

# Microbial community dynamics in a humic lake: differential persistence of common freshwater phylotypes

Ryan J. Newton,<sup>1</sup> Angela D. Kent,<sup>2,3†</sup> Eric W. Triplett<sup>4</sup> and Katherine D. McMahon<sup>1,2\*</sup>

<sup>1</sup>Microbiology Doctoral Training Program, University of Wisconsin-Madison, Madison, WI 53706, USA.

<sup>2</sup>Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, WI 53706, USA.

<sup>3</sup>Center for Limnology, University of Wisconsin-Madison, Madison, WI 53706, USA.

<sup>4</sup>Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611-0700, USA.

## Summary

In an effort to better understand the factors contributing to patterns in freshwater bacterioplankton community composition and diversity, we coupled automated ribosomal intergenic spacer analysis (ARISA) to analysis of 16S ribosomal RNA (rRNA) gene sequences to follow the persistence patterns of 46 individual phylotypes over 3 years in Crystal Bog Lake. Additionally, we sought to identify linkages between the observed phylotype variations and known chemical and biological drivers. Sequencing of 16S rRNA genes obtained from the water column indicated the presence of phylotypes associated with the *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, *Proteobacteria*, *TM7* and *Verrucomicrobia* phyla, as well as phylotypes with unknown affiliation. Employment of the 16S rRNA gene/ARISA method revealed that specific phylotypes varied independently of the entire bacterial community dynamics. *Actinobacteria*, which were present on greater than 95% of sampling dates, did not share the large temporal variability of the other identified phyla. Examination of phylotype relative abundance patterns (inferred using ARISA fragment relative fluorescence) revealed a strong correlation between the dominant phytoplankton succession and the relative abundance patterns of the

majority of individual phylotypes. Further analysis revealed covariation among unique phylotypes, which formed several distinct bacterial assemblages correlated with particular phytoplankton communities. These data indicate the existence of unique persistence patterns for different common freshwater phylotypes, which may be linked to the presence of dominant phytoplankton species.

## Introduction

Bacterioplankton communities are integrally involved in the biogeochemical processes underpinning freshwater ecosystems (Cotner and Biddanda, 2002). In humic lakes, which receive an exceptionally large input of allochthonous (terrestrially derived) organic matter, bacterioplankton play a critical role in determining the flux of nutrients between the terrestrial and aquatic environment. Humic lakes are further characterized by moderate phytoplankton productivity, but high bacterial metabolism (Wetzel, 2001). In addition, these lakes contain a simplified food web due to a lack of planktivorous fish. The combination of continuous nutrient input and a general lack of higher trophic levels result in a system dominated by microbial activity. Significant research efforts have focused on the contribution of bacterial communities as single entities to ecosystem functions in these lakes (Wetzel, 2001); however, the factors influencing bacterial community composition (BCC), and in particular the dynamics of individual community members, remain relatively unknown.

Recently, several studies have examined BCC in the epilimnion of freshwater lakes and reservoirs. As a result of these studies, a core group of bacterial phylotypes common to freshwater has emerged (Zwart *et al.*, 2002). Of the ubiquitous bacterial phylotypes, the *acl* clade of *Actinobacteria* and the *Beta* I and II clades of *Betaproteobacteria* are generally reported as numerically dominant (Hahn, 2003; Warnecke *et al.*, 2004; Simek *et al.*, 2005), with a single clade often comprising 30–50% of the total bacterial cells in the water column. In addition, evidence suggests BCC is significantly impacted by brief intense protistan grazing periods (Kent *et al.*, 2004; Pernthaler *et al.*, 2004; Simek *et al.*, 2005) and may be influenced by the composition and accessibility of

Received 29 July, 2005; accepted 4 November, 2005. \*For correspondence. E-mail [tmcmahon@engr.wisc.edu](mailto:tmcmahon@engr.wisc.edu); Tel. (+1) 608 263 3137; Fax (+1) 608 262 5199. †Present address: Department of Natural Resources and Environmental Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

autochthonous nutrients made available during phytoplankton blooms in both freshwater (Eiler and Bertilsson, 2004) and marine systems (Pinhassi *et al.*, 2004; Rooney-Varga *et al.*, 2005).

Although the existence of common epilimnetic freshwater bacterial phylotypes is apparent, we do not yet understand the extent of variation, and associated ecological drivers of change, in these dominant groups or of other freshwater bacterial community members over time scales of seasons or years. Yannarell and colleagues (2003) showed that lakes experience quite dramatic changes in BCC within- and between-years, regardless of lake trophic status. Similarly, several other researchers have shown that BCC varies greatly between lake types (Methe and Zehr, 1999; Lindstrom, 2000; Yannarell and Triplett, 2004). However, in none of these studies were temporally persistent and variable organisms identified. In an attempt to reconcile the notion that particularly common freshwater taxa exist with the observation that overall BCC is highly variable, we set out to identify and distinguish between populations that are characteristically variable and those that tend to recur over seasonal and annual time scales.

A focus on identifying dynamic and persistent bacterial populations should lead to a greater understanding of factors influencing bacterial community structure in lakes. In addition, an examination of factors that may be regulating not only the microbial community dynamics, but the dynamics of individual members of that community will lead to an increased understanding of potential freshwater microbial-mediated processes linked to specific organisms. Therefore, we designed a multiyear study of microbial populations in a single lake, with a sampling frequency designed to capture the pace of change in BCC, in an effort to provide insight into the ecology of both common and transient freshwater bacteria.

An initial study of Crystal Bog Lake revealed annual but dynamic BCC patterns and a significant correlation between bacterial phylotype richness and eukaryotic plankton succession (Kent *et al.*, 2004). However, the identity of individual phylotypes making up the bacterioplankton community and their population dynamics were not described. The aim of this study was to investigate patterns in humic freshwater BCC and diversity at a fine-scale taxonomic level over 3 years, to identify the persistent and dynamic bacterial community members, and to explore chemical and biological factors influencing individual community member dynamics. This was accomplished by coupling automated ribosomal intergenic spacer analysis (ARISA) fingerprinting to 16S ribosomal RNA (rRNA) gene clone library analysis (Brown *et al.*, 2005) to examine BCC at multiple levels of taxonomic resolution during a long-term and intense sampling effort.

## Results

### Bacterial community composition

Clone library analysis of Crystal Bog Lake identified representatives of six bacterial phyla (*Actinobacteria*, *Bacteroidetes*, *Firmicutes*, *Proteobacteria*, *TM7* and *Verrucomicrobia*), including members of the classes *Alpha*-, *Beta*-, *Delta*- and *Gammaproteobacteria*, and three clones with unclassified phylogenetic affiliation (Figs 1–4). The *Betaproteobacteria* class contained the most representatives at each assigned operational taxonomic unit (OTU) (Table 1), while only one representative at each OTU assignment was detected for the *Deltaproteobacteria* class and *Firmicutes* and *TM7* phyla (Table 1).

Thirteen freshwater-specific clades (Glockner *et al.*, 2000; Zwart *et al.*, 2002; Warnecke *et al.*, 2004) were identified in Crystal Bog Lake, which included members of the *Actinobacteria*, *Bacteroidetes*, *Alpha*- and *Betaproteobacteria*, and *Verrucomicrobia*. All phylotypes classified in the *Actinobacteria* phylum were members of the previously described freshwater-specific aCl-B or soil and freshwater-specific soil II–III clade (Fig. 1). Sequences affiliated with the soil II–III clade clustered with previously defined peat bog clones. Many of the Crystal Bog Lake *Betaproteobacteria* 16S rRNA gene sequences were affiliated with the freshwater clades Beta I, II, III and IV (Fig. 2). The remaining *Betaproteobacteria* clones were closely related to *Janthinobacterium*, *Ralstonia* or *Burkholderia*-type bacteria. The majority of *Bacteroidetes*-related clones were members of the previously defined freshwater clades CF I and CF III (Glockner *et al.*, 2000). All of the *Alphaproteobacteria* clones were identified as belonging to the freshwater clades Alpha I, II, III and IV and most of the *Verrucomicrobia* clones belonged to the FukuN18 freshwater clade (Figs 3 and 4). The remaining

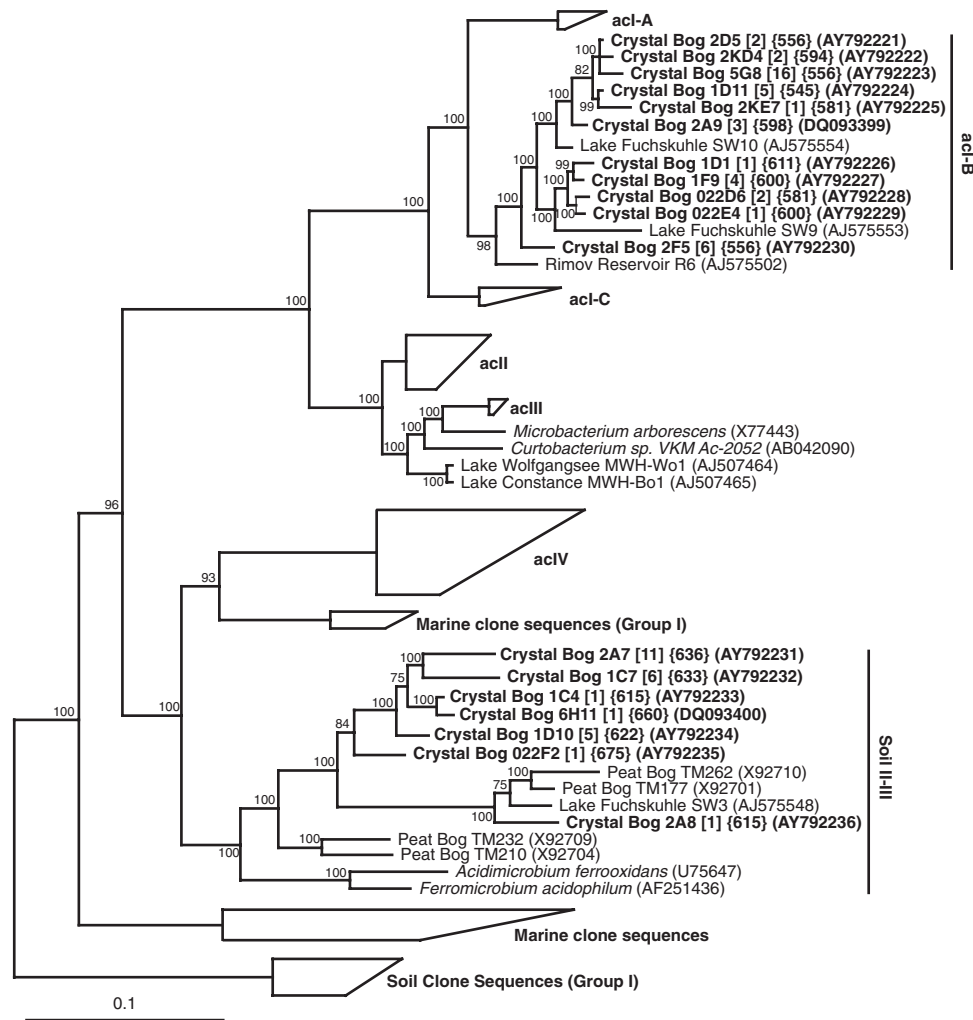
**Table 1.** Number of OTU assignments by phylum.

