



# Draft Genome Sequences of Five *Paenibacillus* Species of Dairy Origin

Atinuke M. Olajide,<sup>a</sup> Shu Chen,<sup>b</sup> Gisèle LaPointe<sup>a</sup>

<sup>a</sup>CRIFS, Department of Food Science, University of Guelph, Guelph, Ontario, Canada

<sup>b</sup>Agriculture and Food Laboratory, Laboratory Services Division, University of Guelph, Guelph, Ontario, Canada

**ABSTRACT** *Paenibacillus* species are important spoilage organisms in the dairy industry. The genomes of five *Paenibacillus* species which were isolated from dairy products in Canada were sequenced using the Illumina MiSeq platform. Draft genomes ranging from 5.1 to 7.1 Mb with GC contents of 49 to 53% were generated.

*Paenibacillus* is a genus of Gram-positive psychrotolerant spore-forming bacteria associated with fluid milk spoilage by causing off flavors and curdling (1). They, together with *Bacillus* species, occur regularly in dairy farm environments, processing facilities, and pasteurized milk (2). Their spore-forming nature confers resistance to high-temperature treatments, so they survive pasteurization (3). Recently, some *Paenibacillus* species were isolated from samples obtained from a cheese processing plant in Ontario, Canada. Five isolates identified as *Paenibacillus* spp. with matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) (Bruker MBT Compass 1829023 library v9) according to the manufacturer’s manual or with sequencing of the 16S rRNA gene (4) were selected for whole-genome sequencing. *Paenibacillus campinasensis* 3CS1, *P. macerans* 3CT49, *P. validus* 2CS3, and *P. woosongensis* 12CR55 were isolated from cheese curds, and *P. timonensis* 12ME58 was isolated from raw milk. While 3CS1, 3CT49, and 12CS3 were obtained after a spore pasteurization treatment (80°C/12 min), 12ME58 and 12CR55 were obtained after a laboratory pasteurization treatment (63°C/30 min). All isolates but 3CT49 were obtained from aerobic incubation on brain heart infusion agar (30°C/2 days). Isolate 3CT49 grew under anaerobiosis on reinforced clostridium agar (37°C/3 days) after enrichment in reinforced *Clostridium* medium for 7 days at 37°C.

After genomic DNA from one colony per species (picked from fresh plates) was obtained using the Ultraclean microbial DNA isolation kit (Qiagen, Mississauga, ON, Canada), the concentration and quality were determined via Qubit and Nanodrop, respectively. Sequencing libraries were constructed from the DNA samples using the Nextera DNA Flex library preparation kit (Illumina Canada, Vancouver, BC) following the manufacturer’s protocol. Sequencing was performed on the MiSeq platform using a MiSeq V2 reagent kit (Illumina) and paired-end cycles according to the manufacturer’s instructions. Raw sequence reads were filtered and trimmed using the MiSeq sequencer system software v3.1 and FastQC v1.0.0. A total of 17,657,241 reads were obtained, and the average length after trimming was 250 bp.

Sequences with a quality score of  $\geq 30$  were assembled into contigs and scaffolds via *de novo* assembly following an overlap-layout-consensus method using the SPAdes v3.9.0 genome assembler (5). Genome assembly quality was evaluated using the QUAST tool (from the SPAdes pipeline) to generate summary statistics (e.g.,  $N_{50}$ , largest contig length, GC content, and number of Ns per 100 kbp) (Table 1). Gene prediction, naming, and annotation were done using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.10 (6). Default parameters were used for all software unless other-

**Citation** Olajide AM, Chen S, LaPointe G. 2020. Draft genome sequences of five *Paenibacillus* species of dairy origin. Microbiol Resour Announc 9:e00971-20. <https://doi.org/10.1128/MRA.00971-20>.

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

**Copyright** © 2020 Olajide 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 Atinuke M. Olajide, [atinuke@uoguelph.ca](mailto:atinuke@uoguelph.ca), or Gisèle LaPointe, [glapoint@uoguelph.ca](mailto:glapoint@uoguelph.ca).

