AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Genome Sequence and Methylation Pattern of *Haloterrigena salifodinae* BOL5-1, an Extremely Halophilic Archaeon from a Bolivian Salt Mine

Priya DasSarma,[a] Brian P. Anton,[b] Satyajit L. DasSarma,[a] Hedvig A. L. von Ehrenheim,[a] Fabiana L. Martinez,[a,c] Daniel Guzmán,[d] Richard J. Roberts,[b] Shiladitya DasSarma[a,e]

[a]Department of Microbiology and Immunology, Institute of Marine and Environmental Technology, University of Maryland, Baltimore, Maryland, USA
[b]New England Biolabs, Ipswich, Massachusetts, USA
[c]Instituto de Investigaciones para la Industria Química, Consejo Nacional de Investigaciones Científicas y Técnicas, Universidad Nacional de Salta, Salta, Argentina
[d]Centro de Biotecnología, Faculty of Sciences and Technology, Universidad Mayor de San Simón, Cochabamba, Bolivia
[e]Blue Marble Space Institute of Science, Seattle, Washington, USA

ABSTRACT The halophilic archaeon *Haloterrigena salifodinae* BOL5-1 was isolated from a Bolivian salt mine and sequenced using single-molecule real-time sequencing. The GC-rich genome was 5.1 Mbp, with a 4.2-Mbp chromosome and 5 plasmids ranging from 96 to 281 kbp. The genome annotation was incorporated into HaloWeb (https://halo.umbc.edu), and the methylation patterns were incorporated into REBASE (http://tools.neb.com/genomes/view.php?seq_id=99167&list=1).

Halophilic microbes capable of surviving extreme conditions are of interest for biotechnology and astrobiology (1–12). Our recent focus has been on high elevation and subsurface hypersaline environments which yield polyextremophilic varieties. In this announcement, isolation of an extremely halophilic archaeon, *Haloterrigena salifodinae* BOL5-1, is reported, together with the first complete genome sequence for this species.

Pink salt was sampled from a remote salt mine in the Department of Tarija, O'Connor Province, Bolivia (21°24′19.73″S, 64°07′51.52″W), at 1,230 m elevation, where temperatures range from −10 to 37°C. The salt samples were dissolved in CM^+ medium and grown with shaking at 220 rpm at 37°C as previously described (13, 14). Enrichment cultures were plated onto CM^+ agar plates, and a pigmented isolate, *H. salifodinae* BOL5-1, was purified by three rounds of streaking.

Nucleic acids were extracted using standard methods (14), and sequencing was performed using the Sequel platform (PacBio, Menlo Park, CA). A SMRTbell library was prepared from 5 μg unsheared BOL5-1 genomic DNA, size selected on the BluePippin system (Sage Science, Beverly, MA) with a lower limit of 15 kb, purified for three rounds with AMPure beads (Pacific Biosciences) at 0.45×, and sequenced on one single-molecule real-time (SMRT) cell with the Sequel binding kit version 3.0 with 20-h collection and 2-h preextension times. The sequencing subreads were filtered and assembled *de novo* using the microbial assembly pipeline under SMRTLink version 9.0.0.92188 with default parameters. The 121,750 mapped subreads (mean length, 14,289 bp; coverage, 340×) resolved into six polished, circular contigs.

The assembled *H. salifodinae* BOL5-1 genome sequence comprised 5,087,240 bp (GC content, 63.4%) and included a circular chromosome (4,180,318 bp; GC content, 64.7%) and the plasmids pHTS280.6 (280,619 bp; GC content, 61.9%), pHTS220 (220,397 bp; GC content, 62.2%), pHTS171 (171,484 bp; GC content, 64.2%), pHTS138 (138,030 bp; GC content, 55.7%), and pHTS96 (96,392 bp; GC content, 58.4%). The genes were predicted first using GeneMark HMM (15), analyzed further with HaloWeb version r1613245396

Citation DasSarma P, Anton BP, DasSarma SL, von Ehrenheim HAL, Martinez FL, Guzmán D, Roberts RJ, DasSarma S. 2021. Genome sequence and methylation pattern of *Haloterrigena salifodinae* BOL5-1, an extremely halophilic archaeon from a Bolivian salt mine. Microbiol Resour Announc 10:e00275-21. https://doi.org/10.1128/MRA.00275-21.

Editor Kenneth M. Stedman, Portland State University

Address correspondence to Shiladitya DasSarma, sdassarma@som.umaryland.edu.

