JOURNAL OF BACTERIOLOGY, Mar. 2011, p. 1483–1484 Vol. 193, No. 6
0021-9193/11/$12.00 doi:10.1128/JB.01515-10

Complete Genome Sequences for the Anaerobic, Extremely Thermophilic Plant Biomass-Degrading Bacteria *Caldicellulosiruptor hydrothermalis*, *Caldicellulosiruptor kristjanssonii*, *Caldicellulosiruptor kronotskyensis*, *Caldicellulosiruptor owensensis*, and *Caldicellulosiruptor lactoaceticus*[▽]

Sara E. Blumer-Schuette,[1] Inci Ozdemir,[1] Dhaval Mistry,[1] Susan Lucas,[2] Alla Lapidus,[3] Jan-Fang Cheng,[2] Lynne A. Goodwin,[5] Samuel Pitluck,[2] Miriam L. Land,[4] Loren J. Hauser,[4] Tanja Woyke,[2] Natalia Mikhailova,[2] Amrita Pati,[2] Nikos C. Kyrpides,[2] Natalia Ivanova,[2] John C. Detter,[5] Karen Walston-Davenport,[5] Shunsheng Han,[5] Michael W. W. Adams,[6] and Robert M. Kelly[1]*

Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina 27695-7905[1]; DOE Joint Genome Institute, Walnut Creek, California 94598[2]; Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111-2497[3]; BioEnergy Science Center and Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831[4]; Los Alamos National Laboratory, Bioscience Division B-6, Genome Science, Joint Genome Institute, Los Alamos, New Mexico 87545[5]; and Department of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia 30602-7229[6]

Received 17 December 2010/Accepted 23 December 2010

The genus *Caldicellulosiruptor* contains the most thermophilic, plant biomass-degrading bacteria isolated to date. Previously, genome sequences from three cellulolytic members of this genus were reported (*C. saccharolyticus*, *C. bescii*, and *C. obsidiansis*). To further explore the physiological and biochemical basis for polysaccharide degradation within this genus, five additional genomes were sequenced: *C. hydrothermalis*, *C. kristjanssonii*, *C. kronotskyensis*, *C. lactoaceticus*, and *C. owensensis*. Taken together, the seven completed and one draft-phase *Caldicellulosiruptor* genomes suggest that, while central metabolism is highly conserved, significant differences in glycoside hydrolase inventories and numbers of carbohydrate transporters exist, a finding which likely relates to variability observed in plant biomass degradation capacity.

Members of the genus *Caldicellulosiruptor* are asporogenic, plant biomass-degrading, hydrogen-generating members of the order *Clostridiales* (13, 17). The genus is globally distributed: species have been isolated from terrestrial geothermal hot springs in Russia (15, 18, 20), Iceland (3, 16), Yellowstone National Park in the United States (9), and New Zealand (17) and, in one case, from solar-heated mud flats in Owens Lake, CA (11). With optimal growth temperatures ranging from 70 to 78°C, the genus *Caldicellulosiruptor* contains the most thermophilic microorganisms capable of biological cellulose hydrolysis known. While 16S rRNA phylogeny indicates a very close relationship within the eight species studied thus far (94.8% to 99.4% identity), microbiological analysis indicated that the genus is more physiologically divergent than previously thought (2). Three complete genome sequences are currently available for this genus, those of *C. saccharolyticus* (19), *C. bescii* (12), and *C. obsidiansis* (5), all of which are capable of crystalline cellulose hydrolysis. To further probe the plant biomass-degrading capacity among *Caldicellulosiruptor* species, genome sequences of five additional members of this genus, including some hemicellulolytic but less cellulolytic members, were completed: those of *C. hydrothermalis*, *C. kristjanssonii*, *C. kronotskyensis*, *C. owensensis*, and *C. lactoaceticus* (draft phase). This also provides additional geographical diversity for examining *Caldicellulosiruptor* physiology.

All members of the *Caldicellulosiruptor* genus have similarly sized chromosomes, approximately 2.4 to 2.97 Mb. Their genomes are also A+T rich, ranging from 35 to 36% G+C. Similarly to *C. bescii* (4, 12), *C. kristjanssonii* also possesses an extrachromosomal element, the 15.9-kb plasmid pCALKR01. There are no similarities, however, between pATHE01/pATHE02 and pCALKR01, suggesting that there are no ubiquitous *Caldicellulosiruptor* plasmids. The other six genomes do not contain any extrachromosomal elements, although *C. hydrothermalis* possesses an inverted region that could exist outside the chromosome.

Sequencing strategy. All general aspects of library construction and sequencing can be found at http://www.jgi.doe.gov/. The genomes of *Caldicellulosiruptor* species were sequenced at the U.S. Department of Energy Joint Genome Institute (JGI) using a combination of Illumina (1) and 454 (14) technologies similar to the sequencing strategy for *C. obsidiansis* (5). Sequencing data were assembled either with VELVET (21) or by being converted into a Phrap assembly. The Phred/Phrap/Consed software package was used for sequence assembly and quality assessment (6–8) in the finishing process. Illumina data were used to correct base errors and increase consensus quality using Polisher software (A. Lapidus, unpublished data). After the shotgun stage, reads were assembled with parallel Phrap

* Corresponding author. Mailing address: Department of Chemical and Biomolecular Engineering, North Carolina State University, EB-1, 911 Partners Way, Raleigh, NC 27695-7905. Phone: (919) 515-6396. Fax: (919) 515-3465. E-mail: rmkelly@eos.ncsu.edu.

▽ Published ahead of print on 7 January 2011.

