9-1-2021

Draft Genome Sequences of Staphylococcus sp. Strain CWZ226, of Unknown Origin, and Pseudomonas sp. Strain CVAP#3, Antagonistic to Strain CWZ226

Brianna L. McCall
Smith College

J. A.C. Vriezen
Smith College, jvriezen@smith.edu

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequences of *Staphylococcus* sp. Strain CWZ226, of Unknown Origin, and *Pseudomonas* sp. Strain CVAP#3, Antagonistic to Strain CWZ226

Brianna L. McCall,[a]* J. A. C. Vriezen[a]

[a]Smith College, Department of Biological Sciences, Northampton, Massachusetts, USA

ABSTRACT Many *Staphylococcus* and *Pseudomonas* species, such as *Staphylococcus aureus* and *Pseudomonas aeruginosa*, are opportunistic human pathogens. However, *Pseudomonas* species are also known to produce bioactive compounds. Here, we report on the genome sequences of a *Pseudomonas* isolate and a *Staphylococcus* species of unknown origin that it inhibits.

Many non-*aureus* or non-*aeruginosa* *Staphylococcus* and *Pseudomonas* species are known emerging pathogens (1–6). However, *Pseudomonas* strains are also known to produce toxic compounds such as cyclic lipopeptides (7, 8).

A bacterial colony (CVAP#3) was isolated from a soil sample (sampled at 42°19′04.01″N, 72°38′25.60″W, at an altitude of 41.76 m, on 21 November 2014) that had been suspended in 1× phosphate-buffered saline (PBS), plated onto 10% tryptic soy agar (TSA), and incubated for 48 h at 25°C. CVAP#3 inhibited a lawn of *Staphylococcus* (CWZ226) on TSA at 25°C overnight. CWZ226 is of unknown origin and is maintained in a culture collection at Smith College.

In order (i) to determine the phylogenetic position of strains CVAP#3 and CWZ226 and (ii) to identify loci for bioactive compounds against *Staphylococcus*, the genomes of both strains were sequenced. To isolate DNA, the Promega Wizard genomic DNA kit was employed (catalog number A1120). The Nextera DNA Flex library preparation kit (catalog number 20018704) was used for library construction. Sequencing was performed at Smith College using the MiSeq reagent kit (cartridge) v3 for 150 cycles (catalog number MS-102-3001) or 600 cycles (catalog number MS-102-3003); this resulted in 5,050,320 and 956,980 raw reads for the paired-end reads of 75-bp and 300-bp sequences, respectively, for the CVAP#3 genome. Similarly, two sequencing runs for the *Staphylococcus* genome resulted in 2,386,984 and 2,600,230 paired-end reads of 300 bp.

Both sets of paired-end reads were assembled in SPAdes v3.11.1 (9) in careful mode with the -kmer set at 73 (CVAP#3) or default (CWZ226) and a Phred offset of 33. Contigs smaller than 700 bp were removed. Metagenome binning in PATRIC v3.6.9 (10) resulted in 45 contigs and 31 contigs in one bin with 33.1× and 506× coverage for CVAP#3 and CWZ226, respectively. Comprehensive genome analysis and annotation of the CVAP#3 genome in PATRIC resulted in a total genome length of 6,255,897 bp, an N_{50} value of 301,990 bp, and an L_{50} value of 7. Completeness is 99.1%, the GC content is 58.7%, and a total of 58 tRNAs, 1 rRNA, 1,395 hypothetical proteins, and 4,510 known proteins were found. The analysis of the CWZ226 genome resulted in a total length of 2,557,926 bp, an N_{50} value of 209,069 bp, and an L_{50} value of 5. Completeness is 99.9%, the GC content is 36.1%, and a total of 57 tRNAs, 6 rRNAs, 520 hypothetical proteins, and 1,854 known proteins were found.

The 16S sequences were used as queries in the NCBI genomic database (11). The highest level of identity was with *Pseudomonas fluorescens* (99.9%) (12) and *Staphylococcus simulans* (99.7%) (4) for strains CVAP#3 and CWZ226, respectively. Strict analysis of the *Pseudomonas* genome with antiSMASH v6.0.0 (13) resulted in seven gene clusters putatively involved in the production of secondary metabolites. Five of these gene clusters are putatively involved in the

Citation McCall BL, Vriezen JAC. 2021. Draft genome sequences of *Staphylococcus* sp. strain CWZ226, of unknown origin, and *Pseudomonas* sp. strain CVAP#3, antagonistic to strain CWZ226. Microbiol Resour Announc 10:e00688-21. https://doi.org/10.1128/MRA.00688-21.

Editor Catherine Putonti, Loyola University Chicago

Address correspondence to J. A. C. Vriezen, jvriezen@smith.edu.

* Present address: Brianna L. McCall, Sound Fertility Care, Seattle, Washington, USA.

