Vol. 45: 263–275, 2006 | AQUATIC MICROBIAL ECOLOGY
Aquat Microb Ecol | Published December 21

High resolution genetic diversity studies of marine *Synechococcus* isolates using *rpoC1*-based restriction fragment length polymorphism

Martin Mühling[1, 2,*], Nick J. Fuller[2], Paul J. Somerfield[1], Anton F. Post[3], William H. Wilson[1], David J. Scanlan[2], Ian Joint[1], Nicholas H. Mann[2]

[1]Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth PL1 3DH, Devon, UK
[2]Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, Midlands, UK
[3]InterUniversity Institute for Marine Science, Coral Beach, 88103 Eilat, Israel

ABSTRACT: *Synechococcus*-specific polymerase chain reaction (PCR) primers were developed and used to amplify fragments of the RNA polymerase core subunit gene *rpoC1* of a set of 23 marine isolates from different oceanographic locations. Restriction fragment length polymorphism (RFLP) analysis of the PCR products differentiated between most strains and all phylogenetic clades, thus allowing the screening of large sample sets at high genetic resolution. The method was used to analyse changes in the genetic diversity of marine *Synechococcus* strains through an annual cycle at 100 m depth in the Gulf of Aqaba, Red Sea, and along the depth profile of a sample from June 1999. RFLP analysis of 36 clones of each of the 8 *rpoC1* clone libraries screened showed that the *Synechococcus* population during summer and autumn was dominated by 1 to 3 genetically different clones, while maximum genetic richness was found during spring and winter. Analysis of the depth profile during the period of stratification revealed differences in relative abundance of particular RFLP types among depths (10, 30, 100 m), although the same genotypes dominated the assemblage throughout. Multivariate analysis of environmental variables and *Synechococcus* assemblage composition showed that *Synechococcus* genetic diversity was most closely, and significantly, correlated to a combination of temperature and NO_3^- ($\rho = 0.645$, $p = 0.038$). Phylogenetic analysis of the marine *rpoC1* nucleotide sequences of the 40 different RFLP types identified in this study, and all *Synechococcus rpoC1* sequences in the databases, demonstrates that the new PCR primers amplify *rpoC1* gene fragments from all MC-A and some *Cyanobium* lineages of marine *Synechococcus* strains and are also useful for the identification of potentially novel lineages.

KEY WORDS: *Synechococcus* spp. · Marine · *rpoC1* · Genetic diversity · Population dynamics

INTRODUCTION

Cyanobacteria belonging to the genus *Synechococcus* are widely distributed and abundant in marine environments from tropical to polar regions, and under oligotrophic, mesotrophic and eutrophic conditions (Gradinger & Lenz 1995, Partensky et al. 1996, 1999). The dominant *Synechococcus* group found in the open ocean and coastal regions (Cluster A) possesses phycoerythrin as the major light-harvesting protein which can be spectrally diverse due to variations in the ratio of the phycoerythrin chromophores, phycourobilin (PUB) and phycoerythrobilin (PEB) (Wood et al. 1999). Marine Cluster A *Synechococcus* have already been shown, in both natural populations and cultured isolates, to be genetically diverse (Ferris & Palenik 1998, Rocap et al. 2002, Scanlan & West 2002, Fuller et al. 2003), comprising at least 10 phylogenetically distinct lineages. Recent efforts to map the spatial distribution of marine *Synechococcus* genotypes have aimed to identify whether these specific lineages show coherence with respect to the *in situ* niche that they occupy, and to identify the environmental parameters that dictate these distributions (Fuller et al. 2003, 2005). Since

*Email: mmuh@pml.ac.uk

the genetic signature, and hence physiology, of cultured isolates reflects the environmental niche occupied by that isolate *in situ* this is an important step in defining meaningful genotype–phenotype characteristics for this genus.

Studies on the genetic population structure of marine *Synechococcus* strains have often used the 16S rRNA gene as a molecular marker. In 2 recent studies, Fuller et al. (2003, 2005) used dot-blot hybridisation of polymerase chain reaction (PCR)-amplified 16S rDNA fragments with clade-specific probes to demonstrate that strains comprising a single *Synechococcus* clade were dominant throughout the water column during an annual cycle in the Gulf of Aqaba, Red Sea, in 1999. However, these studies targeted only those clades previously recognised from cultured isolates or environmental sequences (Fuller et al. 2003) and may thus have 'missed' some diversity. The 16S rDNA approach may also miss more subtle higher resolution changes in population structure due to the fact that a large part of the 16S rRNA gene is relatively conserved in the whole bacterial phylum.

In contrast, it has been demonstrated that molecular markers such as RuBisCO (Chen et al. 2004), the 16S-23S ribosomal DNA internal transcribed spacer sequences (ITS: Ernst et al. 2003, Rocap et al. 2002), or the *ntcA* gene (Penno et al. 2006) provide higher genetic resolution. Another marker gene that has been shown to have a high genetic resolution for marine cyanobacterial diversity is the RNA polymerase core subunit gene *rpoC1* (Palenik 1994, Toledo & Palenik 1997, Ferris & Palenik 1998, Toledo et al. 1999, Ma et al. 2004). For example, 16S rDNA nucleotide sequences of *Synechococcus* Isolates WH7805 and WH8103 show 1.4% sequence divergence along a 1400 bp fragment, whilst a 612 bp fragment of their *rpoC1* genes differs in 17% of their bases (Palenik 1994). To assess whether different gene loci reveal the same general trends in genetic diversity in natural populations, and to increase the genetic resolution of marine *Synechococcus* population-genetics data, we use and improve primers encoding this gene marker. The new method presented here uses restriction fragment length polymorphism (RFLP) to screen *rpoC1* fragment clone libraries constructed from environmental DNA extracted from seawater. Samples were taken from 100 m depth 8 times during an annual cycle, and from a depth profile during June 1999, in the Gulf of Aqaba.

MATERIALS AND METHODS

Strains. Details of strains utilised, the sampling site and flow cytometry analysis have been described by Fuller et al. (2003, 2005) (see summary in Table 1).

Sampling. Samples were collected from the Gulf of Aqaba (Stn A: 29° 28′ N, 34° 55′ E) between March 1999 and January 2000: 5 l of seawater collected from a depth of 100 m were filtered onto a Supor®-450 filter (Gelman Sciences) with a 0.45 µm pore size, and total environmental DNA was isolated from the filters following the method of Fuller et al. (2003).

Preparation of *rpoC1* clone libraries and RFLP analysis. Eight *Synechococcus* clone libraries were prepared from PCR-amplified *rpoC1* fragments using 2 primer sets. The first set was composed of *rpoC1*-N5 (5'-GGNCCNKCBAARGAYTGGGA) and the C-terminal (reverse) primer (*rpoC1*-C) of Palenik (1994). To compensate for potential differences in the PCR conditions between reactions, 2 independent PCRs were carried out and the reactions were subsequently pooled. Aliquots of the combined PCR products were used as templates for 2 independent re-PCRs with a nested primer set (*rpoC1*-39F: 5'-GGCATYGTYTGYGAGCGYTG; *rpoC1*-462R: 5'-CGYAGRCGCTTGRTCAGCTT) designed to be biased to marine *Synechococcus* isolates (see 'Results'). The PCR reactions with both primer sets were carried out in 50 µl volumes containing 2 mM $MgCl_2$, 0.2 mM dNTPs and 2.5 U of Taq DNA polymerase (GIBCO BRL™, Life Technologies). PCRs with primer set *rpoC1*-N5/C contained each primer in a 1 µM concentration and 20 µg of bovine serum albumin (BSA) while the re-PCR with primers *rpoC1*-39F/462R contained the primers in a 20 nM concentration and no BSA. Both PCRs used the same cycle protocol: following an initial denaturation step of 5 min at 95°C, 30 PCR cycles were performed (95°C for 60 s, 51°C for 60 s, 72°C for 60 s) followed by a final extension step at 72°C for 10 min. PCR products of the 2 combined independent re-PCRs were cloned into TOPO TA-vector PCR2.1 (Invitrogen).

