environmental microbiology

society for applied microbiology sfam

Environmental Microbiology (2015) doi:10.1111/1462-2920.12692

Environmental rather than spatial factors structure bacterioplankton communities in shallow lakes along a > 6000 km latitudinal gradient in South America

Caroline Souffreau,[1]* Katleen Van der Gucht,[2] Ineke van Gremberghe,[2] Sarian Kosten,[3,4] Gissell Lacerot,[5] Lúcia Meirelles Lobão,[6] Vera Lúcia de Moraes Huszar,[7] Fabio Roland,[6] Erik Jeppesen,[8,9] Wim Vyverman[2] and Luc De Meester[1]

[1]*Laboratory of Aquatic Ecology, Evolution and Conservation, University of Leuven, Leuven, Belgium.*
[2]*Laboratory of Protistology and Aquatic Ecology, Ghent University, Gent, Belgium.*
[3]*Department of Aquatic Ecology and Environmental Biology, Institute for Water and Wetland Research, Radboud University Nijmegen, Nijmegen, The Netherlands.*
[4]*Aquatic Ecology and Water Quality Management Group, Wageningen University, Wageningen, The Netherlands.*
[5]*Functional Ecology of Aquatic Systems, CURE, Universidad de la República, Rocha, Uruguay.*
[6]*Laboratory of Aquatic Ecology, Universidade Federal de Juiz de Fora, Juiz de Fora, Brazil.*
[7]*Museu Nacional, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.*
[8]*Department of Bioscience and the Arctic Centre, Aarhus University, Silkeborg, Denmark.*
[9]*Sino-Danish Centre for Education and Research, Beijing, China.*

Summary

Metacommunity studies on lake bacterioplankton indicate the importance of environmental factors in structuring communities. Yet most of these studies cover relatively small spatial scales. We assessed the relative importance of environmental and spatial factors in shaping bacterioplankton communities across a > 6000 km latitudinal range, studying 48 shallow lowland lakes in the tropical, tropical*i* (isothermal subzone of the tropics) and tundra climate regions of South America using denaturing gradient gel electrophoresis. Bacterioplankton community composition (BCC) differed significantly across regions. Although a large fraction of the variation in BCC remained unexplained, the results supported a consistent significant contribution of local environmental variables and to a lesser extent spatial variables, irrespective of spatial scale. Upon correction for space, mainly biotic environmental factors significantly explained the variation in BCC. The abundance of pelagic cladocerans remained particularly significant, suggesting grazer effects on bacterioplankton communities in the studied lakes. These results confirm that bacterioplankton communities are predominantly structured by environmental factors, even over a large-scale latitudinal gradient (6026 km), and stress the importance of including biotic variables in studies that aim to understand patterns in BCC.

Received 30 June, 2014; revised 24 October, 2014; accepted 24 October, 2014. *For correspondence. E-mail caroline.souffreau@bio.kuleuven.be; Tel. +32 1637 3550; Fax +32 1632 4575.

Introduction

An important goal in ecology is to understand the distribution of organisms and the processes underlying these distributions. Gaining knowledge about distribution patterns of bacteria is of particular importance because bacteria may well comprise the majority of the earth's biodiversity and perform processes that are critical to sustain life on earth. In the past decade, an increasing amount of research has focused on whether microbial communities share patterns of distribution and diversity similar to those of macroscopic organisms and, more specifically, to what extent they show a biogeographical signal (e.g. Horner-Devine *et al*., 2004; Beisner *et al*., 2006; Green and Bohannan, 2006; Martiny *et al*., 2006; Ramette and Tiedje, 2007b; Van der Gucht *et al*., 2007; Hanson *et al*., 2012; Soininen, 2012). The traditional hypothesis among microbiologists, 'Everything is everywhere, but the environment selects' (Baas-Becking, 1934), presumes ubiquity based on the high dispersal rates and high local population densities of microorganisms. This hypothesis has been challenged in a number of studies where it is suggested that at least some microbial taxa can exhibit geographical isolation and specific distribution patterns (Cho and Tiedje, 2000; Papke *et al*., 2003; Whitaker *et al*., 2003; Ramette and Tiedje, 2007a).

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd

However, several surveys of bacterial community structure focusing on similar environments across large geographical scales have reported environmental gradients to be more important than geographical distance (a proxy for dispersion of microorganisms) in shaping community structure in microbial communities (e.g. Fierer and Jackson, 2006; Van der Gucht *et al*., 2007; King *et al*., 2010; Redford *et al*., 2010; De Bie *et al*., 2012; Wang *et al*., 2013).

The relative importance of local environmental and regional processes in determining community composition is a key theme in microbial community ecology (Hanson *et al*., 2012; Lindström and Langenheder, 2012). Metacommunity theory provides a framework to understand and investigate the ecological processes underlying the observed patterns in community composition over space and time (Leibold, 1998; Chase and Leibold, 2002; Leibold *et al*., 2004). Metacommunities are sets of local communities linked by dispersal of potentially interacting species (Leibold *et al*., 2004; Holyoak *et al*., 2005). Both local and regional ecological processes can shape the local community structure within metacommunities (Leibold *et al*., 2004). Regional processes emphasize dispersal dynamics and include dispersal limitation, neutral processes and mass effects, while local processes imply the selection of species by the local abiotic and biotic conditions, referred to as species sorting (Leibold *et al*., 2004). Under the species sorting paradigm, the presence of a species in a habitat is not limited by dispersal, but only by the local conditions while dispersal is not so high that it influences species composition by mass effects (Leibold, 1998; Chase and Leibold, 2002; Leibold *et al*., 2004).

The relative importance of local and regional processes on metacommunity structure can be assessed by statistically relating community composition, local environmental conditions, and spatial and historical information. Because spatial distance often covers gradients in environmental conditions, it is important to take into account the effect of spatially structured environmental variables when disentangling the relative importance of space and environment in explaining variation in metacommunity structure. Variation partitioning (Borcard *et al*., 1992) has often been used to provide information about the fraction of variation in a community dataset explained by purely spatial signals, purely environmental signals and the combined effect of space and environment (due to spatially structured environmental variables). Using this method, field studies have shown that freshwater bacterioplankton metacommunities are primarily structured by local abiotic and biotic conditions, including pH, temperature and conductivity, concentrations of dissolved organic carbon (DOC), nitrogen and phosphorus compounds, lake depth and lake area, and abundances of zooplankton and heterotrophic nanoflagellates (e.g. Beisner *et al*., 2006; Langenheder and Ragnarsson, 2007; Van der Gucht *et al*., 2007; Wu *et al*., 2007; Logue and Lindström, 2010; Schiaffino *et al*., 2011; De Bie *et al*., 2012; Langenheder *et al*., 2012). However, most of the freshwater bacterioplankton metacommunity studies accounting explicitly for the relative importance of space, environment and spatially structured environmental factors using variation partitioning have been performed over relatively small spatial scales, except for the study of Van der Gucht and colleagues (2007) covering a > 2500 km north-south gradient in Europe and the study of Schiaffino and colleagues (2011) across a > 2100 km transect from Argentinean Patagonia to Maritime Antarctica. Both studies used denaturing gradient gel electrophoresis (DGGE) to characterize the bacterioplankton community structure, and while Van der Gucht and colleagues (2007) observed only a minor impact of spatial distance, Schiaffino and colleagues (2011) detected that both spatial and environmental factors control bacterioplankton community composition over latitude. Studying metacommunity processes at larger spatial scales is rarely done but is key to improve our insight in the scale dependency of ecological processes underlying observed metacommunity patterns.

In the present study, we assessed the bacterial community composition in 48 shallow lakes along a latitudinal gradient that ranged from the tropics to the near-Antarctic in South America (5–55°S; *c.* 6200 km) (Fig. S1; Table S1) to test for spatial and environmental correlates of bacterial community composition along a semi-continental gradient. The lakes were similar in morphology and altitude, varied as much as possible in trophic state within each climate zone and were sampled once during the dry season (tropical*i* and tropical lakes) or in summer (tundra lakes) between August 2005 and February 2006. We used a 16S rRNA gene-based fingerprinting technique, DGGE, to determine bacterial community structure. Fingerprints are banding patterns where each band is translated to an operational taxonomic unit (OTU). Our specific goals were (i) to analyse to which degree spatial structure within and among regions influences bacterial community composition (BCC); and (ii) to identify the environmental factors that explain variation in BCC.

