Environmental Microbiology (2009) **11**(9), 2463–2472 doi:10.1111/j.1462-2920.2009.01977.x

Evidence for structuring of bacterial community composition by organic carbon source in temperate lakes

Stuart E. Jones,[1] Ryan J. Newton[2] and Katherine D. McMahon[3,4]*

[1]*Limnology and Marine Sciences Program,*
[2]*Microbiology Doctoral Training Program, and*
Departments of [3]*Civil and Environmental Engineering and* [4]*Bacteriology, University of Wisconsin-Madison, 1415 Engineering Dr., Madison, WI 53706, USA.*

Summary

Water entering lakes from the surrounding watershed often delivers large amounts of terrestrial-derived dissolved organic carbon (DOC) that can contribute to aquatic bacterial production. However, research suggests that phytoplankton-derived DOC is more labile than its terrestrial counterpart, owing to microbial processing of terrestrial-derived DOC along its flow path to surface waters. The ratio of water colour (absorbance at 440 nm) to chlorophyll *a* has been suggested as a simple measure of the relative contribution of terrestrial and aquatic primary production to aquatic secondary production. To explore the correlation between primary DOC source and the occurrence of bacterial taxonomic groups, we conducted a survey of bacterial 16S rRNA gene composition in 15 lakes positioned along a water colour : chlorophyll *a* gradient. Our goal was to identify bacterial taxa occurrence patterns along the colour : chlorophyll *a* gradient that may indicate a competitive advantage for bacterial taxa using terrestrial or aquatic carbon. We observed a large number of bacterial taxa occurrence patterns suggestive of carbon substrate niche partitioning, especially when relatively highly resolved taxonomic groups were considered. Our survey supports the hypothesis that bacterial taxa partition along a carbon substrate source gradient and highlights carbon source–bacterial interactions that should be explored further.

Received 8 February, 2009; accepted 25 May, 2009. *For correspondence. E-mail tmcmahon@engr.wisc.edu; Tel. (+1) 608 890 2836; Fax (+1) 608 262 5199.

Introduction

Approximately 1.9 Pg of carbon is exported from terrestrial to aquatic ecosystems by fluvial processes each year (Cole *et al*., 2007). Therefore, recipient aquatic ecosystems have the potential to process large amounts of terrestrially derived (allochthonous) carbon. The fate of terrestrial carbon reaching aquatic ecosystems (burial, mineralization or incorporation into biomass) has important implications for the global carbon cycle (Cole *et al*., 2007) and production of biomass in aquatic systems (Kritzberg *et al*., 2004; Carpenter *et al*., 2005). Biological processes acting on allochthonous carbon can be observed at ecosystem (Cole *et al*., 1994; Hanson *et al*., 2003) and organismal scales (Kritzberg *et al*., 2004; Carpenter *et al*., 2005). As the primary consumers of dissolved organic carbon (DOC), bacteria drive these processes by using terrestrial subsidies for both secondary production and generation of energy (Bergstrom and Jansson, 2000; Kritzberg *et al*., 2004). Previous work has not resolved whether different bacterial taxa specialize in the metabolism of allochthonous DOC or if all bacteria augment their metabolism with the terrestrial carbon subsidy. We might expect the ability to process allochthonous carbon, which is generally thought to be recalcitrant due to its aromatic and reduced nature (Hessen and Tranvik, 1998), to be taxonomically limited. The cost of enzymatic machinery to process allochthonous carbon likely imposes strong selection for organisms able to utilize the recalcitrant resource (Hessen and Tranvik, 1998). Direct evidence of the potential for bacteria to specialize in the use of a particular carbon substrate was presented by Cottrell and Kirchman (2000), who used fluorescent *in situ* hybridization linked with microautoradiography to demonstrate selective use of multiple organic carbon substrates by different bacterial taxa in the ocean. These findings suggest that in order to adequately predict rates of carbon mineralization or burial in aquatic ecosystems, ecologists must understand the influence of DOC source on both bacterial community composition and metabolism.

Despite strong evidence of differential substrate use by bacterial taxa, only a few studies have attempted to identify freshwater bacterial operational taxonomic units

(OTUs) that are associated uniquely with terrestrial- or aquatic-derived (autochthonous) organic matter. Time-series observations (Crump *et al*., 2003; Haukka *et al*., 2005; Kritzberg *et al*., 2006; Nelson, 2009), mesocosm experiments (Eiler *et al*., 2003; Judd *et al*., 2006) or spatial surveys (Lindstrom, 2000) have been used to demonstrate that changes in DOC source drive changes in the bacterial community. However, most previous studies were not able to explicitly relate the occurrence of an OTU in a lake to the dominant DOC source. Linking specific bacterial OTUs with autochthonous or allochthonous carbon will allow us to predict the occurrence of particular bacterial taxa in lakes of a given trophic status.

Researchers hoping to identify specific taxa responsible for processing autochthonous DOC, characterized by labile, small and oxidized organic molecules (Biddanda and Benner, 1997), have often focused on bacterial communities present during and after phytoplankton blooms. For example, diatom blooms were induced in mesocosms to identify bacterial taxa that responded to the associated increase in autochthonous DOC availability (Riemann *et al*., 2000). Similarly, 16S ribosomal RNA (rRNA) gene sequences revealed that *Bacteroidetes* and *Proteobacteria* phyla responded to cyanobacterial blooms in four Swedish lakes (Eiler and Bertilsson, 2004). Eiler and Bertilsson (2004) suggested that their most frequently recovered sequence clusters were the bacterial groups selected for during cyanobacterial blooms, which are assumed to provide an enrichment of autochthonous DOC. However, attempts to identify taxa that preferentially use terrestrial carbon are less common (Judd *et al*., 2006; Perez and Sommaruga, 2006).

Recently the colour : chlorophyll *a* ratio (CtCh) was suggested as a proxy for the extent of allochthony (terrestrial support) of a lake ecosystem, with 'colour' being defined by light absorbance at 440 nm (Carpenter *et al*., 2005; Pace *et al*., 2007; Solomon *et al*., 2008; Weidel *et al*., 2008). The CtCh ratio is a useful metric because of its ease of measurement and widespread availability in the literature. Water colour is a measure of chromophoric dissolved organic matter (CDOM). CDOM is assumed to be primarily derived from terrestrial sources (Hessen and Tranvik, 1998) and therefore approximates concentrations of terrestrial DOC in an aquatic ecosystem. Chlorophyll *a* is used as a measure of phytoplankton biomass and is used here as a proxy for aquatic DOC production.

