Draft Genome Sequence of *Shewanella decolorationis* S12, a Dye-Degrading Bacterium Isolated from a Wastewater Treatment Plant

Meiying Xu,[a,d]* Yun Fang,[a,b,c,d] Jun Liu,[e] Xingjuan Chen,[a,d] Guoping Sun,[a,d] Jun Guo,[a,d] Zhengshuang Hua,[e] Qichao Tu,[f] Liyou Wu,[f] Jizhong Zhou,[f] Xueduan Liu[b,c]

Guangdong Institute of Microbiology, Guangzhou, Guangdong, China[a]; School of Minerals Processing and Bioengineering, Central South University, Changsha, Hunan, China[b]; Key Laboratory of Biometallurgy of Ministry of Education, Changsha, Hunan, China[c]; State Key Laboratory Cultivation Base, Province & Ministry Co-constructed South China Applied Microbiology Laboratory, Guangzhou, Guangdong, China[d]; State Key Laboratory of Biocontrol and Guangdong Provincial Key Laboratory of Plant Resource, School of Life Sciences, Sun Yat-Sen University, Guangzhou, China[e]; Institute for Environmental Genomics, and Department of Botany and Microbiology, Stephenson Research & Technology Center, University of Oklahoma, Norman, Oklahoma, USA[f]

* Present address: Meiying Xu, Guangdong Institute of Microbiology, Guangzhou, Guangdong, China.

M.X. and Y.F. contributed equally to this work.

Shewanella decolorationis is a valuable microorganism for degrading diverse synthetic textile dyes. Here, we present an annotated draft genome sequence of *S. decolorationis* S12, which contains 4,219 protein-coding genes and 86 structural RNAs. This information regarding the genetic basis of this bacterium can greatly advance our understanding of the physiology of this species.

Address correspondence to Meiying Xu, xumy@gdim.cn.

The genus *Shewanella* is famous for bioremediation of environments contaminated with toxic organic and inorganic chemicals and has been found in water, sediment, and surfaces of shale and sandstone (1). Recent studies have demonstrated that *Shewanella decolorationis* is a powerful species to degrade diverse manufactured textile dyes, including azo and anthraquinone dye, which are carcinogenic to humans (2–4). *S. decolorationis* S12 was isolated in 2002 from activated sludge of a textile-printing wastewater treatment plant in Guangzhou, China (3). Whole-genome sequences of *S. decolorationis* have not been available, which is a limitation to the genetic study of this useful species.

Genomic DNA extracted from *S. decolorationis* S12 was sequenced on an Illumina Miseq sequencer according to Illumina's recommendations. A total of 16,692,489 paired-end 250-base reads were generated. The draft genome was assembled using Velvet (5) and Newbler, containing 77 contigs of >250 bp each after manual curation, with a total of 4,843,449 bases, an N_{50} of 155.5 kbp, and a G+C content of 47.1%. The draft genome sequence of strain S12 was annotated using Blast2GO (6) and the NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP) (7, 8).

The draft genome sequence contains 4,305 genes, including 4,219 predicted coding sequences (CDS), 9 rRNA genes, and 77 tRNA genes. A total of 3,184 CDS were classified using the Cluster of Orthologous Groups (COG) database (9). The most abundant groups of proteins were those that are involved in signal transduction (COG category T [9.5%]), amino acid metabolism and transport (COG category E [9.0%]), transcription (COG category K [8.5%]), and energy production and conversion (COG category C [7.8%]). The genome sequence of *S. decolorationis* S12 serves as a basis for further investigation of the molecular basis of its potential in dye degradation.

Currently, the genus *Shewanella* contains 45 type species, and 28 strains from 15 type species have been sequenced for the whole genomes. Unlike these species from natural environments, *S. decolorationis* S12 was isolated from a human-altered and highly contaminated system, and thus its genetic evolution and cell responses might be different. Comparative genomics analysis showed that *S. decolorationis* S12 was closest to *Shewanella* sp. strain ANA-3 (89% coding sequence similarity), followed by *Shewanella* sp. strain MR-7 (88%) and *Shewanella* sp. strain MR-4 (88%). Further analysis will provide a better understanding of the *S. decolorationis* genome and shed more light on genomic evolution of the *Shewanella* genus.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under accession number AXZL00000000. The version described in this paper is the first version, AXZL01000000.

ACKNOWLEDGMENTS

Science and the Guangdong Province Computational Science Innovative Research Team.

REFERENCES

1. **Rodionov DA, Novichkov PS, Stavrovskaya ED, Rodionova IA, Li X, Kazanov MD, Ravcheev DA, Gerasimova AV, Kazakov AE, Kovaleva GY, Permina EA, Laikova ON, Overbeek R, Romine MF, Fredrickson JK, Arkin AP, Dubchak I, Osterman AL, Gelfand MS.** 2011. Comparative genomic reconstruction of transcriptional networks controlling central metabolism in the *Shewanella* genus. BMC Genomics **12**(Suppl 1):S3. doi: 10.1186/1471-2164-12-S1-S3.
2. **Hong Y, Xu M, Guo J, Xu Z, Chen X, Sun G.** 2007. Respiration and growth of *Shewanella decolorationis* S12 with an azo compound as the sole electron acceptor. Appl. Environ. Microbiol. **73**:64–72.
3. **Xu M, Guo J, Cen Y, Zhong X, Cao W, Sun G.** 2005. *Shewanella decolorationis* sp. nov., a dye-decolorizing bacterium isolated from activated sludge of a waste-water treatment plant. Int. J. Syst. Evol. Microbiol. **55**:363–368.
4. **Xu M, Guo J, Zeng G, Zhong X, Sun G.** 2006. Decolorization of anthraquinone dye by *Shewanella decolorationis* S12. Appl. Microbiol. Biotechnol. **71**:246–251.
5. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.
6. **Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M.** 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics **21**:3674–3676.
7. **Klimke W, Agarwala R, Badretdin A, Chetvernin S, Ciufo S, Fedorov B, Kiryutin B, O'Neill K, Resch W, Resenchuk S, Schafer S, Tolstoy I, Tatusova T.** 2009. The National Center for Biotechnology Information's protein clusters database. Nucleic Acids Res. **37**:D216–D223.
8. **Pruitt KD, Tatusova T, Klimke W, Maglott DR.** 2009. NCBI reference sequences: current status, policy and new initiatives. Nucleic Acids Res. **37**:D32–D36.
9. **Tatusov RL, Galperin MY, Natale DA, Koonin EV.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28**:33–36.