

# 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

**April 2013**



# Contents

<b>List of Figures</b>	<b>v</b>
<b>List of Tables</b>	<b>vii</b>
<b>List of Acronyms</b>	<b>x</b>
<b>1 The Polar Front</b>	<b>1</b>
1.1 Abstract . . . . .	1
1.2 Introduction . . . . .	2
1.3 Methods . . . . .	3
1.3.1 Sampling and metagenomic sequencing . . . . .	3
1.3.2 Phylogenetic analysis of metagenomic data . . . . .	6
1.3.2.1 BLAST comparison to RefSeq database . . . . .	6
1.3.2.2 OTU abundances and variance between zones . . . . .	6
1.3.2.3 Fragment recruitment to verify Operational Taxonomic Unit (OTU) identification . . . . .	7
1.3.2.4 Additional samples to test “polynya hypothesis” . . . . .	8
1.3.3 Functional analysis of metagenomic data . . . . .	8
1.3.3.1 BLAST comparison to KEGG database . . . . .	8
1.3.3.2 Analysis of functional potential . . . . .	9
1.3.3.3 Taxonomic decomposition . . . . .	9
1.4 Results . . . . .	9
1.4.1 Metagenomic sequencing . . . . .	9
1.4.2 Phylogenetic analysis of metagenomic data . . . . .	10
1.4.3 Fragment recruitment to verify OTU identification . . . . .	15
1.4.4 Additional samples to test alternative “polynya hypothesis” . . . . .	15
1.4.5 Functional analysis of metagenomic data . . . . .	15
1.5 Discussion . . . . .	19
1.5.1 Taxonomic groups differentiating the zones . . . . .	19
1.5.1.1 GSO-EOSA-1 . . . . .	19
1.5.1.2 Ammonia-oxidizing Crenarchaeota . . . . .	21
1.5.1.3 Cyanobacteria . . . . .	21
1.5.1.4 SAR11 and SAR116 clades . . . . .	22
1.5.1.5 Bacteroidetes . . . . .	23
1.5.1.6 Rhodobacterales . . . . .	23
1.5.1.7 Alteromonadales . . . . .	24
1.5.1.8 Verrucomicrobia . . . . .	24
1.5.2 Functional capacities differentiating the zones . . . . .	24
1.5.2.1 Conclusions . . . . .	28
<b>References</b>	<b>29</b>



# List of Figures

1.1	Map showing sites of seawater samples used in the Polar Front study . . . . .	4
1.2	Summing relative abundances across size fractions . . . . .	7
1.3	Rank-abundance curves for OTUs in each zone and size fraction . . . . .	12
1.4	Contribution of OTUs to variance between the North and South zones . . . . .	14
1.5	Read recruitment to reference genomes . . . . .	16
1.6	Tree of GSO-EOSA-1 related 16S rRNA genes . . . . .	20
1.7	Taxonomic decomposition of KEGG modules . . . . .	26



# List of Tables

1.1	Details of samples used in Polar Front study . . . . .	5
1.2	Additional samples used to test polynya hypothesis . . . . .	8
1.3	Twenty most abundant OTUs . . . . .	11
1.4	Highest-contributing OTUs to the difference between the North and South zones . . . . .	13
1.5	Contributions of KEGG modules to variance between the North and South zones . . . . .	17
1.6	Contributions of KEGG ortholog groups to variance between the North and South zones . . . . .	18



# List of Acronyms

- AAD** Australian Antarctic Division.
- AAP** Aerobic Anoxygenic Phototroph.
- ACC** Antarctic Circumpolar Current.
- ANOSIM** Analysis of Similarities.
- AOA** Ammonia-Oxidizing Archaea.
- AZ** Antarctic Zone.
- CASO** Climate of Antarctica and the Southern Ocean.
- CEAMARC** Collaborative East Antarctic Marine Census.
- DFAA** Dissolved Free Amino Acids.
- DMSP** dimethylsulfoniopropionate.
- DOC** Dissolved Organic Carbon.
- DOM** Dissolved Organic Matter.
- FISH** Fluorescence *In Situ* Hybridization.
- GAAS** Genome relative Abundance and Average Size.
- GOS** Global Ocean Sampling expedition.
- GSO-EOSA-1** Gammaproteobacterial Sulfur Oxidizer-EOSA-1.
- HMW** High Molecular Weight.
- HNLC** High Nutrient, Low Chlorophyll.
- KEGG** Kyoto Encyclopedia of Genes and Genomes.
- MGI** Marine Group I Crenarchaeota.
- MMPA** methylmercaptopropionate.
- NZ** North Zone.
- OTU** Operational Taxonomic Unit.
- PF** Polar Front.
- PFZ** Polar Frontal Zone.

**POM** Particulate Organic Matter.

**SAF** Subantarctic Front.

**SAZ** Subantarctic Zone.

**SIMPER** Similarity Percentages.

**SO** Southern Ocean.

**STF** Subtropical Front.

**SZ** South Zone.

# Chapter 1

## The Polar Front as a major biogeographic boundary in the Southern Ocean

Sections of this chapter have been previously published in 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. (2013). Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. *Environmental Microbiology*, 15:1318–1333.

### 1.1 Abstract

The Polar Front (PF) is a major oceanographic feature of the Southern Ocean (SO), dividing surface water masses with distinct physicochemical characteristics. To confirm and characterise the PF's role as a biogeographic barrier in the SO, sixteen metagenomic samples were collected on a latitudinal transect from waters near Hobart, Australia ( $44^{\circ}$  S) to near the Mertz Glacier, Antarctica ( $67^{\circ}$  S). The samples were shotgun sequenced, and the community composition and functional potential of the captured microorganisms were analysed. The SAR11 and SAR116 clades and the cyanobacterial genera *Prochlorococcus* and *Synechococcus* were strongly overrepresented north of the PF. Conversely, the Gammaproteobacterial Sulfur Oxidizer-EOSA-1 (GSO-EOSA-1) complex, the phyla Bacteroidetes and Verrucomicrobia and order Rhodobacterales were overrepresented in waters south of the PF. Functions enriched south of the PF included a range of transporters, sulphur reduction and histidine degradation to glutamate, while branched-chain amino acid transport, nucleic acid biosynthesis and methionine salvage were overrepresented north of the PF. The taxonomic and functional characteristics suggested a shift of primary production from cyanobacteria in the north to eukaryotic phytoplankton in the south, and reflected the different trophic states of the two regions. Overall, this study supported the PF as a defining biogeographic feature of the SO.

## 1.2 Introduction

The surface waters of the SO consist of zones separated by circumpolar fronts, the locations of which vary with time and longitude (Whitworth, 1980; Orsi *et al.*, 1995; Sokolov and Rintoul, 2002). From north to south, the major fronts are the Subtropical Front (STF), the Subantarctic Front (SAF) and the PF (??). The SAF and PF are associated with the Antarctic Circumpolar Current (ACC), the world's largest ocean current and a defining oceanographic feature of the SO. Anthropogenic climate change may be driving the warming and freshening of the ACC (Böning *et al.*, 2008). It may also be shifting the ACC and associated fronts poleward; the mean path of the current has moved ~50 km south since the 1950s (Gille, 2002). If this migration continues, by 2100 it will have displaced a volume of water south of the ACC approximately equivalent to the Arctic Ocean (Fyfe and Saenko, 2005).

Climate change may also be affecting the surface regions defined by these fronts. The major surface zones of the SO are the Subantarctic Zone (SAZ), between the STF and SAF; the Polar Frontal Zone (PFZ), between the SAF and the PF; and the Antarctic Zone (AZ), between the PF and the Antarctic continent (??). These zones have different physicochemical properties, such as density, salinity, temperature and nutrient concentrations (Sokolov and Rintoul, 2002), and the fronts represent stepwise transitions in these properties (Whitworth and Nowlin, 1987). However, as a consequence of climate change, waters on the poleward side of the ACC have become warmer and more saline, while those to the north cooler and fresher (Böning *et al.*, 2008).

The PF has been suggested to be a major biogeographic boundary in the distribution and abundance of both zooplankton (Chiba *et al.*, 2001; Hunt *et al.*, 2001; Esper and Zonneveld, 2002; Ward *et al.*, 2003) and bacterioplankton (Selje *et al.*, 2004; Abell and Bowman, 2005; Giebel *et al.*, 2009; Weber and Deutsch, 2010). The effects of climate change on the physical oceanography of the SO may therefore have global ecological significance, particularly as the SO is a major site for sequestration of anthropogenic CO<sub>2</sub> (Sabine *et al.*, 2004; Mikaloff Fletcher *et al.*, 2006) through both physicochemical processes and the microbe-driven "biological pump" of CO<sub>2</sub> fixation (Thomalla *et al.*, 2011), potentially forming a feedback loop (Cox *et al.*, 2000). However, the microbial assemblages that inhabit the SO are generally poorly understood, and their diversity and functional capacity poorly characterised (reviewed in Murray and Grzymski, 2007). A community-level understanding is needed to predict the effects of a shifting ACC on the distribution, abundance and ecosystem functions of microorganisms.

Large scale metagenome surveys have not previously been performed in the SO. Compared to traditional environmental microbiological methods such as culture-based or DGGE surveys, metagenomic studies are able to offer both deeper (capture of sequence from rare taxa) and broader (larger sample sizes with greater statistical significance) data on the taxonomic makeup and functional capacities of marine microbial communities (e.g. Rusch *et al.*, 2007; Angly *et al.*, 2006; Dinsdale *et al.*, 2008). Metagenome sampling is also better suited to large-scale biogeographic studies, as it allows large numbers of samples to be taken across broad spatial and temporal scales with uniform sampling, storage and processing, compared to e.g. culture-based methods where technical replicability is more difficult.

This project is part of a metagenomic survey of the SO begun in the austral summer of 2006, based on the sampling design of the Global Ocean Sampling expedition (GOS) (Rusch *et al.*, 2007). The GOS sampling strategy involves size fractionation of plankton assemblages by passing seawater through a 20  $\mu\text{m}$  prefilter then capturing planktonic biomass on sequential 3.0  $\mu\text{m}$ , 0.8  $\mu\text{m}$  and 0.1  $\mu\text{m}$  filter membranes. As well as allowing deeper sequencing of metagenomes and thus better representation of low-abundance taxa (Rusch *et al.*, 2007), this approach allows comparison to other samples collected using the GOS methods (e.g. Brown *et al.*, 2012). This study will thus focus on the 0.1–20  $\mu\text{m}$  fraction of the sampled microbial assemblages.

## 1.3 Methods

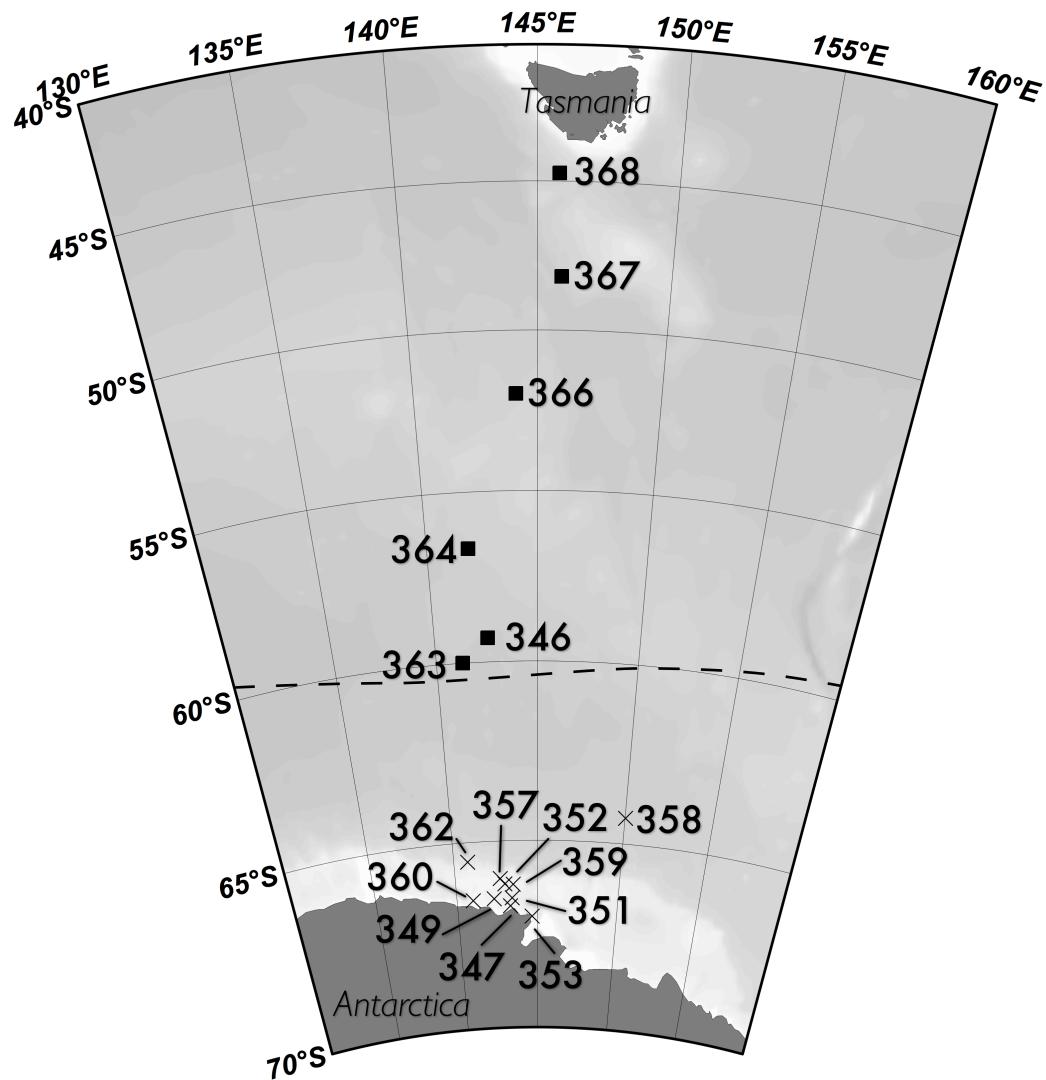
### 1.3.1 Sampling and metagenomic sequencing

Sampling<sup>1</sup> was conducted on board the RSV *Aurora Australis* during Australian Antarctic Division (AAD) cruise V3 Collaborative East Antarctic Marine Census (CEAMARC) / Climate of Antarctica and the Southern Ocean (CASO) from 13 December 2007–25 January 2008. This cruise occupied the SR3 latitudinal transect from waters near Hobart, Australia ( $\sim 44^\circ \text{S}$ ) to the Mertz Glacier polynya, Antarctica ( $\sim 67^\circ \text{S}$ ) within a longitudinal range of  $\sim 140$ – $150^\circ \text{E}$ . Sixteen samples were obtained within this range (Fig. 1.1).

