

## Genome Sequence of the Edible Cyanobacterium *Arthrospira* sp. PCC 8005<sup>∇</sup>

P. J. Janssen,<sup>1</sup>\* N. Morin,<sup>1</sup># M. Mergeay,<sup>1</sup> B. Leroy,<sup>2</sup> R. Wattiez,<sup>2</sup> T. Vallaey,<sup>3</sup> K. Waleron,<sup>4</sup> M. Waleron,<sup>4</sup> A. Wilmotte,<sup>5</sup> P. Quillardet,<sup>6</sup> N. Tandeu de Marsac,<sup>6</sup> E. Talla,<sup>7</sup> C.-C. Zhang,<sup>7</sup> and N. Leys<sup>1</sup>

Molecular and Cellular Biology—Unit of Microbiology, Institute for Environment, Health and Safety, Belgian Nuclear Research Centre SCK•CEN, B-2400 Mol, Belgium<sup>1</sup>; Department of Proteomics and Microbiology, University of Mons, B-7000 Mons, Belgium<sup>2</sup>; Laboratoire ECOLAG, UMR CNRS-UMII 5119, Université Montpellier II, cc 093, Place E. Bataillon, F-34095 Montpellier Cedex 5, France<sup>3</sup>; Intercollegiate Faculty of Biotechnology, University of Gdańsk and Medical University of Gdańsk, Kladki 24, Gdańsk, Poland<sup>4</sup>; Centre for Protein Engineering, Institute of Chemistry B6, Sart Tilman, University of Liège, B-4000 Liège, Belgium<sup>5</sup>; Unité des Cyanobactéries, CNRS-URA2172, Institut Pasteur, F-75015, Paris, France<sup>6</sup>; and Laboratoire de Chimie Bactérienne, CNRS-UPR9043, Aix-Marseille Université, 31 chemin Joseph Aiguier, 13402 Marseille Cedex 20, France<sup>7</sup>

Received 2 February 2010/Accepted 24 February 2010

**We determined the genome sequence of *Arthrospira* sp. PCC 8005, a cyanobacterial strain of great interest to the European Space Agency for its nutritive value and oxygenic properties in the Micro-Ecological Life Support System Alternative (MELiSSA) biological life support system for long-term manned missions into space.**

*Arthrospira* is a genus of nonheterocystous filamentous cyanobacteria that typically reside in alkaline lakes (9). Although they clearly form a separate taxonomic unit (4), various *Arthrospira* species were originally included in the genus *Spirulina*, and some are still being commercialized under that name. *Arthrospira* species have a high protein content, are rich in essential fatty acids, and produce a variety of minerals, vitamins, and nutritional pigments (3). Consequently, they have a long history of human consumption and are used worldwide as feedstock in agriculture and aquaculture (5). Besides their nutritional value, *Arthrospira* species have clinical relevance as an antioxidant (7) and are used in biohydrogen technology (1). The PCC 8005 strain was selected by the European Space Agency (ESA) as an oxygen producer and as a nutritional end product of the Micro-Ecological Life Support System Alternative (MELiSSA) life support system (6).

Whole-genome shotgun sequencing of strain PCC 8005 was performed using 454 pyrosequencing technology (amounting to 400,000 reads) and Sanger sequencing (up to 96,000 longer reads), leading to a final assembly of 119 contigs into 16 scaffolds representing 6,279,260 bases with an average G+C content of 44.7%. These scaffolds were processed by the MaGe annotation platform (8) and predicted 5,856 protein-coding sequences (CDSs) and 176 genes encoding RNA.

The genome of *Arthrospira* sp. PCC 8005 displays the highest overall synteny with the genomes of *Arthrospira maxima* CS-328, with 4,618 proteins as bidirectional best hits (BDBHs),

and *Lyngbya* PCC8106, with 3,083 BDBHs. The MaGe annotation system assigned 63% of the CDSs to one or more functional COGs (clusters of orthologous groups) and reported 1,704 conserved hypothetical and 884 hypothetical proteins. Functional annotation by MaGe showed that many genes belonging to the same pathway are dispersed over the genome.

The PCC 8005 genome is highly repetitive in nature, with more than 300 kb present as tandem sequences, and contains four clustered, regularly interspaced short palindromic repeats (CRISPRs), which may provide a cellular defense against phages and plasmids (2). The genome also contains at least 140 complete insertion elements (ISs) belonging to various families and eight copies of a putative genomic island.

