Draft Genome Sequence of *Erwinia dacicola*, a Dominant Endosymbiont of Olive Flies

Anne M. Estes,a,b David J. Hearn,b Suvarna Nadendla,a Elizabeth A. Pierson,c Julie C. Dunning Hotoppa,d

a Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, USA
b Department of Biological Sciences, Towson University, Towson, Maryland, USA
c Department of Horticultural Sciences, Texas A&M University, College Station, Texas, USA
d Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, Maryland, USA

**ABSTRACT** *Erwinia dacicola* is a dominant endosymbiont of the pestiferous olive fly. Its genome is similar in size and GC content to those of free-living *Erwinia* species, including the plant pathogen *Erwinia amylovora*. The *E. dacicola* genome encodes the metabolic capability to supplement and detoxify the olive fly’s diet in larval and adult stages.

*Erwinia dacicola* Oroville was sequenced to determine its potential beneficial role in supplementing the diet of the tephritid olive fly, *Bactrocera oleae*. *E. dacicola* is the dominant bacterium that resides in specialized digestive system structures within the olive fly (1, 2) and is found in the vast majority of wild olive flies sampled from Greece, Italy, and the United States (3, 4). Most laboratory olive fly populations reared on artificial diets are not colonized by *E. dacicola* (3) and are not as healthy as wild olive flies (3, 5). To date, a draft *E. dacicola* genome sequence (6) and transcriptome (7) are available from populations of olive flies in Greece. European and U.S. olive flies belong to haplotypes hTA and hTB, respectively (4). The U.S. population most likely was founded by European olive flies, providing an interesting population for comparison.

*E. dacicola* has not been successfully cultured from the olive fly (1, 7, 8); therefore, DNA for the *E. dacicola* genome was isolated from four sets of 6 to 10 sterilely isolated olive fly esophageal bulbs. The bulbs were homogenized, and the bacteria were separated from host tissues using centrifugation. DNA was extracted using the DNeasy (Qiagen) Gram-negative protocol, and the 16S rRNA was amplified, cloned, and sequenced to verify that the DNA was solely from *E. dacicola* without contaminants. DNA was subsequently amplified using an Illustra GenomiPhi version 2 kit, and ~8 μg was used to construct tagged, multiplexed libraries for sequencing on an Illumina Genome Analyzer II at the Arizona Genomics Institute at the University of Arizona, generating a total of 25,876,346 paired 75-bp reads. Except where otherwise specified, default parameters were used for all software. Sequences were assembled into contigs using *de novo* assembly in ABySS-pe version 1.2.5 with 4 different kmer sizes (40, 45, 50, 55) with a minimum contig length of 300 bp. The K55 assembly was selected since it had a genome size closest to that of the genome of *E. amylovora* and is the most complete *E. dacicola* genome sequenced to date. A total of 1,039 contigs spanning 2,858,157 bp were produced with an N50 value of 5,472 bp, maximum scaffold size of 79 kbp, 52.2% GC content, and >300× sequencing depth. The Institute for Genome Sciences (IGS) Annotation Engine (9) using Glimmer version 3.02 (10) identified 4,033 open reading frames (ORFs). Both the genome size and GC content are more similar to those of free-living bacteria than those of other intracellular bacteria found in other insects.

The genome encodes the ability to supplement amino acids and vitamins missing from the olive fruit on which the larvae feed. Similar to other draft genomes, the
encoded potential of *E. dacicola* suggests that this endosymbiont may supplement and detoxify the fly diet.

**Data availability.** The reads (Sequence Read Archive number [SRP155530](https://doi.org/10.1128/AEM.00778-09)), assembly, and annotation can be accessed through BioProject number [PRJNA288714](https://doi.org/10.1128/genomeA.00896-16) and BioSample number [SAMN03836967](https://doi.org/10.1128/genomeA.00896-16). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number [LJAM00000000](https://doi.org/10.1128/genomeA.00896-16). The version described in this paper is version LJAM02000000.

**ACKNOWLEDGMENTS**

We thank Leland S. Pierson for graciously providing access to laboratory equipment and disposable supplies. Hannah Burrack collected the insects. Yeisoo Yu and other members of the Arizona Genomics Institute DNA sequencing center sequenced the samples and assembled the reads. Sean Daugherty and Michelle Giglio ran the IGS Annotation Engine on the sequenced reads.

