AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

# Genome Sequence of *Halobacterium* sp. Strain BOL4-2, Isolated and Cultured from Salar de Uyuni, Bolivia

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

[a]Institute of Marine and Environmental Technology, Department of Microbiology and Immunology, University of Maryland, Baltimore, Maryland, USA
[b]New England Biolabs, Ipswich, Massachusetts, USA
[c]Centro de Biotecnología, Faculty of Sciences and Technology, Universidad Mayor de San Simón, Cochabamba, Bolivia

**ABSTRACT** *Halobacterium* sp. strain BOL4-2 was isolated from an Andean salt flat, Salar de Uyuni, in Bolivia. Single-molecule real-time (SMRT) sequencing revealed a 2.4-Mbp genome with a 2.0-Mbp chromosome and four plasmids (2 to 299 kb). Its isolation from an environment experiencing multiple extremes makes the strain interesting for astrobiology.

An extreme halophile, strain BOL4-2, was isolated from Salar de Uyuni, Department of Potosí, Bolivia, the world's largest salt flat and an environment remarkable for its high elevation, cold temperatures, and UV radiation exposure (1–4). Such environments are of significant interest to the astrobiology community due to their multiple extremes (4–11).

For isolation, salt crust was sampled at the surface of Salar de Uyuni (20°33′28.58″S, 67°12′29.56″W), 3,647 m above sea level, in March 2015. Typical conditions there include pH values of 7.3 to 7.6, ≥28% (wt/vol) NaCl concentration, and temperatures ranging from −15 to 22°C. Salt samples were dissolved in complete medium plus trace metals ($CM^+$), and cells were grown at 37°C with shaking at 220 rpm (product number 4230; Innova, New Brunswick, NJ, USA) as described previously (12, 13). The enrichment culture was plated on $CM^+$ agar plates and purified by three rounds of streaking.

Nucleic acids were extracted using standard methods (14–16). Briefly, after cell lysis by osmotic shock, the lysate was extracted three times with phenol saturated with Tris-HCl (pH 8), followed by dialysis and treatment with RNase. Single-molecule real-time (SMRT) sequencing was performed using the Sequel platform (Pacific Biosciences [PacBio], Menlo Park, CA). Genomic DNA (2 $\mu$g) was randomly sheared with a Megaruptor (Diagenode, Denville, NJ) using the 40-kb setting, and a SMRTbell sequencing library was prepared from the sheared DNA using a modified version of the manufacturer's protocol (17). No size selection was performed. The library was sequenced on a single SMRT cell (Sequel binding kit v3.0 and Sequel sequencing plate v3.0) with 10-h collection and 2-h preextension times. Sequencing reads were filtered (quality scores of ≥0.7) and assembled (2,485,291 subreads [mean length, 4,516 bp]) separately with HGAP4 (using FALCON override ovlp_DBsplit_option = -s50, yielding 7 contigs) and Microbial Assembly analysis (using default parameters, yielding 16 contigs). Both sets included 5 contigs with high levels of coverage (~4,000×), which formed the final assembly. Three contigs were circularized automatically with Microbial Assembly analysis, a fourth (pBOL4-2_299.7) was circularized from an HGAP4 contig and checked by manual analysis of read structure mapped back to the circularization point, and the fifth (pBOL4-2_2.1) manifested as a linear concatemer in both assemblies, with the final sequence inferred as the monomer. Methylation patterns were determined using Base Modification analysis using default parameters. All PacBio programs were run under the SMRT Link v6.0.0.47841 environment.

**Editor** Kenneth M. Stedman, Portland State University



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

**Received** 25 October 2021
**Accepted** 5 November 2021
**Published** 2 December 2021

Genome annotation was performed in house using GeneMark.hmm v2 (18) and EMBOSS (19). The genome was also processed through the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (build 3190) (20).

The 2,428,492-bp genome consists of a large circular chromosome (1,989,773 bp [GC content, 68.1%]) and four plasmids, i.e., pBOL4-2_299.7 (299,663 bp [GC content, 59.1%]), pBOL4-2_105.9 (105,947 bp [GC content, 60%]), pBOL4-2_49.8 (49,847 bp [GC content, 61.5%]), and pBOL4-2_2.1 (2,144 bp [GC content, 61.5%]). A single rRNA operon, including 16S rRNA and 47 tRNA genes, are present. The genome sequence was submitted to GenBank, and the taxonomy was determined by NCBI Taxonomy (21).