We conducted a survey of bacterial 16S rRNA genes in 15 lakes in Wisconsin, USA, along a CtCh gradient to identify bacterial OTUs associated with terrestrial- or aquatic-derived DOC. Spatial surveys have been successfully conducted to identify environmental variables that may be structuring limnetic bacterial assemblages using community fingerprinting techniques (Yannarell and Triplett, 2005; Newton *et al*., 2007) or reverse-line blot hybridizations (Lindstrom *et al*., 2005). Here we used 16S rRNA gene cloning and sequencing to determine the phylogenetic affiliation of OTUs associated with particular DOC regimes. We hypothesized that the relative contributions of terrestrial and aquatic sources to a lake's DOC pool would influence bacterial community structure. We predicted that lakes at the extremes of the CtCh gradient should harbour distinct bacterial assemblages and provide clues as to which bacterial OTUs preferentially use terrestrial- or aquatic-derived DOC.

Results

The lakes selected for this study represented a broad range of CtCh ratios (0.047–0.430) with chlorophyll *a* ranging from 4.2 to 31.2 $\mu g\ l^{-1}$ and colour from 0.49 to 4.86 m^{-1}. The ranges of other environmental parameters were also broad and spanned the ranges of the majority of lakes that have been studied in Wisconsin, USA (Table 1). Notably, our hypothesized link between relative contribution of terrestrial- and aquatic-derived carbon and DOC quality or availability appears to hold in our set of lakes. The CtCh ratio was significantly correlated with specific UVA absorbance (SUVA-254; $n = 15$, $r = 0.56$, $P < 0.05$), a measure of DOC aromaticity or recalcitrance.

Multivariate statistical tools were used to compare community composition across lakes and to relate observed differences to measured environmental variables. We detected moderately strong linkages between bacterial community structure in a lake and the relative contribution of autochthonous and allochthonous organic matter (Table 2). The strongest observed relationship between community ordination sample scores and an environmental driver was between the first axis from the Unifrac principal coordinates analysis (PCA) and CtCh (Fig. 1A, Table 2). In addition, colour explained a significant portion of the variation along the first axis from the correspondence analysis (CA) conducted using OTUs (Fig. 1B, Table 2), and chlorophyll *a* explained a significant portion of the variation along the second axis. Weaker correlations with pH and Secchi depth were also observed. However, the second axis only explained 15% variation while the first axis captured 39%.

At the broadest level of taxonomic organization, we recovered sequences from nine phylogenetic groups in our data set. These nine groups corresponded to phyla and proteobacterial classes. The most commonly recovered phylum was *Actinobacteria*, followed by the *Bacteroidetes* phylum, and the *Beta* class of *Proteobacteria* (Fig. 2); clones from these groups were observed in all lakes. Clones from *Cyanobacteria* (12 out of 15 lakes), *Alphaproteobacteria* (12/15) and *Verrucomicrobia* (11/15) were recovered in the majority of lakes. 16S rRNA gene sequences from the *Gamma* and *Delta* classes of *Proteo-*

Table 1. Lake physical and chemical parameters.

Lake	Lat (N)	Long (W)	Area (ha)	pH	DOC (mg l^{-1})	TP (μg l^{-1})	Chla (μg l^{-1})	Colour (m^{-1})[a]	CtCh
Little Arbor Vitae (LA)	45.91	89.62	216	7.2	3.9	9.3	10.5	0.49	0.047
Salmo Pond (SA)	43.12	89.69	1	7.6	2.0	11.6	7.3	0.37	0.049
Delton (DE)	43.60	89.78	108	8.6	4.6	17.6	31.2	1.54	0.049
Cox Hollow (CH)	43.01	90.11	39	8.5	4.0	12.2	13.9	0.83	0.060
Twin Valley (TW)	43.03	90.09	62	8.5	3.5	9.2	11.9	0.73	0.061
Green (GR)	43.79	89.04	2973	8.6	5.4	8.7	7.4	0.48	0.064
Mendota (ME)	43.10	89.41	3938	8.6	5.8	83.0	11.2	0.75	0.067
Trout (TR)	46.03	89.66	1608	7.3	3.7	4.5	4.2	0.31	0.075
Brandy (BR)	45.91	89.70	45	6.6	4.3	6.1	7.2	0.66	0.092
Little Trout (LT)	46.06	89.86	407	6.6	6.4	5.2	6.6	0.78	0.119
Clear (CL)	42.80	88.98	33	8.3	5.6	9.5	5.6	0.78	0.139
Ike Walton (IW)	46.03	89.81	576	6.2	3.4	4.2	4.7	1.19	0.253
Mirror (MI)	43.57	89.81	55	8.2	3.1	30.4	4.5	1.35	0.300
Red Cedar (RC)	42.98	88.98	145	7.2	11.8	11.6	5.2	2.07	0.398
Hook (HO)	42.94	89.34	51	5.4	12.2	19.9	11.3	4.86	0.43

a. Colour is defined as wavelength-specific absorbance at 440 nm, with units of absorbance per metre (Cuthbert and Del Giorgio, 1992).
All values are seasonal averages from single spring, summer and fall measures, collected as described previously (Yannarell and Triplett, 2005). Lakes are ranked in order of increasing colour : chlorophyll *a* ratio (CtCh). Lat, latitude; Long, longitude; DOC, dissolved organic carbon; TP, total phosphorus; Chla, chlorophyll *a*.

bacteria, as well as the *Chloroflexi* and *Firmicutes* phyla were observed in less than 5 of the 15 libraries.

We found few OTUs that occurred in both allochthonous lakes (CtCh > 0.25) and autochthonous lakes (CtCh < 0.08, 13 OTUs) or both allochthonous lakes and intermediate lakes (CtCh 0.09–0.14, 3 OTUs) (Fig. 3). More OTUs (24) had representative sequences recovered from at least one lake in each of the three CtCh categories (Fig. 3). Finally, Unifrac analysis suggested that each CtCh category represented a distinct community (Pairwise Unifrac significance test, Bonferroni corrected).