Phylum	Clade <sup>a</sup>	Species <sup>b</sup>	AFL <sup>c</sup>
<i>Actinobacteria</i>	2	7	13
<i>Bacteroidetes</i>	5	11	12
<i>Firmicutes</i>	1	1	1
<i>Proteobacteria</i>	18	34	52
<i>Alpha</i>	4	9	8
<i>Beta</i>	7	14	25
<i>Delta</i>	1	1	1
<i>Gamma</i>	6	10	18
<i>TM7</i>	1	1	1
<i>Verrucomicrobia</i>	2	2	6
Unknowns	3	3	3
Total	32	59	88

**a.** Clades were determined by the branching patterns obtained following phylogenetic tree construction and have sequence identity = 90%.

**b.** Species are defined as 16S rRNA sequence groups sharing = 97% gene identity.

**c.** AFL = ARISA fragment length.



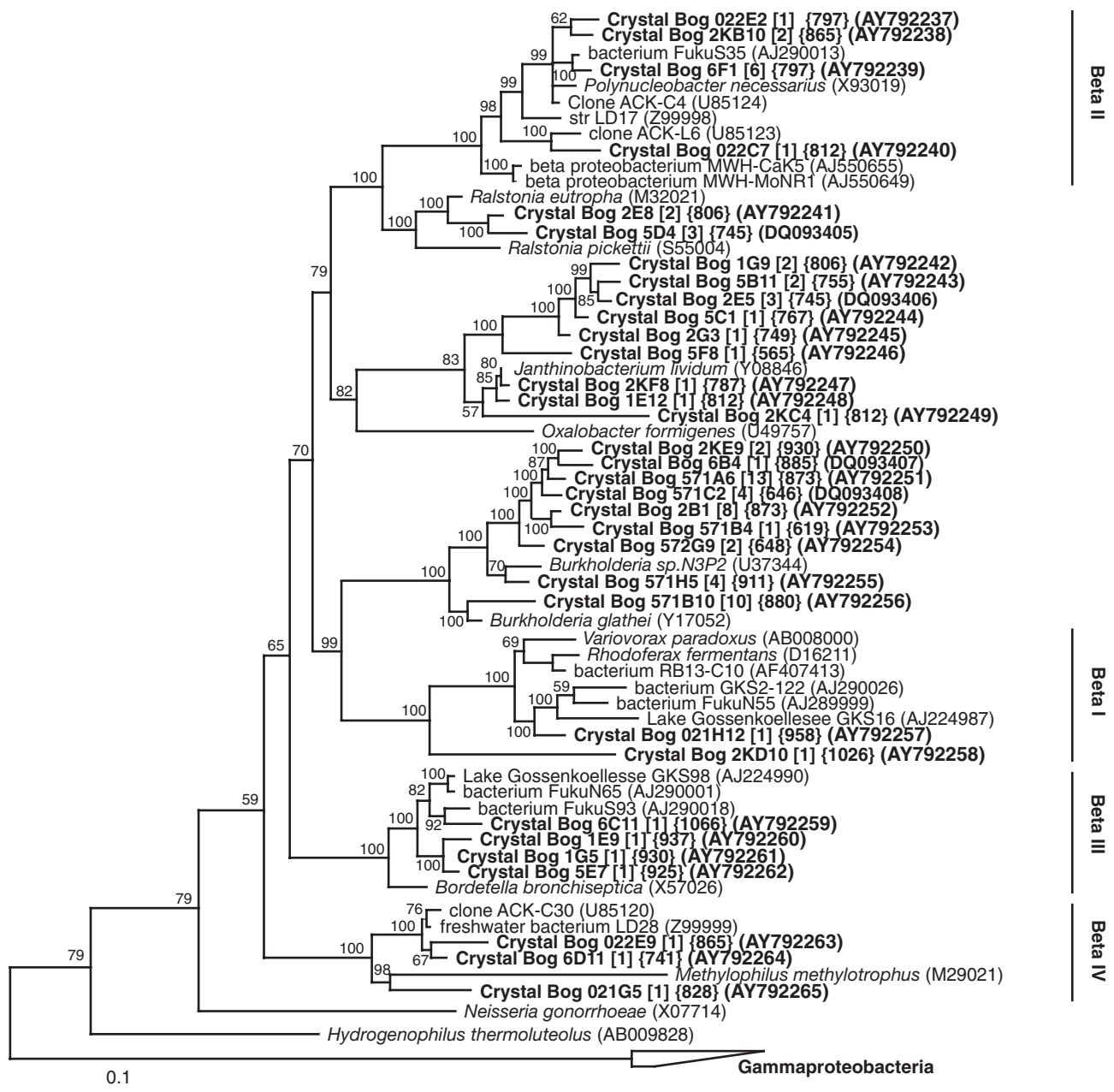
**Fig. 1.** Unrooted consensus phylogram depicting a subset of common *Actinobacteria* freshwater clades based on nearly full-length (> 1300 bp) 16S rRNA gene sequences. Relationships were determined by Bayesian analysis (software, MrBayes 3.0) using a 50% base frequency mask with 100 000 generations, yielding 8000 trees following 20 000 generations of burnin. Nodes with posterior probability values of >50% are indicated. The unique (< 97% 16S rRNA gene sequence identity or unshared AFL) sequences identified from Crystal Bog Lake are shown in bold with their corresponding AFL in curly brackets. The number of clones represented by each depicted sequence is shown in square brackets and the GenBank accession number of all sequences is indicated in parentheses. The scale bar indicates 0.1 changes per site.

identified clones were not affiliated with any previously described freshwater clades. Following alignment and tree inference, three singleton clones Crystal Bog 2E1, Crystal Bog 2KA12, and Crystal Bog 021B9 did not cluster with any previously defined phyla (Fig. 4). Clone Crystal Bog 2E1 was most closely affiliated with 16S rRNA gene sequences from the *TM6* and *TM7* phyla (Fig. 4). Clones Crystal Bog 2KA12 and Crystal Bog 021B9 were most closely affiliated with each other and are loosely affiliated with members of the *Verrucomicrobia* phylum (Fig. 4).

#### Clone library analysis

Four Crystal Bog clone libraries produced 289 16S rRNA gene sequences and their corresponding ARISA fragment

lengths (AFLs, measured as the number of nucleotides amplified with primers 1406F and 23SR). The coverage of the largest clone library (170 sequences from 3 year pooled DNA) as calculated based on the species OTU (97% 16S rRNA gene sequence identity) by Good's Clone Coverage (Good, 1953) was 89% and as estimated by the  $C_{ace}$  statistic was 88% (Kemp and Aller, 2004). The probability of drawing a new sequence from this library at the species level on the next draw was 1.7% (Clayton and Frees, 1987). The  $S_{chao1}$  diversity estimate (Chao, 1987) of this library predicted 185 unique species sequences. Because the remaining three libraries were constructed with a subset of the total dates and one library was screened by AFL prior to sequencing, they were not included in the clone library coverage analyses. The final



