Draft Genome Sequence of *Bacillus* sp. FMQ74, a Dairy-Contaminating Isolate from Raw Milk

Mira Okshevsky,a Viduthalai R. Regina,a Ian P. G. Marshall,b Lars Schreiber,b Rikke L. Meyer,c
Interdisciplinary Nanoscience Center (iNANO), Aarhus University, Aarhus, Denmarka; Center for
Geomicrobiology, Section for Microbiology, Department of Bioscience, Aarhus University, Aarhus, Denmarkb; Microbiology Section, Department of Bioscience, Aarhus University, Aarhus, Denmarkc

ABSTRACT Representatives of the genus *Bacillus* are common milk contaminants that cause spoilage and flavor alterations of dairy products. *Bacillus* sp. FMQ74 was isolated from raw milk on a Danish dairy farm. To elucidate the genomic basis of this strain’s survival in the dairy industry, a high-quality draft genome was produced.

*Bacillus* spp. are Gram-positive bacteria commonly associated with spoilage and flavor alterations of finished dairy products. They are found in raw milk and subsequent stages of milk processing (1, 2) and can cause food poisoning in humans (3, 4). The ability of *Bacillus* spp. to form communities called biofilms greatly contributes to their success as dairy contaminants. Biofilms are difficult to eradicate from food-processing surfaces due to their recalcitrance to antimicrobials and common cleaning procedures (5, 6). A survey of milk contaminants on 49 dairy farms in Denmark revealed that the FMQ74 operational taxonomic unit (OTU) was present in raw milk on 29 farms. *Bacillus* sp. FMQ74 was therefore selected for sequencing due to its ubiquitous presence as a milk contaminant.

*Bacillus* strain FMQ74 was isolated from raw milk. The milk was heated to 63°C for 30 min, diluted 1:10 with tryptic yeast extract broth, and incubated for 24 h at 30°C. The enriched milk was serially diluted, inoculated onto tryptic soy broth agar plates, and incubated at 30°C. The single colony was restreaked six times prior to genomic DNA extraction using the GeneJet genomic DNA extraction kit (Thermo Scientific). A sequencing library was prepared using the Nextera XT kit (Illumina, USA). Genome sequencing was performed using the Illumina MiSeq platform with a paired-end 300-bp MiSeq reagent kit version 3, resulting in ca. 4.4 million sequencing reads representing 1.2 Gbp and an approximately 250× coverage. Reads were quality inspected using FastQC version 0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and were quality and adapter trimmed using Trimmomatic version 0.36 (7). Quality-based trimming used a 4-bp sliding window and minimum average quality score ≥20. Reads >200 bp were assembled using SPAdes 3.9.0 (8) with the parameters: –careful –k 21, 33, 55, 77, 99, 127, resulting in 65 contigs (4,167,316 bp). Only contigs with G+C content between 20% and 65% and coverage between 10× and 3,500× (determined using BBmap version 35.82 [https://sourceforge.net/projects/bbmap/] ) were retained, resulting in 52 contigs representing 99.9% of the original assembly. Contigs identified as contamination by 16S comparison with the SILVAngs database release 128 (9) were removed, leaving 47 contigs in the final draft genome assembly. This assembly was manually augmented with 849 bp of the 16S gene PCR amplified from genomic DNA and Sanger sequenced.

The draft genome sequence *Bacillus* sp. FMQ74 has a total length of 4,159,532 bp, an average G+C content of 43.3%, and an N50 length of 345,666 bp. The genome was
estimated to be 99.4% complete compared to the single-gene marker set for the genus *Bacillus* via CheckM 1.0.7 (10). Prokka 1.12-beta (11) identified 4,204 protein-coding sequences, 13 rRNA sequences, and 84 tRNA sequences. This genome contains the complete epsA-O operon encoding the major polysaccharide of the *Bacillus subtilis* biofilm extracellular matrix (12), the tapA-sipW-tapA operon encoding amyloid protein (13–15), biofilm regulatory proteins *sinL* and *sinR* (16), as well as the poly-γ-DL-glutamic acid biosynthesis genes required for submerged biofilm formation in *B. subtilis* (17).

