AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Complete Genome Sequence of *Vibrio* sp. Strain 2521-89, a Close Relative of *Vibrio cholerae* Isolated from Lake Water in New Mexico, USA

Kevin Liang,[a] Fabini D. Orata,[a] Nathan S. Winkjer,[b] Lori A. Rowe,[b] Cheryl L. Tarr,[b] Yan Boucher[a]

Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada[a]; National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, USA[b]

ABSTRACT *Vibrio* sp. strain 2521-89 is an environmental isolate from lake water in New Mexico, USA. Average nucleotide identity, *in silico* DNA-DNA hybridization, and core genome single-nucleotide polymorphism (SNP)-based phylogenetic analysis suggest that this may be a potentially novel species that is closely related to *Vibrio cholerae*.

Received 20 July 2017 **Accepted** 26 July 2017 **Published** 31 August 2017

Citation Liang K, Orata FD, Winkjer NS, Rowe LA, Tarr CL, Boucher Y. 2017. Complete genome sequence of *Vibrio* sp. strain 2521-89, a close relative of *Vibrio cholerae* isolated from lake water in New Mexico, USA. Genome Announc 5:e00905-17. https://doi.org/10.1128/genomeA.00905-17.

Address correspondence to Yan Boucher, yboucher@ualberta.ca.

The *Vibrio* genus contains over 140 different species, including *Vibrio cholerae*, the causative agent of cholera. This species alone harbors over 200 different serogroups based on the O antigen, several of which are associated with human infections (1). Here, we describe a potentially novel *Vibrio* species, *Vibrio* sp. strain 2521-89, which is a close relative of *Vibrio cholerae* but is markedly divergent from it.

Genomic DNA was extracted using an ArchivePure DNA cell/tissue kit (5 PRIME) according to the manufacturer's instructions. Sequencing libraries were prepared using the RSII SMRTbell template preparation kit 1.0 (PacBio) and sequenced with the P6v2 single-molecule real-time (SMRT) sequencing platform (PacBio). The complete genome has a total size of 4,123,863 bp and consists of two chromosomes (3,017,468 bp and 1,106,395 bp). The DNA G+C content was determined to be 47.2 mol%. Genome annotation was done using RAST 2.0 (2) and the Prokaryotic Genome Annotation Pipeline (PGAP) (https://www.ncbi.nlm.nih.gov/genome/annotation_prok/). PGAP annotations identified 3,814 coding sequences, 31 rRNA sequences (11 5S rRNA, 10 16S rRNA, and 10 23S rRNA sequences), and 104 tRNA sequences. *Vibrio* sp. 2521-89 contains genes encoding the transcriptional activator ToxR protein, one of the major regulators of virulence genes in *V. cholerae* (3), but it lacks both genes encoding the cholera toxin and the toxin-coregulated pilus cluster.

Species identity was determined by calculating average nucleotide identity (ANI) and *in silico* digital DNA-DNA hybridization (dDDH) between *Vibrio* sp. 2521-89 and *V. cholerae* N16961, *Vibrio metoecus* OP3H, or *Vibrio mimicus* CAIM 602 using JSpecies 1.2.1 (4) and GGDC 2.1 (5), respectively. *Vibrio* sp. 2521-89 has an ANI value of 95.1% with *V. cholerae* N16961 and values ranging from 84.8% to 87.9% with other closely related *Vibrio* species. Its dDDH value with *V. cholerae* N16961 is 61.8% (58.9% to 64.6% range, with 95% confidence interval), compared to less than 32.6% with other *Vibrio* species. Despite the ANI value being borderline to the 95% species cutoff (6), a dDDH value well below the 70% species cutoff (7) suggests that *Vibrio* sp. 2521-89 is potentially a novel species. Parsnp 1.2 (8) was used to perform core-genome single-nucleotide polymorphism (SNP)-based phylogenetic analysis of *Vibrio* sp. 2521-89 and closely related vibrios. This analysis showed the novel strain grouping strongly in a

monophyletic cluster with *V. cholerae* isolates, albeit displaying an unusually long branch length, possibly representing a speciation event.

Accession number(s). The complete genome sequence of *Vibrio* sp. 2521-89 has been deposited in GenBank under the accession numbers CP022353 (chromosome I) and CP022352 (chromosome II). The versions described in this paper are the first versions, CP022353.1 and CP022352.1, respectively.

ACKNOWLEDGMENTS

We thank Paul Kirchberger (University of Alberta) for assistance with the core genome phylogeny. We also acknowledge technical assistance from Dhwani Batra and Monica Santovenia (Centers for Disease Control and Prevention).

This work was supported by the Natural Sciences and Engineering Research Council of Canada, the Integrated Microbial Biodiversity program of the Canadian Institute for Advanced Research (to Y.B.), graduate student scholarships from Alberta Innovates–Technology Futures and the Bank of Montreal Financial Group (to F.D.O.), and federal appropriations to the Centers for Disease Control and Prevention through the Advanced Molecular Detection Initiative (to N.S.W., L.A.R., and C.L.T.).

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

REFERENCES

1. Safa A, Nair GB, Kong RYC. 2010. Evolution of new variants of *Vibrio cholerae* O1. Trends Microbiol 18:46–54. https://doi.org/10.1016/j.tim.2009.10.003.
2. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.
3. DiRita VJ, Parsot C, Jander G, Mekalanos JJ. 1991. Regulatory cascade controls virulence in *Vibrio cholerae*. Proc Natl Acad Sci U S A 88:5403–5407. https://doi.org/10.1073/pnas.88.12.5403.
4. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 106:19126–19131. https://doi.org/10.1073/pnas.0906412106.
5. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60. https://doi.org/10.1186/1471-2105-14-60.
6. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81–91. https://doi.org/10.1099/ijs.0.64483-0.
7. Colston SM, Fullmer MS, Beka L, Lamy B, Gogarten JP, Graf J. 2014. Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test case. mBio 5:e02136-14. https://doi.org/10.1128/mBio.02136-14.
8. Treangen TJ, Ondov BD, Koren S, Phillippy AM. 2014. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol 15:524. https://doi.org/10.1186/s13059-014-0524-x.