



Whole-Genome Sequences of *Bacillus* and *Paenibacillus* sp. Strains Isolated from Honey in Japan

 Kayo Okumura,^{a,*} Mariko Okamoto,^b  Daisuke Takamatsu^{b,c}

^aDepartment of Veterinary Medicine, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido, Japan

^bDivision of Infectious Animal Disease Research, National Institute of Animal Health, National Agriculture and Food Research Organization, Tsukuba, Ibaraki, Japan

^cThe United Graduate School of Veterinary Sciences, Gifu University, Gifu, Gifu, Japan

ABSTRACT Knowledge about bacterial species in bee environments is essential for maintaining healthy honeybee colonies. Therefore, we performed whole-genome sequence analysis on bacteria isolated from honey harvested in Japan. This study reports the genomic sequences of the five bacterial strains identified.

Although honeybees are essential livestock for pollination and providing honey, research on infectious diseases affecting honeybees is not comprehensive. Previously, we isolated bacteria from honey in Japan (1) and performed whole-genome shotgun sequencing to investigate honeybee pathogens. However, obtaining complete genomic sequences is key to comprehensively understanding the toxicity and genetic profiles of the bacteria. In this study, we performed complete genomic analysis of five representative strains. Here, we report the complete genomic data for *Bacillus paralicheniformis* and *Paenibacillus* spp. found in honey.

Bacteria isolated from honey and stored as glycerol stocks were subsequently recovered and their genomic DNA purified (1). Bacteria were cultured on Columbia agar with 5% defibrinated sheep blood, harvested, and suspended in 550 μ L Tris-EDTA buffer. The cells were treated with approximately 10 mg/mL lysozyme and 50 U/mL mutanolysin (Sigma-Aldrich, USA) at 37°C and lysed with 1% sodium dodecyl-sulfate. The lysates were mixed with once with phenol, thrice with phenol-chloroform-isoamyl alcohol (PCI), and once with chloroform to extract DNA and remove the proteins. After ethanol precipitation, 100 μ g/mL RNase (Nippon Gene, Japan) treatment, PCI and chloroform extraction, ethanol precipitation, and a 70% ethanol rinse, the purified DNA was suspended in 10 mM Tris-HCl (pH 8.5) and stored at –20°C.

Whole-genome sequencing was performed using the PacBio RS II platform combined with a SMRT cell 8Pac ver. 3 and a DNA polymerase binding kit P6 (Pacific Biosciences, USA). For Illumina sequencing, the NovaSeq 6000 (Illumina, USA) system was used with 150-bp paired-end read settings. Samples were prepared using the TruSeq DNA PCR free kit (Illumina, USA) per the manufacturer's instructions. Read quality control, raw read filtering, and *de novo* sequence assembly were performed using HGAP ver. 3 software (Pacific Biosciences) (2) and polished using Quiver (2) with default settings. The assembly accuracy was validated by mapping Illumina reads obtained in a previous study (1) using Pilon ver. 1.21 (3) with default settings, with 100.00% coverage (except for *Paenibacillus azoreducens* J34TS1; 99.97%). After assembly, a self-dot plot was created using UGENE ver. 35.0 (<http://ugene.net>) to verify the plasmid and chromosome circularization by checking the contig end overlap and the 5' and 3' end connections. Analysis was performed with BUSCO ver. 3.0 software (4), using the bacteria_odb9 database to validate the completeness of genome assembly; for all genomes, quality scores of >98.7% were obtained. The general features of each genome are listed in Table 1.

Few genomes of *P. azoreducens* and *Paenibacillus dendritiformis* have been registered. Although NCBI contains genomic information on approximately 100 *B. paralicheniformis*

Editor Steven R. Gill, University of Rochester School of Medicine and Dentistry

Copyright © 2022 Okumura et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Daisuke Takamatsu, p1013dt@affrc.go.jp.

*Present address: Kayo Okumura, Management Department of Biosafety, Laboratory Animal, and Pathogen Bank, National Institute of Infectious Diseases, Higashimurayama, Tokyo, Japan.

The authors declare no conflict of interest.

Received 21 August 2022

Accepted 15 October 2022

Published 27 October 2022

TABLE 1 General features of the bacterial genomes presented in this study^a

Species	Isolate	Total no. of subreads	N_{50} (bp)	Total no. of subread bases	Size (bp)	GC content (%)	Coverage (x)	BioSample accession no.	SRA accession no.
<i>Bacillus paralicheniformis</i>	J25TS1	78,163 (P), 37,219,662 (I)	16,453	708,855,100 (P), 5,620,168,962 (I)	4,394,709	45.9	129 (P), 1,152 (I)	SAMD000425495	DRX320385 (P), DRX321129 (I)
	J36TS2	74,813 (P), 22,269,482 (I)	15,585	651,721,606 (P), 3,362,691,782 (I)	4,404,119	45.9	116 (P), 678 (I)	SAMD000425496	DRX320386 (P), DRX321130 (I)
	J36TS2 plasmid ^b				200,793	36.5	159 (P)		
	J41TS8	90,978 (P), 22,392,292 (I)	15,975	914,167,224 (P), 3,381,236,092 (I)	4,501,066	45.8	171 (P), 699 (I)	SAMD000425497	DRX320387 (P), DRX321131 (I)
<i>Paenibacillus azoreducens</i>	J34TS1	96,372 (P), 36,125,506 (I)	13,745	840,311,025 (P), 5,454,951,406 (I)	7,316,152	48.0	96 (P), 676 (I)	SAMD000425498	DRX320388 (P), DRX321132 (I)
<i>Paenibacillus dendritiformis</i>	J27TS7	130,145 (P), 32,933,880 (I)	15,307	1,359,398,548 (P), 4,973,015,880 (I)	6,547,390	54.7	124 (P), 690 (I)	SAMD000425499	DRX320389 (P), DRX321133 (I)

^a The total number of subreads and subread bases and the coverage depth values were obtained by PacBio (P) and Illumina (I) sequencing. The N_{50} size, and GC content values were obtained by PacBio sequencing.
^b *B. paralicheniformis* J36TS2 has a single plasmid, pBP1.

strains, complete genome sequences are absent for honey-derived strains. The data obtained in this study are valuable for understanding bacteria in bee environments to promote future research.

Data availability. The whole-genome sequences have been deposited at DDBJ/GenBank under BioProject accession number [PRJDB12645](#) and BioSample accession numbers [SAMD00425495](#) to [SAMD00425499](#) (Table 1). The raw sequence reads have been deposited in the DDBJ DRA/NCBI SRA, and the accession numbers are listed in Table 1.

ACKNOWLEDGMENTS

This study was supported by a Grant-in-Aid for Scientific Research (A) (JP21H04757) from the Japan Society for Promotion of Science. We declare that we have no conflicts of interest.

All authors meet the MRA authorship criteria. K.O., M.O., and D.T. were responsible for the organization and coordination of the analysis. D.T. was the chief investigator and was responsible for the data analysis. All authors contributed to the writing of the final manuscript.

REFERENCES

1. Okamoto M, Kumagai M, Kanamori H, Takamatsu D. 2021. Antimicrobial resistance genes in bacteria isolated from Japanese honey, and their potential for conferring macrolide and lincosamide resistance in the American foul-brood pathogen *Paenibacillus larvae*. *Front Microbiol* 12:667096. <https://doi.org/10.3389/fmicb.2021.667096>.
2. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
3. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
4. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol* 35:543–548. <https://doi.org/10.1093/molbev/msx319>.