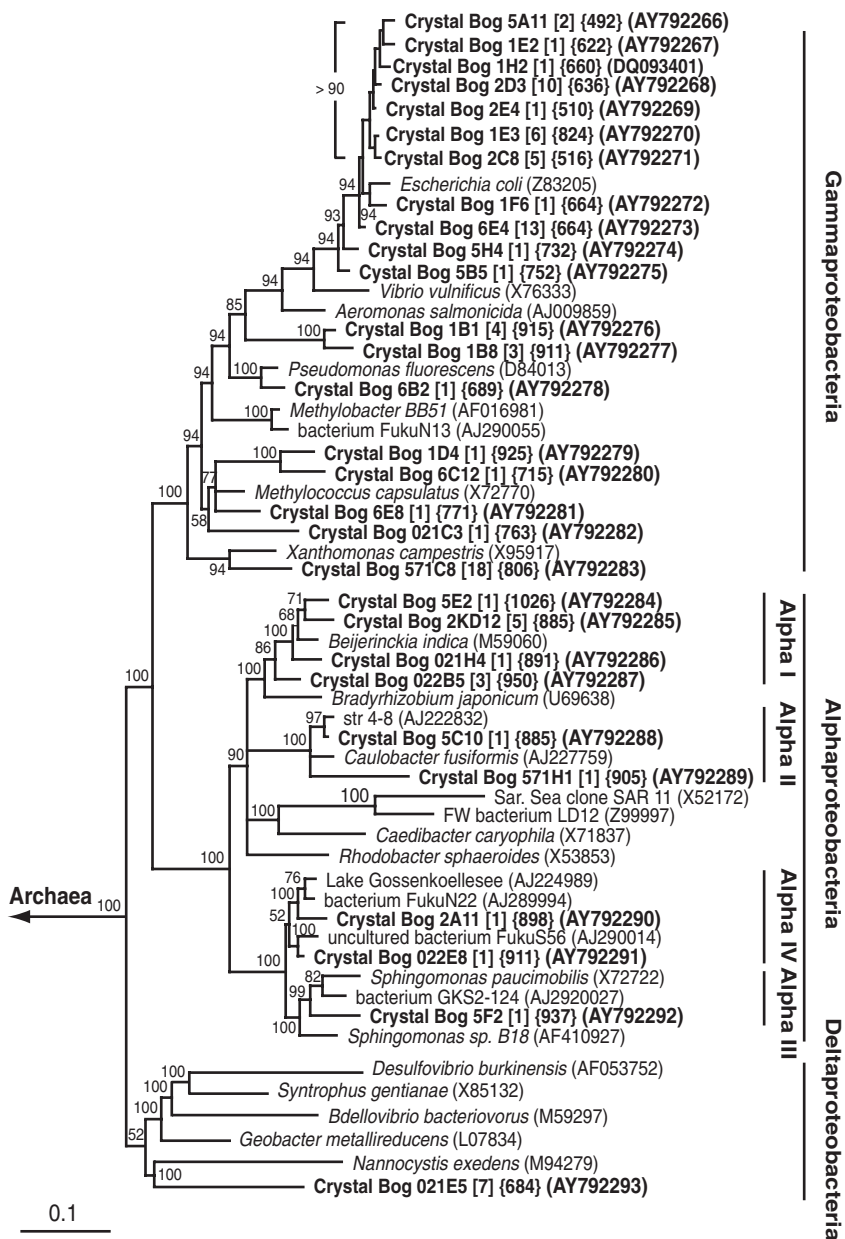
**Fig. 2.** Unrooted consensus phylogram depicting a subset of common *Betaproteobacteria* freshwater clades based on nearly full-length (> 1300 bp) 16S rRNA gene sequences. Relationships were determined by Bayesian analysis (software, MrBayes 3.0) using a 50% base frequency mask with 100 000 generations, yielding 7000 trees following 30 000 generations of burnin. Nodes with posterior probability values of >50% are indicated. The unique (< 97% 16S rRNA gene sequence identity or unshared AFL) sequences identified from Crystal Bog Lake are shown in bold with their corresponding AFL in curly brackets. The number of clones represented by each depicted sequence is shown in square brackets and the GenBank accession number of all sequences is indicated in parentheses. The scale bar indicates 0.1 changes per site.

library of 96 clones, which was screened for unidentified AFLs prior to sequencing, returned six clones containing an AFL we had not yet identified.

#### Community composition dynamics

ARISA fingerprints were obtained from samples collected

during the ice-off season for 3 years in Crystal Bog Lake. These fingerprints contained 126 different ARISA fragments (based on fragment length) and a total of 3041 ARISA fragments summed across all 68-sample dates during the 3 year sampling period. Sixty-five (52%) of the unique ARISA fragments and 2341 (77%) of the total ARISA fragments were assigned a taxonomic identity

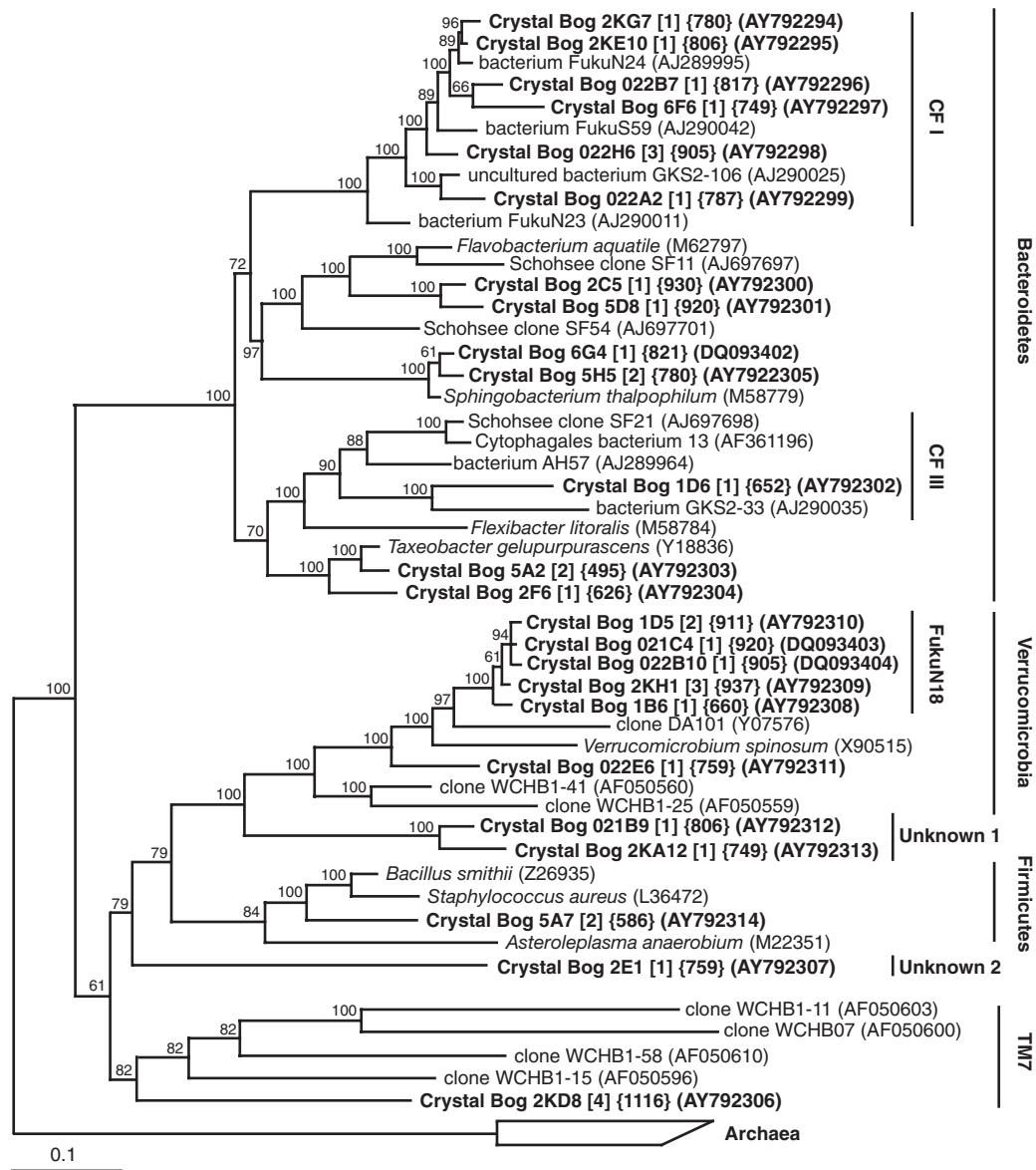


**Fig. 3.** Unrooted consensus phylogram depicting a subset of *Alpha*-, *Delta*- and *Gammaproteobacteria* based on nearly full-length (> 1300 bp) 16S rRNA gene sequences. Four common freshwater *Alphaproteobacteria* clades are illustrated. Relationships were determined by Bayesian analysis (software, MrBayes 3.0) using a 50% base frequency mask with 100 000 generations, yielding 8000 trees following 20 000 generations of burnin. Nodes with posterior probability values of >50% are indicated. The unique (< 97% 16S rRNA gene sequence identity or unshared AFL) sequences identified from Crystal Bog Lake are shown in bold with their corresponding AFL in curly brackets. The number of clones represented by each depicted sequence is shown in square brackets and the GenBank accession number of all sequences is indicated in parentheses. The scale bar indicates 0.1 changes per site.

based on matching to 16S rRNA gene sequences, representing 83% of the total fluorescence from all ARISA profiles.

The BCC of Crystal Bog Lake was quite dynamic. Less than 20% of the AFLs were present on more than 50% of the sampling dates, but nearly 50% of the AFLs were present in at least one sample in each of the 3 years (Fig. 5). Although the overall BCC changes quite rapidly during a year, some members of the bacterial community did not share the dynamic behaviour of the majority. The *Actinobacteria*, especially members of the *aci*-B clade, were more consistently present than any

other clade. Four AFLs associated with the *aci*-B clade were present on more than 90% of the 68 sampling dates (Fig. 5). As a clade, *aci*-B contributed more than 25% of the total fluorescence units on more than 60% of the sampling dates while no other clade contributed greater than 25% on more than two sampling dates (data not shown). In contrast, the majority of phylotypes associated with the *Beta*- or *Gammaproteobacteria* showed substantial presence/absence variability. Generally the AFLs from these two classes of *Proteobacteria* were present on no more than 50% of the sampling dates (Fig. 5).



**Fig. 4.** Unrooted consensus phylogram depicting a subset of *Bacteroidetes*, *TM7*, *Verrucomicrobia* and *Firmicutes* phyla based on nearly full-length (> 1300 bp) 16S rRNA gene sequences. Two freshwater *Bacteroidetes* and one freshwater *Verrucomicrobia* clade is depicted. Sequences not belonging to a known phylum are labelled as Unknown. Relationships were determined by Bayesian analysis (software, MrBayes 3.0) using a 50% base frequency mask with 100 000 generations, yielding 7500 trees following 25 000 generations of burnin. Nodes with posterior probability values of >50% are indicated. The unique (< 97% 16S rRNA gene sequence identity or unshared AFL) sequences identified from Crystal Bog Lake are shown in bold with their corresponding AFL in curly brackets. The number of clones represented by each depicted sequence is shown in square brackets and the GenBank accession number of all sequences is indicated in parentheses. The scale bar indicates 0.1 changes per site.

