

Genome Announcement

Complete genome sequence of *Bacillus subtilis* BS16045 isolated from Gochujang

SaeBom Jeon, Jun Heo, and Tai-Boong Uhm*

Department of Biological Sciences, Chonbuk National University, Jeonju 54896, Republic of Korea

고추장에서 분리된 *Bacillus subtilis* BS16045의 유전체 서열 분석

전새봄 · 허 준 · 엄태봉*

전북대학교 자연과학대학 생명과학과

(Received December 26, 2016; Revised December 27, 2016; Accepted December 28, 2016)

Bacillus subtilis BS16045 was isolated from Gochujang, a Korean red chili paste, in order to get a starter strain that can be used for preservation of the fermented foods. We report the whole genome sequence of *B. subtilis* BS16045, which contains 4,165,121 bp with a G+C content of 43.6%. We also confirmed the set of antibiotic genes producing surfactin, kanosamine, bacillaene, plipastatin, subtilosin A, and bacilysin, which are related to antifungal and antibacterial activities. These results indicate that *B. subtilis* BS16045 could be a potential starter strain for solving contamination by food-borne pathogens in the soybean products factory.

Keywords: *Bacillus subtilis*, antimicrobial activity, complete genome sequence, traditional fermented food

Naturally fermented foods without starter strain are frequently exposed to the risk of food-borne pathogens responsible for food poisoning. We have focused on the selection of *Bacillus* species indigenous to traditional Korean foods for the safety of soybean products and isolated *Bacillus subtilis* BS16045 from Gochujang. The strain exhibits the inhibition activities against pathogens including *Bacillus cereus*, *Escherichia coli*, *Micrococcus luteus*, and *Candida albicans*. *B. subtilis* is a Gram-positive,

spore-forming and rod-shaped bacteria. This species is generally recognized as safety (GRAS) organism for the fermentation of soybean products that are considered as the essential sauces of authentic Korea cuisine (Yeo *et al.*, 2012). It can produce various types of antibiotics that may act as the food preservatives (Stein, 2005). Recently, whole genome sequencing and RNA-sequencing of industrially important *B. subtilis* strains enabled us to give the entire picture of the gene-related fermentation progress (Kunst *et al.*, 1997; Brinsmade *et al.*, 2014; Kamada *et al.*, 2015). To confirm the antipathogenic factors, we conducted whole genome sequencing of the strain BS16045.

The genomic DNA was extracted from cells grown at the early stationary phase using a Wizard genomic DNA isolation kit (Promega). The whole genome was constructed *de novo* using Pacific Biosciences (PacBio) sequencing data. Sequencing analysis was performed by Chunlab Inc. All the sequences were assembled using CLC Genomic Workbench (CLCbio) and SMRT analysis version 2.3.0 (Pacific Biosciences). The alignment showed about 206 x coverage. The complete genome sequence of BS16045 is composed of a single circular chromosome. Coding DNA sequences (CDSs), tRNA and CRISPRs were predicted by using Prodigal (Hyatt *et al.*, 2010) and tRNAscan-SE. rRNAs (Lowe and Eddy, 1997) and other non-coding RNAs were searched using Rfam 12.0 database

*For correspondence. E-mail: tbuhm@chonbuk.ac.kr;
Tel.: +82-63-270-3439; Fax: +82-63-270-3362

Table 1. Genome features of *Bacillus subtilis* BS16045

Features	Chromosome
Genome size (bp)	4,165,121
GC content (%)	43.6
rRNA genes	30
tRNA genes	86
Protein coding genes	4,209
Protein coding genes with function prediction	4,113

