Microbial Ecology and Biogeography

— of the —

Southern Ocean

David Wilkins

Submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy.

School of Biotechnology and Biomolecular Sciences
University of New South Wales, Sydney

March 2013

Contents

List of Figures iii

List of Tables v

List of Acronyms vii

Acknowledgements ix

The advection effect 1
- Summary . . . 1
- Introduction . . . 1
- Methods . . . 2
  - Sampling . . . 2
  - DNA extraction . . . 3
  - Sequencing . . . 4
  - Taxonomic assignment . . . 4
  - Physicochemical and spatial distances . . . 4
  - Generation of advection distance matrix . . . 5
  - Ordination of distance matrices and comparison to water masses . . . 6
  - Testing of advection effect . . . 6
- Results . . . 6
  - Sequencing and taxonomic assignment . . . 6
  - Environment and distance effects . . . 9
- Discussion . . . 11

References 13

List of Figures

1 Map showing sites of samples used in the advection study 2
2 OTU assignments in the advection study. 9
3 nMDS of advective distances between samples. 10
4 nMDS of advective distances between samples. 10
5 dbRDA ordination of relationship between environment and community. 12

List of Tables

1 Full sample data for advection study 7
1 (cont.) Full sample data for advection study. 8
2 Correlations between dbRDA axes and physicochemical variables 11

List of Acronyms

PASSAGE 2 Pattern Analysis, Spatial Statistics and Geographic Exegesis version 2.

QIIME Quantitative Insights Into Microbial Ecology.

AABW Antarctic Bottom Water.

AAIW Antarctic Intermediate Water.

ANOSIM Analysis of Similarities.

AZ Antarctic Zone.

CDW Circumpolar Deep Water.

CMS Connectivity Modeling System.

CTD Conductivity, Temperature and Depth.

dbRDA Distance-based Redundancy Analysis.

DCM Deep Chlorophyll Maximum.

distLM Distance-based Linear Models.

ECCO Estimating the Circulation and Climate of the Ocean.

LCD Lower Circumpolar Deep.

nMDS Non-Metric Multidimensional Scaling.

OTU Operational Taxonomic Unit.

PF Polar Front.

PFZ Polar Frontal Zone.

SAF Subantarctic Front.

SAMW Subantarctic Mode Water.

SO Southern Ocean.

SOSE Southern Ocean State Estimate.

UCD Upper Circumpolar Deep.

Acknowledgements

The advection effect as a driver of microbial biogeography

Sections of this chapter have been previously published in

Summary

Introduction

The central goal of microbial biogeography is to understand how the distribution and abundance of microorganisms are shaped by their physical context. The Baas Becking hypothesis — that "*everything is everywhere*, but, *the environment selects* (Baas Becking, 1934; de Wit and Bouvier, 2006)" — posits that the rapid dispersal of microorganisms means microbial community structure is determined entirely by environmental selection. This stands in contrast to macroorganism biogeography, which has long been recognised as being under the control of historical (in addition to contemporary environmental) factors, particularly spatial influences such as barriers to dispersal. Microbial biogeography studies have begun to show that historical factors may also shape the distribution of microorganisms (Martiny *et al.*, 2006), e.g. a correlation between spatial and genetic distance (a "distance effect") in fluorescent *Pseudomonas* strains in soils (Cho and Tiedje, 2000). This study, among others (Ramette and Tiedje, 2007; Storch and Sizling), also demonstrated the importance of taxonomic resolution in describing such biogeographic patterns. Other studies have found that dispersal potential varies between microbial species, leading to different or absent distance effects (Bissett *et al.*, 2010). When combined with contemporary environmental selection ("environment effect"), distance effects explain some but not all variation between microbial communities, and the mechanism(s) by which a distance effect arises are not always clear (Hanson *et al.*, 2012).

In the ocean, several recent studies have found that microbial communities can be endemic to hydrographically distinct water masses. Surveys in the Arctic (Galand *et al.*, 2009) and North Atlantic (Agogué *et al.*, 2011) oceans have found that bacterial assemblages within the same water mass can be similar across a range of thousands of kilometres, but assemblages can differ between water masses across a range of hundreds of meters. Water masses are defined by their distinct physicochemical properties, so such patterns do not directly imply the existence of factors beyond environmental selection. However, in some cases a water mass-community relationship has been shown to persist even when environment effects are statistically controlled for (Hamilton *et al.*, 2008; Hamdan *et al.*, 2013).

During the study on the biogeographic effect of the Polar Front (PF) described earlier in this volume, it was found that microbial communities in surface waters of the Mertz Glacier region, a site of deep water formation, were very similar to those at the bottom of the water column, despite the very different environmental conditions (see "**??**"). One hypothesis explaining both this observation and water mass endemicity is that microbial assemblages are influenced by the advection (physical transport) of cells by ocean currents. Higher dispersal rates cause the microbial community composition at a given site to increasingly resemble the dispersed colonisers, and less reflect local environmental selection and stochastic effects such as genetic drift (Hanson *et al.*, 2012). Hence, it would be expected that locations that are closely connected by advection (e.g. those within the same water mass, or different levels of the water column at a site of deep water formation) would have more similar compositions than those that are not, even when the environment effect is accounted for. Indeed, advection is often

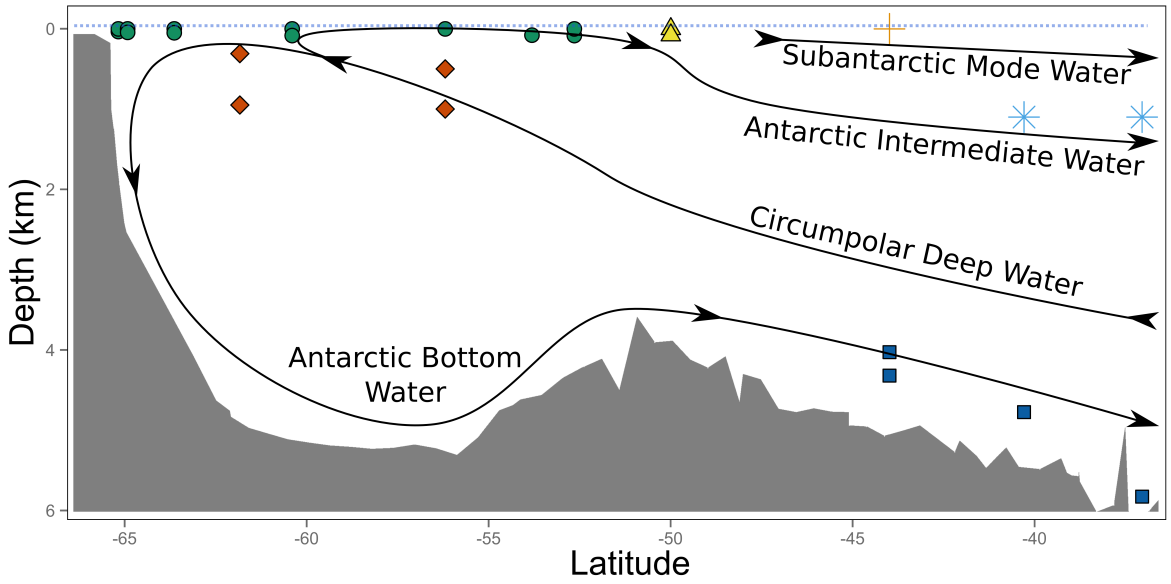

Figure 1: Antarctic Intermediate Waters (AAIW), light blue stars; Subantarctic Mode Water (SAMW), orange crosses; Antarctic Bottom Water (AABW), dark blue squares; Antarctic Zone (AZ), green circles; Polar Frontal Zone (PFZ), yellow triangles; Circumpolar Deep Water (CDW), red diamonds; sea surface, blue dashed horizontal line. Bathymetry is an approximate representation for 115° E, and is indicative only.

invoked to explain observations of microbial diversity or abundance which do not seem attributable to environmental selection (e.g. Sul *et al.*, 2013; Ghiglione *et al.*, 2012; Giebel *et al.*, 2009; Lauro *et al.*, 2007). The exchange of very small volumes of water between marine microbial mesocosms has been found to greatly reduce their β-diversity even under consistent environmental conditions (Declerck *et al.*, 2013). This suggests that advection of even small numbers of cells could have a large homogenizing effect independent of environmental selection. However, the existence of a relationship between advection and community composition that is independent of environment and distance effects has not been directly tested.

