AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements

Genome Sequence of the Multiple-Protease-Producing Strain *Geobacillus thermoleovorans* N7, a Thermophilic Bacterium Isolated from Paniphala Hot Spring, West Bengal, India

Sucharita Bose,[a] Trinetra Mukherjee,[a] Urmimala Sen,[a] Chayan Roy,[b] Moidu Jameela Rameez,[b] Wriddhiman Ghosh,[b] Subhra Kanti Mukhopadhyay[a]

Department of Microbiology, The University of Burdwan, Burdwan, West Bengal, India[a]; Bose Institute, Centenary Campus, Department of Microbiology, Kolkata, West Bengal, India[b]

Here, we present the draft genome sequence of *Geobacillus thermoleovorans* strain N7 (MCC 3175), isolated from Paniphala Hot Spring, West Bengal, India, which contains genes that encode several industrially and medically important thermostable enzymes like neutral protease, xylose isomerase, rhamnogalacturonan acetylesterase, nitrate and nitrite reductase, L-asparaginase, glutaminase, and RNase P.

Received 2 September 2016 **Accepted** 8 September 2016 **Published** 27 October 2016

Citation Bose S, Mukherjee T, Sen U, Roy C, Rameez MJ, Ghosh W, Mukhopadhyay SK. 2016. Genome sequence of the multiple-protease-producing strain *Geobacillus thermoleovorans* N7, a thermophilic bacterium isolated from Paniphala Hot Spring, West Bengal, India. Genome Announc 4(5):e01202-16. doi:10.1128/genomeA.01202-16.

Address correspondence to Subhra Kanti Mukhopadhyay, microskm@gmail.com.

Geobacillus thermoleovorans strain N7 (MCC 3175) is a rod-shaped, Gram-positive, aerobic, thermophilic bacterium, isolated from Paniphala Hot Spring (63 ± 1°C, pH 7.6 ± 0.2), located at Barabani, Asansol, West Bengal (23°45′33″N, 86°58′54″E). This strain belongs to the family *Bacillaceae* under the phylum *Firmicutes*. The strain grew on a nutrient agar plate at (60 ± 1°C) with an incubation period of 16 to 20 h.

The purified genomic DNA of N7 was sequenced on an Ion PGM sequencer (1) using a 318 chip. Raw reads totaling 14,977,65 bp were assembled using SPAdes version 3.8.0 into 226 contigs with a coverage of 52%. A genome comprising 3,400,891 bp was obtained, with the largest contig size being 1,88,606 bp and the smallest being 200 bp, with a GC content of 52.4%. The genome was annotated and analyzed using NCBI PGAP (http://www.ncbi.nlm.nih.gov/genome/annotation_prok), RAST server version 2.0 (2), and PATRIC (3). NCBI PGAP predicted 3,530 genes, with 3,403 coding sequences, 36 rRNA genes, and 86 tRNA genes with 5 noncoding RNA genes. Genes encoding for rhamnogalacturonan acetylesterase (EC 3.1.1.86), an enzyme responsible for deacetylation of the hairy region of pectin (4), and xylose isomerase (EC 5.3.1.5), which is capable of xylose and glucose isomerization, have been found. Assimilatory nitrate reductase (EC 1.7.99.4) and nitrite reductase (EC 1.7.1.4) convert nitrate to nitrite and finally to ammonia. Three kinds of protease, i.e., serine protease, metalloprotease, and neutral protease, are also present in N7. The genome sequence of *G. thermoleovorans* CCB_US3_UF5, closest member of N7 (99% identity) does not have rhamnogalacturonan acetylesterase and neutral protease (5). The bacterial neutral proteases (pH around 7) are very much in use in food industry because they generate less bitterness in hydrolyzed food proteins in comparison to animal proteases (6). Wet lab experiments have already been done for the confirmation of the enzymatic activity of these industrially important enzymes. Apart from that, several genes of medically useful enzymes have been found, namely, L-asparaginase, glutaminase, and RNase P. L-asparaginase and glutaminase have anticancer activity and are used to treat acute lymphoblastic leukemia (7). As reported, an engineered RNase P ribozyme variant has been used as a potential antiviral agent in reducing human cytomegalovirus gene expression and growth (8). Xenobiotic-degrading genes encoding for enzymes like catechol-2,3-dioxygenase (EC 1.13.11.2) and nitrilotriacetate monooxygenase (EC 1.14.14.10), which are responsible for the degradation of catechol and nitrilotriacetate, have been found. Genes for fructose and mannose metabolism, maltose and maltodextrin utilization, and fatty acid biosynthesis and degradation have also been found.