The *Halobacterium* sp. strain BOL4-2 genome contained 2,555 encoded proteins, with a calculated mean pI value of 4.91, a highly acidic proteome characteristic of haloarchaea (22–24). The genome encodes all 799 conserved core haloarchaeal groups (cHOGs) (25), 10 Orc/Cdc6 proteins, a TATA-binding protein, and 7 transcription factor B (TFB) proteins (26–29). Also encoded are the retinal proteins bacteriorhodopsin, halorhodopsin, and sensory rhodopsin 2 (30) and buoyant gas vesicle nanoparticles (31, 32). The methylated DNA motifs (with methylated base underlined) recorded in the REBASE database include CTAG (m4C), GTCACG (m4C), CGAYNNNNNNGTRC/GYACNNNNNNRTCG (m6A), and AGCANNNNNNCTG/CAGNNNNNNTGCT (m6A) (33).

**Data availability.** The *Halobacterium* BOL4-2 genome sequence has been deposited in GenBank with the accession numbers CP070332.1, CP070334.1, CP070335.1, CP070336.1, and CP070337.1. Raw data are available in the NCBI Sequence Read Archive (SRA) with the accession number SRX10292502.

## ACKNOWLEDGMENTS

Work in the DasSarma laboratory is supported by NASA Exobiology grant 80NSSC17K0263 and NIH grant AI139808. F.L.M. was supported by the Fulbright Fellowship Program, and D.G. thanks the Swedish International Development Cooperation Agency (ASDI) for supporting his work.

