OPEN

The ISME Journal (2013) 7, 1507–1518

www.nature.com/ismej

ORIGINAL ARTICLE

Effects of environmental variation and spatial distance on *Bacteria*, *Archaea* and viruses in sub-polar and arctic waters

Christian Winter[1,2], Blake Matthews[3] and Curtis A Suttle[4,5,6]

[1]*Department of Earth and Ocean Sciences, The University of British Columbia, Vancouver, British Columbia, Canada;* [2]*Department of Marine Biology, University of Vienna, Vienna, Austria;* [3]*Department of Aquatic Ecology, Swiss Federal Institute of Aquatic Science and Technology (EAWAG), Dübendorf, Switzerland;* [4]*Department of Earth, Ocean and Atmospheric Sciences, the Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada;* [5]*Department of Microbiology and Immunology, the Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada and* [6]*Department of Botany, the Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada*

We investigated the influence of environmental parameters and spatial distance on bacterial, archaeal and viral community composition from 13 sites along a 3200-km long voyage from Halifax to Kugluktuk (Canada) through the Labrador Sea, Baffin Bay and the Arctic Archipelago. Variation partitioning was used to disentangle the effects of environmental parameters, spatial distance and spatially correlated environmental parameters on prokaryotic and viral communities. Viral and prokaryotic community composition were related in the Labrador Sea, but were independent of each other in Baffin Bay and the Arctic Archipelago. In oceans, the dominant dispersal mechanism for prokaryotes and viruses is the movement of water masses, thus, dispersal for both groups is passive and similar. Nevertheless, spatial distance explained 7–19% of the variation in viral community composition in the Arctic Archipelago, but was not a significant predictor of bacterial or archaeal community composition in either sampling area, suggesting a decoupling of the processes regulating community composition within these taxonomic groups. According to the metacommunity theory, patterns in bacterial and archaeal community composition suggest a role for species sorting, while patterns of virus community composition are consistent with species sorting in the Labrador Sea and suggest a potential role of mass effects in the Arctic Archipelago. Given that, a specific prokaryotic taxon may be infected by multiple viruses with high reproductive potential, our results suggest that viral community composition was subject to a high turnover relative to prokaryotic community composition in the Arctic Archipelago.

The ISME Journal (2013) **7,** 1507–1518; doi:10.1038/ismej.2013.56; published online 4 April 2013
Subject Category: microbial population and community ecology
Keywords: metacommunity framework; ocean; prokaryotes; viruses; dispersal

Introduction

The sentence 'Everything is everywhere, but, the environment selects' (Baas Becking, 1934) has inspired microbial ecologists for decades. The statement implies that there are no boundaries for microbial dispersal, perhaps, due to their large populations and small size, and that observed patterns are primarily a result of spatial environmental heterogeneity. Today, there is broad consensus that biogeographic patterns exist for microbes, but it is still debated how dispersal interacts with environmental heterogeneity to produce these patterns (Hughes Martiny *et al.*, 2006).

The metacommunity framework of community ecology provides a broad context for exploring the role of dispersal and the local environment (Holyoak *et al.*, 2005), and is increasingly being adopted by microbial ecologists (Logue and Lindström, 2008; Langenheder and Székely, 2011). A metacommunity is defined as a set of local communities with multiple and potentially interacting taxa that are linked by dispersal (Hanski and Gilpin, 1991; Holyoak *et al.*, 2005). The framework presents four scenarios that lead to different ecological interpretations of the effects of environmental and spatial gradients on community composition (Leibold *et al.*, 2004). (1) Species sorting is equivalent to the Baas Becking hypothesis (Baas Becking, 1934), suggesting

Correspondence: C Winter, Department of Marine Biology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria.
E-mail: cwjournals@mac.com
Received 10 September 2012; revised 1 March 2013; accepted 4 March 2013; published online 4 April 2013

that local community composition at a given site is only controlled by environmental variation (for example, nutrient concentration, grazing pressure) and not by the spatial distribution of sites. (2) In the mass effects scenario community composition is influenced by both environmental conditions and spatial distance among sites that are connected by dispersal. In this case, divergent environments can lead to different communities, but dispersal is strong enough between sites that it maintains taxa that are not favored by the local environment. This may lead to taxa occurring in locations in which they are not well adapted and to a poor match between community composition and the local environment. (3) The neutral scenario assumes that taxa have identical niches, and that spatial variation in community composition does not occur because of spatially structured environmental conditions, but because of stochastic processes of immigration, emigration and extinction of taxa at local scales. A neutral scenario can be diagnosed by observing relationships between community composition and spatial distance independent of spatially structured environmental variation. However, such patterns could also arise from unmeasured, spatially structured variation in the environment. (4) The patch dynamics scenario focuses on colonization-extinction dynamics among identical patches that can be occupied or empty, and predicts that community composition will not vary with the environment. Here, we focus on the species sorting and mass effects scenarios, because the assumptions of the neutral and patch dynamics scenarios are unlikely to apply over broad spatial and environmental scales sampled in the current study.

Studies of large-scale patterns of microbial community composition vary in their support for the role of dispersal in structuring microbial communities. For example, Van der Gucht *et al.* (2007) examined lakes separated by 100–2500 km, and found that over a broad range of dispersal rates local environmental parameters explained more of the variance in bacterial community composition than did spatial distance. Furthermore, Soininen *et al.* (2011) have shown that the influence of geographic distance on the community composition of zooplankton, phytoplankton and prokaryotes (prokaryotes here refers to the domains *Bacteria* and *Archaea*, no phylogenetic relationship is implied) increases with increasing size of the organisms. Others argue that environments with identical conditions can harbor different microbial communities because of different founder assemblages (Langenheder *et al.*, 2006).

The ocean is vertically structured with overlaying water masses that may move in different directions and are distinguished by density differences stemming from salinity and temperature differences. These density differences form barriers to mixing and hence, prokaryotic dispersal, and can lead to different prokaryotic communities over small vertical distances (Reinthaler *et al.*, 2006; Agogué *et al.*, 2008; Varela *et al.*, 2008; Galand *et al.*, 2009b, 2010). In contrast to dispersal in terrestrial environments or for motile macroscopic organisms, the dominant dispersal mechanism for prokaryotes and viruses in the ocean is by movement of water masses, although prokaryotes (Cho and Hwang, 2011) and viruses (Snyder *et al.*, 2007) have also been found to disperse in aerosols; hence dispersal for prokaryotes and viruses is passive and similar.

Although viruses are the most abundant biological entities on the planet (Suttle, 2007), few authors have addressed distribution patterns of viruses over broad spatial scales and environmental gradients. Some studies suggest that high dispersal rates of viruses could prevent local adaptation (Breitbart and Rohwer, 2005; Short and Suttle, 2005; Snyder *et al.*, 2007); whereas, other authors find biogeographical patterns for viruses (Angly *et al.*, 2006; Desnues *et al.*, 2008; Kunin *et al.*, 2008; Held and Whitaker, 2009). However, with few exceptions (for example, Held and Whitaker (2009)), these studies ignore host distribution as a critical resource affecting the distribution of viruses, even though the composition of the host community is probably the most relevant underlying factor influencing the composition of a virus community (Rodriguez-Valera *et al.*, 2009; Winter *et al.*, 2010).

