

Insight into the First Draft Genome Sequence of the Genus *Sediminibacillus*, *Sediminibacillus halophilus* Strain NSP9.3

Rinku Dey,^a Kamal Krishna Pal,^a Dharmesh Sherathia,^a Bhoomika Sukhadiya,^a Trupti Dalsania,^a Ilaxi Patel,^a Kinjal Savsani,^a Manesh Thomas,^a Sejal Vanpariya,^a Mona Mandaliya,^a Rupal Rupapara,^a Priya Rawal,^a Sucheta Ghorai,^a Sharmila Bhayani,^a Abhi Shah,^a Anil Kumar Saxena^b

Microbiology Section, Directorate of Groundnut Research, Junagadh, Gujarat, India^a; Division of Microbiology, Indian Agricultural Research Institute, New Delhi, India^b

We report the 3.98-Mbp first draft genome sequence of *Sediminibacillus halophilus* strain NSP9.3, a moderate halophile isolated from a seasonal salt marsh of the Great Rann of Kutch, India. Exploring the genome of this organism will facilitate the understanding of the mechanism(s) of osmotolerance and survival in differential osmolarity.

Received 27 November 2013 Accepted 2 December 2013 Published 9 January 2014

Citation Dey R, Pal KK, Sherathia D, Sukhadiya B, Dalsania T, Patel I, Savsani K, Thomas M, Vanpariya S, Mandaliya M, Rupapara R, Rawal P, Ghorai S, Bhayani S, Shah A, Saxena AK. 2014. Insight into the first draft genome sequence of the genus *Sediminibacillus*, *Sediminibacillus halophilus* strain NSP9.3. *Genome Announc.* 2(1):e01133-13. doi:10.1128/genomeA.01133-13.

Copyright © 2014 Dey et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Rinku Dey, rinku_dy@yahoo.co.in.

In the recent past, genomes of a number of halophiles, thriving in the saline environments of the Rann of Kutch, India, were sequenced to understand the mechanism(s) of osmotolerance (1–4). *Sediminibacillus halophilus* strain NSP9.3 (16S rRNA, GenBank accession no. JX518263), a moderately halophilic bacterium, was isolated from a seasonal salt marsh of the Great Rann of Kutch, India. This bacterium grows optimally at a 5% NaCl (range, 0 to 20%) concentration in growth medium at 37°C and 7.5 pH. The genome of *Sediminibacillus halophilus* strain NSP9.3 was sequenced to gain insight into its osmotolerance mechanism(s).

The whole genome of *Sediminibacillus halophilus* strain NSP9.3 (G+C content, 44.05%) was sequenced at Macrogen, Inc., South Korea, through Sequencher Tech Pvt., Ltd., Ahmedabad, India, using a Roche 454 genome sequencer (GS FLX), by both shotgun and 3-kb mate-paired library sequencing. Sequencing of shotgun and mate-paired libraries generated 782,771, 137,390, and 122,992 reads of 432,228,177, 60,630,498, and 53,887,735 bases with average read lengths of 552, 441, and 438 bp, respectively.

De novo assembly of reads using GS De Novo Assembler v 2.6 (5) gave approximately 134-fold coverage, with 10 scaffolds of 3,985,996 bp and 25 contigs of 3,983,638 bp. N_{50} scaffold lengths of 693,230 bp (smallest, 2,877 bp, and largest, 1,699,278 bp) and N_{50} contig lengths of 276,706 bp (smallest, 2,877 bp, and largest, 572,364 bp) were obtained. All assembly data were deposited in the DDBJ/EMBL/GenBank nucleotide sequence database.

The genome was annotated by using the RAST server (6), Glimmer 3 (7, 8), GeneMark (9, 10), the KEGG database (11), tRNAscan-SE (12), RNAmmer (13), and Signal P4.1 (14).

Using the different software tools, we predicted 4,232 coding sequences (CDS), 83 RNA-encoding genes (72 tRNAs and 11 rRNAs), 436 subsystems, and 60 signal peptides. Among the CDS, 2,445 are not in the subsystem (1,010 nonhypothetical, 1,435 hypothetical), whereas 1,787 CDS (1,679 nonhypothetical, 108 hypothetical) are in the subsystem. RAST annotation found involve-

ment of 119 genes in stress responses. These genes included 30 in osmotic stress (2 in osmoregulation, 4 in ectoine biosynthesis and regulation, 23 in choline and betaine uptake and betaine biosynthesis, and 1 in synthesis of osmoregulatory periplasmic glucans), 40 in oxidative stress (9 in protection from reactive oxygen species [ROS], 20 in oxidative stress, 2 in the glutathione: redox cycle, 5 in redox-dependent regulation of nuclear processes, 2 in glutathione: nonredox reactions, and 2 in glutaredoxin reactions), 4 in cold shock, 19 in heat shock, and 26 in no subcategory. In addition, 12 genes in potassium homeostasis, 9 in glycerol and glycerol-3-phosphate uptake and utilization, 8 in mannitol utilization, 307 in amino acids and derivatives, 479 in carbohydrates utilization, etc. were identified. Moreover, a number of genes associated with different ABC transporters (map02010), like glycine betaine/proline (ProX, ProW, and ProV), osmoprotectant (OpuBC and OpuBA), phosphate (PstA, PstB, PstC, and PstS), branched-chain amino acids (LivF, LivG, LivH, LivK, and LivM), etc. and those in two-component systems (map02020) like salt stress-degradative enzymes (DegS and DegU), etc., have been mapped to KEGG pathways.

