Complete Genome Sequence of *Lysinibacillus* sp. Strain SGAir0095, Isolated from Tropical Air Samples Collected in Singapore

Anthony Wong,[a] Ana Carolina M. Junqueira,[b] Akira Uchida,[a] Rikky W. Purbojati,[a] James N. I. Houghton,[a] Caroline Chénard,[a] Megan E. Clare,[a] Kavita K. Kushwaha,[a] Alexander Putra,[a] Nicolas E. Gaultier,[a] Balakrishnan N. V. Premkrishnan,[a] Vineeth Kodengil Vettath,[a] Cassie E. Heinle,[a] Daniela I. Drautz-Moses,[a] Stephan C. Schuster[a]

[a]Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore
[b]Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

ABSTRACT *Lysinibacillus* sp. strain SGAir0095 was isolated from tropical air samples collected in Singapore, and its complete genome was sequenced with a hybrid strategy using single-molecule real-time sequencing and short reads. The genome consists of one chromosome of 4.14 Mbp and encompasses 3,885 protein-coding genes, 39 rRNAs, and 101 tRNAs.

*L*ysinibacillus spp. are rod-shaped, Gram-positive bacteria belonging to the family *Bacillaceae*. *Lysinibacillus* was first proposed as a genus distinct from *Bacillus* by Ahmed et al. (1) in 2007, as its peptidoglycan was found to contain lysine, aspartic acid, alanine, and glutamic acid. Since then, the reclassification of several *Bacillus* species to *Lysinibacillus* has been proposed (2–4). *Lysinibacillus* spp. have been found to be associated with plants (5–7) and also have been isolated from air in Makkah, Saudi Arabia (8), which contrasts with its natural soil habitat (1). Here, we report the complete genome sequence of another *Lysinibacillus* sp. collected from the air, which may aid future metagenomics studies and provide insight into its presence in the air microbiome.

Lysinibacillus sp. strain SGAir0095 was isolated from air collected in Singapore (global position system coordinates 1.347654N, 103.685240E) using an Andersen type air sampler (IUL S.A., Spain). The air was impacted onto marine agar (Becton Dickinson, USA), and subsequent isolation of the colonies was done on Trypticase soy agar (TSA; Becton Dickinson) at 30°C. Prior to DNA extraction, a single colony was cultured in lysogeny broth (Merck, USA) at 30°C overnight. Genomic DNA was purified using the Wizard genomic DNA purification kit (Promega, USA), following the manufacturer's recommended protocol. Library preparation was performed with the SMRTbell template prep kit 1.0 (Pacific Biosciences, USA), followed by single-molecule real-time (SMRT) sequencing on the Pacific Biosciences RS II platform. The assembly was further supplemented with a 300-bp paired-end run to generate short reads on a MiSeq platform (Illumina, USA). The whole-genome shotgun libraries of short fragments were constructed using the TruSeq Nano DNA library preparation kit (Illumina), following the manufacturer's protocol.

Subsequent analysis was performed using default software parameters, unless otherwise specified. Quality control was done with the PreAssembler filter version 1 from the Hierarchical Genome Assembly Process (HGAP) version 3 (9) and with Cutadapt version 1.8.1 (10) for PacBio and MiSeq reads, respectively. A total of 109,403 subreads from SMRT sequencing with an N_{50} length of 10,466 bp were used for *de novo* assembly using HGAP version 3 (9) implemented in the PacBio SMRT

Analysis package version 2.3.0. The assembly was polished using Quiver (9) and error corrected using Pilon version 1.16 (11) and 716,265 MiSeq paired-end reads. The final assembly consisted of a circularized 4,142,194-bp chromosomal contig, with a coverage of 191.716-fold and a G+C content of 36.68%. Species evaluation was also performed with Microbial Species Identifier based on the genome-wide average nucleotide identity (ANI) method (12) using ANICalculator, with default parameters, against a database of 6,387 bacterial RefSeq genomes created using text filter for: "type, synonym type, proxytype" and subsequent "getorf -find 3" option. The closest identified species was *Lysinibacillus sinduriensis* BLB-1 (82.74%), with an alignment fraction of 38%. However, via 16S rRNA sequencing, the closest hit was *Bacillus* sp. strain B13 (99.47%).