Here, we aim at understanding the processes controlling community composition of *Bacteria*, *Archaea* and viruses in marine environments from sub-polar to arctic conditions. Abiotic environmental variation was assessed using water temperature and salinity, and biotic environmental variation was measured as prokaryotic and viral abundance, and the relative abundance of *Bacteria*. The Baas Becking hypothesis (Baas Becking, 1934) suggests that prokaryotic community composition would follow a species sorting scenario, leading to relationships between prokaryotic community composition and environmental variation but not with spatial distance (for example, Winter *et al.* (2008)). As most viruses in the ocean likely infect prokaryotes (for example, Li and Dickie (2001); Baudoux *et al.* (2007); Payet and Suttle (2008)) and given the importance of host community composition for viruses, host community composition is considered to be a biotic environmental parameter. We hypothesized that virus community composition would be explained primarily by variation in prokaryotic community composition. Such a pattern would be equivalent to a species sorting scenario for viruses as prokaryotic community composition is hypothesized to vary independently of spatial distance.

Materials and methods

Study sites and sampling

Thirteen stations were visited during a cruise from Halifax (Nova Scotia, Canada) to Kugluktuk (Nunavut, Canada) aboard *CCGS Louis S. St-Laurent* from 6–24 July 2007 (Figure 1). Water samples were retrieved with Niskin bottles mounted on a rosette

frame that carried sensors for depth, conductivity and temperature (Sea-Bird). The study sites were grouped into three areas, representing a gradient from subpolar to arctic: the Labrador Sea (LS1–LS3; $n=16$), Baffin Bay (BB1–BB5; $n=15$), and the Arctic Archipelago (AA1–AA5; $n=24$; Figure 1). Water samples (10 l) were retrieved from 2–6 depths per station ranging between 5 and 1000 m (Supplementary Figure S1A). We considered all stations to be part of one metacommunity and the three sampling areas (LS, BB and AA; Figure 1) constituted the local communities linked by dispersal.

Prokaryotic and viral abundance

Duplicate 1.8-ml samples were fixed with glutaraldehyde (0.5% final concentration) in the dark for 15 min, flash-frozen in liquid nitrogen and stored at $-80\,^{\circ}C$ until analysis. Upon thawing, the samples were stained with SYBR Green I (Invitrogen, Burlington, ON, Canada). Prokaryotic and viral abundance were determined on a FACSCalibur (BD Biosciences, Mississauga, ON, Canada) flow cytometer (Marie *et al.*, 1999; Brussaard, 2004) and the data analyzed using BD Cell Quest Pro (version 4.0.1; BD Biosciences). The abundance of prokaryotes and viruses is given as the average of duplicate measurements.

Relative abundance of Bacteria determined by catalyzed reporter deposition fluorescence in situ hybridization

Duplicate samples (20–50 ml depending on depth) were fixed with formaldehyde (2% final concentration) and stored at $4\,^{\circ}C$ for up to 24 h. Subsequently, the samples were filtered over 0.22 μm pore-size

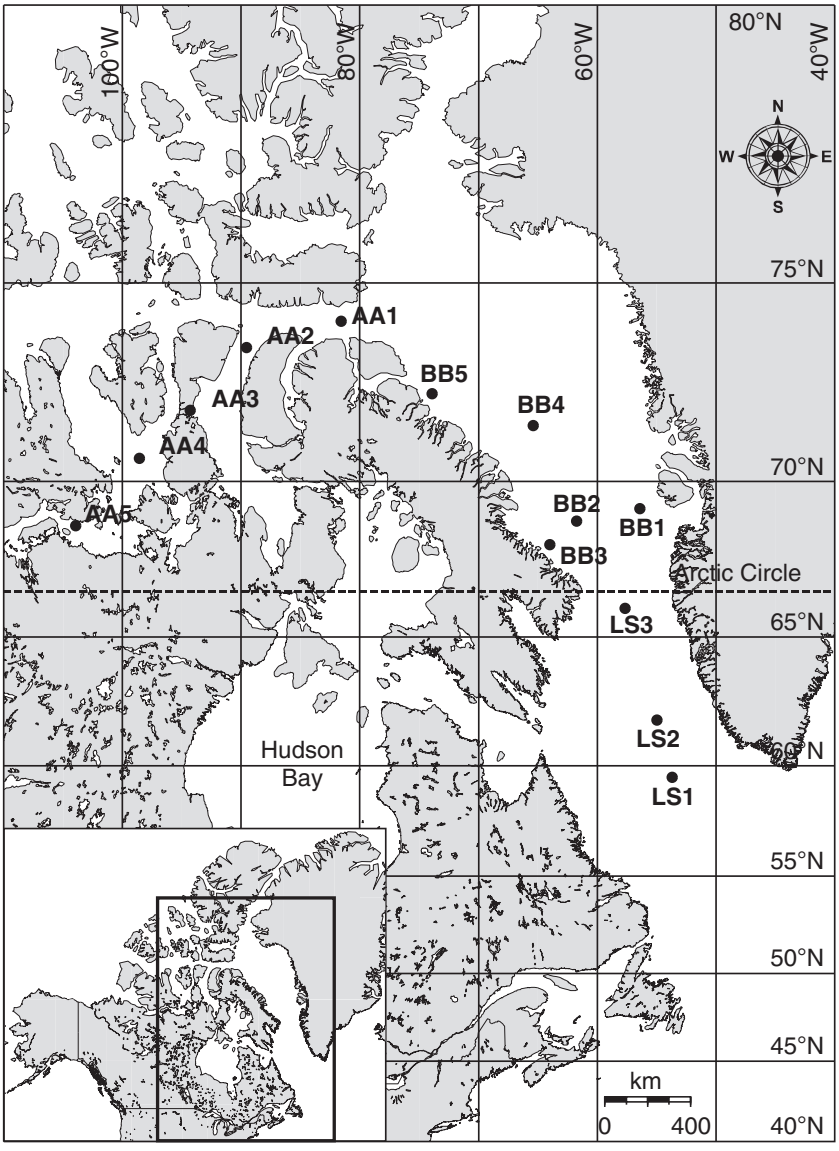

Figure 1 Location of sampling stations. The figure shows the sampling locations in the Labrador Sea (LS1–LS3), Baffin Bay (BB1–BB5) and the Arctic Archipelago (AA1–AA5). The area shown in the main map is marked by the rectangle in the inset.

filters (cyclopore polycarbonate, 25 mm diameter, cat. no. 7060-2502; Whatman) to collect the prokaryotes. The filters were air dried and stored at −20 °C until analysis. Permeabilization, hybridization and mounting of filter slices on slides was done as previously described (Teira *et al.*, 2004). To determine the relative abundance of *Bacteria* we used the nucleic acid probe Eub338 (targeting *Bacteria*; 5′-GCT GCC TCC CGT AGG AGT-3′; Amann *et al.*, 1990) and on selected samples Non338 (antisense probe; 5′-ACT CCT ACG GGA GGC AGC-3′; Wallner *et al.*, 1993) to assess non-specific background hybridization. More than 250 DAPI (4′,6-diamidino-2-phenylindole) stained cells in a minimum of 20 fields of view were counted per sample with an OlympusProvis AX70 microscope (Olympus, Richmond Hill, ON, Canada) equipped with a 100-Watt mercury lamp and appropriate filters for DAPI and Alexa488 (Invitrogen). Data are presented as the average of duplicate measurements.

