Current Research in Microbial Sciences 2 (2021) 100074

Contents lists available at ScienceDirect

Current Research in Microbial Sciences

journal homepage: www.sciencedirect.com/journal/current-research-in-microbial-sciences

Draft genome sequence of *Bradyrhizobium* sp. strain Oc8 isolated from *Crotalaria ochroleuca* nodule

Verónica Berriel [a], María A. Morel [b,c], Carla V. Filippi [b], Jorge Monza [b,*]

[a] Centro de Aplicaciones de Tecnología Nuclear en Agricultura Sostenible, Departamento de Suelos y Aguas, Facultad de Agronomía, Universidad de la República, Av. Garzón 809, Montevideo PC 12.900, Uruguay
[b] Departamento de Biología Vegetal, Laboratorio de Bioquímica, Universidad de la República, Av. Garzón 809, Montevideo, PC 12.900, Uruguay
[c] Laboratorio de Microbiología del Suelo (LMS), Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo, PC 11.400, Uruguay

ARTICLE INFO

Keywords:
Draft genome
Bradyrhizobium
Biological nitrogen fixation

ABSTRACT

In this study, we report the draft genome sequence of *Bradyrhizobium* sp. strain Oc8, a rhizobium isolated from *Crotalaria ochroleuca*, efficient in *C. ochroleuca, C. juncea, C. spectabilis*, and *Cajanus cajan*. The whole genome of the strain Oc8 contains 46 scaffolds, 8,283,342 bp, and 63.27% of GC content. *Bradyrhizobium* sp. Oc8 is an effective nitrogen-fixing bacterium with potential use as an inoculant for legumes used as cover crops and green manures.

Rhizobia are Gram-negative bacteria belonging to alpha and beta-proteobacteria that establish nitrogen-fixing symbiosis with legumes. This association makes legumes self-sufficient in nitrogen (N) and important in ecological and economic terms (Lorite et al., 2018). The use of legumes as cover crops offers advantages for the environment and agriculture since they contribute N to the ecosystem through biological fixation (Berriel et al., 2020), increasing soil productivity and the yield of the subsequent cash crops (Mahama et al., 2016). *Crotalaria ochroleuca, C. juncea, C. spectabilis* and *Cajanus cajan*, used as cover crops associated with specific rhizobia have a potential to fix N (Oliveira et al., 2007; Pereira et al., 2016; Berriel et al., 2020). These tropical forage legumes are nodulated by a relatively large group of rhizobia (Jorrin et al., 2021), and so their agronomic evaluation should include the rhizobia present in the soil.

In this study, we report the draft genome of *Bradyrhizobium* sp. Oc8 strain, isolated from a nodule of *C. ochroleuca* grown in soil of Uruguay (34.6 S, 55.6 W). Rhizobia isolation was carried out using the nodule squash technique (Gaunt et al., 2004) after surface sterilization according to Batista et al. (2015). A drop of the resulting suspension was subsequently spread onto YEM agar medium (Vincent, 1970) and incubated at 28 °C for 4–5 days. Strain Oc8 was obtained by picking a single colony from the agar plate. The isolated strain was checked for its ability to nodulate its host plant *C. ochroleuca, C. juncea, C. spectabilis* and *C. cajan* as described by Batista et al. (2015).

Oc8 strain was grown in a liquid YEM medium with 180 rpm orbital shaking for 24 h at 27 °C. Subsequently, genomic DNA was extracted using QIAamp DNA Micro Kit (QIAGEN, Germany). Whole-genome sequencing (Novaseq-Illumina, paired-end, PE, 2 × 151 bp) was performed at Macrogen (Korea). Sequencing quality was visually inspected using FastQC (Andrews, 2010) and Trimmomatic (v0.36) (Bolger et al., 2014) was used to discard/trim low-quality reads, keeping 94.35% of the initial PE (i.e., 10,379,129 PE reads). Unicycler (v0.4.7) was used for de novo contig assembly (Wick et al., 2017) yielding 63 contigs. After that, SSPACE (v2.1) was used for scaffolding (Boetzer et al., 2011). Assembly statistics, for both contig and scaffold level assemblies, were obtained using QUAST (v5.0.20) (Gurevich et al., 2013). While scaffolding generated a significant fragmentation reduction, it had no impact on main assembly metrics as the largest contig length, N50 or L50. Thus, the generated draft genome comprises 46 scaffolds, covering 8283,342 bp (largest contig: 1882,916 bp; N50: 537,804 bp; L50: 5; N's per 100,000 bp: 1.03). The GC content was estimated at 63.27%. Blastn (v2.5.0, Altschul et al., 1990) was locally run, with the NCBI RefSeq virus database (v5), in order to check for potential viral (phage) contamination. In addition, PlasmidFinder (v2.0.1, default parameters,