**Accession number(s).** This draft genome sequence has been deposited at DDBJ/ENA/GenBank under the accession number MOEO00000000. The version described here is version MOEO01000000.

**ACKNOWLEDGMENTS**

We are grateful to Britta Poulsen for Illumina MiSeq sequencing.

This study was funded by Futuremilq Innovation Consortium (Danish Ministry for Higher Education and Science grant no. 11-118502), Aarhus University Graduate School of Science and Technology, Danish National Research Foundation (grant no. DNRF104), ERC Advanced Grant MICROENERGY (grant no. 294200), and Marie Curie IIF fellowship “ATP_adapt_low_energy” (both European Union seventh Framework Programs).

**REFERENCES**

1. Gopal N, Hill C, Ross PR, Beresford TP, Fenelon MA, Cotter PD. 2015. The prevalence and control of *Bacillus* and related spore-forming bacteria in the dairy industry. Front Microbiol 6:1418. https://doi.org/10.3389/fmicb.2015.01418.

2. Kalogridou-Vassiliadou D. 1992. Biochemical activities of *Bacillus* species isolated from flat sour evaporated milk. J Dairy Sci 75:2681–2686. https://doi.org/10.3168/jds.S0022-0302(92)78030-8.

3. Logan NA. 2012. *Bacillus* and relatives in foodborne illness. J Appl Microbiol 112:417–429. https://doi.org/10.1111/j.1365-2672.2011.05204.x.

4. Pavic S, Brett M, Petric N, Lastre D, Smoljanovic M, Atkinson M, Kovacic A, Cetinic E, Ropac D. 2005. An outbreak of food poisoning in a kindergarten caused by milk powder containing toxigenic *Bacillus subtilis* and *Bacillus licheniformis*. Arch Lebensmittelhygiene 56:20–22.

5. Brooks JD, Flint SH. 2008. Biofilms in the food industry: problems and potential solutions. Int J Food Sci Technol 43:2163–2176. https://doi.org/10.1111/j.1365-2651.2008.01839.x.

6. Kumar CG, Anand SK. 1998. Significance of microbial biofilms in food industry: a review. Int J Food Microbiol 42:9–27.

7. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

8. 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.

9. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596. https://doi.org/10.1093/nar/gks1219.

10. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.

11. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

12. Branda SS, González-Pastor JE, Dervyn E, Ehrlich SD, Losick R, Kolter R. 2004. Genes involved in formation of structured multicellular communities by *Bacillus subtilis*. J Bacteriol 186:3970–3979. https://doi.org/10.1128/JB.186.12.3970-3979.2004.

13. Terra R, Stanley-Wall NR, Cao G, Lazazzera BA. 2012. Identification of *Bacillus subtilis* SipW as a bifunctional signal peptidase that controls surface-adhered biofilm formation. J Bacteriol 194:2781–2790. https://doi.org/10.1128/JB.06780-11.

14. Stöver AG, Driks A. 1999. Regulation of synthesis of the *Bacillus subtilis* transition-phase, spore-associated antibacterial protein TasA. J Bacteriol 181:5476–5481.

15. Stöver AG, Driks A. 1999. Control of synthesis and secretion of the *Bacillus subtilis* protein YqXm. J Bacteriol 181:7065–7069.

16. Chu F, Kearns DB, McLoon A, Chai Y, Kolter R, Losick R. 2008. A novel regulatory protein governing biofilm formation in *Bacillus subtilis*. Mol Microbiol 68:1117–1127. https://doi.org/10.1111/j.1365-2958.2008.06201.x.

17. Stanley NR, Lazazzera BA. 2005. Defining the genetic differences between wild and domestic strains of *Bacillus subtilis* that affect poly-γ-DL-glutamic acid production and biofilm formation. Mol Microbiol 57:1143–1158. https://doi.org/10.1011/j.1365-2958.2005.04746.x.