



**BRUNA DANIELA ORTIZ LOPEZ**

**IDENTIFICATION AND PHYLOGENY OF SYMBIOTIC  
RHIZOBIA FROM FOREST LEGUMINOUS SPECIES**

**LAVRAS – MG**

**2023**

**BRUNA DANIELA ORTIZ LOPEZ**

**IDENTIFICATION AND PHYLOGENY OF SYMBIOTIC RHIZOBIA FROM  
FOREST LEGUMINOUS SPECIES**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Ciência do Solo, para obtenção do título de Doutora.

Prof (a). Dr (a). Fatima Maria de Souza Moreira

Orientador (a)

**LAVRAS – MG**

**2023**

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca  
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Lopez, Bruna Daniela Ortiz.

Identification and phylogeny of symbiotic rhizobia of forest  
species / Bruna Daniela Ortiz Lopez. - 2023.

82 p. : il.

Orientador(a): Fatima Maria de Souza Moreira.

Tese (doutorado) - Universidade Federal de Lavras, 2023.

Bibliografia.

1. Biodiversity. 2. Taxonomy. 3. Phylogeny. I. Moreira, Fatima  
Maria de Souza. II. Título.

**BRUNA DANIELA ORTIZ LOPEZ**

**IDENTIFICAÇÃO E FILOGENIA DE RIZÓBIOS SIMBIÓTICOS DE ESPÉCIES  
FLORESTAIS**

**IDENTIFICATION AND PHYLOGENY OF SYMBIOTIC RHIZOBIA OF FOREST  
SPECIES**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Ciência do Solo, para obtenção do título de Doutora.

Aprovada em 31 de janeiro de 2023.

Dra. Elaine Martins da Costa (UFPI)

Dra. Rafaela Simão Abrahão Nóbrega (UFRB)

Dr. Rubens Manoel dos Santos (UFLA)

Dra. Silvia Maria de Oliveira Longatti (UFLA)

Prof (a). Dr (a). Fatima Maria de Souza Moreira  
Orientador (a)

**LAVRAS – MG**

**2023**

*Aos meus pais Margarita e Assis pelo apoio e  
dedicação e por nunca me deixarem desistir.*

*Ao meu irmão Diego pelo exemplo de ser humano  
que sempre me fez melhorar como pessoa.*

*Dedico*

## **AGRADECIMENTOS**

À Universidade Federal de Lavras, especialmente ao Departamento de Ciência do Solo.

Ao CNPq, pela concessão da bolsa de doutorado e demais recursos financeiros para o desenvolvimento da minha pesquisa. À CAPES, CNPq e Fapemig pelos recursos financeiros destinados aos programas de Pós-Graduação e projetos de pesquisa.

À professora Dra. Fatima Maria de Souza Moreira, pela orientação, dedicação e paciência que me ajudaram a concluir esta etapa acadêmica da minha vida.

A todos os professores e funcionários do DCS/UFLA, especialmente os professores Júnior César Avanzi (Chefe do Departamento de Ciência do Solo), Bruno Montoani Silva (Coordenador do Programa de Pós-Graduação em Ciência do Solo) e Dirce de Cássia Corrêa Macedo (Secretaria Integrada da ESAL).

A todos os colegas de departamento, pelos momentos de aprendizado e descontração durante esse período. Quero agradecer especialmente aos amigos Davi, Fernandes, Paulo, Flávia Romam, Ana Paula, Rafaella, Lucas Lenin e Juliano. Vocês foram sensacionais!

À toda minha família pelo apoio, palavras e gestos de carinho. Aos meus pais, Assis e Margarita pelo amor e dedicação em todas as etapas da minha vida, a minha madrinha que mesmo de longe preza pelo meu sucesso e ao meu irmão Diego, exemplo que sigo desde a infância.

**MUITO OBRIGADA!**

*“Reserve o direito de pensar, mesmo se você estiver errado, é melhor do que não pensar em nada.”  
(Hipátia de Alexandria)*

## RESUMO

O Brasil é conhecido por sua grande diversidade vegetal e potencial biotecnológico. A simbiose rizóbio-leguminosa é uma importante fonte de nitrogênio (amônia) na biosfera e a identificação dos rizóbios que promovem a melhor fixação no seu hospedeiro é uma área do conhecimento muito valiosa para gerar alternativas a fertilizantes nitrogenados minerais. Várias novas espécies de bactérias simbióticas fixadoras de nitrogênio já foram descritas, principalmente pertencentes aos gêneros *Bradyrhizobium* e *Rhizobium*. No entanto, não há identificação em nível de espécie de estirpes que nodulam e fixam N<sub>2</sub> em simbiose com as leguminosas forestais nativas *Machaerium nycitans* (Vell.) Benth. e *Platypodium elegans* (Voegel.). Pretendeu-se, neste trabalho, caracterizar genética e fenotipicamente estirpes isoladas dessas duas espécies com capacidades nodulíferas e fixadoras de nitrogênio atmosférico. Além disso foram apresentadas, dentre elas, as novas espécies dos gêneros *Bradyrhizobium* e *Rhizobium*. Para isso foi realizada a extração do DNA das estirpes UFLA01-814, UFLA01-839, UFLA01-860, UFLA01-1127, UFLA01-1134 e UFLA01-1164 para posterior sequenciamento do genoma de cada uma. A análise genômica contou com a comparação do genoma das estirpes do estudo com genomas dos gêneros *Bradyrhizobium* e *Rhizobium* disponíveis no NCBI (National Center for Biotechnology Information) por ANI (average nucleotide identity) e TYGS (Type genome server). A partir das evidências encontradas e toda a caracterização das estirpes é possível comprovar a existência de quatro novas espécies de rizóbios, sendo três do gênero *Bradyrhizobium* e um do gênero *Rhizobium*.

**Palavras-chave:** Biodiversidade. Taxonomia. Filogenia. Bactérias fixadoras de nitrogênio.



## ABSTRACT

Brazil is known for its great plant diversity and biotechnological potential. The rhizobium-legume symbiosis is an important source of fixed nitrogen (ammonia) in the biosphere and the identification of rhizobia that promote better fixation in their host is a very valuable area of knowledge to generate alternatives to mineral nitrogen fertilizers. Several new species of symbiotic nitrogen-fixing bacteria have already been described, mainly belonging to the genera *Bradyrhizobium* and *Rhizobium*. However, there is no identification at the species level of strains that nodulate and fix N<sub>2</sub> in symbiosis with the native forest legume *Machaerium nycitans* (Vell.) Benth. and *Platypodium elegans* (Voegel.). The aim of this work was to genetically and phenotypically characterize strains isolated from forest species (*Machaerium nycitans* and *Platypodium elegans*) with noduliferous and atmospheric nitrogen fixing capacities. In addition, new species of the genera *Bradyrhizobium* and *Rhizobium* were presented, among them. For this purpose, DNA extraction of strains UFLA01-814, UFLA01-839, UFLA01-860, UFLA01-1127, UFLA01-1134 and UFLA01-1164 was performed for subsequent genome sequencing of each one. The genomic analysis included the comparison of the genome of the study strains with genomes of the *Bradyrhizobium* and *Rhizobium* genera available at the NCBI (National Center for Biotechnology Information) by ANI (average nucleotide identity) and TYGS (Type genome server). From the evidence found and the entire characterization of the strains, it is possible to prove the existence of four new species of rhizobia, three of the genus *Bradyrhizobium* and one of the genus *Rhizobium*.

**Keywords:** Biodiversity. Taxonomy. Phylogeny. Nitrogen fixing bacteria.

## SUMÁRIO

<b>PRIMEIRA PARTE .....</b>	<b>11</b>
1. INTRODUÇÃO .....	11
2. REFERENCIAL TEÓRICO .....	12
2.1 Microrganismos fixadores de nitrogênio atmosférico.....	12
2.2 Leguminosas florestais .....	12
2.3 Técnicas de identificação bacteriana.....	14
3. REFERÊNCIAS.....	16
<b>SEGUNDA PARTE - ARTIGOS.....</b>	<b>19</b>
ARTIGO 1 .....	20
<i>Bradyrhizobium versatile</i> sp. nov., <i>Bradyrhizobium cerradense</i> sp. nov. and <i>Bradyrhizobium nyctitae</i> , nitrogen-fixing bacterial strains isolated from a native leguminous tree from Cerrado ecoregion in Brazil.....	20
ARTIGO 2 .....	64
<i>Rhizobium elegans</i> sp. nov. nitrogen-fixing bacteria strain isolated from nodules of <i>Platypodium elegans</i> (Voegel).....	64

## **PRIMEIRA PARTE**

## 1. INTRODUÇÃO

A absorção de nitrogênio pelas plantas pode ser obtida pela disponibilidade que há nos solos, seja nas formas inorgânicas como nitrato e amônio (COURTY et.al., 2015) e orgânicas dissolvidas como aminoácidos, proteínas e ureia (JONES, et.al., 2005; NASHOLM et al., 1998). No entanto, as leguminosas conseguem formar interações mutualísticas com bactérias específicas que vão fixar nitrogênio atmosférico, transformá-lo numa forma assimilável e em troca a planta fornece fotossintatos para a o estabelecimento dessas bactérias.

Esse processo denominado fixação biológica de nitrogênio (FBN) consiste na conversão de nitrogênio atmosférico ( $N_2$ ) em amônia ( $NH_3$ ) por microrganismos que possuem a capacidade de cortar a tripla ligação de  $N_2$  pela enzima nitrogenase. Esse processo permite a disponibilidade de nitrogênio para as plantas, reduzindo o uso de fertilizantes minerais e promovendo a sustentabilidade dos ecossistemas. Os rizóbios fazem parte desse grupo de microrganismos diazotróficos e promovem a FBN mais eficiente devido à formação de nódulos nas raízes das leguminosas.

Existe uma ampla gama de dizotróficos com representantes de arqueobacterias, cianobactérias, bactérias gram-negativas e gram-positivas com diversidade morfológica, fisiológica e genética. No entanto, ainda há muito por ser explorado, já que menos de 5% dos microrganismos do solo foram identificados. No solo há uma diversidade de espécies de bactérias presentes, entretanto a FBN ocorre através de uma relação mutualística estabelecida entre a planta e o microrganismo. As bactérias fixadoras nodulíferas são classificadas em: actinobacterias gram-positivos (*Frankia*) e os rizóbios (gram-negativas) (MOREIRA e SIQUEIRA 2006).

As espécies leguminosas florestais vêm ganhando visibilidade em projetos de recuperação de áreas degradadas, em reflorestamento e restauração ambiental (MIRA et. al., 2022; FIORE et. al., 2019). A inoculação das mesmas poderia ser um atributo favorável para esses projetos devido ao fator ecológico que a fixação biológica possui (PRAGYA e POONAM, 2022). Para isso, a combinação de planta-microrganismo ideal deve ser encontrada e a melhor maneira é a partir do isolamento e identificação de microrganismos provenientes dos nódulos dessas espécies florestais.

## 2. REFERENCIAL TEÓRICO

### 2.1 Microrganismos fixadores de nitrogênio atmosférico

Os microrganismos que possuem a enzima nitrogenase são chamados dizotróficos, e são representados por muitas eubactérias diferentes e algumas arqueias metanogênicas, que podem estabelecer algum tipo de interação com a planta e dessa forma disponibilizar o N para absorção. A enzima nitrogenase é composta de duas proteínas, a Fe-proteína e a MoFe-proteína, além de um cofator FeMo que reduz o N<sub>2</sub> para NH<sub>3</sub>. Após a fixação do N, o NH<sub>3</sub> se transforma em NH<sub>4</sub>, que é convertido em glutamina e armazenado nos plastídios, essa glutamina também pode ser convertida em amidas e ureides, sendo dessa forma transportadas para a parte aérea das plantas (BUCHANAN et.al., 2015).

Dentro do grupo das Alpha-proteobacterias, os gêneros conhecidos de espécies nodulíferas compreendem: *Rhizobium*, *Bradyrhizobium*, *Paraburkholderia*, *Sinorhizobium*, *Mesorhizobium*, *Allorhizobium*, *Azorhizobium* e *Cupriavidus* (Moreira et al., 2006). *Bradyrhizobium*, que possui 82 espécies descritas, é um dos gêneros de rizóbios conhecidos por serem eficientes na FBN de espécies leguminosas de grãos como soja (ROCHA et al. 2021; COSTA et al. 2020; ZILLI et al. 2021), feijão-caupi (AYALEW et al. 2021; MICHEL et al., 2020; SOARES et al. 2014); leguminosas forrageiras (DE SÁ et al. 2019; OMARA et al. 2018), leguminosas para adubação verde (ARAÚJO et al., 2017; DA SILVA et al. 2021) e leguminosas florestais (LOPEZ et al. 2021; COSTA et al. 2021, 2018). Apesar das espécies superiores terem muitos relatos de simbioses com o gênero *Bradyrhizobium*, foram identificadas estirpes do gênero *Rhizobium* capazes de nodular espécies florestais, sendo uma delas *Platypodium elegans* (LOPEZ et al., 2021). Com exceção do inoculante para *Calliandra houstoniana* e *C. surinamensis*, não há inoculante com bactérias do gênero *Rhizobium* recomendadas para espécies florestais. Além disso não há espécies identificadas a nível de espécie que nodulam a espécie *Platypodium elegans* e *Machaerium nyctitans*.

### 2.2 Leguminosas florestais

O uso inadequado do solo tem gerado modificação das condições físicas, químicas e biológicas que podem levar a degradação do mesmo. Nesse processo os ciclos biogeoquímicos se vêm prejudicados, pois o sistema se encontra em desequilíbrio. A ausência de material orgânico impede que as funções do solo atuem de maneira sustentável, sendo necessária intervenção. O principal fator notado em áreas degradadas é a baixa disponibilidade de nutrientes, principalmente o nitrogênio, de forma que o uso de técnicas que aumentem sua

disponibilidade é requerido. A utilização de espécies florestais leguminosas pode ser uma alternativa para o problema, pois podem obter nitrogênio a partir da fixação biológica, além de translocar os nutrientes a camadas mais profundas do solo (SILVA 2013) e pode acelerar o processo de recuperação do ecossistema em áreas degradadas (RESENDE et al., 2013).

*Machaerium nyctitans* (Vell.) Benth. É uma espécie arbórea, perene, da família das leguminosas, comumente conhecida por “Jacarandá Bico-de-pato”, “Pau-de-óleo” ou “tatajubamateiga”. Segundo o banco de dados “Plantas do mundo online”, é uma espécie nativa do Brasil e da Argentina (POWO, 2022). No Brasil, destaca-se na região de Floresta Ombrófila Densa e Floresta Estacional Semidecidual, classificada como ativadora secundária inicial de sucessão (MEIRA JUNIOR et al. 2015), sendo recomendada para restauração ecológica (BARBOSA et al. 2017) e para a recuperação de áreas degradadas pela mineração (MEIRA JUNIOR et al. 2015). Além do uso na recuperação de áreas, essa espécie pode ser utilizada no paisagismo e em diversos objetos de carpintaria, como móveis e instrumentos musicais, por ser uma madeira resistente e sua seiva ter propriedades curativas (LEWIS et al. 2005). Essa espécie arbórea pode atingir até 40 metros de altura e possui uma copa ampla e espalhada com uma densa folhagem verde-escura. A madeira de *Macherium nyctitans* é altamente valorizada por sua durabilidade e resistência a cupins, tornando-se uma escolha popular para construção e carpintaria na região. *M. nyctitans* também é conhecido por suas propriedades medicinais, e sua casca e folhas têm sido tradicionalmente usadas pelos indígenas para tratar várias doenças, incluindo diarreia, febre e infecções de pele. No entanto, como muitas outras espécies da floresta tropical, o *M. nyctitans* está atualmente ameaçado pelo desmatamento e perda de habitat devido a atividades humanas, como extração de madeira, agricultura e mineração. Esforços de conservação são necessários para proteger esta espécie e seu habitat para as gerações futuras.

*Platypodium elegans* (Vogel) é uma espécie florestal, da família das leguminosas, de origem na América central e sul e encontrada em países como Panamá, Colômbia, Venezuela, Bolívia, Paraguai e uma ampla distribuição pelo Brasil. Possui diversos usos como madeira para construção civil, carpintaria e decoração (LEWIS 2005). É uma espécie recomendada para restauração de áreas degradadas por se tratar de uma espécie pioneira e por ser resistente a condições de alta luminosidade tanto durante a germinação quanto no decorrer do seu desenvolvimento. *Platypodium elegans*, também conhecido como "Faveiro", "Uruvalheiro", “Jacarandá do campo”, “Amendoim bravo” dentre outros, é comumente encontrado em florestas tropicais, pântanos e outros habitats de zonas úmidas. Pode atingir até 25 metros de altura e tem uma aparência distinta com folhas grandes e brilhantes que podem crescer até 60

cm de comprimento. A árvore também produz flores pequenas e perfumadas, de cor amarela ou branca, seguidas por vagens de sementes longas e achatadas. A madeira de *Platypodium elegans* é altamente valorizada por sua durabilidade e é comumente usada para construção e fabricação de móveis. As folhas e a casca da árvore também têm sido usadas na medicina tradicional para tratar várias doenças, incluindo febre, infecções e problemas digestivos. Assim como *M. nycitans*, *Platypodium elegans* está enfrentando ameaças de perda de habitat e desmatamento devido a atividades humanas, como extração de madeira, agricultura e mineração. Esforços de conservação são necessários para proteger esta espécie e seu habitat, e várias organizações estão trabalhando para promover práticas florestais sustentáveis e proteger os ecossistemas da floresta tropical.

Embora não existam estirpes aprovadas para a produção de inoculantes para *Machaerium nycitans* e *Platypodium elegans*, estudos têm demonstrado a capacidade dessa espécie de nodular (FARIA et al., 1998; LOPEZ et al. 2021). Além disso, essas espécies tiveram estirpes isoladas de seus nódulos e identificadas a nível de gênero a partir de sequenciamento do gene 16S rRNA e os genes *housekeeping gyrB* e *atpD*. Desta forma foi possível identificar tais estirpes bacterianas que nodulam *Machaerium nycitans* como pertencentes ao gênero *Bradyrhizobium* e que nodulam *Platypodium elegans* como pertencentes ao gênero *Rhizobium* (LOPEZ et al. 2021).

### 2.3 Técnicas de identificação bacteriana

A taxonomia de bactérias oriundas de nódulos de leguminosas tem mudado nos últimos 30 anos. Em primeiro momento eram realizadas análises fenotípicas como testes bioquímicos, nutricionais e sorológicos bem como as características morfológicas de colônias em placa de petri contendo meio YMA para separar os grupos (KAWAKA E MUOMA 2020). A microscopia também foi uma aliada na distinção de bactérias seja por coloração de gram ou absorção de vermelho Congo.

Depois evoluiu para a análise genética a partir do descobrimento da Taq polimerase, capaz de auxiliar na reação em cadeia da polimerase (PCR). Essa análise permitiu amplificar genes de interesse (16S rRNA e *Housekeeping genes*) e a partir disso levar o DNA bacteriano a sequenciamento de primeira geração (MOREIRA E SIQUEIRA 2006). O gene 16S rRNA é considerado um marcador genético modelo pois está presente em todos os organismos, é altamente conservado e possui aproximadamente 1500 pb de comprimento o que fornece informação suficiente para a taxonomia (CAPUTO et al. 2019). No entanto, as sequências do

gene 16S rRNA nem sempre permitem a discriminação de espécies (FERRAZ HELENE et al. 2022; HAQUE et al. 2017). O uso de *Housekeeping genes* para melhorar essa discriminação ganhou espaço pois o grau de divergência de sequências é maior que com 16S rRNA (FERRAZ HELENE et al. 2022; HAQUE et al. 2017).

O sequenciamento de genes envolvidos na simbiose e fixação de N<sub>2</sub> como *nodC* e *nifH* tem se mostrado importantes para a caracterização das bactérias e para auxiliar no entendimento das relações evolutivas desses microrganismos, a partir de transferências entre estirpes (LAU 2014).

Atualmente, ferramentas e técnicas moleculares modernas melhoraram consideravelmente a identificação dessas bactérias. Marcadores genéticos moleculares são mais sensíveis e precisos além de ter melhorado a disponibilidade de sequenciamento em bases de dados para distinção de espécies bacterianas intimamente relacionadas e detectam maior diversidade em comparação com métodos fenotípicos. Além disso hoje temos o sequenciamento do genoma completo das bactérias, que é um aliado na distinção entre espécies já identificadas melhorando a acurácia das comparações. Esse tipo de sequenciamento foi facilitado pelas novas técnicas chamadas de segunda e terceira geração. Genomas completos de *Bradyrhizobium*, *Rhizobium*, *Sinorhizobium*, *Mesorhizobium* e *Azorhizobium*, entre outros, foram sequenciados e disponibilizadas para uso público o que favorece a comparação e identificação até nível de espécie.

Todos os métodos e técnicas taxonômicas têm suas próprias fragilidades no estudo da diversidade e filogenia das bactérias de modo que a recomendação é alinhar diversas técnicas e assim obter uma identificação mais robusta. Estudos mostraram que a combinação de técnicas clássicas descritivas em conjunto com métodos modernos de sequenciamento molecular resultam na identificação precisa de novos táxons (BERG et al. 2020).