In order to determine whether OTUs demonstrated a preference for different levels of allochthony, we conducted hierarchical clustering of OTU occurrence patterns. We also looked for generalized patterns of occurrence at broader taxonomic scales (the lineage or phylum/class level). Divergent clusters of OTUs were evident based upon their occurrence patterns in the three lake types defined by CtCh (Fig. 4). The occurrence patterns of various freshwater lineages and phylogenetic affiliations of sequences from each OTU are presented in Fig. 5. The majority of bacterial phyla were represented in each cluster, but distinct lineages within each phylum seemed to occur in a given CtCh category or pair of CtCh categories. Many of the observed OTUs were representative of previously observed freshwater lineages. However, others did not fall within those freshwater lineages, based upon insertion into a reference freshwater 16S rRNA tree using the ARB parsimony insertion tool (Ludwig *et al.*, 2004), and could only be identified at the phylum or class level.

Our results were robust to the taxonomic resolution used to define OTUs (Table 3). We saw strong agreement in the relationships among lake bacterial communities when using increasingly more refined sequence identify cut-offs. This observation weakened at the extremes (95% and 99.5%), but remained strong at traditional definitions of bacterial 'species' (97–99% 16S sequence identity; Stackebrandt and Goebel, 1994; Stackebrandt and Ebers, 2006). In addition, agreement between

Table 2. R^2 for linear regressions between correspondence analysis (CA) and Unifrac PCA axes and environmental variables.

Environmental variable	Unifrac PCA Axis 1	Unifrac PCA Axis 2	CA[a] Axis 1	CA Axis 2
Lake area (ha)	0.09	0.09	0.05	–0.07
Secchi depth (m)	–0.03	0.21*	0.16	0.15
pH	0.23*	–0.03	0.17	0.14
Chlorophyll *a* (μg l^{-1})	0.19	0.15	–0.08	0.28*
Colour	0.21	0.08	0.34*	–0.06
DOC (mg l^{-1})	0.11	–0.05	0.03	–0.08
Total phosphorus (μg l^{-1})	–0.04	0.07	–0.07	0.0
Colour : chlorophyll *a*	0.45*	0.08	0.43*	–0.06

a. Correspondence analysis was conducted using 97% identity OTU definitions as described for Fig. 1B.
The asterisk (*) indicates a significant regression at $\alpha = 0.05$.

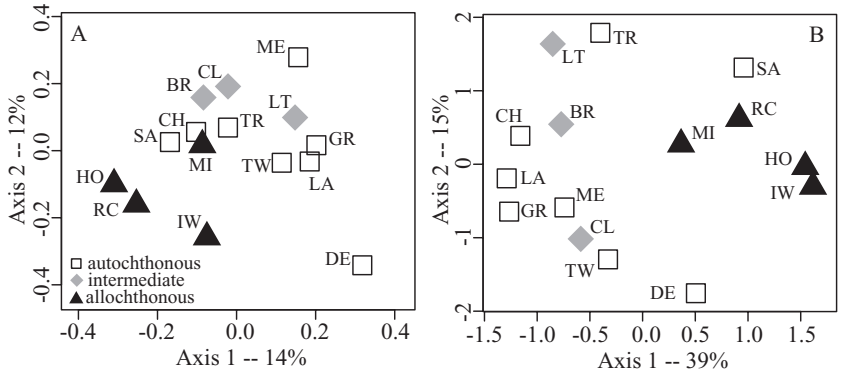

Fig. 1. A. Principal coordinates analysis (PCA) ordination of the bacterial communities from 15 north temperate lakes. The PCA was created using Unifrac distances (see *Experimental procedures*).
B. Correspondence analysis ordination of the bacterial communities. The correspondence analysis was based on relative recovery rates of sequences in 168 operational taxonomic units defined at 97% sequence identity using DOTUR.
Bacterial communities are coded by the colour : chlorophyll *a* ratio of the lake from which they were collected (autochthonous, < 0.08; intermediate, 0.09–0.14; allochthonous, > 0.25). Lake names and abbreviations can be found in Table 1. Axis labels indicate per cent variation explained.

environmental correlations with the Unifrac PCA (no OTU definition) and 97% identity OTU CA suggests that our findings are not likely to be sensitive to the per cent sequence identity cut-off (above ~97%) used to define OTUs.

Discussion

The relative contributions of terrestrial- and aquatic-derived organic matter to a lake determine whether it is a net source or sink of CO_2 to the atmosphere (Hanson *et al*., 2003) and can strongly influence the secondary productivity of the lake (Carpenter *et al*., 2005). We hypothesized that the CtCh ratio would be a strong driver of bacterial community composition and predicted that distinct assemblages would be associated with high- and low-CtCh lakes respectively. Our results support this hypothesis, and suggest that the relative contributions of terrestrial and aquatic carbon to the DOC pool represent an important selective force on bacterial community composition. The CtCh ratio appears to be a helpful metric for characterizing relative contributions of these two organic matter sources to lake secondary production (Carpenter *et al*., 2005) and structuring of bacterial community composition.

In reality, all lake ecosystems are supported by some combination of terrestrial- and aquatic-derived DOC. It is likely that niche partitioning between carbon substrate sources occurs within a single lake. Previous research suggests that autochthonous carbon, approximated by chlorophyll *a* concentration in this study, is more labile and

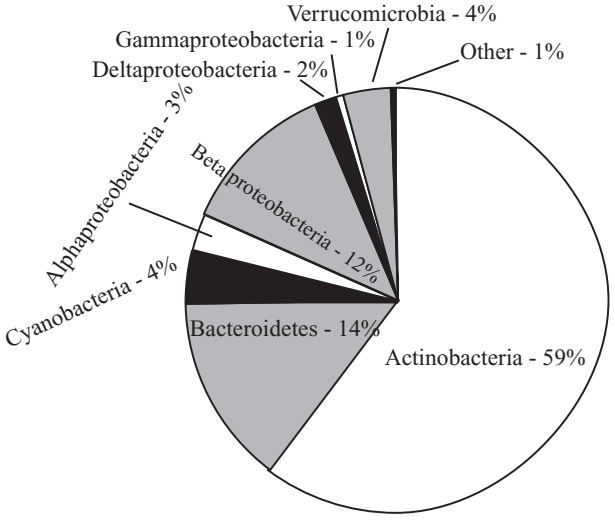

Fig. 2. Proportion of all clones ($n = 1094$) designated by phyla, based on the Classifier function of RDP (Wang *et al*., 2007).

