



UvA-DARE (Digital Academic Repository)

Colourful coexistence : a new solution to the plankton paradox

Stomp, M.

Publication date
2008

[Link to publication](#)

Citation for published version (APA):

Stomp, M. (2008). *Colourful coexistence : a new solution to the plankton paradox*. [Thesis, fully internal, Universiteit van Amsterdam].

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 3

Diversity and phylogeny of Baltic Sea picocyanobacteria inferred from their ITS and phycobiliprotein operons

Abstract

Picocyanobacteria of the genus *Synechococcus* span a range of different colours, from red strains rich in phycoerythrin (PE) to green strains rich in phycocyanin (PC). Here, we show that coexistence of red and green picocyanobacteria in the Baltic Sea is widespread. The diversity and phylogeny of red and green picocyanobacteria was analysed using three different genes: 16S rRNA-ITS, the *cpeBA* operon of the red PE pigment, and the *cpcBA* operon of the green PC pigment. Sequencing of 209 clones showed that Baltic Sea picocyanobacteria exhibit high levels of microdiversity. The partial nucleotide sequences of the *cpcBA* and *cpeBA* operons from the clone libraries of the Baltic Sea revealed two distinct phylogenetic clades: one clade containing mainly sequences from cultured PC-rich picocyanobacteria, while the other contains only sequences from cultivated PE-rich strains. A third clade of phycourobilin (PUB) containing strains of PE-rich *Synechococcus* spp. did not contain sequences from the Baltic Sea clone libraries. These findings differ from previously published phylogenies based on 16S rRNA gene analysis. Our data suggest that, in terms of their pigmentation, *Synechococcus* spp. represent three different lineages occupying different ecological niches in the underwater light spectrum. Strains from different lineages can coexist in light environments that overlap with their light absorption spectra.

This chapter is based on the paper: Haverkamp T, SG Acinas, M Doeleman, M Stomp, J Huisman, LJ Stal (2008) Diversity and phylogeny of Baltic Sea picocyanobacteria inferred from their ITS and phycobiliprotein operons. *Environmental Microbiology* **10**: 174-188.

Introduction

Picocyanobacteria of the *Synechococcus* group span a range of different colours, depending on their pigment composition (Wood 1985; Olson *et al.* 1990; Pick 1991; Vörös *et al.* 1998; Stomp *et al.* 2007). Picocyanobacteria with high concentrations of the pigment phycoerythrin (PE) absorb green light effectively, and have a red appearance. Picocyanobacteria with high concentrations of phycocyanin (PC) absorb red light effectively, and have a blue-green colour. Recent competition models and laboratory experiments showed that red picocyanobacteria win the competition in green light, green picocyanobacteria win in red light, while red and green picocyanobacteria can coexist in white light by partitioning of the light spectrum (Stomp *et al.* 2004). This matches their distribution patterns. Red picocyanobacteria are dominant components of the *Synechococcus* group in open ocean waters (Li *et al.* 1983; Platt *et al.* 1983; Campbell & Carpenter 1987; Campbell & Vaulot 1993), where green and particularly blue light penetrate deeply into the water column. Moreover, red picocyanobacteria can have two different bilin pigments known as phycoerythrobilin (PEB) and phycourobilin (PUB), which both bind to the apoprotein phycoerythrin. The absorption peak of PUB is shifted slightly further to the blue part of the spectrum, and picocyanobacteria with a high PUB/PEB ratio are typically dominant in oligotrophic regions of the oceans where blue light prevails (Olson *et al.* 1990; Wood *et al.* 1998; Toledo *et al.* 1999). In addition, some strains are able to modify their pigmentation through the synthesis of PE with two alternative chromophores, PEB and PUB (Type IV CA; Everroad *et al.* 2006). Green picocyanobacteria dominate in turbid waters, where red light prevails (Stomp *et al.* 2007). Coexistence of red and green picocyanobacteria can be found in waters of intermediate colouration, including coastal seas and many freshwater lakes (Pick 1991; Vörös *et al.* 1998; Murrell and Loes 2004; Katano *et al.* 2005; Mózes *et al.* 2006; Stomp *et al.* 2007).

The genus *Synechococcus* is polyphyletic. Several clusters have been identified, based on photosynthetic pigmentation, nitrogen requirements, motility and salinity (Herdman *et al.* 2001). In marine environments, *Synechococcus* spp. are dominated by members of cluster 5. *Synechococcus* cluster 5 is divided in two sub-clusters, 5.1 and 5.2. Both sub-clusters consist of isolates from the ocean as well as from coastal origin. Members of cluster 5.1 typically have a red colour. They produce PE as their main photosynthetic pigment, have a GC content between 55-62%, and require elevated salt levels for growth. In contrast, members of cluster 5.2 have a green appearance. They produce the pigment PC but lack PE, have a GC content between 63- 66%, and are often able to grow without elevated salt requirements (Herdman *et al.* 2001).

Freshwater picocyanobacteria are often assigned to *Cyanobium*, a genus closely related to *Synechococcus*. *Cyanobium* is only known from freshwater and brackish environments (Crosbie *et al.* 2003; Ernst *et al.* 2003). It contains PC as its main photosynthetic pigment and possesses a high GC content (66-71%). *Cyanobium* is composed of clusters that are distinguished by salt-tolerance and GC content (Herdman *et al.* 2001).

The phylogenetic tree of picocyanobacteria is not always consistent with their pigmentation type. Some strains isolated from marine and freshwater environments produce PE, but are related to the cluster according to sequence information of their 16S rRNA gene, the ribosomal internally transcribed spacer (ITS) region, and the *rpoC1* gene (Crosbie *et al.* 2003; Ernst *et al.* 2003; Everroad & Wood 2006). Conversely, most members of *Synechococcus* cluster 5.1 are rich in PE, but PC-rich isolates were obtained from the Red Sea. Although the genomic GC content of one of these isolates, strain RS9917, (64%) is within the range of *Cyanobium*, it is unknown whether this is also the case for the other strains of that clade (VIII) of cluster 5.1 (Fuller *et al.* 2003).



Figure 3.1 The sampling stations S298, S300, S314 and S320 along the East-West transect from the Gulf of Finland to the Baltic Sea during the CYANO-cruise in 2004.

Here, we studied natural communities of picocyanobacteria from the Baltic Sea by constructing clone libraries of partial sequences of the 16S rRNA-ITS, *cpeBA* and *cpcBA* operons. The latter two encode for the pigments PE and PC, respectively. Earlier studies suggested that the phylogeny of *cpcBA* of freshwater picocyanobacteria correlated with pigmentation (Neilan *et al.* 1995; Robertson *et al.* 2001; Crosbie *et al.* 2003). Our results demonstrate that a phylogeny based on the operons encoding for phycocyanin and phycoerythrin in picocyanobacteria differs from earlier phylogenies based on the 16S rRNA-ITS operon.

