

Improved Draft Genome Sequence of Probiotic Strain *Lactobacillus gasseri* K7

Primož Treven, Aljoša Trmčič,* Bojana Bogovič Matijašič, Irena Rogelj

University of Ljubljana, Biotechnical Faculty, Institute of Dairy Science and Probiotics, Domžale, Slovenia

* Present address: Aljoša Trmčič, Department of Food Science, Cornell University, Ithaca, New York, USA.

***Lactobacillus gasseri* K7 is an isolate from infant feces and has *in vitro* and *in vivo* established probiotic properties. Here, we report the improved version of the draft genome sequence, which comprises 8 scaffolds (13 contigs), a total length of 1.99 Mb, and 1,841 predicted protein-coding sequences.**

Received 26 June 2014 Accepted 3 July 2014 Published 24 July 2014

Citation Treven P, Trmčič A, Bogovič Matijašič B, Rogelj I. 2014. Improved draft genome sequence of probiotic strain *Lactobacillus gasseri* K7. *Genome Announc.* 2(4):e00725-14. doi:10.1128/genomeA.00725-14.

Copyright © 2014 Treven 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 Primož Treven, primoz.treven@bf.uni-lj.si.

Lactobacillus gasseri is an autochthonous microorganism that colonizes the gastrointestinal tract, oral cavity, and vagina in humans and animals. Strains of *L. gasseri* species are frequently declared to be probiotics (1). *L. gasseri* K7 is an isolate from infant feces and has established probiotic properties *in vitro* and *in vivo* (2–4). The gasserins K7 A (GenBank accession no. EF392861) and K7 B (GenBank accession no. AY307382) produced by this strain are among the first characterized bacteriocins of human-derived probiotic bacteria (5), which are particularly interesting due to a wide range of activities against Gram-positive bacteria, including *Clostridium difficile* and *Clostridium perfringens* (6, 7).

The *L. gasseri* K7 genome was sequenced using 454 Titanium GS FLX+ pyrosequencing (Roche, Branford, CT). The obtained reads were assembled using Newbler version 2.6. The resulting 51 contigs (maximum length, 207,105 bp; minimum length, 101 bp; N_{50} contig size, 86,593; mean coverage, 29 \times) were ordered against the complete genome sequence of *L. gasseri* ATCC 33323 using ABACAS (8). Gap closure was performed by Sanger sequencing of gap-closing PCR products obtained with primers generated with the Primer3 software (9). Gap closure resulted in a final 13 contigs, and the relationships against the reference genome of *L. gasseri* ATCC 33323 were maintained in the GenBank submission by the inclusion of an A Golden Path (AGP) file. Annotation and gene prediction were performed by the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) (10) and the IMG-ER platform (11), which was also used to compare the genome of the *L. gasseri* K7 strain with other genomes. Before final submission, genome annotation was manually curated to correct the nomenclature of bacteriocins and insertion sequences. Artemis (12) and the IMG-ER platform were used for manual curation.

The submitted genome sequence of *L. gasseri* K7 comprises 8 scaffolds (13 contigs), with a total length of 1,993,970 bp, 34.81% G+C content, and 1,841 predicted protein-coding sequences (CDSs). It includes 10 rRNA gene operons and 55 predicted tRNA genes. Compared to *L. gasseri* ATCC 33323, part (278,000 bp) of the *L. gasseri* K7 genome sequence is inverted. The results of a BLAST search against a plasmid-specific database available on the

PATRIC website (13) indicate that contig005 is part of the plasmid sequence. PHAST analysis (14) revealed one complete (ϕ ADH), one questionable (ϕ ADH), and two incomplete prophage regions. One incomplete and 19 complete repeats of insertion sequence ISLga1 (IS30 family) were identified by IS Finder (15) and a BLAST search. One clustered regularly interspaced short palindromic repeat (CRISPR) element (37 bp) with 14 spacers was detected using CRISPRFinder (16). Besides the previously described bacteriocins gasserin K7 A and K7 B (5), analysis with BAGEL3 (17) revealed two additional putative coding genes for bacteriocin helveticin-J and for gasserin A-like bacteriocin (class IIc). *L. gasseri* K7 has 103 genes with no homologs in other publicly available *L. gasseri* genomes. Among the unique genes, we found 4 putative amino acid ABC transporters, 3 different putative glucosidases, a putative arabinose transporter, and 4 different putative glycosyltransferases, which suggests that the strain is adapted to specific sugar- and amino acid-rich environments, such as the infant intestinal tract.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [ASRG00000000](https://www.ncbi.nlm.nih.gov/ASRG00000000/). The version described in this paper is version [ASRG02000000](https://www.ncbi.nlm.nih.gov/ASRG02000000/).

