



## UvA-DARE (Digital Academic Repository)

### Shades of red and green : the colorful diversity and ecology of picocyanobacteria in the Baltic Sea

Haverkamp, T.H.A.

**Publication date**  
2008

[Link to publication](#)

#### **Citation for published version (APA):**

Haverkamp, T. H. A. (2008). *Shades of red and green : the colorful diversity and ecology of picocyanobacteria in the Baltic Sea*. [Thesis, externally prepared, Universiteit van Amsterdam]. Netherlands Institute of Ecology (NIOO) - Royal Netherlands Academy of Arts and Sciences.

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

# Chapter 5

*Phenotypic and genetic diversification of Pseudanabaena spp. (Cyanobacteria)*

Thomas H. A. Haverkamp<sup>1</sup>

Silvia G. Acinas<sup>1,2</sup>

Jef Huisman<sup>3</sup>

Lucas J. Stal<sup>1,3</sup>

<sup>1</sup> Department of Marine Microbiology, Netherlands Institute of Ecology, NIOO-KNAW, P.O.Box 140, 4400 AC Yerseke, The Netherlands

<sup>2</sup> Present addresses: Institut de Ciències del Mar, CMIMA-CSIC, 08003, Barcelona, Spain.

<sup>3</sup> Aquatic Microbiology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS Amsterdam, The Netherlands.



Accepted for publication in: The ISME journal.

Running Title: Diversity of *Pseudanabaena* populations.

Subject Category: Microbial population and community ecology

Keywords: *Pseudanabaena* / Cyanobacteria/ Complementary chromatic adaptation/ Microbial diversity/ Biogeography/ Evolutionary diversification

## Abstract

*Pseudanabaena* species are poorly known filamentous bloom-forming cyanobacteria closely related to *Limnothrix*. Because of their small size, the importance of *Pseudanabaena* has been overlooked and they have not been recognized as dominant organisms in blooms. We isolated 28 *Pseudanabaena* strains from the Baltic Sea and the Albufera de Valencia (Spain). By combining phenotypic and genotypic approaches the phylogeny, diversity, biogeography and evolutionary diversification of these isolates were explored. Analysis of the *in vivo* absorption spectra of the *Pseudanabaena* strains revealed two coexisting pigmentation phenotypes: (i) phycocyanin-rich strains, and (ii) strains containing both phycocyanin and phycoerythrin. Strains of the latter phenotype were all capable of complementary chromatic adaptation (CCA). About 65kb of each of the *Pseudanabaena* genomes, containing the 16 and 23S rRNA genes, the ribosomal intergenic spacer ITS-1, the *cpcBA* operon encoding phycocyanin, and the intergenic spacer (IGS) between *cpcA* and *cpcB*, were sequenced. In addition, the presence of *nifH*, one of the structural genes of nitrogenase, was investigated. Sequence analysis of ITS and *cpcBA*-IGS allowed for the differentiation between *Pseudanabaena* isolates exhibiting high levels of microdiversity. This multi-locus sequencing approach revealed specific clusters for the Baltic Sea, the Albufera de Valencia, and a mixed cluster with strains from both ecosystems. The latter comprised exclusively CCA phenotypes. The phylogenies of the 16 and 23S rRNA genes are consistent, but analysis of other loci indicated loss of substructure, suggesting that recombination between these loci has occurred. Population genetic analyses of the phycocyanin genes suggest an evolutionary diversification of *Pseudanabaena* through purifying selection.

## Introduction

Cyanobacteria are the dominant component of the phytoplankton in many freshwater and marine environments where they may form nuisance blooms (Chorus and Bartram, 1999; Huisman *et al.*, 2005; Granéli and Turner, 2006). The attention is usually towards the larger species that form aggregates and possess gas vesicles that make them buoyant and therefore accumulate at the surface. These species are either toxic or fix N<sub>2</sub> or both and may cause serious environmental and socio-economical problems. However, often it is not recognized that smaller cyanobacteria exceed the larger species in terms of biomass and activity and therefore play a major role in the ecosystem dynamics. Several reports mentioned the occurrence of the tiny filamentous *Pseudanabaena* in cyanobacterial blooms in brackish and freshwater ecosystems (Vasconcelos and Pereira, 2001; Stal *et al.*, 2003; Gkelis *et al.*, 2005; Zwart *et al.*, 2005; Kim *et al.*, 2006; Willame *et al.*, 2006). Nevertheless, *Pseudanabaena* spp. have been overlooked, perhaps because they have not been recognized as a dominant component of the phytoplankton. As a consequence, *Pseudanabaena* is a poorly known cyanobacterium.

*Pseudanabaena* species are non-heterocystous cyanobacteria belonging to the order of Oscillatoriales. The family of *Pseudanabaenaceae* is characterized by simple trichomes with a width less than 4 µm. The cells are longer than wide, possess parietal thylakoids, contain polar gas vesicles, and the cross walls are conspicuously constricted (Castenholz *et al.*, 2001; Komárek, 2003). Some strains display complementary chromatic adaptation (CCA). This process allows these organisms to regulate the ratio of the accessory photosynthetic pigments phycocyanin (PC) and phycoerythrin (PE), which helps them to adapt to the prevailing light spectrum (reviewed in Kehoe and Gutu (2006)) thereby favoring their persistence in competition against other species (Stomp *et al.*, 2004; Stomp *et al.*, in press). Most cultured strains reveal gliding motility and some are capable of anaerobic N<sub>2</sub>-fixation (Rippka and Herdman, 1992).

Morphologically, *Pseudanabaena* resembles *Limnothrix* making their identification difficult. The main differences are the somewhat wider cells (1- 6 µm) and the less distinct constriction of the cross walls in *Limnothrix* (Castenholz *et al.*, 2001). Although they are rarely recognized as dominant organisms, *Pseudanabaena* as well as *Limnothrix* species occur and form blooms in eutrophic water bodies and occasionally dominate the phytoplankton (Mayer *et al.*, 1997; Rucker *et al.*, 1997; Zwart *et al.*, 2005). *Limnothrix* is typically found in meso- to eutrophic freshwater ecosystems, whereas *Pseudanabaena* is more widely distributed and occurs in diverse aquatic environments (Castenholz *et al.*, 2001; Zwart *et al.*, 2005; Diez *et al.*, 2007). Based on the 16S rRNA gene, *Pseudanabaena* and *Limnothrix* form a monophyletic cluster within the cyanobacteria (Zwart *et al.*, 2005; Willame *et al.*, 2006). However, the available molecular data of the *Pseudanabaena* / *Limnothrix* group is scarce and consists of a limited number of environmental sequences and a few isolates. This limited data does not resolve the phylogeny of the *Pseudanabaena* / *Limnothrix* group.

Here we present a multi-phasic phenotypic and genotypic approach to explore the evolutionary diversification of *Pseudanabaena* strains isolated from two distant geographical regions in Europe, the Baltic Sea, a large brackish basin in the North, and Albufera de Valencia,

a coastal lagoon in the South of Europe. The Baltic Sea is one of the largest bodies of brackish water in the world. It is a eutrophic system that develops blooms of the conspicuous  $N_2$ -fixing heterocystous cyanobacteria *Aphanizomenon* and *Nodularia*, of which the latter is toxic (Stal *et al.*, 2003). However, the dominant component of the cyanobacterial community in the Baltic Sea consists of a colorful mixture of unicellular picocyanobacteria of the *Synechococcus* group and the tiny filamentous *Pseudanabaena* (Stal *et al.*, 2003; Stomp *et al.*, 2007; Haverkamp *et al.*, 2008). The Albufera de Valencia is a highly eutrophic coastal freshwater lagoon in Spain that is fed by streams, rivers and irrigation channels carrying fertilizer of the surrounding rice fields. The Albufera is characterized by dense water blooms of cyanobacteria among which *Pseudanabaena* spp. is a dominant group of organisms (Romo and Miracle, 1994; Villena and Romo, 2003). This is the first report on the diversity, biogeography and evolution of *Pseudanabaena* and a crucial step towards an understanding of the ecology of this remarkable but hitherto neglected organism.

## Materials and methods

### *Isolation, cultivation and strain collection*

The *Pseudanabaena* strains used in this study were isolated between 1995 and 2004 (Table 5.1). Strain CCY9508 was isolated in 1995 from the Baltic Sea (Bornholm Sea). The Spanish strains were isolated in 1997 from Albufera de Valencia. For the isolation of these strains, water was pre-filtered through 2  $\mu\text{m}$  mesh plankton net. Subsequently, the filtrate was spread onto 0.7% agarose medium in Petri dishes. Strain CCY9508 was isolated on a mixture consisting of 1/3 volume ASNIII + 2/3 volume BG11 medium with a salinity of 12.2‰ (Rippka *et al.*, 1979). The Spanish isolates were isolated on the freshwater BG11 medium. Single colonies were picked from the agarose plates and repeatedly transferred until axenic monoclonal strains were obtained. The other Baltic Sea strains were isolated from samples collected at various stations in the Gulf of Finland (from 59.1 °N 22.2 °E to 60.0 °N 26.2 °E) during a research cruise in July 2004. Water samples were collected from defined depths using a rosette sampler. *Pseudanabaena* strains were isolated using two different approaches. In one approach, water samples were fractionated, first using one layer and subsequently two layers of plankton net (20  $\mu\text{m}$  mesh) under gentle vacuum. This filtrate was successively filtered through 5  $\mu\text{m}$ , 1  $\mu\text{m}$  and 0.45  $\mu\text{m}$  membrane filters. Finally, the 0.45  $\mu\text{m}$  filter was transferred to a sterile 10 cm Petri dish filled with a mixture of 4/5 parts BG11 and 1/5 parts ASNIII medium, containing  $\text{NH}_4\text{Cl}$  ( $0.05 \text{ g}\cdot\text{l}^{-1}$ ) as the nitrogen source. In the other approach, water was filtered through 20  $\mu\text{m}$  plankton net and subsequently diluted to extinction in 96 deep well microtiterplates (Nunc, Inc.) containing a mixture of 4/5 parts BG11 and 1/5 parts ASNIII medium containing  $\text{NH}_4\text{Cl}$  ( $0.05 \text{ g}\cdot\text{l}^{-1}$ ). In both approaches the cells were first grown under a light regime of  $10 \mu\text{mol photons}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  for two weeks. Subsequently, the light intensity was increased to  $20 \mu\text{mol photons}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  and the cultures were incubated for another 6 weeks. Trichomes growing on the filters or at the surface of the wells were picked and transferred to solid media. In order

**Table 5.1.** Characteristics of *Pseudanabaena* strains used in this study, year and location of isolation and PCR amplification for *nifH* genes.

