Korean Journal of Microbiology (2018) Vol. 54, No. 4, pp. 465-467
DOI https://doi.org/10.7845/kjm.2018.8088

pISSN 0440-2413
eISSN 2383-9902

Complete genome sequence of *Niabella ginsenosidivorans* BS26^T, a ginsenoside-converting bacterium, isolated from compost

Young-Woo Lee[1], Muhammad Zubair Siddiqi[1,2], Qing-Mei Liu[1,2], Dae-Cheol Kim[1], and Wan-Taek Im[1,2]*

[1]*Department of Biotechnology, Hankyong National University, Anseong 17579, Republic of Korea*
[2]*AceEMzyme Co., Ltd., Academic Industry Cooperation, Anseong 17579, Republic of Korea*

퇴비에서 분리한 진세노사이드 전환능력이 있는 *Niabella ginsenosidivorans* BS26^{T}의 유전체 서열 분석

이영우[1] · 시디키 무하마드 주베르[1,2] · 류청매[1,2] · 김대철[1] · 임완택[1,2]*

[1]국립한경대학교 농업생명과학대학 생명공학과, [2]주식회사 에이스엠자임

(Received November 20, 2018; Revised November 30, 2018; Accepted December 6, 2018)

An orange-colored, rod-shaped strain, designated *Niabella ginsenosidivorans* BS26^T, was isolated from compost. Strain BS26^T showed the ability to convert major ginsenosides to minor ginsenosides, and its whole genome was sequenced. The whole genome of *N. ginsenosidivorans* BS26^T consists of a single circular chromosome of 5,627,734 bp with 44.48% G + C content. Based on the complete genome sequence of strain BS26^T, we found several glycosides hydrolase-encoding genes that might involve in the conversion of major ginsenosides into minor ginsenoside and deliberate its strong pharmacological effects.

Keywords: *Niabella ginsenosidivorans*, complete genome, compost, glycoside hydrolase, PacBio RS II

The genus *Niabella* was first described by Kim *et al*. (2007). Species of this genus are non-flagellated, non-spore-forming, and Gram negative. The DNA G + C content ranges between 42.0 mol% and 47.5 mol%. Currently, the genus comprises of 10 recognized species with published names (LPSN, http://www.bacterio.net/niabella.html), which were commonly isolated from various sources such as soil, medicinal leeches, and lake water (Kim *et al*., 2007; Siddiqi and Im, 2016).

To identify a ginsenoside transforming positive bacterium, a Gram-negative bacterium, *N. ginsenosidivorans* BS26^T, was isolated from compost (decayed feedstuff) in the Republic of Korea. *N. ginsenosidivorans* BS26^T was orange coloured, non-spore forming and non-motile rod bacterium. Based on the production of minor ginsenosides from major ginsenosides (Yi *et al*., 2015; Siddiqi *et al.*, 2017), strain *N. ginsenosidivorans* BS26^T was selected for a whole genome study to identify the target functional genes. Whole genome sequence analysis showed more than 40 glycoside hydrolases that may involve in the conversion of ginsenosides. This strain is available from the host institute and from two culture collections (= KACC 16620^T = JCM 18199^T).

The genomic DNA of *N. ginsenosidivorans* BS26^T was extracted and purified with the genomic DNA extraction kit (Biofact), and was sequenced using the Pacific Biosciences RSII platform. And a library was constructed according to the Pacific Biosciences RSII Sequencing method manual. The general features about the complete genome sequencing and library construction are available at the JGI website (https://

***For correspondence.** E-mail: wandra@hknu.ac.kr;
Tel.: +82-31-670-5335; Fax: +82-31-670-5339

www.jgi.doe.gov). PacBio SMRT Analysis (version 2.3.0) was used with default options for sequence reads assembly, and the protein coding sequences (CDSs) were predicted by Glimmer 3.02 (Delcher *et al*., 1999). The genome sequence was annotated through the NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAP; http://www.ncbi.nlm.nih.gov/books/NBK174280/). The tRNAs and rRNAs were predicted by using rRNAmmer and tRNAscan-SE, respectively.

