

Draft genome sequence of *Methylobacterium aquaticum* LEGMi-203a, isolated from root nodules of *Pithecellobium hymenaeifolium*

Valeria Castro-Camacho,¹ Rubí Robles-Azor,¹ Luciana Rodríguez-Burdock,¹ Keilor Rojas-Jimenez,¹ Bradd Mendoza-Guido²

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We report the draft genome of *Methylobacterium aquaticum* LEGMi-203a, a root nodule isolated from *Pithecellobium hymenaeifolium*. Genomic analysis supports its classification as *M. aquaticum*, and annotated nitrogen fixation and nodulation genes underscore its possible functional capabilities as a symbiont in tropical plants.

KEYWORDS symbiosis, nitrogen fixation, Costa Rica

The genus *Methylobacterium* includes species known for plant growth promotion and facilitation of symbiosis within root nodules (1). *M. aquaticum* has shown the ability to utilize methanol, a by-product of plant metabolism, as a carbon source, emphasizing its relevance in symbiotic interactions (2). Genomic analysis has revealed the evolutionary and functional diversity of *Methylobacterium*, highlighting their plant-associated traits (3).

We obtained the isolate in August 2021 from root nodules of the native legume *P. hymenaeifolium* (Michiguiste) in Costa Rica (10°00'50"N 84°07'53"W). We sampled four plant nodules, surface-sterilized, macerated, and serially diluted them (10⁻¹–10⁻²) in Peptone Yeast Extract (PY) broth. Aliquots (50 µL) were plated on PY agar supplemented with cycloheximide (40 mg/L) and incubated at 28°C for up to 3 weeks. Colonies were repeatedly streaked on PY agar and preserved in PY broth with 20% glycerol at –80°C. A loopful of the bacterial strain was then collected from the solid culture, resuspended in 250 µL PBS, and genomic DNA was extracted using the DNEasy PowerSoil Kit (Qiagen, USA) according to the manufacturer's protocol.

Libraries were prepared using the Illumina DNA Prep kit, following the manufacturer's protocol. Genomic DNA was randomly sheared, end repaired, A-tailed, and ligated with Illumina adapters. Adapter-ligated fragments were PCR-amplified, size-selected, purified, and assessed using Qubit, real-time PCR, and Bioanalyzer. Quantified libraries were pooled, and paired-end whole-genome sequencing (150 bp reads) was performed on the Illumina NovaSeq 6000 platform at Novogene Co. (CA, USA).

Sequencing produced 13,714,316 raw reads that were assessed for quality using fastp v0.20.1 (4), with average quality threshold of 30 and assembled using SPAdes v3.15.4 (5) with k-mer values of 21, 33, 55, and 77. Functional annotation was performed using PGAP v6.10 (6). Default parameters were used for all software except where otherwise noted.

We performed phylogenomic analysis employing two strains from each species in *Methylobacterium* clade C1 (7) as reference, and three *M. aquaticum* genomes available in the NCBI database (Fig. 1) to confirm taxonomy. We used an in-house pipeline that selects all the single-copy core genes (SCCGs) in the pangenome of the data set applying a MCL inflation value of 2 (8). The SCCGs were aligned with MAFFT v7.397 (9) and used to generate a phylogenomic tree with FastTree v2.1.10 employing the General Time

Editor J. Cameron Thrash, University of Southern California, Los Angeles, California, USA

Address correspondence to Keilor Rojas-Jimenez, keilor.rojas@ucr.ac.cr, or Bradd Mendoza-Guido, bradd.mendoza@ucr.ac.cr.

The authors declare no conflict of interest.