(Griffiths-Jones *et al.*, 2003). Each of the predicted CDSs were annotated and classified into different groups based on the orthologous group using EggNOG 4.1 (Jensen *et al.*, 2008), Swiss-prot, SEED and KEGG databases. The basic genome statistics are provided in Table 1. The complete genome size is 4,165,121 bp with a G+C content of 43.6%. As a result of gene prediction, this genome contains 4,209 CDSs, 86 tRNA, and 30 rRNA genes. We also confirmed various genes related to antipathogenic activity in the genome of BS16045. These genes are related to the synthesis of surfactin (BS16045_00398–BS16045_00401), kanosamine (BS16045_01127–BS16045_01129), bacillaene (BS16045_01814–BS16045_01829), plipastatin (BS16045_01946–BS16045_01950), subtilosin A (BS16045_03925–BS16045_03932), and bacilysin (BS16045_03964–BS16045_03970). Compared to the genome of the type strain *B. subtilis* 168 (accession number AL009126.3), the BS16045 genome lacks the genes for the synthesis and immunity of sublancin, an antimicrobial glycopeptide (Kunst *et al.*, 1997). The antibiotics synthesized by the expression of these genes would take a role in both antifungal and antibacterial activities (Paik *et al.*, 1998). Especially, clear inhibition zones toward several *B. cereus* strains (KACC10097, KACC12672, and KACC12682) indicated that *B. subtilis* BS16045 could be a potential starter for solving contamination problem by *B. cereus* in the soybean products factory.

Nucleotide sequence accession numbers

The *Bacillus subtilis* BS16045 is available at KCCM 43240. The complete genome sequence of *Bacillus subtilis* BS16045 was deposited in GenBank under the accession No. CP017112.

적 요

한국 발효식품의 보존성을 높이기 위한 스타터 균주를 얻기 위하여 고추장에서 *Bacillus subtilis* BS16045를 분리하였다. *B. subtilis* BS16045에 대한 유전체 분석을 실시하였으며, G+C 비율이 43.6%인 4,165,121 bp 크기의 염기서열을 얻었다. 또한 이 유전체로부터 항진균 및 항균 활성에 연관이 있는 surfactin, kanosamine, bacillaene, plipastatin, subtilosin A, bacilysin 생산 유전자들을 확인하였다. 이러한 결과들을 통해 *B. subtilis* BS16045는 장류 제조시설에서 유해균의 오염문제를 해결할 수 있는 스타터로 이용될 수 있을 것으로 보인다.

Acknowledgements

This research was supported by High value-added Food Technology Development Program, of the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries (IPET), and the Ministry for Food, Agriculture, Forestry, and Fisheries of Republic of Korea (313037-03).

References

- Brinsmade, S.R., Alexander, E.L., Livny, J., Stettner, A.I., Segrè, D., Rhee, K.Y., and Sonenshein, A.L. 2014. Hierarchical expression of genes controlled by the *Bacillus subtilis* global regulatory protein CodY. *Proc. Natl. Acad. Sci. USA* **111**, 8227–8232.
- Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A., and Eddy, S.R. 2003. Rfam: an RNA family database. *Nucleic Acids Res.* **31**, 439–441.
- Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 1.
- Jensen, L.J., Julien, P., Kuhn, M., von Mering, C., Muller, J., Doerks, T., and Bork, P. 2008. eggNOG: automated construction and annotation of orthologous groups of genes. *Nucleic Acids Res.* **36**, 250–254.
- Kamada, M., Hase, S., Fujii, K., Miyake, M., Sato, K., Kimura, K., and Sakakibara, Y. 2015. Whole-genome sequencing and comparative genome analysis of *Bacillus subtilis* strains isolated from non-salted fermented soybean foods. *PLoS One* **10**, 1.
- Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.O., Azevedo, V., Bertero, M.G., Bessleres, P., Bolotin, A., Borriss, R., *et al.* 1997. The complete genome sequence of the

Gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249–256.

- Lowe, T.M. and Eddy, S.R.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964.
- Paik, S.H., Chakicherla, A., and Hansen, J.N.** 1998. Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168. *J. Biol. Chem.* **273**, 23134–23142.
- Stein, T.** 2005. *Bacillus subtilis* antibiotics: structures, syntheses and specific functions. *Mol. Microbiol.* **56**, 845–857.
- Yeo, I.C., Lee, N.K., and Hahm, Y.T.** 2012. Genome sequencing of *Bacillus subtilis* SC-8, antagonistic to the *Bacillus cereus* group, isolated from traditional Korean fermented-soybean food. *J. Bacteriol.* **194**, 536–537.