Korean Journal of Microbiology (2022) Vol. 58, No. 3, pp. 205-207
DOI https://doi.org/10.7845/kjm.2022.2054

pISSN 0440-2413
eISSN 2383-9902

Draft genome sequences of *Sabulilitoribacter multivorans* KCTC 32326^T and *Sabulilitoribacter arenilitoris* KCTC 52401^T

Ji-Sung Oh and Dong-Hyun Roh*

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

Sabulilitoribacter multivorans KCTC 32326^{T}와 *Sabulilitoribacter arenilitoris* KCTC 52401^{T}의 유전체 염기서열 분석

오지성 · 노동현*

충북대학교 대학원 생명시스템학과

(Received August 22, 2022; Revised September 1, 2022; Accepted September 1, 2022)

The draft genome sequences of *Sabulilitoribacter multivorans* KCTC 32326^T and *Sabulilitoribacter arenilitoris* KCTC 52401^T were determined using Illumuna Hiseq X-ten platform. The assembled genome of *Sabulilitoribacter multivorans* KCTC 32326^T was composed of 13 contigs with a total length of 3,278,864 bp and G + C content was 32.7%. The assembled genome of *Sabulilitoribacter arenilitoris* KCTC 52401^T comprises 41 contigs with a total length of 3,491,608 bp and the genomic DNA G + C content was 32.0%. Two genomes showed differences in carotenoid biosynthesis, nitrogen metabolism and carbohydrate-active enzymes. These genome information of type strains will be expected to be helpful for classification of other isolated *Sabulilitoribacter* strains.

Keywords: *Sabulilitoribacter multivorans* KCTC 32326^T, *Sabulilitoribacter arenilitoris* KCTC 52401^T, draft genome sequence

The genus *Sabulilitoribacter*, belong to the family *Flavobacteriaceae*, was first proposed by Park *et al.* (2013) and currently comprises 2 recognized species, including *Sabulilitoribacter multivorans* and *S. arenilitoris*. The members of the genus *Sabulilitoribacter* were isolated from sand of seashore, and were characterized by Gram-stain-negative, aerobic, catalase- and oxidase-positive, non-motile, rod shaped. Herein, we report the draft genome sequences and annotations of *S. multivorans* KCTC 32326^T and *S. arenilitoris* KCTC 52401^T.

Sabulilitoribacter multivorans KCTC 32326^T and *S. arenilitoris* KCTC 52401^T were obtained from the Korean Collection for Type Cultures (KCTC) and revival and routinely cultured on marine agar 2216 (Difco) at 30°C for 3 days. Genomic DNA was extracted using MagAttract® HMW DNA kit (Qiagen) according to the manufacturer's instructions. The draft genome sequencing was performed on the Illumina Hiseq X-ten platform with TruSeq Nano DNA (350 bp insert size) library by Macrogen Inc. Raw reads were qualified by FastQC (version 0.11.5) and were assembled by SPAdes (version 3.13.0) (Bankevich *et al.*, 2012) or Platanus-allee (version 2.2.2) (Kajitani *et al.*, 2019). Orthologous average nucleotide identity (OrthoANI) and 16S rRNA gene sequence similarity were calculated using the Orthologous Average Nucleotide Identity Tool (http://www.ezbiocloud.net/tools/orthoani) and the online pairwise sequence alignment tool for the taxonomy (http://www.ezbiocloud.net/tools/pairAlign), respectively (Yoon *et al.*, 2017). The genome annotation was conducted using NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova *et al.*, 2016), and additional function of the predicted genes were conducted by BlastKOALA with KEGG database (Kanehisa

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

Table 1. Genomic features of *Sabulilitoribacter multivorans* KCTC 32326^T and *Sabulilitoribacter arenilitoris* KCTC 52401^T

Property	Value	
	S. multivorans KCTC 32326^T	*S. arenilitoris* KCTC 52401^T
Genome assembly		
Assemble method	SPAdes (version 3.15.0)	Paltanus-allee (version 2.2.2)
Genome coverage	147.9X	149.3X
Sequencing technology	Illumina Hiseq X-ten	Illumina Hiseq X-ten
Genome features		
Genome size (bp)	3,278,864	3,491,608
Number of scaffolds	11	37
Number of contigs	13	41
G + C content (%)	32.7	32.0
Protein-coding genes (CDSs)	2,886	2,945
rRNA genes (5S, 16S, 23S)	1, 2, 1	1, 1, 1
tRNA genes	38	34
ncRNA genes	4	4
Pseudogenes	8	56
Accession number (GenBank)	JAKKDV000000000	JAKKDU000000000

et al., 2016) and EggNOG 5.0 (Huerta-Cepas *et al.*, 2018). The dbCAN2 meta server was used for carbohydrate-active enzyme (CAZyme) annotation (Zhang *et al.*, 2018). A Bacterial Pan Genome Analysis pipeline (BPGA) was used to find the number of core gene (Chaudhari *et al*., 2016).

