Draft genome sequence of *Arthrospira platensis* C1 (PCC9438)

Supapon Cheevadhanarak1,2,3* Kalyanee Paithoonrangsrarid1, Peerada Prommeenate1, Warunee Kaewngam1, Apiluck Musigkain3, Somvong Tragoonrung2, Satoshi Tabata7, Takakazu Kaneko7, Jeerayut Chaijaruwanich6, Duangjai Sangsrakru2, Sithichoke Tangphatsornruang5, Juntima Chanprasert6, Sissades Tongsima3, Kanthida Kusommano3, Wattana Jeamton2, Sudarat Dulsawat1, Amornpan Klanchui5, Tayvich Vorapreeda1, Vasunun Chumchua3, Chiraphan Khannapho1, Chinea Thammarongtham1, Vethachai Plengvidhya5, Sanjukta Subudhi5, Apiradee Hongsthong1, Marasri Ruengjitchatchawalya2,3, Asawin Meechai4, Jittisak Senachak1, and Morakot Tanticharoen5

1 Biochemical Engineering and Pilot Plant Research and Development Unit, National Center for Genetic Engineering and Biotechnology at King Mongkut’s University of Technology Thonburi, Bangkhuntien, Bangkok 10150, Thailand
2 School of Bioresources and Technology, King Mongkut’s University of Technology Thonburi, Bangkhuntien, Bangkok 10150, Thailand
3 Pilot Plant Development and Training Institute, King Mongkut’s University of Technology Thonburi, Bangkhuntien, Bangkok 10150, Thailand
4 Department of Chemical Engineering, King Mongkut’s University of Technology Thonburi, Bangkok 10140, Thailand
5 National Center for Genetic Engineering and Biotechnology, Pathumthani 12120, Thailand
6 Department of Computer Science, Faculty of Science, Chiangmai University, Chiangmai, Thailand
7 Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Japan

* Corresponding author: Supapon Cheevadhanarak, supapon.che@kmutt.ac.th

*Arthrospira platensis* is a cyanobacterium that is extensively cultivated outdoors on a large commercial scale for consumption as a food for humans and animals. It can be grown in moniculture under highly alkaline conditions, making it attractive for industrial production. Here we describe the complete genome sequence of *A. platensis* C1 strain and its annotation. The *A. platensis* C1 genome contains 6,089,210 bp including 6,108 protein-coding genes and 45 RNA genes, and no plasmids. The genome information has been used for further comparative analysis, particularly of metabolic pathways, photosynthetic efficiency and barriers to gene transfer.

**Introduction**

*Arthrospira platensis* is a cyanobacterium that contains large amounts of proteins, vitamins, lipids and pigments [1]. It is widely used as a human food and an animal feed. In addition, its extracts can enhance the immune system and promote health [1,2]. As the natural habitat is soda lakes, *Arthrospira* spp are cultivated under highly alkaline conditions in open ponds on a large commercial scale. This condition can minimize and sometimes prevent the culture from contamination [3]. Unlike many plant food products whose nutritional value rapidly deteriorates at high temperatures, the nutritional value of *Arthrospira* products is maintained even when the cells are processed at high temperatures [4]. In contrast to many cyanobacteria, there is no report of toxicity of *Arthrospira* for humans, animals or environments [4].

The genome sequences of *Arthrospira* spp. have been the subject of immense interest due to the beneficial properties of these organisms in the biotechnology and environmental fields [5,6]. *A. platensis* C1 is the fifth complete genome report for a member of the genus *Arthrospira*. *A. platensis* C1 has long been used as a laboratory strain for physiological and molecular studies due to its non-gliding property, which enables single colonies formation. This property facilitates studies at the molecular level and strain improvement, particularly, the development of a transformation system. Currently, a successful transformation system for *Arthrospira* has not yet been established.
Thus the genome sequences may help to identify barriers responsible for the instability of the transformants. Here, we present a summary classification and a set of features of *A. platensis* C1 together with the complete genomic sequence and its annotation.

