environmental microbiology

Environmental Microbiology (2014) **16**(6), 1566–1578 doi:10.1111/1462-2920.12287

Unveiling abundance and distribution of planktonic *Bacteria* and *Archaea* in a polynya in Amundsen Sea, Antarctica

Jong-Geol Kim,[1] Soo-Je Park,[2] Zhe-Xue Quan,[3] Man-Young Jung,[1] In-Tae Cha,[1] So-Jeong Kim,[1] Kyoung-Ho Kim,[4] Eun-Jin Yang,[5] Young-Nam Kim,[5,6] Sang-Hoon Lee[5] and Sung-Keun Rhee[1]*

[1]*Department of Microbiology, Chungbuk National University, 12 Gaeshin-dong, Heungduk-gu, Cheongju 361-763, South Korea.*

[2]*Department of Biology, Jeju National University, 102 Jejudaehakno, Jeju, 690-756, Rep. of Korea.*

[3]*Department of Microbiology and Microbial Engineering, School of Life Sciences, Fudan University, Shanghai 200433, P. R. China.*

[4]*Department of Microbiology, Pukyong National University, Busan 608-737, South Korea.*

[5]*Division of Polar Climate Research, Korea Polar Research Institute, Incheon 406-840, South Korea.*

[6]*Marine Ecosystem Management Team, Korea Marine Environment Management Corp. (KOEM), Haedong Building, Samseong-dong71, Gangnam-gu, Seoul 135-870, South Korea.*

Summary

Polynyas, areas of open water surrounded by sea ice, are sites of intense primary production and ecological hotspots in the Antarctic Ocean. This study determined the spatial variation in communities of prokaryotes in a polynya in the Amundsen Sea using 454 pyrosequencing technology, and the results were compared with biotic and abiotic environmental factors. The bacterial abundance was correlated with that of phytoplankton, *Phaeocystis* spp. and diatoms. A cluster analysis indicated that the bacterial communities in the surface waters of the polynya were distinct from those under the sea ice. Overall, two bacterial clades, *Polaribacter* (20–64%) and uncultivated *Oceanospirillaceae* (7–34%), dominated the surface water in the polynya while the *Pelagibacter* clade was abundant at all depths (7–42%). The archaeal communities were not as diverse as the bacterial communities in the polynya, and marine group I was dominant (> 80%). Canonical correspondence analysis indicated that the oceanographic properties facilitated the development of distinct prokaryotic assemblages in the polynya. This analysis of the diversity and composition of the psychrophilic prokaryotes associated with high phytoplankton production provides new insights into the roles of prokaryotes in biogeochemical cycles in high-latitude polynyas.

Received 12 November, 2012; accepted 12 September, 2013. *For correspondence. E-mail rhees@chungbuk.ac.kr; Tel. (+82) 43 261 2300; Fax (+82) 43 264 9600.

Introduction

Most of the pelagic communities in polar seas are covered with seasonal sea ice, but ice-free sections known as polynyas develop regularly because of the non-uniform melting of the sea ice pack (Barber and Massom, 2007). It has been suggested that the upwelling of warm circumpolar deep water leads to the melting of sea ice around the polynyas in the Amundsen Sea (Jenkins *et al.*, 2010). The strong effects of solar radiation, nutrients and temperature in polynyas facilitate the production of major phytoplankton blooms (Smith and Gordon, 1997). Indeed, the polynyas in the Amundsen Sea are regarded as one of the most biologically productive regions (reaching up to 5.54 Tg C $year^{-1}$) in the world's oceans, and they are an ecological hotspot in the Antarctic Ocean (Arrigo and van Dijken, 2003). This high productivity combined with the formation of Antarctic bottom water contributes to the formation of organic-rich deep water and biological carbon sequestration (Siegenthaler and Wenk, 1984). Polynyas also release heat and moisture to the atmosphere and are sites for the exchange of greenhouse and ozone-depleting gases (N_2O, CH_4, DMS, etc.) between the atmosphere and the ocean in the Antarctic (Miller and DiTullio, 2007). Annual variations in the opening times and size of polynyas are associated with climate and current changes in polar regions, which also affect the structure and function of microbial communities in polynyas.

Bacteria are abundant and experience major seasonal variations in polar oceans, where they are involved with many biogeochemical cycles and food webs (Comeau *et al.*, 2011; Ghiglione and Murray, 2011). The activity levels of these cold-adapted bacteria are associated with

the remineralization of fixed carbon, which contributes to the efficiency of carbon sequestration at high latitudes. The associations of heterotrophs, including *Cytophaga-Flavobacteria-Bacteroides*, *Gammaproteobacteria* and *Alphaproteobacteria*, with phytoplankton blooms have been reported in polar oceans (Claire Horner-Devine *et al*., 2003). Several groups of planktonic *Archaea*, including *Thaumarchaeota* IA [Marine Group I (MGI)] (Mincer *et al*., 2007) and *Euryarchaeota* (Marine Group II, III and IV) (Galand *et al*., 2009; Alonso-Sáez *et al*., 2011), have been discovered in various oceanic regions, including polar oceans (Church *et al*., 2003; Galand *et al*., 2009; Alonso-Sáez *et al*., 2011).

A detailed analysis of the diversity and abundance of prokaryotes in polar oceans is essential for understanding the biogeographical compositions of microbial communities. In Antarctica, several multidisciplinary studies have been conducted in polynyas in the Mertz Glacier, Ross Sea and Cosmonaut regions (Bowman and McCuaig, 2003; Ayton *et al*., 2010). However, the assemblages of prokaryotes found in polynyas remain largely unknown because of a lack of comprehensive analyses of the diversity, distribution and abundance of *Bacteria* and *Archaea*. Therefore, next-generation sequencing techniques based on 16S ribosomal RNA (rRNA) genes facilitate the investigation of archaeal and bacterial communities in various environments, and the coverage obtained using these methods far exceeds that in previous clone libraries (Comeau *et al*., 2011; Qian *et al*., 2011).

There have been no extensive descriptions of the microbial communities found in polynyas in the Southern Antarctic Ocean. Distinct psychrophilic prokaryotes might be associated with ice retreat and the high level of primary production observed in these areas. In this study, the spatial variations in psychrophilic prokaryote communities were described and compared with biotic and abiotic environmental factors, thereby providing new insights into the roles of psychrophilic prokaryotes in biogeochemical cycles, including carbon remineralization in high-latitude polynyas. Twelve samples were obtained from different depths at three stations within a polynya in the Amundsen Sea, Antarctica (Fig. 1). Five samples were also included from ice-free open ocean and surface waters below the sea ice coverage, which differed greatly from the polynyas in terms of the light intensity and temperature. Pyrosequencing was used to obtain a detailed taxonomy of the prokaryotic assemblages in the Antarctic polynya, including low-abundance taxa. The average read length obtained in this study (200–300 bp) was much higher than those reported in previous studies (ca 100 bp) of polar oceans (Galand *et al*., 2009; Ghiglione and Murray, 2011), which facilitated a more accurate taxonomic assignment, with a higher phylogenetic resolution in the community analysis.

Results and discussion

Oceanographic data and microbial abundance

The vertical profiles of the salinity, temperature, oxygen and fluorescence at the sampling stations are shown in Fig 2 (three polynya stations) and Supplementary Fig. S1 (other stations). There were dynamic vertical changes in the environmental variables measured in the surface waters (between 0 and 100 m). At the polynya stations, i.e. ST13 (centre), ST8 (margin close to sea ice) and ST11 (ice shelf in the vicinity of the Getz Ice Shelf), the maximum chlorophyll *a* (chl. *a*) fluorescence was recorded at shallow depths (< 30 m), and the maximum chl. *a* concentration was 4–10 mg m^{-3}. Deep (ca 200 m) fluorescence (i.e. chl. *a*) penetration was observed at the polynya ice shelf station (Fig. 2). The highest chlorophyll concentration was recorded at the polynya centre station, which was completely open. The oceanographic data (i.e. low temperature and halocline stratification) indicated the relatively recent opening of the polynya margin station. The oxygen concentration was > 200 μmol kg^{-1}, and there was no oxygen minimum zone. The temperatures of the surface water at the polynya centre and at the ice shelf station were higher than those at the sea ice stations. The intensities of fluorescence at the two sea ice stations (ST17 and ST26) and an ice-free open ocean station (ST1) were lower than those at the three polynya stations (Supplementary Fig. S1). The presence of a steep halocline was obvious only at ST8, ST26 and ST1. Relatively warm (ca 0°C) water was detected below 400 m at the polynya and sea ice stations, which might have been due to the intrusion of Antarctic circumpolar deep water (Fig. 2).