Several genes encoding desaturase were identified, including those required for the production of  $\beta$ -carotene and the two essential fatty acids linoleic acid (LA-18:2<sup>Δ9,Δ12</sup>) and  $\gamma$ -linolenic acid (GLA-18:3<sup>Δ6,Δ9,Δ12</sup>). Genome data confirm the inability of PCC 8005 to fix nitrogen, as it lacks essential *nif* genes, and indicate that nitrogen metabolism in PCC 8005 follows classic routes utilizing nitrate, nitrite, urea, and ammonium, with NtcA as the global nitrogen regulator. An intact *nthPAB* locus encoding a nitrile hydratase and its regulator would allow strain PCC 8005 to utilize nitriles (R—C≡N) as the sole source of nitrogen. Hydrogen production in strain PCC 8005 is governed by a soluble bidirectional hydrogenase encoded and activated by intact *hox* and *hyp* loci. No *hup* genes for a separate uptake hydrogenase were detected.

The PCC 8005 genome sequence and its curated annotation are important assets to better understand the physiology and metabolic potential of *Arthrospira* and will open up new opportunities in the functional genomics of this species.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project was deposited at DDBJ/EMBL/GenBank under accession no. ADDH00000000. The version described here

\* Corresponding author. Mailing address: Molecular and Cellular Biology—Unit of Microbiology, Institute for Environment, Health and Safety, Belgian Nuclear Research Centre SCK•CEN, Boeretang 200, B-2400 Mol, Belgium. Phone: 32-14-33-2129. Fax: 32-14-31-4793. E-mail: pjanssen@sckcen.be.

# Joint first authors.

<sup>∇</sup> Published ahead of print on 16 March 2010.

is the first version, ADDH01000000. MaGe annotation data will become publicly available from the ArthroScope project site within 6 months of the GenBank submission date.

We thank C. Médigue and D. Vallenet of Génoscope (Centre National de Séquençage, Evry, France) for essential advice and training in genome annotation using the MaGe platform.

This work was supported by the European Space Agency ESA/ESTEC under MELGEN-2 contract no. 20397/06/NL/SFe in the frame of the MELiSSA project. A.W. is a research associate of the FRS-FNRS.

This Genome Announcement is dedicated to the lasting memory of our colleague Larissa Hendrickx (deceased 3 March 2008), who initiated the PCC 8005 genome project as part of MELiSSA.

#### REFERENCES

1. Ananyev, G., D. Carrieri, and G. C. Dismukes. 2008. Optimization of metabolic capacity and flux through environmental cues to maximize hydrogen production by the cyanobacterium "*Arthrospira* (*Spirulina*) *maxima*". *Appl. Environ. Microbiol.* **74**:6102–6113.
2. Barrangou, R., and P. Horvath. 2009. The CRISPR system protects microbes against phages, plasmids. *Microbe* **4**:224–230.
3. Belay, A. 2002. The potential application of *Spirulina* (*Arthrospira*) as a nutritional and therapeutic supplement in health management. *J. Am. Nutra. Assoc.* **5**:27–48.
4. Castenholz, R. W., R. Rippka, M. Herdmann, and A. Wilmotte. 2001. Form-genus I. *Arthrospira* Stizenberger 1852, p. 542–543. In D. R. Boone, R. W. Castenholz, and G. M. Garrity (ed.), *Bergey's manual of systematic bacteriology*, vol. 1. Springer, New York, NY.
5. Habib, M. A. B., M. Pariv, T. C. Huntington, and M. R. Hasan. 2008. A review on culture, production and use of spirulina as food for humans and feeds for domestic animals and fish. *FAO Fisheries and Aquaculture Circular* no. 1034. Food and Agriculture Organisation (FAO) of the United Nations, Rome, Italy.
6. Hendrickx, L., H. De Wever, V. Hermans, F. Mastroleo, N. Morin, A. Wilmotte, P. Janssen, and M. Mergeay. 2006. Microbial ecology of the closed artificial ecosystem MELiSSA (Micro-Ecological Life Support System Alternative): reinventing and compartmentalizing the Earth's food and oxygen regeneration system for long-haul space exploration missions. *Res. Microbiol.* **157**:77–86.
7. Karkos, P. D., S. C. Leong, C. D. Karkos, N. Sivaji, and D. A. Assimakopoulos. 2008. *Spirulina* in clinical practice: evidence-based human applications. *Evid. Based Complement. Alternat. Med.*
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. Vonshak, A. 1997. *Spirulina*: growth, physiology and biochemistry, p. 43–66. In A. Vonshak (ed.), *Spirulina platensis* (*Arthrospira*): physiology, cell-biology and biotechnology. Taylor & Francis, London, United Kingdom.