### 3. REFERÊNCIAS

- ARAÚJO, J.; FLORES-FÉLIX, J.D.; IGUAL, J.M.; PEIX, A.; GONZÁLEZ-ANDRÉS, F.; DÍAZ-ALCÁNTARA, C.A.; VELÁZQUEZ, E. *Bradyrhizobium cajani* sp. nov. isolated from nodules of *Cajanus cajan*. **Int J Syst Evol Microbiol.** 67:2236–2241, 2017. <https://doi.org/10.1099/ijsem.0.001932>
- AYALEW, T.; YOSEPH, T. PETRA, H.; CADISCH, G. Yield response of field-grown cowpea varieties to *Bradyrhizobium* inoculation. **Agronomy Journal** 113:3258-3268, 2021. <https://doi.org/10.1002/agj2.20763>
- BARBOSA, L.M.; SHIRASUNA, R.T.; DE LIMA, F.C.; ORTIZ, P.R.T.; BARBOSA, K.C.; BARBOSA, T.C. **Lista de espécies indicadas para restauração ecológica para diversas regiões do Estado de São Paulo**. Instituto de Botânica, São Paulo. 2017.
- BERG, G. et.al. Microbiome definition re-visited: old concepts and new challenges. **Microbiome.** 8:103, 2020. <https://doi.org/10.1186/s40168-020-00875-0>
- BUCHANAN, B. B.; GRUISSEM, W.; JONES, R. L. **Biochemistry e molecular biology of plants**. 2 ed. Pondicherry: Wiley Blackwell, 2015.
- CAPUTO, A.; FOURNIER, P.E.; RAOULT, D. Genome and pan-genome analysis to classify emerging bacteria. **Biol Direct** 14:5–5, 2019. <https://doi.org/10.1186/s13062-019-0234-0>
- COSTA, E.M.; GUIMARÃES, A.A.; CARVALHO, T.S.; RODRIGUES, T. L.; RIBEIRO, P.R.A.; LEBBE, L.; WILLEMS, A.; MOREIRA, F.M.S. *Bradyrhizobium forestalis* sp. nov., an efficient nitrogen-fixing bacterium isolated from nodules of forest legume species in the Amazon. **Arch Microbiol.** 200:743–752, 2018. <https://doi.org/10.1007/s00203-018-1486-2>
- COSTA, E.M.; RIBEIRO, P.R.A.; CARVALHO, T.S.; VICENTIN, R.P.; BALSANELLI, E.; SOUZA, E.M.; LEBBE, L.; WILLEMS, A.; MOREIRA, F.M.S. Efficient Nitrogen Fixing Bacteria Isolated from Soybean Nodules in the Semi arid Region of Northeast Brazil are Classified as *Bradyrhizobium brasilense* (Symbiovar Sojae). **Curr Microbiol.** 77:1746–1755, 2020. <https://doi.org/10.1007/s00284-020-01993-6>
- COSTA, E.M.; LIMA, W.; RIBEIRO, P.R.A.; MOREIRA, F.M.S. Acid and high-temperature tolerant *Bradyrhizobium* spp. Strains from Brazilian soils are able to promote *Acacia mangium* and *Stizolobium aterrimum* growth. **Symbiosis** 83:65–78, 2021. <https://doi.org/10.1007/s13199-020-00732-6>
- COURTY, P. E.; SMITH, P.; KOEGEL, S.; REDECKER, D.; WIPF, D. Inorganic nitrogen uptake and transport in beneficial plant root-microbe interactions. **Critical Reviews in Plant Sciences.** 34, 4– 16, 2015. <https://doi.org/10.1080/07352689.2014.897897>
- DE SÁ, O.A.A.L.; RIBEIRO, P.R.A.; RUFINI, M.; CRUVINEL, I.A.F.; CASAGRANDE, D.R.; MOREIRA, F.M.S. Microsymbionts of forage peanut under different soil and climate conditions belong to a specific group of *Bradyrhizobium* strains. **Appl Soil Ecol.** 143:201–212, 2019. <https://doi.org/10.1016/j.apsoil.2019.07.018>

- FARIA, S.M.; DE LIMA, H.C. Additional studies of the nodulation status of legume species in Brazil. **Plant and Soil**. 200:185–192, 1998.
- FERRAZ HELENE, L.C.; KLEPA, M.S.; HUNGRIA, M. New insights into the taxonomy of bacteria in the genomic era and a case study with rhizobia. **Int J Microbiol**. 2022:4623713, 2022. <https://doi.org/10.1155/2022/4623713>
- FIGOIRE, N.V.; FERREIRA, C.C.; DZEDZEJ, M.; MASSI, K.G. Monitoramento da Restauração de Plantio de Mudanças em Área de Preservação Permanente do Bioma Mata Atlântica Sudeste, Brasil. **Florestas**. 10 , 768, 2019. <https://doi.org/10.3390/f10090768>
- KAWAKA, F.; MUOMA, J. Distribution and phenotypic characteristics of common bean (*Phaseolus vulgaris* L.) nodulating bacteria in diverse soils. **Acta Agric Scand Section B Soil Plant Sci**. 70:564–571, 2020. <https://doi.org/10.1080/09064710.2020.1807595>
- HAQUE, A.; ENGEL, J.; TEICHMANN, S.A.; LÖNNBERG, T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. **Genome Med**. 9:75, 2017. <https://doi.org/10.1186/s13073-017-0467-4>
- JONES, D. L.; HEALEY, J. R.; WILLETT, V. B.; FARRAR, J. F.; HODGE, A. Dissolved organic nitrogen uptake by plants—An important N uptake pathway? **Soil Biology & Biochemistry**. 37, 413– 423, 2005.
- LAU M.C.Y.; CAMERON, C.; MAGNABOSCO, C.; BROWN, C.T.; SCHILKEY, F.; GRIM, S.; HENDRICKSON, S.; PULLIN, M.; SHERWOOD, L.B.; VAN HEERDEN, E.; KIEFT, T.L.; ONSTOTT, T.C. Phylogeny and phylogeography of functional genes shared among seven terrestrial subsurface metagenomes reveal N-cycling and microbial evolutionary relationships. **Front Microbiol**. 2014. <https://doi.org/10.3389/fmicb.2014.00531>
- LEWIS, G.P.; SCHRIRE, B.; MACKINDER, B.; LOCK, M. Legumes of the world. Royal Botanic Gardens, Kew, Richmond, Surrey. 2005.
- LOPEZ, B.D.O.; TEIXEIRA, A.F.S.; MICHEL, D.C.; GUIMARÃES, A.A.; COSTA, A.M.; COSTA, J.S.; PEREIRA, M.S.; DUARTE, B.L.M.; MOREIRA, F.M.S. Genetic and symbiotic characterization of rhizobia nodulating legumes in a mining área in southeast Brazil. **Sci Agric**. 79(2):e20200238, 2021. <https://doi.org/10.1590/1678-992X-2020-0238>
- MEIRA JUNIOR, M.S.; PEREIRA, I.M.; MACHADO, E.L.M.; MOTA, S.L.L.; OTONI, T.J.O. Espécies potenciais para recuperação de áreas de floresta estacional semidecidual com exploração de minério de ferro na serra do espinhaço. **Biosci J** 31(1)283-295, 2015.
- MIRA, E.; ROUSTEAU, A.; TOURNEBIZE, R.; LABBOUZ, L.; ROBERTO, M.; EVETTE, A. A. Conservação e Restauração de Florestas Riparianas ao longo das Margens dos Rios do Caribe Usando Leguminosas. **Sustentabilidade**. 14, 3709, 2022. <https://doi.org/10.3390/su14073709>
- MICHEL, D.C.; GUIMARÃES, A.A.; COSTA, E.M.; CARVALHO, T.S.; BALSANELLI, E.; WILLEMS, A.; SOUZA, E.M.; MOREIRA, F.M.S. *Bradyrhizobium uaiense* sp. nov., a new highly efficient cowpea symbiont. **Arch microbiol**. 202:1135–1141, 2020. <https://doi.org/10.1007/s00203-020-01827-w>

MOREIRA, F.M.S.; SIQUEIRA, J.O.; BRUSSAARD, L. **Soil Biodiversity in Amazonian and other Brazilian Ecosystem**. CABI Publishing, London, UK, 2006.

NASHOLM, T.; EKBLAD, A.; NORDIN, A.; GIESLER, R.; HOGBERG, M.; HOGBERG, P. Boreal forest plants take up organic nitrogen. **Nature**. 392, 914–916, 1998.

<https://doi.org/10.1038/31921>

OMARA, A.E.; EL-GAAFAREY, T. Alleviation of salinity stress effects in forage cowpea (*Vigna unguiculata* L.) by *Bradyrhizobium* sp. Inoculation. **Microbiol Res J Int**. 23(3):1-16, 2018.

POWO (Plants of the World Online). *Machaerium nycitans* (Vell.) Benth. Disponível em: <https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:505530-1>. Acesso em 12 de julho de 2022

PRAGYA, D.; POONAM, J. Restoration of Forests: Human Concern. **International Journal for Research in Applied Sciences and Biotechnology**. 9(3), 85–89, 2022.

<https://doi.org/10.31033/ijrasb.9.3.15>

RESENDE, A. S.; CHAER, G. M.; CAMPELLO, E. F. C.; SILVA, A. P.; LIMA, K. D. R.; CURCIO, G. R. Uso de leguminosas arbóreas na recuperação de áreas degradadas. In: **Tópicos em Ciências do Solo**, Viçosa: SBCS, v. 8, p. 71-92, 2013.

ROCHA, L.B.; DA COSTA, E.M.; COSTA, G.S.; DA SILVA, J.K.P.; RIBEIRO, E.C.; BORGES, J.F.; DA SOUSA, A.P.M.; MOREIRA, F.M.S. *Bradyrhizobium brasilense* as an efficient soybean microsymbiont in two contrasting soils of the southwestern region of Piauí (Cerrado biome). **Rev Bras Cienc Agrar**. 16(3):e212, 2021.

<https://doi.org/10.5039/agraria.v16i3a212>

SILVA, J.S.; OLIVEIRA, D.P.; RUFINI, M.; SILVA JUNIOR, C.L.; BAPTISTA, M.V.B.D.; SILVA, L.C.; ARAGÃO, O.O.S.; PEREIRA, T.A.; MOREIRA, F.M.S. New strains of *Bradyrhizobium* enrich plant biomass nitrogen content in *Crotalaria* for use as a green manure. **Bragantia**. 80:e4021, 2021. <https://doi.org/10.1590/1678-4499.20200469>

SILVA, A. B.; LIRA JUNIOR, M. A.; DUBEUX JUNIOR, J. C. B.; FIGUEIREDO, M. V. B.; VICENTIN, R. P. Estoque de serapilheira e fertilidade do solo em pastagem degradada de *Brachiaria decumbens* após implantação de leguminosas arbustivas e arbóreas forrageiras. **Revista Brasileira de Ciência do Solo**. Viçosa, v. 37, n. 2, p. 502-511, 2013.

SOARES, B.L.; FERREIRA, P.A.A.; OLIVEIRA-LONGATTI, S.M.; MARRA, L.M.; RUFINI, M.; ANDRADE, M.J.B.; MOREIRA, F.M.S. Cowpea symbiotic efficiency, pH and aluminum tolerance in nitrogen-fixing bacteria. **Sci Agric**. 71(3)171-180, 2014.

<https://doi.org/10.1590/S0103-90162014000300001>

ZILLI, J.E.; PACHECO, R.S.; GIANLUPPI, V.; SMIDERLE, O.J.; URQUIAGA, S.; HUNGRIA, M. Biological N<sub>2</sub> fixation and yield performance of soybean inoculated with *Bradyrhizobium*. **Nutr Cycl Agroecosyst**. 119:323–336, 2021.

<https://doi.org/10.1007/s10705-021-10128-7>

**SEGUNDA PARTE - ARTIGOS**

**ARTIGO 1**

***Bradyrhizobium versatile* sp. nov., *Bradyrhizobium cerradense* sp. nov. and *Bradyrhizobium nycitiae*, nitrogen-fixing bacterial strains isolated from a native leguminous tree from Cerrado ecoregion in Brazil**

Artigo formatado nas normas da revista Archives in Microbiology

Bruna Daniela Ortiz Lopez<sup>1</sup>, Daniele Cabral Michel<sup>1</sup>, Lucas Lenin Resende de Assis<sup>1</sup>, Fatima Maria de Souza Moreira<sup>1\*</sup>

<sup>1</sup> Setor de Biologia, Microbiologia e Processos Biológicos do Solo, Departamento de Ciência do Solo, Universidade Federal de Lavras, Campus UFLA, 37200-000, Lavras, Minas Gerais, Brazil.

\*Corresponding author: Tel.: +55 35 3829 1254; fax: +55 35 3829 1251. E-mail address: fmoreira@ufla.br (F.M. de Souza Moreira).

## Abstract

Several new species of symbiotic nitrogen-fixing bacteria have already been described, mainly belonging to the genus *Bradyrhizobium*. However, there is no identification at the species level of strains that nodulate and fix N<sub>2</sub> in symbiosis with the native legume *Machaerium nycitans* (Vell.) Benth. In our study, the average nucleotide identity (ANI) between these four strains and 67 type strains of already described species showed values of 95.53, 95.52, 94.57 and 94.09% for strains UFLA01-814 and UFLA01-860, UFLA01-839 and UFLA01-1164 with the closest species: *Bradyrhizobium elkanii* USDA 76<sup>T</sup>, *Bradyrhizobium tropiciagri* CNPSo 1112<sup>T</sup>, *Bradyrhizobium daqingense* CCBAU 15774<sup>T</sup>, respectively. The strain UFLA01-839 is resistant to all the antibiotics tested, different from the other strains and *B. elkanii* USDA76<sup>T</sup>. Therefore, the strains UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164 represent three new species: *B. versatile* sp. nov. (type strain UFLA01-839<sup>T</sup>), *B. cerradense* sp. nov. (type strain UFLA01-860<sup>T</sup>), *B. nycitiae* sp. nov. (type strain UFLA01-1164<sup>T</sup>), respectively.

**Keywords:** taxonomy, tree species, biodiversity, phylogeny

## Introduction

The process of biological nitrogen fixation (BNF) is crucial for ecosystem sustainability. It is mediated by microorganisms capable of transforming atmospheric nitrogen (N<sub>2</sub>) into ammonia (NH<sub>3</sub>). Ammonia, in turn, is assimilated by plants, allowing for a reduction in the use of mineral fertilizers. Among the bacteria responsible for this transformation, also called diazotrophic microorganism, the bacteria commonly called rhizobia stand out because they establish symbiosis with legume species, forming root and exceptionally stem nodules, including several forest species (Faria et al. 1998; Moreira et al. 1993). *Bradyrhizobium*, which has 82 described species, is one of the rhizobia genera known to be efficient in the BNF of leguminous species of grains such as soybean (Rocha et al. 2021; Costa et al. 2020; Zilli et al. 2021), cowpea (Ayalew et al. 2021; Michel et al., 2020; Soares et al. 2014); forage legumes (De Sá et al. 2019; Omara et al. 2018), legumes for green manuring (Araújo et al., 2017; Da Silva et al. 2021) and forest legumes (Lopez et al. 2021; Costa et al. 2021, 2018).

*Machaerium nycitans* (Vell.) Benth. it is an arboreal, perennial species. According to the “Plants of the world online” database, it is a native species of Brazil and Argentina (Powo, 2022). In Brazil, it stands out in the region of Dense Ombrophylous Forest and Semideciduous Seasonal Forest, classified as an initial secondary activator of succession (Meira Junior et al. 2015), being recommended for ecological restoration (Barbosa et al. 2017) and for the recovery of degraded areas by mining (Meira Junior et al. 2015). In addition to its use for recovering areas, this species can be used in landscaping and for various carpentry objects such as furniture and musical instruments because it is a resistant wood, and its sap has healing properties (Lewis et al. 2005).

Although there are no strains approved for the production of inoculants for *Machaerium nycitans* are available, several studies have demonstrated the ability of this species to nodulate (Faria et al., 1998; Lopez et al. 2021). Besides, no isolated strains of *M. nycitans* have been identified at the species level. Based on previous studies with the 16S rRNA gene and housekeeping genes *gyrB* and *atpD*, it was possible to identify bacterial strains that nodulate *Machaerium nycitans* as belonging to the genus *Bradyrhizobium* (Lopez et al. 2021). Therefore, the

present study aimed to identify at the species level using advanced techniques of genomic analysis strains of *Bradyrhizobium* that nodulate *Machaerium nycitans*: UFLA01-814, UFLA01-860, UFLA01-1164, and UFLA01-839, that was efficient in axenic conditions.

## **Material and methods**

### **Origin of strains**

Phylogenetic analysis by sequencing the 16S rRNA, *gyrB* and *atpD* genes indicated that strains UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164 isolated from nodules of *Machaerium nycitans* in a nursery could be new species (Lopez et al. 2021). These strains are deposited in the plant growth promoting bacterial collection of the Sector of Biology, Microbiology and Soil Biological Processes/Soil Science Department at the Federal University of Lavras, Brazil.

### **Sequencing and quality checking of genomes**

The strains UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164 were grown in culture medium 79 (Fred and Waskman 1928) for subsequent DNA extraction. The genomic DNA of the four strains was extracted using the Wizard® Genomic DNA Purification Kit, following the manufacturer's recommendations. The quality and concentration of the extracted DNA was analyzed using a Colibri spectrometer. Samples above 100 ng/microliter were lyophilized and sent for genome sequencing. Paired reads (2 × 250 bases) were sequenced with the MiSeq Reagent kit 500v2 (Illumina) on the MiSeq platform (Illumina) (Table 1). The sequencing data were analyzed in the KBase (Arkin AP et al. 2018) online platform as described below: genome import, sequence assembly using SPADES 3.15.3 (Bankevich et al. 2012) (Table 1), genome quality verification with CheckM (Parks et al. 2015) (Table 2) and annotation of the genomes with Rast (Aziz et al, 2008; Overbeek et al, 2014; Brettin 2015).

### **Genomic comparison (ANI)**

Average nucleotide identity (ANI), a pairwise comparison between two or more genomic sequences, has been an alternative to the DNA-DNA hybridization technique with greater robustness and accuracy for species identification (Konstantinidis and Tiedje, 2005; Michel et al. 2020, 2021). Among the options available for this comparison, an online platform was used (Rodriguez-R and Konstantinidis 2016), which uses the ANI calculator to estimate the average nucleotide identity using the best hits (unidirectional ANI) and the best reciprocal hits (bidirectional ANI) between two genomic sets (Goris et al. 2007). This algorithm allows for the comparison of up to 45 different sequences at a time, which facilitates the comparison of strains with all 67 type species of the same genus, whose genomes are available at the NCBI. The platform generates an identity matrix, in which the percentages of similarity between sequences are observed. Based on previous publications, values below 96% similarity have been used to classify the study sequences as species (Richter and Rosselló-Morá 2009; Avontuur, 2022).

### Multilocus Sequence Analysis (MLSA) of housekeeping genes and symbiotic genes

From the housekeeping genes studied, only the *recA* gene was found in all the 14 type strains without genomes information. The *rpoB* and *gyrB* genes were found in 10 *Bradyrhizobium* type strains and *glnII* and *atpD* in 11 type strains. Thus, two phylogenetic trees based on housekeeping genes were constructed. In these trees were also included those housekeeping genes extracted from the genomes of the other 67 type strains. Based on these sequences of the housekeeping genes and those of the strains UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164, a phylogenetic tree was constructed for each gene in the MEGA10 software. From the species for which it was possible to obtain the sequences of all housekeeping genes, the concatenated phylogenetic analysis was continued using the SeaView software that included the *gyrB-recA-rpoB* and *atpD-glnII-recA* genes. In addition, for the construction of phylogenetic trees, the sequences of the *nifH* and *nodC* genes of 59 and 53 type strains, respectively, and of the strains UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164 were obtained from their genomes. The sequences of all genes studied (housekeeping genes, *nifH* and *nodC*) were aligned using MUSCLE (Edgar 2004). The phylogenetic trees were constructed using the statistical methods neighbor-joining (NJ) (Saitou and Nei 1987) and maximum likelihood (ML) (Felsenstein 1981), with the General Time Reversible (GTR) model with Gamma correction value and % of sites invariant (G + I) with bootstrap of 1000 replications for *gyrB-rpoB-recA* and *atpD-glnII-recA* trees and Tamura 3 model (Tamura 1992) with Gamma correction value and % of sites invariant (G + I) and bootstrap of 1000 repetitions for trees of *nifH* and *nodC*. Models tested in MEGA X and chosen according to the lowest BIC (Bayesian Information Criterion) value (Nei and Kumar, 2000; Kumar et al., 2018) (model details in table S2, S3, S4 and S5).

### GTDB.tk and TYGS

The genomic sequences of strains UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164 were submitted to GTDB.tk v1.7.0 analysis (Chaumeil et al., 2019), which is a software toolkit capable of assigning objective taxonomic classification to genomes of bacteria, based on genome-wide taxonomy databases (Parks et al., 2020; 2018).

Furthermore, these sequences were uploaded to TYGS (Type Genome Server), a free bioinformatics platform available at <https://tygs.dsmz.de>, to perform genome-wide taxonomic analysis (Meier-Kolthoff JP et al., 2022; 2019). For the determination of the closest type strain genomes, first, the comparison of all study genomes with all type strain genomes available in the TYGS database was done by means of the MASH algorithm, a fast approximation of intergenomic relatedness (Ondov et al. al., 2016) and the ten strain types with the smallest MASH distances chosen by the genome itself. Second, an additional set of ten closely related type strains were determined from the 16S rRNA gene sequences. These were extracted from the genomes using RNAmmer (Lagesen and Hallin 2007) and each sequence was subsequently BLASTed (Cmacho et al., 2009) against the 16S rDNA gene sequence of each of the currently 18452 type strains available in the TYGS database. This was used as a proxy to find the top 50 matching type strains (according to bitscore) for each genome and subsequently calculated the precise distances using the Genome BLAST Distance Phylogeny (GBDP) approach under the 'coverage' algorithm and the d5 distance formula (Meier-Kolthoff JP et al, 2013). These distances were finally used to determine the 10 closest strain-type genomes for each of the genomes.



Paired comparison of genomic sequences was conducted using GBDP and precise intergenomic distances inferred under the 'trimming' algorithm and d5 distance formula (Meier-Kolthoff JP et al., 2013). 100 distance replicates were calculated each. Digital DNA-DNA Hybridization (dDDH) values and confidence intervals were calculated using the recommended GGDC 3.0 settings (Meier-Kolthoff JP et al., 2022; 2013). The resulting intergenomic distances were used to infer a balanced minimal evolution tree with branching support via FASTME 2.1.6.1 including SPR post-processing (Lefort V et al., 2015). Branching support was inferred from 100 pseudo-bootstrap replicates each. Trees were rooted at the midpoint (Farris JS, 1972) and visualized with PhyD3 (Kreft L et al., 2017).

Type-based grouping of species using a 70% dDDH radius around each of the 45 type strains was done as previously described (Meier-Kolthoff and Göker M, 2019). The resulting groups are shown in Table 3. Subspecies grouping was done using a dDDH threshold of 79% as previously introduced (Meier-Kolthoff et al., 2014).

### Phenotypical tests

The growth of strains UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164, were evaluated in culture medium 79 at different concentrations of NaCl (w/v) (0.25, 0.50, 0.75 and 1%), different pH levels (4.0, 5.5, 6.8, 8.0, 9.0 and 10) and at different temperatures (15, 20, 28, 34, 37 and 40°C). Furthermore, strains susceptibility to antibiotics was tested using antibiogram discs containing: nalidixic acid 30 µg mL<sup>-1</sup>, amoxicillin 10 µg mL<sup>-1</sup>, ampicillin 10 µg mL<sup>-1</sup>, azithromycin 15 µg mL<sup>-1</sup>, cefuroxime 30 µg mL<sup>-1</sup>, clarithromycin 15 µg mL<sup>-1</sup>. Chloraphenicol 30 µg mL<sup>-1</sup>, doxycillin 30 µg mL<sup>-1</sup>, gentamicin 10 µg mL<sup>-1</sup>, rifampicin 30 µg mL<sup>-1</sup>, tetracycline 30 µg mL<sup>-1</sup>, vancomycin 30 µg mL<sup>-1</sup>, neomycin 30 µg mL<sup>-1</sup>, kanamycin 30 µg mL<sup>-1</sup>. The ability of the strains to assimilate different carbon (L-arabinose, D-arabinose, L-asparagine, citric acid, glycerol, D-fructose, glycine, D-glucose, glutamine, L-glutamic acid, malic acid, lactose, maltose, mannitol, methionine, sodium lactate and sucrose) and nitrogen sources (L-asparagine, L-arginine, casein hydrolyzate, L-cysteine, glycine, L-glutamic acid, tryptophan and l-methionine) was evaluated in medium 79 modified according to Costa et al. (2017). As a complementary phenotypic test, the API-BioMérieux test was performed, which allows us to quickly perform 20 different biochemical tests with each of the strains. Cell optical density turbidity was adjusted to the 5.0 McFarland standard.

## Results and discussion

### Multilocus Sequence Analysis of housekeeping genes

According to the MLSA carried out with the genes *gyrB*, *recA* and *rpoB* (Fig. 1), the strains UFLA01-814 and UFLA01-860 are positioned in a separate clade from all *Bradyrhizobium* species included in the phylogeny. The more close species were *Bradyrhizobium australafricanum* WSM 4400<sup>T</sup> and *Bradyrhizobium ferriligni* CCBAU 51502<sup>T</sup> with 98.67% and 97.95% similarity, respectively. The strain UFLA01-839 was phylogenetically separated from *Bradyrhizobium uaiense* UFLA03-164<sup>T</sup> and *Bradyrhizobium acacia* 10BB<sup>T</sup> with 95.68% and 96.41% similarity, respectively. The strain UFLA01-1164 is positioned in a separate clade, whose closest species is

*Bradyrhizobium daqingense* CCBAU 15774<sup>T</sup> with 95.94% similarity. This positioning showed high statistical support with 67% of bootstrap (Table 4).

According to the MLSA carried out with the genes *atpD*, *glnII* and *recA* (Fig. 2), the strains UFLA01-814 and UFLA01-860 have 100% similarity to each other and they are positioned in a separate clade from all already described species from *Bradyrhizobium* genus and from the other three UFLA strains. This positioning showed high statistical support with 99% of bootstrap. The closest species from the UFLA 01-814 was *B. australafricanum* WSM 4400<sup>T</sup> 98.71%, *B. brasilense* UFLA03-321<sup>T</sup> 97.09% and *B. pachyrhizi* PAC 48<sup>T</sup> 96.81 % similarity. The strain UFLA01-839 positioned separately from the species already described included in the phylogenetic tree with 79% of bootstrap. The closest species from UFLA01-839 was *Bradyrhizobium tropiciagri* CNPSo 1112<sup>T</sup> 97.81% similarity. The strain UFLA01-1164 is also in a separate clade whose closest species are *Bradyrhizobium americanum* CMVU 44<sup>T</sup> 98.55% and *Bradyrhizobium daqingense* CCBAU 15774<sup>T</sup> 96.53% similarity (Table 5).

In recent years, the description of species, mainly in the genus *Bradyrhizobium*, has used a greater number of housekeeping genes (five or six) for a more accurate positioning of potentially new strains (Costa et al. 2020; Michel et al. 2021). Although, studies using five housekeeping genes, for example, have already pointed out the need to review the species delimitation parameter within the genus *Bradyrhizobium*, showing that the MLSA analysis can present values above 97% even for different species (Klepa et al. 2021). For this reason, associated with MLSA, genomic comparisons, such as ANI has been used (De Lajudie 2019).

