AMERICAN
SOCIETY FOR
MICROBIOLOGY

Microbiology®
Resource Announcements

Check for
updates

Chromosomal Sequence of *Lactobacillus brevis* Oregon-R-modENCODE Strain BDGP6, a Lactic Acid Bacterium Isolated from the Gut of *Drosophila melanogaster*

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

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

ABSTRACT *Lactobacillus brevis* Oregon-R-modENCODE strain BDGP6 was isolated from the gut of *Drosophila melanogaster* for functional host-microbial interaction studies. The bacterial chromosome is a single circular DNA molecule of 2,785,111 bp with a G+C content of 46%.

The gut microbiome of *Drosophila melanogaster* is dominated by species from the genera *Lactobacillus* and *Acetobacter* (reviewed in references 1–3). *Lactobacillus* species produce vitamin B_2, which is important for larval development (4, 5), increases amino acid metabolism, and promotes larval growth under conditions of nutrient scarcity (6). The first draft sequence of *Lactobacillus brevis* from *Drosophila melanogaster*, published in 2014, consisted of 117 contigs (3). We report here the complete bacterial chromosomal sequence, consisting of a single circular DNA molecule.

Lactobacillus brevis Oregon-R-modENCODE strain BDGP6 was isolated from the gut of *Drosophila melanogaster* as a single colony from lactobacillus MRS agar plates and cultured at 37°C for 16 to 18 h in MRS broth without shaking, and an aliquot was used for 16S V1 and V4 PCR (7) and sequence identification (reviewed in reference 8). DNA was isolated by cetyltrimethylammonium bromide (CTAB)/NaCl and phenol-chloroform extractions, followed by cesium chloride banding and isopropanol precipitation (9); then, it was sent to the National Center for Genome Resources (NCGR; Santa Fe, NM) for whole-genome sequencing using Pacific Biosciences (Menlo Park, CA) long-read sequencing on the RS II instrument (10). A single-molecule real-time (SMRT) cell library was constructed with 5 to 10 μg of unsheared DNA, size selected to >7 kb (BluePippin, Sage Science) using the PacBio 20-kbp protocol, and sequenced on one SMRT cell using P6 polymerase and C4 chemistry with 6-h movie times. Quality control filtering was performed via the PacBio SMRT Analysis Portal with smrtanalysis_2.3.0.140936.p5.167094 software. Default parameters were used for all software tools, except when otherwise noted.

Sequencing yielded 22,365 reads totaling 319,828,877 bp (chromosomal coverage, >70-fold), with a filtered mean read length of 14,300 bp and an N_{50} value of 20,694 bp. A *de novo* assembly was constructed using the HGAP2 protocol from SMRT Analysis v2.0 (11, 12). The final contig was manually trimmed and reviewed to produce a single circular chromosome. The sequence quality was assessed using minimap2 (13). The percentage of reads mapping to the assembly is 82.84%, and the estimated coverage is 72.58×. The remaining reads map to likely plasmids not characterized here. Annotations were predicted using the RAST v2.0 (annotation scheme, ClassicRAST) tool (14) and the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.2 (15).

The PGAP annotation predicted 2,676 protein-coding genes, 60 pseudogenes, 5 rRNA operons, and 71 tRNAs, 70 with canonical anticodon triplets and 1 undetermined (tRNA-OTHER). Of the 2,676 protein-coding genes, 332 are contained within candidate prophages. Like other *L. brevis* strains, BDGP6 contains integrated likely prophages. There are nine copies, namely, five complete, ranging in size from 38,511 to 49,964 bp, and four partial copies, ranging in size from 4,871 bp to 13,432 bp. Together, the prophage regions

