Biogeographic partitioning of Southern Ocean picoplankton revealed by metagenomics

David Wilkins[1], Federico M. Lauro[1], Timothy J. Williams[1], Matthew Z. Demaere[1], Mark V. Brown[1,2], Jeffrey M. Hoffman[3], Cynthia Andrews-Pfannkoch[3], Jeffrey B. Mcquaid[3], Martin J. Riddle[4], Stephen R. Rintoul[5], Ricardo Cavicchioli[1,*]

[1] *School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, New South Wales, 2052, Australia.*

[2] *Evolution and Ecology Research Centre, The University of New South Wales, Sydney, New South Wales, 2052, Australia.*

[3] *J. Craig Venter Institute, 9704 Medical Center Drive, Rockville, MD, 20850, USA.*

[4] *Australian Antarctic Division, Channel Highway, Kingston, Tasmania, 7050, Australia.*

[5] *CSIRO Marine and Atmospheric Research, and Centre for Australian Weather and Climate Research - A partnership of the Bureau of Meteorology and CSIRO, and CSIRO Wealth from Oceans National Research Flagship, and the Antarctic Climate and Ecosystems Cooperative Research Centre, Castray Esplanade, Hobart, Tas, 7001, Australia.*

** To whom correspondence should be addressed: Ricardo Cavicchioli, School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, NSW, 2052, Tel. (+61 2) 9385 3516, Fax. (+61 2) 9385 2742, E-mail. r.cavicchioli@unsw.edu.au*

Summary

We performed a metagenomic survey (6.6 Gbp of 454 sequence data) of Southern Ocean picoplankton during the austral summer of 2007–2008, examining the genomic signatures of communities across a latitudinal transect from Hobart (44°S) to the Mertz Glacier, Antarctica (67°S). Operational taxonomic units (OTUs) of the SAR11 and SAR116 clades and the cyanobacterial genera *Prochlorococcus* and *Synechococcus* were strongly overrepresented north of the Polar Front (PF). Conversely, OTUs of the Gammaproteobacterial Sulfur Oxidizer-EOSA-1 (GSO-EOSA-1) complex, the phyla Bacteroidetes and Verrucomicrobia and order Rhodobacterales were characteristic of 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 statuses of the two regions. The study provides a new level of understanding about Southern Ocean microbial communities, describing the contrasting taxonomic and functional characteristics of microbial assemblages either side of the PF.

Introduction

The SO plays a critical role in sustaining marine life around the globe. Upwelling of nutrient-rich Circumpolar Deep Water (CDW) returns nutrients transported to the deep ocean by the sinking of organic matter (Rath *et al.*, 1998) and supports 75% of global ocean primary production north of 30°S. Surface waters at high southern latitudes remain cold (< 3°C) year-round but undergo extreme seasonal variations in sea-ice cover, light levels and day length. Primary production and

biomass are high in summer and very low in winter. Bacteria are abundant in the SO despite the low temperatures and seasonal variability in productivity and are a major route for carbon flow (Hessen *et al.*, 2004).

The SO is composed of several zones separated by circumpolar fronts, the locations of which vary temporally and with longitude (Whitworth III, 1980; Orsi *et al.*, 1995; Sokolov and Rintoul, 2002). The fronts separate regions with different physiochemical properties, such as density, salinity, temperature and nutrient concentrations (Sokolov and Rintoul, 2002). Hydrographic, bathythermographic and satellite altimetry data have been used to determine the frontal structure of the Antarctic Circumpolar Current (ACC) south of Australia (Sokolov and Rintoul 2002). From north to south, the major fronts are the Subtropical Front (STF), the Subantarctic Front (SAF), the Polar Front (PF) and the southern ACC front (SACCF). Each of these fronts consists of multiple branches (Sokolov and Rintoul, 2002; Sokolov and Rintoul, 2009b, a). The Subantarctic Zone (SAZ) lies between the STF and SAF, the Polar Frontal Zone (PFZ) lies between the SAF and the PF, and the Antarctic Zone (AZ) lies between the PF and the Antarctic continent.

The PF has been suggested to be a major biogeographical 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). However, the microbial assemblages that characterize Antarctic waters are generally poorly understood, and their diversity and functional capacity are not well characterized (Murray and Grzymski, 2007). Large scale metagenome surveys have not previously been performed.

Recent anthropogenic climate change may be driving the warming and freshening of the ACC (Boning *et al.*, 2008) and shifting it and its fronts poleward (Fyfe and Saenko, 2005;

Biastoch *et al.*, 2009). A community-level understanding is required to effectively understand the main components and dynamics of the microbial food web in the SO and thereby predict the effects of a shifting ACC on the distribution and abundance of plankton. The oceanic changes may have global ecological significance as the SO performs many ecosystem functions, including significant sequestration of anthropogenic CO_2 (Sabine *et al.*, 2004; Mikaloff Fletcher *et al.*, 2006) through both physiochemical processes and the "biological pump" of CO_2 fixation (Thomalla *et al.*, 2011).

In the austral summer of 2006 we initiated a metagenome program based on the sampling design of the Global Ocean Sampling expedition (Rusch *et al.*, 2007), aimed at providing a baseline to monitor microbial communities in the Australian section of the SO. To date we have sampled SO water from 73°E to 150°E and 44°S to 68°S at depths from the surface to ~ 3700 m. In this study, we present data for plankton assemblages passed through a 20 µm prefilter and captured onto sequential 0.1 µm, 0.8 µm and 3.0 µm filters~~, representing surface communities north and south of the PF~~. This fractionation method allows deeper sequencing of the community to improve identification of low-abundance taxa, and provides additional context for interpreting sequence data, for example particle attachment and trophic status (Lauro *et al.*, 2011). The samples were collected in summer 2007/2008 on the SR3 transect (Sokolov and Rintoul, 2002) of the Climate of Antarctica and the Southern Ocean (CASO), and Collaborative East Antarctica Marine Census (CEAMARC) projects during the International Polar Year (IPY) program (Fig. 1). We assessed the taxonomic and functional profiles of microbial communities from either side of the PF, thereby contributing important new information about the microbial ecology of the SO and defining the microbial communities most influenced by the effects of the PF forming a biogeographical barrier.

Results and discussion

Overview of taxonomic biogeography

6.6 Gbp of 454 sequence data representing picoplankton in the size range 0.1 – 3.0 µm was obtained from 16 samples. After removal of low-quality reads, 454 sequencing yielded 157,507 to 597,689 reads per sample (mean 354,399) of lengths ranging from 100 to 606 bp (mean 378). The proportion of reads in each sample which yielded matches to RefSeq ranged from 25% to 85% (mean 62%). The most abundant OTUs in each sample are given in Table 1 and a full list of OTU abundances in Table S2. It is important to note that the taxonomic assignments are confident assignments (*i.e.* above an E-value threshold of 1.0×10^{-3}) that are limited to the sequences of organisms available in the database and identified by MINSPEC as part of the minimal species set.

