



Complete Genome Sequence of *Lactobacillus plantarum* Oregon-R-modENCODE Strain BDGP2 Isolated from *Drosophila melanogaster* Gut

Kenneth H. Wan, Charles Yu,* Soo Park, Ann S. Hammonds, Benjamin W. Booth, Susan E. Celniker

Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

ABSTRACT *Lactobacillus plantarum* Oregon-R-modENCODE strain BDGP2 was isolated from the gut of *Drosophila melanogaster* for functional host microbial interaction studies. The complete genome comprised a single circular genome of 3,407,160 bp, with a G+C content of 44%, and four plasmids.

Lactobacillus plantarum increases amino acid metabolism and promotes larval growth in *Drosophila melanogaster* under nutrient-scarce conditions (1). Previously published *L. plantarum* strains include one derived from a laboratory fly strain (KP) and one from wild-caught flies (DF) (2). Here, we report the genome sequence of a strain of *L. plantarum*, Oregon-R-modENCODE strain BDGP2 associated with a sequenced *D. melanogaster* host, to facilitate mechanistic studies of host-microbiome interactions.

L. plantarum Oregon-R-modENCODE strain BDGP2 was isolated from a fecal swab. Bacteria were streaked onto Difco de Man, Rogosa, and Sharpe (MRS) agar (3) agar (BD catalog number 288210) plates, single colonies were amplified, and an aliquot was used for 16S V1 and V4 PCR (4) and sequence identification (5). DNA was isolated (6), and whole-genome sequencing was performed by the National Center for Genome Resources (NCGR), Santa Fe, New Mexico, USA, using Pacific Biosciences (PacBio, Menlo Park, CA, USA) long-read sequencing on the RS II instrument (7). A single-molecule real-time (SMRT) cell library was constructed with 5 to 10 μ g of DNA using the PacBio 20-kb protocol and sequenced on one SMRT cell using P6 polymerase and C4 chemistry with 6-h movie times. Sequencing yielded 82,902 reads with a filtered mean read length of 11,874 bp, totaling 984,453,362 bp (>200-fold coverage of the chromosome and plasmids). The files generated by the PacBio instrument were used for a *de novo* assembly constructed using the HGAP2 protocol from SMRT Analysis version 2.0 (8, 9). The final contigs were manually trimmed and reviewed to produce a single circular chromosome and four plasmids. Annotations of protein-coding open reading frames and noncoding RNAs were predicted using the Rapid Annotation of microbial genomes using Subsystems Technology tool (10) and the GenBank annotation pipeline (11).

The chromosomal genome annotation predicts 3,148 protein-coding genes, 131 pseudogenes, 5 rRNA operons, a single 5S rRNA, and 85 tRNAs (78 with canonical anticodon triplets that base pair with codons for amino acids, 1 suppressor tRNA [tRNA-Sup-TTA], 1 selenocysteine-specific tRNA [tRNA-SeC-TCA], 1 undetermined tRNA [tRNA-OTHER], and 4 pseudo-tRNA genes). Of the 3,148 protein-coding genes, 328 are contained within candidate prophages. Like other *L. plantarum* strains, ours contains integrated likely prophages (Phage_Finder; Omic Tools). There are seven copies, six ranging in size from 36,225 to 46,235 bp and in sequence similarity from 45 to 70% and a partial copy of only 13,544 kb. Together they constitute 7.5% of the genome. Two cornerstone proteins highly conserved in prophages are the large terminase subunit

Received 13 September 2017 Accepted 18 September 2017 Published 12 October 2017

Citation Wan KH, Yu C, Park S, Hammonds AS, Booth BW, Celniker SE. 2017. Complete genome sequence of *Lactobacillus plantarum* Oregon-R-modENCODE strain BDGP2 isolated from *Drosophila melanogaster* gut. Genome Announc 5:e01155-17. <https://doi.org/10.1128/genomeA.01155-17>.

Copyright © 2017 Wan et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Susan E. Celniker, secelniker@lbl.gov.

