AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

Check for updates

| Environmental Microbiology | Announcement

Whole-genome sequencing of three extremophile *Bacillus* sp. strains isolated from the Atacama Desert

Nicole T. Cavanaugh,[1] Girish Kumar,[2] Matteo Couto Frignani,[1] Nathan Thewedros,[1] Marcello Twahirwa,[1] Carlos Riquelme,[3] André O. Hudson,[2] Yunrong Chai,[1] Veronica Godoy-Carter[1]

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT The Atacama Desert is home to bacteria that use biofilms as a means of protecting themselves against the harsh environment. To inform research regarding this survival mechanism, we cultured and sequenced the genomes of three *Bacillus* sp. isolates from Atacama Desert soil.

KEYWORDS extremophiles, environmental microbiology, biofilms, genomes, genomics

Downloaded from https://journals.asm.org/journal/mra on 14 October 2026 by 146.83.153.245.

The Atacama Desert is one of the most ancient, most highly irradiated, and driest deserts in the world (1). Despite this, it is home to a variety of bacterial life. Bacteria in this region are shown to have developed adaptations to survive this harsh, irradiated environment such as high levels of pigment production and biofilm formation (2). In the past, most research addressing survival mechanisms in microbes from the Atacama have focused on archaea. We are studying how eubacteria have developed adaptations to survive in such a desert environment.

To study how biofilms impact bacterial survival, we started by culturing and isolating a variety of bacterial colonies from desert soil and screening for biofilm-forming isolates. Bacterial colonies were purified after plating 100 µL of a soil water slurry on Reasoner's 2A (R2A) agar medium and incubating plates at 28°C for 48–72 h. To identify the isolates, we performed 16S rRNA gene sequencing on the V3/4 variable regions using the following primers: 5′-CCTACGGGNGGCWGCAG-3′ and 5′-GACTACHVGGGTATCTAATCC-3′. Using National Society for Biotechnology Information (NCBI) BLASTN version 2.15.0, three *Bacillus* sp. were identified (3, 4). The *Bacillus* genus contains Gram-positive bacteria that are found throughout the environment, many of which are known to produce biofilms (5–10). To test biofilm formation, strains were grown at 37°C to OD600 = 1.0 in Luria–Bertani (LB) broth, and 2 µL of culture was spotted on LB agar plates supplemented with 1% glycerol and 0.1 mM MnSO4 (LBGM, biofilm-inducing medium) (11). Colony biofilms were incubated for 48 h (Fig. 1).

These *Bacillus* sp. were selected for further genomic analysis. Genomic DNA was isolated from 2-mL LB cultures prepared from single colonies using the Wizard Genomic DNA Purification Kit (Promega). DNA libraries were prepared using the Nextera XT library preparation kit (Illumina). Libraries were sequenced using the Illumina MiSeq System using the MiSeq Reagent V3 kit (2 × 300 cycles) (Illumina) at the Genomics Lab, Rochester Institute of Technology. Raw reads were trimmed using Trimmomatic Galaxy version 0.38.1 (13). The quality of the trimmed reads was assessed using FastQC version 0.74+galaxy0 (14). Trimmed reads over 100 bp were assembled *de novo* using Unicycler version 0.5.0+galaxy1 (15). Quality analysis was performed using Quast version 5.2.0+galaxy1 (16). Quality assessments are outlined in Table 1. All programs used were run at default parameters unless otherwise noted. Analysis was performed using the Prokaryotic Genome Analysis Pipeline (PGAP) version 6.6 to reveal the read length and total number of genes, protein coding sequences, rRNAs, and tRNAs in each assembly (Table 1) (17–19).

Editor David A. Baltrus, The University of Arizona, Tucson, Arizona, USA

Address correspondence to Veronica Godoy-Carter, v.godoycarter@northeastern.edu.

The authors declare no conflict of interest .

Received 1 July 2024
Accepted 25 October 2024
Published 16 December 2024

Copyright © 2024 Cavanaugh et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Downloaded from https://journals.asm.org/journal/mra on 14 October 2026 by 146.83.153.245.