ANOSIM was performed to test for statistically significant differences between the OTU profiles of the zones. The analysis showed that the zones harbor significantly different microbial communities ($R = 0.451$, $p < 0.004$). SIMPER was performed in order to identify the contribution of individual OTUs to the differences between the zones. The analysis found that no single OTU contributed more than 2.9% of variance (Fig. 2), 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 3.0 µm fraction, 37% by the 0.8 µm fraction, and 9% by the 0.1 µm fraction (Table S3, supporting information).

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 spp., SAR11 and SAR116 were on average more abundant in the NZ (Fig. 2). Some groups, such as the Alteromonadales, had variable relative representation depending on size fraction.

GSO-EOSA-1

The Gammaproteobacterial Sulfur Oxidizer-EOSA-1 (GSO-EOSA-1) cluster, represented in RefSeq by the OTUs *Candidatus* Vesicomyosocius okutanii strain HA and *Ca.* 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 µm size fractions respectively, compared to 1.1%, 0.84% and 0.30% in the NZ. The contribution to variance of this group was highest in the 0.1 µm size fraction, followed by 0.8 µm and 3.0 µm (Fig. 2). 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). The majority of 16S rRNA genes from our metagenome with best BLASTN matches to *Ca.* R. magnifica and *Ca.* V. okutanii clustered with ARTIC96BD-19 in a neighbour-joining phylogenetic tree (Fig. S1, supporting information), indicating this is the dominant GSO-EOSA-1 representative. 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)~~(Williams *et al.* (2012)~~. We assembled our GSO-EOSA-1 affiliated reads and identified several contigs containing ORFs with high identity to aerobic respiration genes, including cytochrome C oxidase subunits, from *Ca.* R. magnifica and *Ca.* V. okutanii (Fig. S2, supporting information). 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 input in AZ waters due to the lack of available light, both from seasonal darkness and ice cover (Grzymski *et al.*, 2012). The high relative abundance of GSO-EOSA-1 we detected 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.

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) and are the only representatives in the reference database of the Ammonia Oxidizing Archaea (AOA). The contribution of OTUs of *C. symbiosum* to the AOA signature was low (Table S3, supporting information). 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. 2). 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 µm fraction; 0.8 µm: 11%; 3.0 µm: 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), 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 (Kalanetra *et al.*, 2009).

Cyanobacteria

OTUs of the cyanobacterial genera *Prochlorococcus* and *Synechococcus* were overrepresented in the NZ in all size fractions (Fig. 3). 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 (Table S2, supporting information), consistent with *Synechococcus* species' average cell diameter of approximately 0.9 µm. The high abundance of both cyanobacterial genera on the 3.0 µm fraction was unexpected given the cell diameter, although it has previously been reported (Lauro *et al.*, 2011)~~(Lauro *et al.*, 2011a)~~ and attributed to aggregation (Lomas and Moran, 2011)~~{Lomas, 2011 #712}~~. Samples 367 and 368 were separated from the other samples north of

Formatted: Tab stops: 1.27 cm, Left + 2.54 cm, Left + 3.81 cm, Left + 5.08 cm, Left + 6.35 cm, Left + 7.62 cm, Left + 8.89 cm, Left + 10.16 cm, Left + 11.43 cm, Left + 14.12 cm, Left

191 the PF by the Subtropical Front (STF). While the STF was not a significant boundary on the
192 assemblage level, it may mark a significant biogeographical boundary for these cyanobacteria.
193 *Synechococcus* and *Prochlorococcus* together represent a large proportion of both phytoplankton
194 abundance and carbon fixation in temperate and tropical waters, in many regions contributing
195 more than half of total primary production (Liu *et al.*, 1997; Liu *et al.*, 1998; André *et al.*, 1999).
196 The role of the STF in determining the latitudinal range of *Synechococcus* and *Prochlorococcus* is
197 therefore important, as it will affect models of ocean productivity under changing climactic
198 conditions, and warrants further investigation.
199

200 Despite the high abundance of cyanobacteria north of the STZ, they were also a significant
201 feature of the SAZ; for example, *Synechococcus* sp. CC9902 composed 3-5% of the 0.8 µm
202 fraction in SAZ samples. Our results extend the latitudinal distribution of both *Prochlorococcus*
203 and *Synechococcus* to include presence at very low abundance as far south as the Antarctic coast.
204 *Prochlorococcus* have been reported to be restricted to tropical and subtropical waters within 40º
205 of latitude (Partensky *et al.*, 1999), and to be a negligible (Ghiglione and Murray, 2011) or ~~an~~
206 undetectable ~~component~~ (Gryzymski *et al.* 2012) component of marine picoplankton in Antarctic
207 waters. However, our findings are consistent with findings of a logarithmic relationship of
208 cyanobacterial numbers with temperature, where cyanobacteria were found at concentrations of
209 $10^3 – 10^4$ cells per litre even in the coldest waters, approximately four orders of magnitude less
210 than in waters around Tasmania (Marchant *et al.*, 1987). Cyanophage proteins have also been
211 detected in a metaproteomic analysis of Antarctic Peninsula coastal surface waters (Williams *et*
212 *al.* 2012). The depth of shotgun metagenome sequence coverage of our samples is likely to have
213 contributed to the detection of these cyanobacteria.
214

SAR11 and SAR116 clades

Ca. Pelagibacter ubique HTCC1062 is a good representative of total SAR11 abundance in our study, as it is a member of the SAR11 phylotype which is most abundant in SO waters 22(Brown *et al.*, 2012). *Ca.* P. 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. 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 DOC concentrations due to low primary productivity (Giovannoni *et al.*, 2005; Alonso and Pernthaler, 2006), such as the high-nutrient, low-chlorophyll (HNLC) SAZ region of the SO. Overall, our findings are consistent with reports that SAR11 is ubiquitous in the world's oceans (*e.g.* (Mary *et al.*, 2006; Carlson *et al.*, 2009) and more abundant north of the ACC (Giebel *et al.*, 2009).

OTUs of *Ca.* Puniceispirillum marinum from the SAR116 clade were a moderate contributor to variance between the NZ and SZ with higher abundance in the NZ (Fig. 2). A genomic analysis reported *Ca.* P. marinum IMCC1322 to be a metabolic generalist with genes for aerobic CO fixation, C_1 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 (determined using fluorescence *in situ* hybridisation) was reported to be higher in more productive waters where SAR11 numbers were lower (Topping *et al.*, 2006). However, our analysis across an extended latitudinal transect indicates that overall SAR11 and SAR116 have similar biogeographic distributions.

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 µm, 0.8 µm and 3.0 µm fractions respectively; SZ: 2.3%, 9.8% and 9.1%) and significant contributors to variance between the NZ and SZ (Fig. 2). Flavobacteria have been previously reported to compose the majority of both Bacteroidetes (Murray and Grzymski, 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 and Particulate Organic Matter (DOM and POM) (Kirchman, 2002), marine Flavobacteria are major components of marine aggregates (Rath *et al.*, 1998; Crump *et al.*, 1999; Zhang *et al.*, 2007). The higher abundance of Flavobacteria OTUs on the 0.8 µm and 3.0 µm fractions indicates their association with particulate matter.

