


Complete genome sequence of *Cellulosilyticum lentocellum* WCF-2 isolated from cow dung

Jun Heo, Jaehong You, InCheol Park, Byeong-Hak Han, Soon-Wo Kwon, and Jae-Hyung Ahn* 

Agricultural Microbiology Division, National Institute of Agricultural Sciences, Rural Development Administration, Wanju 55365, Republic of Korea

소 분변에서 분리된 *Cellulosilyticum lentocellum* WCF-2의 유전체 염기서열 분석

허준 · 유재홍 · 박인철 · 한병학 · 권순우 · 안재형* 

농촌진흥청 국립농업과학원 농업미생물과

(Received September 9, 2019; Revised September 18, 2019; Accepted September 18, 2019)

An anaerobic bacterial strain WCF-2 was isolated from cow dung in finding cellulose-degrading bacteria for use as silage additives. Strain WCF-2 showed a higher cellulolytic activity than *Cellulosilyticum lentocellum* DSM 5427^T, the closest relative of strain WCF-2 (98.2% of 16S rRNA gene sequence similarity). We sequenced the complete genome of strain WCF-2 and compared it with that of *C. lentocellum* DSM 5427^T. The OrthoANI value between the two strains was 97.9% thus strain WCF-2 was identified as *C. lentocellum*. The genome size of strain WCF-2 was 4,779,774 bp with a G + C content of 34.4%, 4,154 coding genes (CDS), 54 pseudo genes, and 142 RNA genes. Strain WCF-2 harbored seven cellulase genes, five of which showed low similarities with those of *C. lentocellum* DSM 5427^T.

Keywords: *Cellulosilyticum lentocellum*, cellulolytic activity, genome sequence

A cellulolytic anaerobic bacterium, strain WCF-2 was isolated from cow dung in finding cellulose-degrading bacteria for use as silage additives. The strain WCF-2 showed a higher cellulolytic activity than *Cellulosilyticum lentocellum* DSM 5427^T, the

closest relative of strain WCF-2 (98.2% of 16S rRNA gene similarity). Currently, the genus *Cellulosilyticum* includes two species, *C. lentocellum*, which was isolated from estuarine mud bank of a river receiving paper-mill and domestic effluent (Murray *et al.*, 1986), and *C. ruminicola*, which was isolated from the rumen content of yak (Cai and Dong, 2010). They are anaerobic, hydrolyse cellulose and xylan, and produce acetate as one of the major fermentation products (Cai and Dong, 2010). They have been often reported to be associated with digestive tracts of animals (Meehan and Beiko, 2014; Guevarra *et al.*, 2015).

The complete genome sequence of strain WCF-2 was obtained from 1.5 Gb PacBio RS II platform data (Pacific Biosciences) and 1.4 Gb Illumina HiSeq platform data (Illumina) at the Macrogen Inc. *De novo* assembly was performed based on the PacBio data using Canu (Version 1.4) (Koren *et al.*, 2017) and the genome coverage was 156.0×. Then the assembly was error-corrected based on the Illumina HiSeq data using Pilon (version 1.21) (Walker *et al.*, 2014). The genome was annotated by NCBI prokaryotic genome annotation pipeline (Version 4.9) (Tatusova *et al.*, 2016). The genome sequence has been deposited in NCBI GenBank database (<http://www.ncbi.nlm.nih.gov/>) under accession number CP034675. The strain has also been

*For correspondence. E-mail: hyungz@korea.kr;
Tel.: +82-63-238-3045; Fax: +82-63-238-3834

Table 1. Genome features of *Cellulosilyticum lentocellum* WCF-2

| Attribute | Value |
|----------------------|-----------------|
| Genome size (bp) | 4,779,774 |
| GC content (%) | 34.4 |
| No. of contigs | 1 (CP034675) |
| Total genes | 4,350 |
| Protein-coding genes | 4,154 |
| Pseudogenes | 54 |
| rRNAs (5S, 16S, 23S) | 36 (13, 11, 12) |
| tRNAs | 102 |
| Other RNAs | 4 |

deposited in the Korea Agricultural Culture Collection under KACC number 92181P.

