Picocyanobacterial community structure of freshwater lakes and the Baltic Sea revealed by phylogenetic analyses and clade-specific quantitative PCR

Patricia Sánchez-Baracaldo, Barbara A. Handley and Paul K. Hayes

School of Biological Sciences, Woodland Road, University of Bristol, Bristol BS8 1UG, UK

Phylogenetic relationships among picocyanobacteria from the Syn/Pro clade sensu Sánchez-Baracaldo et al. (2005) were determined using small subunit (ssu) rDNA sequences from novel culture isolates together with environmental samples from the Baltic Sea and seven freshwater lakes. The picocyanobacterial community comprised members of previously identified clades and of two previously undescribed clades. The number of well-supported clades suggests that freshwater picocyanobacterial communities encompass much greater diversity than is found in marine systems. To allow the quantification of community structure and temporal succession, clade-specific ssu rDNA TaqMan assays were designed and implemented. These assays were used to assess picocyanobacterial community structure in two lakes over an annual cycle in 2003/4, and in a small number of Baltic Sea samples collected in July 2003. In the lake-water samples, picocyanobacteria were found to be scarce during most of the year, with members of each clade reaching their peak abundance over a relatively short period during the summer (June to September), although representatives of the Cyanobium clade also developed an autumn peak extending towards the end of October. All four freshwater clades were present in the Baltic Sea, but their distribution was patchy over relatively short spatial scales. The use of molecular tools for describing and quantifying community structures reveals previously unexplored complexity in the phytoplankton and will facilitate the development of a more sophisticated understanding of community dynamics at the base of the food chains in lakes.

INTRODUCTION

Picoplanktonic cyanobacteria contribute between 5 and 80% of the total primary production in lakes, depending on the season, water chemistry and hydrography (Stockner et al., 2000). For such an important group of organisms our current understanding of their ecology, diversity and taxonomy in freshwaters is still limited. Studies of picocyanobacterial diversity and community structure in lakes are needed to improve our understanding of the dominant ecological processes at the base of the food web. There is convincing evidence for both marine and freshwater environments that individual water bodies can support a diverse community of picoplanktonic cyanobacteria (e.g. Stockner et al., 2000). Autotrophic picoplankton (APP), in marine environments, is dominated by a single clade of picocyanobacteria encompassing diverse strains assigned to the genera Synechococcus and Prochlorococcus (Sánchez-Baracaldo et al., 2005), whereas in freshwater environments the APP community is dominated by organisms generally described as Synechococcus-type (Weisse, 1993; Padisák et al., 2003).

Picocyanobacteria lack sufficient readily discernible morphological features to provide a robust taxonomic separation into distinct genera and species, and differences based on pigmentation (e.g. Postius & Ernst, 1999) can be misleading (Ernst et al., 2003). In such morphologically depauperate groups, knowledge of phylogenetic relationships is crucial to understanding lineage diversification and is therefore useful when studying community structures. Most phylogenetic studies of picocyanobacterial communities have focused on the relationships among marine picoplankton (i.e. Synechococcus and Prochlorococcus) using sequence data derived from the small subunit (ssu) rDNA,
the rDNA internal transcribed spacer (ITS-1), the phyco-
biliprotein encoding genes (cpc and cpe) and others
(Urbach et al., 1998; Rocap et al., 2002; Ting et al., 2002;
Fuller et al., 2003; Steglich et al., 2003). A smaller number of
studies have looked at the phylogenetic relationships of non-
marine picocyanobacteria and have been primarily based
on ssu rDNA, rDNA ITS-1 and cpcBA-IGS (inter-
genic spacer) sequences (Ernst et al., 2003; Crosbie et al.,
2003). Although ssu rDNA sequences seem to resolve
phylogenetic relationships within relatively closely related
strains, they do not provide well-supported resolution at
deep branches (Honda et al., 1999; Litvaitis, 2002).
However, because of the amount of ssu rDNA sequence
data already publicly available in GenBank, this locus
represents a useful resource for determining the phylo-
genetic and taxonomic relationships of sequences derived from
organisms that have not yet been isolated into culture. A
recent phylogenomic study, in which a backbone phylo-
genetic tree was calculated using 36 concatenated slowly
evolving genes from 14 cyanobacterial genomes and a
larger number of taxa with a more limited amount of data
(ssu rDNA, rpoC and morphology), resolved the deep
branches within the Cyanobacteria (Sánchez-Baracaldo et
al., 2005). This study suggested that the clade containing the
picocyanobacteria, i.e. the Synechococcus/
Prochlorococcus/Cyanobium (Syn/Pro) clade, branches deep in
the Cyanobacteria, just after Gloeobacter, the deepest branch.
Character evolution studies, using parsimony
reconstruction, also suggest that the early Earth cyanobac-
teria evolved in terrestrial and/or freshwater environments
(Sánchez-Baracaldo et al., 2005), as opposed to marine
environments as previously postulated (Honda et al., 1999;
Crosbie et al., 2003). It is likely that lineages within the
Syn/Pro clade exhibit some morphological and ecological
traits present in the earliest cyanobacterial lineages, such as
small cell diameter and free-living planktonic habit
(Sánchez-Baracaldo et al., 2005).

