HAL open science

Specific eukaryotic plankton are good predictors of net community production in the Western Antarctic Peninsula

Yajuan Lin, Nicolas Cassar, Adrian Marchetti, Carly Moreno, Hugh Ducklow, Zuchuan Li

To cite this version:

Yajuan Lin, Nicolas Cassar, Adrian Marchetti, Carly Moreno, Hugh Ducklow, et al.. Specific eukaryotic plankton are good predictors of net community production in the Western Antarctic Peninsula. Scientific Reports, 2017, 7, pp.14845. 10.1038/s41598-017-14109-1 . hal-02577611

HAL Id: hal-02577611

https://hal.science/hal-02577611v1

Submitted on 14 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

SCIENTIFIC REPORTS

OPEN

Specific eukaryotic plankton are good predictors of net community production in the Western Antarctic Peninsula

Received: 16 May 2017
Accepted: 3 October 2017
Published online: 01 November 2017

Yajuan Lin[1,4], Nicolas Cassar[1,4], Adrian Marchetti[2], Carly Moreno[2], Hugh Ducklow[3] & Zuchuan Li[1]

Despite our current realization of the tremendous diversity that exists in plankton communities, we have little understanding of how this biodiversity influences the biological carbon pump other than broad paradigms such as diatoms contributing disproportionally to carbon export. Here we combine high-resolution underway O_2/Ar, which provides an estimate of net community production, with high-throughput 18 S ribosomal DNA sequencing to elucidate the relationship between eukaryotic plankton community structure and carbon export potential at the Western Antarctica Peninsula (WAP), a region which has experienced rapid warming and ecosystem changes. Our results show that in a diverse plankton system comprised of ~464 operational taxonomic units (OTUs) with at least 97% 18 S identity, as few as two or three key OTUs, i.e. large diatoms, Phaeocystis, and mixotrophic/phagotrophic dinoflagellates, can explain a large majority of the spatial variability in the carbon export potential (76–92%). Moreover, we find based on a community co-occurrence network analysis that ecosystems with lower export potential have more tightly coupled communities. Our results indicate that defining plankton communities at a deeper taxonomic resolution than by functional groups and accounting for the differences in size and coupling between groups can substantially improve organic carbon flux predictions.

The biological carbon pump, fueled by nutrient delivery to the sunlit layers of the ocean, is believed to be a function of multiple processes, including net community production (NCP), where NCP is equal to gross primary production minus community respiration. Plankton communities regulate NCP through their influence on the fate of organic matter production, which can be either recycled or exported to depth depending on phytoplankton size and density[1–3], as well as microbial and grazing activities[4]. This is generally predicted and modeled to be a function of Plankton Functional Types (PFTs). PFTs refer to collections of organisms based on explicit biogeochemical roles or requirements. Conventional PFTs include picoautotrophs, diazotrophs (N_2-fixers), calcifiers (e.g., coccolithophorids), and silicifiers (e.g., diatoms)[5–7]. These PFTs are expected to contribute differently to carbon export production and carbon export efficiency. For example, most diatoms are believed to be efficient carbon exporters to depth because of their large size and heavy silica shells, which make them sink faster[8].

In this study, we combine high-frequency NCP estimates[9,10] with in-depth molecular taxonomic measurements[11–13] to explore the relationship between the carbon export potential and plankton community assemblages at the Western Antarctic Peninsula (WAP). High-throughput sequencing has recently been shown to be a powerful tool to explore the influence of plankton community structure on carbon export in the oligotrophic oceans[14]. We apply similar molecular tools at the WAP, where large spatial variability in iron and light availability drives steep gradients in plankton diversity and NCP[15,16]. Furthermore, the WAP may represent the "canary in a coal mine" of the Southern Ocean, a region that has a disproportionate influence on global climate[17]. The WAP region

[1]Division of Earth and Ocean Sciences, Nicholas School of the Environment, Duke University, Durham, NC, 27708, USA. [2]Department of Marine Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27514, USA. [3]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, 10964, USA. [4]Université de Brest - UMR 6539 CNRS/UBO/IRD/Ifremer, Laboratoire des sciences de l'environnement marin – IUEM, Rue Dumont D'Urville, Plouzané, 29280, France. Correspondence and requests for materials should be addressed to Y.L. (email: yajuan.lin@univ-brest.fr)

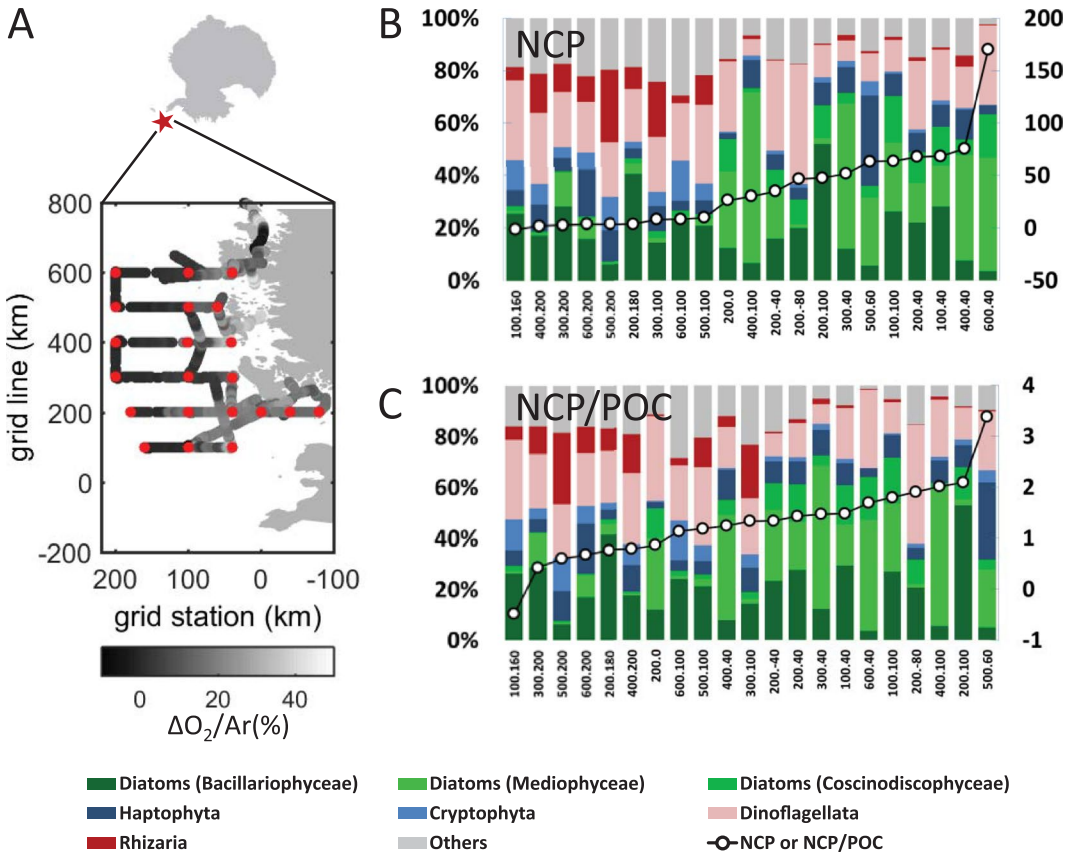


Figure 1. Eukaryotic plankton community composition of major taxonomic groups versus NCP. (**A**) Location of the Western Antarctica Peninsula and sampling sites (red dots) projected on Palmer LTER grid (x axis – grid station, y axis – grid line, in km). Samples were taken across a large gradient in biological O2. Stations (labeled as "grid line. grid station") are ranked in ascending values of NCP (black circles) (**B**) in mmol C m^{-2} day^{-1}, and (**C**) NCP/POC in m day^{-1}. The relative abundance of different taxa (Phylum to Division) in % of 18 S reads are represented in different colors. Centric diatoms (Mediophyceae and Coscinodiscophyceae) are in light greens, whereas pennate diatoms (Bacillariophyceae) are in dark green. Less dominant taxa (<5% in total abundance) are combined into 'others' in grey. The maps in this figure are generated by MATLAB R2014a (http://www.mathworks.com/).

