Microbiology (2002), **148**, 453–465 Printed in Great Britain

Genotyping of axenic and non-axenic isolates of the genus *Prochlorococcus* and the OMF-'*Synechococcus*' clade by size, sequence analysis or RFLP of the Internal Transcribed Spacer of the ribosomal operon

Wassila Laloui,[1]† Katarzyna A. Palinska,[1]‡ Rosmarie Rippka,[1] Frédéric Partensky,[2] Nicole Tandeau de Marsac,[1] Michael Herdman[1] and Isabelle Iteman[1]

Author for correspondence: Isabelle Iteman. Tel: +33 1 4568 8416. Fax: +33 1 40613042. e-mail: iiteman@pasteur.fr

[1] Unité des Cyanobactéries (CNRS URA 2172), Département de Biochimie et Génétique Moléculaire, Institut Pasteur, 28 Rue du Dr Roux, 75724 Paris Cedex 15, France

[2] Station Biologique de Roscoff (CNRS UPR 9042), BP 74, 29682 Roscoff Cedex, France

PCR amplicons of the Internal Transcribed Spacer (ITS) of the *rrn* operon of three axenic OMF (oceanic, marine and freshwater) strains of '*Synechococcus*' (WH7803, PCC 7001 and PCC 6307, respectively) differ greatly in length from that of the axenic *Prochlorococcus marinus* subsp. *pastoris* PCC 9511^T, although these four cyanobacteria cluster relatively closely in phylogenetic trees inferred from 16S rRNA gene sequences. The ITSs of three strains (PCC 9511^T, PCC 6307 and PCC 7001) were sequenced and compared with those available for strains *Prochlorococcus* MED4 (CCMP 1378) and MIT9313 from genome sequencing projects. In spite of large differences in length, sequence and mean DNA base composition, conserved domains important for transcriptional antitermination and folding of the rRNA transcripts were identified in all ITSs. A new group-specific primer permitted ITS amplification even with non-axenic isolates of *Prochlorococcus* and one OMF-'*Synechococcus*' strain. *Prochlorococcus* isolates of the high-light-adapted clade (HL) differed from representatives of the low-light-adapted clade (LL) by the length of their ITS. Restriction fragment length polymorphism (RFLP) of the ITS amplicons revealed three subclusters among the HL strains. Size, sequence data and RFLP of the ITS amplicons will therefore be valuable markers for the identification of different *Prochlorococcus* genotypes and for their discrimination from other cyanobacterial relatives with which they often co-exist in oceanic ecosystems.

Keywords: *Prochlorococcus marinus* subsp. *pastoris* PCC 9511^T, cyanobacteria, group-specific ITS primers, oceanic, coastal marine and freshwater '*Synechococcus*', secondary structure of the *rrn* operon

INTRODUCTION

Members of the genus *Prochlorococcus* were first discovered in 1988 with the aid of flow cytometry and this taxon was formally described by Chisholm *et al.* (1992) on the basis of the non-axenic type strain *Prochlorococcus marinus* CCMP 1375^T, isolated from the Sargasso Sea and also known as strain SS120. With cell diameters between 0·5 and 0·8 μm, and an abundance of up to 5×10^5 cells ml^{-1}, these oxygen-evolving photosynthetic prokaryotes occur throughout the euphotic zone, from surface layers to depths of 150–200 m, and play an important role for global primary production in tropical and temperate oceans (for a review, see Partensky *et al.*, 1999). Although members of the cyanobacterial phylogenetic lineage (Palenik & Haselkorn, 1992; Urbach *et al.*, 1992; Turner, 1997), these

†**Present address:** Unité de Recherche en Génomique Végétale (INRA/CNRS), Genoscope, Centre National de Séquençage, 2, Rue Gaston Crémieux, 91057 Evry Cedex, France.

‡**Present address:** Carl von Ossietzky University, Geomicrobiology, PO Box 2503, 26111 Oldenburg, Germany.

Abbreviations: DV, divinyl; HL, high-light-adapted clade; LL, low-light-adapted clade; ITS, Internal Transcribed Spacer; OMF, oceanic, marine and freshwater; RFLP, restriction fragment length polymorphism.

The GenBank accession numbers for the ITS amplicon sequences reported in this paper are AF387607 (PCC 9511^T), AF387610 (PCC 6307), AF387608 (PCC 7001) and AF387609 (PCC 7941).

oxyphototrophs differ from more typical cyanobacteria by lacking phycobilisomes and synthesizing divinyl (DV) derivatives of chlorophylls *a* and *b* (Partensky *et al*., 1999). Trace amounts of a novel type of phycoerythrin, or the genes encoding this phycobiliprotein, have been identified in some members (Hess *et al*., 1996, 1999; Penno *et al*., 2000).

Based on their preferential habitats, surface waters or deep layers receiving high or low irradiance, respectively, DV-chlorophyll *b*/*a* ratios and phylogenetic position deduced from 16S rRNA gene analyses, three subgroups have been recognized among isolates and environmental samples of *Prochlorococcus* (Morel *et al*., 1993; Partensky *et al*., 1993; Moore *et al*., 1995, 1998; Palenik, 1994; Urbach *et al*., 1998; West & Scanlan, 1999; West *et al*., 2001; see also Partensky *et al*., 1999). Two of these (HLI and HLII) are group members of the 'high-light-adapted clades' (HL) that proliferate in surface layers of the oceans and have a very low DV-chlorophyll *b*/*a* ratio. The third, loosely defined as the 'low-light-adapted clade' (LL), is composed of genetically more diverse ecotypes that exhibit higher DV-chlorophyll *b*/*a* ratios and occur preferentially at depths of more than 100 m. However, the vertical positioning is influenced by the degree of stratification of the water columns, more extensive mixing leading to the co-existence of both HL and LL ecotypes at the same depth (Moore *et al*., 1998; Ferris & Palenik, 1998). In addition, the relative abundance and depth distribution of members of the HLI and HLII subclades (West & Scanlan, 1999) may differ strikingly, depending on the geographical region (West *et al*., 2001).

The axenic *Prochlorococcus marinus* subsp. *pastoris* strain PCC 9511^{T} has a low mean DNA base composition (32 mol % G + C) and a small genome (2 Mbp) (Rippka *et al*., 2000). However, as inferred from 16S rDNA sequence analysis, the closest phylogenetic relatives of *Prochlorococcus* isolates or environmental clones (Urbach *et al*., 1998; West & Scanlan, 1999; West *et al*., 2001) are some unicellular cyanobacteria of the genetically diverse '*Synechococcus*' group (Waterbury & Rippka, 1989; Turner, 1997; Honda *et al*., 1999; Herdman *et al*., 2001). These are all of relatively small diameter (0·8–1 μm), span a mean DNA base composition in the range of 55–70 mol % G + C (Herdman *et al*., 1979b, 2001; Waterbury *et al*., 1986; Waterbury & Rippka, 1989) and (where known) have a genome size of about 4 Mbp (Herdman *et al*., 1979a), though that of the oceanic '*Synechococcus*' WH8102 is only 2·4 Mbp as shown by the genome sequencing project. '*Synechococcus*' strains that cluster most tightly with members of *Prochlorococcus*, and were thus assigned to the 'marine picophytoplankton clade' (Urbach *et al*., 1998), have all been isolated from the open ocean (Waterbury *et al*., 1986; Waterbury & Rippka, 1989) where they often co-exist with *Prochlorococcus* populations (Chisholm *et al*., 1988; Li & Wood, 1988; Olson *et al*., 1990; Shimada *et al*., 1996; Partensky *et al*., 1999). In contrast, '*Synechococcus*' strains that are positioned at the root of the former clade (Urbach *et al*., 1998; West & Scanlan, 1999; Rippka *et al*., 2000; West *et al*., 2001) are of freshwater origin (Kane *et al*., 1997; Postius & Ernst, 1999), or were isolated from a coastal marine environment (Rippka *et al*., 2000). To distinguish the ensemble of these related '*Synechococcus*' strains (or environmental clones) from other members of the '*Synechococcus*' group, that in 16S rDNA trees are more remote and dispersed in several distinct phylogenetic lineages (Turner, 1997; Honda *et al*., 1999; Herdman *et al*., 2001; Robertson *et al*., 2001), we shall here refer to them collectively as members of the oceanic, marine and freshwater (OMF)-'*Synechococcus*' clade.

For a variety of reasons, precise identification of cyanobacteria by classical bacteriological criteria is difficult (Rippka *et al*., 1979; Castenholz & Waterbury, 1989), but a number of different PCR-based strategies for the genetic characterization of strains, or environmental samples, are now available. Identification of *Prochlorococcus* genotypes has been achieved by different molecular approaches (Palenik, 1994; Ferris & Palenik, 1998; Scanlan *et al*., 1996; Urbach *et al*., 1998; West *et al*., 1999, 2001), including single-cell hybridization with labelled probes and detection of the fluorescent signals by light microscopy or flow cytometry (West *et al*., 2001; Worden *et al*., 2000). However, some of these methods are labour-intensive or are based on relatively conserved RNA polymerase or 16S rRNA gene sequences and thus may not provide sufficient resolution to distinguish between the different *Prochlorococcus* genotypes with precision. Furthermore, even discrimination of some members of *Prochlorococcus* from representatives of the OMF-'*Synechococcus*' clade may be difficult (West *et al*., 2001). Thus, more specific tools for the distinction within and between these two groups of organisms that often share the same ecological niche (Chisholm *et al*., 1988; Olson *et al*., 1990; Campbell & Vaulot, 1993; Partensky *et al*., 1999) are still desirable.