#### Biological, chemical, and physical drivers of bacterial community composition

The physical/chemical parameters measured in this study by themselves or in combination did not significantly explain the BCC change observed in 2002 (Kent *et al.*, 2004). Likewise, the temporal dynamics of individual bacterioplankton community members (assessed by

AFL relative fluorescence) were not significantly correlated to the changes of any single measured chemical and physical factor, or any combination thereof (data not shown).

Phytoplankton community succession and heterotrophic nanoflagellate (HNF) abundance were closely monitored during 2002 (Kent *et al.*, 2004). The dynamics of dominant assemblages (regimes) are described here

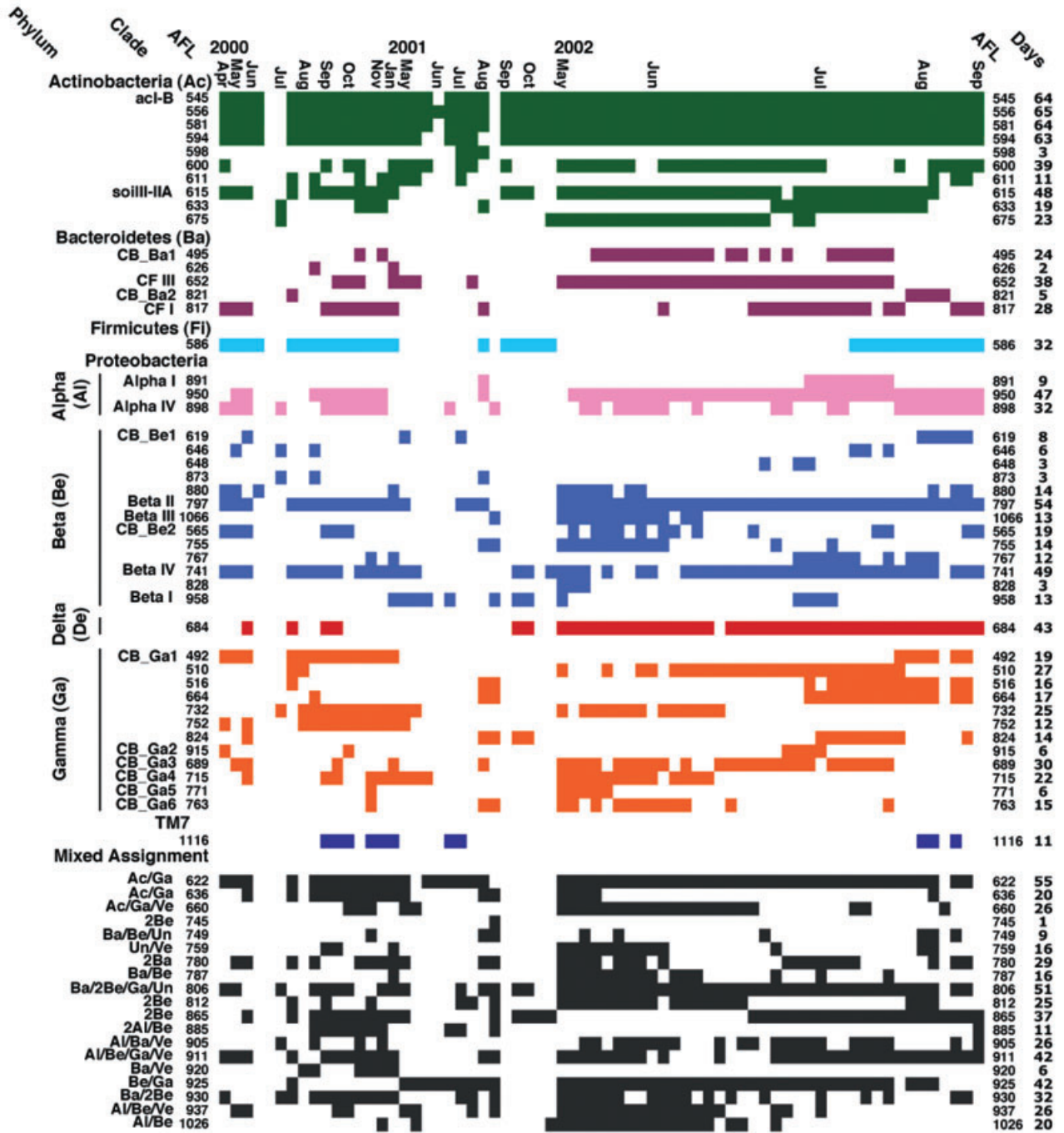


Fig. 5. Three year plot by sample date showing the presence/absence of all AFLs associated with clades identified in Crystal Bog Lake. The presence of a coloured box indicates that the AFL was present on that sample date. The months and years listed across the top row correspond to the first sampled date within that month/year. The phyla, clades and AFLs are listed to the left of the respective plot row. All AFLs listed below each clade designation belong to that clade. The AFL and the total number of sample dates on which the AFL was present are listed to the right of the corresponding plot row. AFLs assigned to more than one clade are listed separately at the bottom as mixed assignments. Phylogenetic affiliation not listed: *Verrucomicrobia* (Ve) and unknown (Un).

as follows: the Chrysophyte *Dinobryon* and dinoflagellate *Peridiniopsis* co-dominated (in terms of biovolume) the phytoplankton population during a significant increase in abundance of HNFs during the mid-spring season; *Cryp-*

*tomonas*, a motile unicellular photosynthetic alga, dominated during late spring; the dinoflagellates *Gymnodinium fuscum*, *Peridinium limbatum* and *Peridinium cinctum* co-dominated during early summer; and the two *Peridinium*

species alone dominated during mid-summer. In addition, the unicellular bristled Chrysophyte, *Mallomonas*, showed a significant increase in total biovolume during the end of early summer and beginning of mid-summer, although it was not the dominant phytoplankton community member at any time during that period.

Calculation of the Pearson product–moment correlation coefficient revealed significant correlations between individual bacterial phylotypes (assessed by AFL relative fluorescence) and individual dominant phytoplankton regimes over the course of 2002 (Table 2). The majority of identified AFLs (41 out of 65) exhibited strong correlations ( $P \leq 0.001$ ) to the dynamics of at least one phytoplankton/HNF regime in 2002. Canonical correspondence analysis (CCA) using individual phytoplankton species biovolume as explanatory variables illustrates the relationships between individual AFLs and particular phytoplankton regimes (Fig. 6). Notably, AFLs assigned to the *Beta* and *Gammaproteobacteria* are associated with the intense bacterivory period that included blooms of *Peridiniopsis* and *Dinobryon*, while the majority

of *Actinobacteria* AFLs do not appear to be influenced by any of the measured phytoplankton taxa (Fig. 6). Altogether, 69% of the AFL–phytoplankton relationship is explained by the first two CCA axes, and the relationship is significant ( $P = 0.01$ ).

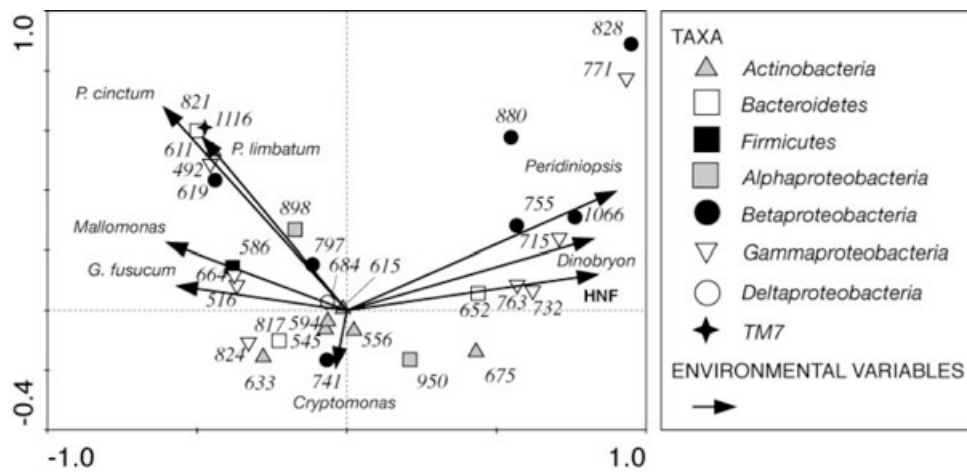
Several groups of covarying phylotypes related to individual phytoplankton/HNF regimes became apparent from these analyses (Table 2 and Fig. 6). An analysis of similarity (ANOSIM) with groups defined by the strongest correlation to a phytoplankton regime (listed in bold, Table 2) confirmed the significance of these covarying assemblages ( $R$ -value = 0.8,  $P$ -value < 0.001). Although the taxonomic composition of the bacterial community comprising the assemblages varied greatly, a few trends emerged. The *acl*-B clade of *Actinobacteria* relative abundance was negatively correlated to the presence of flagellate grazers, which indicates the *acl*-B clade was a less significant part of the community during this intense bacterivory period. On the other hand, a large number of phylotypes from clades in the *Betaproteobacteria*, *Bacteroidetes* and *Gammaproteobacteria*

**Table 2.** Pearson product–moment correlation values between bacterial phylotype relative abundance and algal phylotype biovolume<sup>a</sup> or HNF abundance.

