AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Complete Genome Sequence of *Acetobacter pomorum* Oregon-R-modENCODE Strain BDGP5, an Acetic Acid Bacterium Found in the *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 *Acetobacter pomorum* Oregon-R-modENCODE strain BDGP5 was isolated from *Drosophila melanogaster* for functional host-microbe interaction studies. The complete genome is composed of a single chromosomal circle of 2,848,089 bp, with a G+C content of 53% and three plasmids of 131,455 bp, 19,216 bp, and 9,160 bp.

Received 25 October 2017 **Accepted** 30 October 2017 **Published** 30 November 2017

Citation Wan KH, Yu C, Park S, Hammonds AS, Booth BW, Celniker SE. 2017. Complete genome sequence of *Acetobacter pomorum* Oregon-R-modENCODE strain BDGP5, an acetic acid bacterium found in the *Drosophila melanogaster* gut. Genome Announc 5:e01333-17. https://doi.org/10.1128/genomeA.01333-17.

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

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

In *Drosophila*, *Acetobacter* spp. are among the major commensals of the gut microbiota and contribute to larval growth (reviewed in reference 1). Furthermore, monocolonization of *Drosophila* with *Acetobacter* species significantly reduced host development (2). The first draft sequence of *Acetobacter pomorum* from *Drosophila*, published in 2014, consisted of 137 contigs (3). We report here the sequence of the complete genome, consisting of a single chromosome and three plasmids.

A. pomorum Oregon-R-modENCODE strain BDGP5 was isolated from homogenized dissected guts. Bacteria were streaked onto Nutrient broth agar (BD catalog number 213000) plates, single colonies were amplified in culture, and an aliquot was used for 16S V1 and V4 PCR (4) and sequence identification (reviewed in reference 5). DNA for sequencing was isolated (6), and whole-genome DNA sequencing was performed by the National Center for Genome Resources (NCGR) (Santa Fe, NM), using Pacific Biosciences (PacBio) long reads on the RSII instrument (7). A single-molecule real-time (SMRT) cell library was constructed with 5 to 10 μg of DNA using the PacBio 20-kbp protocol. The library was sequenced using P6 polymerase and C4 chemistry with 6-h movie times. Sequencing yielded a total of 52,913 reads with filtered mean read length of 12,181 bp, totaling 644,585,547 bp (>150-fold coverage). A *de novo* assembly was constructed using the Hierarchical Genome Assembly Process (HGAP2) protocol from SMRT Analysis version 2.0 (8–10). The final contigs were manually trimmed and reviewed to produce a single circular chromosome and three plasmids. Annotations of protein-coding open reading frames and noncoding RNAs (ncRNAs) were predicted using the Rapid Annotations using Subsystems Technology (RAST) tool (11) and the GenBank annotation pipeline (12).

The chromosomal genome annotation predicted 2,858 genes in total, with 2,782 protein-coding genes, 77 RNA genes, including 5 rRNA operons, 57 tRNA genes, 1 transfer-messenger RNA (tmRNA) gene, 3 noncoding RNA (ncRNA) genes, and 120 pseudogenes. Of the 2,782 protein-coding genes, 46 are contained within two partial cryptic prophages (20 kb, 24 genes; 17.7 kb, 22 genes). The first is characterized by a number of genes encoding Mu-like FluMu proteins (13). The candidate prophages are only 1.3% of the genome. Additionally, the genome contains three plasmids, pApBDGP5A (151,013 bp), pApBDGP5B (9,160 bp), and pApBDGP5C (19,216 bp), with predicted G+C contents of 52%, 54%, and 54%, respectively. The plasmids contain 203

candidate protein-coding genes. Among them are genes essential for conjugation (*traG* [pApBDGP5A]), plasmid replication (*repA* [pApBDGP5A and pApBDGP5C]), and toxin-antitoxin (TA) modules associated with stable plasmid inheritance at cell division, including *relE/stbE-relB/stbD* (pApBDGP5A) *doc-phd*, *yoeB-yefM,* and *vapC-B* (pApBDGP5C) (reviewed in reference 14).

Intriguingly, the bacterial chromosome also contains toxin-antitoxin modules for *doc-phd*, *mazF-E*, *vapC-B*, and *hicA-B*, thought in the chromosomal case to be important for bacterial stress response (reviewed in reference 15).

Our sequence has significant similarity to *Acetobacter pasteurianus*, being 98.8% identical by average nucleotide identity (ANI) (16). Despite the sequence similarity, phenotypically, our strain belongs to the *A. pomorum* species based on its ability to grow in *n*-propanol and glycerol (17).

Accession number(s). The complete chromosome and three plasmid sequences of *A. pomorum* Oregon-R-modENCODE strain BDGP5 are deposited in GenBank under the accession numbers CP023657 (chromosome), CP023658 (plasmid pApBDGP5A), CP023659 (plasmid pApBDGP5B), and CP023660 (plasmid pApBDGP5C).

ACKNOWLEDGMENTS

We thank J. B. Brown, J.-H. Mao, A. Snijders, S. Langley, and N. 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 the Lawrence Berkeley National Laboratory under U.S. Department of Energy contract number 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. 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.
3. 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.
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.
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. 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.
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. Chaisson MJ, Tesler G. 2012. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. BMC Bioinformatics 13:238. https://doi.org/10.1186/1471-2105-13-238.
11. 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.
12. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufo S, Li W. 2013. Prokaryotic Genome Annotation Pipeline. The NCBI handbook, 2nd ed. National Center for Biotechnology Information, Bethesda, MD.
13. Morgan GJ, Hatfull GF, Casjens S, Hendrix RW. 2002. Bacteriophage mu genome sequence: analysis and comparison with Mu-like prophages in *Haemophilus*, *Neisseria* and *Deinococcus*. J Mol Biol 317:337–359. https://doi.org/10.1006/jmbi.2002.5437.
14. Holčík M, Iyer VN. 1997. Conditionally lethal genes associated with bacterial plasmids. Microbiology 143:3403–3416. https://doi.org/10.1099/00221287-143-11-3403.
15. Coussens NP, Daines DA. 2016. Wake me when it's over—bacterial toxin-antitoxin proteins and induced dormancy. Exp Biol Med 241:1332–1342. https://doi.org/10.1177/1535370216651938.
16. Luo C, Rodriguez-R LM, Konstantinidis KT. 2013. A user's guide to quantitative and comparative analysis of metagenomic datasets. Methods Enzymol 531:525–547. https://doi.org/10.1016/B978-0-12-407863-5.00023-X.
17. Sokollek SJ, Hertel C, Hammes WP. 1998. Description of *Acetobacter oboediens* sp. nov. and *Acetobacter pomorum* sp. nov., two new species isolated from industrial vinegar fermentations. Int J Syst Bacteriol 48:935–940. https://doi.org/10.1099/00207713-48-3-935.