AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Genome Sequence of *Bacillus subtilis* natto VK161, a Novel Strain That Produces Vitamin K_2

Dylan Parks,[a] In-Hwa Chung,[b] In kyu Lee,[b] Eui-Joong Kim,[c] Sarbjeet Niraula,[a] Woo-Suk Chang[a]

[a]Department of Biology, University of Texas, Arlington, Texas, USA
[b]SungWun Pharmacopia/Bio Co., Ltd., Seoul, Republic of Korea
[c]GF Fermentech, Inc., Sejong, Republic of Korea

ABSTRACT *Bacillus subtilis* strain natto VK161 was selected for its high production of vitamin K_2. Its genome was sequenced and annotated in the Department of Energy-Joint Genome Institute (DOE-JGI) annotation pipeline. It resulted in a chromosome of 4,073,396 bp, which is composed of 4,332 protein-coding genes, 23 rRNA genes, and 77 tRNA genes.

Bacillus subtilis natto is a Gram-positive, aerobic, spore-forming bacterium closely related to the laboratory strain *B. subtilis* Marburg 168, which was the first sequenced genome of the *B. subtilis* family (1). Although there are several natto strains, only the *B. subtilis* natto BEST 195 strain has been sequenced completely (2). *B. subtilis* natto is known to produce poly-γ-glutamate (γ-PGA), which gives a characteristic slimy texture to the fermented soybean product called natto (3). In addition, its potential applications in therapeutics (4) and probiotics (5) have been reported in the scientific community.

Bacillus subtilis natto VK161 was selected for its high production of menaquinone, known as vitamin K_2, using chemical mutagenesis with *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine (NTG). Briefly, the parental strain *B. subtilis* natto KCCM 12027 obtained from the Korean Culture Center of Microorganisms (KCCM) was preexposed to NTG. We screened for vitamin K_2 production through the NTG-induced mutagenesis library and found that a single mutant colony, subsequently named *B. subtilis* natto VK161, produced the highest level of vitamin K_2 (ca. 120 mg/liter). Here, we report the genome sequence of *B. subtilis* natto VK161. Strain VK161 was cultured in Luria-Bertani (LB) medium overnight at 37°C with shaking at 200 rpm. Genomic DNA (gDNA) of *B. subtilis* natto VK161 was isolated using the G-spin genomic extraction kit (iNtRON Biotechnology). The amount of gDNA (>250 ng/μl) and its quality (260/280 ratio, ≈1.8) were determined using a NanoDrop spectrophotometer (Thermo Fisher Scientific). Gel electrophoresis was also used to inspect gDNA, confirming that a band was above 10 kb of a DNA ladder marker. Sequencing was performed at the University of Texas at the Austin Genomic Sequencing and Analysis Facility (GSAF) on an Illumina MiSeq platform with paired-end (PE) 2 × 300-bp run specifications. Sequencing produced 2,260,886 reads for approximately 170× coverage. Raw reads were quality filtered, and adapter sequences were subsequently removed using BBDuk, one of the BBTools developed at the U.S. Department of Energy-Joint Genome Institute (DOE-JGI), using the parameters ktrim=r, k = 23, mink = 11, hdist = 1, qtrim=r, trimq = 10, ftm = 5, maq = 25, and minlen = 100. Filtered read quality was ensured by inspection using FastQC (6). Default parameters were used for all other software programs, unless otherwise specified. Assembly of filtered reads was performed using SPAdes 3.10.1 (7) and assembly quality assessed via QUAST 4.6.1 (8). The genome assembly contained 96 contigs (≥1,000 bp) totaling 4.1 Mbp, with an N_{50} value of 78,653 bp.

Citation Parks D, Chung I-H, Lee IK, Kim E-J, Niraula S, Chang W-S. 2019. Genome sequence of *Bacillus subtilis* natto VK161, a novel strain that produces vitamin K_2. Microbiol Resour Announc 8:e00444-19. https://doi.org/10.1128/MRA.00444-19.