**Classification and features**

Historically, the classification of the *Arthrospira* and *Spirulina* genera [Figure 1] was a subject of controversy. For the commercial strain, *Arthrospira* or *Spirulina* was used interchangeably. Both *Arthrospira* and *Spirulina* are similar in morphological characters; cylindrical, multicellular, filamentous cyanobacteria with an open, left-handed helical shape [Table 1]. They both belong to the Phylum *Cyanobacteria*, Order *Oscillatoriales* and Family *Oscillatoriaceae* [13]. However, they can be differentiated by the presence of cell septa: *Arthrospira* possess septa, whereas *Spirulina* do not [14].

| MIGS ID | Property                        | Term                        | Evidence Code |
|---------|---------------------------------|-----------------------------|---------------|
|         | Domain                          | *Bacteria*                  | TAS [9]       |
|         | Phylum                          | *Cyanobacteria*             | TAS [10,11]   |
|         | Class                           | *Cyanobacteria*             | TAS [11,12]   |
|         | Current Classification           | Order *Oscillatoriales*     | TAS [13]      |
|         |                                 | Family *Oscillatoriaceae*   | TAS [13]      |
|         |                                 | Genus *Arthrospira*         | TAS [11]      |
|         |                                 | Species *Arthrospira platensis* C1 | TAS [11]  |
|         | Gram stain                      | Negative                    | TAS [14]      |
|         | Cell shape                      | Spiral                      | TAS [14]      |
|         | Motility                        | None                        |               |
|         | Sporulation                     | None                        |               |
|         | Temperature range                | 20 – 40                     | TAS [14]      |
|         | Optimum temperature             | 30-35                       | TAS [14]      |
| MIGS-22 | Relationship to Oxygen          | Aerobic                     | TAS [3]       |
|         | Carbon source                   | Phototroph, Mixotroph       | TAS [3]       |
|         | Energy source                   | Phototroph                  | TAS [3]       |
| MIGS-6  | Habitat (EnvO)                  | Fresh water                 | TAS [3]       |
| MIGS-6.1| Temperature                     | 20 - 40                     | TAS [3]       |
| MIGS-6.2| pH                             | 8.0 -10.0                   | TAS [3]       |
| MIGS-6.3| Salinity                       | 0.06                        | TAS [3]       |
| MIGS-10 | Extrachromosomal elements       | None                        | TAS [6]       |
| MIGS-11 | Estimated Size                  | 6.08 Mb                     | IDA           |
| MIGS-14 | Known Pathogenicity             | None                        | NAS           |
| MIGS-15 | Biotic Relationship             | Free living                 | NAS           |
| MIGS-4  | Geographic Location             | Not reported                | NAS           |
| MIGS-4.1| Latitude                        | Not reported                | NAS           |
| MIGS-4.2| Longitude                       | Not reported                | NAS           |
| MIGS-4.3| Depth                          | Not reported                | NAS           |
| MIGS-4.4| Altitude                       | Not reported                | NAS           |

Evidence codes - IDA: Inferred from Direct Assay (first time in publication); TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [15]. If the evidence code is IDA, then the property was directly observed for a living isolate by one of the authors or an expert mentioned in the acknowledgements.
Figure 1 The phylogenetic tree of 51 cyanobacterial concatenated ribosomal proteins. The main topology is in agreement with earlier inferences of the phylogeny of this taxon with the 16s rRNA based on the GTR+G+I substitution model [7]. The tree is built using the Neighbor-Joining method and 1,000 re-samplings to calculate bootstrap values. *A. platensis* C1 was clustered together with other strains in the order *Oscillatoriales* and was clearly separated from related species in the order *Nostocales*. The conserved, concatenated ribosomal protein phylogenetic tree indicated the monophyly of this *Arthospira* genus.
**Chemotaxonomy**

*Arthrospira platensis* C1 or *Arthrospira* sp. PCC 9438, as designated based on its morphology by Pasteur Institute, Paris, France, was originally classified as *Spirulina platensis* C1 [1,3]. This reclassification was in agreement with the presence of γ-linolenic acid (GLA) in the fatty acid profile, a chemotaxonomic marker of *Arthrospira*, while GLA is absence in *Spirulina* [16-18].

