Complete Genome Sequence of *Pediococcus pentosaceus* Strain GDIAS 001

Maopeng Song,[a] Yongfei Wang,[a] Xianyong Ma,[b] Dun Deng[b]

[a]Department of Bioengineering, School of Life Sciences, Jinan University, Guangzhou, China
[b]State Key Laboratory of Animal Breeding, South China Key Laboratory of Animal Nutrition and Feed, Ministry of Agriculture, Institute of Animal Science, Guangdong Academy of Agricultural Sciences, Animal Breeding and Nutrition Public Laboratory, Poultry Meat Quality, Quality Control and Evaluation Engineering Technology Center, Guangzhou, China

ABSTRACT *Pediococcus pentosaceus* strain GDIAS 001 was isolated from a tapioca sample in Guangzou, China. The genome of GDIAS 001 was assembled using single-molecule real-time (SMRT) sequencing, and it contains 1 chromosome of 1.83 Mbp and 1,835 protein-coding genes, 71 RNA genes, and 56 tRNA genes.

*P*ediococcus pentosaceus is a member of the lactic acid bacteria and facultative anaerobic Gram-positive bacteria (1). *P. pentosaceus* is usually found in food and feed materials. Those strains can improve food flavor (2) and taste (3), especially for meat and cheese (4). Some *P. pentosaceus* strains are also effective in resisting pathogenic bacteria (5, 6). Thus, *P. pentosaceus* has broad application prospects in the food, medicine, and feed industries.

Here, *Pediococcus pentosaceus* strain GDIAS 001 was isolated from tapioca. It was cultured on MRS agar plates (7) at 30°C. Six passages of GDIAS 001 were stored in 20% glycerol at −80°C. Then, the genomic DNA of GDIAS 001 was prepared and sequenced with the following steps. First, a single colony was cultured in MRS liquid culture. Second, 10-ml overnight cultures were used to prepare genomic DNA following the manual instructions for a bacterial DNA kit (Omega Bio-Tek, USA). The genomic DNA of GDIAS 001 was monitored using both 1% agarose gels and NanoDrop One (Thermo Fisher Scientific). Lastly, qualified DNA was used to construct a genomic DNA library with the SMRTbell template prep kit 1.0-SP v3 (Pacific Biosciences, USA), and then sequencing was conducted with PacBio RS II single-molecule real-time (SMRT) sequencing technology (Pacific Biosciences).

A total of 1.46 Gb of SMRT sequencing data (subreads) were obtained. These data were self-corrected, and preliminary assembly was done using Falcon v0.3+ with default settings (default parameters were used for all software unless otherwise specified) (8). The consensus sequence was corrected using Genomic Consensus software v2.3.3 and Spira v0.9.9.23. Lastly, the bacterial genome was circularized using Circlator v1.5.1 (9).

After filtering, 61,653 reads with an N_{50} value of 7,499 bp were assembled using the Hierarchical Genome Assembly Process (HGAP) pipeline of SMRT Analysis v2.3.0 (10) with genome coverage of 164×. The genome of *P. pentosaceus* GDIAS 001 consists of one circular chromosome of 1,831,351 bp with a 37.71% average DNA G+C content. Glimmer v3.0 (http://ccb.jhu.edu/software/glimmer/index.shtml) and Prodigal v2.6.3 (https://github.com/hyattpd/prodigal/releases/) were used to retrieve the related coding gene (11). tRNA genes were predicted using tRNAscan-SE v1.3.1 (12). rRNA genes were analyzed using RNAmmer v1.2 (13). Small nuclear RNAs (snRNAs) were predicted using BLAST against the Rfam database v10.0 (14). A total of 1,835 putative protein-coding genes and 71 RNA genes, including 15 rRNA (5 5S rRNAs, 5 16S rRNAs, and 5 23S rRNAs) and 56 tRNA genes, were predicted. The Virulence Factors Database (VFDB)

(http://www.mgc.ac.cn/VFs) (15) was used to detect virulence factors, and 266 putative virulence factors were found in the genome of GDIAS 001. Proteins were predicted using SignalP v4.1 (16), and 35 secretory proteins were presumed to contain signal peptides. The Antibiotic Resistance Genes Database v1.1 (http://ardb.cbcb.umd.edu/) (17) was used to evaluate the genomic sequences for antimicrobial resistance (AMR) genes, and three AMR genes (resistance to trimethoprim, elfamycin, and rifampin) were identified. The genomic sequence of GDAIS 001 was annotated by the NCBI PGAP (18) and deposited in the GenBank database.