The Southern Ocean (SO) is composed of several water masses, which are physicochemically distinct but linked by circulation (see "**??**" for a full description; see also Figure 1). This study aimed to determine whether advection shapes the community structure of bacteria and archaea, independent of environment and distance effects. By sampling each of the SO water masses (depths from surface to ∼6 km), dissimilarity between microbial communities over a large spatial distance (∼3000 km) and range of environments could be determined, in order to test whether advection played a role in shaping their composition.

Methods

Sampling

Sampling[1] was conducted on board the RSV *Aurora Australis* during cruise V3 from January 20th–February 7th 2012. This cruise occupied a latitudinal transect from waters north of Cape Poinsett, Antarctica (65° S) to south of Cape Leeuwin, Australia (37° S) within a longitudinal range of 113–115° E.

Sampling was performed as described in Wilkins *et al.* (2012a), with sites and depths selected to provide coverage of all major SO water masses. At each surface station, ∼250–560 L of seawater was pumped from ∼1.5–2.5 m depth. At some surface stations, an additional sample was taken from the Deep Chlorophyll Maximum (DCM), as determined by chlorophyll fluorescence measurements taken

[1]Sampling was performed by David Wilkins, Timothy J. Williams and Sheree Yau.

from a Conductivity, Temperature and Depth (CTD) cast at each station. Samples of mesopelagic and deeper waters (∼120–240 L) were also collected at some stations using Niskin bottles attached to the CTD. Sampling depths were selected based on temperature, salinity and dissolved oxygen profiles to capture water from the targeted water masses. Profiles were generated on the CTD descent, and samples collected on the ascent at the selected depths. Deep water masses were identified by the following criteria: Circumpolar Deep Water (CDW) = oxygen minimum (Upper Circumpolar Deep (UCD)) or salinity maximum (Lower Circumpolar Deep (LCD)); Antarctic Bottom Water (AABW) = deep potential temperature minimum; Antarctic Intermediate Water (AAIW) = salinity minimum (Foldvik and Gammelsrød, 1988). Surface zones were identified relative to the major fronts of the SO, which are marked by strong latitudinal gradients in temperature and salinity (Sokolov and Rintoul, 2002; Orsi *et al.*, 1995). The Antarctic Zone (AZ) lies south of the PF (∼51° S at the time of sampling), while the Polar Frontal Zone (PFZ) lies between the PF and the Subantarctic Front (SAF). Subantarctic Mode Water (SAMW) overlays AAIW north of the SAF. In total, 25 samples from the AZ, PFZ, SAMW, AAIW, CDW and AABW were collected for this study (fig. 1).

Seawater samples were prefiltered through a 20 µm plankton net, then filtrate was captured on sequential 3.0 µm 0.8 µm and 0.1 µm 293 mm polyethersulfone membrane filters (Pall, Port Washington, USA), and immediately stored at −20 °C (Rusch *et al.*, 2007; Ng *et al.*, 2010).

DNA extraction

DNA extraction was performed using a modified version of the phenol-chloroform method described in Rusch *et al.* (2007). Samples were thawed in a 37 °C water bath. Half of the storage buffer (∼10 mL) was decanted into a clean 50 mL centrifuge tube. If the volume decanted was less than 10 mL, the difference was made with sterile water (Sigma-Aldrich, St. Louis, USA). An equal volume of 50% sucrose lysis buffer (50 mM TRIS-HCl, 40 mM EDTA, 0.75 M Sucrose, pH 8) was added such that the final concentration was 25% sucrose lysis buffer. A small pinch of lysozyme (Sigma-Aldrich, St. Louis, USA) (final concentration ∼2.5 mg/mL) and 1 mL TRIS-EDTA (10 mM TRIS, 1 mM EDTA, pH 8) was added.

The filter membrane was removed from the storage tube and cut in half aseptically. One half was returned to the storage tube, which was refrozen at −80 °C. The remaining half was cut in half again, and one quarter-filter placed atop the other such that the biomass (filtrand) on each piece was facing outwards. Keeping the filters together, they were cut into very fine (∼3 mm by 10 mm) strips, which were placed in the 50 mL centrifuge tube containing the buffer and lysozyme mixture. This tube was mixed by gentle inversion, then tapped such that all filter strips collected at the bottom of the tube and were covered by lysis buffer. The tube was then incubated in a 37 °C shaking water bath at 275 RPM for 30–60 min.

200 µL of 20 mg/mL Proteinase K (Sigma-Aldrich, St. Louis, USA) was added to the tube, which was mixed by gentle inversion. The tube was gently tapped such that all filter strips collected at the bottom covered by lysis buffer. The tube was then subjected to three freeze-thaw cycles, each cycle consisting of 20–30 min in a −80 °C freezer followed by 20–30 min in a 55 °C water bath. After the final complete thaw, 200 µL of 20 mg/mL Proteinase K and 2 mL of 10% SDS (Sigma-Aldrich, St. Louis, USA) were added to the tube. The tube was mixed by gentle inversion then gently tapped such that all filter strips collected at the bottom covered by lysis buffer. It was then incubated in a 55 °C shaking water bath at 175 RPM for two hours.

The supernatant was pipetted from the tube using a genomic tip and split evenly into two new 50 mL centrifuge tubes. An equal volume of buffer-saturated (10 mM TRIS HCl, 1 mM EDTA, pH 8) phenol (Sigma-Aldrich, St. Louis, USA) was added to each of the tubes, which were mixed by gentle inversion. The mixtures were then fractionated in a fixed-angle rotor centrifuge for 15 min at 3700 RPM at room temperature. The bottom layer of each tube was removed by pipette into a new 50 mL centrifuge tube. Each of these two tubes was then made to 50 mL with sterile water (Sigma-Aldrich, St. Louis, USA). After mixing by gentle inversion, each 50 mL mixture was then split evenly into two new 50 mL centrifuge tubes, resulting in four tubes each containing 25 mL of mixture. These tubes were then made to 50 mL with 1-propanol (Sigma-Aldrich, St. Louis, USA). The mixtures were homogenised by gentle inversion and incubated at 4 °C overnight.

Following incubation, the tubes were centrifuged using a fixed-angle rotor for 30 min at 7500 RPM and room temperature. The majority of the supernatant was removed by decanting, and the tubes left

to sit until the remaining supernatant (∼1 mL) collected at the bottom over the precipitated pellet. The pellet was then resuspended by gentle pipetting with a genomic tip, and the suspension placed in a new 1.5 mL microcentrifuge tube (four tubes total). These tubes were then centrifuged in a microcentrifuge for 10 minutes at 13,000 RPM and room temperature. The supernatant was removed by pipette and the tubes placed in a 37 °C heat block with the lids opened and covered by a sterile KimWipe (Kimberly-Clark, Irving, USA) for 10 min, or longer if the supernatant did not evaporate completely in that time. 93.75 µL of TRIS-EDTA was added to each tube, and the tubes were incubated at 4 °C for one hour to allow the DNA pellet to redissolve.

After this incubation, the pellets were gently pipetted with a genomic tip to ensure complete resuspension. The suspensions from all four tubes were combined, and an additional 750 µL of TRIS-EDTA added. This was then split evenly into two new 1.5 mL microcentrifuge tubes (∼562.5 µL per tube).

750 µL of buffer-saturated phenol was added to each tube, and the tubes mixed gently by inversion until a visible emulsion formed. Phase separation was performed by centrifugation for 5 min at 13,000 RPM and room temperature. The upper (aqueous) phase was removed to a new 1.5 mL microcentrifuge tube using a genomic tip.

750 µL of phenol-chloroform-isoamyl alcohol (25:24:1) mixture (Sigma-Aldrich, St. Louis, USA) was added to each tube, and the tubes mixed by gentle inversion until a visible emulsion formed. Phase separation was performed by centrifugation for 5 min at 13,000 RPM and room temperature. The upper (aqueous) phase was removed to a new 1.5 mL microcentrifuge tube using a genomic tip.

