720	MK682709	MK682698	MK682731	MK682742
<i>B. frederickii</i> CNPSO 3426 ^T	SPQS0000000	MK682688	MK682721	MK682710	MK682699	MK682732	MK682743
<i>B. frederickii</i> CNPSO 3431		MK682689	MK682722	MK682711	MK682700	MK682733	MK682744
<i>B. frederickii</i> CNPSO 3443		MK682690	MK682723	MK682712	MK682701	MK682734	MK682745
<i>B. frederickii</i> CNPSO 3446		MK682691	MK682724	MK682713	MK682702	MK682735	MK682746
<i>B. frederickii</i> CNPSO 3447		MK682692	MK682725	MK682714	MK682703	MK682736	MK682747
<i>B. tropiciagri</i> CNPSO 1112 ^T	LFLZ01000014.1	FJ391048	HQ634890	FJ391168	HQ634909.1	HQ259540.1	KP234520
<i>B. embrapense</i> CNPSO 2833 ^T	LFIP02000007.1	GQ160500	HQ634891	HQ634899	HQ634910.1	KP234518.2	KP234521
<i>B. viridifuturi</i> SEMIA 690 ^T	SAMN03890369	KR149131.1	KR149134.1	KR149140.1	SAMN03890369	KR149137.1	LGTB01000005
<i>B. elkani</i> USDA 76 ^T	KB900701.1	AY599117.1	AM418800	AY591568.1	AM295348	AB094963	NZ_KB900701.1
<i>B. retamae</i> Ro19 ^T	NZ_LLYA01000174.1	KC247108.1	KF896204.1	KF962711.1	LLYA01000113.1	KF670138	KC247112.1
<i>B. diazoefficiens</i> USDA 110 ^T	NC_004463.1					NC_004463	BA000040
<i>B. canariense</i> LMG 22265 ^T	NZ_NAFC00000000.1	AY386765.1	FM253220	AY591553.1	FM253263	EU818926	AJ560653.1
<i>B. betae</i> LMG 21987 ^T	JGI Project Id: 1205813	AB353733.1	FM253217	AB353734.1	FM253260	-	-
<i>B. japonicum</i> USDA 6 ^T	NC_017249.1	AF169582	AM418801	AM182158	AM295349	HM047126	AP012206
<i>B. daqingense</i> CCBAU 15774 ^T	JGI Project Id: 1041378	HQ231301.1	JX437669.1	HQ231270.1	JGI Project Id: 1041378	HQ231323.1	HQ231326.1
<i>B. yuanmingense</i> LMG 21827 ^T		AY386780.1	FM253226	AM168343	FM253269	EU818927.1	FMAE01000014.1
<i>B. liaoningense</i> LMG 18230 ^T	Project ID: 1052895	AY386775.1	FM253223	AY591564.1	FM253266	EU818925	GU263466.1
<i>B. iriomotense</i> EK05 ^T		AB300995	AB300997	AB300996	HQ587646.1	AB300998.1	AB301000.1
<i>B. cytisi</i> CTAW11 ^T		GU001594.1	JN186292.1	GU001575.1	JN186288.1	GU001618.1	EU597844.2
<i>B. huanghuaihaiense</i> CCBAU 23303 ^T	JGI Project Id: 1041381	HQ231639.1	JX437672.1	HQ231595.1	HQ587647.1	HQ231551	KF962706.1
<i>B. rifense</i> CTAW71 ^T		GU001604.1	KC569466.1	GU001585.1	KC569468.1	GU001627.1	EU597853.2
<i>B. arachidis</i> CCBAU 051107 ^T	FPBQ01000022.1	HM107251.1	JX437675.1	HM107233.1	FPBQ01000038.1	HM107283	HM107267.1
<i>B. ingae</i> BR 10250 ^T	-	KF927067.1	KF927079.1	KF927061.1	KF927073.1	KF927085.1	KF927054.1
<i>B. manausense</i> BR 3351 ^T	NZ_LJYG01000088.1	KF785986.1	KF786000.1	KF785992.1	NZ_LJYG01000045.1	LJYG010000 13.1	KF786002.1

<i>B. neotropiale</i> BR 10247 ^T	NZ_LSEF01000078.1	KJ661700.1	KJ661707.1	KJ661714.1	NZ_LSEF01000121.1	NZ_LSEF01000026.1	KJ661727.1
<i>B. ottawaense</i> OO99 ^T	CP029425.1	HQ587750.1	HQ873179.1	HQ587287.1	NZ_CP029425.1	JN186287.1	HQ587980.1
<i>B. subterraneum</i> 58-2-1 ^T		KM378484.1	KX661396.1	KM378397	KM378349.1	KM378289.1	-
<i>B. kavangense</i> 14-3 ^T		KM378446.1	KX661397.1	KM378399.1	KM378311.1	KM378254.1	KT033402.1
<i>B. guangdongense</i> CCBAU 51649 ^T	CP030052.1	KC509023.1	KC509072.1	CP030052.1	CP030051.1	CP030052.1	-
<i>B. guangxiense</i> CCBAU 53363 ^T	NZ_CP022220.1	KC509033.1	KC509082.1	NZ_CP022220.1	NZ_CP022219.1	KC509140.1	-
<i>B. stylosanthis</i> BR 446 ^T	LVEM0000000.1	KU724148	KU724151	KU724163	KU724166	KU724157.1	KU724160.1
<i>B. shewense</i> ERR11 ^T	NZ_FMAI01000001.1	JQ809893.1	NZ_FMAI01000013.1	NZ_FMAI01000022.1	JQ810006.1	-	-
<i>B. vignae</i> 7-2 ^T	NZ_RDQF01000031.1	KM378443.1	KX683216.1	KM378374.1	RDQF01000055.1	KM378251.1	KT362339.1
<i>B. jicamae</i> PAC 68 ^T	NZ_LLXZ01000113.1	FJ428204	HQ873309.1	HM047133.1	NZ_LLXZ01000027.1	HM047127	AB573869.1
<i>B. lablabi</i> CCBAU 23086 ^T	LLYB01000064.1	GU433498.1	JX437670.1	GU433522.1	LLYB01000057.1	GU433546	U433565.1
<i>B. pachyrhizi</i> PAC 48 ^T	NZ_LFIQ01000053.1	FJ428201.1	HQ873310.1	HM047130.1	NZ_LFIQ01000076.1	HM047124	NZ_LFIQ01000053.1
<i>B. paxllaeri</i> LMTR 21 ^T	NZ_MAXB01000110.1	KF896169	KF896195	JX943617	NZ_MAXB01000001.1	DQ085619.2	KF896160.1
<i>B. icense</i> LMTR 13 ^T	NZ_CP016428.1	KF896175	KF896201	JX943615	-	KF896161	KF896159.1
<i>B. valentinum</i> LmjM3 ^T	NZ_LLXX01000153.1	JX518575	LLXX01000044.1	JX518589.2	-	KF806461	JX514897.2
<i>B. ferriligni</i> CCBAU 51502 ^T		KJ818099.1	KJ818102.1	KJ818112.1	MG811655.1	KJ818108.1	KJ818109.1
<i>B. denitrificans</i> LMG 8443 ^T		HM047121.1	FM253239.1	EU665419.1	FM253282.1	HM047125.1	-
<i>B. oligotrophicum</i> LMG 10732 ^T	NC_020453.1	JQ619233.1	KF962697.1	JQ619231.1	NC_020453.1	-	-
<i>B. ganzhouense</i> RITF806 ^T		JX277110.1	KP420022.1	JX277144.1	-	JX277110.1	JX292035.1
<i>B. lupini</i> USDA 3051 ^T		KM114862.1	JN685748.1	KM114866.1	-	-	KM114864.1
<i>B. centrosemae</i> A9 ^T		KX012940.1	-	KC247145.1	-	KC247139.1	KC247134.1
<i>B. americanum</i> CMVU44 ^T		KX012942.1	-	KC247141.1	-	KC247135.1	KC247130.1
<i>B. cajani</i> AMBPC1010 ^T		KY349442.1	-	KY349440.1	-	-	KY349444.1
<i>B. brasilense</i> UFLA03-321 ^T	MPVQ01000056.1	MPVQ00000000	KF452827.1	KT793142.1	KF452879.1	-	-
<i>B. mercantei</i> SEMIA 6399 ^T	MKFI00000000	KX690621.1	KX690623.1	KX690615.1	-	KX690625.1	-
<i>B. centrolobii</i> BR 10245 ^T	NZ_LUUB00000000.1	KX527991.1	KX528004.1	KX527954.1	KF983827.3	LUUB01000073.1	NZ_LUUB01000057.1
<i>B. macuxiense</i> BR 10303 ^T	LNCU01000070.1	KX527995.1	KX528008.1	KX527958.1	KX527969.1	NZ_LNCU01000039.1	LNCU01000070.1
<i>B. sacchari</i> BR 10280 ^T	LWIG01000060.1	KX065099.1	-	KX065095.1	-	LWIG01000010.1	KF196792.2
<i>B. namibiense</i> 5-10 ^T		KM378440.1	KX661393.1	KM378377.1	KM378306.1	KM378249.1	KX661399.1
<i>B. forestalis</i> INPA54B ^T	PGVG00000000.1	PGVG00000000	KF452831.1	KF452867.1	-	PGVG00000000.1	PGVG00000000.1
<i>B. ripae</i> WR4 ^T		MF593086.1	MF593094.1	MF593090.1	MF593098.1	-	-
<i>B. nitroreducens</i> TSA1 ^T	NZ_LFJC01000003.1						