Assessing bacterial and archaeal community composition using terminal restriction fragment length polymorphism (T-RFLP)

Sampling and nucleic acid extraction. Samples (1 l) were filtered through 0.22 μm pore-size filters (Durapore membrane, GVWP04700, 47 mm diameter; Millipore, Etobicoke, ON, Canada), flash-frozen in liquid nitrogen and stored at −80 °C until analysis. Upon thawing, nucleic acids were extracted using a PowerSoil DNA Isolation kit (cat. no. 12888-100; Mo Bio, Carlsbad, CA, USA) using the alternative lysis protocol involving two heating steps to 70 °C for 5 min. The extracts had a volume of 100 μl in solution S6 (contains no EDTA) and subsamples were used directly in subsequent PCR amplifications.

PCR amplification. The primer pair 27 F (5′-AGA GTT TGA TCC TGG CTC AG-3′) and 1492 R (5′-GGT TAC CTT GTT ACG ACT T-3′) was used to amplify a 1484-bp long fragment (*E. coli* numbering position 27–1511) of the bacterial 16S rRNA gene (Lane, 1991). Primers 21 F (5′-TTC CGG TTG ATC CYG CCG GA-3′) and 958 R (5′-YCC GGC GTT GAM TCC AAT T-3′) were used to amplify a 956-bp long fragment (*E. coli* numbering position 21–977) of the archaeal 16S rRNA gene (DeLong, 1992). The forward primers were fluorescently labeled on the 5′-end with 6-carboxyfluorescein (Thermo Scientific, Waltham, MA, USA) and the reverse primers with the fluorescent dye VIC (Applied Biosystems, Burlington, ON, Canada). The volume of DNA extracts used in the PCR reactions was adjusted to obtain adequate levels of PCR products and varied between 1–3 μl. All PCR chemicals were supplied by Invitrogen and cycling was performed in a Hybaid PCR Express cycler (ESBE Scientific, Markham, ON, Canada). Each 50 μl PCR reaction contained 5 μl of 10 × Taq buffer (200 mM Tris-HCl (pH 8.4), 500 mM KCl), 1.5 μl of a 50 mM $MgCl_2$ solution, 1 μl of 10 mM deoxyribonucleotide triphosphate mix, 2.5 μl of each 10 μM primer solution and 0.25 μl of 5 units μl^{-1} Platinum Taq DNA polymerase (cat. no. 10966-034). The PCR reactions started with an initial denaturation at 95 °C for 5 min followed by 30 cycles with denaturation at 95 °C for 1 min, annealing at 55 °C for 1 min and elongation at 72 °C for 1 min. The final elongation step was performed at 72 °C for 30 min to prevent the formation of artifacts (Janse *et al.*, 2004) followed by a hold at 4 °C. Subsequently, the PCR fragments were cleaned and concentrated using a QIAquick PCR purification kit (cat. No. 28106; Qiagen, Toronto, ON, Canada) according to the manufacturer's instructions resulting in a final volume of 28 μl in elution buffer (Qiagen). Giving absolute concentrations of fluorescently labeled PCR products is not possible due to auto-fluorescence of the products; thus, we used standard agarose gel electrophoresis to size and standardize the concentration of products in the individual reactions in comparison with a molecular mass standard (Low DNA Mass Ladder, cat. no. 10068-013; Invitrogen) to prepare for subsequent restriction digestion and T-RFLP analysis.

Restriction digestion and T-RFLP analysis. Each 15 μl restriction digest contained 1.5 μl of 10 × NEBuffer 4 (200 mM Tris-acetate, 500 mM potassium acetate, 100 mM magnesium acetate, 10 mM Dithiothreitol, pH 7.9), 0.15 μl of 100 μg ml^{-1} bovine serum albumin, 0.5 μl of 20 units μl^{-1} *Hha*I restriction enzyme (cat. no. R0139S; New England Biolabs, Ipswich, MA, USA), and 10 μl of fluorescently labeled PCR products. The digests were incubated in a PCR cycler at 37 °C for 12 h, followed by 65 °C for 20 min to inactivate the restriction enzyme, and a final hold at 4 °C. Five μl of the restriction digests were directly used in subsequent T-RFLP analysis. The samples were analyzed on an ABI 3730 DNA analyzer (Applied Biosystems) equipped with 50 cm long capillaries running POP7 polymer (cat. no. 4363929; Applied Biosystems). The fluorescent signals from the primers of the bacterial and archaeal digests were collected and the peaks sized using Peak Scanner (version number 1.0; Applied Biosystems) against an internal size standard (Genescan 1200 LIZ, cat. no. 4379950; Applied Biosystems). The peak patterns were translated into a binary data matrix (presence vs absence) for further statistical analysis.

Assessing viral community composition using randomly amplified polymorphic DNA PCR (RAPD-PCR)

Sample preparation and nucleic acid extraction. Samples (9 l) were subjected to sequential tangential-flow filtration using a Vivaflow 200 device with either 0.22 μm pore-size or 100 kDa molecular weight cutoff (polyethersulfone, cat. no. VF20P7; Sartorius Stedim Biotech) to remove cells and to concentrate viruses from the filtrate, respectively. The viral concentrates

(30–50 ml) were flash-frozen in liquid nitrogen and stored at −80 °C until analysis. Upon thawing, 15 ml of the viral concentrate was further concentrated to 300–500 µl using an Amicon Ultra-15 centrifugal filtration device (cat. no. UFC903096; Millipore) in a swing-out rotor at 3000 *g* and 4 °C. Subsequently, nucleic acids were extracted from 200 µl of the viral concentrate using a QIAmp MinElute Virus Spin Kit (cat. no. 57704; Qiagen) according to the manufacturer's instructions. The extracts had a volume of 30 µl in AVE buffer (Qiagen).

RAPD-PCR and analysis. Every sample was subjected to two RAPD-PCR reactions with either primer CRA-22 (5′-CCG CAG CCA A-3′) or OPA-13 (5′-CAG CAC CCA C-3′; Neilan 1995). Two microliter of viral DNA extracts yielded adequate levels of PCR products for the RAPD-PCR reactions. PCR chemicals were from Invitrogen and cycling was performed in a HYBAID PCR Express cycler (ESBE Scientific). Each 50 µl PCR reaction contained 5 µl of 10 × Taq buffer (200 mM Tris-HCl (pH 8.4), 500 mM KCl), 1.5 µl of a 50-mM $MgCl_2$ solution, 1 µl of 10 mM deoxyribonucleotide triphosphate mix, 5 µl of a 10-µM solution of either primer CRA-22 or OPA-13 and 0.25 µl of 5 units μl^{-1} Platinum Taq DNA polymerase (cat. no. 10966-034). Initial denaturation was at 94 °C for 10 min followed by 30 cycles with denaturation at 94 °C for 30 s, annealing at 35 °C for 3 min, and elongation at 72 °C for 1 min. The final elongation step was performed at 72 °C for 30 min (Janse *et al.*, 2004) followed by a hold at 4 °C. Subsequently, 5 µl of each RAPD-PCR reaction were loaded onto 2% agarose gels run in 1 × TBE buffer (89 mM Tris-HCl (pH 8.3), 89 mM boric acid, 2 mM EDTA, pH 8.3) at 3 V cm^{-1} electrode distance for 120 min. Gels were stained with SYBR Gold (1:10 000 dilution of stock solution; Invitrogen) for 30 min and electronic images were obtained using an electronic gel documentation system. The band patterns were translated into a binary data matrix for further statistical analysis (Winter and Weinbauer, 2010).