Received 12 March 2021
Accepted 9 April 2021
Published 6 May 2021

TABLE 1 Motifs containing the methylated bases ^{m6}A and ^{m4}C

Motif[a]	Modification type	No. of sites in genome	% detected	Mean IPD ratio[b]	No. of predicted ORFs[c]
GGRC**A**G	^{m6}A	4,340	99.8	5.6	Unknown
CGTG**A**YC	^{m6}A	3,713	100	5.1	Unknown
C**A**TTC	^{m6}A	3,903	99.7	5.0	16,755 (M.Hsa51II)
GAGA**A**G	^{m6}A	3,192	99.7	4.9	Unknown
ACGA**C**GC	^{m4}C	4,012	88.5	3.3	Unknown
CTAG	^{m4}C	4,856	74.5	3.2	18,220 (M.Hsa51I)

[a] Locations of methylated bases are in bold for the top strand and underlined for the bottom strand.
[b] IPD, interpulse duration.
[c] ORFs, open reading frames.

(16) and EMBOSS version 6.6.0.0 (17), and finally deposited in NCBI, where the genome was reannotated using NCBI's Prokaryotic Genome Annotation Pipeline (PGAP) build 3190 (18). DTU Health Tech Feature Extract version 1.2 (19) was used to converge the GeneMark and GenBank annotations, and the genome sequence and annotation were made publicly available on HaloWeb (https://halo.umbc.edu).

The BOL5-1 genome contained 4,729 protein genes, plus 4 rRNA operons and 54 tRNA genes. The 16S RNA sequence and average nucleotide identity were used for taxonomic analysis at GenBank. The proteome was highly acidic (2), with a calculated mean pI value of 4.64, and nearly all of the core haloarchaeal orthologous groups (cHOGs) were present (20–22). The BOL5-1 genome contained expanded gene families, e.g., Orc/Cdc6, TATA-binding, and TFB protein genes (23), and a gene cluster for gas vesicle nanoparticles (24, 25) on the chromosome. The genome also encoded transposase genes, suggesting the presence of ISH elements (26).

Methylated DNA motifs and the methyltransferases (MTases) were identified using the Pacific Biosciences base modification analysis protocol under SMRTLink version 9.0.0.92188 using default parameters and were deposited in REBASE (Table 1) (27).

Data availability. The *H. salifodinae* BOL5-1 genome sequence has been deposited in GenBank under the accession numbers CP069188 through CP069193. The raw data are available under the BioSample accession number SAMN17385152.

ACKNOWLEDGMENTS

The DasSarma laboratory was supported by NASA grant 80NSSC17K0263 and NIH grant AI139808. F.L.M. was supported by the Fulbright Fellowship Program and D.G. by the Swedish International Development Cooperation Agency (ASDI).

B.P.A. and R.J.R. work for New England Biolabs, a company that sells research reagents, including restriction enzymes and DNA methyltransferases, to the scientific community.

REFERENCES

1. DasSarma S, DasSarma P. 2017. Halophiles. *In* Encyclopedia of life science. John Wiley & Sons, Hoboken, NJ. https://doi.org/10.1002/9780470015902.a0000394.pub4.
2. DasSarma S, DasSarma P. 2015. Halophiles and their enzymes: negativity put to good use. Curr Opin Microbiol 25:120–126. https://doi.org/10.1016/j.mib.2015.05.009.
3. DasSarma P, DasSarma S. 2018. Survival of microbes in Earth's stratosphere. Curr Opin Microbiol 43:24–30. https://doi.org/10.1016/j.mib.2017.11.002.
4. DasSarma S, DasSarma P, Laye VJ, Schwieterman EW. 2020. Extremophilic models for astrobiology: haloarchaeal survival strategies and pigments for remote sensing. Extremophiles 24:31–41. https://doi.org/10.1007/s00792-019-01126-3.
5. DasSarma S, Schwieterman EW. 2018. Early evolution of purple retinal pigments on Earth and implications for exoplanet biosignatures. Int J Astrobiol:1–10. https://doi.org/10.1017/S1473550418000423.
6. DasSarma P, Antunes A, Simões MF, DasSarma S. 2020. Earth's stratosphere and microbial life. Curr Issues Mol Biol 38:197–244. https://doi.org/10.21775/cimb.038.197.
7. DasSarma P, Antunes A, Simões MF, DasSarma S. 2020. Earth's stratosphere and microbial life, p 197–244. *In* Antunes A (ed), Astrobiology: current, evolving and emerging perspectives. Caister Academic Press, Poole, United Kingdom. https://doi.org/10.21775/9781912530304.08.
8. Carrier BL, Beaty DW, Meyer MA, Blank JG, Chou L, DasSarma S, Des Marais DJ, Eigenbrode JL, Grefenstette N, Lanza NL, Schuerger AC, Schwendner P, Smith HD, Stoker CR, Tarnas JD, Webster KD, Bakermans C, Baxter BK, Bell MS, Benner SA, Bolivar Torres HH, Boston PJ, Bruner R, Clark BC, DasSarma P, Engelhart AE, Gallegos ZE, Garvin ZK, Gasda PJ, Green JH, Harris RL, Hoffman ME, Kieft T, Koeppel AHD, Lee PA, Li X, Lynch KL, Mackelprang R, Mahaffy PR, Matthies LH, Nellessen MA, Newsom HE, Northup DE, O'Connor BRW, Perl SM, Quinn RC, Rowe LA, Sauterey B, Schneegurt MA, Schulze-Makuch D, Scuderi LA, Spilde MN, Stamenkovic V, Torres Celis JA, Viola D, Wade BD, et al. 2020. Mars extant life: what's next? Conference report. Astrobiology 20:785–814. https://doi.org/10.1089/ast.2020.2237.
9. DasSarma P, Anton BP, DasSarma SL, Martinez FL, Guzman D, Roberts RJ, DasSarma S. 2019. Genome sequences and methylation patterns of *Natrinema versiforme* BOL5-4 and *Natrinema pallidum* BOL6-1, two extremely halophilic archaea from a Bolivian salt mine. Microbiol Resour Announc 8:e00810-19. https://doi.org/10.1128/MRA.00810-19.

