JB Accepts, published online ahead of print on 7 January 2011
J. Bacteriol. doi:10.1128/JB.01513-10

Genome Announcement

Genome sequence of the plant pathogenic bacterium *Dickeya dadantii* 3937

Jeremy D. Glasner[1], Ching-Hong Yang[2], Sylvie Reverchon[3], Nicole Hugouvieux-Cotte-Pattat[3], Guy Condemine[3], Jean-Pierre Bohin[4], Frédérique Van Gijsegem[5], Shihui Yang[2], Thierry Franza[5], Dominique Expert[5], Guy Plunkett III[6], Michael J. San Francisco[7], Amy O. Charkowski[8], Béatrice Py[9], Kenneth Bell[10], Lise Rauscher[5], Pablo Rodriguez-Palenzuela[11], Ariane Toussaint[12], Maria C. Holeva[10], Sheng Yang He[13], Vanessa Douet[9], Martine Boccara[14], Carlos Blanco[15], Ian Toth[10], Bradley D. Anderson[1], Bryan S. Biehl[11], Bob Mau[1], Sarah M. Flynn[10], Frédéric Barras[9], Magdalen Lindeberg[16], Paul Birch[10], Shinji Tsuyumu[17], Xiangyang Shi[18], Michael Hibbing[1], Mee-Ngan Yap[8], Mathilde Carpentier[14], Elie Dassa[19], Masahiro Umehara[17], Jihyun F. Kim[20], Michael Rusch[1], Pritin Soni[1], George F. Mayhew[1], Derrick E. Fouts[21], Steven R. Gill[21], Frederick R. Blattner[6], Noel T. Keen[18], and Nicole T. Perna[6]

[1] Genome Center, University of Wisconsin – Madison, 425 Henry Mall, Madison, WI 53706, USA

[2] Department of Biological Sciences, University of Wisconsin-Milwaukee, WI 53211, USA

[3] Microbiologie, Adaptation et Pathogenie, UMR 5240 CNRS, Université Lyon 1 Villeurbanne 69622, INSA-Lyon, 69621 Villeurbanne Cedex, France

[4] UMR 8576 CNRS/U.S.T.L., Batiment C9, 59655 Villeneuve d'Ascq Cedex, France.

[5] Laboratoire Interactions Plantes-Pathogènes, UMR 217 INRA/INA-PG/UPMC, 16 rue Claude Bernard, 75005 Paris, France

[6] Department of Genetics and Genome Center of Wisconsin, University of Wisconsin – Madison, 425 Henry Mall, Madison, WI 53706, USA

[7] Department of Biological Sciences, Texas Tech University, Lubbock, TX 79409, USA

[8] Department of Plant Pathology, University of Wisconsin – Madison, Madison, WI 53706, USA

[9] Université de la Méditerrannée, Laboratoire de Chimie Bacterienne, CNRS UPR9043, 13402 Marseille Cedex 20, France.

[10] Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK

[11] Departamento de Biotecnología-UPM, E.T.S. Ingenieros Agrónomos, E-28040 Madrid, Spain

[12] Service de Conformation de Macromolécules Biologiques et de Bioinformatique SCMBB), Université Libre de Bruxelles, Bruxelles, Belgium

[13] MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824-1312, USA

[14] Atelier de Bioinformatique, Université Paris 6, 75005 Paris, France

[15] CNRS UMR 6026, Universite de Rennes I, 35042 Rennes cedex, France

[16] Magdalen Lindeberg, Department of Plant Pathology, Cornell University, Ithaca, NY 14853, USA

[17] Institute for Molecular Biology and Biotechnology, Shizuoka University, Shizuoka #422-8529, Japan

[18] Department of Plant Pathology, University of California, Riverside, CA 92521, USA

19 Unité des Membranes Bactériennes CNRS URA 2172, Département de Microbiologie Fondamentale et Médicale, Institut Pasteur, 75724 Paris Cedex 15, France