The complete genome of *N. ginsenosidivorans* BS26^T consists of one circular chromosome of 5,627,734 bp with 44.48% G + C content. Out of the 4,756 predicted genes, 4,704 were protein-coding genes (CDS), and 49 were RNAs. Moreover, 85 pseudogenes were also identified (Fig. 1). The majority of the protein-coding genes (98.81%) were assigned with a putative function, while the remaining predicted genes were annotated as hypothetical or conserved hypothetical proteins. The genome statistics are presented in Table 1. Analysis of the complete genome sequence showed many glycoside hydrolase-encoding genes including 14 β- glucosidases, 11 α-glucosidases, 8 β- xylosidases, 8 α-arabinofuranosidases, and 1 β-arabinofuranosidase, which may be responsible for its ability to convert ginsenosides. In addition, the genome also encoded a nitrite reductase (NADH) small subunit, mercury resistance protein (MerC), quinone reductase or related Zn-dependent oxidoreductase, putative oxidoreductase, ferredoxin-NADP reductase, sulfatase, sulfite reductase, sulfate degradation, Clp endopeptidase (endonuclease/exonuclease), thiol-disulfide oxidoreductase, protein involved in L-lysine biosynthesis, and multiple antibiotic resistance proteins. The presence of these

Table 1. General features of *Niabella ginsenosidivorans* BS26^T

Features	Chromosome
Genome size (bp)	5,627,734
DNA coding region (bp)	4,908,857
G + C content (%)	44.48
Total genes	4,756
Pseudo genes	85
Coding sequences (CDSs)	4,704
Number of rRNA genes (5S, 16S, 23S)	6 (2, 2, 2)
Number of tRNA genes	43