## Genomic comparison

### Average Nucleotide Identity

The ANI technique determines the evolutionary distance between a pair of lineages from the similarity values of homologous genomic regions shared between the tested genomes (Konstantinidis and Tiedje 2005; Kim et al. 2014). The cut-off values considered to determine whether the species are new is below 95-96% (Goris et al. 2007; Konstantinidis and Tiedje 2005; De Lajudie et al., 2019; Avontuur et al., 2019). The strains UFLA01-814 and UFLA01-860 have greater similarity with the species *Bradyrhizobium elkanii* USDA 76<sup>T</sup> (95.53% and 95.51%) and *Bradyrhizobium brasilense* UFLA 03-321<sup>T</sup> (95.13% and 95.12%), respectively. For strains UFLA01-839 and UFLA01-1164, ANI values below 95% were obtained. Among all the described species of *Bradyrhizobium*, strain UFLA01-839 had greater similarity with *Bradyrhizobium tropiciagri* CNPSo 1112<sup>T</sup> with percentage of 94.57% and strain UFLA01-1164 had greater similarity with *Bradyrhizobium daqingense* CCBAU 15774<sup>T</sup> at 94.09%. More details in table 6.

The content of G + C in the DNA of strains UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164 were 63.76 mol%, 63.98 mol%, 63.68 mol% and 63.78 mol%, respectively. The percentage agrees with the range of 62 to 66 mol% found in species of the genus *Bradyrhizobium* (Kuykendall et al. 1992; Xu et al. 1995; Chahboune et al. 2011; Ramírez-Bahena et al. 2013, 2016).

### Comparison of phylogeny obtained by MLSA and ANI

The elaboration of phylogenetic analysis using MLSA has been used for a long time to distinguish bacterial species from each other and to present evidence of new species. With the advancement of molecular biology and the wide availability of genomic analyses, it was possible to observe interspecific differences with greater accuracy and precision, since it considers the complete molecular content of the bacteria to the detriment of a small portion obtained with some concatenated genes. So why would it be necessary to use outdated techniques to analyze new species? Currently, there are 67 type strains of the *Bradyrhizobium* genus with their sequenced genome available in international databases. However, 82 species of this genus have already been described. To obtain a comparison regarding these species without their sequenced genome, it is still necessary to use the MLSA. Furthermore, it is important to point out that, based on the results obtained, the closest similarity of the strains studied in the present work corroborates the highest ANI values found, so that the results obtained by MLSA are still valid, with the exception of the concatenated tree of *atpD-glnII-recA* genes. On the other hand, if we considered only MLSA, it would be more likely to exclude certain strains from the description of species, since the values obtained can reach little more than 97% similarity and, in order to maintain a safety interval, values below 95% are usually accepted. Therefore, the use of both techniques is favorable to obtain a more complete and accurate picture.

#### **Phylogenetic analysis of *nifH* and *nodC* genes**

The phylogeny of genes related to biological nitrogen fixation (*nifH*) and nodulation (*nodC*) generally differs from the phylogeny of the 16S rRNA and housekeeping genes. As shown in Fig. 3 and 4, the studied strains are positioned in clades separate from those observed in Fig. 1 and 2. This occurs due to the horizontal transfer of genes located in “chromosomal islands” in the evolutionary process. This transfer commonly occurs between bacterial species and may have a geographic influence, that is, bacteria recruit genetic apparatus to maintain their cellular functions induced by the host and depending on the environment (Provorov et al., 2022). The transfer of symbiotic genes to bacteria adapted to local soil conditions may allow bacteria that were previously incompatible for nitrogen nodulation/fixation to become legume symbionts (Andrews et al. 2018). In a study with in vitro tests of symbiotic gene transfer, it was possible to observe post-transfer modifications in the DNA of bacteria that do not have the nodulation and fixation capacity to adapt to the new mutation by increasing the frequency of nitrogen-fixing variants in a population not fixative (Moura et al. 2020). The phylogeny of symbiotic genes may reveal symbiovars within species, as demonstrated by Costa et al. (2020) with *Bradyrhizobium* species.

Forest soils may harbor non-symbiotic species of the genus *Bradyrhizobium* (VanInsberghe et al., 2015) as well as rhizobia that do not need nod factors to induce nodulation (Giraud et al., 2007). This may be the case of the UFLA01-814 strain, for which it was not possible to detect the *nodC* gene.

#### **Type-based species and subspecies clustering**

The taxonomic identification of the closest strains is found in Table 3. In summary, the grouping yielded 35 species groups and the provided query strains were assigned to 3 of them. Additionally, strains were located in 3 of the 37 subspecies groups. For strains UFLA01-814 and UFLA01-860, a greater proximity to the species *Bradyrhizobium elkanii* USDA76 was observed. For the UFLA01-839 strain, a closer proximity to the species *Bradyrhizobium tropiciagri* SEMIA 6148 was observed. As for the strain UFLA01-1164, a greater proximity to the species

*Bradyrhizobium daqingense* CGMCC 1.10947 was observed. This result corroborates what was found in the GTDB.tk, ANI and MLSA of the present study. Graphical representation from the phylogenetic trees of the 16S rRNA gene and the complete genome can be seen in figure 5 and 6.

### Phenotypical test

The main characteristics that differentiate the strains of the present study and *Bradyrhizobium elkanii* USDA76<sup>T</sup> are shown in table 7. All strains grow at temperatures from 20 to 34°C. With the exception of strain UFLA01-1164, which grows from pH 5.5 onwards, the other strains grow in a wide pH range (4.0 to 10.0). Strains UFLA01-814, UFLA01-860 and UFLA01-839 are tolerant to salinity up to 0.75% as well as USDA76<sup>T</sup>, while strain UFLA01-1164 tolerates up to 0.5% of NaCl. Although UFLA01-814 and UFLA01-860 are 100% similar according to ANI, they behaved differently in relation to carbon and nitrogen sources, confirming the fact that they are not clones. Strain UFLA01-814 was grown in a medium containing glycerol, different from UFLA01-860. The strain UFLA01-860 grew in a medium containing glycine and methionine as a source of C, different from UFLA01-814. Strain UFLA01-860 also grows on a medium containing L-methionine and hydrolyzed casein as nitrogen sources, different from UFLA01-814. Strain UFLA01-839 is resistant to all tested antibiotics, whereas UFLA01-814 and UFLA01-860 are sensitive to neomycin, as well as *B. elkanii* USDA 76<sup>T</sup> (Delamuta JRM et al, 2015), and rifamycin. Strain UFLA01-1164 is sensitive to ampicillin, amoxicillin and cefuroxime. More details are found in the species description. Detailed images of bacteria from samples from UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164 are shown in figure 7.

Description of *Bradyrhizobium versatile* sp. nov.

*Bradyrhizobium versatile* (ver.sa.ti.le N.L. neut. adj. versatile for its ability to resist all tested antibiotics).

The cells are gram-negative rods (approximately  $2.8 \times 0.7 \mu\text{m}$ ) and are aerobic and non-spore forming. Colonies are greater than 1 mm in diameter at 5 days, are circular, and produce an alkaline reaction in medium 79 with mannitol as a carbon source and bromothymol blue as an indicator. The type strain tolerates up to 0.75% NaCl. It grows at pH levels of 4.0 to 10.0 and at temperatures ranging from 20 to 34°C, with optimal growth at 28°C. It is resistant to ampicillin ( $10 \mu\text{g mL}^{-1}$ ), amoxicillin ( $10 \mu\text{g mL}^{-1}$ ), doxycycline ( $30 \mu\text{g mL}^{-1}$ ), nalidixic acid ( $30 \mu\text{g mL}^{-1}$ ), tetracycline ( $30 \mu\text{g mL}^{-1}$ ), neomycin ( $30 \mu\text{g mL}^{-1}$ ), chloramphenicol ( $30 \mu\text{g mL}^{-1}$ ), clarithromycin ( $15 \mu\text{g mL}^{-1}$ ), cefuroxime ( $30 \mu\text{g mL}^{-1}$ ), gentamicin ( $10 \mu\text{g mL}^{-1}$ ) kanamycin ( $30 \mu\text{g mL}^{-1}$ ), vancomycin ( $30 \mu\text{g mL}^{-1}$ ), rifamycin ( $30 \mu\text{g mL}^{-1}$ ) and azithromycin ( $15 \mu\text{g mL}^{-1}$ ). Regarding carbon sources, it assimilates mannitol, lactose, sucrose, L-glutamic and glutamic acid; weakly assimilates malic acid, sucrose and glucose. Does not assimilate sodium lactate, glycerol, maltose, fructose, citric acid, L-methionine and glycine. Regarding nitrogen sources, it assimilates hydrolyzed L-cysteine, L-asparagine, tryptophan, glutamic acid and casein. Weakly assimilates L-methionine. Does not assimilate glycine. The G+C content of UFLA 01-839<sup>T</sup> is 63.98 mol%.

The type strain UFLA 01-839<sup>T</sup> was isolated from effective nodules of *Machaerium nycitans* (Vell.) Benth. seedlings, from Minas Gerais, Brazil. The Genbank/NCBI accession numbers of the genome is xxxx. The housekeeping genes were extracted from the UFLA01-839<sup>T</sup> type strain genome.

Description of *Bradyrhizobium cerradense* sp. nov.

*Bradyrhizobium cerradense* (cer.ra.den.se. N.L. neut. Adj. cerradense of Cerrado, referring to the fact that strains were isolated from Cerrado ecoregion in southeastern Brazil).

The cells are gram-negative rods (approximately  $3.0 \times 0.9 \mu\text{m}$ ) and are aerobic and non-spore forming. Colonies are greater than 1 mm in diameter at 5 days, are circular, and produce an alkaline reaction in medium 79 with mannitol as a carbon source and bromothymol blue as an indicator. The type strain tolerates up to 0.75% NaCl. It grows at pH levels of 4.0 to 10.0 and at temperatures ranging from 20 to 34°C, with optimal growth at 28°C. It is resistant to ampicillin ( $10 \mu\text{g mL}^{-1}$ ), amoxicillin ( $10 \mu\text{g mL}^{-1}$ ), doxycycline ( $30 \mu\text{g mL}^{-1}$ ), nalidixic acid ( $30 \mu\text{g mL}^{-1}$ ), tetracycline ( $30 \mu\text{g mL}^{-1}$ ), chloramphenicol ( $30 \mu\text{g mL}^{-1}$ ), clarithromycin ( $15 \mu\text{g mL}^{-1}$ ), cefuroxime ( $30 \mu\text{g mL}^{-1}$ ), gentamicin ( $10 \mu\text{g mL}^{-1}$ ), kanamycin ( $30 \mu\text{g mL}^{-1}$ ), vancomycin ( $30 \mu\text{g mL}^{-1}$ ), and azithromycin ( $15 \mu\text{g mL}^{-1}$ ), poorly resistant to neomycin ( $30 \mu\text{g mL}^{-1}$ ) and rifamycin ( $30 \mu\text{g mL}^{-1}$ ), forming a small halo. As for carbon sources, it assimilates glucose, mannitol, lactose, sucrose, L-glutamic, glutamic acid and glycine; weakly assimilates fructose, malic acid, sucrose and L-methionine. Does not assimilate sodium lactate, glycerol, maltose, citric acid. Regarding nitrogen sources, it assimilates hydrolyzed L-cysteine, L-asparagine, tryptophan, glutamic acid, L-methionine and casein. Does not assimilate glycine and L-methionine. The G+C content of UFLA 01-860<sup>T</sup> is 63.68 mol%.

The type strain UFLA 01-860<sup>T</sup> was isolated from effective nodules of *Machaerium nyctitans* (Vell.) Benth. seedlings, from Minas Gerais, Brazil. The Genbank/NCBI accession numbers of the genome is xxxx. The housekeeping genes were extracted from the UFLA01-860<sup>T</sup> type strain genome.

Description of *Bradyrhizobium nyctitae* sp. nov.

*Bradyrhizobium nyctitae* (nyc.ti.tae. N.L. neut. adj. nyctitae of the leguminous tree specie from which the strain was isolated from its nodules, *Machaerium nyctitans*).

The cells are gram-negative rods (approximately  $4.4 \times 0.35 \mu\text{m}$ ) and are aerobic and non-spore forming. Colonies are greater than 1 mm in diameter at 5 days, are circular, and produce an alkaline reaction in medium 79 with mannitol as a carbon source and bromothymol blue as an indicator. The type strain tolerates up to 0.50% NaCl. It grows at pH levels of 5.5 to 10.0 and at temperatures ranging from 20 to 34°C, with optimal growth at 28°C. It is resistant to doxycycline ( $30 \mu\text{g mL}^{-1}$ ), nalidixic acid ( $30 \mu\text{g mL}^{-1}$ ), tetracycline ( $30 \mu\text{g mL}^{-1}$ ), neomycin ( $30 \mu\text{g mL}^{-1}$ ), chloramphenicol ( $30 \mu\text{g mL}^{-1}$ ), clarithromycin ( $15 \mu\text{g mL}^{-1}$ ), gentamicin ( $10 \mu\text{g mL}^{-1}$ ), kanamycin ( $30 \mu\text{g mL}^{-1}$ ), vancomycin ( $30 \mu\text{g mL}^{-1}$ ), rifamycin ( $30 \mu\text{g mL}^{-1}$ ) and azithromycin ( $15 \mu\text{g mL}^{-1}$ ). It is sensitive to ampicillin ( $10 \mu\text{g mL}^{-1}$ ), amoxicillin ( $10 \mu\text{g mL}^{-1}$ ) and cefuroxime ( $30 \mu\text{g mL}^{-1}$ ). Regarding carbon sources, it assimilates fructose, L-asparagine, sucrose, glucose. Weakly assimilates lactose, glutamic acid. Does not assimilate glycerol, maltose, malic acid, citric acid, L-methionine, glycine and L-glutamine. The G+C content of UFLA 01-1164<sup>T</sup> is 63.87 mol%.

The type strain UFLA 01-1164<sup>T</sup> was isolated from effective nodules of *Machaerium nyctitans* (Vell.) Benth. seedlings, from Minas Gerais, Brazil. The Genbank/NCBI accession numbers of the genome is xxxx. The housekeeping genes were extracted from the UFLA01-1164<sup>T</sup> type strain genome.

**Acknowledgements** We thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Process: 142473/2018-9), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (CAPES/PROEX AUXPE593/2018) and the Fundação de Amparo e Pesquisa de Minas Gerais (Fapemig) (CAG-RED- O0330- 16) for financial support and for granting scholarships. This research is associated with the Brazilian Instituto Nacional de Ciência e Tecnologia Biodiversidade do Solo (Soil Biodiversity/INCT-CNPq). The authors would also like to thank the Fapemig, Capes, CNPq and Finep (Financiadora de Estudos e Projetos) for maintaining the electron microscopy laboratory.

## References

- Andrews M, De Meyer S, James EK, Stepkowski T, Hodge S, Simon MF, Young JPW (2018) Horizontal Transfer of Symbiosis Genes within and Between Rhizobial Genera: Occurrence and Importance. *Genes (Basel)* 9(7):321. <https://doi.org/10.3390/genes9070321>
- Araújo J, Flores-Félix JD, Igual JM, Peix A, González-Andrés F, Díaz-Alcántara CA, Velázquez E (2017) *Bradyrhizobium cajani* sp. nov. isolated from nodules of *Cajanus cajan*. *Int J Syst Evol Microbiol.* 67:2236–2241. <https://doi.org/10.1099/ijsem.0.001932>
- [Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, et al. \(2018\) KBase: The United States Department of Energy Systems Biology Knowledgebase. \*Nat Biotechnol\* 36:566–569. <https://doi.org/10.1038/nbt.4163>](https://doi.org/10.1038/nbt.4163)
- Avontour JR, Palmer M, Beukes CW, Chan WY, Coetzee MPA, Blom J, Stepkowski T, Kyrpides NC, Woyke T, Shapiro N, Whitman WB, Venter SN, Steenkamp ET (2019) Genome-informed *Bradyrhizobium* taxonomy: where to from here? *Syst. Appl. Microbiol.* 42(4):427-439. <https://doi.org/10.1016/j.syapm.2019.03.006>
- Avontuur JR, Palmer M, Beukes CW, Chan WY, Tasiya T, Van Zyl E, Coetzee MPA, Stepkowski T, Venter SN, Steenkamp ET (2022) *Bradyrhizobium altum* sp. nov., *Bradyrhizobium oropedii* sp. nov. and *Bradyrhizobium acaciae* sp. nov. from South Africa show locally restricted and pantropical *nodA* phylogeographic patterns. *Mol. Phylogenetics Evol.* 167 107338, ISSN 1055-7903. <https://doi.org/10.1016/j.ympev.2021.107338>
- Ayalew T, Yoseph T, Petra H, Cadisch G (2021) Yield response of field-grown cowpea varieties to *Bradyrhizobium* inoculation. *Agronomy Journal* 113:3258-3268. <https://doi.org/10.1002/agj2.20763>
- Aziz RK, Bartels D, Best AA, et al. (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19(5):455-77. <https://doi.org/10.1089/cmb.2012.0021>
- Barbosa LM, Shirasuna RT, de Lima FC, Ortiz PRT, Barbosa KC, Barbosa TC (2017) Lista de espécies indicadas para restauração ecológica para diversas regiões do Estado de São Paulo. Instituto de Botânica, São Paulo

- Brettin T, Davis J, Disz T, et al. (2015) RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5: 8365. <https://doi.org/10.1038/srep08365>
- Chahboune R, Carro L, Peix A, Barrijal S, Velázquez E, Bedmar EJ (2011) *Bradyrhizobium cytisi* sp. nov., isolated from effective nodules of *Cytisus villosus*. *Int J Syst Evol Microbiol* 61:2922–2927. <https://doi.org/10.1099/ijs.0.027649-0>
- Costa EM, Guimarães AA, Vicentin RP, Ribeiro PRA, Leão ACR, Balsanelli E, Lebbe L, Aerts M, Willems A, Moreira FMS (2017) *Bradyrhizobium brasilense* sp. nov., a symbiotic nitrogen-fixing bacterium isolated from Brazilian tropical soils. *Arch Microbiol* 199:1211–1221. <https://doi.org/10.1007/s00203-017-1390-1>
- Costa EM, Guimarães AA, Carvalho TS, Rodrigues TL, Ribeiro PRA, Lebbe L, Willems A, Moreira FMS (2018) *Bradyrhizobium forestalis* sp. nov., an efficient nitrogen-fixing bacterium isolated from nodules of forest legume species in the Amazon. *Arch Microbiol* 200:743–752. <https://doi.org/10.1007/s00203-018-1486-2>
- Costa EM, Ribeiro PRA, Carvalho TS, Vicentin RP, Balsanelli E, Souza EM, Lebbe L, Willems A, Moreira FMS (2020) Efficient Nitrogen-Fixing Bacteria Isolated from Soybean Nodules in the Semi-arid Region of Northeast Brazil are Classified as *Bradyrhizobium brasilense* (Symbiovar Sojae). *Curr Microbiol* 77:1746–1755. <https://doi.org/10.1007/s00284-020-01993-6>
- Costa EM, Lima W, Ribeiro PRA, Moreira FMS (2021) Acid and high-temperature tolerant *Bradyrhizobium* spp. Strains from Brazilian soils are able to promote *Acacia mangium* and *Stizolobium aterrimum* growth. *Symbiosis* 83:65–78. <https://doi.org/10.1007/s13199-020-00732-6>
- Darling AE, Mau B, Perna NT (2010) progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS One* 5(6): e11147. <https://doi.org/10.1371/journal.pone.0011147>
- De Lajudie PM, Andrews M, Ardley J, Eardly B, Jumas-Bilak E, Kuzmanovic N, Lassale F, Lindström K, Mhamdi R, MartínezRomero E, Moulin L, Mousavi SA, Nesme X, Peix A, Putawska J, Steenkamp E, Stepkowski T, Tian CF, Vinuesa P, Wei G, Willems A, Zilli J, Young P (2019) Minimal standards for the description of new genera and species of rhizobia and agrobacteria. *Int J Syst Evol Microbiol* 69(7):1852–1863
- De Sá OAAL, Ribeiro PRA, Rufini M, Cruvinel IAF, Casagrande DR, Moreira FMS (2019) Microsymbionts of forage peanut under different soil and climate conditions belong to a specific group of Bradyrhizobium strains. *Appl Soil Ecol* 143:201–212. <https://doi.org/10.1016/j.apsoil.2019.07.018>
- Delamuta JRM, Ribeiro RA, Ormeño-Orrillo E, Parma MM, Melo IS, Martínez-Romero E, Hungria M (2015) *Bradyrhizobium tropiciagri* sp. nov. and *Bradyrhizobium embrapense* sp. nov., nitrogen fixing symbionts of tropical forage legumes. *Int J Syst Evol Microbiol* 65:4424–4433. <https://doi.org/10.1099/ijsem.0.000592>
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Faria SM, De Lima HC (1998) Additional studies of the nodulation status of legume species in Brazil. *Plant and Soil* 200:185–192

- Felsenstein, J (1981) Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach. *J Mol Evol* 17:368-376
- Fred EB, Waskman SA (1928) Laboratory manual of general microbiology: with special reference to the microorganisms of the soil. McGraw-Hill, New York.
- Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E, Avarre JC, et al. (2007). Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science* 316:1307–1312. <https://doi.org/10.1126/science.1139548>
- Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM (2007) DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57:81–91. <https://doi.org/10.1099/ijs.0.64483-0>
- Kim M, Oh HS, Park SC, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 64:346–351. <https://doi.org/10.1099/ijs.0.059774-0>
- Klepa MS, Helene LCF, O’Hara G, Hungria M (2021) *Bradyrhizobium agreste* sp. nov., *Bradyrhizobium glycinis* sp. nov. and *Bradyrhizobium diversitatis* sp. nov., isolated from a biodiversity hotspot of the genus *Glycine* in Western Australia. *Int J Syst Evol Microbiol* 71:004742. <https://doi.org/10.1099/ijsem.0.004742>
- Konstantinidis KT, Tiedje JM (2005) Genomic insights that advance the species definition for prokaryotes. *PNAS* 7:2567–2572
- Kuykendall LD, Saxena B, Devine TE, Udell SE (1992) Genetic diversity in *Bradyrhizobium japonicum* Jordan 1982 and a proposal for *Bradyrhizobium elkani* sp. nov. *Can J Microbiol* 38:501–505
- Lewis GP, Schrire B, Mackinder B, Lock M (2005) Legumes of the world. Royal Botanic Gardens, Kew, Richmond, Surrey
- Lopez BDO, Teixeira AFS, Michel DC, Guimarães AA, Costa AM, Costa JS, Pereira MS, Duarte BLM, Moreira FMS (2021) Genetic and symbiotic characterization of rhizobia nodulating legumes in a mining área in southeast Brazil. *Sci Agric* 79(2):e20200238. <https://doi.org/10.1590/1678-992X-2020-0238>
- Meira Junior MS, Pereira IM, Machado ELM, Mota SLL, Otoni TJO (2015) Espécies potenciais para recuperação de áreas de floresta estacional semidecidual com exploração de minério de ferro na serra do espinhaço. *Biosci J* 31(1):283-295
- Michel DC, Guimarães AA, Costa EM, Carvalho TS, Balsanelli E, Willems A, Souza EM, Moreira FMS (2020) *Bradyrhizobium uaiense* sp. nov., a new highly efficient cowpea symbiont. *Arch microbiol* 202:1135–1141. <https://doi.org/10.1007/s00203-020-01827-w>
- Michel DC, Costa EM, Guimarães AA, Carvalho TS, Caputo PSC, Willems A, Moreira FMS (2021) *Bradyrhizobium campsiandrae* sp. nov., a nitrogen-fixing bacterial strain isolated from a native leguminous tree from the Amazon adapted to flooded conditions. *Arch Microbiol* 203:233–240. <https://doi.org/10.1007/s00203-020-02022-7>