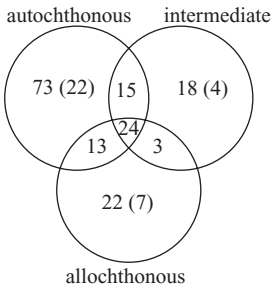

Fig. 3. A Venn diagram of overlap in occurrence of 168 DOTUR operational taxonomic units (OTUs) defined at 97% sequence identity. Each circle represents a set of lakes within the defined colour : chlorophyll *a* (CtCh) categories (autochthonous, < 0.08, intermediate, 0.09–0.14, allochthonous, > 0.25). Numbers in parentheses are OTUs observed in only one CtCh category, but observed in more than one clone library (i.e. more than one lake). Note that the number of lakes in each category differ: eight autochthonous lakes, three intermediate lakes and four allochthonous lakes.

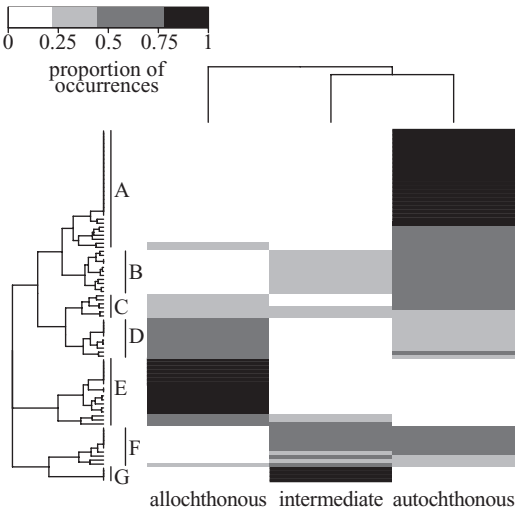

Fig. 4. Two-dimensional hierarchical clustering of 97% sequence identity operational taxonomic units (OTUs) defined with DOTUR. Only OTUs containing more than one sequence are included in the analysis. Rows are OTUs and columns are colour : chlorophyll *a* (CtCh) categories. The number of lakes in the CtCh > 0.25 (allochthonous), CtCh 0.09–0.14 (intermediate) and CtCh < 0.08 (autochthonous) groups were 4, 3 and 8 respectively. The colour of each element of the OTU-by-CtCh matrix indicates the proportion of occurrences for that OTU in that set of lakes.

used with higher growth efficiency by bacteria (Biddanda *et al*., 2001). However, the availability of algal-derived carbon is highly variable in time. The efficiency of bacterial growth on allochthonous carbon may be lower (Hobbie, 1988; Eiler *et al*., 2003), but terrestrial carbon is often more consistently available (Wetzel, 2001; Hanson *et al*., 2006). If members of a freshwater bacterial lineage were able to persist on terrestrial-derived carbon, the individuals could adapt a slow, but continuous growth strategy in contrast to a rapid and short-lived growth strategy that would be required by dependence on algal-derived carbon. Thus, we predict that taxa with the ability to use terrestrial carbon would be abundant in lakes with high CtCh, while lakes with low CtCh should be enriched in those preferring algal carbon.

Our sampling procedure (< 100 clones per lake) likely identified only the abundant members of each lake's bacterial community. Although we cannot rule out the possibility that other relatively rare freshwater lineages were also present, it is likely that our recovery of many of the abundant populations is representative of those organisms that have some competitive advantage in the sampled system. Therefore, despite the relatively low sample coverage for each lake, our study was capable of identifying the central CtCh tendencies for which we would expect to find various freshwater bacterial lineages, with the caveat that we could not delineate each lineages' full natural range of CtCh.

In our survey, few distributional patterns could be observed at the phylum level (Fig. 5). This should not be surprising; if ecologists studying macro-scale eucaryotes had attempted to identify generalizations in the diets of all *Chordates*, they would have met with little success. The common freshwater bacterial phyla and classes (*Actinobacteria*, *Bacteroidetes* and *Betaproteobacteria*) were observed in most hierarchical clusters defined by OTU occurrence (Fig. 5). At more refined levels of phylogeny, we noted some suggestive trends. Many of the lineages we recovered have been observed in association with algal blooms or nutrient-enriched conditions (Eiler and Bertilsson, 2004; Simek *et al*., 2005), but few have been hypothesized to preferentially use terrestrial carbon. We were able to identify bacterial lineages that may have competitive advantages in lakes primarily supported by DOC of aquatic origin or lakes primarily supported by DOC of terrestrial origin (Figs 4 and 5 and discussed individually below). These findings set the stage for further research using culture-dependent and culture-independent methods to test the hypothesized ecophysiological specializations presented here.

Major freshwater groups

Previous efforts to determine the DOC preferences of the freshwater cosmopolitan lineages of *Actinobacteria* have produced conflicting results. One study suggested a growth response to algal-derived carbon (Pernthaler *et al*., 2001) while others report growth responses to enrichment of humic or terrestrial organic matter (Haukka *et al*., 2005). These contrasting results might be attributable to intra-phylum differences in substrate use preferences, since the studies cited above were conducted using more coarsely defined OTUs. Recent studies by Newton and colleagues (2007) and Warnecke and colleagues (2004) have developed a well-defined phylogeny for freshwater OTUs of the *Actinobacteria* phylum at a comparatively refined level (~98% 16S rRNA sequence identity), which we used here. In our data set, the majority

Table 3. Correlation between results for varied per cent sequence identity cut-offs used in furthest neighbour OTU definition in DOTUR.

	95%	96%	97%	98%	99%	99.5%
95%	1.00					
96%	0.91	1.00				
97%	0.82	0.96	1.00			
98%	0.79	0.94	0.99	1.00		
99%	0.76	0.90	0.95	0.96	1.00	
99.5%	0.71	0.88	0.94	0.96	0.98	1.00

The values correspond to Mantel R statistics between Bray–Curtis pairwise distance matrices constructed using occurrence patterns across all 15 lakes. Distance matrices were calculated using relative clone recovery rates.

