

Draft Genome Sequence of *Bacillus* sp. Strain NSP2.1, a Nonhalophilic Bacterium Isolated from the Salt Marsh of the Great Rann of Kutch, India

Rinku Dey,^a Kamal Krishna Pal,^a Dharmesh Sherathia,^a Trupti Dalsania,^a Kinjal Savsani,^a Ilaxi Patel,^a Bhoomika Sukhadiya,^a Mona Mandaliya,^a Manesh Thomas,^a Sucheta Ghorai,^a Sejal Vanpariya,^a Rupal Rupapara,^a Priya Rawal,^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

The 5.52-Mbp draft genome sequence of *Bacillus* sp. strain NSP2.1, a nonhalophilic bacterium isolated from the salt marsh of the Great Rann of Kutch, India, is reported here. An analysis of the genome of this organism will facilitate the understanding of its survival in the salt marsh.

Received 28 September 2013 Accepted 2 October 2013 Published 24 October 2013

Citation Dey R, Pal KK, Sherathia D, Dalsania T, Savsani K, Patel I, Sukhadiya B, Mandaliya M, Thomas M, Ghorai S, Vanpariya S, Rupapara R, Rawal P, Saxena AK. 2013. Draft genome sequence of *Bacillus* sp. strain NSP2.1, a nonhalophilic bacterium isolated from the salt marsh of the Great Rann of Kutch, India. *Genome Announc.* 1(5):e00909-13. doi:10.1128/genomeA.00909-13.

Copyright © 2013 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.

The genomes of a number of species of *Bacillus* inhabiting the Little Rann and Great Rann of Kutch, India, have been sequenced recently with the aim of understanding their mechanism(s) of osmotolerance (1–3). *Bacillus* sp. strain NSP2.1 (16S rRNA GenBank accession no. JF802192), a nonhalophilic and endospore-forming bacterium isolated from the salt marsh of the Great Rann of Kutch, India, grows optimally without NaCl in the growth medium (NaCl range, 0 to 4.5%), at 37°C, and a pH 7.5. The genome of *Bacillus* sp. NSP2.1 was sequenced with the aim of understanding its survival mechanism(s) in the salt marsh.

The whole genome of NSP2.1 was sequenced by both shotgun and mate-paired library sequencing using the Roche 454 genome sequencer (GS FLX) at Macrogen Inc., South Korea, through Sequencher Tech Pvt. Ltd., Ahmedabad, India. In shotgun sequencing, 769,707 reads of 347,675,293 bases were obtained (average read length, 451 bp). The sequencing of mate-paired libraries generated 140,804 reads of 63,691,257 bp (average read length, 452 bp) and 140,783 reads of 64,402,285 bp (average read length, 457 bp).

We used the GS *de novo* assembler version 2.6 (4) for assembling the reads, with a coverage of 85-fold. The genome assembly of *Bacillus* sp. NSP2.1 (G+C content of 53.99%) contains 25 scaffolds and 107 contigs of 5,521,528 bp and 5,426,897 bp, respectively, with average lengths of 220,861 bp and 50,710 bp, respectively. An N₅₀ scaffold length of 376,716 bp (smallest, 1,808 bp; largest, 1,854,053 bp), and an N₅₀ contig length of 90,389 bp (smallest, 775 bp; largest, 380,501 bp) were obtained. All assembly data were deposited in the DDBJ/EMBL/GenBank nucleotide sequence database.

The draft genome of *Bacillus* sp. NSP2.1 was annotated by the RAST server (5), Glimmer 3 (6, 7), GeneMark (8, 9), the KEGG database (10), tRNAScan-SE (11), RNAmmer (12), and Signal P4.1 (13) for predicting subsystems, coding sequences (CDS), tRNA and rRNA genes, signal peptides, etc.

Using the different softwares, we predicted 5,425 coding se-

quences (CDSs), with 4,655,526 bp in the CDSs. There are 126 RNA-encoding genes (124 tRNA and 2 rRNA genes), 420 subsystems, and 329 signal peptides. Among the CDSs, 3,459 are not in the subsystem (1,405 nonhypothetical, 2,054 hypothetical), whereas 1,966 CDSs (1,871 nonhypothetical, 95 hypothetical) are in the subsystem. RAST annotation also revealed the involvement of 103 genes in stress responses: 10 in osmotic stress (1 in osmoregulation, 9 in choline and betaine uptake), 36 in oxidative stress (5 in protection from reactive oxygen species, 18 in oxidative stress, 2 in glutathione:nonredox reactions, 7 in redox-dependent regulation of nucleus processes, 1 in the glutathione:redox cycle, 2 in glutaredoxins, and 1 in glutathionylspermidine and trypanothione), 5 in cold shock, 15 in heat shock, 15 in detoxification, and 22 in no subcategory. A number of genes associated with ABC transporters (map02010), including osmoprotectants (OpuBC, OpuBB, and OpuBA), glycerol-3-phosphate uptake, and that of two-component systems (map02020), such as those involved in the response to K⁺-limitation and K⁺-transport, and salt stress-degrading enzymes, have been mapped.

Our laboratory is exploring the genome of *Bacillus* sp. NSP2.1 further to know the mechanisms of survival for this nonhalophilic bacterium in salty settings.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [AVBJ00000000](https://www.ncbi.nlm.nih.gov/nuccore/AVBJ00000000). The version described in this paper is version AVBJ01000000.

ACKNOWLEDGMENTS

The work was carried out in the subproject “Diversity analysis of *Bacillus* and other pre-dominant genera in extreme environments and its utilization in agriculture” of the National Agricultural Innovation Project (NAIP) and 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 also thank the national director and coordinators, NAIP, and the

directors of the Directorate of Groundnut Research, Junagadh and NBALM, Mau, for help during the course of this study. We thank Macro-gen 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. doi: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 halophile isolated from a salt pan of the Little Rann of Kutch, India. *Genome Announc.* 1(5):e00816-13. doi: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 the salt marsh of the Great Rann of Kutch, India. *Genome Announc.* 1(5):e00835-13. doi:10.1128/genomeA.00835-13.
4. 454 Life Sciences Corporation. 2011. 454 sequencing system software manual, version 2.6. 454 Life Sciences Corporation, Branford, CT.
5. 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. doi:10.1186/1471-2164-9-75.
6. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* 23:673–679.
7. Salzberg SL, Delcher AL, Kasif S, White O. 1998. Microbial gene identification using interpolated Markov models. *Nucleic Acids Res.* 26: 544–548.
8. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genome. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 29:2607–2618.
9. Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26:1107–1115.
10. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. 2004. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32: 277–280.
11. 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.
12. 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.
13. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8:785–786.