

# Draft Genome Sequence of the *Bactrocera oleae* Symbiont “*Candidatus Erwinia dacicola*”

Frances Blow,<sup>a</sup> Anastasia Gioti,<sup>b</sup> David Starns,<sup>a</sup> Michael Ben-Yosef,<sup>c</sup> Zohar Pasternak,<sup>d</sup> Edouard Jurkevitch,<sup>d</sup> John Vontas,<sup>b,e</sup> Alistair C. Darby<sup>a</sup>

Institute of Integrative Biology, University of Liverpool, Liverpool, Merseyside, United Kingdom<sup>a</sup>; Institute of Molecular Biology & Biotechnology, Foundation for Research & Technology Hellas, Heraklion Crete, Greece<sup>b</sup>; Department of Entomology, The Hebrew University of Jerusalem, Rehovot, Israel<sup>c</sup>; Department of Plant Pathology and Microbiology, Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot, Israel<sup>d</sup>; Department of Crop Science, Agricultural University of Athens, Athens, Greece<sup>e</sup>

“*Candidatus Erwinia dacicola*” is a *Gammaproteobacterium* that forms a symbiotic association with the agricultural pest *Bactrocera oleae*. Here, we present a 2.1-Mb draft hybrid genome assembly for “*Ca. Erwinia dacicola*” generated from single-cell and metagenomic data.

Received 20 July 2016 Accepted 26 July 2016 Published 15 September 2016

Citation 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(5):e00896-16. doi:10.1128/genomeA.00896-16.

Copyright © 2016 Blow 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 Alistair C. Darby, [acdarby@liv.ac.uk](mailto:acdarby@liv.ac.uk).

The association between *Bactrocera oleae* (*Tephritidae*) and a bacterial symbiont was first discovered in 1909 (1). Several studies have since identified this organism as “*Candidatus Erwinia dacicola*” (*Enterobacteriaceae*) (2–4). “*Ca. Erwinia dacicola*” plays a role in nutrient provisioning, particularly during juvenile development in unripe olives, where it is essential for larval survival (5–8). However, due to a lack of genomic resources and the inability to culture “*Ca. Erwinia dacicola*,” the metabolic basis of its association with *B. oleae* remains elusive. We present a draft of the “*Ca. Erwinia dacicola*” genome sequence that will inform future investigations into the functional and evolutionary foundations of the symbiosis.

Multiple single-cell (eight) and metagenomic (two) libraries were used to generate a draft hybrid genome assembly of “*Ca. Erwinia dacicola*.” Single-cell libraries were prepared from the guts of adult female flies collected in Heraklion, Greece, stained with CellTracker deep red, and sorted on a Sony SH800. Genomic DNA was amplified using the REPLI-g kit (Qiagen) and was validated as “*Ca. Erwinia dacicola*” by amplification of the 16S rRNA gene, followed by digestion with the restriction enzyme PstI (9). Shotgun libraries were then prepared with the NEBNext Ultra DNA kit (New England Biosciences) and sequenced on an Illumina MiSeq sequencer at the Centre for Genomic Research, University of Liverpool, United Kingdom. Metagenomic shotgun and mate-pair (2- to 6-kb) libraries were prepared from gastric ceca dissected from third-instar larvae isolated in Israel from unripe olives and from a mixture of ripe and unripe olives, respectively. DNA for the shotgun library was extracted using an adapted cetyltrimethylammonium bromide (CTAB) method (10) with additional bead beating and lysozyme digestion, and the library was prepared with the Ovation rapid DR multiplex system (NuGen). The mate-pair library was prepared with the gel-free NexteraMate protocol from DNA extracted with the Chemagic DNA bacteria kit (Chemagen). Both libraries were prepared and sequenced on

an Illumina MiSeq sequencer by LGC Genomics GmbH (Berlin, Germany).

Reads were assembled with SPAdes version 3.7.1 (11) in single-cell mode. Using Blobology (12), contigs identified as belonging to other organisms based on coverage and G+C content were excluded, and the 12,519,932 reads that mapped back to putative “*Ca. Erwinia dacicola*” contigs were extracted and reassembled with SPAdes. The result was a 2.1-Mb assembly comprising 333 scaffolds (>500 bp) at ~1,000× coverage, with an  $N_{50}$  of 9,998 bp. The assembly was assessed as in reference 13 and was found to be 92% complete in comparison to free-living bacteria and 100% complete in comparison to the endosymbiotic bacteria of aphids and tsetse flies, *Buchnera aphidicola* and *Wigglesworthia glossinidia*, respectively. Its G+C content (53.5%) is similar to that observed in other members of the *Erwinia* genus (14) and higher than that of other vertically transmitted endosymbiotic bacteria (15). The genome contains 2,407 protein-coding genes and 28 RNA-coding genes, based on annotation with PROKKA version 1.5.2 (16).