- Moreira FMS, Gillis M, Pot B, Kersters K, Franco AA (1993) Characterization of rhizobia isolated from different divergence groups of tropical Leguminosae by comparative polyacrylamide gel electrophoresis of their total proteins. *Syst Appl Microbiol* 16:135–146
- Moura GGD, Remigi P, Masson-Boivin C, Capela D (2020) Experimental Evolution of Legume Symbionts: What Have We Learnt? *Genes* 11(3):339. <https://doi.org/10.3390/genes11030339>
- Omara AE, El-Gaafarey T (2018) Alleviation of salinity stress effects in forage cowpea (*Vigna unguiculata* L.) by *Bradyrhizobium* sp. Inoculation. *Microbiol Res J Int* 23(3):1-16
- Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42(Database issue):D206-14. <https://doi.org/10.1093/nar/gkt1226>
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25(7):1043-55. <https://doi.org/10.1101/gr.186072.114>
- POWO (Plants of the World Online). *Machaerium nyctitans* (Vell.) Benth. Disponível em: <https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:505530-1>. Accessed 12 July 2022
- Provorov NA, Andronov EE, Kimeklis AK, Onishchuk OP, Igolkina AA and Karasev ES (2022) Microevolution, speciation and macroevolution in rhizobia: Genomic mechanisms and selective patterns. *Front Plant Sci* 13:1026943. <https://doi.org/10.3389/fpls.2022.1026943>
- Ramírez-Bahena MH, Flores-Félix JD, Chahboune R, Toro M, Velázquez E, Peix A (2016) *Bradyrhizobium centrosemae* (symbiovar centrosemae) sp. nov., *Bradyrhizobium americanum* (symbiovar phaseolarum) sp. nov. and a new symbiovar (tropici) of *Bradyrhizobium viridifuturi* establish symbiosis with *Centrosema* species native to America. *Syst Appl Microbiol* 39(6):378-383. <https://doi.org/10.1016/j.syapm.2016.06.001>
- Ramírez-Bahena MH, Chahboune R, Peix A, Velázquez E (2013) Reclassification of *Agromonas oligotrophica* into the genus *Bradyrhizobium* as *Bradyrhizobium oligotrophicum* comb. nov. *Int J Syst Evol Microbiol* 63:1013–1016. <https://doi.org/10.1099/ijss.0.041897-0>
- Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. *PNAS* 106(45):19126–19131
- Rocha LB, Da Costa EM, Costa GS, Da Silva JKP, Ribeiro EC, Borges JF, Da Sousa APM, Moreira FMS (2021) *Bradyrhizobium brasilense* as an efficient soybean microsymbiont in two contrasting soils of the southwestern region of Piauí (Cerrado biome). *Rev Bras Cienc Agrar* 16(3):e212. <https://doi.org/10.5039/agraria.v16i3a212>
- Rodríguez-R LM, Konstantinidis KT (2016) The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints* 4:e1900v1. <https://doi.org/10.7287/peerj.preprints.1900v1>

- Saitou N, Nei M (1987) The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol Biol Evol* 4(4):406-425
- Silva JS, Oliveira DP, Rufini M, Silva Junior CL, Baptista MVBD, Silva LC, Aragão OOS, Pereira TA, Moreira FMS (2021). New strains of *Bradyrhizobium* enrich plant biomass nitrogen content in *Crotalaria* for use as a green manure. *Bragantia* 80:e4021. <https://doi.org/10.1590/1678-4499.20200469>
- Soares BL, Ferreira PAA, Oliveira-Longatti SM, Marra LM, Rufini M, Andrade MJB, Moreira FMS (2014) Cowpea symbiotic efficiency, pH and aluminum tolerance in nitrogen-fixing bacteria. *Sci Agric* 71(3):171-180. <https://doi.org/10.1590/S0103-90162014000300001>
- Tamura K (1992) Estimation of the Number of Nucleotide Substitutions When There Are Strong Transition-Transversion and G+C-Content Biases. *Mol Biol Evol* 9(4):678-687. <https://doi.org/10.1093/oxfordjournals.molbev.a040752>
- VanInsberghe D, Maas KR, Cardenas E, Strachan CR, Hallam SJ, Mohn WW (2015) Non-symbiotic *Bradyrhizobium* ecotypes dominate North American forest soils. *ISME J* 9:2435–2441. <https://doi.org/10.1038/ismej.2015.54>
- Xu LM, Ge C, Cui Z, Li J, Fan H (1995) *Bradyrhizobium liaoningense* sp. nov., Isolated from the Root Nodules of Soybeans. *Int J Syst Bacteriol* 45(4):706-711. <https://doi.org/10.1099/00207713-45-4-706>
- Zilli JE, Pacheco RS, Gianluppi V, Smiderle OJ, Urquiaga S, Hungria M (2021) Biological N<sub>2</sub> fixation and yield performance of soybean inoculated with *Bradyrhizobium*. *Nutr Cycl Agroecosyst* 119:323–336. <https://doi.org/10.1007/s10705-021-10128-7>

## TABLES AND FIGURES

Table 1. Data resulting from the assembly of genomes by SPADES of the studied strains

<b>Statistics without reference</b>	<b>UFLA01-814</b>	<b>UFLA01-839</b>	<b>UFLA01-860</b>	<b>UFLA01-1164</b>
# contigs	389	289	401	224
# contigs ( $\geq 0$ bp)	389	289	401	224
# contigs ( $\geq 1000$ bp)	372	284	383	219
# contigs ( $\geq 10000$ bp)	242	195	252	158
# contigs ( $\geq 100000$ bp)	6	16	7	12
# contigs ( $\geq 1000000$ bp)	0	0	0	0
Largest contig	181184	246470	190903	282764
Total length	9113603	8937499	9235839	7674511
Total length ( $\geq 0$ bp)	9113603	8937499	9235839	7674511
Total length ( $\geq 1000$ bp)	9101277	8934061	9223514	7670727
Total length ( $\geq 10000$ bp)	8450528	8443068	8596767	7368107
Total length ( $\geq 100000$ bp)	795571	2333791	898755	1742517
Total length ( $\geq 1000000$ bp)	0	0	0	0
N50	41570	51623	40833	65415
N75	24734	31129	22536	34248
L50	71	49	71	39
L75	141	103	147	79
GC (%)	63.76	63.98	63.68	63.78
<b>Mismatches</b>				
# N's	194	304	0	207
# N's per 100 kbp	2.13	3.4	0	2.7
<b>Predicted genes</b>				
# predicted genes (unique)	8855	8729	9000	7637
# predicted genes ( $\geq 0$ bp)	8735 + 120 part	8631 + 98 part	8849 + 151 part	7552 + 85 part
# predicted genes ( $\geq 300$ bp)	7572 + 111 part	7528 + 87 part	7654 + 135 part	6490 + 80 part

# predicted genes (>= 1500 bp)	1076 + 16 part	1038 + 10 part	1084 + 20 part	896 + 5 part
# predicted genes (>= 3000 bp)	107 + 2 part	93 + 1 part	102 + 5 part	80 + 0 part

Table 2. CheckM v1.0.18: tabular representations of the phylogenetic marker completeness and contamination. CheckM generates clade-specific marker gene sets for each bin and reports the taxonomic resolution possible for each bin in the "Marker Lineage" column. Users may want to look at the "Marker Lineage" column to see what MAGs were classified with, for example, the "d\_\_Bacteria" or "d\_\_Archaea" marker sets. Instances where a broad (domain-level) marker set is used compared to a marker set from specific lineage (e.g. c\_\_Alphaproteobacteria) can help one contextualize (and evaluate) the genome completeness and contamination estimates.

Bin Name	Marker Lineage	# Genomes	# Markers	# Marker Sets	0	1	2	3	4	5+	Completeness	Contamination
UFLA01-814_02-21.contigs	f__Bradyrhizobiaceae	47	693	296	2	687	4	0	0	0	99.8	0.69
UFLA01-839_02-18.contigs	f__Bradyrhizobiaceae	47	693	296	0	688	5	0	0	0	100.0	0.76
UFLA01-860_02-20.contigs	f__Bradyrhizobiaceae	47	693	296	1	687	5	0	0	0	99.89	0.83
UFLA01-1164_02-14.contigs	f__Bradyrhizobiaceae	47	693	296	1	688	4	0	0	0	99.66	0.78

Table 3. Pairwise comparisons of user genomes vs. type strain genomes

Strain	Subject strain	dDDH (d0, in %)	C.I. (d0, in %)	dDDH (d4, in %)	C.I. (d4, in %)	dDDH (d6, in %)	C.I. (d6, in %)	G+C content difference (in %)
UFLA01-814	UFLA01-860	99,8	[99.5 - 99.9]	100	[99.9 - 100.0]	99,9	[99.8 - 100.0]	0,08
UFLA01-814	<i>Bradyrhizobium elkanii</i> USDA 76	60,5	[56.8 - 64.1]	63,3	[60.4 - 66.1]	62,4	[59.1 - 65.6]	0,04
UFLA01-860	<i>Bradyrhizobium elkanii</i> USDA 76	60	[56.3 - 63.5]	63,1	[60.2 - 65.9]	61,9	[58.6 - 65.1]	0,04
UFLA01-814	<i>Bradyrhizobium australafricanum</i> WSM 4400	64,6	[60.8 - 68.2]	62,3	[59.4 - 65.1]	65,9	[62.5 - 69.1]	0,23

<i>Bradyrhizobium australafricanum</i>								
UFLA01-860	WSM 4400	64	[60.2 - 67.6]	62,1	[59.2 - 64.9]	65,3	[62.0 - 68.5]	0,16
<i>Bradyrhizobium brasilense</i>								
UFLA01-860	UFLA03-321T	63,8	[60.0 - 67.4]	61,4	[58.5 - 64.1]	65	[61.6 - 68.2]	0,25
<i>Bradyrhizobium brasilense</i>								
UFLA01-814	UFLA03-321T	64,7	[60.9 - 68.3]	61,3	[58.4 - 64.1]	65,8	[62.4 - 69.0]	0,17
UFLA01-814	<i>Bradyrhizobium pachyrhizi</i> PAC48	63,9	[60.1 - 67.5]	59,7	[56.9 - 62.5]	64,7	[61.3 - 67.9]	0,01
UFLA01-860	<i>Bradyrhizobium pachyrhizi</i> PAC48	63,1	[59.3 - 66.7]	59,7	[56.8 - 62.4]	64	[60.6 - 67.2]	0,08
<i>Bradyrhizobium tropiciagri</i>								
UFLA01-839	SEMIA 6148	59,1	[55.5 - 62.7]	58,5	[55.7 - 61.3]	60,2	[56.9 - 63.3]	0,49
<i>Bradyrhizobium daqingense</i>								
UFLA01-1164	CGMCC 1.10947	66,2	[62.3 - 69.8]	57,1	[54.3 - 59.8]	66,1	[62.7 - 69.3]	0,05

Table 4. Similarity analysis of concatenated gene sequences (*gyrB*, *recA* and *rpoB*) with higher values found among the study strains and sequences of *Bradyrhizobium* type strains available at the NCBI

Stains	UFLA01-814 <sup>T</sup>	UFLA01-839 <sup>T</sup>	UFLA01-860 <sup>T</sup>	UFLA01-1164 <sup>T</sup>
UFLA01-814		94,93590931	100	84,60289295
UFLA01-839 <sup>T</sup>	94,93590931		94,93590931	85,00994603

---

<b>UFLA01-860<sup>T</sup></b>	100	94,93590931		84,60289295
<b>UFLA01-1164<sup>T</sup></b>	84,60289295	85,00994603	84,60289295	
<b><i>B. viridifuturi</i> SEMIA 690<sup>T</sup></b>	<b>95,54691442</b>	95,36267762	<b>95,54691442</b>	85,02966127
<b><i>B. daqingense</i> CCBAU 15774<sup>T</sup></b>	84,39555281	84,80260589	84,39555281	<b>95,93478814</b>
<b><i>B. pachyrhizi</i> PAC68<sup>T</sup></b>	<b>97,09873471</b>	94,98664898	<b>97,09873471</b>	84,65363263
<b><i>B. tropiciagri</i> CNPSo 1112<sup>T</sup></b>	<b>95,46504715</b>	<b>95,28081036</b>	<b>95,46504715</b>	84,947794
<b><i>B. elkanii</i> USDA 76<sup>T</sup></b>	<b>97,77414655</b>	94,94187782	<b>97,77414655</b>	84,60886146
<b><i>B. brasiliense</i> UFLA 03-321<sup>T</sup></b>	<b>96,81485651</b>	94,70277078	<b>96,81485651</b>	84,36975443
<b><i>B. australafricanum</i> WSM 4400<sup>T</sup></b>	<b>98,66716584</b>	94,43306808	<b>98,66716584</b>	84,10005173
<b><i>B. ferriligni</i> CCBAU 51502<sup>T</sup></b>	<b>97,94955956</b>	93,56433014	<b>97,94955956</b>	83,23131379
<b><i>B. mercantei</i> SEMIA 6399<sup>T</sup></b>	<b>95,55402552</b>	<b>95,36978872</b>	<b>95,55402552</b>	85,03677237
<b><i>B. ripae</i> WR4<sup>T</sup></b>	<b>95,66329997</b>	94,97416067	<b>95,66329997</b>	84,64114432
<b><i>B. uaiense</i> UFLA01-164<sup>T</sup></b>	93,11193815	<b>95,68317163</b>	93,11193815	83,18597488
<b><i>B. septentrionale</i> 1S1<sup>T</sup></b>	93,84050471	<b>95,28064102</b>	93,84050471	83,91454144
<b><i>B. acaciae</i> 10BB<sup>T</sup></b>	93,83779695	<b>96,40903042</b>	93,83779695	83,91183367

---

Table 5. Similarity analysis of concatenated gene sequences (*atpD*, *glnII* and *recA*) with higher values found among the study strains and sequences of *Bradyrhizobium* type species available at the NCBI

	UFLA01-814	UFLA01-839	UFLA01-860	UFLA01-1164
<b>UFLA01-814</b>		94,36857122	100	84,81492884
<b>UFLA01-839<sup>T</sup></b>	94,36857122		94,36857122	85,37059569
<b>UFLA01-860<sup>T</sup></b>	100	94,36857122		84,81492884
<b>UFLA01-1164<sup>T</sup></b>	84,81492884	85,37059569	84,81492884	
<b><i>B. daqingense</i> CCBAU 15774<sup>T</sup></b>	84,46493581	85,02060267	84,46493581	<b>96,53039096</b>
<b><i>B. pachyrhizi</i> PAC 48<sup>T</sup></b>	<b>96,80706182</b>	93,33159265	<b>96,80706182</b>	83,77795027
<b><i>B. embrapense</i> CNPSo 2833<sup>T</sup></b>	<b>95,2612616</b>	93,55026499	<b>95,2612616</b>	83,99662261
<b><i>B. tropiciagri</i> CNPSo 1112<sup>T</sup></b>	94,22060857	<b>97,81107943</b>	94,22060857	85,22263305
<b><i>B. americanum</i> CMVU 44<sup>T</sup></b>	84,92777931	85,48344617	84,92777931	<b>98,54715831</b>
<b><i>B. elkanii</i> USDA 76<sup>T</sup></b>	94,13979435	<b>95,31060162</b>	94,13979435	85,14181883
<b><i>B. brasilense</i> UFLA 03-321<sup>T</sup></b>	<b>97,092805</b>	93,61733583	<b>97,092805</b>	84,06369345
<b><i>B. ivorense</i> CI-1B<sup>T</sup></b>	93,05749866	<b>95,46574633</b>	93,05749866	84,05952313
<b><i>B. quebecense</i> 66S1<sup>T</sup></b>	<b>96,35386911</b>	94,64287251	<b>96,35386911</b>	85,08923012
<b><i>B. altum</i> Pear77<sup>T</sup></b>	<b>95,99434584</b>	94,28334924	<b>95,99434584</b>	84,72970685
<b><i>B. australafricanum</i> WSM 4400<sup>T</sup></b>	<b>98,71076141</b>	94,44636053	<b>98,71076141</b>	84,89271815

Table 6. ANI with the species closest to the study strains. Values generated by the ANI calculator from the website: <http://enve-omics.ce.gatech.edu/g-matrix/index>

	<b>UFLA01-814</b>	<b>UFLA01-839</b>	<b>UFLA01-860</b>	<b>UFLA01-1164</b>
<b>UFLA01-814</b>		90,22	<b>99,99*</b>	81,95
<b>UFLA01-839</b>	90,22		90,20	81,78
<b>UFLA01-860</b>	<b>99,99*</b>	90,20		81,85
<b>UFLA01-1164</b>	81,95	81,78	81,85	
<i>Bradyrhizobium tropiciagri</i> CNPSo 1112 <sup>T</sup>	<b>94,86**</b>	<b>94,57**</b>	<b>94,8**</b>	81,69
<i>Bradyrhizobium elkanii</i> USDA 76 <sup>T</sup>	<b>95,53**</b>	90,16	<b>95,51**</b>	81,96
<i>Bradyrhizobium brasilense</i> UFLA 03-321 <sup>T</sup>	<b>95,13**</b>	90,08	<b>95,12**</b>	81,84
<i>Bradyrhizobium daqingense</i> CCBAU 15774 <sup>T</sup>	81,91	81,77	81,83	<b>94,09**</b>
<i>Bradyrhizobium viridifuturi</i> SEMIA 690 <sup>T</sup>	89,85	<b>92,20**</b>	89,56	81,83



<i>Bradyrhizobium mercantei</i> SEMIA 6399 <sup>T</sup>	90,46	<b>91,12**</b>	90,47	81,79
<i>Bradyrhizobium forestalis</i> INPA 54B <sup>T</sup>	82,12	81,92	82,16	<b>88,07**</b>
<i>Bradyrhizobium cajani</i> AMBPC1010 <sup>T</sup>	82,35	82,08	82,31	<b>88,08**</b>
<i>Bradyrhizobium australiafricanum</i> WSM4400 <sup>T</sup>	95,27		95,27	

(\*) Value indicating that the species is the same between the species being compared.

(\*\*) Species with the closest identity.

Table 7. Results of Analytical profile index (API) tests on strains UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164 and *Bradyrhizobium elkanii* USDA 76T

Tests	Reactions/enzymes	UFLA01-839	UFLA01-1164	UFLA01-814	UFLA01-860	<i>B. elkanii</i> USDA76 <sup>T</sup>
<b>NO3</b>	NO3 (reduction of nitrates into nitrites)	-	-	-	+	+
<b>TRP</b>	TRP (indole formation (TRyptophan))	-	-	-	-	-
<b>GLU</b>	GLU (fermentation (GLUcose))	+	+	-	-	-
<b>ADH</b>	ADH (Arginine DiHydrolase)	+	+	-	-	-
<b>URE</b>	URE (UREase)	+	+	+	+	+

---

<b>ESC</b>	ESC (Hydrolysis ( $\beta$ -glucosidase) (ESCulin))	+	+	+	+	+
<b>GEL</b>	GEL (Hydrolysis (protease) (GELatin))	+	+	-	-	-
<b>PNG</b>	PNG ( $\beta$ -galactosidase (Para-NitroPhenyl- $\beta$ DGalactopyranosidase))	+	+	-	-	-
<b>GLU</b>	GLU (assimilation (GLUcose))	+	+	-	-	-
<b>ARA</b>	ARA (assimilation (ARAbinose))	-	+	-	-	-
<b>MNE</b>	MNE (assimilation (ManNosE))	+	+	-	-	-
<b>MAN</b>	MAN (assimilation (MANnitol))	+	+	-	-	-
<b>NAG</b>	NAG (Assimilation (N-Acetyl-Glucosamine))	+	+	-	-	-
<b>MAL</b>	MAL (assimilation (MALtose))	+	+	-	-	-
<b>GNT</b>	GNT (Assimilation (Potassium GlucoNato))	+	+	-	-	-
<b>CAP</b>	CAP (assimilation (CAPrate acid))	+	-	-	-	-
<b>ADI</b>	ADI (assimilation (ADIpate acid))	-	-	-	-	-
<b>MLT</b>	MLT (Assimilation (MaLaTo))	+	+	-	-	-
<b>CIT</b>	CIT (assimilation (Trisodium Citrate))	+	+	-	-	-
<b>PAC</b>	PAC (Assimilation (Phenyl Acid Acetate))	-	-	-	-	-

---

(+) Positive result, (-): negative result

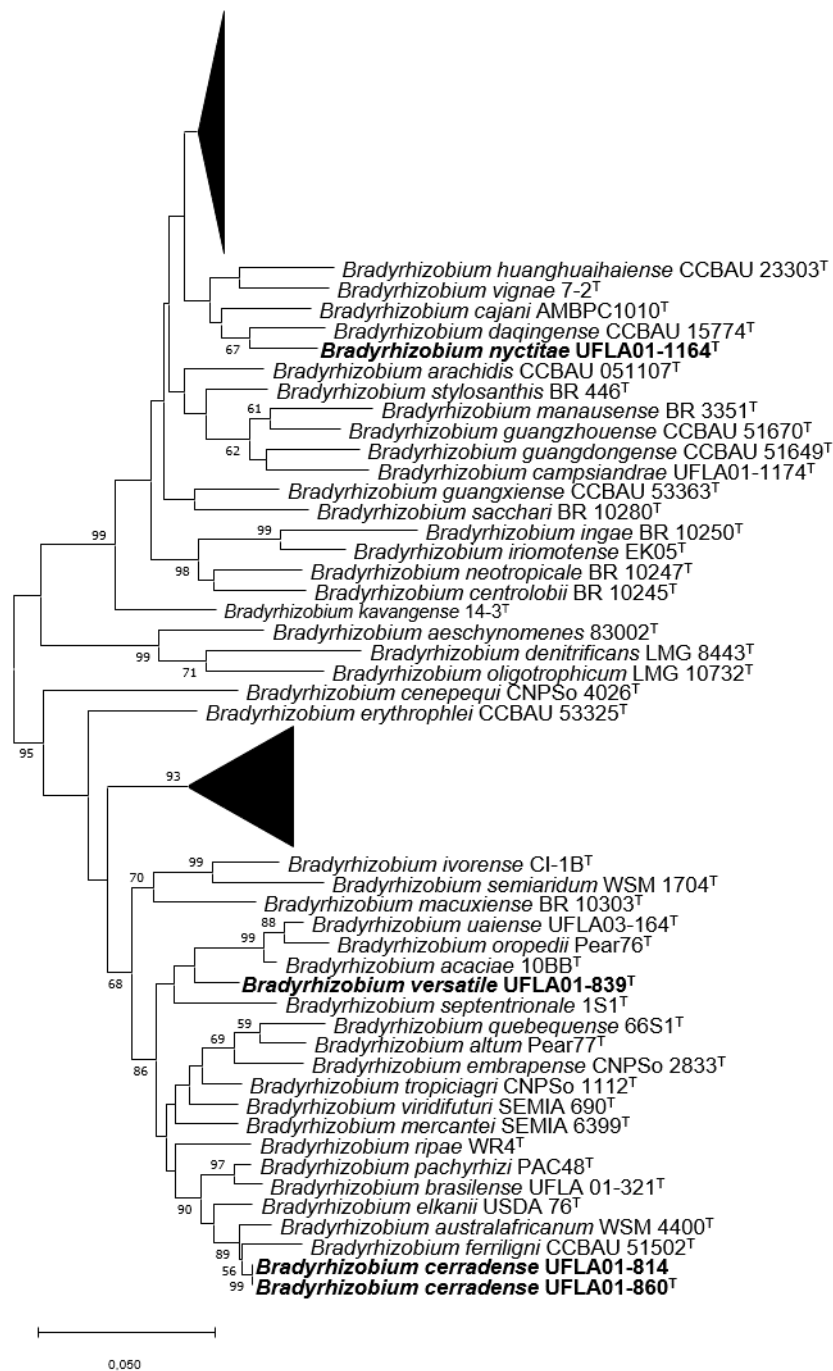


Fig. 1 Maximum likelihood phylogeny based on the concatenated genes (*gyrB*, *recA* and *rpoB*), showing the relationship between strains of the new species (shown in bold) and type strains belonging to the genus *Bradyrhizobium*. Bootstrap values greater than 50% are indicated on nodes. NCBI accession numbers are given in parentheses. Polygons are groups of compressed *Bradyrhizobium* species listed in a supplementary file (S1).

