

First Draft Genome Sequence from a Member of the Genus *Agrococcus*, Isolated from Modern Microbialites

Richard Allen White III,^a Christopher J. Grassa,^b Curtis A. Suttle^{a,b,c,d}

Department of Microbiology & Immunology, University of British Columbia, Vancouver, British Columbia, Canada^a; Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada^b; Department of Earth, Ocean & Atmospheric Science, University of British Columbia, Vancouver, British Columbia, Canada^c; Canadian Institute for Advanced Research^d

† For this virtual institution, see <http://www.cifar.ca/>.

We report the first draft genome sequence from a member of the genus *Agrococcus*, isolated from cold thrombolitic microbialites within Pavilion Lake, British Columbia, Canada. The draft genome assembly for *Agrococcus pavilionensis* strain RW-1 has a size of 2,878,403 bp with a G+C content of 72.56%.

Received 7 May 2013 Accepted 21 May 2013 Published 27 June 2013

Citation White RA, III, Grassa CJ, Suttle CA. 2013. First draft genome sequence from a member of the genus *Agrococcus*, isolated from modern microbialites. *Genome Announc* 1(4):e00391-13. doi:10.1128/genomeA.00391-13.

Copyright © 2013 White et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Curtis A. Suttle, suttle@science.ubc.ca.

Groth et al. (1) first described the genus *Agrococcus* based on two strains of *Agrococcus jenensis* isolated from soil and the surface of sandstone. The genus *Agrococcus* is classified within the family *Microbacteriaceae*, members of which exhibit the unusual feature of having diamminobutyric acid within the cell wall, which imparts a distinctive yellow color (1). The genus has eight described species that have been isolated from a wide range of environments, including air (2), a coal mine (3), cheese (4), cold desert soil (5), forest soil (6), a medieval wall painting (7), dried seaweed (8), and the phyllosphere of potato plants (9). The genetic richness and diversity of this genus is poorly known, with little genetic information available other than for the 16S rRNA genes.

The source of *Agrococcus pavilionensis* strain RW-1 is the oligotrophic (3.3 $\mu\text{g liter}^{-1}$ total phosphorus) Pavilion Lake (50.86677°N, 121.74191°W), which lies in Marble Canyon near Lillooet, British Columbia, Canada. Pavilion Lake harbors a diverse array of modern microbialites, which are contemporary biogenically derived carbonate structures (10, 11). Pavilion Lake microbialites consist mainly of clotted and nonlayered thrombolitic structures (10) that occur in the permanently cold (4 to 8°C) water deeper than 5 m (11). *A. pavilionensis* was isolated from a cabbage-shaped thrombolite collected at a depth of 20 m. DNA was extracted using Qiagen QIAamp followed by Qiagen MinElute cleanup columns. The Illumina MiSeq library was constructed using the Lucigen NxSeq library prep kit without final PCR enrichment.

Whole-genome shotgun sequencing was completed using Illumina MiSeq in the 250-bp paired-read format. A partial flow cell obtained 2.89 million raw reads with 713,936,519 bp of raw sequence. Paired reads were error corrected and connected using AllPaths-LG (version 44837) (12). In the data set, 31mers were counted using Jellyfish (version 1.1.10) (13). Reads that contained 31mers with a multiplicity of >1,250 were partitioned for *de novo* assembly. The partitioned reads were assembled using Celera as-

sembler 7.0 (14). The high-copy reads assembled as a single contig of 1,427 bp in length representing a high-copy plasmid. The remaining reads assembled as 50 contigs summing 2,878,403 bp (N_{50} length, 133,224; N_{90} length, 31,609; G+C content, 72.56%). The 16S rRNA gene sequence was confirmed by Sanger sequencing and was found to have 99.99% identity to the 16S rRNA gene predicted from the draft genome.

Annotation was conducted on the RAST server using the Glimmer 3 option (15) and predicted 2,506 protein-coding genes, including 48 noncoding RNA genes and 126 predicted SEED subsystem features. The potential to metabolize a wide range of carbon compounds is predicted from the genome, including D-ribose, fructose, lactate, glycerate, chitin, deoxyribose, and deoxynucleoside catabolism. Genes related to those encoding the phosphate (Pho) regulon for high-affinity uptake of phosphate and cold shock proteins were also found.

