

Draft Whole-Genome Sequence of Urease-Producing *Sporosarcina koreensis*

Sara Abdul Majid,^a Michael F. Graw,^{b*} Hanh Nguyen,^{b*} Anthony G. Hay^b

Department of Research, Weill Cornell Medical College in Qatar, Doha, Qatar^a; Department of Microbiology, Cornell University, Ithaca, New York, USA^b

* Present address: Michael F. Graw, College of Earth, Ocean, and Atmospheric Science, Oregon State University, Corvallis, Oregon, USA; Hanh Nguyen, Department of Soil and Water Science, University of Florida, Gainesville, Florida, USA.

Urease-producing microbes are of significance due to their potential application in biocement production. *Sporosarcina koreensis* Q1 is a urease-producing bacterium belonging to the phylum *Firmicutes*. Here, we present the draft whole-genome sequence of *S. koreensis* Q1, isolated from a barchan sand dune in Qatar.

Received 28 January 2016 Accepted 3 February 2016 Published 17 March 2016

Citation Abdul Majid S, Graw MF, Nguyen H, Hay AG. 2016. Draft whole-genome sequence of urease-producing *Sporosarcina koreensis*. *Genome Announc* 4(2):e00096-16. doi:10.1128/genomeA.00096-16.

Copyright © 2016 Abdul Majid 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/).

Address correspondence to Anthony G. Hay, agh5@cornell.edu.

Urease-producing microbes have been extensively studied for their potential application in biocement production for soil stabilization (1), aquaculture (2), and engineering works (3, 4). Biocementation is a technology relying on microbiologically induced calcite precipitation (MICP) via pH changes brought on by the hydrolysis of urea and release of ammonium by urease-producing microorganisms. *Sporosarcina* spp. are heterotrophic, Gram-positive bacteria that belong to the *Planococcaceae* family, *Bacilli* order, of the phylum *Firmicutes*. There are 16 species identified to date (<http://www.bacterio.net>) (5). A number of these species have been documented as urease-producers including *S. ginsengisoli* (6) and *S. pasteurii* (7). A urease-producing strain of *Sporosarcina pasteurii* was previously sequenced and published in *Genome Announcements* (8), but the sequence was removed from NCBI due to contamination and is no longer accessible (<http://www.ncbi.nlm.nih.gov/nucore/AYOX00000000.1>).

Strain Q1 was isolated from the sand of a barchan dune located south of Doha city in the State of Qatar (25.009450°, 51.340490°) in April of 2013. The sand was used as an inoculum in nitrogen free minimal medium supplemented with glucose and urea as sole carbon and nitrogen sources, respectively. The strain produced urease activity on urease indicator medium and was able to form water-stable aggregates (>0.5 mm) when mixed with fine sand in the presence of calcium and urea.

Initial characterization was done via partial 16S rRNA gene sequencing; the strain was found to be 98% identical to *Sporosarcina koreensis* strain F73. Library construction was performed using the Nextera XT kit (Illumina, USA). After library construction, the sample was spiked with 50% PhiX and sequenced on an Illumina HiSeq2500 instrument using 250 bp reads according to the manufacturer's protocol at the Center of Biotechnology in Cornell University, Ithaca, New York.

The genome was assembled with Velvet (9) using the Velvet Optimizer add-on. The final assembly used a k-mer length of 157 and a coverage cutoff of 5.53. The assembly produced 4,320,341 bp of sequence across 426 contigs with an N_{50} of

72,066 bp, a longest sequence of 212,353 bp, and a G + C content of 44.5%. A total of 51 putative rRNA genes were identified using the BLASTn_rRNA script within WebMGA (10); 10 of these resulted in highly significant (e-values < 10^{-100}), high identity (99% similarity) alignments against *Sporosarcina luteola* (NR114283.1), *saromensis* (NR114249.1), and *koreensis* (NR043526.1) partial 16S rRNA genes within the NCBI 16S rRNA database. The draft genome was annotated using the Rapid Annotations using Subsystem Technology (RAST) v2.0. RAST identified 4,413 coding sequences spanning 439 subsystems, dominated by the following features: amino acids and derivatives (14%), carbohydrates (9%), and cofactors, vitamins, prosthetic groups, and pigments (9%). One of the subsystems includes the seven urease genes (3 structural *UreaA*, *UreaB*, *UreaC*, and 4 accessory *UreaD*, *UreaE*, *UreaF*, *UreaG*). We anticipate that this genome sequence will be of value to those studying MICP and/or related *Sporosarcina* species.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession no. [LQYA00000000](https://www.ncbi.nlm.nih.gov/nuccore/LQYA00000000). The version described in this paper is the first version, [LQYA01000000](https://www.ncbi.nlm.nih.gov/nuccore/LQYA01000000).

FUNDING INFORMATION

This work, including the efforts of Anthony G. Hay, was funded by Qatar National Research Fund (QNRF) (6-059-2-023).

This publication was made possible by National Priorities Research Program grant 6-059-2-023 from The Qatar National Research Fund. The statements made herein are solely the responsibility of the authors.

REFERENCES

1. Stabnikov V, Naeimi M, Ivanov V, Chu J. 2011. Formation of water-impermeable crust on sand surface using biocement. *Cem Concr Res* 41: 1143–1149. <http://dx.doi.org/10.1016/j.cemconres.2011.06.017>.
2. Chu J, Ivanov V, Stabnikov V, Li B. 2013. Microbial method for construction of an aquaculture pond in sand. *Géotechnique* 63:871–875. <http://dx.doi.org/10.1680/geot.SIP13.P.007>.
3. Van Tittelboom K, De Belie N, De Muynck W, Verstraete W. 2010. Use

- of bacteria to repair cracks in concrete. *Cem Concr Res* **40**:157–166. <http://dx.doi.org/10.1016/j.cemconres.2009.08.025>.
4. De Muynck W, De Belie N, Verstraete W. 2010. Microbial carbonate precipitation in construction materials: a review. *Ecol Eng* **36**:118–136. <http://dx.doi.org/10.1016/j.ecoleng.2009.02.006>.
 5. Euzéby JP. 1997. List of bacterial names with standing in nomenclature: a folder available on the internet. *Int J Syst Bacteriol* **47**:590–592. <http://dx.doi.org/10.1099/00207713-47-2-590>.
 6. Achal V, Pan X, Fu Q, Zhang D. 2012. Biomineralization based remediation of As(III) contaminated soil by *Sporosarcina ginsengisoli*. *J Hazard Mater* **201–202**:178–184. <http://dx.doi.org/10.1016/j.jhazmat.2011.11.067>.
 7. Ahmed Zoheir FT. 2013. Urease activity and induction of calcium carbonate precipitation by *Sporosarcina pasteurii* NCIMB 8841. *J Appl Sci Res* **9**:1525–1533.
 8. Tiwari PK, Joshi K, Rehman R, Bhardwaj V, Shamsudheen KV, Sivasubbu S, Scaria V. 2014. Draft genome sequence of urease-producing *Sporosarcina pasteurii* with Potential Application in Biocement production. *Genome Announc* **2**(1):. <http://dx.doi.org/10.1128/genomeA.01256-13>.
 9. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
 10. Wu S, Zhu Z, Fu L, Niu B, Li W. 2011. WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* **12**:444. <http://dx.doi.org/10.1186/1471-2164-12-444>.