The *rpoC1* fragments of 36 clones of each of the libraries were amplified using PCR with Primer Pair *rpoC1*-39F/462R under the same conditions as described above, and were screened by RFLP using an initial digest with *Bst*UI (60°C), followed by a double digest with *Hae*III and *Bcl*I in the presence of added salts at 37°C. The selection of enzymes was based on *in silico* digests of known *rpoC1* fragments using the Webcutter 2.0 software (Heiman 1997; available at www.firstmarket.com/cutter/cut2.html). The different size fragments were separated electrophoretically on a 12% (w/v) polyacrylamide gel. The gels were stained with SybrGold (40 min at RT in TAE buffer, in the dark), followed by de-staining for 10 min in TAE buffer. The DNA banding pattern was observed using UV light. The nucleotide sequence of the 403 bp *rpoC1* fragment was determined for at least 1 member of each of the 40 different RFLP types detected during the study. Twenty different clones of RFLP Type S19 and 5

different clones of RFLP Type S12 were sequenced to ensure that clones of these 2 abundant RFLP patterns are indeed the same in terms of their nucleotide sequence; the greatest difference between clones with the same RFLP pattern was found for RFLP Type S19 with a maximum of 4 bases over the length of the 403 bp fragment.

Good's nonparametric coverage estimator was used to calculate the coverage obtained for the clone libraries: $C = 1 - (n_i/N) \times 100$, where N = total number of clones screened and n_i = number of RFLP types that occurred only once among the 36 clones tested (Good 1953, Good & Toulmin 1956).

Sequence analysis. Double-stranded plasmid DNAs containing the *rpoC1* fragments were sequenced bi-directionally using an ABI 373A automated sequencer (Applied Biosystems) or on a LI-COR (LI-COR) automated sequencer, and both strands were sequenced in full. Sequences are deposited at EMBL under Accession Nos. AJ621015 to AJ621023 (isolates), AJ621508 to AJ621532 (RFLP types occurring only at 30 and 100 m depth), AJ584713 to AJ584749 (RFLP types occurring at 10 m depth; Mühling et al. 2005).

Phylogenetic analysis. All *rpoC1* sequences that cover the 403 bp fragment amplified by Primers 39F/462R were retrieved from the NCBI sequence database. Additionally, several sequences of the combined set of *rpoC1* sequences were used to screen the metagenomic libraries of marine microbial communities (e.g. Venter et al. 2004, Tringe et al. 2005) for the presence of *Synechococcus rpoC1* gene fragments. Multiple sequence alignments of the nucleotide sequences were performed using the ARB programme (Ludwig et al. 2004). Phylogenetic trees were inferred using ARB or PHYLIP. Trees were calculated from a nucleotide alignment of *rpoC1* gene fragments (403 bases) using the neighbor-joining method within ARB, with Jukes-Cantor corrections and a maximum frequency filter (Ludwig et al. 2004). The nucleotide sequence of the *rpoC1* fragment of freshwater strain *Synechocystis* sp. PCC6803 was used as an outgroup. The confidence of branch points was determined by 3 separate analyses (maximum likelihood, neighbor-joining, maximum parsimony), with multifurcations indicating branch points that were collapsed, until supported in all 3 analyses. Bootstrap confidence values were based on 100 simulated trees using the neighbor-joining method within PHYLIP Version 3.6 (Felsenstein 2002; available at: http://evolution.genetics.washington.edu/phylip.html).

Statistical analysis. Nonparametric multivariate analyses (Clarke 1993, Clarke & Gorley 2006) were conducted on untransformed standardised (percentage) data of *Synechococcus rpoC1* RFLP types using Primer Version 6 (www.primer-e.com). Similarities between samples were calculated using the Bray-Curtis coefficient, and ordinated using nonmetric multidimensional scaling (MDS). Relationships between resemblance (similarity/distance) matrices were tested using the non-parametric Mantel test RELATE, which calculates the significance of a Spearman rank correlation (ρ) between corresponding elements of 2 matrices by permutation. In addition to resemblance matrices derived from data, model matrices were used in RELATE to test for serial pattern (representing samples ordered in a linear sequence) and seasonal pattern (representing samples ordered in a cyclical pattern).

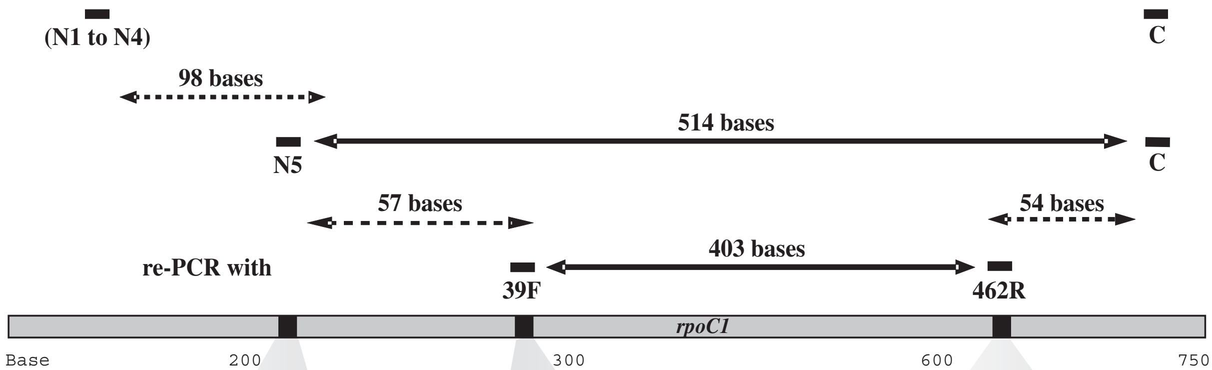
The BIOENV (biota and/or environment matching) routine searches through a set of variables to find subsets that best match the pattern in a predefined resemblance matrix. The best match is defined as the subset of variables that produce a resemblance matrix with the highest rank correlation with the predefined one. Normalised Euclidean distance was the measure used to define resemblance in terms of environmental variables.

RESULTS

Design of marine *Synechococcus*-specific *rpoC1* primer set

Alignment of *rpoC1* nucleotide sequences from several marine *Synechococcus* and *Prochlorococcus* strains retrieved from the NCBI sequence database allowed the design of a 384-fold degenerate forward primer (*rpoC1*-N5). This primer is located 98 bp downstream of forward primer *rpoC1*-N1 previously developed by Palenik (1994). (For schematic representation of primer locations within the *rpoC1* gene see Fig. 1). This primer was used for PCR amplification of *rpoC1* fragments from genomic DNA in combination with the reverse primer *rpoC1*-C of Palenik (1994). A second primer pair was designed nested to the first: forward primer *rpoC1*-39F is located 57 bp downstream of primer *rpoC1*-N5, and reverse primer *rpoC1*-462R is located 54 bp upstream of primer *rpoC1*-C. This nested primer pair was used in a re-PCR with an aliquot of the PCR product obtained with primer pair *rpoC1*-N5/-C as template. The expected sizes of the PCR-amplified *rpoC1* fragments were calculated to be 514 and 403 bp (excluding the primer sequences) for Primer Pairs *rpoC1*-N5/-C and *rpoC1*-39F/-462R, respectively.

Special emphasis was put on designing primers specific for *Synechococcus rpoC1* sequences as the genus *Prochlorococcus* usually outnumbers *Synechococcus* spp. in the oligotrophic regions of the oceans, and PCR amplification with unspecific primers would thus lead to *Prochlorococcus-rpoC1*-dominated clone libraries.