The phylogenetic tree of cyanobacteria was reconstructed with evolutionary information embedded in conserved, concatenated ribosomal proteins. *A. platensis* C1 was positioned into the genus *Arthrospira* with 100% of the bootstrapping value. The closest strain is *Arthrospira sp.* PCC 8005 with 97.43% sequence identity, whereas the other strains of this genus share 94.93-96.58% sequence identity. The nearest related order is *Nostocales* with approximately 70% sequence identity.

**Genome sequencing and annotation**

**Genome project history**

*Arthrospira* genome sequencing projects have been carried out in 5 research centers: Genoscope, France; the DOE Joint Genome Institute, USA; the National Institute of Technology and Evaluation (NITE), Japan; the University of Applied Sciences, Switzerland, and King Mongkut’s University of Technology Thonburi (KMUTT), Thailand, using various *Arthrospira* strains (*A. platensis* PCC8005, *A. maxima* CS-328, *A. platensis* NIES-39, *A. platensis* Paraca and *A. platensis* C1 (this study), respectively). In this study, the genome of *A. platensis* C1 has been sequenced, and the results provide data that can be used for the further study of its biological functions.

The genome project is deposited in the GenBank Database (NCBI ID 67617 and accession number AFXD00000000). DNA sequencing and finishing were performed in collaboration between the Genome Institute, BIOTEC-NSTDA, Thailand and Kazusa DNA Research Institute, Japan. The genome assembly and annotation steps were performed in collaboration between KMUTT, the Genome Institute, BIOTEC-NSTDA, Chiang Mai University, Thailand and Kazusa DNA Research Institute. The summary of the project information is shown in Table 2.

| MIGS-ID | Property                     | Term                                                                 |
|---------|------------------------------|----------------------------------------------------------------------|
| MIGS-31 | Finishing quality            | Draft                                                                |
| MIGS-28 | Libraries used               | Genomic libraries: Sanger (one each of 2 and 5 kb library), standard |
|         |                              | 454 pyrosequence library and BAC library                           |
| MIGS-29 | Sequencing platforms         | 454 and Sanger Technology                                            |
| MIGS-30 | Assemblers                   | Newbler version 2.3, Phrap                                              |
| MIGS-31.2 | Sequence coverage            | 28×                                                                 |
| MIGS-31.3 | Contigs                     | 63                                                                  |
| MIGS-32 | Gene calling method          | Glimmer 3.0                                                           |
| GOLD ID |                              | Gi09635                                                              |
| NCBI project ID |                      | 67617                                                               |
| Database: IMG-GEBA |                | 2507262036                                                         |
| Source material identifier |                   | PCC9438                                                              |
| Project relevance |                  | Biotechnological                                                      |

Table 2 Genome sequencing project information
Growth conditions and DNA isolation

A. platensis strain C1 was obtained from Prof. Avigad Vonshak (Algal Biotechnology, Ben-Gurion University of the Negev, Israel). The cells were routinely grown at 35 °C with constant shaking at 150 rpm in Zarrouk’s medium [19], under aeration of 1% (v/v) CO₂ and illumination of 100 μE m⁻² s⁻¹ from fluorescent lamps. For stock cultures, cells were transferred into new Zarrouk’s medium [19] every two weeks.

For the shotgun genome library preparation, a total DNA sample was sheared by nebulization with nitrogen at 45 psi for 1 min and purified using a QIAGEN Purification kit (QIAGEN, Valencia, CA). DNA fragments with sizes greater than 300 bp were selected by Ampure Bead purification (Beckman Coulter, Brea, CA) and blunt-ended with T4 DNA polymerase and T4 polynucleotide kinase (454 Life Sciences, Roche, Branford, CT). After DNA purification, 454-library adaptors were added to both ends of the fragments, which were immobilized using Library Immobilization Beads for further DNA sequencing.

