

Genome Sequence of *Magnetospirillum magnetotacticum* Strain MS-1

Matthew D. Smalley, Georgi K. Marinov, L. Elizabeth Bertani, Gilberto DeSalvo

Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California, USA

Here, we report the genome sequence of *Magnetospirillum magnetotacticum* strain MS-1, which consists of 36 contigs and 4,136 protein-coding genes.

Received 18 February 2015 Accepted 23 February 2015 Published 2 April 2015

Citation Smalley MD, Marinov GK, Bertani LE, DeSalvo G. 2015. Genome sequence of *Magnetospirillum magnetotacticum* strain MS-1. *Genome Announc* 3(2):e00233-15. doi:10.1128/genomeA.00233-15

Copyright © 2015 Smalley et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Gilberto DeSalvo, gilberto@caltech.edu.

Magnetospirillum magnetotacticum strain MS-1 (1) is an alphaproteobacterial representative of magnetotactic bacteria. Magnetotactic bacteria are a unique group of prokaryotes characterized by their ability to orient themselves along the lines of a magnetic field (magnetotaxis), conferred by the presence of specialized organelles called magnetosomes (2, 3). Magnetosomes are a rare example of a lipid-bounded intracellular organelle in prokaryotes and are of great interest both from a purely biological perspective and because of their numerous potential technological applications (4, 5).

An assembly of the MS-1 genome was previously deposited in GenBank (assembly ID GCA_000166875.1). However, the annotation of that assembly revealed that it is heavily contaminated with sequences from an unidentified *Methylobacter* strain. It contains two sets of rRNAs, one of which is a best match to *Methylobacterium* rRNAs, and >60% of the predicted open reading frames (ORFs) also had a best match in the genome of one of the sequenced *Methylobacterium* strains. We therefore carried out *de novo* whole-genome sequencing of the MS-1 strain. We generated ~1,000× coverage of the MS-1 genome using 2 × 150-bp Illumina HiSeq reads, with an average fragment length of ~700 bp. The reads were error corrected using BayesHammer (6) and assembled using SPAdes (version 3.1.1) (7), IDBA (version 1.1.1) (8), and Edena (version 3.131028) (9). Contigs with a coverage of ≤200× and length ≤200 bp were filtered out, and the assemblies were merged with CISA (version 1.3) (10). After filtering out potential contaminants (i.e., contigs that were close matches to the human and other eukaryotic genome sequences), the final assembly consists of 4,523,935 bp in 36 contigs, with an N_{50} of 541,957 and N_{90} of 102,810. The largest contig is 1,344,907 bp in length, and the G+C content is 63.5%.

The annotation of protein-coding genes was carried out using RAST (11). Ribosomal RNAs were annotated using RNAmmer 1.2 (12), and tRNAscan-SE (version 1.21) (13) was used to find tRNA genes. Additional noncoding RNA genes and elements were annotated using Infernal (version 1.1.1) (14) and the Rfam database (version 12.0) (15).

A total of 4,136 protein-coding genes, 48 tRNAs, and a single set of rRNAs were predicted. The vast majority of the genes have their best BLAST match in the genomes of related magnetotactic bacteria, such as *Magnetospirillum* sp. strain AMB-1, indicating the absence of contamination in the assembly.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in GenBank under the accession no. JXSL00000000. The version described in this paper is the first version, JXSL01000000. The raw sequencing reads are available from the Short Read Archive under the accession no. SRR1765663.

ACKNOWLEDGMENTS

This work was supported by the Millard and Muriel Jacobs Genetics and Genomics Laboratory and the Division of Biology and Biological Engineering, located at the California Institute of Technology, Pasadena, CA.

REFERENCES

- Maratea D, Blakemore RP. 1981. *Aquaspirillum magnetotacticum* sp. nov., a magnetic *Spirillum*. *Int J Syst Bacteriol* 31:452–455. <http://dx.doi.org/10.1099/00207713-31-4-452>.
- Blakemore R. 1975. Magnetotactic bacteria. *Science* 190:377–379. <http://dx.doi.org/10.1126/science.170679>.
- Blakemore RP. 1982. Magnetotactic bacteria. *Annu Rev Microbiol* 36:217–238. <http://dx.doi.org/10.1146/annurev.mi.36.100182.001245>.
- Alphandéry E. 2014. Applications of magnetosomes synthesized by magnetotactic bacteria and magnetosomes in medicine: a mini-review. *Front Bioeng Biotechnol* 2:5. <http://dx.doi.org/10.3389/fbioe.2014.00005>.
- Lefèvre CT, Wu LF. 2013. Evolution of the bacterial organelle responsible for magnetotaxis. *Trends Microbiol* 21:534–543. <http://dx.doi.org/10.1016/j.tim.2013.07.005>.
- Nikolenko SI, Korobeynikov AI, Alekseyev MA. 2013. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* 14(Suppl 1):S7. <http://dx.doi.org/10.1186/1471-2164-14-S1-S7>.
- Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Prjibelsky A, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, McLean J, Lasken R, Clingenpeel SR, Woyke T, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling genomes and mini-metagenomes from highly chimeric reads, p 158–170. *In* Deng M, Jiang R, Sun F, Zhang X (ed), *Research in computational molecular biology*, vol 7821. Springer, Heidelberg, Germany.
- Peng Y, Leung HCM, Yiu SM, Chin FYL. 2010. IDBA—a practical iterative de Bruijn graph *de novo* assembler, p 426–440. *In* Berger B (ed), *Research in computational molecular biology*, vol 6044. Springer, Heidelberg, Germany.
- Hernandez D, François P, Farinelli L, Østerås M, Schrenzel J. 2008. *De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res* 18:802–809. <http://dx.doi.org/10.1101/gr.072033.107>.
- Lin S-H, Liao Y-C. 2013. CISA: contig integrator for sequence assembly of bacterial genomes. *PLoS One* 8:e60843. <http://dx.doi.org/10.1371/journal.pone.0060843>.

11. 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, Parrello 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. <http://dx.doi.org/10.1186/1471-2164-9-75>.
12. Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS Web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* 33:W686–W689. <http://dx.doi.org/10.1093/nar/gki366>.
13. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. <http://dx.doi.org/10.1093/nar/gkm160>.
14. Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29:2933–2935. <http://dx.doi.org/10.1093/bioinformatics/btt509>.
15. Burge SW, Daub J, Eberhardt R, Tate J, Barquist L, Nawrocki EP, Eddy SR, Gardner PP, Bateman A. 2013. Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res* 41:D226–D232. <http://dx.doi.org/10.1093/nar/gks1005>.