Antarctic Science 25(6), 741–751 (2013) © Antarctic Science Ltd 2013 doi:10.1017/S0954102013000229

Regional variability in eukaryotic protist communities in the Amundsen Sea

CHRISTIAN WOLF[1], STEPHAN FRICKENHAUS[1], ESTELLE S. KILIAS[1], ILKA PEEKEN[1,2] and KATJA METFIES[1]

[1]*Alfred Wegener Institute for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany*
[2]*MARUM - Centre for Marine Environmental Sciences, University of Bremen, Leobener Straße, 28359 Bremen, Germany*
Christian.Wolf@awi.de

Abstract: We determined the composition and structure of late summer eukaryotic protist assemblages along a west–east transect in the Amundsen Sea. We used state-of-the-art molecular approaches, such as automated ribosomal intergenic spacer analysis (ARISA) and 454-pyrosequencing, combined with pigment measurements via high performance liquid chromatography (HPLC) to study the protist assemblage. We found characteristic offshore and inshore communities. In general, total chlorophyll *a* and microeukaryotic contribution were higher in inshore samples. Diatoms were the dominant group across the entire area, of which *Eucampia* sp. and *Pseudo-nitzschia* sp. were dominant inshore and *Chaetoceros* sp. was dominant offshore. At the most eastern station, the assemblage was dominated by *Phaeocystis* sp. Under the ice, ciliates showed their highest and haptophytes their lowest abundance. This study delivers a taxon detailed overview of the eukaryotic protist composition in the Amundsen Sea during the summer 2010.

Received 3 September 2012, accepted 15 February 2013, first published online 16 April 2013

Key words: ARISA, HPLC, microbial diversity, next-generation-sequencing, phytoplankton

Introduction

The Pacific sector of the Southern Ocean, and especially the Amundsen Sea, are the least studied oceanic regions in the world (Griffiths 2010). Severe ice conditions year-round and the geographic remoteness make sampling in this area very difficult. The biodiversity of the Amundsen Sea, especially of the coastal and shelf areas, is almost unknown (Kaiser *et al*. 2009). Recently, scientists began to highlight the diversity and distribution of isopods and phytoplankton in this isolated region (Kaiser *et al*. 2009, Fragoso & Smith 2012). Gravalosa *et al*. (2008) concentrated on the distribution of coccolithophores and showed that their dispersion is restricted north of the Polar Front. Fragoso & Smith (2012) focused their study areas near the coast and delivered an overview of the phytoplankton assemblages in this area. They revealed diatom dominated assemblages in offshore areas of the Amundsen Sea. However, they used pigment based and microscopic analysis and thereby, the taxonomical resolution was not very detailed. So far, no comprehensive survey of the whole eukaryotic protist spectrum in the Amundsen Sea exists.

In the course of the controversially conducted debate about the "everything is everywhere" hypothesis (Lachance 2004), many studies focused on the biogeography of protists (Finlay 2002, Finlay & Fenchel 2004). Our recent study, focusing on the distribution of eukaryotic protists along a transect from New Zealand to the coast of Antarctica, revealed distinct biogeographical patterns, defined by the oceanic fronts (Wolf *et al.* unpublished). These patterns were driven by strong environmental gradients and included different large-scale water masses. To complement our knowledge about the biogeography of protists in the Pacific sector of the Southern Ocean, their distribution has to be highlighted on a smaller, more regional scale. Narrow environmental differences within a large-scale water mass have to be investigated.

Most investigations of eukaryotic protist composition and distribution in the Southern Ocean mainly used traditional microscopic and pigment extraction based methods (Ishikawa *et al*. 2002, Wright *et al*. 2009). However, microscopic surveys have difficulties in identifying small cells and pigment analysis only targets autotrophic cells. Here, molecular tools are advantageous. Few investigations in the Southern Ocean used molecular approaches, such as denaturing gradient gel electrophoresis (DGGE) (Gast *et al*. 2004) or 18S rRNA gene cloning and sequencing (Lopez-Garcia *et al.* 2001). The automated ribosomal intergenic spacer analysis (ARISA) approach provides a quick overview of the diversity and facilitates the comparison of different samples. It is well established for investigations of prokaryotic diversity (Smith *et al*. 2010), and we successfully implemented it for the analysis of eukaryotic phytoplankton diversity. The newly emerging 454-pyrosequencing approach (e.g. of the V4 region of the 18S rRNA gene) allows assessing microbial communities with high-resolution, based on sufficient deep taxon sampling (Margulies *et al.* 2005, Stoeck *et al*. 2010), regardless of cell size and nutrition.

The objective of this study is to determine the composition of late summer eukaryotic protist assemblages in the Amundsen Sea, south of the southern boundary of the Antarctic

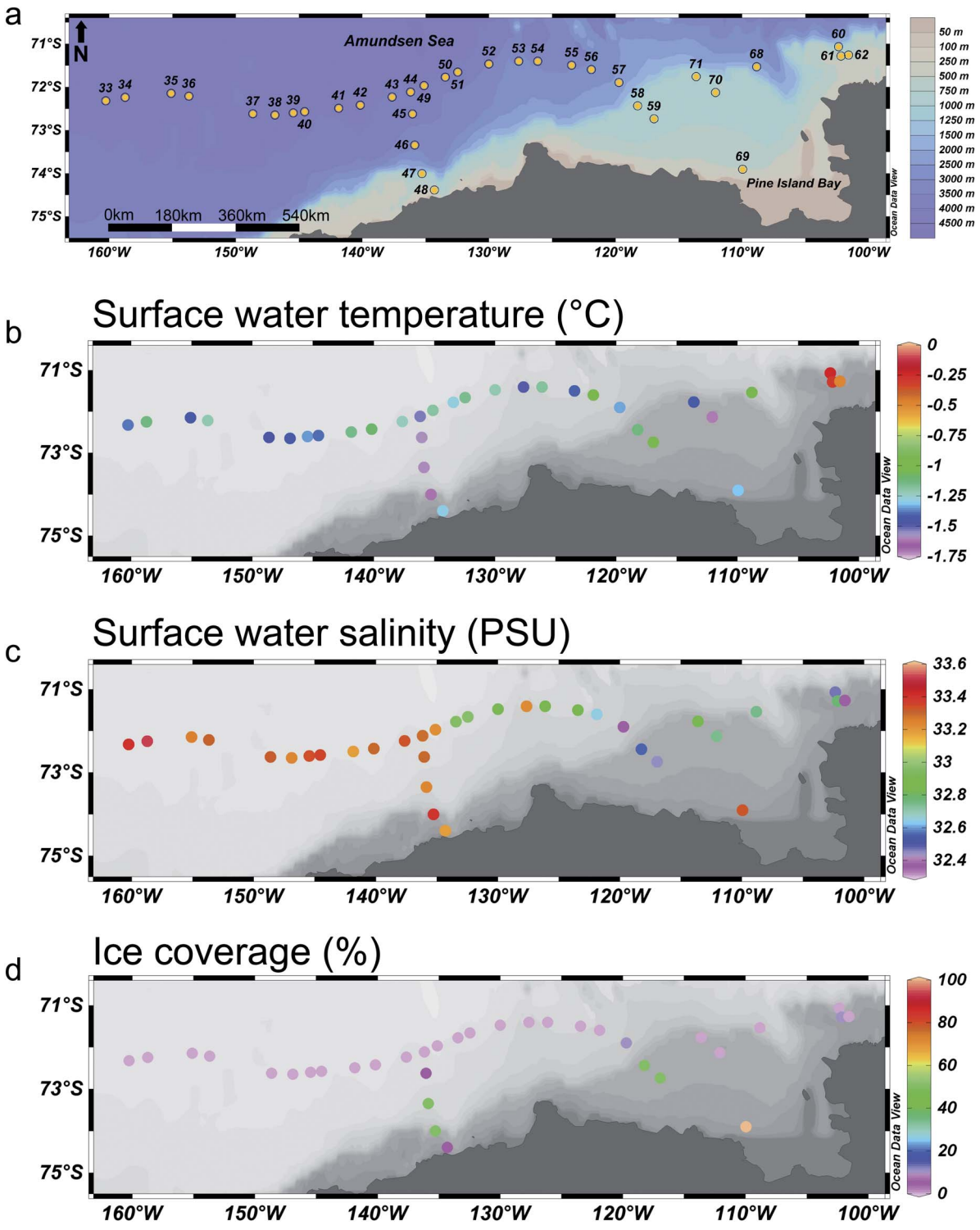

Fig. 1. Study area and environmental conditions. **a.** Location of surface water samples and water depth. **b.** Surface water temperature. **c.** Surface water salinity. **d.** Ice coverage.