For paired-end library preparation, a total DNA sample was sheared by nebulization with nitrogen at 10 psi for 2 min to obtain 2.2 kb fragments with hairpin adaptors added to both ends of the DNA fragments. The DNA fragments with hairpin adaptors were immobilized using Library Immobilization Beads and ligated to paired-end adaptors. DNA libraries were linked to the capture beads and amplified by emulsion PCR (Roche Applied Science protocol).

Genome sequencing and assembly

The genome of A. platensis C1 was sequenced using a hybrid method between the 454 Life Sciences technology on the Genome Sequencer (GS) FLX System and BigDye Terminator v3.1 Cycle sequencing. Pyrosequencing reads were assembled using the Newbler de novo sequence assembly software version 2.0.0 (Roche). The Phred/Phrap/Consed software package [20] was used for sequence assembly and quality assessment in the finishing process. The remaining gaps between contigs were closed by custom primer walk or PCR amplification and then editing in Consed. The final assembly contains 739,684 reads from pyrosequencing and 45,959 reads from Sanger sequencing, resulting in 28× coverage of the genome. Employing A. maxima CS-328 contigs [21] as a reference, the A. platensis C1 circular genome of 6.08 Mb total size with 1 scaffold and 63 gaps has been constructed.

Table 3  Genome statistics

| Attribute                              | Value     | % of Total |
|----------------------------------------|-----------|------------|
| Genome size (bp)                       | 6,089,210 | 100.00%    |
| DNA coding region (bp)                 | 4,951,337 | 81.31%     |
| G+C content (bp)                       | 2,651,568 | 44.68%     |
| DNA scaffolds                          | 1         | 100.00%    |
| Total genes                            | 6,153     | 100.00%    |
| Protein coding genes                   | 6,108     | 100.00%    |
| RNA genes                              | 45        | 0.73%      |
| Genes with protein function prediction | 3,757     | 61.06%     |
| Genes with enzymes                     | 952       | 15.47%     |
| Genes with Transporter Classification  | 345       | 5.61%      |
| Genes assigns to KEGG pathways         | 1,012     | 16.45%     |
| Genes assigned to KEGG Orthology (KO)  | 1,837     | 29.86%     |
| Genes assigned to COGs                 | 3,459     | 56.22%     |
| Genes assigned to Pfam                 | 3,529     | 57.35%     |
| Genes assigned to TIGRfam              | 1,180     | 19.18%     |
| Genes assigned to InterPro             | 4,244     | 68.97%     |
| Genes assigned in paralog clusters     | 1,048     | 17.03%     |
| Genes assigned with signal peptides    | 570       | 9.26%      |
| Genes assigned with transmembrane proteins | 1,094 | 17.78%  

http://standardsingenomics.org
**Arthospira platensis C1 (PCC9438)**

**Genome annotation**
In agreement with the result from the Integrated Microbial Genomes Expert Review (IMG-ER) platform [22], all the genes in the *A. platensis* C1 genome were identified using the GLIMMER 3.0 program in our Microbial Inhouse Annotation Pipeline. Initial criteria for automated functional assignment required a minimum of 50% residue identity with over 80% length match for BLASTP alignments to the NCBI nonredundant database, InterPro, SwissProt, SignalP, COG, and KEGG databases. The tRNAscan-SE tool [23] was used to find tRNA genes, whereas ribosomal RNAs were found by using the tool RNAmmer [24]. Additional gene prediction analysis and functional annotation were performed within IMG-ER platform [22] and a round of manual curation, including confirmation with proteomic data [25,26].

**Genome properties**
The genome of *A. platensis* C1, with a total of 6.08 Mbp (6,089,210 base pairs), contains 44.68% G+C (Table 3) and, in agreement with the findings of Fujisawa et al. (2010) [6], no plasmid DNA. Our results confirm the presence of a single genome in *A. platensis* C1. The *A. platensis* C1 circular genome of 6.08 Mbp was compared with the *A. maxima* CS-328 and *A. platensis* NIES-39 genomes [6]. A total of 6,153 open reading frames (orfs) were predicted. Of these, 3,757 were annotated as coding for known protein functions and 45 for RNA genes (6 for rRNA and 39 for tRNA). The distribution of genes into COGs is presented in Table 4.

