# Complete Genome of a *Methanosarcina mazei* Strain Isolated from Sediment Samples from an Amazonian Flooded Area

Diego Assis das Graças,a Rommel Thiago Jucá Ramos,a Ana Carolina Vieira Araújo,b Ramiro Zahlouth,a Adriana Ribeiro Carneiro,a Thiago Souza Lopes,a Rafael Azevedo Baraúna,a Vasco Azevedo,c Maria Paula Cruz Schneider,a Vivian Helena Pellizari,b Artur Silvaa

Institute of Biological Sciences, Federal University of Pará, Belém, PA, Brazila; Institute of Oceanography, University of São Paulo, São Paulo, Brazilb; Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, MG, Brazilc

***Methanosarcina mazei* is a strictly anaerobic methanogen from the *Methanosarcinales* order, which is known for its broad catabolic range among methanogens and is widespread throughout diverse environments. The draft genome of the strain presented here was cultivated from sediment samples collected from the Tucuruí hydroelectric power station reservoir.**

Address correspondence to Artur Silva, asilva@ufpa.br.

Methane is the second most important greenhouse gas, and approximately 80% is produced by biogenic sources (1). The only organisms able to produce methane are methanogenic archaea (2). Methanogens are found in several environments, such as rumen, rice fields, sea and freshwater sediments, and flooded areas (2). Although the methanogenesis pathway is well known, some aspects of the core genome, genome evolution, and shared genes are still unclear.

*Methanosarcina mazei* is a methanogenic archaeon capable of metabolizing several substrates, including $CO_2$, acetate, and methylated compounds (2). This species is found in several environments and was detected in abundance in the reservoir of the fourth largest hydroelectric dam in the world (3). To characterize the genome of this species, a sediment sample from the Tucuruí hydropower station reservoir was inoculated in mineral media (4) supplemented with acetate and methanol and was maintained in an $H_2$:$CO_2$ (80:20) atmosphere to enrich and cultivate *M. mazei*. The enrichment was conducted at 30°C under standard anaerobic conditions (5). After several molecular and cellular analyses, such as transmission electron microscopy, denaturing high-performance liquid chromatography, 16S rRNA gene sequencing, and fluorescence *in situ* hybridization, total DNA was extracted from a nonpure culture of *M. mazei*, amplified using phi29 DNA polymerase (BioLabs), and used as a source template for genome sequencing.

The genome was first sequenced using a SOLiD System V3 with a mate-paired library, which yielded 24,405,103 and 24,399,268 reads (50 bp) for the R3 and F3 tags, respectively. A second round of sequencing was performed using the SOLiD 5500 XL platform with a mate-paired library, resulting in a total of 113,588,848 reads (60 bp) for each tag (F3 and R3). All reads were filtered by Quality Assessment software (6), whereby reads with an average quality score below Phred 20 were removed.

The reads were assembled by Velvet (7) and Edena (8), and the redundant sequences were removed by Simplifier (9), resulting in a total of 16,811 contigs. The scaffolds were produced by mapping these contigs against the *M. mazei* GO1 (AE008384) strain using BLAST (BLASTn) software. We used the Graphical contig Analyzer for All Sequencing Platforms software (G4ALL) to manually curate and generate the genome scaffold with gaps. Many of these gaps were closed using CLC Genomics Workbench software after recursive mapping of the contigs against the draft genome (10, 11). The complete genome sequence of *M. mazei* TUC01 contained 3,427,949 bp with a GC content of 42.5%. The genome was annotated by RAST (12), and 3,252 coding DNA sequences (CDS) were predicted.

This is the first archaeal genome from a Brazilian environment, and it provides biological information that will be helpful for more in-depth analyses and for obtaining more insight into the ecology and genome evolution of the *Methanosarcinales* order.

**Nucleotide sequence accession number.** The genome sequence obtained in this study has been deposited in the GenBank database under accession number CP004144.1.

## REFERENCES

1. **Wuebles DJ, Hayhoe K.** 2002. Atmospheric methane and global change. Earth Sci. Rev. **57**:177–210.
2. **Garcia J-L, Patel BKC, Ollivier B.** 2000. Taxonomic, phylogenetic, and Ecological diversity of methanogenic *Archaea*. Anaerobe **6**:205–226.
3. **Graças DA, Miranda PRO, Baraúna RA, McCulloch JÁ, Ghilardi R, Jr, Schneider MPC, Silva A.** 2011. Microbial diversity of an anoxic zone of a

hydroelectric power station reservoir in Brazilian Amazonia. Microb. Ecol. **62**:853–861.

4. **Zinder SH, Koch M.** 1984. Non-aceticlastic methanogenesis from acetate: acetate oxidation by a thermophilic syntrophic coculture. Arch. Microbiol. **138**:263–272.

5. **Wolfe RS.** 2011. Techniques for cultivating methanogens. Methods Enzymol. **494**:1–22.

6. **Ramos RT, Carneiro AR, Baumbach J, Azevedo VAC, Schneider MPC, Silva A.** 2011. Analysis of quality raw data of second generation sequencers with quality assessment software. BMC Res. Notes **4**:130.

7. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.

8. **Hernandez D, François P, Farinelli L, Osterås M, Schrenzel J.** 2008. De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. Genome Res. **18**:802–809.

9. **Ramos RTJ, Carneiro AR, Azevedo V, Schneider MP, Barh D, Silva A.** 2012. Simplifier: a web tool to eliminate redundant NGS contigs. Bioinformation **8**:996–999.

10. **Cerdeira LT, Carneiro AR, Ramos RTJ, de Almeida SS, D'Afonseca V, Schneider MPC, Baumbach J, Tauch A, McCulloch JA, Azevedo VAC, Silva A.** 2011. Rapid hybrid *de novo* assembly of a microbial genome using only short reads: Corynebacterium pseudotuberculosis I19 as a case study. J. Microbiol. Methods **86**:218–223.

11. **Tsai IJ, Otto TD, Berriman M.** 2010. Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. Genome Biol. **11**:R41.

12. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9**:75.