Complete genome sequence data of a broad-spectrum antipathogen, *Bacillus amyloliquefaciens* KCTC 18343P, isolated from *Makgeolli*, Korean traditional rice wine

Eun-Hee Park, Hyunsu Sim, Myoung-Dong Kim*

Division of Food Biotechnology and Biosystems Engineering, Kangwon National University, Chuncheon, 24341, Republic of Korea

**Abstract**

*Bacillus amyloliquefaciens* KCTC18343P (≡MBE1283) isolated from *Makgeolli*, Korean traditional rice wine, strongly inhibits the growth of food and plant pathogens. A complete genome sequence of *B. amyloliquefaciens* KCTC18343P is presented in this report. The genome is 3,979,925 bp in size and harbors 3856 genes. The BioProject has been deposited at DDBJ/EMBL/GenBank. The GenBank accession numbers are PRJNA301202 for the BioProject, NZ_CP013727 for the chromosome, and NZ_CP013728 for the plasmid.

© 2020 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Data description

*Bacillus amyloliquefaciens* strains are reported to have antimicrobial activity exerted by the secondary metabolites, including bacillomycin D, fengycin, surfactin, subtilin, and subtilosin A [7,8]. The *B. amyloliquefaciens* strain KCTC18343P isolated from *Makgeolli*, Korean traditional rice wine, exhibited strong antimicrobial activity against the food and plant pathogens *Aspergillus terreus*,

*Corresponding author.
E-mail address: y82345@gmail.com (M.-D. Kim).
Aspergillus flavus, Staphylococcus aureus, Escherichia coli O157:H7, Bacillus cereus, Salmonella Typhimurium, and Listeria monocytogenes, which is consistent with previous reports [1,11]. We previously reported that B. amyloliquefaciens KCTC18343P is also antagonistic to the human pathogen Cryptococcus neoformans [11]. This strain has been deposited as KCTC18343P in the Korean Collection for Type Cultures (KCTC). The corresponding BioProject has been deposited at DDBJ/EMBL/GenBank. The GenBank accession number are PRJNA301202 for the BioProject, CP013727 for the chromosome, and NZ_CP013728 for the plasmid.

A total of 512,192,070 paired-end reads with an average read length of 5843 base pairs (bp) were obtained. The reads were assembled de novo into two contigs using a Hierarchical Genome Assembly Process (HGAP). The genome of strain KCTC18343P consisted of a single circular chromosome and one circular plasmid with a total length of 3,979,925 bp and 46.5% G + C content. In total, 3856 gene-coding regions, 27 rRNA operons, and 86 tRNAs were identified (Table 1). A genome-wide comparative analysis of two B. amyloliquefaciens strains, KCTC18343P and DSM7, revealed that the genome of KCTC18343P contains the gene clusters srf and dhb, which encode enzymes involved in biosynthesis of lipopeptide surfactin and bacillibactin, respectively [3,13]. Nine genes associated with macrolactin H, bacillaene, fengycin, difficidin, bacillibactin, bacilysin, mersacidin, surfectin, and butirosin A were identified by antiSMASH [2]. The gene cluster bacillibactin of B. amyloliquefaciens KCTC18343P corresponded to that of B. subtilis 168 (similarity 100%). The gene cluster srf of B. amyloliquefaciens KCTC18343P was 82% similar to that of B. velezensis FZB42. The complete genome of B. amyloliquefaciens KCTC18343P showed the highest Average Nucleotide Identity (ANI) value of 99.41% with those of B. amyloliquefaciens Y14 (99.41%) and B. amyloliquefaciens JRS8 (99.27%; Table 2).

Plasmid pBABEL01 is a 13,003 bp circular molecule with a G + C content of 42.6% and harbors three components, replication initiator protein, double-strand origin, and single-strand origin [9], which are found in rolling-circle replication plasmids. Bacillus amyloliquefaciens KCTC18343P isolated from Makgeolli (Korean rice wine) exhibits strong and broad antimicrobial activity against pathogens. The data on genome sequence of strain KCTC18343P can be used to search and characterize biotechnology-relevant enzymes and gene clusters. Gene clusters srf and dhb which contribute to broad antipathogenic activity were identified by whole-genome analysis.