In a survey of about 320 axenic cyanobacterial strains of the Pasteur Culture Collection of Cyanobacteria (PCC; Institut Pasteur, Paris, France), it was observed that the different sizes of the Internal Transcribed Spacer (ITS) obtained by PCR amplification correlated well (I. Iteman, unpublished results) with particular strain clusters in genetically diverse 'genera' such as '*Synechococcus*' or '*Synechocystis*' (Waterbury & Rippka, 1989; Rippka & Herdman, 1992; Herdman *et al*., 2001). Knowing that the ITS is less conserved than the 16S rRNA gene in cyanobacteria (Wilmotte, 1994; Neilan *et al*., 1997; Otsuka *et al*., 1999; Iteman *et al*., 2000), we have examined the value of this domain for the genotyping of eight isolates of *Prochlorococcus* (one axenic, seven non-axenic) by PCR amplification, sequencing and/or RFLP of the ITS regions. For comparison, four '*Synechococcus*' strains of the OMF-'*Synechococcus*' clade (three axenic, one non-axenic) were also included in this study, one of which (PCC 6307) has previously been proposed as the type strain of

Cyanobium gracile gen. nov., sp. nov. Rippka & Cohen-Bazire 1983 (Komárek *et al.*, 1999; Herdman *et al.*, 2001). However, since the new genus *Cyanobium* has not yet been validated under the Bacteriological Code of Nomenclature and may in the future also be applicable to some of the other representatives of the OMF-'*Synechococcus*' clade, this nomenclatural change has been avoided here.

METHODS

Cyanobacterial strains. The non-axenic *Prochlorococcus* strains NATL2, OLI 36 FJA, GP2 (RCC296), SB (RCC295), TAK 9803-2 (RCC264) and NATL1-MIT (RCC277) were obtained from the Culture Collection of the Station Biologique, Roscoff, France (RCC). If still available from the latter collection, the corresponding RCC numbers are indicated in parentheses. The non-axenic *Prochlorococcus marinus* CCMP 1375^T (= SS120) was kindly provided by Dr Lisa Moore [Massachusetts Institute of Technology (MIT), Cambridge, MA, USA]. The axenic strains *Prochlorococcus marinus* subsp. *pastoris* PCC 9511^T (Rippka *et al.*, 2000) and '*Synechococcus*' PCC 6307 and PCC 7001 were from the PCC (Rippka & Herdman, 1992). '*Synechococcus*' WH 7803 (axenic) was received from Dr Debbie Lindell (Steinitz Marine Biology Lab, Eilat, Israel). '*Synechococcus*' TAK RED (non-axenic) was fortuitously selected from the same primary culture that gave rise to *Prochlorococcus* TAK 9803-2.

Cultivation of strains. All cultures, except strains PCC 6307 and PCC 7001, were grown at 18–20 °C in liquid medium PCR-Tu_2 (Rippka *et al.*, 2000). Strain PCC 6307 was grown at 23 °C in medium BG11 (Rippka *et al.*, 1979). Strain PCC 7001 was cultivated in a mixture (1:1) of media BG11 and ASNIII (Rippka *et al.*, 1979). For all strains, white light was supplied by fluorescent tubes providing during a light/dark cycle of 14/10 h a photosynthetic photon flux density (PPFD) of about 20 µm photons m^{-2} s^{-1} (measured with a LICOR LI-185B quantum/radiometer/photometer equipped with a LI-190SB quantum sensor). The neon tubes were Duro-Lite (USA) for all strains, except PCC 6307 and PCC 7001, for which illumination was with Universal White neon tubes. The strain of *Microcystis aeruginosa*, PCC 7941, used in one control experiment, was grown in medium $BG11_0$ (Rippka *et al.*, 1979) supplemented with $NaNO_3$ (2 mM) and $NaHCO_3$ (10 mM) under the light and temperature regimes described for strains PCC 6307 and PCC 7001. Several bacterial contaminants that also served for control PCR reactions were isolated from *Prochlorococcus* strains MED4 and SS120 on plates of medium PCR-Tu_2 supplemented with glucose (0·2 %, w/v), Casamino acids (0·002 %, w/v) and Luria Broth (5 %, v/v) (Sambrook *et al.*, 1989) and solidified with washed agar (0·8 %, w/v) (Rippka *et al.*, 2000). Incubation was at 20 °C in the dark. The purity of the axenic *Prochlorococcus* and '*Synechococcus*' strains was routinely checked in the appropriate growth medium with the same supplements as described for the isolation of the bacterial contaminants, but solidified with Difco Bacto agar (1 %, w/v).

PCR amplification of the ITS region. Amplification of the ITS regions of *Prochlorococcus* and '*Synechococcus*' strains was performed on cells (1–2 ml; OD_{750} about 0·1 for *Prochlorococcus* and 0·5–1·0 for '*Synechococcus*') harvested by centrifugation (15 min at 20 °C and 20000 ***g***), washed twice in 150 mM NaCl and resuspended in 50–200 µl sterile H_2O or TE buffer (10 mM Tris, pH 8·0, 0·1 mM EDTA). Aliquots of the cell suspensions (1–10 µl) were either used directly for the PCR reactions or after lysis by five alternating cycles of freezing in liquid nitrogen and thawing at 50 °C. The respective standard forward and reverse primers for amplification were 322 (5′-TGTACACACCGCCCGTC-3′) and 340 (5′-CTCTGTG-TGCCTAGGTATCC-3′), previously employed for cyanobacteria (Wilmotte *et al.*, 1993; Iteman *et al.*, 1999, 2000). Primer 322 initiates amplification at a region near the end of the 16S rDNA on the RNA-like strand (positions 1338–1354 in *Synechocystis* PCC 6803; *Escherichia coli* numbering 1391–1407); primer 340 is complementary to a region on the opposite strand at the beginning of the 23S rDNA (positions 26–45 in both *Synechocystis* PCC 6803 and *E. coli*). For the numbering of the *E. coli rrn* sequences, see Brosius *et al.* (1981). Alternatively, forward primer 600, designed in this study (5′-CACCTCCTAACAGGGAGACA-3′) and corresponding to the last 10 nt of the 16S rDNA (*E. coli* numbering 1533–1542) and the first 10 nt of the ITS on the RNA-like strand in the *rrn* operon of *Prochlorococcus* PCC 9511, was used in combination with reverse primer 340. The higher specificity of the latter set of primers for the *Prochlorococcus* and '*Synechococcus*' strains was tested by attempting ITS amplification on a cell suspension (1 µl, OD_{750} about 0·5) of the cyanobacterium *Microcystis aeruginosa* PCC 7941 and lysates (10 µl) of several bacterial contaminants prepared by alternating cycles of freezing in liquid nitrogen and thawing as described above. PCR amplifications with primers 322 and 340 served as positive controls.

The PCR mixture contained 10 µl *Taq* commercial buffer (10 ×) supplemented with $MgCl_2$ to a final concentration of 2 mM, 1–10 µl cells or lysates as templates, 150 µM each dNTP, 500 ng each primer and 2·5 U *Taq* polymerase (Promega). The total reaction volume was 50 or 100 µl. Incubation of the reactions was initially performed in a Perkin-Elmer Cetus Gene Amp 9700 Thermal Cycler using the following programme. After an initial cycle consisting of 5 min at 95 °C, 1 min at 55 °C and 30 s at 72 °C, 30 cycles of amplification were started (1 min at 95 °C, 1 min at 55 °C and 1 min at 72 °C). The reaction was terminated by a cycle of 5 min at 72 °C. For later PCR reactions in a Stratagene Robocycler Gradient 40, the programme was modified: 1 cycle consisting of 5 min at 94 °C, 1 min at 55 °C and 30 s at 72 °C; 30 cycles of 30 s at 94 °C, 30 s at 55 °C and 30 s at 72 °C; the final cycle was 1 min at 95 °C, 1 min at 55 °C and 5 min at 72 °C. Storage of the reactions in both programmes was at 4 °C. A negative control without template was included in some experiments. The PCR products were analysed by electrophoresis on 1·5 % (w/v) agarose gels (Litex LSL 4000, FMC BioProducts Europe) in 1 × TBE (Tris-borate-EDTA) buffer (Sambrook *et al.*, 1989), stained with ethidium bromide (0·5 µg ml^{-1}) and photographed under UV light. The length of the DNA fragments was estimated by comparison with a 100 bp DNA ladder (Amersham Pharmacia Biotech).

Cloning and sequencing of the ITS region. The PCR products of *Prochlorococcus* PCC 9511^T, '*Synechococcus*' strains PCC 6307 and PCC 7001, and *Microcystis aeruginosa* PCC 7941, obtained with primers 322 and 340, were sequenced after cloning or directly on the PCR products as described by Iteman *et al.* (2000).

Alignment and analyses of the ITS sequences. Alignment of the ITS sequences of *Prochlorococcus* PCC 9511^T and '*Synechococcus*' PCC 6307 and PCC 7001, together with those of *Prochlorococcus* MED4 and MIT9313, available from the genome sequencing projects, were made manually using Genedoc v. 2.5 (http://www.psc.edu/biomed/genedoc/) and by reference to the cyanobacterial ITS alignment and second-

Table 1. Origin and genotypes of *Prochlorococcus* and '*Synechococcus*' strains examined

The two strains for which the genome sequences are available (*Prochlorococcus* MED4 and MIT9313) and whose ITS sequences were aligned to those determined in this study are also included in this table.

