Korean Journal of Microbiology (2022) Vol. 58, No. 4, pp. 314-317
DOI https://doi.org/10.7845/kjm.2022.2073

pISSN 0440-2413
eISSN 2383-9902

Draft genome sequence of a *Flavobacterium chungbukense* CS100^T isolated from soil

Ji-Sung Oh and Dong-Hyun Roh*

Department of Biological Sciences and Biotechnology, Chungbuk National University, Cheongju 28644, Republic of Korea

토양에서 분리된 *Flavobacterium chungbukense* CS100^{T}의 유전체 염기서열

오지성 · 노동현*

충북대학교 대학원 생명시스템학과

(Received October 24, 2022; Revised December 6, 2022; Accepted December 7, 2022)

The draft genome sequence of *Flavobacterium chungbukense* CS100^T isolated from soil was determined using Illumuna Hiseq X-ten platform. The assembled genome consists of 18 scaffolds with a total length of 5,355,850 bp with N50 values of 851,609 bp. The genomic DNA G + C content was 33.5%. The draft genome encoded 4,436 protein-coding genes, 8 rRNA genes, 52 tRNA genes, 3 non-coding RNA genes and 53 pseudogenes. The genome sequence of this type strain in the genus *Flavobacterium* will be used for reference to taxonomical classification based on the genome. Additionally, several useful genes in the genome related to antimicrobials, moisturizer and pigments production, and the degradation of aromatic compounds and biopolymers might be used for industrial applications.

Keywords: *Flavobacterium chungbukense* CS100^T, draft genome sequence, taxonomical classification, type strain

The genus *Flavobacterium* belongs to the family *Flavobacteriaceae*, phylum *Bacteriodetes.* This genus was originally described by Bergey *et al.* (1923) and reclassified by the phylogenetic tree based on the 16S rRNA sequences (Bernardet *et al*., 1996). After reclassification, the genus was extensively emended with following characteristics; Gram-negative, aerobic, motile by gliding, yellow-pigmented rod-shaped bacteria with menaquinone-6 as the major respiratory quinone (Bernardet *et al*., 1996). Thereafter, it was additionally emended that the genus contained phosphatidylethanolamine as the major polar lipid, iso-$C_{15:0}$, iso-$C_{15:1}$ G, iso-$C_{15:0}$ 3-OH, iso-$C_{16:0}$ 3-OH and iso-$C_{17:0}$ 3-OH as the predominant fatty acids and 30–52 mol% of DNA G + C content, and that cells were non-motile or may be motile by gliding. Type species in the genus is *Flavobacterium aquatile*. This genus currently consists of 268 recognized species validly published with correct names (Parte *et al.*, 2020). Strain of the genus *Flavobacterium* have been isolated in a variety of environments, such as soil, sediment, fresh- and sea-water, Antarctica, a glacier, diseased fish and activated sludge. Ecologically, there has been reported that members of the genus *Flavobacterium* are related to degrade complex organic chemicals in soil and play a role in the turnover of biopolymers (Bernardet and Bowman, 2006).

Due to the universal presence and mix of conserved and variable regions that permit species-specific identification in bacterial strains, and the correlation with DNA-DNA reassociation values, 16S rRNA gene has been widely used for phylogenetic identification of bacteria (Coenye and Vandamme, 2003). However, as the number of available genomic sequence data increases, taxonomical classification using genomes is becoming more demanding (Achtman and Wagner, 2008). In

***For correspondence.** E-mail: dhroh@chungbuk.ac.kr;
Tel.: +82-43-261-3368; Fax: +82-43-264-9600

this paper, we describe the draft genome sequence of a type strain *F. chungbukense* CS100^T isolated from soil of Korea (Lim *et al.*, 2011).

For the extraction of genomic DNA, *F. chungbukense* CS100^T was incubated in Tryptic soy broth (TSB, Difco) at 25°C for 3 days and then the genomic DNA was extracted using MagAttract® HMW DNA kit (Qiagen) according to the manufacturer's instructions. The sequencing *F. chungbukense* CS100^T genome was performed by Macrogen Inc. as the Illumina Hiseq X-ten platform with TruSeq Nano DNA (350 bp insert size) library. Raw reads were qualified by FastQC (version 0.11.5) and were assembled by SPAdes (version 3.13.0). Genome completeness and contamination were analyzed with CheckM (Version 1.0.18). The genome annotation was conducted using NCBI Prokaryotic Genome Annotation Pipeline, and additional functions of the predicted genes were conducted by BlastKOALA with KEGG database, RAST server with SEED database and PathoSystems Resource Integration Center server. AntiSMASH 6.0 (https://antismash.secondarymetabolites.org/) was used to detect and characterize biosynthetic gene clusters for secondary metabolite.