The genome features of *S. multivorans* KCTC 32326^T and *S. arenilitoris* KCTC 52401^T are shown in Table 1. The draft genome of *S. multivorans* KCTC 32326^T contained 11 scaffolds with a total length of 3,278,864 bp (N50 value, 766,194 bp). The G + C content was 32.7%, and 2,886 protein-coding genes, 4 rRNA genes, 38 tRNA genes, 4 non-coding RNA genes and 8 pseudo genes were predicted. The draft genome of *S. arenilitoris* KCTC 52401^T was composed of 37 scaffolds with a total length of 3,491,608 bp (N50 value, 204,683 bp). The G + C content was 32.0%, and 2,945 protein-coding genes, 3 rRNA genes, 34 tRNA genes, 4 non-coding RNA genes and 56 pseudo genes were predicted. Average nucleotide identity (ANI) value and 16S rRNA gene sequence similarity value between *S. multivorans* KCTC 32326^T and *S. arenilitoris* KCTC 52401^T were 79.3% and 96.9%, respectively. The number of core genes in both strains was 1,863 and a unique gene of each strain accounted for about one-third.

While both genomes contained carotenoid biosynthesis related genes such as 15-cis-phytoene synthase CrtB, phytoene desaturase CrtI and β-carotene 3-hydroxylase CrtZ, lycopene β-cyclase CrtY, β-carotene/zeaxanthin 4-ketolase CrtW and zeaxanthin glucosyltransferase CrtX were found only in *S. arenilitoris* KCTC 52401^T. In the nitrogen metabolism, both genomes commonly contained nitrous-oxide reductase NosZ, but ferredoxin-nitrate reductase NarB and nitrite reductase (NADH) large subunit NirB were found only in *S. multivorans* KCTC 32326^T. As a result of CAZyme annotation, *S. multivorans* KCTC 32326^T involved 138 CAZymes (4 auxiliary activities, 3 carbohydrate-binding modules, 11 carbohydrate esterases, 55 glycoside hydrolases, 49 glycosyl transferases and 16 polysaccharide lyases), and *S. arenilitoris* KCTC 52401^T involved 149 CAZymes (1 auxiliary activity, 2 carbohydrate-binding modules, 8 carbohydrate esterases, 84 glycoside hydrolases, 53 glycosyl transferases and 1 polysaccharide lyases). Both genomes common encoded biopolymeric degradation related genes such as exodeoxyribonuclease III XthA, β-glucosidase BglX, α-glucosidase MalZ, β-galactosidase LacZ, xylose isomerase XylA and xylose kinase XylB. α-Amylase AmyA, pullanase PulA and xylan 1,4-β-xylosidase XynB were only found in *S. multivorans* KCTC 32326^T, and chitinase was only found in *S. arenilitoris* KCTC 52401^T. These genomic data might be useful for the genetic classification of the genus *Sabulilitoribacter*, and for study the ecological status and organic material circulation of the marine *Flavobacteriaceae*.

Nucleotide sequence accession numbers

The draft genome sequence of *Sabulilitoribacter multivorans* KCTC 32326^T and *Sabulilitoribacter arenilitoris* KCTC 52401^T has been deposited to GenBank under the accession number JAKKDV000000000 and JAKKDU000000000, respectively. The version described in this paper are JAKKDV010000000 and JAKKDU010000000.

적 요

Sabulilitoribacter multivorans KCTC 32326^{T}와 *Sabulilitoribacter arenilitoris* KCTC 52401^{T}의 유전체 초안을 Illumina Hiseq X-ten platform을 사용하여 결정하였다. *Sabulilitoribacter multivorans* KCTC 32326^{T}의 조립된 유전체는 전체 길이 3,278,864 bp의 13개 contig로 구성되었고 G + C 함량은 32.7% 이었다. *Sabulilitoribacter arenilitoris* KCTC 52401^{T}의 조립된 유전체는 전체 길이 3,491,608 bp의 41개 contig로 구성되었고 G + C 함량은 32.0% 이었다. 이들 유전체는 carotenoid 생합성, 질소 대사 및 탄수화물 활성 효소 들에서 차이를 보였다. 이러한 type 균주의 유전 정보는 다른 분리된 *Subulilitoribacter* 균의 분류에 도움이 될 것으로 기대된다.

Acknowledgments

This research was supported by Chungbuk National University Korea National University Development Project (2021).

Conflict of Interest

The authors have no conflict of interest to report.

References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, *et al.* 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477.

Chaudhari NM, Gupta VK, and Dutta C. 2016. BPGA-an ultra-fast pan-genome analysis pipeline. *Sci. Rep.* **6**, 24373.

Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, *et al.* 2018. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314.

Kajitani R, Yoshimura D, Okuno M, Minakuchi Y, Kagoshima H, Fujiyama A, Kubokawa K, Kohara Y, Toyoda A, and Itoh T. 2019. Platanus-allee is a *de novo* haplotype assembler enabling a comprehensive access to divergent heterozygous regions. *Nat. Commun.* **10**, 1702.

Kanehisa M, Sato Y, and Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* **428**, 726–731.

Park S, Jung YT, Lee JS, Lee KC, and Yoon JH. 2013. *Sabulilitoribacter multivorans* gen. nov., sp. nov., a polysaccharide-degrading bacterium of the family *Flavobacteriaceae* isolated from seashore sand. *Antonie van Leeuwenhoek* **104**, 973–981.

Tatusova T, DiCuccio M, Badretdin A, Chetvernin 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.

Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, and Chun J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA and whole genome assemblies. *Int. J. Syst. Evol. Microbiol.* **67**, 1613–1617.

Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, and Yin Y. 2018. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **46**, W95–W101.