



ELSEVIER

Contents lists available at ScienceDirect

## Data in brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)

## Data Article

# 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



## ARTICLE INFO

## Article history:

Received 6 January 2020

Received in revised form 11 February 2020

Accepted 14 February 2020

Available online 22 February 2020

## Keywords:

Complete genome

*Bacillus amyloliquefaciens*

Antipathogenic activity

Rice wine

## 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 subtilisin 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](mailto:y82345@gmail.com) (M.-D. Kim).

Specification Table

|                                |  |
|--------------------------------|--|
| Subject                        | Genetics   |
| Specific subject area          | Genomics and Molecular Biology   |
| Type of data                   | Table  |
| How data were acquired         | The complete genome sequence was determined using the Pac-Bio RSII(P6–C4).   |
| Data format                    | Raw and Analyzed   |
| Parameters for data collection | The reads of <i>B. amyloliquefaciens</i> genome by Pac-Bio were assembled <i>de novo</i> into two contigs using HGAP (Hierarchical Genome Assembly Process).                                   |
| Description of data collection | Whole-genome sequencing followed by genome assembly  |
| Data source location           | <i>B. amyloliquefaciens</i> KCTC18343P was isolated from Korean rice wine produced in Kangwon-do, Republic of Korea.   |
| Data accessibility             | Data are available within this article, and the genome sequence of <i>B. amyloliquefaciens</i> KCTC18343P is available in GenBank under the accession number NZ_CP013727.                      |
| Related research article       | Sim, H. S. and Kim, M. D., Antipathogenic activity of <i>Bacillus amyloliquefaciens</i> isolated from Korean traditional rice wine. <i>Microbiol. Biotechnol. Lett.</i> (2016), 44(1), 98–105. |

#### Value of the Data

- *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, surfactin, 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 has lower starch-degrading activity and lower resistance to high temperature than other *B. amyloliquefaciens* strains [11]. We compared the  $\alpha$ -amylase amino acid sequences of *B. amyloliquefaciens* KCTC18343P and *B. amyloliquefaciens* KCTC13012 and found that the length and position of the active sites are notably different [4,6].

**Table 1**  
Genomic features of six different *Bacillus amyloliquefaciens* strains.

|            | Length (bp) | G + C content (%) | Predicted ORFs | rRNA genes | tRNA genes | Reference |
|------------|-------------|-------------------|----------------|------------|------------|-----------|
| KCTC18343P | 3,979,925   | 46.5              | 3725           | 27         | 86         | –         |
| DSM7       | 3,980,199   | 46.1              | 3870           | 30         | 94         | [3]       |
| CMW1       | 3,897,775   | 45.9              | 3762           | 64         | 5          | [9]       |
| MT45       | 3,897,521   | 46.1              | 3752           | 24         | 81         | [13]      |
| JRS8       | 4,090,896   | 46.3              | 4006           | 12         | 1          | [1]       |
| Y14        | 4,107,280   | 46.1              | 3623           | 33         | 100        | [8]       |

**Table 2**  
Average nucleotide identity values among six different *Bacillus amyloliquefaciens* strains.

|                | 1   | 2     | 3     | 4     | 5     | 6     |
|----------------|-----|-------|-------|-------|-------|-------|
| 1. KCTC 18343P | 100 | 94.13 | 93.90 | 95.04 | 99.27 | 99.41 |
| 2. DSM7        | –   | 100   | 98.37 | 98.78 | 94.49 | 95.13 |
| 3. CMW1        | –   | –     | 100   | 61.18 | 56.00 | 62.14 |
| 4. MT45        | –   | –     | –     | 100   | 57.70 | 95.12 |
| 5. JRS8        | –   | –     | –     | –     | 100   | 56.20 |
| 6. Y14         | –   | –     | –     | –     | –     | 100   |

*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<sup>TM</sup>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 (P6–C4) 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](http://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, *Bioresour. Technol.* 101 (2010) 8822–8827.
- [4] X.H. Chen, A. Koumoutsis, R. Scholz, R. Borriss, 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* 6B, *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.