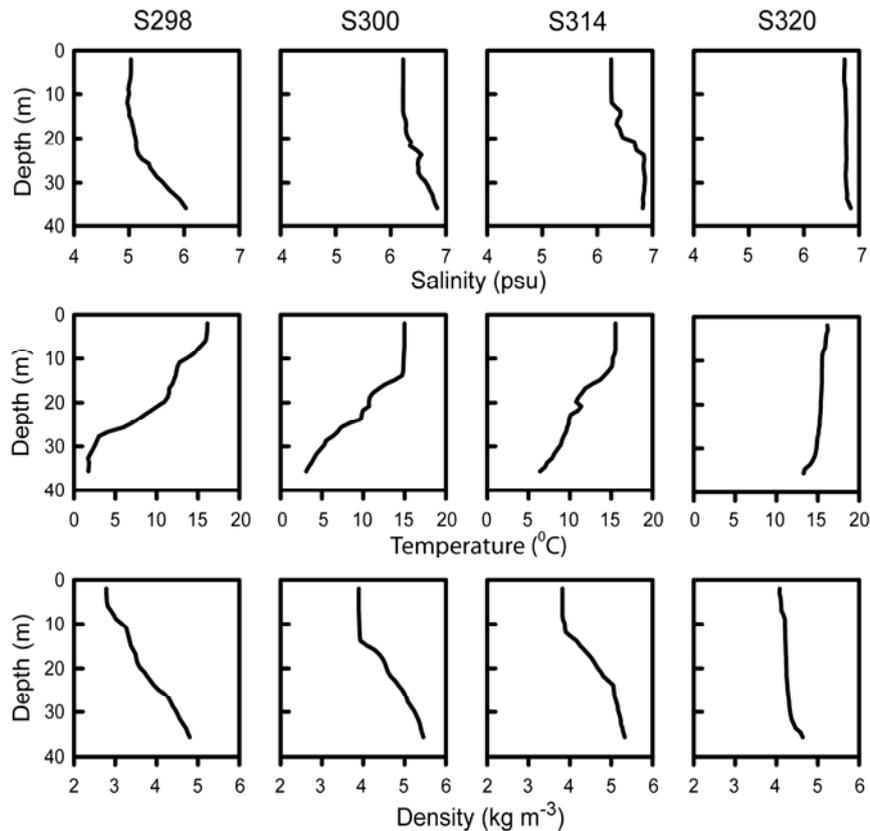


Figure 3.2 Vertical profiles of salinity, temperature, and density at stations S298, S300, S314 and S320

Results

Environmental conditions

Stratification: In July 2004, water was sampled at 4 stations in the Gulf of Finland and Baltic Sea proper (Figure 3.1). Station S298 had a triple thermal stratification at 5, 10 and 20 m depth (Figure 3.2). The density profile of this station revealed that the upper 5 m was well mixed, while density gradually increased with depth below this shallow surface-mixed layer. Station S300 showed a clear surface-mixed layer of ~15 m depth. Station S314 had a slightly shallower surface-mixed layer, with a thermocline and pycnocline at 10-12 m depth. Both stations S300 and S314 had a subtle secondary stratification at ~21 m depth. Station S320 was not stratified, but showed nearly homogeneous vertical profiles of salinity, temperature, and density up to 30 m depth (Figure 3.2).

Underwater light spectra: The underwater light spectrum of natural waters largely depends on light attenuation by water itself, by the ‘background turbidity’ caused by dissolved organic

matter (known as gilvin in the optics literature) and inanimate suspended particles (tripton, like sediment and detritus), and by phytoplankton species present in the water column (Kirk 1994). Water absorbs strongly in the red wavelengths, whereas gilvin and tripton are responsible for rapid attenuation of blue wavelengths. In the Baltic Sea, light absorption in the blue and the red end of the spectrum is of a similar magnitude. At all 4 stations, this yielded an underwater light spectrum that narrowed to green wavelengths with increasing depth (Figure 3.3a).

The light absorption spectra of a red and a green strain of Baltic Sea picocyanobacteria are depicted in Figure 3.3b as an example to illustrate how they are tuned to the underwater light spectrum. PE-rich strains have an absorption peak at ~ 560 nm, and hence absorb green light effectively. PC-rich strains have an absorption peak at ~ 625 nm, and absorb orange-red light effectively. Chlorophyll peaks were also clearly visible in the absorption spectra at 440 nm (Soret band) and 680 nm.

We found euphotic depths of 10.5 m at station S298, 15.3 m at station S300, and 20.3 m at stations S314 and S320, where the euphotic depth is defined as the depth at which the irradiance (PAR, 400-700 nm) equals 1% of the surface irradiance. Hence, the background turbidity of the surface water decreased from the Eastern towards the Western part of the Gulf of Finland.

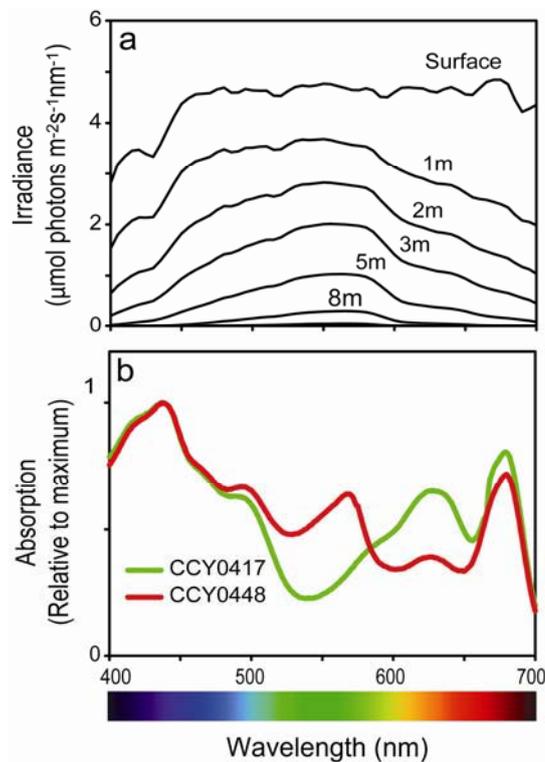


Figure 3.3 Comparison of the underwater light spectrum and the light absorption spectra of PE-rich and PC-rich picocyanobacteria. (a) Underwater light spectra measured at station S320 in the Gulf of Finland (Baltic Sea). The spectrum narrows to the green waveband with increasing depth. Underwater light spectra at the three other stations were similar. (b) Absorption spectra of the PC-rich strain CCY0417 and the PE-rich strain CCY0448 isolated from the Gulf of Finland (Baltic Sea). Absorption spectra are scaled to their maximum value.

Nutrients: Dissolved inorganic nitrogen and phosphorus were measured in water samples from the surface (0 m) and from 30 m depth (Table 3.1). At all stations, nitrogen and phosphorus concentrations were lower at the surface than at depth. Nitrate and nitrite concentrations at the surface were at or below the detection limit of 0.01 μM . At station S320, the phosphorus concentration at the surface was also below the detection limit. At all stations, the N:P ratios were well below the Redfield ratio of 16 (Table 3.1), indicating that nitrogen was relatively more limiting for phytoplankton growth than phosphorus.