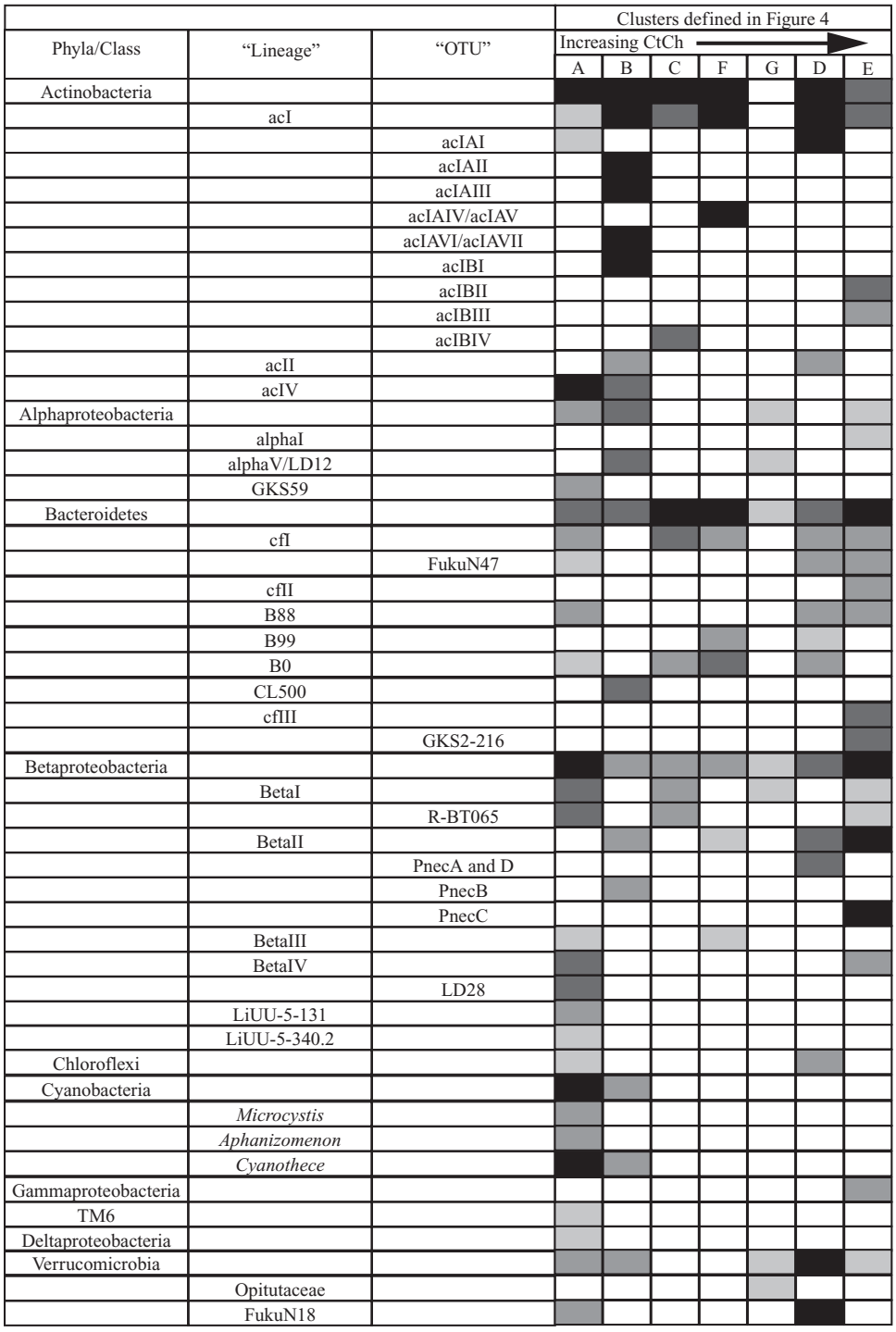

Fig. 5. Occurrence patterns of bacterial phyla and previously defined freshwater lineages. Columns of occurrence correspond to clusters in Fig. 4 (defined based on OTU occurrence) and are ordered from left to right by increasing CtCh ratio. The intensity of the shaded boxes indicate the number of clones affiliated with a given freshwater lineage that was present in that cluster. We defined a 'lineage' as a monophyletic group more refined than the class level, but broader than 97% 16S rRNA sequence identity. 'OTU' was defined as a group of 16S rRNA sequences sharing at least 97% sequence identity.

of *acIA* OTUs appeared to prefer lakes with low to intermediate CtCh while two OTUs within *acIB* predominantly occurred in lakes with high CtCh (Fig. 5). Thus, we hypothesize that OTUs within the *acIA* lineage tend to occur in more autochthonous lakes and OTUs within the *acIB* lineage prefer more allochthonous systems. Lineages *acII* and *acIV* tended to occur in lakes on the lower end of the CtCh range we sampled, suggesting that they

may be associated with autochthonous carbon sources, as well.

Relative to the *Actinobacteria*, freshwater *Bacteroidetes* are less studied and their phylogeny is poorly resolved. In general, OTUs within the lineages *cfII* and *cfIII* occurred in lakes with high CtCh. However, lineages previously observed by Eiler and Bertilsson (2007) during cyanobacterial blooms (*B0*, *B88*, *B99*) and CL500 tended to occur in lakes we categorized as 'intermediate' and 'low' with respect to CtCh. Agreement between our observations in multiple Wisconsin lakes and previous findings by Eiler and Bertilsson (2007) supports our hypotheses about carbon substrate preferences for these bacterial lineages.

Occurrence patterns of lineages and OTUs within freshwater *Betaproteobacteria* in our survey also agreed well with previous results. R-BT065 sequences were recovered from lakes with low-CtCh, which is supported by findings reported by Simek and colleagues (2005). *LiUU* clusters, initially described and observed during cyanobacterial blooms (Eiler and Bertilsson, 2004), were observed in lakes with low CtCh. Finally, OTUs related to *Polynucleobacter B* (*PnecB*) were recovered from low-CtCh lakes, but representative sequences from *PnecC*, *PnecA* and *PnecD* were found in high-CtCh lakes. These *Polynucleobacter* distribution patterns are similar to those observed by Hahn (2003) and Hahn and colleagues (2005). Operational taxonomic units associated with the *betaIII* and *betaIV* lineages were observed in both high- and low-CtCh lakes, suggesting that a more refined phylogenetic analysis of the lineage may be needed in order to resolve patterns in their distribution. Alternatively, the lineage could be populated by organisms that are generalists with respect to carbon substrate (Mou *et al*., 2008).