Results

Geographical and environmental heterogeneity within and among climatic regions

Overall, the spatial scales covered within each of the three climatic regions (tropical*i*, tropical and tundra) were relatively comparable (Fig. 1A). Tropical*i* is an isothermal subzone in the tropics, which has a smaller annual

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd,

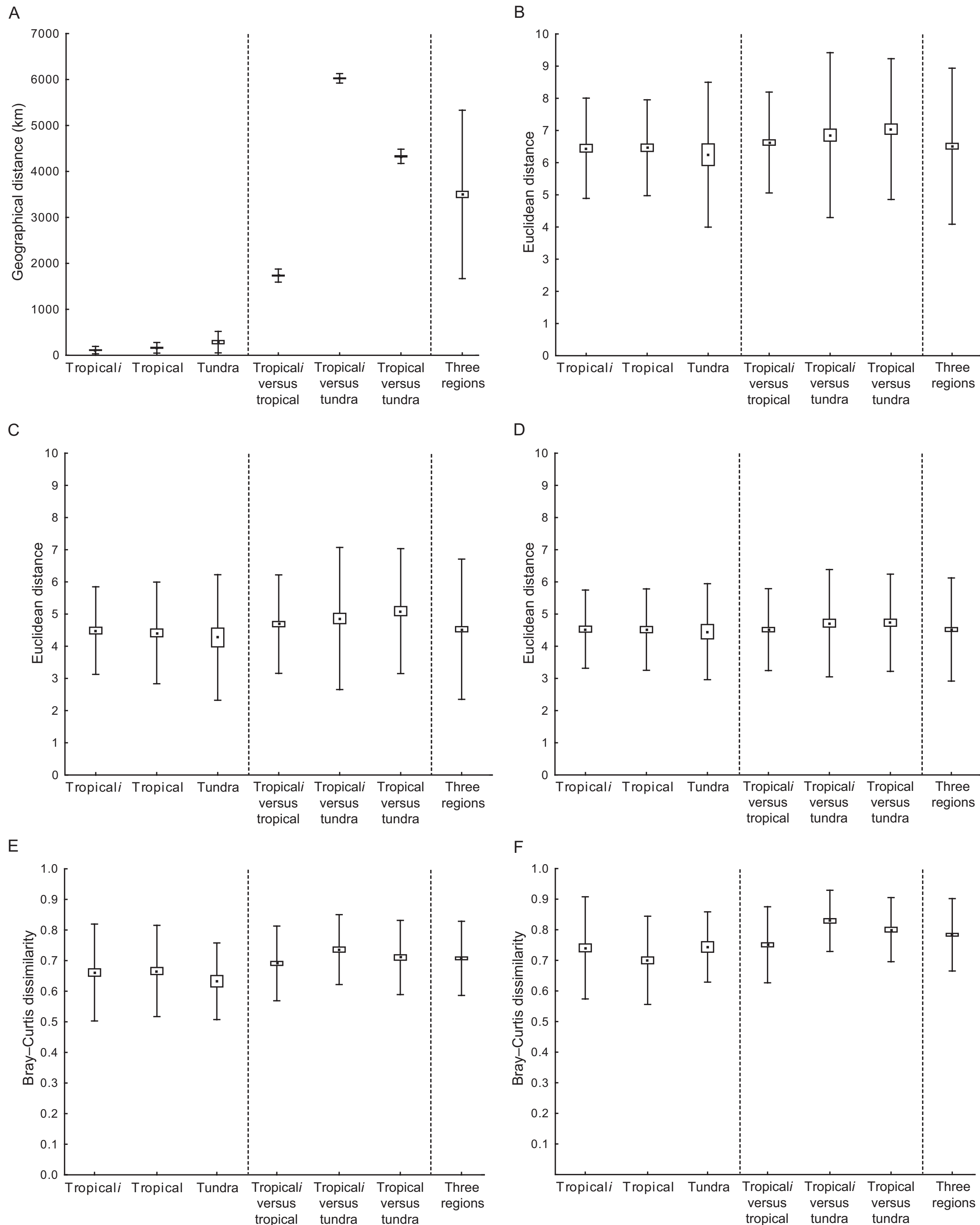

Fig. 1. (A) Mean geographical distance, (B) mean environmental heterogeneity based on 22 abiotic and biotic variables, (C) mean environmental heterogeneity based on 11 abiotic variables, (D) mean environmental heterogeneity based on 11 biotic variables, (E) mean Bray–Curtis OTU dissimilarity based on PA and (F) mean Bray–Curtis OTU dissimilarity based on RA among lakes within each climatic region (tropical*i*, tropical and tundra), among the tropical*i* and tropical regions (tropical*i* versus tropical), tropical*i* and tundra regions (tropical*i* versus tundra), tropical and tundra regions (tropical versus tundra), and among the three regions (three regions) using all within-region (tropical*i*, tropical and tundra) or among-region (two regions, three regions) pairwise comparisons. Boxes represent ± standard error, whiskers represent ± standard deviation.

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*

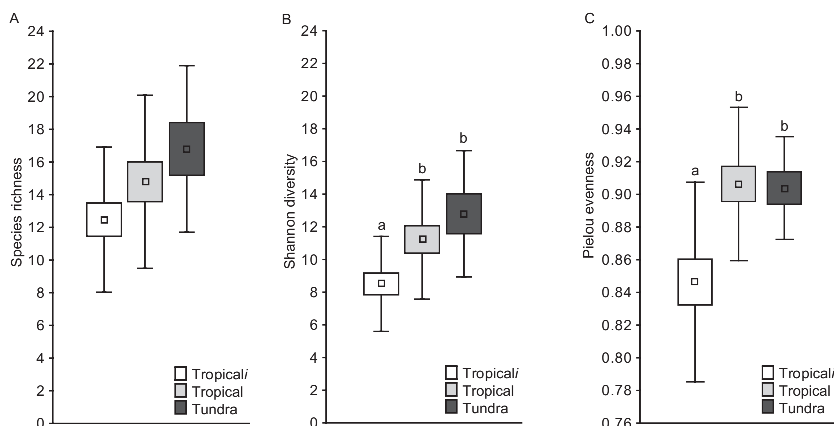

Fig. 2. (A) Mean OTU bands richness, (B) Shannon diversity (C) and Pielou evenness of the bacterial communities within each of the climatic regions: tropical*i* ($n = 19$), tropical ($n = 19$) and tundra ($n = 10$). Boxes represent ± standard error, whiskers represent ± standard deviation. Different letter codes represent significantly different groups at $P < 0.05$.

temperature range than the tropical zone. Within each climatic region, the mean geographical distance among lakes ranged from 114 km (tropical*i* region), 164 km (tropical) to 287 km (tundra) (Fig. 1A). Mean geographical distance between the tropical*i* and tropical lakes was 1735 km, between the tropical*i* and tundra lakes 6026 km, and between the tropical and tundra lakes 4330 km (Fig. 1A). Mean geographical distance between the tropical*i*, tropical and tundra lakes was 3500 km (Fig. 1A). Within the three climatic regions (tropical*i*, tropical and tundra), mean environmental heterogeneity among lakes and the variability in this heterogeneity were comparable (Fig. 1B), also when analysing the abiotic data (Fig. 1C) and biotic data (Fig. 1D) separately. The level of environmental heterogeneity among lakes did not show a tendency towards an increase with increasing spatial scale, as mean environmental heterogeneity among lakes of two regions, and among lakes of the three regions, were similar to the mean within-region environmental heterogeneity (Fig. 1B–D).

Bacterial community similarity among and within climatic regions

A total of 70 different OTUs were detected from the 48 study lakes. Sixty-four OTUs were recorded in the tropical*i* lakes, 59 in the tropical lakes and 61 in the tundra lakes. The total number of OTUs found in one lake ranged from 8 to 26. Within each of the three climatic regions (tropical*i*, tropical and tundra), there was a relatively high mean Bray–Curtis dissimilarity among lakes, with high variability in dissimilarity within regions for both presence-absence (Fig. 1E) and relative abundance data (RA; Fig. 1F). Mean Bray–Curtis dissimilarities tended to be comparable for the three climatic regions and over higher spatial scales (Fig. 1E and F). OTU richness did not differ significantly among climatic regions [one-way analysis of variance (ANOVA), $P = 0.0792$], but there was a main effect of climatic region on Shannon diversity (one-way ANOVA: $P = 0.0053$) and Pielou evenness (one-way ANOVA: $P = 0.0012$) (Fig. 2). The tropical*i* lakes had a significantly lower Shannon diversity compared with the tropical [Tukey honestly significant difference (HSD) test: $P = 0.0471$] and tundra lakes (Tukey HSD test: $P = 0.00689$) (Fig. 2). Similarly, the tropical*i* lakes had a significantly lower Pielou evenness compared with the tropical (Tukey HSD test: $P = 0.00204$) and tundra lakes (Tukey HSD test: $P = 0.0155$) (Fig. 2). Although 76% of the OTUs were found in all three geographical regions, there was a significant overall differentiation in BCC between the regions, as shown by the results of redundancy analysis (RDA) and permutational multivariate analysis of variance (perMANOVA) ($P < 0.05$; Table 1), both based on RA and presence-absence data (PA). RDA and perMANOVA between individual regions revealed significant differences in BCC between tropical and tropical*i*

Table 1. *P*-values (in *italic*) and adjusted R^2-values (between brackets, in %) of the comparisons of similarity in BCC with RDA (Hellinger-transformed) and perMANOVA (Bray–Curtis distance) between the different climatic regions based on RA and PA.

	RA		PA	
Compared regions	RDA	perMANOVA	RDA	perMANOVA
Three regions	*0.001* (6.3%)	*0.0002* (8.6%)	*0.001* (4.7%)	*0.0002* (7.7%)
Tropical-tropical*i*	*0.001* (2.2%)	*0.0272* (2.9%)	*0.008* (1.8%)	*0.0306* (2.9%)
Tropical-tundra	*0.001* (4.7%)	*0.0003* (8.6%)	*0.002* (3.4%)	*0.0019* (6.7%)
Tropical*i*-tundra	*0.001* (6.7%)	*0.0002* (10.5%)	*0.001* (5.4%)	*0.00001* (9.9%)

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*

lakes, between tropical and tundra lakes, and between tropical*i* and tundra lakes ($P < 0.05$; Table 1).

The results of the similarity percentage (SIMPER) analyses of RA identifying the OTUs that contribute most strongly to the dissimilarity between the geographical regions are given in Table S2; parts of these bands were excised and sequenced. Bacterioplankton communities in tropical*i* lakes differed from those of the other lakes mainly because of higher relative abundance of a member of *Synechococcus* (DGGE 51.0), a member of the *Alphaproteobacteria* (DGGE 34.8), a member of the *Betaproteobacteria* (DGGE 52.2) and a member of the *Actinobacteria* (DGGE 53.4). In the lakes situated in the tropical region, we found a higher average abundance of a member of the *Actinobacteria* (DGGE 62.4) and a member of the *Betaproteobacteria* (DGGE 28.7), and a lower abundance of members of the genus *Synechococcus* (DGGE 41.0, 51.0). The samples taken in tundra lakes distinguished themselves from the other samples by a higher average relative abundance of a member of the genus *Flavobacterium* (DGGE 70.6) and of *Acinetobacter lwoffii* (DGGE 36.3). Some species exhibited a relative decrease in abundance from north to south (DGGE 53.4, 77.0, 34.8, 52.2, 78.1, 46.8) or from south to north (DGGE 70.6, 36.3).