(High Performance Software, LLC). Possible misassemblies were corrected with gapResolution (C. Han, unpublished data), Dupfinisher (10), or sequencing of cloned bridging PCR fragments. Gaps between contigs were closed by editing in Consed, by PCR, and by Bubble PCR primer walks. The creation of additional reactions and shatter libraries was necessary to close gaps and to raise the quality of the finished sequences. Final assemblies are based on a minimum of 30× coverage of the genome.

Nucleotide accession numbers. GenBank accession numbers for *Caldicellulosiruptor* genomes announced here are as follows: *C. hydrothermalis*, CP002219; *C. kristjanssonii*, CP002326; *C. kronotskyensis*, CP002330; *C. lactoaceticus*, AEKD00000000; and *C. owensensis*, CP002216.

This research was supported by a grant (DE-PS02-06ER64304) from the BioEnergy Science Center (BESC), Oak Ridge National Laboratory, a U.S. Department of Energy (DOE) Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

REFERENCES

1. **Bennett, S.** 2004. Solexa Ltd. Pharmacogenomics **5:**433–438.
2. **Blumer-Schuette, S. E., D. L. Lewis, and R. M. Kelly.** 2010. Phylogenetic, microbiological, and glycoside hydrolase diversities within the extremely thermophilic, plant biomass-degrading genus *Caldicellulosiruptor*. Appl. Environ. Microbiol. **76:**8084–8092.
3. **Bredholt, S., J. Sonne-Hansen, P. Nielsen, I. M. Mathrani, and B. K. Ahring.** 1999. *Caldicellulosiruptor kristjanssonii* sp. nov., a cellulolytic, extremely thermophilic, anaerobic bacterium. Int. J. Syst. Bacteriol. **49:**991–996.
4. **Clausen, A., M. J. Mikkelsen, I. Schroder, and B. K. Ahring.** 2004. Cloning, sequencing, and sequence analysis of two novel plasmids from the thermophilic anaerobic bacterium Anaerocellum thermophilum. Plasmid **52:**131–138.
5. **Elkins, J. G., et al.** 2010. Complete genome sequence of the cellulolytic thermophile *Caldicellulosiruptor obsidiansis* OB47^T. J. Bacteriol. **192:**6099–6100.
6. **Ewing, B., and P. Green.** 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. **8:**186–194.
7. **Ewing, B., L. Hillier, M. C. Wendl, and P. Green.** 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:**175–185.
8. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
9. **Hamilton-Brehm, S. D., et al.** 2010. *Caldicellulosiruptor obsidiansis* sp. nov., an anaerobic, extremely thermophilic, cellulolytic bacterium isolated from Obsidian Pool, Yellowstone National Park. Appl. Environ. Microbiol. **76:**1014–1020.
10. **Han, C. S., and P. Chain.** 2006. Finishing repeat regions automatically with Dupfinisher, p. 141–146. *In* H. R. Arabnia and H. Valafar (ed.), Proceedings of the 2006 International Conference on Bioinformatics & Computational Biology. CSREA Press, Las Vegas, NV.
11. **Huang, C. Y., B. K. Patel, R. A. Mah, and L. Baresi.** 1998. *Caldicellulosiruptor owensensis* sp. nov., an anaerobic, extremely thermophilic, xylanolytic bacterium. Int. J. Syst. Bacteriol. **48:**91–97.
12. **Kataeva, I. A., et al.** 2009. Genome sequence of the anaerobic, thermophilic, and cellulolytic bacterium "*Anaerocellum thermophilum*" DSM 6725. J. Bacteriol. **191:**3760–3761.
13. **Ludwig, W., K.-H. Schleifer, and W. Whitman.** 2009. Revised road map to the phylum *Firmicutes*, p. 1–14. *In* P. de Vos et al. (ed.), Bergey's manual of systematic bacteriology, vol. 3. Springer, New York, NY.
14. **Margulies, M., et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.
15. **Miroshnichenko, M. L., et al.** 2008. *Caldicellulosiruptor kronotskyensis* sp. nov. and *Caldicellulosiruptor hydrothermalis* sp. nov., two extremely thermophilic, cellulolytic, anaerobic bacteria from Kamchatka thermal springs. Int. J. Syst. Evol. Microbiol. **58:**1492–1496.
16. **Mladenovska, Z., I. M. Mathrani, and B. K. Ahring.** 1995. Isolation and characterization of *Caldicellulosiruptor lactoaceticus* sp. nov., an extremely thermophilic, cellulolytic, anaerobic bacterium. Arch. Microbiol. **163:**223–230.
17. **Rainey, F. A., et al.** 1994. Description of *Caldicellulosiruptor saccharolyticus* gen. nov., sp. nov: an obligately anaerobic, extremely thermophilic, cellulolytic bacterium. FEMS Microbiol. Lett. **120:**263–266.
18. **Svetlichnyi, V. A., T. P. Svetlichnaya, N. A. Chernykh, and G. A. Zavarzin.** 1990. *Anaerocellum thermophilum* gen. nov sp. nov: an extremely thermophilic cellulolytic eubacterium isolated from hot springs in the Valley of Geysers. Microbiology **59:**598–604.
19. **van de Werken, H. J., et al.** 2008. Hydrogenomics of the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. Appl. Environ. Microbiol. **74:**6720–6729.
20. **Yang, S. J., et al.** 2010. Classification of '*Anaerocellum thermophilum*' strain DSM 6725 as *Caldicellulosiruptor bescii* sp. nov. Int. J. Syst. Evol. Microbiol. **60:**2011–2015.
21. **Zerbino, D. R., and E. Birney.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18:**821–829.