A range of physicochemical data were recorded by integrated instruments on the RSV *Aurora Australis* including water column depth, water temperature, salinity, fluorescence and meteorological conditions (Table 1.1). These data were used to locate the PFZ based on a surface temperature gradient of  $\sim 1.35^\circ \text{C}$  across a distance of 45–65 km, placing the PF at approximately  $59.70^\circ \text{S}$ , consistent with previous descriptions (Moore *et al.*, 1999; Sokolov and Rintoul, 2002). Samples were accordingly grouped into “North” and “South” zones (Table 1.1). The North Zone (NZ) represents waters from the Subtropical, Subantarctic and PFZ regions, while the South Zone (SZ) represents the AZ.

---

<sup>1</sup>Sampling was performed by Jeffrey M. Hoffman and Jeffrey B. McQuaid.



**Figure 1.1:** Sites of seawater samples used in this study. Squares indicate surface samples from the North Zone; crosses samples from the South Zone. The dashed line represents the Polar Front.

**Table 1.1:** Sampling time, location and physicochemical properties of samples used in this study. All data were retrieved from underway instruments aboard the RSV *Aurora Australis*.

Sample	Zone	Date	Latitude	Longitude	Depth (m)	Temperature (°C)	Salinity (PSU)	Fluorescence ( $\mu\text{g L}^{-1}$ )
346	North	20/12/2007	-59.31	142.59	2	2.9	33.75	0.3
347	South	23/12/2007	-66.02	142.74	2	0.6	34.20	4.0
349	South	27/12/2007	-66.57	142.32	1.5	-1.3	34.40	2.3
351	South	28/12/2007	-66.56	143.43	1.5	-0.6	34.30	1.3
352	South	29/12/2007	-66.77	143.32	2.5	-0.8	34.30	3.1
353	South	30/12/2007	-67.05	144.68	2	-1.8	34.40	0.3
357	South	05/01/2008	-66.17	143.02	2	-0.4	34.15	2.5
358	South	09/01/2008	-64.30	150.03	2	0	33.55	0.5
359	South	12/01/2008	-66.19	143.53	2	-0.2	34.21	2.5
360	South	13/01/2008	-66.58	141.02	2	-0.7	34.04	6.2
362	South	19/01/2008	-65.54	140.83	2	0.7	32.20	0.5
363	North	22/01/2008	-60.00	141.31	2	3.3	33.77	0.1
364	North	23/01/2008	-56.70	141.88	2	4	33.70	0.5
366	North	24/01/2008	-52.02	144.14	2	7.6	33.84	0.3
367	North	25/01/2008	-48.25	145.90	2	11	34.43	0.2
368	North	26/01/2008	-44.72	145.78	2	14.8	34.96	1.3

At each station, ~250–560 L of seawater was pumped from ~1.5–2.5 m below the sea surface into drums stored at ambient temperature on deck. Seawater samples were prefiltered through a 20 µm plankton net, then biomass 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<sup>2</sup> was performed at the J. Craig Venter Institute (Rockville, USA) using a modified phenol-chloroform method as described in Rusch *et al.* (2007). Shotgun sequencing was performed on a GS20 FLX Titanium platform (Roche, Branford, USA) also at the J. Craig Venter Institute as described in Lauro *et al.* (2011). Perfect duplicate reads were removed using sffToCa (Celera Corporation, Alameda, USA), read ends were trimmed with LUCY (Chou and Holmes, 2001), and the bottom 8% and top 3% of reads ordered by length were removed as described in Lauro *et al.* (2011)<sup>3</sup>.

### 1.3.2 Phylogenetic analysis of metagenomic data

#### 1.3.2.1 BLAST comparison to RefSeq database

A subset of the RefSeq microbial (bacterial and archaeal) genome database (release 41, retrieved 31 May 2010 from <ftp://ftp.ncbi.nih.gov/refseq/release/>) was prepared by excluding sequences with the words "shotgun", "contig", "partial", "end" or "part" in their headers. This is necessary to ensure only full and complete genomes yielded matches, as the Genome relative Abundance and Average Size (GAAS) software tool used for later processing uses this assumption to estimate environmental genome sizes (following Angly *et al.*, 2009).

The metagenomic reads from each sample were compared against this database using TBLASTX, with default parameters except for: E-value threshold  $1.0 \times 10^{-3}$ , cost to open gap 11, cost to extend gap 1, masking of query sequence by SEG masking with lookup table only. Operational Taxonomic Units (OTUs) in this study are thus defined as the best species match in RefSeq by these parameters. The outputs of all TBLASTX searches against RefSeq were processed by MINSPEC, and hits not belonging to the minimal OTU sets were removed.

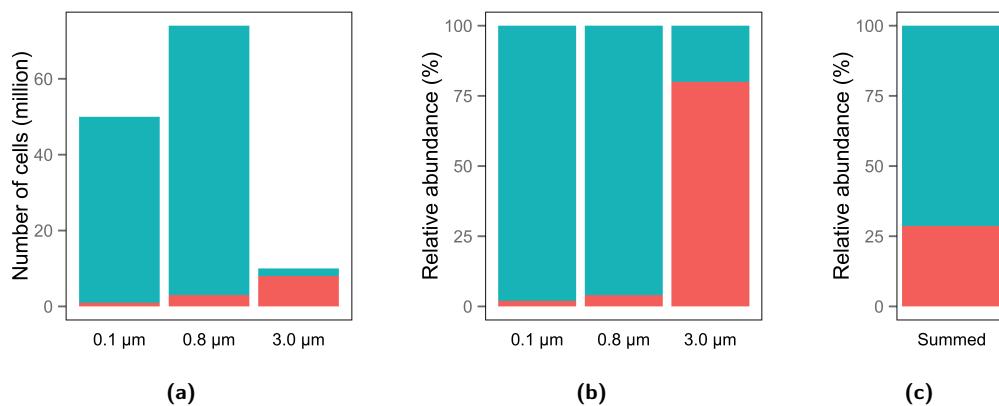
#### 1.3.2.2 OTU abundances and variance between zones

The relative OTU abundances for each sample were determined using the PERL script GAAS (Angly *et al.*, 2009). GAAS estimates the relative abundance of OTUs from the number and quality of BLAST matches ("hits") to each OTU, taking into account differences in genome size. This approach takes advantage of all sequencing reads in a shotgun metagenome, in contrast with marker-gene bases approaches (e.g. 16S), and allows relative abundances to be estimated on a "one genome per cell" basis, increasing the accuracy of these estimates. GAAS was run with the default settings. To normalise for reads which did not yield acceptable hits, the relative abundances for each sample were scaled by that sample's effective BLAST hit rate.

A difficulty with the size fractionation approach is that relative abundances cannot be summed across fractions. This is because each fraction almost certainly represents a

<sup>2</sup>DNA extraction was performed by Cynthia Andrews-Pfannkoch and others at the J. Craig Venter Institute.

<sup>3</sup>Read quality filtering was performed by Federico M. Lauro and Matthew Z. DeMaere.



**Figure 1.2:** A visual demonstration of the error introduced by summing OTU relative abundances across size fractions. The red OTU composes only 9% of the *absolute* number of cells captured (a). However, it has a high *relative* abundance (80%) in the 3.0  $\mu\text{m}$  fraction (b). A metagenomic analysis without cell counts for each size fraction is only able to determine these relative abundances. If the relative abundances for each fraction are simply summed, the red OTU appears to compose 29% of the community, more than three times the correct value (c). Without cell counts for each fraction, relative abundances cannot be summed between the fractions.

different absolute number of cells. If the relative abundances are summed, the apparent abundance of some OTUs could be distorted (visually demonstrated in Fig. 1.2). Thus, an OTU profile was generated for each sample by encoding the scaled relative abundance of each OTU from each size fraction as a separate variable. This allows samples to be compared on the basis of community composition without introducing systematic biases.

To test the hypothesis that the oceanic zones harbour significantly different communities, Analysis of Similarities (ANOSIM) with 999 permutations was performed on a standardised, log-transformed Bray-Curtis resemblance matrix of OTU profiles. Similarity Percentages (SIMPER) analysis was performed to identify the contribution of individual OTUs to differences between the zones. All statistical procedures were performed in PRIMER 6 as described by Clarke and Warwick (2001).

### 1.3.2.3 Fragment recruitment to verify OTU identification

Fragment recruitment to reference genomes was visualised as a precaution against unintended sequencing bias and to ensure the OTU annotation pipeline was functioning as intended. Plots of read recruitment depth on reference genomes were generated using a custom script in R and PERL<sup>4</sup>. Samples from both the SZ and NZ were randomly selected and their 0.1  $\mu\text{m}$  fractions used as sources of recruited reads. Three high-abundance and three low-abundance OTUs were randomly selected as the reference genomes. As shotgun sequencing of a metagenome should be essentially random, it was predicted that these plots would show fairly even coverage of the reference genomes. Genomes where all recruited reads were concentrated in a few small regions would be likely candidates

<sup>4</sup>This script, along with other accessory scripts used in this project, is open source and available at <https://github.com/wilcox/blast-tools>.

**Table 1.2:** Dates and locations of additional samples used to falsify the hypothesis that the Mertz Glacier polynya was responsible for the observed biogeographic partitioning.

Sample	Date	AAD Voyage Number	Latitude	Longitude
388	20/10/2008	V1 2008–09	−63.81	115.16
398	22/10/2008	V1 2008–09	−64.80	112.37
390	30/10/2008	V1 2008–09	−64.82	80.72
392	13/12/2008	V2 2008–09	−64.18	76.45
393	15/12/2008	V2 2008–09	−55.26	74.25

for OTUs which were not present, but shared a small genomic region with an OTU which was present, or possibly indicative of non-random priming during shotgun sequencing.

#### 1.3.2.4 Additional samples to test “polynya hypothesis”

Because many samples from the SZ were taken in the region of the Mertz Glacier polynya, an alternative hypothesis for any difference observed between the SZ and NZ would be that the difference represented the effect of the polynya, rather than the PF. To test this, metagenomes from an additional set of samples collected<sup>5</sup> in a different project were obtained. These samples were obtained in waters south of the PF on different voyages and in a different longitudinal range (Table 1.2), where the influence of the Mertz Glacier polynya could not account for any observed difference between the zones. Sampling, DNA extraction and sequencing were<sup>6</sup> performed as described above, although only metagenomes from the 0.1–0.8 µm size fraction were available. The resulting metagenomic reads were compared to RefSeq and weighted relative abundances calculated with GAAS as described above.

Two ANOSIM tests were performed on the standardised and log-transformed Bray-Curtis resemblance matrices of the samples’ OTU abundance profiles and similarly prepared matrices representing only the 0.1 µm fractions of the SZ and NZ samples. In the first ANOSIM test (“polynya hypothesis”), all samples from the Mertz Glacier polynya region (i.e. all SZ samples except 358) were grouped together, while the samples from the open ocean were placed in a separate group. In the second test (“PF hypothesis”), the samples were grouped by their location relative to the PF.

### 1.3.3 Functional analysis of metagenomic data

#### 1.3.3.1 BLAST comparison to KEGG database

In order to identify functional differences between the zones, the set of metagenomic reads from each sample was compared against the Kyoto Encyclopedia of Genes and Genomes (KEGG) GENES database (retrieved 2 July 2010 from <ftp://ftp.genome.jp/pub/kegg/genes/fasta/genes.pep>) with BLASTX, with default parameters except for: maximum number of database sequence alignments 10; E-value threshold  $1.0 \times 10^{-3}$ ; gap opening penalty 11; gap extension penalty 1; masking of query sequence by SEG masking for lookup table only.

<sup>5</sup>Collection was performed by Ricardo Cavicchioli, Federico M. Lauro and Mark V. Brown.

<sup>6</sup>TODO who did this?

### 1.3.3.2 Analysis of functional potential

Genes identified by BLASTX were aggregated to KEGG ortholog groups according to the KEGG Orthology schema (<ftp://ftp.genome.jp/pub/kegg/genes/ko>, retrieved Mar 29 2011), and ortholog group abundances calculated for each sample. Following Coleman and Chisholm (2010), a read was considered a hit to a given ortholog group if the top three hits for that read (or all hits if fewer than three total hits) were to genes from the same ortholog group, and had bit scores > 40. If the bit score difference between any two top hits was greater than 30, only the hits above this difference were considered.

Ortholog group counts were then used to calculate the abundance of KEGG modules. Because many ortholog groups are members of more than one module, the abundance  $a_m$  of each module  $m$  was calculated as

$$a_m = \sum_{K=1}^n \frac{C_K}{M_K}$$

where  $n$  is the number of ortholog groups  $K$  belonging to module  $m$ ,  $C_K$  is the number of hits to ortholog group  $K$ , and  $M_K$  is the total number of modules to which  $K$  belongs. To account for differences in sequencing depth between samples, module abundances were scaled to 500,000 reads per sample. To test the hypothesis that the NZ and SZ harbour significantly different functional potential, one-way ANOSIM with 999 permutations was performed as above on a standardised, log-transformed Bray-Curtis distance resemblance matrices of the module and ortholog group profiles. SIMPER was performed as above to identify modules which contributed highly to the variation in functional potential between the two zones.