Clade <sup>b</sup>	AFL	HNF	Per.	Din.	Cryp.	Gym.	Mal.	P. cin.	P. lim.
Beta IV	741	<b>-0.51</b>	–	–	–	–	–	–	–
CF I	817	<b>-0.60</b>	–	–	–	–	–	–	–
Soil II–III	675	<b>0.51</b>	–	–	–	–	–	–	–
CB_Ga1	732	<b>0.63</b>	0.53	–	–	–	–	–	–
CB_Ga6	763	<b>0.59</b>	0.52	–	–	–	–	–	–
<i>acl</i> -B	545	–	<b>-0.64</b>	-0.56	–	–	–	–	–
<i>acl</i> -B	556	–	<b>-0.53</b>	–	–	–	–	–	–
<i>acl</i> -B	594	-0.60	<b>-0.68</b>	–	–	–	–	0.54	–
CF III	652	0.77	<b>0.83</b>	0.79	–	-0.54	–	-0.55	–
CB_Ga4	715	0.81	<b>0.86</b>	0.83	–	–	–	–	–
CB_Be2	755	–	<b>0.64</b>	–	–	–	–	–	–
CB_Ga5	771	0.51	<b>0.86</b>	0.74	–	–	–	–	–
Beta IV	828	–	<b>0.63</b>	–	–	–	–	–	–
CB_Be1	880	–	<b>0.74</b>	0.58	–	–	–	–	–
Beta III	1066	0.70	<b>0.78</b>	0.64	–	–	–	–	–
Alpha I	950	–	–	–	<b>0.61</b>	–	-0.53	–	–
Delta	684	–	–	–	–	<b>0.81</b>	0.70	–	–
CB_Ga1	516	–	–	–	–	0.58	<b>0.84</b>	–	–
Firm.	586	–	–	–	–	0.56	<b>0.60</b>	0.55	–
<i>acl</i> _B	611	–	–	–	–	–	<b>0.68</b>	0.58	0.54
Soil II–III	615	–	–	–	–	–	<b>0.68</b>	–	–
Soil II–III	633	–	–	–	–	–	<b>0.55</b>	–	–
CB_Ga1	664	–	–	–	–	0.63	<b>0.84</b>	0.58	–
CB_Ga1	824	–	–	–	–	0.68	<b>0.71</b>	–	–
CB_Ga1	492	–	–	–	–	–	–	<b>0.74</b>	0.60
CB_Be1	619	–	–	–	–	–	–	<b>0.68</b>	0.63
Beta II	797	–	–	–	–	–	–	<b>0.65</b>	–
CB_Ba2	821	–	–	–	–	–	–	<b>0.82</b>	0.67
Alpha IV	898	–	–	–	–	–	–	<b>0.74</b>	–
TM7	1116	–	–	–	–	–	–	<b>0.65</b>	–

**a.** In the interest of clarity, correlation coefficients are presented only for correlations that were significant at a level of  $P < 0.001$ ,  $N = 38$ . The strongest correlations for each clade are in bold text. Per., *Peridiniopsis*; Din., *Dinobryon*; Cryp., *Cryptomonas*; Gym., *Gymnodinium*; Mal., *Mallomonas*; P. cin., *Peridinium cinctum*; P. lim., *Peridinium limbatum*.

**b.** Clades were determined by the branching patterns obtained following phylogenetic tree construction and have sequence identity  $\geq 90\%$ . Clade groupings are listed in Fig. 5. See trees (Figs 1–4) for freshwater clade identification.



**Fig. 6.** CCA biplot showing individual AFLs and their relationship to food web variables during 2002. Phytoplankton (biovolume) and nanoflagellate (abundance) explanatory variables are represented by black arrows (eigenvectors) that indicate the direction of increase for each variable. The length of each arrow indicates the degree of correlation with the ordination axes. Note that the *Cryptomonas* eigenvector is not strongly correlated with the first two ordination axes.

exhibited significant positive correlations during this same period. Over the course of the summer phytoplankton succession, nearly all of the examined AFLs with a significant correlation to a single phytoplankton regime showed a significant positive relationship, which indicates that certain bacterioplankton favour specific phytoplankton regimes. These assemblages exhibited a wide range of taxonomically diverse bacteria, but contained a large number of *Gammaproteobacteria* phylotypes (Table 2).

## Discussion

The advent of microbial fingerprinting techniques has allowed microbial ecologists to carry out more efficient analysis of microbial community composition and dynamics during intense and/or long-term sampling efforts. However, the basis of these techniques makes it difficult to obtain truly quantitative data from multiple phylotypes without prohibitive effort. Yet, sampling at appropriate temporal scales may be critical for the identification of ecological patterns related to BCC (e.g. Yannarell *et al.*, 2003). It is also recognized that examining solely presence-absence data may significantly hinder the identification of ecologically relevant trends in community analysis (Yannarell and Triplett, 2005). Therefore, others have used the relative abundances of individual AFLs to examine the change of that AFL across multiple samples (Hewson and Fuhrman, 2004; Brown *et al.*, 2005; Yannarell and Triplett, 2005). We also considered these relative abundance data when examining links between changes in the bacterial community and other chemical or biological parameter dynamics.

## Community composition

All phyla identified in Crystal Bog Lake, except for the TM7 phylum, were previously identified in other freshwater studies (Zwart *et al.*, 2002). Furthermore, 13 of the identified clades were formerly recognized as containing mostly freshwater members. Included in this freshwater group were the *acl*-B clade of *Actinobacteria*, the CF I and CF III clades of *Bacteroidetes*, several clades of the *Beta*- and *Alphaproteobacteria*, and the FukuN18 clade of *Verucomicrobia* (Figs 1–4), the majority of which were seen in all three sampling years (see Fig. 5), providing additional evidence to the hypothesis that members of these clades represent a substantial cosmopolitan component of lake bacterial communities (Zwart *et al.*, 2002). Although the majority of 16S rRNA gene sequences retrieved from Crystal Bog Lake were affiliated with freshwater-specific clades, many were from bacteria not belonging to recognized freshwater clades. As extensive phylogenetic surveys of bacteria in freshwater are relatively limited, some of these sequences may represent as yet unidentified freshwater-specific bacteria. For example, the soil II–III clade of *Actinobacteria* contains sequences obtained from bog lakes and numerous soil environments (Warnecke *et al.*, 2004). Our clone libraries contained 16S rRNA gene sequences from the soil II–III clade of *Actinobacteria* that formed a distinct monophyletic cluster with sequences from other humic lakes and bogs, suggesting the existence of humic lake-specific populations (Fig. 1). Burkert and coworkers hypothesized that the prevalence of *Actinobacteria* in humic lakes may be due to an ability of these organisms to break down humic acid containing compounds, a trait seen in many terrestrial *Actinobacte*

*ria*. If the freshwater subset of the soil II–III clade is truly specific to humic lakes as the sequence record suggests, then it is quite possible that these organisms are specialized to use the allochthonous humic compounds found in bog lakes. Alternatively, sequences that fell outside of known freshwater clades, such as the *Burkholderia*-related and enteric-related sequences may represent transient lake community members more typically associated with environments outside of the lake water column (e.g. the surrounding sphagnum mat or sediments) that are periodically transported into the lake by rainfall events or wildlife activity. The fairly large number of these sequences obtained in the clone libraries indicates Crystal Bog Lake may receive a large and continual flux of allochthonous bacteria (see below).

Several 16S rRNA gene sequences retrieved from Crystal Bog Lake, including all those from *TM7* and *Deltaproteobacteria*, were not closely (< 92% identity) related to any other sequences in the NCBI GenBank databases (19, May 2005). Additionally, three 16S rRNA gene sequences did not affiliate with any known phyla, indicating they may represent unrecognized bacterial phyla (GenBank Accession numbers, AY792312–AY792314). However, all three of these unique 16S rRNA gene sequences were obtained only once in the clone libraries. Although extensive chimera detection methods were employed, the possibility that these sequences are artifacts of PCR reactions cannot be dismissed. Additional sequence collection and phylogenetic analyses are required to determine if these clones are derived from novel phyla.

The *Betaproteobacteria* exhibited the greatest richness at all OTU definitions (Table 1). The large number of unique *Betaproteobacteria* taxa inhabiting freshwater may be indicative of the diverse metabolic composition of this group (Madigan *et al.*, 2002; Burkert *et al.*, 2003), which might allow phylogenetically similar taxa to occupy separate niches within the same physical space. Alternatively, this high level of observed richness may be due to populations containing multiple compositionally diverse *rrn* operons (Klappenbach *et al.*, 2000; Acinas *et al.*, 2004). In freshwater mesocosm and isolation studies, members of the *Betaproteobacteria* were observed to respond quickly to nutrient additions (Burkert *et al.*, 2003; Hahn, 2003; Simek *et al.*, 2005) and are thought to be high nucleic acid containing bacteria (Simek *et al.*, 2005). Bacteria capable of responding quickly to nutrient concentration fluctuations would be predicted to contain a larger number of *rrn* operons and have higher nucleic acid concentration than slow growers or those that respond less quickly to nutrient additions (Klappenbach *et al.*, 2000). As the internally transcribed spacer (ITS) length is not necessarily conserved among multiple *rrn* operons within a single 16S rRNA phylotype, the large diversity of *Betap-*

*roteobacteria* phylotypes might actually represent a smaller number of organisms with multiple divergent *rrn* operons. Further investigations are required to adequately test these hypotheses and lie outside the scope of this project.