DV1	GGA CCA TCC AAA GAC TGG GA	GG**A** ATT GTC TGT GA**A** **A**G**A** TG	AA**A** CT**C** AT**T** AA**A** **A**G**G** **A**TA **A**G
P.marinus	GGC CCG TCT AAA GAT TGG GA	GG**A** ATT GTT TGC GAG CGC TG	AAG CT**T** ATC AA**A** CGT CTA **A**G
MIT9313	GGC CCT TC**A** AAG GAT TGG GA	GGC ATT GTC TGT GAG CGT TG	AAG **T**T**A** AT**T** AAG CGT **T**TG CG
PCC7942	GGT CCT GCG AAG GAC TGG GA	GGC ATT GTC TGC GAG CGT TG	AAG **T**T**A** ATT AA**A** CGT **T**TG CG
PCC7002	GGT CCT GCG AAG GAC TGG GA	GGC ATT GTC TGC GAG CGT TG	AAG **T**T**A** ATC AA**A** CGC CTG CG
WH7803		GGC ATC GTT TGT GA**A** CGT TG	AAG **T**TG ATC AAG CGT CTG CG
WH8018		GGC ATC GTT TGT GAG CGT TG	AAG **T**TG ATC AAG CGC CTG CG
WH8102	GGT CCC TCC AAG GAC TGG GA	GGC ATT GTC TGT GAG CGC TG	AAG CTG ATC AAG CGT CTG CG
C129	GGT CCC TCC AAG GAT TGG GA	GGC ATC GTC TGT GA**A** CGC TG	AAG CTG ATC AAG CGT CTG CG
WH8020	GGT CCC TCT AAG GAT TGG GA	GG**G** ATT GTT TGT GAG CG**G** TG	AAG **T**TG AT**T** AAG CGC CT**T** CG
WH8011	GGT CCC TCC AAG GAC TGG GA	GGC ATT GTC TGT GAG CGC TG	AAG CTG ATC AAG CGT CTG CG
WH8112	GGT CCC TCC AAG GAC TGG GA	GGC ATT GTC TGT GAG CGC TG	AAG CTG ATC AAG CGT CTG CG
WH8113	GGT CCC TCC AAG GAC TGG GA	GGC ATT GTC TGT GAG CGC TG	AAG CTG ATC AAG CGT CTG CG
CC9902	GGC CCT TCT AAG GAC TGG GA	GGC ATT GTT TGT GA**A** CGT TG	AAG CTG ATC AA**A** CGT CT**T** CG
CC9301	GGT CCC TCC AAG GAC TGG GA	GGC ATT GTC TGT GAG CGC TG	AAG CTG ATC AAG CGT CTG CG
CC9703	GGT CCC TCC AAG GAC TGG GA	GGC ATC GTC TGT GAG CGC TG	AAG CTG ATC AAG CGT CTG CG
CC9702	GGT CCC TCC AAG GAC TGG GA	GGC ATT GTC TGT GAG CGC TG	AAG CTG ATC AAG CGT CTG CG
CC9705	GGC CCG TCC AAA GAT TGG GA	GGC ATC GTT TGT GA**A** CGC TG	AAG **T**TG ATC AAG CGT CTG CG
CC9704	GGT CCC TCT AAG GAC TGG GA	GGC ATC GTC TGC GAG CGT TG	AAG CTG ATC AAG CGC CTG CG
CC9615	GGC CCC TCC AAG GAC TGG GA	GGC ATC GTC TGC GAG CGT TG	AAG CTG ATC AAG CGC CTG CG
CC9701	GGC CCT TCC AAA GAC TGG GA	GGC ATC GTT TGT GA**A** CGC TG	AAG **T**TG ATC AAG CGT CTG CG
CC9305-3	GGT CCC TCC AAA GAC TGG GA	GGC ATT GTC TGT GAG CGC TG	AAG CTG ATC AAG CGC **T**TG CG
CC9317	GGC CCG TCC AAG GAC TGG GA	GGC ATT GTC TGT GAG CGC TG	AAG CTG ATC AAG CGC **T**TA CG
CC9318	GGC CCG TCC AAG GAC TGG GA	GGC ATC GTC TGT GAG CGT TG	AAG CT**A** AT**T** AAG CGC **T**TG CG
CC9311	GGT CCC TCT AAG GAT TGG GA	GGC ATT GTT TGT GAG CGC TG	AAG **T**TG AT**T** AAG CGC CTA CG
CC9617	GGT CCC TCT AAG GAT TGG GA	GGC ATC GTT TGT GAG CGC TG	AAG **T**TG AT**T** AAG CGC CTG CG
CC9616	GGT CCC TCC AAG GAC TGG GA	GGC ATC GTC TGC GAG CGT TG	AAG CTG ATC AAG CGC CTG CG
CC9605	GGC CCG TCC AAA GAT TGG GA	GGC ATC GTT TGT GA**A** CGC TG	AAG **T**TG ATC AAG CGT CTG CG
PCC6307	GGC CCC TCC AAG GAC TGG GA	GGC ATC GTC TGT GAG CGC TG	AAG CTG ATC AAG CGC CTG CG
SS9401	GGC CCG TC**A** AAA GAC TGG GA	GGC ATC GTT TGC GA**A** CGC TG	AAG **T**TG ATC AAG CGT CTG CG
AMINO ACID	G P A/S K D W E	G I V C E R C	K L T/I K R L R
PRIMER	GGN CCN KCB AAR GAY TGG GA	GGC ATY GTY TGY GAG CGY TG	AAG CTG AYC AAG CGY CTR CG

Fig. 1. Schematic representation of primer locations within *rpoC1* gene and nucleotide sequence of priming sites in representative *Synechococcus* strains and the 3 *Prochlorococcus rpoC1* gene fragments available at the time of the primer design: DV1 (Accession No. Z11159), *P. marinus* Strain LG (Z11160), MIT9313 (NC_005071). *Synechococcus* Strain PCC7942 is a freshwater isolate, and PCC7002 is a representative *rpoC1* sequence of a MC-C (*Synechococcus* Cluster 3) isolate. Actual nucleotide sequence of Reverse Primer *rpoC1*-462R is reverse complement of that shown. Nucleotides in boldface indicate mismatches with the primers. Amino acid sequence refers only to *rpoC1* of majority of *Synechococcus* sequences

Therefore, degenerate bases were avoided in the primer sequences where possible. Fig. 1 also shows an alignment of the nucleotide sequence of the priming sites within the *rpoC1* gene of *Synechococcus* strains. Mismatches between the primers and the *rpoC1* sequences are mainly G·T mismatches. These, however, should not affect the annealing of the primers, as G·T mismatches have been shown to form stable hydrogen-bonded wobble pairs (Allawi & SantaLucia 1997). Furthermore, the mismatches do not occur within the last three 3′-end bases which are thought to be critical for the polymerase to begin extending the DNA strand.

The primer pairs were tested on a set of 23 *Synechococcus* strains (Table 1) including 19 recent isolates (Fuller et al. 2003) as well as 7 *Prochlorococcus* strains (EQPAC1, NATL1, NATL2, SS120, TAK9803, MIT9303, PCC9511) and *Escherichia coli* Strain DH5α. The oceanographic origins of the *Synechococcus* strains were highly diverse and included the Pacific and Atlantic Oceans, the Mediterranean Sea and the North Sea (Table 1). PCR-amplification using *rpoC1*-N5/-C was successful for all templates tested. Re-amplification with the nested PCR-primer set *rpoC1*-39F/-462R resulted in products for all the *Synechococcus* strains tested and for *Prochlorococcus* Strain MIT9303, but none of the other *Prochlorococcus* strains or of *E. coli*. However, *rpoC1* sequences of neither *Prochlorococcus* Strain MIT9303 nor any other *Prochlorococcus* strain were detected within any of the clone libraries screened (see 'Discussion'). The re-amplification yielded only specific products, i.e. single bands as judged by electrophoresis on agarose gels. To confirm that the PCR products were indeed *rpoC1* fragments, and to determine their phylogenetic position, the PCR products were cloned and their nucleotide sequences determined. Phylogenetic analysis of these sequences indicates the presence of 9 or more clusters among the 23 strains (Fig. 2).

We also searched the environmental metagenomic libraries for *Synechococcus* sequences. The analysis revealed 3 clones (Accession Nos. AACY01168569, AACY01087037, AACY01256525) from the Sargasso Sea sample (Venter et al. 2004) with sequence similarity to the *rpoC1* sequences used for the search. However, the latter 2 clones extend only 58 and 337 bases, respectively, downstream of primer *rpoC1*-39F, and thus do not cover the site of the reverse primer. The nucleotide sequence at the priming sites in the 3 environmental DNA clones was homologous to that of the PCR primers (as far as the sequences extended to the priming sites).

Alignment of *rpoC1* sequences (including the new sequences) was also the basis for the selection of restriction endonucleases to distinguish strains or, at least, phylogenetic clades by RFLP. A combination of 3 enzymes (*Bst*UI, *Hae*III, *Bcl*I) was found to distinguish between most strains, and all phylogenetic clades (Fig. 2). Subsequent RFLP analysis of the cloned *rpoC1* fragments confirmed the *in silico* restriction digest analysis: the 23 strains were resolved into 9 different RFLP types, their restriction fragment pattern reflecting their phylogenetic positions (Fig. 2).