The concentrations of nitrate + nitrite and dissolved silicate were in the ranges of 14.29–17.43 μmol l^{-1} and 79.42–91.27 μmol l^{-1}, respectively, which increased with depth at the three polynya stations (Supplementary Table S1). The nitrate + nitrite concentrations in the surface waters of the sea ice stations were higher than those of the polynya stations. However, the silicate concentrations in the surface waters of the sea ice stations were lower than those of the polynya stations. This might have been related to differences in the phytoplankton abundance and composition between the polynya and sea ice stations. In fact, the abundance and composition of the phytoplankton communities differed significantly among the polynya, sea ice and ice-free open ocean stations (Supplementary Table S2). *Phaeocystis antarctica* was dominant in the polynya stations while diatom was dominant in the polynya and ice-free open ocean. The prokaryotic abundance tended to decrease from the surface water to the deep water. The abundance of *Bacteria* relative to *Archaea* was significantly higher in the polynya stations than the sea ice stations (Supplementary Table S1). The

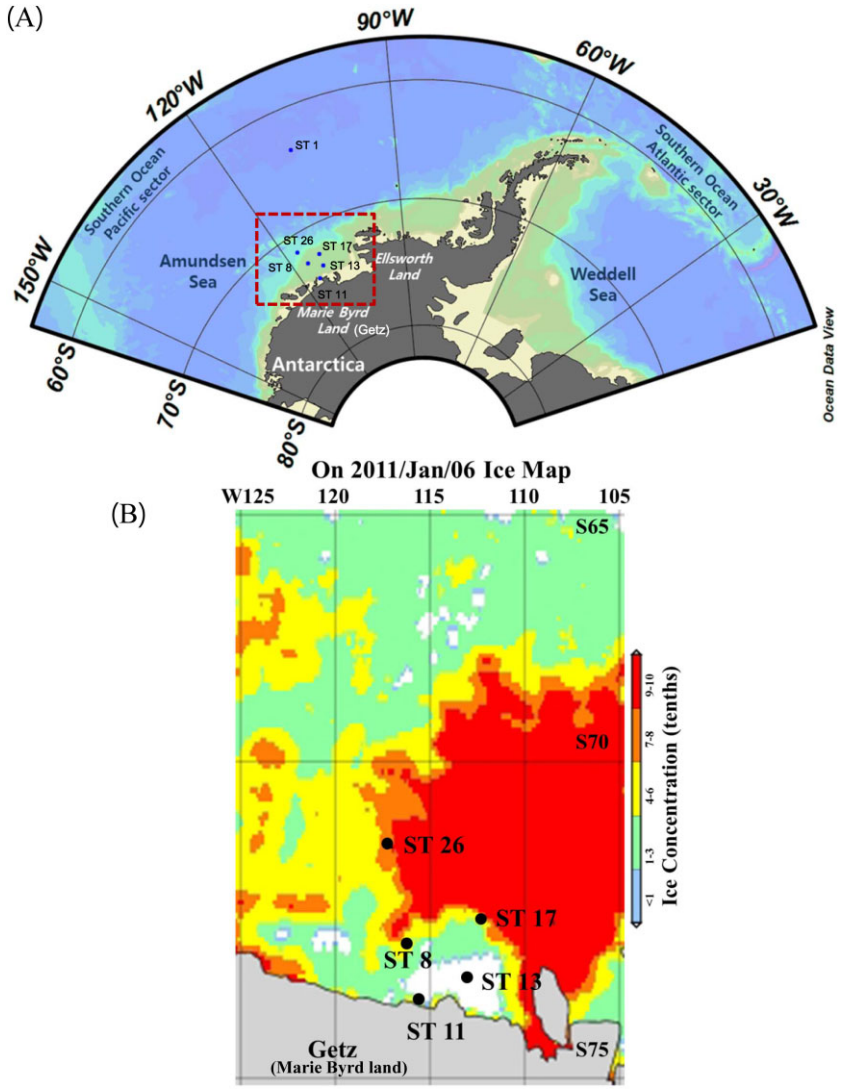

Fig. 1. (A) Map of the Amundsen Sea in Antarctica. The boxed area in A is magnified in B and shows the ice concentration data. The sampling stations are indicated with numbered black dots, i.e. the polynya [ST13 (centre), ST8 (margin, close to sea ice) and ST11 (ice shelf, in the vicinity of the Getz Ice Shelf)], sea ice (stations 17 and 26) and ice-free open (station 1) sampling stations.

bacterial abundance was highly correlated with *Phaeocystis* spp. and diatoms ($R > 0.9$, $P < 0.05$), which indicates that there may be a close coupling between the phytoplankton and prokaryote communities in Antarctic polynyas, as previously reported (Piquet *et al.*, 2011).

Based on the vertical variation in the oceanographic data, the following depth samples were selected for prokaryotic community analysis: surface (5–30 m) and deep (200–1000 m) waters at the polynya stations (as indicated in Fig. 2). Samples from sea ice stations (surface waters: 2–60 m) and ice-free open ocean stations (surface water: 10 m) far from the frozen Amundsen Sea (as indicated in Supplementary Fig. S1) were included for comparison.

Estimation of richness and diversity

The 454 pyrosequencing platform produced approximately 383 000 (ca 225 975 *Bacteria* and 157 025 *Archaea*) raw reads for 17 samples from six stations. After removing the low-quality reads, 71 748 bacterial and 57 568 archaeal reads remained, which were used for further analyses. After trimming the primer sequences, the average length of the bacterial reads was ca 300 bp, and that of the archaeal reads was ca 200 bp. Sequences related to plastids were recovered frequently (16–27%) from surface water samples (ST1_10 m, ST11_200 m, ST17_2 m, ST17_60 m and ST26_40 m) that were not included in this study.

The bacterial and archaeal diversities (Shannon index) were noticeably higher in deep waters (≥ 200 m) than surface waters (5–30 m) at polynya stations, where the highest diversity occurred at a depth of ca 200 m (Table 1). The bacterial and archaeal diversities in the sea ice station surface waters were higher than those in the polynya station surface waters. Overall, there were similar depth-wise increases in the diversity indices, the numbers of operational taxonomic units (OTUs) and the