### 1.3.3.3 Taxonomic decomposition

To link modules with a high contribution to variance or otherwise of interest to specific OTUs (“taxonomic decomposition”), a script was written in PERL to decompose each module to its constituent ortholog groups, then each ortholog group to each representative gene sequence in the KEGG GENES database. Because each KEGG GENES sequence is associated with a particular organism, this allowed the abundance of each module to be decomposed into the relative contribution from each of these organisms. Functional contributions could then be putatively assigned to OTUs, including those which were not identified in the taxonomic analysis, as the database included gene sequences for organisms for which a full genome was not available. To aid interpretation of these results, the taxonomic decompositions for some modules were aggregated to higher taxonomic ranks, such as class or family.

## 1.4 Results

### 1.4.1 Metagenomic sequencing

6.6 Gbp of sequence data representing microorganisms in the size range 0.1–3.0 µm was obtained from 16 samples. After removing low-quality reads, 157,507–597,689 reads re-

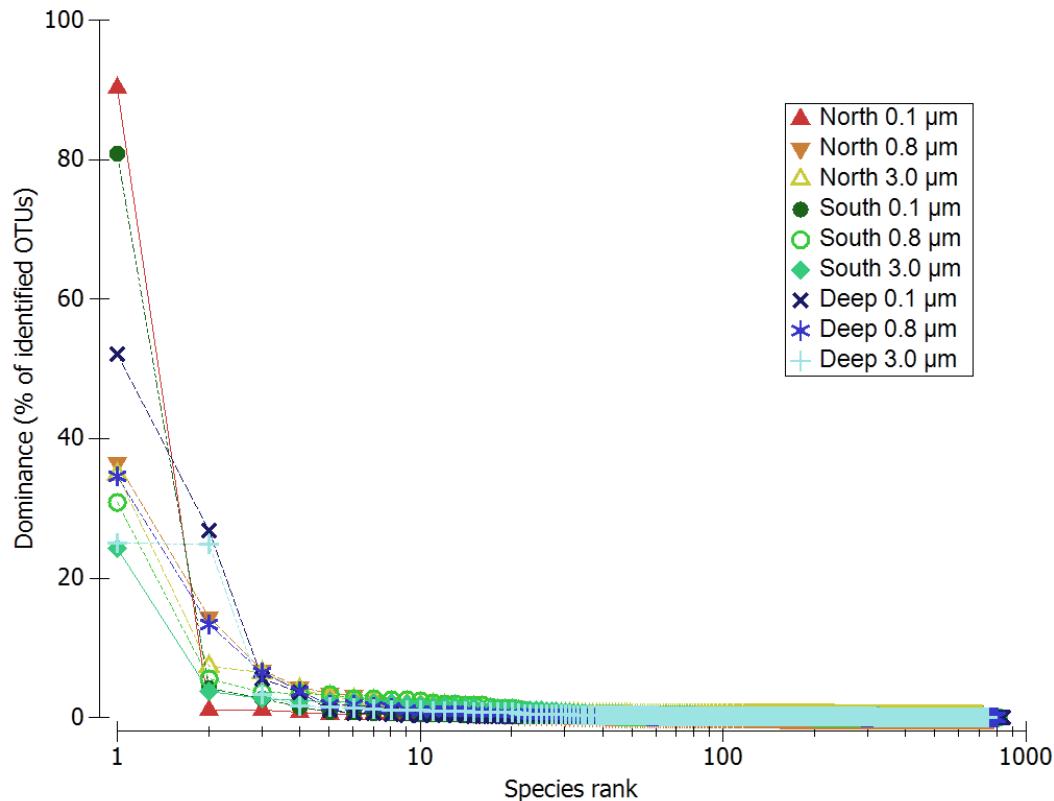
mained per sample (mean 354,399) with lengths ranging from 100 to 606 bp (mean 378 bp).

### **1.4.2 Phylogenetic analysis of metagenomic data**

The proportion of reads in each sample that yielded hits to RefSeq ranged from 25–85% (mean 62%). The most abundant OTUs in each sample are given in Table 1.3 and a full list of OTU abundances in the supplementary material (*PF-all-OTUs.txt*). All samples and size fractions exhibited very low OTU evenness (Fig. 1.3).

**Table 1.3:** Relative abundances (as percentages) of the twenty most abundant OTUs identified in this study, in each zone and size fraction.

OTU	North			South		
	0.1 µm	0.8 µm	3.0 µm	0.1 µm	0.8 µm	3.0 µm
"Candidatus Pelagibacter ubique" HTCC1062	61.76	25.00	23.87	58.85	22.40	17.61
<i>Nitrosopumilus maritimus</i> SCM1	0.01996	0.01438	0.009508	1.076	1.309	1.210
"Candidatus Ruthia magnifica" str. Cm ( <i>Calyptogena magnifica</i> )	0.6699	0.6458	0.5484	2.987	2.616	1.025
<i>Roseobacter</i> sp. OCh114	0.3125	2.932	1.588	0.4477	3.994	2.657
<i>Synechococcus</i> sp. CC9902	0.1081	9.837	4.973	0.0007484	0.004156	0.09733
<i>Silicibacter pomeroyi</i> DSS-3	0.2578	2.286	1.154	0.3070	2.505	1.576
<i>Gramella forsetii</i> strain KT0803	0.2412	1.210	1.755	0.4993	2.347	1.890
"Candidatus Vesicomyosocius okutanii" strain HA	0.4634	0.4642	0.2078	1.970	1.807	0.2174
<i>Robiginitalea biformata</i> strain HTCC2501	0.2751	1.099	1.297	0.4722	1.878	1.405
<i>Flavobacterium psychrophilum</i> strain JIP02/86	0.1718	0.8409	1.224	0.4316	1.960	1.598
<i>Synechococcus</i> sp. CC9311	0.03014	4.624	4.409	0.0007221	0.002778	0.02764
"Candidatus Puniceispirillum marinum" IMCC1322	0.6444	2.077	1.267	0.3586	1.377	0.7109
<i>Silicibacter</i> sp. TM1040	0.2274	1.652	0.8738	0.2709	1.803	1.233
<i>Jannaschia</i> sp. DFL-12	0.1776	1.378	0.7350	0.2443	1.692	0.8009
<i>Zunongwangia profunda</i> strain SM-A87	0.1522	0.7487	1.059	0.2968	1.410	1.204
<i>Colwellia</i> sp. 34H	0.02345	0.3636	2.736	0.05207	0.5140	1.041
<i>Coraliomargarita akajimensis</i> strain DSM 45221	0.03698	0.07573	0.1197	0.1154	1.543	1.680
<i>Jannaschia</i> sp. CCS1	0.1173	0.9344	0.4784	0.1711	1.230	0.8239
<i>Pseudoalteromonas atlantica</i> strain T6c	0.01251	0.4772	1.993	0.02270	0.4089	1.132
<i>Saccharophagus degradans</i> strain 2-40	0.06532	0.4325	0.5429	0.1289	1.072	0.8663
<i>Flavobacterium johnsoniae</i> strain UW101	0.08822	0.4220	0.6141	0.2034	0.9389	0.8578
<i>Capnocytophaga ochracea</i> strain DSM 7271	0.1143	0.4830	0.5399	0.2314	0.8815	0.6814
<i>Marinomonas</i> sp. MWYL1	0.03777	0.2529	0.3026	0.1514	1.300	0.7006
<i>Cellvibrio japonicus</i> strain Ueda107	0.05884	0.3080	0.3231	0.1155	0.9917	0.4713
<i>Marinobacter hydrocarbonoclasticus</i> VT8	0.04093	0.2889	0.3883	0.08418	0.7195	0.3848
<i>Pseudoalteromonas haloplanktis</i> strain TAC125	0.01389	0.2505	0.8896	0.03427	0.3561	0.6530
<i>Teredinibacter turnerae</i> strain T7901	0.05665	0.3051	0.3081	0.1138	0.9174	0.5127
<i>Acinetobacter baumannii</i> strain SDF	0.004886	0.007187	0.4073	0.006260	0.04218	1.459

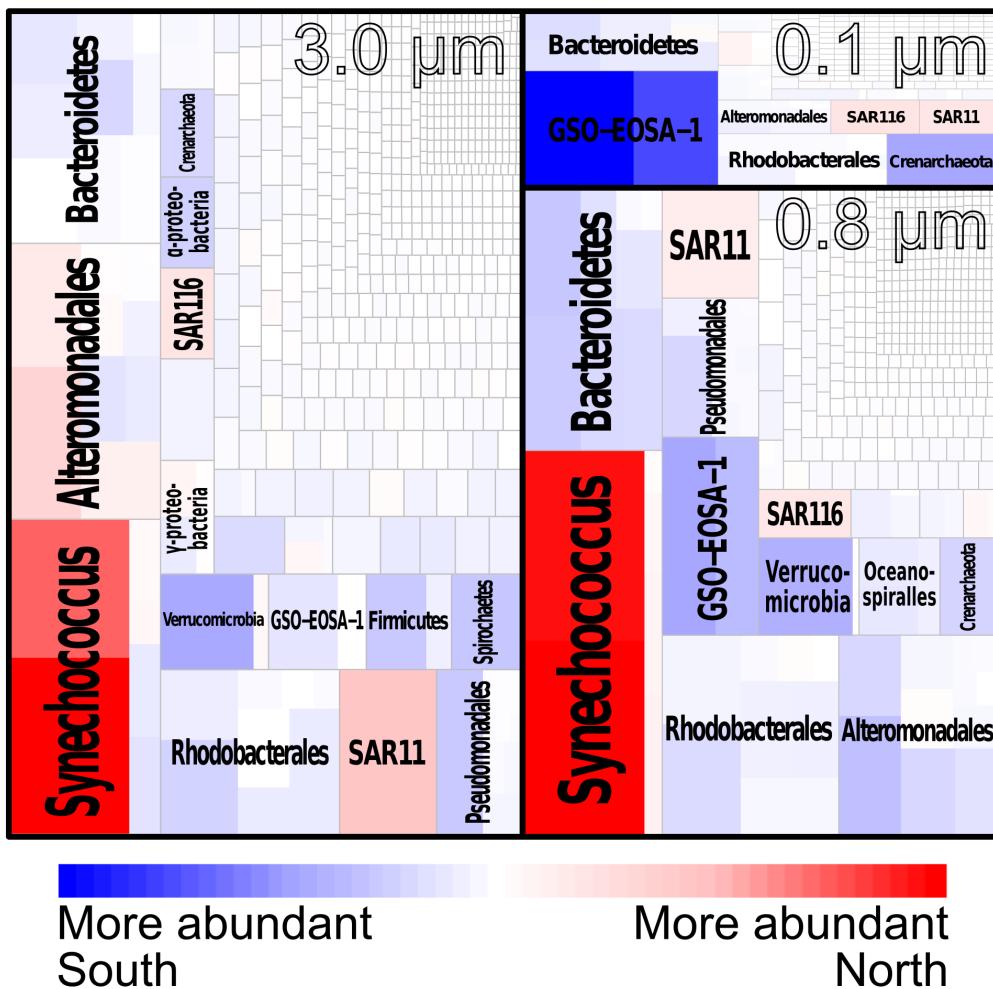


**Figure 1.3:** Rank-abundance curves for OTUs identified in each zone and size fraction. The dominance of a given OTU is its relative abundance as a percentage of all identified OTUs. The x-axis is scaled logarithmically. Generated using PRIMER 6.

ANOSIM analysis showed that the zones harbor significantly different microbial communities ( $R = 0.45, p = 0.004$ ). SIMPER was performed in order to identify the contribution of individual OTUs to the difference between the NZ and SZ. The results for the highest contributors are provided in Table 1.4, and are graphically summarised for all OTUs in Fig. 1.4.

**Table 1.4:** The thirty OTUs with the highest contributions to the difference between the NZ and SZ. Abundances are zonal averages and have been standardised and log-transformed. As each OTU on each size fraction was encoded as a separate variable in the SIMPER analysis, the size fraction is given after each OTU name.

OTU	Abundance South	Abundance North	Contribution to variance (%)
<i>Synechococcus</i> sp. CC9311 0.8 µm	0.00	1.08	2.88
<i>Synechococcus</i> sp. CC9902 0.8 µm	0.00	1.04	2.81
<i>Synechococcus</i> sp. CC9311 3.0 µm	0.01	0.98	2.59
<i>Synechococcus</i> sp. CC9902 3.0 µm	0.04	0.76	2.03
" <i>Candidatus Pelagibacter ubique</i> " HTCC1062 3.0 µm	1.97	2.40	1.97
" <i>Candidatus Ruthia magnifica</i> " str. Cm ( <i>Calyptogena magnifica</i> ) 0.1 µm	0.82	0.25	1.57
<i>Colwellia</i> sp. 34H 3.0 µm	0.34	0.66	1.32
" <i>Candidatus Ruthia magnifica</i> " str. Cm ( <i>Calyptogena magnifica</i> ) 0.8 µm	0.74	0.25	1.32
" <i>Candidatus Pelagibacter ubique</i> " HTCC1062 0.8 µm	2.32	2.48	1.32
" <i>Candidatus Vesicomyosocius okutanii</i> " strain HA 0.1 µm	0.62	0.18	1.20
<i>Coraliomargarita akajimensis</i> strain DSM 45221 0.8 µm	0.48	0.04	1.13
<i>Coraliomargarita akajimensis</i> strain DSM 45221 3.0 µm	0.49	0.06	1.10
<i>Roseobacter</i> sp. OCh 114 0.8 µm	1.01	0.81	1.08
<i>Pseudoalteromonas atlantica</i> strain T6c 3.0 µm	0.38	0.54	1.08
" <i>Candidatus Vesicomyosocius okutanii</i> " strain HA 0.8 µm	0.57	0.19	1.04
<i>Acinetobacter baumannii</i> strain SDF 3.0 µm	0.45	0.18	0.95
<i>Gramella forsetii</i> strain KT0803 0.8 µm	0.72	0.43	0.94
<i>Marinomonas</i> sp. MWYL1 0.8 µm	0.46	0.11	0.92
<i>Roseobacter</i> sp. OCh 114 3.0 µm	0.76	0.54	0.91
<i>Flavobacterium psychrophilum</i> strain JIP02/86 0.8 µm	0.63	0.32	0.89
<i>Silicibacter pomeroyi</i> DSS-3 0.8 µm	0.75	0.69	0.86
<i>Brachyspira hyodysenteriae</i> strain WA1 3.0 µm	0.47	0.19	0.84
" <i>Candidatus Ruthia magnifica</i> " str. Cm ( <i>Calyptogena magnifica</i> ) 3.0 µm	0.34	0.21	0.82
<i>Pseudoalteromonas haloplanktis</i> strain TAC125 3.0 µm	0.22	0.33	0.77
<i>Robiginitalea bifomata</i> strain HTCC2501 0.8 µm	0.61	0.40	0.74
<i>Nitrosopumilus maritimus</i> SCM1 0.1 µm	0.27	0.01	0.72
<i>Gramella forsetii</i> strain KT0803 3.0 µm	0.59	0.59	0.71
<i>Lysinibacillus sphaericus</i> strain C3-41 3.0 µm	0.29	0.02	0.71
<i>Nitrosopumilus maritimus</i> SCM1 0.8 µm	0.25	0.01	0.70
<i>Silicibacter</i> sp. TM1040 0.8 µm	0.59	0.55	0.69