Statistical analysis

Statistical analyses were performed using Mathematica (version 9; Wolfram Research, Oxfordshire, UK). Variation partitioning based on redundancy analysis was used to determine the fraction of variation in bacterial, archaeal and viral community composition that was explained by environmental parameters, spatial distance and spatially correlated environmental parameters (Legendre and Legendre, 1998). Before analysis, binary matrices were transformed by principal coordinates analysis based on Jaccard distances (Legendre and Legendre, 1998). The coordinates of the sampling stations representing spatial distance were subjected to a principal components analysis and abiotic (temperature, salinity) and biotic (prokaryotic and viral abundance, relative abundance of *Bacteria*) environmental parameters were centered on their respective means. Redundancy analysis involves the multivariate linear regression of response variables (for example, principal coordinates of bacterial community composition) on explanatory variables (for example, environmental parameters) and in case of partial redundancy analysis correcting for covariables (for example, spatial distance). The fitted values of the response variables are subsequently transformed using a principal components analysis where the sum of the canonical eigenvalues represents the variation explained by the model. By choosing different combinations of parameters as explanatory variables and covariables we calculated the fractions of variation in prokaryotic and viral community composition due to abiotic and biotic environmental parameters as well as spatial distance. Statistical significance of redundancy and partial redundancy analysis was tested by 10 000 random permutations of residuals under a full model (Legendre and Legendre, 1998).

Partial Mantel tests (Mantel, 1967; Legendre and Legendre, 1998) were used to quantify the relationships of prokaryotic and viral community composition with each other corrected for environmental parameters and spatial distance. For that, distance matrices from the presence-absence matrices of the T-RFLP and RAPD-PCR data were calculated using Jaccard distances. For abiotic and biotic environmental parameters the Euclidean distance was used as calculated from data where each parameter was standardized to a mean of zero and unity variance. Spatial distance was calculated based on the coordinates of the sampling stations using a spheroidal model of Earth and used directly as a distance matrix in the Mantel tests. The Mantel statistic (r_M) was calculated as the Spearman rank correlation coefficient between the upper non-diagonal values of the distance matrices being compared. Here, we further corrected the Mantel statistics for the potential influence of environmental parameters and geographic distance. This was done by calculating the partial Mantel statistic (r_{PM}), equivalent to a first-order partial correlation analysis conducted on three distance matrices (Smouse *et al.*, 1986). Statistical significance of partial Mantel statistics was tested based on 10 000 bootstrap replicates (Legendre and Legendre, 1998; Manly, 2007). Generally, the results of the statistical tests were assumed significant at P-values $\leqslant 0.05$. In case of multiple comparison tests, the α-level indicating significant results was corrected for the number of comparisons using the Bonferroni method (Rohlf and Sokal, 1994).

Results

Variation partitioning of bacterial, archaeal and viral community composition

Detailed results of variation partitioning including testing for statistical significance are found in Supplementary Tables S5–S10. Generally, the results

obtained from the forward primers for *Bacteria* and *Archaea* corresponded well with the corresponding reverse primers (Figure 2). In contrast, the results for viral communities detected by the primers OPA-13 and CRA-22 were more differentiated (Figure 2). The fraction of unexplained variation in bacterial, archaeal and viral community composition tended to increase from the Labrador Sea, to Baffin Bay and the Arctic Archipelago and varied between 67–78% for the entire sampling region (Figure 2).

A significant fraction of the variation in bacterial community composition in the Labrador Sea, Baffin Bay, the Arctic Archipelago, as well as the entire sampling region was explained by environmental parameters (Figure 2). Also, for the entire sampling region, a significant fraction of the variation in bacterial community composition was explained by biotic (prokaryotic and viral abundance, relative abundance of *Bacteria*) environmental parameters alone (Supplementary Table S11). Spatial distance had no statistically significant effect on bacterial community composition in either sampling area or the entire region. Variation of archaeal community composition could not be explained by the available environmental parameters or spatial distance in the Labrador Sea. However, in Baffin Bay, the Arctic