Less commonly recovered phyla

Members of the freshwater *Alphaproteobacteria*, *Cyanobacteria* and *Verrucomicrobia* were recovered at a lower frequency across our clone libraries. *Cyanobacteria* could be expected to co-vary with phytoplankton production, as both groups are photoautotrophic, and therefore occur in lakes with lower CtCh. *Cyanobacterial* OTUs were indeed found in low-CtCh lakes. Much less is known about freshwater distribution patterns of *Alphaproteobacteria* and *Verrucomicrobia*. In our current study, the former were observed across the CtCh gradient, suggesting that further taxonomic resolution of these phyla in freshwater systems is necessary. *Verrucomicrobia* were not recovered often, but were also seen in clusters across the CtCh gradient. The four phyla/classes with the lowest recovery rates were *Chloroflexi*, *TM6*, *Deltaproteobacteria* and *Gammaproteobacteria*. *TM6* and *Deltaproteobacteria* were observed towards the lower end of the CtCh gradient, and *Gammaproteobacteria* were only recovered from the high-CtCh lakes. Operational taxonomic units within *Chloroflexi* were observed at both extremes of the CtCh gradient. The low recovery rates of these groups may suggest that they are not actual residents of lake ecosystems and are only detected due to transport via the atmosphere or run-off (Lindstrom and Bergstrom, 2004; Jones and McMahon, 2009), that our molecular techniques are biased against recovery of sequences from these phylogenetic groups, or that members of *Chloroflexi*, *TM6*, *Deltaproteobacteria* and *Gammaproteobacteria* maintain low abundance in lakes.

Our findings were generally robust to changes in the cut-off level used to define OTUs, which we varied between 95% and 99.5% 16S rRNA gene sequence identity (Table 3). Some have proposed that two groups sharing less than 99% identity at this locus can be regarded as distinct species (Stackebrandt and Ebers, 2006). When we used groupings defined for freshwater lineages based on 97–99% identity and actual phylogenetic structure (R.J. Newton, S.E. Jones, A. Eiler, S. Bertilsson, and K.D. McMahon, unpublished), we generally observed discrete occurrences, often in a single CtCh cluster (data not shown). We propose that when considering carbon substrate use, a ~99% 16S rRNA gene sequence identity cut-off may be sufficient to define ecologically relevant taxonomic units. Our successful identification of coherent distribution patterns for bacterial phylogenetic groups defined at such a resolved level implies that a new 'working' taxonomy of freshwater lineages is needed. The most recent attempt to synthesize all available freshwater 16S rRNA gene sequences from freshwater was conducted in 2002 (Zwart *et al*., 2002) and clusters were defined at ~95% sequence identity during this meta-analysis. Our findings suggest this lower level of phylogenetic resolution may not be sufficient to perceive important microecological patterns.

Our data reveal surprisingly coherent patterns of OTU distribution among common freshwater lineages and bring us closer to being able to parameterize their niches. Further research, perhaps using carbon stable isotopes or microautoradiography, is needed to make direct links between substrate use and phylogenetic affiliation. Technology and methods currently under development, such as nano secondary ion mass spectrometry (Li *et al*., 2008) or stable isotope probing (Radajewski *et al*., 2000), are ideal tools to directly test the hypotheses presented in this article. The use of radioactive labelled organic carbon sources has also been used to make linkages between phylogenetic affiliations and substrate use (Cottrell and Kirchman, 2000), but this will only be feasible if suitably specific probes can be designed to distinguish the ecologically relevant OTUs via fluorescent *in situ* hybridization.

Experimental procedures

Field sampling

Fifteen lake epilimnia in Wisconsin, USA were selected along a CtCh gradient (Table 1). These lakes were sampled using an integrated water column sampler at the deepest point in the lake between 16 and 31 July 2002 as part of a 30-lake survey conducted by Yannarell and Triplett (2005). Samples were used to chemically characterize the lakes (Table 1) and for collection of bacterial biomass, as previously described. Briefly, a 250–500 ml subsample of water from each sample was vacuum filtered through a 0.2-mm-pore-size polyethersulfone membrane filter (Pall Supor-200) to capture bacterial biomass. Filters were placed into cryovials, frozen immediately in liquid nitrogen, and transported back to the laboratory where they were stored at –80°C. For analysis of chlorophyll *a*, absorbance at 665 and 750 nm was measured using a Kontron 9300 spectrophotometer before and after acidification of methanol-extracted chlorophyll *a*. Colour of 0.45 μm lake water filtrate was calculated as the 440 nm specific absorbance coefficient using a Kontron 9300 spectrophotometer and a 1 cm cuvette. Colour was calculated using the method outlined by Cuthbert and Del Giorgio (1992):

$$a_{440} = 2.303 \times (\text{absorbance at } 440\,\text{nm} \div 0.01\,\text{m}).$$

DNA extraction and clone library construction

Total DNA was extracted by Yannarell and Triplett (2005) from filters with a FastPrep Spin DNA purification kit (QBiogene) following the SPIN protocol. We amplified 16S rRNA genes and the corresponding 16S–23S internal transcribed spacer (ITS) regions from the resulting DNA pool using PCR (Eppendorf Mastercycler) with previously published conditions (Newton *et al*., 2007) and 8F (5′-AGAGTTTGATCMTGGCTCAG-3′; bacteria specific, 16S rRNA gene) and 23SR (5′-GGGTTBCCCCATTCRG-3′; bacteria specific, 23S rRNA gene) primers. We gel purified the PCR products using standard methods specified in the Novagen SpinPrep gel DNA purification protocol (EMD Biosciences).

We constructed one clone library for each of the 15 lakes using standard protocols of the Invitrogen Topo TA cloning system with TOP10 chemically competent cells. After cloning, 96 inserts from each library were amplified directly from cells using previously published PCR conditions (Newton *et al*., 2007) and M13 Forward (5′-GTAAAACGACGGCCAG-3′) and M13 Reverse (5′-CAGGAAACAGCTATGAC-3′) primers. The PCR products were purified using an Agencourt AMPure kit as recommended by the manufacturer's protocol.

Sequence analysis

Sequencing was conducted using an ABI Prism BigDye terminator sequencing kit (PE Applied Biosystems) and the bacteria-specific 8F primer under standard PCR sequencing conditions. We purified the sequencing PCR products with an Agencourt CleanSeq kit. All sequences were manually checked for quality using 4Peaks v. 1.7.2 and sequence chromatograms. Following quality control, 1094 sequences remained from the 15 libraries (58–88 sequences per lake). The sequences were then aligned using the ARB software package (Ludwig *et al*., 2004) containing a publicly available 16S rRNA gene ARB database (accessed January 2002; Hugenholtz, 2002) supplemented with freshwater bacterial 16S rRNA gene sequences (Zwart *et al*., 2002; Newton *et al*., 2006; 2007). Sequences were initially aligned using the FAST_ALIGNER ARB tool before the alignment was heuristically adjusted using primary and secondary rRNA structure as a guide. Sequences used in this study have been deposited to the GenBank database (Accession No. EU117556–EU117629, EU117640–EU117650, EU117689–EU117748, EU117760–EU117989 and FJ916087–FJ916806).