#### *BCC variation over time*

Previous work illustrated the highly variable nature of lake BCC within- and between-years (Yannarell *et al.*, 2003). The majority of Crystal Bog Lake community members (assessed by AFL) are quite dynamic, yet >70% are present at some time during all 3 years (Fig. 5). As described above, community members represented during short continuous intervals, illustrated by the CB\_Ba2 and TM7 phylotypes and numerous AFLs within other clades (Fig. 5) may indicate the presence of transient unsustainable populations that are occasionally transported into the lake (Warnecke *et al.*, 2004). Another plausible explanation is that these populations are not numerous and therefore regularly fall below detection levels (Yannarell and Triplett, 2004). In either case, the observed temporal variation in BCC suggests that many community members are significantly influenced by the multiple ecological drivers known to affect these freshwater communities (Nold and Zwart, 1998; Crump *et al.*, 2003; Kent *et al.*, 2004; Yannarell and Triplett, 2005), of which food web dynamics are proposed to be the dominant factor in this lake (Kent *et al.*, 2004).

The *acl-B* clade of *Actinobacteria* is a clear exception to the otherwise continual variation in BCC over time (Fig. 5). Warnecke and colleagues (2004) suggest that the *acl Actinobacteria* clade does not constitute a transient component of lake communities originating from soil in the catchment, but represents a unique pelagic freshwater lineage capable of sustaining growth in the lake. The data presented here support their argument, because it is unlikely that bacteria washing in from the surrounding catchment would be found on nearly all 68 sampling dates across 3 years. In addition, *acl Actinobacteria* sequences have been obtained in the majority of clone libraries constructed from freshwater and are often the numerically dominant member of freshwater bacterial communities (Glockner *et al.*, 2000; Warnecke *et al.*, 2005). Taking into account the prevalence of this group noted in numerous studies, and the persistence seen in this study, it appears the *acl* clade of *Actinobacteria* possesses a significant and sustainable competitive advantage over most bacteria in the freshwater system. Pernthaler and colleagues (2001) demonstrated that the small size of freshwater *Actinobacteria* led to decreased grazing upon this clade by some bacterivorous protists. Recently, Warnecke and coworkers demonstrated a correlation between increased UV solar radiation and the per cent abundance of clade

acl, which suggests another possible mechanism for the prevalence and persistence of these organisms. However, it seems unlikely that the small cell size and possible increased resistance to UV solar radiation are the only factors contributing to the success of this clade in freshwater. In addition to the aforementioned traits, the ability of acl to inhabit a diverse suite of lake types (represented by differences in trophic status, hydrology, elevation, chemistry, etc.) suggest possible underlying factor(s) that are unique to freshwater lakes, and specifically exploited by these organisms.

Burkert and colleagues (2003) using FISH identified the Beta II clade as the numerically dominant clade in their humic lake samples. Four 16S rRNA gene sequences were attributed to the Beta II clade in Crystal Bog Lake. One of these four phylotypes had an AFL of 797 base pairs, for which the ITS length is the same as the cultured *Polynucleobacter* strains studied by Hahn and coworkers (Hahn *et al.*, 2005). The corresponding ITS lengths of bacteria sampled from two different continents indicate that this phylotype may have a cosmopolitan distribution in freshwater. Furthermore, the Beta II clade was detected on more sampling dates than any other non-*Actinobacterial* clade (Fig. 5). This degree of persistence over time supports the conclusions of previous studies, based on coarser scales of temporal resolution, that the Beta II clade is a common pelagic resident in freshwater systems (Burkert *et al.*, 2003; Hahn, 2003) and may (like the acl clade) exploit a particular niche that is intrinsic to freshwater systems.

Although Crystal Bog Lake contained many sequences from the *Gammaproteobacteria* phylum, in general these phylotypes were detected on <50% of the sampling dates. The extreme temporal variability within this phylum (Fig. 5) suggests these bacteria may be transient community members washing in from the surrounding landscape, or often present at levels below detection limits. Although there are few *Gammaproteobacteria* 16S rRNA gene sequences recognized as freshwater-specific (Zwart *et al.*, 2002), members of this phylum often make up a smaller but still significant portion of the bacterioplankton community (Pernthaler *et al.*, 2004; Simek *et al.*, 2005).

Recent studies have shown that members of the *Bacteroidetes* phylum represent a large percentage of the bacterial community in lakes, especially during grazing periods (Pernthaler *et al.*, 2004). This trend is seemingly due to the distinctive filamentous morphology assumed by these bacteria, which significantly increases their resistance to grazing by protistan bacterivores. The greatest number of AFLs that we did not identify occurred during the intense 2002 bacterivory period (data not shown), suggesting that our clone libraries may not have adequately sampled the diversity of *Bacteroidetes* community members present on these dates. Furthermore, filamen-

tous bacteria were enriched during this period (Kent *et al.*, 2004). If members of the filamentous LD2 *Bacteroidetes* clade (Pernthaler *et al.*, 2004) were highly prevalent during intense grazing periods, then it is quite possible that these bacteria were part of the unidentified mid-spring 2002 population. The lack of detection of these organisms may be due to the use of universal bacterial primers with mismatches to many members of this phylum (O'Sullivan *et al.*, 2004).

#### Food web interactions

The data collected in 2000 and 2001 indicated a correlation between the change in BCC and the change in dominant phytoplankton regime (Kent *et al.*, 2004). However, little evidence existed for a similar relationship between the measured chemical/physical parameters and BCC dynamics. Although the data suggested a relationship between the phytoplankton and bacterioplankton communities, it was also apparent that an increased sampling effort would be needed to perceive this relationship more accurately. Thus, to examine a more relevant temporal scale for the phytoplankton–bacterioplankton relationship, the 2002 samples were taken more frequently than at the previous biweekly pace.

The majority of individual community phylotype persistence patterns (assessed by AFL relative fluorescence) were highly correlated to the phytoplankton succession in 2002 (Fig. 6). Within the overall BCC pattern, AFL assemblages demonstrated unique patterns correlated to individual phytoplankton regimes (Table 2). The persistence pattern of several AFLs associated with the acl-B clade of *Actinobacteria* was negatively correlated with the intense bacterivory period (Table 2). This relative reduction during an intense bacterivory period may indicate effective grazing on this clade. However, several studies including controlled mesocosm feeding experiments have shown that *Actinobacteria*, which are generally very small, are less grazed upon than their freshwater counterparts (Pernthaler *et al.*, 2001; Simek *et al.*, 2005). On the other hand, these same studies showed a significant decrease in total *Actinobacterial* cells during increased bacterivory. A significant decrease in total cells combined with the increase in abundance of filamentous organisms seen during this period may be the cause of the negative correlation between *Actinobacteria* and flagellate grazers. Because no members of the acl clade of *Actinobacteria* have been cultured, this clade's ecophysiology remains unknown. As described above, Burkert and colleagues (2003) conjecture that acl clade members could possess attributes similar to the related soil *Actinomycete* group, which produces peroxidases capable of breaking down recalcitrant compounds such as humic acids, one of the most abundant carbon sources in humic lakes. This

hypothesis suggests that variation in *Actinobacteria* populations would be uncoupled to phytoplankton succession and the corresponding unique autochthonous carbon sources made available during each phytoplankton interval. Our data support this hypothesis, as the AFLs associated with the acl clade of *Actinobacteria* were generally not correlated to changes in the non-mixotrophic phytoplankton regimes (Table 2 and Fig. 6).

The majority of AFL assemblages (grouped by significant correlation patterns, Table 2) contained AFLs from a wide variety of the represented phyla. This observation may indicate that community assemblages of taxonomically diverse organisms are maintained in this lake and that these assemblages are selected for during the different phytoplankton regimes. Upon closer inspection, the *Beta*- and *Gammaproteobacteria* community dynamics show the strongest correlation to phytoplankton succession (Fig. 6), suggesting that as a whole the *Proteobacteria* phylum is most closely tied to phytoplankton dynamics. Members of this phylum seem to have the ability to grow quickly during shifts in nutrient availability (Burkert *et al.*, 2003), which may be brought about by phytoplankton succession. Thus, it appears that the *Actinobacteria* acl-B clade and the *Betaproteobacteria* Beta II clade, two of the most abundant and ubiquitous freshwater phylotypes, inhabit different ecological niches within the water column.

## Conclusions

The planktonic bacterial community of Crystal Bog Lake consisted of both persistent and transient populations. These contrasting population dynamics were divided among different bacterial phylotypes. The acl-B clade of *Actinobacteria* was by far the most prevalent phylotype over the 3 year study, and showed a significant negative correlation to the intense bacterivory period. Furthermore, this clade's dynamics were seemingly uncoupled to the changes in phytoplankton regime, which may indicate a preference for allochthonous nutrient sources. In contrast, phylotypes in the *Gammaproteobacteria* class exhibited extremely variable presence/absence patterns suggesting a transient existence in the lake. Phytoplankton and grazer communities provide 'bottom-up' and 'top-down' pressures respectively, which influence bacterial communities. In particular, the *Proteobacteria* phylum contributed heavily to unique bacterial assemblages that were selected for during the phytoplankton community transitions. The increase in relative abundance of the majority of phylotypes associated with the unique phytoplankton regimes may indicate elevated growth rates linked to the availability of autochthonous algal-derived nutrients for these organisms. Controlled community manipulation experiments will be needed to further examine the drivers of persistent and transient bacterial community members

as well as the phytoplankton community – bacterioplankton community relationship in humic lakes.

## Experimental procedures

### Study sites and sample collection

Crystal Bog Lake is a shallow humic lake located in the Northern Highlands State Forest in Vilas County, Wisconsin (89° 36' W long, 46° N lat). It is part of the North Temperate Lakes Long-Term Ecological Research program (Magnuson *et al.*, 1997). Detailed limnological data for this lake and sampling procedures have been described previously (Kent *et al.*, 2004). The physical/chemical data collected for this study included: total chlorophyll, dissolved organic carbon, ammonia, nitrate/nitrite, total oxygen, pH, total nitrogen, total phosphorus, total particulate matter and water temperature.