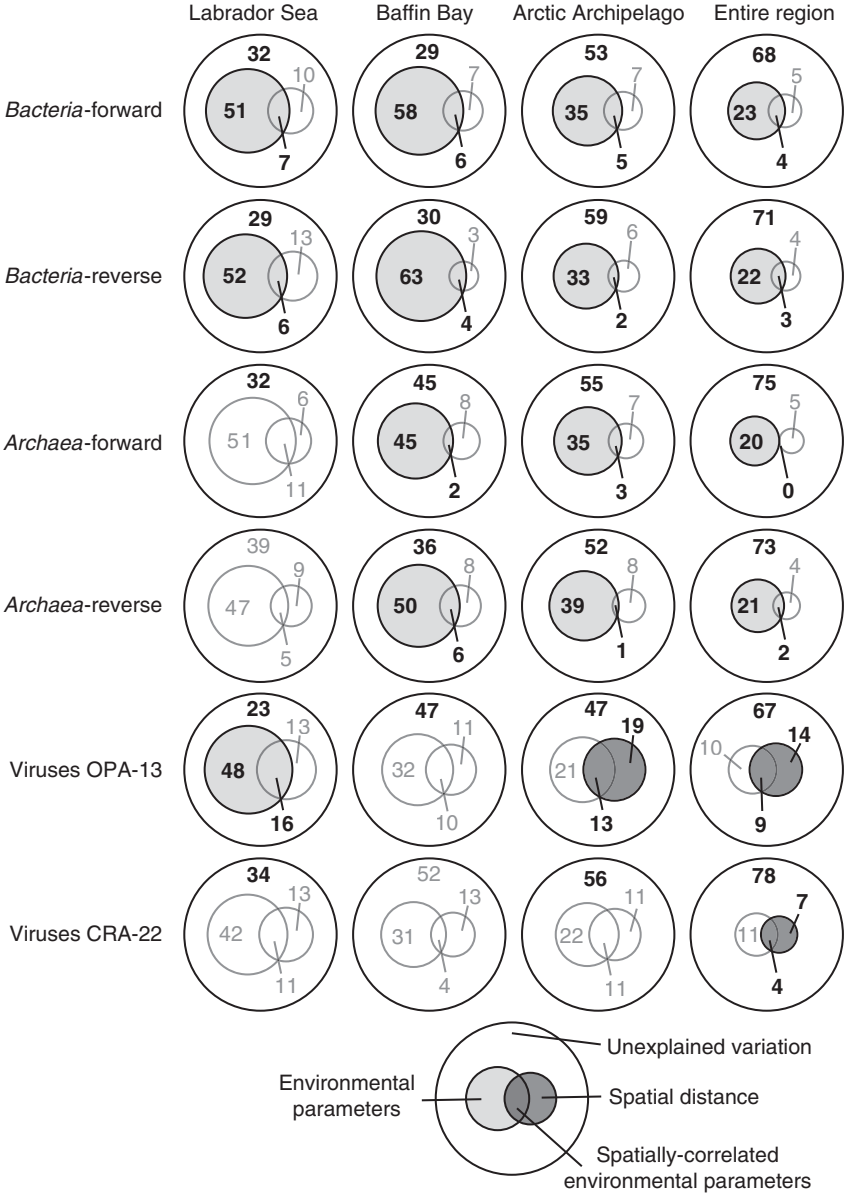

Figure 2 Variation partitioning of bacterial, archaeal and viral community composition. The figure depicts the fractions (%) of variation in bacterial, archaeal and viral community composition that are explained by environmental parameters (temperature, salinity, prokaryotic and viral abundance, relative abundance of *Bacteria*), spatial distance, spatially correlated environmental parameters or remain unexplained for the Labrador Sea, Baffin Bay, the Arctic Archipelago and the entire sampling region in the form of Venn diagrams. The outer circles represent 100% of variation and the surface areas for each fraction correspond to their size given by the numbers. Statistically insignificant values are shown in gray and the circles are not shaded (see Supplementary Tables S5–S10).

Archipelago and the entire sampling region environmental parameters explained a significant fraction of the variation in archaeal community composition (Figure 2), with no detectable effect of spatial distance. Virus community composition in the Labrador Sea obtained by the primer OPA-13 was well explained by environmental parameters but was not affected by spatial distance (Figure 2). Especially, abiotic (temperature, salinity) environmental parameters alone explained a significant fraction of the variation in viral community composition detected by the primer OPA-13 in the Labrador Sea (Supplementary Table S11). In Baffin Bay, viral community composition detected by the primer OPA-13 could not be explained by either environmental parameters or spatial distance. However, in the Arctic Archipelago and for the entire sampling region, viral community composition obtained by the primer OPA-13 was significantly affected by spatial distance but not by environmental parameters (Figure 2). In contrast, neither available environmental parameters nor spatial distance explained any variation in viral community composition obtained by the primer CRA-22 in the Labrador Sea, Baffin Bay or the Arctic Archipelago (Figure 2). However, for the entire region, viral

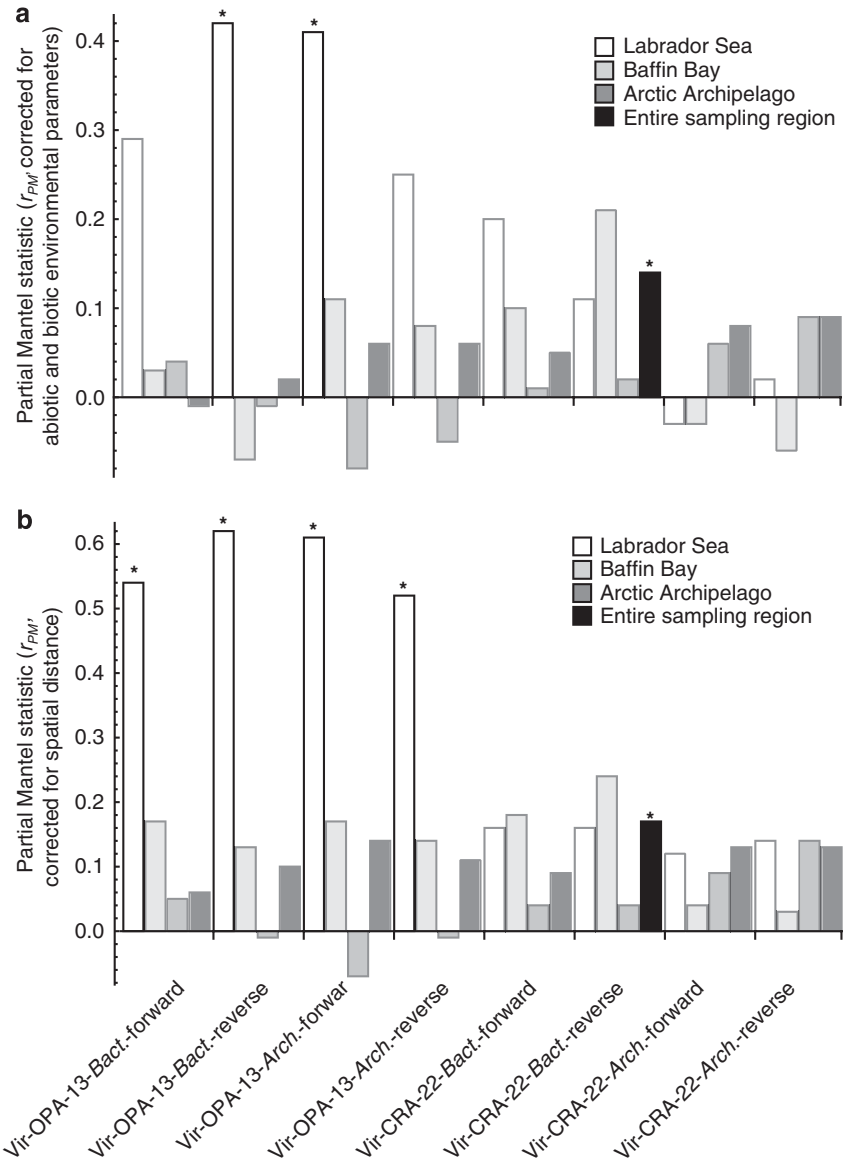

Figure 3 Relationships between viral and prokaryotic community composition. The figure shows the partial Mantel statistics (r_{PM}) calculated between viral community composition (primers OPA-13 and CRA-22) and bacterial as well as archaeal community composition corrected for (**a**) environmental parameters (temperature, salinity, prokaryotic and viral abundance, relative abundance of *Bacteria*) and (**b**) spatial distance. Statistically significant results are marked with '*', statistically insignificant results are grayed out (Bonferroni-corrected: $P \leqslant 0.0125$; Supplementary Tables S3–S4).

community composition obtained by the primer CRA-22 was significantly affected by spatial distance but not by environmental variation (Figure 2).

The fraction of variation that is explained by spatially correlated environmental parameters (Figure 2) cannot be tested for statistical significance, as these results derive from the difference in the variations explained by two different models that do not share a structural relationship (Supplementary Tables S5–S10). However, for *Bacteria* and *Archaea* the fraction of variation explained by spatially correlated environmental parameters was very small compared with the variation explained by environmental parameters (Figure 2; 0–17% of the variation explained by environmental parameters). In contrast, spatially correlated environmental parameters explained a much higher fraction of variation in viral community composition (4–16% of the total variation) compared with the other significant fractions (Figure 2; 33–68% of the variation explained by either environmental parameters or spatial distance).