**Figure 1.4:** Contribution of OTUs to variance between the North and South zones, and differential abundance of OTUs from each size fraction between the two zones. Each coloured (red or blue) rectangle represents an OTU identified through analysis of BLAST matches between SO metagenome data and the RefSeq database. The area of each rectangle as a proportion of the total plot area corresponds to that OTU's contribution to the total variance between the two zones. The colour of each rectangle corresponds to difference in relative abundance of that OTU between the zones, with blue indicating a higher relative abundance south of the PF, and red a higher abundance north of the PF. OTUs from clades or taxonomic ranks of interest have been grouped, with labels in bold and groups separated by gray lines. Groups and OTUs with a low contribution to variance which were not grouped are unlabeled. OTUs from each size fraction have also been grouped, with labels in black outline and size fractions separated by thick black lines. The data used to generate this figure are given in the supplementary material (PF-OTUs-SIMPER.csv).

SIMPER analysis found that no single OTU contributed more than 2.9% of variance, and 74% of variance was contributed by OTUs with a contribution less than 1%. There was also a large difference in the contribution to variance of the three size fractions, with approximately 52% of all variance contributed by OTUs from the 3.0 μm fraction, 37% by the 0.8 μm fraction, and 9% by the 0.1 μm fraction. Notably, OTUs within several taxonomic groups that had high contribution to variance covaried in their relative representation in the NZ and SZ. For example, Bacteroidetes and GSO-EOSA-1 representatives were on average more abundant in the SZ; while *Prochlorococcus* and *Synechococcus* species, SAR11

and SAR116 were on average more abundant in the NZ (Fig. 1.4). Some groups, such as the Alteromonadales, had variable relative representation depending on size fraction.

#### 1.4.3 Fragment recruitment to verify OTU identification

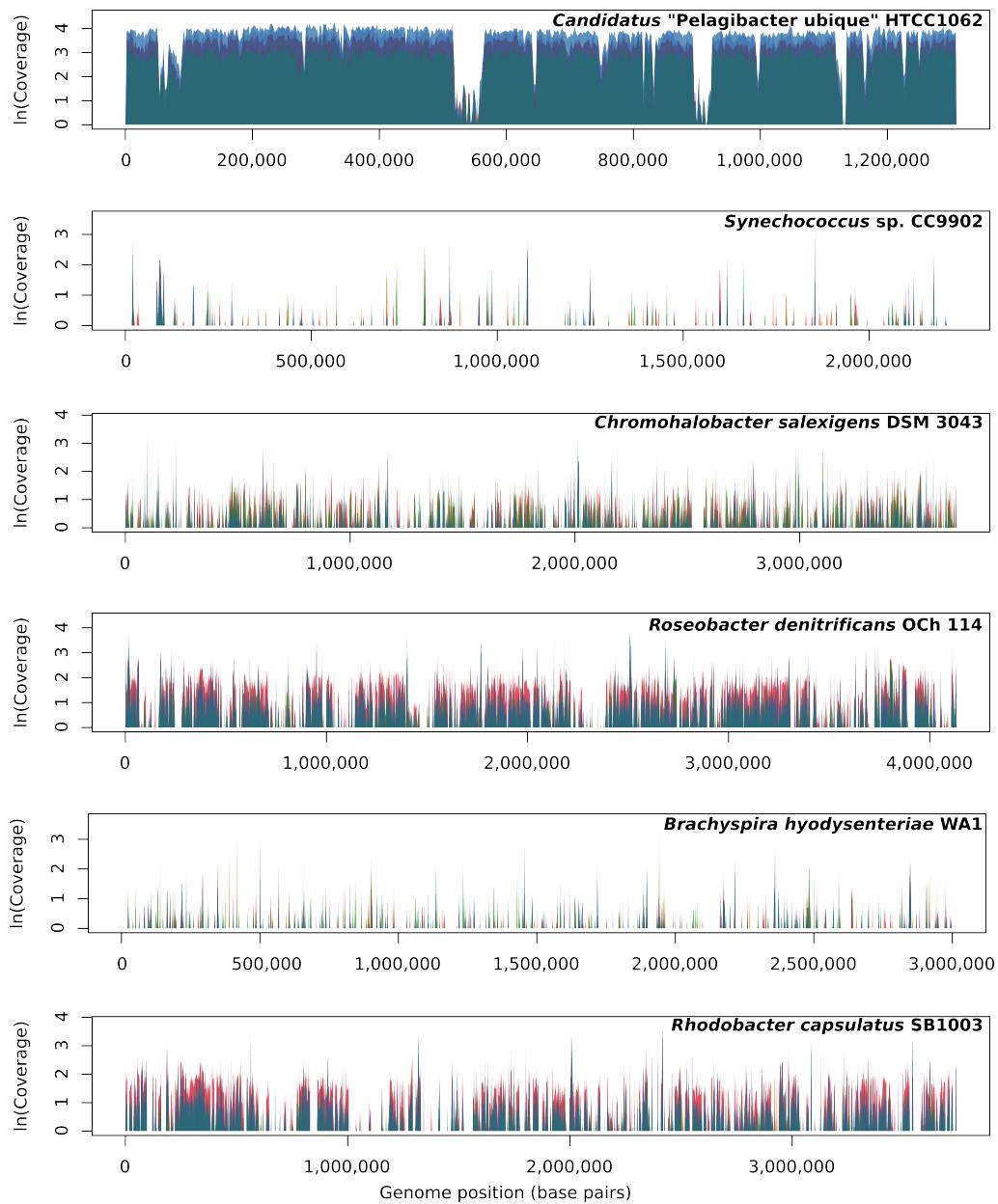
As predicted, read recruitment to reference genomes produced fairly even coverage, especially in the high-abundance OTUs. This indicates that OTU identification was accurate and unlikely to be confounded by small regions of genomic identity. Fig. 1.5 gives examples of read recruitment plots to three high-abundance and three low-abundance OTUs.

#### 1.4.4 Additional samples to test alternative “polynya hypothesis”

The ANOSIM analysis strongly supported the “PF hypothesis” ( $R = 0.44$ ,  $p = 0.002$ ), i.e. that the PF was primarily responsible for the difference observed between the NZ and SZ, over the “polynya hypothesis” ( $R = 0.29$ ,  $p = 0.005$ ) that the influence of the Mertz Glacier polynya was responsible. This also provides evidence that the PF effect is robust over a longitudinal range and over time.

#### 1.4.5 Functional analysis of metagenomic data

ANOSIM analysis of the samples’ KEGG ortholog group and module profiles revealed that the zones had significantly different functional potential (ortholog group:  $R = 0.64$ ,  $p = 0.001$ ; module:  $R = 0.87$ ,  $p = 0.001$ ). SIMPER was performed on the profiles in order to identify the specific functional differences between the zones. The highest-contributing modules are given in Table 1.5, and a complete list in the supplementary material (PF-modules-SIMPER.csv). The highest-contributing ortholog groups are given in Table 1.6, and a complete list in the supplementary material (PF-ortholog-groups-SIMPER.csv). No single ortholog group or module contributed more than 2.2% of the variance. There was a strong trend for ortholog groups and modules with higher contributions to variance to be overrepresented in the NZ in the 3.0  $\mu\text{m}$  fraction but the SZ in the smaller fractions, indicating that the functional diversity of each zone was strongly segregated by size fraction.



**Figure 1.5:** Examples of read recruitment (fragment recruitment) plots of metagenomic reads to reference genomes. Coloured regions indicate coverage for each of six example samples: pink sample 346; orange 347; red 358; green 361; blue 362. The even distribution of recruitment sites across the genome suggests that OTUs identifications are relatively accurate; an OTU identification resulting from a small region of genomic identity would generate a plot with read recruitment concentrated in small islands.

**Table 1.5:** The thirty KEGG modules with the highest contributions to the difference between the NZ and SZ. Abundances are zonal averages and have been standardised and log-transformed.

KEGG module	Abundance South	Abundance North	Contribution to variance (%)
Photosystem II	0.42	0.57	2.21
Complex I (NADH dehydrogenase), NADH dehydrogenase I/diaphorase subunit of the bidirectional hydroge-nase	0.01	0.24	1.80
Photosystem I	0.43	0.34	1.70
Pyrimidine deoxyribonucleotide biosynthesis, CDP/CTP → dCDP/dCTP,dTDP/dTTP	0.51	0.66	1.16
Histidine degradation, histidine → N-formiminoglutamate → glutamate	0.42	0.31	1.14
Methionine salvage pathway	0.29	0.43	1.14
sn-Glycerol 3-phosphate transport system	0.29	0.16	1.11
Complex I (NADH dehydrogenase), NADH dehydrogenase I	1.08	1.05	1.06
Branched-chain amino acid transport system	0.79	0.83	0.96
Dipeptide transport system	0.14	0.02	0.95
Adenine nucleotide biosynthesis, IMP → ADP/dADP,ATP/dATP	0.62	0.74	0.95
Glycine betaine/proline transport system	0.66	0.56	0.94
Sulfur reduction, sulfate → H <sub>2</sub> S	0.54	0.44	0.91
Simple sugar transport system	0.46	0.39	0.90
Peptides/nickel transport system	0.99	0.98	0.89
Ribosome, eukaryotes	0.26	0.27	0.89
Multiple sugar transport system	0.55	0.55	0.86
Type II general secretion system	0.21	0.21	0.82
Sulfonate/nitrate/taurine transport system	0.45	0.37	0.82
Guanine nucleotide biosynthesis, IMP → GDP/dGDP,GTP/dGTP	0.72	0.82	0.81
RNA polymerase II, eukaryotes	0.11	0.20	0.76
Histidine biosynthesis, PRPP → histidine	0.94	0.86	0.76
Putrescine transport system	0.18	0.09	0.72
Leucine biosynthesis, pyruvate → 2-oxoisovalerate → leucine	1.29	1.37	0.71
C5 isoprenoid biosynthesis, non-mevalonate pathway	0.70	0.77	0.71
Leucine degradation, leucine → acetoacetate + acetyl-CoA	0.64	0.59	0.71
Thiamine transport system	0.13	0.05	0.69
Spliceosome, 35S U5-snRNP	0.18	0.20	0.68
Cytochrome b6f complex	0.14	0.12	0.67
Menaquinone biosynthesis, chorismate → menaquinone	0.25	0.27	0.66

**Table 1.6:** The thirty KEGG ortholog groups with the highest contribution to the difference between the NZ and SZ. Abundances are zonal averages and have been standardised and log-transformed. As each ortholog group on each size fraction was encoded as a separate variable in the SIMPER analysis, the size fraction is given after each ortholog group name.

KEGG ortholog group	Abundance South	Abundance North	Contribution to variance (%)
Hypothetical protein 3.0 µm	0.11	0.24	0.26
Hypothetical protein 0.8 µm	0.68	0.57	0.24
Ribonucleoside-diphosphate reductase alpha chain [EC:1.17.4.1] 0.8 µm	0.17	0.24	0.15
DNA polymerase III subunit alpha [EC:2.7.7.7] 0.8 µm	0.25	0.19	0.14
Hypothetical protein 0.1 µm	0.26	0.24	0.12
Proline dehydrogenase / delta 1-pyrroline-5-carboxylate 0.8 µm	0.10	0.04	0.12
Aminomethyltransferase [EC:2.1.2.10] 0.8 µm	0.25	0.19	0.12
Ribonucleoside-diphosphate reductase alpha chain [EC:1.17.4.1] 3.0 µm	0.02	0.08	0.12
Sarcosine oxidase, subunit alpha [EC:1.5.3.1] 0.8 µm	0.22	0.17	0.12
Integrator complex subunit 6 3.0 µm	0.07	0.05	0.11
Multicomponent Na <sup>+</sup> :H <sup>+</sup> antiporter subunit D 0.8 µm	0.11	0.05	0.11
Glutamine synthetase [EC:6.3.1.2] 0.8 µm	0.24	0.19	0.11
Pyruvate dehydrogenase E1 component [EC:1.2.4.1] 0.8 µm	0.15	0.10	0.11
Cobaltochelatase CobN [EC:6.6.1.2] 0.8 µm	0.11	0.06	0.11
Formate dehydrogenase, alpha subunit [EC:1.2.1.2] 0.8 µm	0.15	0.10	0.11
DNA-directed RNA polymerase subunit beta [EC:2.7.7.6] 3.0 µm	0.03	0.08	0.11
Glutamate synthase (NADPH/NADH) large chain [EC:1.4.1.13 1.4.1.14] 0.8 µm	0.25	0.22	0.11
Dimethylglycine dehydrogenase [EC:1.5.99.2] 0.8 µm	0.17	0.14	0.11
Flagellin 0.8 µm	0.06	0.10	0.10
DNA-directed RNA polymerase subunit beta [EC:2.7.7.6] 3.0 µm <sup>7</sup>	0.03	0.08	0.10
Photosystem II PsbA protein 0.8 µm	0.01	0.06	0.09
Aldehyde dehydrogenase (NAD+) [EC:1.2.1.3] 0.8 µm	0.17	0.13	0.09
Glutamate synthase (NADPH/NADH) large chain [EC:1.4.1.13 1.4.1.14] 3.0 µm	0.02	0.07	0.09
Thymidylate synthase (FAD) [EC:2.1.1.148] 0.8 µm	0.02	0.06	0.09
Topoisomerase IV subunit A [EC:5.99.1.-] 0.8 µm	0.11	0.07	0.09
DNA mismatch repair protein MutS 0.8 µm	0.13	0.08	0.09
Glutamate dehydrogenase [EC:1.4.1.2] 0.8 µm	0.07	0.03	0.09
DNA polymerase I [EC:2.7.7.7] 0.1 µm	0.12	0.11	0.09
GTP-binding protein 0.8 µm	0.26	0.21	0.09
GTP-binding protein 3.0 µm	0.03	0.07	0.09

## 1.5 Discussion

### 1.5.1 Taxonomic groups differentiating the zones

#### 1.5.1.1 GSO-EOSA-1

The GSO-EOSA-1 cluster, represented in RefSeq by “*Candidatus Vesicomyosocius okutanii*” strain HA and “*Candidatus Ruthia magnifica*” strain Cm. (*Calyptogena magnifica*) (Walsh *et al.*, 2009), made a large contribution to variance between the NZ and SZ, with higher abundance in the SZ: relative abundances of GSO-EOSA-1 in the SZ were 5.2%, 3.4% and 0.25% in the 0.1, 0.8 and 3.0  $\mu\text{m}$  size fractions respectively, compared to 1.1%, 0.84% and 0.30% in the NZ (Table 1.3). The contribution to variance of this group was highest in the 0.1  $\mu\text{m}$  size fraction, followed by 0.8  $\mu\text{m}$  and 3.0  $\mu\text{m}$  (Table 1.4). This pattern most likely represents a small cell size and lack of association with particulate matter.