* Corresponding author.
E-mail address: monzajorge@gmail.com (J. Monza).

https://doi.org/10.1016/j.crmicr.2021.100074
Received 14 August 2021; Received in revised form 1 October 2021; Accepted 1 October 2021
Available online 15 October 2021

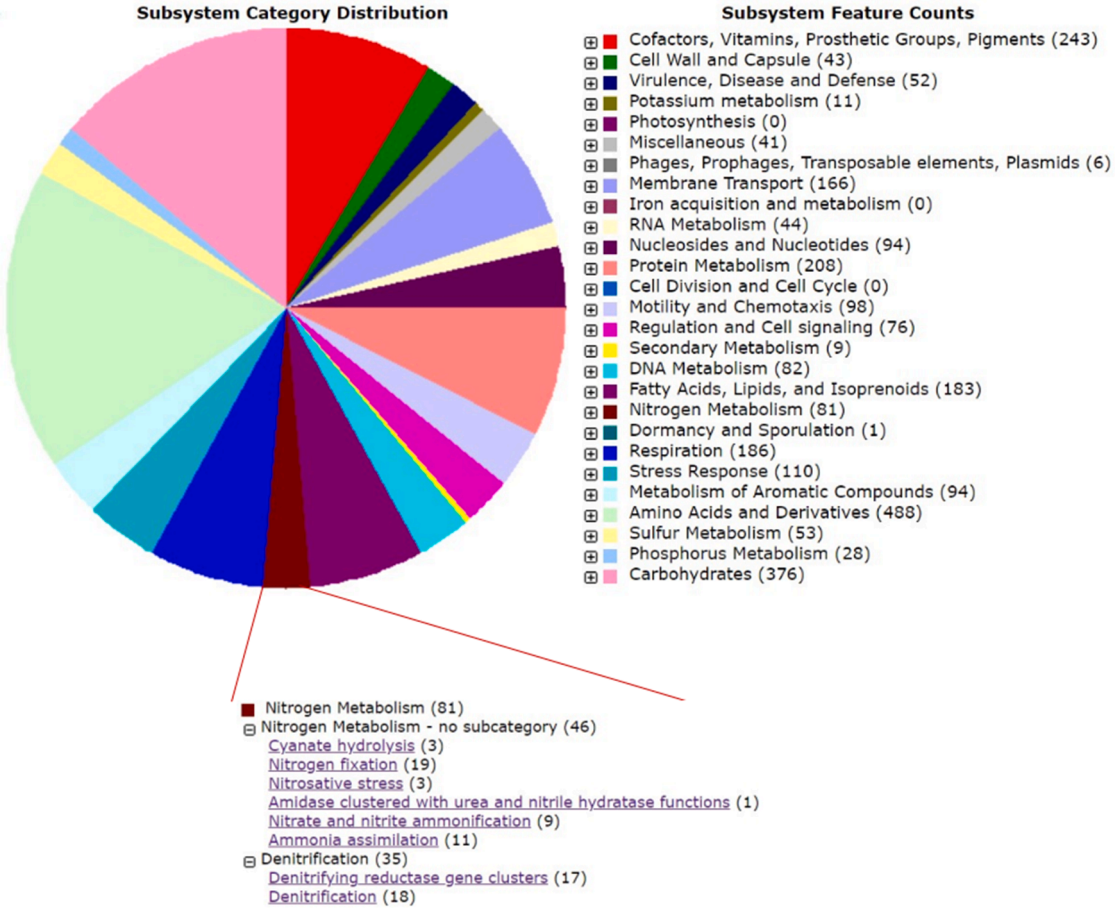

Fig. 1. Bacterial genome representation showing subsystem category distribution of coding sequences (CDS) from strain Oc8, generated through RASTtk pipeline. The number of CDS in the subsystem is shown in brackets.