Isolation Details			Growth	Morphological Characteristics			CCA <sup>e</sup> Absorption Ratio			<i>nifH</i> PCR	
Strain	Origin	Year	Medium <sup>a</sup>	Cell Size (µm) <sup>b</sup>	PGV <sup>c</sup>	Mo-tility <sup>d</sup>	570 / 625 <sup>f</sup>	680 / 570 <sup>g</sup>	680 / 625 <sup>h</sup>	Amplifi-cation	
CCY9701	Albufera de Valencia	1997	BG 11	6.1x1.5	+	-	-	0.75	1.22	0.91	+
CCY9702	Albufera de Valencia	1997	BG 11	6.7x1.4	+	-	ND	ND	ND	ND	+
CCY9703	Albufera de Valencia	1997	BG 11	5.5x1.5	+	+	+	1.07	0.9	0.96	ND
CCY9704	Albufera de Valencia	1997	BG 11	6.0x1.6	+	-	+	1.02	1.07	1.1	ND
CCY9705	Albufera de Valencia	1997	BG 11	4.3x1.6	+	+	+	1.18	0.96	1.14	ND
CCY9709	Albufera de Valencia	1997	BG 11	4.7x1.5	+	+	+	1.06	1.01	1.08	ND
CCY9710	Albufera de Valencia	1997	BG 11	3.5x1.4	-	+	-	0.82	1.21	0.99	ND
CCY9712	Albufera de Valencia	1997	BG 11	4.9x1.5	+	+	-	0.67	1.22	0.81	ND
CCY9714	Albufera de Valencia	1997	BG 11	4.9x2.0	+	+	-	0.69	1.41	0.97	ND
CCY9715	Albufera de Valencia	1997	BG 11	4.9x1.4	+	-	+	1.18	1.08	1.27	ND
CCY9508	Baltic Sea	1995	1/3A+2/3B	6.1x1.5	+	+	+	0.67	1.51	1.01	ND
CCY0471	Baltic Sea	2004	1/5A+4/5B	3.2x1.4	+	+	+	1.04	1.48	1.53	+
CCY0472	Baltic Sea	2004	1/5A+4/5B	5.2x1.2	+	+	+	1.17	1.52	1.78	+
CCY0473	Baltic Sea	2004	1/5A+4/5B	2.5x1.5	+	+	+	1.28	1.12	1.43	+
CCY0474	Baltic Sea	2004	1/5A+4/5B	2.8x1.5	+	+	+	1.49	1.18	1.75	-
CCY0475	Baltic Sea	2004	1/5A+4/5B	2.9x2.0	+	+	+	1.52	0.88	1.34	-
CCY0476	Baltic Sea	2004	1/5A+4/5B	2.2x1.6	+	+	-	0.62	1.45	0.89	-
CCY0477	Baltic Sea	2004	1/5A+4/5B	2.6x2.0	+	+	+	1.29	0.95	1.23	+
CCY0478	Baltic Sea	2004	1/5A+4/5B	2.7x1.6	+	+	+	1.21	1.08	1.31	-
CCY0479	Baltic Sea	2004	1/5A+4/5B	2.6x1.9	+	+	+	1.4	0.92	1.28	-
CCY0480	Baltic Sea	2004	1/5A+4/5B	2.4x1.3	+	+	-	0.74	1.42	1.05	-
CCY0481	Baltic Sea	2004	1/5A+4/5B	2.1x1.9	+	+	+	1.52	0.96	1.47	-
CCY0482	Baltic Sea	2004	1/5A+4/5B	2.3x1.2	+	+	-	0.55	1.92	1.06	-
CCY0483	Baltic Sea	2004	1/5A+4/5B	2.4x1.9	+	+	+	1.46	0.89	1.31	-
CCY0484	Baltic Sea	2004	1/5A+4/5B	2.4x1.9	+	+	-	0.6	1.91	1.14	-
CCY0485	Baltic Sea	2004	1/5A+4/5B	2.3x1.3	+	+	-	0.59	2.51	1.5	-
CCY0486	Baltic Sea	2004	1/5A+4/5B	2.6x1.5	+	+	+	1.17	1.26	1.47	-
CCY0488	Baltic Sea	2004	1/5A+4/5B	3.0x1.3	+	+	-	0.69	1.62	1.11	-

<sup>a</sup>. Growth media description can be found in Rippka et al., 1979.[49] (1/3A +2/3B: 1 part ASNIII medium + 2 parts BG11 medium)

<sup>b</sup>. Cell size based on the average length x width of at least 40 cells.

<sup>c</sup>. PGV: Polar Gas Vesicles.

<sup>d</sup>. Motility on agarose plates.

<sup>e</sup>. CCA: Complementary Chromatic Adaptation.

<sup>f</sup>. Ratio between absorption at 570 nm (phycoerythrin) and 625 nm (phycocyanin).

<sup>g</sup>. Ratio between absorption at 680 nm (Chla) and 570 nm (phycoerythrin).

<sup>h</sup>. Ratio between absorption at 680 nm (Chla) and 625 nm (phycocyanin).

ND. Not determined.

to obtain monoclonal axenic strains, trichomes were repeatedly transferred. Once obtained, monoclonal axenic strains were maintained in their specific growth medium in the Culture Collection Yerseke (CCY) (Table 5.1).

### *Morphology and microscopy.*

Cells were collected from exponential or stationary liquid cultures and were fixed in a mixture of 1% (w/v) formaldehyde and 0.05% (w/v) glutaraldehyde and subsequently stored at -80°C until they were analyzed (Biegala *et al.*, 2003). Microscope slides were prepared by covering them with a thin layer of 1% (w/v) molten agarose (50 °C) (Sigma-Aldrich) that was allowed to solidify shortly before cells were applied. The slides were examined using a Zeiss Axiophot microscope equipped with a ProgRes C10 plus digital imaging system (JENOPTIK Laser, Optik, Systeme GmbH). The images were subsequently processed using ProgRes CapturePro2.0 software (JENOPTIK Laser, Optik, Systeme GmbH). From at least 30 cells in each culture the width and length were measured.

### *In vivo absorption spectra*

Exponential or stationary liquid cultures grown under white light (20  $\mu\text{mol photons}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ ) were used for determination of the *in vivo* absorption spectra. Spectra were measured from 400 to 750 nm using a Varian Cary 100 Bio equipped with an integrating sphere DRA-CA-3300. Distilled water was used as reference.

### *Determination of CCA*

In order to test for the capacity of CCA the strains were cultured on solid (agarose) media. Each strain was inoculated in two Petri dishes (Greiner Bio-One) that were incubated under a different color of light. Green light was obtained using Lee filter no. 124 (dark green) and red light through Lee filter no. 26 (red). The incident white light intensity was 100  $\mu\text{mol photons}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ . In order to document CCA, the cultures were photographed after two weeks of growth under monochromatic light. Subsequently, the cultures were changed to the other color of monochromatic light and incubated another two weeks after which they were documented again. A change from green to reddish / black phenotype and *vice versa*, was taken as evidence for CCA.

The strains CCY9703 and CCY9710 were also cultured in liquid medium. Three liquid cultures of each strain were grown under white light conditions (14  $\mu\text{mol photons}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ ) until stationary phase was reached. Subsequently, the three cultures were grown under white, green or red light. After two weeks the cultures were sampled and the *in vivo* absorption spectra were measured as described above.

### *DNA isolation*

For DNA extraction cells were collected from exponentially growing or stationary phase cultures. Briefly, two ml of culture was centrifuged in a table top centrifuge (Eppendorf type 5424) at 10000 r.p.m. for 1 minute at room temperature. The supernatant was removed and

**Table 5.2.** List of the primers used for PCR and sequencing in this study.

Primer Name	Target	Sequence (5'-3')	Reference
Bact1055F <sup>a</sup>	16S rRNA gene	AATGGCTGTCGTCAGCTCGT	Garcia-Martinez et al., 1999
PitsE-cyanR <sup>a</sup>	23S rRNA gene	CTCTGTGTGCCAAGGTATC	Ernst et al., 2003
129F <sup>b</sup>	23S rRNA gene	CYGAATGGGRVAACC	Hunt et al, 2006
2241R <sup>b</sup>	23S rRNA gene	ACCGCCCCAGTHAAACT	Lane et al, 1992*
cpcAR <sup>c</sup>	Phycocyanin operon	TAGTGTAACGACGGCCAGTTGYYTKGCGACATGGA	Robertson et al., 2001
cpcBF <sup>c</sup>	Phycocyanin operon	TAGCAGGAAACAGCTATGACGTGGTGTARGGGAAYTT	Robertson et al., 2001
nifH1	nifH	TGYGAYCCNAARGCNGA	Zani et al., 2000
nifH2	nifH	ADNGCCATCATYTCNCC	Zani et al., 2000
nifH3	NifH (internal primer)	ATRTTRTTNGCNGCRTA	Zani et al., 2000
nifH4	NifH (internal primer)	TTYTAYGGNAARGGNGG	Zani et al., 2000

<sup>a</sup>. These primers were also used to obtain the complete ITS-1 sequence.

<sup>b</sup>. These primers were used to amplify the complete 23S rRNA gene from the *Pseudanabaena* strains.

<sup>c</sup>. These primers were used to amplify the subunits beta and alpha of the phycocyanin operon plus the intergenic spacer (IGS) between both genes.