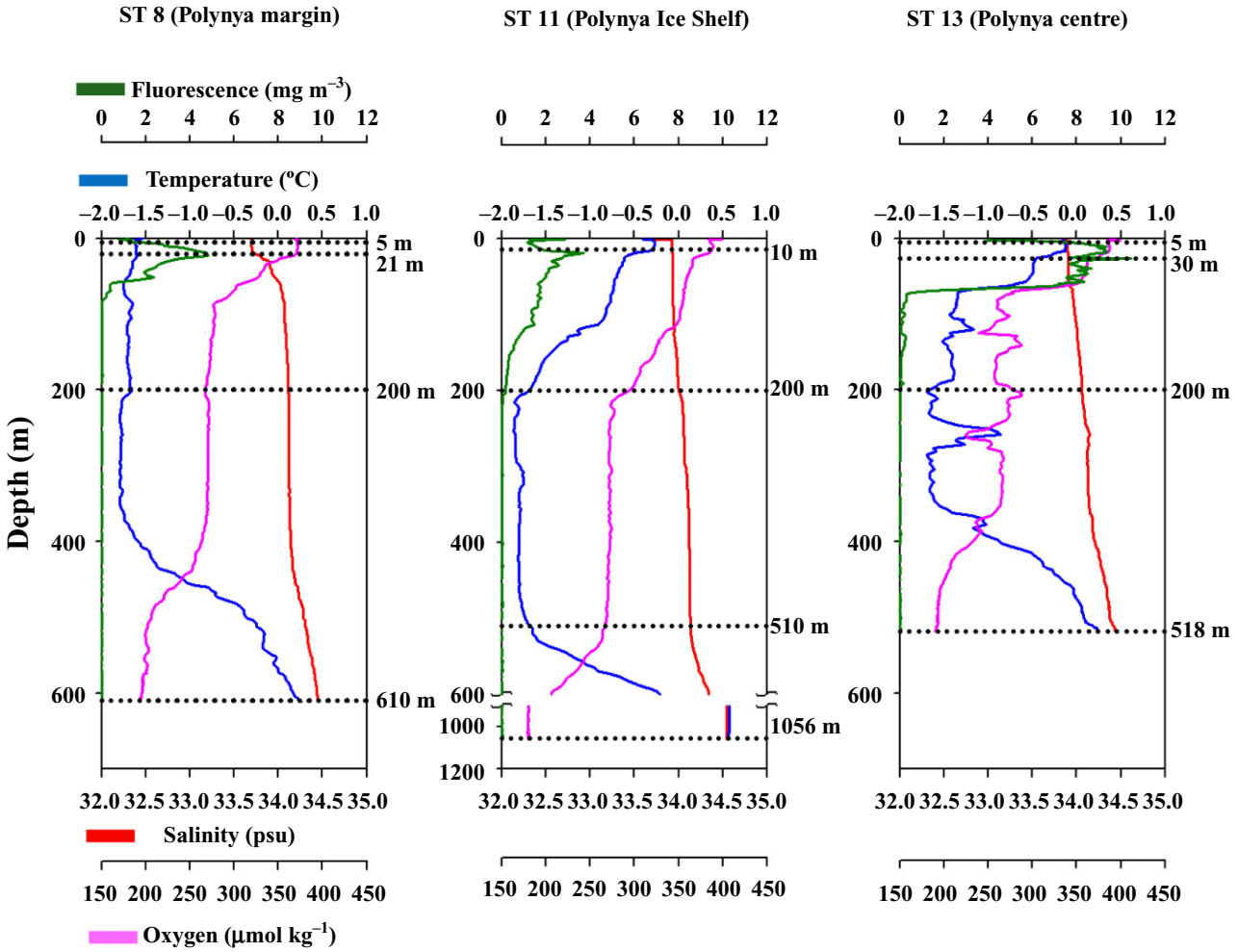

Fig. 2. Vertical profiles of the physicochemical characteristics (fluorescence, temperature, salinity and dissolved oxygen) throughout the water mass of the polynya, i.e. ST8 (polynya margin close to sea ice), ST11 (ice shelf in the vicinity of the Getz Ice Shelf) and ST13 (polynya centre). The dashed lines indicate the depth of the water sample used in the analysis.

abundance-based coverage estimator (ACE) and Chao1 estimators for *Bacteria* and *Archaea* at the polynya centre and polynya margin stations. Most of the rarefaction curves of the samples were saturated, although some stations required more sequencing effort to obtain the degree of saturation for bacteria (Supplementary Fig. S2). Good's coverage was high, with an average of 0.97 based on all of the samples (Table 1).

Comparison of the 16S amplicons among samples

The similarity/dissimilarity of the 16S rRNA gene sequences was measured among different samples using the weighted UniFrac clustering method. Bacterial communities from the deep waters at the polynya stations formed a cluster (Fig. 3). The clustering of bacterial communities in the surface waters from the polynya centre and margin stations was obvious ($P < 0.3$ with other stations). A *t*-test based on the distances of the UniFrac data suggested that the bacterial communities in the surface waters from the polynya ice shelf station were too distinct to be clustered with the other polynya stations ($P = 0.09$). The difference might have been related to the deep penetration of chl. *a* at the polynya ice shelf station. The bacterial communities at the sea ice stations were similar, with the exception of ST26_40 m, which was divided by a strong halocline (see Supplementary Fig. S1) and was quite different from the others. It is arguable at this stage that the prokaryotic communities detected in the surface waters of the polynya centre and margin stations were typical of the polynyas in the Southern Antarctic Ocean. The archaeal communities were separated into two clear clusters where the cut-off value was at a 20% similarity distance. The surface water sample from the ice-free open ocean station was distinct from all of the others.

Table 1. Estimates of phylotype richness, diversity and coverage for the bacterial and archaeal assemblages.

	Bacteria						*Archaea*					
Sample	Reads	OTU	ACE	Chao 1	Shannon	Good	Reads	OTU	ACE	Chao 1	Shannon	Good
ST8 5 m	5442	79	175	143	2.60	0.96	3036	13	57	27	0.90	0.99
ST8 21 m	4525	93	181	152	2.83	0.96	5364	27	73	52	1.45	0.99
ST8 200 m	2906	165	272	247	4.04	0.93	4412	17	28	23	1.00	1.00
ST8 610 m	6274	115	147	147	3.65	0.97	4236	34	43	45	1.81	0.99
ST11 10 m	3361	76	131	118	2.40	0.97	4710	6	13	11	0.05	1.00
ST11 200 m	3038	153	263	225	3.99	0.94	3342	26	154	65	1.18	0.99
ST11 510 m	4899	189	332	291	4.19	0.92	3404	19	27	25	1.02	0.99
ST11 1056	2330	67	73	78	3.48	0.99	3499	12	28	21	0.49	0.99
ST13 5 m	3984	62	191	118	2.61	0.97	3365	13	47	28	0.41	0.99
ST13 30 m	4872	61	210	134	2.42	0.97	2640	11	37	26	0.29	0.99
ST13 250 m	4919	213	472	362	4.28	0.90	3679	27	50	37	1.72	0.99
ST13 518 m	6558	133	384	291	3.55	0.93	5465	56	235	133	1.67	0.97
ST17 2 m	3988	113	202	177	3.59	0.96	1815	15	27	22	0.30	0.99
ST17 60 m	3418	89	178	170	3.08	0.96	3072	17	61	29	0.57	0.99
ST26 2 m	4178	109	227	175	3.37	0.95	1023	22	24	24	1.29	1.00
ST26 40 m	3645	69	166	109	2.47	0.97	2505	12	21	16	1.10	1.00
ST1 10 m	3411	83	152	130	3.14	0.97	2001	24	50	39	1.03	0.99

An OTU was defined as containing sequences with ≤ 3% difference. The diversity indices were calculated using resampled data. The resampling of the subsamples with the lowest number of reads (1023 reads) was repeated 100 times, and the average values are presented.Good, Good's coverage.

The principal components analysis analysis supported the clustering results (Supplementary Fig. S3).

Bacterial community composition

Pyrosequencing of the 16S rRNA gene sequences showed that the *Bacteria* were dominated by heterotrophic bacteria, i.e. *Bacteroidetes*, *Gammaproteobacteria* and *Alphaproteobacteria*, in the polynya (Fig. 4 and Supplementary Fig. S4), which agreed with previous analyses of Antarctic oceans using molecular approaches (Abell and Bowman, 2005; Gentile *et al*., 2006). There were obvious differences in the bacterial community compositions of the surface water and deep water in the polynya. High proportions of bacterial reads (> 98%) from the surface waters could be classified into specific phyla at all stations. Deep waters collected from the polynya stations contained a high proportion of unclassified *Bacteria* (10–20%) (Fig. 4). In addition, minor amounts of sequences related to *Actinobacteria*, *Deltaproteobacteria* and unclassified *Bacteria* were obtained from the surface waters, whereas the relative abundances of these groups increased up to 9%, 23% and 18%, respectively, in deep waters.

The phylum *Proteobacteria* comprised 26–84% of the total reads from all samples. *Alphaproteobacteria* and

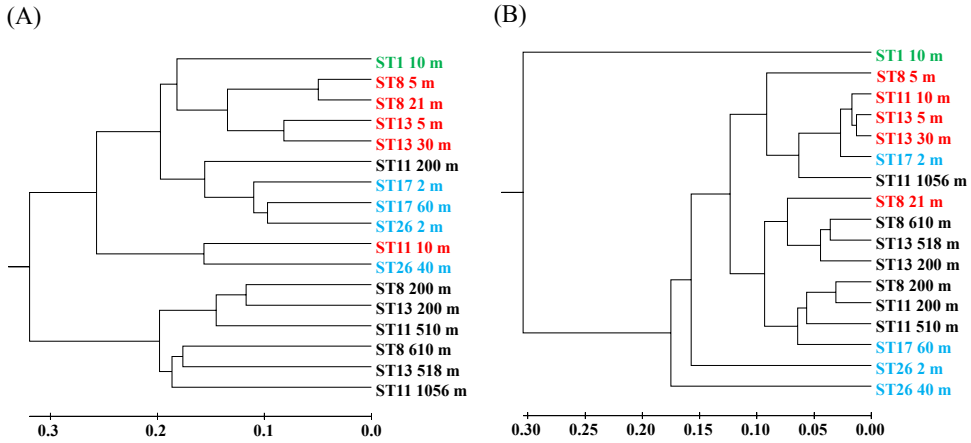

Fig. 3. UniFrac clustering relationships estimated from (A) bacterial and (B) archaeal OTU reads. The clusters are indicated by the colour of the text: green, ice-free open ocean; red, surface water from the polynya; blue, surface water from sea ice; black, deep water from the polynya.