Phylogenetic and statistical analyses

The ARB alignment was used for maximum likelihood phylogenetic reconstruction using the Genetic Algorithm for Rapid Likelihood Inference (GARLI; Zwickl, 2006) and the Cyberinfrastructure for Phylogenetic Research website (CIPRES, http://www.phylo.org). The resulting phylogenetic tree was used as input to Unifrac (Lozupone and Knight, 2005). Unifrac allows coding of the tips of a phylogenetic reconstruction by environment or environmental characteristics. The user can then test whether the proportions of non-shared branch length across defined environments are less than what is expected by random chance (determined by permutation of environment labels). In addition, the fraction of shared branch length between pairs of environments can be used as input for further multivariate analyses, such as principal components analysis.

Although ecologically appropriate levels of phylogenetic resolution for bacteria are uncertain (Schloss and Handelsman, 2005), we defined three operational levels of taxonomy for clarity. For our broadest level of taxonomy we used classically defined bacterial phyla and *Proteobacteria* classes; 'lineages' were considered as monophyletic groups more refined than the phylum or class level, but broader than 97% 16S rRNA sequence identity. Finally, OTU was used to refer to groups of 16S rRNA sequences sharing at least 97% sequence identity. We used DOTUR (Schloss and Handelsman, 2005) to define OTUs at 97% sequence identity, here referred to as 'OTUs'. A pairwise distance matrix was calculated in ARB using the alignment of the 1094 sequences and a 50% base frequency filter to remove highly variable positions for input to DOTUR. Operational taxonomic units were assigned to bacterial phyla and proteobacterial classes using RDP (Wang *et al*., 2007) and more refined affiliation of OTUs was achieved using a 16S rRNA gene ARB database supplemented with freshwater bacterial 16S rRNA gene sequences (Zwart *et al*., 2002; Newton *et al*., 2006; 2007) and the ARB parsimony insertion tool (Ludwig *et al*., 2004). Relative clone recoveries of the OTUs in each lake were ordinated using CA and The R Statistics Package (R Development Core Team 2007) in order to visualize relative differences among lake bacterial communities. In addition, linear regression was used to test for statistically significant relationships between ordination axes and environmental characteristics of the lakes.

We also evaluated the sensitivity of our findings to the per cent sequence identity cut-off we used to define OTUs. We determined relative DOTUR OTU recoveries in each lake at six sequence per cent identity cut-offs (95, 96, 97, 98, 99 and

99.5). To compare these data sets, we calculated pairwise Bray–Curtis distance across lakes for each data set (created at each of the listed sequence per cent cut-offs). We then calculated pairwise Mantel R statistics to quantify agreement among the six Bray–Curtis distance matrices. A Mantel R approaching 1.0 suggests strong correlation or correspondence between the data sets created with two different sequence per cent cut-offs.

Finally, we used hierarchical clustering to evaluate relationships between occurrence patterns of OTUs across three CtCh categories (eight lakes < 0.08, three lakes between 0.09 and 0.14, and four lakes > 0.25; referred to as autochthonous, intermediate and allochthonous lakes for clarity). These CtCh categories were heuristically defined based upon natural breaks in the gradient observed in this study. Occurrence patterns and clustering of the OTUs along with clustering of the CtCh categories were visualized using the heatmap.2 function in The R Statistics Package (R Development Core Team 2007). Only OTUs populated by two or more individual sequences were considered for the hierarchical clustering analyses. Membership in OTU clusters determined by the hierarchical clustering analysis was used to identify putative ranges of occurrence along the CtCh gradient for freshwater OTUs. The OTU clusters were ranked from the lowest to the highest CtCh by summing the CtCh values of the lakes from which each clone was recovered and then dividing by the total number of clones in a given CtCh OTU cluster, i.e. an average CtCh for a given OTU cluster weighted by clone recovery.

Acknowledgements

The authors gratefully acknowledge technical help from Tracey Cadkin and staff of the University of Wisconsin-Madison Center for Limnology and Trout Lake Station. We thank Tony Yannarell for graciously allowing us to use sample collected during his PhD dissertation work. We thank Jay Lennon for providing thoughtful comments on earlier versions of the manuscript. This work was supported by funding from the University of Wisconsin-Madison Graduate School, and from the National Science Foundation to the North Temperate Lakes Long-term Ecological Research Site (DEB-0217533 and DEB-0822700) and Microbial Observatory (MCB-0702395).

References

Bergstrom, A.K., and Jansson, M. (2000) Bacterioplankton production in humic Lake Ortrasket in relation to input of bacterial cells and input of allochthonous organic carbon. *Microb Ecol* **39:** 101–115.

Biddanda, B., and Benner, R. (1997) Carbon, nitrogen, and carbohydrate fluxes during the production of particulate and dissolved organic matter by marine phytoplankton. *Limnol Oceanogr* **42:** 506–518.

Biddanda, B., Ogdahl, M., and Cotner, J. (2001) Dominance of bacterial metabolism in oligotrophic relative to eutrophic waters. *Limnol Oceanogr* **46:** 730–739.

Carpenter, S.R., Cole, J.J., Pace, M.L., Van de Bogert, M., Bade, D.L., Bastviken, D., *et al.* (2005) Ecosystem subsidies: terrestrial support of aquatic food webs from C-13 addition to contrasting lakes. *Ecology* **86:** 2737–2750.

Cole, J.J., Caraco, N.F., Kling, G.W., and Kratz, T.K. (1994) Carbon-dioxide supersaturation in the surface waters of lakes. *Science* **265:** 1568–1570.

Cole, J.J., Prairie, Y.T., Caraco, N.F., McDowell, W.H., Tranvik, L.J., Striegl, R.G., *et al.* (2007) Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems* **10:** 171–184.

Cottrell, M.T., and Kirchman, D.L. (2000) Natural assemblages of marine proteobacteria and members of the *Cytophaga–Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl Environ Microbiol* **66:** 1692–1697.

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.

Cuthbert, I.D., and Del Giorgio, P. (1992) Toward a standard method of measuring color in freshwater. *Limnol Oceanogr* **37:** 1319–1326.

Eiler, A., and Bertilsson, S. (2004) Composition of freshwater bacterial communities associated with cyanobacterial blooms in four Swedish lakes. *Environ Microbiol* **6:** 1228–1243.

