JOURNAL OF BACTERIOLOGY, Aug. 2010, p. 4256–4257 Vol. 192, No. 16
0021-9193/10/$12.00 doi:10.1128/JB.00588-10

GENOME ANNOUNCEMENT

Genome Sequence of a Cellulose-Producing Bacterium, *Gluconacetobacter hansenii* ATCC 23769▽

Prashanti R. Iyer,[1] Scott M. Geib,[2] Jeff Catchmark,[3] Teh-hui Kao,[4] and Ming Tien[4]*

Chemical Biology, the Huck Institutes of the Life Sciences, the Pennsylvania State University, University Park, Pennsylvania[1]; USDA-ARS, Pacific Basin Agricultural Research Center, Hilo, Hawaii[2]; Department of Agricultural and Biological Engineering, the Pennsylvania State University, University Park, Pennsylvania[3]; and Department of Biochemistry and Molecular Biology, the Pennsylvania State University, University Park, Pennsylvania[4]

Received 24 May 2010/Accepted 1 June 2010

The Gram-negative bacterium *Gluconacetobacter hansenii* is considered a model organism for studying cellulose synthesis. We have determined the genome sequence of strain ATCC 23769.

Plants produce cellulose, an unbranched chain of β-1,4-linked glucose units, as a structural polysaccharide. It is the most abundant polymer on earth, recently receiving much interest due to its potential use as a feedstock for bioethanol. Bacteria also produce cellulose. Among these, *Gluconacetobacter hansenii* (previously named *Acetobacter xylinus*) (4) has been extensively characterized and is a model system for cellulose biosynthesis (1, 2, 7). *G. hansenii* produces extracellular cellulose that is devoid of lignin or hemicellulose, making it an excellent source for pure cellulose. A lack of a completely sequenced genome for this organism has been a limiting factor in identifying other key proteins involved in cellulose synthesis.

The whole-genome sequencing of *G. hansenii* ATCC 23769 was performed using the 454 FLX-Titanium pyrosequencing technology (5). A combinatorial sequencing approach using 489,201 reads obtained from the shotgun library and 195,088 reads from an 8-kb pair end library (3) produced a total of 221,294,116 bp. These reads were assembled using the Newbler assembler, producing 88 large contigs (>500 bp) and a chromosome-sized scaffold of 3,646,142 bp with an average coverage of ×50.5. This scaffold contained exclusively chromosomal DNA and no plasmid sequences. The gaps in the large scaffold were filled by primer walking and subsequent sequencing of the PCR products. The resulting high-quality draft assembly, consisting of a large scaffold with 71 contigs, was annotated using the Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) service of the National Institute of Biotechnology Information (NCBI).

The chromosomal sequence of *G. hansenii* 23769 contains 3,547,122 bp, with a G+C content of 59%. The genome contains 3,351 genes, of which 3,308 are protein-encoding genes, accounting for 84% of the genome. There are 43 genes for tRNAs and 2 rRNA loci. The genes encoding proteins involved in cellulose synthesis are in an operon consisting of *acsAB* (GXY_04277), *acsC* (GXY_04282), and *acsD* (GXY_04292), as previously shown by Saxena et al. (7). Interestingly, there are two additional copies of *acsAB*, GXY_08864 and GXY_14452, which share 69% and 72% sequence identity, respectively, with the *acsAB* genes in the operon; the deduced amino acid sequences are 40% and 46% identical, respectively, with that deduced from *acsAB* in the operon. There are also two additional copies of *acsC*, GXY_08869 and GXY_014472, which share 72% and 65% DNA sequence identity, respectively, with the *acsC* gene in the operon; the deduced amino acid sequences share 28% and 30% amino acid identity, respectively, with that deduced from *acsC*. *acsAB* (GXY_08864) and *acsC* (GXY_08869) are only 17 bp apart, less than the distance (66 bp) between the *acsAB* and *acsC* genes in the operon. *acsAB* (GXY_14452) and *acsC* (GXY_14472) are separated by 3,299 bp, with three genes in between. However, a*csD* is present only in the operon, not duplicated elsewhere in the genome. The genome also contains three genes encoding diguanylate cyclase, as previously reported by Tal et al. (8). Diguanylate cyclase catalyzes the formation of cyclic di-GMP, a second messenger in bacteria that functions as an allosteric activator of cellulase synthase AcsAB (6).

Nucleotide sequence accession number. The high-quality draft of the *G. hansenii* ATCC 23769 genome sequence has been deposited in the GenBank under accession number CM000920 and project identification number 43711.

This material is based upon work supported as part of the Center for LignoCellulose Structure and Function, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under award number DE-SC0001090.

DNA sequencing was performed at the Genomics Core Facility of the Huck Institutes of the Life Sciences, the Pennsylvania State University, University Park, PA.

* Corresponding author. Mailing address: 408 Althouse Laboratory, Department of Biochemistry and Molecular Biology, the Pennsylvania State University, University Park, PA 16802. Phone: (814) 863-1165. Fax: (814) 863-7024. E-mail: mxt3@psu.edu.

▽ Published ahead of print on 11 June 2010.

REFERENCES

1. **Benziman, M., C. H. Haigler, R. M. Brown, Jr., A. R. White, and K. M. Cooper.** 1980. Cellulose biogenesis: polymerization and crystallization are coupled processes in Acetobacter xylinum. Proc. Natl. Acad. Sci. U. S. A. **77:**6678–6682.
2. **Deinema, M. H., and L. P. Zevenhui.** 1971. Formation of cellulose fibrils by

Gram-negative bacteria and their role in bacterial flocculation. Arch. Mikrobiol. **78:**42–57.

3. **Edwards, A., H. Voss, P. Rice, A. Civitello, J. Stegemann, C. Schwager, J. Zimmermann, H. Erfle, C. T. Caskey, and W. Ansorge.** 1990. Automated DNA sequencing of the human HPRT locus. Genomics **6:**593–608.
4. **Lisdiyanti, P., R. R. Navarro, T. Uchimura, and K. Komagata.** 2006. Reclassification of Gluconacetobacter hansenii strains and proposals of Gluconacetobacter saccharivorans sp. nov. and Gluconacetobacter nataicola sp. nov. Int. J. Syst. Evol. Microbiol. **56:**2101–2111.
5. **Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. T. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. I. Alenquer, T. P. Jarvie, K. B. Jirage, J. B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. G. Yu, R. F. Begley, and J. M. Rothberg.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.
6. **Ross, P., H. Weinhouse, Y. Aloni, D. Michaeli, P. Weinberger-Ohana, R. Mayer, S. Braun, E. de Vroom, G. A. van der Marel, J. H. van Boom, and M. Benziman.** 1987. Regulation of cellulose synthesis in Acetobacter xylinum by cyclic diguanylic acid. Nature **325:**279–281.
7. **Saxena, I. M., K. Kudlicka, K. Okuda, and R. M. Brown, Jr.** 1994. Characterization of genes in the cellulose-synthesizing operon (*acs* operon) of Acetobacter xylinum: implications for cellulose crystallization. J. Bacteriol. **176:** 5735–5752.
8. **Tal, R., H. C. Wong, R. Calhoon, D. Gelfand, A. L. Fear, G. Volman, R. Mayer, P. Ross, D. Amikam, H. Weinhouse, A. Cohen, S. Sapir, P. Ohana, and M. Benziman.** 1998. Three *cdg* operons control the turnover of cyclic di-GMP in Acetobacter xylinum: genetic organization and occurrence of conserved domains in isoenzymes. J. Bacteriol. **180:**4416–4425.