

Article title

Draft Genome of *Rhizobium oryzzhabitans* associated with yellow green alga *Vaucheria bursata*

Authors

Gerardo Laureano¹, Vignesh Lal², Laurel Mitchell², Elizabeth Santillan Olea², Joseph Tovar², Ashley Scoles² & Alok Arun^{2#}

Affiliations

¹Inter American University of Puerto Rico, Barranquitas, 00794-517, Puerto Rico, USA

²Department of Biological Sciences, California State University Stanislaus, Turlock, California, USA

Running title

Rhizobium oryzzhabitans MAG from *Vaucheria bursata*.

Corresponding author's email address

aarun1@csustan.edu

Abstract

We present a draft metagenome-assembled genome (MAG) of a *Rhizobium* species derived from the sequenced genome of *Vaucheria bursata*. The MAG is 89.8% complete, with 4,289 ORFs and 2,002 annotated genes. Investigating its function will enhance understanding of bacteria-algae interactions and their role within the *Vaucheria* microbiome.

Announcement

The exploration of metagenomes-assembled genomes (MAGs) provides a robust approach to identifying new taxa from uncultivable prokaryotes. Macroalgae are known to play essential roles in shaping microbial communities through their biological and

chemical interactions (1). Also, the *Rhizobium* family has been showed to have endophytic relation with rice species (2). This parallel suggests potential common mechanisms in how both macroalgae and higher plants influence and interact with beneficial microbial partners. *Rhizobium*, a genus of nitrogen-fixing bacteria, is well studied for its symbiotic interactions with legumes and its potential to enhance plant growth and soil fertility. In our current study, a novel metagenome-assembled genome (MAG) of *Rhizobium oryzae* was obtained and assembled from the genomic DNA sequencing data of the freshwater alga *Vaucheria bursata*. Our study supports an earlier finding suggesting the presence of endophytic bacteria in *Vaucheria sp.* (3 & 4).

Cultures of *Vaucheria bursata* vegetative filaments were obtained from the UTEX culture collection of algae at UT-Austin (Austin, USA). They were grown in 3N Bold Modified medium under a photo cycle of 16L:8D at 25°C. Total genomic DNA was extracted from filaments using the commercial kit NucleoSpin Plant II from Macherey-Nagel, according to the manufacturer's instructions. A standard 350 base pair (bp) DNA library was prepared, and the Illumina NovaSeq sequencing platform was used to sequence the genome using paired-end reads (2x150bp). A total of 167,750,552 reads providing a sequencing coverage of 90X were obtained. All software was used in their default parameter unless otherwise specified. Low-quality read trimming and quality verification were performed using Fastp v0.23.4 (5). MetaWrap v1.3 pipeline was applied to construct the MAG (6). The MetaWrap classify module, GTDb-tk version 2.4.0, MIGA analysis all showed the MAG to have a strong similarity *Rhizobium oryzae* (7 & 8). For GTDb-tk, it showed to have similarity with *Agrobacterium oryzae*, which was reclassified to *Rhizobium oryzae*. All taxonomic classification analyses were able to provide identification at a species level. The quality of the MAG was assessed by CheckM2, with a completeness of 89.8% and contamination of 0.78% (9). BUSCO analysis was also performed using the lineage Rhizobium/Agrobacterium Group resulting in a genome completeness of 80.7% (i.e., single copy = 80.69%, duplicates = 0.1%) (10).

58 The *Rhizobium oryzae* MAG showed a total genome length of 3,461,440 bp
 59 (N50=388,377 bp), distributed across 162 contigs, with a GC content of 59.40% Gene
 60 prediction and annotation were performed using Prokka in the European Galaxy web
 61 server platform (11). Using InterProScan v5.59, eggNOG Mapper v2.18, KofamScan
 62 v1.30, functional annotation was performed (12-14). A total of 4,289 ORFs were
 63 predicted, of which 3,147 were annotated and the remaining were hypothetical proteins
 64 (Table 1).