Strain/genome sequence	Origin/depth	Genotype*	Reference†
Prochlorococcus			
Axenic			
Prochlorococcus PCC 9511^T	Sargasso Sea, 120 m‡	HLI	1, 2, 3, 4, 5
Non-axenic			
Prochlorococcus NATL2	North Atlantic, 30 m	HLI	4, 5
Prochlorococcus GP2	West Pacific, 150 m	HLII	4, 5, 6
Prochlorococcus SB	Sugura Bay, Japan, 40 m	HLII	4, 5, 7,
Prochlorococcus TAK 9803-2	Takapoto atoll, Pacific, 20 m	HLII	4, 5
Prochlorococcus OLI 36 FJA	Equatorial Pacific, 100 m	Unknown	8
Prochlorococcus NATL1-MIT	North Atlantic, 30 m	LL	4, 5
Prochlorococcus CCMP 1375^T (SS120)	Sargasso Sea, 120 m	LL	1, 2, 4, 5
'***Synechococcus***'			
Axenic			
'*Synechococcus*' PCC 7001	Intertidal mat, New York, USA	OMF-SYN	1, 9,
'*Synechococcus*' PCC 6307	Lake water, Wisconsin, USA	OMF-SYN	1, 2, 9
'*Synechococcus*' WH7803	North Atlantic, surface	OMF-SYN	10, 11
Non-axenic			
'*Synechococcus*' TAK RED	Same as TAK 9803-2	Unknown	This study
Genome sequences			
Prochlorococcus MED4 (CCMP 1378)	Mediterranean Sea, 5 m	HLI	2, 3, 4, 5
Prochlorococcus MIT9313	North Atlantic, 135 m	LL	3, 4, 5, 11, 12

* The distinction between HLI and HLII is according to West & Scanlan (1999) and West *et al*. (2001). OMF-SYN, OMF-'*Synechococcus*'.

† Only references appropriate for the description of the isolates or assignment to phylogenetic clusters (genotypes) are provided. 1, Chisholm *et al*. (1992); 2, Urbach *et al*. (1998); 3, Rippka *et al*. (2000); 4, Partensky *et al*. (1999); 5, West & Scanlan (1999); 6, Shimada *et al*. (1996); 7, Shimada *et al*. (1995); 8, F. Partensky, unpublished; 9, Rippka *et al*. (1979); 10, Waterbury *et al*. (1986); 11, West *et al*. (2001); 12, Moore *et al*. (1998).

‡ The origin of the axenic strain PCC 9511^T is somewhat doubtful (see Rippka *et al*., 2000).

ary structure models of the rRNA transcripts proposed by Iteman *et al*. (2000).

Analyses of the 16S rDNA–ITS overlapping regions. The sequences of the 16S rDNA–ITS overlapping regions of *Prochlorococcus* MED4 and MIT9313, and '*Synechococcus*' WH8102, located on contigs 26, 478 and 52, respectively, were obtained by consultation of the genome sequencing projects (http://www.jgi.doe.gov/JGI_microbial/html/prochlorococcus/prochlo_pickastrain.html and http://www.jgi.doe.gov/JGI_microbial/html/synechococcus/synech_homepage.html).

The GenBank accession numbers of additional 16S rDNA, ITS or 16S rDNA–ITS sequences used to investigate the specificity of primer 600 are as follows: *Synechococcus* PCC 6301, X00436; *Synechocystis* PCC 6803, D90916; *Spirulina* PCC 6313, X75045; *Arthrospira* PCC 8005, X70769; *Arthrospira* PCC 7345, X75044; environmental clone ENV WH7B, AJ007374; *Spirulina subsalsa* FACHB351, AF329394; all these sequences contain both the 16S rDNA and the ITS region. For the following strains two sequences were necessary to cover the 16S rDNA–ITS junction: *Trichodesmium* NIBB1067, X70767 (16S rDNA) and X72871 (ITS); *Anabaena* (*Nostoc*) PCC 7120, X59559 (16S rDNA) and AF180968 (ITS); *Nodularia* BCNOD 9427, AJ224447 (16S rDNA) and AJ224448 (ITS). For strains *Pseudanabaena* PCC 7403 and PCC 7409, and '*Chlorogloeopsis*' PCC 7518 (X687580; 16S rDNA), the 16S rDNA sequence and/or corresponding ITS were kindly provided by Gerard Van der Auwera and Annick Wilmotte (see also Iteman *et al*., 2000).

The sequences of the 16S rDNA–ITS junctions of *Arthrospira* OUQDS6 (AF329393), FACHB438 (AF329392) and FACHB439 (AF329391) are identical to that of *Arthrospira* PCC 7345 and, as for the corresponding sequence of *Nostoc* PCC 73102 (that shares 100 % sequence identity in this domain with strain *Anabaena* PCC 7120), are not included in Fig. 4, but were counted in the total of 24 cyanobacterial 16S rDNA–ITS junctions compared (see Discussion). The 16S rDNA and ITS sequences of *Nostoc* PCC 73102 are both located on contig 647, available from the genome sequencing project (http://www.jgi.doe.gov/JGI_microbial/html/nostoc/nostoc_homepage.html).

Restriction endonuclease digestion of the ITS amplicons. After amplification with primer pair 600 and 340, the PCR products (10–20 µl) of the *Prochlorococcus* strains were digested with 5 U of the following restriction enzymes: *Alu*I, *Dde*I, *Eco*RI, *Eco*RV, *Hae*III, *Hin*fI and *Hin*dIII (Gibco-BRL,

Life Technologies), according to the instructions of the manufacturer. The DNA fragments were separated by electrophoresis for 4–6 h at 60 V on 2·5 % (w/v) agarose gels (Metaphor, FMC, BioProducts) in 1 × TBE buffer (Sambrook *et al*., 1989). Staining, DNA size markers and visualization of the restriction fragments were as described above.

RESULTS

Strain description

The origin and phylogenetic assignments of the axenic and non-axenic strains of *Prochlorococcus* and '*Synechococcus*' examined in this study, or whose published ITS sequences were used for comparative analyses (*Prochlorococcus* strains MED4 and MIT9313), are described in Table 1. '*Synechococcus*' WH7803 is a representative of the open ocean (Waterbury *et al*., 1986) and a close relative of '*Synechococcus*' WH7805 (Urbach *et al*., 1998; West *et al*., 2001); both, like the impure strain TAK RED, have a high phycoerythrin content (Waterbury *et al*., 1986; R. Rippka, unpublished observation). Strains '*Synechococcus*' PCC 6307 and PCC 7001, more distantly related to the former oceanic strains (Urbach *et al*., 1998; Rippka *et al*., 2000), are of freshwater origin or isolated from a coastal marine environment, respectively, and neither of them synthesize phycoerythrin (Waterbury & Rippka, 1989; Herdman *et al*., 2001).

Amplification of the ITS region with primers 322 and 340

With primers 322 and 340, a single PCR amplicon of different size was observed for the axenic strains *Prochlorococcus* PCC 9511^T and '*Synechococcus*' PCC 6307, PCC 7001 and WH7803 (Fig. 1a; Table 2). After subtraction of a total of 200 bp contributed by the end of the 16S rDNA (150 bp) and the beginning of the 23S rDNA (50 bp), the sizes of the ITS regions amounted to about 550 bp for *Prochlorococcus* PCC 9511^T and to about 1000, 1100 and 800 bp, respectively, for '*Synechococcus*' PCC 6307, PCC 7001 and WH7803. However, amplifications with these primers performed on the non-axenic strains of *Prochlorococcus* and '*Synechococcus*' TAK RED yielded one major and one or more minor PCR products. The more intense bands observed