Circumpolar Current. We used state of the art molecular approaches, such as ARISA and 454-pyrosequencing, and high-performance liquid chromatography (HPLC). Furthermore, we want to assess the impact of different environmental conditions on the biogeography of protists, within this oceanic region. The pigment and ARISA analysis provide an overview of differences in structure and diversity of the whole investigated area and the 454-pyrosequencing of selected samples gives more detailed information about the species composition, dominant representatives and the

https://doi.org/10.1017/S0954102013000229 Published online by Cambridge University Press

distribution of the rare biosphere (phylotypes with an abundance < 1% of total sequences) in the observed area.

Materials and methods

Location and sampling

A total of 34 surface-water samples were taken on a regular basis (*c.* every 40 km) during the RV *Polarstern* cruise ANT XXVI/3 between 12 February and 22 March 2010 in the Amundsen Sea (Fig. 1a) along a west–east transect from the eastern Ross Sea to the western Bellingshausen Sea (within 71.06–74.39°S and 160.27–101.58°W). All surface water samples were collected using the ship pumping system (membrane pump), located at the bow at 8 m depth below the surface. For the determination of pigments, 4 l samples were immediately filtered onto 25 mm Whatman GF/F filters and stored at -80°C until further analysis in the laboratory. For molecular analysis, 2 l samples were immediately fractionated, by filtering them on Isopore Membrane Filters (Millipore, USA) with a pore size of 10 μm, 3 μm and 0.2 μm. Filters were stored at -80°C until analysis in the laboratory.

Pigment analysis (HPLC)

Samples were measured using a Waters HPLC-system, equipped with an auto sampler (717 plus), pump (600), PDA (2996), a fluorescence detector (2475) and EMPOWER software. For analytical preparation, 50 μl internal standard (canthaxanthin) and 1.5 ml acetone were added to each filter sample and then homogenized for 20 sec in a Precellys® tissue homogenizer. After centrifugation, the supernatant liquid was filtered through a 0.2 μm PTFE filter (Rotilabo) and placed in Safe-Lock Tubes (Eppendorf, Germany). An aliquot (100 μl) was transferred to the auto sampler (4°C). Just prior to analysis, the sample was premixed with 1 M ammonium acetate solution in the ratio 1:1 (v/v) in the auto sampler and injected onto the HPLC-system. The pigments were analysed by reverse-phase HPLC, using a VARIAN Microsorb-MV3 C8 column (4.6 x 100 mm) and HPLC-grade solvents (Merck, Germany). Solvent A consisted of 70% methanol and 30% 1 M ammonium acetate and solvent B contained 100% methanol. The gradient was modified after Barlow *et al.* (1997). Eluting pigments were detected by absorbance (440 nm) and fluorescence (Ex: 410 nm, Em: > 600 nm). Pigments were identified by comparing their retention times with those of pure standards and algal extracts. Additional confirmation for each pigment was done by comparing their absorbance spectra between 390 and 750 nm with the library of the standards. Pigment concentrations were quantified based on peak areas of external standards, which were spectrophotometrically calibrated using extinction coefficients published by Bidigare (1991) and Jeffrey *et al.* (1997). For correction of experimental losses and volume changes, the concentrations of the pigments were normalized to the internal standard canthaxanthin. Phytoplankton group composition was calculated applying the CHEMTAX program and input ratios of Mackey *et al.* (1996). To estimate the various size classes of the phytoplankton, the following groups were combined: prasinophytes and pelagophytes for picoplankton (< 2 μm), haptophytes and cryptophytes for nanoplankton (2–20 μm), and dinoflagellates and diatoms for microplankton (20–200 μm). Their respective contribution to total biomass is based on their CHEMTAX derived chlorophyll *a* (chl *a*) concentration.

DNA extraction

The DNA was extracted with the E.Z.N.A.TM SP Plant DNA Kit (Omega Bio-Tek, USA). At the beginning, the filters were incubated with lysis buffer. All further steps

Table I. Summary of recovered 454-pyrosequencing reads, quality filtering and number of OTUs (operational taxonomic units). Samples are arranged from west–east.

	Sample						
	41	47	51	57	70	69	62
Total 454-reads	29 807	24 109	34 767	49 355	33 020	63 836	43 222
Average length (bp)	332	333	345	360	383	370	393
Acceptable length*	21 339	17 332	26 562	36 695	26 687	49 221	36 520
Quality filtering:							
More than one N	142	67	135	309	228	406	386
Chimeras	754	504	484	1464	1278	2096	1265
Incorrect forward primer	302	204	280	185	159	327	393
Singletons	1007	853	1377	3223	1730	3583	3025
Non-target organisms	1131	1485	2175	2273	5913	14619	3462
Total filtered reads	18 003	14 219	22 111	29 241	17 379	28 190	27 989
OTUs (97% identity)	927	893	1161	1593	1219	1687	1554
Abundant OTUs**	12	15	11	11	8	13	12
Rare OTUs**	915	878	1150	1582	1211	1674	1542

*reads with a minimum length of 300 bp and a maximum length of 670 bp.
**abundant OTU = number of reads ≥ 1% of total reads, otherwise it is rare.

https://doi.org/10.1017/S0954102013000229 Published online by Cambridge University Press

were performed as described in the manufacturer's instructions. At the end, the DNA was eluted in 60 μl of elution buffer and the extracts were stored at -20°C until further analysis. DNA concentration was measured with a NanoDrop 1000 (Thermo Fisher Scientific, USA) (average DNA concentration: 23 ng μl^{-1}).

PCR amplification, ARISA

An equal volume of extracted DNA of each size fraction (> 10 μm, 3–10 μm and 0.2–3 μm) from each sample was pooled. The ITS1 (internal transcribed spacer) region was amplified in triplicates using the primer-set 1528F (5'-GTA GGT GAA CCT GCA GAA GGA TCA-3') (modified after Medlin *et al*. (1988)) and ITS2 (5'-GCT GCG TTC TTC ATC GAT GC-3') (White *et al*. 1990). The 1528F primer was labelled at the 5'-end with the dye 6-FAM (6-carboxyfluorescein). The PCR (polymerase chain reaction) mixtures contained 1 μl of DNA extract, 1 x HotMaster Taq Buffer containing 2.5 mM Mg^{2+} (5 Prime, USA), 0.8 mM dNTP-mix (Eppendorf, Germany), 0.2 μM of each Primer and 0.4 U of HotMaster Taq DNA polymerase (5 Prime, USA) in a final volume of 20 μl. Reactions were carried out in a Mastercycler (Eppendorf, Germany) under the following conditions: an initial denaturation at 94°C for 3 min, 35 cycles of denaturation at 94°C for 45 sec, annealing at 55°C for 1 min and extension at 72°C for 3 min, and a final extension at 72°C for 10 min. PCR fragments were separated by capillary electrophoresis on an ABI Prism 310 Genetic Analyser (Applied Biosystems, USA).

