Title: Whole-Genome Sequencing of a *Janthinobacterium sp.* Isolated from the Patagonian Desert

Running Title: WGS of Extremophile *Janthinobacterium sp.*

Authors: Nicole T Cavanaugh[1], Girish Kumar[2], Alicyn Reverdy Pearson[1], Julia Colbert[1], Carlos Riquelme[3], André O Hudson[2], Yunrong Chai[1], Veronica Godoy-Carter[1*]

Affiliations:
[1]Northeastern University College of Science, Department of Biology, Boston, MA
[2]Rochester Institute of Technology College of Science, Thomas H. Gosnell School of Life Sciences, Rochester, NY
[3]Universidad de Antofagasta, Departamento de Biotecnología, Antofagasta, Chile
[*]Corresponding Author: v.godoycarter@northeastern.edu

Abstract:
Janthinobacterium is a genus of Gram-negative environmental bacteria that survive extreme conditions by forming biofilms and producing pigments. *Janthinobacterium* sp. LS2A, an extremophile isolated from soil in the Chilean Patagonia, contains seven known biosynthetic gene clusters, including the purple pigment violacein, which may aid in its survival in harsh environments.

Announcement:
Janthinobacterium is a genus of Gram-negative bacteria that is widespread in cold climate soil and water (1, 2, 3). Cells of *Janthinobacterium* are rod-shaped and produce a purple pigment called violacein (4). Violacein is of high biotechnological significance due to its known antiviral, antibacterial, antimycotic, antitumor, algicidal, and antioxidant activities (5, 6, 7, 8). The genus also contains extremophiles that are tolerant to cold temperatures 2-28 ºC and UV exposure (9, 10, 11). *Janthinobacterium sp*. are known to form biofilms which help them survive these harsh conditions (4).

The goal of this project was to study extremophiles in Patagonia, Chile and the mechanisms that help them survive in an extreme environment, including pigment production and biofilm formation. Bacteria were isolated after plating 100 µL of a soil/water slurry from Laguna Sarmiento on Reasoner's 2A (R2A) agar medium plates followed by incubation at 28 ºC for 48-72 hours. To test biofilm formation, LS2A was grown at 30 ºC to OD600 = 1.0 in R2A broth and 2 µL of culture was spotted on a dry R2A plate, which was incubated for 96 hours before imaging (Figure 1A). As a first approximation to identify closely related species, we performed 16S rRNA gene sequencing on the V3/4 variable regions using the following primers: 5'-CCTACGGGNGGCWGCAG-3' and 5'-GACTACHVGGGTATCTAATCC-3'. The isolate was determined to be *Janthinobacterium sp*. using National Society for Biotechnology Information (NCBI) BLASTN version 2.15.0 (12, 13). The sequence of the V3/4 16S rRNA region was used, along with a variety of 16S sequences from *Janthinobacterium* sp. that were present in the

RefSeq database, to create a multiple sequence alignment and a guide tree using ClustalOmega version 1.2.4 (Figure 1B) (14).

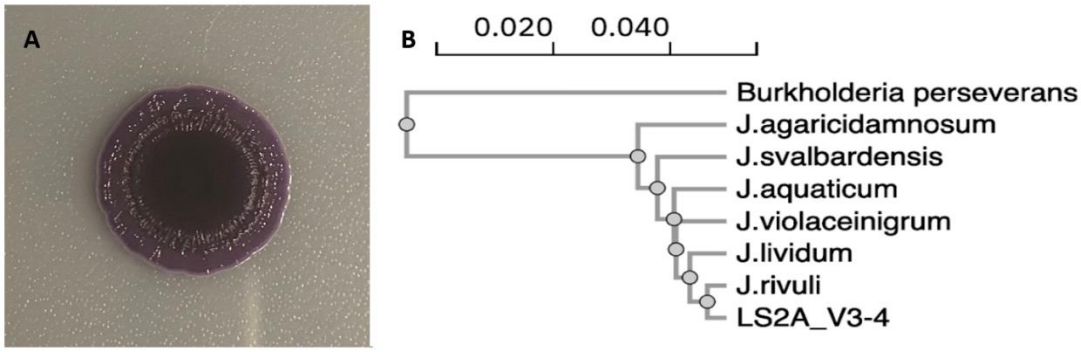

Figure 1: **Janthinobacterium sp.** ***LS2A forms biofilms and its sequence demonstrates it belongs to the genus. 1a. Biofilm grown from*** **Janthinobacterium sp.** ***LS2A. 1b. Phylogenetic Tree****. A phylogenetic tree relating the LS2A V3/4 16S rRNA regions to other* Janthinobacterium sp*. in the RefSeq database. From top to bottom:* Burkholderia preserverans *NR_179094.1,* J. agaricidamonosum *NR_114134.1,* J. svalbardensis *NR_132608.1,* J. aquaticum *NR_170539.1,* J. violaceinigrum *NR_170541.1,* J. lividum *NR_026365.1*, J.rivuli *NR_170540.1, and* Janthinobacterium sp. *LS2A.*