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

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

constitute 8.6% of the genome sequence. At the nucleotide level, they have little sequence similarity except for two copies that share 92% similarity (BLASTN v2.0MP-WashU) (16) for about 10 kb. Two cornerstone proteins highly conserved in prophages, the large terminase subunit and the portal protein (reviewed in reference 17), are encoded by four complete prophage copies. Based on its growth in MRS broth supplemented with antibiotics, strain BDGP6 is resistant to tetracycline (20 μg/ml), spectinomycin (50 μg/ml), kanamycin (50 μg/ml), gentamicin (20 μg/ml), and chloramphenicol (0.5 μg/ml), likely aided by 39 genes encoding ribosomal protection, efflux transporter, and antibiotic-modifying proteins. KEGG Automatic Annotation Server (18) pathway analysis showed that the genome encodes 56 enzymes involved with B vitamin metabolism and biosynthesis. In addition to vitamin B_2 (riboflavin), *L. brevis* BDGP6 encodes a vitamin B_5 (pantothenate) transporter, a cystine transporter (cystine is converted to two cysteines under high pH), and the biosynthetic pathway (five genes encoding the enzymes pantothenate kinase [CoaA], coenzyme A [CoA] biosynthesis bifunctional protein [CoaBC], phosphopantetheine adenylyltransferase [coaD], and dephospho-CoA kinase [coaE]) to convert these two molecules into CoA, an essential component of fatty acid synthesis and energy production.

Data availability. The chromosome sequence of *Lactobacillus brevis* Oregon-R-modENCODE strain BDGP6 is available under the GenBank accession number CP024635. The sequencing reads are available under the SRA accession number SRR12450050.

REFERENCES

1. Douglas AE. 2018. The Drosophila model for microbiome research. Lab Anim (NY) 47:157–164. https://doi.org/10.1038/s41684-018-0065-0.
2. Wong CNA, Ng P, Douglas AE. 2011. Low-diversity bacterial community in the gut of the fruitfly Drosophila melanogaster. Environ Microbiol 13:1889–1900. https://doi.org/10.1111/j.1462-2920.2011.02511.x.
3. Newell PD, Chaston JM, Wang Y, Winans NJ, Sannino DR, Dobson AJ, Kagle J, Douglas AE. 2014. In vivo function and comparative genomic analyses of the Drosophila gut microbiota identify candidate symbiosis factors. Front Microbiol 5:576. https://doi.org/10.3389/fmicb.2014.00576.
4. Fridmann-Sirkis Y, Stern S, Elgart M, Galili M, Zeisel A, Shental N, Soen Y. 2014. Delayed development induced by toxicity to the host can be inherited by a bacterial-dependent, transgenerational effect. Front Genet 5:27. https://doi.org/10.3389/fgene.2014.00027.
5. Wong AC-N, Dobson AJ, Douglas AE. 2014. Gut microbiota dictates the metabolic response of Drosophila to diet. J Exp Biol 217:1894–1901. https://doi.org/10.1242/jeb.101725.
6. Broderick NA, Lemaitre B. 2012. Gut-associated microbes of Drosophila melanogaster. Gut Microbes 3:307–321. https://doi.org/10.4161/gmic.19896.
7. 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.
8. Slatko BE, Kieleczawa J, Ju J, Gardner AF, Hendrickson CL, Ausubel FM. 2011. "First generation" automated DNA sequencing technology. Curr Protoc Mol Biol 96:7.2.1–7.2.28. https://doi.org/10.1002/0471142727.mb0702s96.
9. 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.
10. Korlach 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.
11. 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.
12. Koren S, Harhay GP, Smith TPL, 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.
13. Li H. 2018. minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191.
14. 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.
15. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufo S, Li W. 2013. Prokaryotic Genome Annotation Pipeline. In Beck J, Benson D, Coleman J, Hoeppner M, Johnson M, Maglott D, Mizrachi I, Morris R, Ostell J, Pruitt K, Rubinstein W, Sayers E, Sirotkin K, Tatusova T (ed), The NCBI Handbook [Internet]. National Center for Biotechnology Information, Bethesda, MD.
16. Lopez R, Silventoinen V, Robinson S, Kibria A, Gish W. 2003. WU-Blast2 server at the European Bioinformatics Institute. Nucleic Acids Res 31:3795–3798. https://doi.org/10.1093/nar/gkg573.
17. 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.
18. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35:W182–W185. https://doi.org/10.1093/nar/gkm321.