

Complete Genome Sequence of *Pediococcus pentosaceus* Strain SL4

Shruti Harnal Dantoft,^a Eliza Maria Bielak,^a Jae-Gu Seo,^b Myung-Jun Chung,^b Peter Ruhdal Jensen^a

Systems Biotechnology Group, Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark^a; R&D Center of Cell Biotech Co., Ltd., Gimpo, South Korea^b

***Pediococcus pentosaceus* SL4 was isolated from a Korean fermented vegetable product, kimchi. We report here the whole-genome sequence (WGS) of *P. pentosaceus* SL4. The genome consists of a 1.79-Mb circular chromosome (G+C content of 37.3%) and seven distinct plasmids ranging in size from 4 kb to 50 kb.**

Received 20 November 2013 Accepted 26 November 2013 Published 26 December 2013

Citation Dantoft SH, Bielak EM, Seo J-G, Chung M-J, Jensen PR. 2013. Complete genome sequence of *Pediococcus pentosaceus* strain SL4. *Genome Announc.* 1(6):e01106-13. doi:10.1128/genomeA.01106-13.

Copyright © 2013 Dantoft et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Peter Ruhdal Jensen, prj@bio.dtu.dk.

The genus *Pediococcus* is a member of the family *Lactobacillaceae* and is classified as a lactic acid bacterium (LAB), because the main metabolic end product it produces is lactic acid (1). Like other members of the LAB family, *Pediococcus pentosaceus* is a Gram-positive, nonmotile, facultative anaerobe and is acid tolerant. *P. pentosaceus*, along with *P. acidolactici*, is often used as a starter culture in natural and controlled fermentations (2), being a member of indigenous microflora. The strain SL4 was isolated from a traditional Korean fermented vegetable product, kimchi. Based on the 16S RNA sequence, it was taxonomically classified as *P. pentosaceus* and deposited in the Korean Collection for Type Cultures (KCTC) under the accession number KCTC 10297. The strain was shown to inhibit the growth of *Listeria monocytogenes* and *Staphylococcus aureus* via production of a bacteriocin.

The whole genome of *P. pentosaceus* SL4 was sequenced using the Illumina HiSeq 2000 platform (Macrogen, South Korea), yielding 65,886,608 paired-end reads with an average length of 101 bp, amounting to 6,654,547,408 bp and an average genome coverage of $\times 3,700$. A draft assembly was created by assembling the reads *de novo* into 148 contigs, using the CLC Genomics Workbench v. 6.5 (CLC Bio, Aarhus, Denmark). The average contig size was 13,074 kb, with the largest contig being 439,694 kb. The assembled contigs were annotated using the Rapid Annotations using Subsystems Technology (3).

The contigs were analyzed by blast searching at NCBI (4) against the sequence of the closest neighbor of *P. pentosaceus* SL4, *P. pentosaceus* ATCC 25745 (GenBank accession number CP000422). Contigs were further assembled into larger scaffolds, and PCR was used to close the gaps and to obtain a complete circular map of the SL4 chromosome. The assembled map of *P. pentosaceus* SL4 represents a 1,789,138-bp genome, with an average G + C content of 37.3%. This circular genome was annotated by the NCBI Prokaryotic Genome Annotation Pipeline v. 2.2 (released 2013) (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/) (5) upon submission of the genome sequence. A total of 1,817 genes were identified, accounting for 1,709 coding sequences (CDS), 41 pseudogenes, 1 CRISPR array, 15 rRNAs, 51 tRNAs, and 1 noncoding RNA (ncRNA).

The complete genome sequences of SL4 and the type strain

ATCC 25745 were aligned using NCBI blast, which showed that 92% of the sequence of SL4 aligned with that of the type strain with 99% sequence identity, and a further analysis of the differences between the strains is warranted. Interestingly, SL4 possesses a large number of plasmids, which poses a challenge in trying to engineer the strain. The plasmid sequences and their roles in the strain are currently under investigation in our lab. A strain of *P. pentosaceus*, IE-3 (accession numbers CAHU01000001 to CAHU01000091), isolated from effluent from the dairy industry (6) has recently been sequenced. The present study combined with the previous genome sequencing efforts could be used to improve our understanding of the effects of diverse ecological niches on genome diversity in this industrially important species.

Nucleotide sequence accession number. The complete genome sequence of *P. pentosaceus* SL4 has been deposited in GenBank under accession number [CP006854](https://www.ncbi.nlm.nih.gov/nuclink/CP006854).

ACKNOWLEDGMENTS

This work was supported by funding from Cell Biotech Co., Ltd., Gimpo, South Korea.

We acknowledge Rolf Sommer Kaas for conducting contamination checks on the sequencing data. We thank Regina Áris Schürmann for excellent technical work.

REFERENCES

1. Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, Shakhova V, Grigoriev I, Lou Y, Rohksar D, Lucas S, Huang K, Goodstein DM, Hawkins T, Plengvidhya V, Welker D, Hughes J, Goh Y, Benson A, Baldwin K, Lee JH, Díaz-Muñoz I, Dosti B, Smeianov V, Wechter W, Barabote R, Lorca G, Altermann E, Barrangou R, Ganesan B, Xie Y, Rawsthorne H, Tamir D, Parker C, Breidt F, Broadbent J, Hutkins R, O'Sullivan D, Steele J, Unlu G, Saier M, Klaenhammer T, Richardson P, Kozyavkin S, Weimer B, Mills D. 2006. Comparative genomics of the lactic acid bacteria. *Proc Natl. Acad. Sci. U. S. A.* 103:15611–15616.
2. Papagianni M, Anastasiadou S. 2009. Pediocins: the bacteriocins of pediococci. Sources, production, properties and applications. *Microb. Cell Fact.* 8:3. doi:10.1186/1475-2859-8-3.
3. 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.
4. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.
 5. Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyripides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2005. Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. *Omics* 12:137–141.
 6. Midha S, Ranjan M, Sharma V, Kumari A, Singh PK, Korpole S, Patil PB. 2012. Genome sequence of *Pediococcus pentosaceus* strain IE-3. *J. Bacteriol.* 194:4468.