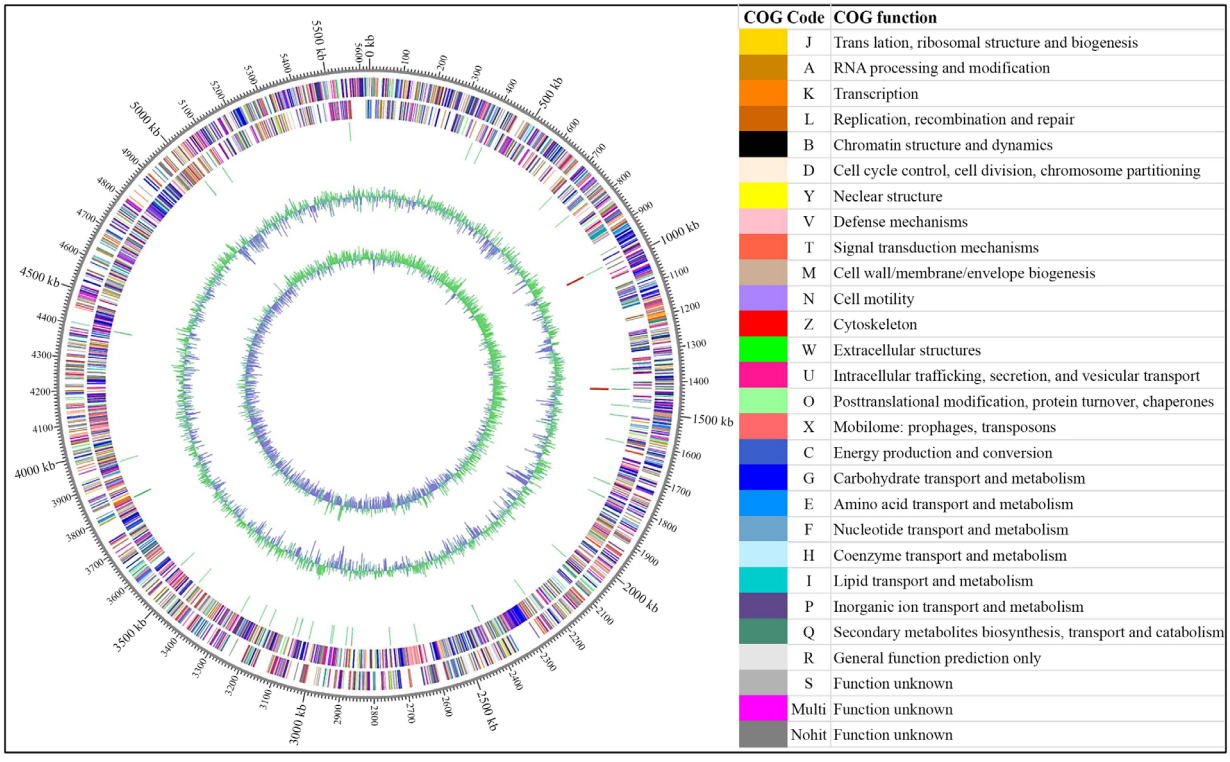

Fig. 1. Graphical map of the chromosome of *Niabella ginsenosidivorans* BS26^T. The rings from the outside to the center show the following: genes on the forward strand (colored by COG category), genes on the reverse strand (colored by COG category), RNA genes (tRNAs, green; rRNAs, red; other RNAs, black), GC content, and GC skew.

genes shows the importance of the bacterial group represented by this strain for the cycling of organic and inorganic elements.

The availability of the whole genome sequence of *N. ginsenosidivorans* BS26^T will allow further functional and comparative analyses to understand the genomic traits well involved in the conversion of plant secondary metabolites as described by Siddiqi *et al*. (2017).

Nucleotide sequence accession number

The complete genome sequence of *Niabella ginsenosidivorans* BS26^T has been deposited at DDBJ/EMBL/NCBI GenBank under accession number CP015772.

적 요

퇴비로부터 분리한 *Niabella ginsenosidivorans* BS26^T 균주의 유전체서열을 분석하였다. 균주 BS26^{T}의 유전체는 G + C 비율이 44.48%이며, 4,800개의 유전자와 4,704개의 단백질 코딩 유전자, 85개의 위유전자 그리고49개의 RNA유전자를 포함한 단일 원형 염색체로 구성되었으면 그 크기는 5,627,734 bp였다. 균주 BS26^{T}는 인삼사포닌의 당 분해에 관여하는 여러 타입의 글라이코시다제 유전자를 가지고 있었다. 이러한 유전체 분석은 주요 진세노사이드 전환에 관여하는 유전자 특징을 이해하는데 큰 기여가 되었다.

Acknowledgements

This research was supported by the project for the survey and excavation of Korean indigenous species of the National Institute of Biological Resources (NIBR) and the Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ012283) of the Rural Development Administration, Republic of Korea.

References

Delcher AL, Harmon D, Kasif S, White O, and Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27**, 4636-4641.

Kim BY, Weon HY, Yoo SH, Hong SB, Kwon SW, Stackebrandt E, and Go SJ. 2007. *Niabella aurantiaca* gen. nov., sp. nov., isolated from a greenhouse soil in Korea. *Int. J. Syst. Evol. Microbiol*. **57**, 538-541.

Siddiqi MZ, Cui CH, Park SK, Han NS, Kim SC, and Im WT. 2017. Comparative analysis of the expression level of recombinant ginsenoside-transforming β-glucosidase in GRAS hosts and mass production of the ginsenoside Rh2-Mix. *PLoS One* **12**, 1371-1385.

Siddiqi MZ and Im WT. 2016. *Niabella aquatica* sp. nov., isolated from lake water. *Int. J. Syst. Evol. Microbiol*. **66**, 2774-2779.

Yi KJ, Im WT, Kim DW, Liu QM, and Kim SK. 2015. *Niabella ginsenosidivorans* sp. nov., isolated from compost. *J. Microbiol.* **53**, 762-766.