* Present address: Charles Yu, Genentech, Inc., South San Francisco, California, USA.

and the portal protein (12). These candidate prophages, including the partial copy, encode a large terminase subunit protein, and all but one of our candidate prophages encode a portal protein. In addition, the genome contains four plasmids, pLtBDGP2A (61,448 bp), pLtBDGP2B (58,689 bp), pLtBDGP2C (41,234 bp), and pLtBDGP2D (13,055 bp). Plasmids pLtBDGP2B and pLtBDGP2C contain candidate genes for the plasmid replication initiation proteins RepE and RepA, respectively. Plasmids pLtBDGP2A and pLtBDGP2D show sequence similarity to *L. plantarum* strain DF plasmid unnamed1 (GenBank accession number CP013754) and *L. plantarum* strain TMW 1.277 (palm wine isolate) plasmid pL1277-4 (accession number CP017367), respectively.

Accession number(s). The complete genome sequences of the chromosome and four plasmids of *L. plantarum* Oregon-R-modENCODE strain BDGP2 have been deposited in GenBank under the accession numbers [CP023174](#) (chromosome) and [CP023175](#) (pLtBDGP2A), [CP023176](#) (pLtBDGP2B), [CP023177](#) (pLtBDGP2C), and [CP023178](#) (pLtBDGP2D) (plasmids).

ACKNOWLEDGMENTS

We thank J. B. Brown, J.-H. Mao, A. Snijders, S. Langley, and Nancy Bonini for scientific discussions. We thank Faye Schilkey, Jennifer Jacobi, and Nicholas Devitt of the NCGR for PacBio sequencing. The Oregon-R-modENCODE fly strain, initially obtained from the Bloomington *Drosophila* Stock Center (25211), has been cultured at the BDGP for 6 years.

This work was supported by the Laboratory Directed Research and Development Program of Lawrence Berkeley National Laboratory under U.S. Department of Energy contract DE-AC02-05CH11231.

REFERENCES

1. Broderick NA, Lemaitre B. 2012. Gut-associated microbes of *Drosophila melanogaster*. *Gut Microbes* 3:307–321. <https://doi.org/10.4161/gmic.19896>.
2. Petkau K, Fast D, Duggal A, Foley E. 2016. Comparative evaluation of the genomes of three common *Drosophila*-associated bacteria. *Biol Open* 5:1305–1316. <https://doi.org/10.1242/bio.017673>.
3. de Man JC, Rogosa M, Sharpe ME. 1960. A medium for the cultivation of lactobacilli. *J Appl Bacteriol* 23:130–135. <https://doi.org/10.1111/j.1365-2672.1960.tb00188.x>.
4. Yu Z, Morrison M. 2004. Comparisons of different hypervariable regions of *rrs* genes for use in fingerprinting of microbial communities by PCR-denaturing gradient gel electrophoresis. *Appl Environ Microbiol* 70:4800–4806. <https://doi.org/10.1128/AEM.70.8.4800-4806.2004>.
5. Slatko BE, Kieleczawa J, Ju J, Gardner AF, Hendrickson CL, Ausubel FM. 2011. “First generation” automated DNA sequencing technology. *Curr Protoc Mol Biol* Chapter 7:Unit 7.2. <https://doi.org/10.1002/0471142727.mb0702s96>.
6. Wilson K. 2001. Preparation of genomic DNA from bacteria. *Curr Protoc Mol Biol* Chapter 2:Unit 2.4. <https://doi.org/10.1002/0471142727.mb0204s56>.
7. Korfach J, Bjornson KP, Chaudhuri BP, Cicero RL, Flusberg BA, Gray JJ, Holden D, Saxena R, Wegener J, Turner SW. 2010. Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol* 472:431–455. [https://doi.org/10.1016/S0076-6879\(10\)72001-2](https://doi.org/10.1016/S0076-6879(10)72001-2).
8. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
9. Koren S, Harhay GP, Smith TP, Bono JL, Harhay DM, McVey SD, Radune D, Bergman NH, Phillippy AM. 2013. Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome Biol* 14:R101. <https://doi.org/10.1186/gb-2013-14-9-r101>.
10. 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>.
11. Tatusova T, DiCuccio M, Badretdin A, Chetvermin V, Ciufu S, Li W. 2013. Prokaryotic genome annotation pipeline. *In* The NCBI Handbook, 2nd ed. National Center for Biotechnology Information, Bethesda, MD. <https://www.ncbi.nlm.nih.gov/books/NBK174280>.
12. Casjens S. 2003. Prophages and bacterial genomics: what have we learned so far? *Mol Microbiol* 49:277–300. <https://doi.org/10.1046/j.1365-2958.2003.03580.x>.