Further analysis of the genome, including functional and biochemical measurements, will be used to understand the possible roles of *A. pavilionensis* in the highly diverse microbial community contained within the Pavilion Lake microbialites. This is the first draft genome for the genus *Agrococcus*, which will provide a template for many further phylogenetic, comparative genomic, metagenomic, and functional studies of this widely distributed genus.

Nucleotide sequence accession numbers. This Whole-Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [ASHR00000000](https://www.ncbi.nlm.nih.gov/nuccore/ASHR00000000). The version described in this paper is version [ASHR01000000](https://www.ncbi.nlm.nih.gov/nuccore/ASHR01000000).

ACKNOWLEDGMENTS

We thank Sugandha Dandekar (Uma) and Hemani Wijesuriya at the UCLA Sequencing & Genotyping Core for the excellent sequencing data and Donnie Reid and the divers and boat operators from the Pavilion Lake Research Project who collected the sample from which the bacterial strain was isolated.

Financial support was provided by the MARS LIFE Project (9F052-10-

0176), funded by the Canadian Space Agency, and the Natural Science and Engineering Council of Canada.

REFERENCES

- Groth I, Schumann P, Weiss N, Martin K, Rainey FA. 1996. *Agrococcus jenensis* gen. nov., sp. nov., a new genus of actinomycetes with diaminobutyric acid in the cell wall. *Int. J. Syst. Bacteriol.* 46:234–239.
- Zlamala C, Schumann P, Kämpfer P, Rosselló-Mora R, Lubitz W, Busse HJ. 2002. *Agrococcus baldri* sp. nov., isolated from the air in the “Virgilkapelle” in Vienna. *Int. J. Syst. Evol. Microbiol.* 52:1211–1216.
- Dhanjal A, Kaur I, Suresh K, Schumann P, Cameotra SS, Pukall R, Klenk H, Mayilraj S. 2011. *Agrococcus carbonis* sp. nov., isolated from soil of a coal mine. *Int. J. Syst. Bacteriol.* 61:1253–1258.
- Bora N, Vancanneyt M, Gelsomino R, Swings J, Brennan N, Cogan TM, Larpin S, Desmaures N, Lechner FE, Kroppenstedt RM, Ward AC, Goodfellow M. 2007. *Agrococcus casei* sp. nov., isolated from the surfaces of smear-ripened cheeses. *Int. J. Syst. Evol. Microbiol.* 57:92–97.
- Mayilraj S, Suresh K, Schumann P, Kroppenstedt RM, Saini HS. 2006. *Agrococcus lahaulensis* sp. nov., isolated from a cold desert of the Indian Himalayas. *Int. J. Syst. Evol. Microbiol.* 56:1807–1810.
- Zhang JY, Liu XY, Liu SJ. 2010. *Agrococcus terreus* sp. nov. and *Micrococcus terreus* sp. nov., isolated from forest soil. *Int. J. Syst. Evol. Microbiol.* 60:1897–1903.
- Wieser M, Schumann P, Martin K, Altenburger P, Burghardt J, Lubitz W, Busse HJ. 1999. *Agrococcus citreus* sp. nov., isolated from a medieval wall painting of the chapel of castle Herberstein (Austria). *Int. J. Syst. Bacteriol.* 49(Pt 3):1165–1170.
- Lee SD. 2008. *Agrococcus jejuensis* sp. nov., isolated from dried seaweed. *Int. J. Syst. Evol. Microbiol.* 58:2297–2300.
- Behrendt U, Schumann P, Ulrich A. 2008. *Agrococcus versicolor* sp. nov., an actinobacterium associated with the phyllosphere of potato plants. *Int. J. Syst. Bacteriol.* 58:2833–2838.
- Laval B, Cady SL, Pollack JC, McKay CP, Bird JS, Grotzinger JP, Ford DC, Bohm HR. 2000. Modern freshwater microbialite analogues for ancient dendritic reef structures. *Nature* 407:626–629.
- Lim DSS, Laval BE, Slater G, Antoniadis D, Forrest AL, Pike W, Pieters R, Saffari M, Reid D, Schulze-Makuch D, Andersen D, McKay CP. 2009. Limnology of Pavilion Lake B.C.—characterization of a microbialite forming environment. *Fund. Appl. Limnol.* 173:329–351.
- Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl. Acad. Sci. U. S. A.* 108:1513–1518.
- Marçais G, Kingsford C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27:764–770.
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, et al. 2001. The sequence of the human genome. *Science* 291:1304–1351.
- 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.