Long-term monitoring in the mesotrophic Lake Constance
indicates potential shifts in the composition of spring and
summer picocyanobacterial communities (Gaedke &
Weisse, 1998); however, the lack of distinctive morphol-
ogies for different taxa within the genus Synechococcus has
allowed only limited differentiation within such communi-
cies (Postius & Ernst, 1999). Phylogenetic studies have
shown that Synechococcus is polyphyletic, i.e. is not a
natural taxon (Honda et al., 1999; Wilmette & Herdman,
2001; Robertson et al., 2001), and the continuing use of
this generic designation in ecology can lead to confusion
and a lack of consistency between studies. The recent
application of molecular ecology techniques has revealed
substantial diversity in freshwater picocyanobacterial
communities (Postius & Ernst, 1999; Rocap et al., 2002;
Ernst et al., 2003; Becker et al., 2004). The use of Taq
nuclease assays has allowed sensitive detection of genotypes
in natural microbial communities, in which single
cocytotypes, based on ssu rDNA ITS-1 sequences, have been
monitored in both space and time in Lake Constance
time PCR has also allowed quantification of microcystin-
encoding genes and cell concentrations in temperate lakes
(Vaitomaa et al., 2003; Rinta-Kanto et al., 2005; Schober &
Kurmayer, 2006), detection of toxic Nodularia in the Baltic
Sea (Koskenniemi et al., 2007) and differences in
distribution of Prochlorococcus ecotypes in the Sargasso
Sea and Atlantic Ocean (Ahlgren et al., 2005; Johnson et
al., 2006). Clade-specific oligonucleotide probes have revealed
structuring of the Synechococcus community within a
stratified water column in the Red Sea (Fuller et al.,
2003). Furthermore, phylogenetically informed use of
molecular techniques, such as dot-blot hybridization and
fluorescence in situ hybridization, has revealed the distribution
of ecotypes and lineages of marine
Prochlorococcus and Synechococcus in the Mediterranean
Sea and Atlantic Ocean (Garczarek et al., 2007; Zwirglmaier
et al., 2007).

To better understand freshwater picocyanobacterial diversity,
we have used a phylogenetic approach combined with
quantitative Taq nuclease assays (TaqMan PCR) to assess
community structure in seven English post-glaciation and
man-made lakes. Phylogenetic analyses of ssu rDNA
sequences from the marine Synechococcus and
Prochlorococcus lineages, novel cultured freshwater isolates
and ssu rDNA clone libraries derived from environmental
DNA were used to identify well-supported clades of
freshwater picocyanobacteria. Clade-specific TaqMan
quantitative PCR assays were used to assess the abundance
of the members of four picocyanobacterial clades in
freshwater APP communities. Samples for these analyses
were collected from the Baltic Sea in the summer of 2003
and over an annual cycle from two lakes in the UK. We
show that individual water bodies can support a diverse
range of picocyanobacteria and that the community
structure is temporally and spatially variable. We also
show that the brackish Baltic Sea supports a picocyanobac-
terial community that has elements of both the
freshwater and marine floras and that it is patchy in
structure over relatively small spatial scales.

**METHODS**

**Cultures.** Cultures of Baltic Sea picocyanobacteria were established
from samples collected from the upper mixed layer of the water
column in June and July 1998 (Table 1). Samples were pre-filtered
through 20 μm mesh plankton netting, then cells were concentrated
onto 0.6 μm polycarbonate filters (Nucleopore) and either smeared
onto BG11 culture medium (Rippka et al., 1979) solidified with 0.5%
(w/v) agarose or suspended in 1 ml of either BG11 or filter-sterilized
seawater in 2 ml screw-cap tubes. Cells on solidified medium were
stored under black netting on the open bench until return to the UK.
Liquid cell suspensions were initially stored in tanks of running
surface seawater and subsequently 50 μl aliquots of 10^6, 10^2 and
10^0 dilutions in BG11 were spread onto BG11 agarose. In the
laboratory all plates were incubated at 15 °C under an incident
irradiance of 50 μmol m^{-2} s^{-1} with a 12:12 light/dark cycle.
Picocyanobacteria were streaked to single colonies to ensure clonality
and then maintained in liquid culture. Cultures of picocyanobacteria
Table 1. Picocyanobacteria cultures and clones and their sampling locations

Sampling locations: bs/BS, Baltic Sea; pp/PP, Priest Pot; ew/EW, Esthwaite Water; ap/AP, Abbot’s Pool; cwp/CWP, Cotswold Water Park lakes. The number of times each sequence was found is shown in parentheses: where there is no number the sequence was encountered only once.