Genes with multiple resistance to some heavy metals like cobalt, cadmium, chromium, and mercury and antimicrobials like polymyxin and melittin, have been found, which were present in strain N7 but not in its closest neighbor *G. thermoleovorans* CCB_US3_UF5 (5). Genes for the uptake and biosynthesis of some compatible solutes like glycine, myo-inositol, l-proline, sarcosine, and trehalose were also found. Enzymes of the central metabolic pathways (i.e., glycolysis, TCA cycle, pentose phosphate pathway, glyoxylate cycle, and oxidative phosphorylation) were also found, which indicates the bacterium's aerobic mode of respiration.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number MDCP00000000. The version described in this paper is the first version, MDCP01000000.

ACKNOWLEDGMENTS

S.B. was supported by a UGC-State Fund Fellowship. S.K.M. was funded by a UGC-sponsored Major Research Project grant (MRP-MAJOR-MICR-2013-7783). T.M. received a fellowship from DST-INSPIRE (IF140017). C.R. and M.J.R. received fellowships from UGC (GOI).

FUNDING INFORMATION

This work, including the efforts of Sucharita Bose, was funded by University Grants Commission (UGC) State Fund. This work, including the efforts of Trinetra Mukherjee, was funded by DST-INSPIRE (IF140017). This work, including the efforts of Chayan Roy and Moidu Jameela Rameez, was funded by UGC (GOI). This work, including the efforts of Subhra Kanti Mukhopadhyay, was funded by UGC-sponsored major research project (MRP-MAJOR-MICR-2013-7783).

REFERENCES

1. **Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schoen M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT, Bustillo J.** 2011. An integrated semiconductor device enabling non-optical genome sequencing. Nature **475:**348–352. http://dx.doi.org/10.1038/nature10242.
2. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9:**75. http://dx.doi.org/10.1186/1471-2164-9-75.
3. **Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJC, Seung YH, Zhang C, Zhang Y, Sobral BW.** 2014. PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res **42:**D581–D591. http://dx.doi.org/10.1093/nar/gkt1099.
4. **Searle-van Leeuwen MJF, van den Broek LAM, Schols HA, Beldman G, Voragen AGJ.** 1992. Rhamnogalacturonan acetylesterase: a novel enzyme from *Aspergillus aculeatus*, specific for the deacetylation of hairy (ramified) regions of pectins. Appl Microbiol Biotechnol **38:**347–349. http://dx.doi.org/10.1007/BF00170084.
5. **Muhd Sakaff MK, Abdul Rahman AY, Saito JA, Hou S, Alam M.** 2012. Complete genome sequence of the thermophilic bacterium *Geobacillus thermoleovorans* CCB_US3_UF5. J Bacteriol **194:**1239. http://dx.doi.org/10.1128/JB.06580-11.
6. **Rao MB, Tanksale AM, Ghatge MS, Deshpande VV.** 1998. Molecular and biotechnological aspects of microbial proteases. Microbiol Mol Biol Rev **62:**597–635.
7. **Ramya LN, Doble M, Rekha VP, Pulicherla KK.** 2012. L-asparaginase as potent anti-leukemic agent and its significance of having reduced glutaminase side activity for better treatment of acute lymphoblastic leukaemia. Appl Biochem Biotechnol **167:**2144–2159. http://dx.doi.org/10.1007/s12010-012-9755-z.
8. **Yang Z, Vu GP, Qian H, Chen YC, Wang Y, Reeves M, Zen K, Liu F.** 2014. Engineered RNase P ribozymes effectively inhibit human cytomegalovirus gene expression and replication. Viruses **6:**2376–2391. http://dx.doi.org/10.3390/v6062376.