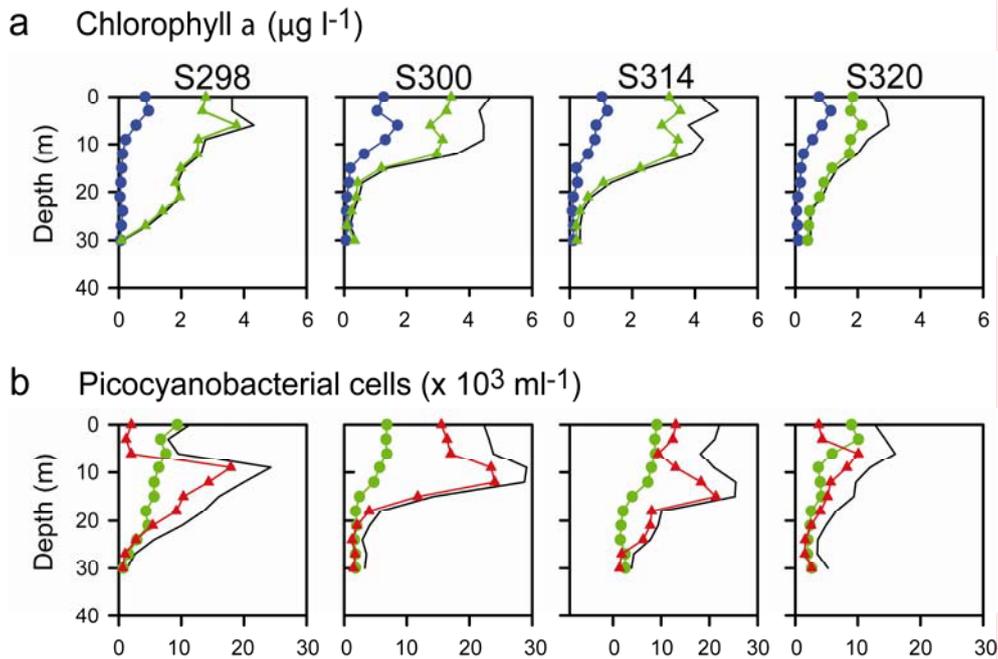


Figure 3.4 Vertical profiles of chlorophyll *a* and picocyanobacteria at stations S298, S300, S314, and S320. **(a)** Concentration of chlorophyll *a* in the large size fraction (blue dots; $> 20 \mu\text{m}$) and in the small size fraction (green triangles; $< 20 \mu\text{m}$). Total concentration of chlorophyll *a* is shown in black. **(b)** Concentration of PC-rich picocyanobacteria (green dots) and PE-rich picocyanobacteria (red triangles). In black is shown the total number of picocyanobacterial cells.

Distribution of picocyanobacteria

Chlorophyll *a* was measured in two size fractions, a small size fraction ($< 20 \mu\text{m}$) and a large size fraction ($> 20 \mu\text{m}$). Microscopic examination indicated that the small size fraction contained mainly picocyanobacteria ($< 2 \mu\text{m}$) and also small filaments of *Pseudanabaena* spp., consistent with earlier studies (Albertano *et al.* 1997; Stal & Walsby 2000; Stal *et al.* 2003). The large size fraction was dominated by the filamentous, N_2 -fixing cyanobacteria *Nodularia spumigena*, *Anabaena* spp. and *Aphanizomenon flos-aquae*, which were mainly concentrated in the upper 10 m (Figure 3.4). Picocyanobacteria were mainly distributed over the upper 15-20 m at

stations S300, S314 and S320 and even down to 30 m at station S298. The small size fraction represented 70-80% of the total chlorophyll *a* in the upper 10 m, and even more than 90% of the total chlorophyll *a* below 10 m (Figure 3.4). Red and green picocyanobacteria were counted by flow cytometry, on the basis of their size and pigment composition. The depth distributions revealed that red and green picocyanobacteria coexisted throughout the upper 30 m (Figure 3.4). The cell numbers of the green picocyanobacteria showed a gradual decline with depth, while the red picocyanobacteria formed a subsurface maximum. At stations S298, S300 and S314, the subsurface maximum of the red picocyanobacteria was at the euphotic depth. At station S320, which lacked a clear stratification pattern (Figure 3.2), the subsurface maximum at ~8 m was less pronounced (Figure 3.4).

Table 3.1 Nutrient concentration ($\mu\text{mol L}^{-1}$) at the sampling stations.

Station	PO_4^{3-}		NO_3^-		NO_2^-		NH_4^+		N:P	
	0m	30m	0m	30m	0m	30m	0m	30m	0m	30m
S298	0.02	0.28*	0	0.01*	0	0.01*	0.18	0.17*	9	0.68*
S300	0.18	0.85	0	1.11	0	0.29	0.21	0.86	1.17	2.66
S314	0.09	0.32	0	0.32	0	0.08	0.11	1.56	1.22	6.13
S320	0	0.19	0.03	0.09	0	0	0.07	1.3	n.d.	7.32

*At station S298, deep samples were from 20 m instead of 30m.

The 16S rRNA and ITS region

The diversity of picocyanobacteria was assessed by sequencing environmental clone libraries containing PCR fragments with a part of the 16S rRNA gene and the internally transcribed spacer between the 16S and 23S rRNA genes (ITS). At all 4 stations, samples were taken at 3 and 12 m depth, where both PC-rich and PE-rich picocyanobacteria were abundant (Figure 3.4). The samples were size fractionated, to separate the small cyanobacteria ($< 20 \mu\text{m}$) from the larger phytoplankton. This yielded a total of 8 samples, from which DNA was extracted and PCR amplified using oligonucleotide primers specific for cyanobacteria. We sequenced the last 400 bases of the 16S rRNA gene and the complete ITS of 74 clones, and compared these sequences against existing databases (NCBI, RDP-II) (Table S1 and Figure S1 in the Online Supplementary Material). One clone appeared to be from the filamentous heterocystous cyanobacterium *Anabaena flos-aquae* (99% similarity to the 16S rRNA sequence; AJ630422), and was therefore not further considered.

The vast majority of clones (65 out of 74) exhibited high sequence similarity (96 to 99%) to several closely related *Synechococcus* strains (LM94, BO8807 and *S. rubescens*), which belong to freshwater group B (Crosbie *et al.* 2003; Ernst *et al.* 2003, Table S1 and Figure S1 in the Online Supplementary Material). This is consistent with earlier studies, which have shown

that strains of group B are more than 99% similar at the 16S-rRNA level (Crosbie *et al.* 2003), and more than 95% similar at the ITS sequence (Ernst *et al.* 2003). The remaining clones displayed high sequence similarity (96 to 98%) to other freshwater *Synechococcus* strains (Table S1 and Figure S1 in Online Supplementary Material). One of our clones sequences (TH320-12-6) had a 99% similarity to the 16S rRNA gene of *Synechococcus* strain MH305 (Crosbie *et al.* 2003). The ITS sequence of this clone was completely disparate from other clones, except for the tRNA genes. The position of clone TH320-12-6 in our phylogenetic analysis confirms this by placing the sequence close to the root of the tree with low bootstrap support (Figure S1 in Online Supplementary Material). We observed large variations in ITS length and GC content in our clone libraries, consistent with earlier studies (Laloui *et al.* 2002; Rocap *et al.* 2002; Ernst *et al.* 2003; Chen *et al.* 2006).

Comparison of the clone libraries from 3 m and 12 m depth, using the program Web-Libshuff (Singleton *et al.* 2001), revealed that there was no significant difference between the libraries obtained from the two sampling depths ($P > 0.05$). We therefore assumed that the libraries from 3 m and 12 m depth have the same composition, and they were lumped in our diversity analysis. The diversity in the clone libraries was analysed using the program DOTUR that calculates several diversity estimators and can be used to create rarefaction curves and similarity plots (Schloss & Handelsman 2005). Rarefaction was used to determine the diversity structure within the 16S rRNA gene - ITS clone library (Figure 3.5a, Table 3.2). These results indicate a high degree of microdiversity in our clone library, suggesting that many of the sequences belong to the same or closely related “species”. When the similarity was further reduced, the number of OTUs continued to decrease until all clones merged into a single OTU at 73% similarity (Figure 3.5b).

Because the ITS region is highly variable, we also tested the diversity within our library by using only the sequences encoding part of the 16S rRNA gene (487 bp). This revealed that 68% of the partial 16S rRNA sequences fall into the 99% clusters (Table 3.2). Several diversity estimators were calculated, such as the Shannon-Weaver and Simpson diversity indices, Good’s Coverage, and the Chao and ACE richness estimates (Good 1953; Chao & Lee 1992; Magurran 1988). Assuming a 99% similarity criterion, the Chao and ACE richness estimates indicated a species richness of 37 and 36, respectively (Table 3.2).