The genome sequence of strain WCF-2 was composed of a single circular chromosome. The genome size was 4,779,774 bp with a G + C content of 34.4%, 4,154 coding genes (CDS), 54 pseudo genes and 142 RNA genes. The genomic features of strain WCF-2 are summarized in Table 1.

Genomic similarities between strain WCF-2 and type strains of the genus *Cellulosilyticum* were calculated using the ANI calculator at the EZBioCloud website (www.ezbiocloud.net) (Yoon *et al.*, 2017). The OrthoANI value between strain WCF-2 and *C. lentocellum* DSM 5427^T (CP002582) was 97.9% while it was 73.5% for strain WCF-2 and *C. ruminicola* JCM 14822^T (BBCG00000000). Because OrthoANI values of 95–96% was recommended as cut-off for species demarcation (Lee *et al.*, 2016; Chun *et al.*, 2018), strain WCF-2 was identified as *C. lentocellum*.

We found that strain WCF-2 has seven genes the products of which were annotated as cellulase (locus tags: EKH87_00115, EKH87_05715, EKH87_06845, EKH87_11290, EKH87_11980, EKH87_15865, and EKH87_17100) while *C. lentocellum* DSM 5427^T has six cellulase-coding genes. Two cellulases of strain WCF-2 showed high similarities of amino acid sequence (more than 94%) with two cellulases of *C. lentocellum* DSM 5427^T, while the remaining five genes showed low similarities (lower than 16.0%) with the other cellulases of *C. lentocellum* DSM 5427^T.

적 요

사일리지 제조에 사용하기 위한 섬유소 분해균을 탐색하는

중 절대혐기성 세균인 WCF-2 균주를 선발하였다. WCF-2 균주는 16S rRNA 유전자 염기서열 유사도가 가장 높은(98.2%) 표준균주인 *Cellulosilyticum lentocellum* DSM 5427^T보다 높은 섬유소 분해 활성을 나타내었다. WCF-2 균주의 전체 유전체 염기서열을 분석하고 이를 *C. lentocellum* DSM 5427^T와 비교하였을 때 두 균주의 OrthoANI 값은 97.9%로 나타나 WCF-2를 *C. lentocellum*으로 동정하였다. WCF-2 균주의 유전체 크기는 4,779,774 bp이고 G + C 함량은 34.4%였으며 4,154개의 단백질 암호화 유전자 및 142개의 RNA 암호화 유전자를 보유하고 있었다. 또한 WCF-2 균주는 7개의 cellulase를 보유하고 있었으며 이 중 5개는 *C. lentocellum* DSM 5427^T의 cellulase와 낮은 유사도를 나타내었다.

Acknowledgements

This study was carried out with the support (PJ013596) of National Institute of Agricultural Sciences, Rural Development Administration, Republic of Korea.

References

- Cai, S and Dong X.** 2010. *Cellulosilyticum ruminicola* gen. nov., sp. nov., isolated from the rumen of yak, and reclassification of *Clostridium lentocellum* as *Cellulosilyticum lentocellum* comb. nov. *Int. J. Syst. Evol. Microbiol.* **60**, 845–849.
- Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, et al.** 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **68**, 461–466.
- Guevarra, RB, Kim J, Nguyen SG, and Unno T.** 2015. Comparison of fecal microbial communities between white and black pigs. *J. Appl. Biol. Chem.* **58**, 369–375.
- Koren, S, Walenz BP, Berlin K, Miller JR, Bergman NH, and Phillippy AM.** 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* **27**, 722–736.
- Lee I, Kim YO, Park SC, and Chun J.** 2016. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int. J. Syst. Evol. Microbiol.* **66**, 1100–1103.
- Meehan CJ and Beiko RG.** 2014. A phylogenomic view of ecological specialization in the *Lachnospiraceae*, a family of digestive tract-associated Bacteria. *Genome Biol. Evol.* **6**, 703–713.
- Murray WD, Hofmann L, Campbell NL, and Madden RH.** 1986. *Clostridium lentocellum* sp. nov., a cellulolytic species from river sediment containing paper-mill waste. *Syst. Appl. Microbiol.*

8, 181–184.

Tatusova T, DiCuccio M, Badretdin A, Chetvermin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, and Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **44**, 6614–6624.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, and Earl AM. 2014.

Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**, e112963.

Yoon SH, Ha SM, Lim J, Kwon S, and Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek* **110**, 1281–1286.