Relative importance of environmental and spatial variables

When considering the three regions together, variation partitioning between the selected environmental and spatial variables on RA revealed that pure spatial variables (i.e. after the removal of environment-related variation in space) explained 3% (2% on PA) of the total variance in BCC, while pure environmental variables explained 7% (3%), and a common environment-region effect explained 8% (6%) of the total variance (R^2-adjusted; Fig. 3A; Fig. S3A). A large amount of variation (82% and 89% respectively) remained unexplained. Based on RA, the environmental variables that significantly contributed to explain the overall BCC patterns after correction for the spatial effect were total phosphorus concentration, pH and abundances of pelagic cladocerans (Table 2). When based on PA, only the abundances of pelagic cladocerans contributed significantly to the pure environmental signal (Table 2). Abundances of pelagic cladocerans were significantly affected by climatic region (one-way ANOVA: $P = 0.032$), being significantly higher in the tundra region than in the tropical region (Tukey HSD test: $P = 0.025$) (Fig. S2).

Tropicali and tropical zones. Restricting our analysis to the tropical*i* and tropical regions, environmental and spatial variables together explained over 20% (10% on PA) of the total variance in BCC (Fig. 3B; Fig. S3B). Pure spatial variables did not explain a significant part of the total variance in BCC, while pure environmental variables explained 11% (6%), and a common environment-region effect explained 8% (5%) of the total variance. The environmental variables that significantly explained the BCC patterns after correcting for space based on RA were abundances of pelagic cladocerans, conductivity and Secchi depth (Table 2). For PA, only abundances of pelagic cladocerans and conductivity contributed significantly to the pure environmental signal (Table 2).

Tropicali zone. Pure spatial variables did not explain a significant part of the total variation in BCC, while 8% (7%) of the total variance was explained by pure environmental variables, and 19% (9%) by a common effect (Fig. 3C; Fig. S3C). However, 69% (82%) of the variation remained unexplained. For both relative abundance and PA, none of the selected environmental variables significantly explained the BCC pattern when corrected for spatial structure (Table 2).

Tropical zone. After removal of environment-related variation, space explained 7% (5% on PA) of the variation in the data (Fig. 3D; Fig. S3D). Of the total variance, 20% (10%) was explained by pure environmental variables, and 0% (3%) was explained by a combined effect of space and environment. Abundance of pelagic cladocerans was the only factor that significantly contributed to the pure environmental signal (Table 2).

Tundra zone. After removal of environment-related variation, space no longer explained any variation in the data (Fig. 3E; Fig. S3E). Of the total variance, 23% (11%) was explained by pure environmental variables, and 5% (4%) was explained by a combined effect of space and environment. The environmental variables that significantly contributed to the pure environmental signal were chlorophyll a concentration (RA and PA) and abundances of cyclopoid copepods (relative abundance) (Table 2).

Influence of abiotic and biotic variables on the combined effect of space and environment. Partitioning the variation in BCC based on three sets of explanatory variables (abiotic, biotic and spatial variables) showed that the combined effects of space and environment (Fig. 3; Fig. S3) were predominantly due to spatially related abiotic variables (Figs S3 and S4). For the three regions together (Fig. S4A), the tropical*i* and tropical zone (Fig. S4B) and the tropical zone (Fig. S4D), and for the RA of the tropical*i* zone (Fig. S4C) and the Tundra zone (Fig. S4E), 2.8–5.8% of the total variation was explained by a combined

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*

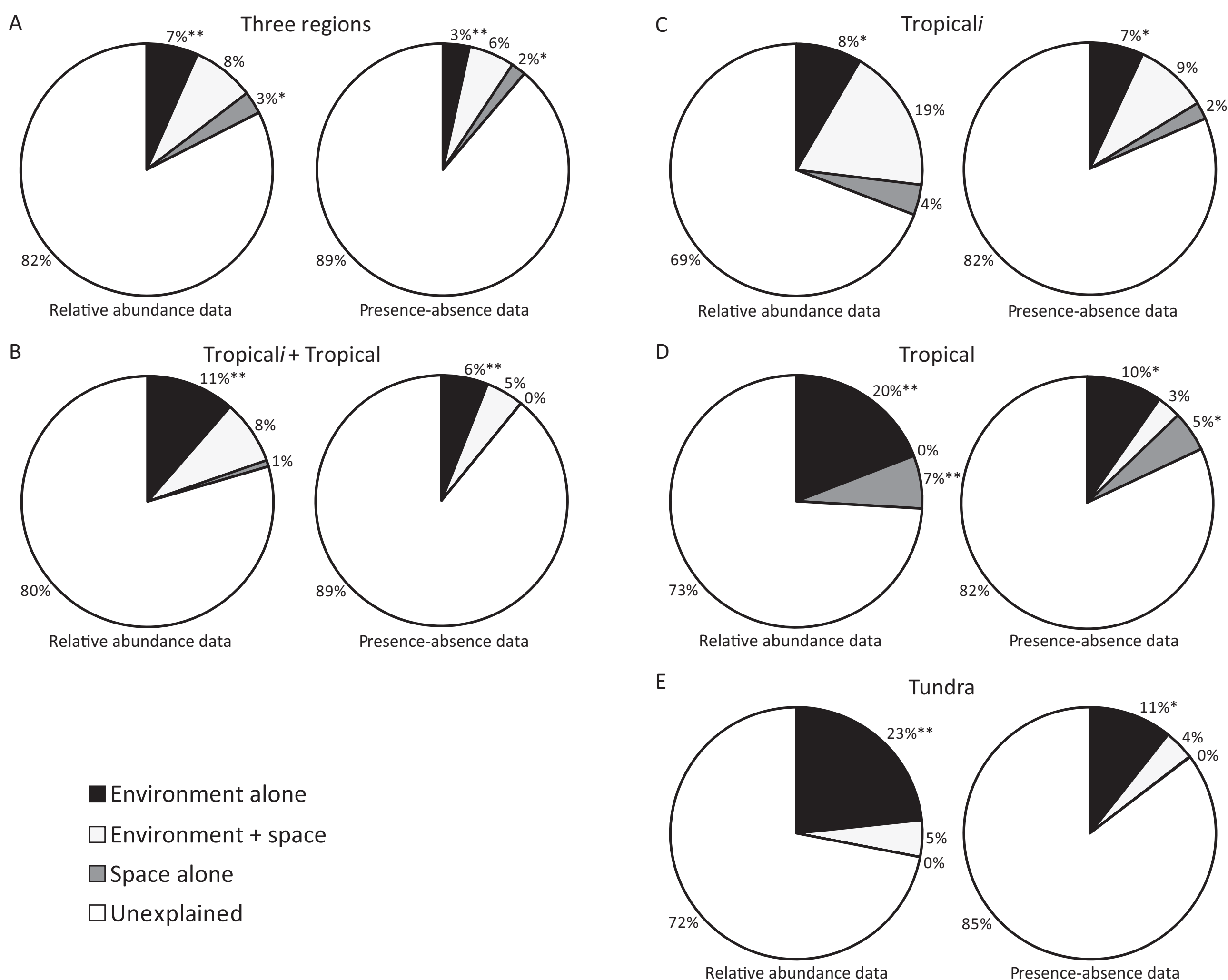

Fig. 3. Variation partitioning of the bacterial community composition. Shown are the results for (A) the three regions analysed together, (B) the regions tropical*i* and tropical together, (C) tropical*i* ($n = 19$), (D) tropical ($n = 19$) and (E) tundra ($n = 10$). Asterisks indicate percentages explaining a statistically significant part of the variation (*$P < 0.05$; **$P < 0.01$).

effect of abiotic and spatial variables, against 0–1.9% by a combined effect of biotic and spatial variables.

Discussion

The present survey of shallow lakes along a latitudinal gradient in South America shows that the bacterioplankton community composition (BCC) differs significantly among the tundra, tropical and tropical*i* climate zones included in the study, despite that 76% of the OTUs were found in all three geographical regions. Significant latitudinal variation in BCC within similar habitats has been observed before for soil communities (Fierer and Jackson, 2006), marine benthic bacteria (Zinger *et al.*, 2011) and freshwater bacterioplankton (Yannarell and Triplett, 2005; Van der Gucht *et al.*, 2007; Schiaffino *et al.*, 2011), but not for marine bacterioplankton (Pommier *et al.*, 2007; Zinger *et al.*, 2011). Similarly to our results, Schiaffino and colleagues (2011) and Van der Gucht and colleagues (2007) observed that a relatively high percentage of the OTUs were present in all sampled regions (60% and 85%, respectively, against 76% in the present study). In our study, the dominant discriminating taxa belonged to common phylogenetic groups of freshwater bacteria, including the *Betaproteobacteria* and *Actinobacteria* (Allgaier and Grossart, 2006; Niu *et al.*, 2011). However, because of the short sequence length, higher-resolution conclusions cannot be drawn.

In previous studies, the latitudinal variation in BCC has been mainly linked to variations in local environmental factors (Yannarell and Triplett, 2005; Fierer and Jackson, 2006; Van der Gucht *et al.*, 2007; Schiaffino *et al.*, 2011) and landscape characteristics (Yannarell and Triplett,

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*

Table 2. Overview of the environmental variables significant in explaining the variation in BCC without (marginal effect) and with (conditional effect) correction for spatial effects, for the different regions and BCC data types (relative abundance and PA).