has experienced drastic changes[18,19] over the last 50 years warming at an unprecedented rate[20]. Community production at the bottom of the food chain, and local and remote upper trophic levels (e.g., krill, fish, penguins and other seabirds, seals, and whales) are likely to be affected by this rapid warming and associated sea ice loss[21].

Using DNA sequencing, we first characterized community structure at a taxonomic level comparable to PFTs and High Performance Liquid Chromatography (HPLC) pigment-based definitions (Fig. 1, Supplementary Table S1). At this level of phylum to division, the sequences were grouped into diatoms, Haptophyta, Cryptophyta, Dinoflagellata, among others. Based on relative contributions, we found that diatoms were the dominant components of the plankton population during our survey, accounting for an average of 41% of the total reads. In order to account for the influence of biomass on NCP, we normalized NCP to Particulate Organic Carbon (POC) concentration in the mixed layer to identify "High Biomass, Low NCP" and "Low Biomass, High NCP" regime[10,22]. Cassar *et al.* (2015)[10] hypothesized based on theoretical considerations that the POC-normalized NCP may, under some conditions, provide insight into the apparent sinking rate and turn over time of the organic carbon inventory in the mixed layer. We found that the relative abundance of diatoms was positively correlated with NCP ($r^2 = 0.35$, $p = 0.005$) but that there was no significant correlation with NCP/POC. Diatoms were further divided into centrics and pennates. Near the coast, on the continental shelf, and in the open ocean, the relative abundance (%) of centrics averaged 36.8 ± 16.2, 22.9 ± 26.5, and 5.6 ± 4.7; whereas pennates averaged 14.2 ± 8.2, 23.9 ± 15.5, and 22.2 ± 11.9, respectively. While there were no significant offshore-onshore trends in the relative abundance of pennate diatoms, the relative abundance of centric diatoms was significantly higher near the coast compared to the open ocean (non-parametric Mann-Whitney U test, $P < 0.01$). A positive correlation was observed between the relative abundance of centric diatoms and NCP ($r^2 = 0.47$, $p = 0.0006$), as well as centric diatoms and NCP/POC ($r^2 = 0.20$, $p = 0.04$). Surprisingly, the relative abundance of pennate diatoms was not significantly correlated with NCP or NCP/POC, which may be explained by the difference in size between centrics and pennates in the region (see discussion below). However, should this relation hold in other regions of the Southern Ocean, a correlation between HPLC pigment estimates of total diatom abundance and NCP may not be observed. This could

explain why earlier studies found no clear correlation between HPLC diatom relative abundance and NCP in the WAP region[15] and in the Subantarctic Zone south of Australia[10].

Little is known about the carbon export potential of particular eukaryotic plankton species. The vast array of plankton attributes (e.g., cell biovolume, elemental ratios, and degree of silicification for diatoms) and the associated food web dynamics (grazing, aggregation, and fecal pellet export) are likely to be first-order controls on the carbon export potential. In order to gain insight into how NCP relates to the presence of specific eukaryotic plankton, we further characterized plankton assemblages at the OTU level (i.e., species to genus) based on a 97% DNA similarity threshold (Fig. 2). In total, 464 OTUs were defined in the surveyed region. The top 20 OTUs, a small portion (4.3%) of the total diversity, accounted for 78% of the overall sequence read abundance. This is consistent with the hyperdominance of few taxa in oceanic plankton ecosystems as recently reported by the Tara Oceans plankton survey[11]. Based on the relative abundance of these 20 OTUs, a stochastic search variable selection (SSVS) analysis identified which OTUs were the best predictors of NCP and NCP/POC (Supplementary Table S2). Stepwise regression analyses confirmed the top variables identified by SSVS (Supplementary Table S3). In the resulting multiple linear regression model based on the SSVS selected variables, three OTUs accounted for 92% of the NCP variance (Fig. 2A, Supplementary Table S3), and only two OTUs explained 76% of the NCP/POC variance (Fig. 2B, Supplementary Table S3). Both are significantly higher than the variance explained by the relative abundance at the higher taxonomic level (Supplementary Table S1), e.g. bulk diatoms explained 35% and 16% of the NCP and NCP/POC variances, respectively. Interestingly, the relative abundance of the most dominant OTU (OTU_4) (Supplementary Fig. S2), a pennate diatom belonging to the genus *Fragilariopsis*, was not a good predictor of the geographic variability of NCP ($r^2 = 0.12$, $p > 0.1$) or NCP/POC ($r^2 = 0.05$, $p > 0.1$). Many *Fragilariopsis* species are ice algae and the high abundance in some areas may be due to sea ice melt seeding rather than new cell growth. Our results suggest that PFTs may not be good predictors of NCP in some regions and that the intra-PFT variability may be greater than inter-PFT in terms of NCP.

The relative abundances of two diatoms with distinct biogeography, i.e. *Stellarima* (OTU_8) in a north bloom near Palmer Station and *Proboscia* (OTU_6) dominating the south ice edge bloom (Fig. 2; Supplementary Fig. S1), were both positively correlated with NCP but not NCP/POC. This suggests that these OTUs are associated with high production but relatively less efficient export systems. In contrast, the haptophyte *Phaeocystis* (OTU_3) was not correlated with NCP, but positively correlated with NCP/POC. *Phaeocystis* accounted for only 9.0% of the total 18 S reads on average, but co-dominated (34.4%) with diatoms (36.1%) at station (600.040) that had the highest NCP/POC (Fig. 2). *Phaeocystis* has been reported to be a major contributor to high particle flux out of the euphotic zone in the Amundsen Sea Polynya[23] and in the Ross Sea[24,25]. It has been observed to form massive aggregates at the colony stage and sink rapidly out of the ocean surface and thus has been considered to be a more efficient exporter of carbon compared to diatoms[24].