Genomic DNA was isolated from a 2-mL R2A liquid culture using the Wizard Genomic DNA Purification Kit (Promega, USA) following manufacturer's instructions for Gram-negative bacteria. After quantification using a NanoDrop spectrophotometer, DNA libraries were prepared using the Nextera XT library preparation kit (Illumina) and sequenced using the Illumina MiSeq System at the Genomics Lab, Rochester Institute of Technology. Raw reads were trimmed using Trimmomatic Galaxy version 0.38.1 (15). The trimmed reads were assembled *de novo* using Unicycler Galaxy version 0.5.0+galaxy1 (16). Quality analysis was performed using Quast Galaxy version 5.2.0+galaxy1 (17) and FastQC Galaxy version 0.74+galaxy0 (18). All programs listed above were run using default parameters unless stated otherwise. The assembly consists of 4,575,745 reads assembled into 29 contigs with a read length of 6,251,303 base pairs (bp) containing 667.6 Mbp total. The N50 is 540,701 bp and the GC content is 62.69%. Analysis by the Prokaryotic Genome Analysis Pipeline (PGAP) version 6.6 revealed that the LS2A genome contained 5,664 total genes, 5,548 protein coding sequences, 3 rRNAs operons, and 77 tRNAs (19-21).

The genome assembly was analyzed using antibiotics and secondary metabolite analytics shell (antiSMASH) version 7.1.0 (22). A summary of the biosynthetic gene clusters (BGC's) for putative antibiotics and secondary metabolites predicted by the software are outlined in Table 1. antiSMASH detected seven gene clusters that encode unique metabolites, including violacein.

Region	Type	Start	End	Most Similar To	Similarity Score
1.1	RiPP-like	364,750	372,021	burkholderic acid	low (<10%)
1.2	Indole	533,201	556,218	violacein	100%
1.3	RiPP-like	830,949	841,878		
1.4	Acyl Amino Acids	936,131	1,031,826	O-antigen	14%
1.5	Terpene	1,087,968	1,109,729		
3.1	RiPP-like	254,757	266,376		
7.1	Arylpolyene	1	32,414	APE Ec	36%

Table 1: Predicted BGC's in LS2A*. A summary of biosynthetic gene clusters in* Janthinobacterium *sp. LS2A identified by antiSMASH.*

Data Availability:

The WGS projects for Janthinobacterium sp. LS2A have been deposited in GenBank under BioProject number PRJNA1071647, BioSample number SAMN39706722, and SRA number SRR27839980. Results from PGAP analysis can be found using accession number JAZHPB000000. The following sequences were used to construct Figure 1 and were retrieved from GenBank: NR_179094.1, NR_114134.1, NR_132608.1, NR_170539.1, NR_170541.1, NR_026365.1, NR_170540.1,

The NCBI BLASTN webserver was used to identify the LS2A 16s rRNA gene sequence. All programs found on the Galaxy platform can be found at usegalaxy.org. Clustal Omega can be found on the EMBL European Bioinformatics Institute webpage.

Acknowledgements:

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. was supported by an HHMI Inclusive Excellence grant and the Northeastern Global Experiential Office. J.C. was supported by NU Undergraduate Research and Fellowships Office PEAK Awards. N.C. was supported by the NSF Graduate Research Fellowship Program (1938052). A.R.P. was supported by the Northeastern University Provost Dissertation Completion Fellowship.

References:

[1] Chernogor, L., Bakhvalova, K., Belikova, A., & Belikov, S. (2022). Isolation and Properties of the Bacterial Strain Janthinobacterium sp. SLB01. *Microorganisms*, *10*(5), 1071. https://doi.org/10.3390/microorganisms10051071

[2] Gong, X., Skrivergaard, S., Korsgaard, B. S., Schreiber, L., Marshall, I. P. G., Finster, K., & Schramm, A. (2017). High quality draft genome sequence of Janthinobacterium psychrotolerans sp. Nov., isolated from a frozen freshwater pond. *Standards in Genomic Sciences*, *12*(1), 8. https://doi.org/10.1186/s40793-017-0230-x

[3] Schloss, P. D., Allen, H. K., Klimowicz, A. K., Mlot, C., Gross, J. A., Savengsuksa, S., McEllin, J., Clardy, J., Ruess, R. W., & Handelsman, J. (2010). Psychrotrophic Strain of Janthinobacterium lividum from a Cold Alaskan Soil Produces Prodigiosin. *DNA and Cell Biology*, *29*(9), 533–541. https://doi.org/10.1089/dna.2010.1020

[4] Pantanella, F., Berlutti, F., Passariello, C., Sarli, S., Morea, C., & Schippa, S. (2007). Violacein and biofilm production in Janthinobacterium lividum. *Journal of Applied Microbiology*, *102*(4), 992–999. https://doi.org/10.1111/j.1365-2672.2006.03155.x