“*Ca. R. magnifica*” and “*Ca. V. okutanii*” are chemoautotrophic endosymbionts of deep-sea bivalves (Kuwahara *et al.*, 2007; Newton *et al.*, 2007) and are thus unlikely to be present in open ocean surface waters. However, GSO-EOSA-1 representative ARCTIC96BD-19 has recently been reported at high abundance in Antarctic coastal waters (Ghiglione and Murray, 2011; Grzymski *et al.*, 2012). To investigate the correct taxonomic placement of the GSO-EOSA-1 representative in this study, reads with identity to “*Ca. R. magnifica*” and “*Ca. V. okutanii*” were compared to the 16S rRNA gene of the SUP05 GSO-EOSA-1 isolate identified by Walsh *et al.* (2009) (GenBank accession GQ351268.1) by BLASTN with default settings and an E-value threshold of  $10^{-3}$ . 16S rRNA sequences from a reference set of endosymbiotic (“*Ca. R. magnifica*”, “*Ca. V. okutanii*”, *Solemya reidi* gill symbiont (GenBank accession L25709.1), *Bathymodiolus septemtierum* thioautotrophic gill symbiont (GenBank accession AB036709.1), *Vesicomya gigas* gill symbiont (GenBank accession AF035726.1) and *Ridgeia piscesae* endosymbiont (GenBank accession JX570608.1)) and planktonic (SUP05, ARCTIC96BD-19 (GenBank accession AF354606.1)) GSO-EOSA-1 representatives were obtained. A neighbour-joining phylogenetic tree was constructed using ARB (Ludwig *et al.*, 2004), using *Escherichia coli* IHE3034 (GenBank accession CP001969.1) as a gammaproteobacterial outgroup to root the tree. The GSO-EOSA-1 16S rRNA gene affiliated metagenome reads were then placed in this tree by the “ARB parsimony” insertion method, with alignments manually verified. The majority of 16S rRNA genes from all samples and size fractions clustered with ARTIC96BD-19 (Fig. 1.6), indicating this is the dominant GSO-EOSA-1 representative.



**Figure 1.6:** Neighbour-joining tree of GSO-EOSA-1-like 16S rRNA gene sequences from the metagenomes in this study. Sequences labeled in black text are reads from the metagenomes. Red labels are 16S rRNA gene sequences from Gammaproteobacterial Sulfur Oxidizers (GSO) and other Gammaproteobacteria. Bootstrap confidence values for each node are indicated as percentages. The tree was constructed using ARB (Ludwig *et al.*, 2004).

Single-cell genomic analysis of ARCTIC96BD-19 from global mesopelagic waters indicates the lineage is probably mixotrophic, able to couple carbon fixation to oxidation of reduced sulphur compounds as well as assimilate organic carbon (Swan *et al.*, 2011). GSO-EOSA-1 cytochrome C oxidase (CoxII) has been identified in a winter metaproteome of Antarctic Peninsula coastal waters, suggesting the capacity for aerobic respiration (Williams *et al.*, 2012). Taken together, this evidence suggests the GSO-EOSA-1 representative in Antarctic coastal waters is a versatile chemolithoautotroph capable of aerobic respiration.

It has been proposed that during the winter months, chemolithoautotrophy is dominant over photoautotrophy as the major carbon fixation process in AZ waters due to the lack of available light, both from seasonal darkness and ice shading (Grzymski *et al.*, 2012). The high relative abundance of GSO-EOSA-1 in SZ compared to NZ waters may therefore represent the remnants of an annual winter increase in population in the marginal ice zone which does not occur in the open ocean.

#### 1.5.1.2 Ammonia-oxidizing Crenarchaeota

*Nitrosopumilus maritimus* SCM1 and *Cenarchaeum symbiosum* are chemolithoautotrophic, nitrifying members of the Marine Group I Crenarchaeota (MGI) (Preston *et al.*, 1996; Walker *et al.*, 2010). These species were the only representatives in the reference database used in this study of the Ammonia-Oxidizing Archaea (AOA). The contribution of *C. symbiosum* to total AOA abundance was low (0–0.015% across all fractions and zones, see supplementary material (PF-a11-OTUs.txt)). As *C. symbiosum* is a sponge symbiont (Preston *et al.*, 1996) and given the poor representation of AOA in RefSeq, it is likely this OTU has attracted sequences originating from planktonic AOA and *C. symbiosum* itself is not present. AOA were moderate contributors to variance between the NZ and SZ, and were overrepresented in the SZ in all size fractions (Fig. 1.4). As with the GSO-EOSA-1 cluster, MGI have been proposed to be abundant chemolithoautotrophs and therefore major drivers of winter carbon fixation in Antarctic coastal waters (Grzymski *et al.*, 2012; Williams *et al.*, 2012).

Sample 353 had a particularly high relative abundance of *N. maritimus* OTUs (7.5% of the 0.1  $\mu\text{m}$  fraction; 0.8  $\mu\text{m}$ : 11%; 3.0  $\mu\text{m}$ : 12%). This sample was taken closer to the Antarctic continent (3.7 km) than any other, in relatively shallow (180 m) waters 17.6 km from the Mertz Glacier. The high abundance of ammonia oxidizers may reflect an input of ammonia from terrestrial sources (e.g. penguin guano), or resuspension of benthic sediments in which MGI are abundant (Bowman and McCuaig, 2003) by near-shore turbulence and iceberg scouring. Breakdown of water column stratification has been previously suggested as a cause of increased AOA abundance in Antarctic coastal surface waters (Kalantra *et al.*, 2009).

#### 1.5.1.3 Cyanobacteria

OTUs assigned to the cyanobacterial genera *Prochlorococcus* and *Synechococcus* were overrepresented in the NZ in all size fractions (Fig. 1.4). The mean relative abundance of cyanobacteria in samples 367 and 368, the two northernmost samples, was strikingly higher than the mean abundance across all other samples in the NZ. *Synechococcus* sp.

CC9902 alone composed greater than 22% of the 0.8 µm fraction in these samples, consistent with *Synechococcus* species' cell diameters of up to 0.9 µm (Waterbury *et al.*, 1985). The high abundance of both cyanobacterial genera on the 3.0 µm fraction has previously been reported (Lauro *et al.*, 2011) and may be attributable to aggregation (Lomas and Moran, 2011).

Samples 367 and 368 were separated from the other samples north of the PF by the STF. While the STF was not a significant boundary on the assemblage level (ANOSIM,  $p > 0.05$ ), it may mark a significant biogeographic boundary for these cyanobacteria. *Synechococcus* and *Prochlorococcus* together represent a large proportion of both phytoplankton abundance and carbon fixation in temperate and tropical waters, in many regions contributing more than half of total primary production (Liu *et al.*, 1997, 1998; André *et al.*, 1999). The role of the STF in determining the latitudinal range of *Synechococcus* and *Prochlorococcus* is therefore important, as it will affect models of ocean productivity under climate change, and warrants further investigation. Despite the high abundance of cyanobacteria north of the STF, they were also a significant feature of the SAZ; for example, *Synechococcus* sp. CC9902 composed 3–5% of the 0.8 µm fraction in SAZ samples.

These results confirm previous reports (Marchant *et al.*, 1987; Ghiglione and Murray, 2011) that cyanobacteria occur as far south as the Antarctic coast, albeit at very low abundances ( $10^3$ – $10^4$  cells/L, Marchant *et al.* (1987)), and contradicts the assumption that *Prochlorococcus* is restricted to tropical and subtropical waters within 40° of latitude (Partensky *et al.*, 1999). Cyanophage proteins have also been detected in a metaproteomic analysis of Antarctic Peninsula coastal surface waters (Williams *et al.*, 2012).

#### 1.5.1.4 SAR11 and SAR116 clades

“*Candidatus Pelagibacter ubique*” HTCC1062 was the most abundant OTU across all samples and fractions (NZ average: 62%, 25% and 24% of the 0.1 µm, 0.8 µm and 3.0 µm fractions respectively; SZ: 59%, 22% and 18%) and one of the most significant contributors to variance between the NZ and SZ (Fig. 1.4). “*Ca. P. ubique*” HTCC1062 is a good representative of total SAR11 abundance in this study, as it is a member of the SAR11 phylotype most abundant in SO waters (Brown *et al.*, 2012). The high abundance of SAR11 in the 0.1 µm fraction is consistent with the small size of SAR11 cells (Rappé *et al.*, 2002). The higher representation in the NZ may reflect the competitiveness of SAR11 members in regions with low Dissolved Organic Carbon (DOC) concentrations due to low primary productivity (Giovannoni *et al.*, 2005; Alonso and Pernthaler, 2006), such as the High Nutrient, Low Chlorophyll (HNLC) SAZ. Overall, these findings are consistent with reports that SAR11 is ubiquitous in the world’s oceans (Mary *et al.*, 2006; Carlson *et al.*, 2009) and more abundant north of the ACC (Giebel *et al.*, 2009).

OTUs of “*Candidatus Puniceispirillum marinum*” from the SAR116 clade were a moderate contributor to variance between the NZ and SZ with higher abundance in the NZ (Fig. 1.4). A genomic analysis reported “*Ca. P. marinum*” IMCC1322 to be a metabolic generalist with genes for aerobic CO fixation, C1 metabolism and a “*Ca. P. ubique*”-like dimethylsulfoniopropionate (DMSP) demethylase, suggesting SAR116 and SAR11 occupy similar ecological niches (Oh *et al.*, 2010). In the Scotia Sea, SAR116 abundance (as determined with Fluorescence *In Situ* Hybridization (FISH)) was reported to be higher in more

productive waters where SAR11 numbers were lower (Topping *et al.*, 2006). However, this analysis across an extended latitudinal transect indicates that overall SAR11 and SAR116 have similar biogeographic distributions.

#### 1.5.1.5 Bacteroidetes

OTUs of the phylum Bacteroidetes, in particular members of the class Flavobacteria, were found to be abundant (NZ average: 1.2%, 5.0% and 6.9% of the 0.1  $\mu\text{m}$ , 0.8  $\mu\text{m}$  and 3.0  $\mu\text{m}$  fractions respectively; SZ: 2.3%, 9.8% and 9.1%) and significant contributors to variance between the NZ and SZ (Fig. 1.4). Flavobacteria have been previously reported to compose the majority of both Bacteroidetes (Murray and Grzynski, 2007) and total planktonic biomass (Abell and Bowman, 2005) in the SO, as well as being abundant in sea ice (Brown and Bowman, 2001). As heterotrophic degraders of High Molecular Weight (HMW) compounds in the form of both Dissolved Organic Matter (DOM) and Particulate Organic Matter (POM) (Kirchman, 2002), marine Flavobacteria are major components of marine aggregates (Rath *et al.*, 1998; Crump *et al.*, 1999; Zhang *et al.*, 2007) and are associated with blooming phytoplankton. This has been confirmed in Antarctic waters, with a positive relationship between Flavobacteria abundance and chl *a* fluorescence (Williams *et al.*, 2013). The higher abundance of Flavobacteria OTUs on the 0.8  $\mu\text{m}$  and 3.0  $\mu\text{m}$  fractions reflects their association with particulate matter. Similar size partitioning of SO Flavobacteria has previously been reported (Abell and Bowman, 2005).

The higher abundance of OTUs of Flavobacteria in the SZ may reflect an input of cells from melting sea ice (Brown and Bowman, 2001), but is probably driven primarily by the higher rates of primary productivity in the south and the role of the Flavobacteria as degraders of HMW DOM and POM. Because deposition of marine snow is a major route for sequestration of fixed carbon in the ocean (e.g. Hessen *et al.*, 2004), the Flavobacteria that associate with this particulate matter represent a remineralising shunt, decreasing carbon export by this route.