75 µL of 3 M sodium acetate (pH 8) and 750 µL of 1-propanol was added to each tube. The tubes were centrifuged at 13,000 RPM and room temperature for 30 min to precipitate the DNA. The supernatant was removed by pipetting, and 100 µL of 70% ethanol added. The tubes were centrifuged again at 13,000 RPM and room temperature for 5 min. The supernatant was removed by pipetting and the DNA pellet dried in a 37 °C heat block. The DNA was dissolved overnight in 40–200 µL of TRIS-EDTA, depending on the expected yield.

Sequencing

Tag pyrosequencing was performed by Research and Testing Laboratory (Lubbock, USA) on a GS FLX+ platform (Roche, Branford, USA), using a modification of the standard 926F/1392R primers targeting the V6–V8 hypervariable regions of bacterial and archaeal 16S rRNA genes (926wF: AAA-CTY-AAA-KGA-ATT-GRC-GG, 1392R: ACG-GGC-GGT-GTG-TRC). The additional wobble bases have been found to amplify a greater range of environmental bacteria and archaea, particularly Euryarchaeota, than the standard primers (Federico M. Lauro, personal communication). Denoising, chimera removal and trimming of poor quality read ends were performed by the sequencing facility.

Taxonomic assignment

Using Quantitative Insights Into Microbial Ecology (QIIME) 1.6.0 (Caporaso *et al.*, 2010), tag pyrosequencing reads were clustered at the 97% sequence similarity level against the SILVA database of rRNA sequences (release 108, eukaryote and chloroplast sequences removed) (Quast *et al.*, 2013), with non-clustering reads discarded. QIIME was used to assign to each Operational Taxonomic Unit (OTU) cluster a description representing the most detailed lineage common to at least 90% of the clustered reads, with ranks labelled "uncultured" or "other" ignored. To generate a taxonomic profile for each sample, the relative abundances of reads assigned to each OTU in each size fraction were encoded as variables. To account for the reads discarded during clustering, abundances in each size fraction were standardised by the proportion of reads retained from that fraction. The abundances were square root transformed and Bray-Curtis dissimilarity indices between samples calculated in PRIMER 6 (PRIMER-E, Lutton, UK).)

Physicochemical and spatial distances

Environmental data were collected from CTD casts at each sample site. Pressure, dissolved oxygen concentration and water temperature measurements were collected with CTD instruments. Salinity and concentrations of dissolved phosphate, nitrate and silicate were obtained from hydrochemical

analysis of seawater samples collected in Niskin bottles during CTD casts (Rosenberg and Rintoul, 2012). These samples were collected at discrete depths, and the hydrochemical sample closest to the depth of the relevant biological sample was selected. The exceptions were samples 32 and 33 (49.5° S, 115° E), for which nitrate concentrations were not available, and sample 29 (53.2° S, 115° E) for which phosphate concentration was not available. In these cases, a reading from the appropriate depth was substituted from the nearest available cast (50.0° S, 115° E for samples 32 and 33; 53.8° S, 115° E for sample 29). Pressure values were $log(x + 1)$ transformed to reduce right-skew (Clarke and Gorley, 2006) and the combined instrument and hydrochemical data were used to create environmental profiles for each sample. The variables were normalised and a Euclidean distance matrix generated in PRIMER 6.

Distance-based Linear Models (distLM) multivariate analysis (Legendre and Anderson, 1999) was performed to confirm the selection of physicochemical variables and explore their relationship with taxonomic composition. In PRIMER 6, all possible combinations of variables ("BEST selection") were explored by distLM, and the models (sets of variables) that best fit the taxonomic dissimilarity matrix (adjusted R^2 as the fitness measure) were selected. The relationship between the resulting model and the taxonomic dissimilarity between samples was visualised by Distance-based Redundancy Analysis (dbRDA) ordination.

To generate a spatial distance matrix, pairwise ellipsoidal distances between samples (including difference in depth) were calculated using INVERS3D (National Oceanic and Atmospheric Administration, Silver Springs, USA).

Generation of advection distance matrix

Erik van Sebille contributed to writing the following paragraph.

Advection distances between the sites were computed using three-dimensional velocity data from a hydrodynamic numerical ocean model in combination with a Lagrangian trajectory toolset[2]. The ocean model used was the Southern Ocean State Estimate (SOSE) (Mazloff *et al.*, 2010), a numerical model of the SO based on the Estimating the Circulation and Climate of the Ocean (ECCO) machinery (Wunsch and Heimbach, 2007) and constrained by a large set of in situ and remote-sensed observations. SOSE has been validated in the SO (Cerovečki *et al.*, 2011; Firing *et al.*, 2011). Here, the five-day averaged three-dimensional velocity fields for the period January 2005–December 2007 were used, on a 1/6° horizontal resolution and with 42 vertical levels. The Connectivity Modeling System (CMS) (Paris *et al.*, 2013) 1.1 was used to integrate virtual Lagrangian particles within the SOSE velocity fields. For each site, 100 particles were released every 5 days (total of 2.2×10^4 per site). The particles were released at the latitude and depth of the site, evenly spaced in a 1° longitudinal line centred at the site longitudes. The particles were then advected for 100 years, looping through the 3 years of available velocity fields as described in van Sebille *et al.* (2012). Three-dimensional locations of the particles were saved every 5 days.

The trajectory of each particle was analysed to detect encounters between particles and sample sites. An encounter was defined as the vector between any two consecutive 5 day particle locations intersecting a box bounded by ±0.2° of latitude, ±0.5° of longitude and ±50 m of depth from a sample site. Only the first encounter between any particle and sample was counted. Four pairs of samples (10/11; 12/13; 16/17; 21/22), where a DCM sample was taken directly below a surface sample within the mixed layer, were too close to act as separate particle release sites. For these samples, simulated particle releases were performed for only one of the pair, and the generated encounters were attributed to both. For all samples, the mean time in seconds between a particle being released from one sample and encountering another was calculated. Pairwise advective distance between samples was defined as the mean of the two directional mean times between each sample in the pair. This metric was selected to ensure advective flows which may be of high biological relevance, such as a small number of particles quickly transported between sites, were appropriately weighted when paired with flows of lower biological relevance, such as a large number of particles transported between sites over decades. To ensure the results were robust to the choice of metric, subsequent statistical tests were repeated with pairwise advective distance redefined as the mean time for all pairwise encounters. The pairwise distance between the surface/DCM samples discussed above was set to zero. For pairs of samples that

[2]Modelling was performed by Erik van Sebille.

did not yield mutual encounters (47 pairs, all including at least one AABW sample; see Results), this was set to the maximum run time of the simulation (100 years). Subsequent statistical tests were rerun with the DCM and AABW samples excluded to ensure these constraints were not unduly influencing the results (see "Testing of advection effect", below).

Ordination of distance matrices and comparison to water masses

Ordinations of the taxonomic, environmental and advection distance matrices were produced by Non-Metric Multidimensional Scaling (nMDS) in PRIMER 6. Analysis of Similarities (ANOSIM) was also performed in PRIMER 6 to test for statistically significant differences between water masses in each of these three factors. Right-tailed p-values for each test were computed using 999 random label permutations of one of the test matrices.

Testing of advection effect

Mantel tests were performed using Pattern Analysis, Spatial Statistics and Geographic Exegesis version 2 (PASSAGE 2) (Rosenberg and Anderson, 2011). To test for and quantify distance and environment effects, partial Mantel tests were performed comparing the taxonomic matrix to the spatial then environmental matrices, with the remaining matrix held constant. To test the hypothesis that advection shapes SO microbial assemblages independent of distance and environment effects, a partial Mantel test was performed comparing the taxonomic and advection matrices, with both the spatial and environmental matrices held constant. Right-tailed p-values for all tests were calculated using 999 random label permutations of one the test matrices.

To ensure that the result was not unduly influenced by the samples to which the 100 year ceiling was applied (all AABW, see Results) and those for which particle releases were not simulated (samples 11, 13, 17 and 22), the test was repeated with these samples removed.

To confirm that the advection effect was directional, i.e. that "upstream" sites were acting as sources of diversity to "downstream" sites, SOURCETRACKER (Knights *et al.*, 2011) was used to identify sources of OTUs in each sample. Each sample was sequentially designated a sink, with the remaining samples as potential sources, and the most probable proportion of OTUs originating from each potential source determined over 100 randomised trials per sample. Spearman's rank correlation was then calculated between the SOURCETRACKER predicted source proportions and particle encounter source proportions for each sample pairwise, with right-tailed p-value determined by permutation.

