AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of *Virgibacillus* sp. Strain AGTR, Isolated from Hypersaline Lake Acıgöl in Turkey

Meryem Menekse Kılıc,[a,b,c] Nurgul Balci,[d] Nevin Gul Karaguler[a,b]

[a]Istanbul Technical University, Faculty of Science and Letters, Department of Molecular Biology and Genetics, Istanbul, Turkey
[b]Istanbul Topkapi University, Plato Vocational School, Medical Laboratory Techniques Program, Istanbul, Turkey
[c]Istanbul Technical University, Dr. Orhan Öcalgiray Molecular Biology, Biotechnology, and Genetics Research Center, Istanbul, Turkey
[d]Istanbul Technical University, Department of Geological Engineering, Geomicrobiology and Biogeochemistry Laboratory, Istanbul, Turkey

ABSTRACT *Virgibacillus* sp. strain AGTR, which is a haloalkaliphilic microorganism, was isolated from a sediment sample collected in hypersaline Lake Acıgöl in Turkey. It has the potential to produce biotechnologically essential proteases. Here, the whole-genome sequence and its annotations are reported.

The use of enzymes from halophiles in industrial applications is not limited to their stability at high salt concentrations, as these extremozymes are also tolerant to high temperatures or low temperatures and alkali pH and are stable in the presence of organic solvents (1–4). *Virgibacillus* sp. strain AGTR is a Gram-positive, endospore-forming, motile, and halophilic microorganism that was isolated from hypersaline Lake Acıgöl (the salinity rate varies between 5.8% and 13% [wt/vol]). Sediment samples were collected from hypersaline Lake Acıgöl in Turkey (37.837100N, 29.861E). Sediment samples were diluted, and 100 μL diluted sample was plated on protease activity screening plates, which are nutrient agar (NA) plates containing 1% (wt/vol) skim milk and 10% NaCl (wt/vol). The plates were incubated at 30°C for 7 days. Colonies were isolated by repeated streaking (three times) on fresh screening plates. The isolate showing the maximum hydrolysis zone was selected for whole-genome sequencing. The isolate was cultivated in 100 mL nutrient broth (NB) (with 10% NaCl) at 30°C for 24 h, and DNA isolation was carried out with a MoBio kit (catalog number 12888-50). One hundred nanograms of genomic DNA was used to create sequencing libraries with the TruSeq Nano DNA low-throughput library preparation kit (catalog number 20015964; Illumina). Quality control, in terms of size distribution and quantity of the libraries, was performed by using a 2100 Bioanalyzer (Agilent Technologies, USA). Sequencing by synthesis (SBS) was performed using the HiSeq Rapid SBS kit v2 (catalog number FC-402-4023; Illumina). The Illumina HiSeq 2500 platform was used for sequencing. The raw sequence data were checked with FastQC (5), in terms of sequence quality and the presence of any adapter sequences. A total of 4,625,053 raw reads were obtained.

Raw paired-end sequencing data were used for *de novo* assembly with the Unicycler genome assembler v0.4.8 (6). The assembled genome size was 4,708,499 bp, with a GC content of 36.66%, which is in good accordance with values for previously published *Virgibacillus* genomes. The assembled genome was annotated using Prokka v1.13 (7). The annotation pipeline with Prokka uses several external tools, such as Prodigal (8) for coding sequences, RNAmmer v1.2 (9) for rRNA genes, ARAGORN (10) for tRNA genes, and Infernal v1.1 (11) for noncoding RNAs. The draft genome was predicted to have 4,536 coding sequences, 64 tRNA genes, and 5 rRNA sequences. Assembled genome features are indicated in Table 1. The average nucleotide identity (ANI) was calculated by using EzBioCloud TruBac ID software v1 (12), which revealed 99.44% similarity to *Virgibacillus marismortui*. Default parameters were used for all software unless otherwise noted.

Editor Frank J. Stewart, Montana State University

Address correspondence to Nevin Gul Karaguler, karaguler@itu.edu.tr.

The authors declare no conflict of interest.

Received 5 June 2022
Accepted 16 August 2022
Published 31 August 2022

TABLE 1 Genome features of *Virgibacillus* sp. strain AGTR

Feature	Finding
Total no. of contigs	215
Total length (bp)	4,708,499
Length of largest contig (bp)	209,032
N_{50} (bp)	50,902
N_{75} (bp)	27,280
L_{50}	27
L_{75}	58
No. of coding genes	4,536
Mean sequencing coverage (×)	291.70

Data availability. The whole-genome sequence of *Virgibacillus* sp. strain AGTR was submitted to GenBank under BioProject accession number PRJNA701885, BioSample accession number SAMN17916045, and nucleotide accession number JAJERH000000000. The raw data were deposited in the Sequence Read Archive (SRA) under SRA accession number SRP376786.

ACKNOWLEDGMENTS

This study was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) (grant 113Y464 to N.B.) and the ITU Scientific Research Projects Division (project 38445).

REFERENCES

1. Huo YY, Rong Z, Jian SL, Xu CD, Li J, Xu XW. 2017. A novel halotolerant thermoalkaliphilic esterase from marine bacterium *Erythrobacter seohaensis* SW-135. Front Microbiol 8:2315. https://doi.org/10.3389/fmicb.2017.02315.
2. Noby N, Hussein A, Saeed H, Embaby AM. 2020. Recombinant cold-adapted halotolerant, organic solvent-stable esterase (estHlJ) from *Bacillus halodurans*. Anal Biochem 591:113554. https://doi.org/10.1016/j.ab.2019.113554.
3. Borba TM, Machado TB, Brandelli A, Kalil SJ. 2018. Thermal stability and catalytic properties of protease from *Bacillus* sp. P45 active in organic solvents and ionic liquid. Biotechnol Prog 34:1102–1108. https://doi.org/10.1002/btpr.2672.
4. Wang M, Ai L, Zhang M, Wang F, Wang C. 2020. Characterization of a novel halotolerant esterase from *Chromohalobacter canadensis* isolated from salt well mine. 3 Biotech 10:430. https://doi.org/10.1007/s13205-020-02420-0.
5. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
6. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
7. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
8. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.
9. 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. https://doi.org/10.1093/nar/gkm160.
10. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32:11–16. https://doi.org/10.1093/nar/gkh152.
11. Kolbe DL, Eddy SR. 2011. Fast filtering for RNA homology search. Bioinformatics 27:3102–3109. https://doi.org/10.1093/bioinformatics/btr545.
12. Ha S-M, Kim CK, Roh J, Byun J-H, Yang S-J, Choi S-B, Chun J, Yong D. 2019. Application of the whole genome-based bacterial identification system, TrueBac ID, using clinical isolates that were not identified with three matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) systems. Ann Lab Med 39:530–536. https://doi.org/10.3343/alm.2019.39.6.530.