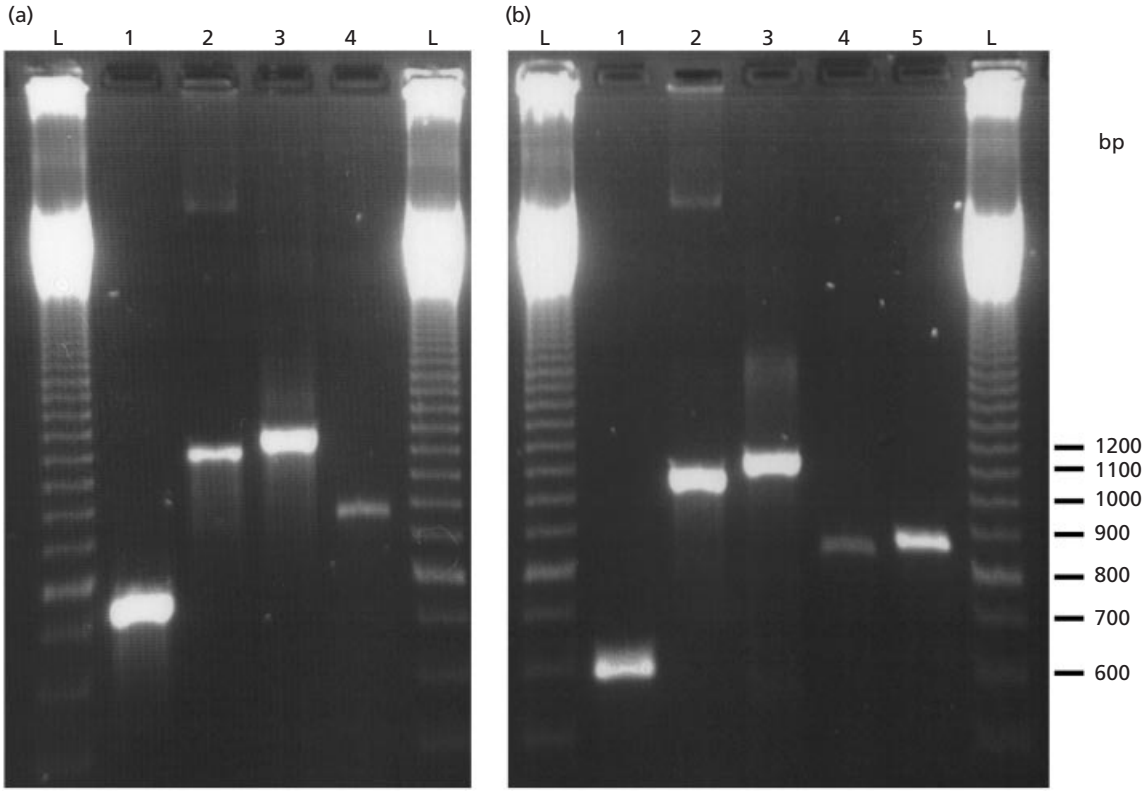

Fig. 1. Agarose gels showing the PCR products obtained by amplification of the ITS regions with two different sets of primers. (a) ITS amplicons with primer set 322 and 340 obtained with four axenic strains of *Prochlorococcus* and '*Synechococcus*'. Lanes: 1, *Prochlorococcus* PCC 9511^T; 2, '*Synechococcus*' PCC 7001; 3, '*Synechococcus*' PCC 6307; 4, '*Synechococcus*' WH 7803. (b) PCR products after amplification of the ITS regions with primer set 600 and 340. Lanes: 1, *Prochlorococcus* PCC 9511^T; 2, '*Synechococcus*' PCC 7001; 3, '*Synechococcus*' PCC 6307; 4, '*Synechococcus*' WH7803; 5, '*Synechococcus*' TAK RED (non-axenic). Lanes L, 100 bp molecular mass ladder.

Table 2. Sizes of the PCR products obtained for *Prochlorococcus* and '*Synechococcus*' isolates with two different sets of primers (322 and 340; 600 and 340) and groupings of *Prochlorococcus* strains based on size and RFLP profiles of the ITS amplicons obtained with primer set 600 and 340

Strain	Genotype*	Size (bp) of PCR amplicon or ITS (in parentheses) with primers:†		ITS RFLP cluster	Enzyme and length (bp) of restriction fragments‡			
		322 and 340	600 and 340		*Alu*I	*Dde*I	*Hin*fI	*Hae*III
'*Synechococcus*'								
PCC 6307	OMF-SYN	1200 (1010)[a]	1050 (1010)[a]	ND	ND	ND	ND	ND
PCC 7001	OMF-SYN	1300 (1065)[a]	1100 (1065)[a]	ND	ND	ND	ND	ND
WH7803	OMF-SYN	1000 (800)	880 (820)	ND	ND	ND	ND	ND
TAK RED	Unknown	1000§ (800)	890 (830)	ND	ND	ND	ND	ND
Prochlorococcus								
PCC 9511^T	HLI	750 (548)[a]	600 (548)[a]					
NATL2	HLI	750§ (550)	600 (540)	I.1	170, 140 (d), 80	260, 130, 80	210 (d)	400, 140
OLI 36 FJA	Unknown	750§ (550)	600 (540)					
TAK 9803-2	HLII	750§ (550)	600 (540)	I.2A				400, 140
GP2	HLII	750§ (550)	600 (540)	I.2B	170, 150, 140, 80	380, 80	290, 200, 80	
SB	HLII	750§ (550)	600 (540)					380, 230
SS120	LL	850§ (650)	700 (640)	II	240, 190, 150, 70	310, 130, 90, 70	380, 200, 130	500, 130
NATL1-MIT	LL	850§ (650)	700 (640)					

ND, Not determined.

* The distinction between HLI and HLII is according to West & Scanlan (1999) and West *et al*. (2001). OMF-SYN, OMF-'*Synechococcus*'.

† Since the amplicons include part of the adjacent 16S rDNA and 23S rDNA, the true ITS sizes (indicated in parentheses) were either determined from the ITS sequences ([a]) or calculated by subtracting 200 bp from the amplicons if using primers 322 and 340 and 60 bp if using primers 600 and 340.

‡ Fragments indicated with (d) for 'doublets' had a slightly stronger fluorescence and correspond most likely to two different restriction fragments of very similar size (see Results).

§ Although minor bands were generated by the contaminants with primers 322 and 340, only the sizes of the major PCR products are given.

for five of the impure *Prochlorococcus* strains (Table 2) were similar in size (PCR product about 750 bp = ITS of 550 bp) to the ITS amplicon of the axenic *Prochlorococcus* PCC 9511^T (see Fig. 1a), but they were larger (PCR product about 850 bp = ITS of 650 bp) for strains SS120 and NATL1-MIT (Table 2). The major PCR product of the impure strain '*Synechococcus*' TAK RED amounted, as for the axenic '*Synechococcus*' WH7803 (Fig. 1a; Table 2), to an ITS size of about 800 bp (Table 2).

The minor bands repeatedly observed with primer set 322 and 340 were undoubtedly the result of bacterial contamination. This was confirmed on cell lysates of five bacterial strains isolated from two different *Prochlorococcus* cultures and for which ITS sizes of 280, 580, 680, 700 and 900 bp could be determined (data not shown). As shown by our analyses by NCBI BLASTN 2.1.3 (http://www.ncbi.nlm.nih.gov/blast/), the region complementary to primer 322 is well conserved among many prokaryotes, but that complementary to primer 340 corresponds to a signature sequence typical of the 23S rRNA gene in cyanobacteria and chloroplasts. Hence, the primer set 322 and 340 is relatively specific for cyanobacteria (Wilmotte *et al*., 1993; Iteman *et al*., 2000), However, as evidenced by the positive amplification results with the contaminants, conservation of both primer regions seems also to be typical of the marine bacteria associated with the impure *Prochlorococcus* and '*Synechococcus*' strains examined.

Amplification of the ITS region with primers 600 and 340

Given that the standard set of primers (322 and 340) amplified the ITS of the contaminants, forward primer 322 was replaced by primer 600, whose oligonucleotide sequence was designed on the basis of the 16S rDNA and ITS sequences of the axenic *Prochlorococcus* PCC 9511^T (see below). Amplifications with this new set of primers performed on the axenic *Prochlorococcus* PCC 9511^T and '*Synechococcus*' strains PCC 6307, PCC 7001 and WH7803 gave ITS amplicons of the expected size (compare Fig. 1a, b; see also Table 2). After subtraction of the 60 bp contributed by the end of the 16S rDNA (10 bp) and the beginning of the 23S rDNA (50 bp), the length of the ITS regions amounted, respectively, to about 540, 1000, 1050 and 820 bp, which is in good agreement with the estimates based on amplifications with primers 322 and 340 (Fig. 1a). No additional bands were observed for the impure strain '*Synechococcus*' TAK RED (Fig. 1b) and the non-axenic *Prochlorococcus* isolates (Table 2), demonstrating the higher specificity of the new set of primers. Furthermore, PCR reactions performed at the same temperature (55 °C) as that used for the *Prochlorococcus* and '*Synechococcus*' strains,

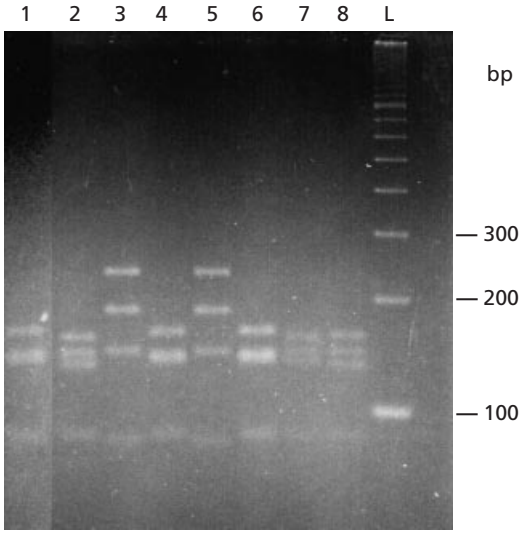

Fig. 2. RFLP of the ITS amplicons of one axenic and seven non-axenic *Prochlorococcus* strains obtained with primers 600 and 340 and digested with *Alu*I. Lanes: 1, PCC 9511^T (axenic); 2, TAK 9803-2; 3, NATL1-MIT; 4, NATL2; 5, SS120; 6, OLI 36 FJA; 7, SB; 8, GP2; L, 100 bp molecular mass ladder. Note that for convenience of presentation, the digest of the ITS amplicon of the axenic strain PCC 9511^T (lane 1), examined on the same gel but not located next to the other *Prochlorococcus* strains, was added using Adobe Photoshop.

gave no amplicons with cells of the cyanobacterium *Microcystis aeruginosa* PCC 7941 and lysates of the bacterial strains isolated from the impure *Prochlorococcus* cultures (data not shown).