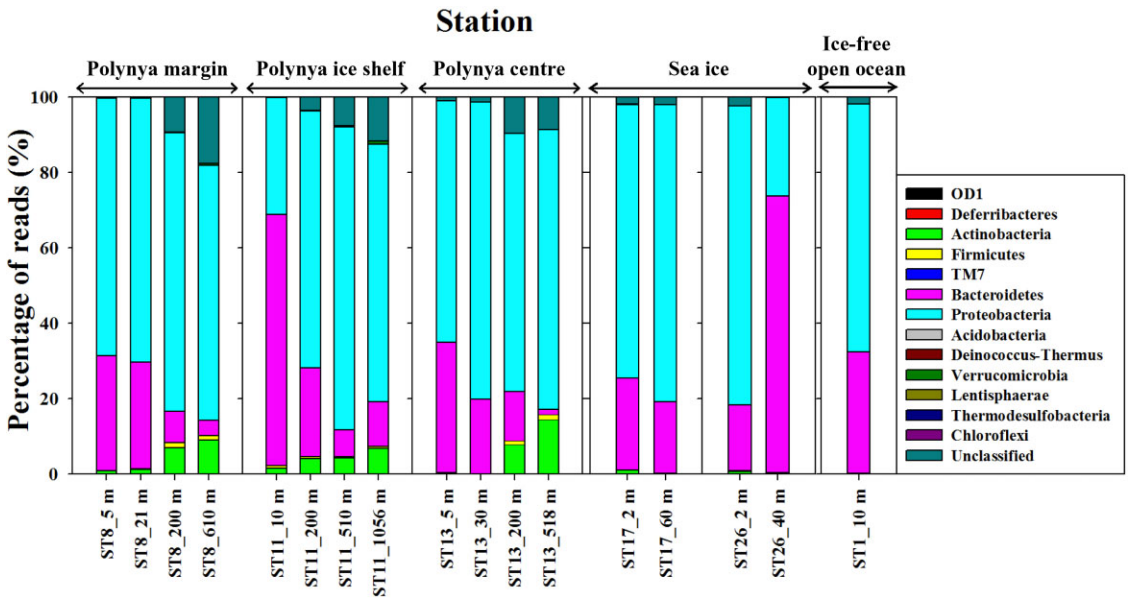

Fig. 4. Stacked column graphs showing the relative abundance of bacterial reads at the phylum level.

Gammaproteobacteria were the most abundant the subphyla of the *Proteobacteria* (Supplementary Fig. S4). The majority of *Alphaproteobacteria* reads were classified as SAR11 (Morris *et al*., 2005), i.e. 25–62% and 7–42% of the total proteobacterial and bacterial reads respectively. The SAR11 clade of *Alphaproteobacteria* comprises very small marine heterotrophs, which are found throughout the oceans where they account for about 25% of all microbial cells (Morris *et al*., 2002). The subgroup SAR11_1a sequence reads were dominant at all depths (Supplementary Fig. S5). Subgroup SAR11_1a is represented by *Pelagibacter ubiqua*, which is known to be oligotrophic and unresponsive to eutrophic conditions and light (Giovannoni *et al*., 2005; Alonso and Pernthaler, 2006; Stingl *et al*., 2007). The relative abundance of minor subgroups, i.e. SAR11_1b and SAR2, increased with the depth in the SAR11 clade.

The *Gammaproteobacteria* produced three major groups of reads: phylotypes affiliated with uncultivated *Oceanospirillaceae*, SUP05 (with > 99% 16S rRNA intragene sequence similarity) and a SAR92-like group (Fig. 5). The sequences related to the uncultivated *Oceanospirillaceae* were abundant only in surface waters (7–34%), whereas the SUP05 sequences were abundant in deep waters at the polynya stations (6–16%). The SUP05 sequences were affiliated with lineages represented by the environmental clone 'ARCTIC96BD-19' in the sulphur-oxidizing cluster (Swan *et al*., 2011) (with 99% 16S rRNA gene similarity), which is found in extensive oxygen minimum zones (OMZs) (Paulmier and Ruiz-Pino, 2009). However, no OMZs were present in the study stations, as shown in Fig. 2 (the O_2 concentrations were > 200 μM). In agreement with this, a recently isolated member of the SUP05/Arctic96BD-19 clade from the North Pacific gyre was inferred to be an aerobic, sulphur-oxidizing heterotroph (Marshall and Morris, 2013). The strictly anaerobic *Epsilonproteobacteria* detected in an OMZ water at Rod Bay station during a strong phytoplankton bloom (Gentile *et al*., 2006) were not observed in this study. Intriguingly, SUP05 comprised up to 23% of sequences from surface waters at the sea ice stations where relatively low abundances of *Polaribacter* and uncultivated *Oceanospillilaceae* were observed. The SAR92-like group was detected from the surface waters of the polynya and the sea ice station. Previously, Ghiglione and colleagues (2012) found that the SAR92-like group was dominant in surface waters close to the sea coast in Antarctica, and proteorhodopsin-coding gene sequences from this group are frequently represented in the Arctic Ocean (Cottrell and Kirchman, 2009).

The *Deltaproteobacteria* clusters, SAR324 and *Nitrospina*, were dominant in the deep waters of the polynya, in addition to the SUP05 cluster (Fig. 5). The SAR324 cluster, another typical deep water cluster, has also been reported from the Arctic, the Antarctic and North Atlantic (López-García *et al*., 2001; Ghiglione *et al*., 2012). The SAR324 clusters have recently been implicated as mixotrophic, based on evidence from single-cell genome sequencing (Swan *et al*., 2011). Sequences related to the *Nitrospina* cluster, which mediate nitrite oxidation, were detected frequently in deep waters at the ALOHA station (Mincer *et al*., 2007) and in the Arctic Ocean (Alonso-Saez *et al*., 2010). Nitrite-oxidizing bacteria (NOB) or their sequences have not been obtained from