\* This primer was revised recently by Hunt et al, 2006.[50]

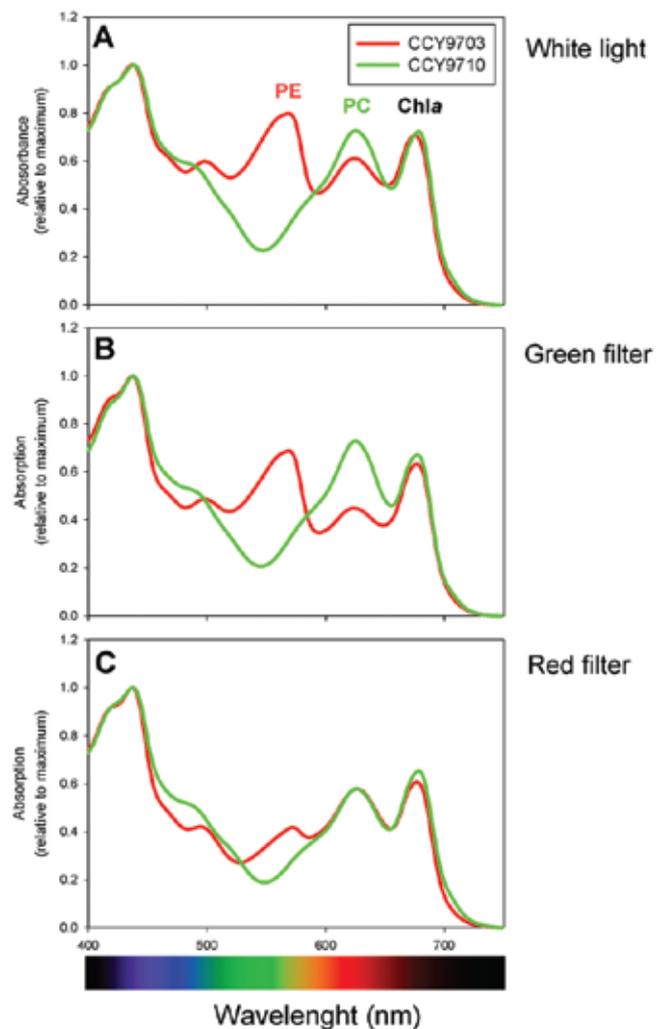
the cell pellet was re-suspended in lysis buffer provided by the Powersoil DNA extraction kit (MoBio). DNA extraction was performed following the instructions of the manufacturer. DNA quantity and quality were checked by running agarose gels as well as spectrophotometrically (Nanodrop ND1000).

### *PCR and sequencing*

PCR reactions were performed using the PCR primers listed in Table 5.2. The B1055F and PitsE-cyanR were used to amplify the end of the 16S rRNA gene plus the ITS-1. The 23S rRNA gene was amplified with 129F and 2241R primers and the phycocyanin operon (*cpcBA*) with *cpcAR* and *cpcBF* primers. Finally *nifH* PCR amplification were performed by the set of primers described in Table 5.2. Each reaction contained 0.2 mM of each dNTP, 2 mM MgCl<sub>2</sub>, 5 or 10 pmol of each primer, 1 µl template DNA (5-10 ng), 1x PCR buffer and 1 units of HotStarTaq (Qiagen). MQ-grade H<sub>2</sub>O was added to a final volume of 30 µl. The PCR reactions were carried out in a GeneAmp System 2700 thermocycler (Applied Biosystems). The PCR program used to amplify the 16S rRNA-ITS-1 region, the 23S rRNA gene and the *cpcBA* operon were as follows: a hot start at 94°C of 15 minutes, followed by 30 cycles of 1 minute 94°C, 1 minute of 55°C and 1 minute of 72°C. Following the last cycle an elongation step of 10 minutes at 72°C was applied. The PCR program for amplification of the *nifH* gene was according the nested PCR protocol as described by Zani *et al.* (2000) with modifications. In brief, the first PCR using primers nifH 1 and 2 was started with a hot start of 15 minutes at 96°C and followed by 35 cycles of 1 minute 94°C, 1 minute at 57°C, 1 minute at 72°C, followed by 10 minutes of elongation at 72°C. One µl of the PCR product was then used in the second PCR using the primers nifH 3 and 4. The PCR program started with a 15 minute hot

start at 94°C, followed by 35 cycles of 1 minute 94°C, 1 minute at 54°C, 1 minute at 72°C, followed by 10 minutes of elongation at 72°C.

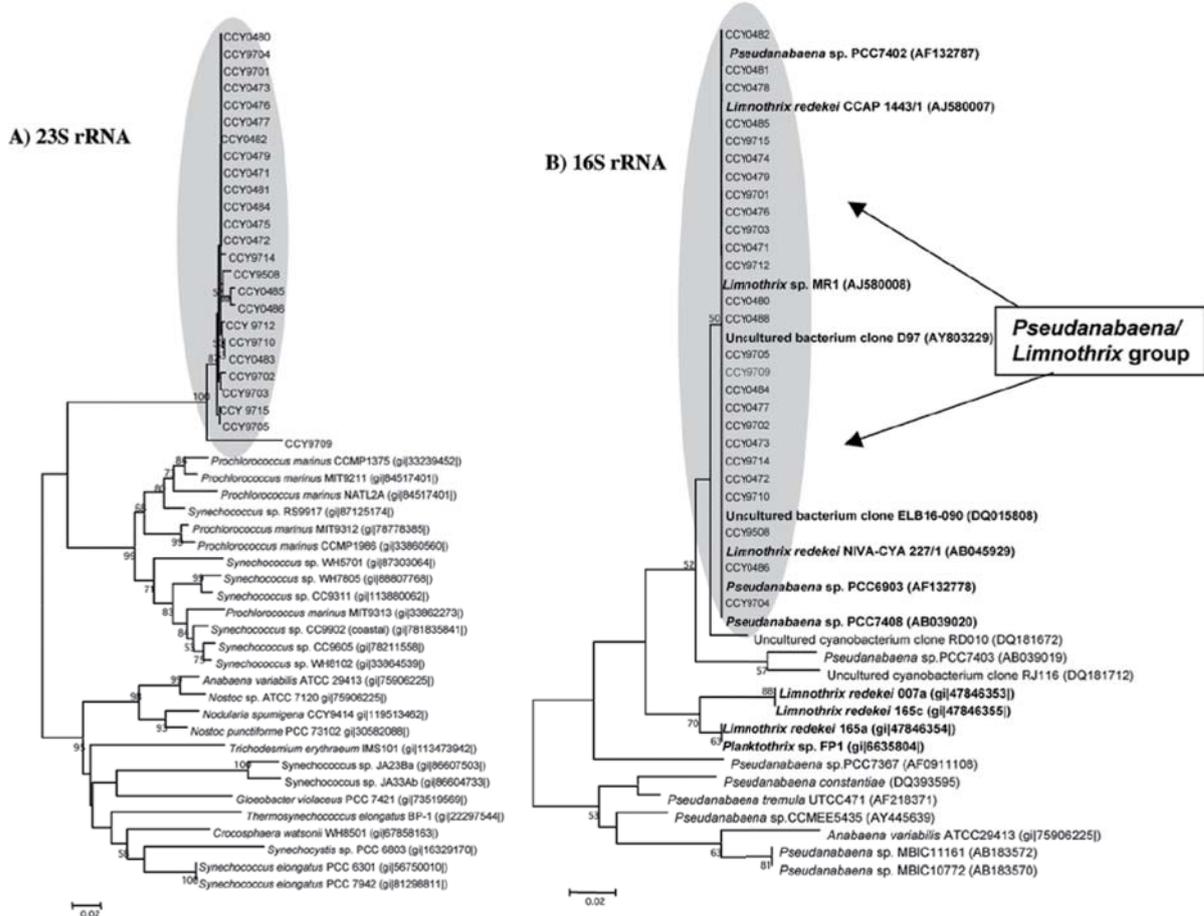
The DNA clean & concentrator-5 kit (Zymogram) was used in order to remove primer dimers from the PCR reactions following the instructions of the manufacturer. DNA concentration of the purified PCR products was checked spectrophotometrically (Nanodrop, ND1000). For sequencing reactions 3.5 µl of the purified and concentrated PCR product served as template using 10 µM of the sequencing primer and the Big Dye Terminator v1.1 cycle sequencing kit (Applied Biosystems) following the manufacturer's instructions. Sequencing primers were the same forward and the reverse primers as were used for PCR amplification (Table 5.2). Sequence products were analyzed using a 3130 Genetic Analyzer (Applied Biosystems). Sequences were edited manually using ChromasPro V 1.41 (Technelysium Pty Ltd) and manually checked for errors in base calling. Only high quality sequences were included in the final dataset. The sequences were deposited in GenBank under the following accession numbers: 16S rRNA (EU025781-EU025806), ITS-1 region (EU119301-EU119325), 23S-129F (EU025807-EU025831), 23S-2241R (EU025756-EU025780) and the *cpcBA* operon (EU119326-EU119352).



**Figure 5.1.** The effect of light colour on the pigment composition of *Pseudanabaena* strains CCY9703 and CCY9710. Three cultures of each strain were grown under white light until the cultures reached the stationary phase. The cultures were then transferred and grown under white (A), green (B) or red (C) light. Strain CCY9703 shows a decreased PE absorption under the influence of red light and an increased PC absorption indicating its capacity for complementary chromatic adaptation (CCA). Strain CC9710 showed a decrease in PC absorption under influence of red light (PC-rich strain).

### Phylogenetic analysis

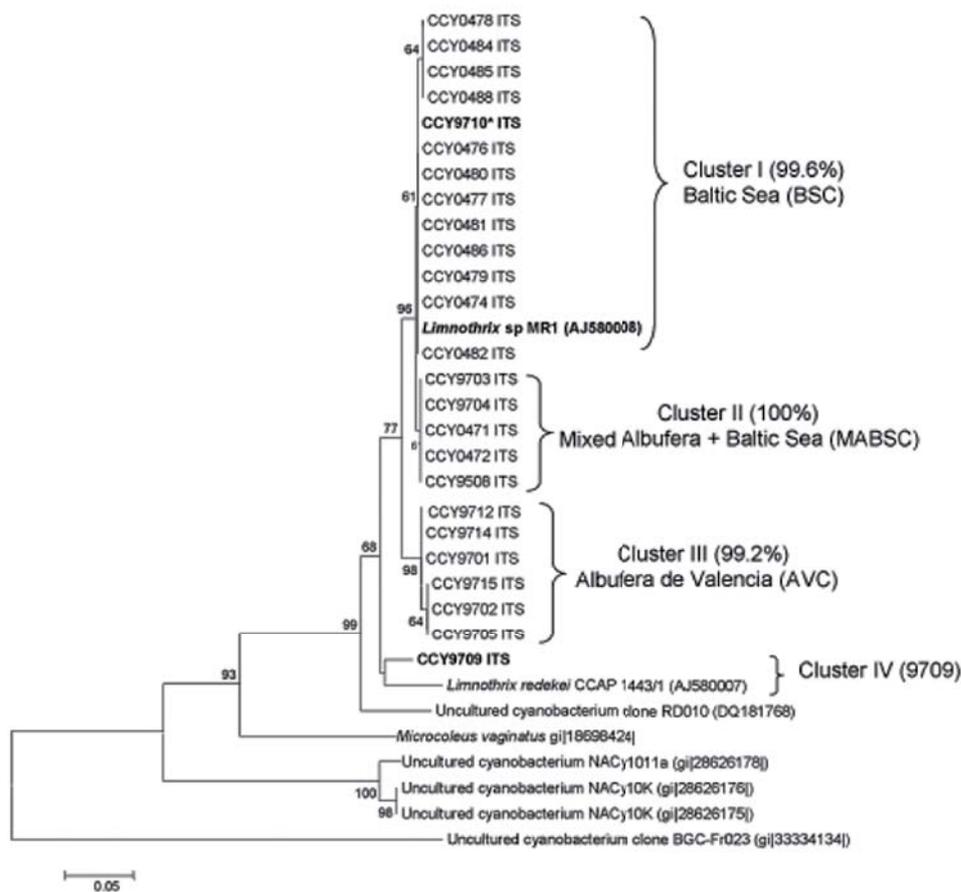
All *Pseudanabaena* sequences obtained from this study were aligned using CLUSTALW integrated into the package BioEdit (Hall, 1999). These sequences were aligned against sequences closely related to *Pseudanabaena/Limnothrix* group previously identified by BLASTN comparison from GenBank and other references identified in previous published studies. Sequence comparison and phylogenetic analyses of the partial sequencing of 16S rRNA, 23S rRNA, ITS-1, *cpcBA* and IGS were performed using the software MEGA3.1 (Kumar *et al.*, 2004). Neighbor-joining with Jukes–Cantor correction and 1000 bootstraps was used to build the corresponding phylogenetic trees. Partial amino acid sequences of the coding region of the *cpcBA* operon were also used in the phylogenetic analyses performed with the neighbor-joining method as well as with maximum parsimony. Maximum parsimony was used with the close-neighbor-interchange search algorithm with random tree addition using 100 bootstraps.