PCR amplification, 454-pyrosequencing

Seven samples were sequenced (Table I). For each fraction of a sample, we amplified *c.* 670 base pair (bp) fragments of the 18S rRNA gene, containing the highly variable V4-region, using the primer-set 528F (5'-GCG GTA ATT CCA GCT CCA A-3') and 1055R (5'-ACG GCC ATG CAC CAC CAC CCA T-3') (modified after Elwood *et al*. (1985)). The PCR mixtures were composed as described previously for ARISA. Reaction conditions were as follows: an initial denaturation at 94°C for 3 min, 30 cycles of denaturation at 94°C for 45 sec, annealing at 59°C for 1 min and extension at 72°C for 3 min, and a final extension at 72°C for 10 min. An equal volume of PCR reaction of each size fraction from each sample was pooled and purified with the MinElute PCR purification kit (Qiagen, Germany) following the manufacturer's instructions. Pyrosequencing was performed on a Genome Sequencer FLX system (Roche, Germany) by GATC Biotech AG (Germany).

Data analysis, ARISA

Electropherograms were analysed using the GeneMapper Software v4.0 (Applied Biosystems, USA). Peaks with a size smaller than 50 bp (corresponding to primer and primer dimer peaks) were removed from the dataset. To remove the background noise and to get sample-by-binned-OTU (operational taxonomic unit) tables, the data were binned using the binning scripts, according to Ramette (2009), for R (R Development Core Team 2008). The resulting sample-by-binned-OTU tables were transformed into presence/absence matrices and the distances between the samples were calculated, using the Jaccard index implemented in the R package vegan (Oksanen *et al*. 2011), which was also used in the following steps. MetaMDS (maximum random starts of 300) plots were computed. Clusters were determined using the hclust function in R. To test, whether the resulting clusters differ significantly, an ANOSIM was performed. A Euclidean distance matrix with the normalized environmental parameters was calculated. The correlation between the ARISA distance matrix and the environmental distance matrix was tested with a Mantel test (10 000 permutations), implemented in the R package ade4 (Dray & Dufour 2007). A principal component analysis (PCA) with the environmental parameters and the HPLC size fractions was performed (R package ade4).

Data analysis, 454-pyrosequencing

Raw sequence reads were processed to obtain high quality reads. The forward primer 528F, used for the sequencing, attaches *c.* 25 bp upstream of the V4 region, which has in general a length of *c.* 230 bp (Nickrent & Sargent 1991). Reads with a length under 300 bp were excluded from further analysis to assure inclusion of the whole hyper variable V4 region in the analysis and to get rid of short reads. Unusually long reads that were greater than the expected amplicon size (> 670 bp) and reads with more than one uncertain base (N) were removed. Remaining reads were checked for chimeric sequences with the software UCHIME 4.2.40 (Edgar *et al*. 2011) and all reads considered being chimeric were excluded from further analysis. The high quality reads of all samples were clustered into OTUs at the 97% similarity level using the software Lasergene 10 (DNASTAR, USA). Subsequently, reads not starting with the forward primer were manually removed. Consensus sequences of each OTU were generated, which reduced the amount of sequences to operate with and attenuated the influence of sequencing errors and uncertain bases. The 97% similarity level has been shown to be the most suitable to reproduce original eukaryotic diversity (Behnke *et al*. 2011) and has the effect of bracing most of the sequencing errors (Kunin *et al*. 2010). Furthermore, known intragenomic SSU polymorphism levels can range to 2.9% in dinoflagellate species (Miranda *et al*. 2012). Operational taxonomic units comprised of only one sequence (singletons) were removed. The consensus sequences were aligned into a reference alignment obtained from SILVA (see below) using the software HMMER 2.3.2

https://doi.org/10.1017/S0954102013000229 Published online by Cambridge University Press

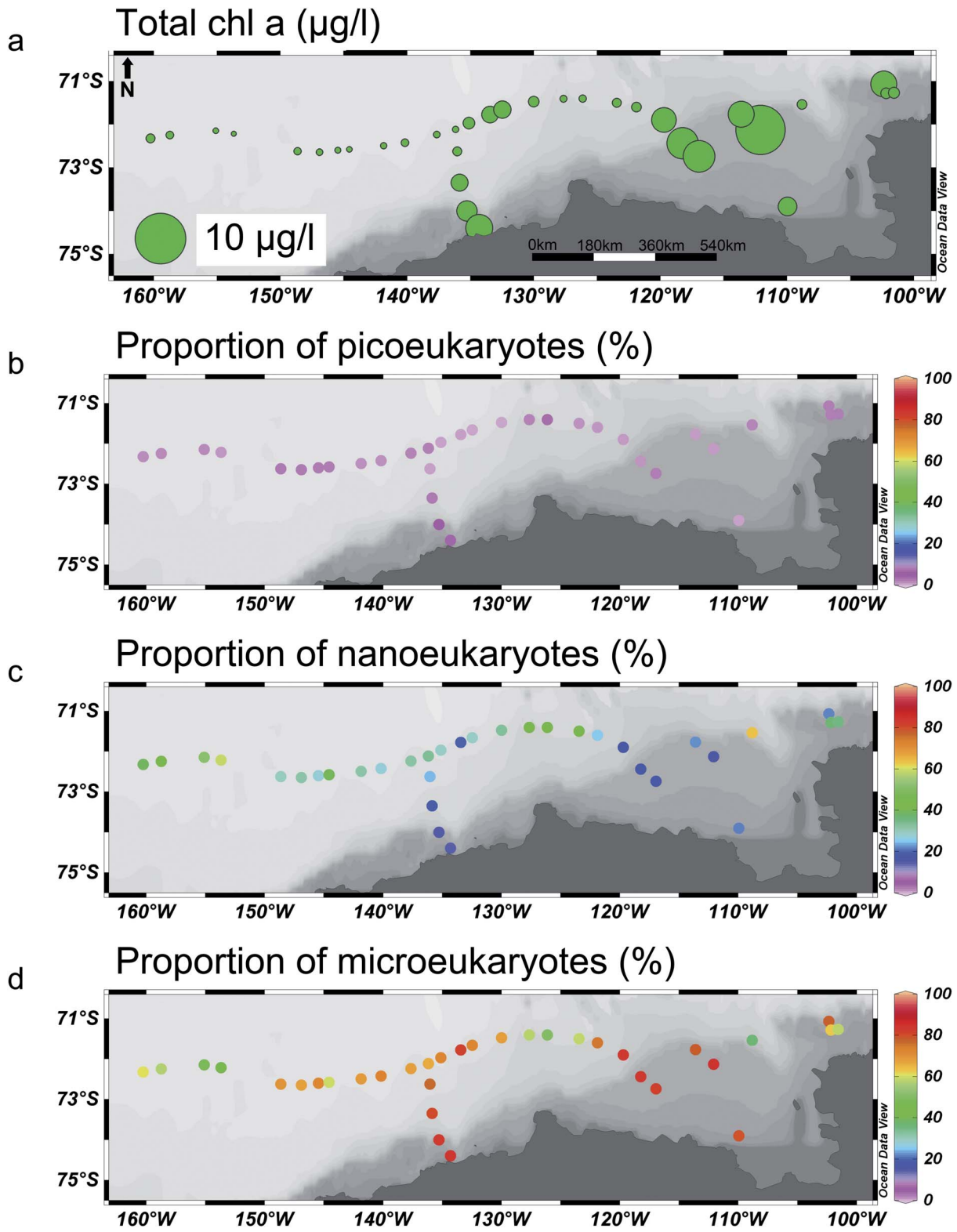

Fig. 2. Total chlorophyll *a* (chl *a*) concentration and size class distribution of total chl *a* based on CHEMTAX identification of the various algae classes. **a.** Total chl *a*. **b.** Proportion of picoeukaryotes. **c.** Proportion of nanoeukaryotes. **d.** Proportion of microeukaryotes.