Results

Sequencing and taxonomic assignment

After trimming, denoising and chimera removal, the 25 samples (each with three separately sequenced size fractions) yielded 1,008,963 pyrosequencing reads of length 251–561 bp (mean 426 bp). Individual fractions yielded 3,687–52,192 reads (mean 13,453). After clustering against the SILVA database, 2,295–30,760 (mean 9,618) reads per fraction were retained for taxonomic assignment.

1417 unique OTUs were identified across all fractions of all samples. The Chao 1 statistic was calculated, and estimated OTU were under-sequenced by 0–50% across all samples (mean 26%) (table 1). In all water masses, decreasing numbers of reads yielded OTU assignments as size fraction increased (fig. 2). This probably reflects an increasing number of eukaryotic cells (which were excluded from OTU assignment) on the higher fractions.

Table 1: Full location, summary of taxonomic assignments and full physicochemical data for each of the 25 samples in this study. Units are given in column headers. Water mass abbreviations are Antarctic Intermediate Waters (AAIW); Subantarctic Mode Water (SAMW); Antarctic Bottom Water (AABW); Antarctic Zone (AZ); Polar Frontal Zone (PFZ); Circumpolar Deep Water (CDW).

Sample	Fraction (µm)	Date	Latitude (°)	Longitude (°)	Depth (m)	Water mass	OTU count	Chao 1	Pressure (dbar)	Dissolved O_2 (µmol/L)	Temperature) (°C)	Phosphate (µmol/L)	Nitrate (µmol/L)	Silicate (µmol/L)	Salinity (PSU)
10	0.1	2012-01-20	−65.17	113.1	35	AZ	431	570	36	348.1	−1.125	1.82	26.73	60.6	33.6
10	0.8	2012-01-20	−65.17	113.1	35	AZ	444	593	36	348.1	−1.125	1.82	26.73	60.6	33.6
10	3.0	2012-01-20	−65.17	113.1	35	AZ	148	293	36	348.1	−1.125	1.82	26.73	60.6	33.6
11	0.1	2012-01-20	−65.17	113.1	2	AZ	433	655	2	365.7	−0.798	1.66	25.13	57.8	33.4
11	0.8	2012-01-20	−65.17	113.1	2	AZ	241	422	2	365.7	−0.798	1.66	25.13	57.8	33.4
11	3.0	2012-01-20	−65.17	113.1	2	AZ	145	202	2	365.7	−0.798	1.66	25.13	57.8	33.4
12	0.1	2012-01-21	−64.92	113.3	2	AZ	294	387	2	356.4	−0.358	1.70	25.87	55.0	33.7
12	0.8	2012-01-21	−64.92	113.3	2	AZ	68	79	2	356.4	−0.358	1.70	25.87	55.0	33.7
12	3.0	2012-01-21	−64.92	113.3	2	AZ	158	232	2	356.4	−0.358	1.70	25.87	55.0	33.7
13	0.1	2012-01-21	−64.92	113.3	45	AZ	259	363	46	314.9	−1.604	2.05	30.17	65.8	34.2
13	0.8	2012-01-21	−64.92	113.3	45	AZ	299	418	46	314.9	−1.604	2.05	30.17	65.8	34.2
13	3.0	2012-01-21	−64.92	113.3	45	AZ	115	160	46	314.9	−1.604	2.05	30.17	65.8	34.2
16	0.1	2012-01-22	−63.64	113.3	2	AZ	245	338	2	355.9	0.500	1.62	24.60	47.1	33.8
16	0.8	2012-01-22	−63.64	113.3	2	AZ	210	307	2	355.9	0.500	1.62	24.60	47.1	33.8
16	3.0	2012-01-22	−63.64	113.3	2	AZ	135	222	2	355.9	0.500	1.62	24.60	47.1	33.8
17	0.1	2012-01-22	−63.64	113.3	50	AZ	169	219	50	319.1	−1.352	2.06	29.69	63.3	34.2
17	0.8	2012-01-22	−63.64	113.3	50	AZ	227	315	50	319.1	−1.352	2.06	29.69	63.3	34.2
17	3.0	2012-01-22	−63.64	113.3	50	AZ	199	290	50	319.1	−1.352	2.06	29.69	63.3	34.2
18	0.1	2012-01-24	−61.84	113.5	310	CDW	474	666	314	186.8	1.909	2.35	34.09	83.2	34.6
18	0.8	2012-01-24	−61.84	113.5	310	CDW	466	676	314	186.8	1.909	2.35	34.09	83.2	34.6
18	3.0	2012-01-24	−61.84	113.5	310	CDW	136	172	314	186.8	1.909	2.35	34.09	83.2	34.6
19	0.1	2012-01-24	−61.84	113.5	950	CDW	138	205	962	202.4	1.624	2.18	31.59	95.1	34.7
19	0.8	2012-01-24	−61.84	113.5	950	CDW	124	153	962	202.4	1.624	2.18	31.59	95.1	34.7
19	3.0	2012-01-24	−61.84	113.5	950	CDW	214	315	962	202.4	1.624	2.18	31.59	95.1	34.7
21	0.1	2012-01-25	−60.40	115.0	2	AZ	287	385	2	335.8	2.462	1.75	26.62	16.2	33.9
21	0.8	2012-01-25	−60.40	115.0	2	AZ	217	296	2	335.8	2.462	1.75	26.62	16.2	33.9
21	3.0	2012-01-25	−60.40	115.0	2	AZ	153	240	2	335.8	2.462	1.75	26.62	16.2	33.9
22	0.1	2012-01-25	−60.40	115.0	85	AZ	318	458	86	336.4	1.724	1.96	28.52	24.7	33.9
22	0.8	2012-01-25	−60.40	115.0	85	AZ	298	418	86	336.4	1.724	1.96	28.52	24.7	33.9
22	3.0	2012-01-25	−60.40	115.0	85	AZ	323	420	86	336.4	1.724	1.96	28.52	24.7	33.9
25	0.1	2012-01-27	−56.19	115.0	500	CDW	297	423	506	187.7	2.296	2.39	35.09	72.9	34.5
25	0.8	2012-01-27	−56.19	115.0	500	CDW	595	736	506	187.7	2.296	2.39	35.09	72.9	34.5
25	3.0	2012-01-27	−56.19	115.0	500	CDW	307	396	506	187.7	2.296	2.39	35.09	72.9	34.5
26	0.1	2012-01-27	−56.19	115.0	1000	CDW	257	316	1012	190.1	2.107	2.23	32.90	80.7	34.7
26	0.8	2012-01-27	−56.19	115.0	1000	CDW	438	510	1012	190.1	2.107	2.23	32.90	80.7	34.7
26	3.0	2012-01-27	−56.19	115.0	1000	CDW	380	584	1012	190.1	2.107	2.23	32.90	80.7	34.7
27	0.1	2012-01-27	−56.19	115.0	2	AZ	368	491	2	324.8	4.159	1.64	25.32	9.60	33.8
27	0.8	2012-01-27	−56.19	115.0	2	AZ	318	454	2	324.8	4.159	1.64	25.32	9.60	33.8
27	3.0	2012-01-27	−56.19	115.0	2	AZ	288	394	2	324.8	4.159	1.64	25.32	9.60	33.8

Continued on following page.

Table 1: (cont.) Full sample data for advection study.