The phycocyanin operon

We included known *pcBA* sequences in our alignment for comparison with 68 clones that we obtained from the Baltic Sea. The lengths of the sequences available in GenBank ranged from 320 bp to almost 500 bp (excluding the intergenic spacer, IGS), complicating phylogenetic analysis of the *pcBA* genes. We decided to remove sequences shorter than 380 bp (IGS excluded) from our alignment to avoid incorrect topologies (Nei *et al.* 1998; Tamura *et al.* 2004). This approach gave a more robust phylogenetic tree of the *pcBA* gene. Figure 3.6 shows the phylogenetic tree we obtained for the partial *pcBA* gene sequences. Many of the picocyanobacteria of the Baltic Sea are closely related to the known groups A, B, H, and I

(Robertson *et al.* 2001; Crosbie *et al.* 2003; Table S2 in Online Supplementary Material), confirming results based on the 16S rRNA-ITS operon. The Baltic Sea Group 3 is probably a novel taxon within the picocyanobacteria, since these sequences form a monophyletic group that separates with a long branch and with good bootstrap support from the other sequences. We cannot exclude that the other Baltic Sea groups might also represent unique groups although the branch lengths separating these sequences from known sequences are small. Hence, this might as well represent microdiversity between the clusters.

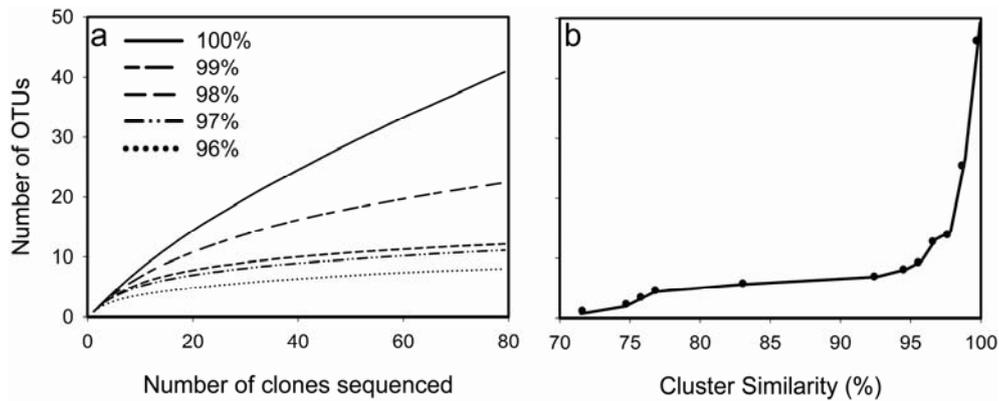


Figure 3.5 Diversity patterns of the Baltic Sea picocyanobacteria using 16S rRNA-ITS sequences. (a) Rarefaction curves of the number of observed OTUs at 100, 99, 98, 97 and 96 % similarity cut-offs. (b) Number of OTUs plotted against different cluster cut-off values in 1.0% increments for sequences grouped into similarity clusters.

There are also some striking differences between the *cpcBA* phylogeny and the existing 16S rRNA phylogenies (Crosbie *et al.* 2003; Fuller *et al.* 2003). First, the *cpcBA* phylogeny separated most picocyanobacteria with a green phenotype from picocyanobacteria with a red phenotype, although there were a few red strains within the green clusters (Figure 3.6; Figure S2 in Online Supplementary Material). Second, in contrast to the 16S rRNA phylogeny, in the *cpcBA* phylogeny green picocyanobacteria isolated from marine environments (e.g., strains RS9917 and WH5701) clustered with green freshwater picocyanobacteria. Third, the green *Cyanobium* strain CCY9201 (previously known as BS4) and the red *Cyanobium* strain CCY9202 (previously known as BS5), which were nearly identical according to the 16S rRNA-ITS phylogeny (Crosbie *et al.* 2003; Ernst *et al.* 2003), were completely separated in the *cpcBA* phylogeny. Fourth, the *cpcBA* phylogeny revealed that phycourobilin (PUB)-producing picocyanobacteria form a distinct cluster within the red picocyanobacteria.

The *cpcBA* phylogeny pointed at a close correlation between pigment phenotype and GC content (Figure 3.6). PC-rich isolates had GC-contents higher than 60%, while most PE-rich isolates had GC contents less than 60% although there were a few exceptions. The difference in GC content between the *cpcBA* sequences was mainly caused by higher GC content at the third codon position, resulting in synonymous mutations in most of the codons investigated. Likewise, the *cpcBA* phylogeny pointed at a close correlation between pigment

phenotype and the effective number of codons (ENC). The ENC number represents a measure for the codon usage bias (Comeron & Aguade 1998). An ENC number of 20 means that only one codon is used for each amino acid, while an ENC number of 61 indicates that all codons are used equally often and in that case there is no bias in codon usage (Wright 1990). PC-rich isolates had a low ENC number in the range of 23-32, while almost all PE-rich isolates had a high ENC number ranging from 33 to 45 (Figure 3.6). Interestingly, PE-rich strains with a GC content exceeding 60% and an ENC-number below 33 clustered with the PC-rich strains.

The phycoerythrin operon

PCR amplification of the *cpeBA* operon encoding the pigment phycoerythrin resulted in 68 clones (for primers see Everroad & Wood 2006). The number of *cpeBA* sequences available in existing databases such as GenBank was limited to 37 full-length sequences of different cyanobacteria and red algae. BLASTn searches using the nucleotide sequences of all our *cpeBA* clones returned only one of two different top hits, marine *Synechococcus* strains WH7803 (X72961) and WH8102 (BX569694) (Table S3 in Online Supplementary Material). Our sequences showed only 81% to 90% similarity with these two sequences. BLASTp searches using our *cpeBA* sequences as query were done using the CPE-A and the CPE-B protein coding sequences. Both fragments showed the highest similarity with the CPE-A (range 86 to 93%) and CPE-B (91 to 97%) proteins from the marine *Synechococcus* strain WH7805 (Table S3 in Online Supplementary Material).

We performed a phylogenetic analysis using our Baltic Sea partial *cpeBA* nucleotide sequences and those recovered from existing databases. Analysis of the phenotypes revealed that all cultured strains within the *cpeBA* phylogeny were PE-rich strains with a GC content between 53 and 63 % and a ENC number ranging from 30 to 45 (Figure S3 in Online Supplementary Material). The *cpeBA* phylogeny yielded two major groups (Figure 3.7). Again these two groups matched the pigmentation of picocyanobacteria. The first group was formed by *cpeBA* genes from freshwater and marine *Synechococcus* strains producing PEB only, while the second group consisted of marine strains producing both PUB and PEB. This topology was consistent with the *cpcBA* phylogeny, where the PUB-producing picocyanobacteria formed a distinct cluster (Figure 3.6). All *cpeBA* sequences that we obtained from the Baltic Sea were constrained within the PEB group (Figure 3.7). These Baltic Sea sequences were separated into two major clades, one clade comprising the clusters 1 and 2, and the other clade formed by clusters 3 and 4. Comparison of the overall similarity at the amino acid level showed that the similarity within each of these two clades is more than 98%, while the similarity between the two clades is only 86.6%.