(Eddy 2011). Subsequently, taxonomical affiliation was determined by placing the consensus sequences into a reference tree, containing about 1200 high quality sequences of Eukarya from the SILVA reference database (SSU Ref 108), using the software pplacer 1.0 (Matsen *et al*. 2010). The compiled reference database is available

https://doi.org/10.1017/S0954102013000229 Published online by Cambridge University Press

on request in ARB-format. OTUs assigned to fungi and metazoans were excluded from further analysis. Rarefaction curves were computed using the freeware program Analytic Rarefaction 1.3. The dataset generated in this study has been deposited at GenBank's Short Read Archive (SRA) under Accession No. SRA057133.

Results

Environmental conditions

The investigated area showed a very heterogeneous setting, in terms of water depth, surface temperature, surface salinity and ice coverage (Fig. 1a–d). Samples 33–46 and 49–57 were lying offshore, with water depths (Fig. 1a) from 1969 m (sample 57) to 4334 m (sample 34). Samples 47 and 48 (polynya) and 58–71 (Pine Island Bay) were inshore, with water depths from 398 m (sample 69) to 714 (sample 48). Sample 60 was lying over the continental slope and showed a greater depth (1447 m).

The surface water temperature ranged between -1.63°C (sample 47) and -0.24°C (sample 60) (Fig. 1b). Hence, the temperature only varied weakly, but showed significantly higher values at the most eastern sample sites (60–62), located at the transition to the Bellingshausen Sea.

Surface water salinity (Fig. 1c) showed values between 32.35 PSU (sample 57) and 33.51 PSU (sample 34). In general, salinity was higher in the western part of the transect and declined eastwards. Among the eastern samples, sample 69 showed a very high salinity (33.32 PSU).

Most samples were located near the ice-edge (Fig. 1d) with no ice. At samples 45–48, we crossed an ice field to reach a polynya, with a high spatial variability of the ice cover (5–50%). Samples 57 and 58 were taken in an ice field and showed an ice coverage of 10–50%. Sample 69 was obtained in a region with 100% ice cover.

Structure/diversity overview

We used a combination of HPLC and ARISA to assess the impact of different environmental conditions on the structure/diversity of the plankton assemblages in the sub-polar region.

High performance liquid chromatography

Total chl *a* concentrations (Fig. 2a) along the entire transect ranged between 0.11 $\mu g\,l^{-1}$ (sample 36) and 9.58 $\mu g\,l^{-1}$ (sample 70). In general, the highest chl *a* concentrations occurred in samples lying inshore. The chl *a* concentrations in these areas always exceeded 1 $\mu g\,l^{-1}$. However, the majority of samples (21) showed chl *a* concentrations lower than 0.5 $\mu g\,l^{-1}$. All these samples, except for sample 68, were taken offshore.

The contribution of the three size classes (picoeukaryotes (0.2–2 μm), nanoeukaryotes (2–20 μm) and microeukaryotes

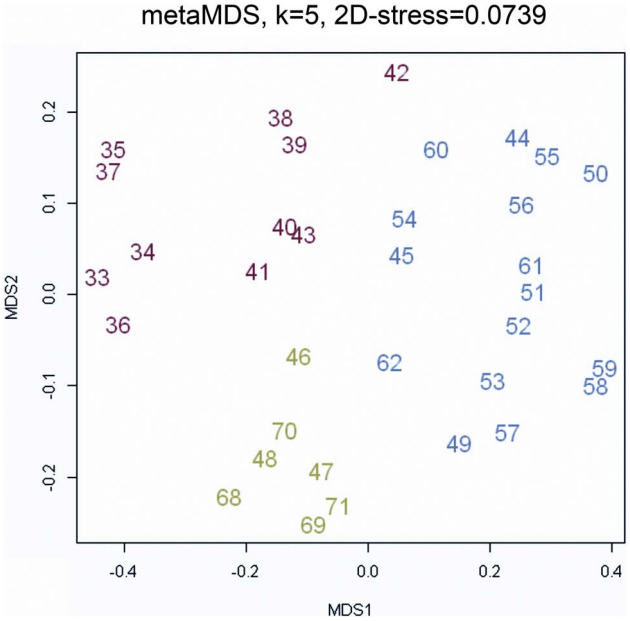

Fig. 3. MDS plot based on Jaccard distances of all 34 samples, gained via ARISA profiles. Colours of the samples indicate the three groups (red = group A, blue = group B, green = group C).

(> 20 μm)) to total chl *a* showed that picoeukaryotes (Fig. 2b) did not significantly contribute to phytoplankton biomass throughout the entire transect. The highest contribution of picoeukaryotes occurred in samples 47 (4%), 48 (3.2%), 54 (6.5%) and 62 (7.7%), mainly samples lying inshore. In all other samples, picoeukaryotes did not exceed a contribution of 1.9%. In general, nanoeukaryotes showed the highest contribution to total chl *a* in offshore samples (Fig. 2c). In these areas, they contributed up to 58.5% (sample 36). In offshore samples, they accounted for 33% ± 10% of chl *a* on average, whereas in inshore samples they only accounted for 25% ± 15% on average. Sample 68, with a contribution of nanoeukaryotes of 63%, presented as an outlier, just as for total chl *a* concentration. The lowest contribution of nanoeukaryotes (*c.* 14%) was shown by the two polynya samples (samples 47 and 48). Microeukaryotes were always the dominant size class (Fig. 2d), except for samples 33, 35, 36 and 68, where nanoeukaryotes were dominant. Microeukaryotes contributed 35.8–84.5% to total chl *a*, in which the highest values occurred generally in inshore samples (except sample 68). In these areas, they contributed 73% ± 15% on average, whereas in offshore ocean samples they accounted for 66% ± 11% on average.

Automated ribosomal intergenic spacer analysis

The fragment length analysis of the ITS1 region of all 34 surface water samples resulted in 97 different fragments with a length of 50–432 bp, of which 16 only occurred in one sample (unique fragments). The number of fragments

https://doi.org/10.1017/S0954102013000229 Published online by Cambridge University Press

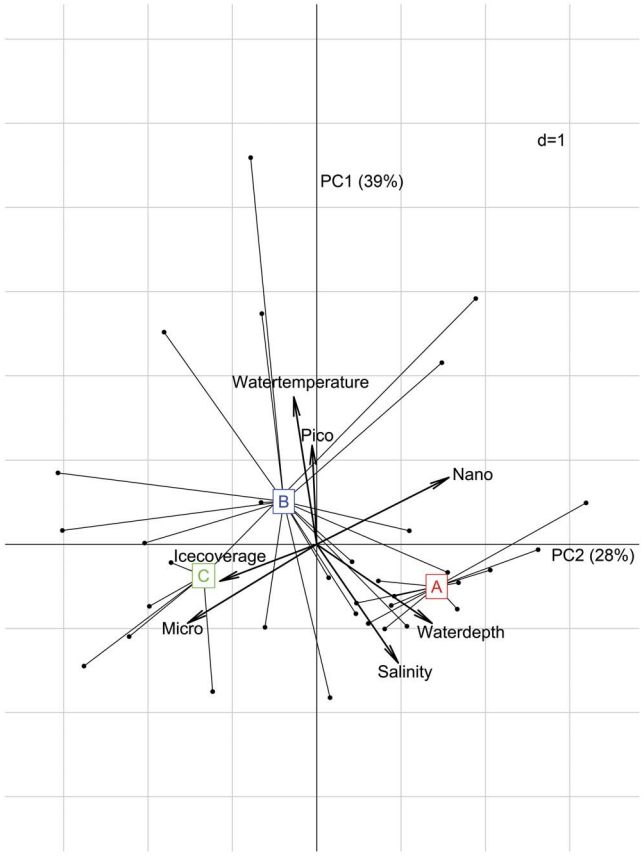

Fig. 4. Principal component analysis of environmental conditions and HPLC size fractions with plotted ARISA groups (A, B and C). Both axes are explaining 67% of the variance (PC1: 39%, PC2: 28%). Group A shows greater water depths, higher salinities, and a higher contribution of nanoeukaryotes. Group B is characterized by lower salinities and the highest picoeukaryotic contribution. Group C shows a higher ice coverage and a high contribution of microeukaryotes. d = axis scaling factor.