<i>B. ampicarpae</i> 39S1MB ^T		KP768605	-	KP768721	-	-	-
<i>B. guangzhouense</i> CCBAU 51670 ^T		KC509008	-	KC509254	-	-	-
<i>B. zhanjiangense</i> CCBAU 51778 ^T		KC508861	-	KC509263	-	-	-
<i>B. nanningense</i> CCBAU 53390 ^T		KC509028	-	KC509274	-	-	-
<i>B. algeriense</i> RST89 ^T	PYCM01000201.1	FJ264924.1	-	FJ264927.1	-	FJ348678.1	-

Table S3. Nucleotide sequence identity percentage (% NI) between two representative strains of *Bradyrhizobium frederickii* and other *Bradyrhizobium* type strains. Similarities based on the sequence identity single and concatenated housekeeping genes (*atpD*, *dnaK*, *gap*, *glnII*, *gltA*, *gyrB*, *pnp*, *recA*, *rpoB* and *thrC*).

Species	Gene										
	<i>atpD</i> (1,373 bp)	<i>dnaK</i> (1,795 bp)	<i>gap</i> (992 bp)	<i>glnII</i> (985 bp)	<i>gltA</i> (1,283 bp)	<i>gyrB</i> (2,405 bp)	<i>pnp</i> (2,120 bp)	<i>recA</i> (1,076 bp)	<i>rpoB</i> (4,125 pb)	<i>thrC</i> (1,422 pb)	MLSA (17,576 pb)
Within CNPSo 3424 and 3426 ^T	97.8	99.1	99.0	97.6	99.1	98.3	98.6	98.5	99.3	99.4	97.8
Between CNPSo strains and:											
<i>B. algeriense</i>	88.9-89.1	91.7-91.8	88.8-89.4	89.0	89.7-89.8	86.2	91.2-91.7	90.8-91.3	91.8-92.1	83.5-83.8	88.9-89.1
<i>B. arachidis</i>	96.4-96.5	96.4	94.5-95.3	94.1-94.2	94.8-95.4	91.9-92.3	93.5-93.7	95.-95.1	96.9-97.2	91.8-91.9	96.4-96.5
<i>B. betae</i>	94.1-94.4	91.8-92.2	94.8-95.3	94.7-95.3	94.7-94.8	92.3	94.3-94.7	93.9-94.4	95.5-95.8	91.2	94.1-94.4
<i>B. brasiliense</i>	90.8-90.9	95.4-95.5	89.9-90.3	90.0-90.2	89.8-90.1	87.9	94.5-94.7	92.6-92.8	93.1-93.4	85-85.1	90.8-90.8
<i>B. canariense</i>	90.2-90.6	96.3-96.4	93.2-93.6	94.1-94.8	94.3-94.5	91.4-91.5	93.8	93.5-94	95.4-95.6	91.4-91.5	90.2-90.6
<i>B. centrolobii</i>	91.8-92.0	95.4-95.7	92.4-92.9	94.4-94.9	93.0	92.0-92.2	92.5	93.6-94	95.4-95.5	89.1-89.2	91.8-92
<i>B. daqingense</i>	95.9-96.6	96.1-96.3	95.0-95.3	94.0-95.2	93.6-93.9	91.5	94.2-94.4	94.7-95.1	97-97.4	91.6-92	95.9-96.6
<i>B. diazoefficiens</i>	96.4-96.9	96.7-97.0	93.8-94.3	95.1-95.8	95.4-95.5	92.0	95.1-95.3	94.4-94.6	96.9-97.3	92.2-92.3	96.4-96.9
<i>B. elkanii</i>	91.1-91.3	92.2-92.5	90.0-90.4	90.5-90.6	89.4-89.5	88.2-88.3	92.2-92.3	93.3	93.3-93.6	84.3-84.4	91.1-91.3
<i>B. embrapense</i>	90.9-91.0	91.7-92.0	89.8-90.2	89.9-90.3	89.2-89.5	87.8-87.9	91.5-91.6	92.6-92.8	93-93.3	83.4-83.6	90.9-91
<i>B. forestalis</i>	92.3-92.4	95.9-96.1	94.8-95.5	94.2-94.4	93.1-93.4	92.3-92.4	93.4-93.5	95.3-95.7	96.4-96.5	93	92.3-92.4
<i>B. guangdongense</i>	95.8-96.1	95.4-95.7	93.2-93.4	94.4-95.0	92.9	91.7-91.8	94.7-94.9	94.3-94.7	95.6-95.8	88.7-88.8	95.8-96.1
<i>B. guangxiense</i>	95.6-96.0	96.4-96.6	94.8-95.1	95.1-95.9	95.4-95.7	91.8	94.5-94.7	95.3-95.7	97.1-97.5	94.1-94.5	95.6-96
<i>B. huanghuaihaiense</i>	96.4-96.7	96.5-96.7	95.2-95.5	95.4-95.9	93.6	92.3-92.6	94.6-94.9	94.2-94.6	97.1-97.2	92.5-92.8	96.4-96.7
<i>B. icense</i>	88.9-89.3	91.8-92.1	88.7-89.1	89.4-89.9	89.4-89.7	86.3-86.6	90.1-90.5	91-91.1	91.9-92.2	82.4-82.6	89.3-89.9
<i>B. japonicum</i>	91.6-91.9	96.3-96.4	94.4-94.8	95.1-96.0	94.7-94.8	92.4-92.6	94.8-95.2	93.8-94.3	96.6-96.9	91.6-91.8	91.6-96.9
<i>B. jicamae</i>	90.7-90.9	91.7-92.2	89.4-90.0	89.0-89.5	89.0-89.4	86.9	91.2-91.4	91-91.4	92.1-92.4	84.3	90.7-90.9
<i>B. lablabi</i>	90.0-90.4	92.3-92.7	90.7-91.2	89.2	89.7-90.1	87.2-87.5	90.5-90.7	91.2-91.5	92-92.2	83.8-84	90-90.4
<i>B. liaoningense</i>	95.6-96.1	96.3	95.1-95.5	95.3-96.0	93.2-93.3	93.4-93.5	94.9-95.2	95.6-96	96.3-96.6	92.2-92.3	95.6-96.1
<i>B. macxiense</i>	91.6	81.7-81.9	89.0-89.2	90.2	87.9-88.3	87.6-88.0	91.3-91.4	92.1-92.3	92.9-93.2	83.3-83.6	91.6
<i>B. manausense</i>	95.9-96.4	95.2	94.0-94.3	93.4-93.9	95.0-95.1	91.4-91.5	95.2-95.6	93.8-94.4	95.2-95.3	88.4-88.5	95.9-96.4
<i>B. mercantei</i>	91.0-91.8	92.5-92.7	89.4-89.7	90.4-90.5	89.6-89.8	88.1-88.3	91.8-91.9	92.2-92.4	92.6-92.8	84.8-85	91-91.8
<i>B. neotropicale</i>	95.8-96.2	96.8-97.1	95.5-96.2	95.7-96.1	94.3-94.1	92.3-92.5	92.8	95.9-96	95.2-95.4	89.6	95.8-96.2

