The ISME Journal (2010) 4, 729–738
© 2010 International Society for Microbial Ecology All rights reserved 1751-7362/10 $32.00
www.nature.com/ismej

ORIGINAL ARTICLE

Species sorting affects bacterioplankton community composition as determined by 16S rDNA and 16S rRNA fingerprints

Jürg B Logue and Eva S Lindström
Department of Ecology and Evolution/Limnology, Uppsala University, Uppsala, Sweden

To understand the mechanisms determining community composition, it is essential to distinctively unravel the importance of local from that of regional processes. In this effort, the mechanisms underlying bacterioplankton community assembly were analysed in eight lakes of short water residence time (WRT) during a four-season sampling campaign. Bacterioplankton community composition (BCC) was determined using terminal-restriction fragment length polymorphism (t-RFLP) on the 16S rRNA gene (16S rDNA) and 16S rRNA. The relationship between similarity in BCC between a lake and its major inlet on the one hand and cell import per cell production rate from the inlet to the lake epilimnion on the other was used as a measure of the importance of cell dispersal (mass effects) for community assembly. Low similarities in BCC between lakes and their inlets were observed even at short WRTs, and the degree of similarity correlated better with the environmental conditions in lakes and streams than with cell import per cell production rates. Thus, mass effects seemed less important for local lake BCC in comparison to environmental habitat characteristics (species sorting). Analyses of 16S rDNA and 16S rRNA community fingerprints yielded similar results, indicating that species-sorting dynamics exerted an equally important effect on both the abundant and active fraction within the studied bacterioplankton communities.
The ISME Journal (2010) **4,** 729–738; doi:10.1038/ismej.2009.156; published online 4 February 2010
Subject Category: microbial population and community ecology
Keywords: 16S rDNA; 16S rRNA; biogeography; mass effects; metacommunity; species sorting

Introduction

Central to community ecology is the comprehension of the mechanisms underlying community assembly; that is, to grasp the processes affecting the distribution, abundance, demography and interactions of populations. Set off by the discourse between Clements (1916) and Gleason (1926) on plant succession, a debate evolved into a dispute on whether communities assemble because of deterministic (niche, competition) or stochastic (neutral, chance) processes (Diamond, 1975; Connor and Simberloff, 1979; Hubbell, 2001). More recently, community ecology moved from focusing on a single-scale to placing emphasis on a multi-scale perspective, acknowledging dynamics and interactions at both local and regional scales (Ricklefs, 1987; Wilson, 1992; Leibold *et al.*, 2004; Harrison and Cornell, 2008).

The metacommunity framework (Leibold *et al.*, 2004; Holyoak *et al.*, 2005) represents one of the latest additions to the suite of community ecology theories and frameworks and aims at shedding light on the question of to what extent a local community is shaped by local and/or regional processes. The framework comprises four models. Each model describes a set of local communities that are linked by dispersal and highlights different aspects of spatial community dynamics. Moreover, they differ pertaining to the importance of environmental heterogeneity, dispersal, stochasticity and colonization-competition trade-offs in structuring local communities. (Leibold *et al.*, 2004). This study centres on only two models: species sorting and mass effects. The species-sorting model, which has much in common with traditional niche theory, assumes environmental heterogeneity to produce community differences. Dispersal is assumed to occur frequently enough to compensate for species loss but not so frequent as to cause homogenization. Higher levels of dispersal, however, cause mass effects that affect local community composition, thereby, at least partly, overriding the effects of local processes.

The mechanisms by which microbial communities assemble are poorly understood, yet it has become apparent in recent years that microbial

Correspondence: JB Logue, Department of Ecology and Evolution/Limnology, Uppsala University, Norbyvägen 18D, Uppsala SE-752 36, Sweden.
E-mail: jb.logue@ebc.uu.se
Received 13 July 2009; revised 17 December 2009; accepted 21 December 2009; published online 4 February 2010

assemblages do differ in composition and diversity in different locations (Martiny *et al.*, 2006). At present, there is a lively debate on the relative importance of local and regional forces exerting influence on microbial community assembly (see Logue and Lindström, 2008 for an overview) and on whether or not microbial diversity shows explainable patterns of distribution comparable to such of macroorganisms. But for all that, studies analysing the importance of regional spatial effects in relation to local processes for microbial community assembly are, still, rare (Logue and Lindström, 2008). In any case, Lindström *et al.* (2006), Nelson *et al.* (2009) and Crump *et al.* (2007) found signatures of mass effects in freshwater systems of short water residence time (WRT). Other findings, however, indicate that bacterioplankton community composition (BCC) in lakes is primarily driven by species-sorting processes (Van der Gucht *et al.*, 2007; Jones and McMahon, 2009). Langenheder and Ragnarsson (2007) then again identified both species-sorting and regional processes at work in affecting bacterioplankton communities in rock pools. Thus, it becomes apparent that observations are not consistent and little conceptual understanding does exist in explaining patters of microbial community assembly (Logue and Lindström, 2008).

Studies of microbial biogeography rely on molecular fingerprinting approaches based on, for instance, the 16S ribosomal RNA (rRNA) gene (henceforward termed 16S rDNA) to follow changes in community composition. Given the accruing evidence that not all cells in bacterial assemblages are metabolically active (see del Giorgio and Gasol, 2008 for review), it can be questioned whether this approach yields relevant information from an ecosystem point of view. In contrast, 16S rRNA-based approaches do provide clues on cellular activity (see, for example, Schaechter *et al.*, 1958; Bremer and Dennis, 1987) and may thus provide more relevant information from an ecosystem perspective. It has been shown that community composition can differ depending on whether 16S rDNA or 16S rRNA is targeted (see, for example, Teske *et al.*, 1996; Schäfer *et al.*, 2001; Troussellier *et al.*, 2002). Moreover, compositional shifts of bacterial assemblages can be detected more readily and more reproducibly with 16S rRNA rather than with 16S rDNA (Mahmood and Prosser, 2006; Hoshino and Matsumoto, 2007). Further, if 16S rRNA analyses do indeed better reflect the active community component, variations in such fingerprints should correspond to contemporary changes in BCC, whereas 16S rDNA fingerprints may also be the results of historical processes.

The primary objective of this study was to analyse the relative importance of mass effects as opposed to species-sorting processes in structuring bacterioplankton assemblages. We hypothesize that (1) bacterioplankton communities in lakes of short WRT are assembled by mass effects due to high cell import rates, (2) the importance of mass effects depends on season as a consequence of changes in water flow and thus changes in cell import rates and (3) the importance of mass effects depends on whether the total community (16S rDNA) or its active fraction (16S rRNA) is targeted, with the active fraction potentially being more prone to species-sorting dynamics. To test for the abovementioned hypotheses, bacterioplankton assemblages of eight lakes and their main inlets were sampled on four occasions differing in water flow regime using terminal-restriction fragment length polymorphism (t-RFLP).