Sample	Fraction (µm)	Date	Latitude (°)	Longitude (°)	Depth (m)	Water mass	OTU count	Chao 1	Pressure (dbar)	Dissolved O_2 (µmol/L)	Temperature) (°C)	Phosphate (µmol/L)	Nitrate (µmol/L)	Silicate (µmol/L)	Salinity (PSU)
29	0.1	2012-01-28	−53.81	115.0	80	AZ	281	394	80	324.4	4.399	1.66	24.78	7.80	33.8
29	0.8	2012-01-28	−53.81	115.0	80	AZ	261	377	80	324.4	4.399	1.66	24.78	7.80	33.8
29	3.0	2012-01-28	−53.81	115.0	80	AZ	271	403	80	324.4	4.399	1.66	24.78	7.80	33.8
30	0.1	2012-01-29	−52.65	115.0	85	AZ	217	341	86	330.8	3.517	1.73	26.27	15.7	33.8
30	0.8	2012-01-29	−52.65	115.0	85	AZ	244	358	86	330.8	3.517	1.73	26.27	15.7	33.8
30	3.0	2012-01-29	−52.65	115.0	85	AZ	269	393	86	330.8	3.517	1.73	26.27	15.7	33.8
31	0.1	2012-01-29	−52.65	115.0	2	AZ	392	486	2	332.2	3.941	1.70	26.15	15.4	33.8
31	0.8	2012-01-29	−52.65	115.0	2	AZ	317	470	2	332.2	3.941	1.70	26.15	15.4	33.8
31	3.0	2012-01-29	−52.65	115.0	2	AZ	332	425	2	332.2	3.941	1.70	26.15	15.4	33.8
32	0.1	2012-01-30	−49.99	115.0	2	PFZ	297	407	2	306.9	7.412	1.40	22.74	4.00	33.9
32	0.8	2012-01-30	−49.99	115.0	2	PFZ	269	337	2	306.9	7.412	1.40	22.74	4.00	33.9
32	3.0	2012-01-30	−49.99	115.0	2	PFZ	225	304	2	306.9	7.412	1.40	22.74	4.00	33.9
33	0.1	2012-01-30	−49.99	115.0	80	PFZ	325	422	80	300.3	7.061	1.39	23.02	4.20	34.0
33	0.8	2012-01-30	−49.99	115.0	80	PFZ	352	547	80	300.3	7.061	1.39	23.02	4.20	34.0
33	3.0	2012-01-30	−49.99	115.0	80	PFZ	326	473	80	300.3	7.061	1.39	23.02	4.20	34.0
38	0.1	2012-02-03	−43.99	115.0	2	SAMW	383	494	2	279.1	13.02	0.59	4.540	1.40	34.7
38	0.8	2012-02-03	−43.99	115.0	2	SAMW	215	220	2	279.1	13.02	0.59	4.540	1.40	34.7
38	3.0	2012-02-03	−43.99	115.0	2	SAMW	177	181	2	279.1	13.02	0.59	4.540	1.40	34.7
39	0.1	2012-02-03	−43.99	115.0	4320	AABW	520	682	4400	217.5	0.8497	2.30	32.92	127	34.7
39	0.8	2012-02-03	−43.99	115.0	4320	AABW	158	215	4400	217.5	0.8497	2.30	32.92	127	34.7
39	3.0	2012-02-03	−43.99	115.0	4320	AABW	129	152	4400	217.5	0.8497	2.30	32.92	127	34.7
40	0.1	2012-02-03	−43.99	115.0	4028	AABW	77	102	4100	216.9	0.8503	2.29	32.94	126	34.7
40	0.8	2012-02-03	−43.99	115.0	4028	AABW	156	203	4100	216.9	0.8503	2.29	32.94	126	34.7
40	3.0	2012-02-03	−43.99	115.0	4028	AABW	42	48	4100	216.9	0.8503	2.29	32.94	126	34.7
44	0.1	2012-02-05	−40.29	115.0	4775	AABW	248	303	4866	217.9	0.8716	2.28	33.15	130	34.7
44	0.8	2012-02-05	−40.29	115.0	4775	AABW	405	497	4866	217.9	0.8716	2.28	33.15	130	34.7
44	3.0	2012-02-05	−40.29	115.0	4775	AABW	370	442	4866	217.9	0.8716	2.28	33.15	130	34.7
45	0.1	2012-02-05	−40.29	115.0	1100	AAIW	601	779	1112	199.4	4.321	2.12	31.29	33.9	34.4
45	0.8	2012-02-05	−40.29	115.0	1100	AAIW	150	151	1112	199.4	4.321	2.12	31.29	33.9	34.4
45	3.0	2012-02-05	−40.29	115.0	1100	AAIW	118	124	1112	199.4	4.321	2.12	31.29	33.9	34.4
46	0.1	2012-02-07	−37.05	115.0	1100	AAIW	166	176	1112	201.7	5.233	2.20	32.12	39.5	34.4
46	0.8	2012-02-07	−37.05	115.0	1100	AAIW	227	312	1112	201.7	5.233	2.20	32.12	39.5	34.4
46	3.0	2012-02-07	−37.05	115.0	1100	AAIW	251	360	1112	201.7	5.233	2.20	32.12	39.5	34.4
47	0.1	2012-02-07	−37.05	115.0	5827	AABW	138	248	5952	217.2	1.030	2.29	33.00	129	34.7
47	0.8	2012-02-07	−37.05	115.0	5827	AABW	379	509	5952	217.2	1.030	2.29	33.00	129	34.7
47	3.0	2012-02-07	−37.05	115.0	5827	AABW	106	123	5952	217.2	1.030	2.29	33.00	129	34.7

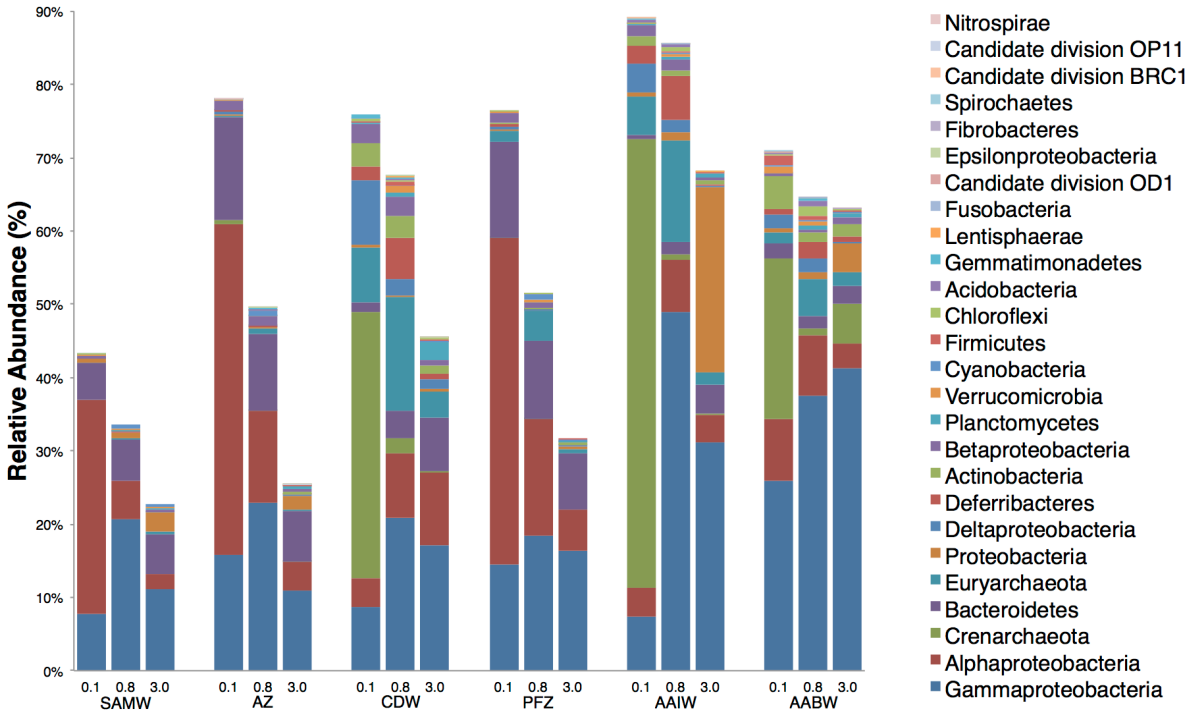

Figure 2: Taxonomic assignments for each sampled water mass. Water masses (x-axis): Subantarctic Mode Water (SAMW); Antarctic Zone (AZ); Circumpolar Deep Water (CDW); Polar Frontal Zone (PFZ); Antarctic Intermediate Water (AAIW); Antarctic Bottom Water (AABW). Size fractions are given in µm. All OTUs aggregated to phylum, except for members of the Proteobacteria which were aggregated to class when known. Relative abundance is percentage of all reads assigned to a given taxonomic group and has been scaled to account for unassigned reads.