Table 1. *Synechococcus* strains analysed in this study. Data on 16S rDNA clades from Fuller et al. (2003). *rpoC1* clade nominations based on 16S rDNA phylogeny. na: not available

Strain	Geographic origin	16S rDNA and *rpoC1* clade nomination	Source of *rpoC1* data
CC9311	Pacific	I	Toledo & Palenik (1997)
WH8103	Atlantic	III	Palenik (1994)
Minos2	Mediterranean	III	This study
Minos12	Mediterranean	III	This study
Max42	Sargasso	III	This study
WH7803	Atlantic	V	Palenik (1994)
UW01	na	V	This study
WH8018	Atlantic	VI	This study
Minos1	Mediterranean	X	This study
Minos11	Mediterranean	X	This study
NS01[a]	North Sea	–[a]	This study
Almo3	Mediterranean	I	This study
RS9901	Gulf of Aqaba	IX	This study
RS9902	Gulf of Aqaba	II	This study
RS9903	Gulf of Aqaba	II	This study
RS9904	Gulf of Aqaba	II	This study
RS9905	Gulf of Aqaba	III	This study
RS9906	Gulf of Aqaba	VIII	This study
RS9907	Gulf of Aqaba	II	This study
RS9908	Gulf of Aqaba	II	This study
RS9911	Gulf of Aqaba	II	This study
RS9912	Gulf of Aqaba	II	This study
RS9914	Gulf of Aqaba	VIII	This study

[a]Strain NS01 has been shown to belong to genus *Cyanobium* by Fuller et al. (2003) and Rippka et al. (2001)

Depth profile of *Synechococcus* diversity at Stn A in the Gulf of Aqaba

The RFLP approach was subsequently used to analyse the genetic structure of the *Synechococcus* population in the the Gulf of Aqaba. Pro-

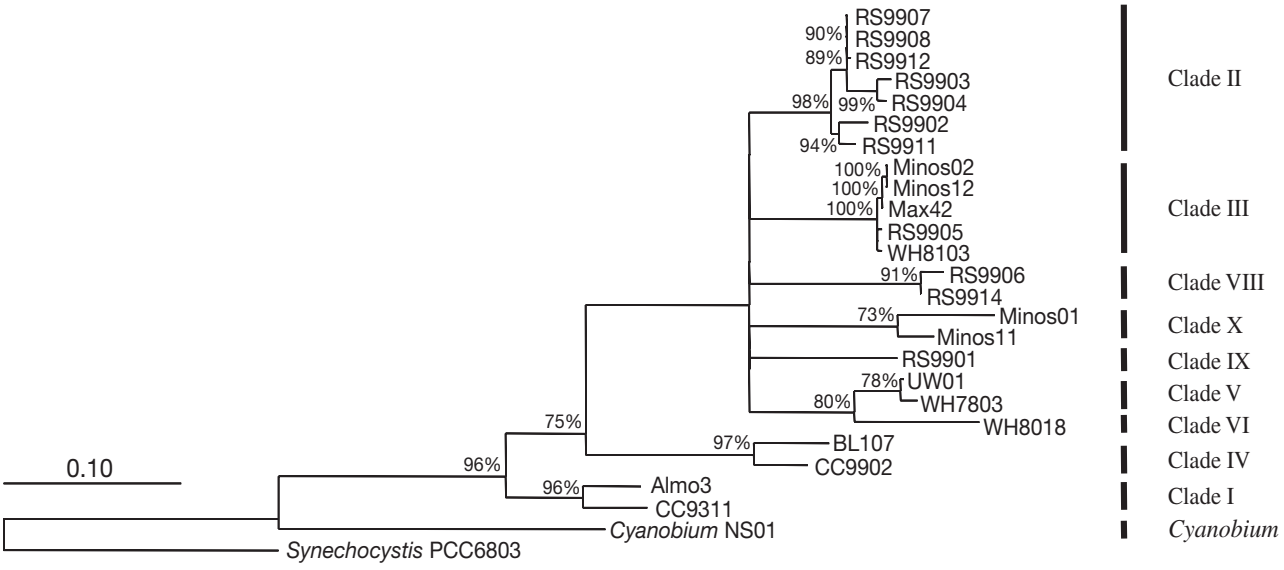


Fig. 2. *Synechococcus* isolates. Phylogenetic analysis of nucleotide sequences of *rpoC1* gene fragments. Tree calculated from nucleotide alignment of *rpoC1* gene fragments (403 bases) using neighbor-joining method within ARB (Ludwig et al. 2004). Freshwater strain *Synechocystis* sp. PCC6803 used as outgroup. Confidence of branch points was determined by 3 separate analyses (maximum likelihood, neighbor-joining, maximum parsimony), with multifurcations indicating branch points that were collapsed until supported in all 3 analyses. Percentages on branches indicate bootstrap values (calculated using neighbor-joining method; 100 replicates), but those <70% have been omitted. Bars to right of tree indicate clades defined by 16S rDNA phylogeny (Rocap et al. 2002, Fuller et al. 2003) and strains with identical RFLP pattern

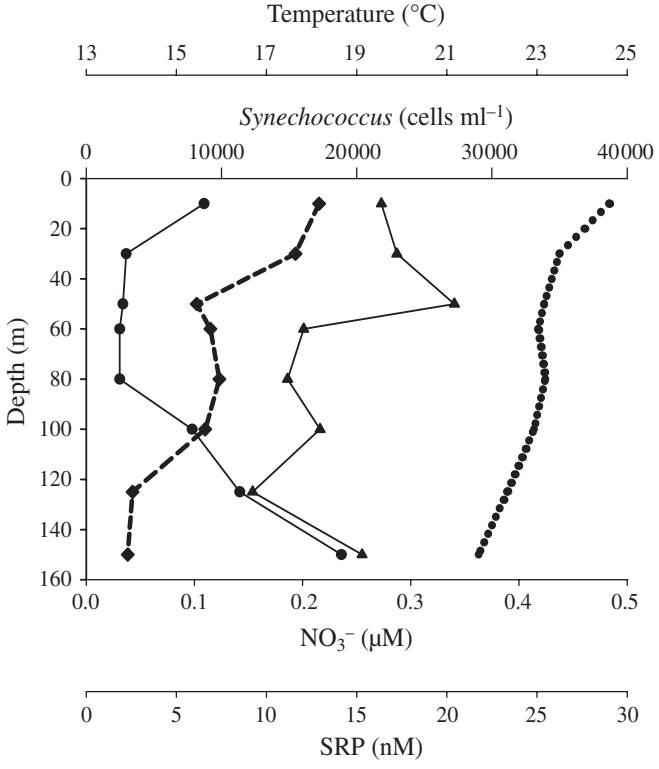


Fig. 3. Vertical profiles of water column on 14 June 1999. (◆) *Synechococcus* spp. cell number, (••••••) temperature, (●) nitrate, (▲) soluble reactive phosphate (SRP)

files of temperature from samples taken on 14 June 1999 revealed a stratified water column at Stn A (Fig. 3), located at the northern tip of the Gulf of Aqaba. The temperature profile identified a shallow surface warming down to about 30 m, then a surface mixed layer (SML) down to about 80 m. *Synechococcus* cell density declined with increasing depth, probably in response to lower temperatures and irradiance rather than a decrease in nutrient supply, as both soluble reactive phosphate (SRP) and nitrate concentrations were similar at 10 and 100 m depth (Fig. 3). There were differences in the genetic diversity between samples from the 3 different depths (Table 2). Although the same RFLP types dominated at each depth (RFLP Types S4, S12 and S19), the relative abundance of these clones found in each of the depth libraries varied. For example, RFLP Type S12 was found in one-third of the clones tested in the clone libraries from 10 and 30 m depth, while it constituted half of the clones in the clone library prepared from the 100 m sample. In contrast, the relative abundance of RFLP Type S19 within the clone libraries decreased significantly from 10 m (41%) to 30 m (33%) and 100 m (25%) (Table 2). The 3 dominant RFLP types (S4, S12 and S19) were also the only clones detected in all libraries from different depths (Table 2). There were more unique RFLP types at 10 m (4) than at 30 m (2) or 100 m (1) depth, indicat-

Table 2. *Synechococcus* RFLP types detected in depth profile for 14 June 1999. Occurrence (N) within clone libraries; 36 clones of each of the *rpoC1* clone libraries were screened by RFLP. pn: potentially novel clade; –: RFLP type not detected in this depth sample

RFLP type	Clade	10 m	30 m	100 m
S3	II	2	–	–
S4	pn	4	8	8
S12	III	11	12	18
S18	pn	1	–	–
S19	pn	15	13	9
S27	pn	1	1	–
S28	X	1	–	–
S29	pn	1	–	–
S31	pn	–	1	–
S33	II	–	2	–
S34	V	–	–	1
Total no. clones analysed		36	36	36
No. different RFLP types		8	8	6	4
No. unique RFLP types		4	4	2	1

ing that an increase in sample size would lead to many more types being discovered at 10 m compared to deeper depths.