Materials and methods

Study sites and sampling

Eight lakes and their respective main inlets were included in the sampling scheme. The selected lakes are all situated in Uppland, which is a region in the southeastern part of central Sweden, and are of rather similar physico-chemical nature (Table 1). Neither lakes nor inlets are headwater systems. Sampling took place on four occasions throughout 2007: 1–2 March (winter), 11–12 April (spring), 3–4 July (summer), and 7–8 November (autumn). One depth-integrated epilimnion (lake) and one water column (inlet) sample was taken on each site and date. Epilimnion samples were collected at the centre of the lake, approximately 100 m away from the inlet. Water column samples were taken at the mid cross-section and at an intermediate depth of the inlet stream. Depth-integrated epilimnion sampling was conducted as follows: in winter and summer, when lakes were stratified, water was collected from every $\frac{1}{2}$ or 1 m over the epilimnion water column, whereas in spring and autumn, when lakes were fully mixed, water was sampled every 1 or 2 m over the entire water column. Water samples were mixed in a bucket and—on site—divided into three subsamples: one for water chemistry, the second for bacterial production and the third for bacterial abundance and community measurements. Subsamples were kept cold and in the dark for 2 to 8h until processing in the laboratory.

Respective WRT estimates were taken from Brunberg and Blomqvist (1998) and are based on lake volumes, drainage areas as well as area-specific runoff data (yearly average). They are theoretical in nature, estimating the average time for a water molecule to travel from lake inlet to outlet; that is, lakes with short WRT receive greater amounts of water per lake volume unit when compared with lakes with a long WRT. Water flow was determined by multiplying the size of the drainage area of the inlet stream with its specific runoff at the day of sampling. Drainage areas for both lakes and inlet streams were determined using topographical maps. Specific runoff data from the day of sampling were obtained from hydrological stations located near the lakes' drainage areas (Swedish Meteorological and Hydrological Institute).

Table 1 Physico-chemical and biological characteristics of the eight dimictic study lakes for the four sampling occasions

Site	Abbreviation	Location North	Location East	Lake area (km²)	Drainage area (km²)	Lake volume (m³ × 10⁶)	Max. depth (m)	WRT (days)	Date	Water flow (m³ s⁻¹)	pH	Chl-a (μg l⁻¹)	Tot-P (μg l⁻¹)	Tot-N (mg l⁻¹)	TOC (mg l⁻¹)	Abs_{436} (5 cm⁻¹)	$Abs_{250/365}$	Bray–Curtis similarities in BCC between lake and inlet 16S rDNA	Bray–Curtis similarities in BCC between lake and inlet 16S rRNA	Cell import per cell production rate
Fibysjön[M]	FS	59°54′	17°22′	0.47	19.20	0.59	1.85	49	March	0.15	6.52	0.00	23.72	1.16	26.46	0.37	4.65	54.78	46.64	55.29
									April	0.20	6.67	2.11	24.10	1.13	25.27	0.39	4.70	28.31	42.03	1.30
									July	0.03	7.11	7.73	24.10	1.00	22.21	0.30	4.99	23.99	29.05	0.56
									November	0.07	7.22	3.12	18.30	0.92	19.89	0.23	5.17	23.95	32.51	4.99
Lumpen[M]	Lum	59°58′	17°17′	0.25	3.93	0.31	1.90	164	March	0.03	6.47	0.62	14.81	1.15	29.23	0.43	4.52	73.88	65.99	26.96
									April	0.04	6.73	4.06	19.85	1.40	28.77	0.42	4.73	60.41	67.19	0.67
									July	0.01	7.30	18.98	28.36	1.07	25.84	0.33	5.09	21.28	25.49	0.17
									November	0.01	7.24	11.25	14.04	0.99	21.43	0.20	5.73	27.88	26.15	0.39
Norrsjön A[E]	NSA	59°52′	18°00′	0.29	31.60	0.30	3.20	16	March	0.12	6.66	0.62	44.24	2.28	25.32	0.31	5.53	42.98	39.47	105.31
									April	0.14	7.33	10.55	41.92	2.18	22.40	0.28	5.01	38.86	41.64	6.13
									July	0.01	7.85	54.83	58.57	1.63	24.13	0.15	6.17	25.41	17.56	0.21
									November	0.01	7.60	19.33	23.72	1.07	19.71	0.10	7.25	38.59	30.31	0.64
Norrsjön N[M]	NSN	59°49′	17°55′	0.17	6.86	0.34	3.00	40	March	0.03	6.75	1.56	33.79	1.40	21.51	0.25	5.38	75.50	70.21	11.90
									April	0.03	7.13	6.56	31.85	1.26	18.40	0.21	5.21	73.48	58.15	1.17
									July	0.00	7.53	5.27	18.68	0.78	16.46	0.09	6.64	29.45	28.33	0.09
									November	0.00	7.57	16.17	19.85	0.84	16.42	0.06	8.16	27.49	22.65	0.20
Ramsjön[M]	RS	59°50′	17°13′	0.17	6.82	0.41	4.20	117	March	0.05	6.37	0.31	24.49	1.57	31.42	0.48	4.62	70.58	61.05	31.15
									April	0.07	6.98	2.46	31.07	1.36	24.69	0.40	4.59	75.54	66.22	0.94
									July	0.01	7.36	115.29	41.53	1.19	22.42	0.30	4.91	21.41	31.29	0.58
									November	0.03	7.01	7.19	33.40	1.22	19.06	0.24	5.08	30.66	26.72	2.49
Siggeforasjön[O]	Sig	59°58′	17°09′	0.76	20.30	3.20	11.00	180	March	0.16	6.34	0.00	9.39	0.80	19.85	0.26	4.76	71.21	72.28	7.83
									April	0.21	6.35	0.70	12.88	0.81	19.15	0.26	4.69	69.36	63.45	0.74
									July	0.03	6.86	6.68	10.16	0.68	16.71	0.18	5.14	49.58	49.52	0.63
									November	0.07	6.74	1.87	7.84	0.57	14.59	0.14	5.48	55.55	58.69	1.56
Stensjön[M]	SS	60°03′	17°49′	0.08	6.61	0.08	2.20	20	March	0.02	6.54	0.62	19.85	1.73	29.50	0.37	6.01	35.75	36.19	27.29
									April	0.03	7.05	13.71	25.65	1.50	25.56	0.33	5.03	28.02	32.79	1.53
									July	0.00	7.00	10.55	22.94	0.93	19.18	0.21	5.33	29.39	38.42	0.18
									November	0.00	7.39	7.73	12.10	0.65	14.63	0.12	6.04	29.22	33.81	0.97
Velången[M]	Vel	60°07′	17°28′	1.69	32.90	1.90	2.50	102	March	0.25	6.47	0.00	19.46	2.69	47.03	0.59	4.95	77.69	83.05	11.73
									April	0.34	6.84	4.69	21.39	2.05	39.00	0.51	4.90	71.09	74.52	1.42
									July	0.05	6.90	9.84	25.65	1.38	35.49	0.40	5.33	59.64	53.99	0.52
									November	0.15	7.52	5.27	12.49	1.28	29.88	0.23	6.28	68.61	45.18	1.87