Received 7 July 2025

Accepted 22 September 2025

Published 13 October 2025

Copyright © 2025 Castro-Camacho et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

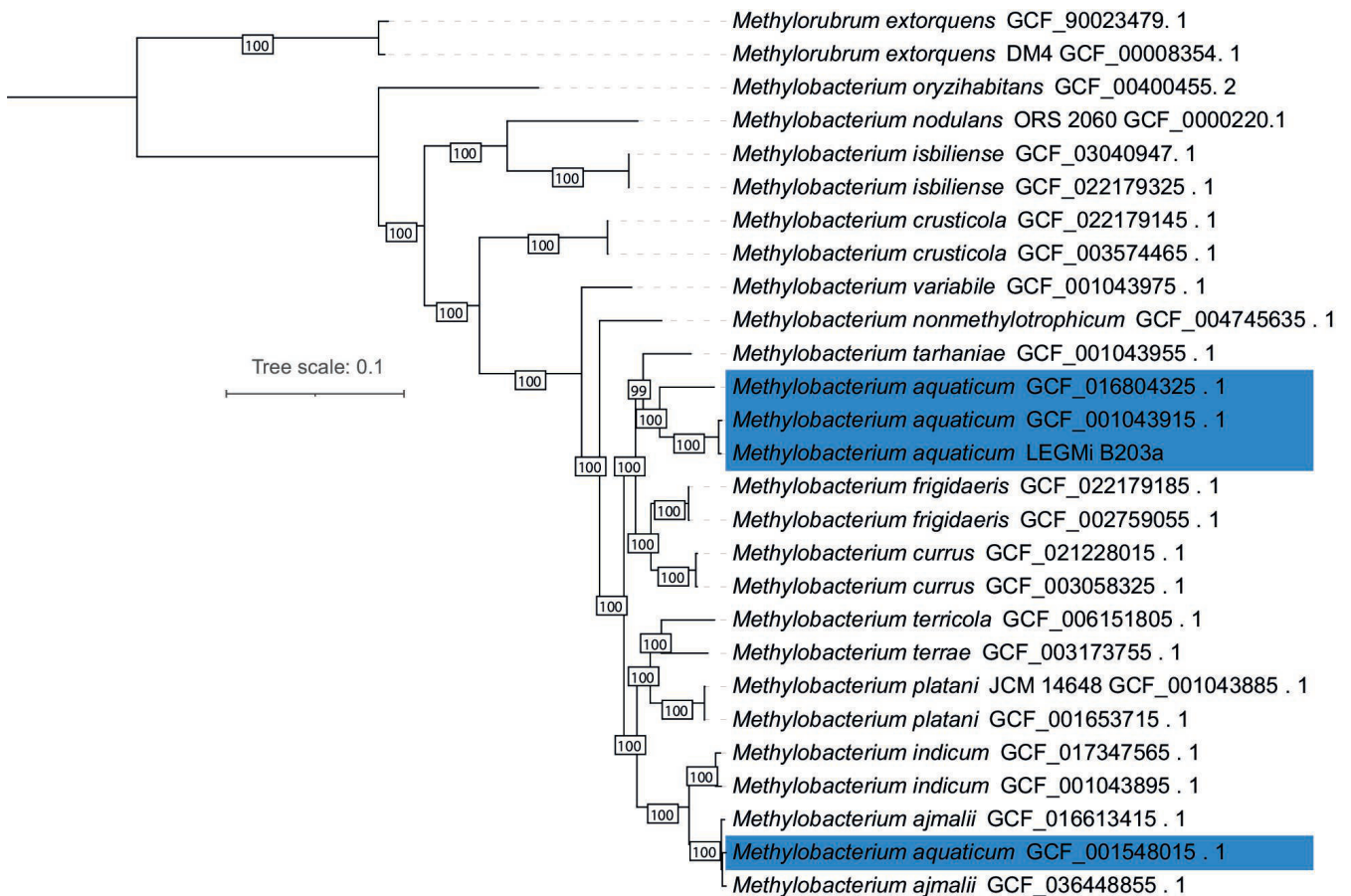


FIG 1 Phylogenomic tree of 27 genomes closely related to *M. aquaticum*, based on clade C1. The tree was constructed with FastTree v2.1.11 (10) using amino acid sequences and 891 single-copy core genes. Visualization and edition were conducted in iTOL v6 (13). Bootstrap values are shown in the nodes.

Reversible model and 1,000 bootstrap resamplings (10). Average nucleotide identity (ANI) was evaluated with pyANI (11) and digital DNA-DNA hybridization (dDDH) with the TYGS web server (12).