Seasonal changes in *Synechococcus* population in Gulf of Aqaba

Flow cytometry analysis indicated that *Synechococcus* cell numbers were high during spring and winter but low during summer and autumn months (Fig. 4A). There were reduced phosphate and nitrate concentrations between May and October (Fig. 4A), which seems likely to have been (at least partially) responsible for the reduced *Synechococcus* cell numbers (see below).

The RFLP method was subsequently applied to test whether the observed seasonal changes in *Synechococcus* population size in the Gulf of Aqaba were accompanied by changes in genetic diversity. To cover an annual cycle, 8 *rpoC1* clone libraries were prepared from samples collected from 100 m depth between March 1999 and January 2000. Screening of the clone libraries by RFLP showed that there was high genetic diversity in the spring (March, April), which then declined towards summer and autumn (June to October) when only 3 or fewer RFLP types (out of 5 abundant RFLP types: S4, S12, S19, S25, S37) dominated the population (Fig. 4B, Table 3). In January the population was again genetically very diverse (Fig. 4B). Of the 36 clones from each of the libraries screened by RFLP there were between 1 (14 June) and 9 (19 April) RFLP types that were found only once. The coverage of these libraries thus varied between ca. 97% (14 June) and 75% (19 April) for these samples, although it was ca. 82% or above in all other cases.

Besides the 3 abundant RFLP types (S4, S12, S19) that dominated the *Synechococcus* assemblage in spring and summer (March to August) only 2 others (S25, S37) were found in more than 2 of the clone libraries. All other RFLP types were found in 1 or 2 libraries only and, in most cases, only in low numbers. RFLP types S25 and 37, which were not detected in any of the surface clone libraries (Mühling et al. 2005, Table 3), dominated the *Synechococcus* assemblage at 100 m depth from August to October (Table 3). The pattern of occurrence of the 5 abundant genotypes within the clone libraries indicated a succession in the dominance of these *Synechococcus* RFLP types over the annual cycle (Table 3). MDS also revealed evidence of a cyclical change through the year in the genetic composition of the *Synechococcus* population (Fig. 5). In this context a perfect cyclical (or seasonal) pattern would be represented by a 2D plot with samples arranged in a circular pattern and zero stress. Although stress values are low (<0.1), indicating excellent representations of multivariate relationships in 2D (Clarke 1993), they are not zero, but evidence of seasonal changes in assemblage structure is apparent. RELATE analyses, based on the underlying similarity matrices, are a more objective

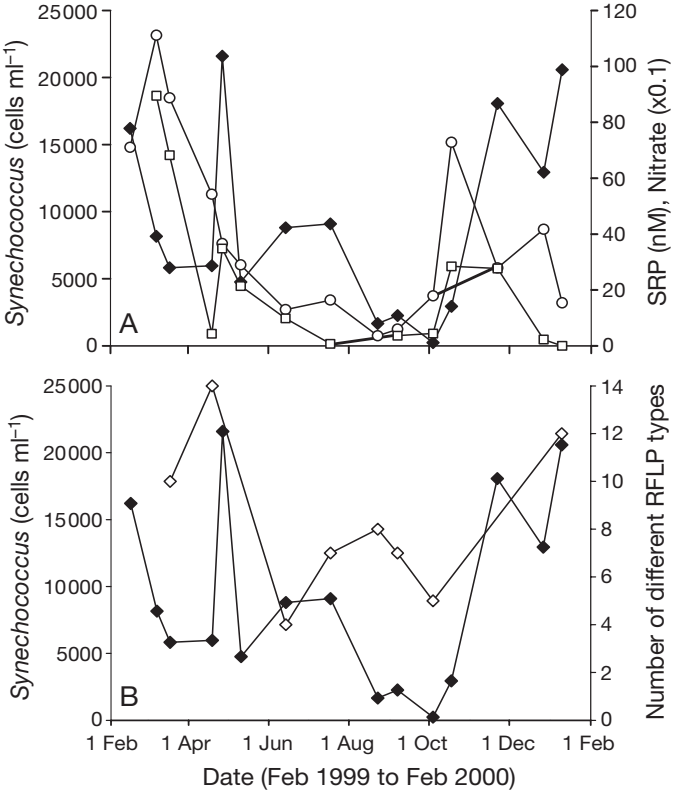


Fig. 4. *Synechococcus* spp. changes in abundance over annual cycle at 100 m depth in comparison with (A) soluble reactive phosphate (SRP, ○) and nitrate (□), and (B) genetic diversity as number of different RFLP types detected in clone libraries (◇). (◆) *Synechococcus* cell numbers

Table 3. Annual cycle at 10 and 100 m depth; data are shown only for those RFLP types that occurred in more than two of the eight 100 m depth libraries tested; Others: number of other RFLP types detected among remaining clones (number in parentheses) in same sample; data of 10 m *rpoC1* clone libraries from Mühling et al. (2005). –: RFLP type not detected

Sampling date	RFLP type S4	S12	S19	S25	S37	Others
10 m						
18 Mar 99	4	–	–	–	–	9 (19)[a]
19 Apr 99	7	9	6	–	–	8 (14)
14 Jun 99	4	11	15	–	–	5 (6)
18 Jul 99	–	5	31	–	–	0 (0)
23 Aug 99	–	15	19	–	–	1 (2)
07 Sep 99	1	19	11	–	–	4 (5)
04 Oct 99	3	14	17	–	–	1 (2)
10 Jan 00	4	–	3	–	–	15 (29)
100 m						
18 Mar 99[b]	13	1	–	–	–	8 (22)
19 Apr 99	–	8	6	1	–	11 (21)
14 Jun 99	8	18	9	–	–	1 (1)
18 Jul 99	1	11	19	–	–	4 (5)
23 Aug 99	–	1	25	3	3	4 (4)
07 Sep 99	–	1	8	8	14	4 (6)
04 Oct 99	–	10	–	–	13	3 (13)
10 Jan 00	6	3	6	–	–	9 (21)

[a]Only 23 clones were screened for this clone library (see Mühling et al. 2005)
[b]Lower depth sample for March 1999 was obtained from 80 not 100 m depth

method for examining patterns. A seriation test (ρ = 0.278, p = 0.067) and a cyclicity test (ρ = 0.287, p = 0.019) indicated a significant seasonal pattern in the dynamics of the RFLP types at 100 m, rather than a simple temporal autocorrelation. While patterns in RFLP types at 10 m (Fig. 5A; data from Mühling et al. 2005) and at 100 m (Fig. 5B) are significantly related (ρ = 0.489, p = 0.037) they are not identical, with further RELATE tests indicating significant temporal autocorrelation (seriation test ρ = 0.425, p = 0.023) in assemblage structure based on relative abundances at 10 m.

These multivariate patterns are driven by changes in relative abundances of RFLP types between sampling dates. During early spring (March 1999) the population was dominated by RFLP Type S4 (Table 3), whose dominance decreased towards June when RFLP Type S12 became the dominant genotype. During the summer months of July and August RFLP Type S19 was the dominant genotype, being replaced during autumn (September, October) by RFLP Types S25 and S37. In winter (January) the *Synechococcus* population was genetically very diverse (Fig. 4B) and none of the RFLP types seemed to dominate the population, although RFLP Type S4 was again one of the more abundant clones within the population (Table 3).

Multivariate analysis was employed to test which environmental parameter(s) might be responsible for the observed seasonal succession of *Synechococcus* genotypes. BIOENV was used to determine the subset of measured environmental variables (temperature, salinity, SRP, NO_3^-, NO_2^- and total organic nitrogen) which most closely matched the pattern in *Synechococcus* genetic composition. A combination of temperature and NO_3^- provided the closest match (ρ = 0.645, p = 0.038).