<table>
<thead>
<tr>
<th>Sampling location</th>
<th>Sampling dates*</th>
<th>Environmental ssu rDNA sequences</th>
<th>ssu rDNA sequences from cultures§</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Short†</td>
<td>Long‡</td>
</tr>
<tr>
<td>Baltic Sea (55°14′–59°04′N, 13°35′–20°06′E)</td>
<td>24/06–02/07/1998</td>
<td>bsA1-5 (x 11), bsA3-2, bsA3-6, bsL4-3 (x 2)</td>
<td>B9801-type (x 1), B9802-type (x 4), B9803-type (x 1)</td>
</tr>
<tr>
<td>Baltic Sea (54°28′–59°29′N, 10°43′–15°48′E)</td>
<td>01–11/07/1998</td>
<td>bsA1-5 (x 5), bsA3-2 (x 3); bsL4-3, bsL28-2</td>
<td>9801-type (x 4), 9802-type (x 5), 9803-type (x 2)</td>
</tr>
<tr>
<td>Baltic Sea (58°10′–58°35′N, 18°34′–19°60′E)</td>
<td>23–25/06/1999</td>
<td>bsA1-3 (x 3), bsL4-3</td>
<td></td>
</tr>
<tr>
<td>Priest Pot (54°22′N 02′59′W)</td>
<td>18/06/2002</td>
<td>pp9-1 (x 4)</td>
<td>EW7-2</td>
</tr>
<tr>
<td>Esthwaite Water (54°21′N 02′58′W)</td>
<td>22/07/2003</td>
<td></td>
<td>EW2, EW8 and EW151</td>
</tr>
<tr>
<td>Abbots Pool (51°27′N 02′40′W)</td>
<td>22/07/2002</td>
<td></td>
<td>AP10-7; AP10-8</td>
</tr>
<tr>
<td>CWP 9 (51°39′N 01′53′W)</td>
<td>17/06/2002–18/07/2002</td>
<td>cwp9-1 (x 4); cwp9.2-2</td>
<td>CWP9.2-7; CWP9.2-33</td>
</tr>
<tr>
<td>CWP 31 (51°39′N 01′57′W)</td>
<td>16/04/2002</td>
<td>cwp31</td>
<td>CWP123.2-1; CWP123.2-2</td>
</tr>
<tr>
<td>CWP 123 (51°41′N 01′43′W)</td>
<td>12/08/2002</td>
<td></td>
<td>CWP124.1-1; CWP124.1-10</td>
</tr>
<tr>
<td>CWP 124 (51°41′N 01′43′W)</td>
<td>18/07/2002</td>
<td>cwp124.3</td>
<td></td>
</tr>
</tbody>
</table>

*Day/month/year.
†Short sequences were 362–424 bp in length and were derived from amplification products generated using primer pair 16S.19F and 16S.402R (Table 2).
‡Long sequences were 1362–1430 bp in length and were derived from amplification products generated using primer pairs Syn3 and Syn5 (Table 2) or Cyano5’ and Cyano3’ (Beard et al., 1999).
§Sequences from cultures were 1405–362 bp in length and were derived from amplification products generated using primer pairs Cyano5’ and Cyano3’ (Beard et al., 1999) or 16S.19F and 16S.402R or CYA395F and CYA781(a)/CYA781(b) (Nübel et al., 1997).
|| Cultures isolated in June and October 2000.

isolated from Esthwaite Water (English Lake District) were obtained from Dr Jackie Parry, University of Lancaster.

Samples of natural communities. Picocyanobacteria in pooled samples from the upper mixed layer of the Baltic Sea collected in June 1999 were pre-filtered through 20 μm plankton netting, and then concentrated on to 0.45 μm pore-size membrane filters (Gelman Supor 450) until the filters blocked; filters were stored in 1 ml DNA extraction buffer (10 mM Tris, 0.5 mM NaCl, 100 mM EDTA) at −20 °C. For all other environmental samples (for sampling locations see Tables 1 and 4) surface water was prefiltered sequentially through 30 μm mesh plankton netting and 5 μm pore-size cellulose nitrate membranes (Whatman) prior to collecting cells from a known volume of water (380 ml to 5 l) on 0.45 μm pore-size cellulose acetate membranes ( Pall Corporation). Cells immobilized on the membrane filters were rinsed in situ with 1 ml DNA extraction buffer and stored at −20 °C.

DNA extraction. DNA from picocyanobacterial cultures was extracted using lysozyme and proteinase K lysis followed by phenol/chloroform/isoamyl alcohol extraction as described by Beard et al. (1999). DNA from microbial communities stored frozen on filters in DNA extraction buffer was isolated by phenol/chloroform extraction (Kerkhof & Ward, 1993; Voytek & Ward, 1995). For quantitative PCR (see below) a 4 mm diameter punch was used to excise one or more 50.3 mm² sections from filters. Filter fragments were incubated with occasional shaking in 400 μl of 5 % (w/v) Chelex-100 (sodium form; 100–200 mesh; Bio-Rad) for 30 min at 100 °C, vortexed at high speed and the debris pelleted by centrifugation for 2 min (Becker et al., 2002). DNA-containing supernatants were stored frozen in aliquots at either −20 °C or −80 °C.