in each sample was 26 on average, ranging from nine (sample 50) to 49 (sample 68). The ordination analysis based on the ARISA profiles (Fig. 3) clustered the samples in three groups. Group A includes samples 33–42, group B contains samples 44 and 45 and 49–62, and group C includes samples 46–48 and 68–71. The three groups show significantly different ARISA profiles (ANOSIM, R = 0.637, $P = 0.001$). Groups A and B consist of offshore samples and represent the western and eastern part of the transect, respectively. Group C consists of samples collected inshore. Samples 58–62 fall into group B, although they were located over the shelf.

The ARISA profiles distances are significantly correlated with the distances of environmental conditions profiles (Mantel test, r = 0.142, $P = 0.023$). Figure 4 shows the PCA of the environmental conditions and the HPLC size fractions with the three ARISA groups plotted in. The two axes are explaining 67% of the total variance. Group C

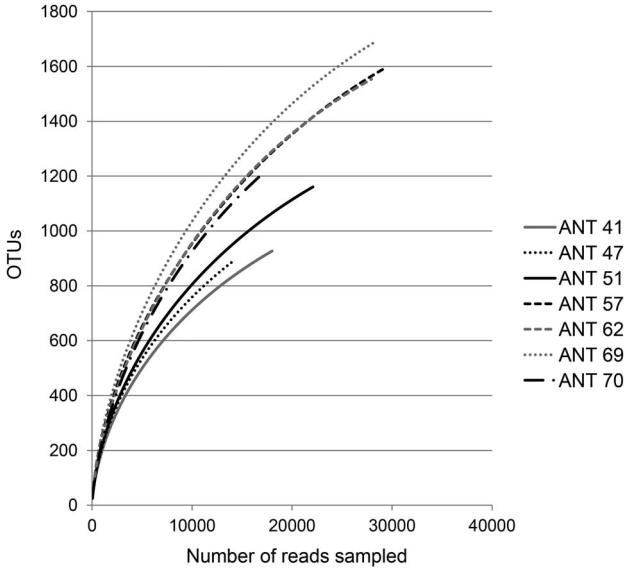

Fig. 5. Rarefaction analysis for each of the seven sequenced samples based on clustering at the 97% similarity level.

is mainly separated from group A and B by higher microeukaryotic contribution and a higher ice coverage. Group A is primarily separated from group B by higher salinities, lower temperatures, and a higher nanoeukaryotic contribution. Group B shows the highest water temperatures and the highest contribution of picoeukaryotes.

Detailed community structure

To obtain detailed taxonomic information about the community, we sequenced seven samples (samples 41, 47, 51, 57, 62, 69 and 70), spanning the entire transect and including all three ARISA groups. Three samples (samples 41,

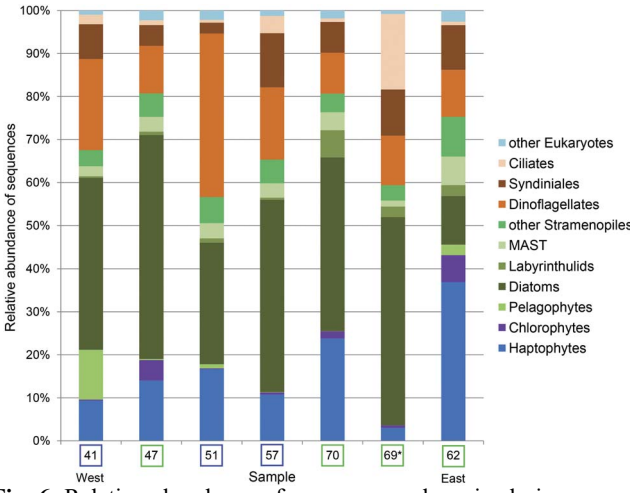

Fig. 6. Relative abundance of sequence reads, gained via 454-pyrosequencing, assigned to major taxonomic groups. Blue encircled samples = "offshore", green encircled samples = "inshore", * = 100% ice coverage.

https://doi.org/10.1017/S0954102013000229 Published online by Cambridge University Press

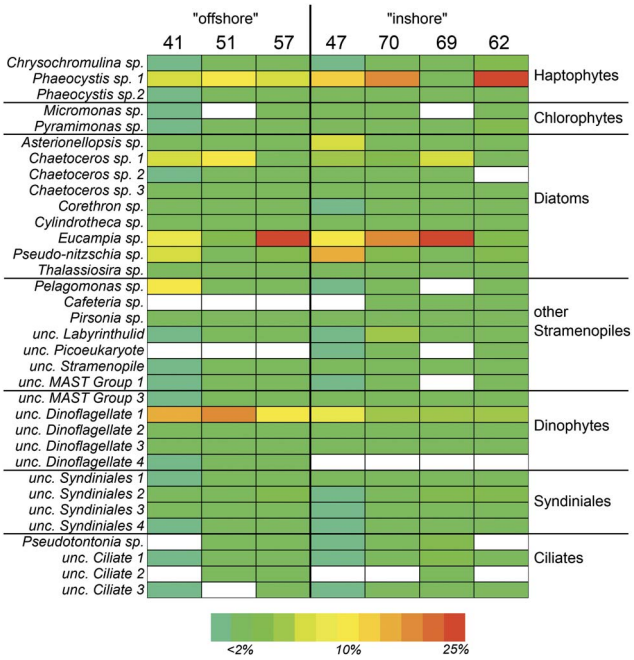

Fig. 7. Colour-coded matrix plot, illustrating the relative read abundance of abundant OTUs (operational taxonomic units) (abundance ≥ 1%, at least in one sample) in the seven sequenced samples. White boxes indicate the absence of the respective OTU.

51 and 57) were taken in open ocean waters and four samples (samples 47, 62, 69 and 70) were taken in inshore waters.

454-pyrosequencing

The summary of recovered 454-pyrosequencing reads is shown in Table I. In total, 278 116 sequence reads were obtained from 454-pyrosequencing, of which 77.1% had an acceptable length (300–670 bp). After the quality filtering, 56.5% of the total reads were left for analysis. The number of analysed reads ranged between 14 219 (sample 47) and 29 241 (sample 57). Subsequent to the clustering, 4044 different OTUs could be observed. The number of OTUs for each sample (Table I) ranged between 893 (sample 47) and 1687 (sample 69), at which only 0.7% (sample 57 and 70) to 1.7% (sample 47) were abundant (number of reads ≥ 1% of total reads). The proportion of unique OTUs (i.e. OTUs occurring in one sample only) was 36%.

The rarefaction curves (Fig. 5) show that none of the samples demonstrates saturation. However, the stacking of the curves suggests that samples 41, 47 and 51 harboured the lowest diversity.