The sizes of the bands observed after amplification of the impure *Prochlorococcus* strains with primer set 600 and 340 (Table 2) were in accordance with those determined for the major PCR products obtained with primer pair 322 and 340 (Table 2). Again it was possible to distinguish strains SS120 and NATL1-MIT from the other six *Prochlorococcus* isolates by their larger ITS size (about 640 versus 540 bp). The latter two strains correspond to members of the LL clade whereas, like the axenic strain PCC 9511^T, four of the remaining five (NATL2, TAK 9803-2, GP2 and SB; Table 2) are representatives of the HL clade (Urbach *et al*., 1998; West & Scanlan, 1999; West *et al*., 2001; see also Table 2). The phylogenetic position of strain OLI 36 FJA is unknown, but based on its ITS size (about 550 bp) would also be assignable to the HL clade. The size of the ITS of the impure strain of '*Synechococcus*' TAK RED (830 bp; see Fig. 1b and Table 2) proved again similar to that of the axenic '*Synechococcus*' WH7803 (820 bp; see Fig. 1a, b and Table 2).

Restriction fragment length polymorphism (RFLP) of the ITS amplicons

The ITS amplicons of the seven non-axenic *Prochlorococcus* isolates and of the axenic strain PCC 9511 obtained with primer set 600 and 340 were digested with seven different restriction enzymes. Although no cutting sites could be identified for any of the ITS amplicons with the enzymes *Eco*RI, *Eco*RV and *Hin*dIII (not shown), RFLP was observed with the remaining four enzymes tested (Table 2). A representative gel of amplicons restricted with *Alu*I is shown in Fig. 2. Depending on the restriction enzyme employed, the RFLP patterns were composed of two to four bands between 70 and 500 bp (Table 2) and it was possible to distinguish three RFLP subclusters (I.1, I.2A and I.2B) among the strains that yielded a small ITS amplicon (about 600 bp; Table 2). Subcluster I.1 regroups the HLI strains PCC 9511^T and NATL2 (West & Scanlan, 1999; Rippka *et al*., 2000) and includes OLI 36 FJA, whose phylogenetic position was previously unknown. Apart from fragments of short length (less than 70 bp) that were not detected (and thus the sum of the fragments may not add up perfectly to the size of the amplicons), the bands visualized by gel electrophoresis for the ITS amplicons of members of RFLP cluster I.1 are in good agreement with those expected on the basis of the restriction sites identified by DNA Strider 1.2 (Christian Marck, CEA, France) for the ITS sequence of *Prochlorococcus* MED4 (from the genome sequencing project) and that of *Prochlorococcus* PCC 9511^T (see below). The latter analysis also confirmed the existence of restriction fragments of similar size (139 and 141 bp for *Alu*I and 207 and 209 bp for *Hinf*I) that are only revealed as a single band ('doublets', Table 2) of slightly stronger ethidium bromide fluorescence on the gels (see also Fig. 2, after restriction of the amplicons with *Alu*I).

Strains of RFLP subclusters I.2A and I.2B (TAK9803-2, GP2 and SB) correspond to the HLII clade (West & Scanlan, 1999; West *et al*., 2001), but strain TAK9803-2 (subcluster I.2A) differs by having the same *Hae*III RFLP profile as strains of subcluster I.1, whereas strains GP2 and SB (subcluster I.2B) share a distinct banding pattern with this enzyme (Table 2). In contrast, strains *Prochlorococcus* SS120 and NATL1-MIT, that resulted in a larger ITS amplicon (about 700 bp; Table 2) and correlate with the LL clade (Urbach *et al*., 1998; West & Scanlan, 1999; Rippka *et al*., 2000), could not be further subdivided by ITS RFLP with the number of enzymes examined (Table 2).

Sequences and alignments of the ITS

The sequences of the ITS regions were determined for *Prochlorococcus* PCC 9511^T and two of the '*Synechococcus*' strains (PCC 6307 and PCC 7001). They were compared to those of *Prochlorococcus* MED4 and MIT9313, using as a reference the cyanobacterial ITS alignment proposed by Iteman *et al*. (2000). In agreement with previous conclusions concerning the possible identity of strains *Prochlorococcus* PCC 9511^T and MED4 (Rippka *et al*., 2000), sequence comparison demonstrated that the ITS of these two strains are identical (data not shown). Consequently, the ITS sequence of strain MED4 was omitted from the alignment shown in Fig. 3. For the reasons discussed below, the alignment

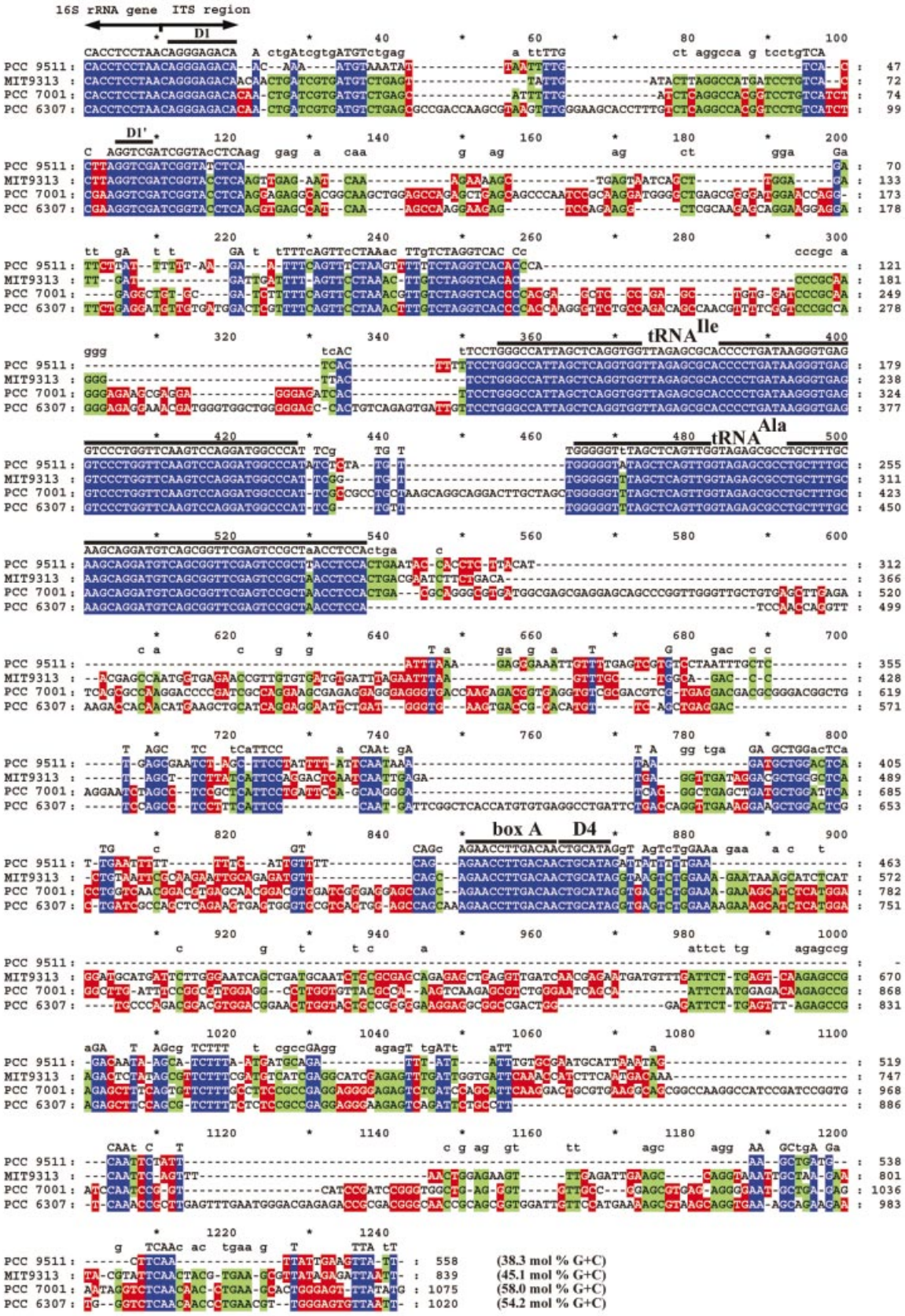

Fig. 3. Alignment of the nucleotide sequences of the end of the 16S rRNA gene and the ITS region of *Prochlorococcus* PCC 9511^{T}, *Prochlorococcus* MIT9313, '*Synechococcus*' PCC 7001 and '*Synechococcus*' PCC 6307. Percentage identity is

also includes the last 10 nt of the 16S rRNA gene. This region, together with that covering the first 10 nt of the ITS, is fully conserved in all sequences compared (Fig. 3), including MED4 (data not shown). In agreement with the approximate relative molecular mass determined for the ITS amplicons (see Fig. 1a, b and Table 2), the ITS sequences for *Prochlorococcus* PCC 9511^T (and MED4), '*Synechococcus*' PCC 6307 and PCC 7001 are 548, 1010 and 1065 bp in length, respectively. Strain *Prochlorococcus* MIT9313, a deeply branching member of the LL clade (Moore *et al*., 1998; West & Scanlan, 1999; West *et al*., 2001), has an ITS of intermediate length (829 bp), compared to the two HLI strains *Prochlorococcus* PCC 9511^T and MED4 and strains '*Synechococcus*' PCC 6307 and PCC 7001 (Fig. 3), but is similar in size to the ITS of 820 and 830 bp determined, respectively, for the oceanic '*Synechococcus*' strains WH7803 (Fig. 1a, b; Table 2) and TAK RED (Fig. 1b; Table 2).