**Figure 5.2.** Neighbor Joining Tree based on comparison of the partial sequences of the 23S rRNA (A, 129F primer) and 16S rRNA (B, Bact1055F primer) genes of the *Pseudanabaena* strains collected in this study. All *Pseudanabaena* strains grouped in a single 99% cluster, for both the 23- and the 16S rRNA genes (shaded area).

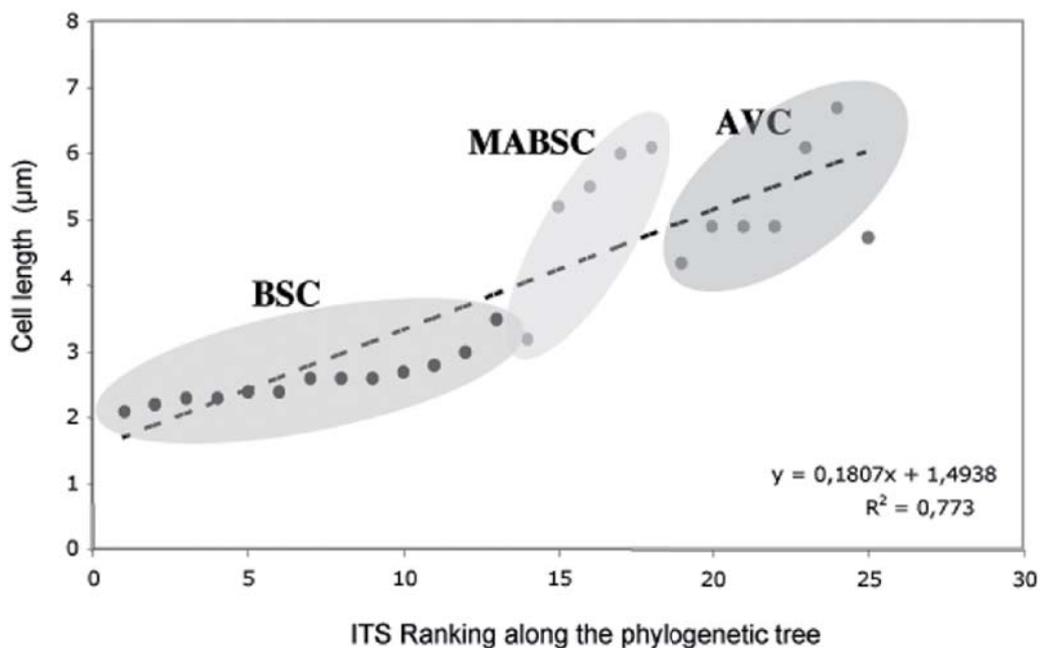
### Population genetic analysis.

The protein coding nucleotide sequences from the *cpcBA* locus aligned by ClustalW were analyzed using DnaSP version 4.0 (Rozas *et al.*, 2003) to calculate the following parameters: (i) synonymous and non-synonymous polymorphic sites, (ii) estimation of Ka/Ks divergence ratio, (iii) The McDonald-Kreitman test (McDonald and Kreitman, 1991) to detect positive or purifying selection and the HKA test based on neutral theory of molecular evolution (Hudson *et al.*, 1987), (iv) degree of genetic differentiation between populations, estimated by *Fst* (Hudson *et al.*, 1992), (v) estimation of the recombination parameter (*R*) and minimum number of recombination events (Hudson *et al.*, 1987).



**Figure 5.3.** Microdiversity of *Pseudanabaena* spp. as revealed by ITS-1 analysis. The ITS allowed further differentiation showing geographical patterns, with clusters specific for the Baltic Sea (Cluster I) and Albufera de Valencia, Spain (Cluster III), but also a cluster with representatives of both locations (Cluster II). Strain CCY9710 from the Albufera de Valencia represents the only exception and clustered with the Baltic Sea strains.

Finally, Selecton version 2.2 (<http://selecton.bioinfo.tau.ac.il>) was used to identify positive and purifying selection at each of the amino acids using a Bayesian inference approach (Stern *et al.*, 2007). The Selecton server automatically calculates the ratio between Ka and Ks ( $\omega$ ) at each codon site using a maximum likelihood (ML) approach. The value of  $\omega$  at each site is translated to a discrete color scale projected onto one of the homologous sequences for each sequence clusters. Colors 1 to 2 (dark and light yellow) indicate  $\omega > 1$  and stand for sites with positive selection while shades of white through magenta (colors 3 through 7) indicated various levels of  $\omega \leq 1$  where Selecton results can be accurately used to infer sites undergoing purifying selection.

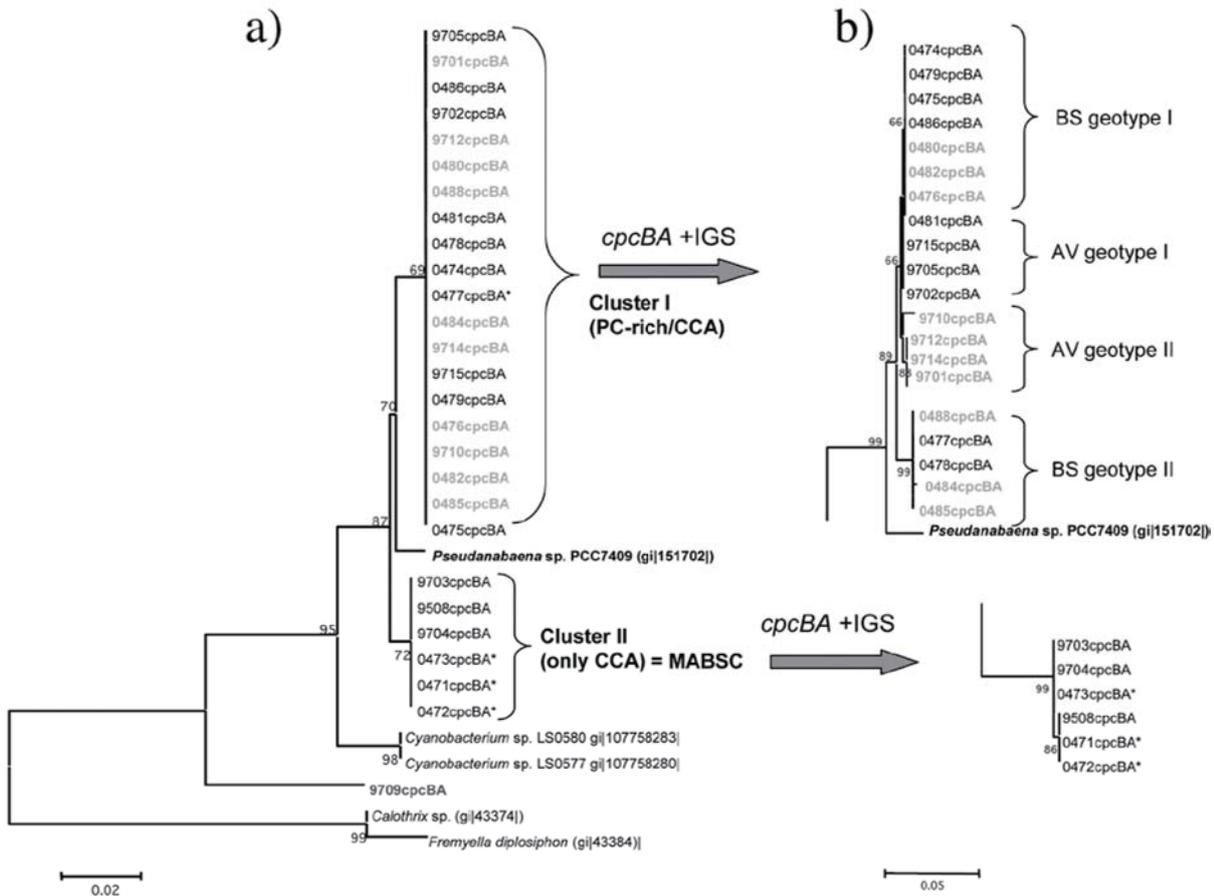


**Figure 5.4.** Correlation between cell length and the ranking of *Pseudanabaena* strains in the phylogenetic tree derived from the ITS-1 sequences. BSC correspond to the Baltic Sea Cluster, MABSC to the Mixed Albufera and Baltic Sea Cluster and AVC to the Albufera de Valencia Cluster. The blue dot corresponds to strain CCY9709.

## Results

*Phenotypic traits: morphological characteristics and photosynthetic pigment composition.*

A total of 28 strains were isolated from two geographically distant locations: Albufera de Valencia, Spain (10 strains isolated in 1997) and the Baltic Sea (1 strains isolated in 1995, 17 strains in 2004). The isolates were assigned to *Pseudanabaena* based on their morphological characteristics, such as cell size, motility on agarose plates and the presence of polar gas vesicles (Table 5.1). The isolates displayed different cell dimensions ranging from 2.1×1.9 to 6.7×1.4 μm (length × width) (Figure 1SM). While cell width remained within a narrow range, cell length varied considerably. All strains, except CCY9710 from Albufera de Valencia, possessed polar gas vesicles (Table 5.1). Gliding motility was observed in most of the strains except for 4 isolates from Albufera de Valencia (Table 5.1).