Region	BCC data type	Significant environmental variables	
		E (marginal effect)	E/S (conditional effect)
Three regions	RA	Temperature, conductivity, abundances of pelagic cladocerans, pH, total phosphorus and $NO_3 + NO_2$ concentrations	Total phosphorus concentration, pH, abundances of pelagic cladocerans
	PA	Temperature, abundances of pelagic cladocerans, conductivity, $NO_3 + NO_2$ and total phosphorus concentrations	Abundances of pelagic cladocerans
Tropical*i* + tropical	RA	Conductivity, abundances of pelagic cladocerans, Secchi depth, temperature	Abundances of pelagic cladocerans, Secchi depth, conductivity
	PA	DOC, temperature, abundances of pelagic cladocerans, conductivity	Abundances of pelagic cladocerans, conductivity
Tropical*i*	RA	DOC, depth, pH	/
	PA	DOC, pH	/
Tropical	RA	Abundances of pelagic cladocerans, Secchi depth, conductivity	Abundances of pelagic cladocerans
	PA	Abundances of pelagic cladocerans, Secchi depth, total phosphorus concentration	Abundances of pelagic cladocerans
Tundra	RA	Chlorophyll a concentration, abundances of cyclopoid copepods, conductivity	Chlorophyll a concentration, abundances of cyclopoid copepods
	PA	Chlorophyll a concentration, pH	Chlorophyll a concentration

E, uncorrected environmental effect; E/S, environmental effect corrected for spatial signal; PA, presence-absence data.

2005), but also to geographical distance (Schiaffino *et al*., 2011; Zinger *et al*., 2011). In the present study, variation partitioning indicates that the variation in BCC over the whole latitudinal gradient is due to local environmental factors (3–7%), spatially related environmental factors (6–8%) and to a lesser extent space (2–3%). A large part of the variation (82–89%) could not be explained by the variables measured, suggesting that other factors such as unmeasured environmental variables and landscape characteristics could impact the variation in BCC. It is important to keep in mind that the BCC was characterized by DGGE, a fingerprint technique that has limitations in describing compositional patterns. DGGE is known to detect the most abundant taxa while ignoring the less abundant and rare taxa (Muyzer and Smalla, 1998; Duarte *et al*., 2012). This technique gives therefore incomplete information on bacterial biogeography. The patterns observed in this study should thus be interpreted to indicate that, even at the large spatial scales considered, the relative importance of the dominant species of the bacterioplankton communities are influenced by environmental conditions rather than space. However, our study does not allow to make strong inferences on rare species.

Not only at the largest spatial scale (from the tropical*i* to tundra region), but also on smaller spatial scales (tropical*i* and tropical region) and within the three climatic zones, the bacterioplankton community composition was to a larger extent determined by local environmental conditions than by spatial factors. The observed importance of local factors is in accordance with previous studies on freshwater bacterioplankton (Beisner *et al*., 2006; Langenheder and Ragnarsson, 2007; Van der Gucht *et al*., 2007; Wu *et al*., 2007; Logue and Lindström, 2010; De Bie *et al*., 2012; Langenheder *et al*., 2012) and indicates that the freshwater bacterioplankton community structure is thus mainly the result of selection by local environmental conditions (species sorting; Chase and Leibold, 2003, Leibold *et al*., 2004), while regional processes such as dispersal limitation or mass effects are of less importance. In metacommunity datasets, such an environmental signal may be the result of a direct effect of the measured environmental variables or affected by environmental variables that are strongly correlated with the measured variables.

In the present dataset, both biotic and abiotic environmental variables were important in explaining variation in BCC, but it is remarkable that when correcting for spatial effects, mainly biotic variables (abundances of pelagic cladocerans, abundances of cyclopoid copepods, chlorophyll a concentrations) remained significant. The importance of biotic variables in structuring freshwater bacterioplankton communities has been shown before (e.g. Langenheder and Jurgens, 2001; Jardillier *et al*., 2004; Van der Gucht *et al*., 2007; Verreydt *et al*., 2012), with significant associations of bacterioplankton abundance and community composition with phytoplankton community composition (e.g. Allgaier and Grossart, 2006; Niu *et al*., 2011), macrophyte abundance (e.g. Wu *et al*., 2007; Ng *et al*., 2010) and densities of antagonists including lytic bacterial viruses, bacterivorous protists and zooplankton (e.g. Sanders *et al*., 1989; Jurgens *et al*., 1994;

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*

Yoshida *et al*., 2001; Weinbauer *et al*., 2007; Lymer *et al*., 2008; Pradeep Ram *et al*., 2014). Additionally, it has been shown that processes affecting biotic variables can indirectly affect BCC through trophic cascades (Verreydt *et al*., 2012). Such indirect effects may either reflect direct interactions (e.g. competition and predation) or a strong impact of the biota on abiotic environmental conditions (Verreydt *et al*., 2012). In the present study, abundances of pelagic cladocerans significantly explained variation in BCC both in the combined dataset and within the tropical region, while abundances of cyclopoid copepods significantly explained variation in BCC within the tundra region. It is well established that predators of bacteria, such as many cladocerans, ciliates and heterotrophic nanoflagellates, or predators on bacterivores, such as cyclopoid copepods, may directly or indirectly have strong effects on bacterial production, abundance and community composition (Jurgens and Jeppesen, 2000; Pernthaler, 2005). In the present study, the impact of the higher trophic level is a key environmental driver that remains significant in explaining variation in BCC when exploring pure environmental effects. This is an important observation, given that many studies on BCC in natural freshwater systems do not quantify biomass or other characteristics of zooplankton and other antagonists. Most studies that do show an effect of antagonists either involved surveys that explicitly monitored these (e.g. Van der Gucht *et al*., 2007; De Bie *et al*., 2012) or involved experiments that specifically tested for such effects under standardized conditions (reviewed in Pernthaler, 2005). Based on our data, we cannot determine whether pelagic cladocerans and cyclopoid copepods influenced BCC directly through predation or indirectly through nutrient regeneration or predation on bacterivores. Because the abundances of ciliates or flagellates – important bacterivores – was not related to BCC, we speculate that the relation of cladocerans and copepods with BCC was caused either by direct predation or through nutrient regeneration rather than through their effect on bacterivores. Additionally, it is notable that the abundance of cladocerans is strongly influenced by the fish community in the present dataset (Kosten *et al*., 2009b). Warmer lakes (tropical*i* and tropical region) contain higher densities of omnivorous fish, diminishing the abundances of cladocerans in these lakes, compared with the tundra lakes (See Fig. 7 in Kosten *et al*., 2009b). The influence of cladocerans on the bacterioplankton could thus be part of a trophic cascade.

Despite the importance of environmental variables in our analysis, spatial variables still explained a significant part of the variation in BCC in the combined dataset and within the tropical region. Spatial signals may reflect five different, non-exclusive processes: (i) dispersal limitation of the bacterial species themselves (Whitaker *et al*., 2003), (ii) dispersal limitation of other organisms influencing BCC, for instance, through trophic cascades (e.g. dispersal limitation of zooplankton; Verreydt *et al*., 2012), (iii) mass effects (source-sink dynamics, with a continuous or substantial influx of organisms that are not self-maintaining in the target environment; Leibold *et al*., 2004), (iv) an influence of unmeasured, spatially structured environmental variables, or (v) an influence of historical contingency related to the geographic position of the habitat (cf. priority effects; Fukami *et al*., 2007; 2010). Mass effects are unlikely to be of importance in our study lakes as they would have resulted in a gradually stronger spatial signal with decreasing spatial scales. Moreover, no water exchange occurred among the studied lakes, nor were they connected to a common river. However, our dataset does not allow us to determine whether the spatial signal observed is due to dispersal limitation of the bacteria themselves or whether it reflects the impact of an unmeasured spatially structured environmental variable (including dispersal limitation of an antagonist). Within the tropical*i* and tundra regions, there was no significant spatial signal, suggesting that the distances covered in these regions (average distance among lakes is 114 km in the tropical*i* and 287 km in the tundra region) do not intrinsically result in dispersal limitation in bacterioplankton, at least at the taxon resolution used in this study. However, within the tropical region (average distance among lakes is 164 km), 5–7% of the variation in BCC was significantly explained by space, but the underlying processes are unclear.

In most of our data subsets, there was also a relatively high combined effect of space and environment in explaining the variation in BCC, especially within the tropical*i* region (9.4–18.5%), indicating the importance of spatially structured environmental variables in bacterioplankton community datasets. Of the environmental variables that significantly explained the variation in BCC based on forward selection, several abiotic variables (temperature, nutrient concentrations, pH and conductivity) were spatially structured and not significant any more after correcting for space, whereas the biotic variables (abundances of pelagic cladocerans and cyclopoid copepods, chlorophyll a concentrations) remained significant in explaining the variation in BCC after correcting for space. The combined effect of space and environment in our dataset was thus mainly the result of abiotic variables. This was confirmed by a variation partitioning using three sets of explanatory variables (abiotic, biotic and spatial variables), showing that 2.8–5.8% of the total variation was explained by a combined effect of abiotic and spatial variables, against 0–1.9% by a combined effect of biotic and spatial variables.

Even though a significant part of the variation in BCC in our datasets could be explained by environmental and/or

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd,

spatial variables included in our analyses, the largest proportion of the variation in BCC (69–89%) remained unexplained. This large fraction of unexplained variance is a recurrent pattern in studies of bacterial metacommunities and of metacommunities in general. This may be due, among others, to unmeasured environmental variables, random and temporal variability, chaotic dynamics (Beninca *et al*., 2008), priority effects (e.g. Fukami *et al*., 2007; Fukami *et al*., 2010) and evolutionary dynamics (Loeuille and Leibold, 2008; Urban *et al*., 2008; Turcotte *et al*., 2012).