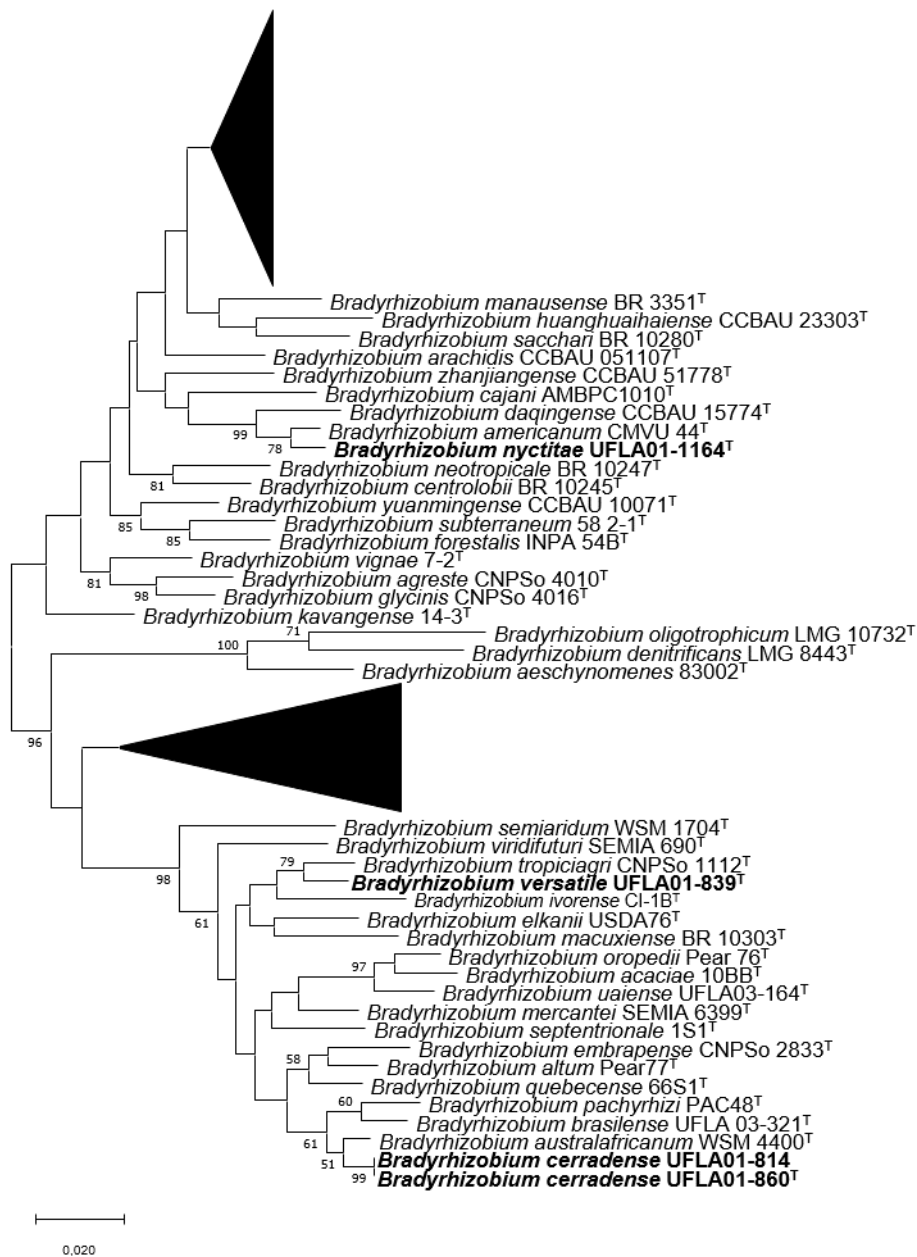


Fig. 2 Maximum likelihood phylogeny based on the concatenated genes (*atpD*, *glnII* and *recA*), showing the relationship between strains of the new species (shown in bold) and type strains belonging to the genus *Bradyrhizobium*. Bootstrap values greater than 50% are indicated on nodes. NCBI accession numbers are given in parentheses. Poligons in black are groups of compressed *Bradyrhizobium* type strains listed in a supplementary file (S1).

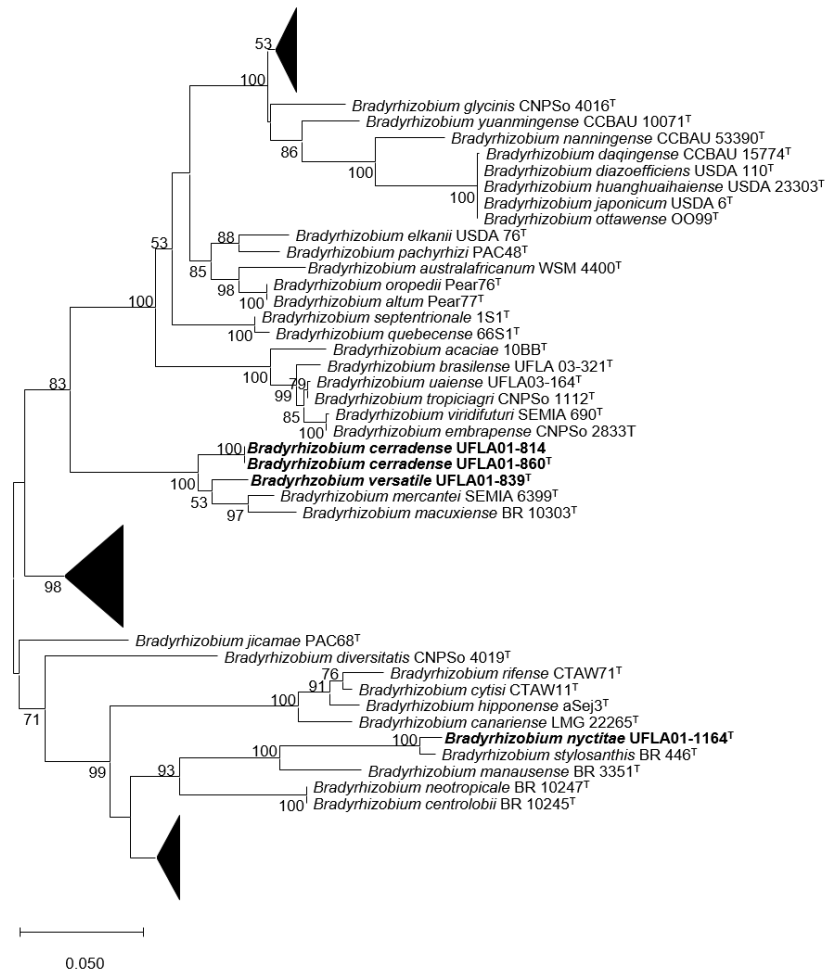


Fig. 3 Maximum likelihood phylogeny based on the *nifH* gene (880 bp), showing the relationship between the strains of the new species (shown in bold) and type strains belonging to the genus *Bradyrhizobium*. Bootstrap values greater than 50% are indicated on nodes. NCBI accession numbers are given in parentheses. Polygons are groups of compressed *Bradyrhizobium* species listed in a supplementary file (S1).

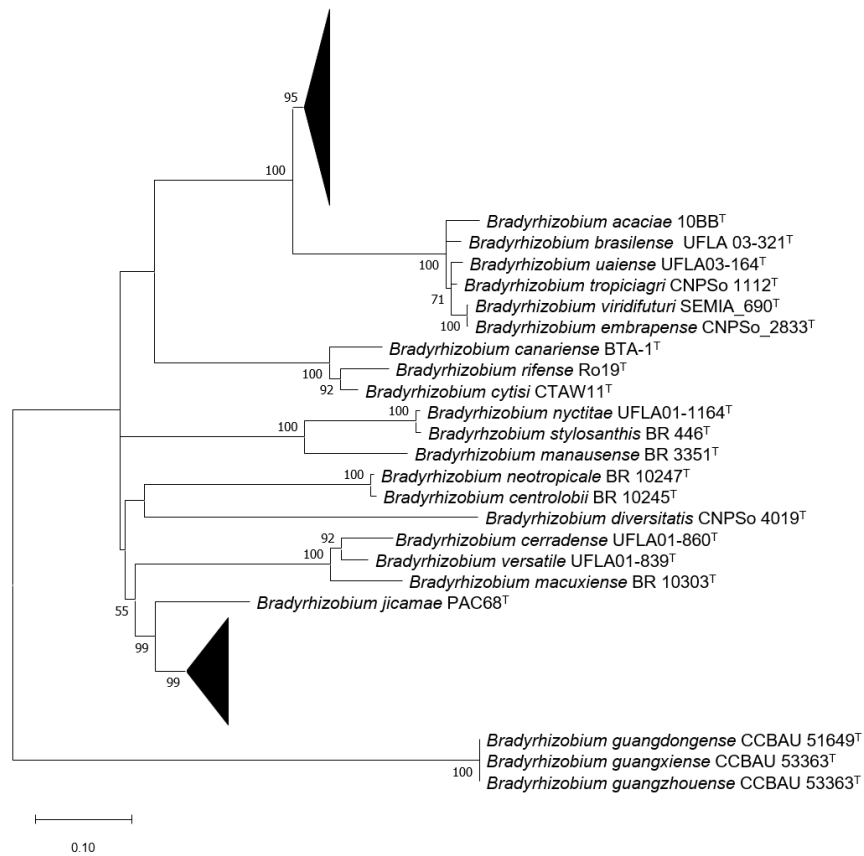


Fig. 4 Maximum likelihood phylogeny based on the *nodC* gene (808 bp), showing the relationship between the strains of new species (shown in bold) and type strains belonging to the genus *Bradyrhizobium*. Bootstrap values greater than 50% are indicated on nodes. NCBI accession numbers are given in parentheses. I, II and III are groups of compressed *Bradyrhizobium* species listed in a supplementary file (S2). Code of type strains in supplementary material (S2)

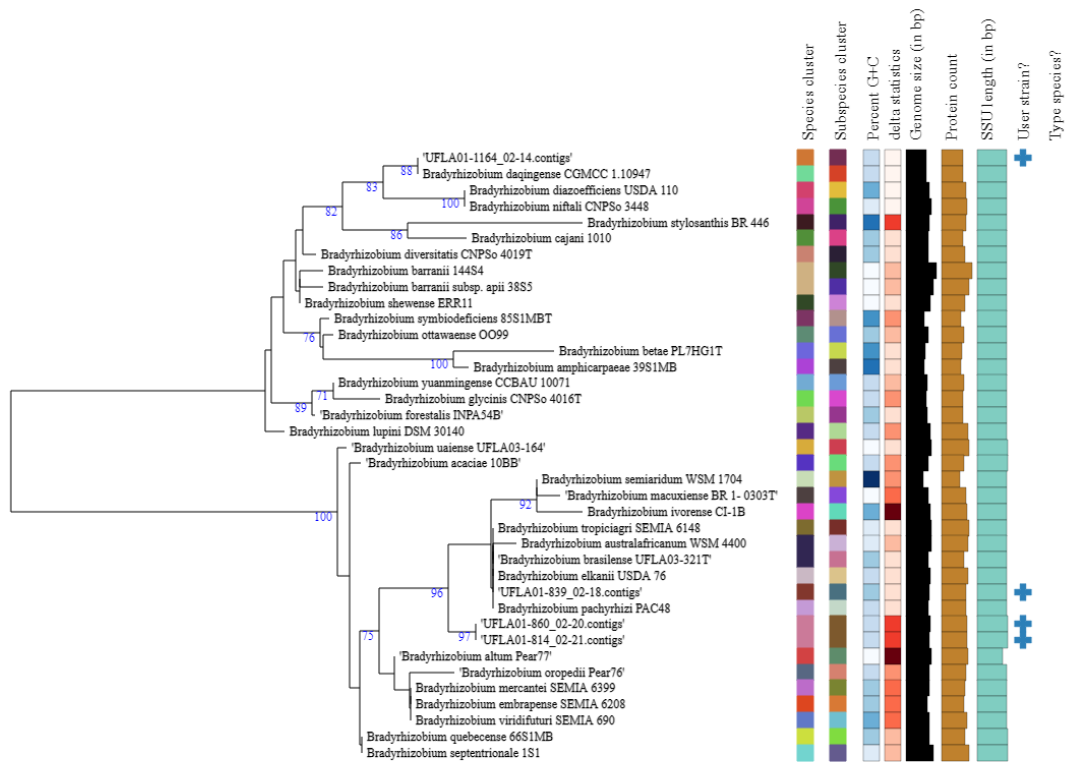


Fig 5. Tree inferred with FastME 2.1.6.1 (Lefort V et al., 2015) from GBDP distances calculated from 16S rDNA gene sequences. The branch lengths are scaled in terms of GBDP distance formula  $d_5$ . The numbers above branches are GBDP pseudo-bootstrap support values > 60 % from 100 replications, with an average branch support of 49.0 %. The tree was rooted at the midpoint (Farris JS, 1972).



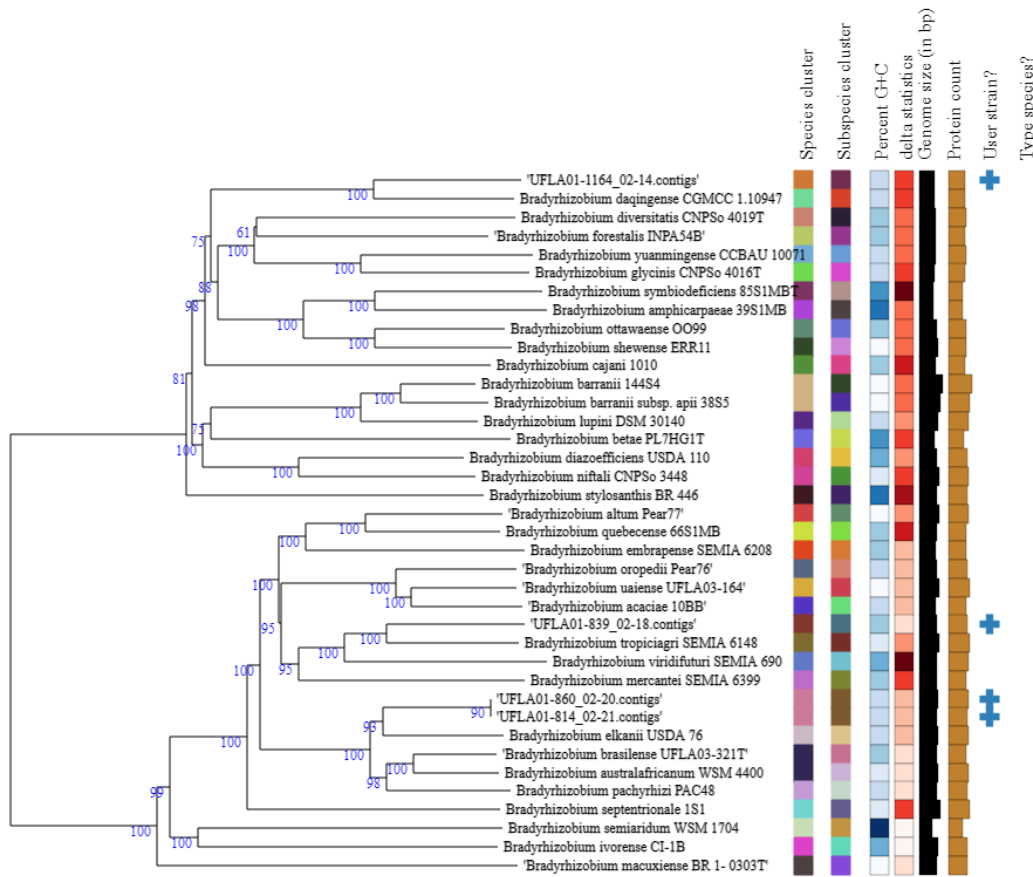


Fig 6. Tree inferred with FastME 2.1.6.1 (Lefort V et al., 2015) from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d5. The numbers above branches are GBDP pseudo-bootstrap support values > 60 % from 100 replications, with an average branch support of 95.7 %. The tree was rooted at the midpoint (Farris JS, 1972).

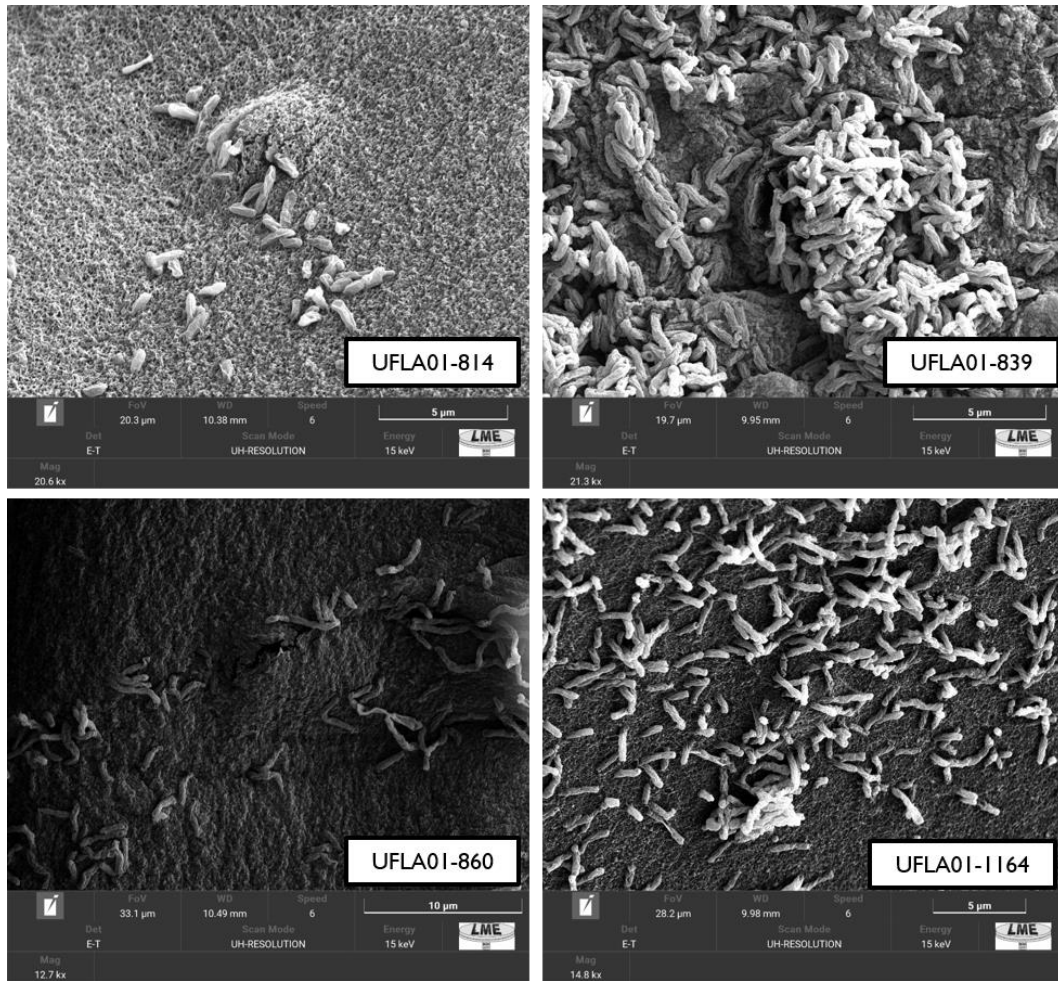


Fig. 7 Electromicrograph generated by scanning electron microscopy of samples of bacterial colonies of strains UFLA01-814, UFLA01-839, UFLA01-860 and UFLA01-1164

Table S1. List of housekeeping genes access codes of the type species used for analysis

Type species	Assembly genomes	Housekeepings genes					
		<i>rpoB</i>	<i>dnaK</i>	<i>glnII</i>	<i>gyrB</i>	<i>recA</i>	<i>atpD</i>
<i>B._viridifuturi_SEMIA_690T</i>	ASM123827v1	KU724169 .1	KR149128	KR149131	KR149134	KR149140	NZ_LGTB01000 039.1
<i>B._arachidis_CCBAU_051107T</i>	ASM1529170v1	JX437682	JX437668	HM10725 1	JX437675	HM107233	HM107217.1
<i>B._betae_LMG_21987T</i>	ASM893211v1	FM253260	FM253303. 1	AB353733	FM253217	AB353734	FM253129.1
<i>B._canariense_LMG_22265T (BTA-IT)</i>	ASM1940266v1	FM253263	AY923047	AY38676 5	FM253220	FM253177	FM253135.1
<i>B._cytisi_CTAW11T</i>	ASM812351v1	JN186288	JQ945184	GU00159 4	KF532653.1	GU001575	GU001613.1
<i>B._daqingense_CCBAU_15774T</i>	ASM783020v1	JX437676	KF962684	HQ23130 1	JX437669	HQ231270	HQ231289.1
<i>B._denitrificans_LMG_8443T (SZCCT0094T **)</i>	ASM1812952v1	FM253282	KF962685	HM04712 1	FM253239	FM253196	FM253153.1
<i>B._elkanii_USDA_76T</i>	ASM37914V1	AM29534 8	AY328392	AY59911 7	AM418800.1	AY591568	AY386758.1
<i>B._ganzhouense_RITF806T</i>	Nd	nd	KP420023	JX277110	KP420022	JX277144	JX277182.1
<i>B._huanghuaihaiense_CCBAU_23303T</i>	ASM783063v1	LM994169	JX437665	HQ23163 9	JX437672	HQ231595	HQ231682.1
<i>B._icense_LMTR_13T</i>	ASM169338v1	<b>Kbase genome</b>	KF896182	KF896175	KF896201	JX943615	KF896192.1
<i>B._ingae_BR_10250T</i>	Nd	KF927073	KF927055	KF927067	KF927079	KF927061	KY753593.1

<i>B. iriomotense_EK05T</i> (SZCCT0346 **)	ASM1813076v1	HQ587646	JF308944	AB300995	AB300997	AB300996	LM994395.1
<i>B. japonicum_USDA_6T</i>	ASM28437v1	AM29534 9	AM182120	HQ58787 5	AB070586.1	AM182158	AM418753.1
<i>B. jicamae_PAC68T</i>	ASM144039v1	HQ587647	JF308945	FJ428204	HQ873309	HM047133	FJ428211.1
<i>B. lablabi_CCBAU_23086T</i>	ASM144047v1	JX437677	KF962687	GU43349 8	JX437670	GU433522	GU433473.1
<i>B. liaoningense_LMG_18230T</i> (SZCCT0400 **)	ASM1813082v1	FM253266	AY923041	AY38677 5	FM253223	FM253180.1	AY386752.1
<i>B. manausense_BR_3351T</i>	ABySSBR3351v1	<b>Kbase genome</b>	KF786001	KF785986	KF786000	KF785992	NZ_LJYG01000 004.1
<i>B. neotropiale_BR_10247T</i>	ASM164169v1	KF983829	KJ661693	KJ661700	KJ661707	KJ661714	NZ_LSEF01000 046.1
<i>B. oligotrophicum_LMG_10732T</i>	ASM34480v1	KF962713	KF962688	JQ619233	KF962697	JQ619231	JQ619232.1
<i>B. oligotrophicum_LMG_10732T</i>	ASM227813v2	HQ587518	JF308816	HQ58775 0	HQ873179	HQ587287	HQ455212.1
<i>B. pachyrhizi_PAC48T</i>	ASM118924v1	HQ587648	JF308946	FJ428201	HQ873310	HM047130	FJ428208.1
<i>B. paxllaeri_LMTR_21T</i>	ASM169351v2	KP308154	AY923038	KF896169	KF896195	JX943617	KF896186.1
<i>B. retamae_Ro19T</i>	ASM144041v1	KF962714	KF896184	KC247108	KF896204	KF962711	KC247101.1
<i>B. rifense_CTAW71T</i>	ASM812342v1	KC569468	JQ945187	GU00160 4	KC569466	GU001585	GU001617.1
<i>B. yuanmingense_CCBAU_10071T</i>	IMG-taxon 2617270741	FM253269	FM253312. 1	AY38678 0	FM253226	FM253183	AY386760.1
<i>B. embrapense_CNPSO_2833T</i>	ASM118923v2	HQ634910	KP234519	GQ16050 0	HQ634891	HQ634899	HQ634875.1

<i>B._guangdongense_CCBAU 51649T</i>	ASM411497v1	KC509318	KC508964	KC509023	KC509072	KC509269	KC508916.1
<i>B._guangxiense_CCBAU_53363T</i>	ASM411491v1	KC509328	KC508974	KC509033	KC509082	KC509279	KC508926.1
<i>B._kavangense14-3T</i>	Nd	KM37831 1	KR259949	KM37844 6	KX661397.1	KM378399	KX661392.1
<i>B._valentinum_LmjM3T</i>	ASM144040v1	<b>Kbase genome</b>	<b>Kbase genome</b>	JX518575	NZ_LLXY01000 166.1	JX518589	JX518561.2
<i>B._vignae_7-2T</i>	ASM411442v1	KM37830 8	KR259951	KM37844 3	KX683216.1	KM378374	KX683215.1
<i>B._subterraneum_58_2-1T</i>	Nd	KM37834 9	KP308157	KM37848 4	KX661396.1	KM378397	KX661391.1
<i>B._erythrophlei_CCBAU_53325T (GAS401 **)</i>	IMG-taxon 2695420919 annotated assembly	MG81165 4.1	nd	KF114693	KF114717	KF114669	nd
<i>B._ferriligni_CCBAU_51502T</i>	Nd	MG81165 5.1	nd	KJ818099	KJ818102	KJ818112	nd
<i>B._diazoefficiens_USDA_110T</i>	ASM164267v1	<b>Kbase genome</b>	<b>Kbase genome</b>	<b>kbase genome</b>	<b>kbase genomes</b>	<b>kbase genomes</b>	<b>kbase genome</b>
<i>B._tropiciagri_CNPSO 1112T</i>	ASM118984	HQ634909	FJ391008	FJ391048	HQ634890	FJ391168	FJ390968.1
<i>B._mercantei SEMIA 6399T</i>	ASM198263v1	<b>Kbase genome</b>	KX690617	KX69062 1.1	KX690623	KX690615.1	NZ_MKFI01000 006.1
<i>B._forestalis INPA 54B T</i>	ASM279524v1	<b>Kbase genome</b>	<b>Kbase genome</b>	<b>kbase genome</b>	KF452831.1	KF452867.1	KF452722.1
<i>B._americanum CMVU 44T</i>	Nd	nd	nd	KX01294 2.1	nd	KC247141.1	KC247125.1
<i>B._cajani AMBPC1010T</i>	ASM975966v1	<b>Kbase genome</b>	<b>Kbase genome</b>	KY34944 2.1	NZ_WQNE0100 0054.1	KY349440.1	NZ_WQNE0100 0013.1

