Draft Genome Sequence of *Lactobacillus paracasei* DmW181, a Bacterium Isolated from Wild *Drosophila*

Austin J. Hammer,[a] Amber Walters,[a] Courtney Carroll,[a] Peter D. Newell,[b] John M. Chaston[a]

Department of Plant and Wildlife Sciences, Brigham Young University, Provo, Utah, USA[a]; Department of Biological Sciences, SUNY-Oswego, Oswego, New York, USA[b]

ABSTRACT The draft genome sequence of *Lactobacillus paracasei* DmW181, an anaerobic bacterium isolate from wild *Drosophila* flies, is reported here. Strain DmW181 possesses genes for sialic acid and mannose metabolism. The assembled genome is 3,201,429 bp, with 3,454 predicted genes.

*L*actobacillus paracasei* is a Gram-positive anaerobic bacterium that is part of the lactic acid bacteria (LAB) group that have been studied for, among other reasons, use as probiotics (1). LAB are also commonly associated with fruit flies (2). In the present study, we sequenced and analyzed the genome of *L. paracasei* DmW181 derived from wild *Drosophila* flies.

Wild *Drosophila* flies were collected from a household kitchen in Ithaca, NY, USA (42.427481°N 76.463983°W). Flies were homogenized, and the homogenate was plated on modified de Man-Rogosa-Sharpe (MRS) medium (3). The 16S rRNA gene from a single colony was analyzed by Sanger sequencing, and a preliminary taxonomic assignment as *Lactobacillus casei* or *L. paracasei* was made. DNA was extracted using the Qiagen DNeasy blood and tissue kit (Qiagen, Hilden, Germany). Using NEBNext double-stranded DNA (dsDNA) fragmentase (NEB, Ipswich, MA, USA), DNA was fragmented and adaptors were ligated using NEBNext UltraII DNA library prep kit components, as instructed by the manufacturer (NEB). A magnetic bead size selection achieved an insert size of 1,100 nucleotides, and the library was sequenced on an Illumina HiSeq 2500 with chemistry for 250-bp paired-end reads. A total of 5,290,771 reads passed quality filtering and were used for genome assembly. Reads were assigned to one of five separate bins, each representing 200× genome coverage, and assembled into contigs using Velvet 1.2.10 (4), with k-mer lengths varying between 185 and 207. For each bin, a single assembly that maximized N_{50} was selected and used to assemble the final genome, as in our previous work (5). The final assembly contained 3,201,429 nucleotides in 127 contigs, with a maximum contig length of 515,598 bp and N_{50} of 72,562 bp. A total of 3,454 putative genes were predicted by the NCBI Prokaryotic Genome Annotation Pipeline. When we compared the shotgun genome assembly with *L. casei* and *L. paracasei* isolates available in the JSpeciesWS Web server (6), the highest similarity was 99.9% with *L. paracasei* Lpp46, so we named this strain *L. paracasei* DmW181.

Preliminary genome analysis and annotation with RAST version 2.0 (7) were performed to evaluate unique genetic features of *L. paracasei* DmW181 that could influence interactions with *Drosophila*. Comparative genomic analysis revealed genes unique to *L. paracasei* DmW181, relative to the 2 *Lactobacillus casei* isolates in RAST, *L. casei* ATCC 334 and *Lactobacillus casei* BL23. Sialic acid metabolism genes were found, including a utilization regulator and genes that participate at the interface of mannose and sialic acid metabolism. Previous research has shown that sialic acid could be

important for the growth of some bacteria that colonize the mammalian gut (8–10). We suggest that future work could test if sialic acid influences colonization of the *Drosophila* gut by *Lactobacillus* or other bacteria.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. NDXH00000000. The version described in this paper is version NDXH01000000.

REFERENCES

1. Holzapfel WH, Haberer P, Geisen R, Björkroth J, Schillinger U. 2001. Taxonomy and important features of probiotic microorganisms in food and nutrition. Am J Clin Nutr 73:365S–373S.
2. Broderick NA, Lemaitre B. 2012. Gut-associated microbes of *Drosophila melanogaster*. Gut Microbes 3:307–321. https://doi.org/10.4161/gmic.19896.
3. Newell PD, Douglas AE. 2014. Interspecies interactions determine the impact of the gut microbiota on nutrient allocation in *Drosophila melanogaster*. Appl Environ Microbiol 80:788–796. https://doi.org/10.1128/AEM.02742-13.
4. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.
5. Newell PD, Chaston JM, Wang Y, Winans NJ, Sannino DR, Wong AC, 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.
6. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a Web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929–931. https://doi.org/10.1093/bioinformatics/btv681.
7. 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. https://doi.org/10.1186/1471-2164-9-75.
8. Koropatkin NM, Cameron EA, Martens EC. 2012. How glycan metabolism shapes the human gut microbiota. Nat Rev Microbiol 10:323–335. https://doi.org/10.1038/nrmicro2746.
9. Etzold S, Kober OI, Mackenzie DA, Tailford LE, Gunning AP, Walshaw J, Hemmings AM, Juge N. 2014. Structural basis for adaptation of lactobacilli to gastrointestinal mucus. Environ Microbiol 16:888–903. https://doi.org/10.1111/1462-2920.12377.
10. McDonald ND, Lubin JB, Chowdhury N, Boyd EF. 2016. Host-derived sialic acids are an important nutrient source required for optimal bacterial fitness *in vivo*. mBio 7:e02237-15. https://doi.org/10.1128/mBio.02237-15.