**Figure 5.5.** (a) Phylogenetic relationships of *Pseudanabaena* strains using 150 amino acids of the phycocyanin operon (*cpcBA*) genes. The green labeled strains represent PC-rich *Pseudanabaena* spp. All other strains contained both phycocyanin and phycoerythrin and are capable of complementary chromatic adaptation (CCA). Asterisks indicate *nifH* positive strains. Cluster II consists entirely of CCA strains, and is identical to Cluster MABSC in the ITS-based phylogenetic tree (see Figure 3). (b) Phylogenetic relationships derived from the *cpcBA* operon concatenated with the Intergenic Spacer (IGS) (about 600 bp). This revealed different sub-clusters ('geotypes') from the Baltic Sea (BS) and Albufera de Valencia (AV) within Cluster I.

The presence of the major light-harvesting pigments chlorophyll *a*, phycocyanin (PC) and phycoerythrin (PE) was determined by *in vivo* absorption spectra, from which the ratios PE:PC, Chl*a*:PE and Chl*a*:PC were calculated (Table 5.1). *Pseudanabaena* isolates were divided into strains that have both PE and PC pigments (63% of all strains) and those with only PC (37%) as their major pigment (Figure 5.1). All strains possessing both PE and PC pigments were capable of complementary chromatic adaptation (CCA). PC-rich strains reveal an absorption peak at ~625 nm and, hence, harvest orange-red light effectively. The proportion of strains capable of CCA that were isolated from the Baltic Sea was 66%, slightly higher than the 55% for the Albufera de Valencia. The majority of strains that were positive for CCA possessed a ratio of the absorption at 570 and 625 nm of approximately 1 or more when incubated in white light. When incubated under red or green light cultures changed pigmentation towards green and red, respectively (Figure 5.1b and 5.1c and Figure 2SM). Chlorophyll *a* showed absorption peaks at 440 nm (Soret band) and 680 nm. The absorption peaks of the three major light harvesting pigments were at the same wavelengths in all isolates but their relative heights varied substantially (Table 5.1).

### The phylogeny of *Pseudanabaena* revealed from their 16- and 23S rRNA genes

In order to determine the phylogenetic relationships of the *Pseudanabaena* isolates, the 16- and 23S rRNA genes were partly sequenced and analyzed. Because the 23S rRNA gene offers a higher phylogenetic resolution, the start and end of the 23S rRNA gene were both sequenced



**Figure 5.6.** Selection results for the *cpcBA* locus on 20 sequences from Cluster I and 6 sequences from Cluster II using the E8 model (Stern et al., 2007). Different levels of purifying selection are colored in shades of magenta through 150 amino acids of the *cpcBA* locus. The scale bar indicates selection ranging from positive (yellow colors 1 and 2) to purifying selection (colors 3 through 7). No positive selection is observed. The star symbols represent the non-synonymous sites (amino acids changes) between Cluster I and Cluster II.

using the forward 129F and reverse 2241R primers, respectively (Table 5.2). This resulted in two products with a sequence length of 440bp (129F) and 527 bp (2241R) respectively. The sequences obtained by using these primers gave similar phylogenetic relationships, although a larger number of polymorphisms were observed when using primer 129F (data not shown). Figure 5.2 depicts the phylogeny based on neighbor-joining analysis and compares the 23S rRNA (Figure 5.2a) with the 16S rRNA gene tree topologies (Figure 5.2b) obtained using the primers 129F (440bp) and Bact1055F (400bp) respectively. In both trees all isolates grouped into a single 99% similarity cluster indicating congruency in the phylogeny of the 16- and 23S rRNA genes. Moreover, the 23S rRNA gene displayed more polymorphisms than the 16S rRNA gene, which was virtually identical in virtually all isolates. The phylogenetic analysis of the 16S rRNA gene confirmed that all isolates belong to the *Pseudanabaena/Limnothrix* group with 99% cluster similarity (Zwart *et al.*, 2005; Willame *et al.*, 2006). From BLAST searches against the GenBank database we observed that 42% of the isolates were 100% identical to *Pseudanabaena* sp. PCC6903 (AM709632). The other strains possess high similarity (99%) to *Pseudanabaena* sp. Itu24s9 (AM259269), which originated from the Finnish freshwater Lake Tuusulanjarvi (Table 1SM). Only strain CCY9709 from Albufera de Valencia exhibited a higher divergence based on the 23S rRNA sequence (Figure 5.2a) showing a 94% similarity with *Pseudanabaena* sp. PCC6903 (Table 1SM).

### *Microdiversity within Pseudanabaena strains: ITS sequencing analysis*

The partial sequencing of the 16- and 23S rRNA genes was not sufficient to resolve the phylogeny of the *Pseudanabaena/Limnothrix* group. Therefore the internal transcribed spacer (ITS) region located between the 16- and 23S rRNA genes was sequenced. All *Pseudanabaena* ITS sequences revealed the same structure and contained two tRNA genes, tRNA<sup>Ile</sup> and tRNA<sup>Ala</sup> (data not shown). However, there was a high nucleotide divergence and length variability among the ITS sequences. The phylogenetic analysis of the ITS revealed a higher level of differentiation within the *Pseudanabaena/Limnothrix* group (Figure 5.3), which constrained most of the strains in three major 99% clusters (“microdiversity clusters”). This indicates a high level of microdiversity among the *Pseudanabaena* isolates. Unfortunately, no ITS sequences related to *Pseudanabaena* species have been published to date, and therefore our sequences clustered with the two known *Limnothrix* ITS sequences.

The ITS analysis uncovered biogeographical patterns with clusters containing isolates specific for the Baltic Sea, isolates specific for the Albufera de Valencia, and a mixed cluster containing isolates from both locations. Most of the Baltic Sea strains grouped in the Baltic Sea Cluster (BSC) with 99.6% similarity. BSC contained 12 isolates from the Baltic Sea but also one strain from the Albufera de Valencia (CCY9710) and *Limnothrix* sp. MR1 (isolated from Lake Loosdrecht, The Netherlands (Zwart *et al.*, 2005)). The Mixed Albufera and Baltic Sea Cluster (MABSC) contained 5 identical sequences retrieved from isolates from both locations. The fact that strains originating from such distant locations and with 7 years between their isolation possess identical ITS sequences is remarkable. The Albufera de Valencia Cluster (AVC) comprised 6 isolates from Albufera de Valencia possessing 99.2% similarity. Strain CCY9709 grouped in a fourth cluster together with *Limnothrix redekei* CCAP 1443/1. Furthermore,

unique insertion sequences (IS) were found in some clusters. A unique 36 bp IS was found in the MABSC. In CCY9709 three specific IS of 23, 6 and 5 base pairs were found. The position of the isolates in the ITS phylogenetic tree showed a close relationship with cell length (Figure 5.4). The cell length of all BSC strains, including CCY9710 from Albufera de Valencia, ranged from 2.1 - 3.5  $\mu\text{m}$ . The MABSC cluster was characterized by cells that are slightly longer, ranging from 3.2 - 6.1  $\mu\text{m}$ . Finally, the AVC cluster comprised the strains with the longest cells, ranging from 4.3 to 6.7  $\mu\text{m}$ . Strain CCY9709 possessed cells of 4.9  $\mu\text{m}$  long and grouped in an independent monophyletic cluster with *Limnothrix redekei* CCAP 1443/1 which, however, has considerably longer cells (6-10  $\mu\text{m}$ , personal communication of the Culture Collection of Algae and Protozoa (CCAP)).

### *Phylogeny of the phycocyanin operon: correlation of the cpcBA gene clusters with light absorption spectra*

A major limitation of the use of ribosomal genes as molecular markers is the impossibility to attribute eco-physiological traits to these genes. Therefore, we also sequenced the phycocyanin operon (*cpcBA*). *CpcBA* encodes the two subunits of phycocyanin, which is part of the phycobilisome, the major light harvesting complex of cyanobacteria. The main goal was to assign sequence clusters of *cpcBA* to the *in vivo* light absorption spectra of the isolated strains in order to reveal ecologically different populations (ecotypes) within *Pseudanabaena* (Figure 5.5). The *cpcBA* locus has been widely used for the study of cyanobacterial diversity and phylogeny and it is therefore suitable for our purpose (Ivanikova *et al.*, 2007; Six *et al.*, 2007; Haverkamp *et al.*, 2008). In addition, the intergenic spacer (IGS) between *cpcB* and *cpcA* (*cpcBA*-IGS) was sequenced (Figure 3SM and Figure 5.5b) in order to explore whether coding and non-coding regions of the *cpcBA* operon display different evolutionary rates resulting in different phylogenies.

Phylogenetic analysis of the partial sequences of *cpcBA* encoding for 150 amino acids, displayed two well-supported clusters (Figure 5.5a) with similarities higher than 99% ("microdiversity clusters"). Cluster I comprises sequences with 99.9% similarity and grouped 20 isolates from both locations. Half of these isolates were rich in phycocyanin (PC-rich) and unable to perform CCA, while the other 10 strains were capable of CCA. This cluster is closely related to *Pseudanabaena* sp. PCC7409 that is capable of CCA. Cluster II contained only strains capable of CCA, both from the Baltic Sea and the Albufera de Valencia and possess 100% sequence similarity. Half of the strains of Cluster II investigated for *nifH* PCR amplification possessed *nifH* encoding for dinitrogenase reductase, which is a component of nitrogenase. Among Cluster I there was only one strain (CCY0477) that possessed *nifH* in a total of 20 isolates. *NifH* was found in strains originating from both environments. *CpcBA* cluster II in Figure 5.5 corresponds to the MABSC cluster of ITS sequences in Figure 5.4. This hints to the presence of a conserved and coherent lineage within *Pseudanabaena*. Again the CCA positive strain CCY9709 displayed the most divergent position in the *cpcBA* phylogeny.