The *M. aquaticum* LEGMi-203a strain genome is 7,778,826 bp long with 7,913 putative coding sequences. The assembly resulted in 151 contigs, an N50 value of 150114, L50 of 17, a GC content of 69.8%, and an average coverage of 237.511. PGAP identified genes linked to nitrogen fixation and nodulation, including *nifU*-like domains, the transcriptional regulator *FixJ*, and the N-related dehydratase *MaoC*. Our isolate appeared related to the *M. aquaticum* type strain (DSM16371¹) (ANI 98.9%, dDDH 83.3%) (Fig. 1).

ACKNOWLEDGMENTS

This work was conducted as part of the University of Costa Rica's School of Biology Prokaryotic Genomics course.

This study was funded by the University of Costa Rica (projects B9-204 and C0-524) and by the Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones de Costa Rica (MICITT)-Consejo Nacional para Investigaciones Científicas y Tecnológicas (CONICIT) (project FI-036B-19).

We acknowledge the Centro Nacional de Alta Tecnología (CNCA-CENAT) for providing access to the computing cluster where all bioinformatic analyses were performed.

AUTHOR AFFILIATIONS

¹Escuela de Biología, Universidad de Costa Rica, San José, Costa Rica

²Instituto de Investigaciones en Salud, Universidad de Costa Rica, San José, Costa Rica

AUTHOR ORCID*s*

Keilor Rojas-Jimenez  <http://orcid.org/0000-0003-4261-0010>

Bradd Mendoza-Guido  <http://orcid.org/0009-0005-3214-4599>

AUTHOR CONTRIBUTIONS

Valeria Castro-Camacho, Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft | Rubí Robles-Azor, Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft | Luciana Rodríguez-Burdock, Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft | Keilor Rojas-Jimenez, Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – review and editing | Bradd Mendoza-Guido, Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

This whole-genome sequence project has been deposited in GenBank under the accession no. [ASM5040874v1](https://doi.org/10.1093/asm/asm5040874v1). The raw sequencing reads were deposited under the accession number [SRX26808260](https://doi.org/10.1093/asm/srx26808260).

REFERENCES

- Gallego V, García MT, Ventosa A. 2005. *Methylobacterium hispanicum* sp. nov. and *Methylobacterium aquaticum* sp. nov., isolated from drinking water. *Int J Syst Evol Microbiol* 55:281–287. <https://doi.org/10.1099/ijs.0.63319-0>
- Sy A, Giraud E, Jourand P, Garcia N, Willems A, de Lajudie P, Prin Y, Neyra M, Gillis M, Boivin-Masson C, Dreyfus B. 2001. Methylophilic *Methylobacterium* bacteria nodulate and fix nitrogen in symbiosis with legumes. *J Bacteriol* 183:214–220. <https://doi.org/10.1128/JB.183.1.214-220.2001>
- Jourand P, Renier A, Rapior S, Miana de Faria S, Prin Y, Galiana A, Giraud E, Dreyfus B. 2005. Role of methylophilicity during symbiosis between *Methylobacterium nodulans* and *Crotalaria podocarpa*. *Mol Plant Microbe Interact* 18:1061–1068. <https://doi.org/10.1094/MPMI-18-1061>
- Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- 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* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>
- Alessa O, Ogura Y, Fujitani Y, Takami H, Hayashi T, Sahin N, Tani A. 2021. Comprehensive comparative genomics and phenotyping of *Methylobacterium* species. *Front Microbiol* 12:740610. <https://doi.org/10.3389/fmicb.2021.740610>
- Mendoza-Guido B, Rivera-Montero L, Barrantes K, Chacon L. 2025. Plasmid and integron-associated antibiotic resistance in *Escherichia coli* isolated from domestic wastewater treatment plants. *FEMS Microbiol Lett* 372:fnaf041. <https://doi.org/10.1093/femsle/fnaf041>
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>
- Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods* 8:12–24. <https://doi.org/10.1039/C5AY02550H>
- Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M. 2022. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* 50:D801–D807. <https://doi.org/10.1093/nar/gkab902>
- Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49:W293–W296. <https://doi.org/10.1093/nar/gkab301>