The diversity estimators showed that the diversity in the *cpcBA* and *cpeBA* library is low compared to the 16S rRNA-ITS library (Table 3.2). This might be attributed to inherent differences in variability between these libraries, but also to differences in length between the 16S rRNA-ITS and the *cpcBA* and *cpeBA* sequences. The number of OTUs was rather similar

for the *cpcBA* and *cpeBA* operons. According to the Chao-1 and ACE richness estimates and the Shannon and Simpson diversity indices, however, the diversity at the *cpeBA* operon encoding for phycoerythrin was slightly higher than the diversity at the *cpcBA* operon encoding for phycocyanin (Table 3.2).

Table 3.2 Diversity estimators for the clone libraries of the 16S rRNA-ITS, 16S rRNA, *cpcBA* operon and *cpeBA* operon, with and without intergenic spacers. The number of Operational Taxonomic Units (OTUs) is shown at 100%, 99% and 97% similarity cut-off values. The coverage is expressed as defined by Good (1953). The Chao-1 richness, ACE richness, Shannon diversity index and Simpson diversity index use 99% similarity cut-off values. Numbers within parentheses for the Chao-1 and ACE richness estimators are 95% confidence intervals.

Gene	Number of Clones	OTUs (100/99/97%)	Good's Coverage (%)	Chao-1	S-ACE	Shannon index	Simpson index (1/D)
16S-ITS complete	73	40/22/11	86.3	37 (26-86)	36 (26-66)	2.64	10.90
16S without ITS	73	19/6/1	95.9	9 (6-31)	14 (7-79)	0.89	1.85
<i>cpcBA</i> operon	68	24/11/8	92.7	21 (13-63)	16 (12-37)	1.52	2.76
<i>cpcBA</i> without IGS	68	20/10/8	94.1	16 (11-48)	13 (11-30)	1.49	2.75
<i>cpeBA</i> operon	68	24/11/5	91.8	26 (14-79)	28 (14-107)	1.85	5.52
<i>cpeBA</i> without IGS	68	24/12/6	91.8	27 (15-80)	23 (14-70)	2.01	6.66

Discussion

Colourful coexistence of red and green picocyanobacteria

Our results show that PC-rich and PE-rich picocyanobacteria coexist in the Baltic Sea, where they are approximately equally abundant players in the cyanobacterial community (Figure 3.4). This confirms earlier results of Stomp *et al.* (2004, 2007). PC-rich picocyanobacteria were slightly more abundant in the upper 5 m of the water column, while PE-rich picocyanobacteria were more dominant at 5-15 m depth. This vertical distribution matches the underwater light spectrum, since green light penetrates more deeply into the Baltic Sea than red light (Figure 3.3a). Remarkably, the PC-rich and PE-rich picocyanobacteria maintained their vertical distribution even in waters with a nearly homogeneous temperature and density profile (Station S320, Figures 3.2 and 3.4). Since picocyanobacteria lack buoyancy regulation, this

indicates that local growth rates of PE-rich and PC-rich populations exceeded the rate of vertical mixing by hydrodynamic processes (Huisman *et al.* 1999b).

Sequencing of 209 clones revealed that picocyanobacteria of the Baltic Sea exhibit high levels of microdiversity. Approximately 46-54% of the OTUs present in each clone library were constrained at 99% similarity clusters (micro-clusters; Figure 3.5, Table 3.2). Such high levels of microdiversity have also been detected by many previous studies of marine microbial communities and other natural bacterial populations (Acinas *et al.* 2004; Lopez-Lopez *et al.* 2005; Pommier *et al.* 2007; Rusch *et al.* 2007). The high microdiversity of *Synechococcus* spp. genes found in our clone libraries may reflect local adaptive radiation of picocyanobacteria which allows them to proliferate under a wide range of different conditions in the Baltic Sea.

Phylogeny of red and green picocyanobacteria

Our results show that a phylogeny based on the *pcpA* gene (phycocyanin) and *pcrA* gene (phycoerythrin) differs from a phylogeny based on 16S rRNA gene sequences. This is especially clear for the *pcpA* dataset, where clustering of the different phylotypes largely matched the pigment composition of the picocyanobacteria (see also Robertson *et al.* 2001; Crosbie *et al.* 2003). This is exemplified by the green CCY9201 (previously known as BS4) and red CCY9202 (previously known as BS5) strains used in the competition experiments of Stomp *et al.* (2004). On the basis of their ITS sequences, these two strains are more than 99% similar (Ernst *et al.* 2003), whereas their *pcpA* gene sequences are well separated (Figure 3.6), where the green strain clusters in the group of PC-rich picocyanobacteria while the red strain clusters in the group of PE-rich picocyanobacteria (Figure 3.6). The few sequences of red strains that cluster with the *pcpA* operons of green isolates can be explained by horizontal gene transfer (HGT).

Another example is the placement of the PC-rich marine isolate RS9917. This strain forms a distinct cluster with other PC-rich isolates within the marine picocyanobacteria based on the 16S rRNA gene sequences (Fuller *et al.* 2003). According to our phylogenetic analysis, the partial *pcpA* sequences of strain RS9917 clusters with the *pcpA* sequences of PC-rich freshwater picocyanobacteria. This could have been caused by HGT of the *pcpA* operon of a freshwater picocyanobacterium. Likewise, clustering of similar pigmentation types is also evident from the placement of PUB/PEB-producing marine *Synechococcus* in both the *pcpA* and *pcrA* phylogeny. The marine strain WH7805 produces PEB, but in contrast to other PE-rich marine *Synechococcus* strains it is not capable of producing PUB (Fuller *et al.* 2003). In the *pcrA* and *pcpA* phylogenetic trees, strain WH7805 clustered separately from the PUB-producing marine *Synechococcus* strains. Only strains that produce PUB might possess the capacity of chromatic adaptation of type IV. We have not retrieved any sequences in our Baltic Sea clone libraries that are related to PUB-producing picocyanobacteria.

Overall, our phylogenetic analyses extend earlier findings of Robertson *et al.* (2001) and Crosbie *et al.* (2003), who showed that the *pcpA* operon separates PE-rich and PC-rich

picocyanobacterial isolates from freshwater lakes. In our analysis, we included picocyanobacteria from brackish waters and marine ecosystems, and studied not only the *pcBA* operon but also the *peBA* operon. This revealed three distinct groups of picocyanobacteria separated in line with their pigmentation, namely PUB/PEB, PEB, and PC producing strains. All are members of the monophyletic clade formed by *Synechococcus* and *Cyanobium*.

Correlations with GC content and ENC number

Differences in pigmentation in the *pcBA* phylogeny correlated with the ENC number and the GC content of the sequences. PC-rich picocyanobacteria had higher GC contents and lower ENC numbers than PE-rich picocyanobacteria (Figure 3.6). One possible explanation for differences in GC content in PE-rich and PC-rich picocyanobacteria is that it may reflect differences in expression levels of the *pcBA* gene. In fact, highly expressed genes in *Prochlorococcus* strain MED4 had higher GC contents compared to low expressed genes (Banerjee & Ghosh 2006). A PE-rich cyanobacterial phycobilisome has one disk of PC proteins while containing multiple disks of PE proteins. A PC-rich phycobilisome usually has several disks of PC. The higher demand for phycocyanin might require a higher expression level and, hence a higher GC content of the *pcBA* operon. Alternatively, it could also be that the genomes of PC-rich picocyanobacteria have a higher GC-content. We tested this hypothesis by analyzing the GC-content of the protein-coding genes of the genome sequences of *Synechococcus* spp. present in GenBank. This showed that the overall GC-content of the protein-coding genes of the PC-rich *Synechococcus* strains WH5701 and RS9917 is higher compared to those of the PE-rich picocyanobacteria (Table S4 in Online Supplementary Material). This would contradict the theory that higher expression levels cause the higher GC-content in the *pcBA* operons of PC-rich *Synechococcus* spp. It also confirms the placement of RS9917 among the freshwater picocyanobacteria in our phylogenetic analysis and that it is unlikely that this is caused by HGT of phycobiliprotein genes.