The phylogeny of the concatenated *cpcBA*-IGS sequences revealed biogeographical patterns that were not shown in the *cpcBA* phylogeny alone (Figure 5.5b). The tree topology of the

concatenated *cpcBA*-IGS phylogeny was consistent with the topology of the *cpcBA* phylogeny without the IGS region. Moreover, the heterogeneity and length variability of the IGS sequences allowed a higher level of differentiation and revealed different sub-clusters from specific geographic locations (“geotypes”). Four geotypes were discerned within cluster I (PC-rich/CCA), two of them originated from the Baltic Sea and the other two from Alfabera de Valencia (Figure 5.5b). In all geotypes, specific base pair signatures were assigned (Figure 3SM). Complete IGS sequence length ranged from 104 to 212 bp. The strains of Cluster I (PC-rich/CCA) possessed an IGS of 153 bp, while the IGS sequences of cluster II (only CCA) were shorter and possessed only 104 base pairs. Strain CCY9709 had the longest (212 bp) and most divergent IGS (Figure 3SM). Moreover, Cluster II (only CCA) comprised different IGS sequences compared to Cluster I-IGS and contained several deletions (in total 49 base pairs). Nonetheless, Cluster II-IGS sequences were conserved and only a few specific signatures were found. For example, the T at position 120 or the C at position 132 are representative for 2 of the 3 Baltic Sea strains that possess *nifH* (Figure 3SM).

### *Evolutionary forces operating on the phycocyanin operon (cpcBA) genes*

Based on the analysis of the *cpcBA* operon, the evolutionary processes shaping the phycocyanin genes among *Pseudanabaena* spp. were investigated. Partial sequencing of 450 nucleotides constrained 26 strains in two major lineages, Cluster I (20 isolates) and Cluster II (6 isolates) (Figure 5.5a). The number of polymorphisms within and between both lineages was examined. A total of 16 single nucleotide polymorphisms (SNPs) were found between both lineages and therefore the nucleotide divergence observed between both clusters was 3.82%. The number of synonymous (silent) nucleotide substitutions per synonymous site ( $K_s = 0.0809$ ) was 12.3 fold higher than the non-synonymous substitutions (amino acid-changing) per non-synonymous site ( $K_a = 0.00655$ ). The ratio of non-synonymous and synonymous nucleotide substitutions in protein-coding genes ( $K_a/K_s$ ) may give important clues about the selection pressure on and evolution of protein coding genes (McDonald and Kreitman, 1991). The ratio  $K_a/K_s$  observed for *cpcBA* was 0.074. Ratios of  $K_a/K_s < 1$  indicate that purifying selection takes place. Hence, this is the case for the *cpcBA* clusters I and II. Purifying selection means that non-synonymous vs. synonymous mutations is favored in the direction of the latter. As a consequence, over time, slightly deleterious non-synonymous mutations are continuously removed from the population leaving only synonymous mutations. In addition, the possibility to detect different types of selection (positive or purifying) on specific amino acids within the *cpcBA* operon was explored (Stern *et al.*, 2007). Positive selection was not detected in any of the amino acids, which indicates that all amino acids must have been under different levels of purifying selection (Figure 5.6). Furthermore, the degree of genetic differentiation between both clusters (*Pseudanabaena* subpopulations) was also explored by estimating the fixation index  $F_{st}$ , which indicates the amount of gene flow between populations (Hudson *et al.*, 1992). Values of  $F_{st}$  range between 0-1. A value of 0 indicates that the populations share the same alleles, while a value of 1 shows that the populations are fixed for different alleles. The  $F_{st}$  value for the *cpcBA* loci was 0.886, which indicated high gene flow compared with other described bacterial populations (Whitaker *et al.*, 2003; Miller *et al.*, 2006).

Finally, the number of amino acid changes (non-synonymous sites) was assessed for the 150 codons analyzed. Only two amino acid changes were detected, yielding a 1.4% amino acid divergence. In both cases, the identity of the amino acids differed systematically between Cluster I and Cluster II of the *Pseudanabaena cpcBA* lineages. At position 42 at the end of the beta subunit (*cpcB*), a serine (S, alcohol polar R-group) found in cluster I was exchanged in Cluster II by an alanine (A, aliphatic R-group, a hydrophobic molecule). Moreover, at position 100, at the beginning of the alpha subunit (*cpcA*) a threonine (T, a hydrophilic and hydroxyl-containing amino acid) in cluster I was replaced by alanine (A) in Cluster II. However, neither of these two amino acids were found to be under positive selection (white color in Figure 5.6).

## Discussion

### *Pseudanabaena/Limnothrix group: agreement of phenotype with genetic data*

*Pseudanabaena* is morphologically similar to *Limnothrix*. The large variation of cell lengths observed among the strains isolated in this study suggests that *Pseudanabaena* shows a high level of plasticity. Accordingly, cell length is not very useful for distinguishing *Pseudanabaena* from *Limnothrix*.

Based on 16S rRNA gene sequences, earlier studies showed that *Pseudanabaena* and some strains of *Limnothrix* cluster with *Pseudanabaena* spp., including one isolate assigned to *Limnothrix redekei*. This cluster is commonly referred to as the *Pseudanabaena/Limnothrix* group (Gkelis *et al.*, 2005; Willame *et al.*, 2006). This group also comprised several environmental sequences as well as other isolates including the type strain *Pseudanabaena* PCC7408, strains belonging to *Limnothrix redekei* (Van Goor, Meffert) (*Limnothrix* sp. MR1 from Lake Loosdrecht, *L. redekei* CCAP 1443/1 and *L. redekei* CCAP 227/1) and several other isolates from Lake Loosdrecht, The Netherlands (Zwart *et al.*, 2005; Willame *et al.*, 2006). A second cluster, comprising only *Limnothrix redekei* strains isolated from Lake Kastoria (Greece), has also been observed (Gkelis *et al.*, 2005; Willame *et al.*, 2006). Based on partial sequencing of the 16- and 23S rRNA genes all *Pseudanabaena* strains isolated in this study cluster with the *Pseudanabaena/Limnothrix* group. Neither 16- nor 23S rRNA gene analysis could distinguish between *Pseudanabaena* and *Limnothrix*, confirming previous reports (Zwart *et al.*, 2005; Willame *et al.*, 2006). However, ITS analysis of the *Pseudanabaena* isolates shows a high nucleotide divergence and size variability revealing a higher level of differentiation in this group. For instance, strains belonging to the AV Cluster and strain CCY9709 all possessed cells that were almost 3 times longer than wide (Figure 5.4), a characteristic that is considered typical for *Limnothrix*. This could indicate that these strains are indeed closely related to *Limnothrix redekei*. Nonetheless, in order to confirm these observations, it is necessary to increase the number of ITS sequences of *Limnothrix redekei* and *Pseudanabaena*. Moreover, it will be necessary to re-evaluate the morphological basis on which the genera *Limnothrix* and *Pseudanabaena* are separated.

### *Biogeography of Pseudanabaena*

The distribution of *Pseudanabaena* is clearly more widespread than previously anticipated. Members of this genus have been reported from a variety of different environments including freshwater lakes (Gkelis *et al.*, 2005; Zwart *et al.*, 2005; Kim *et al.*, 2006; Willame *et al.*, 2006), brackish environments (Stal *et al.*, 2003), hot springs (Castenholz *et al.*, 2001) as well as from epilithic cyanobacterial communities of beach rock (Heron Island, Great Barrier Reef) (Diez *et al.*, 2007). Our study revealed several lineages of *Pseudanabaena* based on a multilocus sequence typing approach using 5 loci of 28 isolates. The average genetic divergence of each of the markers varied from 1% in the small (SSU) and large subunit (LSU) ribosomal RNA genes to 3.5% in ITS-1 and 3.82% in *cpcBA*. The Baltic Sea is geographically distant from the Albufera de Valencia. Moreover, they represent quite different habitats (i.e., brackish

vs. freshwater) and it is therefore not surprising to find different *Pseudanabaena* genotypes at these two locations. Endemic clusters (“geotypes”) were detected in the Baltic Sea (BS Cluster) and Albufera de Valencia (AV Cluster) based on the analysis of ITS sequences (Figure 5.3). The analysis of the *cpcBA*-IGS operon revealed even a better differentiation of these geotypes (Figures 5.5b, 3SM).

Yet, the Baltic Sea is connected to the North Sea, and via the Atlantic Ocean and the Mediterranean Sea it ultimately links to Albufera de Valencia. This connection could allow the long-range dispersal of *Pseudanabaena* genotypes, although also other mechanisms for long-range dispersal could be envisioned (such as transport by birds). Salinity differences between the Baltic Sea and the Albufera de Valencia are minor, and organisms might adapt quickly to somewhat higher or lower salt levels. Indeed, a coherent and “cosmopolitan” monophyletic cluster, comprising members from both locations was found for several loci (Cluster MABSC by ITS and Cluster II by *cpcBA* genes or *cpcBA*-IGS analysis) suggesting a conserved and homogeneous lineage within *Pseudanabaena* (Figures 5.3, 5.5). This lineage clustered the strains isolated from the two distant locations, even though the isolation at the two locations was separated by 7 years. This could point to global dispersal of these *Pseudanabaena* strains. However, more *Pseudanabaena* strains from other environments and locations should be included in order to confirm the possibility of global dispersal. Based on analysis of the genes of the phycocyanin operon and the light absorption spectra associated with them, it is proposed that this lineage or subpopulation of *Pseudanabaena* (Cluster MABSC by ITS and Cluster II by *cpcBA* genes or *cpcBA*-IGS analysis) represents an “ecotype” that possesses the ability of CCA. Interestingly, this ecotype also possesses the *nifH* gene in at least half of their members and therefore might be capable of N<sub>2</sub> fixation (Figure 5.5). This ecotype proposed with the name of PCCA ecotype seems to be widely distributed, and the ability of CCA provides it with a selective advantage.