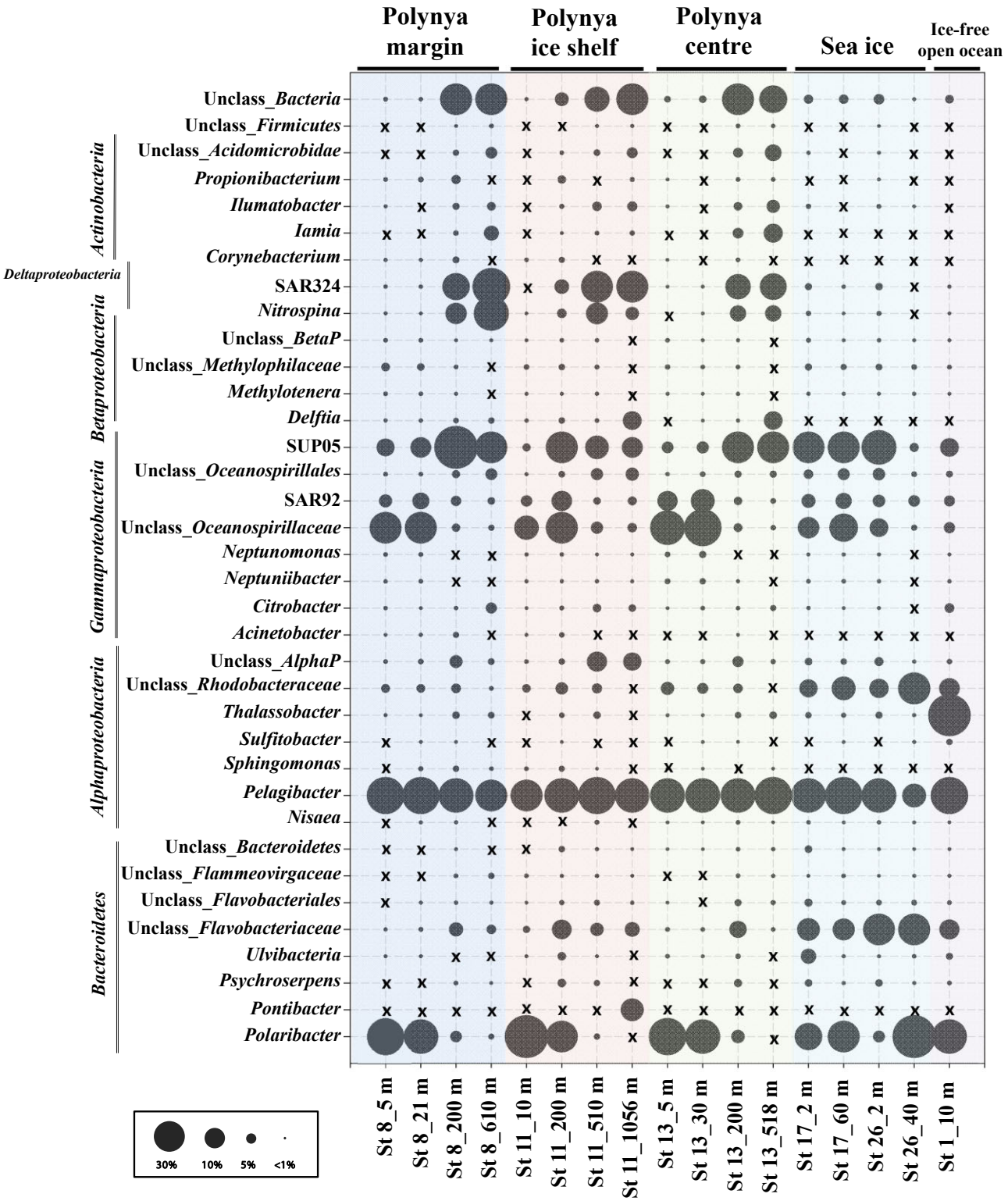

Fig. 5. Taxonomic distribution of the dominant bacterial OTUs based on a 97% similarity cut-off. The most abundant OTUs are compared for each sampling station (total = 36 OTUs), where the circle size indicates the relative average abundance of the OTU in each cluster. OTUs that lacked average relative abundances are indicated with an x. Unclass, unclassified; AlphaP, *Alphaproteobacteria*; GammaP, *Gammaproteobacteria*; BetaP, *Betaproteobacteria*; DeltaP, *Deltaproteobacteria*.

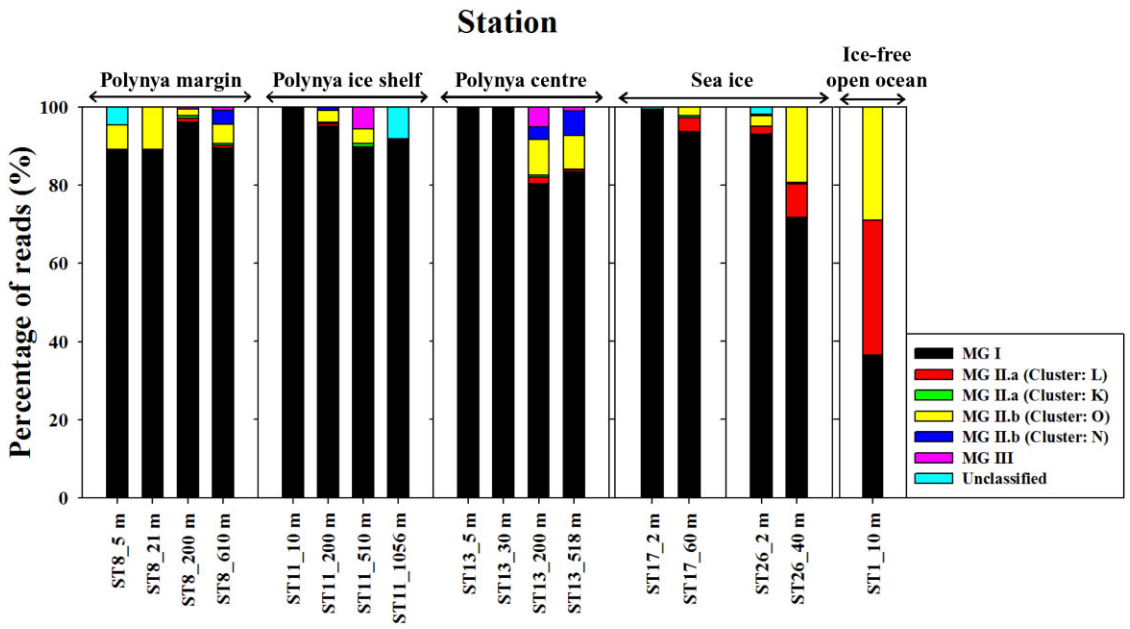

Fig. 6. Stacked column graphs showing the relative abundance of archaeal reads at the phylum and subgroup level.

the Antarctic, so the identification of NOB in Antarctic polynyas demands more sensitive molecular surveys and cultivation studies.

Sequence reads affiliated with the phylum *Bacteroidetes* were abundant in the surface waters at all stations, but they declined sharply with increasing depth (Fig. 4). The phylum *Bacteroidetes* comprised 30–64% of the total bacterial reads in polynya surface waters. Most of the sequences (> 98%) related to *Bacteroidetes* were affiliated with the genus *Polaribacter* (Fig. 5). At the bottoms of the polynya stations, however, the sequences related to *Polaribacter* comprised less than 1% of the total. *Polaribacter* was highly abundant in the surface waters of the polynya (see Fig. 5), and it is a typical cold-adapted genus that has been detected frequently in Antarctic oceans with high temporal and geographical variations (Abell and Bowman, 2005; Gentile *et al.*, 2006). Thus, the abundance of *Polaribacter* and the uncultivated members of the family *Oceanospillilaceae* in the surface waters might be associated with the phytoplankton blooms that occur in polynyas in the Amundsen Sea.

Archaeal community composition in the polynya

The archaeal sequence reads from all stations could be classified into MGI, MGII and MGIII, and only 0.2–8% of the reads were from unclassified phyla (Fig. 6). The high similarity of archaeal communities according to the cluster analysis (see above) may have been due to that *Thaumarchaeota* (MGI) were dominant (70–99%) compared with *Euryarchaeota* (MGII and III) in most of the samples from the polynya and sea ice stations. At the polynya centre and ice shelf stations, the MGI count accounted for almost 100% of the sequences in the surface waters, whereas euryarchaeotal reads were obtained only from the deep waters (up to 10–25%). The most abundant MGI sequences, which represented > 90% of the identified tags, belonged to the cluster affiliated with the cultivated ammonia-oxidizing archaeon (AOA) *Candidatus Nitrosopumilus maritimus* (Könneke *et al.*, 2005), with > 98% 16S rRNA gene sequence similarity. The absence of ammonia-oxidizing bacteria (AOB) sequences from the surface waters suggests that nitrite + nitrate production in the polynya might be mediated by chemoautotrophic AOA from MGI. MGI accounted for ca 1–18% of the total bacterial abundance in the deep water of the polynya [see the quantitative polymerase chain reaction (PCR) data in Supplementary Table S1]. AOB sequences (*Nitrosospira* cluster) were only minor components of the deep waters (about 0.1–0.5% of the total bacterial sequences). AOA, which have a high affinity for ammonium, may be better adapted to the polynyas of the Amundsen Sea where the ammonium concentration ranges from 0 μM to 1.57 μM (Supplementary Table S1) compared with AOB (Martens-Habbena *et al.*, 2009). The distribution patterns of planktonic *Thaumarchaeota* and *Nitrospina* species support possible metabolic interactions between these groups as observed in the ocean's water column (Mincer *et al.*, 2007) and sediments (Park *et al.*, 2010).