The draft genome of *F. chungbukense* CS100^T consisted of 18 scaffolds with a total length of 5,355,850 bp and N50 value of 851,609 bp (Table 1). Genome coverage was 147.3× and the G + C content was 33.5% which was slightly lower than that measured and reported by HPLC (Lim *et al.*, 2011). The CheckM estimation indicated that genome completeness was 99.7% with 0.4% contamination and 0% strain heterogeneity, respectively. The draft genome comprised 4,436 protein-coding genes, 3, 3, 2 rRNA genes (5S, 16S, 23S), and 52 tRNA genes.

A few strains isolated from seawater of the family *Flavobacteriaceae* were known as algicidal bacteria (Adachi *et al.*, 2002). Interestingly, the genome of *F. chungbukense* CS100^T contained the initial synthesis *redP/R* gene for algicide pigment, prodigiosin, as reported in *Halobacillus* sp. Nhm2S1 genome (Oh and Roh, 2021).

Table 1. Genomic features of *Flavobacterium chungbukense* CS100^T

Feature type	Value
Genome size (bp)	5,355,850
Number of sacffolds	18
Scaffold N50	851,609
Genome coverage (X)	147.3
G + C content (%)	33.5
CDS (with protein)	4,436
rRNA genes (5S, 16S, 23S)	3, 3, 2
ncRNA genes	3
Pseudogenes	53
Accession number (GenBank)	JAJGZV000000000

According to the antiSMASH analysis seven secondary metabolite regions were found in the genome of *F. chungbukense* CS100^T. Region 1.1 on scaffold 1 (accession No. JAJGZV010000001.1) has non-ribosomal peptide synthetase and type 1 polyketide synthase cluster. This region contains two genes, *sevB* and *sevA* (locus_tag LL966_RS02975 and _RS02980), for biosynthesis of tripeptide sevadicin (D-Phe-D-Ala-Trp) which had antibacterial activity (Garcia-Gonzales *et al*., 2014). Additionally, there was a similarly known gene cluster related to biosynthesis for bacterial cell surface polysaccharides such as colanic acids that had promising applications in food, cosmetic, and healthcare fields. Region 11.1 and Region 11.2 on scaffold 2 (accession No. JAJGZV010000002.1) had proteusin, a family of post-translationally modified peptides such as bacteriocin and type III polyketide synthase, respectively. Region 12.1 on scaffold 3 (accession No. JAJGZV010000003.1) has specialized polyunsaturated carboxylic acid aryl polyene and organic compound resorcinol gene cluster. Resorcinol is used as a chemical intermediate for pharmaceutic synthesis. In this region from locus_tag LL966_RS11975 to locus_tag LL966_RS121150 showed similar a gene cluster for pigment flexirubin production in *Flavobacterium johnsoniae* UW101 (McBride *et al*., 2009), which were consistent with the pigment of the *F. chungbukense* CS100^T (Lim *et al.*, 2011). Region 12.2 on the same scaffold 3 also has terpene biosynthesis gene cluster. Region 16.1 on scaffold 7 (accession No. JAJGZV010000007.1) has β-lactone containing protease inhibitor gene cluster that is natural product with potential antifungal and antibacterial activity. Finally, region 17.1 on scaffold 8 (accession No. JAJGZV010000008.1) had terpene gene cluster for production of carotenoid.

The genome of *F. chungbukense* CS100^T also contained the genes coding for the degradation of benzoate from 6-hydroxy-cyclohex-1-ene-1-carbonyl-CoA to acetyl-CoA, oxidoreductases (EC 1.1.1-), hydrolases (EC 3.7.1.-), 3-hydroxyl-CoA dehydrogenase (EC 1.1.1.35), acetyl CoA C-acyltransferase (EC 2.3.1.16), glutaryl-CoA dehydrogenase (EC 1.3.8.6), enoyl-

CoA hydratase (EC 4.2.1.17), 3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) and acetyl-CoA C-acetyltransferase (EC 2.3.1.9). It also contained coding 6-steps sequential geraniol degradation genes from 3-hydroxy-3-isohexenylglutaryl-CoA to 3-methylcrotonyl-CoA such as hydroxylmethylglutaryl-CoA lyase (EC 4.1.3.4), acetyl-CoA C-acyltransferase (EC 2.3.1.16), oxidoreductase (EC 1.3.99.-), enoyl-CoA hydratase (EC 4.2.1.17) and 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35). In addition, the genome also contained coding genes for biopolymer degradation such as α-amylase (EC 3.2.1.1), glucan 1,4-α-glucosidase (EC 3.2.1.3) and α-glucosidase (EC 3.2.1.20) related to starch hydrolysis that is reported by Lim *et al.* (2011), and 1,4-β-xylosidase (EC 3.2.1.37) related to xylan decomposition.