10. DasSarma P, Anton BP, DasSarma S, Laye VJ, Guzman D, Roberts RJ, DasSarma S. 2019. Genome sequence and methylation patterns of *Halorubrum* sp. strain BOL3-1, the first haloarchaeon isolated and cultured from Salar de Uyuni, Bolivia. Microbiol Resour Announc 8:e00386-19. https://doi.org/10.1128/MRA.00386-19.
11. Pecher WT, Martínez FL, DasSarma P, Guzmán D, DasSarma S. 2020. 16S rRNA gene diversity in the salt crust of Salar de Uyuni, Bolivia, the world's largest salt flat. Microbiol Resour Announc 9:e00374-20. https://doi.org/10.1128/MRA.00374-20.
12. Pecher WT, Martínez FL, DasSarma P, Guzmán D, DasSarma S. 2020. 16S rRNA gene diversity in ancient gray and pink salt from San Simón salt mines in Tarija, Bolivia. Microbiol Resour Announc 9:e00820-20. https://doi.org/10.1128/MRA.00820-20.
13. Berquist BR, Müller JA, DasSarma S. 2006. Genetic systems for halophilic archaea, p 649–680. *In* Oren A, Rainey F (ed), Methods in microbiology, vol 35. Elsevier Academic Press, San Diego, CA.
14. Ng WL, Yang CF, Halladay JT, Arora P, DasSarma S. 1995. Protocol 25. Isolation of genomic and plasmid DNAs from *Halobacterium halobium*, p 179–184. *In* DasSarma S, Fleischmann EM (ed), Archaea, a laboratory manual: halophiles. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15. 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. https://doi.org/10.1093/nar/29.12.2607.
16. DasSarma SL, Capes MD, DasSarma P, DasSarma S. 2010. HaloWeb: the haloarchaeal genomes database. Saline Syst 6:12. https://doi.org/10.1186/1746-1448-6-12.
17. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 47:W636–W641. https://doi.org/10.1093/nar/gkz268.
18. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.
19. Wenerson R. 2005. FeatureExtract—extraction of sequence annotation made easy. Nucleic Acids Res 33:W567–W569. https://doi.org/10.1093/nar/gki388.
20. Capes MD, DasSarma P, DasSarma S. 2012. The core and unique proteins of haloarchaea. BMC Genomics 13:39. https://doi.org/10.1186/1471-2164-13-39.
21. DasSarma S, Capes M, DasSarma P. 2008. Haloarchaeal megaplasmids, p 3–30. *In* Blum P (ed), Microbial megaplasmids. Springer, Berlin, Germany.
22. Kennedy SP, Ng WV, Salzberg SL, Hood L, DasSarma S. 2001. Understanding the adaptation of *Halobacterium* species NRC-1 to its extreme environment through computational analysis of its genome sequence. Genome Res 11:1641–1650. https://doi.org/10.1101/gr.190201.
23. Capes MD, Coker JA, Gessler R, Grinblat-Huse V, DasSarma SL, Jacob CG, Kim J-M, DasSarma P, DasSarma S. 2011. The information transfer system of halophilic archaea. Plasmid 65:77–101. https://doi.org/10.1016/j.plasmid.2010.11.005.
24. DasSarma S, DasSarma P. 2015. Gas vesicle nanoparticles for antigen display. Vaccines (Basel) 3:686–702. https://doi.org/10.3390/vaccines3030686.
25. DasSarma P, DasSarma S. Gas vesicle nanoparticles, in press.
26. DasSarma S. 2004. Genome sequence of an extremely halophilic archaeon, p 383–399. *In* Fraser CM, Read TD, and Nelson KE (ed), Microbial genomes, Humana Press, Inc., Totowa, NJ.
27. Roberts RJ, Vincze T, Posfai J, Macelis D. 2015. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res 43:D298–D299. https://doi.org/10.1093/nar/gku1046.