There was, moreover, a large variation in the fraction of unexplained variance among our data subsets. The unexplained variation was highest at the larger spatial scales (three or two regions combined) and when using PA, and lowest when looking at RA within individual climatic regions. A higher fraction of explained variation over smaller spatial scales was also observed by Van der Gucht and colleagues (2007), although the environmental variables determined in their study at a small spatial scale differed from those measured at the larger spatial scale. In the present study, the same environmental factors were subjected to forward selection for our different data subsets, and while less (and different) variables were significant within individual regions compared with the larger scale datasets, these variables could explain a larger fraction of the variation in BCC. This might reflect that the environmental and community variation is less complex over smaller than larger spatial scales and can be more accurately caught in a subset of measured environmental variables. Additionally, the pattern of a decrease in the amount of explained variation in BCC may reflect geographic differences in the way environmental drivers impact species occurrences. If environmental drivers structure species occurrence in different ways in different regions, for instance through complex interactions with other environmental drivers, this adds to the noise in the aggregated dataset.

We performed our statistical analyses both on presence-absence and RA. PA give more weight to rare species, whereas RA represent community structure and give more weight to the dominant species. In this study, communities tended to be more similar based on PA compared with RA (Fig. 1E and F), indicating that the variation in community structure among lakes was not only due to changes in species composition, but also to changes in the relative abundance of species. Partitioning of the variation in community structure explained by environmental and spatial variables showed that within each dataset, the relative importance of environmental, spatial and spatially related environmental variables for the explained variation was similar when based on relative abundance and PA, although the total variation explained was always lower when based on PA. A larger proportion of the total variation in community structure could be explained by environment and space when based on RA (17.5–30.8% of the total variation explained) than when based on PA (10.9–18.5% of the total variation explained). This suggests strong effects of environment and space on the abundances of dominant species, e.g. through strong preferences for specific habitats. The possible explanation for the lower amount of explained variation in the PA is that there are more chance effects involved in the detection of relatively rare species, which adds to the unexplained variation component.

Among the three climatic regions studied (tropical*i*, tropical and tundra), there was a difference in the relative contribution of space and environment to the variation in BCC and in the environmental variables significantly explaining this variation. However, one should be cautious in comparing these results as such, as the samples were snapshots in time and the identity and importance of the environmental drivers can fluctuate over time. To assess whether the observed patterns along the latitude gradient are general, there is a need for replicated studies on freshwater bacterioplankton community structure over latitudinal gradients. The same holds for the patterns in diversity observed along the studied latitudinal gradient. The presence of a latitudinal diversity gradient in bacteria is still an open question (Soininen, 2012), and while three studies on marine bacterioplankton reported gradients in diversity over latitude (Pommier *et al*., 2007; Fuhrman *et al*., 2008; Schattenhofer *et al*., 2009), studies on soil bacteria (Fierer and Jackson, 2006) and freshwater bacterioplankton (Humbert *et al*., 2009) failed to find such a gradient. In our study, there was no significant difference in bacterioplankton OTU richness over latitude, but both Shannon diversity and Pielou evenness were significantly lower in the tropical*i* lakes (5.37°S to 6.39°S) compared with the tropical lakes (19.48°S to 22.50°S) and tundra lakes (49.41°S to 54.63°S). Tropical and tundra lakes had thus a more even bacterioplankton community compared with the tropical*i* lakes. Latitudinal gradients in evenness have been studied for several organisms, resulting in positive, negative and non-significant relationships depending on the organism group and the evenness index, with different potential explanations of the observed patterns including the impact of disturbance, productivity and biotic interactions (reviewed in Willig *et al*., 2003). Based on our results, we cannot determine why tropical*i* lakes have relatively more dominating species. The tropical*i* climate zone is characterized by a smaller annual temperature range than the tropical zone, and is likely more stable in terms of environmental conditions, with a higher productivity.

In conclusion, our results confirm the importance of species sorting by local environmental conditions in shaping bacterioplankton communities of shallow lakes.

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*

Even over the large-scale latitudinal gradient studied, which covered 6026 km, the impact of spatial drivers was limited. Moreover, our results emphasize the importance of biotic variables in structuring bacterioplankton metacommunities. Non-inclusion of biotic variables may result in an artificially low environmental signal and a potentially inflated signal of space. The recurrent large proportion of unexplained variation in bacterioplankton community structure indicates that further attention is needed on alternative potential metacommunity-shaping processes, which may include eco-evolutionary dynamics and priority effects, and on alternative approaches to address these processes.

Experimental procedures

Selection and sampling of lakes

The lakes sampled are a subset of the South American Lake Gradient Analysis lakes (Kosten *et al*., 2009a). Forty-eight shallow lakes, located in South America, were sampled and grouped in three categories based on the prevailing climate characteristics (monthly precipitation and temperature) following the Köppen climate classification system (Köppen, 1936), digitized by Leemans and Cramer (1991): the tropical*i* (19 lakes), tropical (19 lakes) and tundra (10 lakes) zone (see Fig. S1 and Table S1 for the geographical location). Tropical*i* is an isothermal subzone in the tropics, which has a smaller annual temperature range than the tropical zone. The lakes were selected to resemble one another as closely as possible in morphology and altitude and to vary as much as possible in trophic state within each climate zone. None of the studied lakes were connected to each other. For a detailed description of the sampled lakes, we refer to Kosten and colleagues (2009a).

Each lake was sampled once during the dry season (tropical*i* and tropical lakes) or in summer (tundra lakes) between August 2005 and February 2006 by the same team. To integrate spatial variability within each lake, water was collected at 20 randomly selected locations per lake and pooled into one sample. To obtain DNA samples of bacteria, 5 l of the pooled sample were fractionated using a mesh of 20 μm, separating bacteria that were free living or attached to small seston particles from organisms attached to large seston particles. The small fraction was then filtered on a 0.2 μm MF-Millipore MCE filter and the filters then stored at –80°C until further analysis. Abiotic local environmental variables measured were transparency (Secchi disc and Sneller depth), pH, temperature, conductivity, concentrations of DOC, nitrate and nitrite, ammonium, soluble reactive phosphorus, total phosphorus, suspended solids and mean lake depth. Sneller depth is the deepest point under water (in centimetres) at which a Secchi disc, lowered in a gray polyvinyl chloride (PVC) tube (8 cm diameter) filled with lake water, can be seen (van de Meutter *et al*., 2005) and has been devised for shallow ponds. Biotic variables determined were chlorophyll a concentration, the biomass and abundance of phytoplankton, pelagic cladocerans, rotifers, calanoids and cyclopoid copepods, the abundances of flagellates and ciliates, and the percentage of lake surface covered by floating plants, emergent vegetation and submerged vegetation. Details on the sampling design and methods are given in Kosten and colleagues (2009a). Abundances of flagellates and ciliates were estimated by direct counts at 1000× magnification with an epifluorescence microscope (Olympus BX60). Water samples were filtered over a black polycarbonate membrane (pore size 0.8 μm, Millipore) and stained with 4′,6-diamidino-2-phenylindole (5 $\mu g\ ml^{-1}$ final concentration, 5 min incubation) following Porter and Feig (1980). At least 50 random fields per slide were counted. Heterotrophic nanoflagellate and ciliate biomasses were estimated based on measurements of cell dimension, geometric formulae to determine volume (Massana *et al*., 1997) and published conversion factors. For heterotrophic nanoflagellate, we used a conversion factor of 220 fg C μm^{-3} (Borsheim and Bratbak, 1987); for ciliates, a conversion factor of 0.19 pg C μm^{-3} (Putt and Stoecker, 1989) was used.

DNA extraction and polymerase chain reaction

DNA was extracted directly using the bead-beating method concomitantly with phenol extraction and ethanol precipitation (Zwart *et al*., 1998). Following extraction, the DNA was purified on a Wizard column (Promega, Madison, WI, USA) according to the manufacturer's instructions. For DGGE analysis, a short 16S rDNA fragment was amplified with eubacterial primers 357F-GC-clamp (5′-CGCCCGCCGCG CCCCGCGCCCGGCCCGCCGCCCCCGCCCCCCTACGGG AGGCAGCAG-3′) and 518R (5′-ATTACCGCGGCTGCTGG-3′). The polymerase chain reaction (PCR) was carried out in a T1-thermocycler (Biometra). Each mixture (50 μl) contained 5 μl of template DNA, each primer at a final concentration of 0.5 μM, each deoxynucleosidetriphosphate at a concentration of 200 μM, 1.5 mM $MgCl_2$, 20 ng of bovine serum albumin, 5 μl of 10× PCR buffer [100 mM Tris-HCl (pH 9), 500 mM KCl], 2.5 U of Taq DNA polymerase (Ampli-Taq, Perkin Elmer) and sterile water (Sigma). After incubation for 5 min at 94°C, a touchdown PCR was performed using 20 cycles consisting of denaturation at 94°C for 1 min, annealing starting at 65°C (the temperature was reduced by 0.5°C for every cycle until the touchdown temperature of 56°C was reached) for 1 min and primer extension at 72°C for 1 min. Five additional cycles were carried out at an annealing temperature of 55°C, followed by a final elongation step for 10 min at 72°C. The presence of PCR products and their concentration were determined by analysing 5 μl of product on 1% (w/v) agarose gels, stained with ethidium bromide and compared with a molecular weight marker (Smartladder; Eurogentec).

DGGE analysis

PCR products obtained with primers 357FGC and 518R were analysed on a 35–70% denaturant DGGE gel as described earlier (Van der Gucht *et al*., 2001). DGGE gels were stained with ethidium bromide and photographed on a UV transillumination table (302 nm) with a CCD camera. The 98 samples were analysed on nine parallel DGGE gels. As standards, we used a mixture of DNA from nine clones, obtained from a clone library of the 16S rRNA genes from Lake Visvijver (Belgium). On every gel, three standard lanes were analysed

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd,

in parallel to the samples. Because these bands are expected to be formed at the same denaturant concentration in the gel, their position was used to compare the patterns formed in different gels. Digitized DGGE images were analysed using the software package BIONUMERICS 1.5 (Applied Maths BVBA, Kortrijk, Belgium). The software performs a density profile through each lane, detects the bands and calculates the relative contribution of each band to the total band signal in the lane after applying a rolling disk as background subtraction. Bands occupying the same position in the different lanes of the gel were first identified by the program and then visually checked. A matrix was compiled based upon the relative contribution of individual bands to the total band signal in each lane.