The euryarchaeotal sequences related to MGII [cluster L (MGII.a)] according to the classification system of Galand

and colleagues (2010) were minor components at the polynya stations, although the abundance of MGII was positively correlated with the chl. *a* concentration in the North Sea (Herfort *et al*., 2007). MGII sequences have been retrieved frequently from surface waters in the Arctic Ocean (Galand *et al*., 2009) and Pacific Ocean (Galand *et al*., 2009; Iverson *et al*., 2012). The MGII sequences were dominant (70%) compared with thaumarchaeotal sequences in the surface waters collected from the ice-free open ocean. Cluster L (MGII.a) was also frequent in the surface waters of the sea ice stations. Cluster O (MGII.b) was observed in most samples, whereas cluster N (MGII.b) and MGIII were observed only in the deep waters at the polynya stations. A recent metagenomic study indicated that MGII members are heterotrophic and phototrophic (proteorhodopsin-containing) in surface waters (Iverson *et al*., 2012). The archaeal counts from exposed surface waters (the polynya and ice-free open ocean stations) were about 10 times lower than those obtained from the waters covered by sea ice (sea ice stations) (see Supplementary Table S1). Because most of the archaeal sequences obtained from the polynya and sea ice stations were related to MGI in *Thaumarchaeota*, the oligotrophic nature of AOA from MGI and their light sensitivity (French *et al*., 2012) might partly explain the comparatively low copy numbers of archaea (MGI) in the polynya.

Primary producers in the Polynya

Polynyas are considered to be crucial areas for the fixation and sequestration of atmospheric carbon in the Antarctic Ocean (Arrigo and van Dijken, 2003). In this study, most of the primary production in the polynya stations may have been attributable to eukaryotic phytoplankton based on the high counts of *P. antarctica* (see Supplementary Table S2) and the lack of sequences related to *Cyanobacteria*. *Phaeocystis antarctica* is a numerically and functionally important component of the polar biogeochemical cycles (Weisse *et al*., 1994; Stoecker *et al*., 1995), and mesoscale blooms of *P. antarctica* appear to be frequent and widespread in polar regions, which was also the case in this study (Supplementary Table S2).

Marine *Cyanobacteria* (*Prochlorococcus* and *Synechococcus*) are common in many temperate coastal regions, as well as subtropical and tropical open ocean waters, but the absence of marine *Cyanobacteria* might be a typical characteristic of high-latitude Antarctic waters (Johnston *et al*., 2005; Ghiglione and Murray, 2011) because it has been reported in other polar oceans (Taton *et al*., 2003; Johnston *et al*., 2005). It is possible that *Cyanobacteria* were underestimated because of the low-binding efficiency of PCR primers with the cyanobacterial 16S rRNA gene, as suggested by Bowman and colleagues (2012). Díez and colleagues (2012) reported that a high diversity of functional cyanobacterial genes (e.g. *nif*H and *het*R) were involved with nitrogen fixation in the Arctic Ocean, despite their failure to detect *Cyanobacteria* based on 16S rRNA gene analyses. Nitrogen was not a limiting nutrient for primary production in the Amundsen Sea (see Supplementary Table S1), so the involvement of highly active diazotrophic *Cyanobacteria* in phytoplankton blooms needs to be tested using activity-based approaches. Although Murray and Grzymski (2007) reported that *Thaumarchaeota* may also provide a significant CO_2 sink in the ocean ecosystem, the contribution of MGI to the carbon cycle could be negligible given their relatively low abundance in the polynya (Supplementary Table S1).

Factors influencing on prokaryotic community structure

In this study, bacteria exhibited significant shifts in their abundance and compositions, which coincided with changes in phytoplankton abundance and environmental variables. The relationships between bacterial assemblages and environmental variables corresponded with the habitat descriptions of the major bacterial phylotypes according to the canonical correspondence analysis (CCA) (Supplementary Fig. S6 and Table S3) and Pearson's correlation coefficient analysis: *Polaribacter* was found mainly in habitats with high oxygen concentrations ($R = 0.70$, $P < 0.002$), and high-fluorescence ($R = 0.91$, $P < 0.001$), uncultivated *Oceanospirillaceae* were also found mainly in habitats with high oxygen concentrations ($R = 0.61$, $P < 0.002$), and *Pelagibacter* appeared to be relatively independent of the environmental variables. The depth, salinity and SiO_2 were significantly positively correlated ($R > 0.52$, $P < 0.05$), and they were negatively correlated with oxygen ($R \leq 0.50$, $P < 0.05$). The distribution of archaeal phylotypes was weakly correlated with changes in the environmental variables according to the CCA.

In conclusion, the present study used pyrosequencing with high coverage and high taxonomic resolution of the microbial communities to detect distinct communities of *Bacteria* and *Archaea* in an Amundsen Sea polynya. The overall vertical trends in the compositions of bacterial communities in the polynya centre and polynya margin stations were consistent. The bacterial community compositions in the surface waters of the polynya were distinct from those of the sea ice stations, which may have been attributable to differences in physicochemical factors, including sunlight, salinity and phytoplankton production promoted by the receding sea ice. The successional stages of the prokaryotic communities associated with sea ice recession could be verified with appropriate spatial replicates and temporal samplings in the future. The spatial variation of the prokaryotic assemblages in

the polynya and the ocean below the sea ice may provide insights that facilitate a better understanding of the biogeochemical processes mediated by psychrophilic prokaryotes in areas of high primary production in high-latitude Antarctic regions.

Experimental procedures

Collection of water samples

Water samples were collected from stations at a polynya (polynya centre and margins), sea ice and ice-free open oceans as part of the Amundsen Sea project 2010 expedition, which was conducted on the Korean icebreaker Araon between 21 December and 23 January 2011 (Fig. 1). Twelve vertical samples were collected from the polynya centre, the polynya margin, the polynya ice shelf stations, four surface water samples from the sea ice stations and surface water from an ice-free open ocean station (Supplementary Table S1). At each location, a rosette water sampler equipped with 10 Niskin bottles (10 l) and a conductivity-temperature-depth unit, which monitored the pressure, temperature, salinity, dissolved oxygen and *in situ* chl. *a* fluorescence, were deployed to collect seawater. Each water sample (4 l) was filtered immediately through a 1.2 μm pore-sized filter (Supor polyethersulfone, Pall Life Sciences, Ann Arbor, MI, USA) to remove any suspended particles and eukaryotes, before being filtered through a 0.22 μm pore-size filter (Supor polyethersulfone; Pall Life Sciences) to capture prokaryotic cells. The filters were preserved at −80°C until DNA extraction, which was conducted after enzymatic lysis and phenol : chloroform purification (Massana *et al*., 1997).

The samples used for nutrient analysis were collected in duplicate and were analysed on board the ship using a Bran and Luebbe model Quatro AA (Seal Analytical, Hampshire, UK). The water samples used to analyse phytoplankton were fixed with glutaraldehyde and Lugol's solution, and enumerated under epifluorescence with an inverted microscope (Nikon type 104, Tokyo, Japan and Olympus IX 70, Tokyo, Japan).

Quantification of prokaryotic 16S rRNA genes

Bacterial and archaeal 16S rRNA gene copies were quantified using a MiniOpticon real-time PCR detection system (BIO-RAD, Hercules, California, USA) and the built-in Opticon Monitor Software version 3.1 (Bio-Rad Laboratories, Hercules, CA, USA). The real-time PCR efficiencies of the bacterial and archaeal 16S rRNA gene assays were 90–95% and 92–96%, respectively, with r^2 values of ≥ 0.99 in all assays. The following thermal cycling parameters were used to amplify all genes: 15 min at 95°C, 40 cycles of 20 s at 95°C, 20 s at 55°C and 20 s at 72°C, and readings were taken between each cycle. Standard curves were used to determine the relationship between a known copy number of genes and the cycle threshold values, as described previously (Park *et al*., 2008). The specificity of real-time PCR was confirmed by analysing the melting curves and checking the size of the PCR products using gel electrophoresis. The bacterial and archaeal 16S rRNA genes were amplified using the archaea-specific primer set 519F-727R and the bacteria-specific primer set Bac518F-Bac786R, as described by Park and colleagues (2010).