nMDS ordination showed that the sampled water masses could be distinguished on the basis of taxonomic distance (fig. 3). This was supported by ANOSIM analysis ($R = 0.77$, $p = 0.001$). While each water mass had a distinct taxonomic profile, some broad differences between surface and deep masses were observed (fig. 2). Surface waters (AZ, PFZ, SAMW) were dominated by representatives of the Alphaproteobacteria, Bacteroidetes and Gammaproteobacteria. The high abundance of Bacteroidetes at the surface reflects their association with phytoplankton, as many species in this lineage specialise in the degradation of high molecular weight products of primary production (Williams *et al.*, 2012). Alphaproteobacteria were represented primarily by the SAR11 clade, abundant in ocean surface communities (Morris *et al.*, 2002) including the SO (Brown *et al.*, 2012), and Roseobacter clades, which have also been associated with degradation of phytoplankton products (Williams *et al.*, 2012; Giebel *et al.*, 2009). The dominant Gammaproteobacterial orders were the Alteromonadales and Oceanospirillales, typical of SO surface waters (Wilkins *et al.*, 2012b). Few archaeal OTUs were detected, consistent with their well-described decline in abundance during summer (Murray *et al.*, 1998; Grzymski *et al.*, 2012). The deep water masses (CDW, AAIW, AABW) were dominated by Crenarchaeota, Euryarchaeota and Gammaproteobacteria, again consistent with previous findings (López-García *et al.*, 2001).

Environment and distance effects

nMDS ordination showed that the sampled water masses clustered well on the basis of environmental distance (fig. 4). This was supported by ANOSIM ($R = 0.84$, $p = 0.001$). A partial Mantel test, comparing the taxonomic to environmental matrices with the spatial matrix held constant, found a correlation of $r = 0.48$ ($p = 0.001$), indicating a strong environment effect.

distLM analysis of the individual physicochemical variables found that considered separately, each of phosphate, silicate, nitrate, oxygen, salinity and pressure explained 17–35% of the taxonomic variance between samples ($p = 0.001$). Temperature had no significant effect on taxonomic composition when considered separately ($p > 0.05$). When all combinations of variables were considered (BEST

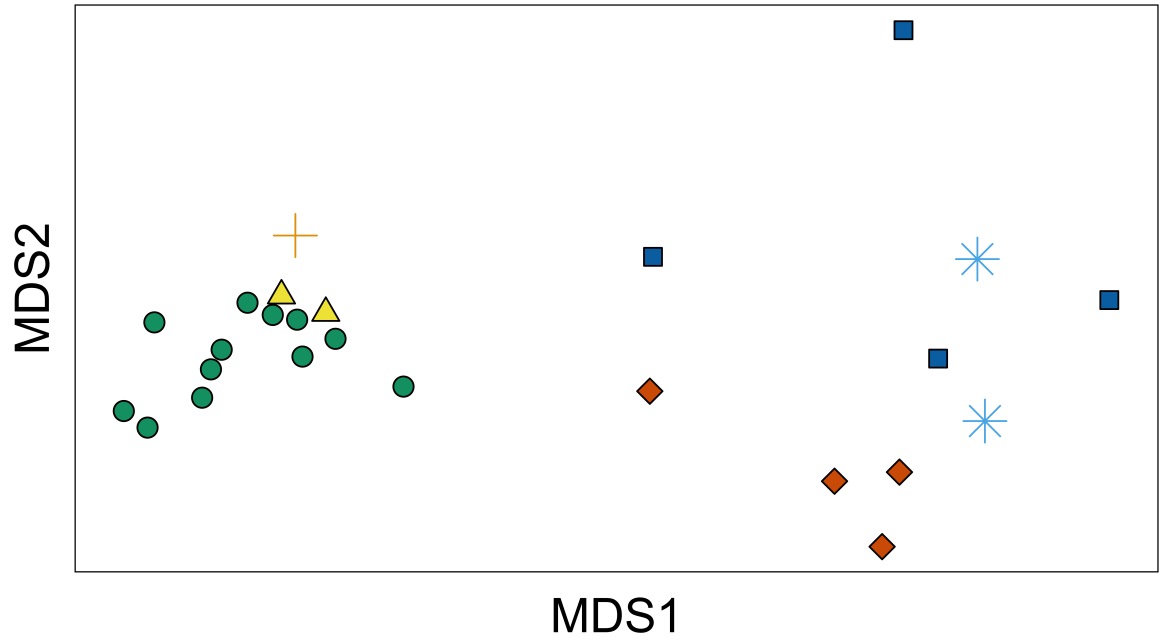

Figure 3: nMDS ordination of the taxonomic distance matrix (2D stress = 0.08). Antarctic Intermediate Waters (AAIW), light blue stars; Subantarctic Mode Water (SAMW), orange crosses; Antarctic Bottom Water (AABW), dark blue squares; Antarctic Zone (AZ), green circles; Polar Frontal Zone (PFZ), yellow triangles; Circumpolar Deep Water (CDW), red diamonds.

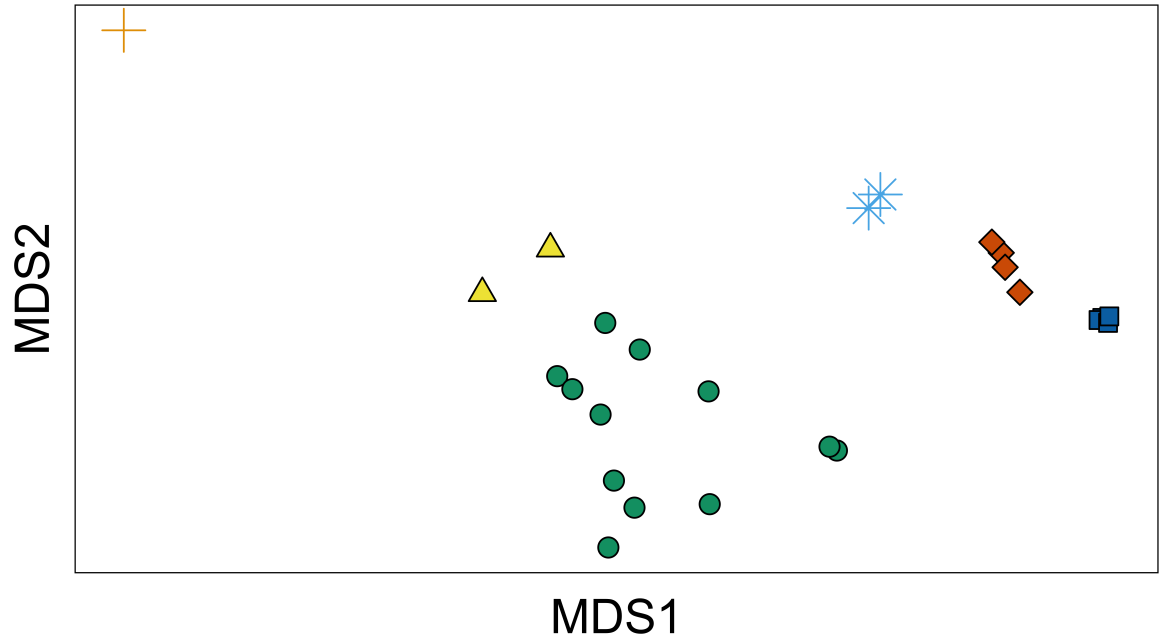

Figure 4: nMDS ordination of the environmental distance matrix (2D stress = 0.02). Antarctic Intermediate Waters (AAIW), light blue stars; Subantarctic Mode Water (SAMW), orange crosses; Antarctic Bottom Water (AABW), dark blue squares; Antarctic Zone (AZ), green circles; Polar Frontal Zone (PFZ), yellow triangles; Circumpolar Deep Water (CDW), red diamonds.

modelling), the best model consisted of all variables with the exception of phosphate (adjusted R^2 = 0.44), with the full set of variables only marginally worse (adjusted R^2 = 0.43). The rejection of phosphate may reflect a redundancy in the measurement of both phosphate and nitrate; these are often held to be constant throughout the ocean at the Redfield ratio of N:P $\sim$16:1 (Anderson and Sarmiento, 1994). However, deviation from this ratio has been observed in the SO and related to iron concentration (which was not measured in this study) and to phytoplankton abundance (Weber and Deutsch, 2010). For these reasons, and because of the marginal effect of discarding phosphate from the variable selection, it was retained when generating the environmental distance matrix.