Two mixotrophic/phagotrophic dinoflagellate OTUs were identified as key negative contributors to the geographical variability of NCP (OTU_238) and NCP/POC (OTU_318) (Fig. 2), suggesting these OTUs are likely heterotrophic. Studies in the WAP have mostly focused on large macrozooplankton grazers such as copepods, krill and salps[26,27]. In contrast to large grazers enhancing NCP by packing organic matter into rapidly-sinking fecal pellets, unicellular protistan microzooplankton are more likely to reduce carbon export through surface remineralization[28]. Microzooplankton are believed to represent an efficient top-down control on phytoplankton growth due to their short generation time and rapid reproductive cycles. With high growth and ingestion rates[29], these small grazers could efficiently terminate an algal bloom or direct organic carbon production to DOC thereby reducing export. OTU_238 belongs to a recently-identified kleptoplastic dinoflagellate group (Supplementary Table S4)[30], known to prey on *Phaeocystis* and retain their chloroplasts for photosynthesis[31]. Such behavior further complicates interpretation of HPLC pigments[32]. This group of dinoflagellates is abundant in Antarctic waters and sea ice[30,32,33]. Our results suggest that they may be important controls on NCP. NCP/POC was negatively correlated to the presence of dinoflagellate OTU_318, which belongs to the genus *Gyrodinium* (Supplementary Table S4) (Fig. 2B). *Gyrodinium* have been observed to ingest diatoms much larger than their own body size and may curtail diatom blooms[34,35]. A negative relationship between HPLC dinoflagellates and NCP/POC was also observed in the Australian sector of the Southern Ocean[10].

At the higher taxonomic level (supergroup), NCP and NCP/POC were also negatively correlated with Rhizaria ($r^2 = 0.29$, $p = 0.01$; $r^2 = 0.20$, $p = 0.04$, respectively), a group of amoeboid heterotrophic protists with members known to feed on diatoms as well as other protists and bacteria[36]. Foraminifera, the protists known to carry autotrophic endosymbionts, are also within this supergroup. The ecology and physiology of Rhizaria are not well understood and are not detected by traditional HPLC analyses due to their lack of pigments. However, the recent discovery of their significant contribution to plankton diversity and abundance worldwide[11,37] suggests that they must serve an important ecological function.

We further investigated the potential impact of phytoplankton species-specific growth rates and size on their relationships to carbon export potential using cultured diatom isolates recently obtained from the WAP region. With the caveat that growth rates measured under laboratory conditions may not necessarily reflect *in situ* growth rates, we found that absolute growth rates alone cannot explain the spatial variability in our observations. The distribution of *F. cylindrus* does not correlate to NCP or NCP/POC, yet its maximum growth rate in culture is significantly higher ($P < 0.05$) than *Proboscia alata* and *Actinocyclus actinochilus* (a strain closely related to *Stellarima*), which belong to or are closely related to the identified key diatom OTUs that are best predictors of NCP. Under varying Fe and light availability growth conditions (Fig. 3), *P. alata* and *A. actinochilus* display comparable reductions in relative growth rates to *F. cylindrus*. Yet, even though the cellular growth rates under the examined conditions were comparable between the three diatoms (i.e., varied by less than 2-fold), they are significantly different in size. The biovolume per cell for *P. alata* and *A. actinochilus* are more than two orders of magnitude larger than that of *F. cylindrus* (Fig. 3). Large size could directly enhance the sinking rates of phytoplankton cells and lead to higher export, or uncouple them from their slow-growing grazers. The extent of decoupling could also result from

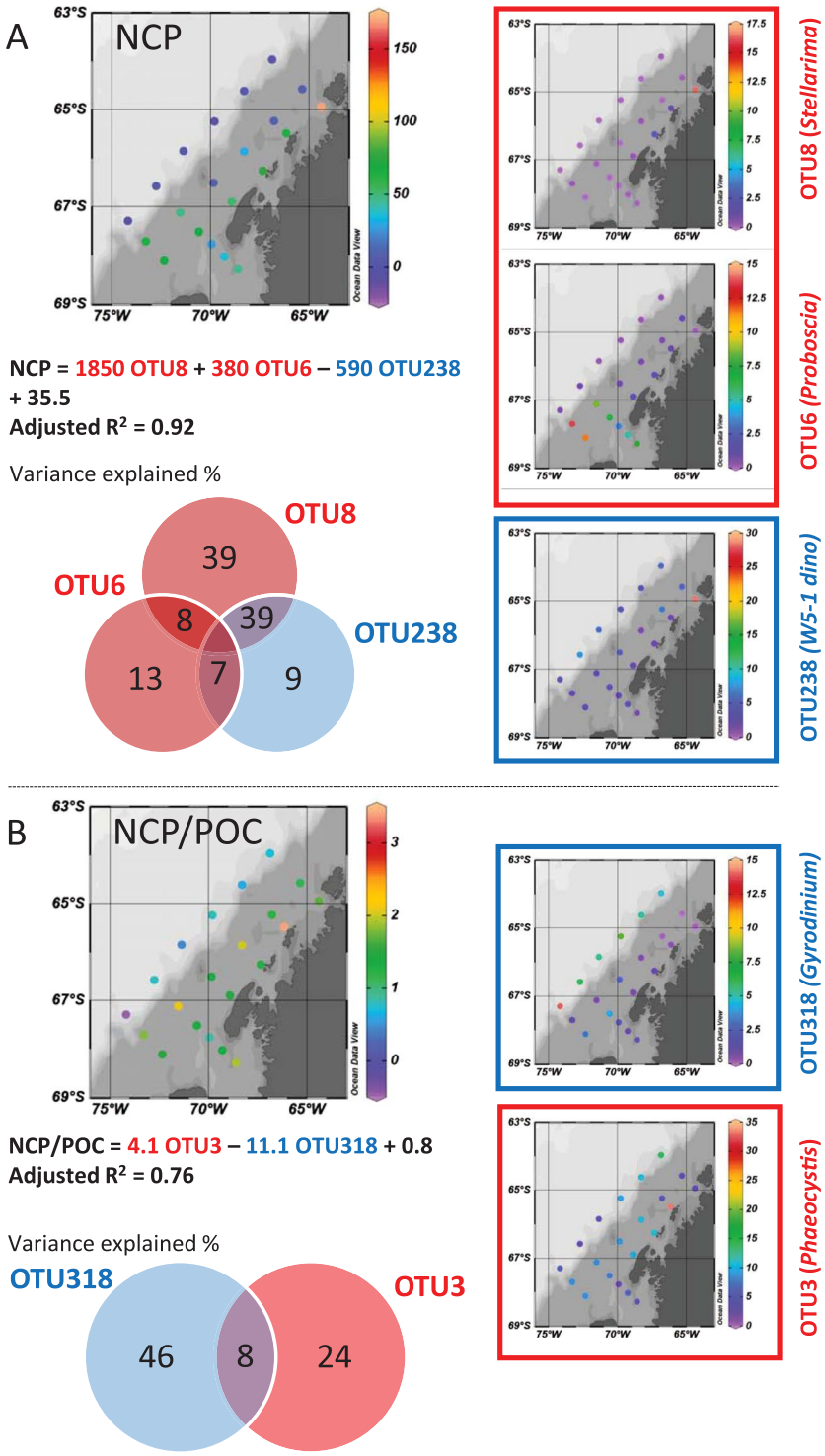


Figure 2. Community composition at the OTU (97%) level explaining carbon export. In a multiple linear regression model, 3 and 2 key OTUs selected by SSVS explained most of the variance in (**A**) NCP in mmol C m^{-2} day^{-1} and (**B**) NCP/POC in m day^{-1} geographical variability, respectively. Each OTU (relative 18 S abundance in %) has distinct biogeography, and its variance partitioning is shown in a Venn diagram. OTUs in red (blue) are positively (negatively) correlated with NCP or NCP/POC. Maps are created with Ocean Data View version 4.6.3.1 (Schlitzer, R., Ocean Data View, http://odv.awi.de, 2015).