<i>B. oligotrophicum</i>	90.0-90.6	91.6-92.1	89.8-90.2	84.4-84.6	89.7-89.9	86.5-86.6	90.4-90.8	89.4-89.6	93.1-93.3	82.8	90-90.6
<i>B. ottawaense</i>	94.5-94.9	97.0-97.1	96.0-96.5	94.4-94.5	95.6-95.8	92.8-93.1	94.9-95.3	95.6-95.9	97.2-97.5	94.4-94.8	94.5-94.9
<i>B. pachyrhizi</i>	91.2-91.6	91.6-92.0	89.8-90.2	89.8-90.0	89.5-89.7	87.8-87.9	91.6-91.8	92.9	93.2-93.5	85.2-85.3	91.2-91.6
<i>B. paxllaeri</i>	89.6	91.9-92.4	89.8-90.2	89.1-89.6	89.0	86.8-86.9	90.5-90.7	90.8-91	92-92.2	84.2	89.6
<i>B. retamae</i>	88.4-89.1	90.6-90.9	90.2-90.6	89.5-90.2	89.0-89.2	86.3-86.5	90.3-90.7	91.5-91.7	91.9-92.2	82.9	88.4-89.1
<i>B. sacchari</i>	95.9-96.1	96.5-96.6	94.5-94.7	95.0-95.5	93.8-93.9	91.8-92.0	94.4-94.6	95.1-95.4	96.3-96.5	91.8-92.1	95.9-96.1
<i>B. stylosanthis</i>	95.6-95.8	96.2-96.3	94.3-94.8	96.0-96.1	95.0-95.1	91.5-91.7	94.5-94.9	95.5-95.7	96.6-96.8	92.8-93	95.6-95.8
<i>B. tropiciagri</i>	91.3-91.5	92.8-93.0	90.4-90.7	90.3-90.5	89.7	87.9-88.1	91.8	92.5	92.9-93.1	94.5-84.6	91.3-91.5
<i>B. valentinum</i>	88.6	90.2-90.5	89.3-89.9	87.9-88.0	89.3-89.9	86.1-85.7	90.7-91	92-92.1	92.2-92.4	82.9-83.1	88.6
<i>B. vignae</i>	92.4-92.8	95.3-95.5	94.7-95.4	94.7-95.3	93.7-94.0	92.2-92.3	93.1-93.3	94.9-95.5	96.1-96.5	91.8-92.1	92.4-92.8
<i>B. viridifuturi</i>	91.3-91.8	92.5-92.8	90.5-90.8	90.0	89.0-89.2	88.2-88.4	91.6-91.7	92.4-92.5	93.1-93.2	84.3-84.4	91.3-91.8