### Phytoplankton and HNF abundance

Phytoplankton enumeration and identification was carried out to species when possible as previously described (Kent *et al.*, 2004). Heterotrophic nanoflagellate cells were stained with DAPI and counted on black 0.2 µm PCTE filters as previously described (Kent *et al.*, 2004).

### Bacterioplankton community fingerprints

Bacterial community composition (BCC) and diversity were assessed using ARISA (Kent *et al.*, 2004). Relative abundance of individual phylotypes was inferred using the fluorescence of each individual peak normalized to total fluorescence within a profile, to account for run-to-run variation during fragment analysis while avoiding the significant distortion associated with presence-absence data transformations, as described previously (Yannarell and Triplett, 2004) and described below.

### Clone library construction

Clone libraries were constructed from 3 years of combined Crystal Bog Lake DNA samples, combined DNA samples from 2000, and combined DNA samples from 2002. Briefly, the 16S rRNA gene and the 16S–23S rRNA ITS region were amplified from pooled environmental DNA samples using primers 8F, 5'-AGAGTTTGATCMTGGCTCAG-3' (bacteria-specific, 16S rRNA gene), and 23SR, 5'-GGGTTBCCCCATTCRG-3' (bacteria-specific, 23S rRNA gene). PCR products were cloned into the pGEM-T Easy vector following the manufacturer's instructions (Promega, cat. #A1380).

### Sequence analysis

Cloned plasmid inserts were amplified directly from cells as described (Vergin *et al.*, 2001) using vector primers. The 16S rRNA gene portion of the cloned DNA was initially sequenced using the ABI Prism BigDye terminator sequencing kit (PE Applied Biosystems) with standard PCR sequencing reaction conditions using the primer 8F. Sequences were assigned

preliminary bacterial phylum associations based on the BLASTN (Altschul *et al.*, 1990, <http://www.ncbi.nih.gov/BLAST/>) and RDP-II Classifier programs (Cole *et al.*, 2003, <http://rdp.cme.msu.edu/classifier/classifier.jsp>). Following classification, all sequences were aligned using the ARB software package (Ludwig *et al.*, 2004) containing a publicly available 16S rRNA gene ARB database, January 2002 (Hugenholtz, 2002) supplemented with freshwater 16S rRNA gene sequences (described by Glockner *et al.*, 2000; Zwart *et al.*, 2002; Warnecke *et al.*, 2004). *Actinobacteria*-related sequences were also independently aligned. Sequences added to the existing ARB database were initially automatically aligned using the FAST\_ALIGNER ARB tool before the alignment was heuristically adjusted using primary and secondary rRNA structure as a guide.

Reference sequences were chosen for further sequencing of the 16S rRNA gene and intergenic spacer region. A total of 132 selected clones were additionally sequenced with the primers 515F, 5'-GTGCCAGCMGCCGCGGTAA-3', 1100F, 5'-CAACGAGCGAGACCCA-3', 1406F, 5'-TGYACACACGCCCCGT-3', 1492R, 5'-GGTTACCTGTTACGACTT-3' and 23SR, 5'-GGGTTBCCCCATTTCRG-3'. All partial and full-length 16S rRNA sequences were edited manually and assembled using the software Sequencher 3.1 (Gene Codes Corporation). Forty-nine clone sequences were identified as putative chimeras by the programs CHIMERA\_CHECK (<http://rdp.cme.msu.edu/>) or BELLEROPHON (Huber *et al.*, 2004, <http://foo.maths.uq.edu.au/~huber/bellerophon.pl>) and were eliminated from further analyses. Eighty-three nearly full-length (> 1300 bp) and an additional 206 partial (> 400 bp) 16S rRNA gene sequences and their corresponding AFLs were acquired and used during all subsequent analyses.

#### Linking fingerprints to phylogeny

Following amplification with vector primers of the 16S rRNA gene and 16S–23S rRNA ITS insert from each clone, the ITS of each clone insert was amplified with primers 1406F and 23SR. The amplified product was then analysed using the methods described previously for community ARISA (Kent *et al.*, 2004). This procedure allowed the matching of the AFL (measured as the number of nucleotides amplified with primers 1406F and 23SR) from an individual clone to the fragment lengths obtained in the 68 ARISA community fingerprints generated over 3 years (procedure recently described by Brown *et al.*, 2005). As 16S rRNA gene sequences were also obtained for each clone, it was then possible to apply multiple hierarchical OTU definitions to each ARISA peak in the fingerprint.

#### Phylogenetic reconstruction

Only nearly complete (> 1300 bp) 16S rRNA gene sequences were used for phylogenetic tree reconstruction. A 50% base frequency filter was calculated on the included sequences to exclude highly variable positions. An alignment of selected Crystal Bog Lake and other reference sequences were exported from ARB into the MrBayes software program v. 3.0 (Ronquist and Huelsenbeck, 2003) for phylogenetic reconstruction using Bayesian inference. A general time

reversible gamma-distributed rates variation model was specified. Three independent Markov Chain Monte Carlo analyses, each starting with random trees for each of four simultaneous chains, were run for 100 000 generations with sampling every 10 generations to create a posterior probability distribution of 10 000 trees. Trees created before chain stabilization were discarded with appropriate burn-in values and a 50% majority-rule tree was calculated. Partial sequences were added to the alignment and the MrBayes analysis was rerun. Placement of the partial sequences in MrBayes was compared with placement of the same partial sequences by the maximum parsimony tool in ARB, while preventing changes in tree topology. Partial sequences were then appropriately added to the final trees so as not to affect final tree topology.

All Crystal Bog Lake 16S rRNA gene sequences were grouped into defined OTUs (clade and species) based on ARB phylogeny and sequence identity determined by the program DOTUR (Schloss and Handelsman, 2005). Species were identified based on a furthest neighbour 97% 16S rRNA gene sequence identity threshold. Clades were identified based on tree topology and consistently share  $\geq 90\%$  16S rRNA gene sequence identity.

#### Community composition data transformations

Presence and absence analysis of ARISA profiles introduces a significant arbitrary bias towards rare taxa (Yannarell and Triplett, 2005). Furthermore, the relative fluorescence produced by a single ARISA peak is highly reproducible across PCR runs (Yannarell and Triplett, 2004) and may be used to compare samples (Hewson and Fuhrman, 2004). These authors concluded that significant valuable information is lost when the relative contribution of each individual peak to the total fluorescence in an ARISA profile is not taken into consideration during data analysis. Therefore, we used relative fluorescence produced by each individual phylotype in all analyses to infer relative abundance of that phylotype in the original sample. We do not utilize the relative fluorescence information as a method to compare the abundance of different phylotypes, but instead use it solely to examine changes in relative contribution of a single phylotype to the community over time. A detailed discussion of sensitivity analyses conducted using the relative fluorescence data transformation can be found elsewhere (Yannarell and Triplett, 2005).

#### Statistical analysis

The Pearson product-moment correlation coefficient ( $r$ ) was calculated for each bacterial phylotype (i.e. unique ARISA fragment) relative fluorescence and the biovolume of the dominant phytoplankton or the abundance of HNFs in Crystal Bog Lake across all sample dates in 2002. The Pearson product-moment correlation coefficient was also calculated for each bacterial phylotype and environmental parameters (total chlorophyll, dissolved organic carbon, total nitrogen, total phosphorus, dissolved oxygen, lake pH, total particulate matter and water temperature) gathered during 2002. A  $P$ -value of  $< 0.001$  was used to establish a significant correlation

between the variables. ARISA fragments associated with more than one clade were excluded from these analyses. The biweekly sampling effort during 2000 and 2001 was not at a sufficient frequency to accurately follow the pace of change in the bacterial community as related to the above-mentioned parameters; these samples were not included in correlation analyses.

The similarity of relative fluorescence patterns from individual phylotypes was evaluated using the Bray-Curtis Index of Similarity (Magurran, 1988). Phylotype assemblages were assigned according to their correlation with unique phytoplankton regimes. The largest *r*-value, indicating the strongest correlation, was used to determine the assemblage assignment for those phylotypes with more than one significant phytoplankton regime correlation (Table 2). An analysis of similarity (ANOSIM) was carried out using rank dissimilarities to test the hypothesis that the relative fluorescence patterns within the assigned assemblages were more similar than the relative fluorescence patterns between assemblages (Clarke, 1993). These multivariate analyses (based on non-metric multidimensional scaling) were performed using the statistical package PRIMER 5 for Windows v. 5.2.7.

ARISA fragment length relative fluorescence data were subjected to CCA using phytoplankton biovolume and HNF abundance as potential explanatory variables. Only those AFLs exhibiting a significant correlation ( $P < 0.001$ ) to a phytoplankton regime (listed in Table 2) were included in the analysis. Canonical correspondence analysis represents individual AFLs as occurring in a theoretical environmental space, or ordination space, which is defined by the potential explanatory variables included in the analysis. The CCA axes represent linear combinations of the protist population data included in the analysis. This approach allows us to explore relationships between bacterial populations and food web factors that may influence the dynamics of bacterial populations of interest. The significance of the set of explanatory variables was tested by a Monte Carlo permutation test (999 permutations). The position of each AFL in ordination space represents its correlation with the explanatory variables. AFLs plotting close to an arrow representing an explanatory variable are strongly correlated with that variable, and populations which plot close together in ordination space have similar responses to the food web factors included in the analysis. Canonical correspondence analysis was carried out using Canoco for Windows, Version 4.51 (Biometris-Plant Research International, 1997–2003). Analysis settings included biplot scaling focusing on interspecies distances, and permutation tests which accounted for the time-series structure of the data.

#### Nucleotide sequences and ARISA profiles

All 16S rRNA gene sequences generated in the current study that were included in the phylogenetic analyses have been submitted to GenBank under Accession numbers AY792221 to AY792314 and DQ093399 to DQ093408.