TABLE 1 Summary of quality assessments for each strain[a]

Isolate	Geographic locations	Elevation (ft)	Most closely related species	Total bases (Mbp)	Assembly size (bp)	Number of reads (FastQC)	Read length	Number of contigs	N50 (bp)	GC content	Total genes	Protein coding sequences	rRNA operons	tRNAs
0102A	22.56.56 S 68.18.53 W	8,340	*B. vallismortis*	813.2	4,192,438	4,406,089	2 × 289	27	350,970	43.87%	4,296	4,067	3	76
0209A	22.31.54 S 68.2.27 W	14,000	*B. vallismortis*	980.4	4,167,904	5,239,849	2 × 285	19	1,048,338	43.98%	4,208	3,989	3	73
0909A	22.31.54 S 68.2.27 W	14,000	*B. mojavensis*	857.1	3,899,559	4,586,376	2 × 286	19	930,776	43.66%	4,061	3,870	3	67

[a]A summary of the identities and quality assessments of each strain 0102A, 0209A, and 0909A. The latitude and longitude and elevation, where each sample was collected, are summarized. For each strain, the total bases and number of assembly size were determined using FastQC Galaxy version 0.74+galaxy0. Read length, number of contigs, N50, and GC content were assessed using Quast Galaxy version 5.2.0+galaxy1. Read length and total genes, protein coding sequences, rRNAs, and tRNAs were determined using the Prokaryotic Genome Analysis Pipeline (PGAP) version 6.6.

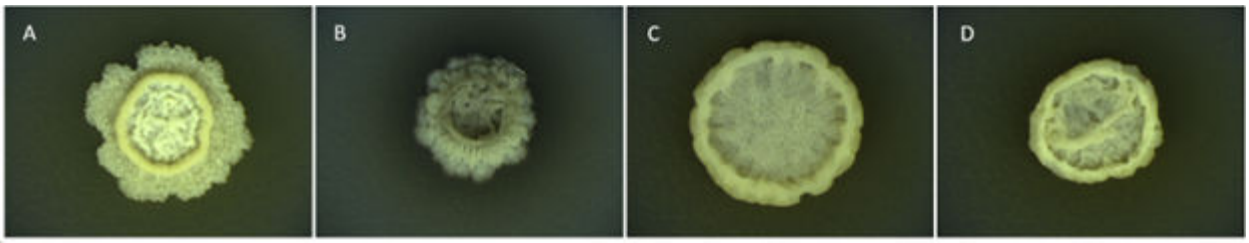


FIG 1 Environmental strains isolated from the Atacama Desert form colony biofilms. Colony biofilms of 0102A (A), 0209A (B), 0909A (C), and 3610 (12), (D) grown on LBGM medium (11) after 48 h of incubation.

Downloaded from https://journals.asm.org/journal/mra on 14 October 2026 by 146.83.153.245.

PGAP filters out reads under 200 bp before assembly and analysis. Further analysis of the *eps* operon in each genome using NCBI BLASTN version 2.14.1+galaxy and taxonomy analysis by Automated Multi Locus Species Tree (autoMLST) showed that these three isolates were most closely related to *Bacillus vallismortis and Bacillus mojavensis* (3, 4, 20).

ACKNOWLEDGMENTS

We would like to thank the National Science Foundation (MCB1651732) and the NU Global Experience Office for their support. Thank you to the RIT Genomics Lab. V.G.-C. was supported by a stipend from an HHMI Inclusive Excellence grant. M.C.F, N.T., and M.T. were supported by NU Undergraduate Research and Fellowships Office PEAK Awards. N.T.C. is supported by the NSF Graduate Research Fellowship Program (1938052).

AUTHOR AFFILIATIONS

[1]Department of Biology, Northeastern University College of Science, Boston, Massachusetts, USA
[2]Thomas H. Gosnell School of Life Sciences, Rochester Institute of Technology College of Science, Rochester, New York, USA
[3]Instituto de Bioinnovacion, Universidad de Antofagasta, Antofagasta, Chile

AUTHOR ORCIDs

André O. Hudson http://orcid.org/0000-0001-5690-4322
Yunrong Chai http://orcid.org/0000-0002-5903-5529
Veronica Godoy-Carter http://orcid.org/0000-0002-1254-9282

AUTHOR CONTRIBUTIONS

Nicole T. Cavanaugh, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review and editing | Girish Kumar, Data curation, Formal analysis, Methodology, Writing – review and editing | Matteo Couto Frignani, Data curation, Investigation, Methodology, Writing – review and editing | Nathan Thewedros, Data curation, Investigation, Methodology, Writing – review and editing | Marcello Twahirwa, Data curation, Investigation, Methodology, Writing – review and editing | Carlos Riquelme, Funding acquisition, Project administration, Resources, Supervision, Writing – review and editing | Yunrong Chai, Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review and editing | Veronica Godoy-Carter, Conceptualization, Data curation, Funding acquisition, Project administration, Resources, Supervision, Writing – review and editing.