The relative abundance of sequences assigned to major protist groups is shown in Fig. 6. Haptophytes showed a read abundance of 9–17% in offshore samples and 14–37% in inshore samples, except in sample 69, where they accounted for only 3%. Chlorophytes occurred in significant amounts only inshore where they composed 1.7–6.3% of the reads. Sample 69 was again an exception, because here chlorophytes only accounted for 0.6% of the reads. Pelagophytes only occurred in great quantities in one offshore sample (sample 41) with 11.6% of the sequence reads. Diatoms were the dominating group in samples 41, 47, 57, 69 and 70 with a read abundance of 40%, 52%, 44.7%, 48.3% and 40.3%, respectively. In samples 51 and 62, they accounted for 28.2% and 11.3% of the reads, respectively. Labyrinthulids occurred in significant amounts only inshore, in samples 62, 69 and 70, where their read abundance accounted for 2.5%, 2.5% and 6.3%, respectively. The read abundance of the marine stramenopiles (MAST) group comprised 1.4–6.6%, whereas the highest abundance occurred in sample 62. Dinoflagellates dominated the sequence assemblage in sample 51 with 38%. In the other samples, they accounted for 9.5–21.2% of the reads. In general, dinoflagellates showed a higher read abundance in offshore than in inshore samples. The highest read abundance of Syndiniales occurred in sample 57 (12.6%). In the other samples, they accounted for 2.5–10.7% of the reads. Ciliates played a minor role in all sequence assemblages, except in sample 69, where they account for 17.6% of the sequences.

Of the 4044 OTUs, 34 were abundant (i.e. abundance > 1%) in at least one sample. A detailed overview of the relative read abundances of the abundant phylotypes is shown in Fig. 7. Three phylotypes were abundant in all seven samples (*Phaeocystis* sp. 1, *Eucampia* sp. and unclassified (unc.) Dinoflagellate 1). They were also among the most abundant phylotypes across the entire transect. The *Phaeocystis* sp. 1 OTU showed the highest read abundance inshore, in samples 70 and 62, with 21% and 26.7%, respectively. However, in sample 69 it was almost rare, with the lowest read abundance of 1.5%. The chlorophytes, represented by *Micromonas* sp. and *Pyramimonas* sp., were only abundant inshore (sample 47 and 62), with 3.7% as highest sequence abundance. We found nine abundant phylotypes among the diatoms. The most abundant was *Eucampia* sp., with a read abundance up to 23.6% (sample 69). Only in sample 47 and 51, the most abundant diatom phylotype was not *Eucampia* sp., but *Pseudo-nitzschia* sp. (13.8%) and *Chaetoceros* sp. 1 (12.6%), respectively. *Pelagomonas* sp., belonging to the pelagophytes, showed a high read abundance offshore, in sample 41 (10.4%), whereas it was nearly rare in all other samples. Among the rest of the "other stramenopiles", the unc. labyrinthulid OTU showed the highest read abundance in sample 70 (4.8%). We found four abundant dinoflagellate phylotypes, of which the unc. Dinoflagellate 1 was the most abundant, with a read abundance ranging from 5.2–19.8%. The highest abundance appeared offshore in sample 51. The other dinoflagellate phylotypes did not exceed a read abundance of 2.1%. Among the abundant Syndiniales phylotypes, the unc. Syndiniales 2 showed the highest read abundance with 3.9% in sample 57. Ciliate phylotypes were only abundant in sample 69, where the unc. Ciliate 1 OTU showed the highest sequence abundance (3.4%). The rare biosphere accounted for 34.2% (sample 47) to 45.8% (sample 62) of all reads.

https://doi.org/10.1017/S0954102013000229 Published online by Cambridge University Press

Discussion

Structure/diversity overview and biogeographical patterns

One aim of this study was to determine the structure and diversity of eukaryotic protist assemblages in the Amundsen Sea and to assess the impact of environmental conditions on their biogeographical patterns. We used a combination of pigment analysis (HPLC) and ARISA to get an overview of the structure/diversity and the biogeographical patterns. The resulting ARISA profiles were linked with the environmental conditions.

Previous biogeographical classifications of surface waters are broad and of larger scale (Spalding *et al*. 2012). For shelf regions the existing classifications are more detailed (Spalding *et al*. 2007). In our previous study we confirmed characteristic protistan assemblages for each large-scale water mass in the Southern Ocean (Wolf *et al.* unpublished). However, it is also important to study more regional patterns, to complement our knowledge about the diversity and biogeography of protists in the Pacific sector of the Southern Ocean.

In general, we observed clear differences of total chl *a* concentrations between the samples taken offshore and inshore. Inshore, the concentrations always exceeded 1 $\mu g\,l^{-1}$. This is congruent with other studies, which observed higher chl *a* concentrations in Antarctic shelf and coastal waters than in open oceanic waters (Hashihama *et al*. 2008, Olguin & Alder 2011). Along the entire transect, sample 70 showed the highest chl *a* concentration with 9.58 $\mu g\,l^{-1}$, indicating a large phytoplankton bloom in this area. The high chl *a* value is not surprising, since recent investigations observed chl *a* concentrations up to 8–14 $\mu g\,l^{-1}$ in the shelf area of the Amundsen Sea (Alderkamp *et al*. 2012, Fragoso & Smith 2012, Mills *et al*. 2012).

The high chl *a* concentrations we observed above the shelf were accompanied by higher proportions of microeukaryotes. Higher chl *a* concentrations were often connected with high abundances of larger cells, like diatoms (Ishikawa *et al*. 2002). The geomorphology in the shelf areas promotes upwelling and mixing and thus, the nutrient availability in this region is higher, which promotes the build-up of biomass and favours larger cells (Irwin *et al*. 2006). Picoeukaryotes were of minor importance throughout the entire transect, which is in contrast to Diez *et al*. (2004), who found out that cells $< 5\mu m$ can contribute up to 80% to total chl *a* in Southern Ocean waters. However, they investigated a different area of the Southern Ocean (Drake Passage) and focused on cells $< 5\mu m$, which include small nanoeukaryotes. In our study, nanoeukaryotes were the counterpart to microeukaryotes in the investigated area. They showed their highest contribution in samples where microeukaryotes were less abundant.

The ARISA profiles generally support the existence of an offshore and an inshore group in the investigated area. The offshore group is split into a western and an eastern part, of which the eastern part was characterized by lower salinities, due to melting ice in this area. Samples 58–62 belong to the second offshore group, although they were taken above the shelf. One explanation could be that these areas are more influenced by open oceanic water. In these areas, Circumpolar Deep Water (CDW) is flowing onto the continental shelf through troughs in the shelf as modified CDW (Alderkamp *et al*. 2012) and may influence the surface layer (upwelling). However, it appears more likely that wind is the major determining factor, influencing the direction of the surface currents.

Detailed community structure

This study delivers the first protist diversity overview gained by molecular data. Previous studies used pigment based techniques and therefore lack deeper taxonomical resolution (Alderkamp *et al*. 2012, Fragoso & Smith 2012, Mills *et al*. 2012).

The most prominent taxonomic group across the entire transect were the diatoms. This group was previously observed to dominate in the Amundsen Sea, especially in the sea ice zones (Alderkamp *et al*. 2012, Fragoso & Smith 2012, Mills *et al*. 2012). The most dominant diatom in the Pine Island Bay was *Eucampia* sp. Garibotti *et al*. (2003) found a large contribution to total diatom biomass of *Eucampia antarctica* (Castracane) Mangin in Marguerite Bay (Antarctic Peninsula). It seems that the conditions in bays may constitute an optimal environment for *Eucampia* to grow. In the Amundsen polynya, we found *Pseudo-nitzschia* sp. as the most dominant diatom, whereas offshore, *Chaetoceros* sp. was generally the dominant diatom. These two genera were previously reported to dominate in waters around Antarctica (Gomi *et al*. 2005).