Carattoli et al., 2014) and plasmidSPAdes (v3.13.1, Bankevich et al., 2012) were used to discard plasmid contamination.

Benchmarking Universal Single-Copy Orthologs (BUSCO, v5.1.2) was used to assess the completeness of the assembly (dataset: bacteria_odb10) (Simão et al., 2015). Of 124 BUSCOs, 123 were complete (123/124, 99.2%), while one was fragmented (shorter than expected). From the 123 complete BUSCOs 121 were single-copy and two were duplicated. Finally, prokka (v1.12) was used for genome annotation (Seemann, 2014), obtaining 7776 predicted coding sequencing (CDS), 3 rRNA and 52 tRNA. Genome annotation was also carried out through Rapid Annotation Using Subsystem Technology (RAST) server (v2.0) (Overbeek et al., 2013). The most abundant subsystem was Amino acids and derivatives, followed by Carbohydrates. CDS related to N metabolism stood out among genes of agricultural importance (Fig. 1). A complete view of the genome was generated using the CGView Server (Fig. 2) (Petkau et al., 2010).

The 16S rRNA gene sequence was extracted from Oc8 genome using RNAmmer (Lagesen and Hallin, 2007) and it was BLASTed (Camacho et al., 2009) against the 16S rRNA gene sequence of each of the currently type strains available in Type Genome Server (TYGS) database (Meier-Kolthoff and Göker, 2019). Additionally, an extended 16S rRNA gene analysis, performed to detect not yet genome-sequenced type strains relevant to the study, was performed via the Genome-to-Genome Distance Calculator (GGDC) web server (Meier-Kolthoff et al., 2013). For maximum likelihood (ML) tree inference, rapid bootstrapping in conjunction with the autoMRE bootstopping criterion (Pattengale et al., 2010) and subsequent search for the best tree was used. For maximum parsimony (MP) tree inference 1000 bootstrapping replicates were used in conjunction with tree-bisection-and-reconnection branch swapping and ten random sequence addition replicates. Since 16S rRNA gene sequences are conserved in *Bradyrhizobium* (Willems et al., 2001), phylogenetic analysis based on two housekeeping genes, *recA* and *ftsA*