The mean DNA base composition of the ITSs also differ considerably, being lowest for *Prochlorococcus* PCC 9511^T (38·3 mol% G+C), intermediate for *Prochlorococcus* MIT9313 (45·1 mol% G+C, as determined from the genome sequencing project) and significantly higher in the two '*Synechococcus*' strains PCC 6307 and PCC 7001 (54·2 and 58·0 mol% G+C, respectively).

In spite of the differences in size and base composition, as well as extensive sequence divergence in the more variable regions of the four ITS sequences aligned, common conserved domains, such as the two tRNA genes (encoding tRNA^{Ile} and tRNA^{Ala}), could be identified (Fig. 3) and also include box A and motifs D1, D1′ and D4 (Fig. 3), the former being important for transcriptional antitermination, the latter three for folding of the rRNA transcripts (Iteman *et al*., 2000).

DISCUSSION

The eight strains of *Prochlorococcus* examined (Table 2) could be subdivided into two ITS clusters by minor size differences (about 550 and 650 bp, corresponding to the HL and LL genotypes, respectively). However, the LL strain *Prochlorococcus* MIT9313, that in phylogenetic trees inferred from 16S rDNA sequences is distinct from the above members of this genus (Moore *et al*., 1998; Urbach *et al*., 1998; Rippka *et al*., 2000; West *et al*., 2001), has a larger ITS (829 bp as revealed by the genome sequence) that is similar in size to that of the oceanic '*Synechococcus*' strains WH7803 and TAK RED (about 820 and 830 bp, respectively, Table 2) and WH8102 (806 bp from the available genome sequence). Both the freshwater and coastal marine isolates '*Synechococcus*' PCC 6307 and PCC 7001 that form a distinct clade in phylogenetic trees (Urbach *et al*., 1998; West & Scanlan, 1999; Rippka *et al*., 2000; West *et al*., 2001) could be distinguished from all the above strains by their very long ITS (1010 and 1065 bp, respectively), which even among bacteria is exceptional. Indeed, a survey of 855 bacterial ITS sequences presently available in GenBank revealed that only representatives of two genera (*Bartonella* and *Bradyrhizobium*) have an ITS of equivalent or slightly greater length.

In contrast to filamentous cyanobacteria that often have more than one *rrn* operon with ITS regions of at least two different sizes (Lu *et al*., 1997; West & Adams, 1997; Neilan *et al*., 1997; Iteman *et al*., 1999, 2000; Barker *et al*., 1999), a single ITS amplicon was observed for all strains of *Prochlorococcus* and '*Synechococcus*'. This may suggest that, like in the genomes of *Prochlorococcus* MED4 and MIT9313 and the oceanic '*Synechococcus*' WH8102, a single *rrn* operon exists in all the strains examined here. However, it is also possible that, at least in some of these isolates, more than one copy of the *rrn* operon is present in their genome, but that the corresponding ITS regions are identical or very similar in size, as is the case for the unicellular cyanobacteria *Synechococcus* PCC 6301 and *Synechocystis* PCC 6803, both of which have two nearly identical ITSs (Williamson & Doolittle, 1983; Tomioka & Sugiura, 1984; Iteman *et al*., 2000). If multiple *rrn* copies exist in the genomes of the strains of *Prochlorococcus* examined, both the sizes and recognition sequences of their ITS seem to be identical, since the sums of the restriction fragments are in fair agreement with the lengths of the ITS amplicons (Table 2) and not larger, as would be the case if there were sequence variations in the recognition sites.

The groupings obtained by RFLP of the ITS amplicons for the eight strains of *Prochlorococcus* examined, correlated well with the three different genotypes (HLI, HLII and LL) recognized on the basis of 16S rRNA gene sequences (Urbach *et al*., 1998; West & Scanlan, 1999; Rippka *et al*., 2000; West *et al*., 2001) and other molecular characterizations (Scanlan *et al*., 1996; Urbach *et al*., 1998; West & Scanlan, 1999; West *et al*., 2001). This clearly demonstrates the overall value of this simple technique. However, it was not possible to discriminate between the two LL representatives SS120 and NATL1-MIT, although, based on 16S rDNA analyses, they are slightly more distant from one another than are the more interrelated HLI and HLII strains (Urbach *et al*., 1998; West & Scanlan, 1999; West *et al*., 2001). Thus, for a higher level of discrimination between these two LL strains, analyses with a wider range of restriction enzymes or ITS sequence information will be required. The low degree of divergence in the 16S rRNA genes between the most distantly related HL and LL representatives of *Prochlorococcus* (97·3% similar-

represented by three colours: blue for 100%, green for 75% and red for 50%, gaps being treated as differences. Conserved motifs D1, D1′, box A and D4 are indicated by reference to the ITS alignment and secondary structural models of the *rrn* transcripts of Iteman *et al*. (2000). The G+C values given at the end of the alignment are those of the ITS, with the last 10 bases of the 16S rDNA excluded.

```
PCC 9511  : CACCTCCTAA-CAGGGAGACA :20
MED4      : CACCTCCTAA-CAGGGAGACA :20
MIT9313   : CACCTCCTAA-CAGGGAGACA :20
PCC 6307  : CACCTCCTAA-CAGGGAGACA :20
PCC 7001  : CACCTCCTAA-CAGGGAGACA :20
WH8102    : CACCTCCTAA-CAGGGAGACA :20
PCC 7941  : CACCTCCTTA-AAGGGAGACC :20
PCC 7409  : CACCTCCTTA-TAGGGAGACC :20
PCC 6301  : CACCTCCTTT-CAGGGAGACC :20
PCC 6803  : CACCTCCTTT-AAGGGAGACC :20
PCC 6313  : CACCTCCTTT-AAGGGAGACC :20
NIBB1067  : CACCTCCTTT-AAGGGAGACC :20
PCC 7403  : CACCTCCTTT-TAGGGAGACC :20
PCC 8005  : CACCTCCTTT-TAGGGAGACC :20
PCC 7120  : CACCTCCTTT-TAGGGAGACC :20
BCNOD9427 : CACCTCCTTT-TAGGGAGACC :20
PCC 7518  : CACCTCCTTT-TAGGGAGACC :20
ENV WH7B  : CACCTCCTTT-TAGGGAGACC :20
PCC 7345  : CACCTCCTTTTTAGGGAGACC :21
FACHB351  : CACCTCCTTTAAAGGGAGACC :21
```

Fig. 4. Alignment of sequences overlapping the 16S rRNA gene and the ITS region of 20 cyanobacteria. The vertical bar on the left of the alignment indicates the cluster of strains sharing 100% identity; boxed motifs denote blocks of conserved domains within all the sequences presented. The sequences shown correspond from top to bottom to the following cyanobacterial strains or environmental clones: *Prochlorococcus* PCC 9511^T, MED4 and MIT9313; '*Synechococcus*' PCC 6307, PCC 7001 and WH8102; *Microcystis* PCC 7941; *Pseudanabaena* PCC 7409; *Synechococcus* PCC 6301; *Synechocystis* PCC 6803; *Spirulina* PCC 6313; *Trichodesmium* NIBB1067; *Pseudanabaena* PCC 7403; *Arthrospira* PCC 8005; *Anabaena* (*Nostoc*) PCC 7120; *Nodularia* BCNOD9427; '*Chlorogloeopsis*' PCC 7518; environmental clone ENV WH7B; *Arthrospira* PCC 7345; and *Spirulina subsalsa* FACHB351. Accession numbers of the sequences used for this alignment are listed in Methods.

ity; West *et al*., 2001), suggests that the three RFLP groupings (I.1, I.2A and I.2B) obtained for the HL strains may correspond to distinctions at the subspecies level. In contrast, the larger ITS size and different RFLP profiles observed for the two LL strains, compared to all HL isolates, would be consistent with them being members of a different species. However, confirmation of these postulated taxonomic relationships needs to await DNA–DNA hybridization studies.

Analyses of the region covering the last 10 bp in the 16S rRNA gene and the first 10 bp in the ITS for a total of 24 cyanobacteria, for which this sequence information is presently available (see Methods), revealed the specificity of the new primer 600. As shown in Fig. 4, only *Prochlorococcus* PCC 9511, MED4 and MIT9313, and strains PCC 6307, PCC 7001 and the oceanic '*Synechococcus*' WH8102 show 100% sequence identity within the complementary target domain. For all other cyanobacteria, three or more mismatches could be identified (Fig. 4) and, as experimentally confirmed for the randomly chosen cyanobacterium *Microcystis aeruginosa* PCC 7941, are sufficient to prevent PCR amplification, if using an annealing temperature of 55 °C. Consequently, with this new group-specific primer the ITS region can be selectively amplified from *Prochlorococcus* and all '*Synechococcus*' strains of the OMF-'*Synechococcus*' clade, even if the cultures are contaminated by several different bacteria. The latter is particularly important for the genus *Prochlorococcus*, since only a single axenic strain (PCC 9511^T) of this genus is presently available (Rippka *et al*., 2000).