[5] Andrighetti-Fröhner, C. R., Antonio, R. V., Creczynski-Pasa, T. B., Barardi, C. R. M., & Simões, C. M. O. (2003). Cytotoxicity and potential antiviral evaluation of violacein produced by Chromobacterium violaceum. *Memórias Do Instituto Oswaldo Cruz*, *98*, 843–848. https://doi.org/10.1590/S0074-02762003000600023

[6] Konzen, M., De Marco, D., Cordova, C. A. S., Vieira, T. O., Antônio, R. V., & Creczynski-Pasa, T. B. (2006). Antioxidant properties of violacein: Possible relation on its biological function. *Bioorganic & Medicinal Chemistry*, *14*(24), 8307–8313. https://doi.org/10.1016/j.bmc.2006.09.013

[7] Durán, N., Justo, G. Z., Ferreira, C. V., Melo, P. S., Cordi, L., & Martins, D. (2007). Violacein: Properties and biological activities. *Biotechnology and Applied Biochemistry*, *48*(3), 127–133. https://doi.org/10.1042/BA20070115

[8] Cai, G., Yang, X., Yu, X., Zheng, W., Cai, R., & Wang, H. (2024). The novel application of violacein produced by a marine *Duganella* strain as a promising agent for controlling *Heterosigma akashiwo* bloom: Algicidal mechanism, fermentation optimization and agent

formulation. *Journal of Hazardous Materials*, *466*, 133548. https://doi.org/10.1016/j.jhazmat.2024.133548

[9] Yang, M., Lu, D., Qin, B., Liu, Q., Zhao, Y., Liu, H., & Ma, J. (2018). Highly efficient nitrogen removal of a coldness-resistant and low nutrient needed bacterium, Janthinobacterium sp. M-11. *Bioresource Technology*, *256*, 366–373. https://doi.org/10.1016/j.biortech.2018.02.049

[10] Mukhia, S., Kumar, A., Kumari, P., & Kumar, R. (2022). Psychrotrophic plant beneficial bacteria from the glacial ecosystem of Sikkim Himalaya: Genomic evidence for the cold adaptation and plant growth promotion. *Microbiological Research*, *260*, 127049. https://doi.org/10.1016/j.micres.2022.127049

[11] Mojib, N., Farhoomand, A., Andersen, D. T., & Bej, A. K. (2013). UV and cold tolerance of a pigment-producing Antarctic Janthinobacterium sp. Ant5-2. *Extremophiles*, *17*(3), 367–378. https://doi.org/10.1007/s00792-013-0525-9

[12] Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389–3402. https://doi.org/10.1093/nar/25.17.3389

[13] Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

[14] Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*, *50*(W1), W276–W279. https://doi.org/10.1093/nar/gkac240

[15] Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

[16] Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology*, *13*(6), e1005595. https://doi.org/10.1371/journal.pcbi.1005595

[17] Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., & Gurevich, A. (2018). Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*, *34*(13), i142–i150. https://doi.org/10.1093/bioinformatics/bty266

[18] *Babraham Bioinformatics—FastQC A Quality Control tool for High Throughput Sequence Data*. (n.d.). Retrieved February 2, 2024, from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

[19] Haft, D. H., DiCuccio, M., Badretdin, A., Brover, V., Chetvernin, V., O'Neill, K., Li, W., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Zheng, C., Thibaud-Nissen, F., Geer, L. Y., … Pruitt, K. D. (2018). RefSeq: An update on prokaryotic genome annotation and curation. *Nucleic Acids Research*, *46*(D1), D851–D860. https://doi.org/10.1093/nar/gkx1068

[20] Li, W., O'Neill, K. R., Haft, D. H., DiCuccio, M., Chetvernin, V., Badretdin, A., Coulouris, G., Chitsaz, F., Derbyshire, M. K., Durkin, A. S., Gonzales, N. R., Gwadz, M., Lanczycki, C. J., Song, J. S., Thanki, N., Wang, J., Yamashita, R. A., Yang, M., Zheng, C., … Thibaud-Nissen, F. (2021). RefSeq: Expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. *Nucleic Acids Research*, *49*(D1), D1020–D1028. https://doi.org/10.1093/nar/gkaa1105

[21] Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., Lomsadze, A., Pruitt, K. D., Borodovsky, M., & Ostell, J. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research*, *44*(14), 6614–6624. https://doi.org/10.1093/nar/gkw569

[22] Blin, K., Shaw, S., Augustijn, H. E., Reitz, Z. L., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B. R., Metcalf, W. W., Helfrich, E. J. N., van Wezel, G. P., Medema, M. H., & Weber, T. (2023).

antiSMASH 7.0: New and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Research*, *51*(W1), W46–W50. https://doi.org/10.1093/nar/gkad344