the differences in the response time of phytoplankton once environments are favorable, a term which is not necessarily related to absolute growth rates[38]. Phytoplankton species can be characterized by two distinct life strategies – a persistence strategy and a boom-and-bust strategy[39]. Species with a persistent life strategy are better at stress

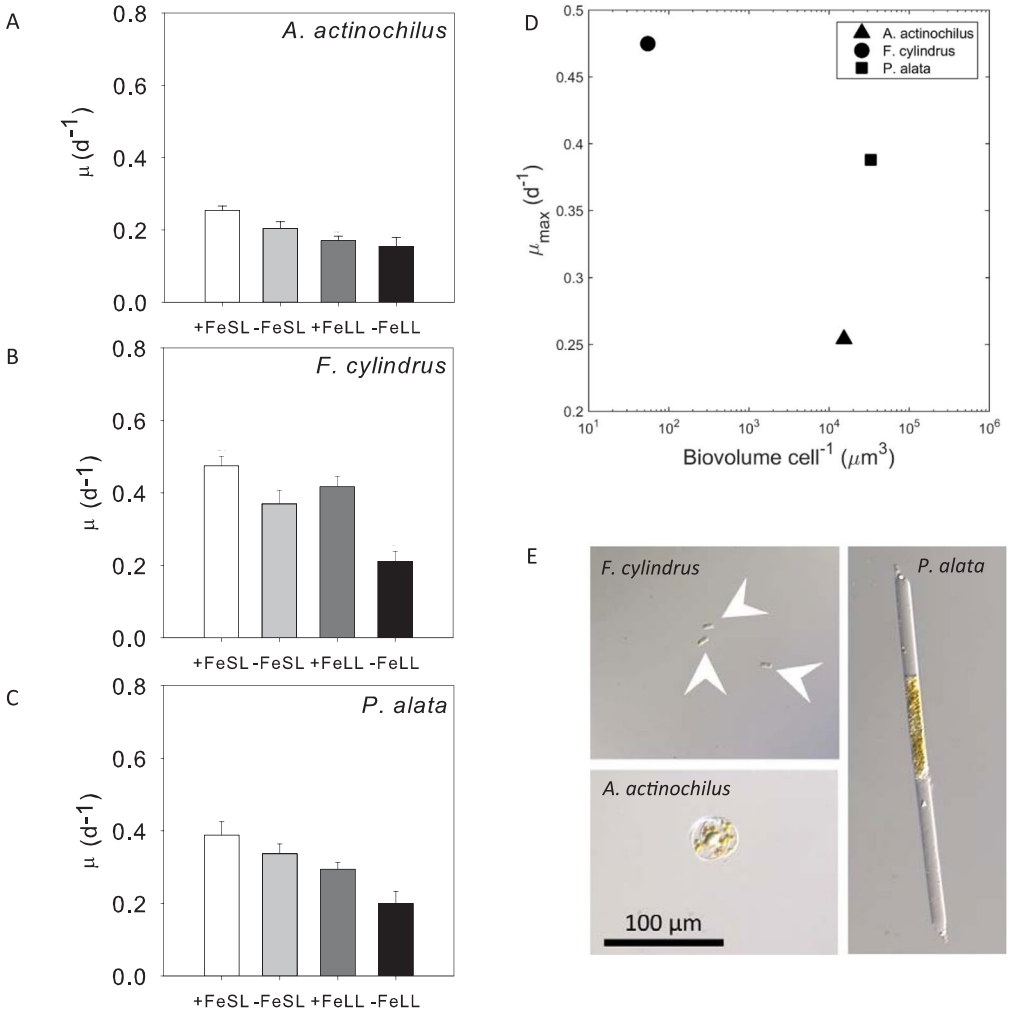
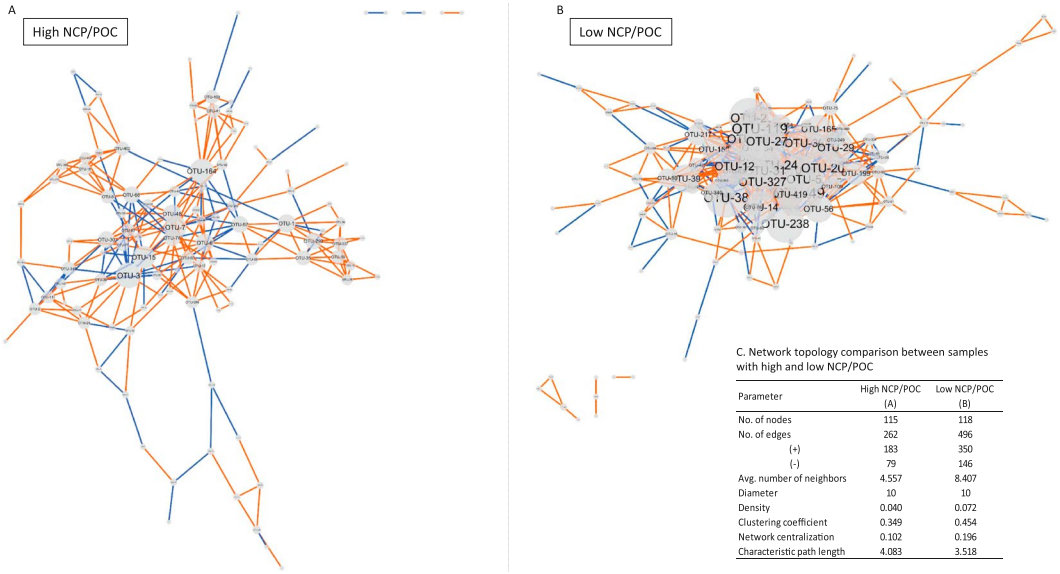


Figure 3. Comparison of specific growth rates and cell biovolume. Three diatom isolates collected from the WAP, *F. cylindrus* (**A**), *P. alata* (**B**), and *A. actinochilus* (**C**) a strain closely related to *Stellarima*, were maintained at 4 °C at either low light (10 μmol photons m^{-2} s^{-1}) or saturating light (90 μmol photons m^{-2} s^{-1}), high iron (1370 nmol L^{-1}) or low iron (3.1 nmol L^{-1}). The controlled culture conditions include high iron and saturating light (+FeSL), low iron and saturating light (−FeSL), high iron and low light (+FeLL), and low iron and low light (−FeLL). *F. cylindrus* has higher maximum growth rates than the other two isolates, but its biovolume is more than two orders of magnitude lower than the other two strains (**C**). Differential interference contrast images showing the cell size difference (**D**). All the images are on the same scale bar with white arrows pointing to *F. cylindrus* cells.

tolerance, such as *F. cylindrus* known to tolerate low temperature, photoinhibition and iron stress[40,41]. Under stable conditions, they dominate the environment in abundance, but with relatively stable growth rates they are under strong grazing control. Conversely, once the environment becomes favorable, boom-and-bust species with high growth rates can quickly outpace grazer growth and contribute to more efficient export.