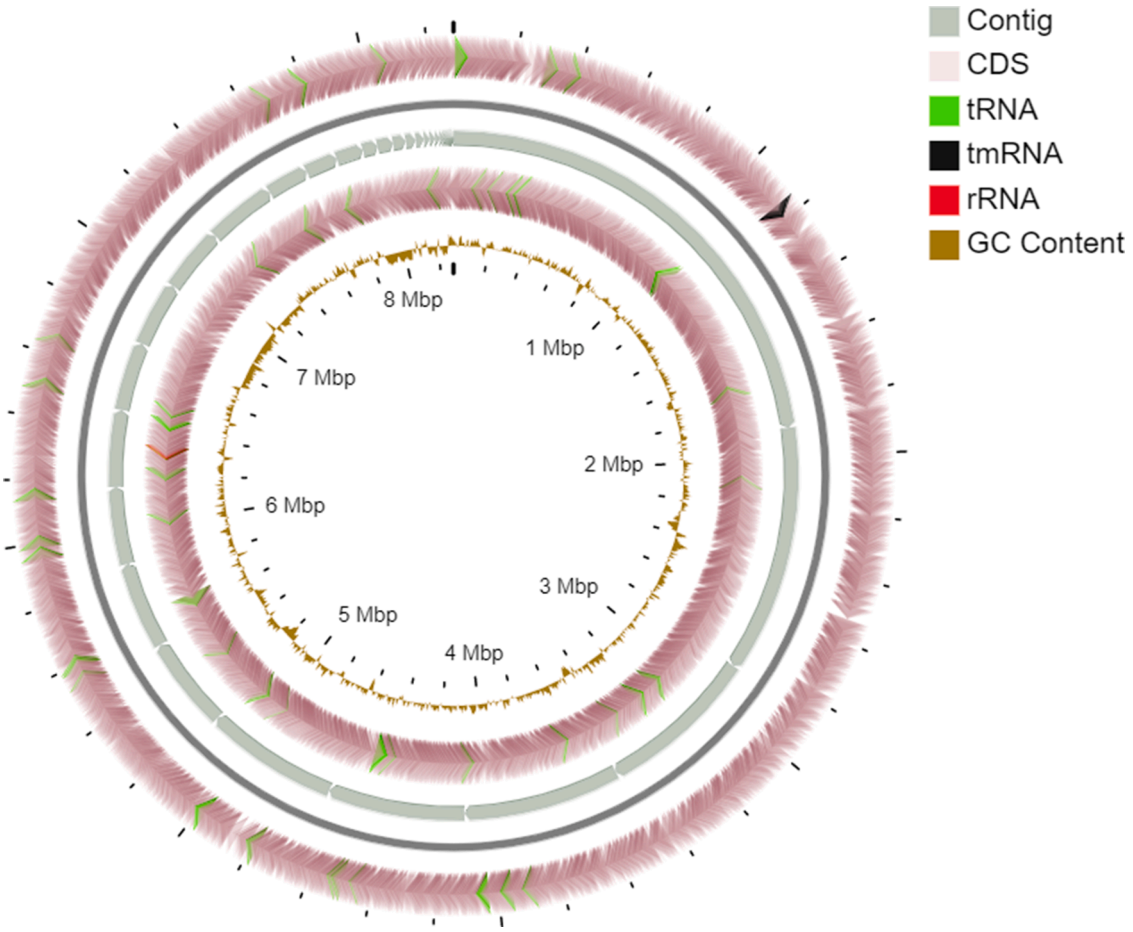

Fig. 2. Circular bacterial genome containing coding sequences (CDS), tRNAs, rRNAs, and GC content skew. The map was generated using the CGView Server beta online software.

(encoding for an actin-like protein involved in prokaryotic cell division) were also conducted, as recommended by Ormeño-Orrillo and Martinez (2019) and Kalita and Malek (2019), respectively. The *recA* and *ftsA* sequences obtained from the genome were compared with *recA* and *ftsA* sequences available in GenBank. Alignment and MP trees were constructed with the MEGA 7 software (Kumar et al., 2016).

The extended 16S rRNA gene-based analysis indicated that the isolate is a *Bradyrhizobium* sp. (Fig. 3.). The ML bootstrapping converged after 950 replicates; the average support was 53.70%. MP analysis yielded the best score of 185 (consistency index 0.56, retention index 0.84) and 50 best trees. The MP bootstrapping average support was 49.35%. Gene comparisons of *recA* and *ftsA* sequences of Oc8 versus publicly available sequences showed lower than 94% and 96.67% of identity percentages, respectively, either with *B. guangzhouense, B. guangdongense, B. diazoefficiens*, and several others *Bradyrhizobium* sp. According to Ormeño-Orrillo and Martinez (2019), nucleotide identities of 98.2% for *recA* could be used as cutoff values to discriminate between described bradyrhizobial species. Kalita and Malek (2019) reported that the *ftsA* sequence similarity range from 80 to 97.4% between *Bradyrhizobium* species (Fig. 4). Based on those reports, the strain Oc8 of *Bradyrhizobium* does not show a close genetic relationship with any *Bradyrhizobium* species.

Although our results showed that the phylogenies of *ftsA* and *recA* were congruent with a possible new species of *Bradyrhizobium*,