<i>B. centrosemae</i> A9T	Nd	nd	nd	KX01294 0.1	nd	KC247145.1	KC247129.1
<i>B. stylosanthis</i> BR 446T	ASM164133v1	KU724166	KU724145	KU72414 8.1	KU724151.1	KU724163.1	<b>kbase genome</b>
<i>B. lupini</i> USDA 3051T	NZ_AMQQ00000000.1 (ASM1419460v1)	nd	nd	KM11486 2.1	MK689366.1	KM114866.1	KU738808.1
<i>B. shewense</i> ERR 11T	IMG-taxon_2617270735	<b>Kbase genome</b>	<b>Kbase genome</b>	<b>kbase genome</b>	NZ_FMAI01000 013.1	NZ_FMAI0100002 2.1	NZ_FMAI01000 019.1
<i>B. sacchari</i> BR 10280T	ASM206809v1	Kbase genome	KX065103. 1	KX06509 9.1	kbase genomes	KX065095.1	KX065107.1
<i>B. centrolobii</i> BR 10245T	ASM164163v1	KF983827. 3	KX527928. 1	KX52799 1.1	KX528004.1	KX527954.1	NZ_LUUB01000 107.1
<i>B. brasilense</i> UFLA 03-321T	ASM196982v1	<b>Kbase genome</b>	<b>Kbase genome</b>	<b>kbase genome</b>	KF452827.1	KT793142.1	KF452730.1
<i>B. macuxiense</i> BR 10303T	ASM154241v1	KX527969 .1	KX527932. 1	KX52799 5.1	KX528008.1	KX527958.1	NZ_LNCU01000 024.1
<i>B. namibiense</i> 5-10T	Nd	KM37830 6.1	KP402058.1	KM37844 0.1	KX661393.1	KM378377.1	KX661387.1
<i>B. nitroreducens</i> TSA1T	ASM277669v1	<b>Kbase genome</b>	LFJC01000 003.1	<b>kbase genome</b>	LFJC01000003.1	LFJC01000003.1	<b>kbase genome</b>
<i>B. ripae</i> WR4T	Nd	MF593098 .1	MF593102. 1	nd	MF593094.1	MF593090.1	nd
<i>B. algeriense</i> RST89T	ASM306228v1	<b>Kbase genome</b>	FJ264922.1	FJ264924. 1	<b>kbase genomes</b>	FJ264927.1	KF956544.1
<i>B. frederickii</i> CNPS0 3426T	ASM457086v1	MK68269 9.1	<b>Kbase genome</b>	MK68268 8.1	MK682721.1	MK682710.1	NZ_SPQS01000 024.1

<i>B. nanningense</i> CCBAU 53390T	ASM411453v1	KC509323 .1	KC508969	KC509028	KC509077.1	KC509274	KC508921.1
<i>B. guangzhouense</i> CCBAU 51670T	ASM411495v1	KC509303 .1	KC508950	KC509008	KC509057.1	KC509254	KC508902.1
<i>B. zhanjiangense</i> CCBAU 51778T	ASM411493v1	KC509312 .1	KC508959	KC509017	KC509066.1	KC509263	KC508911.1
<i>B. symbiodeficiens</i> 85S1MBT	ASM226646v2	KP768667	<b>Kbase genome</b>	KP768609	KP768725	KF615036	KP768551
<i>B. amphicarphaeae</i> 39S1 MBT	ASM226643v2	KP768663	<b>Kbase genome</b>	KP768605	KP768721.1	KF615002.1	KP768547.1
<i>B. uaiense</i> UFLA03-164T	ASM1081187v1	<b>Kbase genome</b>	KF452780.1	<b>kbase genome</b>	KT793133.1	KT793144.1	KF452739.1
<i>B. ivorensis</i> CI-1BT	ERZ807199	KX388393	MK376326	MH75615 7	MH756161	MK376330	CAADFC020000 004.1
<i>B. niftali</i> CNPSo 3448T	ASM457102v1	MK67580 0	<b>Kbase genome</b>	MK67579 1	MK675794	MK675797	NZ_SPQT01000 036.1
<i>B. campsiandrae</i> UFLA01-1174T	ASM1452970v1	<b>Kbase genome</b>	KT793131.1	<b>kbase genome</b>	KT793134.1	KT793146.1	KT793128.1
<i>B. archetypum</i> WSM 1744T	ASM1311483v1	<b>Kbase genome</b>	MK674820	MK86084 7	MK860868	MK863448	MT070745
<i>B. australiense</i> WSM 1791T	ASM1311482v1	<b>Kbase genome</b>	MK674822	MK86084 9	MK860870	MK863450	MT070746
<i>B. murdochi</i> WSM 1741T	ASM47296v1	<b>Kbase genome</b>	MK674817	MK86084 4	MK860865	MK863445	MT070743
<i>B. agreste</i> CNPSo 4010T	ASM1603162v1	<b>Kbase genome</b>	MK674804. 1	MK86083 1.1	MK860852.1	MK863432.1	MT683853.1





(\*\*) Genome of strain that are not type species used for comparison purposes. KBase genome: Sequence extracted from the genome on the KBase platform. nd = Sequence not detected in database

**Table S2. Model test *atpD-glnII\_recA***

Model	#Param	BIC	AICc	lnL	Invariant	Gamma	R	Freq A	Freq T	Freq C	Freq G	A=>T	A=>C	A=>G	T=>A	T=>C	T=>G	C=>A	C=>T	C=>G	G=>A	G=>T	G=>C
GTR+G+I	45	8929,9 79668	8594,1 74211	- 4251,92 679	0,5300 94984	0,8207 763	2,0243 5583	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 2	0,0 1	0,1 2	0,0 2	0,3 1	0,0 6	0,0 1	0,1 9	0,0 7	0,0 7	0,0 3	0,0 7
TN93+G+I	42	8947,3 33793	8633,8 95785	- 4274,80 8055	0,5254 95901	0,7598 17904	2,0492 1149	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,1 1	0,0 3	0,3 1	0,0 5	0,0 3	0,1 9	0,0 5	0,0 7	0,0 3	0,0 5
T92+G+I	39	8949,3 01767	8658,2 34013	- 4289,99 6245	0,5425 41519	0,8040 30616	2,0227 4263	0,1915 41905	0,1915 41905	0,3084 58095	0,3084 58095	0,0 3	0,0 5	0,2 1	0,0 3	0,2 1	0,0 5	0,0 3	0,1 3	0,0 5	0,1 3	0,0 3	0,0 5
GTR+G	44	8955,0 28805	8626,6 78853	- 4269,18 6093	n/a	0,2163 44584	1,8644 6646	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 2	0,0 2	0,1 3	0,0 2	0,2 9	0,0 6	0,0 1	0,1 8	0,0 8	0,0 8	0,0 3	0,0 8
T92+G	38	8967,9 56271	8684,3 45893	- 4304,05 8232	n/a	0,1972 88342	1,9143 1685	0,1915 41905	0,1915 41905	0,3084 58095	0,3084 58095	0,0 3	0,0 5	0,2 1	0,0 3	0,2 1	0,0 5	0,0 3	0,1 3	0,0 5	0,1 3	0,0 3	0,0 5
TN93+G	41	8975,7 56382	8669,7 74814	- 4293,75 4084	n/a	0,2140 11546	1,8578 0471	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,1 3	0,0 3	0,2 9	0,0 5	0,0 3	0,1 7	0,0 5	0,0 8	0,0 3	0,0 5
HKY+G+I	41	8976,5 98355	8670,6 16787	- 4294,17 5071	0,5431 97781	0,8045 24989	2,0616 5434	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,2 2	0,0 3	0,2 1	0,0 5	0,0 3	0,1 3	0,0 5	0,1 4	0,0 3	0,0 5
HKY+G	40	8995,0 50431	8696,5 25615	- 4308,13 5843	n/a	0,1964 81293	1,9536 8458	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,2 1	0,0 3	0,2 5	0,0 3	0,0 2	0,1 5	0,0 3	0,1 3	0,0 3	0,0 5
GTR+I	44	9033,0 69347	8704,7 19395	- 4308,20 6364	0,6329 4532	n/a	1,7198 5845	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 2	0,0 2	0,1 3	0,0 3	0,2 8	0,0 7	0,0 1	0,1 7	0,0 8	0,0 8	0,0 4	0,0 8
TN93+I	41	9059,2 46471	8753,2 64903	- 4335,49 9129	0,6335 00357	n/a	1,6855 8543	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,1 3	0,0 4	0,2 7	0,0 6	0,0 4	0,1 7	0,0 6	0,0 8	0,0 3	0,0 5
T92+I	38	9071,5 76228	8787,9 6585	- 4355,86 821	0,6382 92991	n/a	1,7376 9438	0,1915 41905	0,1915 41905	0,3084 58095	0,3084 58095	0,0 3	0,0 5	0,2 3	0,0 3	0,2 5	0,0 3	0,0 2	0,1 5	0,0 2	0,1 3	0,0 3	0,0 5

K2+G	37	9073,1 16703	8796,9 6401	- 4361,37 3182	n/a	0,2077 13454	1,6883 1441	0,25	0,25	0,25	0,25	0,0 5	0,0 5	0,1 6	0,0 5	0,1 6	0,0 5	0,0 5	0,1 6	0,0 5	0,1 6	0,0 5	0,0 5
K2+G +I	38	9075,1 10162	8791,4 99783	- 4357,63 5177	0,5252 09474	0,7628 5315	2,3526 7225	0,25	0,25	0,25	0,25	0,0 4	0,0 4	0,1 8	0,0 4	0,1 8	0,0 4	0,0 4	0,1 8	0,0 4	0,1 8	0,0 4	0,0 4
HKY +I	40	9099,6 64994	8801,1 40177	- 4360,44 3124	0,6383 56902	n/a	1,7597 2884	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,2	0,0 3	0,2	0,0 5	0,0 3	0,1 2	0,0 5	0,1 3	0,0 3	0,0 5
K2+I	37	9171,0 1226	8894,8 59567	- 4410,32 096	0,6375 21426	n/a	1,5738 6005	0,25	0,25	0,25	0,25	0,0 5	0,0 5	0,1 5	0,0 5	0,1 5	0,0 5	0,0 5	0,1 5	0,0 5	0,1 5	0,0 5	0,0 5
JC+G +I	37	9271,3 13985	8995,1 61293	- 4460,47 1823	0,5363 5733	0,9472 21681	0,5	0,25	0,25	0,25	0,25	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8
JC+G	36	9277,4 30889	9008,7 36194	- 4468,26 5009	n/a	0,2100 55845	0,5	0,25	0,25	0,25	0,25	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8
JC+I	36	9366,4 27774	9097,7 33079	- 4512,76 3452	0,6370 22799	n/a	0,5	0,25	0,25	0,25	0,25	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8
GTR	43	9973,9 68711	9653,0 74575	- 4783,39 078	n/a	n/a	1,1019 8685	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,1	0,0 3	0,2 4	0,0 6	0,0 3	0,1 5	0,1	0,0 7	0,0 4	0,1
TN93	40	9987,2 24146	9688,6 99329	- 4804,22 27	n/a	n/a	1,4509 4104	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 4	0,0 6	0,1	0,0 4	0,2 8	0,0 6	0,0 4	0,1 7	0,0 6	0,0 6	0,0 4	0,0 6
T92	37	10056, 60448	9780,4 51784	- 4853,11 7069	n/a	n/a	1,4308 0033	0,1915 41905	0,1915 41905	0,3084 58095	0,3084 58095	0,0 4	0,0 6	0,1 9	0,0 4	0,1 9	0,0 6	0,0 4	0,1 2	0,0 6	0,1 2	0,0 4	0,0 6
HKY	39	10086, 8011	9795,7 33345	- 4858,74 5911	n/a	n/a	1,4311 6243	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 4	0,0 6	0,1 9	0,0 4	0,1 8	0,0 6	0,0 4	0,1 1	0,0 6	0,1 2	0,0 4	0,0 6
K2	36	10088, 51367	9819,8 1898	- 4873,80 6402	n/a	n/a	1,4132 4526	0,25	0,25	0,25	0,25	0,0 5	0,0 5	0,1 5	0,0 5	0,1 5	0,0 5	0,0 5	0,1 5	0,0 5	0,1 5	0,0 5	0,0 5
JC	35	10270, 96538	10009, 72899	- 4969,76 6986	n/a	n/a	0,5	0,25	0,25	0,25	0,25	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8	0,0 8

Table S3. Model test gyrB\_rpoB\_recA



JC+G	160	46300,0098	44737,46462	-22208,53254	n/a	0,213216536	0,5	0,25	0,25	0,25	0,25	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08
K2+I	161	46449,14024	44876,83165	-22277,21356	0,601808609	n/a	1,66923697	0,25	0,25	0,25	0,25	0,05	0,05	0,16	0,05	0,16	0,05	0,05	0,16	0,05	0,16	0,05	0,05
JC+I	160	47531,72135	45969,17616	-22824,38831	0,601120467	n/a	0,5	0,25	0,25	0,25	0,25	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08
GTR	167	51209,32946	49578,44108	-24622,00295	n/a	n/a	1,34133249	0,187992812	0,160914287	0,311813549	0,339279352	0,03	0,03	0,13	0,03	0,27	0,04	0,02	0,14	0,11	0,07	0,02	0,11
TN93	164	51522,67604	49921,07742	-24796,32885	n/a	n/a	1,48817729	0,187992812	0,160914287	0,311813549	0,339279352	0,03	0,06	0,13	0,04	0,28	0,06	0,04	0,15	0,06	0,07	0,03	0,06
T92	161	51793,84827	50221,53968	-24949,56757	n/a	n/a	1,47898599	0,17445355	0,17445355	0,32554645	0,32554645	0,03	0,06	0,23	0,03	0,26	0,06	0,03	0,11	0,06	0,11	0,03	0,06
K2	160	51908,0843	50345,53912	-25012,56979	n/a	n/a	1,63702887	0,25	0,25	0,25	0,25	0,05	0,05	0,16	0,05	0,16	0,05	0,05	0,16	0,05	0,16	0,05	0,05
HKY	163	51938,19107	50346,35576	-25009,97056	n/a	n/a	1,47879749	0,187992812	0,160914287	0,311813549	0,339279352	0,03	0,06	0,21	0,04	0,19	0,06	0,04	0,16	0,12	0,03	0,06	0,06
JC	159	52964,69915	51411,9174	-25546,76141	n/a	n/a	0,5	0,25	0,25	0,25	0,25	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08

Table S4. Model test nifH

Model	#Param	BIC	AICc	lnL	Invariant	Gamma	R	Freq A	Freq T	Freq C	Freq G	A=>T	A=>C	A=>G	T=>A	T=>C	T=>G	C=>A	C=>T	C=>G	G=>A	G=>T	G=>C
T92+G+I	127	18702,63787	17569,2746	-8657,345066	0,568059732	2,601525143	2,21975122	0,205380683	0,205380683	0,294619317	0,294619317	0,03	0,04	0,21	0,03	0,21	0,04	0,03	0,14	0,04	0,14	0,03	0,04
GTR+G+I	133	18740,30921	17553,42998	-8643,394569	0,566249729	2,752394376	2,30077061	0,229199175	0,181562192	0,279741727	0,309496906	0,02	0,04	0,17	0,02	0,25	0,05	0,03	0,16	0,06	0,13	0,03	0,05



GTR	131	20615,85457	19446,81369	-9592,095969	n/a	n/a	1,62732621	0,229199175	0,181562192	0,279741727	0,309496906	0,03	0,04	0,14	0,04	0,24	0,04	0,03	0,15	0,08	0,12	0,02	0,07
TN93	128	20637,02673	19494,74395	-9619,075135	n/a	n/a	1,85356207	0,229199175	0,181562192	0,279741727	0,309496906	0,03	0,05	0,14	0,04	0,26	0,05	0,04	0,17	0,05	0,15	0,03	0,05
K2	124	20652,4202	19545,81588	-9648,629314	n/a	n/a	1,83237883	0,25	0,25	0,25	0,25	0,04	0,04	0,16	0,04	0,16	0,04	0,04	0,16	0,04	0,16	0,04	0,04
T92	125	20685,05752	19569,53347	-9659,483612	n/a	n/a	1,84555909	0,205380683	0,205380683	0,294619317	0,294619317	0,04	0,05	0,19	0,04	0,19	0,05	0,04	0,13	0,05	0,13	0,04	0,05
HKY	127	20748,49997	19615,13669	-9680,276113	n/a	n/a	1,84670111	0,229199175	0,181562192	0,279741727	0,309496906	0,03	0,05	0,24	0,04	0,18	0,05	0,04	0,12	0,05	0,15	0,03	0,05
JC	123	21243,6084	20145,92387	-9949,687772	n/a	n/a	0,5	0,25	0,25	0,25	0,25	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08

Table S5. Model test nodC

Model	#Param	BIC	AICc	lnL	Invariant	Gamma	R	Freq A	Freq T	Freq C	Freq G	A=>T	A=>C	A=>G	T=>A	T=>C	T=>G	C=>A	C=>T	C=>G	G=>A	G=>T	G=>C
T92+G+I	113	25593,8122	24598,69711	-12186,08803	0,336354907	1,633976771	1,87664693	0,216454802	0,216454802	0,283545198	0,283545198	0,04	0,05	0,19	0,04	0,19	0,05	0,04	0,14	0,05	0,14	0,04	0,05
HKY+G+I	115	25612,33026	24599,61184	-12184,53612	0,335087736	1,620593999	1,87091736	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,05	0,19	0,03	0,19	0,05	0,03	0,15	0,05	0,13	0,04	0,05
TN93+G+I	116	25623,13675	24601,61679	-12184,53389	0,335531048	1,621597489	1,87095354	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,05	0,19	0,03	0,19	0,05	0,03	0,15	0,05	0,13	0,04	0,05
GTR+G+I	119	25624,1014	24576,17729	-12168,79981	0,330736559	1,575086513	1,89112624	0,201230831	0,231678773	0,284039548	0,283050847	0,02	0,05	0,19	0,02	0,19	0,05	0,04	0,15	0,06	0,13	0,04	0,06
T92+G	112	25640,54809	24654,23479	-12214,86144	n/a	0,446148754	1,88431993	0,216454802	0,216454802	0,283545198	0,283545198	0,04	0,05	0,19	0,04	0,19	0,05	0,04	0,14	0,05	0,14	0,04	0,05
HKY+G	114	25662,70308	24658,78629	-12215,128	n/a	0,448211807	1,88134066	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,05	0,19	0,03	0,19	0,05	0,03	0,15	0,05	0,13	0,04	0,05

TN93+G	115	25668,27024	24655,55182	-12212,50611	n/a	0,446609227	1,88343595	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,05	0,18	0,03	0,19	0,05	0,03	0,16	0,05	0,13	0,04	0,05
GTR+G	118	25692,12439	24653,00158	-12208,21677	n/a	0,452647667	1,74002766	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,05	0,16	0,03	0,2	0,04	0,04	0,16	0,06	0,12	0,04	0,06
K2+G+I	112	25722,63743	24736,32413	-12255,90611	0,331296282	1,611560688	1,79304033	0,25	0,25	0,25	0,25	0,04	0,04	0,16	0,04	0,16	0,04	0,04	0,16	0,04	0,16	0,04	0,04
K2+G	111	25765,23308	24787,72165	-12282,60941	n/a	0,461784832	1,80205478	0,25	0,25	0,25	0,25	0,04	0,04	0,16	0,04	0,16	0,04	0,04	0,16	0,04	0,16	0,04	0,04
T92+I	112	25928,00493	24941,69162	-12358,58986	0,378237268	n/a	1,72937309	0,216454802	0,216454802	0,283545198	0,283545198	0,04	0,05	0,18	0,04	0,18	0,05	0,04	0,18	0,05	0,14	0,04	0,05
HKY+I	114	25948,79442	24944,87762	-12358,17367	0,378147706	n/a	1,7195331	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,05	0,18	0,04	0,18	0,05	0,04	0,18	0,05	0,13	0,04	0,05
TN93+I	115	25958,25602	24945,5376	-12357,499	0,378182295	n/a	1,72125651	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,05	0,19	0,04	0,18	0,05	0,04	0,18	0,05	0,13	0,04	0,05
GTR+I	118	25991,74476	24952,62195	-12358,02696	0,378064688	n/a	1,57596224	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,05	0,17	0,04	0,18	0,05	0,04	0,18	0,06	0,12	0,04	0,06
K2+I	111	26056,85165	25079,34021	-12428,41869	0,377953037	n/a	1,66297153	0,25	0,25	0,25	0,25	0,05	0,05	0,16	0,05	0,16	0,05	0,05	0,16	0,05	0,16	0,05	0,05
JC+G+I	111	26477,3607	25499,84927	-12638,67322	0,333265701	1,775820851	0,5	0,25	0,25	0,25	0,25	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08
JC+G	110	26525,49619	25556,7867	-12668,14643	n/a	0,479573847	0,5	0,25	0,25	0,25	0,25	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08
JC+I	110	26779,03198	25810,32249	-12794,91433	0,377773764	n/a	0,5	0,25	0,25	0,25	0,25	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08
T92	111	27579,42137	26601,90993	-13189,70355	n/a	n/a	1,549838	0,216454802	0,216454802	0,283545198	0,283545198	0,04	0,05	0,17	0,04	0,17	0,05	0,04	0,17	0,03	0,15	0,03	0,05
HKY	113	27591,4057	26596,29061	-13184,88478	n/a	n/a	1,54954121	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,06	0,17	0,04	0,17	0,05	0,04	0,17	0,05	0,12	0,04	0,06
GTR	117	27593,72879	26563,40736	-13164,42444	n/a	n/a	1,37530064	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,06	0,16	0,04	0,18	0,05	0,04	0,18	0,07	0,11	0,04	0,08
TN93	114	27602,16631	26598,24952	-13184,85961	n/a	n/a	1,54956687	0,201230831	0,231678773	0,284039548	0,283050847	0,04	0,06	0,17	0,04	0,17	0,05	0,04	0,17	0,05	0,12	0,04	0,06





**ARTIGO 2*****Rhizobium elegans* sp. nov. nitrogen-fixing bacteria strain isolated from nodules of *Platypodium elegans*  
(Voegel.)**

Artigo formatado nas normas da revista Archives in Microbiology

Bruna Daniela Ortiz Lopez<sup>1</sup>, Daniele Cabral Michel<sup>1</sup>, Lucas Lenin Resende de Assis<sup>1</sup>, Raysa Marques Cardoso, Fatima Maria de Souza Moreira<sup>1\*</sup>

<sup>1</sup> Setor de Biologia, Microbiologia e Processos Biológicos do Solo, Departamento de Ciência do Solo, Universidade Federal de Lavras, Campus UFLA, 37200-000, Lavras, Minas Gerais, Brazil.

\*Corresponding author: Tel.: +55 35 3829 1254; fax: +55 35 3829 1251. E-mail address: fmoreira@ufla.br (F.M. de Souza Moreira).

## Abstract

Several new species of symbiotic nitrogen-fixing bacteria have already been described, mainly belonging to the genus *Rhizobium*. However, there is no identification at the species level of strains that nodulate and fix N<sub>2</sub> in symbiosis with the native legume *Platypodium elegans* (Voegel). In our study, the average nucleotide identity (ANI) between these two strains and 94 type strains of already described species showed values of 87,21 and 86,16% for strains UFLA01-1127 and UFLA01-1134 with the closest species *Rhizobium hainanense* CCBAU 57015<sup>T</sup> and 86.99 and 86.96% for strains UFLA01-1127 and UFLA01-1134 with the closest species *Rhizobium tropici* CIAT 899<sup>T</sup>. Therefore, the strains UFLA01-1127 and UFLA01-1134 represent a new species. The recommended name for this new species is *Rhizobium elegans* sp. nov. (type strain UFLA01-1127<sup>T</sup>).

