JOURNAL OF BACTERIOLOGY, Feb. 2010, p. 901–902 Vol. 192, No. 3
0021-9193/10/$12.00 doi:10.1128/JB.01450-09

GENOME ANNOUNCEMENT

Genome Sequence of the Cellulosome-Producing Mesophilic Organism *Clostridium cellulovorans* 743B[▽]

Yutaka Tamaru,[1,2,3]* Hideo Miyake,[1,2,3] Kouichi Kuroda,[4] Akihito Nakanishi,[4] Yujiro Kawade,[1] Kousuke Yamamoto,[1] Masaaki Uemura,[5] Yasuhiro Fujita,[5] Roy H. Doi,[6] and Mitsuyoshi Ueda[4]

Department of Life Science, Mie University Graduate School of Bioresources,[1] Department of Bioinformatics, Mie University Life Science Research,[2] and Laboratory of Applied Biotechnology, Mie University Venture Business Laboratory,[3] 1577 Kurimamachiya, Tsu, Mie 514-8507, Japan; Department of Applied Biochemistry, Kyoto University Graduate School of Agriculture, Kitashirakawa-Oiwake, Sakyo, Kyoto 606-8502, Japan[4]; Sumitomo Corporation, Harumi Triton Square Office Tower Y, 1-8-11 Harumi, Chuo, Tokyo 104-8610, Japan[5]; and Department of Molecular and Cellular Biology, University of California, Davis, One Shields Avenue, Davis, California 95616[6]

Received 4 November 2009/Accepted 17 November 2009

***Clostridium cellulovorans* 743B was isolated from a wood chip pile and is an anaerobic and mesophilic spore-forming bacterium. This organism degrades native substrates in soft biomass such as corn fiber and rice straw efficiently by producing an extracellular enzyme complex called the cellulosome. Here we report the genome sequence of *C. cellulovorans* 743B.**

The biotechnological potential of polysaccharolytic enzymes has resulted in the isolation and characterization of a large number of anaerobic, Gram-positive, spore-forming bacteria, the majority of which have been allocated to the genus *Clostridium*. Among clostridia, the cellulosomes produced by *Clostridium* species are particularly designed for efficient degradation of plant cell wall polysaccharides. The component parts of the multicomponent complex are integrated by virtue of a unique family of integrating modules, the cohesins and the dockerins, whose distribution and specificity dictate the overall cellulosome architecture (3). The cellulosome system in *Clostridium cellulovorans* 743B (ATCC 35296) has been studied extensively for the last 20 years and has resulted in providing basic information about mesophilic cellulosomes. This organism was isolated from a wood chip pile and is an anaerobic spore-forming bacterium whose optimal growth temperature is 37°C (9). It has the ability to utilize cellulose, xylan, pectin, cellobiose, glucose, fructose, galactose, and mannose as carbon sources for growth. Its fermentation products include H_2, CO_2, acetate, butyrate, formate, lactate, and ethanol. When it is grown in the presence of cellulose, electron micrographs have shown that large protuberances are present on its cell surface (4), while small or no protuberances are evident when cells are grown in the presence of glucose or cellobiose (5).

We sequenced a total length of 101,749,598 bp and analyzed 381,514 reads by Genome Sequencer FLX 454./Roche sequencing (8) (GS-FLX version) to highly oversample the genome (20× coverage) and generated 123,892 paired-end sequence tags, to enable the assembly of all tags using the GS De Novo Assembler version 1.1.03.24 (Roche Diagnostics) and the Genome Analyzer II and sequencing kit 36-Cycle Run (Illumina). Finally, we assembled 30 scaffolds (sets of 601 ordered and oriented contigs; total length of 5,123,527 bp) to generate approximately 5.1 Mbp of nearly contiguous *Clostridium botulinum* E3 strain Alaska E43 (accession no. NC_010723) complete genome sequence. We analyzed a number of predicted genes included in the *C. cellulovorans* genome using CRITICA (version 1.05b) (2) and Glimmer 2 (version 2.10) (6) to find regions in proteins with known functions. We annotated and classified according to Gene Ontology (GO) (1). *In silico* Molecular Cloning Genomic Edition ver. 3.0.26 software (In Silico Biology, Co., Ltd., Japan) was used for individual genomic analysis.

