JOURNAL OF BACTERIOLOGY, Oct. 2011, p. 5568–5569 Vol. 193, No. 19
0021-9193/11/$12.00 doi:10.1128/JB.05637-11

Complete Genomic Sequence of the *O*-Desmethylangolensin-Producing Bacterium *Clostridium* rRNA Cluster XIVa Strain SY8519, Isolated from Adult Human Intestine

Shin-ichiro Yokoyama,[1] Kenshiro Oshima,[2] Izumi Nomura,[3] Masahira Hattori,[2] and Tohru Suzuki[3]*

Department of Environmental and Chemical Research, Industrial Technology Center, Gifu Prefectural Government, 47 Kitaoyobi, Kasamatsu-cho, Hashima-gun, Gifu 501-6064, Japan[1]; Graduate School of Frontier Sciences, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan[2]; and The United Graduate School of Agricultural Science, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan[3]

Received 22 June 2011/Accepted 18 July 2011

The *O*-desmethylangolensin-producing *Clostridium* rRNA cluster XIVa strain SY8519 was isolated from the intestinal flora of a healthy human as a key isoflavonoid-metabolizing bacterium. Here, we report the finished and annotated genomic sequence of this organism.

The complete genomic sequence of the *O*-desmethylangolensin (*O*-DMA)-producing bacterium *Clostridium* rRNA cluster XIVa strain SY8519 (7) was determined by a whole-genome shotgun strategy with the Sanger method. Genomic libraries containing 2-kb inserts were constructed with pIS1 plasmids, and 38,400 sequences were generated, providing 9.4-fold coverage from both ends of the genomic clones. Sequence reads were assembled with the Phred-Phrap-Consed program (1). Remaining gaps between contigs were closed by direct sequencing of fosmid clones. Prediction and annotation of protein-coding genes were performed by MiGAP (4).

The genome of strain SY8519 consists of a circular 2,835,737-bp chromosome with a 50.7% GC content and contains 2,619 predicted protein-coding sequences (CDSs) but no plasmid. It also has 4 rRNA operons and 53 tRNA genes. Interestingly, the four rRNA operons of this organism queue up in tandem at 5′-16S-5S-23S-3′. In general, bacterial rRNA operons line up in tandem at 5′-16S-23S-5S-3′. This result suggests that the region would be useful as a PCR marker to identify organisms in the same genus as strain SY8519.

The predicted CDSs were submitted to the Kyoto Encyclopedia of Genes and Genomes Automatic Annotation Server (http://www.genome.jp/tools/kaas/) (3) and the virulence factor database (http://www.mgc.ac.cn/VFs/) (5). We could assign 752 CDSs (29%) to known functions, 102 (4%) as conserved hypothetical genes, and 1,765 (67%) as novel hypothetical genes. In addition to glycolysis/gluconeogenesis pathway-related genes, the genome also possesses genes for butyrate metabolism (*ato*, *buk*, *ptb*, *bcd*, *fad*, and *paaH*). This result corresponds with our previous report that the strain produces lactate and butyrate as end products of glucose fermentation (7). The VFDB result indicates that the strain possesses a gene for a tetracycline resistance protein (*tetW*) and urease and urease accessory protein (*ureG*) genes (CXIVA_10010, CXIVA_25610, and CXIVA_25570, respectively). These results also agree with our previous report that the strain is urease-positive and resistant to aminoglycoside antibiotics (7).

Similarity at the sequence level (1,400 at CDSs) was observed between strain SY8519 and the closely related type strain *Eubacterium rectale* ATCC 33656, which belongs to *Clostridium* cluster XIVa (GenBank accession no. CP001107) (2). However, a reciprocal BLASTP search revealed 1,206 (46.2%) protein-coding genes that are present in strain SY8519 but absent in *E. rectale* ATCC 33656^T. The genome of strain SY8519 is 800 kb shorter than that of *E. rectale* ATCC 33656^T (circular; 3,632,260 bp) (2). Although we could not predict the genes related to daidzein metabolism in this study, such an estimation is in progress in our laboratory. Finally, in addition to the previous report about equol-producing bacteria (6, 8), this is the first report elucidating the complete genomic sequence of a bacterium producing *O*-DMA from daidzein.

Nucleotide sequence accession number. The sequence data of the *Clostridium* strain SY8519 genome have been deposited in GenBank/DDBJ/EMBL under accession number AP012212.

We thank K. Furuya, C. Shindo, H. Inaba, E. Iioka, and Y. Hattori for technical support.

This research was supported by a Grant-in-Aid for Scientific Research on Priority Areas 'Comprehensive Genomics' (M.H.), the global COE project 'Genome Information Big Bang' (M.H. and K.O.), a Grant-in Aid for Scientific Research (C), 20510189 (T.S.), from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and the Iijima Memorial Foundation for the Promotion of Food Science and Technology, 2011 (T.S.).

REFERENCES

1. **Gordon, D., C. Desmarais, and P. Green.** 2001. Automated finishing with Autofinish. Genome Res. **11:**614–625.
2. **Mahowald, M. A., et al.** 2009. Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. Proc. Natl. Acad. Sci. U. S. A. **106:**5859–5864.
3. **Moriya, Y., M. Itoh, S. Okuda, A. C. Yoshizawa, and M. Kanehisa.** 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. **35:**W182–185.
4. **Sugawara, H., A. Ohyama, H. Mori, and K. Kurokawa.** 2009. Microbial Genome Annotation Pipeline (MiGAP) for diverse users, software demonstration S001-1–2. *In* The 20th International Conference on Genome Informatics (Yokohama). Japanese Society for Bioinformatics, Tokyo, Japan.
5. **Yang, J., L. H. Chen, L. L. Sun, J. Yu, and Q. Jin.** 2008. VFDB 2008 release:

* Corresponding author. Mailing address: The United Graduate School of Agricultural Science, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan. Phone and fax: 81-58-293-2991. E-mail: suzuki@gifu-u.ac.jp.

an enhanced web-based resource for comparative pathogenomics. Nucleic Acids Res. **36:**D539–542.
6. **Yokoyama, S., and T. Suzuki.** 2008. Isolation and characterization of a novel equol-producing bacterium from human feces. Biosci. Biotechnol. Biochem. **72:**2660–2666.
7. **Yokoyama, S., T. Niwa, T. Osawa, and T. Suzuki.** 2010. Characterization of an *O*-desmethylangolensin-producing bacterium isolated from human feces. Arch. Microbiol. **192:**15–22.
8. **Yokoyama, S., K. Oshima, I. Nomura, M. Hattori, and T. Suzuki.** 2011. Complete genomic sequence of the equol-producing bacterium *Eggerthella* sp. strain YY7918, isolated from adult human intestine. J. Bacteriol. **193:** 5570–5571.