#### 1.5.1.6 Rhodobacterales

Members of the order Rhodobacterales were abundant (NZ average: 1.2%, 10% and 5.5% of the 0.1  $\mu\text{m}$ , 0.8  $\mu\text{m}$  and 3.0  $\mu\text{m}$  fractions respectively; SZ: 1.6%, 13% and 7.9%) and high contributors to variance, overrepresented in the SZ on all size fractions. As several members of the Roseobacter clade have been shown to have symbiotic relationships with marine eukaryotic algae (Buchan *et al.*, 2005; Wagner-Döbler and Biebl, 2006), and their abundance in the SO has previously been linked to phytoplankton blooms (West *et al.*, 2008; Obernosterer *et al.*, 2011), it is likely that their overrepresentation in the SZ is related to the higher density of phytoplankton in the AZ.

OTUs of *Roseobacter denitrificans* Och114 and *Silicibacter pomeroyi* DSS-3 were consistently the most abundant Roseobacter clade representatives, and indeed among the most abundant OTUs identified in this study (Table 1.3). *R. denitrificans* and *S. pomeroyi* fall within a subclade of Aerobic Anoxygenic Phototroph (AAP) members of the Roseobacter clade (Swingley *et al.*, 2007). These species have diverse mixotrophic metabolisms, with genomic and experimental evidence of photoheterotrophic respiration of organic carbon, fixation of CO<sub>2</sub>, oxidation of CO, oxidation of reduced sulfur compounds, and utilization

of the abundant marine osmolyte DMSP (King, 2003; Moran *et al.*, 2004; Wagner-Döbler and Biebl, 2006; Swingley *et al.*, 2007; Brinkhoff *et al.*, 2008; Howard *et al.*, 2008). This metabolic diversity suggests a complex ecological role, particularly with respect to the capture and release of climatically active gases ( $\text{CO}_2$ , CO, dimethylsulfide) involved in carbon and sulfur cycling.

*R. denitrificans* encodes transporters and metabolic genes for phosphonate, as well as the more usual phosphate (Swingley *et al.*, 2007). The SO south of the PF is not phosphate-limited, but does have a much lower concentration of phosphate relative to nitrate when compared to the nitrate-limiting SAZ (Weber and Deutsch, 2010). Thus, the ability to utilise phosphonate as an alternative phosphorous source may give *R. denitrificans* a competitive advantage in the SZ (Huang *et al.*, 2005).

#### 1.5.1.7 Alteromonadales

Members of the gammaproteobacterial order Alteromonadales were large contributors to variance across the PF. Most Alteromonadales were overrepresented in the SZ, but some were overrepresented in the NZ on the 3.0  $\mu\text{m}$  fraction (Fig. 1.4). *Colwellia psychrerythraea* 34H was one of the most abundant OTUs in the Alteromonadales and exhibited this distribution (NZ average: 0.14%, 2.2% and 16% of the 0.1  $\mu\text{m}$ , 0.8  $\mu\text{m}$  and 3.0  $\mu\text{m}$  fractions respectively; SZ: 0.52%, 5.1% and 10%). *C. psychrerythraea* 34H was isolated from Arctic sediment, grows well at low temperatures and secretes extracellular polysaccharides (Huston *et al.*, 2000; Junge *et al.*, 2003; Methé *et al.*, 2005). Similar to other *Colwellia* species grown under laboratory conditions, *C. psychrerythraea* cells have widths of 0.4–0.8  $\mu\text{m}$  and lengths of 1.5–4.5  $\mu\text{m}$  (Jung *et al.*, 2006). Growth temperature can have a major impact on cell morphology, enzyme secretion and global gene expression in psychrophiles (e.g. Feller and Gerdau, 2003; Junge *et al.*, 2003; Williams *et al.*, 2011; Cavicchioli, 2006; Campanaro *et al.*, 2011). Moreover, marine bacteria can alter their cell dimensions in response to nutrient flux (e.g. Kjelleberg *et al.*, 1987). It is therefore possible that the populations of Alteromonadales captured on the 3.0  $\mu\text{m}$  filters (overrepresented in the NZ) had different physiological properties to those on the 0.1 and 0.8  $\mu\text{m}$  filters (overrepresented in the SZ).

#### 1.5.1.8 Verrucomicrobia

Two representatives of the phylum Verrucomicrobia, *Coraliomargarita akajimensis* and *Akkermansia* sp. Muc-30, were moderate contributors to variance and overrepresented in the SZ (Fig. 1.4). Surprisingly given the small cell size of *C. akajimensis* (Yoon *et al.*, 2007), its contribution to variance increased with size fraction. A global survey reported a similar fractionation pattern, and suggested marine Verrucomicrobia may be predominantly particle attached (Freitas *et al.*, 2012). However, little else is known about the distribution and ecological roles of marine Verrucomicrobia (Freitas *et al.*, 2012).

### 1.5.2 Functional capacities differentiating the zones

A number of modules with transport functions (sn-glycerol 3-phosphate transport system, dipeptide transport system, peptides/nickel transport system, simple sugar transport system, sulfonate/nitrate/taurine transport system) were overrepresented in the SZ

(Table 1.5). As the genomes of copiotrophic bacteria have evolved to have a higher number of narrow-specificity transporters relative to oligotrophic genomes (Lauro *et al.*, 2009), these differences may reflect the higher nutrient availability and thus a dominance of copiotrophs in the SZ. The taxonomic contributors to these modules were varied, although members of the Rhodobacterales were prominent (Fig. 1.7).

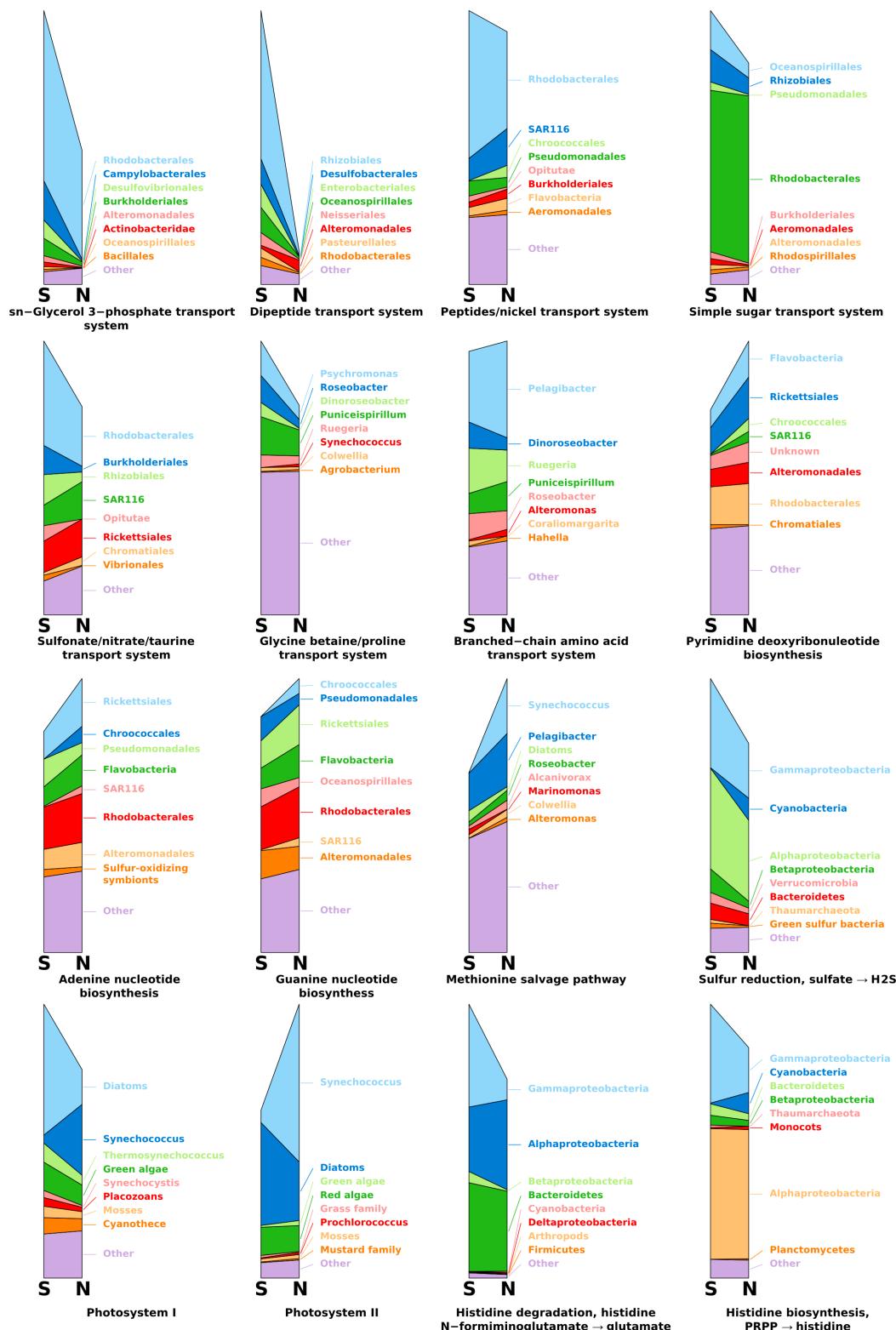
The glycine betaine/proline transport module was also overrepresented in the SZ, though this probably reflects glycine betaine's role as an osmo- and cryoprotectant in the colder SZ waters. This is supported by the major taxonomic contributor to this module, genus *Psychromonas*, which includes several psychrophilic species.

One exception to this pattern was the branched-chain amino acid transport system module, overrepresented in the NZ. The genera *Pelagibacter* and *Puniceispirillum* were major contributors to this module's overabundance in the NZ (Fig. 1.7). As both SAR11 (Giovannoni *et al.*, 2005) and SAR116 (Grote *et al.*, 2011) representatives encode branched-chain amino acid transporters, the abundance of this module is likely to represent taxonomic differences between the zones.

Biosynthesis pathways for all major nucleic acids (pyrimidine deoxyribonucleotide biosynthesis, adenine nucleotide biosynthesis, guanine nucleotide biosynthesis) were consistently high contributors to variance and overabundant in the NZ. This pattern seems inconsistent with the more oligotrophic nature of the NZ, as oligotrophic cells generally have smaller genomes (Lauro *et al.*, 2009) and slower growth rates than copiotrophs, and would therefore be expected to require a lower rate of de novo nucleotide biosynthesis. A possible explanation for this is that SZ cells have higher availability of extracellular DNA as a byproduct of decaying phytoplankton (Lomas and Moran, 2011), which can be imported and salvaged for nucleic acids (Paul *et al.*, 1988) thus reducing the requirement for de novo synthesis. No single taxonomic group contributed a large fraction of the difference in this module (Fig. 1.7), suggesting this is a widespread adaptation.

The methionine salvage pathway module had a large contribution to variance between the zones and was overrepresented north of the PF. This may reflect the higher availability of DMSP in the SZ as a byproduct of blooming eukaryotic algae. DMSP is a major carbon and sulfur source for marine microorganisms, and is commonly assimilated by bacteria through demethylation to methylmercaptopropionate (MMPA), followed by further catabolism to the climatically important compounds dimethylsulfide or methanethiol (review in Curson *et al.*, 2011). However, when DMSP is scarce, MMPA may be derived from methionine through the alternative methionine salvage pathway (Reisch *et al.*, 2011). The genus *Synechococcus*, a noted contributor to marine DMSP uptake and assimilation (Vila-Costa *et al.*, 2006), was a very high contributor to the abundance of this module in the NZ (Fig. 1.7), suggesting *Synechococcus* species may use this route when DMSP is unavailable.

The sulfur reduction module was overrepresented in the SZ, and it is likely that this result is strongly driven by taxonomic differences. While the taxonomic breakdown indicated a large number of genera contributed to the difference, the Gammaproteobacteria were the highest-contributing class (Fig. 1.7). This module also includes the assimilatory sulfate reduction pathway, which is widespread in marine bacteria, but is absent from SAR11, with known representatives reported to lack genes for assimilatory sulfate reduction (cysDNCHIJ) (Tripp *et al.*, 2008). The higher relative abundance of SAR11 in the NZ



**Figure 1.7:** Decomposition of KEGG modules of interest to contributing classes, orders or genera. The left side of each stack (S) indicates the proportion of the module abundance contributed by each class, order or genus in the South Zone, while the right side (N) represents the North Zone. As the contributions are relative and represent unitless module abundances, no axis is given and proportions are not comparable between modules. Contributing classes, orders or genera are arranged in descending order of the difference in the relative contributions between the zones. Only the eight highest contributors for each module are shown, with the remainder collapsed into the “Other” group. The taxonomic ranks to which each module was decomposed are as follows: sn-glycerol 3-phosphate transport, peptide-nickel transport, simple sugar transport and sulfonate/nitrate/taurine transport were decomposed to order; glycine betaine/proline transport and branched-chain amino acid transport to genus; pyrimidine deoxyribonucleotide biosynthesis, adenine nucleotide biosynthesis and guanine nucleotide biosynthesis to order; methionine salvage to genus; sulphur reduction to class; photosystem I and photosystem II to genus; histidine degradation to glutamate and histidine biosynthesis to class.

may therefore contribute to the lower abundance of genes for assimilatory sulfate reduction in that zone.

The sulfur reduction module also included adenylylsulfate reductase (APS reductase, encoded by *aprAB*), an enzyme implicated in sulfite detoxification during heterotrophic growth on organosulfonates (Meyer and Kuever, 2007) (N.B. in recent KEGG releases, *aprA* is no longer included in this module). As the GSO-EOSA-1 representative SUP05 has been found to encode APS reductase, the overabundance of this module may reflect sulfur oxidation through the reverse dissimilatory sulfate reduction pathway (Walsh *et al.*, 2009). Also, Roseobacter clade bacteria are involved in the decomposition of abundant organic sulfur compounds (e.g. DMSP, organosulfonates), and hence have been accorded an important role in marine sulfur cycling (Moran *et al.*, 2007).