Plasmid pBABEL01 is a 13,003 bp circular molecule with a G + C content of 42.6% and harbors three components, replication initiator protein, double-strand origin, and single-strand origin [9], which are found in rolling-circle replication plasmids. Bacillus amyloliquefaciens KCTC18343P isolated from Makgeolli (Korean rice wine) exhibits strong and broad antimicrobial activity against pathogens. The data on genome sequence of strain KCTC18343P can be used to search and characterize biotechnology-relevant enzymes and gene clusters. Gene clusters srf and dhb which contribute to broad antipathogenic activity were identified by whole-genome analysis.

Aspergillus flavus, Staphylococcus aureus, Escherichia coli O157:H7, Bacillus cereus, Salmonella Typhimurium, and Listeria monocytogenes, which is consistent with previous reports [1,11]. We previously reported that B. amyloliquefaciens KCTC18343P is also antagonistic to the human pathogen Cryptococcus neoformans [11]. This strain has been deposited as KCTC18343P in the Korean Collection for Type Cultures (KCTC). The corresponding BioProject has been deposited at DDBJ/EMBL/GenBank. The GenBank accession number are PRJNA301202 for the BioProject, CP013727 for the chromosome, and NZ_CP013728 for the plasmid.

A total of 512,192,070 paired-end reads with an average read length of 5843 base pairs (bp) were obtained. The reads were assembled de novo into two contigs using a Hierarchical Genome Assembly Process (HGAP). The genome of strain KCTC18343P consisted of a single circular chromosome and one circular plasmid with a total length of 3,979,925 bp and 46.5% G + C content. In total, 3856 gene-coding regions, 27 rRNA operons, and 86 tRNAs were identified (Table 1). A genome-wide comparative analysis of two B. amyloliquefaciens strains, KCTC18343P and DSM7, revealed that the genome of KCTC18343P contains the gene clusters srf and dhb, which encode enzymes involved in biosynthesis of lipopeptide surfactin and bacillibactin, respectively [3,13]. Nine genes associated with macrolactin H, bacillaene, fengycin, difficidin, bacillibactin, bacilysin, mersacidin, surfectin, and butirosin A were identified by antiSMASH [2]. The gene cluster bacillibactin of B. amyloliquefaciens KCTC18343P corresponded to that of B. subtilis 168 (similarity 100%). The gene cluster srf of B. amyloliquefaciens KCTC18343P was 82% similar to that of B. velezensis FZB42. The complete genome of B. amyloliquefaciens KCTC18343P showed the highest Average Nucleotide Identity (ANI) value of 99.41% with those of B. amyloliquefaciens Y14 (99.41%) and B. amyloliquefaciens JRS8 (99.27%; Table 2).

Plasmid pBABEL01 is a 13,003 bp circular molecule with a G + C content of 42.6% and harbors three components, replication initiator protein, double-strand origin, and single-strand origin [9], which are found in rolling-circle replication plasmids. Bacillus amyloliquefaciens KCTC18343P isolated from Makgeolli (Korean rice wine) exhibits strong and broad antimicrobial activity against pathogens. The data on genome sequence of strain KCTC18343P can be used to search and characterize biotechnology-relevant enzymes and gene clusters. Gene clusters srf and dhb which contribute to broad antipathogenic activity were identified by whole-genome analysis.
Bacillus amyloliquefaciens KCTC18343P exhibits powerful antimicrobial activity compared to previously reported strains.