Another explanation for the relationship between GC content and pigmentation might come from the environment. Comparative studies suggest that the GC contents of microbial genomes or environmental shotgun libraries vary among habitats of different productivity (Goo *et al.* 2004; Carbone *et al.* 2005; Foerstner *et al.* 2005). For instance, Foerstner *et al.* (2005) observed that the average GC-content of open reading frames (ORFs) from the oligotrophic Sargasso Sea is only 34%, whereas the GC content of ORFs from productive Minnesota soil samples is 61%. These large differences in GC content were not merely an effect of differences in species composition between these two contrasting environments, but remained when the same analysis was focused on phyla present or on genes present in both environments. Extrapolated to the *pcBA* phylogeny, this would mean that the high GC sequences of PC-rich picocyanobacteria come from environments with higher levels of nutrients than the sequences of PE-rich picocyanobacteria that have a lower GC content. This explanation is consistent with the global distribution pattern of picocyanobacteria (e.g.,

Stomp *et al.* 2007), where PC-rich picocyanobacteria dominate in productive lakes and coastal waters while PE-rich picocyanobacteria dominate in the oligotrophic open ocean.

Experimental procedures

Sample collection

Water samples from the Baltic Sea were collected from 12 - 19 July 2004 during a research cruise with the Finnish RV *Aranda*. For the work reported here, we sampled 4 stations (stations S298, S300, S314, S320; Figure 3.1), positioned along an East-West transect from the Gulf of Finland into the Baltic Sea proper (from N 59.1 - 60.0 °N and E 22.2 - 26.2 °E to 59.1 °N 22.2 °E). Samples were taken at 3 m depth intervals from the surface to 30 m depth using a rosette sampler. A Seabird 911 CTD was connected to the rosette sampler, to measure temperature and salinity along these depth profiles. Nutrient concentrations in the water samples were analysed according to standard methods (Grasshoff *et al.* 1983).

Underwater light spectra

Spectra of the incident light and underwater light spectra were measured with a RAMSES-ACC-VIS spectroradiometer (TriOS, Oldenburg, Germany). Light absorption spectra of isolated strains were measured using a Cary 100 Bio equipped with an integrating sphere DRACA-3300, with distilled water as a reference.

Chlorophyll analysis

For chlorophyll *a* analysis, the phytoplankton was divided into two size classes. Total chlorophyll *a* was obtained by filtering 0.5 L on GF/F filters (Whatman, nominal pore size 0.7 µm). Chlorophyll *a* of the large size fraction of phytoplankton was obtained by filtering 1 L on 20 µm nylon mesh (plankton net). Chlorophyll *a* of the small size fraction was calculated as the difference between total chlorophyll *a* and chlorophyll *a* of the large size fraction. This procedure largely discriminates between picoplankton and the larger filamentous cyanobacteria in the Baltic Sea (Stal & Walsby 2000). Chlorophyll *a* was extracted overnight in the dark at room temperature by 96% ethanol and absorption was measured spectrophotometrically at 665 nm. Chlorophyll concentration was calculated using an absorption coefficient of 72.3 ml mg⁻¹ cm⁻¹ (Stal *et al.* 1999).

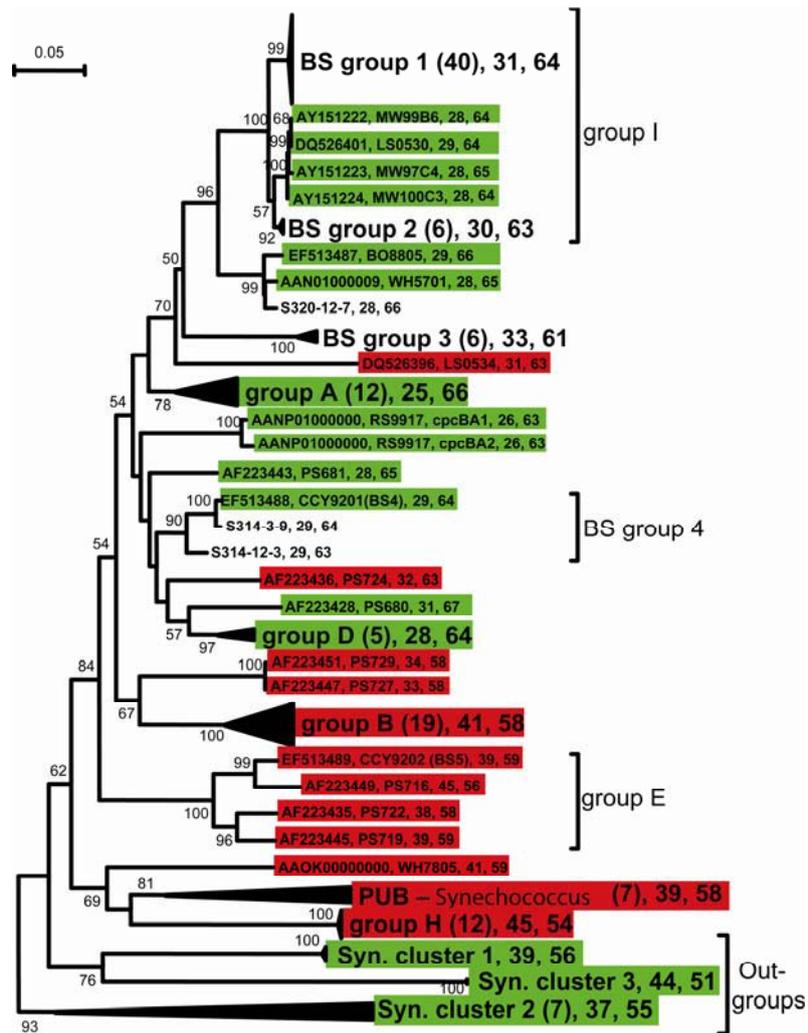


Figure 3.6 Neighbor-joining tree of picocyanobacterial *cpcBA* genes. Clades were condensed to group designations for clarity (Crosbie *et al.* 2003). BS-group designations are assigned to clades formed solely by clone sequences from the Baltic Sea. For condensed groups, the number of *cpcBA* sequences is indicated within brackets. For single sequences, the GenBank accession number and the strain designation are given. For each clade with known isolates, the pigment phenotype is indicated with red (PE-rich) and green (PC-rich). Numbers indicate mean ENC number and mean GC content, resp. The tree was calculated with software MEGA with the neighbor-joining method using the Kimura- 2 parameter model of nucleotide substitution with 1000 replicates (Kumar *et al.* 2004). Bootstrap values (>50%) are shown at the nodes. As out groups were used the *cpcBA* sequences of *Synechococcus* cluster 1 (strains PCC6301, PCC7942, PCC7943), *Synechococcus* cluster 2 (strains PCC6716, PCC6717, *Synechococcus elongates*, JA-2-3b, JA-3-3b), and *Synechococcus* cluster 3 (PCC7002).