This work was partially supported by a fellowship to A.M.E. from the Orchid Society of Arizona. This research was conducted on the National Science Foundation-funded MRI-R2 project number DBI-0959894.

**REFERENCES**

1. Estes AM, Hearn DJ, Bronstein JL, Pierson EA. 2009. The olive fly endosymbiont, *Candidatus Erwinia dacicola*, switches from an intracellular existence to an extracellular existence during host insect development. Appl Environ Microbiol 75:7097–7106. [https://doi.org/10.1128/AEM.00778-09](https://doi.org/10.1128/AEM.00778-09).

2. Capuzzo C, Firrao G, Mazzon L, Squartini A, Girolami V. 2005. *“Candidatus Erwinia dacicola,”* a coevolved symbiotic bacterium of the olive fly *Bactrocera oleae* (Gmelin). Int J Syst Evol Microbiol 55:1641–1647. [https://doi.org/10.1099/ijs.0.63663-0](https://doi.org/10.1099/ijs.0.63663-0).

3. Estes AM, Hearn DJ, Burrack HJ, Rempoulakis P, Pierson EA. 2012. Prevalence of *Candidatus Erwinia dacicola* in wild and laboratory olive fruit fly populations and across developmental stages. Environ Entomol 41:265–274. [https://doi.org/10.1603/EN11245](https://doi.org/10.1603/EN11245).

4. Savio C, Mazzon L, Martínez-Sañudo I, Simonato M, Squartini A, Girolami V. 2012. Evidence of two lineages of the symbiont *Candidatus Erwinia dacicola* in Italian populations of *Bactrocera oleae* (Rossi) based on 16S rRNA gene sequences. Int J Syst Evol Microbiol 62:179–187. [https://doi.org/10.1099/ijs.0.030668-0](https://doi.org/10.1099/ijs.0.030668-0).

5. Estes AM, Nestel D, Belcaro A, Jessup A, Rempoulakis P, Economopoulos AP. 2012. A basis for the renewal of sterile insect technique for the olive fly, *Bactrocera oleae* (Rossi). J App Entomol 136:1–16. [https://doi.org/10.1111/j.1439-0418.2011.01620.x](https://doi.org/10.1111/j.1439-0418.2011.01620.x).

6. Blow F, Gioti A, Starns D, Ben-Yosef M, Pasternak Z, Jurkevitch E, Vontas J, Darby AC. 2016. Draft genome sequence of the *Bactrocera oleae* symbiont *“Candidatus Erwinia dacicola.”* Genome Announc 4:e00896-16. [https://doi.org/10.1128/genomeA.00896-16](https://doi.org/10.1128/genomeA.00896-16).

7. Pavlidi N, Gioti A, Wybouw N, Dermauw W, Ben-Yosef M, Yuval B, Jurkevitch E, Kampouraki A, Van Leeuwen T, Vontas J. 2017. Transcriptomic responses of the olive fruit fly *Bactrocera oleae* and its symbiont *Candidatus Erwinia dacicola* to olive feeding. Sci Rep 7:42633. [https://doi.org/10.1038/srep42633](https://doi.org/10.1038/srep42633).

8. Ben-Yosef M, Aharon Y, Jurkevitch E, Yuval B. 2010. Give us the tools and we will do the job: symbiotic bacteria affect olive fly fitness in a diet-dependent fashion. Proc Biol Sci 277:1545–1552. [https://doi.org/10.1098/rspb.2009.2102](https://doi.org/10.1098/rspb.2009.2102).

9. Galens K, Orvis J, Daugherty S, Creasy HH, Angiuoli S, White O, Wortman J, Mahurkar A, Giglio MG. 2011. The IGS standard operating procedure for automated prokaryotic annotation. Stand Genomic Sci 4:244–251. [https://doi.org/10.4056/sigs.1223234](https://doi.org/10.4056/sigs.1223234).

10. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiotic DNA with Glimmer. Bioinformatics 23: 673–679. [https://doi.org/10.1093/bioinformatics/btm009](https://doi.org/10.1093/bioinformatics/btm009).