PCR amplification of the prokaryotic 16S rRNA genes and pyrosequencing

The V1 and V3 hyper-variable regions of the 16S rRNA genes were amplified from the bacterial and archaeal 16S rRNA genes respectively. PCR amplifications of the 16S rRNA genes were performed using the 8F (5′-AGAGTTTG ATCCTGGCTCAG-3′) and 338R (5′-TGCTGCCTCCCGT AGGAGT-3′) primer set (Ovreas *et al*., 1997; Vickerman *et al*., 2007), and the arch514F (5′-GGTGBCAGCCGCCG CGRKAAHACC-3′) (in this study) and 758R (5′-GGACTA CCCGGGTATCTAATCC-3′) (Juck *et al*., 2000) primer set for *Bacteria* and *Archaea* respectively. PCR was performed using 25 μl 2× PCR Master Mix Solution (Intron, Seongnam, Korea), 1 μM of each primer (final concentration) and ca 10 ng of genomic DNA as the template, and water was added to make a final volume of 50 μl. The following PCR cycles were used: initial denaturation at 94°C for 5 min, 30 cycles at 94°C for 50 s, 55°C for 30 s and 72°C for 50 s, and a final extension at 72°C for 6 min. The amplification products from each sample were purified using a PCR purification Kit (Cosmo4, Seoul, Korea). The DNA was quantified using a spectrophotometer (Nanodrop Technologies, Rockland, DE, USA) and was then mixed in equivalent proportions. Sequencing was performed using an FLX Titanium Genome Sequencer (454 Life Sciences, Branford, CT, USA). Multiplex identifiers (Roche, Basel, Schweiz), an adaptor and a short four-nucleotide sequence (TCAG), which were recognized by the system software and the priming sequences, were used to label the end fragments of the amplification products obtained from the samples by a sequencing provider (NICEM, Seoul, Korea), according to the manufacturer's instructions.

Pyrosequencing data analysis

Based on assessment of the pyrosequencing error rates, the raw reads were processed to remove low-quality sequences according to the following criteria: (i) short reads (< 150 bp) and (ii) reads longer than the expected PCR product size (Huse *et al*., 2007). Taxonomic classification was performed as follows. CLUSTALW was used to align the DNA sequences. After alignment, the sequences were trimmed to remove non-overlapping ends. Bacterial sequence reads were compared with a reference database of known 16S rRNA genes, which were obtained from the Ribosomal Database Project (RDP) databases and were assigned taxonomically based on the RDP classifiers (Cole *et al*., 2007). However, this automated RDP classifier pipeline was not sufficiently precise to classify archaea. Thus, to enhance the precision of the taxonomical affiliation, representative sequences of uncultured archaeal groups were obtained from the NCBI GenBank database and used to construct a phylogenetic tree with the archaeal sequence reads. The phylogenetic trees were constructed according to the neighbour-joining method with Kimura's two-parameter model (Kimura, 1980) for correction using MEGA 3 (Kumar *et al*., 2004). The DNA distance matrices were calculated using the DNADIST program in PHYLIP (version 3.68).

The formation of OTUs was performed using the Taxonomy-Based Clustering (TBC) program (Lee *et al*., 2012). The OTU formation results were transformed to the file format required by the program MOTHUR and were used to calculate rarefaction curves, Shannon index, Good's coverage, ACE and the Chao1 nonparametric richness estimators. Before comparing the diversity calculations, the library size was normalized to the size of the smallest library. The calculations were repeated 100 times using random subsamples of sequences. A 3% dissimilarity level between sequences was used to calculate the diversity estimators.

The microbial community structures in different samples were compared using Fast-UniFrac (Lozupone *et al*., 2006) based on the phylogenetic relationships among representative reads (OTUs) from different samples. The input file was parsed from the OTU output obtained by the program TBC at a similarity of 97%. The number of reads and the sequence ID of a representative read from each OTU were used as the input for the weighted Fast-UniFrac analysis. The trees were constructed using the FastTree program with suitable outgroups (Price *et al*., 2010). The Fast-UniFrac sample clustering results were used to compare the hierarchical relationships among the samples (Lozupone *et al*., 2006). To determine the correlations between the community structures and environmental parameters, detrended correspondence analysis and CCA were conducted using CANOCO version 4.5 (Biometrics, Wageningen, the Netherlands) and XLSTAT (Addinsoft, New York, NY, USA). Pearson's correlation coefficients and *P*-values were calculated using the R STATS package (version 2.15.3) to analyse the linear correlation between different parameters.

Acknowledgements

We thank the crew of the RV *Araon* for their assistance. This study was funded by the Korea Polar Research Institute, Korea (PP13020) and the Mid-Career Researcher Program (NRF-2013R1A2A2A05006754) through the National Research Foundation of Korea. We thank Francisco-Rodriguez-Valera for comments on the manuscript.

References

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

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

Alonso-Saez, L., Galand, P.E., Casamayor, E.O., Pedros-Alio, C., and Bertilsson, S. (2010) High bicarbonate assimilation in the dark by Arctic bacteria. *ISME J* **4:** 1581–1590.

Alonso-Sáez, L., Andersson, A., Heinrich, F., and Bertilsson, S. (2011) High archaeal diversity in Antarctic circumpolar deep waters. *Environ Microbiol Rep* **3:** 689–697.

Arrigo, K.R., and van Dijken, G.L. (2003) Phytoplankton dynamics within 37 Antarctic coastal polynya systems. *J Geophys Res* **108:** 3271. DOI: 10.1029/2002JC001739.

Ayton, J., Aislabie, J., Barker, G.M., Saul, D., and Turner, S. (2010) Crenarchaeota affiliated with group 1.1b are prevalent in coastal mineral soils of the Ross Sea region of Antarctica. *Environ Microbiol* **12:** 689–703.

Barber, D.G., and Massom, R.A. (2007) Chapter 1: the role of sea ice in Arctic and Antarctic polynyas. In *Elsevier Oceanography Series*. Smith, W.O., and Barber, D.G. (eds). Amsterdam, the Netherlands: Elsevier, pp. 1–54.

Bowman, J.P., and McCuaig, R.D. (2003) Biodiversity, community structural shifts, and biogeography of prokaryotes within Antarctic continental shelf sediment. *Appl Environ Microbiol* **69:** 2463–2483.

Bowman, J.S., Rasmussen, S., Blom, N., Deming, J.W., Rysgaard, S., and Sicheritz-Ponten, T. (2012) Microbial community structure of Arctic multiyear sea ice and surface seawater by 454 sequencing of the 16S RNA gene. *ISME J* **6:** 11–20.

Church, M.J., DeLong, E.F., Ducklow, H.W., Karner, M.B., Preston, C.M., and Karl, D.M. (2003) Abundance and distribution of planktonic *Archaea* and *Bacteria* in the waters west of the Antarctic Peninsula. *Limnol Oceanogr* **48:** 1893–1920.

Claire Horner-Devine, M., Leibold, M.A., Smith, V.H., and Bohannan, B.J.M. (2003) Bacterial diversity patterns along a gradient of primary productivity. *Ecol Lett* **6:** 613–622.

Cole, J.R., Chai, B., Farris, R.J., Wang, Q., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., *et al*. (2007) The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. *Nucleic Acids Res* **35:** D169–D172.

Comeau, A.M., Li, W.K., Tremblay, J.E., Carmack, E.C., and Lovejoy, C. (2011) Arctic Ocean microbial community structure before and after the 2007 record sea ice minimum. *PLoS ONE* **6:** e27492.

Cottrell, M.T., and Kirchman, D.L. (2009) Photoheterotrophic microbes in the Arctic Ocean in summer wnd Winter. *Appl Environ Microbiol* **75:** 4958–4966.

Díez, B., Bergman, B., Pedrós-Alió, C., Antó, M., and Snoeijs, P. (2012) High cyanobacterial nifH gene diversity in Arctic seawater and sea ice brine. *Environ Microbiol Rep* **4:** 360–366.

French, E., Kozlowski, J.A., Mukherjee, M., Bullerjahn, G., and Bollmann, A. (2012) Enrichment and characterization of three ammonia-oxidizing *Archaea* from freshwater environments. *Appl Environ Microbiol* **78:** 5773–5780.

Galand, P.E., Casamayor, E.O., Kirchman, D.L., Potvin, M., and Lovejoy, C. (2009) Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing. *ISME J* **3:** 860–869.

Galand, P.E., Gutiérrez-Provecho, C., Massana, R., Gasol, J.M., and Casamayor, E.O. (2010) Inter-annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea (Blanes Bay Microbial Observatory). *Limnol Oceanogr* **55:** 2117–2125.

Gentile, G., Giuliano, L., D'Auria, G., Smedile, F., Azzaro, M., De Domenico, M., and Yakimov, M.M. (2006) Study of bacterial communities in Antarctic coastal waters by a combination of 16S rRNA and 16S rDNA sequencing. *Environ Microbiol* **8:** 2150–2161.

Ghiglione, J.F., and Murray, A.E. (2011) Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton. *Environ Microbiol* **14:** 617–629.