Abbreviations: BCC, bacterioplankton community composition; Chl-a, chlorophyll-a; E, eutrophic; M, mesotrophic; O, oligotrophic; TOC, total organic carbon; Tot-N, total nitrogen; Tot-P, total phosphorus; WRT, water residence time.
Water temperature was averaged and bacterial production calculated for the epilimnion in March and July and for the entire lake in April and November.

Chemical analysis, bacterial abundance and productivity

Chemical analysis. Water chemistry samples were divided anew into subsamples in the laboratory. The first, immediately frozen subsample was analysed for total phosphorus (Tot-P), total nitrogen (Tot-N) and total organic carbon (TOC). TOC was determined on a Shimadzu TOC-5000 total organic analyzer (Columbia, MD, USA) after acidification and purging for evasion of inorganic carbon. Tot-N was analysed by means of oxidative hydrolysis with persulphate/sodiumhydroxide using a standard technique. Tot-P was measured as molybdate-reactive phosphate (Murphy and Riley, 1962) after hydrolysis with potassium peroxodisulphate (Menzel and Corwin, 1965). To determine the absorbance, a second subsample of 100 ml was filtered through a MG C micro-glass fibre filter (1.2 μm pore size; Munktell, Falun, Sweden). The absorbance of the filtrate was then recorded at the three wavelengths of 250 (A_{250}), 365 (A_{365}) and 436 (A_{436}) nm with a PU 8625 UV/VIS spectrophotometer (Philips, DA Best, The Netherlands). For the chlorophyll-a (Chl-a) analysis, a volume of 100 ml was collected onto a pre-weighed GF/C filter (Whatman, Florham Park, NJ, USA), which was subsequently immersed into the ethanol extraction solution (95% $v\,v^{-1}$). Upon boiling and centrifuging, the extract underwent absorption measurements at 665 nm and with correction for suspended solids at 770 nm. The data were not corrected for phaeopigments to include pigments from both living and dead cells (Jespersen and Christoffersen, 1987). Finally, pH measurements were performed with a standard electrode (PW9420; Phillips).

Bacterial abundance. To analyse bacterial abundance, a 20 ml subsample was preserved in formaldehyde (final concentration of 4% $w\,v^{-1}$). Bacterial abundance was enumerated by epifluorescence microscopy of acridine-orange-stained cells (Hobbie *et al.*, 1977). Numbers are based on counts of at least 400 cells in a minimum of 20 microscopic fields.

Bacterial production. Tritiated thymidine incorporation into DNA was used to estimate bacterial production according to Bell (1993). Tritiated thymidine (50 nmol l^{-1}) was added to triplicates of 10 ml lake water immediately upon arrival to the laboratory. The samples were subsequently incubated in darkness at *in situ* temperatures for 60–120min. Growth rates (cells $l^{-1}\,h^{-1}$) were calculated assuming a division rate of 2×10^{18} cells per mol of incorporated thymidine (Bell, 1990).

Bacterioplankton cell budgets were constructed using data on water flow, lake volume, bacterial abundance and bacterial production. Thus, cell import was calculated as inlet water flow × bacterial abundance, whereas in-lake bacterial cell production was computed as bacterial production × lake volume (epilimnion volume during times of stratification, whole-lake volume during times of mixis).

Nucleic acid extraction, reverse transcription and t-RFLP

Bacterial cells were collected onto 0.2 μm membrane filters (Supor Membrane Disc Filters, 47 mm; Pall Corporation, East Hills, NY, USA) filtering 100 ml of pre-sieved (225 μm mesh size) lake and stream water. Pre-sieving was conducted to avoid capture of larger particles, such as zooplankton, and did not significantly alter bacterioplankton abundances (data not shown). Filters were placed into sterile cryogenic vials (Nalgene, Rochester, NY, USA) and stored immediately at −80 °C.

Nucleic acid extraction. Nucleic acid extraction (co-extracting DNA and RNA) was conducted following the protocol # 3 of the Easy-DNA kit (Invitrogen, Carlsbad, CA, USA) with an extra 0.2 g of 0.1 mm silica beads. As an exception, the RNase step was omitted to allow for the co-extraction of RNA. Extracted nucleic acids were sized and quantified by means of agarose (1%) gel electrophoresis, ethidium bromide staining and ultraviolet transillumination. Nucleic acid extracts were divided into two subsamples, one for the analysis of RNA and one for the analysis of DNA.

Reverse transcription. DNA was removed from DNA–RNA extracts by treatment with DNase I (Invitrogen) for 15 min at room temperature in accordance with the manufacturer's instructions. First-strand synthesis of complementary DNA from RNA was conducted using SuperScript II reverse transcriptase (Invitrogen) and random hexamers (Invitrogen). In brief, DNA-free RNA samples were incubated at 65 °C for 5 min with random hexamers (250 ng; Invitrogen) and deoxynucleotide triphosphates (10 mM each; Invitrogen) to ensure melting of RNA secondary structure. Templates were subsequently chilled on ice for 15 min. 5X first-strand buffer, 0.1 M dithiothreitol and, as an exception to the manufacturer's recommendation, T4 gene 32 protein (final concentration of 3 μg μl^{-1}; New England Biolabs, Ipswich, UK) were thereon added and the template was incubated at 25 °C for 2min. T4 gene 32 protein has been shown to increase reverse transcription efficiency (Villalva *et al.*, 2001). Finally, 400 units of SuperScript II reverse transcriptase were added and the template was incubated at 25 °C for 10min. Reverse transcription was performed at 42 °C for 50min followed by an enzyme inactivation step at 70 °C for 15 min. Reverse-transcription negative PCR control reactions (without reverse transcriptase) were run to check for DNA contamination.

t-RFLP. BCC of both 16S rDNA and 16S rRNA subsamples was assessed using t-RFLP (Liu *et al.*, 1997). 16S rDNA and 16S rRNA were PCR amplified from lake and inlet samples applying primers 8F (bacteria specific; 5′-AGRGTTTGATCMTGGCTCAG-3′) (Vergin *et al.*, 1998) and 519R (universal; 5′-GWAT