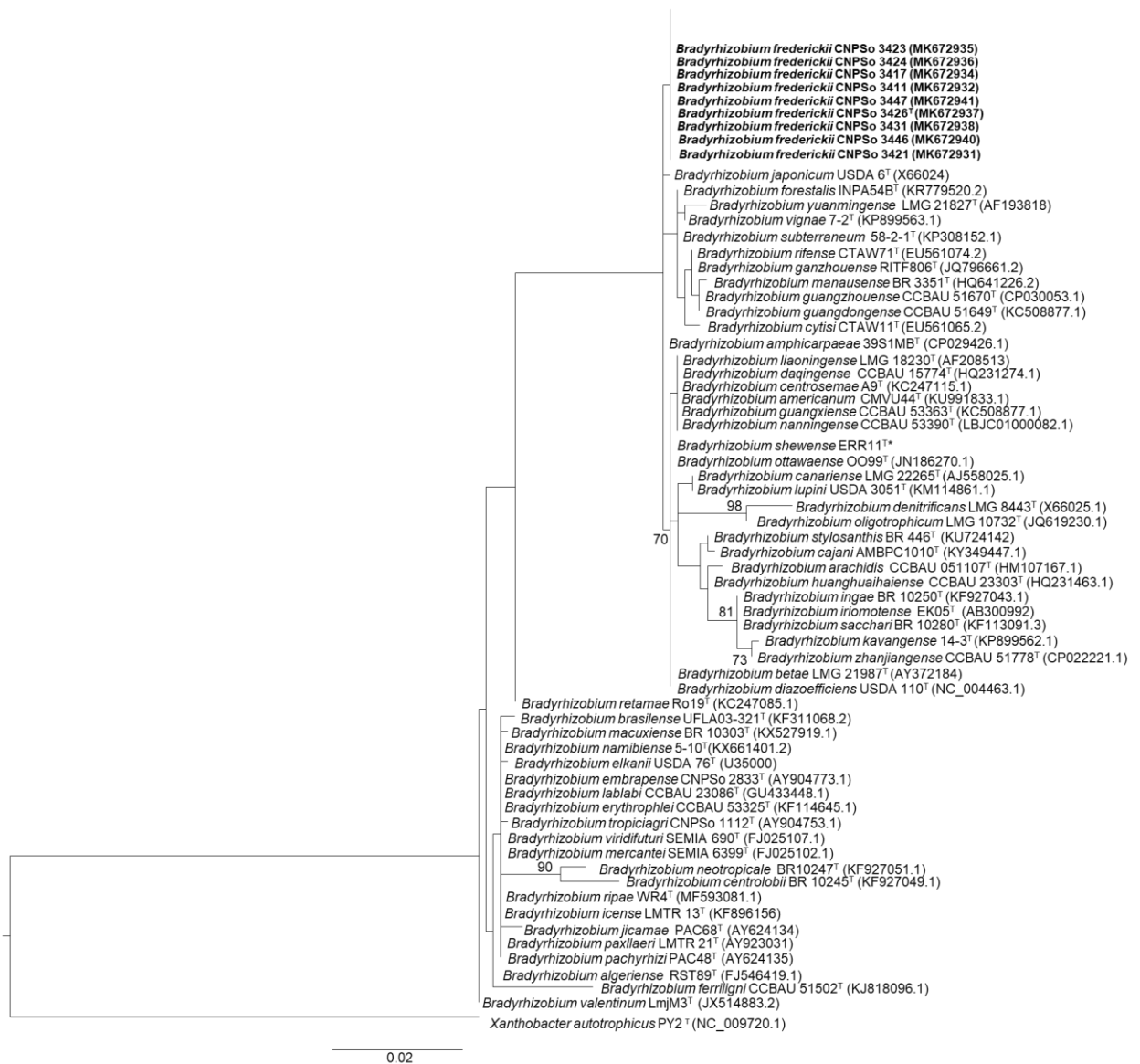


Fig. S1. Maximum Likelihood (ML) phylogeny based on the 16S rRNA aligned sequences (949 bp), showing in bold the phylogenetic position of the novel species *Bradyrhizobium frederickii*. GenBank accession numbers are indicated in parentheses. Bootstrap values greater than 70 % are indicated at the nodes. *Xanthobacter autotrophicus* Py2 was used as outgroup. The scale bar indicates 2 substitutions per 100 nucleotide positions.

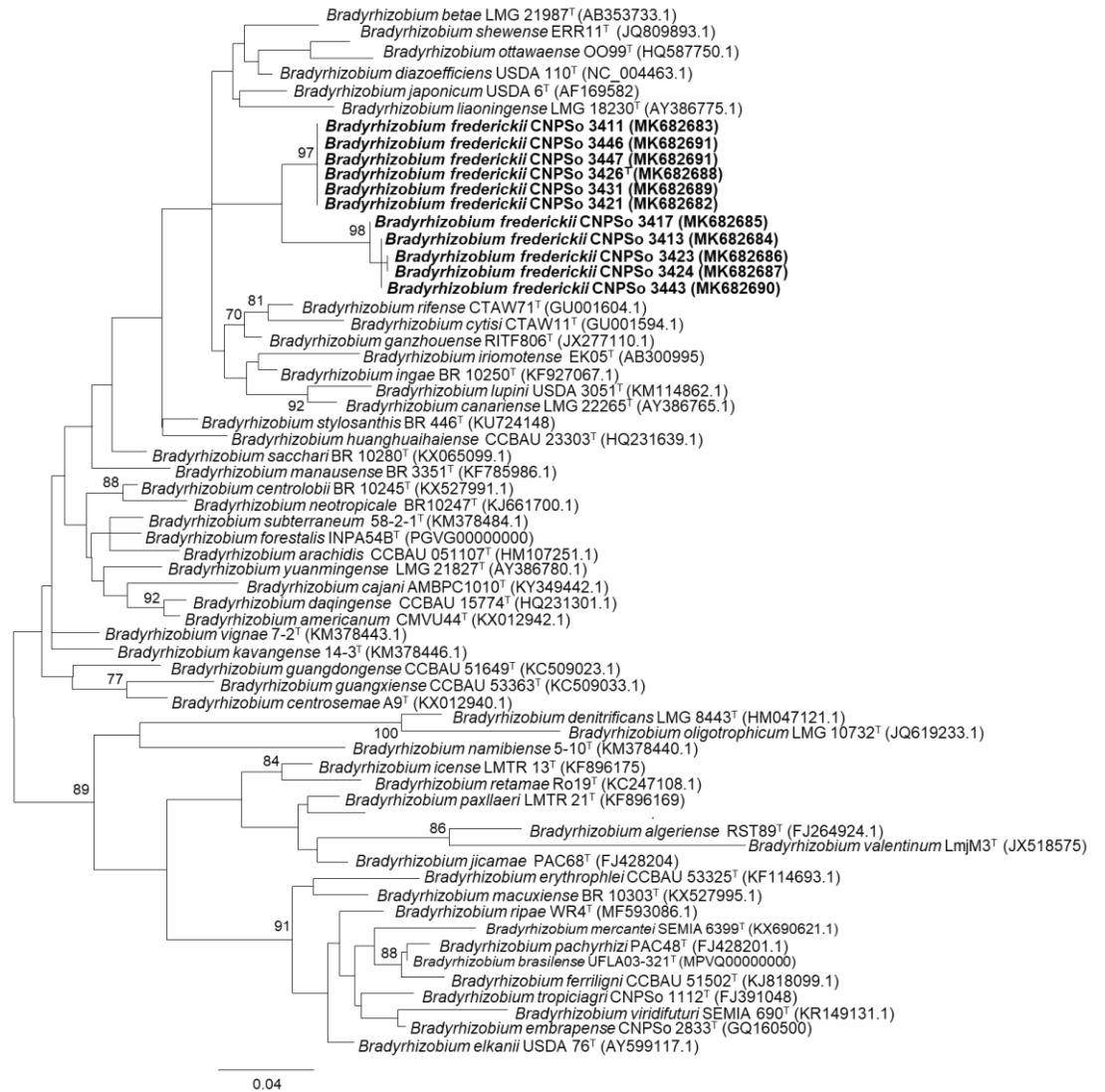


Fig. S2. Maximum Likelihood (ML) phylogeny based on the *glnII* (505 bp) aligned sequences, showing in bold the phylogenetic position of the novel species *Bradyrhizobium frederickii* (in bold) and related *Bradyrhizobium* species. GenBank accession numbers are indicated in parentheses. Bootstrap values greater than 70 % are indicated at the nodes. Bar, 4 substitutions per 100 nucleotide positions.