2. Experimental design, materials, and methods

The genomic DNA of B. amyloliquefaciens KCTC18343P was prepared from cells in the exponential growth phase. Genomic DNA was extracted using a G-DEX™IIc Genomic DNA Extraction Kit (iNtRON, Daejeon, Korea) according to the manufacturer’s instructions. The complete genome of B. amyloliquefaciens KCTC18343P was sequenced using the PacBio RSII (P6eC4) platform (Paciﬁc Biosciences, Menlo Park, CA, USA). The reads were assembled de novo into two contigs (175-fold coverage) using a hierarchical genome assembly process. Prokka was used to predict and subsequently annotate open reading frames [10]. Annotations were computed using eggNOG-mapper based on eggNOG 4.5 orthology data [5]. Pairwise average nucleotide identity between KCTC18343P and other B. amyloliquefaciens strains in the database was determined using the EZBio-Cloud Web server [12]. On antiSMASH bacterial version 5.1.1 (https://antismash.secondarymetabolites.org) was used to identify putative genes in the B. amyloliquefaciens KCTC18343P genome involved in the biological control of phytopathogens.

Acknowledgments

This work was carried out with the support of “Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ009993)” Rural Development Administration, Republic of Korea. We would like to thank Editage (www.editage.co.kr) for English language editing.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2020.105316.
References

[1] A.A. Anthony, O. Marc, H. Badre, L. Yannick, B. Aliain, J. Bernard, F. Patrick, *Bacillus amyloliquefaciens* GAI as a source of potent antibiotics and other secondary metabolites for biocontrol of plant pathogens, Microb. Cell Factories 8 (2009) 1–12.

[2] K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S.Y. Lee, M.H. Medema, T. Weber, antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline, Nucleic Acids Res. 47 (2019) 81–87.

[3] L. Chen, N. Wang, X. Wang, J. Hu, S. Wang, Characterization of two anti-fungal lipopeptides produced by *Bacillus amyloliquefaciens* SH-B10, Bioresearch. Technol. 101 (2010) 8822–8827.

[4] X.H. Chen, A. Koulmoutsi, R. Scholz, R. Borris, More than anticipated - production of antibiotics and other secondary metabolites by *Bacillus amyloliquefaciens* FZB42, J. Mol. Microbiol. Biotechnol. 16 (2009) 14–24.

[5] J. Huerta-Cepas, D. Szklarczyk, K. Forslund, H. Cook, D. Heller, M.C. Walter, T. Rattei, D.R. Mende, S. Sunagawa, M. Kuhn, L.J. Jensen, C. von Mering, P. Bork, eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences, Nucleic Acids Res. 44 (2016) 286–293.

[6] H. Jeong, S.H. Park, S.K. Choi, Genome sequence of antibiotic-producing *Bacillus amyloliquefaciens* strain KCTC 13012, Genome Announc. (2015) 3.

[7] J. Liu, T. Zhou, D. He, X.Z. Li, H. Wu, W. Liu, X. Gao, Functions of lipopeptides bacillomycin D and fengycin in antagonism of *Bacillus amyloliquefaciens* C06 towards *Monilinia fructicola*, J. Mol. Microbiol. Biotechnol. 20 (2011) 43–52.

[8] K.V. Pathak, A. Bose, H. Keharia, Identification and characterization of novel surfactins produced by fungal antagonist *Bacillus amyloliquefaciens* 68, Biotechnol. Appl. Biochem. 61 (2014) 349–356.

[9] J.Q. Qiao, W. Tian da, R. Huo, H.J. Wu, X.W. Gao, Functional analysis and application of the cryptic plasmid pBSG3 harboring the RapQ-PhrQ system in *Bacillus amyloliquefaciens* B3, Plasmid 65 (2011) 141–149.

[10] T. Seemann, Prokka: rapid prokaryotic genome annotation, Bioinformatics 30 (2014) 2068–2069.

[11] H.S. Sim, M.D. Kim, Antipathogenic activity of *Bacillus amyloliquefaciens* isolated from Korean traditional rice wine, Microbiol. Biotechnol. Lett. 44 (2016) 98–105.

[12] S.H. Yoon, S.M. Ha, J. Lim, S. Kwon, J. Chun, A large-scale evaluation of algorithms to calculate average nucleotide identity, Antonie Leeuwenhoek 110 (2017) 1281–1286.

[13] Y. Zhi, Q. Wu, Y. Xu, Genome and transcriptome analysis of surfactin biosynthesis in *Bacillus amyloliquefaciens* MT45, Sci. Rep. 7 (2017) 40976.