The genome was annotated using the NCBI's Prokaryotic Genome Annotation Pipeline (PGAP) version 4.4 (13). A total of 4,083 genes were predicted, consisting of 3,885 protein-coding genes, 13 copies each of 5S, 16S, and 23S rRNAs, 101 tRNAs, 6 noncoding RNAs, and 52 pseudogenes. Using Rapid Annotations using Subsystems Technology (RAST) (14–16) with the ClassicRAST annotation scheme with the "fix frameshift" option set to "yes," functional annotation revealed that 402 and 325 genes were associated with amino acids and their derivatives and carbohydrate metabolism, respectively. These genes may play a role in the altered peptidoglycan that *Lysinibacillus* spp. possess. Interestingly, stress response- and dormancy-related genes (123 and 78 genes, respectively) were also predicted which may allow *Lysinibacillus* sp. strain SGAir0095 to remain dormant during transport in air.

Data availability. The complete genome sequence of *Lysinibacillus* sp. SGAir0095 has been deposited in DDBJ/EMBL/GenBank and the Sequence Read Archive (SRA) under the accession number CP028083 and numbers SRR8894403 and SRR8894404, respectively.

ACKNOWLEDGMENTS

REFERENCES

1. Ahmed I, Yokota A, Yamazoe A, Fujiwara T. 2007. Proposal of Lysinibacillus boronitolerans gen. nov. sp. nov., and transfer of Bacillus fusiformis to Lysinibacillus fusiformis comb. nov. and Bacillus sphaericus to Lysinibacillus sphaericus comb. nov. Int J Syst Evol Microbiol 57:1117–1125. https://doi.org/10.1099/ijs.0.63867-0.
2. Jung MY, Kim JS, Paek WK, Styrak I, Park IS, Sin Y, Paek J, Park KA, Kim H, Kim HL, Chang YH. 2012. Description of Lysinibacillus sinduriensis sp. nov., and transfer of Bacillus massiliensis and Bacillus odysseyi to the genus Lysinibacillus as Lysinibacillus massiliensis comb. nov. and Lysinibacillus odysseyi comb. nov. with emended description of the genus Lysinibacillus. Int J Syst Evol Microbiol 62:2347–2355. https://doi.org/10.1099/ijs.0.033837-0.
3. Rekadwad BN, Gonzalez JM. 2017. Correcting names of bacteria deposited in National Microbial Repositories: an analysed sequence data necessary for taxonomic re-categorization of misclassified bacteria-ONE example, genus Lysinibacillus. Data Brief 13:761–778. https://doi.org/10.1016/j.dib.2017.06.042.
4. Xu K, Yuan Z, Rayner S, Hu X. 2015. Genome comparison provides molecular insights into the phylogeny of the reassigned new genus Lysinibacillus. BMC Genomics 16:140. https://doi.org/10.1186/s12864-015-1359-x.
5. Burkett-Cadena M, Sastoque L, Cadena J, Dunlap CA. 2019. Lysinibacillus capsici sp. nov, isolated from the rhizosphere of a pepper plant. Antonie Van Leeuwenhoek 112:1161. https://doi.org/10.1007/s10482-019-01248-w.
6. Pudova DS, Lutfullin MT, Shagimardanova EI, Hadieva GF, Shigapova L, Toymentseva AA, Kabanov DA, Mardanova AM, Vologin SG, Sharipova MR. 2018. Draft genome sequence data of Lysinibacillus fusiformis strain GM, isolated from potato phyllosphere as a potential probiotic. Data Brief 21:2504–2509. https://doi.org/10.1016/j.dib.2018.11.107.
7. Yu J, Guan X, Liu C, Xiang W, Yu Z, Liu X, Wang G. 2016. Lysinibacillus endophyticus sp. nov., an indole-3-acetic acid producing endophytic bacterium isolated from corn root (Zea mays cv. Xinken-5). Antonie Van Leeuwenhoek 109:1337–1344. https://doi.org/10.1007/s10482-016-0732-3.
8. Papadioti A, Azhar EI, Bibi F, Jiman-Fatani A, Aboushoushah SM, Yasir M, Raoult D, Angelakis E. 2017. 'Lysinibacillus saudimassiliensis' sp. nov., a new bacterial species isolated from air samples in the urban environment of Makkah, Saudi Arabia. New Microbes New Infect 16:25–27. https://doi.org/10.1016/j.nmni.2016.12.011.
9. Chin C-S, 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. https://doi.org/10.1038/nmeth.2474.
10. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10. https://doi.org/10.14806/ej.17.1.200.
11. 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.
12. Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, Kyrpides NC, Pati A. 2015. Microbial species delineation using whole genome sequences. Nucleic Acids Res 43:6761–6771. https://doi.org/10.1093/nar/gkv657.
13. 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.
14. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K,

Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.

15. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.

16. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206–D214. https://doi.org/10.1093/nar/gkt1226.