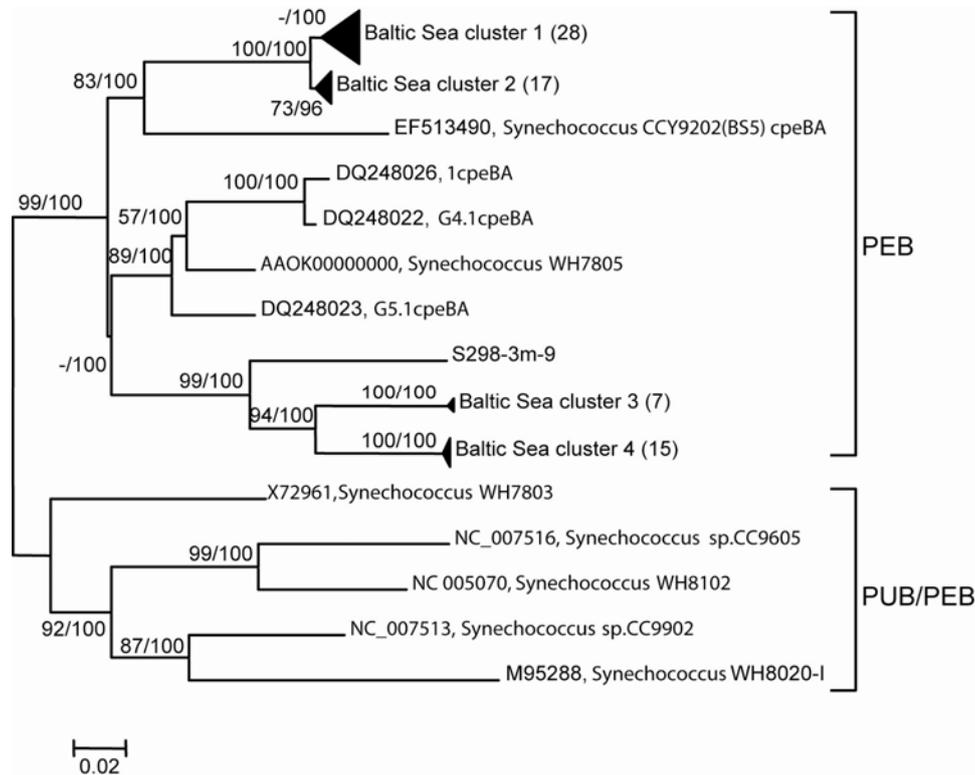


Figure 3.7 Unrooted neighbor-joining tree of picocyanobacterial *cpeBA* genes. Sequences were obtained from the Baltic Sea and from *Synechococcus* strains with sequenced genomes spanning the *cpeBA*-IGS region. Baltic Sea clusters indicate clades formed solely by clone sequences from the Baltic Sea. The number of *cpeBA* clone sequences is indicated within brackets. *Synechococcus* sequences extracted from existing genome sequences or GenBank are in bold. Additional *Synechococcus* sequences from strains used in this study are in italics. The tree revealed that *cpeBA* sequences separated into clades containing PEB and PUB/PEB-producing clades. The Baltic Sea sequences separated into 4 clusters and a single clone (S298-3m-9). Bootstrap values (>50%) based on 1000 replicates are shown at the nodes using distance analysis (first number) and maximum parsimony analyses (second number). A '-' indicates not significant.

Counting red and green picocyanobacteria

The concentrations of red and green picocyanobacteria in the samples were counted by flow cytometry (Jonker *et al.* 1995; Stomp *et al.* 2007), using a Coulter Epics Elite ESP flow cytometer (Beckman Coulter Nederland BV, Mijdrecht, Netherlands) equipped with a green laser (525 nm) and a red laser (670 nm). The flow cytometer distinguished between picocyanobacteria and larger phytoplankton by their size (using side scattering). Red and green picocyanobacteria were distinguished based upon their different fluorescence signals. Cells rich in PE emitted orange light (550-620 nm) when excited by the green laser, whereas cells rich in PC emitted far red light (> 670 nm) when excited by the red laser.

Extraction of nucleic acids

From each station 1 L of seawater from each sampling depth was pre-filtered through 20 µm nylon mesh and collected in polycarbonate bottles that were rinsed by 0.5 M NaOH. The pre-filtered seawater was immediately filtered through 0.2 µm Sterivex filtration units (Millipore) using a peristaltic pump. Subsequently, the Sterivex filters were filled with 2 ml lysis buffer (400 mM NaCl, 20 mM EDTA, 50 mM Tris-HCl [pH 9.0], 0.75 M sucrose) (Massana *et al.* 1997; Moon-van der Staay *et al.* 2001) and stored at -20 °C.

Nucleic acids were extracted as described by Massana *et al.* (1997) with modifications. In brief, lysozyme (final concentration 1 mg ml⁻¹) was added to the Sterivex unit and incubated for 45 min at 37 °C. Subsequently, proteinase-K (final concentration 50 µg ml⁻¹) and sodium dodecyl sulfate (SDS) (1% w/v) were added and incubation was continued overnight at 55 °C. The lysate was recovered from the Sterivex unit by extracting it twice with an equal amount of phenol-chloroform-isoamyl alcohol (25:24:1; pH 8) and once with the same volume of chloroform-isoamyl alcohol (24:1). The extracts were centrifuged (Sigma 4k15 with a swing-out rotor, nr.11156) for 15 min at 1300 rpm and 25 °C. The aqueous phase was transferred to a 15 ml Greiner tube and two volumes of 96% ethanol and 1/10 volume 3 M Na-acetate were added and subsequently incubated for 2 h at -70 °C to precipitate the DNA. Subsequently, the DNA was centrifuged for 20 min at 14000 rpm and 4 °C. The pellet was washed with cold 70% ethanol (-20 °C) and centrifuged for 5 min at 14000 rpm and 4 °C. The supernatant was removed by pipetting and the pellet was air dried. The dry pellet was suspended in 100 µl 10 ml Tris-HCl (pH 8.5). Because the DNA was not PCR grade after this procedure, it was further purified using the Powersoil DNA extraction kit (MoBio Laboratories) following the manufacturer's recommendations.

Primer design

For amplification of part of the 16S rRNA gene and the internal transcribed spacer between the 16S and 23S rRNA genes, we designed oligonucleotide primers that bind to the 5' region of the 23S rRNA sequences of cyanobacteria (Table 3.3). Cyanobacterial 23S rRNA gene sequences were obtained from GenBank and aligned using the Clustal-W program in Bioedit (Thompson *et al.* 1994; Hall 1999). The alignment was imported to Primer Premier software (Premier Biosoft International, version 5.0) and 23S rRNA gene oligonucleotide primers were designed using B1055 as the forward 16S rRNA primer (Singh *et al.* 1998; Zaballos *et al.* 2006; Table 3.3). Primer sequences were checked for their specificity by performing BLASTn searches against the GenBank database.

PCR primers targeting the phycocyanin *pcBA* operons in a wide range of cyanobacteria were available from the literature (Neilan *et al.* 1995; Robertson *et al.* 2001; Crosbie *et al.* 2003). Recently, genome sequences from a variety of picocyanobacteria became available providing the opportunity to design primers that target specifically the *pcBA* genes from *Synechococcus*-like cyanobacteria. Using the Integrated Microbial Genomes database (<http://img.jgi.doe.gov/cgi-bin/pub/main.cgi>), *pcBA* operons were obtained from the

following (un-)finished picocyanobacterial genomes: *Synechococcus* PCC6301 (AP008231), PCC7942 (CP000100), CC9311 (CP000435), CC9605 (CP000110), CC9902 (CP000097), RS9917 (AANP01000000), WH5701 (AANO01000000), WH7805 (AAOK01000000), and WH8102 (BX548020) (Markowitz *et al.* 2006). The *pcpBA* operons M95288 and M95289 from *Synechococcus* strain WH8020 were downloaded from GenBank (Delorimier *et al.* 1993). The full length *pcpBA* operons were aligned in Bioedit using the ClustalW algorithm. The alignment was imported in Primer Premier 5.0 and used to design primers specifically targeting the *pcpBA* genes from the marine cluster B (*Synechococcus* WH5701) (Table 3.3).