65

66 Table 1. Genome features of the *Rhizobium oryzae* MAG.

Parameters	Statistics
Estimated genome size (Mb)	3.88
Number of contigs	163
N50 (bp)	388,377
GC (%)	59.40
Number of ORFs	4,289
Annotated genes	2,002

67

To determine the potential secondary metabolites AntiSmash version 8 was used (15). Secondary metabolites as terpenes and a high confidence score with agrobactin were determined. Agrobactin is known as siderophore and iron transport molecule (16). Based on functional annotation the *Rhizobium oryzae* MAG, has a high probability to be facultatively anaerobic heterotroph, with complete pyruvate oxidation, and TCA cycle and partial glycolysis pathways. The MAG potentially have metabolic capacity with nitrogen pathway associated nitrate fixation. Unraveling the metabolic capacity of the *Rhizobium oryzae* MAG will provide a deeper knowledge in the symbiotic relation between the *Rhizobium* family and macroalgae. Additionally, providing a deeper understanding of the complex microbiome community of *Vaucheria bursata*.

Data availability statement

MAG was deposited at GenBank under the accession number JBNQFZ000000000. The raw sequencing reads of the MAG are available with the accession number SRR32638075 and were deposited in the Bio Project PRJNA1020787.

Acknowledgments

This study was funded by the National Science Foundation NSF-HSI STEM grant (Award No. 1928792) and a Seed Grant from California State University awarded to Alok Arun. We thank Mr. Ramon Rivera Vicens (UIPR) for his insightful feedback on the data. This Course-Based Undergraduate Research Experience (CURE) was part of the BIOL 4840 course at CSU Stanislaus.

References

1. Lavecchia, A., Fosso, B., Engelen, A. H., Borin, S., Manzari, C., Picardi, E., ... & Placido, A. (2024). Macroalgal microbiomes unveil a valuable genetic resource for halogen metabolism. *Microbiome*, 12(1), 47.
2. Zhao, J., Zhao, X., Wang, J., Gong, Q., Zhang, X., & Zhang, G. (2020). Isolation, identification and characterization of endophytic bacterium *Rhizobium oryzae* sp. nov., from rice root with biotechnological potential in agriculture. *Microorganisms*, 8(4), 608.

3. Laureano, G., Bermudez, A. E., Rivera Vicéns, R. E., & Arun, A. (2024). Metagenome-assembled genome of *Pseudomonas* sp. from coenocytic alga *Vaucheria bursata*. *Microbiology Resource Announcements*, 13(11), e00485-24.
4. Ott, D. W. (1979). Endosymbiotic bacteria in *Vaucheria* (Xanthophyceae): association with cytoplasmic microtubules in *Vaucheria sessilis*. *Cytobios*, 24(95-96), 185-194.
5. Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17), i884-i890.
6. Uritskiy, G. V., DiRuggiero, J., & Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome*, 6(1), 1-13.
7. Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2020). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database.
8. Rodriguez-R, L. M., Gunturu, S., Harvey, W. T., Rosselló-Mora, R., Tiedje, J. M., Cole, J. R., & Konstantinidis, K. T. (2018). The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. *Nucleic acids research*, 46(W1), W282-W288.
9. Chklovski, A., Parks, D. H., Woodcroft, B. J., & Tyson, G. W. (2023). CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nature methods*, 20(8), 1203-1212.
10. Seppey, M., Manni, M., & Zdobnov, E. M. (2019). BUSCO: assessing genome assembly and annotation completeness. *Gene prediction: methods and protocols*, 227-245.
11. Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), 2068-2069.
12. Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., ... & Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 30(9), 1236-1240.
13. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). eggNOG-mapper v2: functional annotation, orthology assignments, and

- domain prediction at the metagenomic scale. *Molecular biology and evolution*, 38(12), 5825-5829.
14. Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., & Ogata, H. (2020). KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics*, 36(7), 2251-2252.
15. Blin, K., Shaw, S., Augustijn, H. E., Reitz, Z. L., Biermann, F., Alanjary, M., ... & Weber, T. (2023). antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic acids research*, 51(W1), W46-W50.
16. Ong, S. A., Peterson, T., & Neilands, J. B. (1979). Agrobactin, a siderophore from *Agrobacterium tumefaciens*. *Journal of Biological Chemistry*, 254(6), 1860-1865.