The photosystem II module was overrepresented in the NZ. Given the underrepresentation of cyanobacterial OTUs in the SZ, this may reflect a dominance of primary production by eukaryotic algae south of the PF and cyanobacteria to the north. Decomposition of the taxonomic affiliations of ortholog groups contributing to this module found OTUs of *Synechococcus* and *Prochlorococcus* to be major contributors to the difference (Fig. 1.7). Variation in the photosystem I module, which was marginally overrepresented in the SZ, could largely be attributed to diatoms and other eukaryotic phytoplankton (Fig. 1.7), again supporting a dominance of eukaryotic phytoplankton in SZ primary production. Diatoms have previously been reported at higher abundance south of the PF, and their distribution is likely to be linked to the higher concentration of dissolved silica in that region (Trull *et al.*, 2001). As both eukaryotic phytoplankton and cyanobacteria would be expected to encode both complete photosystems, the differences in module abundance probably reflect the degree of similarity between the photosystem I and II genes in the KEGG database and those found in the sampled environments.

The histidine degradation to glutamate module, which comprises four ortholog groups mediating the degradation of histidine to glutamate via N-formiminoglutamate, was overrepresented in the SZ. The histidine biosynthesis module was also overrepresented in the SZ. While the concentration of dissolved histidine in the SO is generally low (Kawahata and Ishizuka, 2000), blooming eukaryotic phytoplankton (which are more prevalent in the SZ) may deplete nitrate while releasing Dissolved Free Amino Acids (DFAA). As DFAA become available, they are used by bacteria to sense the decaying bloom. Histidine may therefore act as a proxy for DFAA to regulate the expression of bacterial aminopeptidases,

which are involved in lysing diatoms (Bidle and Azam, 2001). The class Bacteroidetes, while a small contributor to the histidine biosynthesis module in the SZ, was a large contributor to histidine degradation (Fig. 1.7), supporting an association between Bacteroidetes and phytoplanktonic bloom products. It is also possible that uptake and degradation of histidine to glutamate (which generates ammonia as a by-product) may function as a limited nitrogen source.

### 1.5.2.1 Conclusions

These results show that there are major taxonomic and functional differences across the PF. The differences in functional potential between the NZ and SZ reflect both their taxonomic profiles and fundamental trophic and ecological differences. In particular, they provide genomic support that the NZ is more oligotrophic than the SZ (Pollard *et al.*, 2002; Giovannoni *et al.*, 2005; Alonso and Pernthaler, 2006; Lauro *et al.*, 2009), and are consistent with the observation that primary production is higher south of the PF (Strutton *et al.*, 2000; Williams *et al.*, 2010).

These results do not exclude the possibility that other major SO fronts, particularly the STF and SAF, are also significant biogeographic boundaries, as has been reported in some previous reports for specific taxonomic groups (e.g. Abell and Bowman, 2005). While the sampling resolution in this study was not sufficient to resolve the effects of other fronts, there are some indications in the data of further structure within the zones. The two samples north of the STF had significantly larger cyanobacterial populations than the remaining NZ samples (see discussion of *Prochlorococcus* and *Synechococcus*, above). Future sampling across these fronts at higher resolution will provide the data necessary to investigate finer biogeographic patterns.

The nature and function of microbial communities in the SO are of global significance. Knowledge of these communities and their biogeographic drivers has relevance for understanding and predicting the long-term effects of climate change. These findings provide a basis for predicting how climate change-driven shifts in the SO may affect microbial communities; in particular, the effects of changes in the nature and location of the ACC on the ecosystem functions of SO microorganisms.

# References

- Abell G. G. J. and Bowman J. P. (2005). Ecological and biogeographic relationships of class Flavobacteria in the Southern Ocean. *FEMS Microbiology Ecology*, 51:265–277.
- Alonso C. and Pernthaler J. (2006). Roseobacter and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environmental Microbiology*, 8:2022–2030.
- André J. M., Navarette C., Blanchot J., and Radenac M. H. (1999). Picophytoplankton dynamics in the equatorial Pacific: Growth and grazing rates from cytometric counts. *Journal of Geophysical Research*, 104:3369–3380.
- Angly F. E., Felts B., Breitbart M., Salamon P., Edwards R. A., Carlson C., Chan A. M., Haynes M., Kelley S., Liu H., Mahaffy J. M., Mueller J. E., Nulton J., Olson R., Parsons R., Rayhawk S., Suttle C. A., and Rohwer F. (2006). The marine viromes of four oceanic regions. *PLoS Biology*, 4:e368.
- Angly F. E., Willner D., Prieto-Davó A., Edwards R. A., Schmieder R., Vega-Thurber R., Antonopoulos D. A., Barott K., Cottrell M. T., Desnues C., Dinsdale E. A., Furlan M., Haynes M., Henn M. R., Hu Y., Kirchman D. L., McDole T., McPherson J. D., Meyer F., Miller R. M., Mundt E., Naviaux R. K., Rodriguez-Mueller B., Stevens R., Wegley L., Zhang L., Zhu B., and Rohwer F. (2009). The GAAS Metagenomic Tool and Its Estimations of Viral and Microbial Average Genome Size in Four Major Biomes. *PLoS Computational Biology*, 5:e1000593.
- Bidle K. D. and Azam F. (2001). Bacterial control of silicon regeneration from diatom detritus: significance of bacterial ectohydrolases and species identity. *Limnology and Oceanography*, 46:1606–1623.
- Böning C. W., Dispert A., Visbeck M., Rintoul S. R., and Schwarzkopf F. U. (2008). The response of the Antarctic Circumpolar Current to recent climate change. *Nature Geoscience*, 1:864–869.
- Bowman J. P. and McCuaig R. D. (2003). Biodiversity, community structural shifts, and biogeography of prokaryotes within Antarctic continental shelf sediment. *Applied and Environmental Microbiology*, 69:2463–2483.
- Brinkhoff T., Giebel H.-A., and Simon M. (2008). Diversity, ecology, and genomics of the Roseobacter clade: a short overview. *Archives of Microbiology*, 189:531–539.
- Brown M. V. and Bowman J. P. (2001). A molecular phylogenetic survey of sea-ice microbial communities (SIMCO). *FEMS Microbiology Ecology*, 35:267–275.
- 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*, 8:595.
- Buchan A., González J. M., and Moran M. A. (2005). Overview of the marine Roseobacter lineage. *Applied and Environmental Microbiology*, 71:5665–5677.

- Campanaro S., Williams T. J., Burg D. W., De Francisci D., Treu L., Lauro F. M., and Caviglioli R. (2011). Temperature-dependent global gene expression in the Antarctic archaeon *Methanococcoides burtonii*. *Environmental Microbiology*, 13:2018–2038.
- Carlson C. A., Morris R., Parsons R., Treusch A. H., Giovannoni S. J., and Vergin K. (2009). Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *The ISME Journal*, 3:283–295.
- Caviglioli R. (2006). Cold-adapted archaea. *Nature Reviews Microbiology*, 4:331–343.
- Chiba S., Ishimaru T., Hosie G. W., and Fukuchi M. (2001). Spatio-temporal variability of zooplankton community structure off east Antarctica (90 to 160°E). *Marine Ecology Progress Series*, 216:95–108.
- Chou H. H. and Holmes M. H. (2001). DNA sequence quality trimming and vector removal. *Bioinformatics*, 17:1093–1104.
- Clarke K. R. and Warwick R. M. *Change in marine communities: an approach to statistical analysis and interpretation*. PRIMER-E, Plymouth, 2nd edition, 2001.
- Coleman M. L. M. and Chisholm S. W. S. (2010). Ecosystem-specific selection pressures revealed through comparative population genomics. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 107:18634–18639.
- Cox P. M., Betts R. A., Jones C. D., Spall S. A., and Totterdell I. J. (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, 408: 184–187.
- Crump B. C., Armbrust E. V., and Baross J. A. (1999). Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia River, its estuary, and the adjacent coastal ocean. *Applied and Environmental Microbiology*, 65:3192–3204.
- Curson A. R. J., Todd J. D., Sullivan M. J., and Johnston A. W. B. (2011). Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes. *Nature Reviews Microbiology*, 9:849–859.
- Dinsdale E. A., Edwards R. A., Hall D., Angly F., Breitbart M., Brulc J. M., Furlan M., Desnues C., Haynes M., Li L., McDaniel L., Moran M. A., Nelson K. E., Nilsson C., Olson R., Paul J., Brito B. R., Ruan Y., Swan B. K., Stevens R., Valentine D. L., Thurber R. V., Wegley L., White B. A., and Rohwer F. (2008). Functional metagenomic profiling of nine biomes. *Nature*, 452:629–632.
- Esper O. and Zonneveld K. A. F. (2002). Distribution of organic-walled dinoflagellate cysts in surface sediments of the Southern Ocean (eastern Atlantic sector) between the Subtropical Front and the Weddell Gyre. *Marine Micropaleontology*, 46:177–208.
- Feller G. and Gerdau C. (2003). Psychrophilic enzymes: hot topics in cold adaptation. *Nature Reviews Microbiology*, 1:200–208.
- Freitas S., Hatosy S., Fuhrman J. A., Huse S. M., Welch D. B. M., Sogin M. L., and Martiny A. C. (2012). Global distribution and diversity of marine *Verrucomicrobia*. *The ISME Journal*, 6:1499–1505.
- Fyfe J. C. and Saenko O. A. (2005). Human-induced change in the Antarctic Circumpolar Current. *Journal of Climate*, 18:3068–3073.
- Ghiglione J. F. and Murray A. E. (2011). Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton. *Environmental Microbiology*, 14:617–629.

- 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*, 11:2164–2178.
- Gille S. T. (2002). Warming of the Southern Ocean Since the 1950s. *Science*, 295:1275–1277.
- Giovannoni S. J., Tripp H. J., Givan S., Podar M., Vergin K. L., Baptista D., Bibbs L., Eads J., Richardson T. H., Noordewier M., Rappé M. S., Short J. M., Carrington J. C., and Mathur E. J. (2005). Genome streamlining in a cosmopolitan oceanic bacterium. *Science*, 309:1242–1245.
- Grote J., Bayindirli C., Bergauer K., Carpintero de Moraes P., Chen H., D'Ambrosio L., Edwards B., Fernández-Gómez B., Hamisi M., Logares R., Nguyen D., Rii Y. M., Saeck E., Schutte C., Widner B., Church M. J., Steward G. F., Karl D. M., DeLong E. F., Eppley J. M., Schuster S. C., Kyrpides N. C., and Rappé M. S. (2011). Draft genome sequence of strain HIMB100, a cultured representative of the SAR116 clade of marine *Alphaproteobacteria*. *Standards in Genomic Sciences*, 5:269–278.
- 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*, 6:1901–1915.
- Hessen D. O., Ågren G. I., Anderson T. R., Elser J. J., and de Ruiter, P.C. (2004). Carbon sequestration in ecosystems: the role of stoichiometry. *Ecology*, 85:1179–1192.
- Howard E. C., Sun S., Biers E. J., and Moran M. A. (2008). Abundant and diverse bacteria involved in DMSP degradation in marine surface waters. *Environmental Microbiology*, 10:2397–2410.
- Huang J., Su Z., and Xu Y. (2005). The evolution of microbial phosphonate degradative pathways. *Journal of Molecular Evolution*, 61:682–690.
- Hunt B. P. V., Pakhomov E. A., and McQuaid C. D. (2001). Short-term variation and long-term changes in the oceanographic environment and zooplankton community in the vicinity of a sub-Antarctic archipelago. *Marine Biology*, 138:369–381.
- Huston A. L., Krieger-Brockett B. B., and Deming J. W. (2000). Remarkably low temperature optima for extracellular enzyme activity from Arctic bacteria and sea ice. *Environmental Microbiology*, 2:383–388.
- Jung S.-Y., Oh T.-K., and Yoon J.-H. (2006). *Colwellia aestuarii* sp. nov., isolated from a tidal flat sediment in Korea. *International Journal of Systematic and Evolutionary Microbiology*, 56:33–37.
- Junge K., Eicken H., and Deming J. W. (2003). Motility of *Colwellia psychrerythraea* Strain 34H at Subzero Temperatures. *Applied and Environmental Microbiology*, 69:4282–4284.
- Kalanetra K. M., Bano N., and Hollibaugh J. T. (2009). Ammonia-oxidizing Archaea in the Arctic Ocean and Antarctic coastal waters. *Environmental Microbiology*, 11:2434–2445.
- Kawahata H. and Ishizuka T. (2000). Amino acids in interstitial waters from ODP Sites 689 and 690 on the Maud Rise, Antarctic Ocean. *Geochemical Journal*, 34:247–261.
- King G. M. (2003). Molecular and Culture-Based Analyses of Aerobic Carbon Monoxide Oxidizer Diversity. *Applied and Environmental Microbiology*, 69:7257–7265.
- Kirchman D. L. (2002). The ecology of *Cytophaga–Flavobacteria* in aquatic environments. *FEMS Microbiology Ecology*, 39:91–100.