The ARISA profiles and ancillary environmental data used to generate Figs 5 and 6 are available in downloadable form in our online database (<http://microbes.limnology.wisc.edu/>) (Jacob *et al.*, 2005).

#### Acknowledgements

The authors wish to thank Anthony Yannarell and Stuart Jones for helpful discussions. A sequence alignment of freshwater *Actinobacteria* 16S rRNA genes was kindly provided by Falk Warnecke and Jakob Pernthaler. Thanks to Martin Hahn for unpublished information about the 16S–23S rRNA intergenic spacer region in *Polynucleobacter* isolates. This research was supported in part by National Science Foundation Grant MCB-9977903 and MCB-0401987 to EWT, and by the National Institutes of Health Biotechnology Training Program Grant 5 T32 G08349 (RJN).

#### References

- Acinas, S.G., Marcelino, L.A., Klepac-Ceraj, V., and Polz, M.F. (2004) Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *J Bacteriol* **186**: 2629–2635.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Brown, M.V., Schwalbach, M.S., Hewson, I., and Fuhrman, J.A. (2005) Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity: development and application to a time series. *Environ Microbiol* **7**: 1466–1479.
- Burkert, U., Warnecke, F., Babenzien, D., Zwirnmann, E., and Pernthaler, J. (2003) Members of a readily enriched beta-proteobacterial clade are common in surface waters of a humic lake. *Appl Environ Microbiol* **69**: 6550–6559.
- Chao, A. (1987) Estimating the population-size for capture recapture data with unequal catchability. *Biometrics* **43**: 783–791.
- Clarke, K.R. (1993) Nonparametric multivariate analyses of changes in community structure. *Aust J Ecol* **18**: 117–143.
- Clayton, M.K., and Frees, E.W. (1987) Nonparametric-estimation of the probability of discovering a New Species. *J Am Stat Assoc* **82**: 305–311.
- Cole, J.R., Chai, B., Marsh, T.L., Farris, R.J., Wang, Q., Kulam, S.A. *et al.* (2003) The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res* **31**: 442–443.
- Cotner, J.B., and Biddanda, B.A. (2002) Small players, large role: microbial influence on biogeochemical processes in pelagic aquatic ecosystems. *Ecosystems* **5**: 105–121.
- Crump, B.C., Kling, G.W., Bahr, M., and Hobbie, J.E. (2003) Bacterioplankton community shifts in an arctic lake correlate with seasonal changes in organic matter source. *Appl Environ Microbiol* **69**: 2253–2268.
- Eiler, A., and Bertilsson, S. (2004) Composition of freshwater bacterial communities associated with cyanobacterial blooms in four Swedish lakes. *Environ Microbiol* **6**: 1228–1243.
- Glockner, F.O., Zaichikov, E., Belkova, N., Denissova, L., Pernthaler, J., Pernthaler, A., and Amann, R. (2000) Comparative 16S rRNA analysis of lake bacterioplankton reveals globally distributed phylogenetic clusters including

- an abundant group of *Actinobacteria*. *Appl Environ Microbiol* **66**: 5053–5065.
- Good, I.L. (1953) The population frequencies of species and the estimation of population parameters. *Biometrika* **40**: 237–264.
- Hahn, M.W. (2003) Isolation of strains belonging to the cosmopolitan *Polynucleobacter necessarius* cluster from freshwater habitats located in three climatic zones. *Appl Environ Microbiol* **69**: 5248–5254.
- Hahn, M.W., Pockl, M., and Wu, Q.L. (2005) Low intraspecific diversity in a *Polynucleobacter* subcluster population numerically dominating bacterioplankton of a freshwater pond. *Appl Environ Microbiol* **71**: 4539–4547.
- Hewson, I., and Fuhrman, J.A. (2004) Richness and diversity of bacterioplankton species along an estuarine gradient in Moreton Bay, Australia. *Appl Environ Microbiol* **70**: 3425–3433.
- Huber, T., Faulkner, G., and Hugenholtz, P. (2004) Bellerophon, a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **10**: 1093.
- Hugenholtz, P. (2002) Exploring prokaryotic diversity in the genomic era. *Genome Biol* **3**: 1–8.
- Jacob, C., Kent, A.D., Benson, B.J., Newton, R.J., Triplett, E.W., and McMahon, K.D. (2005) Biological databases for linking large microbial and environmental datasets. *Proceedings of the 9th World Multiconference on Systematics, Cybernetics and Informatics*, Orlando, FL, USA.
- Kemp, P.F., and Aller, J.Y. (2004) Bacterial diversity in aquatic and other environments: what 16S rDNA libraries can tell us. *FEMS Microbiol Ecol* **47**: 161–177.
- Kent, A.D., Jones, S.E., Yannarell, A.C., Graham, J.M., Lauster, G.H., Kratz, T.K., and Triplett, E.W. (2004) Annual patterns in bacterioplankton community variability in a humic lake. *Microb Ecol* **48**: 550–560.
- Klappenbach, J.A., Dunbar, J.M., and Schmidt, T.M. (2000) rRNA operon copy number reflects ecological strategies of bacteria. *Appl Environ Microbiol* **66**: 1328–1333.
- Lindstrom, E.S. (2000) Bacterioplankton community composition in five lakes differing in trophic status and humic content. *Microb Ecol* **40**: 104–113.
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.
- Madigan, M.T., Martinko, J.M., and Parker, J. (2002) *Brock Biology of Microorganisms*. Upper Saddle River, NJ, USA: Prentice Hall/Pearson Education.
- Magnuson, J.J., Kratz, T.K., Allen, T.F., Armstrong, D.E., Benson, B.J., Bowser, C.J. et al. (1997) Regionalization of long-term ecological research (LTER) on north temperate lakes. *Verh Int Verein Limnol* **26**: 522–528.
- Magurran, A.E. (1988) *Ecological Diversity and its Measurement*. Princeton, NJ, USA: Princeton University Press.
- Methe, B.A., and Zehr, J.P. (1999) Diversity of bacterial communities in Adirondack lakes: do species assemblages reflect lake water chemistry? *Hydrobiologia* **401**: 77–96.
- Nold, S.C., and Zwart, G. (1998) Patterns and governing forces in aquatic microbial communities. *Aquat Ecol* **32**: 17–35.
- O'Sullivan, L.A., Fuller, K.E., Thomas, E.M., Turley, C.M., Fry, C.J., and Weightman, A.J. (2004) Distribution and culturability of uncultivated 'AGG58 cluster' of the *Bacteroidetes* phylum in aquatic environments. *FEMS Microbiol Ecol* **47**: 359–370.
- Pernthaler, J., Posch, T., Simek, K., Vrba, J., Pernthaler, A., Glockner, F.O., et al. (2001) Predator-specific enrichment of *Actinobacteria* from a cosmopolitan freshwater clade in mixed continuous culture. *Appl Environ Microbiol* **67**: 2145–2155.
- Pernthaler, J., Zollner, E., Warnecke, F., and Jurgens, K. (2004) Bloom of filamentous bacteria in a mesotrophic lake: identity and potential controlling mechanism. *Appl Environ Microbiol* **70**: 6272–6281.
- Pinhassi, J., Sala, M.M., Havskum, H., Peters, F., Guadayol, O., Malits, A., and Marrase, C.L. (2004) Changes in bacterioplankton composition under different phytoplankton regimens. *Appl Environ Microbiol* **70**: 6753–6766.
- Ronquist, F., and Huelsenbeck, J.P. (2003) MrBayes 3: bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.
- Rooney-Varga, J.N., Giewat, M.W., Savin, M.C., Sood, S., LeGresley, M., and Martin, J.L. (2005) Links between phytoplankton and bacterial community dynamics in a coastal marine environment. *Microb Ecol* **49**: 163–175.
- Schloss, P.D., and Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71**: 1501–1506.
- Simek, K., Hornak, K., Jezbera, J., Masin, M., Nedoma, J., Gasol, J., and Schauer, M. (2005) Influence of top-down and bottom-up manipulations on the R-BT065 subcluster of  $\beta$ -*Proteobacteria* an abundant group in bacterioplankton of a freshwater reservoir. *Appl Environ Microbiol* **71**: 2381–2390.
- Vergin, K.L., Rappe, M.S., and Giovannoni, S.J. (2001) Streamlined method to analyze 16S rRNA gene clone libraries. *Biotechniques* **30**: 938–944.
- Warnecke, F., Amann, R., and Pernthaler, J. (2004) *Actinobacterial* 16S rRNA genes from freshwater habitats cluster in four distinct lineages. *Environ Microbiol* **6**: 242–253.
- Warnecke, F., Sommaruga, R., Seka, R., Hofer, J., and Pernthaler, J. (2005) Abundances, identity, and growth state of *Actinobacteria* in mountain lakes of different UV transparency. *Appl Environ Microbiol* **71**: 5551–5559.
- Wetzel, R.G. (2001) *Limnology: Lake and River Ecosystems*. San Diego, CA, USA: Academic Press.
- Yannarell, A.C., and Triplett, E.W. (2004) Within- and between-lake variability in the composition of bacterioplankton communities: investigations using multiple spatial scales. *Appl Environ Microbiol* **70**: 214–223.
- Yannarell, A.C., and Triplett, E.W. (2005) Geographic and environmental sources of variation in lake bacterial community composition. *Appl Environ Microbiol* **71**: 227–239.
- Yannarell, A.C., Kent, A.D., Lauster, G.H., Kratz, T.K., and Triplett, E.W. (2003) Temporal patterns in bacterial communities in three temperate lakes of different trophic status. *Microb Ecol* **46**: 391–405.
- Zwart, G., Crump, B.C., Agterveld, M.P.K.V., Hagen, F., and Han, S.K. (2002) Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat Microb Ecol* **28**: 141–155.