

# Genome Sequence of the Alkaline-Tolerant *Cellulomonas* sp. Strain FA1

Michael F. Cohen,<sup>a</sup> Ping Hu,<sup>b</sup> My Vu Nguyen,<sup>b</sup> Nina Kamennaya,<sup>b</sup> Natasha Brown,<sup>c</sup> Tanja Woyke,<sup>c</sup> Nikos Kyrpides,<sup>c</sup> Hoi-Ying Holman,<sup>b</sup> Tamas Torok<sup>a,b</sup>

Department of Biology, Sonoma State University, Rohnert Park, California, USA<sup>a</sup>; Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA<sup>b</sup>; DOE Joint Genome Institute, Walnut Creek, California, USA<sup>c</sup>

**We present the genome of the cellulose-degrading *Cellulomonas* sp. strain FA1 isolated from an actively serpentinizing highly alkaline spring. Knowledge of this genome will enable studies into the molecular basis of plant material degradation in alkaline environments and inform the development of lignocellulose bioprocessing procedures for biofuel production.**

Received 15 May 2015 Accepted 19 May 2015 Published 18 June 2015

**Citation** Cohen MF, Hu P, Nguyen MV, Kamennaya N, Brown N, Woyke T, Kyrpides N, Holman H-Y, Torok T. 2015. Genome sequence of the alkaline-tolerant *Cellulomonas* sp. strain FA1. *Genome Announc* 3(3):e00646-15. doi:10.1128/genomeA.00646-15.

**Copyright** © 2015 Cohen et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Michael F. Cohen, [cohenm@sonoma.edu](mailto:cohenm@sonoma.edu).

The cost feasibility of lignocellulosic crops and wastes as feedstocks for biofuel production is limited by the requirement for processing steps that can include an alkaline pretreatment to unbind lignin from the cell wall, followed by acid or enzyme hydrolysis of polysaccharides, and finally fermentation. One possible means to achieve feasibility is through consolidated bioprocessing, whereby chemical processing is facilitated by concomitant metabolic activity of microbial strains capable of functioning throughout the lignocellulose deconstruction process (1). Here, we report the genomic sequence of *Cellulomonas* sp. strain FA1, which possesses desirable characteristics for consolidated bioprocessing, including facultative anaerobic growth, tolerance to high ethanol concentrations and alkaline conditions (pH ≤ 12.2), and catabolism of plant cell wall polysaccharides to ethanol among other products (M. Cohen, M. V. Nguyen, J. Gray, L. Halberg, F. Bernie, A. Hargens, N. Kamennaya, G. Birarda, K. Sasaki, H. Y. Holman, and T. Torok, unpublished results).

Members of the genus *Cellulomonas* are cellulose-degrading Gram-positive rods or coccobacilli (2). There are currently 27 validly named species of *Cellulomonas* (<http://www.bacterio.net>) and genomes of 8 strains are available in the GenBank database (<https://www.ncbi.nlm.nih.gov/genbank/>). Strain FA1 was isolated from an alkaline switchgrass enrichment culture inoculated with decaying plant material sampled from within rare ultrabasic calcium hydroxide springs (pH 11.5 to 11.9) at The Cedars, Sonoma County, California (38°37'17.8"N, 123°07'57.0"W), a region of active serpentinization (3).

The genome of strain FA1 was sequenced using the PacBio RS sequencing platform (PacBio, Menlo Park, CA). The generated reads were trimmed and assembled *de novo* using HGAP v2.2.0.p1 (<https://github.com/PacificBiosciences/SMRT-Analysis/wiki/SMRT-Pipe-Reference-Guide-v2.2.0>). Gene prediction analysis and manual functional annotation was performed within the Integrated Microbial Genomes (IMG) platform (4). The size of the genome of strain FA1 was found to be 3,990,351 bp, comprising 1 scaffold, with a G+C content of 75%. Strain FA1 contained

3,601 predicted genes, 3,540 putative coding sequences (CDS), 6 rRNAs, and 45 tRNAs.

*In silico* DNA-DNA hybridization (DDH) comparisons with GGDC2.0 (<http://ggdc.dsmz.de/>) (5) between the genome of strain FA1 and other accessible *Cellulomonas* genome sequences in GenBank indicate the closest match to *Cellulomonas flavigena* with an estimated DDH value (high-scoring segment pair length/total length, generalized linear model-based) of 38.5 ± 3.4% while comparisons with other *Cellulomonas* strains gave values from 12.5% to 25.6%. The aligned 16S rRNA gene sequences of these two organisms show 99.3% identity (1,392 bp/1,402 bp).

Strain FA1 has the capacity to metabolize the three major plant cell wall polysaccharides cellulose, hemicellulose, and pectin (M. Cohen, M. V. Nguyen, and T. Torok, unpublished results). Consistent with this, the FA1 genome shows the presence of 82 glycosyl hydrolase and 43 glycosyl transferase homologs, compared to 86 and 44 homologs, respectively, in *C. flavigena*. Four identified Na<sup>+</sup>/H<sup>+</sup> antiporters encoded in the strain FA1 genome, NhaA, NhaD, NhaP, and the six genes of the multisubunit Mrp system, may have a role in the bacterium's tolerance to high pH (6, 7).

**Nucleotide sequence accession number.** This *Cellulomonas* sp. strain FA1 whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [LBMY000000000](https://www.ncbi.nlm.nih.gov/nuclink/LBMY000000000).

## ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy Joint Genome Institute (JGI) Community Science Program. JGI, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231.

## REFERENCES

- Hasunuma T, Okazaki F, Okai N, Hara KY, Ishii J, Kondo A. 2013. A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology. *Bioresour Technol* 135:513–522. <http://dx.doi.org/10.1016/j.biortech.2012.10.047>.
- Stackebrandt E, Schumann P, Prauser H. 2006. The family *Cellulomon-*

- adaceae*, p 983–1001. In Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E (ed), *The prokaryotes*. Springer Verlag, New York, NY.
3. Morrill PL, Kuenen JG, Johnson OJ, Suzuki S, Rietze A, Sessions AL, Fogel ML, Nealson KH. 2013. Geochemistry and geobiology of a present-day serpentinization site in California: The Cedars. *Geochim Cosmochim Acta* 109:222–240. <http://dx.doi.org/10.1016/j.gca.2013.01.043>.
  4. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC. 2009. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 25:2271–2278. <http://dx.doi.org/10.1093/bioinformatics/btp393>.
  5. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. <http://dx.doi.org/10.1186/1471-2105-14-60>.
  6. Krulwich TA, Hicks DB, Ito M. 2009. Cation/proton antiporter complements of bacteria: why so large and diverse? *Mol Microbiol* 74:257–260. <http://dx.doi.org/10.1111/j.1365-2958.2009.06842.x>.
  7. Swartz TH, Ikewada S, Ishikawa O, Ito M, Krulwich TA. 2005. The Mrp system: a giant among monovalent cation/proton antiporters? *Extremophiles* 9:345–354. <http://dx.doi.org/10.1007/s00792-005-0451-6>.