DATA AVAILABILITY

The WGS projects for Bacillus subtilis strains 0102A, 0209A, and 0909A have been deposited in GenBank under BioProject number PRJNA1071648. Strain 0102A is

deposited under BioSample number SAMN39706723 and SRA number SRR27839954. PGAP analysis can be found under accession number JBAIZW000000000. Strain 0209A is deposited under BioSample number SAMN39706724 and SRA number SRR27839953. PGAP analysis can be found under accession number JBAIZV000000000. Strain 0909A is deposited under BioSample number SAMN39706725 and SRA number SRR27839952. PGAP analysis can be found under accession number JBAIZU000000000.
The NCBI BLASTN webserver was used to identify the 0102A, 0209A, and 0909A 16 s rRNA gene sequence. All programs found on the Galaxy platform can be found at usegalaxy.org.

REFERENCES

1. Bull AT, Andrews BA, Dorador C, Goodfellow M. 2018. Introducing the Atacama Desert. Antonie Van Leeuwenhoek 111:1269–1272. https://doi.org/10.1007/s10482-018-1100-2
2. Reverdy A, Hathaway D, Jha J, Michaels G, Sullivan J, McAdoo DD, Riquelme C, Chai Y, Godoy-Carter V. 2024. Insights into the diversity and survival strategies of soil bacterial isolates from the Atacama Desert. Front Microbiol 15:1335989. https://doi.org/10.3389/fmicb.2024.1335989
3. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389
4. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. *BLAST+*: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421
5. Arnaouteli S, Bamford NC, Stanley-Wall NR, Kovács ÁT. 2021. *Bacillus subtilis* biofilm formation and social interactions. Nat Rev Microbiol 19:600–614. https://doi.org/10.1038/s41579-021-00540-9
6. Beauregard PB, Chai Y, Vlamakis H, Losick R, Kolter R. 2013. *Bacillus subtilis* biofilm induction by plant polysaccharides. Proc Natl Acad Sci U S A 110:E1621–E1630. https://doi.org/10.1073/pnas.1218984110
7. Branda SS, González-Pastor JE, Ben-Yehuda S, Losick R, Kolter R. 2001. Fruiting body formation by *Bacillus subtilis*. Proc Natl Acad Sci 98:11621–11626. https://doi.org/10.1073/pnas.191384198
8. Diomandé SE, Nguyen-The C, Guinebretière M-H, Broussolle V, Brillard J. 2015. Role of fatty acids in *Bacillus* environmental adaptation. Front Microbiol 6:813. https://doi.org/10.3389/fmicb.2015.00813
9. Fritze D. 2004. Taxonomy of the genus *Bacillus* and related genera: the aerobic endospore-forming bacteria. Phytopathology 94:1245–1248. https://doi.org/10.1094/PHYTO.2004.94.11.1245
10. Vlamakis H, Chai Y, Beauregard P, Losick R, Kolter R. 2013. Sticking together: building a biofilm the *Bacillus subtilis* way. Nat Rev Microbiol 11:157–168. https://doi.org/10.1038/nrmicro2960
11. Shemesh M, Chai Y. 2013. A combination of glycerol and manganese promotes biofilm formation in *Bacillus subtilis* via histidine kinase KinD signaling. J Bacteriol 195:2747–2754. https://doi.org/10.1128/JB.00028-13
12. Zeigler DR, Prágai Z, Rodriguez S, Chevreux B, Muffler A, Albert T, Bai R, Wyss M, Perkins JB. 2008. The origins of 168, W23, and other Bacillus subtilis legacy strains. J Bacteriol 190:6983–6995. https://doi.org/10.1128/JB.00722-08
13. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
14. Babraham Bioinformatics—FastQC A quality control tool for high throughput sequence data. 2024. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
15. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595
16. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics 34:i142–i150. https://doi.org/10.1093/bioinformatics/bty266
17. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res 46:D851–D860. https://doi.org/10.1093/nar/gkx1068
18. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. Nucleic Acids Res 49:D1020–D1028. https://doi.org/10.1093/nar/gkaa1105
19. 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
20. Alanjary M, Steinke K, Ziemert N. 2019. AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. Nucleic Acids Res 47:W276–W282. https://doi.org/10.1093/nar/gkz282

Downloaded from https://journals.asm.org/journal/mra on 14 October 2026 by 146.83.153.245.