Other than the effects of biomass, size and phytoplankton growth rates, species interactions and particularly grazing are considered regulators of carbon export. To investigate the influence of interactions on the carbon export potential, we performed a co-occurrence network analysis using SparCC (sparse correlation for compositional data[42]) on two equally divided subsets of 18 S OTUs data based on their NCP/POC values (Fig. 4). The topological features of the networks[43] provide insights into putative plankton interactions and niche partitioning of diverse microbes in various environments[14,44,45]. In brief, the average number of connected neighbors, or network density in the normalized version (0 to 1), measures the number of neighbors a node is connected with; the network centralization describes to which extent certain nodes are highly connected to others (e.g., star-shaped network has high centralization); the clustering coefficient measures the degree of clustering between nodes, calculated as ratio of observed vs. maximum number of edges given the number of neighbors; and the characteristic path length is the average shortest path length (number of steps) to connect two nodes. The detailed properties of the low and high NCP/POC networks are summarized in Fig. 4C. Although the sample size is relatively small (i.e., 11 stations for each group), low and high NCP/POC sites are clearly different. Despite having a similar number



C. Network topology comparison between samples with high and low NCP/POC

Parameter	High NCP/POC (A)	Low NCP/POC (B)
No. of nodes	115	118
No. of edges	262	496
(+)	183	350
(-)	79	146
Avg. number of neighbors	4.557	8.407
Diameter	10	10
Density	0.040	0.072
Clustering coefficient	0.349	0.454
Network centralization	0.102	0.196
Characteristic path length	4.083	3.518

Figure 4. The co-occurrence network interactions of eukaryotic phytoplankton and protozoan. Data sets are divided into two groups, one with high NCP/POC (**A**) and one with low NCP/POC (**B**). The edges between different nodes (OTUs) represent strong and significant connections (SparCC correlations with magnitude > 0.75 and p-values > 0.05). Isolated nodes are not shown in the network graph. Blue edges represent negative connections and orange one represent positive connections. The nodes and OTU labels are scaled by their degrees - the number of neighbors they are interacting with. The small groups of nodes (upper right in (**A**) and lower left in (**B**)) are isolated clusters not connected to the central network. (**C**) Table reporting the network topology comparison. Networks are plotted using Cytoscape v3.4.0 (http://www.cytoscape.org).

of nodes, the low NCP/POC communities, compared to the high NCP/POC communities, are characterized by almost double the number of edges (496 vs. 262), both for positive (co-presence) and negative (e.g., mutual exclusion) interactions, as well as higher average number of connected neighbors and higher network density, higher network centralization, higher clustering coefficient, and shorter characteristic path lengths. All these features suggest that the low NCP/POC communities are more tightly coupled than the high NCP/POC communities. This may support the general paradigm that more tightly coupled ecosystems (e.g., coupling between microzookplankton grazers and phytoplankton), are less efficient exporters of carbon. Particularly, microzooplankon grazing could reduce export efficiency through longer food webs, lower trophic efficiencies, and the breakdown of large particles, enhancing organic matter remineralization. In the networks presented in Fig. 4, predator-prey coupling could be positive or negative depending on whether the system is bottom up or top down controlled. In light of the central role grazing plays in carbon fluxes, future network analyses should also include macrozooplankton.

There are several caveats to our study. A lack of steady-state in the mixed-layer carbon budget may blur some of the relationships. Normalizing NCP to POC provides us with a qualitative approach to evaluate how productive the ecosystem is for a given amount of biomass at the ocean surface. However, POC reflects an amalgam of detrital and living organic matter. Our normalization only partly accounts for the multitude of factors that are believed to influence the carbon export potential, including the proportion of detrital matter, aggregation, physiological status, plankton migration and many other processes. These factors vary as a function of the stage of the plankton ecosystem over time (e.g. blooming vs. senescence), which is not taken into account with our normalization. Uncertainty in the ratios of net community O_2 production to CO_2 uptake could affect our oxygen-derived calculations that assume a quotient of 1.4.[46] have found that the PQ at the WAP may be closer to 1. While regional variation in the PQ would affect the statistical analysis, a constant bias in the quotient would not impact our conclusion. Another well-known issue is that the 16 S and 18 S reads are different from taxa-specific cell counts especially considering the variability of rDNA copy numbers per cell. However, total rDNA could possibly be an indicator of taxa-specific biomass as rDNA copy numbers generally increase with cell biovolumes across different phytoplankton species[47,48]. Our network and correlation analyses should also be interpreted with caution because 1) they do not necessarily reflect direct causal ecological relationships, 2) they may result from niche differentiation and interactions other than predation, including mutualism, competition, and parasitism, and 3) could in some cases be fortuitous or symptomatic of other abiotic and biotic factors not captured by our analyses, such as higher trophic level activity (e.g., macrozooplankton grazing), aggregation, and bacterial activity. For example, the low NCP/POC regions are mostly in the stable open ocean environments that are more homogeneous, resulting in less niche differentiation and more correlations between community members[44]. In contrast, more niche partitioning and fewer correlations are anticipated in high NCP/POC regions associated with the dynamic and heterogeneous conditions on the continental shelf and at the ice edge. Finally, a greater proportion of net community carbon production could be in the form of DOC in tightly coupled communities (in lower NCP/POC ecosystems). Overall, it will be important to further investigate these results and validate them with estimates of

in situ growth rates, photophysiology, microzooplankon growth and grazing rates, and relations to other biotic and abiotic components of the ecosystem.

Conclusion

This study explored the rich genomic complexity of polar ecosystems in the context of net community production, further untangling the intricate interplay of plankton community characteristics critical to carbon fluxes at the ocean surface by selecting through statistical inferences candidate key species and interactions. Within the diverse plankton communities of the WAP, as few as 2 to 3 key OTUs can explain a significant proportion of the spatial variability in carbon export potential during our observational period. Microzooplankton (i.e., dinoflagellates and Rhizaria) also arose as important negative predictors of carbon export potential, likely due to enhanced organic matter remineralization at the ocean surface. Using specific OTUs, as opposed to broad functional groups, substantially improved our ability to explain NCP variance from 35% to 92%, and NCP/POC variance from 16% to 76%. These results suggest that the PFT currency needs to be adjusted based on the biogeochemical process or element under study. As an example, diatoms require silica, and their biogeography is strongly delineated by Si availability. However, this functional grouping based on Si requirements does not necessarily translate into a cohesive group when pertaining to carbon export potential[39]. In our study, the most dominant group OTU_4 (i.e., *Fragilariopsis*) was found not to be a good predictor of net community production. Future studies of interannual variability in this and other regions of the Southern Ocean will be critical to assess the generality of our observed patterns. However, if our results are an accurate reflection of other regions of the Southern Ocean, functional group level studies should be interpreted with caution as the relation of organic carbon fluxes to taxonomy is more nuanced than currently depicted.