- Kjelleberg S., Hermansson M., and Mårdén P. (1987). The transient phase between growth and nongrowth of heterotrophic bacteria, with emphasis on the marine environment. *Annual Review of Microbiology*, 41:25–49.
- Kuwahara H., Yoshida T., Takaki Y., Shimamura S., Nishi S., Harada M., Matsuyama K., Takishita K., Kawato M., Uematsu K., Fujiwara Y., Sato T., Kato C., Kitagawa M., Kato I., and Maruyama T. (2007). Reduced Genome of the Thioautotrophic Intracellular Symbiont in a Deep-Sea Clam, *Calyptogena okutanii*. *Current Biology*, 17:881–886.
- Lauro F. M., McDougald D., Thomas T., Williams T. J., Egan S., Rice S., DeMaere M. Z., Ting L., Ertan H., Johnson J., Ferriera S., Lapidus A., Anderson I., Kyrpides N., Munk A. C., Detter C., Han C. S., Brown M. V., Robb F. T., Kjelleberg S., and Cavicchioli R. (2009). The genomic basis of trophic strategy in marine bacteria. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 106:15527–15533.
- Lauro F. M., DeMaere M. Z., Yau S., Brown M. V., Ng C., Wilkins D., Raftery M. J., Gibson J. A., Andrews-Pfannkoch C., Lewis M., Hoffman J. M., Thomas T., and Cavicchioli R. (2011). An integrative study of a meromictic lake ecosystem in Antarctica. *The ISME Journal*, 5:879–895.
- Liu H., Nolla H. A., and Campbell L. (1997). *Prochlorococcus* growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean. *Aquatic Microbial Ecology*, 12:39–47.
- Liu H., Campbell L., Landry M. R., Nolla H. A., Brown S. L., and Constantinou J. (1998). *Prochlorococcus* and *Synechococcus* growth rates and contributions to production in the Arabian Sea during the 1995 Southwest and Northeast Monsoons. *Deep Sea Research Part II: Topical Studies in Oceanography*, 45:2327–2352.
- Lomas M. W. and Moran S. B. (2011). Evidence for aggregation and export of cyanobacteria and nano-eukaryotes from the Sargasso Sea euphotic zone. *Biogeosciences*, 8:203–216.
- Ludwig W., Strunk O., Westram R., Richter L., Meier H., Yadhukumar , Buchner A., Lai T., Steppi S., Jobb G., Förster W., Brettske I., Gerber S., Ginhart A. W., Gross O., Grumann S., Hermann S., Jost R., König A., Liss T., Lüssmann R., May M., Nonhoff B., Reichel B., Strehlow R., Stamatakis A., Stuckmann N., Vilbig A., Lenke M., Ludwig T., Bode A., and Schleifer K.-H. (2004). ARB: a software environment for sequence data. *Nucleic Acids Research*, 32:1363–1371.
- Marchant H. J., Davidson A. T., and Wright S. W. (1987). The distribution and abundance of chroococcoid cyanobacteria in the Southern Ocean. *Proc. NIPR Symp. Polar Biol*, 1: 1–9.
- Mary I., Heywood J. L., Fuchs B. M., Amann R., Tarhan G. A., Burkhill P. H., and Zubkov M. V. (2006). SAR11 dominance among metabolically active low nucleic acid bacterioplankton in surface waters along an Atlantic meridional transect. *Aquatic Microbial Ecology*, 45:107–113.
- Méthé B. A., Nelson K. E., Deming J. W., Momen B., Melamud E., Zhang X., Moult J., Madupu R., Nelson W. C., Dodson R. J., Methe B. A., Nelson K. E., Deming J. W., Momen B., Melamud E., Zhang X., Moult J., Madupu R., Nelson W. C., Dodson R. J., Brinkac L. M., Daugherty S. C., Durkin A. S., DeBoy R. T., Kolonay J. F., Sullivan S. A., Zhou L., Davidsen T. M., Wu M., Huston A. L., Lewis M., Weaver B., Weidman J. F., Khouri H., Utterback T. R., Feldblyum T. V., and Fraser C. M. (2005). The psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 102:10913–10918.
- Meyer B. and Kuever J. (2007). Molecular Analysis of the Diversity of Sulfate-Reducing and Sulfur-Oxidizing Prokaryotes in the Environment, Using *aprA* as Functional Marker Gene. *Applied and Environmental Microbiology*, 73:7664–7679.

- Mikaloff Fletcher S. E., Gruber N., Jacobson A. R., Doney S. C., Dutkiewicz S., Gerber M., Follows M., Joos F., Lindsay K., Menemenlis D., Mouchet A., Müller S. A., and Sarmiento J. L. (2006). Inverse estimates of anthropogenic CO<sub>2</sub> uptake, transport, and storage by the ocean. *Global Biogeochemical Cycles*, 20:GB2002.
- Moore J. K., Abbott M. R., and Richman J. G. (1999). Location and dynamics of the Antarctic Polar Front from satellite sea surface temperature data. *Journal of Geophysical Research*, 104:3052–3073.
- Moran M. A., Belas R., Schell M. A., González J. M., Sun F., Sun S., Binder B. J., Edmonds J., Ye W., Orcutt B., Howard E. C., Meile C., Palefsky W., Goesmann A., Ren Q., Paulsen I., Ulrich L. E., Thompson L. S., Saunders E., and Buchan A. (2007). Ecological Genomics of Marine Roseobacters. *Applied and Environmental Microbiology*, 73:4559–4569.
- Moran M. A., Buchan A., González J. M., Heidelberg J. F., Whitman W. B., Kiene R. P., HenrikSEN J. R., King G. M., Belas R., Fuqua C., Brinkac L., Lewis M., Johri S., Weaver B., Pai G., Eisen J. A., Rahe E., Sheldon W. M., Ye W., Miller T. R., Carlton J., Rasko D. A., Paulsen I. T., Ren Q., Daugherty S. C., Deboy R. T., Dodson R. J., Durkin A. S., Madupu R., Nelson W. C., Sullivan S. A., Rosovitz M. J., Haft D. H., Selengut J., and Ward N. (2004). Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature*, 432:910–913.
- Murray A. E. and Grzemska J. J. (2007). Diversity and genomics of Antarctic marine micro-organisms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362: 2259–2271.
- Newton I. L. G., Woyke T., Auchtung T. A., Dilly G. F., Dutton R. J., Fisher M. C., Fontanez K. M., Lau E., Stewart F. J., Richardson P. M., Barry K. W., Saunders E., Detter J. C., Wu D., Eisen J. A., and Cavanaugh C. M. (2007). The *Calyptogena magnifica* Chemoautotrophic Symbiont Genome. *Science*, 315:998–1000.
- 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*, 4:1002–1019.
- Obernosterer I., Catala P., Lebaron P., and West N. J. (2011). Distinct bacterial groups contribute to carbon cycling during a naturally iron fertilized phytoplankton bloom in the Southern Ocean. *Limnology and Oceanography*, 56:2391–2401.
- Oh H. M., Kwon K. K., Kang I., Kang S. G., Lee J. H., Kim S. J., and Cho J. C. (2010). Complete Genome Sequence of "Candidatus Puniceispirillum marinum" IMCC1322, a Representative of the SAR116 Clade in the Alphaproteobacteria. *Journal of Bacteriology*, 192:3240–3241.
- 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*, 42:641–673.
- Partensky F., Hess W. R., and Vaulot D. (1999). *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiology and Molecular Biology Reviews*, 63:106–127.
- Paul J. H., DeFlaun M. F., and Jeffrey W. H. (1988). Mechanisms of DNA utilization by estuarine microbial populations. *Applied and Environmental Microbiology*, 54:1682–1688.
- Pollard R. T., Lucas M. I., and Read J. F. (2002). Physical controls on biogeochemical zonation in the Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49:3289–3305.

- Preston C. M., Wu K. Y., Molinski T. F., and DeLong E. F. (1996). A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 93:6241–6246.
- Rappé M. S., Connan S. A., Vergin K. L., and Giovannoni S. J. (2002). Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature*, 418:630–633.
- Rath J., Wu K. Y., Herndl G. J., and DeLong E. F. (1998). High phylogenetic diversity in a marine-snow-associated bacterial assemblage. *Aquatic Microbial Ecology*, 14:261–269.
- Reisch C. R., Stoudemayer M. J., Varaljay V. A., Amster I. J., Moran M. A., and Whitman W. B. (2011). Novel pathway for assimilation of dimethylsulphoniopropionate widespread in marine bacteria. *Nature*, 473:208–211.
- 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:e77–e77.
- Sabine C. L., Feely R. A., Gruber N., Key R. M., Lee K., Bullister J. L., Wanninkhof R., Wong C. S., Wallace D. W. R., Tilbrook B., Millero F. J., Peng T.-H., Kozyr A., Ono T., and Rios A. F. (2004). The Oceanic Sink for Anthropogenic CO<sub>2</sub>. *Science*, 305:367–371.
- Selje N. N., Simon M. M., and Brinkhoff T. T. (2004). A newly discovered *Roseobacter* cluster in temperate and polar oceans. *Nature*, 427:445–448.
- Sokolov S. and Rintoul S. R. (2002). Structure of Southern Ocean fronts at 140°E. *Journal of Marine Systems*, 37:151–184.
- Strutton P. G., Griffiths F. B., Waters R. L., Wright S. W., and Bindoff N. L. (2000). Primary productivity off the coast of East Antarctica (80- 150°E): January to March 1996. *Deep Sea Research Part II: Topical Studies in Oceanography*, 47:2327–2362.
- Swan B. K., Martinez-Garcia M., Preston C. M., Sczyrba A., Woyke T., Lamy D., Reinthalier T., Poulton N. J., Masland E. D. P., Gomez M. L., Sieracki M. E., DeLong E. F., Herndl G. J., and Stepanauskas R. (2011). Potential for Chemolithoautotrophy Among Ubiquitous Bacteria Lineages in the Dark Ocean. *Science*, 333:1296–1300.
- Swingley W. D., Sadekar S., Mastrian S. D., Matthies H. J., Hao J., Ramos H., Acharya C. R., Conrad A. L., Taylor H. L., Dejesa L. C., Shah M. K., O'Huallachain M. E., Lince M. T., Blankenship R. E., Beatty J. T., and Touchman J. W. (2007). The Complete Genome Sequence of *Roseobacter denitrificans* Reveals a Mixotrophic Rather than Photosynthetic Metabolism. *Journal of Bacteriology*, 189:683–690.
- Thomalla S. J., Waldron H. N., Lucas M. I., Read J. F., Ansorge I. J., and Pakhomov E. (2011). Phytoplankton distribution and nitrogen dynamics in the southwest Indian subtropical gyre and Southern Ocean waters. *Ocean Science*, 7:113–127.
- Topping J. N., Heywood J. L., Ward P., and Zubkov M. V. (2006). Bacterioplankton composition in the Scotia Sea, Antarctica, during the austral summer of 2003. *Aquatic Microbial Ecology*, 45:229–235.
- Tripp H. J., Kitner J. B., Schwalbach M. S., Dacey J. W. H., Wilhelm L. J., and Giovannoni S. J. (2008). SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature*, 452:741–744.

- Trull T., Rintoul S. R., Hadfield M., and Abraham E. R. (2001). Circulation and seasonal evolution of polar waters south of Australia: implications for iron fertilization of the Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 48:2439–2466.
- Vila-Costa M., Simó R., Harada H., Gasol J. M., Slezak D., and Kiene R. P. (2006). Dimethylsulfoniopropionate Uptake by Marine Phytoplankton. *Science*, 314:652–654.
- Wagner-Döbler I. and Biebl H. (2006). Environmental Biology of the Marine *Roseobacter* Lineage. *Annual Review of Microbiology*, 60:255–280.
- Walker C. B., de la Torre J. R., Klotz M. G., Urakawa H., Pinel N., Arp D. J., Brochier-Armanet C., Chain P. P., Gollabgir A., Hemp J., Hügler M., Karr E. A., Könneke M., Shin M., Lawton T. J., Lowe T., Martens-Habbena W., Sayavedra-Soto L. A., Langefeld D., Sievert S. M., Rosenzweig A. C., Manning G., and Stahl D. A. (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaeota. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 107:8818–8823.
- Walsh D. A., Zaikova E., Howes C. G., Song Y. C., Wright J. J., Tringe S. G., Tortell P. D., and Hallam S. J. (2009). Metagenome of a Versatile Chemolithoautotroph from Expanding Oceanic Dead Zones. *Science*, 326:578–582.
- Ward P., Whitehouse M., Brandon M., Shreeve R., and Woold-Walker R. (2003). Mesozooplankton community structure across the Antarctic Circumpolar Current to the north of South Georgia: Southern Ocean. *Marine Biology*, 143:121–130.
- Waterbury J. B., Willey J. M., Franks D. G., Valois F. W., and Watson S. W. (1985). A cyanobacterium capable of swimming motility. *Science*, 230:74–76.
- Weber T. S. and Deutsch C. (2010). Ocean nutrient ratios governed by plankton biogeography. *Nature*, 467:550–554.
- West N. J., Obernosterer I., Zemb O., and Lebaron P. (2008). Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. *Environmental Microbiology*, 10:738–756.
- Whitworth T. (1980). Zonation and geostrophic flow of the Antarctic Circumpolar Current at Drake Passage. *Deep Sea Research Part I: Oceanographic Research Papers*, 27:497–507.
- Whitworth T., III and Nowlin W. D., Jr. (1987). Water masses and currents of the Southern Ocean at the Greenwich Meridian. *Journal of Geophysical Research*, 92:6462–6476.
- 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. (2013). Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. *Environmental Microbiology*, 15:1318–1333.
- Williams G. D., Nicol S., Aoki S., Meijers A. J. S., Bindoff N. L., Iijima Y., Marsland S. J., and Klocker A. (2010). Surface oceanography of BROKE-West, along the Antarctic margin of the south-west Indian Ocean (30–80°E). *Deep Sea Research Part II: Topical Studies in Oceanography*, 57:738–757.
- Williams T. J., Lauro F. M., Ertan H., Burg D. W., Poljak A., Raftery M. J., and Cavicchioli R. (2011). Defining the response of a microorganism to temperatures that span its complete growth temperature range (-2 °C to 28 °C) using multiplex quantitative proteomics. *Environmental Microbiology*, 13:2186–2203.
- Williams T. J., Long E., Evans F., DeMaere M. Z., Lauro F. M., Raftery M. J., Ducklow H., Grzymski J. J., Murray A. E., and Cavicchioli R. (2012). A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *The ISME Journal*, 6:1883–1900.

- Williams T. J., Wilkins D., Long E., Evans F., DeMaere M. Z., Raftery M. J., and Cavicchioli R. (2013). The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environmental Microbiology*, 15:1302–1317.
- Yoon J., Yasumoto-Hirose M., Katsuta A., Sekiguchi H., Matsuda S., Kasai H., and Yokota A. (2007). *Coraliomargarita akajimensis* gen. nov., sp. nov., a novel member of the phylum 'Verrucomicrobia' isolated from seawater in Japan. *International Journal of Systematic and Evolutionary Microbiology*, 57:959–963.
- Zhang R., Liu B., Lau S. C. K., Ki J.-S., and Qian P.-Y. (2007). Particle-attached and free-living bacterial communities in a contrasting marine environment: Victoria Harbor, Hong Kong. *FEMS Microbiology Ecology*, 61:496–508.