Phylogenetic analysis of different *rpoC1* RFLP types

We detected 35 different RFLP patterns in the samples from 100 m depth (Table 3). A further 3 different RFLP patterns (S27, S28, S29) were detected in the 10 m sample, and another 2 (S31, S33) in the 30 m sample, from 16 June (Table 2). A phylogenetic analysis

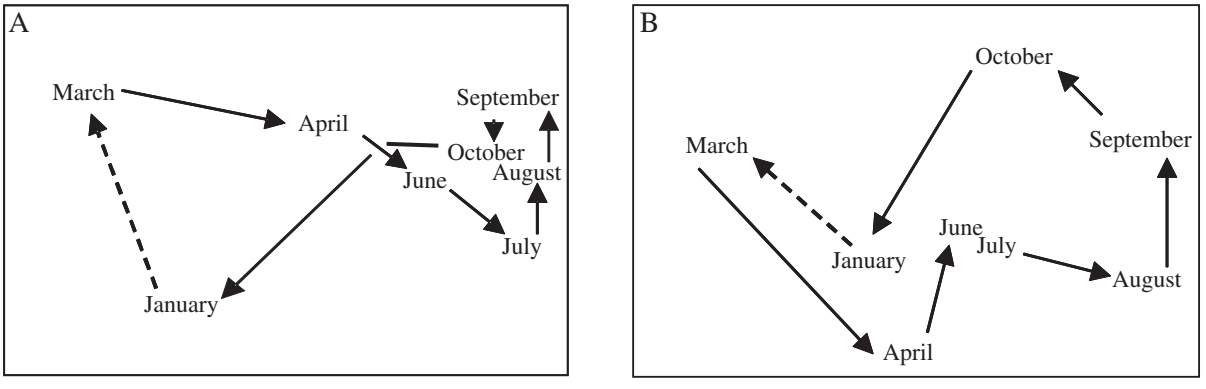


Fig. 5. *Synechococcus* isolates. Nonmetric multidimensional scaling ordination plots of samples from (A) 10 m and (B) 100 m depth. Analysis based on Bray-Curtis similarities calculated from untransformed standardised *Synechococcus* RFLP types. Stress values (Kruskal's Stress Formula 1) were 0.01 and 0.06 for the 10 and 100 m plots, respectively. Each sample is represented by month in which it was collected, and relative distances between months reflect intersample similarities, with samples close together being similar in terms of genetic composition. (Arrows added to assist interpretation, with continuous arrows linking months for which *Synechococcus* composition was analysed and dashed arrows linking beginning and end of a sequence of samples)

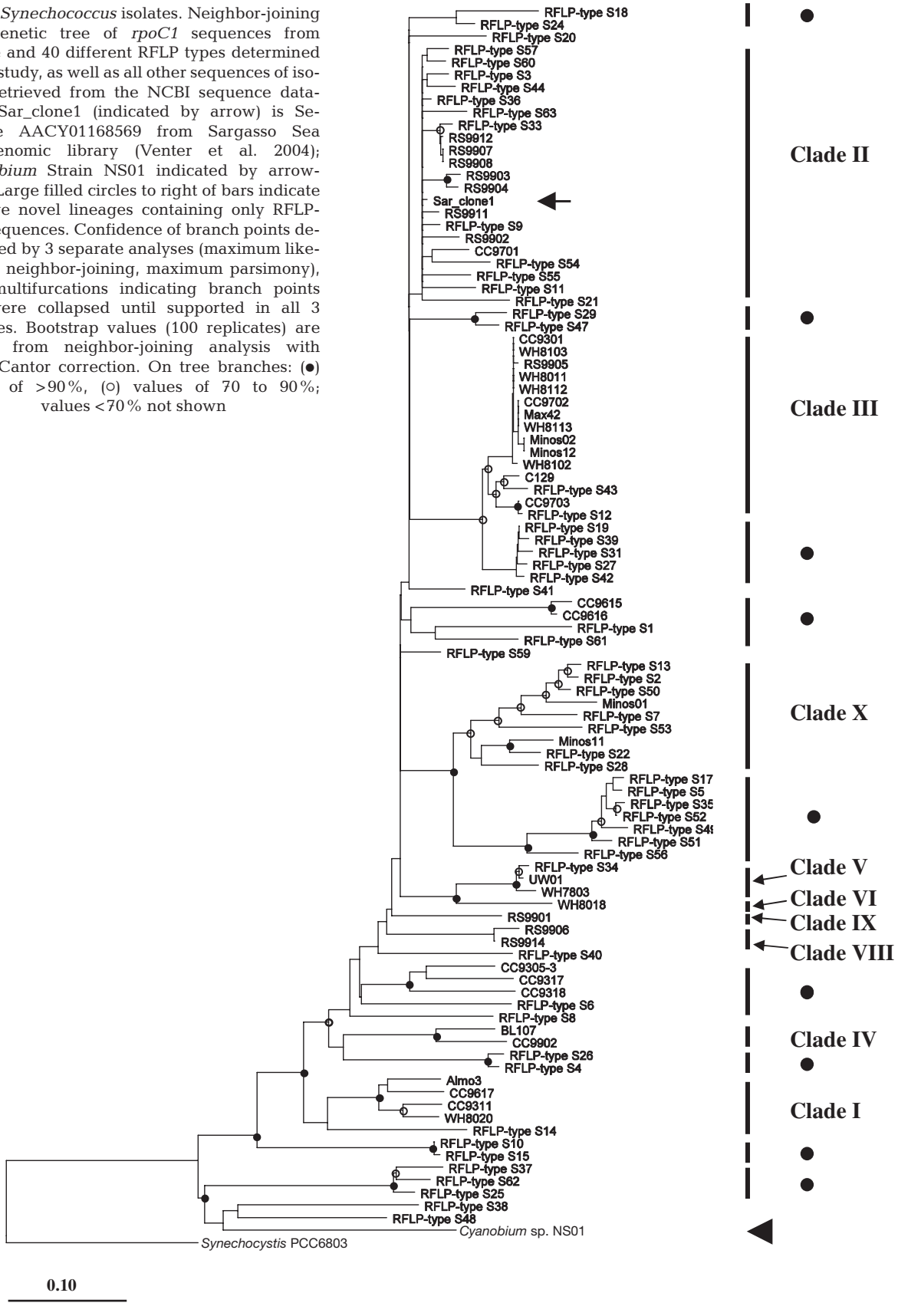


Fig. 6. *Synechococcus* isolates. Neighbor-joining phylogenetic tree of *rpoC1* sequences from marine and 40 different RFLP types determined in this study, as well as all other sequences of isolates retrieved from the NCBI sequence database. Sar_clone1 (indicated by arrow) is Sequence AACY01168569 from Sargasso Sea metagenomic library (Venter et al. 2004); *Cyanobium* Strain NS01 indicated by arrowhead. Large filled circles to right of bars indicate putative novel lineages containing only RFLP-type sequences. Confidence of branch points determined by 3 separate analyses (maximum likelihood, neighbor-joining, maximum parsimony), with multifurcations indicating branch points that were collapsed until supported in all 3 analyses. Bootstrap values (100 replicates) are shown from neighbor-joining analysis with Jukes-Cantor correction. On tree branches: (●) values of >90%, (○) values of 70 to 90%; values <70% not shown

was carried out based on the *rpoC1* nucleotide alignment of these 40 different RFLP types, together with those listed in Table 1 and other representatives of *rpoC1* sequences in the databases, including those from the surface clone libraries (Mühling et al. 2005). The analysis indicates that the *rpoC1* RFLP types from the Gulf of Aqaba represent a wide range of genetic diversity (Fig. 6) with representatives from 9 of the 10 phylogenetic clades identified using 16S rDNA sequence data of 45 isolates (Fuller et al. 2003). Clade VII could not be identified due to the lack of a representative isolate with sequence data for both marker genes. Moreover, several of the sequences were resolved in lineages that did not contain any sequences derived from *Synechococcus* isolates, suggesting that these define novel clades (Fig. 6), e.g. lineages containing RFLP Types S15, S19, S29, S35, whilst there are also other singleton sequences that we cautiously suggest might represent novel lineages, e.g. S14, S20. Again, it should be considered that cultured isolates from a previously identified clade, VII (see Fuller et al. 2003), were not available for sequencing in this study, but are potentially contained within these RFLP types.

DISCUSSION

rpoC1 PCR primers and phylogenetic resolution

The newly developed nested PCR primers proved to be highly specific for *Synechococcus* spp. as, despite being 1 order of magnitude more abundant, no *Prochlorococcus* spp. sequence was detected within any of the clone libraries. This specificity for *Synechococcus* was achieved by putting special emphasis on designing primers with a low degree of degeneracy. As G·T mismatches have been shown to form stable hydrogen-bonded wobble pairs (Allawi & SantaLucia, 1997) they should not affect the annealing of the primers. We therefore decided to include only guanine in the primer sequences rather than having an equal mixture of guanine and adenine (Fig. 1). Furthermore, the mismatches do not occur within the last three 3′-end bases which are thought to be critical for the polymerase to begin extending the DNA strand. Indeed, the fact that isolates such as WH7803, which has a G·T (as well as a G·A) mismatch (Fig. 1), was amplified using the primer set further supports this finding.