Characterization of ssu rDNA sequences. Near-complete ssu rDNA sequences (1430 bp) were amplified using the primer pair Cyan3’ and Cyan43’ (Beard et al., 1999). On the basis of initial results, specific freshwater picocyanobacterial primers, Syn5 and Syn3 (Table 2), were designed to reduce amplification of non-target sequences while still generating near full-length products (1362 bp). For amplification of shorter ssu rDNA fragments (362 bp), cyanobacterial specific primers were used, either 16S.19F and 16S.409R (this study; Table 2) or CYA359F and CYA781(a)/CYA781(b) (Nübel et al., 1997). Amplification reactions (50 μl) contained 10 mM Tris/HCl (pH 9.0), 1.5 mM MgCl₂, 50 mM KCl, 0.2 mM of each dNTP, 0.5 μM of each primer, 1 μl of 75 μg BSA ml⁻¹, 1 unit Supertaq DNA polymerase (HT Biotechnology) and 1 μl genomic DNA. Cycling conditions for long fragments were: initial denaturation at 94 °C for 3 min; 25 cycles of 94 °C for 45 s, 50 °C for 1 min, and 72 °C for 2 min; final extension at 72 °C for 10 min.
Table 2. Oligonucleotide primers and probes used in this study to amplify and quantify ssu rDNA targets.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence (5’–3’)</th>
<th>T_m</th>
<th>Target organisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syn3</td>
<td>TACGACCTCACCCGATGCTACGCTC</td>
<td>63</td>
<td>Cyanobacteria</td>
</tr>
<tr>
<td>Syn5</td>
<td>CAGATGAAACGGTGGGCGGCGTGC</td>
<td>63</td>
<td>Cyanobacteria</td>
</tr>
<tr>
<td>16S.19F</td>
<td>AAG CCT GAC GGA GCA ACG CC</td>
<td>60</td>
<td>Synechococcus</td>
</tr>
<tr>
<td>16S.409R</td>
<td>GGT ATC TAA TCC CTG TCG CTC C</td>
<td>60</td>
<td>Synechococcus</td>
</tr>
<tr>
<td>CLF</td>
<td>CAACCTCAGGCTGATGTCTCCAT</td>
<td>61</td>
<td>Cyanobium gracile cluster/clade I</td>
</tr>
<tr>
<td>CLR</td>
<td>GCAGGGCGCTGCA</td>
<td>59</td>
<td>Clade II</td>
</tr>
<tr>
<td>CII.R</td>
<td>CCCTTACACACTCTAGCCTGTA</td>
<td>59</td>
<td>Clade II</td>
</tr>
<tr>
<td>CII.R</td>
<td>CCCTTACACACTCTAGCCTGTA</td>
<td>59</td>
<td>Clade II</td>
</tr>
<tr>
<td>CIV.F</td>
<td>TCCGGAGGCGGTATTACA</td>
<td>60</td>
<td>Clade IV</td>
</tr>
<tr>
<td>CIV.R</td>
<td>CCCCCACACACTGAGTCTTAC</td>
<td>58</td>
<td>Clade IV</td>
</tr>
</tbody>
</table>

Cycling conditions for shorter fragments were: initial denaturation at 94 °C for 3 min; 30 cycles of 94 °C for 1 min, 60 °C for 1 min and 72 °C for 1.5 min; final extension at 72 °C for 10 min. PCR products were cloned in *Escherichia coli* TOP10 cells using a TA cloning kit (Invitrogen Life Technologies) following the manufacturer’s protocol. Plasmid DNA was purified using the QIAprep Miniprep (Qiagen). Cloned fragments were sequenced commercially on both strands using vector (M13F and R) and internal (16S.19F and 16S.409R) primers (Lark Technologies).

Phylogenetic analyses used both our novel sequences and sequences available from GenBank (http://www.ncbi.nlm.nih.gov/). Because of computational power impediments, maximum-parsimony and maximum-likelihood analyses mostly included only complete sequences of *ssu rDNA*; where included, short sequences were located in the most variable region of the gene. Taxa for inclusion in analyses were selected to provide a broad representation of phylogenetic diversity from the freshwater, brackish, marine and Antarctic lineages of the *Syn/Pro* clade (Sánchez-Baracaldo et al., 2005). To avoid redundancy, we excluded sequences that showed high pair-wise similarities within previously reported clades (Crosbie et al., 2003; Ernst et al., 2003; Fuller et al., 2003). Sequences were aligned manually according to their predicted secondary structure using Se-Al v2.0a11 (Rambaut, 2001). In addition, bootstrap analyses were performed to evaluate branching topologies. For the ML analyses, the estimation of parameters and choice of best-fit model of nucleotide substitution for the data were conducted using the computer program Modeltest 3.0 (Posada & Crandall, 1998), using hierarchical likelihood ratio tests. As a result, the GTR + G + I model of nucleotide substitution was used. These parameters were then used in optimal tree searches, which consisted of 100 random sequence additions and TBR (tree-bisection-reconnection) branch swapping. Clade support was estimated by 10,000 bootstrap replicates with 10-random-addition sequence replicates per bootstrap replicate. Unlike in previous analyses (Ferris & Palenik, 1998; Rocap et al., 2002; Crosbie et al., 2003) *Synechococcus elongatus* PCC 7942 and/or *Synechococcus* PCC 6301 were not used to root the trees because phylogenomic analyses have shown that *S. elongatus* is probably not the sister group of the *Syn/Pro* clade (Sánchez-Baracaldo et al., 2005). It has been shown that the ABRAXAS and ACE *Synechococcus* isolates from the Vestfold Hill Lakes, eastern Antarctica (Vincent et al., 2000), are sister to the marine and freshwater lineages within the *Syn/Pro* clade (Sánchez-Baracaldo et al., 2005); therefore, these strains were used when rooting.