Sequencing and identification of excised DGGE bands

To retrieve phylogenetic information on the DGGE bands, bands of interest were excised and sequenced. Selection of these bands of interest was based on two criteria: (i) within samples, bands with relatively high fluorescence levels (and thus high DNA concentrations) were selected, and (ii) across all samples, it was taken care that the OTU-diversity within the total dataset was covered, so that for each OTU present in the dataset multiple bands were selected if possible. Nucleotide sequences of these DGGE bands were obtained by direct sequencing of DNA from excised DGGE bands as described in Van der Gucht and colleagues (2001). Sequencing was performed with the ABI-Prism sequencing kit (PE-Biosystems) using the primer R519 (5′-GTATTACCG CGGCTGCTG-3′) and an automated sequencer (ABI-Prism 377). Sequences were identified using BLAST (Basic Local Alignment Search Tool) (Altschul *et al.*, 1990) against GenBank and EMBL sequences.

Data analysis

Dataset compilation. Bacterial community structure was based on DGGE. DGGE bands with different positions along the sample lanes in the gel are considered different OTUs. The relative contribution of the band to the total band signal in a sample lane is used as an approximation of the relative abundance of the OTU in that sample. A dataset of 22 environmental variables was compiled from the original 42 environmental parameters based on three selection criteria: (i) strength of correlation among variables based on Spearman rank correlation coefficients and principal component analysis (PCA), (ii) known or presumed importance of variables for bacterioplankton in shallow lake systems based on previous knowledge (e.g. pH, DOC, densities of heterotrophic nanoflagellates) and (iii) the fact that we wanted to retain the diversity in abiotic as well as biotic variable types within our selection (different organism groups for biotic variables, different nutrient types and physical parameters for abiotic variables). Spearman rank correlations coefficients were used to detect the most highly correlated variables (correlation coefficient > 0.75) and were calculated between the untransformed environmental variables using STATISTICA, version 12 (StatSoft). PCA was performed on standardized environmental variables using the function *rda* from the R-package vegan in R 3.0.1 (R Core Team, 2013). The 22 selected variables were depth, temperature, *in vivo* chlorophyll a concentration, Secchi depth, pH, conductivity, concentrations of DOC, nitrate and nitrite, ammonium, soluble reactive phosphorus and total phosphorus, suspended solids, percentage of lake surface covered by floating plants, emergent vegetation and submerged vegetation, absolute abundances of calanoid copepods, cyclopoid copepods, rotifers, pelagic cladocerans, flagellates and ciliates.

The datasets containing the 22 environmental variables, geographical coordinates and bacterial taxon data (relative abundances) of the 48 lakes were reorganized into five subsets based on climatic regions: three separate datasets were made for the tropical*i*, tropical and tundra region, one dataset contained the data from the tropical*i* + tropical regions and one dataset contained all data from the three climatic regions (three regions).

Geographical distance and environmental heterogeneity. Within each individual climatic region, for each pairwise combination of climatic regions and for the three climatic regions combined, pairwise geographical distances between lakes were calculated using the function *rdist.earth* from the R-package fields in R 3.0.1. Within the three individual climatic regions (tropical*i*, tropical and tundra), all pairwise calculations among lakes of the region were taken into account. For the pairwise comparisons of climatic regions (tropical*i* versus tropical, tropical*i* versus tundra, tropical versus tundra) and for the dataset containing the three climatic regions, only the distances between lakes of different climatic regions were taken into account.

Within each individual climatic region, for each pairwise combination of climatic regions and for the three climatic regions combined, pairwise environmental distances between lakes were calculated using the Euclidean distances of the standardized environmental data. Distances were based on the selected 22 environmental variables (see above), and separately on the 11 selected abiotic variables and the 11 selected biotic variables. Variables were standardized using the z-score (mean of 0 and standard deviation of 1) for each dataset and subset separately, and Euclidean distances were calculated using the function *dist* from the R-package stats. Within the three climatic regions tropical*i*, tropical and tundra, all pairwise calculations among lakes were taken into account. For the pairwise comparisons of climatic regions (tropical*i* versus tropical, tropical*i* versus tundra, tropical versus tundra) and for the dataset containing the three climatic regions, only the Euclidean distances between lakes of different climatic regions were taken into account.

Bacterial community structure within and among climatic region. To test whether BCC was significantly affected by climatic region (tropical*i*, tropical and tundra), RDA and perMANOVA were performed on relative abundance and PA. RDA was performed on Hellinger-transformed species data, followed by a permutation test (10 000 permutations), using the functions *rda* and *anova.cca* of the R-package vegan (Oksanen *et al.*, 2013). PerMANOVA was performed on Bray–Curtis community similarity data using the function *adonis* (10 000 permutations) of the same R-package. Analyses were performed on the total dataset (three regions

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*

combined) and on the three pairwise combinations of climatic regions. To identify which OTUs were significantly discriminating among the climate zones, SIMPER analysis (Clarke, 1993) was used on the arcsine-transformed relative abundance values using PRIMER. Within-region dissimilarity in BCC among lakes was assessed by calculating all pairwise Bray–Curtis dissimilarities among lakes within each of the three climatic regions (tropical*i*, tropical and tundra) using the function *bcdist* of the R-package ecodist. Among-region Bray–Curtis dissimilarities were calculated between tropical*i* and tropical lakes, tropical*i* and tundra lakes, tropical and tundra lakes, and between lakes of the three different climatic regions in the same way, and again only comparisons between lakes of different regions were taken into account. Species richness (number of OTUs), Shannon diversity [$\exp(-\Sigma\, p_i \ln p_i)$, with p_i the relative abundance of species *i*] and Pielou evenness (Shannon diversity divided by species richness) were calculated for each of the three climatic regions based on the RA. The Pielou evenness is a relative measure of community evenness, independent of diversity (Jost, 2010). Effects of climatic region on the three diversity indices were analysed by one-way ANOVA followed by a Tukey HSD *post-hoc* test in STATISTICA version 12 (StatSoft) after testing for normality and homogeneity of variances using the Shapiro–Wilk test and Levene's test respectively.

Variation partitioning. To get insight into the relative importance of environmental and spatial variables in explaining the variation in BCC, we assessed the relative contribution of environment (E), space (S), their combined effect (E∩S) and their conditional effects (E/S and S/E) on community variation through variation partitioning (Borcard *et al.*, 1992). For each dataset (tropical*i*, tropical, tundra, tropical*i* + tropical and the three regions together), variation partitioning was performed on Hellinger-transformed relative abundance and presence-absence BCC data. For each of the four datasets, unique environmental and spatial models were constructed *a priori* following the procedure described below, and this separately for the relative abundance and PA. Environmental and spatial models are sets of environmental and spatial variables, respectively, that significantly explain (part of the) variation in the BCC.

To construct the *environmental models* (E), forward selection (Blanchet *et al.*, 2008) was applied to the z-score-transformed 22 environmental variables using the function *forward.sel* from the R-package packfor with a threshold *P*-value of 0.05. Variables significantly contributing to the model were retained for variation partitioning. Before forward selection, we explored the presence of collinearities among the non-transformed 22 variables by computing variance inflation factors (VIFs) using the function *vif.cca* from the R-package vegan. VIFs were in almost all cases < 10. The *spatial models* (S) were constructed based on the original latitude and longitude data and Moran's eigenvector maps (MEM) eigenvectors. MEM eigenvectors were constructed based on the latitude and longitude of the lakes (Borcard and Legendre, 2002; Borcard *et al.*, 2004; Peres-Neto *et al.*, 2006) by the following procedure: a Euclidean distance matrix was constructed from the latitudes and longitudes, and the matrix was truncated using a truncation distance of four times the maximum value of the minimum spanning tree of the Euclidean distance matrix. Based on this truncated distance matrix, principal coordinates were calculated using the function *cmdscale* of the R-package stats. Only principal coordinate eigenvectors with positive eigenvalues were retained. A forward selection procedure (Blanchet *et al.*, 2008) was applied to the spatial dataset containing the principal coordinate eigenvectors with positive eigenvalues (i.e. MEM eigenvectors), and the original latitude and longitude data using the function *forward.sel* from the R-package packfor. The threshold to stop forward selection was the adjusted R^2-value of the spatial RDA model. The spatial variables retained by this forward selection were used as spatial model during variation partitioning.

Variation partitioning of the BCC was performed using the function *varpart* from the R-package vegan, the Hellinger-transformed species data, the forward-selected environmental variables and the forward-selected spatial variables. To assess whether the conditional environmental component (E/S, the pure environmental component without spatial effects) and conditional spatial component (S/E, the pure spatial component without environmental effects) contributed significantly to the model, the significant contribution of each conditional fraction in explaining the variation in community composition was tested with a permutation test (9999 permutations) using the functions *rda* and *anova.cca* of the R-package vegan. To assess which of the selected environmental variables contributed significantly to the conditional environmental model (E/S), a permutation test (9999 permutations) was performed using the function *anova.cca by terms* of the R-package vegan.

To get insight into the contribution of abiotic and biotic variables on the spatially related environmental signal (E∩S), we performed variation partitioning on the relative abundance and PA of the BCC using three sets of explanatory variables: abiotic, biotic and spatial variables. The abiotic and biotic sets contained each 11 of the 22 selected environmental variables (see above). The construction of the abiotic (A), biotic (B) and spatial (S) model was performed using forward selection as described above. Variation partitioning using the three explanatory models, assessment of the significance level of the conditional components [A/(BuS); B/(AuS), S/(AuB)] by RDA and assessment of the variables significantly contributing to the conditional models was performed as described above.