Relationships between prokaryotic and viral community composition

When correcting for the variation of environmental parameters, partial Mantel tests revealed that changes in viral community composition (OPA-13) were significantly linked to bacterial (reverse primer) and archaeal (forward primer) community composition in the Labrador Sea (Figure 3a). Also, when correcting for spatial distance, changes in viral community composition were significantly linked to bacterial and archaeal community composition in the Labrador Sea (Figure 3b). For the entire sampling region, viral community composition (CRA-22) was significantly linked to bacterial community composition (reverse primer) when correcting for environmental parameters or spatial distance (Figure 3a and b). No further statistically significant relationships between prokaryotic and viral community composition were found (Supplementary Tables S3–S4).

Discussion

Different processes determine prokaryotic and viral communities

The most striking result of this study is the strong contrast between the influence of environmental parameters on prokaryotic community composition versus the effect of spatial distance on viral community composition in the Arctic Archipelago and the entire sampling region (Figure 2). This suggests that different processes influence the community assembly of viruses and prokaryotes. Bacterial community composition was explained mainly by changes in environmental conditions, to a lesser extent by spatially correlated environmental parameters, and not at all by spatial distance (Figure 2). These results are consistent with a species sorting scenario in all three areas and along the entire sampling region, where a significant fraction of the variation was explained solely by biotic environmental parameters (Figure 2, Supplementary Table S11). Similarly, we interpret the results of variation partitioning for archaeal community composition in Baffin Bay, the Arctic Archipelago and the entire sampling area as indicative for species sorting similar to the Baas Becking hypothesis (Baas Becking, 1934). In contrast, viruses detected by the primer OPA-13 were affected by environmental parameters only in the Labrador Sea, where it was possible to attribute a significant influence to abiotic environmental parameters (Figure 2, Supplementary Table S11). In addition, changes in the community composition of viruses detected by the primer OPA-13 were related to changes in bacterial and archaeal community composition in the Labrador Sea (Figure 3), suggesting a species sorting scenario for viruses detected by the primer OPA-13 in this sampling area. However, a mass effects scenario is more likely for virus community composition obtained by the primer OPA-13 in the Arctic Archipelago and for both primers in the entire sampling region, because of the strong influence of spatial distance and a relatively large fraction of variation explained by spatially correlated environmental parameters (Figure 2).

In a recent study, Soininen *et al.* (2011) found that the influence of spatial distance increases with increasing size from prokaryotes to zooplankton, but our results suggest the reverse is true for viruses and prokaryotes (Figure 2). One important difference from our study is that we sampled communities over a water column of up to 1000 m deep and not just from the surface layer, where potentially differential dispersal of prokaryotes and viruses in aerosols (Snyder *et al.*, 2007; Cho and Hwang, 2011) might have an effect.

Intuitively, because virus proliferation is host dependent, and infection is usually species- or even strain-specific (Weinbauer, 2004), changes in host community composition should result in shifts in viral community composition. However, changes in viral and prokaryotic communities were only related to each other in the Labrador Sea, and to a much lesser extent also in the entire sampling region, but not in Baffin Bay and the Arctic Archipelago (Figure 3). These results suggest that in the Labrador Sea lytic infection of *Bacteria* and *Archaea* was a source for these viruses, an interpretation that is consistent with a species sorting scenario for viruses detected by the primer OPA-13.

There are several scenarios that could lead to an uncoupling of the processes regulating community composition of viral and prokaryotic communities. For example, viruses infect hosts other than prokaryotes such as eukaryotic phytoplankton, which could affect the patterns at different locations. Also, it is conceivable that some groups of prokaryotes

that are producing viruses are missed by T-RFLP analysis or that some actively produced viruses are not detected by RAPD-PCR in Baffin Bay or the Arctic Archipelago (Figure 3). As well, population dynamics of prokaryotes and viruses are very different, with a single-infected cell producing tens, or even hundreds of viral progeny; hence, a small subset of the prokaryotic population that is susceptible to viral infection could produce a large portion of the viruses (for example, Winter *et al.* (2005)), thereby affecting the relationship between viral and prokaryotic community composition. Based on our data we cannot unambiguously distinguish among these scenarios, although it seems unlikely that hosts other than prokaryotes were responsible for most of the viral production given that they are the dominant cellular organisms in the ocean, or that most viruses had exceptionally broad host ranges. Nevertheless, passive dispersal of prokaryotes and viruses with horizontally moving water masses seems the dominant dispersal mechanism as aerosols can be excluded as a significant source in our case. Thus, in directly connected marine environments (Figure 1), spatial distance also represents a certain degree of temporal variation, because it takes time for a specific water mass to cover the distance between two sampling stations during which the prokaryotic and viral communities may change. Viruses infecting prokaryotes have a high reproductive potential and a specific prokaryotic taxon may be infected by multiple viruses (Winter *et al.*, 2010). Thus, viral community composition may have a high turnover compared with prokaryotic communities, leading to poor (for example, viruses detected by the primer CRA-22 and data from the bacterial reverse primer for the entire sampling region; Figure 3) or no relationships between prokaryotic and viral community composition. Similar results were obtained recently in an experimental study were virus communities displayed a high turnover over time relative to prokaryotic communities (Declerck *et al.*, 2013). Such a scenario also explains the difference in patterns of prokaryotic and viral community composition associated with spatial distance in the Arctic Archipelago and the entire sampling region (Figure 2).

Methodological considerations

It is also important to consider that prokaryotic and viral community compositions were determined using two different fingerprinting approaches (T-RFLP and RAPD-PCR). As data on prokaryotic and viral community compositions were obtained from the same water samples, differences cannot be explained by sampling biases. Nevertheless, to obtain viral concentrates suitable for RAPD-PCR, several filtration steps are necessary. Winget *et al.* (2005) have shown that tangential-flow filtration can result in low recovery efficiencies of viruses. Viral recovery efficiencies decreased from the Labrador Sea (average: 43.4%, s.d.: 19.1%), to Baffin Bay (average: 18.9%, s.d. 18.4%), and the Arctic Archipelago (average: 5.4%, s.d.: 5.0%) and were similar to previously published data (Winget *et al.*, 2005). It is not known if these differences in recovery efficiencies introduce biases into the viral community composition. However, it is evident that in comparison with metagenomic approaches (for example, Angly *et al.* (2006)), RAPD-PCR will not result in a complete characterization of mixed viral communities and similar considerations hold for T-RFLP analyses of prokaryotic communities. Thus, both fingerprinting approaches only recover a subset of the communities. It follows that even if the losses of viruses due to the filtration procedure would be selective, it is reasonable to assume that the same types of viruses will be lost in every sample, making this a systematic and reproducible error. In fact, a number of studies have shown that RAPD-PCR can be used to routinely determine viral community composition (Winget and Wommack, 2008; Helton and Wommack, 2009; Weinbauer *et al.*, 2009). In fact, recently it was shown that RAPD-PCR fingerprinting has greater sensitivity for detecting change in virioplankton communities than a marker gene approach (Jamindar *et al.*, 2012). Winter and Weinbauer (2010) report that viral fingerprints obtained with OPA-13 and CRA-22 represent two different subsets of the viral community consistent with the differences in the results of variation partitioning for the two communities obtained in this study (Figure 2). Indeed, our data show that in Baffin Bay and the Arctic Archipelago the number of viral bands detected by CRA-22 was higher compared with data from primer OPA-13 (Supplementary Figures S1E and S2E–F, Supplementary Table S1), suggesting a change in viral community composition from lower to higher latitudes. The percentage of the viral community that is targeted by CRA-22 and OPA-13 is unknown. However, Winter and Weinbauer (2010) demonstrated that for viral genomes with amplification targets, 67–74% gave one band, 20–25% gave two bands and 6–8% yielded three or more bands when targeted with CRA-22 and OPA-13. We found that the total number of bands for both primers together did not change among areas, similar to the results for the number of prokaryotic phylotypes detected by T-RFLP analysis (Supplementary information and Supplementary Table S1). Assuming that the percentage of virus types resulting in 1, 2, 3 or more bands in RAPD-PCR among the sampling areas is stable, our data suggest that the percentage of the viral community targeted by the two primers together did not vary among locations. Techniques such as T-RFLP analysis are assumed to detect the most abundant prokaryotes, that is a subset of the prokaryotic populations and, thus, can be considered similar to RAPD-PCR; however, even though massively parallel tag sequencing yields additional