PCR and clone library construction

DNA obtained from 3 and 12 m depth of stations S298, S300, S314 and S320 were used to amplify the cyanobacterial 16S rRNA-ITS region, the *pcpBA* and the *pcpBA* operons using the primers listed in Table 3.3. The PCR reaction mixture was composed of 1 μ l of template DNA (1 - 20 ng μ l⁻¹), 2.5 μ l of 10X PCR buffer (Qiagen), 0.5 μ l of 10 mM dNTP's mixture (Roche) and 0.62 units of HotStarTaq DNA polymerase (Qiagen). We added 10 pmol of each forward and reverse primer, except for the 16S rRNA-ITS PCR where 5 pmol was used. Sterile MilliQ grade water was added to a final reaction volume of 25 μ l.

The PCR reactions were run on a GeneAmp System 2700 thermocycler. The program for the 16S-ITS amplification consisted of 15 min hot start at 94 °C; 35 cycles of 1 min at 94 °C; 1 min at 62 °C; and 1 min at 72°C; which was followed by a final elongation step at 72°C for 10 min. For amplification of the *pcpBA* genes the following program was applied: 15 min at 94 °C, 40 cycles of 30 seconds at 94 °C, 30 seconds at 55 °C, and 1.5 min at 72 °C. The final elongation step was 10 min at 72 °C. The same program was used to amplify *pcpBA* except that the elongation step was only 1 min.

PCR-reactions were done in triplicate to decrease variations in amplification (Polz & Cavanaugh 1998). The PCR products of the triplicate reactions were pooled and cloned. Cloning was done using the TOPO TA cloning kit for sequencing (Invitrogen) following the instructions of the manufacturer. For each sample and PCR product 20 clones were picked using sterile toothpicks. The cells were transferred to 200 μ l of sterile LB-Broth and grown overnight. Twenty-five μ l of culture was mixed with 25 μ l of Milli-Q water and heated at 94 °C for 10 minutes. Five μ l of the mixture was used for PCR amplification of the insert using the T3 and T7 primers of the vector. Subsequently, 10 positive PCR reactions were chosen per sample and purified using the DNA Clean & Concentrator (Zymo Research). The DNA concentration was measured using a Nanodrop ND1000 (NanoDrop Technologies) spectrophotometer. The PCR product was sequenced using the Big Dye Terminator v1.1 Cycle sequencing kit (Applied Biosystems) according to the manufacturer's instructions. The clones containing *pcpBA* and *pcpBA* fragments were sequenced using the T3 and T7 primers, while the 16S rRNA-ITS clones were sequenced with the primers B1055, Cya23S-58R2, PITS1 and PITS3 (Table 3.3). Sequencing was done with a 3130 Genetic Analyzer (Applied Biosystems). For each clone, the forward and reverse sequences were manually aligned in

Bioedit and the sequences were checked against GenBank using BLASTn and BLASTp (Altschul *et al.* 1990; McGinnis & Madden 2004). Furthermore, the 16S rRNA clone sequences were compared to the RDP-II database (Cole *et al.* 2005).

Table 3.3 Oligonucleotide primers used in this study.

Primer Name	Target gene	Sequence 5' to 3'	Tm °C	Reference
B1055	16S rRNA	ATGGCTGTCGTCAGCTCGT	66	Zaballos <i>et al.</i> 2006
Cya23S-58r2	23S rRNA	CGTCCTTCATCGCCTCTG	58	This study
PITS 1	ITS	TCAGTTGGTAGAGCGCCTGC	56	Ernst <i>et al.</i> 2003
PITS 3	ITS	GTTAGCGGACTCGAACCGC	65	Ernst <i>et al.</i> 2003
SynpcB-Fw	<i>cpcB</i>	ATGGCTGCTTGCCTGCG	61	This study
SynpcA-Rev	<i>cpcA</i>	ATCTGGGTGGTGTAGGG	50	This study
B3FW	<i>cpeB</i>	TCAAGGAGACCTACATCG	58	Everroad & Wood 2006
SynA1R	<i>cpeA</i>	CAGTAGTTGATCAGRCGCAGGT	64	Everroad & Wood 2006

Diversity calculations and phylogenetic analysis

For the diversity calculations, the clone sequences of the different sampling stations were grouped together. The program DOTUR was used for calculating rarefaction, library coverage, Shannon-Wiener diversity index (H'), Simpson index (D), Chao-1 non-parametric richness estimator and the ACE coverage-based richness estimator (Schloss & Handelsman 2005). Calculations were performed on a Jukes-Cantor corrected distance matrix created with the DNADIST program from the PHYLIP Package (Felsenstein 1989).

Sequences previously identified to be closely related by BLASTn comparison were imported from GenBank into Bioedit and aligned against the clone sequences using ClustalW. Alignments of the 16S rRNA-ITS sequences were done manually in Bioedit by reference of the ITS alignment of the predicted secondary structure models proposed in several papers describing the cyanobacterial ITS sequences (Iteman *et al.* 2000; Laloui *et al.* 2002; Rocap *et al.* 2002; Taton *et al.* 2003). Sequence comparison and phylogenetic analyses were performed using the software MEGA3.1 (Kumar *et al.* 2004). For the 16S rRNA-ITS region the sequences were compared using the neighbor-joining algorithm with Jukes-Cantor correction and 1000 bootstraps. The coding regions of the *cpeBA* and *cpcBA* operon were both used in phylogenetic analyses. Both data sets were separately analysed using the following approach. Phylogenetic analyses were done with the neighbor-joining method as well as with maximum parsimony. Neighbor-joining was performed with the Kimura-2-parameter model for nucleotide evolution with 1000 bootstraps. Maximum parsimony was used with the close-neighbor-interchange

search algorithm with random tree addition using 100 bootstraps. Codon usage in the *pcBA* and *peBA* coding regions was analysed using DnaSP version 4.0 (Rozas *et al.* 2003).

Nucleotide sequence accession numbers

The sequence data reported in this paper have been submitted to the GenBank database under accession numbers: 16S rRNA- ITS clones (EF513279 – EF513350); *pcBA* clones (EF513351 – EF513418); *peBA* clones (EF513418 – EF513486); BO8805 *pcBA* (EF513487); CCY9201 *pcBA* (EF513488); CCY9202 *pcBA* (EF513489); CCY9202 *peBA* (EF513490); *Anabaena*-like 16S-ITS clone TH298-12-6 (EF530539).

Acknowledgements

We thank M. Laamanen for the opportunity to join cruise CYANO-04, and the crew of the research vessel Aranda for help during sampling. We also thank A. Wijnholds-Vreman for carrying out the flow cytometry analyses. We thank C. Everroad and A.M. Wood for sharing with us their *pcBA* primer sequences before publication. We gratefully acknowledge the comments of two anonymous referees. M.S. and J.H. were supported by the Earth and Life Sciences Foundation (ALW), which is subsidised by the Netherlands Organization for Scientific Research (NWO). T.H. and L.J.S. acknowledge support from the European Commission through the project MIRACLE (EVK3-CT-2002–00087).

The Supplementary Online Material of this chapter is available as part of the online article (Haverkamp *et al.* 2008) from <http://www.blackwell-synergy.com>