Data availability. This whole-genome shotgun project, including raw reads of *Pediococcus pentosaceus* GDIAS 001, has been deposited at DDBJ/EMBL/GenBank under the accession number CP046938. The raw reads of *Pediococcus pentosaceus* GDIAS 001 have been deposited under SRA accession number SRR10716342.

REFERENCES

1. Stackebrandt E, Fowler VJ, Woese CR. 1983. A phylogenetic analysis of lactobacilli, *Pediococcus pentosaceus* and *Leuconostoc mesenteroides*. Syst Appl Microbiol 4:326–337. https://doi.org/10.1016/S0723-2020(83)80019-8.
2. Escamilla-Hurtado ML, Valdés-Martínez SE, Soriano-Santos J, Gómez-Pliego R, Verde-Calvo JR, Reyes-Dorantes A, Tomasini-Campocosio A. 2005. Effect of culture conditions on production of butter flavor compounds by *Pediococcus pentosaceus* and *Lactobacillus acidophilus* in semisolid maize-based cultures. Int J Food Microbiol 105:305–316. https://doi.org/10.1016/j.ijfoodmicro.2005.04.014.
3. Kleinschmit DH, Kung L, Jr. 2006. The effects of *Lactobacillus buchneri* 40788 and *Pediococcus pentosaceus* R1094 on the fermentation of corn silage. J Dairy Sci 89:3999–4004. https://doi.org/10.3168/jds.S0022-0302(06)72443-2.
4. Irmler S, Bavan T, Oberli A, Roetschi A, Badertscher R, Guggenbühl B, Berthoud H. 2013. Catabolism of serine by *Pediococcus acidilactici* and *Pediococcus pentosaceus*. Appl Environ Microbiol 79:1309–1315. https://doi.org/10.1128/AEM.03085-12.
5. Lee JY, Hwang KY, Kim HS, Kim K, Sung SI. 2002. Isolation and identification of lactic acid bacteria inhibiting gastro-intestinal pathogenic bacteria of domestic animal. Kor J Microbiol Biotechnol 30:129–134.
6. Muhialdin BJ, Hassan Z, Imdakim MMA, Abdul Kahar FKS, Aween MM. 2012. Malaysian isolates of lactic acid bacteria with antibacterial activity against Gram-positive and Gram-negative pathogenic bacteria. J Feed Res 1:110–116. https://doi.org/10.5539/jfr.v1n1p110.
7. Reuter G. 1985. Elective and selective media for lactic acid bacteria. Int J Food Microbiol 2:55–68. https://doi.org/10.1016/0168-1605(85)90057-1.
8. Chin C-S, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, Dunn C, O'Malley R, Figueroa-Balderas R, Morales-Cruz A, Cramer GR, Delledonne M, Luo C, Ecker JR, Cantu D, Rank DR, Schatz MC. 2016. Phased diploid genome assembly with single-molecule real-time sequencing. Nat Methods 13:1050–1054. https://doi.org/10.1038/nmeth.4035.
9. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. https://doi.org/10.1186/s13059-015-0849-0.
10. 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 10:563–569. https://doi.org/10.1038/nmeth.2474.
11. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.
12. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964. https://doi.org/10.1093/nar/25.5.955.
13. Karin L, Peter H, Andreas RE, Hans-Henrik S, Torbjørn R, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.
14. Akhter S, Aziz RK, Edwards RA. 2012. PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. Nucleic Acids Res 16:126. https://doi.org/10.1093/nar/gks406.
15. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q. 2005. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res 33:325–328. https://doi.org/10.1093/nar/gki008.
16. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786. https://doi.org/10.1038/nmeth.1701.
17. Bo L, Mihai P. 2009. ARDB: antibiotic resistance genes database. Nucleic Acids Res 37:443–447. https://doi.org/10.1093/nar/gkn656.
18. 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.