[20] Korea Research Institute of Bioscience and Biotechnology

[21] The Institute for Genomic Research, Rockville, MD 20850, USA

Corresponding author : Nicole T. Perna

Mailing address : Department of Genetics and Genome Center of Wisconsin, University of Wisconsin – Madison, 425 Henry Mall, Madison, WI 53706, USA

Phone : 608-890-0171

Fax : 608-265-9485

E-mail : ntperna@wisc.edu

Abstract

Dickeya dadantii is a plant pathogenic enterobacterium responsible of the soft rot disease of many plants of economic importance. We present here the sequence of strain 3937, a strain widely used as a model system for research on the molecular biology and pathogenicity of this group of bacteria.

Dickeya dadantii, formerly *Erwinia chrysanthemi* (11), is the causative agent of soft rot disease in a wide range of plants species, including many economically important crops (10). Soft-rot results from maceration of plant tissues following degradation of pectin, the major component of primary cell walls (7). *D. dadantii* is a devastating opportunistic pathogen in storage organs and fleshy tissues particularly when compromised by bruising, excess water, low oxygen levels or high temperature. *D. dadantii* is also associated with systemic infections, vascular disorders, foliar necroses, and latent infections in growing plants. We sequenced and annotated the complete genome of *Dickeya dadanti* strain 3937, a strain widely used as a model system for research on molecular biology and pathogenicity of this group of bacteria. Two whole genome shotgun libraries were prepared in plasmid pHOS2 with target insert sizes of 2-3 Kb and 10-12 Kb. We collected approximately 67,000 dual-end sequences, 67% from small insert clones and 33% from the larger insert library. Sequences were assembled into contigs using the Celera assembler (9), and this assembly was transferred to SeqMan II (LaserGene) for finishing. Primer-walking was employed to close gaps covered by clones available from the shotgun libraries. Remaining gaps were closed by sequencing PCR products generated using primers designed from the ends of assembled and ordered contigs. PCR products spanning each ribosomal RNA operon were sequenced separately to resolve sequence differences between copies. We used Glimmer 2.0 (3) for initial prediction of protein coding regions. We added, deleted and revised endpoints of genes based on comparisons to other genomes and genes and proteins in the Genpept database. tRNA sequences were identified using tRNAscan-SE (8) with additional examination to identify specific tRNAs not distinguishable by their anticodons alone. rRNA genes were identified by comparison to other enterobacterial sequences using BLASTN. Additional functional RNA genes were identified by sequence similarity to known *E. coli* K-12 RNA genes or entries in the RFAM database (5, 6). The *D. dadantii* gene products were annotated using the RAST

automated pipeline (1) and manually augmented by domain experts using ASAP (4). This rich manual annotation is not fully represented in GenBank, but is publicly available in its entirety through ASAP (https://asap.ahabs.wisc.edu/asap/home.php).

The complete circular chromosome of *D. dadantii* 3937 is 4,922,802 base pairs with 57% GC content. There are no extrachromosomal elements. In ASAP, there are 4,543 predicted or known protein-coding genes, 22 rRNA genes organized into seven operons, 75 predicted tRNAs, and 20 noncoding RNA genes. Orthologs of the *D. dadantii* proteins in other sequenced enterobacteria were identified as best-matching proteins from pairwise BLASTP searches. Not surprisingly, the largest number of putative orthologs, 2992, is found in *Pectobacterium atrosepticum*, which is both phylogenetically and phenotypically most similar to *D. dadantii* (2). Interestingly, these account for only 67% of the *P. atrosepticum* and 66% of the *D. dadantii* predicted protein-coding genes, underscoring the differences between these pectinolytic bacteria.

Nucleotide sequence accession number. The sequence has been deposited in Genbank as accession number CP002038.