Quantitative PCR. Four distinct freshwater picocyanobacterial clades were selected for quantitative Taq nuclelease assays (clades I–IV; Fig. 1). Primers (Table 2) were designed to amplify the most variable region of the *ssu rDNA* (between 721 and 811 bp from the 5′-end of the encoded *ssu rRNA* molecule) using Primer Express software version 2.0 (PE Applied Biosystems); for each primer, the number of mismatches with non-target sequences varied between 5 and 7 bp (Table 2). We used two MGB probes (Table 2), one, P100, for picocyanobacterial clades I, II and III and one, P104, specific for clade IV. Both probes were labelled with a 5′ FAM reporter and a non-fluorescent quencher at the 3′-end; the minor groove-binding activity of the quencher at the 3′-end increases the melting temperature (*T_m*), allowing the use of shorter probe sequences (Applied Biosystems). For initial specificity tests and for the construction of standard curves, known concentrations of cloned 390 bp *ssu rDNA* fragments of the following samples were used: cwp31 (clade I); AP10 8 (clade II); cwpv9 1 (clade III) and cwp124 3 (clade IV) (Fig. 1). Primer specificity for each clade was tested using SYBR Green to follow product accumulation in quantitative PCR: 10 µl reactions contained 1 µl of 2 × SYBR Green buffer (with ROX as a passive reference for normalization, Applied Biosystems), 0.3 µl of 10 µM primer (MWG-Biotech), 1 µl of 20 mM dNTP, 1.2 µl of 25 mM MgCl₂, 0.005 µl AmpliTaq Gold DNA polymerase, 2.15 µl water and 4 µl template DNA. For these reactions the template comprised either a single cloned *ssu rDNA*, or mixtures of cloned sequences where the target was present at a concentration of about 10⁶ molecules ml⁻¹ in the presence of each of the three non-target sequences either at the same concentration or at 10 × or 0.1 × relative concentrations. All real-time quantitative PCRs used a 96-well plate format and an ABI 7000 sequence detection system (PE Biosystems) as follows: 50 °C for 2 min, 95 °C for 10 min; 35 cycles of 95 °C for 15 s and 60 °C for 1 min. Quantitative TaqMan assays were performed in a final reaction volume of 17.5 µl. Reactions used to construct standard curves contained 1.57 µl of each primer (10 µM,
MWG-Biotech), 0.7 μl probe (5 μM, Applied Biosystems), 8.8 μl 2× TaqMan Universal PCR Master Mix (containing AmpliTag Gold DNA polymerase, AmpErase UNG, dNTPs with dUTP, passive reference 1, and optimized buffer components; AB Applied Biosystems), 2.86 μl water and 2 μl of the diluted standard template DNA (a five-step 10-fold dilution series providing approximately 10^7 to 10^3 molecules per reaction: actual concentrations differed for the individual clones). In negative controls additional water replaced the template DNA. The molar concentration of template used for the construction of standard curves was calculated from the DNA concentration, determined from the A_260, and the molar mass of the recombinant plasmid carrying the target sequence. For the analysis of natural microbial communities, amplification reactions contained primer, probe and buffer (as above) together with 2.8 μl environmental DNA, 2 μl of the most dilute target DNA standard and 0.06 μl water; the dilute DNA standard was added to each reaction to ensure that all measured template concentrations fell within the range of the standard curve. Cycling parameters were as described above. For all samples and standards, four replicate amplification reactions were used. C_T values for all standards and environmental samples were obtained using the ABI Prism 7000 SDS software, version 1.1 (Applied Biosystems) and the data were exported into Excel, where template concentrations in environmental samples were calculated using the standard curves and from the known proportion of the environmental DNA sample added to the amplification reaction.

**RESULTS AND DISCUSSION**

**Phylogenetic analyses**

Phylogenetic analyses of the sequences obtained from freshwater and Baltic Sea picocyanobacteria (Table 1; Fig. 1) suggest a greater diversity of picocyanobacteria in freshwater than marine environments, which is in keeping with the results of Crosbie et al. (2003). These results are consistent either with a long evolutionary history for freshwater picocyanobacteria or with more rapid speciation associated with the geographical barriers created by the fragmented nature of some freshwater systems, such as glacier-derived lakes. The relative homogeneity within the marine picocyanobacteria could be explained by the
connectivity between different ocean basins giving rise to
global populations, although some authors argue very
strongly that even planktonic marine micro-organisms
display distinct biogeographic differentiation (Medlin,
2007). Our results agree with other analyses in showing
marine Synechococcus as a sister group to Prochlorococcus,
which together form a monophyletic clade, sharing a
relatively recent common ancestor (Honda et al., 1999;
Rocap et al., 2002; Litvaitis, 2002). Both the marine
Synechococcus and Prochlorococcus are themselves natural
groups (Rocap et al., 2002).