We are further exploring the first draft genome of *Sediminibacillus halophilus* NSP9.3 to understand its survival and osmoadaptation strategies.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [AWXX00000000](https://www.ncbi.nlm.nih.gov/nuccore/AWXX00000000). The version described in this paper is version AWXX01000000.

ACKNOWLEDGMENTS

The work was carried out in the subproject “Diversity analysis of *Bacillus* and other predominant genera in extreme environments and its utilization in agriculture” of the National Agricultural Innovation Project (NAIP) and “Exploring bacterial and archaeobacterial diversity in Kutch eco-region of Gujarat for agricultural and industrial applications” under the Application of Microorganisms in Agriculture and Allied Sectors

(AMAAS) of the Indian Council of Agricultural Research (ICAR). We thank ICAR for funding through NAIP and AMAAS.

We thank the directors and coordinators, NAIP, and Directorate of Groundnut Research, Junagadh, and NBAIM, Mau, for help during the course of this study. We also thank Macrogen, Inc., South Korea, and Sequencher Tech Pvt. Ltd., Ahmedabad, India, for the genomic services provided.

REFERENCES

1. Pal KK, Dey R, Sherathia D, Dalsania T, Savsani K, Patel I, Thomas M, Ghorai S, Vanpariya S, Rupapara R, Acharya N, Rawal P, Joshi P, Sukhadiya B, Mandaliya M, Saxena AK. 2013. Draft genome sequence of *Salinibacillus aidingensis* strain MSP4, an obligate halophilic bacterium isolated from a salt crystallizer of the Rann of Kutch, India. *Genome Announc.* 1(4):e00253-13. <http://dx.doi.org/10.1128/genomeA.00253-13>.
2. Pal KK, Dey R, Thomas M, Sherathia D, Dalsania T, Patel I, Savsani K, Ghorai S, Vanpariya S, Sukhadiya B, Mandaliya M, Rupapara R, Rawal P, Saxena AK. 2013. Draft genome sequence of *Bacillus* sp. strain SB47, an obligate extreme halophilic bacterium isolated from a salt pan of the Little Rann of Kutch, India. *Genome Announc.* 1(5):e00816-13. <http://dx.doi.org/10.1128/genomeA.00816-13>.
3. Dey R, Pal KK, Sherathia D, Dalsania T, Savsani K, Patel I, Thomas M, Ghorai S, Vanpariya S, Rupapara R, Rawal P, Sukhadiya B, Mandaliya M, Saxena AK. 2013. Draft genome sequence of *Bacillus* sp. strain NSP9.1, a moderately halophilic bacterium isolated from salt marsh of the Great Rann of Kutch, India. *Genome Announc.* 1(5):e00835-13. <http://dx.doi.org/10.1128/genomeA.00835-13>.
4. Pal KK, Dey R, Thomas M, Sherathia D, Dalsania T, Patel I, Savsani K, Ghorai S, Vanpariya S, Sukhadiya B, Mandaliya M, Rupapara R, Rawal P. 2013. Draft genome sequence of the extremely halophilic *Bacillus* sp. strain SB49, isolated from a salt crystallizer pond of the Little Rann of Kutch, India. *Genome Announc.* 1(5):e00869-13. <http://dx.doi.org/10.1128/genomeA.00869-13>.
5. 454 Life Sciences Corporation. 2011. 454 sequencing system software manual, version 2.6, p 1–228. Life Sciences Corporation, Branford, CT.
6. 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, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.
7. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* 23:673–679. <http://dx.doi.org/10.1093/bioinformatics/btm009>.
8. Salzberg SL, Delcher AL, Kasif S, White O. 1998. Microbial gene identification using interpolated Markov models. *Nucleic Acids Res.* 26: 544–548. <http://dx.doi.org/10.1093/nar/26.2.544>.
9. 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. <http://dx.doi.org/10.1093/nar/29.12.2607>.
10. Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26:1107–1115. <http://dx.doi.org/10.1093/nar/26.4.1107>.
11. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. 2004. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32: 277–280. <http://dx.doi.org/10.1093/nar/gkh063>.
12. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955–964. <http://dx.doi.org/10.1093/nar/25.5.0955>.
13. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108. <http://dx.doi.org/10.1093/nar/gkm160>.
14. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8:785–786. <http://dx.doi.org/10.1038/nmeth.1701>.