AKNOWLEDGMENTS

The project was supported by the Initiative for Future Agriculture and Food Sytems Program of the USDA Cooperative State Research, Education and Extension Service, grant number # 2001-52100-11316 to N.T. Perna, F.R. Blattner and N.T. Keen. We greatly appreciate the technical expertise provided by staff at both The Institute for Genomic Research and the

116 Genome Center of Wisconsin. We would like to thank the following individuals for their
117 contributions to the annotation of the genome: Ravi D. Barabote, Val Burland, Loic
118 Grandemange, Courtney Jahn, Matuoka Keisuke, Hiroyuki Matsumoto and Holly Slater. This
119 work is dedicated to Professor Noel T. Keen.
120

References

1. **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, and O. Zagnitko.** 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics **9**:75.

2. **Bell, K. S., M. Sebaihia, L. Pritchard, M. T. Holden, L. J. Hyman, M. C. Holeva, N. R. Thomson, S. D. Bentley, L. J. Churcher, K. Mungall, R. Atkin, N. Bason, K. Brooks, T. Chillingworth, K. Clark, J. Doggett, A. Fraser, Z. Hance, H. Hauser, K. Jagels, S. Moule, H. Norbertczak, D. Ormond, C. Price, M. A. Quail, M. Sanders, D. Walker, S. Whitehead, G. P. Salmond, P. R. Birch, J. Parkhill, and I. K. Toth.** 2004. Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. *atroseptica* and characterization of virulence factors. Proc Natl Acad Sci U S A **101:**11105-10.

3. **Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res **27:**4636-41.

4. **Glasner, J. D., P. Liss, G. Plunkett, 3rd, A. Darling, T. Prasad, M. Rusch, A. Byrnes, M. Gilson, B. Biehl, F. R. Blattner, and N. T. Perna.** 2003. ASAP, a systematic annotation package for community analysis of genomes. Nucleic Acids Res **31:**147-51.

5. **Griffiths-Jones, S., A. Bateman, M. Marshall, A. Khanna, and S. R. Eddy.** 2003. Rfam: an RNA family database. Nucleic Acids Res **31:**439-41.

6. **Griffiths-Jones, S., S. Moxon, M. Marshall, A. Khanna, S. R. Eddy, and A. Bateman.** 2005. Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res **33:**D121-4.

7. **Hugouvieux-Cotte-Pattat, N., G. Condemine, W. Nasser, and S. Reverchon.** 1996. Regulation of pectinolysis in *Erwinia chrysanthemi*. Annu. Rev. Microbiol. **50:**213-257.

8. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res **25:**955-64.

9. **Myers, E. W., G. G. Sutton, A. L. Delcher, I. M. Dew, D. P. Fasulo, M. J. Flanigan, S. A. Kravitz, C. M. Mobarry, K. H. Reinert, K. A. Remington, E. L. Anson, R. A. Bolanos, H. H. Chou, C. M. Jordan, A. L. Halpern, S. Lonardi, E. M. Beasley, R. C. Brandon, L. Chen, P. J. Dunn, Z. Lai, Y. Liang, D. R. Nusskern, M. Zhan, Q. Zhang, X. Zheng, G. M. Rubin, M. D. Adams, and J. C. Venter.** 2000. A whole-genome assembly of Drosophila. Science **287:**2196-204.

10. **Perombelon, M. C. M.** 2002. Potato diseases caused by soft-rot erwinias: an overview of pathogenesis. Plant Pathol **51:**1-12.

11. **Samson R., Legendre J.B., Christen R., Fischer-Le Saux M., Achouak W., L. Gardan.** 2005. Transfer of *Pectobacterium chrysanthemi* (Burkholder et al. 1953) Brenner *et al.* 1973 and *Brenneria paradisiaca* to the genus *Dickeya* gen. nov. as *Dickeya chrysanthemi* comb. nov. and *Dickeya paradisiaca* comb. nov. and delineation of four novel species, *Dickeya dadantii* sp. nov., *Dickeya dianthicola* sp. nov., *Dickeya dieffenbachiae* sp. nov. and *Dickeya zeae* sp. nov. Int J Syst Evol Microbiol. **55:**1415-1427.