

# Complete Genome Sequence of a Beneficial Plant Root-Associated Bacterium, *Pseudomonas brassicacearum*<sup>∇</sup>

Philippe Ortet,<sup>1,2,3</sup> Mohamed Barakat,<sup>1,2,3</sup> David Lalaoua,<sup>1,2,3</sup> Sylvain Fochesato,<sup>1,2,3</sup> Valérie Barbe,<sup>4</sup> Benoit Vacherie,<sup>4</sup> Catherine Santaella,<sup>1,2,3</sup> Thierry Heulin,<sup>1,2,3</sup> and Wafa Achouak<sup>1,2,3\*</sup>

CEA, DSV, IBEB, SBVME, Lab Ecol Microb Rhizosphere & Environ Extrem (LEMIRE), Saint-Paul-lez-Durance, F-13108, France<sup>1</sup>; CNRS, UMR 6191, Saint-Paul-lez-Durance, F-13108, France<sup>2</sup>; Aix-Marseille Université, Saint-Paul-lez-Durance, F-13108, France<sup>3</sup>; and CEA, DSV, IG, Genoscope, Evry, F-91057, France<sup>4</sup>

Received 25 March 2011/Accepted 4 April 2011

**To shed light on the genetic equipment of the beneficial plant-associated bacterium *Pseudomonas brassicacearum*, we sequenced the whole genome of the strain NFM421. Its genome consists of one chromosome equipped with a repertoire of factors beneficial for plant growth. In addition, a complete type III secretion system and two complete type VI secretion systems were identified. We report here the first genome sequence of this species.**

*P. brassicacearum* has been described as the major root-associated bacteria of *Arabidopsis thaliana* and *Brassica napus* plants (1, 2, 5, 7). It has the ability to suppress plant pathogens (10) by producing antifungal compounds, such as 2,4-diacetylphloroglucinol and cyanide.

To gain more knowledge on the genetic equipment of this highly competitive bacterium in the rhizosphere, we sequenced the genome of the laboratory strain of *P. brassicacearum* NFM421, using a two-step DNA shotgun sequencing strategy that comprises Sanger and pyrosequencing-based approaches, conducted at the Genoscope (8).

The circular chromosome of *P. brassicacearum* NFM421 has 6,843,248 bp and an average G+C content of 60.8%. The genome was submitted to the Genobrowser annotation platform for automated annotation and gene prediction using the annotation tools GeneMark (4), Glimmer (6), and MedP (11). The whole genome was reannotated manually. The chromosome contains 6,097 protein-coding sequences with an average length of 989 bp, 65 tRNAs, and 16 rRNAs. A total of 4,686 proteins (76.9%) were assigned a biological function, while 324 orphan proteins were identified. Protein-coding genes represent 88.2% of the genome. Regulatory genes, including those encoding two-component signal transduction proteins (TCSs) and transcription factors (TFs), were analyzed using P2CS (3) and P2TF (<http://www.p2tf.org>) systems, which identified 178 TCSs and 501 TFs. The genome analysis revealed clusters of hitherto-unknown secondary metabolites, which may contribute to bacterial fitness and competitiveness in the rhizosphere. In addition, the *P. brassicacearum* genome contains a complete type III secretion system and two complete type VI secretion systems.

The large size of *P. brassicacearum* genome is attributable mainly to lateral gene transfer, as acquired genes represent up to 12% of the whole genome. From acquired genes, three prophage regions, at least one confirmed ICE (integrative con-

jugative element) (9), and many genes relevant to plant-bacterium interaction and survival under saprophytic conditions were identified.

A further and deep exploration of the *P. brassicacearum* genome will provide new insights on this bacterium's lifestyle and the molecular mechanism underlying the molecular dialogue with other microorganisms and with plants.

**Nucleotide sequence accession number.** The complete genome sequence of *P. brassicacearum* NFM421 is available in GenBank under accession number CP002585.

This work was supported by ECCOREV and the GENOSCOPE.

## REFERENCES

1. Achouak, W., et al. 2000. Description of *Pseudomonas brassicacearum* sp. nov. and *Pseudomonas thivervalensis* sp. nov., two root-associated bacteria isolated from *Arabidopsis thaliana* and *Brassica napus*. *Int. J. Syst. Evol. Microbiol.* **50**:9–18.
2. Achouak, W., S. Conrod, V. Cohen, and T. T. Heulin. 2004. Phenotypic variation of *Pseudomonas brassicacearum* as a plant root colonization strategy. *Mol. Plant Microbe Interact.* **17**:872–879.
3. Barakat, M., P. Ortet, and D. E. Whitworth. 2011. P2CS: a database of prokaryotic two-component systems. *Nucleic Acids Res.* **39**:D771–D776.
4. Besemer, J., A. Lomsadze, and M. Borodovsky. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* **29**:2607–2618.
5. Chabeaud, P., et al. 2001. Phase-variable expression of an operon encoding extracellular alkaline protease, serine protease homologue and lipase in *Pseudomonas brassicacearum*. *J. Bacteriol.* **183**:2117–2120.
6. Delcher, A. L., K. A. Bratke, E. C. Powers, and S. L. Salzberg. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **23**:673–679.
7. Fromin, N., W. Achouak, J.-M. Thiery, and T. Heulin. 2001. The genotypic diversity of *Pseudomonas brassicacearum* populations isolated from roots of *Arabidopsis thaliana*: influence of plant genotype. *FEMS Microbiol. Ecol.* **37**:21–29.
8. Goldberg, S. M. D., et al. 2006. A Sanger/pyrosequencing hybrid approach for the generation of high-quality draft assemblies of marine microbial genomes. *Proc. Nat. Acad. Sci. U. S. A.* **103**:11240–11245.
9. Mavrodi, D. V., J. E. Loper, I. T. Paulsen, and L. S. Thomashow. 2009. Mobile genetic elements in the genome of the beneficial rhizobacterium *Pseudomonas fluorescens* Pf-5. *BMC Microbiol.* **9**:8.
10. Ross, I., Y. Alami, P. R. Harvey, W. Achouak, and M. Ryder. 2000. Genetic diversity and biological control activity of novel species related to *Pseudomonas* isolated from wheat field soils in South Australia. *Appl. Environ. Microbiol.* **66**:1609–1616.
11. Zhu, H. Q., G. Q. Hu, Z. Q. Ouyang, J. Wang, and Z. S. She. 2004. Accuracy improvement for identifying translation initiation sites in microbial genomes. *Bioinformatics* **20**:3308–3317.

\* Corresponding author. Mailing address: CEA, DSV, IBEB, SBVME, Lab Ecol Microb Rhizosphere & Environ Extrem (LEMIRE), Saint-Paul-lez-Durance, F-13108, France. Phone: (33) 442254961. Fax: (33) 442256648. E-mail: wafa.achouak@cea.fr.

<sup>∇</sup> Published ahead of print on 22 April 2011.