Received 6 July 2021
Accepted 19 August 2021
Published 16 September 2021

production of fengycin and two different siderophores, pseudomonine and pseudopyronine, compounds with antimicrobial and anticancer activity. In addition, two of these gene clusters code for bacteriocins and have domains associated with alkaline phosphatases and endonuclease (DUF692 domain).

Data availability. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank; the accession numbers are JAHMSC000000000 for *Pseudomonas* sp. strain CVAP#3 and JAHMSB000000000 for *Staphylococcus* sp. strain CWZ226. All raw sequencing files can be found in the SRA under accession number SRP323427. The strains used (CVAP#3, MC4100, and CWZ226) can be requested by email (jvriezen@smith.edu).

ACKNOWLEDGMENTS

This study was supported by funding from the Office of Faculty Development and Course Development at Smith College. The funders had no role in the study design, data collection and interpretation, or the decision to submit the work for publication.

Special thanks go to Sarah Sutto-Plunz for isolating the bacterium from soil, Lori Saunders for library construction, Louie Bierwert and Riccardo Racicot for sequencing, Louie Bierwert and Robin Sleith for productive conversations and advice, and Rebecca Wattam at the Biocomplexity Institute at Virginia Tech for assistance with the use of PATRIC. Special thanks go to Christine White Ziegler for providing the infrastructure and intellectual input; without it, this work would not have been done. Thanks also go to Rachel Wright for proofreading. In addition, we thank many students in the courses BIO130 and BIO205 at Smith College.

REFERENCES

1. Aslam B, Wang W, Arshad MI, Khurshid M, Muzammil S, Rasool MH, Nisar MA, Alvi RF, Aslam MA, Qamar MU, Salamat MKF, Baloch Z. 2018. Antibiotic resistance: a rundown of a global crisis. Infect Drug Resist 11:1645–1658. https://doi.org/10.2147/IDR.S173867.
2. Cole K, Atkins B, Llewelyn M, Paul J. 2021. Genomic investigation of clinically significant coagulase-negative staphylococci. J Med Microbiol https://doi.org/10.1099/jmm.0.001337.
3. Goda K, Kenzaka T, Hoshijima M, Yachie A, Akita H. 2021. Toxic shock syndrome with a cytokine storm caused by *Staphylococcus simulans*: a case report. BMC Infect Dis 21:19. https://doi.org/10.1186/s12879-020-05731-y.
4. Naushad S, Barkema HW, Luby C, Condas LAZ, Nobrega DB, Carson DA, De Buck J. 2016. Comprehensive phylogenetic analysis of bovine non-*aureus* staphylococci species based on whole-genome sequencing. Front Microbiol 7:1990. https://doi.org/10.3389/fmicb.2016.01990.
5. Preda M, Mihai MM, Popa LI, Dițu LM, Holban AM, Manolescu LSC, Popa GL, Muntean AA, Gheorghe I, Chifiriuc CM, Popa MI. 2021. Phenotypic and genotypic virulence features of staphylococcal strains isolated from difficult-to-treat skin and soft tissue infections. PLoS One 16:e0246478. https://doi.org/10.1371/journal.pone.0246478.
6. Shields BE, Tschetter AJ, Wanat KA. 2016. *Staphylococcus simulans*: an emerging cutaneous pathogen. JAAD Case Rep 2:428–429. https://doi.org/10.1016/j.jdcr.2016.08.015.
7. de Bruijn I, de Kock MJD, Yang M, de Waard P, van Beek TA, Raaijmakers JM. 2007. Genome-based discovery, structure prediction and functional analysis of cyclic lipopeptide antibiotics in *Pseudomonas* species: cyclic lipopeptides in *Pseudomonas*. Mol Microbiol 63:417–428. https://doi.org/10.1111/j.1365-2958.2006.05525.x.
8. Scales BS, Dickson RP, LiPuma JJ, Huffnagle GB. 2014. Microbiology, genomics, and clinical significance of the *Pseudomonas fluorescens* species complex, an unappreciated colonizer of humans. Clin Microbiol Rev 27:927–948. https://doi.org/10.1128/CMR.00044-14.
9. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
10. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg EK, Olsen GJ, Murphy-Olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Vonstein V, Warren A, Xia F, Yoo H, Stevens RL. 2017. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. Nucleic Acids Res 45:D535–D542. https://doi.org/10.1093/nar/gkw1017.
11. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
12. Price MN, Wetmore KM, Waters RJ, Callaghan M, Ray J, Liu H, Kuehl JV, Melnyk RA, Lamson JS, Suh Y, Carlson HK, Esquivel Z, Sadeeshkumar H, Chakraborty R, Zane GM, Rubin BE, Wall JD, Visel A, Bristow J, Blow MJ, Arkin AP, Deutschbauer AM. 2018. Mutant phenotypes for thousands of bacterial genes of unknown function. Nature 557:503–509. https://doi.org/10.1038/s41586-018-0124-0.
13. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. 2021. AntiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 49:W29–W35. https://doi.org/10.1093/nar/gkab335.