Acknowledgements

We thank the SALGA (South American Lake Gradient Analysis) team for the immense effort in bacterioplankton sampling and collection of data on a broad range of environmental variables in the 48 study lakes. SALGA was financed by The Netherlands Organization for Scientific Research (NWO) Grants W84-549 and WB84-586, and The National Geographic Society Grant 7864-5, in Brazil by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) Grants 480122, 490409 and 311427, and in Uruguay by Programa de Desarrollo de las Ciencias Básicas (PEDECIBA), Maestría en Ciencias Ambientales, Donación Aguas de la Costa S.A. and Banco de Seguros del Estado. In Belgium, this study was financially supported by project Grant

G.0978.10 to W. V. and L. D. M., and project grant KAN 1.5.089.09N to K. V. d. G. of the National Fund for Scientific Research, Flanders (FWO), and by the KU Leuven Research Fund Excellence Center financing PF/2010/07. E. J. was supported by the Centre for Regional change in the Earth System (CRES) supported by the Danish Strategic Research Council, and by the Centre for Informatics Research on Complexity in Ecology (CIRCE) supported by the IDEAS pilot program of Aarhus University. S. K. was supported by NWO-VENI Grant 86312012 of The Netherlands Organization for Scientific Research.

References

Allgaier, M., and Grossart, H.-P. (2006) Seasonal dynamics and phylogenetic diversity of free-living and particle-associated bacterial communities in four lakes in northeastern Germany. *Aquat Microb Ecol* **45:** 115–128.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215:** 403–410.

Baas-Becking, L.G.M. (1934) *Geobiologie of inleiding tot de miliekunde*. Den Haag, the Netherlands: Van Stockkum & Zoon.

Beisner, B.E., Peres Neto, P.R., Lindstrom, E.S., Barnett, A., and Longhi, M.L. (2006) The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology* **87:** 2985–2991.

Beninca, E., Huisman, J., Heerkloss, R., Johnk, K.D., Branco, P., Van Nes, E.H., *et al*. (2008) Chaos in a long-term experiment with a plankton community. *Nature* **451:** 822–U827.

Blanchet, F.G., Legendre, P., and Borcard, D. (2008) Forward selection of explanatory variables. *Ecology* **89:** 2623–2632.

Borcard, D., and Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol Model* **153:** 51–68.

Borcard, D., Legendre, P., and Drapeau, P. (1992) Partialling ou the spatial component of ecological variation. *Ecology* **73:** 1045–1055.

Borcard, D., Legendre, P., Avois-Jacquet, C., and Tuomisto, H. (2004) Dissecting the spatial structure of ecological data at multiple scales. *Ecology* **85:** 1826–1832.

Borsheim, K.Y., and Bratbak, G. (1987) Cell-volume to cell carbon conversion factors for a bacterivorous *Monas sp*. enriched from seawater. *Mar Ecol Prog Ser* **36:** 171–175.

Chase, J.M., and Leibold, M.A. (2002) Spatial scale dictates the productivity-biodiversity relationship. *Nature* **416:** 427–430.

Chase, J.M., and Leibold, M.A. (2003) *Ecological Niches: Linking Classical and Contemporary Approaches*. Chicago, IL, USA: University of Chicago Press.

Cho, J.C., and Tiedje, J.M. (2000) Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. *Appl Environ Microbiol* **66:** 5448–5456.

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

De Bie, T., De Meester, L., Brendonck, L., Martens, K., Goddeeris, B., Ercken, D., *et al*. (2012) Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecol Lett* **15:** 740–747.

Duarte, S., Cassio, F., and Pascoal, C. (2012) Denaturing gradient gel electrophoresis (DGGE) in microbial ecology – insights from freshwaters. In *Gel Electrophoresis – Principles and Basics*. Magdeldin, S. (ed.). pp. 173–196. doi:10.5772/38177 InTech.

Fierer, N., and Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci USA* **103:** 626–631.

Fuhrman, J.A., Steele, J.A., Hewson, I., Schwalbach, M.S., Brown, M.V., Green, J.L., and Brown, J.H. (2008) A latitudinal diversity gradient in planktonic marine bacteria. *Proc Natl Acad Sci USA* **105:** 7774–7778.

Fukami, T., Beaumont, H.J.E., Zhang, X.-X., and Rainey, P.B. (2007) Immigration history controls diversification in experimental adaptive radiation. *Nature* **446:** 436–439.

Fukami, T., Dickie, I.A., Wilkie, J.P., Paulus, B.C., Park, D., Roberts, A., *et al*. (2010) Assembly history dictates ecosystem functioning: evidence from wood decomposer communities. *Ecol Lett* **13:** 675–684.

Green, J., and Bohannan, B.J.M. (2006) Spatial scaling of microbial biodiversity. *Trends Ecol Evol* **21:** 501–507.

Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., and Martiny, J.B.H. (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* **10:** 497–506.

Holyoak, M., Leibold, M.A., and Holt, R.D. (eds) (2005) *Metacommunities: Spatial Dynamics and Ecological Communities*. Chicago, IL, USA: University of Chicago Press.

Horner-Devine, M.C., Lage, M., Hughes, J.B., and Bohannan, B.J.M. (2004) A taxa-area relationship for bacteria. *Nature* **432:** 750–753.

Humbert, J.-F., Dorigo, U., Cecchi, P., Le Berre, B., Debroas, D., and Bouvy, M. (2009) Comparison of the structure and composition of bacterial communities from temperate and tropical freshwater ecosystems. *Environ Microbiol* **11:** 2339–2350.

Jardillier, L., Basset, M., Domaizon, I., Belan, A., Amblard, C., Richardot, M., and Debroas, D. (2004) Bottom-up and top-down control of bacterial community composition in the euphotic zone of a reservoir. *Aquat Microb Ecol* **35:** 259–273.

Jost, L. (2010) The relation between evenness and diversity. *Diversity* **2:** 207–232.

Jurgens, K., and Jeppesen, E. (2000) The impact of metazooplankton on the structure of the microbial food web in a shallow, hypertrophic lake. *J Plankton Res* **22:** 1047–1070.

Jurgens, K., Arndt, H., and Rothhaupt, K.O. (1994) Zooplankton-mediated changes of bacterial community structure. *Microb Ecol* **27:** 27–42.

King, A.J., Freeman, K.R., McCormick, K.F., Lynch, R.C., Lozupone, C., Knight, R., and Schmidt, S.K. (2010) Biogeography and habitat modelling of high-alpine bacteria. *Nat Comm* **1:** 53. doi:10.1038/ncomms1055.

Kosten, S., Huszar, V.L.M., Mazzeo, N., Scheffer, M., Sternberg, L.D.L., and Jeppesen, E. (2009a) Lake and watershed characteristics rather than climate influence

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*

nutrient limitation in shallow lakes. *Ecol Appl* **19:** 1791–1804.

Kosten, S., Lacerot, G., Jeppesen, E., Marques, D.D., van Nes, E.H., Mazzeo, N., and Scheffer, M. (2009b) Effects of submerged vegetation on water clarity across climates. *Ecosystems* **12:** 1117–1129.

Köppen, W. (1936) *Das geographische System der Klimate*. Berlin, Germany: Verlag von Gebrüder Borntraeger.

Langenheder, S., and Jurgens, K. (2001) Regulation of bacterial biomass and community structure by metazoan and protozoan predation. *Limnol Oceanogr* **46:** 121–134.

Langenheder, S., and Ragnarsson, H. (2007) The role of environmental and spatial factors for the composition of aquatic bacterial communities. *Ecology* **88:** 2154–2161.

Langenheder, S., Berga, M., Östman, Ö., and Székely, A.J. (2012) Temporal variation of beta-diversity and assembly mechanisms in a bacterial metacommunity. *ISME J* **6:** 1107–1114.

Leemans, R., and Cramer, W.P. (1991) *The IIASA Database for Mean Monthly Values of Temperature, Precipitation, and Cloudiness on a Global Terrestrial Grid*. Laxenburg, Austria: International Institute fo Applies Systems Analysis.

Leibold, M.A. (1998) Similarity and local co-existence of species in regional biotas. *Evol Ecol* **12:** 95–110.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., *et al.* (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* **7:** 601–613.

Lindström, E.S., and Langenheder, S. (2012) Local and regional factors influencing bacterial community assembly. *Environ Microbiol Rep* **4:** 1–9.

Loeuille, N., and Leibold, M.A. (2008) Evolution in metacommunities: on the relative importance of species sorting and monopolization in structuring communities. *Am Nat* **171:** 788–799.

Logue, J.B., and Lindström, E.S. (2010) Species sorting affects bacterioplankton community composition as determined by 16S rDNA and 16S rRNA fingerprints. *ISME J* **4:** 729–738.

Lymer, D., Lindstrom, E.S., and Vrede, K. (2008) Variable importance of viral-induced bacterial mortality along gradients of trophic status and humic content in lakes. *Freshwater Biol* **53:** 1101–1113.

Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., *et al.* (2006) Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4:** 102–112.

Massana, R., Gasol, J.M., Bjornsen, P.K., Blackburn, N., Hagstrom, A., Hietanen, S., *et al.* (1997) Measurement of bacterial size via image analysis of epifluorescence preparations: description of an inexpensive system and solutions to some of the most common problems. *Sci Mar* **61:** 397–407.

van de Meutter, F., Stoks, R., and De Meester, L. (2005) The effect of turbidity state and microhabitat on macroinvertebrate assemblages: a pilot study of six shallow lakes. *Hydrobiologia* **542:** 379–390.

