Complete Genome Sequence of *Shewanella* sp. Strain Lzh-2, an Algicidal Bacterial Strain Isolated from Lake Taihu, People's Republic of China

Zhenghua Li,[a] Feng Song,[a] Mei Chen[a,b]

[a]Shandong Provincial Key Laboratory of Biophysics, Institute of Biophysics, Dezhou University, Dezhou, People's Republic of China
[b]School of Bioengineering, Qilu University of Technology, Shandong Academy of Sciences, Jinan, People's Republic of China

ABSTRACT *Shewanella* sp. strain Lzh-2 is an algicidal bacterium isolated from surface water samples collected from Meiliang Bay of Lake Taihu in China. Here, we present the complete genome sequence of *Shewanella* sp. Lzh-2. Some functional genes and secondary metabolite gene clusters were predicted.

*S*hewanella sp. strain Lzh-2 was isolated from surface water samples collected from Meiliang Bay of Lake Taihu in China (1). Strain Lzh-2 exhibited significant algicidal activity toward several cyanobacterial species (1). Two algicidal substances (hexahydro-pyrrolo[1,2-*a*]pyrazine-1,4-dione and 2,3-indolinedione) have been purified from the culture of strain Lzh-2, and they both possessed strong algicidal activity against *Microcystis aeruginosa*, which was reported to be one of the most dominant microorganisms in the cyanobacterial blooms of Lake Taihu (1–3). Overall, the strain Lzh-2 and its two algicidal secretions have potential applications in controlling cyanobacterial blooms.

Shewanella sp. Lzh-2 was incubated in beef extract peptone liquid medium (beef extract, 3 g/liter; peptone, 10 g/liter; agar, 15 g/liter [pH 7.4 to 7.6]) and cultivated with a shaking speed of 200 rpm at 28°C for 48 h. Genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method (4). The DNA libraries were prepared using a SMRTbell template prep kit 1.0 (PacBio, USA), and the DNA fragments generated were tested with an Agilent Bioanalyzer 2100. The fragments longer than 10 kb were selected using BluePippin (Sage Science, USA) and sequenced using the PacBio Sequel platform.

The whole-genome sequencing generated 518,801 reads (N_{50}, 9,414 bp), and the reads were then filtered and assembled into a single scaffold using the Hierarchical Genome Assembly Process (HGAP) (5) and Canu (6). The genome was circular and consisted of a 4,634,534-bp chromosome with a GC content of 46.31%. A total of 4,023 open reading frames (ORF) were predicted in the genome using GeneMarkS (7), accounting for 85.85% of the length of the whole genome. Analysis with SignalP 5.0 (http://www.cbs.dtu.dk/services/SignalP/) (8) showed that 388 of the 4,023 genes encoded secretory proteins. Genes encoding carbohydrate-active enzymes (CAZymes) were annotated using the hmmscan tool (9). There were 106 genes encoding predicted carbohydrate enzymes, which comprised 34 glycoside hydrolases, 23 glycosyl transferases, 5 polysaccharide lyases, 23 carbohydrate esterases, 11 auxiliary activities, and 10 carbohydrate-binding modules.

Five secondary metabolite gene clusters were predicted using antiSMASH 5.0 (https://antismash.secondarymetabolites.org/) (10). The five gene clusters belonged to five different types of secondary metabolite clusters, including an aryl polyene type cluster (chromosomes [chr.] 342896 to 386750) synthesizing aryl polyenes (11), a siderophore type cluster (chr. 1663845 to 1675737) synthesizing desferrioxamine E (12), a RiPP-like type cluster (chr. 2627252 to 2638091), a beta-lactone type cluster (chr. 2734421 to 2765790)

synthesizing eicosapentaenoic acid (13), and a ghlE-KS type cluster (chr. 3025122 to 3081331) synthesizing plipastatin (14). The two algicidal compounds, hexahydropyrrolo[1,2-*a*]pyrazine-1,4-dione and 2,3-indolinedione, might be the intermediate metabolites of some of the predicted secondary metabolite gene clusters. However, the biosynthetic pathways of the two compounds in *Shewanella* sp. Lzh-2 have not been characterized.