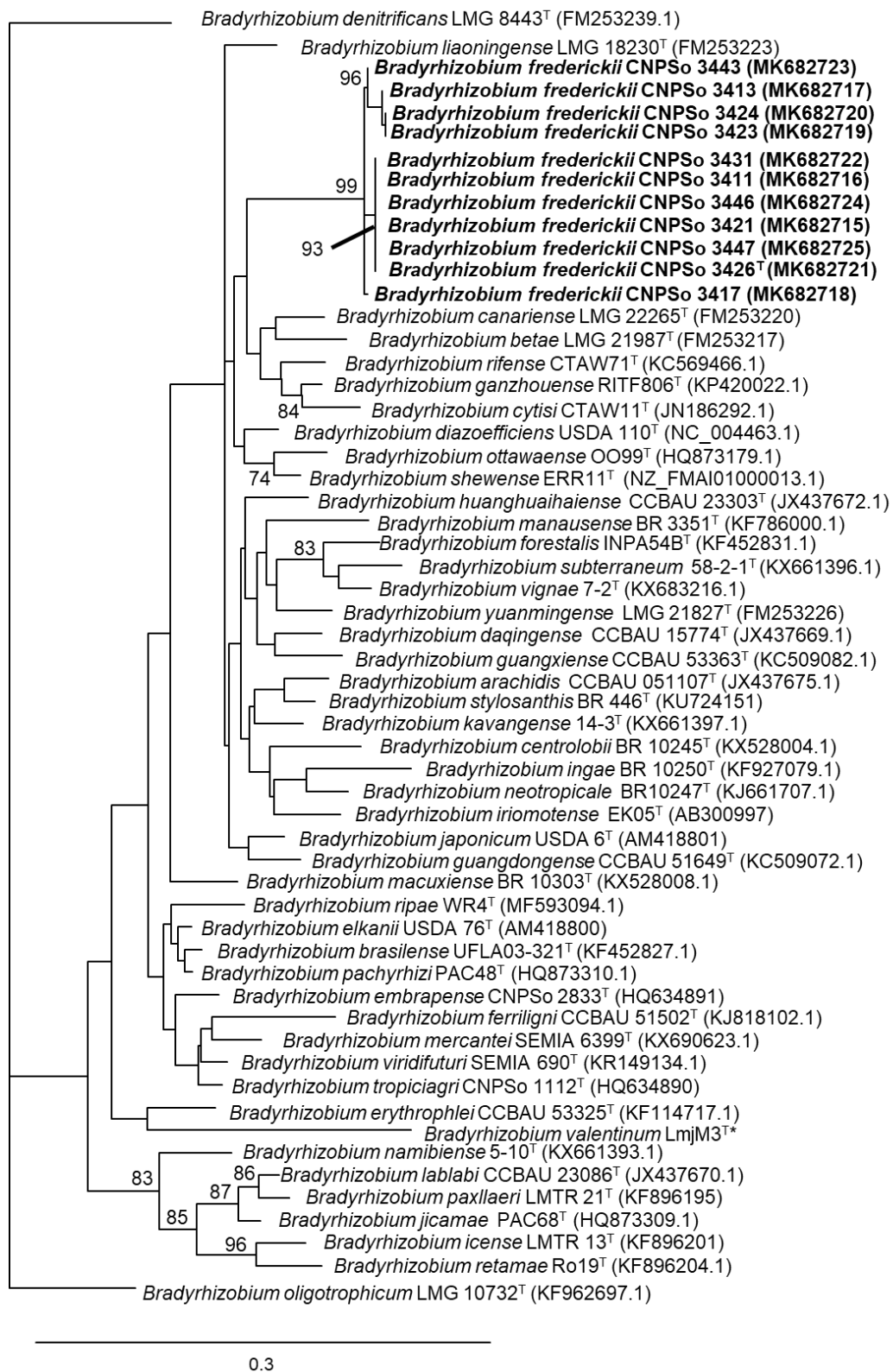


Fig. S3. Maximum Likelihood (ML) phylogeny based on the *gyrB* (562 bp) aligned sequences, showing in bold the phylogenetic position of the novel species *Bradyrhizobium frederickii* (in bold) and related *Bradyrhizobium* species. GenBank accession numbers are indicated in parentheses. Bootstrap values greater than 70 % are indicated at the nodes. Bar, 30 substitutions per 100 nucleotide positions.

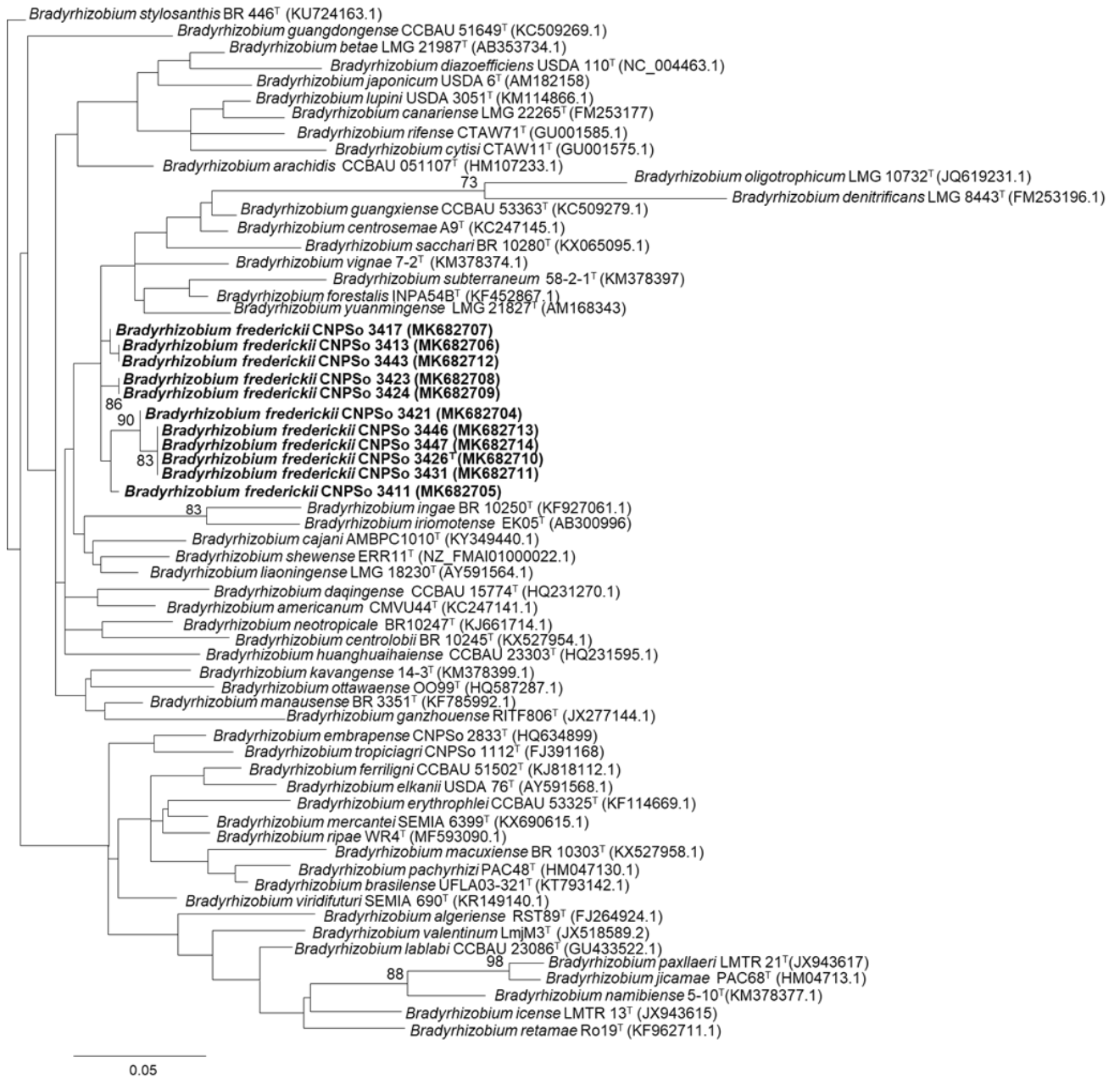


Fig. S4. Maximum Likelihood (ML) phylogeny based on the *recA* (360 bp) aligned sequences, showing in bold the phylogenetic position of the novel species *Bradyrhizobium frederickii* (in bold) and related *Bradyrhizobium* species. GenBank accession numbers are indicated in parentheses. Bootstrap values greater than 70 % are indicated at the nodes. Bar, 5 substitutions per 100 nucleotide positions.