The high specificity of the nested primer pair was probably also responsible for the low yield of PCR product used in a PCR directly with environmental DNA as template. This required the application of a nested PCR approach.

We believe that the newly developed nested PCR approach for *rpoC1* amplifies all phylogenetic lineages of marine *Synechococcus* belonging to the MC-A clade (Cluster 5.1) as well as some that belong to the *Cyanobium* clade for the following reasons. Firstly, phylogenetic analysis of *rpoC1* sequences of the different RFLP types shows that these sequences encompass most of the MC-A clades currently known from 16S rRNA gene sequence analysis and for which *rpoC1* sequences from cultured isolates are available (Fig. 6). Secondly, the use of the new *rpoC1* primers also led to the amplification of the *rpoC1* fragment of *Cyanobium* sp. NS01. 16S rDNA sequencing showed this isolate to cluster with *Cyanobium* sp. PCC7001 (Fuller et al. 2003). This was formerly classified in Marine Cluster B (MC-B) of the genus *Synechococcus*, but was recently re-assigned to the genus *Cyanobium* (Rippka et al. 2001). This genus consists of halotolerant isolates, without phycoerythrin, and a high mean DNA base composition, suggesting that they are genetically distinct members of the 'form-genus' *Synechococcus* (Rippka et al. 2001). The fact that *rpoC1* PCR primers amplify the *rpoC1* fragment from this *Cyanobium* isolate makes it likely that other strains of this clade would also be amplified by these primers.

In this study we did not focus on *Synechococcus* strains from freshwater or estuarine environments but on those occurring in oligotrophic regions of the oceans. Therefore, we did not test the primers for their ability to amplify *rpoC1* fragments of strains belonging to the *Synechococcus* MC-C clade. The fact that Reverse Primer *rpoC1*-462R has 3 mismatches to MC-C Strain PCC7002 (Fig. 1) may indicate that it does not amplify this strain, although 2 of these mismatches are G·T mismatches (see above). Therefore, we are not certain whether the nested primer pair is able to amplify MC-C strains, but recommend modification of Reverse Primer *rpoC1*-462R for this purpose.

Using these newly developed *Synechococcus*-specific *rpoC1* primers, a high genetic resolution of both *Synechococcus* isolates and environmental clones was obtained (Figs. 2 & 6), and the diversity is greater than has previously been observed (Ferris & Palenik 1998, Rocap et al. 2002, Fuller et al. 2003). Unique RFLP patterns (data not shown) were obtained for each phylogenetic clade of *Synechococcus* strains (Fig. 2), demonstrating the high degree of genetic resolution that can be derived using *rpoC1* as a molecular marker. Moreover, several RFLP types form potentially new clades within the marine *Synechococcus* radiation (Fig. 6). Similarly, using *ntcA* as a marker gene, Lindell et al. (2005) and Penno et al. (2006) have also recently found new clusters and extensive microdiversity within the *Synechococcus* phylogeny. Therefore, the use of protein-encoding genes as molecular markers, rather than the 16S rRNA gene, provides higher resolution and reveals more detailed phylogenetic structures.

Seasonal changes in *Synechococcus* population size in the Gulf of Aqaba

The factor responsible for the sharp decline in *Synechococcus* abundance at the end of April is unknown, but the classical paradigm suggests that abiotic factors such as nutrient deficiency are important. Recent data using PstS and *ntcA* as indicators of P and N stress, respectively, suggests that P may be the 'limiting' nutrient at this time (Fuller et al. 2005, Lindell et al. 2005). However, the multivariate statistical analysis of Fuller et al. (2005) also found that different *Synechococcus* clades responded somewhat differently to gradients of N and P, although the correlation was only weak. In addition, biotic factors such as viral infection and grazing may also cause shifts in the abundance and genetic diversity of phytoplankton. A recent study in the Red Sea suggests grazing to be of minor importance in controlling *Synechococcus* populations (Sommer et al. 2002). But what about virus control? Although it was originally suggested that marine *Synechococcus* are resistant to co-occurring viruses (Waterbury & Valois 1993), a later study in an offshore environment demonstrated that the viral lysis rate is similar to the contact rate between virus and host and that *Synechococcus* cells are, therefore, not resistant to the co-occurring viruses (Suttle & Chan 1994).

The importance of virus infection was demonstrated by us (Mühling et al. 2005) in surface waters (10 m depth) of the Gulf of Aqaba during the same sampling period as reported herein. Viral control was unequivocally demonstrated to be the cause of the decline of 3 of the 4 *Synechococcus* spp. peaks during the annual cycle. However, although this was clear for surface waters, the number of *Synechococcus*-infecting viruses at 100 m depth (data not shown) did not show a significant correlation with *Synechococcus* abundance or genetic diversity. This may indicate that the scenario of *Synechococcus* abundance being controlled by virus infection does not apply to the *Synechococcus* assemblage at 100 m depth. The host abundance, however, was above the threshold (10^3 cells ml^{-1}) thought to be required for phage selection pressure to be important (Suttle & Chan 1994, Mann 2003).

Succession of *Synechococcus* genotypes

The genetic diversity of the deep *Synechococcus* spp. population in the Gulf of Aqaba showed a seasonal pattern that reflected the succession of a few dominant clones. As at the surface (Mühling et al. 2005), the genetic diversity of *Synechococcus* was highest during spring and winter, when they were also most abundant. Selection pressure during summer and autumn seems to favour a small subset of the *Synechococcus* genetic pool. At 100 m, the dominant RFLP types during the summer and autumn were S12 and S19. The sequence of the S12 type is phylogenetically most closely related to motile *Synechococcus* strains, e.g. WH8103 (Fig. 6), which form a monophyletic group, as determined by both *rpoC1* and 16S rDNA alleles (Toledo et al. 1999, Fuller et al. 2003). Given the oligotrophic state of the Gulf of Aqaba, it seems likely that motility may also be important for nutrient acquisition in the shallow nutrient gradients that exist within the micro-environment of a cell (Willey & Waterbury 1989). This scenario is further confirmed by the fact that the abiotic factors that appeared most closely related to the *Synechococcus* genetic diversity were a combination of temperature and NO_3^-. Interestingly, temperature and NO_3^- are negatively correlated; while the seawater temperature increases from May (21.5°C) to 23–24°C in the summer and autumn months (June to October) the NO_3^- concentration decreases (Fig. 4A). It may be speculated that the increased seawater temperature leads to an increased metabolism of *Synechococcus* which increases the nutrient stress caused by NO_3^- limitation, making motility and the ability to respond to nutrient-rich micro-environments even more important. Should N supply be NH_4^+ sufficient, as shown by Lindell et al. (2005), then motility would still be advantageous for *Synechococcus* in scavenging the low (<60 nM) ambient concentrations of this compound.

From August to the beginning of October, 2 further RFLP types (S25, S37) dominated the *Synechococcus* composition at 100 m; these were not detected in any of the surface clone libraries (Mühling et al. 2005). Interestingly, these RFLP types cluster within the *Cyanobium* clade (Fig. 6), suggesting that these *Synechococcus* RFLP types lack phycoerythrin. Unless phycoerythrin-containing (yet undiscovered) strains that are genetically closely related to the *Cyanobium* clade exist, this hypothesis would be somewhat surprising, as auxiliary pigments such as phycoerythrin are thought to provide a selective advantage at low light intensities. Chen et al. (2006) recently reported the isolation of 5 *Synechococcus* MC-B strains that contain phycoerythrin, indicating that further efforts to isolate novel *Synechococcus* strains may lead indeed to the discovery of *Cyanobium*-like strains that are able to produce phycoerythrin. However, the bootstrap value (98; Fig. 6) of the branching point that separates RFLP Types S25 and S37 (and S62) from the *Cyanobium* clade may indicate that S25 and S37 are not part of the *Cyanobium* clade proper but of a phylogenetically closely related clade composed of currently undescribed strains that contain phycoerythrin.