### *Genetic diversification of Pseudanabaena populations*

Correlation of the *cpcBA* gene clusters with the light absorption spectra hinted at the coexistence of two *Pseudanabaena* populations with a niche differentiation along the light spectrum (Figure 5.5, Cluster I and II). Cluster II or the “PCCA ecotype” occurs as a group with 99% similarity (“microdiversity cluster”), at a variety of different loci (ITS, *cpcBA* gene and *cpcBA*-IGS analysis). Previous studies indicate that such microdiversity clusters could represent important units of differentiation as ecotypes in natural populations of bacteria (Palys *et al.*, 1997; Moore *et al.*, 1998; Rocap *et al.*, 2003; Konstantinidis and Tiedje, 2005; Lopez-Lopez *et al.*, 2005; Thompson *et al.*, 2005; Cohan, 2006; Polz *et al.*, 2006; Cohan and Perry, 2007), and are often observed in environmental clone libraries (Field *et al.*, 1997; Acinas *et al.*, 2004; Morris *et al.*, 2005; Johnson *et al.*, 2006; Pommier *et al.*, 2007). The microdiversity clusters identified here are correlated with morphological and ecophysiological traits such as cell length and the capacity to perform CCA (Figures 5.4, 5.5), providing further support for the designation as an ‘ecotype’ (Ahlgren and Rocap, 2006; Johnson *et al.*, 2006; Polz *et al.*, 2006; Ward *et al.*, 2006). Although this genetic pattern agrees with the ecotype model for bacterial species (Palys *et al.*, 1997; Cohan, 2002; Gevers *et al.*, 2005; Cohan and Perry, 2007),

other mechanisms causing the genetic diversification of *Pseudanabaena* populations cannot be excluded. Indeed, the loss of substructure in the tree topology occurred when different loci were compared (e.g. ITS vs. *cpcBA*-IGS). For instance, strain CCY9710 from Albufera de Valencia fell into the Baltic Sea Cluster based on its ITS sequence and it shared with all strains a similar cell length. However, this strain grouped with other isolates from Albufera de Valencia when considering the phylogeny based on *cpcBA*-IGS sequences. Similarly, the MABSC cluster was an independent branch in the phylogeny of the *cpcBA* locus while it was sister to the BSC using ITS. Moreover, at least 12 recombination events ( $R_m$ ) were detected at the *cpcBA* locus, emphasizing the importance of homologous recombination (HR) and that this process should be taken into account.

By using the ratio of non-synonymous vs. synonymous fixation as a measure of the level of selective pressure on the phycocyanin operon (*cpcBA*), it was concluded that purifying selection is involved in the evolutionary diversification of *Pseudanabaena* populations. Other population genetic analyses such Tajima's D and the Mc Donald-Kreitman (MK) tests for selection were not significant (data not shown), further supporting our finding that the mutations at *cpcBA* do not deviate from those expected from neutrality and, hence, are not under positive selection. The results show that divergence of the *cpcBA* in Cluster I and II is promoted by purifying selection in both populations.

Recent laboratory experiments investigated the role of complementary chromatic adaptation in the competition of *Pseudanabaena* against red and green *Synechococcus* strains (Stomp *et al.*, in press). The competition experiments showed that *Pseudanabaena* was a strong competitor in fluctuating light environments, provided that it had sufficient time to adjust its pigment composition to the prevailing light spectrum. *Pseudanabaena* can change its pigmentation from red to green, and *vice versa*, within ~7 days. Thus, *Pseudanabaena* benefited from CCA only if fluctuations in underwater light color were slow compared to the time required for CCA, corresponding to slow mixing processes or infrequent storms in their natural habitat (Stomp *et al.*, in press). We hypothesize that phycocyanin-rich strains (PC-rich) in Cluster I have lost the capacity of CCA recently by the loss or presence of dysfunctional genes required to synthesize the phycoerythrin disks in the phycobilisome, or the genes like *rcaE* which is needed for the control of CCA (Terauchi *et al.*, 2004; Kehoe and Gutu, 2006). Knock-out experiments targeting the *rcaE* gene showed that this gene is needed for responsiveness to both red and green light under complementary chromatic adaptation (Terauchi *et al.*, 2004). The loss or the presence of a dysfunctional copy of *rcaE* or other genes involved in the phycobilisomes or CCA may be caused by deleterious mutations, resulting in strains that have lost phycoerythrin and are only able to use phycocyanin and chlorophyll *a* as their main light harvesting pigments. Phycocyanin absorbs photons in the orange-red part of the light spectrum. Accordingly, loss of CCA is likely to be advantageous in moderately turbid waters where orange-red light predominates (Stomp *et al.*, 2007), and during storm periods with rapid mixing when CCA is too slow to track changes in the underwater light spectrum experienced by the entrained *Pseudanabaena* filaments (Stomp *et al.*, in press). Variation in the underwater light spectrum at a range of different time scales could thus have induced genetic divergence between PC-rich

and CCA strains because of differences in fitness. The proposed mechanism of selective sweeps could enable such genetic diversification (Cohan, 2002; Gevers *et al.*, 2005; Cohan, 2006; Cohan and Perry, 2007).

In summary, multi-locus sequencing of 5 independent loci revealed the existence of several lineages or subpopulations within *Pseudanabaena*. The phylogenies of the 16S rRNA genes are consistent, but analysis of the other loci indicated loss of substructure, suggesting recombination between these loci. *Pseudanabaena* isolates exhibited high levels of microdiversity unveiling biogeographical patterns with both local as well as more globally dispersed populations. A conserved *Pseudanabaena* lineage proposed as “PCCA ecotype” was characterized by the capacity of chromatic adaptation and possibility for N<sub>2</sub> fixation. Population genetic analyses of the phycocyanin genes suggest an evolutionary diversification of *Pseudanabaena* through purifying selection.

The isolation of additional *Pseudanabaena/Limnothrix* strains from a variety of different environments is required to further elucidate of the ecology, biogeography and evolution of this enigmatic but understudied group of cyanobacteria.

## Acknowledgements

The authors are grateful to Marije Doeleman for her assistance in sequencing and to Ute Wollenzien for isolation, culturing and maintenance of the *Pseudanabaena* strains. We thank Maayke Stomp and the crew of the RV Aranda for their great help during sampling of the Baltic Sea. Special thanks to Caroline Jenkins for providing us with 0.45 µm filters during the Cyano-cruise in 2004. The research of JH was supported by the Earth and Life Sciences Foundation (ALW), which is subsidized by the Netherlands Organization for Scientific Research (NWO). SGA was supported in the framework of the Flemish-Dutch collaboration for Sea Research (VLANEZO), TH and LJS acknowledge the support from the European Commission through the project MIRACLE (EVK3-CT-2002–00087). This is publication xxxxx of NIOO-KNAW.

## Supplementary Information

The following supplementary material is available for this article:

**Figure 1SM.** Light microscopy images of some representative *Pseudanabaena* isolates from the Baltic Sea and Albufera de Valencia, Spain. Scale bar, 5 $\mu$ m.

**Figure 2SM.** The effect of light color on the pigment composition of *Pseudanabaena* strain CCY9703. Three cultures of strain CCY9703 were grown under different light: white light (black line), green filtered (green line) or red filtered (red line) light, until the cultures reached the stationary phase. Strain CCY9703 shows a decreased PE and an increased PC absorption under red light. Under green light CCY9703 shows some increase in PE absorption and a decrease of PC absorption.

**Figure 3SM.** IGS between the *cpcB* and *cpcA* genes. In the alignment the different signatures for the distinct endemic and cosmopolitan clusters of *Pseudanabaena* spp. strains are indicated with different colors.