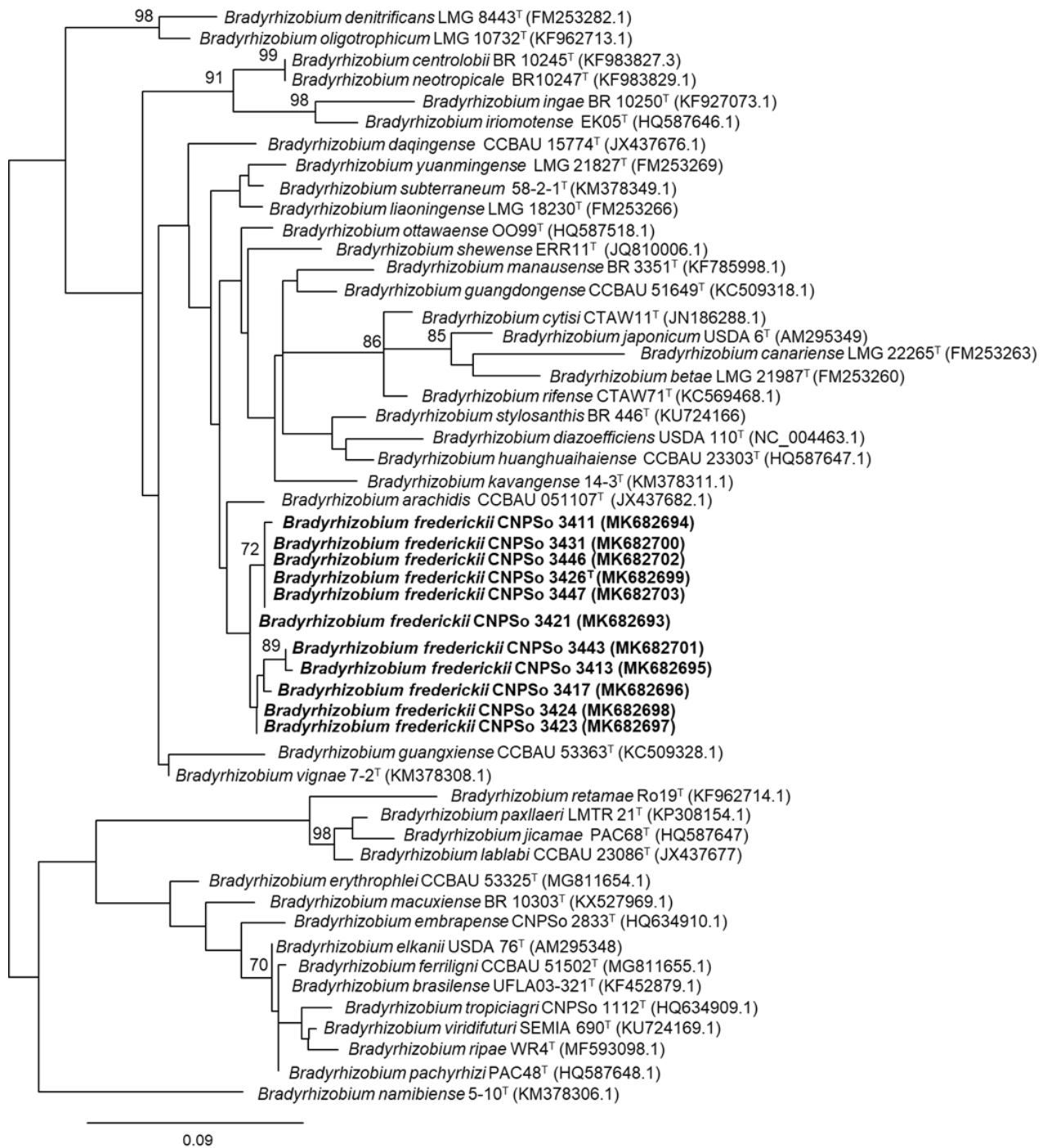


Fig. S5. Maximum Likelihood (ML) phylogeny based on the *rpoB* (371 bp) aligned sequences, showing in bold the phylogenetic position of the novel species *Bradyrhizobium frederickii* (in bold) and related *Bradyrhizobium* species. GenBank accession numbers are indicated in parentheses. Bootstrap values greater than 70 % are indicated at the nodes. Bar, 9 substitutions per 100 nucleotide positions.