The higher abundance of OTUs of Flavobacteria in the SZ may reflect an input of cells from melting sea ice (Brown and Bowman, 2001), the higher rates of primary productivity in the south, and the role of the Flavobacteria as degraders of HMW DOM. 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 remineralizing shunt, which would decrease carbon sequestration by this route.

Rhodobacterales

Members of the order Rhodobacterales were abundant (NZ average: 1.2%, 10% and 5.5% of the 0.1 µm, 0.8 µm and 3.0 µm fractions respectively; SZ: 1.6%, 13% and 7.9%) and high

263 contributors to variance, overrepresented in the SZ on all size fractions. As several members of
264 the Roseobacter clade have been shown to have symbiotic relationships with marine eukaryotic
265 algae (Buchan and Moran, 2005; Wagner-Dobler and Biebl, 2006), and their abundance in the SO
266 has previously been linked to phytoplankton blooms (West *et al.*, 2008; Obernosterer *et al.*,
267 2011), it is likely that their overrepresentation in the SZ is related to the higher density of
268 phytoplankton in the AZ.
269 OTUs of *Roseobacter denitrificans* Och114 and *Silicibacter pomeroyi* DSS-3 were
270 consistently the most abundant Roseobacter clade representatives. *R. denitrificans* and *S.*
271 *pomeroyi* fall within a subclade of Aerobic Anoxygenic Phototrophic (AAP) members of the
272 *Roseobacter* clade (Swingley *et al.*, 2007). These species have diverse mixotrophic metabolisms,
273 with genomic and experimental evidence of photoheterotrophic respiration of organic carbon,
274 fixation of CO_2, oxidation of CO, oxidation of reduced sulfur compounds, and utilization of the
275 abundant marine osmolyte DMSP (King, 2003; Moran *et al.*, 2004; Wagner-Dobler and Biebl,
276 2006; Swingley *et al.*, 2007; Brinkhoff *et al.*, 2008; Howard *et al.*, 2008). This metabolic diversity
277 suggests a complex ecological role, particularly with respect to the capture and release of
278 climatically active gases (CO_2, CO, dimethylsulfide) involved in carbon and sulfur cycling.
279
280 *Alteromonadales*
281
282 Members of the gammaproteobacterial order Alteromonadales were large contributors to
283 variance. Most OTUs were overrepresented in the SZ but some were overrepresented in the NZ
284 on the 3.0 µm fraction (Fig. 2). ~~*Colwellia* *psychrerythraea*~~ ~~sp.~~ 34H~~, type strain of *Colwellia*~~
285 ~~*psychrerythraea*,~~ was one of the most abundant alteromonads and exhibited this distribution (NZ
286 average: 0.14%, 2.2% and 16% of the 0.1 µm, 0.8 µm and 3.0 µm fractions respectively; SZ:

0.52%, 5.1% and 10%). *C. psychrerythraea* 34H is a well-characterised psychrophile isolated from Arctic sediment (Myers, 2000), a lifestyle and environment consistent with its overrepresentation in the SZ on the 0.1 µm and 0.8 µm fractions. However, it is noteworthy that water temperatures at the NZ sample sites (2.9–14 °C) were better suited to *C. psychrerythraea* 34H's optimum growth temperature of 8 °C ~~{Methé, 2005 #6}~~ than those in the SZ (−1.8–0.7 °C). As *C. psychrerythraea* 34H has been shown to synthesise extracellular polysaccharides (Myers, 2000), we speculate that the higher abundance on the 3.0 µm fraction in the NZ may represent aggregation under conditions of optimum growth. ~~*C. psychrerythraea* is a well-characterised psychrophile isolated from Arctic sediment (Methé *et al.*, 2005), a lifestyle and environment consistent with its overrepresentation in the SZ on the 0.1 µm and 0.8 µm fractions. We hypothesized that the 3.0 µm fraction *C. psychrerythraea* hits may represent *Thalassomonas viridans*, a close relative of *Colwellia* identified in temperate waters (Macián *et al.*, 2001) but not represented in RefSeq. However, a comparison of reads with identity to *C. psychrerythraea* 16S rRNA to 16S sequences of *C. psychrerythraea*, *T. viridians* and other proteobacterial species did not support this hypothesis. The *C. psychrerythraea* OTU abundant in the NZ in 3.0 µm fraction may therefore represent an uncharacterized relative or ecotype with a preference for warmer environments.~~

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. 2). Surprisingly given the small cell size of *C. akajimensis* (Yoon *et al.*, 2007), contribution to variance increased with size fraction; a study in the North Sea reported a similar fractionation

Formatted: Font: Not Italic

pattern, and suggested marine Verrucomicrobia may be predominantly particle attached (Trull *et al.*, 2001). However, little else is known about the distribution and ecological roles of marine Verrucomicrobia (Trull *et al.*, 2001).

Overview of functional biogeography

ANOSIM analysis of the samples' KEGG ortholog group and module profiles revealed that the zones had significantly different functional potential (ortholog group: R = 0.642, p < 0.001; module: R = 0.~~819~~871, p < 0.001). SIMPER was performed on the profiles in order to identify the specific functional differences between the zones. No single ortholog group or module contributed more than ~~1.01~~2.2% of the variance, indicating a complex and diverse pattern of functional differences (Table S4 and Table S5, supporting information). 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 µm fraction but the SZ in the smaller fractions, indicating that the functional diversity of each zone was strongly segregated by size fraction.

Functional capacities distinguished by the PF

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 2). 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

Comment [w1]: Double-checked this; still accurate

Comment [w2]: corrected

Comment [w3]: corrected

Field Code Changed

contributors to these modules were varied, although members of the Rhodobacterales were prominent (Fig. 3). 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 comprises several psychrophiles.

~~Two~~ One exception~~s~~ to this pattern ~~were~~ was the branched-chain amino acid transport system ~~and multiple sugar transport~~ module~~s~~, ~~both of which were~~ overrepresented in the NZ. ~~The higher abundance of the multiple sugar module is consistent with the oligotrophic preference for broad specificity transporters, while the relatively lower abundance of the branched-chain amino acid module~~This may reflect a higher availability of more labile dissolved free amino acids (DFAA) and dipeptides in the SZ as byproducts of blooming eukaryotic phytoplankton. Additionally, as the genera *Pelagibacter* and *Puniceispirillum* were major contributors to this module's overabundance in the NZ (Fig. 3), ~~phytoplankton,~~ this may reflect a generalised adaptation to more oligotrophic environments.

Biosynthesis pathways for all major nucleic acids (pyrimidine deoxyribonuleotide 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 (Pop *et al.*, 2004) thus reducing the requirement for *de novo* synthesis. No single taxonomic group

Formatted: Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Font: Italic
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Field Code Changed
Field Code Changed
Field Code Changed

contributed a large fraction of the difference in this module (Fig. 3), suggesting this is a widespread adaptiation.

