

GENOME ANNOUNCEMENTS

Genome Sequence of *Pantoea ananatis* LMG20103, the Causative Agent of *Eucalyptus* Blight and Dieback[∇]

Pieter De Maayer,^{1,2*} Wai Yin Chan,^{1,2} Stephanus N. Venter,^{1,2} Ian K. Toth,^{1,4} Paul R. J. Birch,^{1,4,5} Fourie Joubert,³ and Teresa A. Coutinho^{1,2}

Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria 0002, South Africa¹; Department of Microbiology and Plant Pathology, University of Pretoria, Pretoria 0002, South Africa²; Bioinformatics and Computational Biology Unit, Department of Biochemistry, University of Pretoria, Pretoria 0002, South Africa³; Plant Pathology Programme, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, United Kingdom⁴; and Division of Plant Science, University of Dundee (At SCRI), Invergowrie, Dundee DD2 5DA, United Kingdom⁵

Received 20 January 2010/Accepted 18 March 2010

***Pantoea ananatis* is a Gram-negative plant pathogen that causes disease on a broad range of host plants, including pineapple, maize, rice, onion, melons, and *Eucalyptus*, and has been implicated in several cases of human disease. Here, we report the genome sequence of *P. ananatis* LMG20103 isolated from diseased *Eucalyptus* in South Africa.**

Pantoea ananatis belongs within the *Enterobacteriaceae* family and the genus *Pantoea*, which hosts a number of important plant pathogens and clinically relevant species. It has been isolated from a range of plant hosts worldwide, and its ability to cause disease in several important staple food crops, including maize, onion and rice, has led to its being recognized as an important emerging plant pathogen (3). In South Africa, *P. ananatis* causes severe blight and dieback of *Eucalyptus* hybrid and clone seedlings/cuttings, which affects the pulp and paper industry (4). For this reason, we chose to sequence the genome of the highly virulent *Eucalyptus*-pathogenic strain *P. ananatis* LMG20103. Previous analyses have indicated that this strain lacks genes encoding type II and type III secretion systems (3), which represent two major pathogenicity determinants in most plant- and animal-pathogenic bacteria (7). The genome sequence may resolve how *P. ananatis* LMG20103 causes disease in the absence of these secretion mechanisms. This also constitutes the first genome of a member of the genus *Pantoea* and the first bacterial phytopathogen genome sequenced in Africa.

The genome was sequenced at Inqaba Biotech, South Africa, using a Roche 454 GS20 sequencer. A total of 991,246 reads, 80 to 120 nucleotides in length, yielding 97,152,414 nucleotides were sequenced to give a 20.9-times coverage. A draft assembly of 117 contigs was produced using Newbler Assembler 2.00.00. A scaffolding approach with 12 genomes of closely related members of the *Enterobacteriaceae*, combined with PCR gap closure, was utilized for final genome assembly. Protein-coding sequences (CDSs) were predicted using a combination of Glimmer 3.0 (5) and AMIGene (2). Annotation was per-

formed using the automated annotation engines BASys (9) and MaGe (8) and validated manually against the NCBI (1) and COG databases (6).

The genome of *P. ananatis* LMG20103 consists of a single chromosome ~4.69 Mb in size, with a G+C content of 53.69%. There are 4,272 predicted CDSs with an average length of 969 nucleotides, giving a coding intensity of 88.27%. Seventy-two tRNAs and 6 rRNA operons were identified. Putative functions have been assigned to 73% of the encoded proteins. The presence of potential homologues to genes involved in pathogenesis in other plant- and animal-pathogenic bacteria suggests that the genome sequence of *P. ananatis* LMG20103 will provide considerable insights into the pathogenesis of this organism on *Eucalyptus* and into diseases caused by its plant- and animal-pathogenic relatives within the genus *Pantoea*.

Nucleotide sequence accession number. The genome sequence was deposited in GenBank with the accession number CP001875.

This study was partially supported by the National Research Foundation (NRF), the Tree Protection Co-operative Programme (TPCP), the NRF/Dept. of Science and Technology Centre of Excellence in Tree Health Biotechnology (CTHB), and the THRIP support program of the Department of Trade and Industry, South Africa.

We acknowledge Genoscope, BaSYS, and Gary van Domselaar for their help with the annotation process and Leigh A. Riley at NCBI for assisting with the genome submission.

REFERENCES

1. Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L. Wheeler. 2008. GenBank. *Nucleic Acids Res.* **36**:D25–D30.
2. Bocs, S., S. Cruveiller, D. Vallenet, G. Nuel, and C. Medigue. 2003. AMIGene: annotation of microbial genes. *Nucleic Acids Res.* **31**:3723–3726.
3. Coutinho, T. A., and S. N. Venter. 2009. *Pantoea ananatis*: an unconventional plant pathogen. *Mol. Plant Pathol.* **10**:325–335.
4. Coutinho, T. A., O. Preisig, J. Mergaert, M. C. Cnockaert, K.-H. Riedel, J. Swings, and M. J. Wingfield. 2002. Bacterial blight and dieback of *Eucalyptus* species, hybrids, and clones in South Africa. *Plant Dis.* **86**:20–25.
5. Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg. 1999.

* Corresponding author. Mailing address: Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria 0002, South Africa. Phone: 271 24204430. Fax: 271 24203266. E-mail: Pieter.Demaayer@fab.i.up.ac.za.

[∇] Published ahead of print on 26 March 2010.

- Improved microbial gene identification with Glimmer. *Nucleic Acids Res.* **27**:4636–4641.
6. **Tatusov, R. L., E. V. Koonin, and D. J. Lipman.** 1997. A genomic perspective on protein families. *Science* **278**:631–637.
 7. **Toth, I. K., L. Pritchard, and P. R. J. Birch.** 2006. Comparative genomics reveals what makes an enterobacterial plant pathogen. *Annu. Rev. Phytopathol.* **44**:305–336.
 8. **Vallenet, D., L. Labarre, Z. Rouy, V. Barbe, S. Bocs, S. Cruveiller, A. Lajus, G. Pascal, C. Scarpelli, and C. Médigue.** 2006. MaGe: a microbial genome annotation system supported by synteny results. *Nucleic Acids Res.* **34**:53–65.
 9. **Van Domselaar, G. H., P. Stothard, S. Shrivastava, J. A. Cruz, A. Guo, X. Dong, P. Lu, D. Szafron, R. Greiner, and D. S. Wishart.** 2005. BASys: a web server for automated bacterial genome annotation. *Nucleic Acids Res.* **33**:W455–W459.