TACCGCGGCKGCTG-3′) (Lane *et al.*, 1985). Primer 8F was fluorescently labelled with hexachloro-fluorescein (HEX; Sigma-Aldrich, St Louis, MO, USA) at the 5′ end to allow for fluorescent detection of terminal-restriction fragments. PCR reactions were performed in a 20 μl reaction volume comprising 0.4 U Phusion high-fidelity DNA polymerase (Finnzymes, Espoo, Finland), 5X Phusion HF reaction buffer (Finnzymes), 200 μM of each deoxynucleotide triphosphate (Invitrogen), 200 nM of each primer (Sigma-Aldrich) and 5–10 ng of template DNA. Thermocycling (DNA Engine (PTC-200) Peltier Thermal Cycler; Bio-Rad Laboratories, Hercules, CA, USA) was carried out with an initial denaturation step at 98 °C for 3min, followed by 25 cycles of denaturation at 98 °C for 10 s, annealing at 55 °C for 30 s and primer extension at 72 °C for 30 s, and finalized with a 7-min extension step at 72 °C. Four technical replicates were run per sample, pooled after PCR amplification and purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). Yield and success were checked through agarose (1%) gel electrophoresis in comparison with a Low DNA Mass Ladder (Invitrogen). Fluorescently labelled PCR products were then subjected to parallel digestions with the two restriction enzymes *Hae*III and *Hinf*I (New England Biolabs) for 18h at 37 °C. Each digestion reaction contained approximately 38 ng of DNA, 4 U of each respective restriction enzyme, 1X restriction buffer and sterile water to a final reaction volume of 10 μl. Fluorescently labelled terminal-restriction fragments were separated and detected using an ABI 3730 capillary sequencer running in GeneScan mode (Applied Biosystems, Foster City, CA, USA) together with an internal size standard (GeneScan-500 ROX; Applied Biosystems). t-RFLP peaks were finally determined using GeneMarker v1.70 (SoftGenetics, State College, PA, USA). GeneMarker analysis included fragments with lengths ranging from 50 to 460 bp and an initial peak detection threshold of 90 relative fluorescent units. Peaks of <0.5 bases apart from each other were merged.

Once the original t-RFLP data were obtained, an additional threshold was applied. In doing so, only peaks accounting for $>0.5\%$ of the total peak area of a sample were retained; a procedure applied to correct for unequal quantities in PCR products despite the effort of normalization. Digestion duplicates were merged to a new data set of average values for the relative peak area (relative in relation to the total peak area of the sample) to normalize for differences in capillary electrophoresis load. Digestion duplicates were highly similar to each other with Bray–Curtis similarities ranging from 77% to 98%, indicating the method's high reproducibility. In addition, group means between self ($\text{sample1}_{\text{duplicate1}}$–$\text{sample1}_{\text{duplicate2}}$, and so on) and non-self ($\text{sample1}_{\text{duplicate1}}$-$\text{sample2}_{\text{duplicate1}}$, $\text{sample1}_{\text{duplicate2}}$–$\text{sample2}_{\text{duplicate1}}$, and so on) were highly distinct from each other (Wilcoxon's signed-rank test with continuity correction, $V=8256$, $P<2.2e^{-16}$). At last, the data sets gained from the two parallel restrictions were combined.

Statistical analyses
To test whether mass effects compared with species-sorting processes were of greater importance for structuring bacterioplankton communities at high cell import rates, Spearman's rank correlation analyses were performed, correlating Bray–Curtis similarities in BCC between lakes and their inlets (lake:inlet BCC similarities) to calculated cell import per cell production rates (import from inlet to lake). Because, however, high cell import rates are brought about by high water flow rates that, in turn, can lead to greater similarities in environmental parameters between lakes and their inlets, an alternative scenario was also tested for. This alternative scenario can be specified as follows: high lake:inlet BCC similarities at high water flow rates could be caused by similar lake and inlet environments selecting for similar taxa in lakes and their inlets. Therefore, Spearman's rank correlation analyses were carried out correlating lake:inlet BCC similarities with heterogeneities in environmental variables between lakes and their inlets. Moreover, partial Spearman's rank correlation analyses were conducted to assess the influence of cell import per cell production rate on lake:inlet BCC similarity independent of environment and vice versa. The heterogeneity in environmental variables between lakes and their inlets was calculated as follows: at first, a principal component (PC) analysis was carried out explaining 66% (PC1 = 38% (composed of Chl-a, pH, A_{250}/A_{365} and TOC) and PC2 = 28% (composed of Tot-P and Tot-N)) of the variation in the environmental data set that included the six parameters of pH, Chl-a, Tot-P, Tot-N, TOC and A_{250}/A_{365} for all seasons. Parameters were for this purpose log-transformed (except for pH and A_{250}/A_{365}, with the latter being arcsin square-root transformed) and *z*-score transformed (Legendre and Legendre, 1998). Thereafter, sample scores along PC1 and 2 were extracted, which were then used to calculate absolute Δ lake:inlet values (ABS ($\text{lake}_{\text{samplescore}}$–$\text{inlet}_{\text{samplescore}}$)).

To analyse the potential importance of local environmental (species-sorting) processes for explaining variation in BCC, redundancy analysis (RDA) was performed. This linear, direct gradient analysis method measures the amount of variation in community structure attributable to a set of explanatory variables. Testing for metacommunity models can be carried out by variance partitioning of RDA results (Cottenie, 2005), decomposing the variation in community composition into unique environmental (suggesting species sorting) and spatial components (suggesting, for instance, mass effects). However, analyses of data separated by season were of too low statistical power. Also,

analysing all data together in relation to space was not adequate because of temporal variation in BCC. Therefore, only environmental data went into the RDA. BCC data were square-root transformed, whereas environmental data were treated as mentioned earlier. The significance of the environmental component was evaluated with a Monte Carlo permutation test (999 permutations under the reduced model). Unbiased estimates of the explained variation component were computed according to Peres-Neto *et al.* (2006).

To analyse whether similar trends in community composition were obtained using the two fingerprinting approaches, correlations between 16S rDNA and 16S rRNA similarity matrices were analysed deploying Mantel tests. Similarity matrices (64×64; lake (8) and inlet (8) samples over all four seasons) (Bray–Curtis index of similarity for relative peak area and Sørensen's index of similarity for presence–absence data) were constructed from 16S rDNA and 16S rRNA fingerprints, respectively. The significance of the Mantel tests was assessed through 10 000 permutations. Testing for a significant difference in 16S rDNA and 16S rRNA fingerprints was performed by analysis of similarity (ANOSIM) on a Bray–Curtis distance measure running 10 000 permutations.