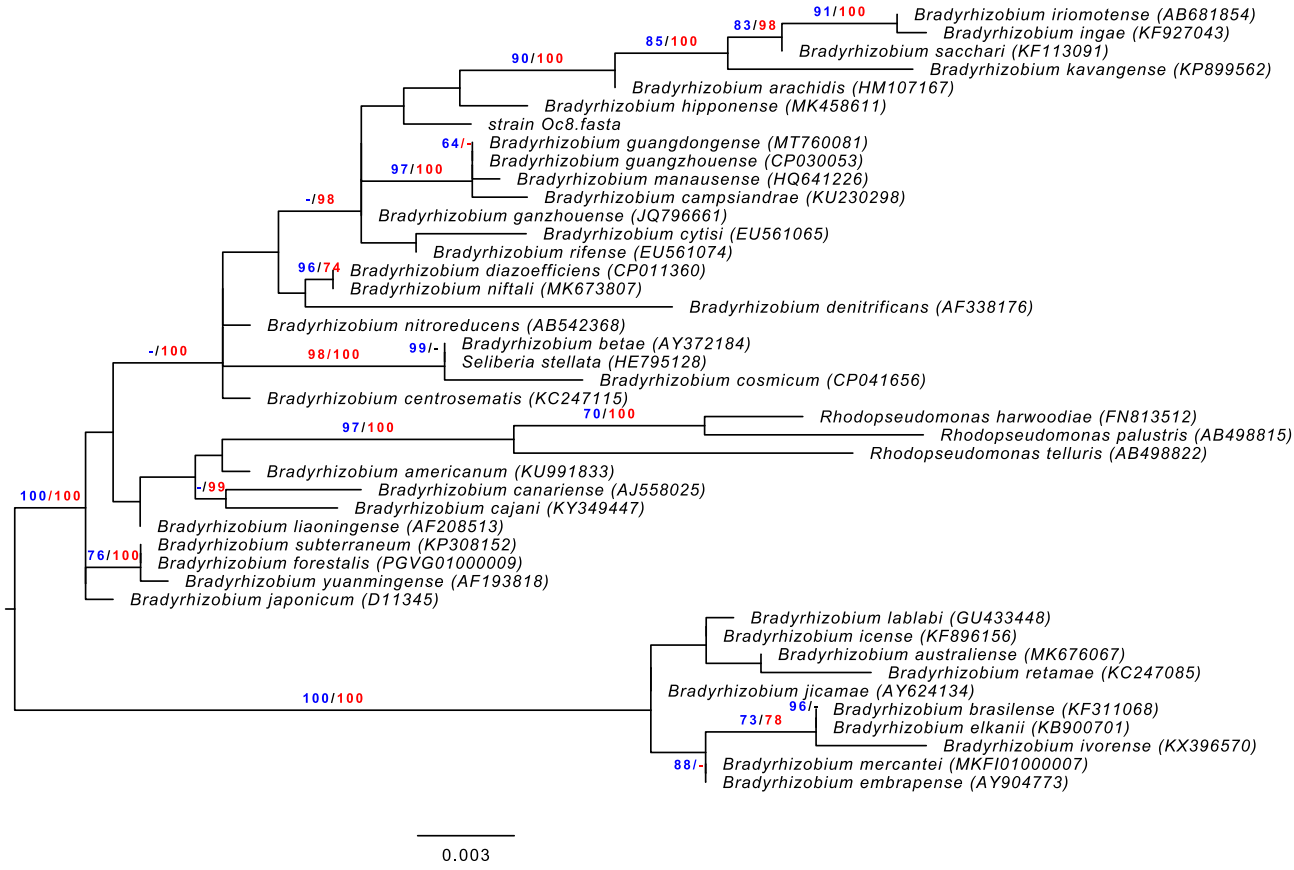

Fig. 3. ML tree inferred from sequences alignments of 16S rRNA sequences under the GTR+GAMMA model and rooted by midpoint-rooting performed via the GGDC web server (Meier-Kolthoff et al., 2013). The branches are scaled in terms of the expected number of substitutions per site. The numbers above the branches are support values when larger than 60% from ML (blue) and MP (red) bootstrapping.

additional analysis should be performed to verify the taxonomic affiliation of strain Oc8. Next, we will determine measures of nucleotide-level genomic similarity, multilocus phylogenetic analysis, and characterization of biochemical and metabolic attributes.

Data availability

The draft genome sequences have been deposited in GenBank under the BioProject accession number PRJNA752993. The version described in this paper is the first version.

Author contributions

Conceptualization, V.B, and J.M.; writing—original draft preparation and visualization, M.A. M., and C.V.F.; writing—review and editing were developed by all authors; supervision, J.M; project administration, V.B. and J.M.; funding acquisition, V.B. All authors have read and agreed to the published version of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