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 P, Anton BP, DasSarma SL, 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.
2. 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.
3. 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.
4. DasSarma S, DasSarma P. 2017. Halophiles. *In* eLS. John Wiley & Sons, Ltd, New York, NY. https://doi.org/10.1002/9780470015902.a0000394.pub4.
5. 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.
6. Reid IN, Sparks WB, Lubow S, McGrath M, Livio M, Valenti J, Sowers KR, Shukla HD, MacAuley S, Miller T, Suvanasuthi R, Belas R, Colman A, Robb FT, DasSarma P, Müller JA, Coker JA, Cavicchioli R, Chen F, DasSarma S. 2006. Terrestrial models for extraterrestrial life: methanogens and halophiles at Martian temperatures. Int J Astrobiol 5:89–97. https://doi.org/10.1017/S1473550406002916.
7. DasSarma P, Coker JA, Huse V, DasSarma S. 2009. Microorganisms, halophiles, industrial applications, p 2769–2777. *In* Flickinger MC (ed), Wiley encyclopedia of industrial biotechnology, bioprocess, bioseparation, and cell technology. John Wiley and Sons, New York, NY.
8. DasSarma P, Simões MF, Antunes A, DasSarma S. 2020. Earth's stratosphere and microbial life. Curr Issues Mol Biol 38:197–244. https://doi.org/10.21775/cimb.038.197.
9. DasSarma S, Schwieterman EW. 2021. Early evolution of purple retinal pigments on Earth and implications for exoplanet biosignatures. Int J Astrobiol 20:241–250. https://doi.org/10.1017/S1473550418000423.
10. 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, Stamenković V, Torres Celis JA, Viola D, Wade BD, Walker CJ, Wiens RC, Williams AJ, Williams JM, Xu J. 2020. Mars extant life: what's next? Conference report. Astrobiology 20:785–814. https://doi.org/10.1089/ast.2020.2237.
11. DasSarma P, Anton BP, von Ehrenheim HAL, Roberts RJ, DasSarma S. 2021. Complete genome sequence of an extremely halophilic archaeon from Great Salt Lake, *Halobacterium* sp. GSL-19. Microbiol Resour Announc 10:e00520-21. https://doi.org/10.1128/MRA.00520-21.
12. Berquist BR, Müller JA, DasSarma S. 2006. Genetic systems for halophilic archaea. Methods Microbiol 35:649–680. https://doi.org/10.1016/S0580-9517(08)70030-8.
13. Ng WV, Kennedy SP, Mahairas GG, Berquist B, Pan M, Shukla HD, Lasky SR, Baliga NS, Thorsson V, Sbrogna J, Swartzell S, Weir D, Hall J, Dahl TA, Welti R, Goo YA, Leithauser B, Keller K, Cruz R, Danson MJ, Hough DW, Maddocks DG, Jablonski PE, Krebs MP, Angevine CM, Dale H, Isenbarger TA, Peck RF, Pohlschroder M, Spudich JL, Jung KW, Alam M, Freitas T, Hou S, Daniels CJ, Dennis PP, Omer AD, Ebhardt H, Lowe TM, Liang P, Riley M, Hood L, DasSarma S. 2000. Genome sequence of *Halobacterium* species NRC-1. Proc Natl Acad Sci U S A 97:12176–12181. https://doi.org/10.1073/pnas.190337797.
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. DasSarma S. 1984. The bacterio-opsin gene in Halobacterium halobium: high-frequency inactivation by insertion sequences and structure of the messenger RNA. PhD thesis. Massachusetts Institute of Technology, Boston, MA. https://mit.primo.exlibrisgroup.com/permalink/01MIT_INST/ejdckj/alma990002112390106761.
16. Fomenkov A, DasSarma P, Kennedy SP, Roberts RJ, DasSarma S. 2021. Complete genome and methylome analysis of the box-shaped halophilic archaeon *Haloarcula sinaiiensis* ATCC 33800. Microbiol Resour Announc 10:e00619-21. https://doi.org/10.1128/MRA.00619-21.
17. Lefoulon E, Vaisman N, Frydman HM, Sun L, Voland L, Foster JM, Slatko BE. 2019. Large enriched fragment targeted sequencing (LEFT-SEQ) applied to capture of *Wolbachia* genomes. Sci Rep 9:5939. https://doi.org/10.1038/s41598-019-42454-w.
18. Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res 33:W451–W454. https://doi.org/10.1093/nar/gki487.
19. 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.
20. 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.
21. Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, Leipe D, Mcveigh R, O'Neill K, Robbertse B, Sharma S, Soussov V, Sullivan JP, Sun L, Turner S, Karsch-Mizrachi I. 2020. NCBI taxonomy: a comprehensive update on curation, resources and tools. Database 2020:baaa062. https://doi.org/10.1093/database/baaa062.
22. Kozlowski LP. 2016. IPC: Isoelectric Point Calculator. Biol Direct 11:55. https://doi.org/10.1186/s13062-016-0159-9.
23. Kennedy SP, Ng WV, Salzberg SL, Hood L, DasSarma S. 2001. Understanding the adaptation of *Halobacterium* species NRC-1 to its extreme environment though computational analysis of its genome sequence. Genome Res 11:1641–1650. https://doi.org/10.1101/gr.190201.
24. 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.
25. 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.
26. DasSarma S, Capes M, DasSarma P. 2008. Haloarchaeal megaplasmids, p 3–30. *In* Schwartz E (ed), Microbial megaplasmids. Springer-Verlag, Berlin, Germany. https://link.springer.com/chapter/10.1007%2F978-3-540-85467-8_1.
27. 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.
28. Berquist BR, Soneja J, DasSarma S. 2005. Comparative genomic survey of information transfer systems in two diverse extremely halophilic archaea, *Halobacterium* sp. strain NRC-1 and *Haloarcula marismortui*, p 151–183. *In* Gunde-Cimerman N, Oren A, Plemenitaš A (ed), Adaptation to life at high salt concentrations in archaea, bacteria, and eukarya. Springer, Dordrecht, Netherlands. https://doi.org/10.1007/1-4020-3633-7.
29. DasSarma S, Kennedy SP, Berquist B, Ng W-LV, Baliga NS, Spudich JL, Krebs MP, Eisen JA, Johnson CH, Hood L. 2001. Genomic perspective on the photobiology of *Halobacterium* species NRC-1, a phototrophic, phototactic, and UV-tolerant haloarchaeon. Photosynth Res 70:3–17. https://doi.org/10.1023/A:1013879706863.
30. 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.
31. DasSarma S, DasSarma P. 2015. Gas vesicle nanoparticles for antigen display. Vaccines (Basel) 3:686–702. https://doi.org/10.3390/vaccines3030686.
32. DasSarma P, DasSarma S. 2021. Gas vesicle nanoparticles. *In* eLS. John Wiley & Sons, Ltd, New York, NY. https://doi.org/10.1002/9780470015902.a0029044.
33. 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.