The *C. cellulovorans* 743B (ATCC 35296) genome consists of 5,123,527 bp. A total of 4,220 polypeptide-encoding open reading frames (ORFs) were identified using CRITICA, while 4,297 ORFs were identified using Glimmer 2. The number of ORFs identical between CRITICA and Glimmer 2 was 2,773. Sixty-three tRNAs and 33 anticodons were also identified using tRNAscan-SE (7). In comparison of the genome sizes among cellulosomal clostridia such as *Clostridium cellulolyticum* H10 (4.07 Mbp) (GenBank accession no. CP001348) and *Clostridium thermocellum* ATCC 27405 (3.84 Mbp) (GenBank accession no. CP000568), the *C. cellulovorans* genome was over 1 Mbp larger than the other genomes. Moreover, the number of predicted genes (4,220 by CRITICA) in the *C. cellulovorans* genome was the largest among them. On the other hand, the G+C content in *C. cellulovorans* was 31.1%, similar to that (30.9%) in *Clostridium acetobutylicum* ATCC 824 (GenBank accession no. AE001437), while the G+C contents in *C. cellulolyticum* and *C. thermocellum* were 37.7% and 39.0%, respectively.

A protein BLAST search against the database of clusters of

* Corresponding author. Mailing address: Department of Life Science, Mie University Graduate School of Bioresources, 1577 Kurimamachiya, Tsu, Mie 514-8507, Japan. Phone and fax: 81-59-231-9560. E-mail: ytamaru@bio.mie-u.ac.jp.

▽ Published ahead of print on 30 November 2009.

orthologous groups (COGs) of proteins indicated that 4,171 genes were annotated by 4,220 predicted coding sequences using CRITICA while 4,098 genes were identified by 4,297 predicted coding sequences using Glimmer 2. On the other hand, a protein BLAST search against the NCBI nr database indicated that 4,184 genes were annotated by 4,220 predicted coding sequences using CRITICA while 4,071 genes were identified by 4,297 predicted coding sequences using Glimmer 2. Interestingly, 57 cellulosomal genes were found in the *C. cellulovorans* genome and coded for not only carbohydrate-active enzymes but also lipases, peptidases, and proteinase inhibitors. Moreover, two novel genes encoding a scaffolding protein were found in the genome. Thus, by examining genome sequences from multiple *Clostridium* species, comparative genomics offers new insight into genome evolution and the way in which natural selection molds functional DNA sequence evolution. Our analysis, coupled with the genome sequence data, will provide a road map for constructing enhanced cellulosome-producing *Clostridium* strains for industrial applications such as biofuel production.

Nucleotide sequence accession number. The *C. cellulovorans* genome sequence and annotation data have been deposited in GenBank under accession number SRA008356.

We thank Sumitomo Corporation, Tokyo, Japan, for financial support of this work. Bioinformatics work was performed under the auspices of the Dragon Genomics Center, Takara Bio Co. Ltd, Japan.

REFERENCES

1. **Ashburner, M., et al.** 2000. Gene Ontology: tool for the unification of biology. Nat. Genet. **25:**25–29.
2. **Badger, J. H., and G. J. Olsen.** 1999. CRITICA: coding region identification tool invoking comparative analysis. Mol. Biol. Evol. **16:**514–524.
3. **Bayer, E. A., R. Lamed, B. White, and H. J. Flint.** 2008. From cellulosomes to cellulosomics. Chem. Rec. **8:**364–377.
4. **Blair, B. G., and K. L. Anderson.** 1998. Comparison of staining techniques for scanning electron microscopic detection of ultrastructural protuberances on cellulolytic bacteria. Biotech. Histochem. **73:**107–113.
5. **Blair, B. G., and K. L. Anderson.** 1999. Regulation of cellulose-inducible structures of *Clostridium cellulovorans*. Can. J. Microbiol. **45:**242–249.
6. **Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:**4636–4641.
7. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
8. **Margulies, M., et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.
9. **Sleat, R., R. A. Mah, and R. Robinson.** 1984. Isolation and characterization of an anaerobic, cellulolytic bacterium, Clostridium cellulovorans sp. nov. Appl. Environ. Microbiol. **48:**88–93.