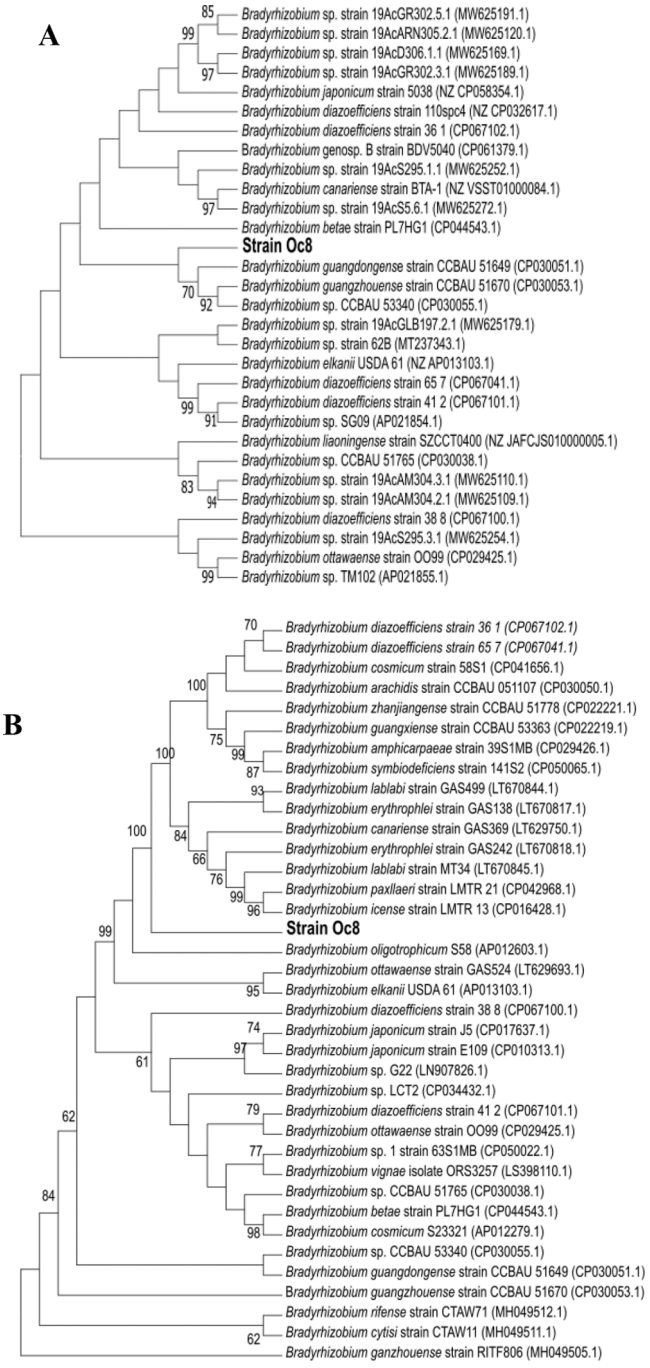

Fig. 4. MP trees inferred from sequences alignments of *recA* (A), and *ftsA* (B) genes using gene sequences selected among the first 100 hit sequences from Blast search and comparison. The numbers in each branch represent bootstrap support values of >60% from 1000 replications.

Acknowledgement

This work was financially supported by la Agencia Nacional de Investigación e Innovación de Uruguay Grant: ANII-FMV_3_2016_1_125492; Doctorado en Biotecnología (Facultad de Ciencias, Universidad de la República) and CAP (Comisión Académica de Postgrado, Universidad de la República). The authors are members of the National Research System (SNI), and Programa de Desarrollo de Ciencias Básicas (PEDECIBA).

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J, 1990. Basic local alignment search tool. J. Mol. Biol. 215 (3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A, 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Computat. Biol.: J. Computat. Mol. Cell Biol. 19 (5), 455–477. https://doi.org/10.1089/cmb.2012.0021.

Batista, L., Irisarri, P., Rebuffo, M., Cuitiño, M., Sanjuán, J., Monza, J., 2015. Nodulation competitiveness as a requisite for improved rhizobial inoculants of *Trifolium pratense*. Biol. Fertil. Soils 51, 1–20. https://doi.org/10.1007/s00374-014-0946-3.

Berriel, V., Monza, J., Perdomo, C.H., 2020. Cover crop selection by jointly optimizing biomass productivity, biological nitrogen fixation, and transpiration efficiency: application to two *Crotalaria* species. Agronomy 10. https://doi.org/10.3390/agronomy10081116.

Boetzer, M., Henkel, C.V., Jansen, H.J., Butler, D., Pirovano, W., 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27, 578–579. https://doi.org/10.1093/bioinformatics/btq683.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. J. Bioinform. 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al., 2009. BLAST+: architecture and applications. BMC Bioinformat. 10, 421. https://doi.org/10.1186/1471-2105-10-421.