### Table 4 Number of genes associated with the general COG functional categories

| Code | Value | % age | Description                                                                 |
|------|-------|-------|-----------------------------------------------------------------------------|
| J    | 161   | 4.22  | Translation, ribosomal structure and biogenesis                             |
| A    | 0     | -     | RNA processing and modification                                              |
| K    | 164   | 4.29  | Transcription                                                               |
| L    | 459   | 12.02 | Replication, recombination and repair                                        |
| B    | 2     | 0.05  | Chromatin structure and dynamics                                            |
| D    | 49    | 1.28  | Cell cycle control, cell division, chromosome partitioning                   |
| Y    | 0     | -     | Nuclear structure                                                            |
| V    | 98    | 2.57  | Defense mechanisms                                                           |
| T    | 337   | 8.82  | Signal transduction mechanisms                                               |
| M    | 234   | 6.13  | Cell wall/membrane/envelope biogenesis                                      |
| N    | 54    | 1.41  | Cell motility                                                                |
| Z    | 0     | -     | Cytoskeleton                                                                 |
| W    | 0     | -     | Extracellular structures                                                     |
| U    | 71    | 1.86  | Intracellular trafficking, secretion, and vesicular transport                |
| O    | 180   | 4.71  | Posttranslational modification, protein turnover, chaperones                 |
| C    | 181   | 4.74  | Energy production and conversion                                             |
| G    | 149   | 3.9   | Carbohydrate transport and metabolism                                        |
| E    | 201   | 5.26  | Amino acid transport and metabolism                                          |
| F    | 66    | 1.73  | Nucleotide transport and metabolism                                          |
| H    | 157   | 4.11  | Coenzyme transport and metabolism                                           |
| I    | 69    | 1.81  | Lipid transport and metabolism                                              |
| P    | 148   | 3.88  | Inorganic ion transport and metabolism                                       |
| Q    | 71    | 1.86  | Secondary metabolite biosynthesis, transport and catabolism                 |
| R    | 525   | 13.75 | General function prediction only                                             |
| S    | 443   | 11.6  | Function unknown                                                             |
|     | 2694  | 43.78 | Not in COG                                                                  |
Table 5. Genome statistics comparison among Arthrospira spp.

| Genome Name      | A. platensis C1 | A. platensis NIES-39 | A. maxima CS-328 | A. platensis Paraca | A. platensis |
|------------------|-----------------|-----------------------|------------------|--------------------|--------------|
| Genome size (bp) | 6,089,210       | 6,788,435             | 6,003,314        | 4,997,563          | 6,145,553    |
| Total genes      | 6,153           | 6,676                 | 5,730            | 5,401              | 5,718        |
| Protein-coding genes | 6,108       | 6,630                 | 5,690            | 5,370              | 5,675        |
| Protein with function prediction | 3,757      | 2,542                 | 3,315            | 3,023              | 3,023        |
| RNA genes        | 45              | 46                    | 40               | 31                 | 43           |
| Enzymes          | 952             | 905                   | 889              | 816                | 882          |
| % Enzymes        | 15.47%          | 13.56%                | 15.51%           | 15.11%             | 15.42%       |
| Transporter Classification | 345          | NA                    | NA               | NA                 | NA           |
| % Transporter Classification | 5.61%       | NA                    | NA               | NA                 | NA           |
| KEGG pathways    | 1,012           | 993                   | 931              | 904                | 954          |
| % KEGG pathways  | 16.45%          | 14.87%                | 16.25%           | 16.74%             | 16.68%       |
| KEGG Orthology (KO) | 1,837       | 1,702                 | 1,623            | 1,492              | 1,658        |
| % KEGG Orthology (KO) | 29.86%     | 25.49%                | 28.32%           | 27.62%             | 29.00%       |
| COGs             | 3,459           | 3,570                 | 3,306            | 3,050              | 3,234        |
| %COGs            | 56.22%          | 53.48%                | 57.70%           | 56.47%             | 56.56%       |
| Pfam             | 3,529           | 3,598                 | 3,564            | 3,431              | 3,526        |
| %Pfam            | 57.35%          | 53.89%                | 62.20%           | 63.53%             | 61.66%       |
| TIGRfam          | 1,180           | 1,213                 | 1,160            | 1,185              | 1,203        |
| %TIGRfam         | 19.18%          | 18.17%                | 20.24%           | 21.94%             | 21.04        |
| InterPro         | 4,244           | 3,969                 | 3,938            | 4,207              | 4,294        |
| %InterPro        | 68.97%          | 59.45%                | 68.73%           | 77.89%             | 75.10%       |
| Signal peptides  | 570             | 1,319                 | 545              | 559                | 1,150        |
| %signal peptides | 9.26%           | 19.76%                | 9.51%            | 10.35%             | 20.11%       |
| Transmembrane proteins | 1,094       | 1,123                 | 1,053            | 1,094              | 1,057        |
| %Transmembrane proteins | 17.78%     | 16.82%                | 18.38%           | 20.26%             | 18.49%       |
| COG clusters     | 1,569           | 1,566                 | 1,491            | 1,489              | 1,472        |
| KOG clusters     | 729             | 723                   | 722              | 737                | 717          |
| Pfam clusters    | 1,740           | 1,732                 | 1,702            | 1,729              | 1,730        |
| TIGRfam clusters | 932             | 936                   | 902              | 924                | 931          |

