Draft Genome Sequence of *Halorubrum* sp. Strain 48-1-W from a Saline Lake in the Novosibirsk Region, Russia

Aleksey S. Rozanov,[a] Aleksandra A. Shipova,[a] Anton V. Korzhuk,[a] Alla V. Bryanskaya,[a] Sergey E. Peltek[a]

[a]Federal Research Center, Institute of Cytology and Genetics of the Siberian Branch of the RAS, Novosibirsk, Russia

ABSTRACT *Halorubrum* sp. strain 48-1-W was isolated from a water sample from a saline lake (Novosibirsk Region, Russia, 54°14′N 78°13′E). The sequenced and annotated genome is 3,584,929 bp and contains 3,506 genes.

The genus *Halorubrum* includes extremely halophilic archaeal species that are widely distributed in high-salinity environments and are a major component of the microbial community of hypersaline environments (1). The genus currently contains 32 validly published species (2, 3).

Halorubrum sp. strain 48-1-W was isolated from a water sample from a saline lake (Novosibirsk Region, Russia, 54°14′N 78°13′E). Isolation and purification of the strain were conducted at 30 to 37°C on agar medium containing 200 g/liter NaCl, 5 g/liter $MgCl_2$, 1 g/liter KCl, 1 g/liter $CaCl_2$, 4 g/liter tryptone, 2 g/liter yeast extract, and 15 g/liter agar. Colonies appeared over the course of a month. We isolated the strain for a collection of biotechnological microorganisms as a source for novel promising objects for biotechnology and bioengineering at the Federal Research Center Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences (RAS).

A single colony of *Halorubrum* sp. 48-1-W was inoculated in 20 ml of liquid medium described above (pH 7.5) and cultivated for 2 weeks with shaking at 200 rpm. Eight milliliters of cell culture was pelleted by centrifugation and resuspended in 75 μl of H_2O. DNA was isolated using a DNA purification kit (Fermentas). The NEBNext Ultra II DNA library prep kit for Illumina (New England BioLabs, USA) was used to create libraries for genome sequencing. Genome sequencing was performed on a MiSeq platform (Illumina), using MiSeq reagent kit version 3 (single-end reads, 150 cycles; Illumina, USA) at the Institute of Molecular and Cellular Biology at the Siberian Branch of the RAS (IMCB SB RAS). A total of 1,980,123 reads were sequenced. The reads were trimmed with Trimmomatic version 0.36 (using options HEADCROP:10, ILLUMINACLIP, and MINLEN:64) (4).

De novo assembly of short reads into contigs was performed using SPAdes version 3.10.1 using default parameters (with option – only-assembler) (5). Contigs shorter than 1,000 bp were deleted. A total of 69 contigs yielded a genome sequence 3,584,929 bp long, with a GC content of 66.07%. Open reading frame (ORF) prediction and automatic annotation were performed using NCBI PGAAP using the default parameters (6). The complete genome sequence contains 3,506 genes, 3,455 coding sequences (CDS), rRNAs (1 5S, 1 16S, and 1 23S), 46 tRNAs, and 2 noncoding RNAs (ncRNAs).

Data availability. The raw data have been deposited at DDBJ/EMBL/GenBank under the accession no. SRR8097410. The draft genome sequence for *Halorubrum* sp. 48-1-W has been deposited at DDBJ/EMBL/GenBank under the accession no. QMIM00000000. The 69 contigs have been deposited under accession no. QMIM01000001 to QMIM01000069.

ACKNOWLEDGMENTS

REFERENCES

1. McGenity TJ, Grant WD. 1995. Transfer of *Halobacterium saccharovorum*, *Halobacterium sodomense*, *Halobacterium trapanicum* NRC 34021 and *Halobacterium lacusprofundi* to the genus *Halorubrum* gen. nov., as *Halorubrum saccharovorum* comb. nov., *Halorubrum sodomense* comb. nov., *Halorubrum trapanicum* comb. nov., and *Halorubrum lacusprofundi* comb. nov. Syst Appl Microbiol 18:237–243. https://doi.org/10.1016/S0723-2020 (11)80394-2.

2. Parte AC. 2014. LSPN—list of prokaryotic names with standing in nomenclature. Nucleic Acids Res 42:D613–D616. https://doi.org/10.1093/nar/gkt1111.

3. Corral P, de la Haba RR, Sánchez-Porro C, Amoozegar MA, Thane Papke R, Ventosa A. 2015. *Halorubrum persicum* sp. nov., an extremely halophilic archaeon isolated from sediment of a hypersaline lake. Int J Syst Evol Microbiol 65:1770–1778. https://doi.org/10.1099/ijs.0.000175.

4. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

5. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

6. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44: 6614–6624. https://doi.org/10.1093/nar/gkw569.