Materials and Methods

We derive NCP from the ratio of oxygen (O_2) to the inert gas argon (Ar). The biological O_2 supersaturation can be estimated from O_2/Ar because O_2 and Ar have similar solubility properties. The O_2/Ar method is suitable to explore the relation of NCP to plankton assemblages because the O_2 residence time in this study is comparable to plankton community turnover time. The residence time of O_2 at the ocean surface is calculated as the MLD divided by the piston velocity. For a mixed layer depth ranging from 11 to 22 m (mean 13 m) in sampled stations, with the gas exchange piston velocity calculated at each station using the weighted 60-day wind speed[49], the O_2 residence time in the mixed layer of 8 to 19 days (average of 11 days) (Supplementary Table S6) is within the 1 week-1 month range of phytoplankton community succession time span at the WAP[50]. We extracted total DNA from 21 stations; PCR amplified and sequenced the hypervariable V4 region of the 18 S marker gene. Based on the DNA phylogenetic information, we examined how community structure influences carbon export potential and efficiency across taxonomic ranks. In order to account for the biomass effect, NCP was normalized to POC to infer the carbon export efficiency[10,22].

Sampling. During the annual Palmer LTER cruise from Jan to early Feb in 2014 aboard the R/V *Laurence M. Gould*, discrete samples for POC, chlorophyll, DNA and RNA were collected at each station in the WAP region. Surface seawater from the vessel seawater supply line was gently vacuum filtered through a 0.2 µm pore size Millipore Supor filters. Each filter was then preserved immediately in 1 ml of RNA-later and stored at −80 °C. Surface seawater temperature, salinity, oxygen, fluorescence, transmissometer beam Cp and NCP were measured underway from the ship's clean flow-through seawater line system.

Underway $NCP_{(O2/Ar)}$ by EIMS. The biological O_2 supersaturation at the ocean surface was calculated as:

$$\Delta(O_2/Ar) = \left[\frac{(O_2/Ar)sample}{(O_2/Ar)sat} - 1\right] \times \ 100\% \tag{1}$$

ΔO_2/Ar was measured underway from the ship's flow-through seawater line by equilibrator inlet mass spectrometry (EIMS) as described in[9]. Assuming steady state and neglecting vertical mixing across the mixed layer boundary, NCP in units of mmol O_2 m^{-2} day^{-1} was calculated as:

$$NCP = k\rho[O_2]sat \ \Delta(O_2/Ar) \tag{2}$$

where ρ is mixed layer density (kg m^{-1}), and k is the gas transfer velocity for O_2 (m day^{-1}) estimated from NCEP/NCAR daily reanalysis winds (Supplementary Table S6) and a wind speed parameterization following[51] and a gas exchange weighting as described in[49]. Mixed layer depth was determined using a density threshold $\Delta\sigma\theta = 0.03$ kg m-3 from CTD profiles following[52] (Supplementary Table S6). NCP for O_2 was then converted to carbon (mmol C m^{-2} day^{-1}) using $O_2/C = 1.4$[53]. NCP is averaged over 10 min at the time point of water collection from the ship's flow-through seawater line.

A recent study has shown that the $NCP_{(O2/Ar)}$ measured in this region is highly correlated ($r^2 = 0.83$) with the NCP calculated from the seasonal DIC drawdown[54].

Assumption of steady-state. Under steady-state or quasi-steady-state, NCP approximates carbon export out of the mixed layer because the inventory change of organic carbon (POC + DOC) in the mixed layer is negligible compared to NCP. In the Southern Ocean, DOC production is a small fraction of NCP[55,56]. In order to assess whether the assumption of steady-state is valid in our study region, we estimated the change in POC inventory over the days prior to ship sampling. To that end, we first estimated the time-derivative of chlorophyll between the time of sampling and 8 days earlier using MODIS ocean color data at 8 days 9 km resolution. We then converted Chl to POC using an empirical relationship derived from *in situ* Chl and POC (bottle measurements) data

measured at the Palmer LTER (see Supplementary Fig. S3). This analysis resulted in only 4 comparisons because of cloud cover and should therefore be interpreted with caution. The estimated time-derivative of the POC inventory over the days prior to ship sampling was on average 18% of the NCP measurements, consistent with a major portion of the NCP being exported out the mixed layer. However, mismatch between NCP and carbon export production has been observed in other regions[57]. Under most conditions, NCP reflects an upper bound on carbon export, or an export potential[58].

DNA extraction and amplicon library construction. In the lab, from each sampled station one filter was first split in two halves, one for DNA extraction and the other for later RNA studies. 500 µl of the RNAlater from the original tube was loaded onto an Amicon 10 k column and the column was spun for 15 minutes at 14,000 g to get rid of most of the RNAlater and concentrate cells suspended in RNAlater. The resulting solution (~50 µl) was then added back to the one half filter. The cells on filter were lysed by bead-beating using 0.2 g of zirconium beads in 400 µl of buffer AP1 (Qiagen). DNA was extracted and purified using Qiagen DNeasy Plant Mini kit following the manufacture's instruction. 18 S (eukaryotic) rDNA gene fragments were amplified by PCR using V4 primer sets: 18 S forward (5′- CCAGCASCYGCGGTAATTCC-3′) and reverse (5′-ACTTTCGTTCTTGAT-3′), modified from[59] to avoid mismatch with haptophytes. Amplicons from each sample were tagged using dual index fusion primers[60] and a "heterogeneity spacer"[61] was used to improve sequencing quality. The "heterogeneity spacer" is a 0–5 bp of spacer added to the index sequence in order to allow different samples be sequenced out of phase and thus improve the amplicon library sequencing quality. Each PCR reaction (25 µl) consisted of 1U of Platinum Taq DNA Polymerase High Fidelity (Invitrogen), 1 × High Fidelity Buffer, 200 µM dNTPs, 2 mM $MgSO_4$, 0.2 µM each primer, and 5–30 ng extracted environmental DNA template. PCR was conducted with an initial activation step at 94 °C for 3 min, followed by 30 three-step cycles consisting of 94 °C for 30 s, 57 °C for 30 s and 72 °C for 1 min, and a final extension step of 72 °C at 10 min. PCR products in triplicates for each sample were then pooled and purified using QIAquick PCR Purification kit, and the concentration of amplicon DNA was quantified using a Qubit dsDNA assay. Amplicons from different locations were pooled in equimolar amounts (final concentration ~10 ng/µl) and submitted to Duke Institute for Genomic Sciences and Policy (IGSP) for sequencing using Illumina MiSeq. 300PE platform. The averaged read count per sample was 105,253 after demultiplexing.

Bioinformatic analysis. Paired-end reads were assembled using PANDAseq[62]. Assembled reads (mean length 424 bp) were further processed including quality filtering and chimera checking using USEARCH, and clustered into 97% OTUs following the UPARSE pipeline[63]. Representative sequences were aligned and assigned taxonomic identity based on the SILVA rRNA database[64] using the software package QIIME[65].