NA: Information not available.
The genome atlas of *A. platensis* C1 is shown in Figure 2. The atlas was created based on one FASTA file containing the nucleotide sequences in one piece. Together with the annotation file, positions of genes were used for the map.

In the genome atlas, the gene annotation, base content, AT and GC skew, percent AT and some structural properties of the DNA are shown. The structural properties are Position Preference, Stacking Energy and Intrinsic Curvature, all of which are related to the flexibility and strength of the DNA molecule [27].

**Genomic comparison**

The genome sequences from all *Arthrospira* strains (*A. maxima* CS-328, *A. platensis* NITE-39, *A. platensis* PCC8005, *A. platensis* Paraca, and *A. platensis* C1) provided the data for a comparative genome analysis of these strains. Based on the genome statistics comparison (Table 5), all of the
Arthrospira spp. genomes revealed highly conserved sequences. Interestingly, the number of signal peptides in *A. platensis* C1 has been reported to be the lowest among the *Arthrospira* spp. Further study of these primary sequences may reveal the importance of signal peptides in protein targeting in this cyanobacterial species compared with other laboratory and commercial strains.

**Insights from the genome sequence**

Analysis of the genome of *Arthrospira* spp. compared to other cyanobacteria confirmed that the *Arthrospira* are non-nitrogen fixing, filamentous cyanobacteria. They are nontoxic, which might be due to the absence of polyketide and non-ribosomal peptide-producing genes.

Genes involved in gliding motility in *A. platensis* NIES-39 (a vigorously motile strain utilizing type IV pili as the major mechanism for gliding) [6] have been compared with those in *A. platensis* C1 (a nonmotile strain). Interestingly, all the genes involved in type IV pili are present in the *A. platensis* C1 genome, however, the lack of gliding ability in *A. platensis* C1 is due to an unknown mechanism. Further studies are needed to elucidate this mechanism of cell motility.

**Cellular defense mechanism**

Like other genomes of *Arthrospira* spp., the genome of *A. platensis* C1 contains highly interspersed repetitive sequences that account for 9% of its genome. Genome comparisons among cyanobacteria revealed unusual genes involved in defense mechanisms, including restriction and modification enzymes, group II introns, insertion elements and CRISPR. These genes are considered to be major barriers for stable transformation. Therefore, these genes have been targeted for the development of a stable gene transformation system for *Arthrospira* spp. Because of the nonmotile property of *A. platensis* C1, single colonies can be selected and used for further strain improvement and genetic manipulation experiments.