data on rare phylotypes, comparisons between biogeographic patterns obtained from data on rare and abundant phylotypes led to similar conclusions (Galand *et al.*, 2009a; Nemergut *et al.*, 2011). These considerations together with our data support the conclusions that the resolution of RAPD-PCR, in the sense of ability to detect changes, is high enough to detect patterns in community composition (for example, Supplementary Figures S2E–F and S7–S8) and appears to be similar to T-RFLP, as widely contrasting resolutions would not allow to detect relationships between prokaryotic and viral community composition (Figure 3).

Conclusions

As initially hypothesized, the data on bacterial community composition suggest a species sorting scenario in our sampling areas and for the entire sampling region. Archaeal community composition could not be explained by environmental variation or spatial distance in the Labrador Sea. However, the archaeal data for Baffin Bay, the Arctic Archipelago and the entire sampling region are indicative of a species sorting scenario consistent with the Baas Becking hypothesis (Baas Becking, 1934; Figure 2). Viral community composition determined by the primer OPA-13 followed a species sorting scenario in the Labrador Sea and showed patterns characteristic of a mass effects scenario in the Arctic Archipelago and the entire sampling region (Figure 2). Although no significant effects of environmental variation or spatial distance on viruses detected by the primer CRA-22 were found in any of the sampling areas, the data for the entire sampling region suggest a mass effects scenario (Figure 2).

In contrast to our initial hypothesis, viral and prokaryotic community compositions were only related with each other in the Labrador Sea and to a much lesser degree in the entire sampling area (Figure 3). Passive dispersal of prokaryotes and viruses within the moving water masses appears to be the dominant mechanism of dispersal. However, because water takes time to move across large distances, spatial distance in directly connected marine environments has a time dimension during which community composition can change. Hence, taking time into account, and that a specific prokaryotic taxon may be infected by multiple viruses (Winter *et al.*, 2010) with high reproductive potential, our results suggest a higher turnover of viral community composition as compared with prokaryotic community composition in the Arctic Archipelago, and for the sampling region as a whole. Such a scenario would explain the lack of a relationship between viral and prokaryotic community compositions, as well as the contrasting patterns between prokaryotes and viruses associated with distance (Figure 2).

Acknowledgements

We are grateful to Eddy Carmack, Fiona McLaughlin, John Nelson and Sarah Zimmermann for their leadership throughout the program. The captains and crews of *CCGS Louis S. St-Laurent* are greatly acknowledged for their assistance. We thank C. Baranyi for sharing his knowledge about statistics. Many thanks go to S. Declerck, C. Baranyi, T. Reinthaler and G. J. Herndl for their comments on an earlier version of the manuscript. This study was financially supported by the TULA Foundation and the Natural Sciences and Engineering Research Council of Canada. Financial and ship time support was also provided by Fisheries and Oceans Canada, the Canadian International Polar Year Program, Canada's Three Oceans and the University of Vienna. Additional financial support was provided by the Austrian Science Fund (FWF): (P24413-B21) to CW.

References

Agogué H, Brink M, Dinasquet J, Herndl GJ. (2008). Major gradients in putatively nitrifying and non-nitrifying Archaea in the deep North Atlantic. *Nature* **456**: 788–791.

Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol* **56**: 1919–1925.

Angly FE, Felts B, Breitbart M, Salamon P, Edwards RA, Carlson C *et al.* (2006). The marine viromes of four oceanic regions. *PLoS Biol* **4**: e368.

Baas Becking LGM. (1934). *Geobiologie of inleiding tot de milieukunde*. W. P. Van Stockum and Zoon: Den Haag.

Baudoux A-C, Veldhuis MJW, Witte HJ, Brussaard CPD. (2007). Viruses as mortality agents of picophytoplankton in the deep chlorophyll maximum layer during IRONAGES III. *Limnol Oceanogr* **52**: 2519–2529.

Breitbart M, Rohwer F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends Microbiol* **13**: 278–284.

Brussaard CPD. (2004). Optimization of procedures for counting viruses by flow cytometry. *Appl Environ Microbiol* **70**: 1506–1513.

Cho BC, Hwang CY. (2011). Prokaryotic abundance and 16S rRNA gene sequences detected in marine aerosols on the East Sea (Korea). *FEMS Microbiol Ecol* **76**: 327–341.

Declerck SAJ, Winter C, Shurin JB, Suttle CA, Matthews B. (2013). Effects of patch connectivity and heterogeneity on metacommunity structure of planktonic bacteria and viruses. *ISME J* **7**: 533–542.

DeLong EF. (1992). Archaea in coastal marine environments. *Proc Natl Acad Sci USA* **89**: 5685–5689.

Desnues C, Rodriguez-Brito B, Rayhawk S, Kelley S, Tran T, Haynes M *et al.* (2008). Biodiversity and biogeography of phages in modern stromatolites and thrombolites. *Nature* **452**: 340–343.

Galand PE, Casamayor EO, Kirchman DL, Lovejoy C. (2009a). Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc Natl Acad Sci USA* **106**: 22427–22432.

Galand PE, Casamayor EO, Kirchman DL, Potvin M, Lovejoy C. (2009b). Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing. *ISME J* **3**: 860–869.

Galand PE, Potvin M, Casamayor EO, Lovejoy C. (2010). Hydrography shapes bacterial biogeography of the deep Arctic Ocean. *ISME J* **4**: 564–576.

Hanski I, Gilpin M. (1991). Metapopulation dynamics: brief history and conceptual domain. *Biol J Linn Soc* **42**: 3–16.

Held NL, Whitaker RJ. (2009). Viral biogeography revealed by signatures in Sulfolobus islandicus genomes. *Environ Microbiol* **11**: 457–466.

Helton RR, Wommack KE. (2009). Seasonal dynamics and metagenomic characterization of estuarine viriobenthos assemblages by randomly amplified polymorphic DNA PCR. *Appl Environ Microbiol* **75**: 2259–2265.