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 dimethylsulphoniopropionate (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 dimethylsulfide or methanethiol (reviewed by Curson *et al.* (2011)). However, when DMSP is scarce, MMPA ~~can~~ 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. 3), suggesting *Synechococcus* species may use this route when DMSP is unavailable.

Field Code Changed

Field Code Changed

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

Field Code Changed

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

Field Code Changed

included in this module). APS reductase is used during sulfur oxidation by autotrophic GSO-EOSA-1 (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). ~~*Roseobacters* and GSO-EOSA-1 were both among the highest contributors to variance and overrepresented in the SZ, suggesting they contributed to the overabundance of this module.~~

~~The photosystem I and II modules were overrepresented in the NZ. As the SZ has on average a higher chlorophyll concentration than the NZ (Moore and Abbott, 2000), this pattern was not expected. The KEGG module for eukaryotic ribosomes was also unexpectedly slightly overrepresented in the NZ. Underrepresentation of plastid sequence in the KEGG database may have contributed to a systematic bias against eukaryotic genes. As our filtration approach used a 20 µm prefilter, cyanobacteria may have been enriched relative to large phototrophic algae. These findings highlight complexities presently associated with interpreting data for marine eukaryotes.~~ 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* spp. to be major contributors to the difference (Fig. 3). Variation in the photosystem I module, which was marginally overrepresented in the SZ, could largely be attributed to diatoms and other eukaryotic photosynthesisers (Fig. 3), 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 photosynthesisers and cyanobacteria would be expected to encode both complete photosystems,

Field Code Changed

Field Code Changed

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

the differences in module abundance probably represent 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 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. 3), supporting an association with 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.

~~Two different KEGG modules representing the NADH dehydrogenase complex (Complex I of the electron transport chain) were very high contributors to variance, but were overrepresented in different zones. KEGG module M00145 ("Complex I (NADH dehydrogenase), NADH dehydrogenase I/diaphorase subunit of the bidirectional hydrogenase") was significantly overrepresented in the NZ, while module M00142 ("Complex I (NADH dehydrogenase), NADH dehydrogenase I") was overrepresented in the SZ. The distinction between these modules in the KEGG orthology is taxonomic: M00145 contains NADH dehydrogenases associated with oxidative phosphorylation in cyanobacteria and chloroplasts, while M00142 comprises the mitochondrial equivalents. The distribution of these modules therefore reflects the taxonomic~~

Formatted: Not Highlight

Field Code Changed

Field Code Changed

Formatted: Highlight

~~differences between the zones, with cyanobacteria overrepresented in the NZ and eukaryotic plankton in the SZ.~~

Biogeographic role of the PF

Our 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). Our findings extend previous work in defining the PF as a strong biogeographic boundary which shapes not only the composition, but also the functional capacity of microbial communities in the SO.

A possible alternative hypothesis for the observed separation is that the samples are partitioned by the continental margin, as all but one of the SZ samples were taken in waters over the Antarctic continental shelf and slope in the vicinity of the Mertz glacier polynya. However, ANOSIM analysis of an alternative grouping of the samples into "polynya" and "open ocean" had poorer support ($R = 0.309$, $p < 0.01$) than the grouping based on the PF. Additional taxonomic profiles for samples taken from the region south of the PF in other seasons (austral summers 06/07, 08/09) and in other sectors of the SO (70-115 °E) also supported the PF as the major discriminator (data not shown). Taken together, this evidence strongly supports the hypothesis that the PF is a major biogeographical boundary in the SO independent of a latitudinal gradient or of the effect of the continental margin and Mertz polynya.

Comment [w4]: This section added little to the analysis

Formatted: Font: (Default) Times, English (U.S.)

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 because of the large oceanic expanse that is involved and the importance of the carbon fixation and nutrient cycling that occurs there. Knowledge of these communities and their biogeographic drivers has relevance for understanding and predicting the long-term effects of environmental change in the region. Our 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 picoplankton.

Experimental procedures

Sampling and DNA sequencing

A volume of ~ 500 L per sample was collected by sequential size fractionation through a 20 µm prefilter directly onto 3.0, 0.8 and 0.1 µm pore sized 293 mm polyethersulfone membrane filters, and cryogenically preserved (Rusch *et al.*, 2007; Ng *et al.*, 2010). DNA extraction (Rusch *et al.*, 2007) and pyrosequencing on GS20 FLX Titanium (Roche, Branford, CT, USA) was performed

at the J. Craig Venter Institute in Rockville, MD, USA as described previously (Lauro *et al.*, 2011). Duplicate reads and reads with many pyrosequencing errors were removed as described previously (Lauro *et al.*, 2011).

Grouping of samples by oceanographic zone

A range of data were recorded on board the RSV *Aurora Australis*, including position, sampling and water column depth, ocean temperature, salinity and fluorescence, and meteorological data (Table S1, supporting information). These were used to locate the PFZ based on a surface temperature gradient ~ 1.35^{o}C across a distance of 45–65 km, placing the PF at approximately -59.70^{o}, consistent with previous descriptions (Moore *et al.*, 1999; Sokolov and Rintoul, 2002). Samples were accordingly grouped into "North" (NZ) and "South" (SZ) zones (Table S1, supporting information). The NZ represents waters from Subtropical, Subantarctic and Polar Frontal Zones, while the SZ represents the AZ.

Comparison to RefSeq database

A subset of the RefSeq microbial (bacteria and archaea) genome database (release 41, retrieved May 31 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 (Angly *et al.*, 2009). Because this database was not expected to contain representative genomes for every species present, Operational Taxonomic Units (OTUs) in this study are defined by the best species match in this database, and may for example represent congeners.

The metagenomic reads from each sample were compared against this database using TBLASTX, with default parameters except for: E-value threshold [-e] 1.0×10^{-3}; cost to open gap [-G] 11; cost to extend gap [-E] 1; masking of query sequence [-F] m S (SEG masking for lookup table only).

Identification of minimal species sets

A computational method to minimise false OTU identifications and increase the accuracy of OTU abundance estimates (MINSPEC) was developed and implemented in PERL. Following the approach of Ye and Doak (2009) to the parsimonious reconstruction of biochemical pathways (MINPATH), MINSPEC computes the smallest set of OTUs sufficient to explain a set of observed high-quality hits against RefSeq (or any other sequence database). The minimal set computation is framed as a linear programming problem and solved with the GNU Linear Programming Kit (GLPK) tool "GLPK linear programming/MIP solver" (GLPSOL) (Free Software Foundation, Boston). This approach eliminates many of the spurious OTU identifications which result from reads with strong identity to more than one OTU. The "minimal species set" is liable to exclude some low-abundance OTUs, but gives more faithful abundance estimates and eliminates many false positives.