High resolution molecular population analysis and statistical evaluation of data

This study indicates the potential for, and provides examples of, the application of high resolution genetic analysis using *rpoC1*-RFLP, in combination with multivariate statistical analysis, for the study of marine *Synechococcus* populations. *Synechococcus* assemblages from different times were compared, and the strains identified that were most likely to be responsible for the observed changes. The pattern of change was examined using multivariate statistical tools. The observed changes show 2 features (Fig. 5): (1) The samples closest to each other in time were most similar; this linear pattern indicates a continuous change in the *Synechococcus* genotypes making up the assemblage. (2) The MDS plot (Fig. 5) shows a cyclic pattern, with the representation for January being close to that for the previous March; this cyclic pattern indicates seasonality in the *Synechococcus* genotypes. Furthermore, correlations to biotic (e.g. viruses; see also Mühling et al. 2005) and abiotic parameters were assessed, and these help to draw conclusions about the underlying biological and environmental drivers of the observed changes in population genotypes. In addition, the subset of the population that contributed most to observed correlations can be identified, as shown in the study by Mühling et al. (2005). This use of multivariate statistics demonstrates that it should be possible to investigate microbial communities beyond simple documentation of genetic diversity at particular times and places using low-resolution genetic data. A promising way forward, leading to understanding of patterns and processes and of the ecological relevance of observed changes, is now possible using high-throughput methods to determine high genetic resolution, coupled with relevant statistical analyses which take the dimensionality of the data into account.

Acknowledgements. We are grateful to Dr. J. Waterbury, Dr. B. Palenik and A. Noordeloos for providing *Synechococcus* DNA or cells of some of the strains used in this study. S. Boulben and F. Le Gall are warmly thanked for purifying and maintaining the Roscoff strains over several years. We also thank A. Rivlin and D. Marie for supplying CTD and flow cytometry data, respectively, from the Red Sea. Dr. K. Zwirglmaier and Dr. D. Scanlan are thanked for providing the *rpoC1* sequence of *Synechococcus* Strains Almo3 and BL107. The genome of strain BL107 has been sequenced with financial support from The Gordon and Betty Moore Foundation. This work was supported by NERC grant GR3/12602 and EU grant MIRACLE (EVK3-CT-2002-00087) and is a contribution towards the core strategic research programme of the Plymouth Marine Laboratory.

LITERATURE CITED

Allawi HT, SantaLucia J Jr (1997) Thermodynamics and NMR of internal G·T mismatches in DNA. Biochemistry 36: 10581–10594

Chen F, Wang K, Kan J, Bachoon DS, Lu J, Lau S, Campbell L (2004) Phylogenetic diversity of *Synechococcus* in the Chesapeake Bay revealed by ribulose-1,5-bisphosphate carboxylase-oxygenase (RuBisCO) large subunit gene (rbcL) sequences. Aquat Microb Ecol 36:153–164

Chen F, Wang K, Kan J, Suzuki MT, Wommack KE (2006) Diverse and unique picocyanobacteria in Chesapeake Bay, revealed by 16S-23S rRNA internal transcribed spacer sequences. Appl Environ Microbiol 72:2239–2243

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. Aust J Ecol 18:117–143

Clarke KR, Gorley RN (2006) Primer v6: user manual/tutorial, Primer-e, Plymouth

Ernst A, Becker S, Wollenzien UIA, Postius C (2003) Ecosystem-dependent adaptive radiations of picocyanobacteria inferred from 16S rRNA and ITS-1 sequence analysis. Microbiology Reading (UK) 149:217–228

Ferris MJ, Palenik B (1998) Niche adaptation in ocean cyanobacteria. Nature 396:226–228

Fuller NJ, Marie D, Partensky F, Vaulot D, Post AF, Scanlan DJ (2003) Clade-specific 16S rDNA oligonucleotides reveal the predominance of a single marine *Synechococcus* clade throughout a stratified water column in the Red Sea. Appl Environ Microbiol 69:2430–2443

Fuller NJ, West NJ, Marie D, Yallop M, Rivlin T, Post AF, Scanlan DJ (2005) Dynamics of community structure and phosphate status of picocyanobacterial populations in the Gulf of Aqaba, Red Sea. Limnol Oceanogr 50:363–375

Good IJ (1953) The population frequencies of species and the estimation of the population parameters. Biometrika 40: 237–264

Good IJ, Toulmin G (1956) The number of new species and the increase of population coverage when a sample is increased. Biometrika 43:45–63

Gradinger R, Lenz J (1995) Seasonal occurrence of picocyanobacteria in the Greenland Sea and central Arctic Ocean. Polar Biol 15:447–452

Lindell D, Penno S, Al-Qutob M, David E, Rivlin T, Lazar B, Post AF (2005) Expression of the nitrogen stress response gene *ntcA* reveals nitrogen-sufficient *Synechococcus* populations in the oligotrophic northern Red Sea. Limnol Oceanogr 50:1932–1944

Ludwig W, Strunk O, Westram R, Richter L and 28 others (2004) ARB: a software environment for sequence data. Nucleic Acids Res 32:1363–1371

Ma Y, Jiao NZ, Zeng YH (2004) Natural community structure of cyanobacteria in the South China Sea as revealed by *rpoC1* gene sequence analysis. Lett Appl Microbiol 39: 353–358

Mann NH (2003) Phages of the marine picophytoplankton. FEMS Microbiol Rev 27:17–34

Mühling M, Fuller NJ, Millard A, Somerfield PJ and 6 others (2005) Genetic diversity of marine *Synechococcus* and co-occurring cyanophage communities: evidence for viral control of phytoplankton. Environ Microbiol 7:499–508

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

Partensky F, Blanchot J, Lantoine F, Neveux J, Marie D (1996) Vertical structure of picophytoplankton at different trophic sites of the tropical northeastern Atlantic Ocean. Deep-Sea Res I 43:1191–1213

Partensky F, Hess WR, Vaulot D (1999) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. Microbiol Mol Biol Rev 63:106–127

Penno S, Lindell D, Post AF (2006) Diversity of *Synechococcus* and *Prochlorococcus* populations determined from DNA sequences of the N-regulation gene *ntcA*. Environ Microbiol 8:1200–1211

Rippka R, Castenholz RW, Herdman M (2001) Form-genus IV. Cyanobium Rippka and Cohen-Bazire 1983. In: Boone DR, Castenmholz RW, Garrity GM (eds) Bergey's manual of systematic bacteriology: the Archaea and the deeply branching and phototrophic bacteria, Vol 1. Springer-Verlag, Berlin, p 498–499

Rocap G, Distel DL, Waterbury JB, Chisholm SW (2002) Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S–23S ribosomal DNA internal transcribed spacer sequences. Appl Environ Microbiol 68:1180–1191

Scanlan DJ, West NJ (2002) Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. FEMS Microbiol Ecol 40:1–12

Sommer U, Berninger UG, Böttger-Schnack R, Cornils A and 8 others (2002) Grazing during early spring in the Gulf of Aqaba and the northern Red Sea. Mar Ecol Prog Ser 239: 251–261

Suttle CA, Chan AM (1994) Dynamics and distribution of cyanophages and their effect on marine *Synechococcus* spp. Appl Environ Microbiol 60:3167–3174

Toledo G, Palenik B (1997) *Synechococcus* diversity in the Californian Current as seen by RNA polymerase (*rpoC1*) gene sequences of isolated strains. Appl Environ Microbiol 63:4298–4303

Toledo G, Palenik B, Brahamsha B (1999) Swimming marine *Synechococcus* strains with widely different photosynthetic pigment ratios form a monophyletic group. Appl Environ Microbiol 65:5247–5251

Tringe SG, von Mering C, Kobayashi A, Salamov AA and 9 others (2005) Comparative metagenomics of microbial communities. Science 308:554–557

Venter JC, Remington K, Heidelberg JF, Halpern AL and 19 others (2004) Environmental genome shotgun sequencing of the Sargasso Sea. Science 304:66–74

Waterbury JB, Valois FW (1993) Resistance to co-occurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. Appl Environ Microbiol 59:3393–3399

Willey JM, Waterbury JB (1989) Chemotaxis towards nitrogenous compounds by swimming strains of marine *Synechococcus* spp. Appl Environ Microbiol 55: 1888–1894

Wood AM, Lipsen M, Coble P (1999) Fluorescence-based characterization of phycoerythrin-containing cyanobacterial communities in the Arabian Sea during the Northeast and early Southwest Monsoon (1994–1995). Deep-Sea Res I 46:1769–1790

Editorial responsibility: Curtis Suttle, Vancouver, British Columbia, Canada

Submitted: January 20, 2006; Accepted: September 18, 2006
Proofs received from author(s): November 20, 2006