Holyoak M, Leibold MA, Mouquet NM, Holt RD, Hoopes MF. (2005). Metacommunities: A framework for large-scale community ecology. In: Holyoak M, Leibold MA, Holt RD (ed) *Metacommunities: spatial dynamics and ecological communities*. University of Chicago Press: Chicago, IL, USA, pp 1–31.

Hughes Martiny JB, 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.

Jamindar S, Polson SW, Srinivasiah S, Waidner L, Wommack KE. (2012). Evaluation of two approaches for assessing the genetic similarity of virioplankton populations as defined by genome size. *Appl Environ Microbiol* **78**: 8773–8783.

Janse I, Bok J, Zwart G. (2004). A simple remedy against artificial double bands in denaturing gradient gel electrophoresis. *J Microb Methods* **57**: 279–281.

Kunin V, He S, Warnecke F, Peterson SB, Martin HG, Haynes M *et al.* (2008). A bacterial metapopulation adapts locally to phage predation despite global dispersal. *Genome Res* **18**: 293–297.

Lane DJ. (1991). 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (ed) *Nucleic acid techniques in bacterial systematics*. John Wiley and Sons: Lincoln, pp 115–176.

Langenheder S, Lindström ES, Tranvik LJ. (2006). Structure and function of bacterial communities emerging from different sources under identical conditions. *Appl Environ Microbiol* **72**: 212–220.

Langenheder S, Székely AJ. (2011). Species sorting and neutral processes are both important during the initial assembly of bacterial communities. *ISME J* **5**: 1086–1094.

Legendre P, Legendre L. (1998). *Numerical Ecology*. Elsevier: Amsterdam.

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.

Li WKW, Dickie PM. (2001). Monitoring phytoplankton, bacterioplankton, and virioplankton in a coastal inlet (Bedford Basin) by flow cytometry. *Cytometry* **44**: 236–246.

Logue JB, Lindström ES. (2008). Biogeography of bacterioplankton in inland waters. *Freshwater Rev* **1**: 99–114.

Manly BFJ. (2007). *Randomization, Bootstrap, and Monte Carlo Methods in Biology*. Chapman and Hall: London, UK.

Mantel N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res* **27**: 209–220.

Marie D, Partensky F, Vaulot D, Brussaard CPD. (1999). Enumeration of phytoplankton, bacteria, and viruses in marine samples. In Robinson JP, Darzynkiewicz Z, Dean PN, Orfao A, Rabinovitch PS, Stewart CC, Tanke HJ, Wheeless LL (eds) *Current protocols in cytometry*. John Wiley and Sons: New York, pp 11.11.1–11.11.15.

Neilan BA. (1995). Identification and phylogenetic analysis of toxigenic cyanobacteria by multiplex randomly amplified polymorphic DNA PCR. *Appl Environ Microbiol* **61**: 2286–2291.

Nemergut DR, Costello EK, Hamady M, Lozupone C, Jiang L, Schmidt SK *et al.* (2011). Global patterns in the biogeography of bacterial taxa. *Environ Microbiol* **13**: 135–144.

Payet JP, Suttle CA. (2008). Physical and biological correlates of virus dynamics in the southern Beaufort Sea and Amundsen Gulf. *J Mar Sys* **74**: 933–945.

Reinthaler T, van Aken H, Veth C, Arístegui J, Robson C, PJleB Williams *et al.* (2006). Prokaryotic respiration and production in the meso- and bathypelagic realm of the eastern and western North Atlantic basin. *Limnol Oceanogr* **51**: 1262–1273.

Rodriguez-Valera F, Martin-Cuadrado A-B, Rodriguez-Brito B, Pašić L, Thingstad TF, Rohwer F *et al.* (2009). Explaining microbial population genomics through phage predation. *Nat Rev Microbiol* **7**: 828–836.

Rohlf FJ, Sokal RR. (1994). *Biometry*. W. H. Freeman and Co.: New York.

Short CM, Suttle CA. (2005). Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. *Appl Environ Microbiol* **71**: 480–486.

Smouse PE, Long JC, Sokal RR. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool* **35**: 627–632.

Snyder JC, Wiedenheft B, Lavin M, Roberto FF, Spuhler J, Ortmann AC *et al.* (2007). Virus movement maintains local virus population diversity. *Proc Natl Acad Sci USA* **104**: 19102–19107.

Soininen J, Korhonen JJ, Karhu J, Vetterli A. (2011). Disentangling the spatial patterns in community composition of prokaryotic and eukaryotic lake plankton. *Limnol Oceanogr* **56**: 508–520.

Suttle CA. (2007). Marine viruses–major players in the global ecosystem. *Nature Rev Microbiol* **5**: 801–812.

Teira E, Reinthaler T, Pernthaler A, Pernthaler J, Herndl GJ. (2004). Combining catalyzed reporter deposition-fluorescence in situ hybridization and microautoradiography to detect substrate utilization by Bacteria and Archaea in the deep ocean. *Appl Environ Microbiol* **70**: 4411–4414.

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.

Varela MM, van Aken HM, Herndl GJ. (2008). Abundance and activity of chloroflexi-type SAR202 bacterioplankton in the meso- and bathypelagic waters of the (sub)tropical Atlantic. *Environ Microbiol* **10**: 1903–1911.

Wallner G, Amann R, Beisker W. (1993). Optimizing fluorescent in situ hybridization with rRNA-targeted

oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* **14**: 136–143.
Weinbauer MG. (2004). Ecology of prokaryotic viruses. *FEMS Microbiol Rev* **28**: 127–181.
Weinbauer MG, Arrieta J-M, Griebler C, Herndl GJ. (2009). Enhanced viral production and infection of bacterioplankton during an iron-induced phytoplankton bloom in the Southern Ocean. *Limnol Oceanogr* **54**: 774–784.
Winget DM, Wommack KE. (2008). Randomly amplified polymorphic DNA PCR as a tool for assessment of marine viral richness. *Appl Environ Microbiol* **74**: 2612–2618.
Winget DM, Williamson KE, Helton RR, Wommack KE. (2005). Tangential flow diafiltration: an improved technique for estimation of virioplankton production. *Aquat Microb Ecol* **41**: 221–232.
Winter C, Bouvier T, Weinbauer MG, Thingstad TF. (2010). Trade-offs between competition and defense specialists among unicellular planktonic organisms: the 'killing the Winner' hypothesis revisited. *Microb Mol Biol Rev* **74**: 42–57.
Winter C, Moeseneder MM, Herndl GJ, Weinbauer MG. (2008). Relationship of geographic distance, depth, temperature, and viruses with prokaryotic communities in the eastern tropical Atlantic Ocean. *Microb Ecol* **56**: 383–389.
Winter C, Smit A, Herndl GJ, Weinbauer MG. (2005). Linking bacterial richness with viral abundance and prokaryotic activity. *Limnol Oceanogr* **50**: 968–977.
Winter C, Weinbauer MG. (2010). Randomly amplified polymorphic DNA reveals tight links between viruses and microbes in the bathypelagic zone of the Northwestern Mediterranean Sea. *Appl Environ Microbiol* **76**: 6724–6732.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)