The G+C content of the ITS (Fig. 3) correlated well with large differences in the mean DNA base composition determined for the total DNA of these organisms (Herdman *et al*., 1979b; Waterbury *et al*., 1986; Rippka *et al*., 2000), or deduced from the genome sequencing projects (for *Prochlorococcus* MED4 and MIT9313). The two HLI strains, *Prochlorococcus* PCC 9511^T and MED4, have an identical ITS with the lowest G+C content (38·3 mol% G+C; see Fig. 3; for MED4, data not shown), reflecting the low mean DNA base composition (32 and 30·9 mol% G+C, respectively) of their total genomic DNA. In contrast, the mean DNA base composition of the ITS for '*Synechococcus*' PCC 6307 and PCC 7001 is much greater (54·2 and 58·0 mol% G+C), which is largely consistent with their higher total DNA base composition (69·7 and 69·5 mol% G+C, respectively). The deeply branching LL strain *Prochlorococcus* MIT9313 has an ITS of an intermediate G+C content (45·1 mol% G+C) and the mean DNA base composition (50·8 mol% G+C) of its genome is, like the length of its ITS (829 bp), closest to oceanic representatives of '*Synechococcus*' (55–62 mol% G+C; Waterbury *et al*., 1986). Thus, in contrast to the 16S rRNA genes whose G+C content spans a very narrow range (53–55 mol% G+C for *Prochlorococcus* and OMF-'*Synechococcus*' strains, compared to 51–59 mol% G+C for the remaining 500 cyanobacterial 16S rDNA sequences presently available in our database), even if the strains differ widely in total mean DNA base composition (32–70 mol% G+C), it seems clear that the ITS region serves as a better indicator of divergence at the genomic level than the more conserved rDNA.

In spite of differences in size and mean DNA base composition, comparison of the new ITS sequences of *Prochlorococcus* PCC 9511^T and the two '*Synechococcus*' strains (PCC 6307 and PCC 7001) with those of *Prochlorococcus* MIT9313 (Fig. 3) and MED4 (data not shown), revealed valuable insights. Several conserved domains important either for transcriptional antitermination (box A) or folding (the complementary domains D1 and D1′ close to the 5′ end of the ITS, together with domain D4 immediately adjacent to box A) of the rRNA transcripts were identified as in other cyanobacteria (Iteman *et al*., 2000) and two tRNA genes encoding tRNA^{Ile} and tRNA^{Ala} were found for all strains. This is in contrast to some other cyanobacterial ITSs that may lack one or both tRNA genes (Iteman *et al*., 2000). As in most other cyanobacterial ITS regions (Iteman *et al*., 2000), both tRNA genes lack the 3′-terminal CCA extension found in many bacteria, although they contain the characteristic subterminal CCA sequence that appears to be typical of cyanobacteria

(Iteman *et al*., 2000). The overall sequence similarity of the ITS sequences, including some of the more variable domains, was also higher among these five strains than that observed with respect to the ITS of other cyanobacterial genera (compare Fig. 3 with Fig. 3 of Iteman *et al*., 2000). However, major differences in sequence are evident in certain regions. *Prochlorococcus* MIT9313 differs from strain PCC 9511 in possessing a long insert (positions 883–1001; Fig. 3) and 12 short inserts situated throughout the ITS. These inserts are also observed in '*Synechococcus*' strains PCC 6307 and PCC 7001, although the latter organisms differ in the positions of additional inserts that further increase the length of their ITS regions beyond that of strain MIT9313: the most significant of these start at alignment positions 44, 64, 181, 258, 304, 336, 589, 746 and 1116 for PCC 6307, and at 138, 157, 258, 304, 443, 556, 649, 691, 1078 and 1132 for PCC 7001. Since strains PCC 6307 and PCC 7001 differ markedly in sequence in certain regions, similar ITS sizes do not necessarily indicate close genetic relationships. However, in the case of *Prochlorococcus* MIT9313 and the oceanic '*Synechococcus*' strain WH7803, for which the phylogenetic relationship is known (West *et al*., 2001), the similarity in length of this domain may be of significance. Indeed, in 16S rRNA gene trees, strain MIT9313 together with *Prochlorococcus* MIT9303, occupies a position that forms the root of the LL *Prochlorococcus* clade and is closest to the oceanic cluster of the OMF-'*Synechococcus*' clade (West *et al*., 2001).

The availability of the three new ITS sequences, together with those of *Prochlorococcus* MED4 and MIT9313, should permit the design of new primers that, in combination with the group-specific primer set 600 and 340, may be used to achieve identification of strains or field samples at different resolution (i.e. at a lower or higher taxonomic rank) by PCR. In addition, probes complementary to discriminatory regions of the ITS may also be designed and labelled to attempt single-cell hybridization and detection by light microscopy or flow cytometry, possibly resulting in a higher specificity of identification than is presently achieved with probes targeting the more conserved 16S rRNA molecule (West *et al*., 2001; Worden *et al*., 2000). However, the search for such specific primers or probes will profit from the determination of additional ITS sequences. Finally, as more complete 16S and 23S rRNA gene and ITS sequences become available, secondary structure analyses will provide valuable insights into evolutionary aspects of these domains in the ribosomal operon of cyanobacteria from oceanic, marine and freshwater environments.

ACKNOWLEDGEMENTS

The authors wish to thank Nyree J. West and David J. Scanlan for having made available a phylogenetic tree of *Prochlorococcus* and oceanic '*Synechococcus*' strains prior to publication. Annick Wilmotte and Gerard Van der Auwera kindly provided unpublished 16S rDNA and ITS sequence data. The work was supported by the Institut Pasteur, the Centre National de la Recherche Scientifique (CNRS, URA 2172) and in part by the EC programme MAST 3-CT97-0128 (PROMOLEC).

REFERENCES

Barker, G. L. A., Hayes, P. K., O'Mahony, S. L., Vacharapiyasophon, P. & Walsby, A. E. (1999). A molecular and phenotypic analysis of *Nodularia* (cyanobacteria) from the Baltic Sea. *J Phycol* **35**, 931–937.

Brosius, J., Dull, T. J., Sleeter, D. D. & Noller, H. F. (1981). Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *J Mol Biol* **148**, 107–127.

Campbell, L. & Vaulot, D. (1993). Photosynthetic picoplankton community structure in the subtropical North Pacific Ocean near Hawaii (station ALOHA). *Deep Sea Res* **40**, 2043–2060.

Castenholz, R. W. & Waterbury, J. B. (1989). Oxygenic photosynthetic bacteria. Group I. Cyanobacteria. Preface. In *Bergey's Manual of Systematic Bacteriology*, pp. 1710–1727. Edited by J. T. Staley, M. P. Bryant, N. Pfenning & J. G. Holt. Baltimore: Williams & Wilkins.

Chisholm, S. W., Olson, R. J., Zettler, E. R., Goericke, R., Waterbury, J. B. & Welschmeyer, N. A. (1988). A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334**, 340–343.

Chisholm, S. W., Frankel, S. L., Goericke, R., Olson, R. J., Palenik, B., Waterbury, J. B., West-Johnsrud, L. & Zettler, E. R. (1992). *Prochlorococcus marinus* nov. gen. nov. sp.: An oxyphototrophic marine prokaryote containing divinyl chlorophyll *a* and *b*. *Arch Microbiol* **157**, 297–300.

Ferris, M. J. & Palenik, B. (1998). Niche adaptation in ocean cyanobacteria. *Nature* **396**, 226–228.

Herdman, M., Janvier, M., Rippka, R. & Stanier, R. Y. (1979a). Genome size of cyanobacteria. *J Gen Microbiol* **111**, 73–85.

Herdman, M., Janvier, M., Waterbury, J. B., Rippka, R., Stanier, R. Y. & Mandel, M. (1979b). Deoxyribonucleic acid base composition of cyanobacteria. *J Gen Microbiol* **111**, 63–71.

Herdman, M., Castenholz, R. W., Iteman, I., Waterbury, J. B. & Rippka, R. (2001). Subsection I. (Formerly *Chroococcales* Wettstein 1924, emend. Rippka, Deruelles, Waterbury, Herdman and Stanier 1979). In *Bergey's Manual of Systematic Bacteriology*, 2nd edn, vol. 1, *The Archaea and the Deeply Branching and Phototrophic Bacteria*, pp. 493–514. Edited by D. R. Boone, R. W. Castenholz & G. M. Garrity. New York, Berlin, Heidelberg: Springer.

Hess, W. R., Partensky, F., van der Staay, G. W. M., Garcia-Fernandez, J. M., Börner, T. & Vaulot, D. (1996). Coexistence of phycoerythrin and a chlorophyll *a*/*b* antenna in a marine prokaryote. *Pro Natl Acad Sci U S A* **93**, 11126–11130.

Hess, W. R., Steglich, C., Lichtlé, C. & Partensky, F. (1999). Phycoerythrins of the oxyphotobacterium *Prochlorococcus marinus* are associated to the thylakoid membranes and are encoded by a single large gene cluster. *Plant Mol Biol* **40**, 507–521.

Honda, D., Yokota, A. & Sugiyama, J. (1999). Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine *Synechococcus* strains. *J Mol Evol* **48**, 723–739.