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., Møller Aarestrup, F., Hasman, H., 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob. Agents Chemother. 58 (7), 3895–3903. https://doi.org/10.1128/AAC.02412-14.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086.

Jorrin, B., Maluk, M., Atoliya, N., Kumar, S.C., Chalasani, D., Tkacz, A., Singh, P., Basu, A., Pullabhotla, S.VSRN, Kumar, M., Mohanty, S.R., East, A.K., Ramachandran, V.K., James, E.K., Podile, A.R., Saxena, A.K., Rao, D.L.N., Poole, P. S., 2021. Genomic diversity of pigeon pea (*Cajanus cajan* L. Millsp.) endosymbionts in India and selection of potential strains for use as agricultural inoculants. Front. Plant Sci. 12. https://www.frontiersin.org/article/10.3389/fpls.2021.680981.

Kalita, M., Malek, W., 2019. The ftsA gene as a molecular marker for phylogenetic studies in Bradyrhizobium and identification of Bradyrhizobium japonicum. J. Appl. *Genet.* 60, 123–126.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054.

Lagesen, K., Hallin, P., 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35, 3100–3108. https://doi.org/10.1093/nar/gkm160. Oxford Univ Press.

Lorite, L., Estrella, M., Escaray, F., Sannazaro, A., Videira-Castro, I., Monza, J., Sanjuán, J., León-Barrios, M., 2018. The rhizobia-lotus symbioses: deeply specific and widely diverse. Front. Microbiol. 9, 2055. https://doi.org/10.3389/fmicb.2018.02055.

Mahama, G.Y., Prasad, P.V., Roozeboom, K.L., Nippert, J.B., Rice, C.W., 2016. Response of maize to cover crops, fertilizer nitrogen rates, and economic return. Agron. J. 108, 17–31. https://doi.org/10.2134/agronj15.0136.

Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.-.P., Göker, M., 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinform. 14, 60. https://doi.org/10.1186/1471-2105-14-60.

Meier-Kolthoff, J.P., Göker, M., 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat. Commun. 10, 2182. https://doi.org/10.1038/s41467-019-10210-3.

Oliveira, F., Guerra, J., Ribeiro, R., Almeida, D., Silva, E., Urquiaga, S., Espindola, J., 2007. The use of sunn hemp as green manure intercropped with taro. Hortic. Bras. 25, 562–566. https://doi.org/10.1590/S0102-05362007000400013.

Ormeño-Orrillo, E., Martínez-Romero, E., 2019. A genomotaxonomy view of the *Bradyrhizobium* genus. Front. Microbiol. 10, 1334. https://doi.org/10.3389/fmicb.2019.01334.

Overbeek, R., Olson, G.D., Pusch, G.J., Olsen, J.J., Davis, T., et al., 2013. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic. Acids. Res. 42, D206–D214. https://doi.org/10.1093/nar/gkt1226.

Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R.P., Moret, B.M.E., Stamatakis, A., 2010. How many bootstrap replicates are necessary? J. Comput. Biol. 17, 337–354. https://doi.org/10.1089/cmb.2009.0179.

Pereira, N., Soares, I., Miranda, F., 2016. Decomposition and nutrient release of leguminous green manure species in the Jaguaribe-Apodi region, Ceará, Brazil. Cienc. Rural 46, 970–975. https://doi.org/10.1590/0103-8478cr20140468.

Petkau, A., Stuart-Edwards, M., Stothard, P., Van Domselaar, G., 2010. Interactive microbial genome visualization with GView. Bioinformatics 26, 3125–3126. https://doi.org/10.1093/bioinformatics/btq588.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351.

Vincent, J.M., 1970. A Manual for the Practical Study of Root-Nodule Bacteria, International Biological Programme Handbook No 15. Blackwell Scientific Publications, Oxford, United Kingdom.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E., 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. 13, e1005595 https://doi.org/10.1371/journal.pcbi.1005595.

Willems, A., Coopman, R., Gillis, M., 2001. Phylogenetic and DNA-DNA hybridization analyses of *Bradyrhizobium* species. Int. J. Syst. Evol. Microb. 51, 111–117. https://doi.org/10.1099/00207713-51-1-111.