ACKNOWLEDGMENTS

This work was supported by the Competence Center for Biotechnological Development and Innovation (CCBDI), Ministry of Higher Education, Science and Technology of Slovenia, the European Regional Development Fund EU, and by P4-0097 Nutrition and Microbial Ecology of GIT, Slovenian Research Agency.

REFERENCES

1. Selle K, Klaenhammer TR. 2013. Genomic and phenotypic evidence for probiotic influences of *Lactobacillus gasseri* on human health. *FEMS Microbiol. Rev.* 37:915–935.
2. Rogelj I, Matijašič BB. 2006. *Lactobacillus gasseri* LF221 and K7—from isolation to application. *Biologica* 61:761–769. <http://dx.doi.org/10.2478/s11756-006-0154-1>.
3. Sagaya FM, Hacin B, Tompa G, Ihan A, Spela S, Cerne M, Hurrell RF,

- Matijasac BB, Rogelj I, Vergeres G. 2014. *Lactobacillus gasseri* K7 modulates the blood cell transcriptome of conventional mice infected with *Escherichia coli* O157 H7. *J. Appl. Microbiol.* 116:1282–1296. <http://dx.doi.org/10.1111/jam.12440>.
4. Bogovič Matijašič B, Rogelj I. 2011. Bacteriocins of probiotics and enteric cytoprotection, p 313–354. *In* Malago JJ, Koninkx JFJG, Marinšek-Logar R (ed), *Probiotic bacteria and enteric infections: cytoprotection by probiotic bacteria*. Springer Verlag, New York, NY.
 5. Zorič Peternel M, Čanžek Majhenič A, Holo H, Nes I, Salehian Z, Berlec A, Rogelj I. 2010. Wide-inhibitory spectra bacteriocins produced by *Lactobacillus gasseri* K7. *Probiotics Antimicrob. Proteins* 2:233–240. <http://dx.doi.org/10.1007/s12602-010-9044-5>.
 6. Bogovič-Matijašič B, Rogelj I. 2000. *Lactobacillus* K7—a new candidate for a probiotic strain. *Food Technol. Biotechnol.* 38:113–119.
 7. Mavrič A, Tompa G, Trmčič A, Rogelj I, Bogovič Matijašič B. 20 May 2014. Bacteriocins of *Lactobacillus gasseri* K7—Monitoring of gassericin K7 A and B genes' expression and isolation of an active component. *Proc. Biochem.* <http://dx.doi.org/10.1016/j.procbio.2014.04.022>.
 8. Assefa S, Keane TM, Otto TD, Newbold C, Berriman M. 2009. ABACAS: algorithm-based automatic contiguation of assembled sequences. *Bioinformatics* 25:1968–1969. <http://dx.doi.org/10.1093/bioinformatics/btp347>.
 9. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 40:e115. <http://dx.doi.org/10.1093/nar/gks596>.
 10. Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. *Omics* 12: 137–141. <http://dx.doi.org/10.1089/omi.2008.0017>.
 11. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res.* 40:D115–D122. <http://dx.doi.org/10.1093/Nar/Gkr1044>.
 12. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. <http://dx.doi.org/10.1093/bioinformatics/16.10.944>.
 13. Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJ, Yoo HS, Zhang C, Zhang Y, Sobral BW. 2014. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res.* 42:D581–D591. <http://dx.doi.org/10.1093/Nar/Gkt1099>.
 14. Zhou Y, Liang YJ, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. *Nucleic Acids Res.* 39:W347–W352. <http://dx.doi.org/10.1093/Nar/Gkr485>.
 15. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 34:D32–D36. <http://dx.doi.org/10.1093/Nar/Gkj014>.
 16. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a Web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35:W52–W57. <http://dx.doi.org/10.1093/Nar/Gkm360>.
 17. de Jong A, van Heel AJ, Kok J, Kuipers OP. 2010. BAGEL2: mining for bacteriocins in genomic data. *Nucleic Acids Res.* 38:W647–W651. <http://dx.doi.org/10.1093/Nar/Gkq365>.