Check for updates

Draft Genome Sequence of "*Candidatus* Spirobacillus cienkowskii," a Pathogen of Freshwater *Daphnia* Species, Reconstructed from Hemolymph Metagenomic Reads

Luana Bresciani,[a,d] Leandro N. Lemos,[b] Nina Wale,[c] Jonathan Y. Lin,[d] Alexander T. Strauss,[e*] Meghan A. Duffy,[c] Jorge L. M. Rodrigues[d,f]

[a]Soil Microbiology Laboratory, Department of Soil Science, Luiz de Queiroz College of Agriculture, University of São Paulo, Piracicaba, São Paulo, Brazil

[b]Cell and Molecular Biology Laboratory, Center for Nuclear Energy in Agriculture, University of São Paulo, Piracicaba, São Paulo, Brazil

[c]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, USA

[d]Department of Land, Air, and Water Resources, University of California, Davis, California, USA

[e]Department of Biology, Indiana University, Bloomington, Indiana, USA

[f]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

ABSTRACT We report here the near-complete genome sequence of "*Candidatus* Spirobacillus cienkowskii," a spiral-shaped, red-pigmented uncultivated bacterial pathogen of *Daphnia* spp. The genome is 2.74 Mbp in size, has a GC content of 32.1%, and contains genes associated with bacterial motility and the production of carotenoids, which could explain the distinctive red color of hosts infected with this pathogen.

The species "*Candidatus* Spirobacillus cienkowskii" is a deep-branching uncultivated *Deltaproteobacteria* pathogen of freshwater daphniids, which are important members of aquatic food webs (1). The most distinctive phenotypic characteristic associated with its infection is the red color in the host hemolymph (2). We used high-throughput sequencing and computational binning approaches to assemble and reconstruct the pathogen genome, previously described only through sequences of 16S rRNA and DNA primase β-subunit (*gyrB*) genes (3).

Daphnia dentifera organisms infected with "*Candidatus* Spirobacillus cienkowskii" were originally collected from Dogwood Lake (38°32′37″N, 87°03′04″W; Greene-Sullivan State Forest, IN). Infections were propagated in *D. dentifera* L6D9 and the "standard" genotype collected from Dogwood Lake and a lake in Barry County, Michigan, respectively. The hemolymph from 43 infected hosts was collected and pooled for DNA extraction using the QIAamp DNA minikit (Qiagen, Germantown, MD) following manufacturer instructions. Metagenomic reads were generated with the Illumina MiSeq platform using paired-end 350-bp sequencing, and a total of 3,257,849 paired-end reads were obtained. Low-quality reads (<100 bp and a Phred score of <30) were filtered using Cutadapt v.1.18 (4), and the genome was assembled using IDBA-UD v.1.1.1 (5). A binning strategy was used to reconstruct the genome by taking into consideration the GC content and coverage of clustering contigs into individual genome populations through MaxBin v.2.0 (6) and manual curation with CheckM v.1.0.5 (7). To validate the genome reconstruction, BLASTN v.2.7.1+ (8) was used to compare nucleotide coding sequences (CDSs) to sequences of "*Candidatus* Spirobacillus cienkowskii" 16S rRNA (GenBank accession number EU220836) and *gyrB* (EU220837) genes deposited in GenBank.

Functional annotation was performed using PATRIC v.3.5.25 (9), Prokka v.1.12 (10), and KAAS v.1.0 (11). The "*Candidatus* Spirobacillus cienkowskii" genome was assembled into 126 contigs (2,739,001 bp with a GC content of 32.1%) with an N_{50} value of 39,228 bp. Quality control of the genome assembly indicated a near-complete genome (91.2% completeness) without contamination (0%). Comparative analysis with available 16S rRNA and *gyrB* genes for "*Candidatus* Spirobacillus cienkowskii" showed 100% identities with our genome bin. A total of 2,553 CDSs, 37 tRNAs, and 1 complete rRNA operon (23S, 16S, and 5S rRNA) were detected in the genome. Of the total number of proteins, 37% had functional assignments (961 proteins), 15% had gene ontology assignments (402), and 46% had FIGfam assignments (1,176). Several genes associated with phytoene production (terpenoid backbone biosynthesis), the colorless precursor of all C_{40} carotenoids (12), and other genes associated with carotenoid synthesis were detected (13). We identified genes associated with flagellar biosynthesis and assembly, which may be used for movement into the host hemolymph or for facilitating transmission to a new host.

Daphnia spp. are key members of lake food webs (14), and pathogen outbreaks reduce the host population growth rate and density and elevate the death rate (15). The "*Candidatus* Spirobacillus cienkowskii" genome will increase our knowledge of host-pathogen interactions. The annotated genome will help microbiologists identify conditions for isolating this ecologically important but as yet uncultivated pathogen.

Data availability. The whole-genome shotgun project of "*Candidatus* Spirobacillus cienkowskii" has been deposited at DDBJ/ENA/GenBank under the accession number QOVW00000000. The version described in this paper is version QOVW01000000. The raw reads were deposited in the Sequence Read Archive (SRA) under the accession number PRJNA450308.

ACKNOWLEDGMENTS

REFERENCES

1. Tessier AJ, Woodruff P. 2002. Cryptic trophic cascade along a gradient of lake size. Ecology 83:1263–1270. https://doi.org/10.1890/0012-9658(2002)083[1263:CTCAAG]2.0.CO;2.
2. Green J. 1959. Carotenoid pigment in Spirobacillus cienkowskii Metchnikoff, a pathogen of *Cladocera*. Nature 183:56–57. https://doi.org/10.1038/183056a0.
3. Rodrigues JLM, Duffy MA, Tessier AJ, Ebert D, Mouton L, Schmidt TM. 2008. Phylogenetic characterization and prevalence of "Spirobacillus cienkowskii," a red-pigmented, spiral-shaped bacterial pathogen of freshwater Daphnia species. Appl Environ Microbiol 74:1575–1582. https://doi.org/10.1128/AEM.02438-07.
4. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.
5. Peng Y, Leung HCM, Yiu SM, Chin FYL. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 28:1420–1428. https://doi.org/10.1093/bioinformatics/bts174.
6. Wu YW, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607. https://doi.org/10.1093/bioinformatics/btv638.
7. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.
8. 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.
9. Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJ, Yoo HS, Zhang C, Zhang Y, Sobral BW. 2014. PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res 42:D581–D591. https://doi.org/10.1093/nar/gkt1099.
10. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
11. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35:W182–W185. https://doi.org/10.1093/nar/gkm321.
12. Schweiggert RM, Carle R. 2016. Carotenoid production by bacteria, microalgae, and fungi, pp 217–240. *In* Kaczor A, Baranska M (ed), Carotenoids: nutrition, analysis and technology. John Wiley & Sons, Chichester, United Kingdom.
13. Guan Z, Xue D, Abdallah II, Dijkshoorn L, Setroikromo R, Lv G, Quax WJ. 2015. Metabolic engineering of Bacillus subtilis for terpenoid production. Appl Microbiol Biotechnol 99:9395–9406. https://doi.org/10.1007/s00253-015-6950-1.
14. Ebert D. 2011. A genome for the environment. Science 331:539–540. https://doi.org/10.1126/science.1202092.
15. Duffy MA, Hall SR. 2008. Selective predation and rapid evolution can jointly dampen effects of virulent parasites on *Daphnia* populations. Am Nat 171:499–510. https://doi.org/10.1086/528998.