Iteman, I., Rippka, R., Tandeau de Marsac, N. & Herdman, M. (1999). Use of molecular tools for the study of genetic relationships of heterocystous cyanobacteria. *Bull Inst Océanogr* (*Monaco*) **special issue 19**, 13–20.

Iteman, I., Rippka, R., Tandeau de Marsac, N. & Herdman, M. (2000). Comparison of conserved structural and regulatory

domains within divergent 16S rRNA–23S rRNA spacer sequences of cyanobacteria. *Microbiology* **146**, 1275–1286.

Kane, M., Maeda, H., Fukunaga, T. & Nishi, K. (1997). Molecular phylogenetic relationship between strains of cyanobacterial picoplankton in Lake Biwa, Japan. *J Mar Biotechnol* **5**, 41–45.

Komárek, J., Kopecky, J. & Cepák, V. (1999). Generic characters of the simplest cyanoprokaryotes *Cyanobium*, *Cyanobacterium* and *Synechococcus*. *Cryptogam Algol* **20**, 209–222.

Li, W. K. W. & Wood, A. M. (1988). Vertical distribution of North Atlantic ultraphytoplankton: Analysis by flow cytometry and epifluorescence microscopy. *Deep Sea Res* **35**, 1615–1638.

Lu, W., Evans, H. E., McColl, M. & Saunders, V. A. (1997). Identification of cyanobacteria by polymorphisms of PCR-amplified ribosomal DNA spacer region. *FEMS Microbiol Lett* **153**, 141–149.

Moore, L. R., Goericke, R. & Chisholm, S. W. (1995). Comparative physiology of *Synechococcus* and *Prochlorococcus*: Influence of light and temperature on growth, pigments, fluorescence and absorptive properties. *Mar Ecol Prog Ser* **116**, 259–275.

Moore, L. R., Rocap, G. & Chisholm, S. W. (1998). Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* **393**, 464–467.

Morel, A., Ahn, Y.-H., Partensky, F., Vaulot, D. & Claustre, H. (1993). *Prochlorococcus* and *Synechococcus*: A comparative study of their optical properties in relation to their size and pigmentation. *J Mar Res* **51**, 617–649.

Neilan, B. A., Stuart, J. L., Goodman, A. E., Cox, P. T. & Hawkins, P. R. (1997). Specific amplification and restriction polymorphisms of the cyanobacterial rRNA operon spacer region. *Syst Appl Microbiol* **20**, 612–621.

Olson, R. J., Chisholm, S. W., Zettler, E. R., Altabet, M. A. & Dusenberry, J. A. (1990). Spatial and temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. *Deep Sea Res* **37**, 1033–1051.

Otsuka, S., Suda, S., Li, R., Watanabe, M., Oyaizo, H., Matsomuto, S. & Watanabe, M. M. (1999). Phylogenetic relationships between toxic and nontoxic strains of the genus *Microcystis* based on 16S to 23S internal transcribed spacer sequence. *FEMS Microbiol Lett* **172**, 15–21.

Palenik, B. (1994). Cyanobacterial community structure as seen from RNA polymerase gene sequence analysis. *Appl Environ Microbiol* **60**, 3212–3219.

Palenik, B. & Haselkorn, R. (1992). Multiple evolutionary origins of prochlorophytes, the chlorophyll *b*-containing prokaryotes. *Nature* **355**, 265–267.

Partensky, F., Hoepffner, N., Li, W. K. W., Ulloa, O. & Vaulot, D. (1993). Photoacclimation of *Prochlorococcus* sp. (Prochlorophyta) strains isolated from the North Atlantic and the Mediterranean Sea. *Plant Physiol* **101**, 285–296.

Partensky, F., Hess, W. R. & Vaulot, D. (1999). *Prochlorococcus*, a key marine photosynthetic prokaryote. *Microbiol Mol Biol Rev* **63**, 106–127.

Penno, S., Campbell, L. & Hess, W. R. (2000). Presence of phycoerythrin in two strains of *Prochlorococcus* isolated from the sub-tropical North Pacific Ocean. *J Phycol* **36**, 723–729.

Postius, C. & Ernst, A. (1999). Mechanisms of dominance: coexistence of picocyanobacterial genotypes in a freshwater ecosystem. *Arch Microbiol* **172**, 69–75.

Rippka, R. & Herdman, M. (1992). Pasteur Culture Collection of Cyanobacterial Strains in Axenic Culture. *Catalogue and Taxonomic Handbook. Vol. 1: Catalogue of Strains*. Paris: Institut Pasteur.

Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M. & Stanier, R. Y. (1979). Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J Gen Microbiol* **111**, 1–61.

Rippka, R., Coursin, T., Hess, W. & 7 other authors (2000). *Prochlorococcus marinus* Chisholm *et al*. 1992 subsp. *pastoris* subsp. nov. strain PCC 9511, the first axenic chlorophyll a_2/b_2-containing cyanobacterium (*Oxyphotobacteria*). *Int J Syst Evol Microbiol* **50**, 1833–1847.

Robertson, B. R., Tezuka, N. & Watanabe, M. M. (2001). Phylogenetic analyses of *Synechococcus* strains (cyanobacteria) using sequences of 16S rDNA and part of the phycocyanin operon reveal multiple evolutionary lines and reflect phycobilin content. *Int J Syst Evol Microbiol* **51**, 861–871.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Scanlan, D. J., Hess, W. R., Partensky, F., Newman, J. & Vaulot, D. (1996). High degree of genetic variation in *Prochlorococcus* (Prochlorophyta) revealed by RFLP analysis. *Eur J Phycol* **31**, 1–9.

Shimada, A., Nishijima, M. & Maruyama, T. (1995). Seasonal appearance of *Prochlorococcus* in Suruga Bay, Japan, in 1992–1993. *J Oceanogr* **51**, 289–300.

Shimada, A., Maruyama, T. & Miyachi, S. (1996). Vertical distributions and photosynthetic action spectra of two oceanic picophytoplankters, *Prochlorococcus marinus* and *Synechococcus* sp. *Mar Biotechnol* **127**, 15–23.

Tomioka, N. & Sugiura, M. (1984). Nucleotide sequence of the 16S–23S spacer region in the *rrn*A operon from a blue-green alga, *Anacystis nidulans*. *Mol Gen Genet* **193**, 427–430.

Turner, S. (1997). Molecular systematics of oxygenic photosynthetic bacteria. *Plant Syst Evol Suppl* **11**, 13–52.

Urbach, E., Robertson, D. L. & Chisholm, S. W. (1992). Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. *Nature* **355**, 267–270.

Urbach, E., Scanlan, D. J., Distel, D. L., Waterbury, J. B. & Chisholm, S. W. (1998). Rapid diversification of marine picophytoplankton with dissimilar light-harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (cyanobacteria). *J Mol Evol* **46**, 188–201.

Waterbury, J. B. & Rippka, R. (1989). Subsection I. Order *Chroococcales* Wettstein 1924, emend. Rippka *et al*., 1979. In *Bergey's Manual of Systematic Bacteriology*, pp. 1728–1746. Edited by J. T. Staley, M. P. Bryant, N. Pfenning & J. G. Holt. Baltimore: Williams & Wilkins.

Waterbury, J. B., Watson, S. W., Valois, F. W. & Franks, D. G. (1986). Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*. *Can Bull Fish Aquat Sci* **214**, 71–120.

West, N. J. & Adams, D. G. (1997). Phenotypic and genotypic comparison of symbiotic and free-living cyanobacteria from a single field site. *Appl Environ Microbiol* **63**, 4479–4484.

West, N. J. & Scanlan, D. J. (1999). Niche-partitioning of *Prochlorococcus* populations in a stratified water column in the eastern North Atlantic ocean. *Appl Environ Microbiol* **65**, 2585–2591.

West, N. J., Schönhuber, W. A., Fuller, N., Amann, R. I., Rippka, R., Post, A. & Scanlan, D. J (2001). Closely related *Prochlorococcus* genotypes show remarkably different depth distributions in two oceanic regions as revealed by *in situ* hybridization using 16S rRNA-targeted oligonucleotides. *Microbiology* **147**, 1731–1744.

Williamson, S. E. & Doolittle, W. F. (1983). Gene for $tRNA^{Ile}$ and

tRNA^{Ala} in the spacer between the 16S and 23S rRNA genes of a blue-green alga: strong homology to chloroplast tRNA genes and tRNA genes of *E. coli rrnD* cluster. *Nucleic Acids Res* **11**, 225–235.

Wilmotte, A. (1994). Molecular evolution and taxonomy of the cyanobacteria. In *The Molecular Biology of Cyanobacteria*, pp. 1–25. Edited by D. A. Bryant. Dordrecht: Kluwer.

Wilmotte, A., Van der Auwera, G. & De Wachter, R. (1993). Structure of the 16S ribosomal RNA of the thermophilic cyanobacterium *Chlorogloeopsis* HTF ('*Mastigocladus laminosus* HTF') strain PCC7518, and phylogenetic analysis. *FEBS Lett* **317**, 96–100.

Worden, A. Z., Chisholm, S. W. & Binder, B. J. (2000). *In situ* hybridization of *Prochlorococcus* and *Synechococcus* (marine cyanobacteria) spp. with rRNA-targeted peptide nucleic acid probes. *Appl Environ Microbiol* **66**, 284–289.

Received 28 June 2001; revised 21 September 2001; accepted 2 October 2001.