All analyses were conducted using R (R Foundation for Statistical Computing: Vienna, Austria; R, 2008) except for ANOSIM, which was performed using PAST (Oslo, Norway) (Hammer *et al.*, 2001).

Results

Table 1 summarizes the physico-chemical properties of the eight dimictic studied lakes. The lakes are all primarily of mesotrophic nature except for two lakes, Norrsjön Almunge and Siggeforasjön, which are eutrophic and oligotrophic, respectively. The extreme Chl-a value measured in Lake Ramsjön (115.29 $\mu g\,l^{-1}$) in July may be due to a bloom of the freshwater Raphidophyte *Gonyostomum*; a phenomenon observed on previous occasions in this lake (Eva S. Lindström, unpublished data). All lakes were ice covered in March and had developed a distinct thermocline in winter (March) as well as in summer (July).

t-RFLP analysis of 16S rDNA and 16S rRNA BCC of the eight studied lakes and their respective inlets on the four sampling occasions yielded a total of 427 unique operational taxonomic units. In all, 380 and 337 operational taxonomic units were recorded throughout the four sampling times for 16S rDNA and 16S rRNA, respectively.

Lake:inlet BCC similarities on the four dates of sampling ranged from 21% to 78% and from 18% to 83% for 16S rDNA and 16S rRNA, respectively. No obvious relationship between WRT and lake:inlet BCC similarities could be observed (Figure 1). Instead, at the shortest WRTs (that is, in lakes with

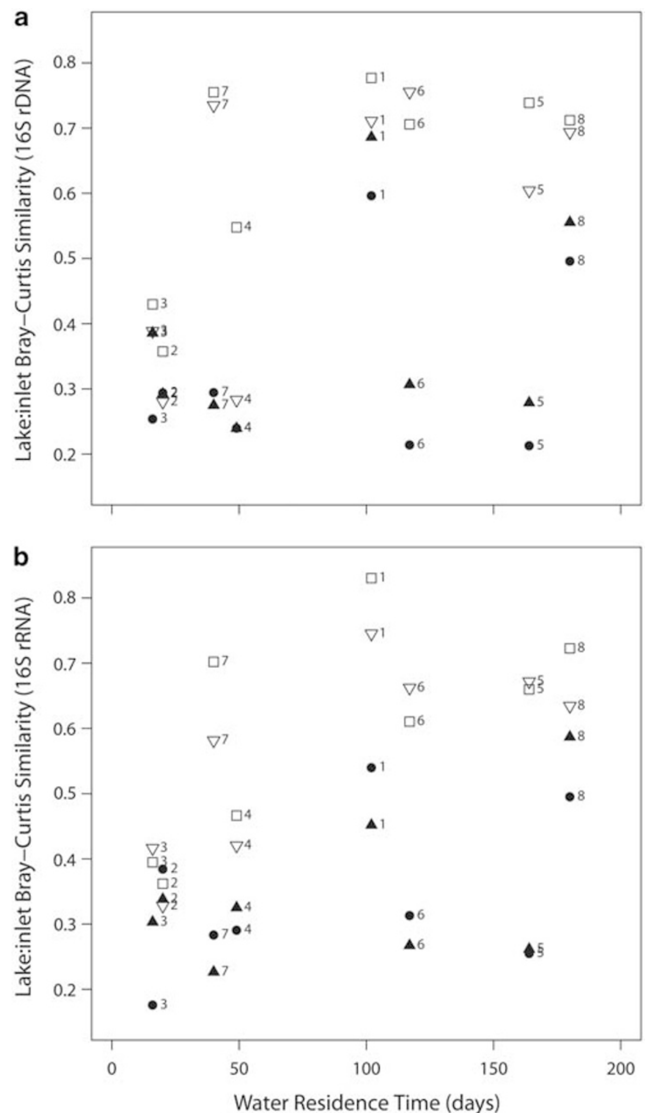

Figure 1 Relationship between lake:inlet Bray-Curtis similarities in 16S rDNA (**a**) and 16S rRNA (**b**) t-RFLP fingerprints and water residence time (WRT) for the eight studied lakes over all four seasons (winter (□), spring (▽), summer (●) and autumn (▲)) (see Table 1 for actual values of both 16S rDNA and 16S rRNA). The eight studied lakes are labelled as follows: 1, Velången (Vel); 2, Stensjön (SS); 3, Norrsjön Almunge (NSA); 4, Fibysjön (FS); 5, Lumpen (Lum); 6, Ramsjön (RS); 7, Norrsjön Norreda (NSN); and 8, Siggeforasjön (Sig).

highest water flow rates on average) lake:inlet BCC similarities were consistently low, but at longer WRTs both very high and also low lake:inlet BCC similarities were found (Table 1, Figure 1). Considering all data from all seasons, lake:inlet BCC similarities for both 16S rDNA and 16S rRNA were significantly positively correlated with cell import per cell production rates (Table 2); that is, higher lake:inlet BCC similarities were recorded at higher cell import per cell production rates. However, lake:inlet BCC similarities were also significantly negatively correlated with local environmental heterogeneity between lakes and their inlets

Table 2 Results from Spearman's rank correlation analyses

	BC 16S rDNA	*BC 16S rRNA*	*Δ PC1*	*Δ PC2*
IM	$P = \mathbf{0.0033}$	$P = \mathbf{0.0030}$	$P = \mathbf{0.0088}$	$P = \mathbf{0.0294}$
	$R = 0.5092$	$R = 0.5128$	$R = -0.4593$	$R = -0.3871$
Δ PC1	$P = \mathbf{0.0004}$	$P = \mathbf{0.0002}$		
	$R = -0.5960$	$R = -0.6254$		
Δ PC2	$P = \mathbf{0.0026}$	$P = \mathbf{0.0043}$		
	$R = -0.5194$	$R = -0.4963$		

Abbreviations: BC, lake:inlet Bray-Curtis similarity in BCC; BCC, bacterioplankton community composition; IM, cell import per cell production rate; Δ PC1, Δ lake:inlet sample scores from PC1; Δ PC2, Δ lake:inlet sample scores from PC2.
Significant P-values are in bold.