Sequences derived from samples collected from freshwaters
fall into a number of both novel and previously described
clades (Fig. 1). We recovered similar clades to previous
studies of the non-marine picocyanobacteria (Ernst et al.,
2003; Crosbie et al., 2003). Cultured isolates from
Esthwaite Water and an environmental sequence from
Cotswold Water Park (CWP) Lake 31 are part of the
Cyanobium gracile cluster (clade I, Fig. 1). A novel cluster
composed of environmental sequences from Abbot’s Pool,
Esthwaite Water and CWP Lakes 9, 123 and 124 form a
well-supported group (clade II, Fig. 1) where BLAST searches
failed to identify any closely related sequences in GenBank.
Environmental sequences from Priest Pot and CWP Lake 9
form another well-supported clade (clade III, Fig. 1), again
with no similar sequences found in GenBank. Finally,
environmental sequences from CWP Lake 124 formed a
cluster with sequences from Mondsee (Austria) and Lake
Biwa (Japan) (clade IV, Fig. 1) that corresponds to group
H described by Crosbie et al. (2003). Of the novel groups
reported here, clades II and III, clade II not only seems to
be more diverse in clone libraries than clade III, it also
seems to be more widely distributed, being found in several
lakes. Organisms from within the C. gracile cluster seem to
be abundant and widespread. Our results show that most
previously described clusters also contain geographically
widespread isolates, and only a few are limited to a single
location, a finding similar to that reported by Crosbie et al.
(2003), who found, for example, cluster B, identified from
deep subalpine lakes, also encompassed isolates from
several Japanese lakes and Lough Neagh in Ireland.

The three cultured picocyanobacterial isolates from the
Baltic Sea do not form a monophyletic group, but cluster
with different clades (Fig. 1). B9801 falls within the C. gracile
cluster (Ernst et al., 2003; Crosbie et al., 2003) that includes a
previously described isolate from the Baltic Sea (BS20, Ernst
et al., 2003), B9802 clusters with freshwater sequences from
Mondsee or group I (Crosbie et al., 2003) that forms a sister
group to both marine and other freshwater lineages within
the Syn/Pro clade, and B9803 forms part of a cluster that
includes the halotolerant WH5701 assigned to Synechococcus
subcluster 5.2 (Fuller et al., 2003; Fig. 1). Sequences derived
from Baltic Sea environmental DNA samples were also
distributed among both marine and freshwater clusters.
Sequence bsL28-2 (Table 1, Fig. 1) falls within the marine
Synechococcus cluster, with most other brackish environ-
mental sequences grouping within what was previously
described as subalpine cluster I (Ernst et al., 2003) or cluster
B (Crosbie et al., 2003). Interestingly, most Baltic Sea-
derived sequences appear more closely related to freshwater
lineages, with only one of marine origin. The presence in the
Baltic Sea of picocyanobacteria related to freshwater and
marine forms is perhaps not surprising given its current
brackish state and its changing salinity over the last 12 000
years (Andrén et al., 2000; Bianchi et al., 2000).

Quantitative analysis of picocyanobacterial community structure

For community structure analysis using TaqMan assays we
selected four well-supported clades from our phylogenetic
analyses (Fig. 1). Oligonucleotide probes and primers were
designed to variable regions within the sequenced region of
the ssu rDNA for each target clade (Table 2). Such variable
regions were identified using an ssu rDNA sequence
alignment that included 58 taxa from the Syn/Pro clade
and representatives of the main cyanobacteria clades shown
by Sánchez-Baracaldo et al. (2005). Because there was only
a small variable region available for probe and primer
design, it was more efficient to use the shorter MGB probes
(Table 2). Probe P100 was designed to target ssu rDNAs of
clades I–III, with P104 specific for clade IV (Fig. 1);
differentiation between clades I–III was achieved through
the specificity of the amplification primers (Table 2).
Although there is only one mismatch between the two
probes, P100 cannot be used to identify clade IV targets
and probe P104 cannot be used to identify clades I–III.
Probe specificity was confirmed using cloned ssu rDNA
fragments for each target clade (data not shown).

Previous studies of picocyanobacterial community structure
have made use of multiple subsamples extracted from
whether a single subsample taken from a population
sample collected on a filter is representative of the whole,
we extracted DNA from three separate punch subsamples
from each of eight replicate filters and then determined the
concentration of clade IV template in each (Table 3). The
results demonstrated that C, values for individual ampli-
fication reactions can deviate from the mean for that
sample by a maximum of 1.07 cycles (±1.9 × under-
or overestimation of the initial template concentration),
but that overall the measured deviation in C, is likely to be
<0.31 (± 0.05) cycles (±1.2 × under- or overestimation
of the initial template concentration). From these measure-
ments we feel confident that single DNA extracts from
population samples are only likely to under- or over-
estimate the abundance of particular organisms by a
maximum of about twofold.

The ability of primer/probe combinations to amplify target
sequences in the presence of non-target DNA was initially
confirmed by amplification of cloned ssu rDNA sequences
from mixtures of equimolar, 10 × excess and 10 × less
concentrated non-target sequences. In all cases the quanti-
tative PCR-based estimates of template concentration were
unaffected by the presence of non-target sequences and recovered the expected value for the target sequence concentration. Furthermore, the \( C_t \) values obtained for the internal standards added to all environmental samples did not change even in the presence of up to a 70-fold excess of non-target template (data not shown).