Sample 62, in contrast, showed a dominance of *Phaeocystis* sp. A dominance of *Phaeocystis antarctica* Karsten in several regions of the Amundsen Sea was previously reported (Alderkamp *et al*. 2012, Mills *et al*. 2012). Arrigo *et al*. (1999) revealed that *Phaeocystis antarctica* dominates where waters are deeply mixed, whereas diatoms dominate in highly stratified waters. Hence, the domination of *Phaeocystis* sp. in sample 62 could be due to more deeply mixed water. Another explanation could be that the succession at the eastern edge of the transect was most advanced (post bloom), due to a longer period free of ice, retraced via Advanced Microwave Scanning Radiometer (AMSR) satellite images (Spreen *et al*. 2008). In polar waters, after a diatom dominated bloom, *Phaeocystis* often dominated the post bloom situation (McMinn & Hodgson 1993).

Sample 69 showed the most extreme ice condition with 100%. Here, the read abundance of *Phaeocystis* was very low. The lack of wind stress, due to the ice coverage, could have caused the water to be highly stratified, and therefore

https://doi.org/10.1017/S0954102013000229 Published online by Cambridge University Press

led to a low *Phaeocystis* abundance (Arrigo *et al*. 1999, Goffart *et al*. 2000). Under the ice, ciliates showed their highest read abundance. This corresponds to other observations of the under-ice community structure, which revealed that heterotrophic biomass was dominated by ciliates (Ichinomiya *et al*. 2007).

In contrast to our previous study, focusing on the distribution of eukaryotic protists across the main oceanic fronts of the Southern Ocean (Wolf *et al.* unpublished), the distribution of OTUs was more even. The proportion of unique OTUs was only half the amount (37%) that it was across the main fronts of the Southern Ocean (76%). This is distinctly visible in the distribution of the abundant biosphere (Fig. 7). There were only a few OTUs, which were not present in all samples (10.1%), whereas in our previous study there were many more (20.4%). Here, in the single large-scale water mass, the observed "rare biosphere" serves as a background population and several species may become abundant when the environmental conditions change. However, in both studies the rarefaction curves suggest that none of the samples have been exhaustively analysed by sequencing. Nevertheless, the rarefaction curves indicate that the highest diversity was observed under the ice (sample 69).

There are some potential biases, which can confound the interpretation of molecular data. The amplification of the different species in an environmental sample can vary and some species might not be captured by the primers used (primer specificity) (Jeon *et al*. 2008, Stoeck *et al*. 2010). The ARISA approach can only serve as an approximate overview of the diversity structure, due to the qualitative character and the identical fragment lengths of some different species (Caron *et al*. 2012). Additionally, the number of rRNA gene copies depends on the cell size and varies between eukaryotes from one to several hundreds (Zhu *et al*. 2005). Especially the dinoflagellates seem to have more rRNA gene copies than the other taxonomical groups and thus might be overrepresented in molecular sequence data. The placement of sequences gained via 454-pyrosequencing has to be interpreted with care, because the length of only *c.* 500–600 bp is affecting the robustness. Therefore, we generally did not classify the OTUs beyond the genus level.

In conclusion, we have shown that within a single water mass protist assemblages differed in dimensions and species composition, according to geographical and environmental conditions. There were two major groups, the offshore and the inshore group. Biomass and microeukaryotes contribution to total chl *a* were highest in the inshore group, whereas in the offshore group the contribution of nanoeukaryotes was the highest across the entire transect. Diatoms were the most prominent protist class, and the diatom species appearing as most abundant differed between the locations. We delivered the first taxon detailed protist diversity overview in the Amundsen Sea during summer.

Acknowledgements

This study was accomplished within the Young Investigator Group PLANKTOSENS (VH-NG-500), funded by the Initiative and Networking Fund of the Helmholtz Association. We thank the captain and crew of the RV *Polarstern* for their support during the cruise. We are grateful to F. Kilpert and B. Beszteri for their bioinformatical support. We also want to thank A. Schroer, A. Nicolaus and K. Oetjen for technical support in the laboratory and Steven Holland for providing access to the program Analytic Rarefaction 1.3. We would like to acknowledge E.M. Nöthig and K. Kohls for their insightful discussions and comments on this manuscript. We also gratefully acknowledge the constructive comments of the reviewers.

References

Alderkamp, A.C., Mills, M.M., van Dijken, G.L., Laan, P., Thuroczy, C.E., Gerringa, L.J.A., de Baar, H.J.W., Payne, C.D., Visser, R.J.W., Buma, A.G.J. & Arrigo, K.R. 2012. Iron from melting glaciers fuels phytoplankton blooms in the Amundsen Sea (Southern Ocean): phytoplankton characteristics and productivity. *Deep-Sea Research II*, **71–76**, 32–48.

Arrigo, K.R., Robinson, D.H., Worthen, D.L., Dunbar, R.B., DiTullio, G.R., VanWoert, M. & Lizotte, M.P. 1999. Phytoplankton community structure and the drawdown of nutrients and CO2 in the Southern Ocean. *Science*, **283**, 365–367.

Barlow, R.G., Cummings, D.G. & Gibb, S.W. 1997. Improved resolution of mono- and divinyl chlorophylls a and b and zeaxanthin and lutein in phytoplankton extracts using reverse phase C-8 HPLC. *Marine Ecology Progress Series*, **161**, 303–307.

Behnke, A., Engel, M., Christen, R., Nebel, M., Klein, R.R. & Stoeck, T. 2011. Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environmental Microbiology*, **13**, 340–349.

Bidigare, R.R. 1991. Analysis of algal chlorophylls and carotenoids. *Geophysical Monograph Series*, **63**, 119–123.

Caron, D.A., Countway, P.D., Jones, A.C., Kim, D.Y. & Schnetzer, A. 2012. Marine protistan diversity. *Annual Review of Marine Science*, **4**, 467–493.

Diez, B., Massana, R., Estrada, M. & Pedros-Alio, C. 2004. Distribution of eukaryotic picoplankton assemblages across hydrographic fronts in the Southern Ocean, studied by denaturing gradient gel electrophoresis. *Limnology and Oceanography*, **49**, 1022–1034.

Dray, S. & Dufour, A.B. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, **22**, 1–20.

Eddy, S.R. 2011. Accelerated profile HMM searches. *Plos Computational Biology*, 10.1371/journal.pcbi.1002195.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C. & Knight, R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, **27**, 2194–2200.

Elwood, H.J., Olsen, G.J. & Sogin, M.L. 1985. The small-subunit ribosomal RNA gene sequences from the hypotrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*. *Molecular Biology and Evolution*, **2**, 399–410.

Finlay, B.J. 2002. Global dispersal of free-living microbial eukaryote species. *Science*, **296**, 1061–1063.

Finlay, B.J. & Fenchel, T. 2004. Cosmopolitan metapopulations of free-living microbial eukaryotes. *Protist*, **155**, 237–244.

Fragoso, G.M. & Smith, W.O. 2012. Influence of hydrography on phytoplankton distribution in the Amundsen and Ross seas, Antarctica. *Journal of Marine Systems*, **89**, 19–29.

GARIBOTTI, I.A., VERNET, M., FERRARIO, M.E., SMITH, R.C., ROSS, R.M. & QUETIN, L.B. 2003. Phytoplankton spatial distribution patterns along the western Antarctic Peninsula (Southern Ocean). *Marine Ecology Progress Series*, **261**, 21–39.

GAST, R.J., DENNETT, M.R. & CARON, D.A. 2004. Characterization of protistan assemblages in the Ross Sea, Antarctica, by denaturing gradient gel electrophoresis. *Applied and Environmental Microbiology*, **70**, 2028–2037.

GOFFART, A., CATALANO, G. & HECQ, J.H. 2000. Factors controlling the distribution of diatoms and Phaeocystis in the Ross Sea. *Journal of Marine Systems*, **27**, 161–175.

GOMI, Y., UMEDA, H., FUKUCHI, M. & TANIGUCHI, A. 2005. Diatom assemblages in the surface water of the Indian sector of the Antarctic Surface Water in summer of 1999/2000. *Polar Bioscience*, **18**, 1–15.