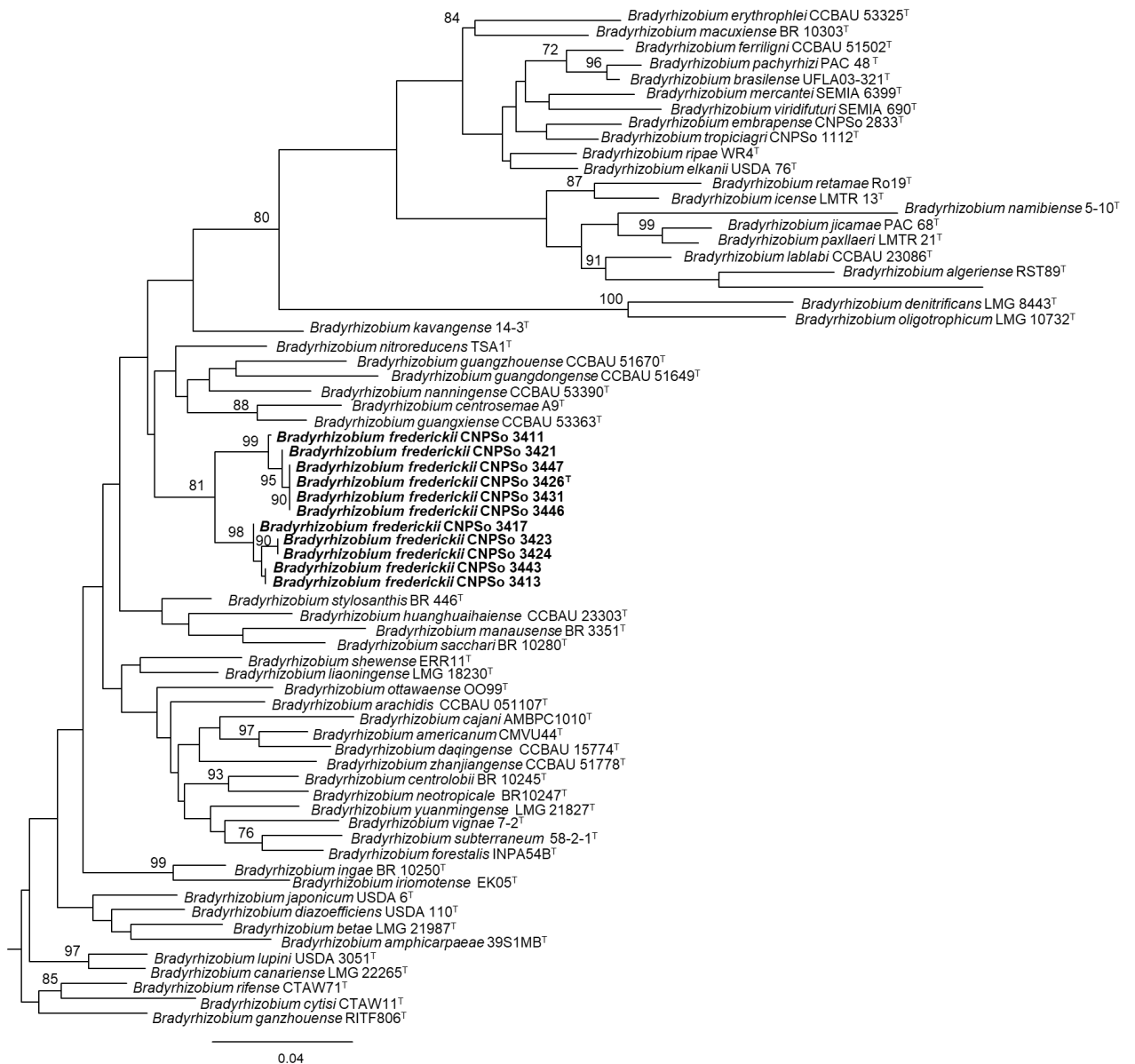


Fig. S6. Maximum Likelihood (ML) phylogeny based on the concatenated aligned sequences of *glnII* (505 pb) and *recA* (408 pb) showing the relationships of the novel species *Bradyrhizobium frederickii* (in bold) with related *Bradyrhizobium* species. GenBank accession numbers are indicated in Table S2. Bootstrap values greater than 70 % are indicated at the nodes. Bar, 4 substitutions per 100 nucleotide positions.

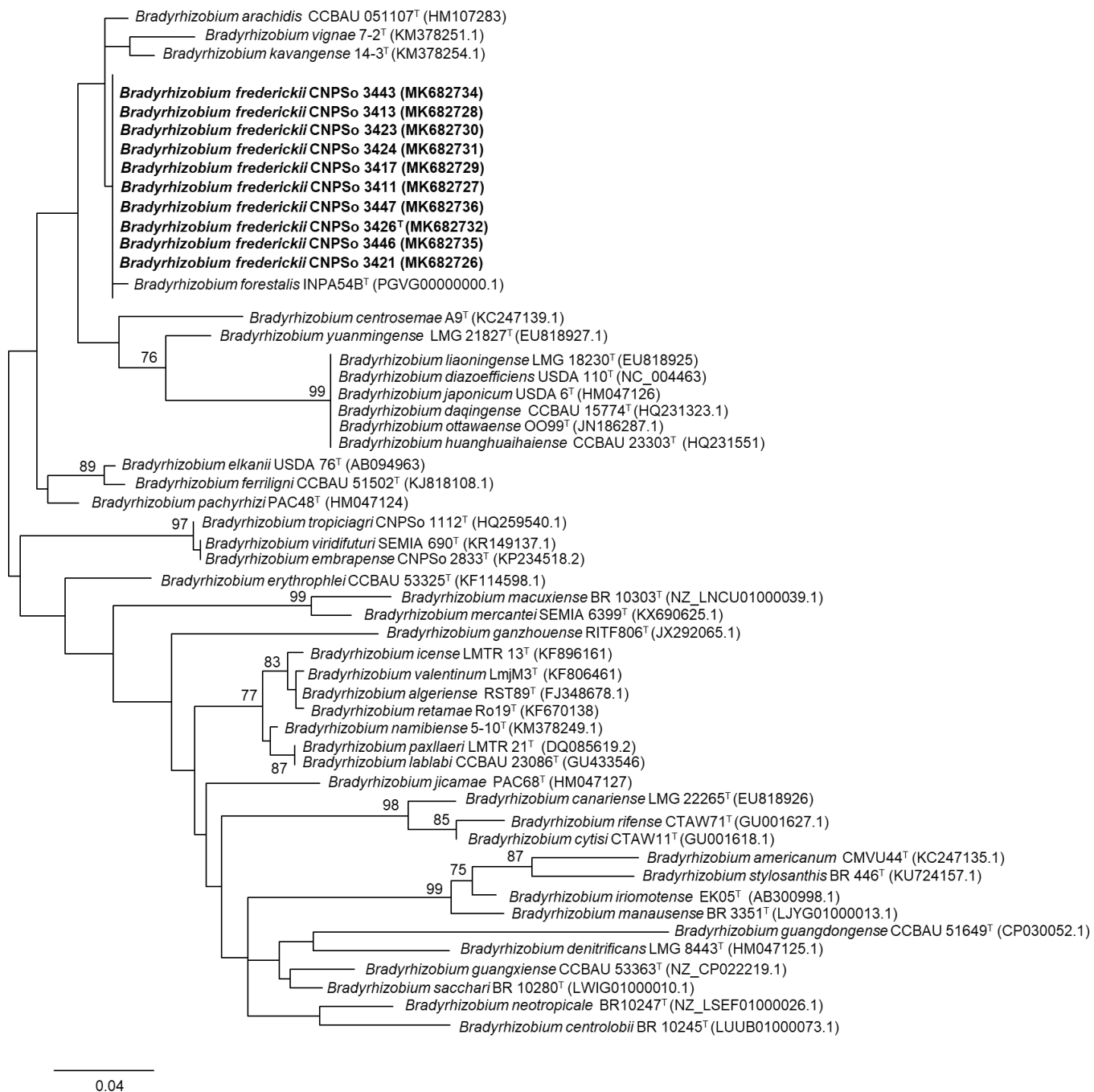


Fig. S7. Maximum Likelihood (ML) phylogenetic tree based on the *nifH* aligned sequences (286 bp) showing the relationships of the novel species *Bradyrhizobium frederickii* (in bold) with other *Bradyrhizobium* species. GenBank accession numbers are indicated in parentheses. Bootstrap values greater than 70 % are indicated at the nodes. Bar, 4 substitutions per 100 nucleotide positions.

BOX-PCR

BOX-PCR

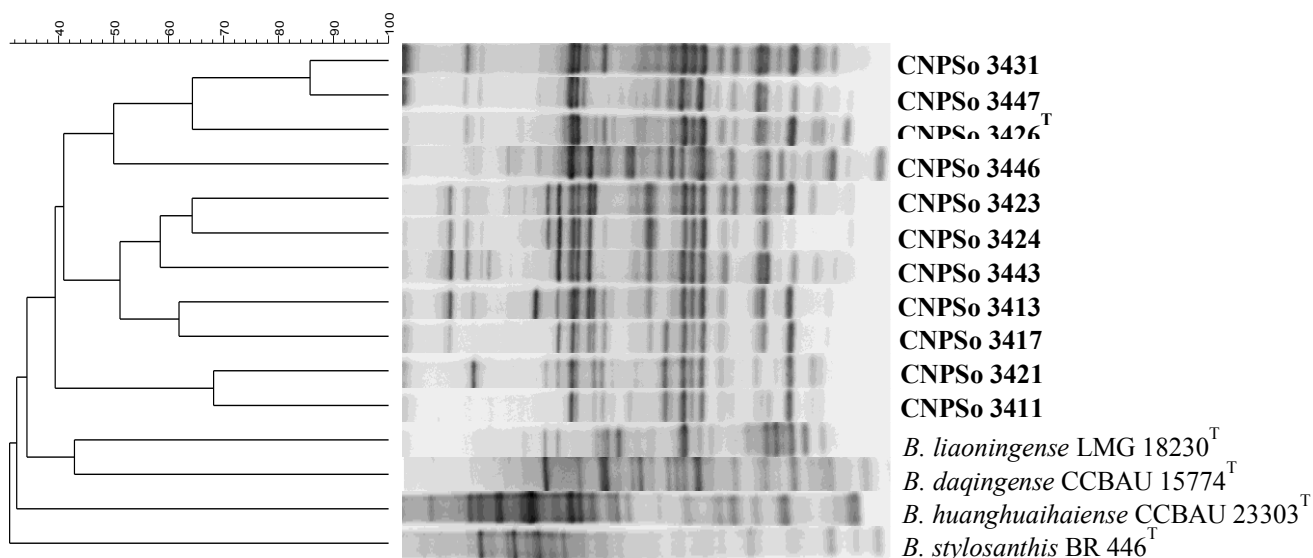


Fig. S8. Cluster analysis of products obtained by BOX- PCR of *Bradyrhizobium frederickii* CNPSo strains (in bold) and closely related *Bradyrhizobium* species. Fingerprinting analysis were performed with Bionumerics software (Applied Mathematics, Kortrijk, Belgium, v.7.6) using the UPGMA algorithm and the Jaccard coefficient with 1 % tolerance.