The dbRDA plot showed the six variables retained by the distLM model (i.e. all except phosphate) structured the samples first along an axis separating surface and deep samples (dbRDA1), strongly related to dissolved oxygen (r = 0.79) (fig. 5). The second axis was best correlated with temperature (r = −0.75). All six retained variables had a moderate correlation with at least one of the first two axes (table 2). As with the nMDS ordinations (TODO ref to all nMDS figures), the water masses were generally well separated by the first two dbRDA axes. Samples from the AABW and AAIW, which were not separated by the first two axes, were clearly separated along the third (table 2), which was best correlated with pressure (r = −0.79). This suggests that these two masses had similar physicochemical properties, and were mainly distinguished by depth, consistent with their common origin in sinking Antarctic Surface Waters (Foldvik and Gammelsrød, 1988).

Table 2: Correlations between dbRDA coordinate axes and physicochemical variables (multiple partial correlations).

Variable	dbRDA1	dbRDA2	dbRDA3	dbRDA4	dbRDA5	dbRDA6
Pressure	− 0.316	− 0.485	− 0.787	− 0.148	− 0.144	− 0.048
Oxygen	0.792	− 0.171	− 0.125	0.069	− 0.567	0.050
Temperature	− 0.025	− 0.750	0.366	0.236	0.180	0.463
Nitrate	− 0.311	0.007	0.325	− 0.636	− 0.561	0.281
Silicate	− 0.303	0.341	− 0.166	0.615	− 0.371	0.498
Salinity	− 0.290	− 0.239	0.311	0.368	− 0.417	− 0.673

A distance effect was detected by comparing the taxonomic and spatial matrices with the environmental matrix held constant (partial Mantel; r = 0.38, p = 0.003). This indicated that a process other than contemporary environmental selection was appreciably affecting variation in microbial community composition.

Discussion

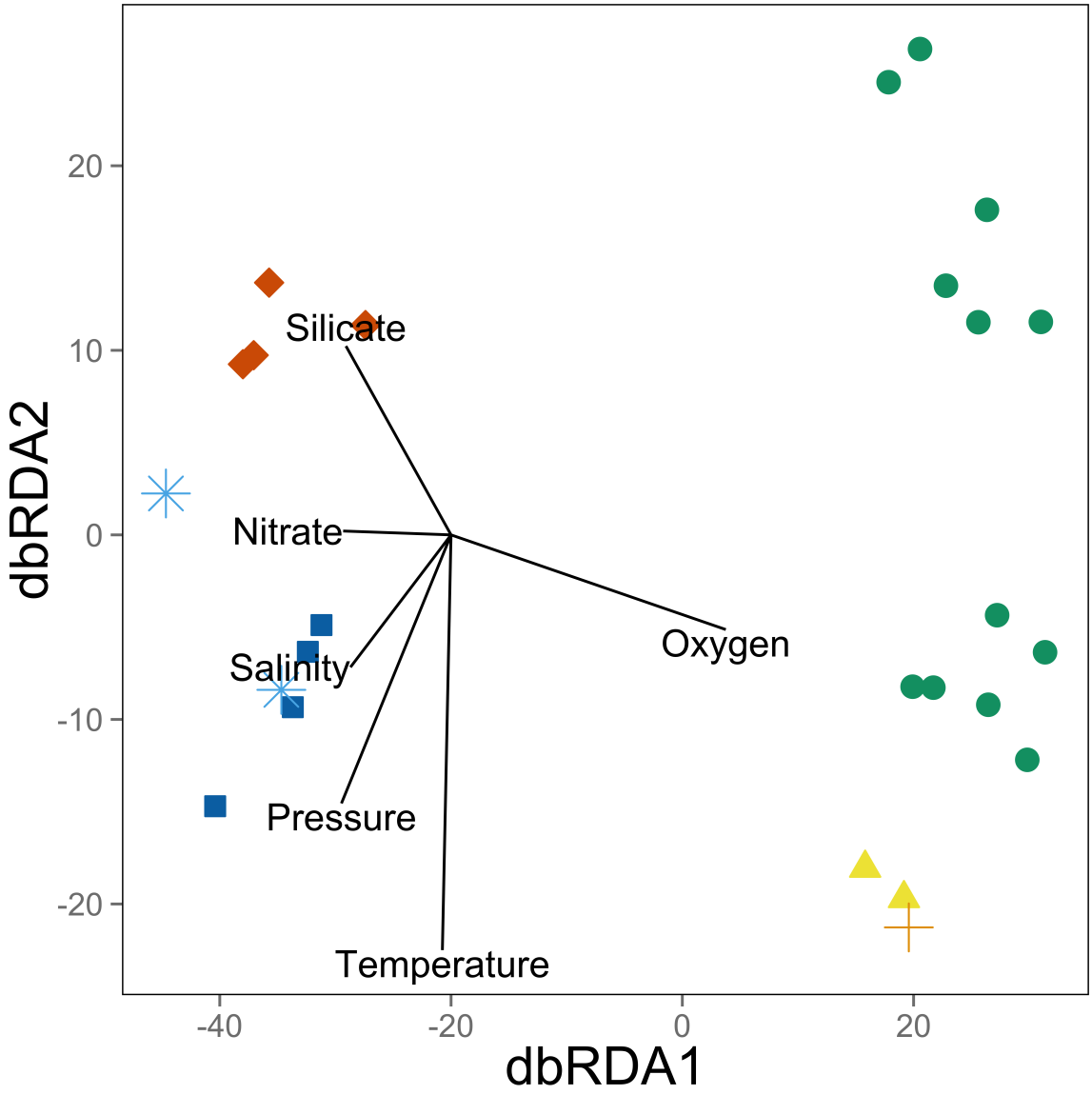

Figure 5: dbRDA ordination of the distLM model describing the relationship between the BEST-selected set of predictor physicochemical variables (pressure, oxygen, temperature, salinity, silicate, and nitrate) and the taxonomic dissimilarity between samples. Vectors represent the effect of each predictor variable on the two visualised axes. Vector length corresponds to the relative size of the effect, while direction represents the correlations to the two displayed axes. The first axis (dbRDA1) captures 64% of fitted and 37% of total variation between the samples' taxonomic profiles; the second (dbRDA2) captures 14% of fitted and 8% of total variation. Antarctic Intermediate Waters (AAIW), light blue stars; Subantarctic Mode Water (SAMW), orange crosses; Antarctic Bottom Water (AABW), dark blue squares; Antarctic Zone (AZ), green circles; Polar Frontal Zone (PFZ), yellow triangles; Circumpolar Deep Water (CDW), red diamonds.

References

Agogué H., Lamy D., Neal P. R., Sogin M. L., and Herndl G. J. (2011). Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. *Molecular Ecology*, 20(2):258–274.

Anderson L. A. and Sarmiento J. L. (1994). Redfield ratios of remineralization determined by nutrient data analysis. *Global Biogeochemical Cycles*, 8(1):65–80.

Baas Becking L. G. M. *Geobiologie Of Inleiding Tot De Milieukunde*. W.P. Van Stockum & Zoon, The Hague, 1934.

Bissett A., Richardson A. E., Baker G., Wakelin S., and Thrall P. H. (2010). Life history determines biogeographical patterns of soil bacterial communities over multiple spatial scales. *Molecular Ecology*, 19(19):4315–4327.

Brown M. V., Lauro F. M., DeMaere M. Z., Muir L., Wilkins D., Thomas T., Riddle M. J., Fuhrman J. A., Andrews-Pfannkoch C., Hoffman J. M., McQuaid J. B., Allen A., Rintoul S. R., and Cavicchioli R. (2012). Global biogeography of SAR11 marine bacteria. *Molecular systems biology*, 8.

Caporaso J. G., Kuczynski J., Stombaugh J., Bittinger K., Bushman F. D., Costello E. K., Fierer N., Pena A. G., Goodrich J. K., and Gordon J. I. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, 7(5):335–336.

Cerovečki I., Talley L. D., and Mazloff M. R. (2011). A comparison of Southern Ocean air-sea buoyancy flux from an ocean state estimate with five other products. *Journal of Climate*, 24:6283–6306.

Cho J.-C. and Tiedje J. M. (2000). Biogeography and degree of endemicity of fluorescent Pseudomonas strains in soil. *Applied and Environmental Microbiology*, 66(12):5448–5456.