**Keywords:** taxonomy, tree species, biodiversity, phylogeny

## Introdução

O nitrogênio em florestas pode vir de diversas fontes como decomposição da matéria orgânica e fixação biológica do nitrogênio e as espécies florestais podem ser fundamentais para esses processos. Em sítios com necessidade de restauração florestal a adição de N<sub>2</sub> ao sistema via inoculação de mudas pode ser uma alternativa econômica e ecologicamente correta. Estudos apontam que em locais afetados por estresse têm aumento substancial da fixação de nitrogênio atmosférico, já que a microbiota local aponta necessidade de resiliência (Barron et al., 2011).

A fixação biológica de nitrogênio é um dos processos mais importantes para a vida na Terra. Consiste na transformação do nitrogênio atmosférico em amônia pela ação da nitrogenase. Bactérias dotadas com esse complexo enzimático, também chamadas bactérias diazotróficas, são responsáveis por adicionar grandes quantidades de nitrogênio ao solo e às plantas. Dentro do grupo dos diazotróficos temos as bactérias simbióticas nodulíferas, que promovem o eficiente fornecimento de N às plantas com formação de nódulos radiculares. Existe uma ampla gama de diazotróficos com representantes de arqueobactérias, cianobactérias, bactérias gram-negativas e gram-positivas com uma ampla diversidade morfológica, fisiológica e genética. Dentro do grupo das Alpha-proteobactérias, os gêneros conhecidos de espécies nodulíferas compreendem: *Rhizobium*, *Bradyrhizobium*, *Paraburkholderia*, *Sinorhizobium*, *Mesorhizobium*, *Allorhizobium*, *Azorhizobium* e *Cupriavidus* (Moreira et al., 2006).

Apesar das espécies leguminosas terem muitos relatos de simbioses com o gênero *Bradyrhizobium* (Sprent et al., 2017), foram identificadas estirpes do gênero *Rhizobium* capazes de nodular espécies florestais, sendo uma delas *Platypodium elegans* (Lopez et al., 2020). Com exceção do inoculante para *Calliandra houstoniana* e *C. surinamensis*, não há inoculante com bactérias do gênero *Rhizobium* recomendadas para espécies florestais (MAPA 2011). Além disso, ainda não existe relato de bactérias diazotróficas que nodulam *P. elegans* identificadas à nível de espécie.

*P. elegans* Vogel é uma espécie florestal de origem na América central e sul e encontrada em países como Panamá, Colômbia, Venezuela, Bolívia, Paraguai e uma ampla distribuição pelo Brasil. Possui diversos usos como madeira para construção civil, carpintaria e decoração (Lewis 2005). É uma espécie recomendada para restauração de áreas

degradadas por se tratar de uma espécie pioneira e por ser resistente a condições de alta luminosidade tanto durante a germinação quanto no decorrer do seu desenvolvimento.

A partir disso o presente estudo tem como objetivo identificar a nível de espécie com uso de técnicas genômicas estirpes do gênero *Rhizobium* que nodulam *Platypodium elegans*: UFLA01-1127<sup>T</sup> e UFLA01-1134.

## **Material e métodos**

### **Origem das estirpes**

A análise filogenética pelo sequenciamento dos genes 16S rRNA, *gyrB* e *atpD* identificou que as estirpes UFLA01-1127 e UFLA01-1134 isoladas de nódulos de *Platypodium elegans* em viveiro poderiam se tratar de novas espécies (Lopez et al. 2021). Essas estirpes estão depositadas no Setor de Biologia, Microbiologia e Processos Biológicos do solo da Universidade Federal de Lavras, Brasil.

### **Sequenciamento e qualidade dos genomas**

As estirpes UFLA01-1127 E UFLA01-1134 foram crescidas em meio de cultura 79 (Fred and Waskman 1928) para posterior extração de DNA. O DNA genômico das estirpes foi extraído usando o Kit Wizard® Genomic DNA Purification, seguindo as recomendações do fabricante. A qualidade e concentração do DNA extraído foi analisada usando um espectrômetro Colibri. As amostras acima de 100 ng/microlitro foram liofilizadas e enviadas para sequenciamento do genoma. Leituras em pares (2 × 250 bases) foram sequenciados com o MiSeq Reagent kit 500v2 (Illumina) na plataforma MiSeq (Illumina) (Tabela 1).

Os dados de sequenciamento foram analisados na plataforma online KBase (Arkin AP et al. 2018) conforme descrito abaixo: importação do genoma, montagem da sequência usando SPADES 3.15.3 (Bankevich et al. 2012) (Tabela 1), verificação da qualidade do genoma com CheckM (Parks et al. 2015) (Tabela 2) e anotação dos genomas com Rast (Aziz et.al., 2008; Overbeek et al, 2014; Brettin 2015). (Tabela 3).

### **ANI (Average Nucleotide Identity)**

O ANI é uma comparação que se faz par a par entre duas ou mais sequências genômicas, tem sido uma alternativa à técnica de hibridização DNA-DNA com maior robustez e acurácia para identificação de espécies (Konstantinidis and Tiedje, 2005; Michel et al. 2021; Michel et al. 2020). Dentre as opções disponíveis para essa comparação foi utilizada plataforma disponível online (Rodriguez-R and Konstantinidis 2016), que utiliza ANI calculator para estimar a identidade média de nucleotídeos usando os melhores acertos (unidirecional ANI) e os melhores acertos recíprocos (bidirecional ANI) entre dois conjuntos genômicos (Goris et al. 2007). Esse algoritmo permite fazer a comparação de até 45 sequências diferentes por vez, o qual facilita a comparação de estirpes com todas as 94 estirpes tipo do mesmo gênero, cujos genomas estão disponíveis no NCBI. A plataforma gera uma matriz de identidade, na qual são observadas as porcentagens de similaridade entre sequências. Baseados em publicações prévias, foram utilizados os valores abaixo de 95~96% de similaridade para classificar as sequências de estudo como espécies novas (Goris et al. 2007; Richter and Rosselló-Móra 2009).

## GTDB.tk e TYGS

As sequências genômicas das estirpes UFLA01-1127 e UFLA01-1134 foram submetidas a análise GTDB.tk v1.7.0 (Chaumeil et al., 2019), que vem a ser um kit de ferramentas de software capaz de atribuir classificação taxonômica objetiva a genomas de bactérias, baseado em bancos de dados de taxonomia por genoma (Parks et al., 2020; 2018).

Além disso, essas sequências foram carregadas no TYGS (Type Genoma Server), uma plataforma de bioinformática gratuita para realizar análise taxonômica baseada no genoma inteiro (disponível em <https://tygs.dsmz.de>) (Meier-Kolthoff JP et al., 2022; 2019). Para a determinação dos genomas de estirpe tipo mais próximo foi feita primeiro, a comparação de todos os genomas do estudo com todos os genomas de estirpe tipo disponíveis no banco de dados TYGS por meio do algoritmo MASH, uma aproximação rápida de parentesco intergenômico (Ondov et al., 2016) e os dez estirpes tipo com as menores distâncias MASH escolhidas pelo próprio genoma. Em segundo lugar, um conjunto adicional de dez estirpes tipo estreitamente relacionadas foi determinado por meio das sequências do gene 16S rRNA. Estes foram extraídos dos genomas usando RNAmmer (Lagesen e Hallin 2007) e cada sequência foi subsequentemente BLASTed (Camacho et al., 2009) contra a sequência do gene 16S rDNA de cada uma das atualmente 18452 estirpes tipo disponíveis no banco de dados TYGS. Isso foi usado como um proxy para encontrar as 50 melhores estirpes tipo correspondentes (de acordo com o bitscore) para cada genoma e, subsequentemente, calculada as distâncias precisas usando a abordagem Genoma BLAST Distance Phylogeny (GBDP) sob o algoritmo 'coverage' e a fórmula de distância d5 (Meier-Kolthoff JP et al., 2013). Essas distâncias foram finalmente usadas para determinar os 10 genomas de estirpe tipo mais próximo para cada um dos genomas.

A comparação pareada de sequências genômicas foi conduzida usando GBDP e distâncias intergenômicas precisas inferidas sob o algoritmo 'trimming' e fórmula de distância d5 (Meier-Kolthoff JP et al., 2013). 100 réplicas de distância foram calculadas cada. Valores digitais de Hibridização DNA-DNA (dDDH) e intervalos de confiança foram calculados usando as configurações recomendadas do GGDC 3.0 (Meier-Kolthoff JP et al., 2022; 2013). As distâncias intergenômicas resultantes foram usadas para inferir uma árvore de evolução mínima balanceada com suporte de ramificação via FASTME 2.1.6.1 incluindo pós-processamento SPR (Lefort V et al., 2015). O suporte de ramificação foi inferido a partir de 100 réplicas pseudo-bootstrap cada. As árvores foram enraizadas no ponto médio (Farris JS, 1972) e visualizadas com PhyD3 (Kreft L et al., 2017).

O agrupamento de espécies baseado em tipo usando um raio dDDH de 70% em torno de cada uma das 45 estirpes tipo foi feito conforme descrito anteriormente (Meier-Kolthoff e Göker M, 2019). O agrupamento de subespécies foi feito usando um limiar de dDDH de 79% conforme introduzido anteriormente (Meier-Kolthoff et al., 2014).

## Multilocus Sequence Analysis (MLSA)

Foram obtidas as sequências dos genes *housekeeping atpD* e *recA* das estirpes tipo do gênero *Rhizobium* já descritas que não possuem genoma sequenciado disponível, ou seja, dentre as 121 estirpes já descritas de *Rhizobium*, foram obtidas sequências dos genes *housekeeping* das 27 que não possuem dados genômicos. Com base nessas sequências dos genes *housekeeping* com aquelas das estirpes UFLA01-1127 e UFLA01-1134 foi construída uma árvore filogenética no software MEGA X. Das espécies que foi possível obter as sequências de

todos os genes housekeeping prosseguiu-se a análise cluster concatenada no software SeaView. As sequências dos genes foram alinhadas utilizando MUSCLE (Edgar 2004). As árvores filogenéticas foram construídas usando os métodos estatísticos de junção de vizinhos (NJ) (Saitou e Nei 1987) e máxima verossimilhança (ML) (Felsenstein 1981), com o modelo General Time Reversible (GTR) com valor de correção gama e % de locais invariantes (G + I) com bootstrap de 1000 repetições para árvore *atpD-recA*. Modelos testados no MEGA X e escolhidos de acordo com o menor valor de BIC (Bayesian Information Criterion) (Nei e Kumar, 2000; Kumar et al., 2018) (detalhes do resultado de teste do modelo em tabela S1).

### Testes fenotípicos

As estirpes UFLA01-1127 e UFLA01-1134 foram cultivadas em meio de cultura 79 em diferentes concentrações de NaCl (p/v) (0.25; 0.50; 1.75 e 1%) e com diferentes níveis de pH (4.0, 5.5, 6.8, 8.0, 9.0 e 10). O crescimento das estirpes em meio de cultura 79 quando incubadas em diferentes temperaturas (15, 20, 28, 34, 37 e 40°C) também foi avaliado. Além disso foi testada sua suscetibilidade a antibióticos a partir da utilização de discos de antibiograma contendo: ácido nalidixico 30 µg mL<sup>-1</sup>, amoxicilina 10 µg mL<sup>-1</sup>, ampicilina 10 µg mL<sup>-1</sup>, azitromicina 15 µg mL<sup>-1</sup>, cefuroxima 30 µg mL<sup>-1</sup>, claritromicina 15 µg mL<sup>-1</sup>, clorafenicol 30 µg mL<sup>-1</sup>, doxicilina 30 µg mL<sup>-1</sup>, gentamicina 10 µg mL<sup>-1</sup>, rifampicina 30 µg mL<sup>-1</sup>, tetraciclina 30 µg mL<sup>-1</sup>, vancomicina 30 µg mL<sup>-1</sup>, neomicina 30 µg mL<sup>-1</sup> e kanamicina 30 µg mL<sup>-1</sup>. Foi verificada a capacidade de as estirpes assimilarem diferentes fontes de carbono (L-arabinose, D-arabinose, L-asparagina, ácido cítrico, glicerol, D-frutose, glicina, D-glicose, glutamina, ácido L-glutâmico, ácido málico, lactose, maltose, manitol, metionina, lactato de sódio e sacarose) e de nitrogênio (L-asparagina, L-arginina, hidrolisado de caseína, L-cisteína, glicina, ácido L-glutâmico, triptofano e L-metionina) em meio 79 modificado segundo Costa e colaboradores (2017). Como teste fenotípico complementar foi feito o teste API-Biomerieux, que permite realizar de forma rápida 20 diferentes testes bioquímicos com cada uma das estirpes. Por fim foram geradas eletromicrografias em microscópio eletrônico de varredura para observação das dimensões das estudadas (Fig 4).

### Resultado e discussão

#### ANI

A técnica ANI determina a distância evolutiva entre um par de linhagens a partir dos valores de similaridade de regiões genômicas homólogas compartilhadas entre os genomas testados (Konstantinidis e Tiedje 2005; Kim et al. 2014). Os valores de corte considerados para determinar se as espécies são novas estão abaixo de 95-96% (Goris et al. 2007; Konstantinidis e Tiedje 2005; De Lajudie et al., 2019; Avontuur et al., 2019). As estirpes UFLA01-1127 e UFLA01-1134 possuem maior similaridade com as espécies *Rhizobium hainanense* CCBAU 57015<sup>T</sup> (87.21% e 86.16%, respectivamente) e *Rhizobium tropici* CIAT 899<sup>T</sup> (86.99% e 86.96%, respectivamente). Todos os valores de ANI ficaram abaixo de 88%, mostrando que as estirpes do presente estudo são uma espécie nova. Ambas as estirpes possuem valor de ANI de quase 100% entre si, demonstrando que se tratam da mesma espécie.

Mais detalhes são apresentados na tabela 3. O conteúdo de G + C no DNA das linhagens UFLA01-1127 e UFLA01-1134 foi de 59,8 mol%.

### **GTDB.tk e TYGS**

A identificação taxonômica das linhagens mais próximas bem como os agrupamentos de espécies e subespécies resultantes estão listados na Tabela 4, enquanto a identificação taxonômica das estirpes de consulta é encontrada na Tabela 1. Resumidamente, o agrupamento rendeu 12 agrupamentos de espécies e as estirpes de consulta fornecidas foram atribuídas a 1 deles. Além disso, as estirpes foram localizadas em 1 dos 12 grupos de subespécies. Para as estirpes UFLA01-1127 e UFLA01-1134, observou-se maior proximidade com a espécie *Rhizobium hainanense* CCBAU 57015<sup>T</sup>. Esse resultado corrobora o que foi encontrado no GTDB.tk, ANI e MLSA do presente estudo. A representação gráfica das árvores filogenéticas do gene 16S rRNA e do genoma completo pode ser vista nas figuras 1 e 2. Os valores de dDDH abaixo de 70% e conteúdo de G+C com porcentagem entre espécies diferentes é menor que 1, o que mostra que as estirpes do presente estudo são espécies novas.

A filtragem de genes parálogos (genes na mesma espécie que surgiram de um gene no ancestral comum por duplicação) ou o relatório de intervalos de confiança para cada valor previsto de dDDH obtidos no TYGS são recursos que vão além dos recursos obtidos com ANI e fornecem informação mais robusta para identificar microrganismos a nível de espécie e até de subespécie (Meier-Kolthoff JP, 2014).

### **Multilocus Sequence Analysis**

De acordo com o MLSA realizado com os genes *atpD* e *recA* (Fig. 3), as estirpes UFLA01-1127 e UFLA01-1134 estão posicionadas em um clado separado de todas as espécies testadas do gênero *Rhizobium* com 53% de bootstrap. Ambas estirpes estão posicionadas juntas com alto suporte estatístico (bootstrap de 100%) e similaridade de 100% mostrando que são da mesma espécie. A espécie mais próxima delas foi *Rhizobium endophyticum* CCGE2052<sup>T</sup> com 88,98% de similaridade. (Tabela 5). É um valor bem abaixo do encontrado em artigos recentes de descrição de novas espécies do gênero *Rhizobium* que avaliaram a similaridade dos genes concatenados *atpD* e *recA*, como *Rhizobium changzhiense* WYCCWR 11279<sup>T</sup> que possui similaridade com *Rhizobium sophorae* CCBAU 03386<sup>T</sup> de 95,40% e *Rhizobium laguerreae* FB T<sup>T</sup> de 93,61% (Zhang 2021) ou *Rhizobium croatiense* 13T<sup>T</sup> que possui similaridade de 98% com *R. sophoriradicis* CCBAU 03470<sup>T</sup> e *Rhizobium redzepovicii* 18T<sup>T</sup> que possui similaridade de 97% com *Rhizobium ecuadorensense* CNPSO 671<sup>T</sup> (RAJNOVIC 2022). Isso mostra que as estirpes UFLA01-1127 e UFLA01-1134 correspondem a uma nova espécie.

### **Análise fenotípica**

As duas estirpes (UFLA01-1127 e UFLA01-1134) crescem em temperaturas de 20 a 37°C e numa ampla faixa de pH (4.0 a 10.0). Ambas estirpes são tolerantes a salinidade de 0.25 a 1% de NaCl no meio de cultura. Embora UFLA01-1127 e UFLA01-1134 sejam 100% semelhantes de acordo com o ANI, elas se comportaram de maneira diferente em relação às fontes de carbono e nitrogênio, confirmando o fato de não serem clones. A estirpe UFLA01-1127 cresceu em meio contendo glicerol como fonte de carbono, diferente de UFLA01-1134. Além disso a estirpe UFLA01-1127 cresceu em meio contendo L-cisteína como fonte de nitrogênio, diferente da UFLA01-1134. Em relação aos antibióticos testados ambas estirpes se comportaram da mesma forma, com sensibilidade

aos antibióticos: Doxyciclina, Tetraciclina, Neomicina, Gentamicina e Rifamicina. Resultados obtidos pelo API 20NE mostraram que ambas as estirpes reagiram aos testes da mesma maneira (Tabela 6). Mais detalhes são encontrados na descrição da espécie. Imagens das bactérias provenientes de amostras contendo colônias de UFLA01-1127 e UFLA01-1134 estão apresentadas na figura 4.

Descrição de *Rhizobium elegans* sp. nov.

*Rhizobium elegans* (e.le.gans. N.L. neut. adj. elegans of the leguminous tree species from which the strain was isolated from its nodules, *Platypodium elegans*).

The cells are gram-negative rods (approximately  $4.4 \times 0.35 \mu\text{m}$ ) and are aerobic and non-spore forming. Colonies are greater than 3 mm in diameter at 2 days, are circular, and produce an acid reaction in medium 79 with mannitol as a carbon source and bromothymol blue as an indicator. After few days produces a lot of mucus. The type strain tolerates up to 1% NaCl. It grows at pH levels of 4.0 to 10.0 and at temperatures ranging from 20 to 37°C, with optimal growth at 28°C. It is resistant to ampicillin ( $10 \mu\text{g mL}^{-1}$ ), amoxicillin ( $10 \mu\text{g mL}^{-1}$ ), nalidixic acid ( $30 \mu\text{g mL}^{-1}$ ), chloramphenicol ( $30 \mu\text{g mL}^{-1}$ ), clarithromycin ( $15 \mu\text{g mL}^{-1}$ ), cefuroxime ( $30 \mu\text{g mL}^{-1}$ ), kanamycin ( $30 \mu\text{g mL}^{-1}$ ), vancomycin ( $30 \mu\text{g mL}^{-1}$ ) and azithromycin ( $15 \mu\text{g mL}^{-1}$ ). It is sensitive to doxycycline ( $30 \mu\text{g mL}^{-1}$ ), tetracycline ( $30 \mu\text{g mL}^{-1}$ ), neomycin ( $30 \mu\text{g mL}^{-1}$ ), gentamicin ( $10 \mu\text{g mL}^{-1}$ ) and rifamycin ( $30 \mu\text{g mL}^{-1}$ ). Regarding carbon sources, it assimilates L-asparagine, glycerol, glucose, lactose, maltose, mannitol, sucrose, L-glutamine, glutamic acid and L-methionine. Weakly assimilates fructose. Does not assimilate glycine and sodium lactate. Regarding nitrogen sources, it assimilates L-asparagine, casein, L-cysteine, glycine, glutamic acid, tryptophan and L-methionine. The G+C content of UFLA 01-1127<sup>T</sup> is 59,8 mol%.

The type strain UFLA 01-1127<sup>T</sup> was isolated from effective nodules of *Platypodium elegans* (Voegel.) seedlings, from Minas Gerais, Brazil. The Genbank/NCBI accession numbers of the genome is xxxx. The housekeeping genes were extracted from the UFLA01-1127<sup>T</sup> type strain genome.

**Acknowledgements** We thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Process: 142473/2018-9), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (CAPES/PROEX AUXPE593/2018) and the Fundação de Amparo e Pesquisa de Minas Gerais (Fapemig) (CAG-RED- O0330- 16) for financial support and for granting scholarships. This research is associated with the Brazilian Instituto Nacional de Ciência e Tecnologia Biodiversidade do Solo (Soil Biodiversity/INCT-CNPq). The authors would also like to thank the Fapemig, Capes, CNPq and Finep (Financiadora de Estudos e Projetos) for maintaining the electron microscopy laboratory.

