

# Draft Genome Sequence of the Purple Photosynthetic Bacterium *Phaeospirillum molischianum* DSM120, a Particularly Versatile Bacterium

K. Duquesne,<sup>a</sup> V. Prima,<sup>a</sup> B. Ji,<sup>b</sup> Z. Rouy,<sup>c</sup> C. Médigue,<sup>c</sup> E. Talla,<sup>b</sup> and J. N. Sturgis<sup>a</sup>

LISM, UMR7255 CNRS, Aix-Marseille University, Marseilles, France<sup>a</sup>; LCB, UMR7283 CNRS, Aix-Marseille University, Marseilles, France<sup>b</sup>; and Genoscope/CEA, Centre National de Séquençage, Evry, France<sup>c</sup>

**Here we present the draft genome sequence of the versatile and adaptable purple photosynthetic bacterium *Phaeospirillum molischianum* DSM120. This study advances the understanding of the adaptability of this bacterium, as well as the differences between the *Phaeospirillum* and *Rhodospirillum* genera.**

The alphaproteobacterium *Phaeospirillum molischianum*, previously named *Rhodospirillum molischianum* (3), is a photosynthetic bacterium which has been extensively studied for its photosynthetic apparatus (1, 2, 4, 6), in particular since it is relatively easy to cultivate and grows rapidly to high cell density. This easy cultivation and widespread occurrence suggest efficient adaptation to the environment, as is indeed evidenced by its capacity for chromatic adaptation. This led us to suspect an extensive collection of environmental sensors.

Whole-genome shotgun sequencing of *P. molischianum* was performed by COGENICS (San Francisco, CA) using two runs of 454 technology, producing 590,359 reads (92.2 Mb of raw sequences representing a coverage of 24-fold). Assembly yielded 291 contigs, 61 of which (sizes of 522 to 416,194 bp) were organized into 14 scaffolds, leading to a draft genome sequence of 3,805,617 bp with a GC content of 61.5%. Pulsed-field electrophoresis allowed us to confirm the genome size and showed that there were no plasmids. Automatic annotation using the MaGe pipeline system (8) predicted 3,862 coding sequences (CDS), 10 fragments of CDS, and 63 genes coding for RNA (6 rRNA genes, 49 tRNA genes, and 8 other miscellaneous RNA genes).

The genome analysis revealed 340 proteins involved in signal transduction mechanisms, including ~60 histidine kinase-type sensors and 65 response regulators (from LuxR, Fis, CheY, and OmpR families). This very high number of two-component systems may explain the adaptability of *P. molischianum* to environment changes. The photosynthetic apparatus was also more complex than expected. The genome contains six pairs of genes encoding the light-harvesting 2 (LH2) antenna, and their expression is regulated by light intensity (6) and probably other parameters detected by the wealth of sensors. Compared to *Rhodocista centenaria*, genes for bacteriochlorophyll (*bch*) and carotenoid (*crt*) biosynthesis, as well as reaction centers and core light-harvesting complexes (*puf* and *puh*) (5), are organized into two photosynthetic gene clusters (*pufCMLAB-bchZYXC-crtEC* and *crtBI-bchP2GFNBHLM-ihA-puhA*) in *P. molischianum*, rather than a single one, as found elsewhere. More surprisingly, *P. molischianum* contains structural gene clusters for two different types of nitrogenases, a Mo-Fe-dependent nitrogenase (Nif cluster) which is also found in *Rhodocista centenaria* (5) and an alternative Fe-Fe nitrogenase (Anf cluster). Only form II of

RubisCO (CbbM), one of the key enzymes in the Calvin cycle of CO<sub>2</sub> fixation, is present.

Finally, genomic comparison of the chromosome gene organization showed little synteny between *P. molischianum* and *Rhodospirillum* strains. This suggests a high genome plasticity of this organism; this could be caused by the large number of repetitive sequences (5.89% of the overall genome) or the high number of transposases (about 81).

In conclusion, we present here the draft genome of *P. molischianum*, which reveals a very adaptable organism with a multitude of environmental sensors and a very plastic genome predisposed to rapid evolution and remodeling.

**Nucleotide sequence accession numbers.** This genome has been deposited in the DDBJ/EMBL/GenBank databases (accession no. CAHP01000001 to CAHP01000061). This first version contains 61 of the original 291 contigs. MaGe annotation data will be made publicly available through the MaGe PhaeoScope website (<http://www.genoscope.cns.fr/agc/mage/phaeoscope>).

## ACKNOWLEDGMENTS

This work was supported by the CNRS, Aix-Marseille University, and ANR-BLAN07-1-183779. B.J. is a Ph.D. student with a China Scholarship Council fellowship.

## REFERENCES

1. Germeroth L, Lottspeich F, Robert B, Michel H. 1993. Unexpected similarities of the B800-850 light-harvesting complex from *Rhodospirillum molischianum* to the B870 light-harvesting complexes from other purple photosynthetic bacteria. *Biochemistry* 32:5615–5621.
2. Gonçalves RP, Bernadac A, Sturgis JN, Scheuring S. 2005. Architecture of the native photosynthetic apparatus of *Phaeospirillum molischianum*. *J. Struct. Biol.* 152:221–228.
3. Imhoff JF, Petri R, Suling J. 1998. Reclassification of species of the spiral-shaped phototrophic purple non-sulfur bacteria of the  $\alpha$ -Proteobacteria: description of the new genera *Phaeospirillum* gen. nov., *Rhodovibrio* gen. nov., *Rhodothalassium* gen. nov. and *Roseospira* gen. nov. as well as transfer of *Rhodospirillum fulvum* to *Phaeospirillum fulvum* comb. nov., of *Rhodospirillum molischianum* to *Phaeospirillum molischianum* comb. nov., of

Received 10 April 2012 Accepted 11 April 2012

Address correspondence to K. Duquesne, [k.duquesne@gmail.com](mailto:k.duquesne@gmail.com).

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JB.00605-12

- Rhodospirillum salinarum* to *Rhodovibrio salexigens*. *Int. J. Syst. Bacteriol.* **48**:793–798.
4. Koepke J, Hu X, Muenke C, Schulten K, Michel H. 1996. The crystal structure of the light-harvesting complex II (B800-850) from *Rhodospirillum molischianum*. *Structure* **4**:581–597.
  5. Lu Y-K, et al. 2010. Metabolic flexibility revealed in the genome of the cyst-forming  $\alpha$ -1 proteobacterium *Rhodospirillum centenum*. *BMC Genomics* **11**:325.
  6. Mascle-Allemand C, Duquesne K, Lebrun R, Scheuring S, Sturgis JN. 2010. Antenna mixing in photosynthetic membranes from *Phaeospirillum molischianum*. *Proc. Natl. Acad. Sci. U. S. A.* **107**:5357–5362.
  7. Reference deleted.
  8. Vallenet D, et al. 2009. MicroScope: a platform for microbial genome annotation and comparative genomics. *Database (Oxford)* **2009**:bap021. doi:10.1093/database/bap021.