To validate this approach and estimate error rates, an assemblage of hypothetical taxa was simulated with varying degrees of overlapping genomic identity and a logarithmic rank-abundance curve. A simulated metagenomic sampling and BLAST search was performed on this set, and the results processed with MINSPEC. Over multiple replications, MINSPEC was consistently able to identify the true set and reject spurious identifications, with false positive and negative rates smaller than 5% (excluding unsampled rare taxa).

The outputs of all TBLASTX searches against RefSeq were processed by MINSPEC, and hits not belonging to the minimal sets were removed.

OTU abundance and variance between zones

The relative OTU abundances for each sample were determined using the PERL script Genome relative Abundance and Average Size (GAAS) (Angly *et al.*, 2009). Briefly, GAAS estimates the relative abundance of OTUs from the number and quality of BLAST hits to each species, taking into account differences in genome size. 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. 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.

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

Assembly of GSO-EOSA-1 contigs

To investigate the physiological potential of the GSO-EOSA representative identified in the SZ, reads with identity to *Ca.* R. magnifica and/or *Ca.* V. okutanii were assembled using the Celera WGS Assembler v6.1 (Myers, 2000)~~(Huson *et al.*, 2001)(Pop *et al.*, 2004)~~. 242 large (> 2 kbp)

Formatted: Indent: First line: 0 cm

contigs were used for ORF prediction by MetaGene (Noguchi *et al.*, 2006), and predicted ORFs compared against the NCBI nr database with blastx (E-value threshold [-e] 1.0×10^{-3}). The resulting annotated contigs were manually analysed for genes of interest.

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 July 2 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 [-b] 10; E-value threshold [-e] 1.0~~x~~ × 10^{-3}; gap opening penalty [-G] 11; gap extension penalty [-E] 1; masking of query sequence [-F] m S (SEG masking for lookup table only).

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 matrix of the module and ortholog group profiles. A functional profile was generated for each sample by summing the scaled abundances of each module from all size fractions, and SIMPER performed as above to identify modules which contributed highly to the variation in functional potential between the two zones. Modules with a high contribution to variance or otherwise of interest were then linked to taxonomy ("taxonomic decomposition") by noting the genus of the organism associated with each gene in the KEGG GENES database and thus calculating the relative contribution of each genus to each module's abundance. This allowed us to putatively assign functional contributions to genera which were not identified in our taxonomic analysis, as the database included gene sequences for organisms for which a full genome was not available.

~~A functional profile was generated for each sample by summing the scaled abundances of each module from all size fractions.~~

~~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 matrix of the module and ortholog group profiles.~~

~~To identify which modules contributed significantly to the variation in functional potential between the two zones, SIMPER was performed as above.~~

Acknowledgments

The authors acknowledge technical support for computing infrastructure and software development from Intersect, and in particular assistance from Joachim Mai, and acknowledge Matthew Lewis from the JCVI for his assistance with DNA sequencing. This work was supported by the Australian Research Council and the Australian Antarctic Division. Funding for sequencing was provided by the Gordon and Betty Moore Foundation to the JCVI.

605 **References**

606 Abell, G.C.J., and Bowman, J.P. (2005) Ecological and biogeographic relationships of class
607 Flavobacteria in the Southern Ocean. *FEMS Microbiol Ecol* **51**: 265-277.

608 Alonso, C., and Pernthaler, J. (2006) Roseobacter and SAR11 dominate microbial glucose uptake
609 in coastal North Sea waters. *Environ Microbiol* **8**: 2022-2030.

610 André, J.-M., Navarette, C., Blanchot, J., and Radenac, M.-H. (1999) Picophytoplankton
611 dynamics in the equatorial Pacific: Growth and grazing rates from cytometric counts. *J Geophys*
612 *Res* **104**: 3369-3380.

613 Angly, F.E., Willner, D., Prieto-Davo, A., Edwards, R.A., Schmieder, R., Vega-Thurber, R. *et al.*
614 (2009) The GAAS metagenomic tool and its estimations of viral and microbial average genome
615 size in four major biomes. *PLoS Comput Biol* **5**: e1000593.

616 Biastoch, A., Boning, C.W., Schwarzkopf, F.U., and Lutjeharms, J.R.E. (2009) Increase in
617 Agulhas leakage due to poleward shift of Southern Hemisphere westerlies. *Nature* **462**: 495-498.

618 Bidle, K.D., and Azam, F. (2001) Bacterial control of silicon regeneration from diatom detritus:
619 Significance of bacterial ectohydrolases and species identity. *Limnol Oceanogr* **46**: 1606-1623.

620 Boning, C.W., Dispert, A., Visbeck, M., Rintoul, S.R., and Schwarzkopf, F.U. (2008) The
621 response of the Antarctic Circumpolar Current to recent climate change. *Nature Geosci* **1**: 864-
622 869.

623 Bowman, J.P., and McCuaig, R.D. (2003) Biodiversity, Community Structural Shifts, and
624 Biogeography of Prokaryotes within Antarctic Continental Shelf Sediment. *Appl Environ*
625 *Microbiol* **69**: 2463-2483.

626 Brinkhoff, T., Giebel, H.-A., and Simon, M. (2008) Diversity, ecology, and genomics of the
627 Roseobacter clade: a short overview. *Arch Microbiol* **189**: 531-539.

628 Brown, M.V., and Bowman, J.P. (2001) A molecular phylogenetic survey of sea-ice microbial
629 communities (SIMCO). *FEMS Microbiol Ecol* **35**: 267-275.

630 Brown, M.V., Lauro, F.M., DeMaere, M.Z., Muir, L., Wilkins, D., Thomas, T. *et al.* (2012)
631 Global Biogeography of SAR11 Marine Bacteria. *Mol Syst Biol*: In Press.

632 Buchan, A., and Moran, M.A. (2005) Overview of the Marine Roseobacter Lineage. *Appl Environ*
633 *Microbiol* **71**: 5665-5677.

634 Carlson, C.A., Morris, R., Parsons, R., Treusch, A.H., Giovannoni, S.J., and Vergin, K. (2009)
635 Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the
636 northwestern Sargasso Sea. *ISME J* **3**: 283-295.

637 Chiba, S., Ishimaru, T., Hosie, G., and Fukuchi, M. (2001) Spatio-temporal variability of
638 zooplankton community structure off east Antarctica (90 to 160°E). *Mar Ecol Prog Ser* **216**: 95-
639 108.

640 Clarke, K.R., and Warwick, R.M. (2001) *Change in Marine Communities: An Approach to*
641 *Statistical Analysis and Interpretation*. Plymoth: PRIMER-E.

642 Coleman, M.L., and Chisholm, S.W. (2010) Ecosystem-specific selection pressures revealed
643 through comparative population genomics. *Proc Natl Acad Sci USA* **107**: 18634-18639.

644 Crump, B.C., Armbrust, E.V., and Baross, J.A. (1999) Phylogenetic analysis of particle-attached
645 and free-living bacterial communities in the Columbia river, its estuary, and the adjacent coastal
646 ocean. *Appl Environ Microbiol* **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. *Nat Rev Microbiol*.

Esper, O., and Zonneveld, K.A. (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. *Mar Micropaleontol* **46**: 177-208.