## Referências bibliográficas

- Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, et al. (2018) KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* 36:566–569. <https://doi.org/10.1038/nbt.4163>
- Aziz RK, Bartels D, Best AA, et al. (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19(5):455-77. <https://doi.org/10.1089/cmb.2012.0021>
- Barron et al., 2011
- Brettin T, Davis J, Disz T, et al. (2015) RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5: 8365. <https://doi.org/10.1038/srep08365>
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10: 421. DOI: 10.1186/1471-2105-10-421
- Costa EM, Guimarães AA, Vicentin RP, Ribeiro PRA, Leão ACR, Balsanelli E, Lebbe L, Aerts M, Willems A, Moreira FMS (2017) *Bradyrhizobium brasilense* sp. nov., a symbiotic nitrogen-fixing bacterium isolated from Brazilian tropical soils. *Arch Microbiol* 199:1211–1221. <https://doi.org/10.1007/s00203-017-1390-1>
- De Lajudie PM, Andrews M, Ardley J, Eardly B, Jumas-Bilak E, Kuzmanovic N, Lassale F, Lindström K, Mhamdi R, MartínezRomero E, Moulin L, Mousavi SA, Nesme X, Peix A, Putawska J, Steenkamp E, Stepkowski T, Tian CF, Vinuesa P, Wei G, Willems A, Zilli J, Young P (2019) Minimal standards for the description of new genera and species of rhizobia and agrobacteria. *Int J Syst Evol Microbiol* 69(7):1852–1863
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Farris JS. Estimating phylogenetic trees from distance matrices. *Am Nat*. 1972;106: 645–667.
- Felsenstein, J (1981) Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach. *J Mol Evol* 17:368-376
- Fred EB, Waskman SA (1928) Laboratory manual of general microbiology: with special reference to the microorganisms of the soil. McGraw-Hill, New York.
- Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM (2007) DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57:81–91. <https://doi.org/10.1099/ijs.0.64483-0>
- Kim M, Oh HS, Park SC, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 64:346–351. <https://doi.org/10.1099/ijs.0.059774-0>
- Konstantinidis KT, Tiedje JM (2005) Genomic insights that advance the species definition for prokaryotes. *PNAS* 7:2567–2572



- Kreft L, Botzki A, Coppens F, Vandepoele K, Van Bel M. PhyD3: A phylogenetic tree viewer with extended phyloXML support for functional genomics data visualization. *Bioinformatics*. 2017;33: 2946–2947. DOI: 10.1093/bioinformatics/btx324
- Lagesen K, Hallin P. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res. Oxford Univ Press*; 2007;35: 3100–3108. DOI: 10.1093/nar/gkm160
- Lefort V, Desper R, Gascuel O. FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol*. 2015;32: 2798–2800. DOI: 10.1093/molbev/msv150
- Lewis GP, Schrire B, Mackinder B, Lock M (2005) *Legumes of the world*. Royal Botanic Gardens, Kew, Richmond, Surrey
- Lopez BDO, Teixeira AFS, Michel DC, Guimarães AA, Costa AM, Costa JS, Pereira MS, Duarte BLM, Moreira FMS (2021) Genetic and symbiotic characterization of rhizobia nodulating legumes in a mining area in southeast Brazil. *Sci Agric* 79(2):e20200238. <https://doi.org/10.1590/1678-992X-2020-0238>
- Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*. 2013;14: 60. DOI: 10.1186/1471-2105-14-60
- Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, Fiebig A, et al. Complete genome sequence of DSM 30083T, the type strain (U5/41T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci*. 2014;9: 2. DOI: 10.1186/1944-3277-9-2
- Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun*. 2019;10: 2182. DOI: 10.1038/s41467-019-10210-3
- Meier-Kolthoff JP, Sardà Carbasse J, Peinado-Olarte RL, Göker M. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acid Res*. 2022;50: D801–D807. DOI: 10.1093/nar/gkab902
- Michel DC, Guimarães AA, Costa EM, Carvalho TS, Balsanelli E, Willems A, Souza EM, Moreira FMS (2020) *Bradyrhizobium uaiense* sp. nov., a new highly efficient cowpea symbiont. *Arch microbiol* 202:1135–1141. <https://doi.org/10.1007/s00203-020-01827-w>
- Michel DC, Costa EM, Guimarães AA, Carvalho TS, Caputo PSC, Willems A, Moreira FMS (2021) *Bradyrhizobium campsiandrae* sp. nov., a nitrogen-fixing bacterial strain isolated from a native leguminous tree from the Amazon adapted to flooded conditions. *Arch Microbiol* 203:233–240. <https://doi.org/10.1007/s00203-020-02022-7>
- Ondov BD, Treangen TJ, Melsted P, et al. Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biol* 2016;17: 1–14. DOI: 10.1186/s13059-016-0997-x
- Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R (2014) The SEED and the Rapid Annotation of microbial genomes

using Subsystems Technology (RAST). *Nucleic Acids Res* 42(Database issue):D206-14.  
<https://doi.org/10.1093/nar/gkt1226>

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25(7):1043-55.  
<https://doi.org/10.1101/gr.186072.114>

POWO (Plants of the World Online). *Machaerium nyctitans* (Vell.) Benth. Disponible em:  
<https://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:505530-1>. Accessed 12 July 2022

Richter PRA, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 106:19126-19131

Rodriguez-R LM, Konstantinidis KT (2016) The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints* 4:e1900v1.  
<https://doi.org/10.7287/peerj.preprints.1900v1>

Saitou N, Nei M (1987) The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol Biol Evol* 4(4):406-425

Ivana Rajnovic, Martha-Helena Ramírez-Bahena, Sanja Kajic, José M. Igual, Álvaro Peix, Encarna Velázquez, Sanja Sikora, *Rhizobium croatiense* sp. nov. and *Rhizobium redzepovicii* sp. nov., two new species isolated from nodules of *Phaseolus vulgaris* in Croatia, *Systematic and Applied Microbiology*, Volume 45, Issue 3, 2022, 126317, ISSN 0723-2020, <https://doi.org/10.1016/j.syapm.2022.126317>

Tamura K (1992) Estimation of the Number of Nucleotide Substitutions When There Are Strong Transition-Transversion and G+C-Content Biases. *Mol Biol Evol* 9(4):678-687.  
<https://doi.org/10.1093/oxfordjournals.molbev.a040752>

Zhang J, Peng S, Andrews M, Liu C, Zhang X (2021) *Rhizobium changzhiense* sp. nov., isolated from effective nodules of *Vicia sativa* L. in North China. *Int J Syst Evol Microbiol* 71:4812.  
<https://doi.org/10.1099/ijsem.0.004724>

## Tabelas e figuras

Tabela 1. SPAdes version: 3.15.3: quality assessment report and summary about genomes

<b>Statistics without reference</b>	<b>UFLA01-1127</b>	<b>UFLA01-1134</b>
# contigs	323	393
# contigs (>= 0 bp)	323	393
# contigs (>= 1000 bp)	299	363
# contigs (>= 10000 bp)	185	196
# contigs (>= 100000 bp)	2	2
# contigs (>= 1000000 bp)	0	0
Largest contig	135033	113141
Total length	6523025	6525000
Total length (>= 0 bp)	6523025	6525000
Total length (>= 1000 bp)	6504731	6502552
Total length (>= 10000 bp)	5961861	5710021
Total length (>= 100000 bp)	240150	216062
Total length (>= 1000000 bp)	0	0
N50	36894	31247
N75	20310	17021
L50	55	65
L75	114	135
GC (%)	59.8	59.8
<b>Mismatches</b>		
# N's	94	99
# N's per 100 kbp	1.44	1.52
<b>Predicted genes</b>		
# predicted genes (unique)	6384	6409
# predicted genes (>= 0 bp)	6247 + 137 part	6274 + 135 part

---

# predicted genes ( $\geq 300$ bp)	5558 + 125 part	5559 + 123 part
# predicted genes ( $\geq 1500$ bp)	678 + 14 part	680 + 12 part
# predicted genes ( $\geq 3000$ bp)	61 + 6 part	64 + 5 part

---

Tabela 2. CheckM v1.0.18: tabular representations of the phylogenetic marker completeness and contamination. CheckM generates clade-specific marker gene sets for each bin and reports the taxonomic resolution possible for each bin in the "Marker Lineage" column. Users may want to look at the "Marker Lineage" column to see what MAGs were classified with, for example, the "d\_\_Bacteria" or "d\_\_Archaea" marker sets. Instances where a broad (domain-level) marker set is used compared to a marker set from specific lineage (e.g. c\_\_Alphaproteobacteria) can help one contextualize (and evaluate) the genome completeness and contamination estimates.

Bin Name	Marker Lineage	# Genomes	# Markers	# Marker Sets	0	1	2	3	4	5+	Completeness	Contamination
UFLA01-1127_02-13.contigs	f__Rhizobiaceae	78	840	354	1	827	12	0	0	0	99.86	1.52
UFLA01-1134_02-12.contigs	f__Rhizobiaceae	78	840	354	1	824	15	0	0	0	99.86	2.05

Tabela 3. ANI with the species closest to the study strains. Values generated by the ANI calculator from the website: <http://enve-omics.ce.gatech.edu/g-matrix/index>

	UFLA01-1127	UFLA01-1134
UFLA01-1127		99,9998
UFLA01-1134	99,9998	
<b><i>Rhizobium hainanense</i> CCBAU 57015<sup>T</sup></b>	<b>87,21</b>	<b>86,16</b>
<i>Rhizobium dioscoreae</i> S-93 <sup>T</sup>	86,80	86,77
<i>Rhizobium freirei</i> PRF 81 <sup>T</sup>	85,83	85,79
<b><i>Rhizobium tropici</i> CIAT 899<sup>T</sup></b>	<b>86,99</b>	<b>86,96</b>
<i>Rhizobium multihospitium</i> CCBAU 83401 <sup>T</sup>	86,80	86,79

*Rhizobium miluonense* CCBAU 41251<sup>T</sup>

85,63

85,58

Tabela 4. Pairwise comparisons of user genomes vs. type strain genomes

Strain	Subject strain	dDDH (d0, in %)	C.I. (d0, in %)	dDDH (d4, in %)	C.I. (d4, in %)	dDDH (d6, in %)	C.I. (d6, in %)	G+C content difference (in %)
UFLA01-1127	UFLA01-1134	100	[100.0 - 100.0]	100	[100.0 - 100.0]	100	[100.0 - 100.0]	0
	<i>Rhizobium hainanense</i>							
UFLA01-1127	CCBAU 57015	48	[44.6 - 51.4]	34,3	[31.9 - 36.9]	44,3	[41.3 - 47.4]	0,15
	<i>Rhizobium hainanense</i>							
UFLA01-1134	CCBAU 57015	48,4	[45.0 - 51.8]	34,3	[31.9 - 36.8]	44,6	[41.6 - 47.6]	0,15
UFLA01-1134	<i>Rhizobium tropici</i> CIAT 899	47,1	[43.7 - 50.5]	33,7	[31.3 - 36.2]	43,4	[40.4 - 46.5]	0,31
UFLA01-1127	<i>Rhizobium tropici</i> CIAT 899	47,5	[44.1 - 50.9]	33,7	[31.3 - 36.2]	43,7	[40.7 - 46.7]	0,31
	<i>Rhizobium multihospitium</i>							
UFLA01-1127	HAMBI 2975	47,1	[43.7 - 50.5]	33,6	[31.2 - 36.1]	43,3	[40.4 - 46.4]	0,01
	<i>Rhizobium multihospitium</i>							
UFLA01-1134	HAMBI 2975	47	[43.6 - 50.4]	33,6	[31.2 - 36.1]	43,3	[40.3 - 46.3]	0,01
UFLA01-1127	<i>Rhizobium dioscoreae</i> S-93	52,2	[48.7 - 55.7]	33,4	[31.0 - 35.9]	47,2	[44.2 - 50.3]	0,14
UFLA01-1134	<i>Rhizobium dioscoreae</i> S-93	52,2	[48.7 - 55.6]	33,4	[31.0 - 36.0]	47,2	[44.2 - 50.2]	0,14

Table 5. Similarity analysis of concatenated gene sequences (*atpD* and *recA*) with higher values found among the study strains and sequences of *Rhizobium* type strains available at the NCBI

Strain	UFLA01-1127	UFLA01-1134
<i>Rhizobium endophyticum</i> CCGE2052	88,97900641	88,97900641
<i>Rhizobium yantingense</i> H66	84,00800336	84,00800336
<i>Rhizobium helanshanense</i> CCNWQTX14	83,87105425	83,87105425
<i>Rhizobium mayense</i> CCGE526	83,41077256	83,41077256
<i>Rhizobium zea</i> CRZM18R	82,95879214	82,95879214
<i>Rhizobium panacihumi</i> DCY116	81,1649116	81,1649116
<i>Rhizobium aquaticum</i> SA-276	81,08778059	81,08778059
<i>Rhizobium sphaerophysae</i> CCNWGS0238	81,05830263	81,05830263

Tabela 6. Results of Analytical profile index (API) tests on strains UFLA01-1127 and UFLA01-1134

Test	UFLA01-1127	UFLA01-1134
NO <sub>3</sub> (redução de nitratos em nitritos)	-	-
TRP (formação de indol (TRiptofano))	-	-
GLU (fermentação (GLUcose))	-	-
ADH (Arginina DiHydrolase)	+	+
URE (UREase)	+	+
ESC (Hidrólise (β-glucosidase) (ESculin))	+	+
GEL (Hidrólise (protease) (GELatina))	-	-
PNG (β-galactosidase (Para-NitroPhenyl-βDGalactopyranosidase))	+	+
GLU (assimilação (GLUcose))	+	+
ARA (assimilação (ARAbinose))	+	+

---

MNE (assimilação (ManNosE))	+	+
MAN (assimilação (MANnitol))	+	+
NAG (assimilação (N-Acetyl-Glucosamine))	+	+
MAL (assimilação (MALtose))	+	+
GNT (assimilação (potássio GlucoNato))	+w	+w
CAP (assimilação (ácido CAPrato))	-	-
ADI (assimilação (ácido ADIpató))	-	-
MLT (assimilação (MaLaTo))	+	+
CIT (assimilação (CITrato de trisódio))	+	+
PAC (assimilação (ácido fenil-ACetato))	-	-

---



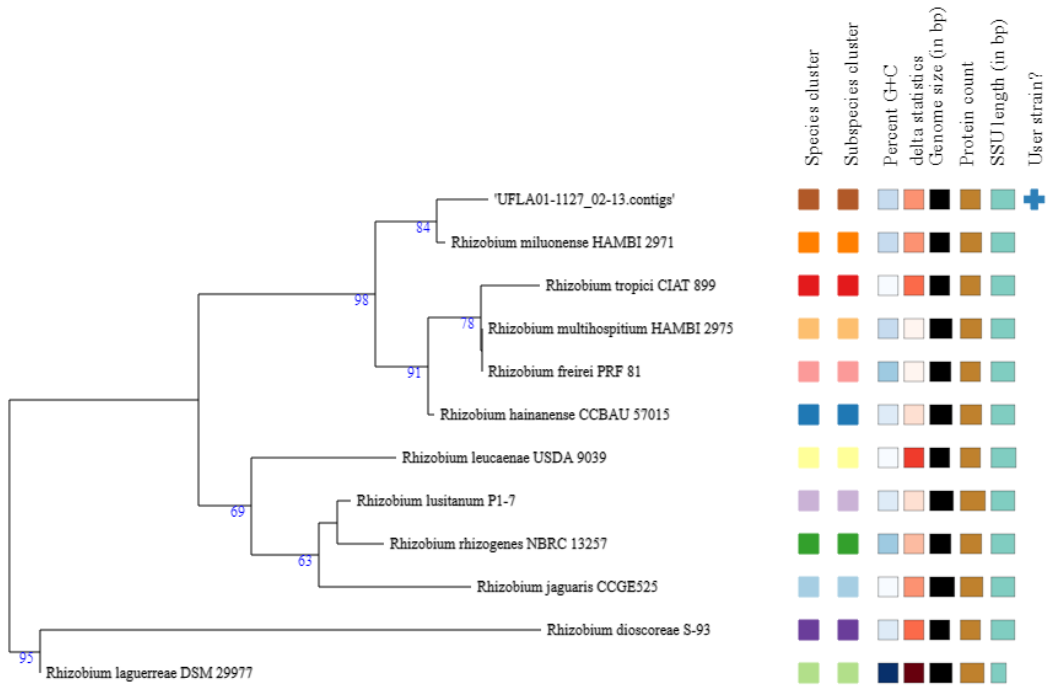


Figura 1. Árvore inferida com FastME 2.1.6.1 (Lefort V et al., 2015) a partir de distâncias GBDP calculadas a partir de seqüências do gene 16S rDNA. Os comprimentos dos ramos são dimensionados em termos da fórmula de distância GBDP  $d_5$ . Os números acima das ramificações são valores de suporte pseudo-bootstrap GBDP > 60% de 100 replicações, com um suporte médio de ramificação de 75,8%. A árvore foi enraizada no ponto médio (Farris JS, 1972).

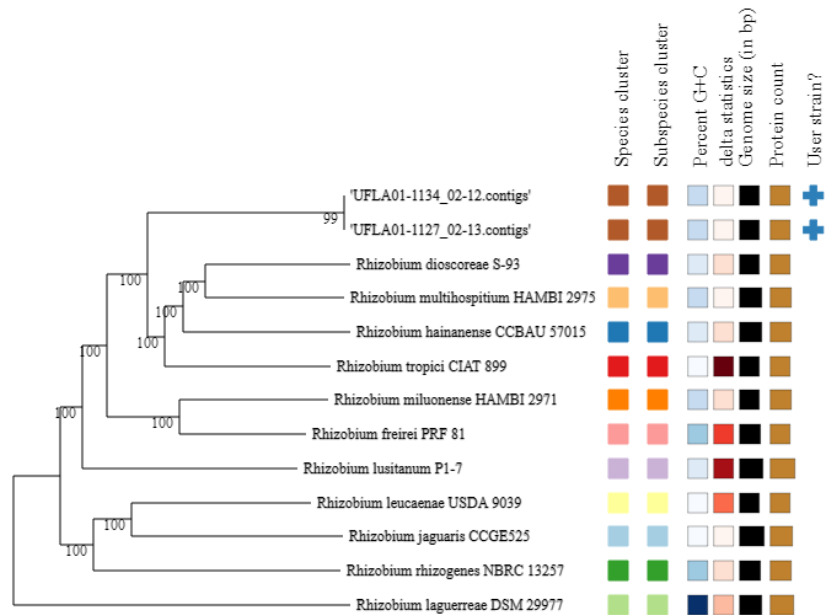


Figura 2. Árvore inferida com FastME 2.1.6.1 (Lefort V et al., 2015) a partir de distâncias GBDP calculadas a partir de seqüências genômicas. Os comprimentos dos ramos são dimensionados em termos da fórmula de

distância GBDP d5. Os números acima das ramificações são valores de suporte pseudo-bootstrap GBDP > 60% de 100 replicações, com um suporte médio de ramificação de 99,9%. A árvore foi enraizada no ponto médio (Farris JS, 1972).

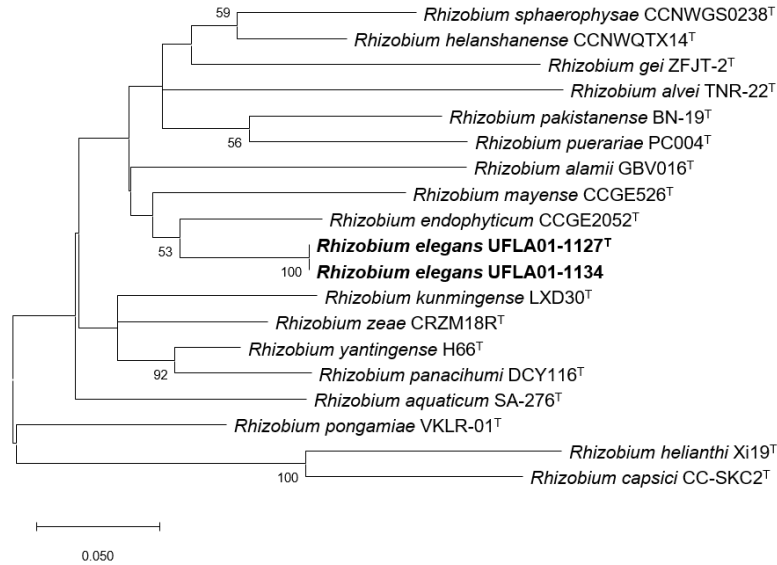


Figura 3. Maximum likelihood phylogeny based on the concatenated genes (*atpD* and *recA*), showing the relationship between strains of the new species (shown in bold) and type strains belonging to the genus *Rhizobium*. Bootstrap values greater than 50% are indicated on nodes. NCBI accession numbers are given in parentheses.

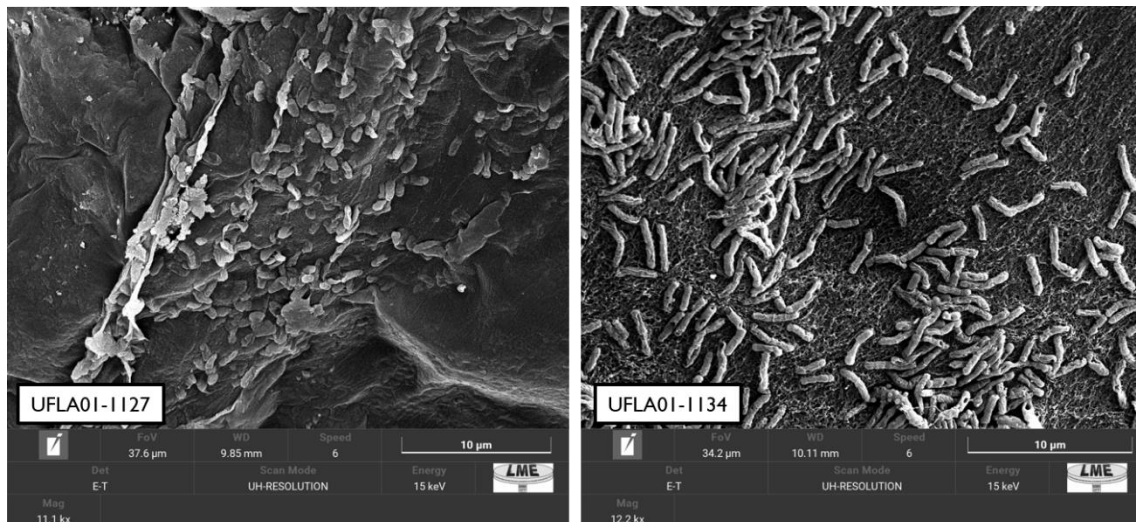


Fig. 4 Eletromicrografia gerada a partir de Microscópio eletrônico de varredura (MEV) no Laboratório Multiusuário do Departamento de Fitopatologia (UFLA) de amostras de colônias bacterianas das estirpes UFLA01-1127 e UFLA01-1134

Table S1. Model test *atpD-recA* concatenated genes

Model	#Param	BIC	AICc	lnL	Invariant	Gamma	R	Freq A	Freq T	Freq C	Freq G	A=>T	A=>C	A=>G	T=>A	T=>C	T=>G	C=>A	C=>T	C=>G	G=>A	G=>T	G=>C
GTR+G+I	45	8929,9 79668	8594,1 74211	- 4251,92 679	0,5300 94984	0,8207 763	2,0243 5583	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 2	0,0 1	0,1 2	0,0 2	0,3 1	0,0 6	0,0 1	0,1 9	0,0 7	0,0 7	0,0 3	0,0 7
TN93+G+I	42	8947,3 33793	8633,8 95785	- 4274,80 8055	0,5254 95901	0,7598 17904	2,0492 1149	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,1 1	0,0 3	0,3 1	0,0 5	0,0 3	0,1 9	0,0 5	0,0 7	0,0 3	0,0 5
T92+G+I	39	8949,3 01767	8658,2 34013	- 4289,99 6245	0,5425 41519	0,8040 30616	2,0227 4263	0,1915 41905	0,1915 41905	0,3084 58095	0,3084 58095	0,0 3	0,0 5	0,2 1	0,0 3	0,2 1	0,0 5	0,0 3	0,1 3	0,0 5	0,1 3	0,0 3	0,0 5
GTR+G	44	8955,0 28805	8626,6 78853	- 4269,18 6093	n/a	0,2163 44584	1,8644 6646	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 2	0,0 2	0,1 3	0,0 2	0,2 9	0,0 6	0,0 1	0,1 8	0,0 8	0,0 8	0,0 3	0,0 8
T92+G	38	8967,9 56271	8684,3 45893	- 4304,05 8232	n/a	0,1972 88342	1,9143 1685	0,1915 41905	0,1915 41905	0,3084 58095	0,3084 58095	0,0 3	0,0 5	0,2 1	0,0 3	0,2 1	0,0 5	0,0 3	0,1 3	0,0 5	0,1 3	0,0 3	0,0 5
TN93+G	41	8975,7 56382	8669,7 74814	- 4293,75 4084	n/a	0,2140 11546	1,8578 0471	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,1 3	0,0 3	0,2 9	0,0 5	0,0 3	0,1 7	0,0 5	0,0 8	0,0 3	0,0 5
HKY+G+I	41	8976,5 98355	8670,6 16787	- 4294,17 5071	0,5431 97781	0,8045 24989	2,0616 5434	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,2 2	0,0 3	0,2 1	0,0 5	0,0 3	0,1 3	0,0 5	0,1 4	0,0 3	0,0 5
HKY+G	40	8995,0 50431	8696,5 25615	- 4308,13 5843	n/a	0,1964 81293	1,9536 8458	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,2 1	0,0 3	0,2 5	0,0 5	0,0 3	0,1 2	0,0 5	0,1 3	0,0 3	0,0 5
GTR+I	44	9033,0 69347	8704,7 19395	- 4308,20 6364	0,6329 4532	n/a	1,7198 5845	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 2	0,0 2	0,1 3	0,0 3	0,2 8	0,0 7	0,0 1	0,1 7	0,0 8	0,0 8	0,0 4	0,0 8
TN93+I	41	9059,2 46471	8753,2 64903	- 4335,49 9129	0,6335 00357	n/a	1,6855 8543	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,1 3	0,0 4	0,2 7	0,0 6	0,0 4	0,1 7	0,0 6	0,0 8	0,0 3	0,0 5
T92+I	38	9071,5 76228	8787,9 6585	- 4355,86 821	0,6382 92991	n/a	1,7376 9438	0,1915 41905	0,1915 41905	0,3084 58095	0,3084 58095	0,0 3	0,0 5	0,2 3	0,0 3	0,2 5	0,0 5	0,0 3	0,1 2	0,0 5	0,1 2	0,0 3	0,0 5
K2+G	37	9073,1 16703	8796,9 6401	- 4361,37 3182	n/a	0,2077 13454	1,6883 1441	0,25	0,25	0,25	0,25	0,0 5	0,0 5	0,1 6	0,0 5	0,1 6	0,0 5	0,0 5	0,1 6	0,0 5	0,1 6	0,0 5	0,0 5
K2+G+I	38	9075,1 10162	8791,4 99783	- 4357,63 5177	0,5252 09474	0,7628 5315	2,3526 7225	0,25	0,25	0,25	0,25	0,0 4	0,0 4	0,1 8	0,0 4	0,1 8	0,0 4	0,0 4	0,1 8	0,0 4	0,1 8	0,0 4	0,0 4
HKY+I	40	9099,6 64994	8801,1 40177	- 4360,44 3124	0,6383 56902	n/a	1,7597 2884	0,1981 78731	0,1849 05078	0,3024 38648	0,3144 77543	0,0 3	0,0 5	0,2 3	0,0 3	0,2 5	0,0 5	0,0 3	0,1 2	0,0 5	0,1 3	0,0 3	0,0 5