**Accession number(s).** This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. MAZB00000000. The version described in this paper is version MAZB01000000 and BioProject no. PRJNA326914.

## ACKNOWLEDGMENTS

We thank Anastasia Kampouraki for technical support and sample collection. The single-cell isolation and sequencing were conducted in the Single Cell Lab at the Centre for Genomic Research, University of Liverpool.

The funders had no role in the study design, data analysis, decision to publish, or preparation of the manuscript.

## FUNDING INFORMATION

This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC), Oxitec Ltd. through an iCASE studentship (BB/K501773/1) awarded to A.C.D., and the ARISTEIA Action of the Opera-

tional Programme Education and Lifelong Learning and is cofunded by the European Social Fund (ESF) and National Resources (code number 4937) to J.V.

## REFERENCES

- Petri L. 1909. Ricerche sopra i batteri intestinali della mosca olearia. Memorie della Regia Stazione di Patologia Vegetale di Roma, Rome, Italy.
- 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. <http://dx.doi.org/10.1099/ijs.0.63653-0>.
- 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. <http://dx.doi.org/10.1128/AEM.00778-09>.
- Sacchetti P, Granchietti A, Landini S, Viti C, Giovannetti L, Belcari A. 2008. Relationships between the olive fly and bacteria. *J Appl Entomol* 132:682–689. <http://dx.doi.org/10.1111/j.1439-0418.2008.01334.x>.
- Hagen KS. 1966. Dependence of the olive fly, *Dacus oleae*, larvae on symbiosis with *Pseudomonas savastanoi* for the utilization of olive. *Nature* 209:423–424. <http://dx.doi.org/10.1038/209423a0>.
- 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. <http://dx.doi.org/10.1098/rspb.2009.2102>.
- Ben-Yosef M, Pasternak Z, Jurkevitch E, Yuval B. 2014. Symbiotic bacteria enable olive flies (*Bactrocera oleae*) to exploit intractable sources of nitrogen. *J Evol Biol* 27:2695–2705. <http://dx.doi.org/10.1111/jeb.12527>.
- Ben-Yosef M, Pasternak Z, Jurkevitch E, Yuval B. 2015. Symbiotic bacteria enable olive fly larvae to overcome host defences. *R Soc Open Sci* 2:150170. <http://dx.doi.org/10.1098/rsos.150170>.
- Estes AM, Segura DF, Jessup A, Wornoayporn V, Pierson EA. 2014. Effect of the symbiont *Candidatus* *Erwinia dacicola* on mating success of the olive fly *Bactrocera oleae* (Diptera: Tephritidae). *Int J Trop Insect Sci* 34:S123–S131. <http://dx.doi.org/10.1017/S1742758414000174>.
- Xin Z, Chen J. 2012. A high throughput DNA extraction method with high yield and quality. *Plant Methods* 8:26. <http://dx.doi.org/10.1186/1746-4811-8-26>.
- Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Pribelsky A, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R. 2013. Assembling genomes and mini-metagenomes from highly chimeric reads. *Res Comput Mol Biol* 7821:158–170.
- Kumar S, Jones M, Koutsovoulos G, Clarke M, Blaxter M. 2013. Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. *Front Genet* 4:237. <http://dx.doi.org/10.3389/fgene.2013.00237>.
- Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu WT, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM. 2013. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499:431–437. <http://dx.doi.org/10.1038/nature12352>.
- Starr MP, Chatterjee AK. 1972. The genus *Erwinia*: enterobacteria pathogenic to plants and animals. *Annu Rev Microbiol* 26:389–426. <http://dx.doi.org/10.1146/annurev.mi.26.100172.002133>.
- Moran NA. 1996. Accelerated evolution and Muller’s ratchet in endosymbiotic bacteria. *Proc Natl Acad Sci U S A* 93:2873–2878. <http://dx.doi.org/10.1073/pnas.93.7.2873>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <http://dx.doi.org/10.1093/bioinformatics/btu153>.