GRAVALOSA, J.M., FLORES, J.A., SIERRO, F.J. & GERSONDE, R. 2008. Sea surface distribution of coccolithophores in the eastern Pacific sector of the Southern Ocean (Bellingshausen and Amundsen seas) during the late austral summer of 2001. *Marine Micropaleontology*, **69**, 16–25.

GRIFFITHS, H.J. 2010. Antarctic marine biodiversity - what do we know about the distribution of life in the Southern Ocean? *PloS ONE*, **5**, e11683.

HASHIHAMA, F., HIRAWAKE, T., KUDOH, S., KANDA, J., FURUYA, K., YAMAGUCHI, Y. & ISHIMARU, T. 2008. Size fraction and class composition of phytoplankton in the Antarctic marginal ice zone along the 140°E meridian during February–March 2003. *Polar Science*, **2**, 109–120.

ICHINOMIYA, M., HONDA, M., SHIMODA, H., SAITO, K., ODATE, T., FUKUCHI, M. & TANIGUCHI, A. 2007. Structure of the summer under fast ice microbial community near Syowa Station, eastern Antarctica. *Polar Biology*, **30**, 1285–1293.

IRWIN, A.J., FINKEL, Z.V., SCHOFIELD, O.M.E. & FALKOWSKI, P.G. 2006. Scaling-up from nutrient physiology to the size-structure of phytoplankton communities. *Journal of Plankton Research*, **28**, 459–471.

ISHIKAWA, A., WRIGHT, S.W., VAN DEN ENDEN, R.L., DAVIDSON, A.T. & MARCHANT, H.J. 2002. Abundance, size structure and community composition of phytoplankton in the Southern Ocean in the austral summer 1999/2000. *Polar Bioscience*, **15**, 11–26.

JEFFREY, S.W., MANTOURA, R.F.C. & BJORNLAND, T. 1997. Data for the identification of 47 key phytoplankton pigments. *In* JEFFREY, S.W., MANTOURA, R.F.C. & WRIGHT, S.W., *eds. Phytoplankton pigments in oceanography*. Paris: UNESCO, 449–559.

JEON, S., BUNGE, J., LESLIN, C., STOECK, T., HONG, S.H. & EPSTEIN, S.S. 2008. Environmental rRNA inventories miss over half of protistan diversity. *Bmc Microbiology*, 10.1186/1471-2180-8-222.

KAISER, S., BARNES, D.K.A., SANDS, C.J. & BRANDT, A. 2009. Biodiversity of an unknown Antarctic sea: assessing isopod richness and abundance in the first benthic survey of the Amundsen continental shelf. *Marine Biodiversity*, **39**, 27–43.

KUNIN, V., ENGELBRETKSON, A., OCHMAN, H. & HUGENHOLTZ, P. 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology*, **12**, 118–123.

LACHANCE, M.A. 2004. Here and there or everywhere? *Bioscience*, **54**, 884.

LOPEZ-GARCIA, P., RODRIGUEZ-VALERA, F., PEDROS-ALIO, C. & MOREIRA, D. 2001. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, **409**, 603–607.

MACKEY, M.D., MACKEY, D.J., HIGGINS, H.W. & WRIGHT, S.W. 1996. CHEMTAX - A program for estimating class abundances from chemical markers: application to HPLC measurements of phytoplankton. *Marine Ecology Progress Series*, **144**, 265–283.

MARGULIES, M., EGHOLM, M. & ALTMAN, W.E. *et al.* 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, **437**, 376–380.

MATSEN, F.A., KODNER, R.B. & ARMBRUST, E.V. 2010. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *Bmc Bioinformatics*, 10.1186/1471-2105-11-538.

MCMINN, A. & HODGSON, D. 1993. Summer phytoplankton succession in Ellis Fjord, Eastern Antarctica. *Journal of Plankton Research*, **15**, 925–938.

MEDLIN, L., ELWOOD, H.J., STICKEL, S. & SOGIN, M.L. 1988. The characterization of enzymatically amplified eukaryotic 16s-like rRNA-coding regions. *Gene*, **71**, 491–499.

MILLS, M.M., ALDERKAMP, A.C., THUROCZY, C.E., VAN DIJKEN, G.L., LAAN, P., DE BAAR, H.J.W. & ARRIGO, K.R. 2012. Phytoplankton biomass and pigment responses to Fe amendments in the Pine Island and Amundsen polynyas. *Deep-Sea Research II*, **71–76**, 61–76.

MIRANDA, L.N., ZHUANG, Y.Y., ZHANG, H. & LIN, S. 2012. Phylogenetic analysis guided by intragenomic SSU rDNA polymorphism refines classification of "Alexandrium tamarense" species complex. *Harmful Algae*, **16**, 35–48.

NICKRENT, D.L. & SARGENT, M.L. 1991. An overview of the secondary structure of the V4-region of eukaryotic small-subunit ribosomal-RNA. *Nucleic Acids Research*, **19**, 227–235.

OKSANEN, J., BLANCHET, F.G., KINDT, R., LEGENDRE, P., O'HARA, R.B., SIMPSON, G.L., SOLYMOS, P., STEVENS, M.H.H. & WAGNER, H. 2011. *vegan: Community Ecology Package*. R package version 1.17-6. http://cran.r-project.org.

OLGUIN, H.F. & ALDER, V.A. 2011. Species composition and biogeography of diatoms in Antarctic and sub-Antarctic (Argentine shelf) waters (37–76°S). *Deep-Sea Research II*, **58**, 139–152.

RAMETTE, A. 2009. Quantitative community fingerprinting methods for estimating the abundance of operational taxonomic units in natural microbial communities. *Applied and Environmental Microbiology*, **75**, 2495–2505.

R DEVELOPMENT CORE TEAM. 2008. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. http://www.R-project.org.

SMITH, J.L., BARRETT, J.E., TUSNADY, G., REJTO, L. & CARY, S.C. 2010. Resolving environmental drivers of microbial community structure in Antarctic soils. *Antarctic Science*, **22**, 673–680.

SPALDING, M.D., AGOSTINI, V.N., RICE, J. & GRANT, S.M. 2012. Pelagic provinces of the world: a biogeographic classification of the world's surface pelagic waters. *Ocean & Coastal Management*, **60**, 19–30.

SPALDING, M.D., FOX, H.E., HALPERN, B.S., MCMANUS, M.A., MOLNAR, J., ALLEN, G.R., DAVIDSON, N., JORGE, Z.A., LOMBANA, A.L., LOURIE, S.A., MARTIN, K.D., MCMANUS, E., RECCHIA, C.A. & ROBERTSON, J. 2007. Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience*, **57**, 573–583.

SPREEN, G., KALESCHKE, L. & HEYGSTER, G. 2008. Sea ice remote sensing using AMSR-E 89-GHz channels. *Journal of Geophysical Research - Oceans*, 10.1029/2005JC003384.

STOECK, T., BASS, D., NEBEL, M., CHRISTEN, R., JONES, M.D.M., BREINER, H.W. & RICHARDS, T.A. 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology*, **19**, 21–31.

WHITE, T.J., BRUNS, T., LEE, S. & TAYLOR, J.W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* INNIS, M.A. *et al.*, *eds. PCR protocols: a guide to methods and applications*. New York: Academic Press, 315–322.

WRIGHT, S.W., ISHIKAWA, A., MARCHANT, H.J., DAVIDSON, A.T., VAN DEN ENDEN, R.L. & NASH, G.V. 2009. Composition and significance of picophytoplankton in Antarctic waters. *Polar Biology*, **32**, 797–808.

ZHU, F., MASSANA, R., NOT, F., MARIE, D. & VAULOT, D. 2005. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *Fems Microbiology Ecology*, **52**, 79–92.