Table 3 Results from partial Spearman's rank correlation analyses testing for mass effects and species-sorting processes

	Variable X	*Constant Z*	R_{XY*Z}	P-*value*
BC 16S rDNA	IM	Δ PC1	0.3300	0.0698
	IM	Δ PC2	0.3760	**0.0371**
	Δ PC1	IM	−0.4737	**0.0071**
	Δ PC2	IM	−0.5032	**0.0039**
BC 16S rRNA	IM	Δ PC1	0.3255	0.0740
	IM	Δ PC2	0.4117	**0.0255**
	Δ PC1	IM	−0.5112	**0.0033**
	Δ PC2	IM	−0.3762	**0.0370**

Abbreviations: BC, lake:inlet Bray–Curtis similarity in BCC; BCC, bacterioplankton community composition; IM, cell import per cell production rate; Δ PC1, Δ lake:inlet sample scores from PC1; Δ PC2, Δ lake:inlet sample scores from PC2.
Significant P-values are in bold.

expressed as Δ lake:inlet PC1 and PC2 sample score values (Table 2); that is, similar environments coincided with similar communities. What is more, cell import per cell production rates were significantly negatively correlated with both Δ PC1 and Δ PC2. Consequently, high lake:inlet BCC similarities at high cell import per cell production rates could either stem from mass effects or from the selection of similar bacterioplankton phylotypes in lakes and their inlets as a result of similar environments (species sorting). Therefore, to separate the effects of the two mechanisms from each other, partial Spearman's rank correlation analyses were performed. Results hereof show that for both types of fingerprints the degree of association between lake:inlet BCC similarity and cell import per cell production rate became insignificant once the effect of environmental heterogeneity (Δ PC1) was controlled for (Table 3). Considering Δ PC2, the association remained significant. However, to support the assumption that mass effects affected BCC, the correlation with cell import would have needed to be entirely independent of the environment; that is, significant for both Δ PCs. Yet on the contrary, the relation between lake:inlet BCC similarity and environmental heterogeneity, for both Δ PC1 and Δ PC2, remained significant once the effect of cell import per cell production rate was removed (Table 3).

Analysing the seasons separately showed only one case of a significant correlation between lake:inlet BCC similarities and cell import per cell production rates (16S rRNA in March, results not shown). This correlation, however, was negative; that is, high lake:inlet BCC similarities were found at low cell import per cell production rates. Thus, analysing seasons separately did not provide any support of mass effects being an important mechanism in structuring lake BCC.

Moreover, RDA analysing the potential importance of species-sorting processes in explaining variation in BCC produced a significant co-variation with local environmental conditions explaining 10.1% and 10.4% of the variation in 16S rDNA and 16S rRNA community structures, respectively.

Mantel tests disclosed similar trends over time and space for 16S rDNA and 16S rRNA fingerprints, independent of whether relative peak area ($R_{MT} = 0.822$, $P < 0.0001$, $n = 64$) or presence–absence ($R_{MT} = 0.781$, $P < 0.0001$, $n = 64$) data measures were used. Also, 68% of all 427 unique operational taxonomic units were detected in both 16S rDNA and 16S rRNA fingerprints, pointing out the great similarities between the two approaches in t-RFLP community patterns.

The high percentage of shared operational taxonomic units between 16S rDNA and 16S rRNA necessitated testing for possible DNA contamination of RNA samples. Reverse-transcription negative PCR control reactions, however, showed no signs of contamination. In addition, Mantel tests on digestion duplicates disclosed higher similarities within 16S rDNA ($R_{MT} = 0.984$) and within 16S rRNA ($R_{MT} = 0.962$) than between 16S rDNA and 16S rRNA. Moreover, ANOSIM showed that there was a statistically significant difference between 16S rDNA and 16S rRNA fingerprints ($R = 0.132$, $P < 0.0001$). Thus, the method's intrinsic variability was smaller than the difference between the two approaches; that is, greatly similar but not identical results were obtained.

Discussion

The aim of this study was to analyse the relative importance of regional versus local processes structuring lake BCC; that is, whether mass effects as opposed to species-sorting processes can be of greater importance in structuring lake bacterioplankton assemblages. We hypothesized that (1) BCC in lakes of short WRT experience mass effects; however, this depends on (2) season and (3) on whether 16S rDNA or 16S rRNA is targeted.

The set of lakes that was selected ranged in WRTs previously proposed by Lindström *et al.* (2006) to be short enough for the communities to experience mass effects. Should the first and second hypothesis

be true, an event of mass effects would bring about high lake:inlet BCC similarities at high numbers of imported bacterioplankton cells, which in turn are a result of short lake WRT and high seasonal water flow. Thus, high lake:inlet BCC similarities should be obtained at short WRTs and lake:inlet BCC similarities should be correlated with rates of cell import per cell production. Considering the third hypothesis, mass effects were expected to be more pronounced when targeting 16S rDNA compared with 16S rRNA. In the latter case, a weaker relationship between cell import and lake:inlet BCC similarities is expected.

Despite sampling lakes of short WRT, both low and high lake:inlet BCC similarities were observed over the entire range of WRT, except for the shortest WRTs when lake:inlet BCC similarities were consistently low (Figure 1). Moreover, variation in lake:inlet BCC similarities was not well explained by variation in cell import per cell production rates. Instead, lake:inlet BCC similarities co-varied better with local environmental heterogeneities between lakes and their inlets (Tables 2 and 3), suggesting that mass effects were of rather minor importance when compared with species-sorting dynamics in explaining bacterioplankton community structure in the lakes studied. This observation was further supported by the significant contribution of environmental variables explaining variation in community structure among the studied lakes in the RDA models. Yet, it cannot be excluded that individual bacterioplankton populations were subject to mass effects. It further cannot be excluded that mass effects could occur in lakes of even shorter WRTs than were included in this study as well as at extremely high water flow events such as during snow melt or heavy rain storms; events that were not sampled in this study. Several previous studies did observe signatures of mass effects exerting influence on a lake's BCC due to high cell import rates (see, for example, Crump *et al.*, 2004; Lindström *et al.*, 2006; Crump *et al.*, 2007; Nelson *et al.*, 2009). Yet, the question has been raised as to at what cell import rates the point differentiating between mass effects and species sorting should come to lie (Logue and Lindström, 2008). Focusing on lakes, Lindström *et al.* (2006) suggested a discrimination point at a WRT of 100 days, whereas results from Nelson *et al.* (2009) suggested a discrimination point at 68 days. As we did not find any signatures of mass effects on BCC in our studied lakes of short WRT (16–180 days), our data do not support conclusions drawn by these researchers. Thus, it remains to be shown whether this point of discrimination in fact could be at much shorter WRTs, such as <2 or <3 days, as has been suggested by Crump *et al.* (2004, 2007) for estuaries and streams, respectively.