Clarke K. R. and Gorley R. N. *PRIMER v6: User Manual / Tutorial*, 1st edition edition, 2006.

Wit R.de and Bouvier T. (2006). '*Everything is everywhere*, but, *the environment selects*'; what did Baas Becking and Beijerinck really say? *Environmental Microbiology*, 8(4):755–758.

Declerck S. A. J., Winter C., Shurin J. B., Suttle C. A., and Matthews B. (2013). Effects of patch connectivity and heterogeneity on metacommunity structure of planktonic bacteria and viruses. *The ISME Journal*, 7(3):533–542.

Firing Y. L., Chereskin T. K., and Mazloff M. R. (2011). Vertical structure and transport of the Antarctic Circumpolar Current in Drake Passage from direct velocity observations. *Journal of Geophysical Research*, 116(C8):C08015.

Foldvik A. and Gammelsrød T. (1988). Notes on Southern Ocean hydrography, sea-ice and bottom water formation. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 67(1-2):3–17.

Galand P. E., Potvin M., Casamayor E. O., and Lovejoy C. (2009). Hydrography shapes bacterial biogeography of the deep Arctic Ocean. *Nature*, 4(4):564–576.

Ghiglione J.-F., Galand P. E., Pommier T., Pedrós-Alió C., Maas E. W., Bakker K., Bertilson S., Kirchmanj D. L., Lovejoy C., Yager P. L., and Murray A. E. (2012). Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 109(43):17633–17638.

Giebel H.-A., Brinkhoff T., Zwisler W., Selje N., and Simon M. (2009). Distribution of *Roseobacter* RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. *Environmental Microbiology*, 11(8):2164–2178.

Grzymski J. J., Riesenfeld C. S., Williams T. J., Dussaq A. M., Ducklow H., Erickson M., Cavicchioli R., and Murray A. E. (2012). A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. *The ISME Journal*, 6(10):1901–1915.

Hamdan L. J., Coffin R. B., Sikaroodi M., Greinert J., Treude T., and Gillevet P. M. (2013). Ocean currents shape the microbiome of Arctic marine sediments. *The ISME Journal*, 7(4):685–696.

Hamilton A. K., Lovejoy C., Galand P. E., and Ingram R. G. (2008). Water masses and biogeography of picoeukaryote assemblages in a cold hydrographically complex system. *Limnology and Oceanography*, pages 922–935.

Hanson C. A., Fuhrman J. A., Horner-Devine M. C., and Martiny J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology*, 10(7):497–506.

Knights D., Kuczynski J., Charlson E. S., Zaneveld J., Mozer M. C., Collman R. G., Bushman F. D., Knight R., and Kelley S. T. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nature methods*, 8(9):761–763.

Lauro F. M., Chastain R. A., Blankenship L. E., Yayanos A. A., and Bartlett D. H. (2007). The unique 16S rRNA genes of piezophiles reflect both phylogeny and adaptation. *Applied and Environmental Microbiology*, 73(3):838–845.

Legendre P. and Anderson M. J. (1999). Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*, 69(1):1–24.

López-García P., López-López A., Moreira D., and Rodríguez-Valera F. (2001). Diversity of free-living prokaryotes from a deep-sea site at the Antarctic Polar Front. *FEMS Microbiology Ecology*, 36(2-3): 193–202.

Martiny J. B. H., Bohannan B. J. M., Brown J. H., Colwell R. K., Fuhrman J. A., Green J. L., Horner-Devine M. C., Kane M., Krumins J. A., Kuske C. R., Morin P. J., Naeem S., Ovreas L., Reysenbach A.-L., Smith V. H., and Staley J. T. (2006). Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, 4(2):102–112.

Mazloff M. R., Heimbach P., and Wunsch C. (2010). An eddy-permitting Southern Ocean state estimate. *Journal of physical oceanography*, 40:880–899.

Morris R. M., Rappé M. S., Connon S. A., Vergin K. L., Siebold W. A., Carlson C. A., and Giovannoni S. J. (2002). SAR11 clade dominates ocean surface bacterioplankton communities. *Nature*, 420(6917): 806–810.

Murray A. E. A., Preston C. M. C., Massana R. R., Taylor L. T. L., Blakis A. A., Wu K. K., and DeLong E. F. (1998). Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters near Anvers Island, Antarctica. *Applied and Environmental Microbiology*, 64(7):2585–2595.

Ng C., DeMaere M. Z., Williams T. J., Lauro F. M., Raftery M., Gibson J. A., Andrews-Pfannkoch C., Lewis M., Hoffman J. M., Thomas T., and Cavicchioli R. (2010). Metaproteogenomic analysis of a dominant green sulfur bacterium from Ace Lake, Antarctica. *The ISME Journal*, 4(8):1002–1019.

Orsi A. H., Whitworth T., and Nowlin W. D. (1995). On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep Sea Research Part I: Oceanographic Research Papers*, 42(5):641–673.

Paris C. B., Helgers J., Sebille E.van, and Srinivasan A. (2013). Connectivity Modeling System: A probabilistic modeling tool for the multi-scale tracking of biotic and abiotic variability in the ocean. *Environmental Modelling and Software*, 42(C):47–54.

Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., Peplies J., and Glöckner F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41(Database issue):D590–6.

Ramette A. and Tiedje J. M. (2007). Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 104(8):2761–2766.

Rosenberg M. and Rintoul S. R. Aurora Australis Marine Science Cruise AU1203 – Oceanographic Field Measurements and Analysis. Technical report, 2012.

Rosenberg M. S. and Anderson C. D. (2011). PASSaGE: pattern analysis, spatial statistics and geographic exegesis. Version 2. *Methods in Ecology and Evolution*, 2(3):229–232.

Rusch D. B., Halpern A. L., Sutton G., Heidelberg K. B., Williamson S., Yooseph S., Wu D., Eisen J. A., Hoffman J. M., Remington K., Beeson K., Tran B., Smith H., Baden-Tillson H., Stewart C., Thorpe J., Freeman J., Andrews-Pfannkoch C., Venter J. E., Li K., Kravitz S., Heidelberg J. F., Utterback T., Rogers Y.-H., Falcón L. I., Souza V., Bonilla-Rosso G., Eguiarte L. E., Karl D. M., Sathyendranath S., Platt T., Bermingham E., Gallardo V., Tamayo-Castillo G., Ferrari M. R., Strausberg R. L., Nealson K., Friedman R., Frazier M., and Venter J. C. (2007). The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biology*, 5(3):e77–e77.

Sokolov S. and Rintoul S. R. (2002). Structure of Southern Ocean fronts at 140°E. *Journal of Marine Systems*, 37(1):151–184.

). The concept of taxon invariance in ecology: do diversity patterns vary with changes in taxonomic resolution? *Folia Geobotanica*, 43:329–344.

Sul W. J., Oliver T. A., Ducklow H. W., Amaral-Zettler L. A., and Sogin M. L. (2013). Marine bacteria exhibit a bipolar distribution. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 110(6):2342–2347.

Sebille E.van, Johns W. E., and Beal L. M. (2012). Does the vorticity flux from Agulhas rings control the zonal pathway of NADW across the South Atlantic? *Journal of Geophysical Research*, 117(C5):C05037.

Weber T. S. and Deutsch C. (2010). Ocean nutrient ratios governed by plankton biogeography. *Nature*, 467(7315):550–554.

Wilkins D., Lauro F. M., Williams T. J., DeMaere M. Z., Brown M. V., Hoffman J. M., Andrews-Pfannkoch C., McQuaid J. B., Riddle M. J., Rintoul S. R., and Cavicchioli R. (2012). Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. *Environmental Microbiology*.

Wilkins D., Yau S., Williams T. J., Allen M. A., Brown M. V., DeMaere M. Z., Lauro F. M., and Cavicchioli R. (2012). Key microbial drivers in Antarctic aquatic environments. *FEMS Microbiology Reviews*.

Williams T. J., Wilkins D., Long E., Evans F., DeMaere M. Z., Raftery M. J., and Cavicchioli R. (2012). The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environmental Microbiology*.

Wunsch C. and Heimbach P. (2007). Practical global oceanic state estimation. *Physica D: Nonlinear Phenomena*, 230(1):197–208.