Ghiglione, J.-F., Galand, P.E., Pommier, T., Pedrós-Alió, C., Maas, E.W., Bakker, K., *et al*. (2012) Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc Natl Acad Sci USA* **109:** 17633–17638.

Giovannoni, S.J., Bibbs, L., Cho, J.C., Stapels, M.D., Desiderio, R., Vergin, K.L., *et al*. (2005) Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* **438:** 82–85.

Herfort, L., Schouten, S., Abbas, B., Veldhuis, M.J., Coolen, M.J., Wuchter, C., *et al*. (2007) Variations in spatial and temporal distribution of *Archaea* in the North Sea in relation to environmental variables. *FEMS Microbiol Ecol* **62:** 242–257.

Huse, S.M., Huber, J.A., Morrison, H.G., Sogin, M.L., and Welch, D.M. (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* **8:** R143.

Iverson, V., Morris, R.M., Frazar, C.D., Berthiaume, C.T., Morales, R.L., and Armbrust, E.V. (2012) Untangling genomes from metagenomes: revealing an uncultured class of marine *Euryarchaeota*. *Science* **335:** 587–590.

Jenkins, A., Dutrieux, P., Jacobs, S.S., McPhail, S.D., Perrett, J.R., Webb, A.T., and White, D. (2010) Observations beneath Pine Island Glacier in West Antarctica and implications for its retreat. *Nat Geosci* **3:** 468–472.

Johnston, A.W.B., Li, Y., and Ogilvie, L. (2005) Metagenomic marine nitrogen fixation feast or famine? *Trends Microbiol* **13:** 416–420.

Juck, D., Charles, T., Whyte, L.G., and Greer, C.W. (2000) Polyphasic microbial community analysis of petroleum hydrocarbon-contaminated soils from two northern Canadian communities. *FEMS Microbiol Ecol* **33:** 241–249.

Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16:** 111–120.

Könneke, M., Bernhard, A.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B., and Stahl, D.A. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437:** 543–546.

Kumar, S., Tamura, K., and Nei, M. (2004) MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* **5:** 150–163.

Lee, J.-H., Yi, H., Jeon, Y.-S., Won, S., and Chun, J. (2012) TBC: A clustering algorithm based on prokaryotic taxonomy. *J Microbiol* **50:** 181–185.

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

Lozupone, C., Hamady, M., and Knight, R. (2006) UniFrac – an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* **7:** 371.

Marshall, K.T., and Morris, R.M. (2013) Isolation of an aerobic sulfur oxidizer from the SUP05/Arctic96BD-19 clade. *ISME J* **7:** 452–455.

Martens-Habbena, W., Berube, P.M., Urakawa, H., Torre, J.R., and Stahl, D.A. (2009) Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* **461:** 976–979.

Massana, R., Murray, A.E., Preston, C.M., and DeLong, E.F. (1997) Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel. *Appl Environ Microbiol* **63:** 50–56.

Miller, L.A., and DiTullio, G.R. (2007) Chapter 5 gas fluxes and dynamics in polynyas. In *Elsevier Oceanography Series*. Smith, W.O., and Barber, D.G. (eds). Amsterdam, the Netherlands: Elsevier, pp. 163–191.

Mincer, T.J., Church, M.J., Taylor, L.T., Preston, C., Karl, D.M., and DeLong, E.F. (2007) Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. *Environ Microbiol* **9:** 1162–1175.

Morris, R., Vergin, K.L., Cho, J.C., Rappe, M.S., Carlson, C.A., and Giovannoni, S.J. (2005) Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-series Study site. *Limnol Oceanogr* **50:** 1687–1696.

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

Murray, A.E., and Grzymski, J.J. (2007) Diversity and genomics of Antarctic marine micro-organisms. *Philos Trans R Soc B Biol Sci* **362:** 2259–2271.

Ovreas, L., Forney, L., Daae, F.L., and Torsvik, V. (1997) Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. *Appl Environ Microbiol* **63:** 3367–3373.

Park, B.J., Park, S.J., Yoon, D.N., Schouten, S., Sinninghe Damste, J.S., and Rhee, S.K. (2010) Cultivation of autotrophic ammonia-oxidizing archaea from marine sediments in coculture with sulfur-oxidizing bacteria. *Appl Environ Microbiol* **76:** 7575–7587.

Park, S.-J., Park, B.-J., and Rhee, S.-K. (2008) Comparative analysis of archaeal 16S rRNA and amoA genes to estimate the abundance and diversity of ammonia-oxidizing archaea in marine sediments. *Extremophiles* **12:** 605–615.

Paulmier, A., and Ruiz-Pino, D. (2009) Oxygen minimum zones (OMZs) in the modern ocean. *Prog Oceanogr* **80:** 113–128.

Piquet, A.M., Bolhuis, H., Meredith, M.P., and Buma, A.G. (2011) Shifts in coastal Antarctic marine microbial communities during and after melt water-related surface stratification. *FEMS Microbiol Ecol* **76:** 413–427.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5:** e9490.

Qian, P.Y., Wang, Y., Lee, O.O., Lau, S.C., Yang, J., Lafi, F.F., *et al*. (2011) Vertical stratification of microbial communities in the Red Sea revealed by 16S rDNA pyrosequencing. *ISME J* **5:** 507–518.

Siegenthaler, U., and Wenk, T. (1984) Rapid atmospheric CO_2 variations and ocean circulation. *Nature* **308:** 624–626.

Smith, W.O., Jr, and Gordon, L.I. (1997) Hyperproductivity of the Ross Sea (Antarctica) polynya during austral spring. *Geophys Res Lett* **24:** 233–236.

Stingl, U., Tripp, H.J., and Giovannoni, S.J. (2007) Improvements of high-throughput culturing yielded novel SAR11

strains and other abundant marine bacteria from the Oregon coast and the Bermuda Atlantic Time Series study site. *ISME J* **1:** 361–371.

Stoecker, D.K., Putt, M., and Moisan, T. (1995) Nano- and microplankton dynamics during the spring Phaeocystis sp. bloom in McMurdo Sound, Antarctica. *J Mar Biol Assoc UK* **75:** 815–832.

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.

Taton, A., Grubisic, S., Brambilla, E., De Wit, R., and Wilmotte, A. (2003) Cyanobacterial diversity in natural and artificial microbial mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a morphological and molecular approach. *Appl Environ Microbiol* **69:** 5157–5169.

Vickerman, M.M., Brossard, K.A., Funk, D.B., Jesionowski, A.M., and Gill, S.R. (2007) Phylogenetic analysis of bacterial and archaeal species in symptomatic and asymptomatic endodontic infections. *J Med Microbiol* **56:** 110–118.

Weisse, T., Tande, K., Verity, P., Hansen, F., and Gieskes, W. (1994) The trophic significance of Phaeocystis blooms. *J Mar Syst* **5:** 67–79.

Supporting information

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

Fig. S1. Vertical profiles of the physicochemical characteristics (fluorescence, temperature, salinity and dissolved oxygen) throughout the water mass of the polynya, i.e. ST17 and ST 26 (sea ice), and ST1 (ice-free open ocean). The dashed lines indicate the depth of the water sample used in the analysis.

Fig. S2. Rarefaction curves for (A) *Bacteria* and (B) *Archaea* for the 97% similarity cut-off value.

Fig. S3. PCA results showing the relatedness of (A) bacterial and (B) archaeal communities.

Fig. S4. Taxonomic composition of reads within the *Proteobacteria* lineage.

Fig. S5. Relative abundance of subgroups of the SAR11 clade.

Fig. S6. Canonical correspondence analysis ordination showing based on the 18 top dominant bacterial OTUs based on 97% similarity cut-off and environmental variables (arrows). Un, unclassified; AlphaP, *Alphaproteobacteria*; GammaP, *Gammaproteobacteria*; DeltaP, *Deltaproteobacteria*.

Table S1. Locations, prokaryote cell abundance and nutrient characteristics for the sampling sites.

Table S2. Phytoplankton community composition at surface chlorophyll maximum of sampling sites. Units represent cells ml^{-1} ($\mu g\ C\ l^{-1}$).

Table S3. Results of canonical correspondence analysis (CCA) of bacterial community structure as determined from the 18 top dominant bacterial OTUs based on 97% similarity cut-off. The highest intra-set correlations are highlighted in bold.