Chances are, though, that this point of discrimination between mass effects and species sorting also differs among system types. Van der Gucht *et al.* (2007) proposed that more oligotrophic lake systems should be rather prone to mass effects dynamics when compared with more eutrophic lakes. Unfortunately, we were not able to test this hypothesis because of the fact that our lakes were all predominantly mesotrophic. However, it is interesting to note that the most oligotrophic lake (Lake Siggeforasjön) did show consistently high lake:inlet BCC similarities despite a relatively long WRT of 180 days. Moreover, as Nelson *et al.* (2009) did not observe mass effects structuring BCC in a headwater lake of extremely short WRT (<5 days) either, it is likely that there is no such thing as a universal point of discrimination between mass effects and species sorting. However, it seems that there are not yet enough data available to consistently determine how high rates of cell import to lakes in nature ought to be to trigger mass effects.

In this context, it might be worth reflecting upon how to detect mass effects in nature. Thus, a concern within this study might be that during periods of stratification the inlet water did not enter the lake's epilimnion but instead the hypolimnion, because of density differences following temperature differences. As a result, lower lake:inlet BCC similarities might have been detected and mass effects thus remained unnoticed. Such a situation could have potentially occurred in July when inlet temperatures were colder than lake temperatures and hence inlet water may not have entered the sampled epilimnion. However, this situation cannot be the sole explanation for not detecting mass effects in general in our studied lakes. Another possible reason for not detecting mass effects might be the fact that having sampled only the main inlet to the lake might have resulted in an underestimation of total cells imported through all streams to the lake. Thus, it is important to consider all possible agents of mass effects; that is, all dispersal routes to a lake. Hence, routes of dispersal other than by streams need to be analysed before conclusions can be drawn about the existence and importance of mass effects influencing microbial community assembly. Among one of the very few studies, Jones and McMahon (2009) analysed the effect of atmospherically dispersed cells on lake BCC. They concluded that the number of bacterial cells that were deposited in lakes through the atmosphere was too low to trigger mass effects. Yet, data on microbial dispersal is virtually inexistent, however crucially needed.

The observed similar results for 16S rDNA and 16S rRNA fingerprints is an interesting finding, considering the emerging pattern that a non-negligible proportion of cells within a microbial community is assumed to be either dormant, damaged or dead (del Giorgio and Gasol, 2008) and that growth rates of bacterial cells are related to cellular RNA content (see, for example, Schaechter *et al.*, 1958; Bremer and Dennis, 1987). Thus, 16S rDNA and 16S rRNA fingerprints are expected to yield different

results, which have been shown in studies conducted in marine and or coastal aquatic ecosystems (Schäfer *et al.*, 2001; Troussellier *et al.*, 2002; Rodriguez-Blanco *et al.*, 2009). Whether the observed similarity in 16S rDNA and 16S rRNA from our study might be a pattern typical for freshwater bacterioplankton communities cannot be said because of a lack of studies in such systems. However, marine environments differ greatly from freshwater ecosystems and are, for example, generally poorer in nutrient content when compared with lakes. As it has been shown that the number of substrate respiring bacterial cells increases with system productivity (del Giorgio and Scarborough, 1995), bacterial production (del Giorgio *et al.*, 1997) and cell growth (Choi *et al.*, 1996; Sherr *et al.*, 1999), it is quite possible that the observed similarity between 16S rDNA and 16S rRNA follows from that the most abundant and dominant bacterioplankton phylotypes detected by t-RFLP in our eight mesotrophic lakes were also active.

Conclusions

Our results show that bacterioplankton communities can be primarily assembled by means of local environmental species-sorting processes rather than by mass effects even in lakes of short WRT. On the basis of the results obtained from this study as well as others (Cottenie, 2005; Yannarell and Triplett, 2005; Van der Gucht *et al.*, 2007; Jones and McMahon, 2009), we conclude that species-sorting processes seem to be important in structuring bacterioplankton communities in lake systems of a wide range of types and water flow. It remains to be observed at what point of discrimination mass effects rather than species-sorting processes are of greater importance. Finally, species-sorting dynamics seemed to be equally important for both the active and total bacterioplankton community; a rather novel finding, suggesting that the most abundant phylotypes detected in the lakes studied may also be the most active ones.

Acknowledgements

We thank Sara Beier and Silke Langenheder for initial help with regard to RNA handling and Xinmei Feng for constructive discussions on t-RFLP analysis. We are thankful to Janne Johansson for assistance in the field and physico-chemical water and bacterial abundance analyses. We are indebted to Silke Langenheder, Lars Tranvik and three anonymous reviewers for constructive comments on the manuscript. The Swedish Meteorological and Hydrological Institute (SMHI) and the Department of Hydrology, Uppsala University, generously provided us with water flow data. This work was financed by grants from the Swedish Research Council (ESL).

References

Bell RT. (1990). An explanation for the variability in the conversion factor deriving bacterial-cell production from incorporation of [h-3] thymidine. *Limnol Oceanogr* **35**: 910–915.

Bell RT. (1993). Estimating production of heterotrophic bacterioplankton via incorporation of tritiated thymidine. In: Kemp PF, Sherr BF, Sherr EB, Cole JJ (eds). *Handbook of Methods in Aquatic Microbial Ecology*. Lewis Publishers: London, pp 495–503.

Bremer H, Dennis PP. (1987). Modulation of chemical composition and other parameters of the cell by growth rate. In: Neidhardt FC, Ingraham JL, Low B, Magasanik B, Schaechter M, Umbarger HE (eds). *Escherichia coli and Salmonella Typhimurium: Cellular and Molecular Biology*. American Society of Microbiology: Washington DC, pp 1527–1542.

Brunberg A-K, Blomqvist P. (1998). *Vatten i Uppsala län 1997*. Upplandsstiftelsen: Uppsala, Sweden.

Choi JW, Sherr EB, Sherr BF. (1996). Relation between presence absence of a visible nucleoid and metabolic activity in bacterioplankton cells. *Limnol Oceanogr* **41**: 1161–1168.

Clements FE. (1916). *Plant Succession. An Analysis of the Development of Vegetation* vol. 242, Carnegie Institute Washington Publication.

Connor EF, Simberloff D. (1979). The assembly of species communities—chance or competition. *Ecology* **60**: 1132–1140.

Cottenie K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Ecol Lett* **8**: 1175–1182.

Crump BC, Hopkinson CS, Sogin ML, Hobbie JE. (2004). Microbial biogeography along an estuarine salinity gradient: combined influences of bacterial growth and residence time. *Appl Environ Microbiol* **70**: 1494–1505.

Crump BC, Adams HE, Hobbie JE, Kling GW. (2007). Biogeography of bacterioplankton in lakes and streams of an arctic tundra catchment. *Ecology* **88**: 1365–1378.

del Giorgio PA, Gasol JM. (2008). Physiological structure and single-cell activity in marine bacterioplankton. In: Kirchman DL (ed). *Microbial Ecology of the Oceans* 2nd edn. Wiley: Hoboken, New Jersey, USA, pp 243–298.

del Giorgio PA, Scarborough G. (1995). Increase in the proportion of metabolically active bacteria along gradients of enrichment in fresh-water and marine plankton—implications for estimates of bacterial-growth and production-rates. *J Plankton Res* **17**: 1905–1924.

del Giorgio PA, Prairie YT, Bird DF. (1997). Coupling between rates of bacterial production and the abundance of metabolically active bacteria in lakes, enumerated using CTC reduction and flow cytometry. *Microb Ecol* **34**: 144–154.