**TaqMan assays to quantify community structures**

The concentration of the picocyanobacterial ssu rDNA genes in natural communities was estimated over an annual cycle for CWP Lakes 9 and 124 during 2003/4, and the Baltic Sea during July 2003. Community DNA extracted from water samples of known volume provided the template for clade-specific assays where the release of the fluorescent dye from the TaqMan probe was used to quantify the accumulation of amplification products. Comparisons of the rate of product accumulation, and hence initial template concentration within community DNA samples, were achieved by evaluating the number of amplification cycles \( (C_t) \) needed for the released fluorescence to reach a threshold value. The efficiency of amplification for the environmentally derived samples and the calibration standards did not differ significantly within a single PCR run, but there was significant variation between runs (data not shown). The inclusion of standards in all runs was essential for robust estimates of template concentrations in the samples derived from nature. Absolute quantification of initial template concentrations was achieved using calibration curves generated from the \( C_t \) values measured for cloned clade-specific templates of known ssu rDNA concentration. All environmental samples were spiked with the most dilute standard, equivalent to a template concentration of between 854 and 1651 ssu rDNA fragments per reaction, to ensure that amplification reactions gave \( C_t \) values falling within the range of the calibration curves. The quantification and detection limit of these assays was dependent on obtaining a \( C_t \) value significantly lower than that for the most dilute standard, which in practice was found to equate to a template concentration of about 100–200 molecules per reaction. From the measured concentration of ssu rDNA molecules in individual assays, the concentration in the original water samples was calculated from (a) the proportion of the total extracted DNA used in the assay (2.8 \( \mu l \) from 400 \( \mu l \); correction factor is \( \times 142.9 \)), (b) the area of the filter from which the DNA was extracted as a proportion of the total filter area on which the community sample was collected (50.3 mm\(^2\) from 1134.1 mm\(^2\); correction factor is \( \times 22.6 \)), and (c) the volume of water filtered. Using this method we were able to follow the development of populations of individual picocyanobacterial clades over an annual cycle (Fig. 2) and to quantify community structures at individual sampling locations in the Baltic Sea (Table 4).

The concentration of picocyanobacteria from all clades in lake surface waters varied over an annual cycle (Fig. 2). Organisms from clades II, III and IV could not be detected in samples collected over winter. For clade I in CWP Lake 9 there were three distinct periods of population development, one in the spring (April/May), one in the summer (July), and a much smaller population maximum in the autumn (September/October). Organisms from clades II, III and IV developed marked population maxima in the summer (July/August) but with some indication of a period of enhanced growth in the autumn (September/October). Clade I could be detected in the water of CWP Lake 124 throughout the year, never falling below a concentration of 3000 ssu rDNA molecules ml\(^{-1}\). The periods of rapid increase suggest minimum population doubling times of between 0.86 and 1.26 days for clade I, 0.91 days for clade II, 1.94 days for clade III and 4.92 days for clade IV; the minimum halving times for the declining populations were 0.86–1.03 days for

**Table 3. Quantitative PCR determination of the abundance of clade IV ssu rDNA from replicate samples of a single population**

<table>
<thead>
<tr>
<th>Replicate</th>
<th>Overall mean ( C_t )*</th>
<th>Max. individual deviation from the mean ( C_t ) (cycles)†</th>
<th>Mean deviation from the overall mean ( C_t ) (cycles)‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24.97</td>
<td>0.72</td>
<td>0.24 (±0.11)</td>
</tr>
<tr>
<td>2</td>
<td>24.58</td>
<td>0.78</td>
<td>0.50 (±0.16)</td>
</tr>
<tr>
<td>3</td>
<td>26.53</td>
<td>0.77</td>
<td>0.42 (±0.11)</td>
</tr>
<tr>
<td>4</td>
<td>25.56</td>
<td>0.23</td>
<td>0.12 (±0.04)</td>
</tr>
<tr>
<td>5</td>
<td>26.46</td>
<td>0.36</td>
<td>0.16 (±0.06)</td>
</tr>
<tr>
<td>6</td>
<td>25.74</td>
<td>1.07</td>
<td>0.68 (±0.15)</td>
</tr>
<tr>
<td>7</td>
<td>25.82</td>
<td>0.25</td>
<td>0.12 (±0.04)</td>
</tr>
<tr>
<td>8</td>
<td>24.39</td>
<td>0.35</td>
<td>0.21 (±0.06)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>0.31 (±0.05)</td>
</tr>
</tbody>
</table>

*Calculated from the 12 individual readings, i.e. four PCR replicates from each of three DNA extractions. †Maximum value by which an individual assay deviates from the overall mean. ‡Numbers in parentheses are 95 % confidence intervals.
clade I, 0.92 days for clade II, 1.18–1.41 days for clade III and 3.71 days for clade IV. Although the trends in population abundance observed here are likely to be robust, the exact numerical values need to be interpreted with caution. Previous studies have demonstrated losses of picocyanobacteria when passing natural samples through a range of pre-filters with different pore sizes (Becker et al., 2002), and thus it is possible that abundances shown (Fig. 2)

**Fig. 2.** Abundance of clades I–IV for 2003/4 obtained by absolute quantification using TaqMan assays based on ssu rDNA sequences.