Eiler, A., and Bertilsson, S. (2007) Flavobacteria blooms in four eutrophic lakes: linking population dynamics of freshwater bacterioplankton to resource availability. *Appl Environ Microbiol* **73:** 3511–3518.

Eiler, A., Langenheder, S., Bertilsson, S., and Tranvik, L.J. (2003) Heterotrophic bacterial growth efficiency and community structure at different natural organic carbon concentrations. *Appl Environ Microbiol* **69:** 3701–3709.

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.L. (2005) Low intraspecific diversity in a *Polynucleobacter* subcluster population numerically dominating bacterioplankton of a freshwater pond. *Appl Environ Microbiol* **71:** 4539–4547.

Hanson, P.C., Bade, D.L., Carpenter, S.R., and Kratz, T.K. (2003) Lake metabolism: relationships with dissolved organic carbon and phosphorus. *Limnol Oceanogr* **48:** 1112–1119.

Hanson, P.C., Carpenter, S.R., Armstrong, D.E., Stanley, E.H., and Kratz, T.K. (2006) Lake dissolved inorganic carbon and dissolved oxygen: changing drivers from days to decades. *Ecol Monogr* **76:** 343–363.

Haukka, K., Heikkinen, E., Kairesalo, T., Karjalainen, H., and Sivonen, K. (2005) Effect of humic material on the bacterioplankton community composition in boreal lakes and mesocosms. *Environ Microbiol* **7:** 620–630.

Hessen, D.O., and Tranvik, L.J. (1998) *Aquatic Humic Substances*. Berlin, Germany: Springer.

Hobbie, J.E. (1988) A comparison of the ecology of planktonic bacteria in fresh and salt-water. *Limnol Oceanogr* **33:** 750–764.

Hugenholtz, P. (2002) Exploring prokaryotic diversity in the genomic era. *Genome Biol* **3:** 1–8.

Jones, S.E., and McMahon, K.D. (2009) Species sorting may explain the apparent minimal effect of immigration on freshwater bacterial community dynamics. *Environ Microbiol* **11:** 905–913.

Judd, K.E., Crump, B.C., and Kling, G.W. (2006) Variation in dissolved organic matter controls bacterial production and community composition. *Ecology* **87:** 2068–2079.

Kritzberg, E.S., Cole, J.J., Pace, M.L., Graneli, W., and Bade, D.L. (2004) Autochthonous versus allochthonous carbon sources of bacteria: results from whole-lake C-13 addition experiments. *Limnol Oceanogr* **49:** 588–596.

Kritzberg, E.S., Langenheder, S., and Lindstrom, E.S. (2006) Influence of dissolved organic matter source on lake bacterioplankton structure and function – implications for seasonal dynamics of community composition. *FEMS Microbiol Ecol* **56:** 406–417.

Li, T., Wu, T.D., Mazeas, L., Toffin, L., Guerquin-Kern, J.L., Leblon, G., and Bouchez, T. (2008) Simultaneous analysis of microbial identity and function using NanoSIMS. *Environ Microbiol* **10:** 580–588.

Lindstrom, E.S. (2000) Bacterioplankton community composition in five lakes differing in trophic status and humic content. *Microb Ecol* **40:** 104–113.

Lindstrom, E.S., and Bergstrom, A.K. (2004) Influence of inlet bacteria on bacterioplankton assemblage composition in lakes of different hydraulic retention time. *Limnol Oceanogr* **49:** 125–136.

Lindstrom, E.S., Kamst-Van Agterveld, M.P., and Zwart, G. (2005) Distribution of typical freshwater bacterial groups is associated with pH, temperature, and lake water retention time. *Appl Environ Microbiol* **71:** 8201–8206.

Lozupone, C., and Knight, R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71:** 8228–8235.

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.

Mou, X., Sun, S., Edwards, R.A., Hodson, R.E., and Moran, M.A. (2008) Bacterial carbon processing by generalist species in the coastal ocean. *Nature* **451:** 708–711.

Nelson, C.E. (2009) Phenology of high-elevation pelagic bacteria: the roles of meteorologic variability, catchment inputs and thermal stratification in structuring communities. *ISME J* **3:** 13–30.

Newton, R.J., Kent, A.D., Triplett, E.W., and McMahon, K.D. (2006) Microbial community dynamics in a humic lake: differential persistence of common freshwater phylotypes. *Environ Microbiol* **8:** 956–970.

Newton, R.J., Jones, S.E., Helmus, M.R., and McMahon, K.D. (2007) Phylogenetic ecology of the freshwater *Actinobacteria* acI lineage. *Appl Environ Microbiol* **73:** 7169–7176.

Pace, M.L., Carpenter, S.R., Cole, J.J., Coloso, J.J., Kitchell, J.F., Hodgson, J.R., and Middelburg, J.J. (2007) Does terrestrial organic carbon subsidize the planktonic food web in a clear-water lake? *Limnol Oceanogr* **52:** 2177–2189.

Perez, M.T., and Sommaruga, R. (2006) Differential effect of algal- and soil-derived dissolved organic matter on alpine lake bacterial community composition and activity. *Limnol Oceanogr* **51:** 2527–2537.

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.

R Development Core Team (2007) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Radajewski, S., Ineson, P., Parekh, N.R., and Murrell, J.C. (2000) Stable-isotope probing as a tool in microbial ecology. *Nature* **403:** 646–649.

Riemann, L., Steward, G.F., and Azam, F. (2000) Dynamics of bacterial community composition and activity during a mesocosm diatom bloom. *Appl Environ Microbiol* **66:** 578–587.

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.M., 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.

Solomon, C.T., Carpenter, S.R., Cole, J.J., and Pace, M.L. (2008) Support of benthic invertebrates by detrital resources and current autochthonous primary production: results from a whole-lake C-13 addition. *Freshwater Biol* **53:** 42–54.

Stackebrandt, E., and Ebers, J. (2006) Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today*. **November:** 152–155.

Stackebrandt, E., and Goebel, B.M. (1994) Taxonomic note: a place for DNA:DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* **44:** 846–849.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73:** 5261–5267.

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.

Weidel, B., Carpenter, S., Cole, J., Hodgson, J., Kitchell, J., Pace, M., and Solomon, C. (2008) Carbon sources supporting fish growth in a north temperate lake. *Aquat Sci* **70:** 446–458.

Wetzel, R.G. (2001) *Limnology: Lake and River Ecosystems*. San Diego, CA, USA: Academic Press.

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.

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.

Zwickl, D.J. (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD Dissertation. The University of Texas at Austin.