Editor David A. Baltrus, University of Arizona

Address correspondence to Woo-Suk Chang, wschang@uta.edu.

Received 18 April 2019
Accepted 29 July 2019
Published 29 August 2019

Functional annotation and gene prediction were performed using the DOE-JGI Microbial Genome Annotation Pipeline (9). JGI's Integrated Microbial Genomes (IMG) system (10) showed that the complete genome of strain VK161 consists of 4,073,396 bp, with 4,332 coding genes (63.9% with predicted functions), 23 rRNA genes, 77 tRNA genes, and an average G+C content of 43.33%. Further annotation and gene prediction were performed using JGI's IMG-Expert Review (IMG ER) software (11).

Data availability. The whole-genome shotgun project of *B. subtilis* natto VK161 has been deposited into NCBI GenBank under the accession number SJSU00000000. The version described in this paper is the first version, SJSU01000000. Raw sequences were deposited in the NCBI SRA database under the accession number SRP192779.

ACKNOWLEDGMENT

We thank the DOE-JGI for providing Microbial Genomics and Metagenomics Workshops for Dylan Parks and Sarbjeet Niraula.

REFERENCES

1. Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, Bertero MG, Bessières P, Bolotin A, Borchert S, Borriss R, Boursier L, Brans A, Braun M, Brignell SC, Bron S, Brouillet S, Bruschi CV, Caldwell B, Capuano V, Carter NM, Choi SK, Cordani JJ, Connerton IF, Cummings NJ, Daniel RA, Denziot F, Devine KM, Düsterhöft A, Ehrlich SD, Emmerson PT, Entian KD, Errington J, Fabret C, Ferrari E, Foulger D, Fritz C, Fujita M, Fujita Y, Fuma S, Galizzi A, Galleron N, Ghim SY, Glaser P, Goffeau A, Golightly EJ, Grandi G, Guiseppi G, Guy BJ, Haga K, et al. 1997. The complete genome sequence of the Gram-positive bacterium *Bacillus subtlis*. Nature 390:249–256. https://doi.org/10.1038/36786.
2. Kamada M, Hase S, Sato K, Toyoda A, Fujiyama A, Sakakibara Y. 2014. Whole genome complete resequencing of *Bacillus subtilis* natto by combining long reads with high-quality short reads. PLoS One 9:e109999. https://doi.org/10.1371/journal.pone.0109999.
3. Inatsu Y, Kimura K, Itoh Y. 2002. Characterization of *Bacillus subtilis* strains isolated form fermented soybean foods in Southeast Asia: comparison with *B. subtilis* (natto) starter strains. Jpn Agric Res Q 36:169–175. https://doi.org/10.6090/jarq.36.169.
4. Dabbagh F, Negahdaripour M, Berenjian A, Behfar A, Mohammadi F, Zamani M, Irajie C, Ghasemi Y. 2014. Nattokinase: production and application. Appl Microbiol Biotechnol 98:9199–9206. https://doi.org/10.1007/s00253-014-6135-3.
5. Hosoi T, Ametani A, Kiuchi K, Kaminogawa S. 2000. Improved growth and viability of lactobacilli in the presence of *Bacillus subtilis* (natto), catalase, or subtilisin. Can J Microbiol 46:892–897. https://doi.org/10.1139/w00-070.
6. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
7. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
8. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
9. Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Palaniappan K, Szeto E, Pillay M, Chen IM, Pati A, Nielsen T, Markowitz VM, Kyrpides NC. 2015. The standard operating procedure of the DOE-JGI microbial genome annotation pipeline (MGAP v.4). Stand Genomic Sci 10:86. https://doi.org/10.1186/s40793-015-0077-y.
10. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG: the integrated microbial genomes database and comparative analysis system. Nucleic Acids Res 40:D115–D122. https://doi.org/10.1093/nar/gkr1044.
11. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC. 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics 25:2271–2278. https://doi.org/10.1093/bioinformatics/btp393.