Fyfe, J.C., and Saenko, O.A. (2005) Human-Induced Change in the Antarctic Circumpolar Current. *J 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. *Environ Microbiol*: 1-13.

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. *Environ Microbiol* **11**: 2164-2178.

Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D. *et al.* (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242-1245.

Grzymski, J.J., Riesenfeld, C.S., Williams, T.J., Dussaq, A.M., Ducklow, H., Erickson, M. *et al.* (2012) A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. *ISME J*.

Hessen, D.O., Agren, 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. *Environ Microbiol* **10**: 2397-2410.

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. *Mar Biol* **138**: 369-381.

Kalanetra, K.M., Bano, N., and Hollibaugh, J.T. (2009) Ammonia-oxidizing Archaea in the Arctic Ocean and Antarctic coastal waters. *Environ Microbiol* **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. *Geochem J* **34**: 247-261.

King, G.M. (2003) Molecular and Culture-Based Analyses of Aerobic Carbon Monoxide Oxidizer Diversity. *Appl Environ Microbiol* **69**: 7257-7265.

Kirchman, D.L. (2002) The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiol Ecol* **39**: 91-100.

Kuwahara, H., Yoshida, T., Takaki, Y., Shimamura, S., Nishi, S., Harada, M. *et al.* (2007) Reduced genome of the thioautotrophic intracellular symbiont in a deep-sea clam, *Calyptogena okutanii*. *Curr Biol* **17**: 881-886.

Lauro, F.M., DeMaere, M.Z., Yau, S., Brown, M.V., Ng, C., Wilkins, D. *et al.* (2011) An integrative study of a meromictic lake ecosystem in Antarctica. *ISME J* **5**: 879-895.

Lauro, F.M., McDougald, D., Thomas, T., Williams, T.J., Egan, S., Rice, S. *et al.* (2009) The genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci U S A* **106**: 15527-15533.

Liu, H., Nolla, H., and Campbell, L. (1997) *Prochlorococcus* growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean. *Aquat Microb Ecol* **12**: 39-47.

690 Liu, H., Campbell, L., Landry, M.R., Nolla, H.A., Brown, S.L., and Constantinou, J. (1998)
691 *Prochlorococcus* and *Synechococcus* growth rates and contributions to production in the Arabian
692 Sea during the 1995 Southwest and Northeast Monsoons. *Deep-Sea Res Pt II* **45**: 2327-2352.
693 Lomas, M.W., and Moran, S.B. (2011) Evidence for aggregation and export of cyanobacteria and
694 nano-eukaryotes from the Sargasso Sea euphotic zone. *Biogeosciences* **8**: 203-216.
695 Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar *et al.* (2004) ARB: a
696 software environment for sequence data. *Nucleic Acids Res* **32**: 1363-1371.
697 Marchant, H.J., Davidson, A.T., and Wright, S.W. (1987) The Distribution and Abundance of
698 Chroococcoid Cyanobacteria in the Southern Ocean. *Proc NIPR Symp Polar Biol* **1**: 1-9.
699 Mary, I., Heywood, J., Fuchs, B., Amann, R., Tarran, G., Burkill, P., and Zubkov, M. (2006)
700 SAR11 dominance among metabolically active low nucleic acid bacterioplankton in surface
701 waters along an Atlantic meridional transect. *Aquat Microb Ecol* **45**: 107-113.
702 Meyer, B., and Kuever, J. (2007) Molecular analysis of the diversity of sulfate-reducing and
703 sulfur-oxidizing prokaryotes in the environment, using aprA as functional marker gene. *Appl*
704 *Environ Microbiol* **73**: 7664-7679.
705 Mikaloff Fletcher, S.E., Gruber, N., Jacobson, A.R., Doney, S.C., Dutkiewicz, S., Gerber, M. *et*
706 *al.* (2006) Inverse estimates of anthropogenic CO 2 uptake, transport, and storage by the ocean.
707 *Global Biogeochem Cy* **20**: 1-16.
708 Moore, J.K., Abbott, M.R., and Richman, J.G. (1999) Location and dynamics of the Antarctic
709 Polar Front from satellite sea surface temperature data. *J Geophys Res* **104**: 3059-3073.
710 Moran, M.A., Belas, R., Schell, M.A., González, J.M., Sun, F., Sun, S. *et al.* (2007) Ecological
711 genomics of marine Roseobacters. *Appl Environ Microbiol* **73**: 4559-4569.
712 Moran, M.A., Buchan, A., Gonzalez, J.M., Heidelberg, J.F., Whitman, W.B., Kiene, R.P. *et al.*
713 (2004) Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment.
714 *Nature* **432**: 910-913.
715 Murray, A.E., and Grzymski, J.J. (2007) Diversity and genomics of Antarctic marine micro-
716 organisms. *Philos T Roy Soc B* **362**: 2259-2271.
717 Myers, E.W. (2000) A Whole-Genome Assembly of Drosophila. *Science* **287**: 2196-2204.
718 Newton, I.L.G., Woyke, T., Auchtung, T.A., Dilly, G.F., Dutton, R.J., Fisher, M.C. *et al.* (2007)
719 The *Calyptogena magnifica* chemoautotrophic symbiont genome. *Science* **315**: 998-1000.
720 Ng, C., DeMaere, M.Z., Williams, T.J., Lauro, F.M., Raftery, M., Gibson, J.a.E. *et al.* (2010)
721 Metaproteogenomic analysis of a dominant green sulfur bacterium from Ace Lake, Antarctica.
722 *ISME J* **4**: 1002-1019.
723 Noguchi, H., Park, J., and Takagi, T. (2006) MetaGene: prokaryotic gene finding from
724 environmental genome shotgun sequences. *Nucleic Acids Res* **34**: 5623-5630.
725 Obernosterer, I., Catala, P., Lebaron, P., and West, N.J. (2011) Distinct bacterial groups
726 contribute to carbon cycling during a naturally iron fertilized phytoplankton bloom in the
727 Southern Ocean. *Limnol Oceanogr* **56**: 2391-2401.
728 Oh, H.-M., Kwon, K.K., Kang, I., Kang, S.G., Lee, J.-H., Kim, S.-J., and Cho, J.-C. (2010)
729 Complete genome sequence of "*Candidatus* Puniceispirillum marinum" IMCC1322, a
730 representative of the SAR116 clade in the Alphaproteobacteria. *J Bacteriol* **192**: 3240-3241.
731 Orsi, H., Whitworth III, T., and Nowlin Jr, W.D. (1995) On the meridional extent and fronts of
732 the Antarctic Circumpolar Current. *Deep-Sea Res Pt I* **42**.

Partensky, F., Hess, W.R., and Vaulot, D. (1999) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Rev* **63**: 106-127.

Pollard, R., Lucas, M., and Read, J. (2002) Physical controls on biogeochemical zonation in the Southern Ocean. *Deep-Sea Res Pt II* **49**: 3289-3305.