Data availability. This whole-genome shotgun project has been deposited in GenBank under the accession no. PRJNA694369. The raw data of genome sequencing have been deposited in the Sequence Read Archive (SRA) database under the accession no. SRP309231. The version described in this paper is the first version.

REFERENCES

1. Li Z, Lin S, Liu X, Tan J, Pan J, Yang H. 2014. A freshwater bacterial strain, *Shewanella* sp. Lzh-2, isolated from Lake Taihu and its two algicidal active substances, hexahydropyrrolo[1,2-a]pyrazine-1,4-dione and 2, 3-indolinedione. Appl Microbiol Biotechnol 98:4737–4748. https://doi.org/10.1007/s00253-014-5602-1.

2. Zhu C, Zhang J, Nawaz MZ, Mahboob S, Al-Ghanim KA, Khan IA, Lu Z, Chen T. 2019. Seasonal succession and spatial distribution of bacterial community structure in a eutrophic freshwater Lake, Lake Taihu. Sci Total Environ 669:29–40. https://doi.org/10.1016/j.scitotenv.2019.03.087.

3. Zhang M, Lu T, Paerl HW, Chen Y, Zhang Z, Zhou Z, Qian H. 2019. Feedback regulation between aquatic microorganisms and the bloom-forming cyanobacterium *Microcystis aeruginosa*. Appl Environ Microbiol 85:e01362-19. https://doi.org/10.1128/AEM.01362-19.

4. Arseneau JR, Steeves R, Laflamme M. 2017. Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. Mol Ecol Resour 17:686–693. https://doi.org/10.1111/1755-0998.12616.

5. Chin CS, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, Dunn C, O'Malley R, Figueroa-Balderas R, Morales-Cruz A, Cramer GR, Delledonne M, Luo C, Ecker JR, Cantu D, Rank DR, Schatz MC. 2016. Phased diploid genome assembly with single-molecule real-time sequencing. Nat Methods 13:1050–1054. https://doi.org/10.1038/nmeth.4035.

6. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.

7. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res 29:2607–2618. https://doi.org/10.1093/nar/29.12.2607.

8. Almagro Armenteros JJ, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol 37:420–423. https://doi.org/10.1038/s41587-019-0036-z.

9. Eddy SR. 2011. Accelerated profile HMM searches. PLoS Comput Biol 7:e1002195. https://doi.org/10.1371/journal.pcbi.1002195.

10. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res 47:W81–W87. https://doi.org/10.1093/nar/gkz310.

11. Cimermancic P, Medema MH, Claesen J, Kurita K, Wieland Brown LC, Mavrommatis K, Pati A, Godfrey PA, Koehrsen M, Clardy J, Birren BW, Takano E, Sali A, Linington RG, Fischbach MA. 2014. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. Cell 158:412–421. https://doi.org/10.1016/j.cell.2014.06.034.

12. Smits TH, Duffy B. 2011. Genomics of iron acquisition in the plant pathogen *Erwinia amylovora*: insights in the biosynthetic pathway of the siderophore desferrioxamine E. Arch Microbiol 193:693–699. https://doi.org/10.1007/s00203-011-0739-0.

13. Allen EE, Bartlett DH. 2002. Structure and regulation of the omega-3 polyunsaturated fatty acid synthase genes from the deep-sea bacterium *Photobacterium profundum* strain SS9. Microbiology (Reading) 148:1903–1913. https://doi.org/10.1099/00221287-148-6-1903.

14. Li XY, Mao ZC, Wang YH, Wu YX, He YQ, Long CL. 2012. ESI LC-MS and MS/MS characterization of antifungal cyclic lipopeptides produced by *Bacillus subtilis* XF-1. J Mol Microbiol Biotechnol 22:83–93. https://doi.org/10.1159/000338530.