By combining a stable transformation system with the advantage of colony-forming ability, we should be able to harness *A. platensis* C1 for many biotechnological applications using gene manipulation and systems biology.

**Transporter Characteristic of *A. platensis* C1**

Because of the ability of cyanobacteria to adapt to extremely different habitats, the relationship between their membrane transporter proteins and their habitats has been a focus of interest. Membrane transporters are proteins that allow cell membranes to deliver essential nutrients, eject waste products, and help the cell sense environmental conditions around it [28].

Interestingly, all *Arthrospira* species contain genes for Na+/H+ antiporters. The NapA-type Na+/H+ antiporter homolog, which is reported to be involved in salt tolerance at alkaline pH in some cyanobacteria, is found in all *Arthrospira* genomes. *Arthrospira* species live in high-alkalinity environments, and there are many alkali transporters in their genomes. These results revealed the relationship between the transporters and the lifestyles and niche adaptations of cyanobacteria.

**Acknowledgements**

This work was supported by the National Center for Genetic Engineering and Biotechnology (BIOTEC), Thailand under the grant numbers P-00-20061 and P-09-00186. The authors would like to thank Prof. Avigad Vonshak (Algal Biotechnology, Ben-Gurion University of the Negev, Israel) for the *A. platensis* C1 strain used in our experiments and Prof. Norio Murata for all his valuable suggestions during this project.

**References**

1. Habib MAB, Pariv M, Huntington TC, Hasan MR. A review on culture, production and use of spirulina as food for humans and feeds for domestic animals and fish. Rome: Food and Agriculture Organisation (FAO) of the United Nations. 2008; 1-33.

2. Richmond A. *Spirulina*, Microalgal Biotechnology. In: Borowitzka M, Borowitzka L, editors. Microalgal Biotechnology. New York: Cambridge University Press; 1987. p. 85-121.

3. Vonshak A. *Spirulina*: Growth, physiology and biochemistry. In: Vonshak A, editor. *Spirulina platensis* (*Arthrospira*). London: Taylor & Francis; 1997, p. 43-65.

4. Henrikson R. *Spirulina* World Food: How this micro algae can transform your health and our planet. Ronore Enterprises, Inc. PO Box 909, Hana, Maui, Hawaii 96718; 2010.
5. Janssen PJ, Morin N, Mergay M, Leroy B, Wattiez R, Vallaeys T, Waleron K, Waleron M, Wilmotte A, Quillardet P, et al. Genome sequence of the edible cyanobacterium Arthrospira sp. PCC 8005. J Bacteriol 2010; 192:2465-2466. PubMed http://dx.doi.org/10.1128/JB.00116-10

6. Fujisawa T, Narikawa R, Okamoto S, Ehira S, Yoshimura H, Suzuki I, Masuda T, Mochimaru M, Takaichi S, Awai K, et al. Genomic structure of an economically important cyanobacterium, Arthrospira (Spirulina) platensis NIES-39. DNA Res 2010; 17:85-103. Med http://dx.doi.org/10.1093/dnares/dsq004

7. Schirrmeister BE, Antonelli A, Bagheri HC. The origin of multicellularity in cyanobacteria. BMC Evol Biol 2011; 11:45. Med http://dx.doi.org/10.1186/1471-2148-11-45

8. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. Nat Biotechnol 2008; 26:541-547. Med http://dx.doi.org/10.1038/nbt1360

9. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 1990; 87:4576-4579. Med http://dx.doi.org/10.1073/pnas.87.12.4576

10. Woese CR, Stackebrandt E, Macke TJ, Fox GE. A phylogenetic definition of the major eu bacterial taxa. Syst Appl Microbiol 1985; 6:143-151. PubMed http://dx.doi.org/10.1016/S0723-2020(85)80047-3

11. McNeill J, Barrie FR, Burdet HM, Demoulin V, Hawksworth DL, Marhold K, Nicolson DH, Prado J, Silva PC, Skog JE, et al. International Code of Botanical Nomenclature, A.R.G. Ganter, Königstein, 2006, p. 1. 