Pop, M., Phillippy, A., Delcher, A.L., and Salzberg, S.L. (2004) Comparative genome assembly. *Brief Bioinform* **5**: 237-248.

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. *Proc Natl Acad Sci U S A* **93**: 6241-6246.

Rappé, M.S., Connon, 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. *Aquat Microb Ecol* **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. *et al.* (2007) The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* **5**: e77.

Sabine, C.L., Feely, R.A., Gruber, N., Key, R.M., Lee, K., Bullister, J.L. *et al.* (2004) The oceanic sink for anthropogenic CO2. *Science* **305**: 367-371.

Selje, N., Simon, M., and Brinkhoff, 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. *J Marine Syst* **37**: 151-184.

Sokolov, S., and Rintoul, S.R. (2009a) The circumpolar structure and distribution of the Antarctic Circumpolar Current fronts. Part 2: Variability and relationship to sea surface height. *J Geophys Res-Oceans*: C11.

Sokolov, S., and Rintoul, S.R. (2009b) The circumpolar structure and distribution of the Antarctic Circumpolar Current fronts. Part 1: Mean circumpolar paths. *J Geophys Res-Oceans*: C11.

Strutton, P., Brian Griffiths, F., Waters, R., Wright, S., and Bindoff, N. (2000) Primary productivity off the coast of East Antarctica (80-150°E): January to March 1996. *Deep-Sea Res Pt II* **47**: 2327-2362.

Swan, B.K., Martinez-Garcia, M., Preston, C.M., Sczyrba, A., Woyke, T., Lamy, D. *et al.* (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. *et al.* (2007) The complete genome sequence of Roseobacter denitrificans reveals a mixotrophic rather than photosynthetic metabolism. *J Bacteriol* **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. *OS* **7**: 113-127.

775 Topping, J.N., Heywood, J.L., Ward, P., and Zubkov, M.V. (2006) Bacterioplankton composition
776 in the Scotia Sea, Antarctica, during the austral summer of 2003. *Aquat Microb Ecol* **45**: 229--
777 235.
778 Tripp, H.J., Kitner, J.B., Schwalbach, M.S., Dacey, J.W.H., Wilhelm, L.J., and Giovannoni, S.J.
779 (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* **452**: 741-
780 744.
781 Trull, T., Rintoul, S.R., Hadfield, M., and Abraham, E.R. (2001) Circulation and seasonal
782 evolution of polar waters south of Australia: implications for iron fertilization of the Southern
783 Ocean. *Deep-Sea Res Pt II* **48**: 2439-2466.
784 Vila-Costa, M., Simó, R., Harada, H., Gasol, J.M., Slezak, D., and Kiene, R.P. (2006)
785 Dimethylsulfoniopropionate uptake by marine phytoplankton. *Science* **314**: 652-654.
786 Wagner-Dobler, I., and Biebl, H. (2006) Environmental biology of the marine Roseobacter
787 lineage. *Annu Rev Microbiol* **60**: 255-280.
788 Walker, C.B., de la Torre, J.R., Klotz, M.G., Urakawa, H., Pinel, N., Arp, D.J. *et al.* (2010)
789 Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in
790 globally distributed marine crenarchaea. *Proc Natl Acad Sci U S A* **107**: 8818-8823.
791 Walsh, D.A., Zaikova, E., Howes, C.G., Song, Y.C., Wright, J.J., Tringe, S.G. *et al.* (2009)
792 Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. *Science*
793 **326**: 578-582.
794 Ward, P., Whitehouse, M., Brandon, M., Shreeve, R., and Woodd-Walker, R. (2003)
795 Mesozooplankton community structure across the Antarctic Circumpolar Current to the north of
796 South Georgia: Southern Ocean. *Mar Biol* **143**: 121-130.
797 Weber, T.S., and Deutsch, C. (2010) Ocean nutrient ratios governed by plankton biogeography.
798 *Nature* **467**: 550-554.
799 West, N.J., Obernosterer, I., Zemb, O., and Lebaron, P. (2008) Major differences of bacterial
800 diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the
801 Southern Ocean. *Environ Microbiol* **10**: 738-756.
802 Whitworth III, T. (1980) Zonation and geostrophic flow of the Antarctic circumpolar current at
803 Drake Passage. *Deep-Sea Res* **27**: 497-507.
804 Williams, G.D., Nicol, S., Aoki, S., Meijers, A.J.S., Bindoff, N.L., Iijima, Y. *et al.* (2010) Surface
805 oceanography of BROKE-West, along the Antarctic margin of the south-west Indian Ocean (30-
806 80E). *Deep-Sea Res Pt II* **57**: 738-757.
807 Williams, T.J., Long, E., Evans, F., Demaere, M.Z., Lauro, F.M., Raftery, M.J. *et al.* (2012) A
808 metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula
809 coastal surface waters. *ISME J* **2**: 1-18.
810 Ye, Y., and Doak, T.G. (2009) A parsimony approach to biological pathway
811 reconstruction/inference for genomes and metagenomes. *PLoS Comput Biol* **5**: e1000465.
812 Yoon, J., Yasumoto-Hirose, M., Katsuta, A., Sekiguchi, H., Matsuda, S., Kasai, H., and Yokota,
813 A. (2007) Coraliomargarita akajimensis gen. nov., sp. nov., a novel member of the phylum
814 'Verrucomicrobia' isolated from seawater in Japan. *Int J Syst Evol Microbiol* **57**: 959-963.
815 Zhang, R., Liu, B., Lau, S.C.K., Ki, J.-S., and Qian, P.-Y. (2007) Particle-attached and free-living
816 bacterial communities in a contrasting marine environment: Victoria Harbor, Hong Kong. *FEMS*
817 *Microbiol Ecol* **61**: 496-508.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Neighbour-joining tree of GSO-EOSA-1-like 16S rRNA gene sequences from the SO. Sequence labeled in black text are reads from our metagenomic dataset; in red text are 16S rRNA gene sequences from Gammaproteobacterial Sulphur Oxidisers (GSO) and other gammoproteobacteria from Genbank. The tree was constructed using ARB (Ludwig et al. 2004).

Table S1. Full sample information including physiochemical parameters.

Table S2. Relative abundances of all OTUs in all samples from all size fractions. Size fraction is given in OTU name, e.g. the column "Mycoplasma genitalium strain G37-08" represents the relative abundance of the *Mycoplasma genitalium* G37 OTU in the 0.8 μm size fraction.

Table S3. Contributions of individual OTUs to variance between the North and South zones. Size fraction is given in OTU name, e.g. the row "Mycoplasma genitalium strain G37-08" represents the contribution to variance of the *Mycoplasma genitalium* G37 OTU in the 0.8 μm size fraction.

Table S4. Genes related to aerobic respiration annotated in scaffolds from assemblies of GSO-EOSA-1 affiliated reads. Assembly of reads with identity to *Ruthia magnifica* or *Ca.* Vesicomyosocius okutanii was performed with WGS-ASSEMBLER (Celera, Alameda) and ORFs

predicted with METAGENE (Noguchi et al. 2006). ORFs were annotated with BLASTN against the NCBI nr database (E-value threshold 1.0×10^{-3}). Only selected matches relevant to aerobic respiration are shown.