Stochastic search variable selection (SSVS). We searched for a subset of predictors, i.e. relative abundance of the top 20 OTUs, best fitting NCP or NCP/POC using Stochastic Search Variable Selection (SSVS)[66]. The predictor variables were mean-centered and variance-scaled (i.e. divided by standard deviation). If the dependent variable y_i (NCP or NCP/POC) is assumed to be a linear function of potential predictors x_i (relative abundance of OTUs), the SSVS can be set up as follow:

$$y_i|\beta,\ \sigma^2 \sim N(x_i\beta + \alpha,\ I\sigma^2) \tag{3}$$

where I is an $p \times p$ identity matrix, x_i is a $1 \times p$ vector, $\beta = (\beta_1, \beta_2, \ldots \beta_p)^T$ is a $p \times 1$ vector, and y_i, α and σ^2 are scalars. i and p represent sample index and number of variables, respectively. α, β and σ^2 are considered unknown, with β_j=0 indicating that variable $x_{i,p}$ is the nonselected predictor. Unknown β has prior distribution of $\left(\beta \sim \prod_{j=1}^{p}\left\{\pi\delta_0\left(\beta_j\right) + (1-\pi)N\left(\beta_j;\ 0,\ \psi_j^2\right)\right\}\right)$, where δ_0 denotes Dirac delta function that places all its mass at zero. Other parameters have prior distributions of $\pi \sim Beta(1,\ 1)$, $\psi_j^{-2} \sim Ga\left(a_j,\ b_j\right)$, $\sigma^{-2} \sim Ga(a,\ b)$, and $\alpha \sim N(\mu_0,\ \sigma_0^2)$, where a, b, a_j, b_j, μ_0, σ_0^2 are constants. The posterior distributions of parameters are approximated using a sequence of random samples from Markov Chain Monte Carlo (MCMC). The MCMC is run with 10000 iterations (It) to ensure convergence with the first 5000 samples discarded (burn-in = 5000). The probability of including predictor β_j is defined as $p\left(\beta_j\right) = \frac{1}{It-B}\sum_{k=B+1}^{It} 1\left(\beta_{j,k} \neq 0\right)$ (4), where B denotes burn-in. OTUs were ranked by $p(\beta_j)$ (Supplementary Table S2) and the top OTUs best prediction NCP and NCP/POC were selected based on a sharp drop of $p(\beta_j)$ after the selected OTUs.

Diatom isolation and identification. Diatoms were isolated from the Western Antarctic Peninsula along the PalmerLTER sampling grid in 2013 and 2014. Diatom species were identified by morphological characterization and 18 S rRNA gene (rDNA) sequencing. DNA was extracted with the DNeasy Plant Mini Kit according to the manufacturer's protocols (Qiagen). Amplification of the nuclear 18 S rDNA region was achieved with standard PCR protocols using eukaryotic-specific, universal 18 S forward and reverse primers. Primer sequences were obtained from[67] and are as follows: 18AF 5′-AACCTGGTTGATCCTGCCAGT-3′ and 18BR 5′-TGATCCTTCTGCAGGTTCACCTAC-3′. PCR products were purified using either QIAquick PCR Purification Kit (Qiagen) or ExoSAP-IT (Affymetrix) and sequenced by Sanger DNA sequencing (Genewiz). Sequences were edited using Geneious Pro 5.6.4 software and BLASTn sequence homology searches were performed against the NCBI nucleotide non-redundant (nr) database to determine species with a cutoff identity of 98%.

Growth conditions, physiological characteristics and biovolumes. Isolates were maintained at 4 °C in constant irradiance at intensities of either 10 µmol photons m^{-2} s^{-1} (low light) or 90 µmol photons m^{-2} s^{-1} (growth saturating light). Light intensities were chosen based on previous culture-based experiments[41,68–70]. Cultures were grown in the synthetic seawater medium, AQUIL, enriched with filter sterilized vitamin and trace metal ion buffer containing 100 µmol L^{-1} EDTA. The growth media also contained 300 µmol L^{-1} nitrate, 200 µmol

L^{-1} silicic acid and 20 μmol L^{-1} phosphate. Premixed Fe-EDTA (1:1) was added separately for total iron concentrations of either 1370 nmol L^{-1} (pFe19) or 3.1 nmol L^{-1} (pFe21.7) to achieve high iron and low iron media, respectively. All media preparation and subsampling were performed under a positive-pressure, trace metal clean laminar flow hood. Cultures were grown in acid-washed 28 mL polycarbonate centrifuge tubes (Nalgene) at least in triplicates ($n \geq 3$) and maintained in exponential phase by dilution. Specific growth rates of successive transfers were calculated from the linear regression of the natural log of *in vivo* chlorophyll a fluorescence using a Turner 10-AU fluorometer during 3–7 days at exponential phase. The error bars of specific growth rates (Fig. 3) were calculated as one standard error of the mean ($n \geq 3$). Growth rate comparison was perfomed using one-way ANOVA and the significance was determined using Holm-Sidak pairwise multiple comparison method.

To estimate biovolumes of each diatom species, frustules were viewed using an Olympus BX61 Upright Wide Field Microscope with the differential interference contrast imaging mode and a 60X/1.42 Oil PlanApo N objective lens. Valve apical length (AL), transapical width (TW), and pervalvar height (PH) dimensions were estimated with Scion Image software (http://scion-image.software.informer.com; June 2015). Diatom biovolumes were calculated according to[71] where pennate diatoms were modeled after elliptic prisms and centric diatoms were modeled after cylinders.

Network analysis. All OTUs were used in the initial dataset. The rare taxa, i.e. OTUs appearing less than three times in at least 20% of the samples, were removed from the dataset. The OTU abundances were then standardized to the median sequencing depth. The dataset was then divided into two equally sized subsets based on their NCP/POC values $\geq$ or $\leq$ the median NCP/POC. For each subset the correlations between OTUs and the p-values were computed using the SparCC method[42], which was specially designed for compositional data. The resulting correlations with magnitude >0.75 and with p-values >0.05 were selected to construct networks for each subset. The two networks were visualized using Cytoscape v3.4.0[72]. In Cyoscape, OTUs are represented as nodes. Correlations are displayed as edges with positive relationships in orange and negative ones in blue. Network properties were computed using the 'Network Analysis' function in Cytoscape, including diameter, density, clustering coefficient, characteristic path length, etc. The topological parameters for each node were also calculated, such as node degree (the number of edges linked to each node), betweenness centrality, and closeness centrality.

Other statistical analyses. Forward stepwise regression was performed in JMP statistical discovery software from SAS. To reduce the number of regressors in the stepwise regression, which increase the possibility of collinearity and over-fitting, only the top 20 most dominant OTUs were included in the analysis. The best model was selected based on Akaike's information criterion correction (AICc). AICc is suitable for small sample size (here $n = 21$), and it reduces over-fitting by greater penalty for increasing number of parameters included in the model. Even though significant in stepwise regression, some OTUs (Supplementary Table S3) were not included in the final multiple linear regression model because they were not selected by SSVS and each OTU only slightly improved the adjusted R^2 of the model (≤ 0.05).

Partition of variation was performed on selected variables vs. NCP or NCP/POC in R (R software, Vienna, Austria) using function "varpart" in the "vegan" package. Briefly, the unique (non-overlapping) or shared (overlapping) fractions of NCP or NCP/POC variance explained by OTUs were calculated by partial regressions and then decomposing the adjusted R^2 [73].

Data availability. Sequences and associated metadata have been deposited in the NCBI Sequence Read Archive under the accession number SRP119815. Other datasets are available from the corresponding author on reasonable request.

