

Draft Genome Sequence of an Obligate and Moderately Halophilic Bacterium, *Thalassobacillus devorans* Strain MSP14, the First Draft Genome of the Genus *Thalassobacillus*

Kamal Krishna Pal,^a Rinku Dey,^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 3.93-Mbp first draft genome sequence of a species of the genus *Thalassobacillus*, *Thalassobacillus devorans* strain MSP14, a moderate but obligate halophile, isolated from a salt crystallizer of the Little Rann of Kutch, India. Exploring the genome of this organism will facilitate understanding the mechanism(s) of its obligate halophilism.

Received 19 November 2013 Accepted 26 November 2013 Published 26 December 2013

Citation Pal KK, Dey R, 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. 2013. Draft genome sequence of an obligate and moderately halophilic bacterium, *Thalassobacillus devorans* strain MSP14, the first draft genome of the genus *Thalassobacillus*. *Genome Announc.* 1(6):e01103-13. doi:10.1128/genomeA.01103-13.

Copyright © 2013 Pal 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 Kamal Krishna Pal, kkpal9426476749@gmail.com.

The genomes of a number of halophilic bacilli, isolated from the salt crystallizers of the Rann of Kutch, Gujarat, India, have been sequenced with a view to understanding the mechanism(s) of osmotolerance (1–4). *Thalassobacillus devorans* strain MSP14 (16S rRNA, GenBank accession no. JX518269), an obligate but moderately halophilic bacterium, was isolated from a salt crystallizer of the Little Rann of Kutch, India. It grows optimally at a concentration of 7.5% NaCl (range, 5 to 15%) in medium at 37°C and pH 7.5. The present genome of MSP14 was sequenced to understand the mechanism(s) of obligate, but moderate, halophilism.

By use of the Roche 454 genome sequencer (GS FLX), the genome of *Thalassobacillus devorans* strain MSP14 (G+C content of 42.97%) was sequenced at Macrogen, Inc., South Korea, through Sequencher Tech Pvt., Ltd., Ahmedabad, India. Both shotgun and 3-kb mate-paired library sequencing were performed. Whereas sequencing of shotgun libraries generated 787,155 reads of 439,717,712 bases (average read length of 558 bp), mate-paired libraries generated 140,683 and 131,751 reads of 62,893,040 and 57,724,648 bases, respectively, with average read lengths of 447 and 438 bp, respectively.

De novo assembly of the reads using GS De Novo Assembler v 2.6 (5) gave approximately 138-fold coverage with 9 scaffolds of 3,935,133 bp and 32 contigs of 3,923,021 bp with average lengths of 437,237 bp and 122,594 bp, respectively. An N_{50} scaffold length of 1,214,989 bp (smallest, 2,007 bp and largest, 1,848,831 bp) and an N_{50} contig length of 260,182 bp (smallest, 1,596 bp and largest 467,343 bp) were obtained. All assembly data were deposited in the DDBJ/EMBL/GenBank nucleotide sequence database.

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

Using the different software tools, we predicted 4,011 coding sequences (CDS), 76 RNA-encoding genes (72 tRNA and 4

rRNA), 464 subsystems, and 31 signal peptides. Among the CDS, 2,097 are not in a subsystem (955 nonhypothetical and 1,142 hypothetical), whereas 1,914 CDS (1,819 nonhypothetical and 995 hypothetical) are in a subsystem. RAST annotation also revealed the association of 123 genes in stress responses: 25 genes in osmotic stress (1 in osmoregulation, 4 in ectoine biosynthesis and regulation, and 20 in choline and betaine uptake and betaine biosynthesis), 44 in oxidative stress (8 in protection from reactive oxygen species [ROS], 21 in oxidative stress, 1 in the glutathione: redox cycle, 7 in redox-dependent regulation of nuclear processes, 2 in glutathione: nonredox reactions, 2 in glutaredoxins, 1 in the CoA disulfide thiol-disulfide redox system, and 2 in cluster-containing glutathione synthetase), 4 in cold shock, 17 in heat shock, 9 in detoxification, and 24 in no subcategory. Similarly, 1,840 genes were mapped to different biochemical pathways of KEGG (map00010 to map05340), including genes in ABC transporters (map02010) and two-component systems (map02020) like transporters for glycine betaine/proline (ProX, ProW, and ProV), osmoprotectants (OpuBC and OpuBA), phosphates (PstA, PstB, PstC, and PstS), salt stress-degradative enzymes (DegS and DegU), etc.

We are exploring the genome of *Thalassobacillus devorans* MSP14 further to unravel the mechanism(s) of its obligate halophilism.

Nucleotide sequence accession numbers. This whole genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [AWXW000000000](https://www.ncbi.nlm.nih.gov/nuccore/AWXW000000000). The version described in this paper is version AWXW01000000.

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 the

Kutch eco-region of Gujarat for agricultural and industrial applications” under 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 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. 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 salt marsh of the Great Rann of Kutch, India. *Genome Announc.* 1(5):e00835-13. doi: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. doi:10.1128/genomeA.00869-13.
5. 454 Life Sciences Corporation. 2011. 454 sequencing system software manual, version 2.8. 454 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.
7. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* 23:673–679.
8. Salzberg SL, Delcher AL, Kasif S, White O. 1998. Microbial gene identification using interpolated Markov models. *Nucleic Acids Res.* 26: 544–548.
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.
10. Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26:1107–1115.
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.
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.
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.
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.