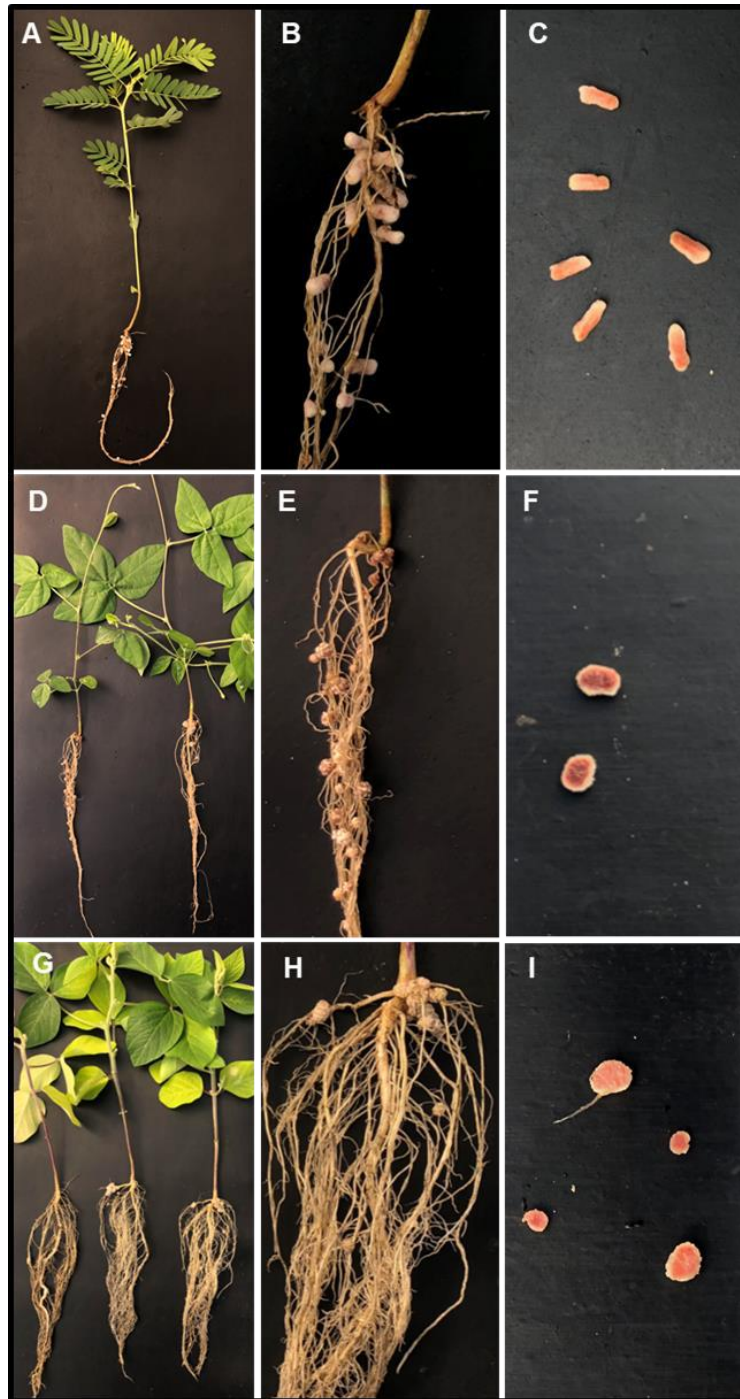


Fig. S9. (A) *Chamaecrista fasciculata*, the original host plant, inoculated with CNPSo 3426^T strain; (B) *C. fasciculata* root with indeterminate nodules formed in association with CNPSo 3426^T; (C) An effective and indeterminate nodules with pink color inside formed during association between *C. fasciculata* and CNPSo 3426^T; (D) Siratro (*Macropitium atropurpureum*), a promiscuous legume widely used in nodulation ability tests, inoculated with CNPSo 3426^T; (E) Siratro root with determinate nodules and healthy plants with green leaves inoculated with CNPSo 3426^T; (F) An effective and determinate nodules with internal pink color formed during association between siratro and CNPSo 3426^T; (G) Soybean (*Glycine max*), a legume with great economic interest, inoculated with CNPSo 3426^T; (H) Soybean root with determinate nodules formed in

association with CNPSo 3426^T; (I) An effective and determinate nodules with internal pink color formed during association between soybean and CNPSo 3426^T.

7. CONSIDERAÇÕES FINAIS

Com base em uma abordagem polifásica completa compreendendo estudos fenotípicos, filogenéticos e genotípicos os resultados indicam que as 11 estirpes simbiotes de *Chamaecrista fasciculata* correspondem a uma nova espécie dentro do gênero *Bradyrhizobium*, para o qual o nome *Bradyrhizobium frederickii* sp. nov. é proposto, com a estirpe CNPSo 3426^T sendo designada como estirpe de referência.

Até o presente momento, os poucos estudos sobre FBN em *Chamaecrista* foram amplos e apenas pontuaram a ocorrência de nódulos, incluindo poucos isolados. Considerando os relatos de que a nodulação em *Chamaecrista* parece ser um evento independente na evolução de simbioses que apresentam FBN, estudos desenvolvidos nesta dissertação podem servir como base para o entendimento de como a “primitiva” Caesalpinioideae evoluiu para as simbioses mais “avançadas” encontradas no clado Mimosoideae e na subfamília Papilionoideae. Dessa forma, o presente trabalho contribui como um dos primeiros estudos que tem como objetivo avaliar a diversidade genética entre as estirpes simbiotes do gênero *Chamaecrista*, destacando a biodiversidade de rizóbios associados a essa leguminosa.

Além disso, os resultados obtidos adicionam mais evidências sobre a alta diversidade genética presente dentro do gênero *Bradyrhizobium*. Em especial, este estudo pode atuar como incentivo para novos trabalhos de diversidade de simbiotes da pouco explorada subfamília Caesalpinioideae, fornecendo, assim, recursos para estudar questões fundamentais sobre a evolução das leguminosas e as origens da simbiose leguminosa-rizóbio.

Por fim, considerando a relevante capacidade de crescimento *in vitro* a 37 °C verificada em todas as representantes de *Bradyrhizobium frederickii*, é possível que o presente trabalho possa contribuir para estudos mais amplos sobre metabolismos de tolerância a temperaturas elevadas, podendo servir como base para futuras aplicações biotecnológicas.