**Table 1SM.** BLAST results of the sequences of 5 loci from *Pseudanabaena* strains used in this study

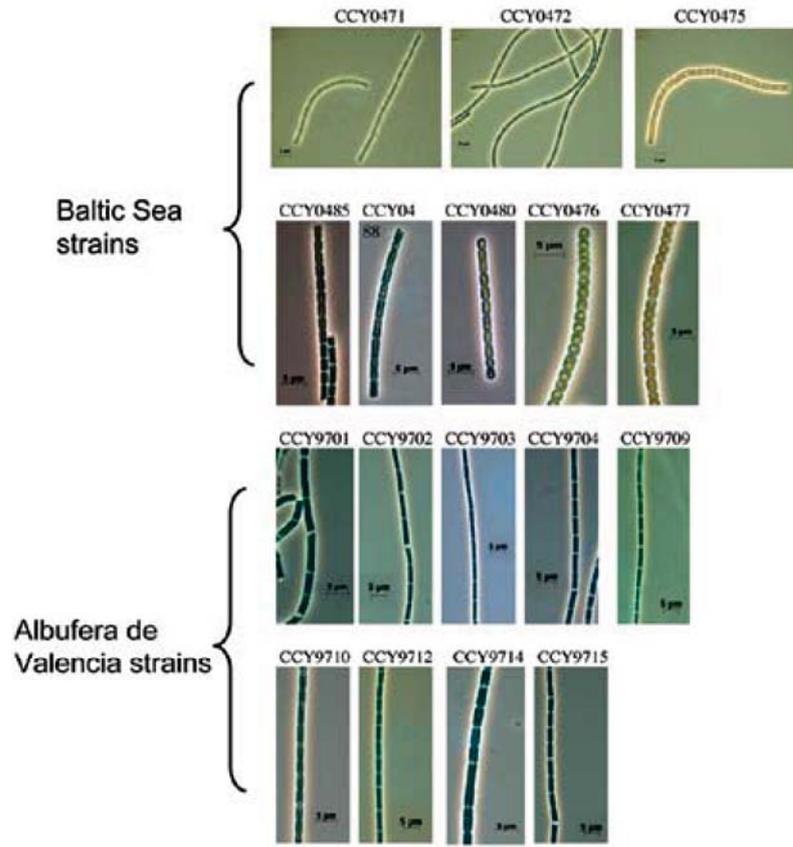


Figure 1SM

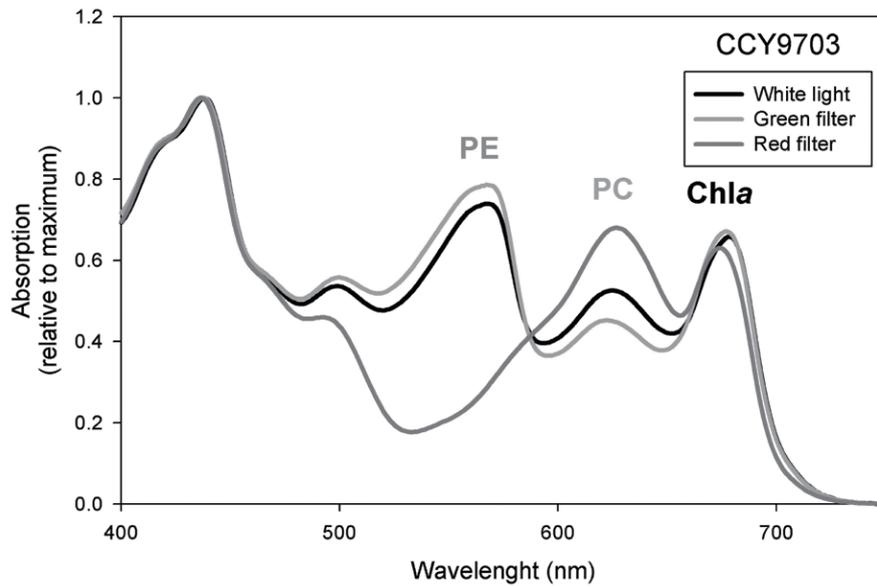


Figure 2SM

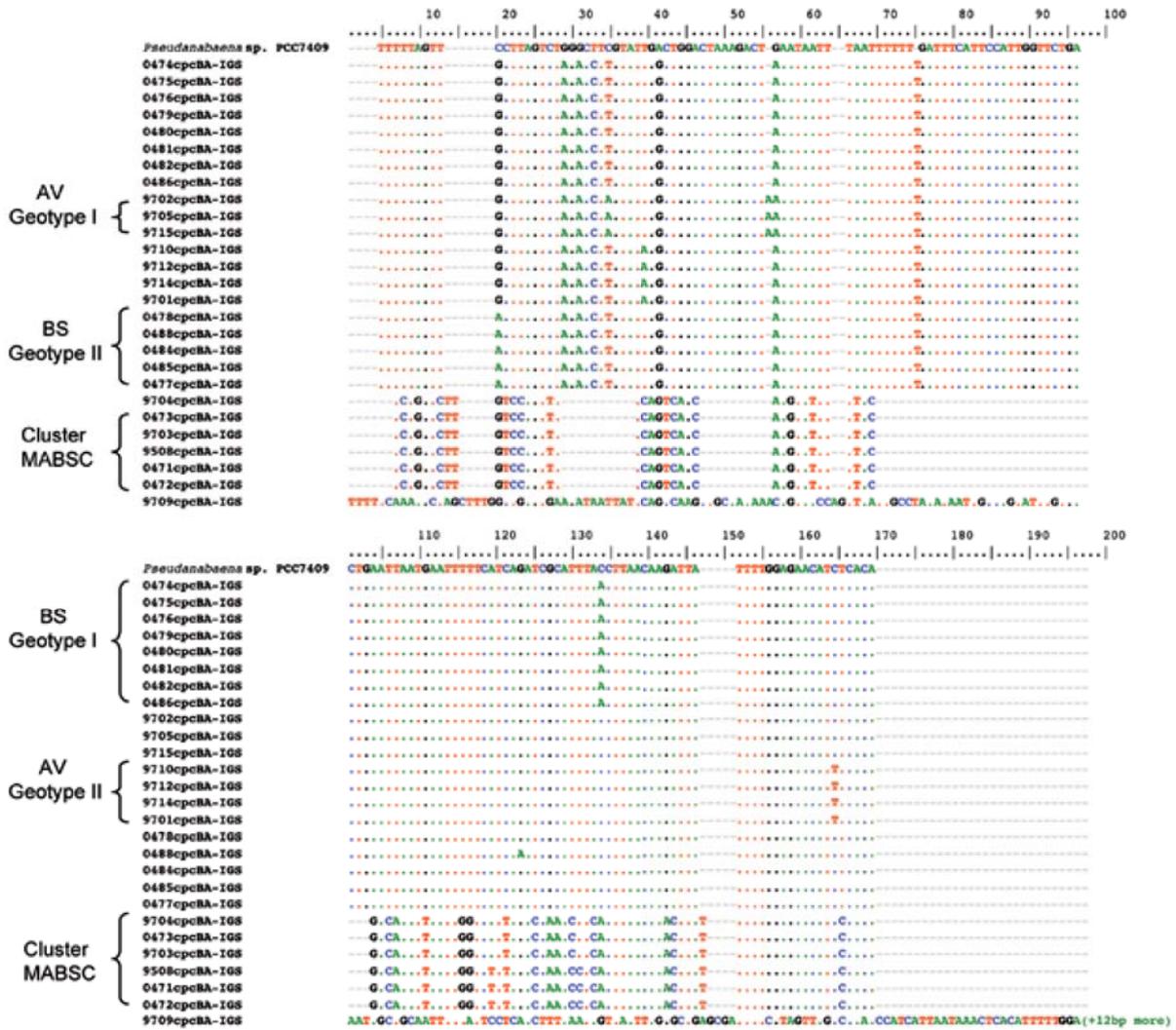


Figure 3SM

**Table 1SM. Blast results of the sequencing of 5 loci from *Pseudanabaena* strains used in this study.**

Strain Name	16S rRNA gene (Bact1055F)		23S rRNA gene (129F primer)		ITS1 (Intergenic Spacer 1)		cpcBA genes plus IGS		
	bp	Closest relative(%) to Genbank /Acc. Nr.	bp	Closest relative(%) to Genbank /Acc. Nr.	bp	Closest relative(%) to Genbank /Acc. Nr.	bp	Closest relative(%) similarity to Genbank /Acc. Nr.	
CCY9702	633	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	323	<i>Limnothrix</i> sp. MR1 (99%)/AJ580008	518	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	497	<i>Limnothrix</i> sp. MR1 (96%)/AJ580008	
CCY9703	553	<i>Pseudanabaena</i> PCC 6301(92%)/AP008231*	458	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	510	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	536	<i>Limnothrix</i> sp. MR1 (99%)/AJ580008	
CCY9704	459	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	471	<i>Synechococcus</i> elongatus PCC 6301(91%)/AP008231*	461	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	536	<i>Limnothrix</i> sp. MR1 (99%)/AJ580008	
CCY9705	646	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	352	<i>Limnothrix</i> sp. MR1 (98%)/AJ580008	522	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	499	<i>Limnothrix</i> sp. MR1 (96%)/AJ580008	
CCY9709	634	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	455	<i>Synechococcus</i> elongatus PCC 6301(88%)/AP008231	549	<i>Trichodesmium</i> eythraeum IMS101 (92%)/CP000393	531	<i>Limnothrix</i> sp. MR1 (100%)/AJ580008 (77% coverage)	
CCY9710	603	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	468	<i>Synechococcus</i> elongatus PCC 6301(91%)/AP008231*	576	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	489	<i>Limnothrix</i> sp. MR1 (99%)/AJ580008	
CCY9712	668	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	471	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	582	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	495	<i>Limnothrix</i> sp. MR1 (96%)/AJ580008	
CCY9714	642	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	472	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	597	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	495	<i>Limnothrix</i> sp. MR1 (96%)/AJ580008	
CCY9715	672	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	472	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	612	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	499	<i>Limnothrix</i> sp. MR1 (96%)/AJ580008	
CCY9708	640	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	450	<i>Synechococcus</i> elongatus PCC 6301(91%)/AP008231*	520	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	536	<i>Limnothrix</i> sp. MR1 (99%)/AJ580008 (65% coverage)	
CCY0471	650	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	473	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	566	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	536	<i>Limnothrix</i> sp. MR1 (99%)/AJ580008 (65% coverage)	
CCY0472	538	<i>Pseudanabaena</i> PCC6903 (100%)(partial seq 322bp)	473	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	542	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	447	<i>Limnothrix</i> sp. MR1 (99%)/AJ580008 (82% coverage)	
CCY0473	477	<i>Pseudanabaena</i> PCC6903 (100%)(partial seq 322bp)	471	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	366	<i>Trichodesmium</i> eythraeum IMS101 (94%)/CP000393	ND, ND	580	<i>Cyanobacterium</i> LS0580 cpcBA (91%)/DC626405 (92% coverage)
CCY0474	508	<i>Pseudanabaena</i> PCC6903 (100%)(partial seq 322bp)	ND, ND	ND, ND	ND, ND	ND, ND	495	<i>Limnothrix</i> sp. MR1 (97%)/AJ580008	
CCY0475	ND, ND	ND, ND	466	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	517	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	ND, ND	608	<i>Pseudanabaena</i> sp. PCC 7409 cpcBA (96%)/M99426
CCY0476	568	<i>Pseudanabaena</i> PCC6903 (100%)(partial seq 322bp)	471	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	615	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	495	<i>Limnothrix</i> sp. MR1 (97%)/AJ580008	
CCY0477	523	<i>Pseudanabaena</i> PCC6903 (100%)(partial seq 322bp)	336	<i>Limnothrix</i> sp. MR1 (100%)/AJ580008	475	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	405	<i>Limnothrix</i> sp. MR1 (96%)/AJ580008	
CCY0478	525	<i>Pseudanabaena</i> PCC6903 (100%)(partial seq 322bp)	ND, ND	ND, ND	ND, ND	ND, ND	495	<i>Limnothrix</i> sp. MR1 (97%)/AJ580008	
CCY0479	543	<i>Pseudanabaena</i> PCC6903 (100%)(partial seq 322bp)	456	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	517	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	444	<i>Limnothrix</i> sp. MR1 (96%)/AJ580008	
CCY0480	611	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	472	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	613	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	489	<i>Limnothrix</i> sp. MR1 (97%)/AJ580008	
CCY0481	625	<i>Pseudanabaena</i> PCC6903 (100%)(partial seq 322bp)	473	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	565	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	406	<i>Limnothrix</i> sp. MR1 (96%)/AJ580008	
CCY0482	666	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	472	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	462	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	482	<i>Limnothrix</i> sp. MR1 (97%)/AJ580008	
CCY0483	ND, ND	ND, ND	376	<i>Limnothrix</i> sp. MR1 (100%)/AJ580008	504	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	ND, ND	ND, ND	
CCY0484	621	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	473	<i>Synechococcus</i> elongatus PCC 6301(92%)/AP008231*	449	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	406	<i>Limnothrix</i> sp. MR1 (96%)/AJ580008	
CCY0485	556	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	351	<i>Limnothrix</i> sp. MR1 (98%)/AJ580008	554	<i>Trichodesmium</i> eythraeum IMS101 (91%)/CP000393	492	<i>Limnothrix</i> sp. MR1 (97%)/AJ580008	
CCY0486	539	<i>Pseudanabaena</i> sp. 1tu24s9 (99%)(partial seq 455bp)	354	<i>Limnothrix</i> sp. MR1 (98%)/AJ580008	478	<i>Trichodesmium</i> eythraeum IMS101 (93%)/CP000393	494	<i>Limnothrix</i> sp. MR1 (97%)/AJ580008	
CCY0488	441	<i>Pseudanabaena</i> PCC6903 (100%)(partial seq 322bp)	ND, ND	ND, ND	ND, ND	ND, ND	401	<i>Limnothrix</i> sp. MR1 (97%)/AJ580008	
								608	<i>Pseudanabaena</i> sp. PCC 7409 cpcBA (96%)/M99426

\* These strains also shown between 99 and 100% identity to *Limnothrix* sp. MR1 (45-47% coverage)  
ND: not determinate