Diamond JM. (1975). Assembly of species communities. In: Cody ML, Diamond JM (eds). *Ecology and Evolution of Communities*. Harvard University Press: Cambridge, MA, pp 342–444.

Gleason HA. (1926). The individualistic concept of the plant association. *Bull Torrey Bot Club* **53**: 7–26.

Hammer Ø, Harper DAT, Ryan PD. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontol Electr* **4**: 9–11.

Harrison S, Cornell H. (2008). Toward a better understanding of the regional causes of local community richness. *Ecol Lett* **11**: 969–979.
Hobbie JE, Daley RJ, Jasper S. (1977). Use of nuclepore filters for counting bacteria by fluorescence microscopy. *Appl Environ Microbiol* **33**: 1225–1228.
Holyoak M, Leibold MA, Holt RD. (2005). *Metacommunities: Spatial Dynamics and Ecological Communities*. The University of Chicago Press: Chicago and London.
Hoshino YT, Matsumoto N. (2007). DNA- versus RNA-based denaturing gradient gel electrophoresis profiles of a bacterial community during replenishment after soil fumigation. *Soil Biol Biochem* **39**: 434.
Hubbell SP. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press: Princeton, NY.
Jespersen AM, Christoffersen K. (1987). Measurements of chlorophyll-a from phytoplankton using ethanol as extraction solvent. *Arch Hydrobiol* **109**: 445–454.
Jones SE, McMahon KD. (2009). Species-sorting may explain an apparent minimal effect of immigration on freshwater bacterial community dynamics. *Environ Microbiol* **11**: 905–913.
Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR. (1985). Rapid-determination of 16s ribosomal-RNA sequences for phylogenetic analyses. *Proc Natl Acad Sci USA* **82**: 6955–6959.
Langenheder S, Ragnarsson H. (2007). The role of environmental and spatial factors for the composition of aquatic bacterial communities. *Ecology* **88**: 2154–2161.
Legendre P, Legendre L. (1998). *Numerical Ecology*. Elsevier: Amsterdam, The Netherlands.
Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF *et al.* (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* **7**: 601–613.
Lindström ES, Forslund M, Algesten G, Bergström AK. (2006). External control of bacterial community structure in lakes. *Limnol Oceanogr* **51**: 339–342.
Liu WT, Marsh TL, Cheng H, Forney LJ. (1997). Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl Environ Microbiol* **63**: 4516–4522.
Logue JB, Lindström ES. (2008). Biogeography of bacterioplankton in inland waters. *Freshwater Rev* **1**: 99–114.
Mahmood S, Prosser JI. (2006). The influence of synthetic sheep urine on ammonia oxidizing bacterial communities in grassland soil. *Fems Microbiol Ecol* **56**: 444–454.
Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL *et al.* (2006). Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4**: 102–112.
Menzel DH, Corwin N. (1965). The measurement of total phosphorus in seawater based on the liberation of organically bound fractions by persulfate oxidation. *Limnol Oceanogr* **10**: 280–282.
Murphy J, Riley JP. (1962). A modified single solution method for the determination of phosphate in natural waters. *Anal Chim Acta* **27**: 31–26.
Nelson CE, Sadro S, Melack JM. (2009). Contrasting the influences of stream inputs and landscape position on bacterioplankton community structure and dissolved organic matter composition in high-elevation lake chains. *Limnol Oceanogr* **54**: 1292–1305.
Peres-Neto PR, Legendre P, Dray S, Borcard D. (2006). Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* **87**: 2614–2625.
R (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria.
Ricklefs RE. (1987). Community diversity—relative roles of local and regional processes. *Science* **235**: 167–171.
Rodriguez-Blanco A, Ghiglione JF, Catala P, Casamayor EO, Lebaron P. (2009). Spatial comparison of total vs active bacterial populations by coupling genetic fingerprinting and clone library analyses in the NW Mediterranean Sea. *Fems Microbiol Ecol* **67**: 30–42.
Schaechter M, Maaloe O, Kjeldgaard NO. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced growth of Salmonella typhimurium. *J Gen Microbiol* **19**: 592–606.
Schäfer H, Bernard L, Courties C, Lebaron P, Servais P, Pukall R *et al.* (2001). Microbial community dynamics in Mediterranean nutrient-enriched seawater mesocosms: changes in the genetic diversity of bacterial populations. *Fems Microbiol Ecol* **34**: 243–253.
Sherr EB, Sherr BF, Sigmon CT. (1999). Activity of marine bacteria under incubated and in situ conditions. *Aquat Microb Ecol* **20**: 213–223.
Teske A, Wawer C, Muyzer G, Ramsing NB. (1996). Distribution of sulfate-reducing bacteria in a stratified fjord (Mariager fjord, Denmark) as evaluated by most-probable-number counts and denaturing gradient gel electrophoresis of PCR-amplified ribosomal DNA fragments. *Appl Environ Microbiol* **62**: 1405–1415.
Troussellier M, Schafer H, Batailler N, Bernard L, Courties C, Lebaron P *et al.* (2002). Bacterial activity and genetic richness along an estuarine gradient (Rhone River plume, France). *Aquat Microb Ecol* **28**: 13–24.
Van der Gucht K, Cottenie K, Muylaert K, Vloemans N, Cousin S, Declerck S *et al.* (2007). The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. *Proc Natl Acad Sci USA* **104**: 20404–20409.
Vergin KL, Urbach E, Stein JL, DeLong EF, Lanoil BD, Giovannoni SJ. (1998). Screening of a fosmid library of marine environmental genomic DNA fragments reveals four clones related to members of the order Planctomycetales. *Appl Environ Microbiol* **64**: 3075–3078.
Villalva C, Touriol C, Seurat P, Trempat P, Delsol G, Brousset P. (2001). Increased yield of PCR products by addition of T4 gene 32 protein to the SMART (TM) PCR cDNA synthesis system. *Biotechniques* **31**: 81–86.
Wilson DS. (1992). Complex interactions in metacommunities, with implications for biodiversity and higher levels of selection. *Ecology* **73**: 1984–2000.
Yannarell AC, Triplett EW. (2005). Geographic and environmental sources of variation in lake bacterial community composition. *Appl Environ Microbiol* **71**: 227–239.