<table>
<thead>
<tr>
<th>Station position</th>
<th>Date*</th>
<th>Clade I ssu mol ml⁻¹</th>
<th>Clade II ssu mol ml⁻¹</th>
<th>Clade III ssu mol ml⁻¹</th>
<th>Clade IV ssu mol ml⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>59° 35’N 23° 37’E</td>
<td>14/07/2003</td>
<td>4.0 \times 10⁶</td>
<td>3.7 \times 10⁵</td>
<td>5.2 \times 10⁴</td>
<td>4.9 \times 10⁴</td>
</tr>
<tr>
<td>59° 41’N 23° 31’E</td>
<td>15/07/2003</td>
<td>3.2 \times 10⁶</td>
<td>ND</td>
<td>ND</td>
<td>3.0 \times 10⁴</td>
</tr>
<tr>
<td>60° 04’N 26° 21’E</td>
<td>16/07/2003</td>
<td>9.7 \times 10⁵</td>
<td>2.7 \times 10⁵</td>
<td>ND</td>
<td>3.3 \times 10⁴</td>
</tr>
<tr>
<td>59° 51’N 24° 50’E</td>
<td>17/07/2003</td>
<td>8.4 \times 10⁵</td>
<td>3.5 \times 10⁵</td>
<td>ND</td>
<td>2.3 \times 10⁴</td>
</tr>
<tr>
<td>59° 35’N 23° 18’E</td>
<td>18/07/2003</td>
<td>3.7 \times 10⁵</td>
<td>ND</td>
<td>ND</td>
<td>1.3 \times 10⁴</td>
</tr>
</tbody>
</table>

ND, Below the limit of detection.

*Day/month/year.
might underestimate natural population sizes. Furthermore, concentrations of ssu rDNA are not equivalent to cell concentrations, because the ssu rDNA copy number per cell is known to be variable (Herdman et al., 1979; Labarre et al., 1989). No attempt was made during the current study to quantify picocyanobacterial abundance by direct cell counts or flow cytometry.

Picocyanobacteria from temperate lakes of all trophic types have seasonal cycles and a variety of successional patterns (Weisse, 1993). While some studies of temperate lakes have shown two peaks of picocyanobacteria abundance, one in spring and one in late summer (Weisse, 1993; Callieri & Stockner, 2000), other lakes lack a spring peak, with only a summer or autumn population maximum (Pick & Agbeti, 1991; Maeda et al., 1992; Hawley & Whitton, 1991). Our data suggest that CWP lakes 9 and 124 exhibit one peak, in summer, for all clades with the exception of clade I. It is difficult to interpret the significance of these observations as we know nothing about the characteristics of the organisms studied here, although differences in apparent doubling times suggest that they may encompass a range of physiologies. Remarkable physiological differences have been demonstrated between closely related Baltic Sea Synechococcus strains that only exhibit 1% sequence divergence in their rDNA ITS-1 (Stomp et al., 2004); therefore, the phylogenetic groupings based on the more highly conserved ssu rDNA sequences may be indicative of a great deal of physiological diversity. Moreover, our results show that there are temporal differences in population development among clades, with considerable differences in the maximum observed concentration of ssu rDNA molecules (Fig. 2, Table 4).

Interestingly, there were differences when comparing cloning and quantitative PCR experiments. PCR-based analyses over an annual cycle detected ssu rDNA molecules from clades that were not identified in clone libraries prepared from the same water body. This is perhaps not surprising and provides an illustration of the need to adopt a number of approaches when attempting to describe diversity in microbial communities. It is also worthy of note that some clades have no cultured representatives, which is almost certainly a function of both the highly variable abundance over an annual cycle and the incubation conditions used for culture establishment.

All freshwater clades studied here seem to form significant populations in the brackish waters of the Baltic Sea. Members of each clade were found at relatively high abundance during the summer of 2003, with ssu rDNA concentrations in the order of $10^5$–$10^6$ ml$^{-1}$ (Table 4). Three of the samples (collected on 14, 15 and 18 July 2003) were from locations <20 km apart, with the remaining sampling locations either about 80 or 170 km from the others. For clades I and IV the observed variation in abundance was probably not significant over these spatial scales, i.e. it is within the twofold error range that could be associated with using just a single DNA extract per sample (see above), but for clades II and III the variation in abundance is significant, with organisms belonging to these clades being undetectable in some samples, but reaching concentrations between 2.7 and $5.2 \times 10^3$ ml$^{-1}$ in others (Table 4). These results illustrate that even in open water environments the picocyanobacterial community structure is patchy over quite small spatial/temporal scales. This patchiness needs to be borne in mind when interpreting field data related to these fast-growing and rapidly disappearing organisms.

Conclusions

Freshwater picocyanobacteria communities encompass a large number of genetically and evolutionarily distinct lineages. The use of quantitative PCR allows the assessment of the contribution of individual clades (groups) to community structure and shows that their abundance varies significantly over relatively short spatial and temporal scales. The dynamic patterns of abundance and distribution shown here seem to support evidence from previous ecological studies. Molecular ecology studies based on novel clades, such as II and III, provide additional and valuable information on the diversity of picocyanobacteria in temperate lakes. Further studies should focus on factors driving changes in picocyanobacterial community structure and on determining both the physiological diversity and the contribution to ecosystem services associated with each of the different clades.

ACKNOWLEDGEMENTS

We would like to thank Annette Richer for assistance with culturing, Sven Becker for advice, Stephen Maberly and Mitzi Dent for sending us water samples from the English Lake District, as well as Andrew Impy and Tim Scoones for assisting with field-work, David Whitaker from Hanson Aggregates for allowing sampling lake CWP 124 in the Costwolds and Tim Colborn for expert assistance with figures. We would like to thank the Centro Internacional de Fı́sica (Bogota – Colombia) for institutional support to P.S.-B. This work was supported by a NERC grant from by the Marine and Freshwater Microbial Biodiversity thematic programme (NER/T/S/2000/01356).

REFERENCES


Edited by: D. J. Scanlan