Based on this genomic information, the *F. chungbukense* CS100^T is expected to play important roles in the degradation of biopolymers and biosynthesis of secondary metabolites. In addition, the genome sequence is also expected to play an important role as a reference for bacterial taxonomical classification.

Nucleotide sequence and strain accession numbers

The draft genome sequence and strain *Flavobacterium chungbukense* CS100^T has been deposited to GenBank and the Korean Culture Center of Microorganisms under the accession number JAJGZV000000000, and KACC 15048^T and JCM 17386^T, respectively.

적 요

토양에서 분리된 *Flavobacterium chungbukense* CS100^T 균주의 초안 유전체 서열을 Illumina Hiseq X-ten platform을 사용하여 결정하였다. 연결 조립된 유전자들의 scaffold는 18개, 전체 길이는 5,355,850 bp, 유전체의 G + C 함량은 33.5% 이었다. 유전체는 4,436개의 단백질 암호와 유전자, 8개의 rRNA 유전자, 52개의 tRNA 유전자, 3개의 non-coding RNA 유전자 및 53개 위유전자(pseudo gene)를 암호화하였다. *Flavobacterium* 속의 이 표준균주 유전체 서열은 유전체를 기반으로 하는 분류학에 대한 기준으로 사용될 것이다. 또한 항균제, 보습제 및 색소 생산, 방향족 화합물 및 생체 고분자 분해와 관련된 유전체의 여러 유용한 유전자가 산업 응용 분야에 사용될 수 있을 것이다.

Acknowledgments

This work was supported by a funding for the academic research program of Chungbuk National University in 2022 and by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1D1A3B04033871).

Conflict of Interest

The authors have no conflict of interest to report.

References

Achtman M and Wagner M. 2008. Microbial diversity and the genetic nature of microbial species. *Nat. Rev. Microbiol.* **6**, 431–440.

Adachi M, Fukami K, Kondo R, and Nishijima T. 2002. Identification of marine algicidal *Flavobacterium* sp. 5 N-3 using multiple probes and whole cell hybridization. *Fish. Sci.* **68**, 713–720.

Bergey DH, Harrison FC, Breed RS, Hammer BW, and Huntoon FM. 1923. Bergey's Manual of Determinative Bacteriology, 1st edn. The Williams & Wilkins Co., Baltimore, Maryland, USA.

Bernardet JF and Bowman JP. 2006. The Genus *Flavobacterium*, pp. 481–531. *In* Dworkin M, Falkow S, Rosenberg E, Schleifer KH, and Stackebrandt E. (eds.), The Prokaryotes. Springer, New York, USA.

Bernardet JF, Segers P, Vancanneyt M, Berthe F, Kersters K, and Vandamme P. 1996. Cutting a Gordian knot: emended classification and description of the genus *Flavobacterium,* emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* nom. nov.(basonym, *Cytophaga aquatilis* Strohl and Tait 1978). *Int. J. Syst. Evol. Microbiol.* **46**, 128–148.

Coenye T and Vandamme P. 2003. Intragenomic heterogeneity between multiple 16S ribosomal RNA operons in sequenced bacterial genomes. *FEMS Microbiol. Lett.* **228**, 45–49.

Garcia-Gonzalez E, Müller S, Ensle P, Süssmuth RD, and Genersch E. 2014. Elucidation of sevadicin, a novel non-ribosomal peptide secondary metabolite produced by the honey bee pathogenic bacterium *Paenibacillus larvae*. *Environ. Microbiol.* **16**, 1297–1309.

Lim CS, Oh YS, Lee JK, Park AR, Yoo JS, Rhee SK, and Roh DH. 2011. *Flavobacterium chungbukense* sp. nov., isolated from soil. *Int. J. Syst. Evol. Microbiol.* **61**, 2734–2739.

McBride MJ, Xie G, Martens EC, Lapidus A, Henrissat B, Rhodes RG, Goltsman E, Wang W, Xu J, Hunnicutt DW, *et al.* 2009. Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Appl. Environ. Microbiol.* **75**, 6864–6875.

Oh JS and Roh DH. 2021. Draft genome sequence of an algicidal bacterium *Halobacillus* sp. Nhm2S1. *Korean J. Microbiol.* **57**, 226–228.

Parte AC, Carbasse JS, Meier-Kolthoff JP, Reimer LC, and Göker M. 2020. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int. J. Syst. Evol. Microbiol.* **70**, 5607–5612.