**Received** 18 August 2020

**Accepted** 21 August 2020

**Published** 10 September 2020

**TABLE 1** Characteristics of draft genome sequences and accession numbers of the five *Paenibacillus* species

Species and strain	Total no. of reads	Size of genome (bp)	No. of contigs	Length of largest contig (bp)	$N_{50}^a$ (bp)	GC content (%)	Coverage (x)	No. of genes	No. of rRNAs <sup>b</sup>	% identity with the 16S rRNA gene (comparison length [bp]) (NCBI accession no.)	GenBank accession no.	Raw read SRA accession no.
<i>P. campinasensis</i> 3CS1	1,839,718	5,362,470	112	765,696	194,562	52.3	86	4,848	116	100.0 (1,577) (AB073187.1)	WOAA000000000	SRX7225308
<i>P. macerans</i> 3CT49	2,993,442	7,085,787	108	644,494	193,195	52.8	106	6,340	120	99.0 (1,358) (AB073196.1)	WNZZ000000000	SRX7225307
<i>P. timonensis</i> 12ME58	3,093,697	5,126,083	88	484,557	211,929	52.6	151	4,735	125	98.0 (1,357) (MK734103.1)	WNZY000000000	SRX7225306
<i>P. validus</i> 2CS3	2,573,113	5,502,307	111	573,538	193,871	52.2	117	5,341	156	99.0 (1,520) (AB073203.1)	WNZX000000000	SRX7225305
<i>P. woosongensis</i> 12CR55	3,558,563	5,683,957	78	1,379,340	475,550	49.6	156	5,122	119	100.0 (1,490) (AY847463.1)	WNZW000000000	SRX7225304

<sup>a</sup>The number of N's per 100 kbp is 0.00 for all genomes, where N is a position that could be occupied by any of the four DNA bases.

<sup>b</sup>No. of rRNAs, sum of tRNAs, rRNAs, and other RNAs.

wise noted. Species identities were confirmed by comparing the available 16S rRNA gene sequences in the NCBI database (<https://www.ncbi.nlm.nih.gov>) using the BLAST service in PATRIC v3.6.6 (<https://www.patricbrc.org/>) with genome sequences obtained (Table 1) due to the lack of available closed genome sequences for these species. These assemblies will help expand the study of *Paenibacillus* species, further explore their genomic diversity, and develop new tools to control spoilage of dairy products caused by these psychrotolerant spore-forming organisms.

**Data availability.** The pure cultures were deposited in the Canadian Research Institute for Food Safety (CRIFS) collection. The genome sequences reported here have been deposited at DDBJ/EMBL/GenBank under the accession numbers listed in Table 1.

## ACKNOWLEDGMENTS

Funding for this research was provided by Genome Canada (Genomic Application Partnership Program [GAPP] round 7) and the NSERC/Dairy Farmers of Ontario Industrial Research Chair in Dairy Microbiology held by G. LaPointe.

We are grateful for the help of Myra Siddiqi in the species identity confirmation.

We declare that there are no potential conflicts of interest to disclose.

## REFERENCES

1. Ivy RA, Ranieri ML, Martin NH, den Bakker HC, Xavier BM, Wiedmann M, Boor KJ. 2012. Identification and characterization of psychrotolerant sporeformers associated with fluid milk production and processing. *Appl Environ Microbiol* 78:1853–1864. <https://doi.org/10.1128/AEM.06536-11>.
2. Huck JR, Woodcock NH, Ralyea RD, Boor KJ. 2007. Molecular subtyping and characterization of psychrotolerant endospore-forming bacteria in two New York State fluid milk processing systems. *J Food Prot* 70: 2354–2364. <https://doi.org/10.4315/0362-028x-70.10.2354>.
3. Ranieri ML, Ivy RA, Mitchell WR, Call E, Masiello SN, Wiedmann M, Boor KJ. 2012. Real-time PCR detection of *Paenibacillus* spp. in raw milk to predict shelf life performance of pasteurized fluid milk products. *Appl Environ Microbiol* 78:5855–5863. <https://doi.org/10.1128/AEM.01361-12>.
4. Weisburg W, Barns S, Pelletier D, Lane D. 1991. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 173:697–703. <https://doi.org/10.1128/jb.173.2.697-703.1991>.
5. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prijibelski 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>.
6. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44: 6614–6624. <https://doi.org/10.1093/nar/gkw569>.