12. Woese CR, Stackebrandt E, Macke TJ, Fox GE. A phylogenetic definition of the major eu bacterial taxa. Syst Appl Microbiol 1985; 6:143-151. PubMed http://dx.doi.org/10.1016/S0723-2020(85)80047-3

13. Castenholz, R.W. 2001. Oxygenic photosynthetic bacteria. In: Boone D.R. and Castenholz R.W. (eds), Bergey’s Manual of Systematic Bacteriology (2nd ed.), Volume 1, Springer-Verlag, New York, pp. 473-600.

14. Tomaselli L, Maria Rosa P, Mario RT. On the correct use of the Spirulina designation. AlgoL Stud 1996; 83:539-548.

15. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000; 25:25-29. Med http://dx.doi.org/10.1038/75556

16. Romano I, Bellitti MR, Nicolaus B, Lama L, Manca MC, Pagnotta E, Gambacorta A. Lipid profile: a useful chemotaxonomic marker for classification of a new cyanobacterium in Spirulina genus. Phytochemistry 2000; 54:289-294. Med http://dx.doi.org/10.1016/S0031-9422(00)00090-X

17. Cohen Z, Margheri MC, Tomaselli L. Chemotaxonomy of cyanobacteria. Phytochemistry 1995; 40:1155-1158. http://dx.doi.org/10.1016/0031-9422(95)00335-5

18. Cohen Z, Vonshak A. Fatty acid composition of Spirulina and Spirulina-like cyanobacteria in relation to their chemotaxonomy. Phytochemistry 1991; 30:205-206. http://dx.doi.org/10.1016/0031-9422(91)84125-C

19. Zarrouk C. Contribution à l'étude d'une cyanophyce'e. Influence de divers facteurs physiques et chimiques sur la croissance et la photosynthèse de Spirulina maxima. Ph. D. Thesis, Université' de Paris. 1966.

20. Phrap and Phred for Windows, MacOS, Linux, and Unix.

21. Arthrospira maxima CS-328 analysis files. http://genome.orl.gov/microbial/amax

22. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics 2009; 25:2271-2278. PubMed http://dx.doi.org/10.1093/bioinformatics/btp393

23. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 1997; 25:955-964. Med http://dx.doi.org/10.1093/nar/25.5.955

24. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and improved detection of transfer RNA genes in genomic sequences. Nucleic Acids Res 2007; 35:3100-3108. Med http://dx.doi.org/10.1093/nar/gkm160

25. Hongsthong A, Sirijuntarut M, Prommeenate P, Lertldaluck K, Porkaew K, Cheevadhanarak S,
Tanticharoen M. Proteome analysis at the subcellular level of the cyanobacterium Spirulina platensis in response to low-temperature stress conditions. *FEMS Microbiol Lett* 2008; **288**:92-101. PubMed [http://dx.doi.org/10.1111/j.1574-6968.2008.01330.x](http://dx.doi.org/10.1111/j.1574-6968.2008.01330.x)

26. Hongsthong A, Sirijuntarut M, Yutthanasirikul R, Senachak J, Kurdrid P, Cheevadhanarak S, Tanticharoen M. Subcellular proteomic characterization of the high-temperature stress response of the cyanobacterium Spirulina platensis. *Proteome Sci* 2009; **7**:33. PubMed [http://dx.doi.org/10.1186/1477-5956-7-33](http://dx.doi.org/10.1186/1477-5956-7-33)

27. Hallin PF, Binnewies TT, Ussery DW. The genome BLASTatlas-a GeneWiz extension for visualization of whole-genome homology. *Mol Biosyst* 2008; **4**:363-371. PubMed [http://dx.doi.org/10.1039/b717118h](http://dx.doi.org/10.1039/b717118h)

28. Ren Q, Paulsen IT. Comparative analyses of fundamental differences in membrane transport capabilities in prokaryotes and eukaryotes. *PLOS Comput Biol* 2005; **1**:e27. PubMed [http://dx.doi.org/10.1371/journal.pcbi.0010027](http://dx.doi.org/10.1371/journal.pcbi.0010027)