Muyzer, G., and Smalla, K. (1998) Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie Van Leeuwenhoek* **73:** 127–141.

Ng, H.-T., Marques, D.D.M., Jeppesen, E., and Sondergaard, M. (2010) Bacterioplankton in the littoral and pelagic zones of subtropical shallow lakes. *Hydrobiologia* **646:** 311–326.

Niu, Y., Shen, H., Chen, J., Xie, P., Yang, X., Tao, M., *et al.* (2011) Phytoplankton community succession shaping bacterioplankton community composition in Lake Taihu, China. *Water Res* **45:** 4169–4182.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., *et al.* (2013) vegan: Community Ecology Package. R package version 2.0-10. http://CRAN.R-project.org/package=vegan

Papke, R.T., Ramsing, N.B., Bateson, M.M., and Ward, D.M. (2003) Geographical isolation in hot spring cyanobacteria. *Environ Microbiol* **5:** 650–659.

Peres-Neto, P.R., Legendre, P., Dray, S., and Borcard, D. (2006) Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* **87:** 2614–2625.

Pernthaler, J. (2005) Predation on prokaryotes in the water column and its ecological implications. *Nat Rev Microbiol* **3:** 537–546.

Pommier, T., Canback, B., Riemann, L., Bostrom, K.H., Simu, K., Lundberg, P., *et al.* (2007) Global patterns of diversity and community structure in marine bacterioplankton. *Mol Ecol* **16:** 867–880.

Porter, K.G., and Feig, Y.S. (1980) The use of DAPI for identifying and counting aquatic microflora. *Limnol Oceanogr* **25:** 943–948.

Pradeep Ram, A.S., Palesse, S., Colombet, J., Thouvenot, A., and Sime-Ngando, T. (2014) The relative importance of viral lysis and nanoflagellate grazing for prokaryote mortality in temperate lakes. *Freshwater Biol* **59:** 300–311.

Putt, M., and Stoecker, D.K. (1989) An experimentally determined carbon-volume ratio for marine oligotrichous ciliates from estuarine and coastal waters. *Limnol Oceanogr* **34:** 1097–1103.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ramette, A., and Tiedje, J.M. (2007a) Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem. *Proc Natl Acad Sci USA* **104:** 2761–2766.

Ramette, A., and Tiedje, J.M. (2007b) Biogeography: an emerging cornerstone for understanding prokaryotic diversity, ecology, and evolution. *Microb Ecol* **53:** 197–207.

Redford, A.J., Bowers, R.M., Knight, R., Linhart, Y., and Fierer, N. (2010) The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ Microbiol* **12:** 2885–2893.

Sanders, R.W., Porter, K.G., Bennett, S.J., and Debiase, A.E. (1989) Seasonal patterns of bacterivory by flagellates, ciliates, rotifers, and cladocerans in a freshwater planktonic community. *Limnol Oceanogr* **34:** 673–687.

Schattenhofer, M., Fuchs, B.M., Amann, R., Zubkov, M.V., Tarran, G.A., and Pernthaler, J. (2009) Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ Microbiol* **11:** 2078–2093.

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*

Schiaffino, M.R., Unrein, F., Gasol, J.M., Massana, R., Balague, V., and Izaguirre, I. (2011) Bacterial community structure in a latitudinal gradient of lakes: the roles of spatial versus environmental factors. *Freshwater Biol* **56:** 1973–1991.

Soininen, J. (2012) Macroecology of unicellular organisms – patterns and processes. *Environ Microbiol Rep* **4:** 10–22.

Turcotte, M.M., Corrin, M.S.C., and Johnson, M.T.J. (2012) Adaptive evolution in ecological communities. *PLoS Biol* **10:** e1001332.

Urban, M.C., Leibold, M.A., Amarasekare, P., De Meester, L., Gomulkiewicz, R., Hochberg, M.E., *et al.* (2008) The evolutionary ecology of metacommunities. *Trends Ecol Evol* **23:** 311–317.

Van der Gucht, K., Sabbe, K., De Meester, L., Vloemans, N., Zwart, G., Gillis, M., and Vyverman, W. (2001) Contrasting bacterioplankton community composition and seasonal dynamics in two neighbouring hypertrophic freshwater lakes. *Environ Microbiol* **3:** 680–690.

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.

Verreydt, D., De Meester, L., Decaestecker, E., Villena, M.-J., Van der Gucht, K., Vannormelingen, P., *et al.* (2012) Dispersal-mediated trophic interactions can generate apparent patterns of dispersal limitation in aquatic metacommunities. *Ecol Lett* **15:** 218–226.

Wang, J., Shen, J., Wu, Y., Tu, C., Soininen, J., Stegen, J.C., *et al.* (2013) Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. *ISME J* **7:** 1310–1321.

Weinbauer, M.G., Hornak, K., Jezbera, J., Nedoma, J., Dolan, J.R., and Simek, K. (2007) Synergistic and antagonistic effects of viral lysis and protistan grazing on bacterial biomass, production and diversity. *Environ Microbiol* **9:** 777–788.

Whitaker, R.J., Grogan, D.W., and Taylor, J.W. (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* **301:** 976–978.

Willig, M.R., Kaufman, D.M., and Steven, R.D. (2003) Latitudinal gradients of biodiversity: pattern, process, scale and synthesis. *Annu Rev Ecol Evol Syst* **34:** 273–309.

Wu, Q.L., Zwart, G., Wu, J., Agterveld, M.P.K.-V., Liu, S., and Hahn, M.W. (2007) Submersed macrophytes play a key role in structuring bacterioplankton community composition in the large, shallow, subtropical Taihu Lake, China. *Environ Microbiol* **9:** 2765–2774.

Yannarell, A.C., and Triplett, E.W. (2005) Geographic and environmental sources of variation in lake bacterial community composition. *Appl Environ Microb* **71:** 227–239.

Yoshida, T., Gurung, T.B., Kagami, M., and Urabe, J. (2001) Contrasting effects of a cladoceran (*Daphnia galeata*) and a calanoid copepod (*Eodiaptomus japonicus*) on algal and microbial plankton in a Japanese lake, Lake Biwa. *Oecologia* **129:** 602–610.

Zinger, L., Amaral-Zettler, L.A., Fuhrman, J.A., Horner-Devine, M.C., Huse, S.M., Welch, D.B.M., *et al.* (2011) Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PLoS ONE* **6:** e24570.

Zwart, G., Huismans, R., van Agterveld, M.P., Van de Peer, Y., De Rijk, P., Eenhoorn, H., *et al.* (1998) Divergent members of the bacterial division Verrucomicrobiales in a temperate freshwater lake. *FEMS Microbiol Ecol* **25:** 159–169.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Fig. S1. Geographical location of the lakes sampled in South America for which bacterioplankton community composition has been characterized, with indication of the prevailing climate following the Köppen (1936) classification. Tropical*i* is an isothermal subzone in the tropics, which has a smaller annual temperature range than the tropical zone. See Table S1 for the geographical coordinates.

Fig. S2. Mean abundance of pelagic cladocerans within each of the climatic regions: tropical*i* ($n = 19$), tropical ($n = 19$) and tundra ($n = 10$). Boxes represent ± standard error, whiskers represent ± standard deviation. Different letter codes represent significantly different groups at $P < 0.05$.

Fig. S3. Variation partitioning of the bacterial community composition based on two sets of explanatory variables (environment and space). R^2-adjusted values are given for (A) the three regions analysed together, (B) the regions tropical*i* and tropical together, (C) tropical*i* ($n = 19$), (D) tropical ($n = 19$), (E) tundra ($n = 10$). Asterisks indicate percentages explaining a statistically significant part of the variation (*$P < 0.05$; **$P < 0.01$). Environmental variables selected by forward selection in the environmental model are given below the Venn diagram. Variables that remained significant after correcting for space are given in bold and the significance level is indicated by asterisks (*$P < 0.05$; **$P < 0.01$). Cond = conductivity; DOC = concentration of dissolved organic carbon; Secc = Secchi depth; T° = temperature; Tot P = total phosphorus concentration; NO3 + NO2 = concentration of nitrates and nitrites; Chl a = concentration of chlorophyll a; Clad = abundances of pelagic cladocerans; Cop = abundances of cyclopoid copepods.

Fig. S4. Variation partitioning of the bacterial community composition based on three sets of explanatory variables (abiotic, biotic and spatial variabels). R^2-adjusted values are given for (A) the three regions analysed together, (B) the regions tropical*i* and tropical together, (C) tropical*i* ($n = 19$), (D) tropical ($n = 19$), (E) tundra ($n = 10$). Asterisks indicate percentages explaining a statistically significant part of the variation (.$P < 0.07$, *$P < 0.05$; **$P < 0.01$). Environmental variables selected by forward selection in the abiotic and biotic models are given below the Venn diagram. Variables that remained significant after correcting for space and the other environmental set of variables (biotic or abiotic respectively) are given in bold and the significance level is indicated by asterisks (*$P < 0.05$; **$P < 0.01$). Cond = conductivity; DOC = concentration of dissolved organic carbon; Secc = Secchi depth; T° = temperature; Tot P = total phosphorus concentration; NO3 + NO2 = concentration of nitrates and nitrites; Chl a = concentration of chlorophyll a;

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*

Clad = abundances of pelagic cladocerans; Cop = abundances of cyclopoid copepods.

Table S1. Geographical coordinates of the 48 lakes for which the bacterioplankton community composition was characterized in this study.

Table S2. Discriminating taxa determined using SIMPER analysis. Differences (< and >) in average abundances of taxa contributing to dissimilarities between geographical regions. A cut-off of a cumulative percentage dissimilarity of 90% was applied. 'Unidentified' designates sequences with no matching record, a low percentage identity or a match with a sequence of an unidentified organism based on the BLAST search algorithm.

© 2014 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*