Table S5. Contributions of KEGG ortholog groups to variance between the North and South zones. Size fraction is given in ortholog group name, e.g. the row "DNA polymerase III subunit alpha [EC:2.7.7.7]-08" represents the contribution to variance of the DNA polymerase III subunit alpha ortholog group in the 0.8 μm size fraction.

Table S6. Contributions of KEGG modules to variance between the North and South zones.

Figure legends

Fig. 1. Sites of samples used in this study. Area depicted is in the Australian sector of the Southern Ocean. Samples are from the North zone (triangles) and South zone (squares). The dashed line gives the approximate position of the Polar Front, associated with a major core of the Antarctic Circumpolar Current (ACC), at the time of the voyage.

Fig. 2. Contribution of OTUs to variance between North and South, 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 total contribution to variance of each size fraction is given as a percentage. Full data are given in Table S3, supporting information.

Fig. 3. 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.

889 **Table 1.** Average OTU abundances for each size fraction and zone

OTU	North Zone			South Zone		
	0.1 µm	**0.8 µm**	**3.0 µm**	**0.1 µm**	**0.8 µm**	**3.0 µm**
Pelagibacter ubique HTCC1062	61.8	25.0	23.9	58.9	22.4	17.6
Synechococcus sp. CC9902	0.11	9.84	4.97	0.00	0.00	0.10
Roseobacter sp. OCh114	0.31	2.93	1.59	0.45	3.99	2.66
Synechococcus sp. CC9311	0.03	4.62	4.41	0.00	0.00	0.03
Ruthia magnifica str. Cm (*Calyptogena magnifica*)	0.67	0.65	0.55	2.99	2.62	1.03
Silicibacter pomeroyi DSS-3	0.26	2.29	1.15	0.31	2.51	1.58
Gramella forsetii strain KT0803	0.24	1.21	1.75	0.50	2.35	1.89

Candidatus Puniceispirillum marinum IMCC1322	0.64	2.08	1.27	0.36	1.38	0.71
Robiginitalea biformata strain HTCC2501	0.28	1.10	1.30	0.47	1.88	1.40
Flavobacterium psychrophilum strain JIP02/86	0.17	0.84	1.22	0.43	1.96	1.60
Silicibacter sp. TM1040	0.23	1.65	0.87	0.27	1.80	1.23
Candidatus Vesicomyosocius okutanii strain HA	0.46	0.46	0.21	1.97	1.81	0.22
Jannaschia sp. DFL-12	0.18	1.38	0.74	0.24	1.69	0.80
Zunongwangia profunda strain SM-A87	0.15	0.75	1.06	0.30	1.41	1.20
Colwellia sp. 34H	0.02	0.36	2.74	0.05	0.51	1.04

Pseudoalteromonas atlantica strain T6c	0.01	0.48	1.99	0.02	0.41	1.13
Jannaschina sp. CCS1	0.12	0.93	0.48	0.17	1.23	0.82
Nitrosopumilus maritimus SCM1	0.02	0.01	0.01	1.08	1.31	1.21
Coraliomargarita akajimensis strain DSM 45221	0.04	0.08	0.12	0.12	1.54	1.68
Flavobacterium johnsoniae strain UW101	0.09	0.42	0.61	0.20	0.94	0.86

890
891

892 * OTU identifications and abundances based on GAAS and MINSPEC analysis of BLAST matches
893 between the SO metagenomic dataset and RefSeq. This table includes the twenty overall most
894 abundant OTUs. A complete of all OTU abundances for all samples and size fractions is available
895 in the supporting information (Table S2). Abundances are relative and expressed as percentages.

896

897 **Table 2.** KEGG modules which contribute highly to variance between the NZ and SZ. A
898 complete list of modules which contribute to variance is given in the supporting information
899 (Table S65).

KEGG module	Average abundance (standardised and log transformed)		Contribution to variance (%)
	North Zone	South Zone	
Photosystem II~~Photosystem II~~	0.57~~0.6~~	0.42~~0.44~~	2.21~~2.18~~
Complex I (NADH dehydrogenase), NADH dehydrogenase I/diaphorase subunit of the bidirectional hydrogenase~~Complex I (NADH dehydrogenase); NADH dehydrogenase I/diaphorase subunit of the bidirectional hydrogenase~~	0.24~~0.26~~	0.01~~0.01~~	1.8~~1.83~~
Photosystem I~~Photosystem I~~	0.34~~0.36~~	0.43~~0.3~~	1.7~~1.2~~
Pyrimidine deoxyribonuleotide biosynthesis, CDP/CTP => dCDP/dCTP,dTDP/dTTP~~Pyrimidine deoxyribonuleotide biosynthesis; CDP/CTP => dCDP/dCTP;dTDP/dTTP~~	0.66~~0.68~~	0.51~~0.53~~	1.16~~1.15~~
Histidine degradation, histidine => N-formiminoglutamate => glutamate~~sn-Glycerol 3-phosphate transport system~~	0.31~~0.17~~	0.42~~0.31~~	1.14~~1.13~~
Methionine salvage pathway~~Methionine salvage pathway~~	0.43~~0.43~~	0.29~~0.29~~	1.14~~1.1~~
sn-Glycerol 3-phosphate transport system~~Histidine degradation; histidine => N-formiminoglutamate =>~~	0.16~~0.33~~	0.29~~0.44~~	1.11~~1.08~~

~~glutamate~~			
Complex I (NADH dehydrogenase), NADH dehydrogenase I~~Aminoacyl-tRNA biosynthesis; prokaryotes~~	1.05~~0.08~~	1.08~~0.14~~	1.06~~1.07~~
Branched-chain amino acid transport system~~Complex I (NADH dehydrogenase); NADH dehydrogenase I~~	0.83~~1.07~~	0.79~~1.1~~	0.96~~1.04~~
Dipeptide transport system~~Glycine betaine/proline transport system~~	0.02~~0.55~~	0.14~~0.67~~	0.95~~1.02~~
Adenine nucleotide biosynthesis, IMP => ADP/dADP,ATP/dATP~~Branched-chain amino acid transport system~~	0.74~~0.86~~	0.62~~0.82~~	0.95~~0.98~~
Glycine betaine/proline transport system~~Adenine nucleotide biosynthesis; IMP => ADP/dADP;ATP/dATP~~	0.56~~0.77~~	0.66~~0.64~~	0.94~~0.97~~
Sulfur reduction, sulfate => H2S~~Dipeptide transport system~~	0.44~~0.02~~	0.54~~0.15~~	0.91~~0.96~~
Simple sugar transport system~~Ribosome; eukaryotes~~	0.39~~0.29~~	0.46~~0.27~~	0.90~~.91~~
Peptides/nickel transport system	0.98	0.99	0.89

The following figures are presented in the order Figure 1-2.