References

1. Alldredge, A. L. & Gotschalk, C. *In situ* settling behavior of marine snow. *Limnol. Oceanogr.* **33**, 339–351 (1988).
2. Guidi, L. *et al.* Relationship between particle size distribution and flux in the mesopelagic zone. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **55**, 1364–1374 (2008).
3. Armstrong, R. A., Lee, C., Hedges, J. I., Honjo, S. & Wakeham, S. G. A new, mechanistic model for organic carbon fluxes in the ocean based on the quantitative association of POC with ballast minerals. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **49**, 219–236 (2001).
4. Steinberg, D. K. & Landry, M. R. Zooplankton and the ocean carbon cycle. *Ann. Rev. Mar. Sci.* **9**, 413–444 (2017).
5. Arrigo, K. R. Marine microorganisms and global nutrient cycles. *Nature* **437**, 349–355 (2005).
6. Falkowski, P. G., Laws, E. A. & Barber, R. T. *et al.* In *Ocean Biogeochemistry: The Role of the Ocean Carbon Cycle in Global* Change (ed. MJR, F.) 99–121 (Springer, 2003).
7. Le Quere, C. *et al.* Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glob. Chang. Biol.* **11**, 2016–2040 (2005).
8. Michaels, A. F. & Silver, M. W. Primary production, sinking fluxes and the microbial food web. *Deep Sea Res. Part A. Oceanogr. Res. Pap.* **35**, 473–490 (1988).
9. Cassar, N. *et al.* Continuous high-frequency dissolved O2/Ar measurements by equilibrator inlet mass spectrometry. *Anal. Chem.* **81**, 1855–1864 (2009).
10. Cassar, N. *et al.* The relation of mixed-layer net community production to phytoplankton community composition in the Southern Ocean. *Global Biogeochem. Cycles* (2015).
11. de Vargas, C. *et al.* Eukaryotic plankton diversity in the sunlit ocean. *Science (80-.).* **348**, 1261605 (2015).
12. Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W. & Huse, S. M. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* **4**, e6372 (2009).
13. Moon-van der Staay, S. Y., De Wachter, R. & Vaulot, D. Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* **409**, 607–610 (2001).
14. Guidi, L. *et al.* Plankton networks driving carbon export in the oligotrophic ocean. *Nature* (2016).
15. Huang, K., Ducklow, H., Vernet, M., Cassar, N. & Bender, M. L. Export production and its regulating factors in the West Antarctica Peninsula region of the Southern Ocean. *Global Biogeochem. Cycles* **26** (2012).

16. Vernet, M. *et al*. Primary production within the sea-ice zone west of the Antarctic Peninsula: I—Sea ice, summer mixed layer, and irradiance. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **55**, 2068–2085 (2008).
17. Sarmiento, J. L., Gruber, N., Brzezinski, M. A. & Dunne, J. P. High-latitude controls of thermocline nutrients and low latitude biological productivity. *Nature* **427**, 56–60 (2004).
18. Ducklow, H. W. *et al*. Marine pelagic ecosystems: the west Antarctic Peninsula. *Philos. Trans. R. Soc. B Biol. Sci.* **362**, 67–94 (2007).
19. Vaughan, D. G. *et al*. Recent rapid regional climate warming on the Antarctic Peninsula. *Clim. Change* **60**, 243–274 (2003).
20. Turner, J. *et al*. Antarctic climate change during the last 50 years. *Int. J. Climatol.* **25**, 279–294 (2005).
21. Saba, G. K. *et al*. Winter and spring controls on the summer food web of the coastal West Antarctic Peninsula. *Nat. Commun.* **5**, (2014).
22. Lam, P. J. & Bishop, J. K. B. High biomass, low export regimes in the Southern Ocean. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **54**, 601–638 (2007).
23. Ducklow, H. W. *et al*. Particle flux on the continental shelf in the Amundsen Sea Polynya and Western Antarctic Peninsula. *Elementa* **3**, (2015).
24. Arrigo, K. R. *et al*. Phytoplankton community structure and the drawdown of nutrients and CO2 in the Southern. *Ocean. Science (80-.).* **283**, 365–367 (1999).
25. Becquevort, S. & Smith, W. O. Aggregation, sedimentation and biodegradability of phytoplankton-derived material during spring in the Ross Sea, Antarctica. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **48**, 4155–4178 (2001).
26. Bernard, K. S., Steinberg, D. K. & Schofield, O. M. E. Summertime grazing impact of the dominant macrozooplankton off the Western Antarctic Peninsula. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **62**, 111–122 (2012).
27. Ducklow, H. *et al*. The marine system of the Western Antarctic Peninsula. *Antarct. Ecosyst. An Extrem. Environ. a Chang. World* 121–159 (2012).
28. Calbet, A. & Landry, M. R. Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnol. Oceanogr.* **49**, 51–57 (2004).
29. Garzio, L. M., Steinberg, D. K., Erickson, M. & Ducklow, H. W. Microzooplankton grazing along the Western Antarctic Peninsula. *Aquat Microb Ecol* **70**, 215–232 (2013).
30. Gast, R. J. *et al*. Abundance of a novel dinoflagellate phylotype in the Ross Sea, Antarctica. *J. Phycol.* **42**, 233–242 (2006).
31. Gast, R. J., Moran, D. M., Dennett, M. R. & Caron, D. A. Kleptoplasty in an Antarctic dinoflagellate: caught in evolutionary transition? *Environ. Microbiol.* **9**, 39–45 (2007).
32. Torstensson, A. *et al*. Physicochemical control of bacterial and protist community composition and diversity in Antarctic sea ice. *Environ. Microbiol.* (2015).
33. Luria, C. M., Ducklow, H. W. & Amaral-Zettler, L. A. Marine bacterial, archaeal and eukaryotic diversity and community structure on the continental shelf of the western Antarctic Peninsula. *Aquat. Microb. Ecol.* **73**, 107–121 (2014).
34. Saito, H., Ota, T., Suzuki, K., Nishioka, J. & Tsuda, A. Role of heterotrophic dinoflagellate Gyrodinium sp. in the fate of an iron induced diatom bloom. *Geophys. Res. Lett.* **33**, (2006).
35. Sherr, E. B. & Sherr, B. F. Heterotrophic dinoflagellates: a significant component of microzooplankton biomass and major grazers of diatoms in the sea. *Mar. Ecol. Ser.* **352**, 187 (2007).
36. Thomsen, H. A., Buck, K. R., Bolt, P. A. & Garrison, D. L. Fine structure and biology of Cryothecomonas gen. nov.(Protista incertae sedis) from the ice biota. *Can. J. Zool.* **69**, 1048–1070 (1991).
37. Biard, T. *et al*. *In situ* imaging reveals the biomass of giant protists in the global ocean. *Nature* **532**, 504–507 (2016).
38. Behrenfeld, M. J. *et al*. Annual boom-bust cycles of polar phytoplankton biomass revealed by space-based lidar. *Nat. Geosci.* (2016).
39. Assmy, P. *et al*. Thick-shelled, grazer-protected diatoms decouple ocean carbon and silicon cycles in the iron-limited Antarctic CircumpolarCurrent. *Proc. Natl. Acad. Sci.* **110**, 20633–20638 (2013).
40. Mock, T. *et al*. Evolutionary genomics of the cold-adapted diatom Fragilariopsis cylindrus. *Nature* 1–5 (2017).
41. Arrigo, K. R. *et al*. Photophysiology in two major Southern Ocean phytoplankton taxa: photosynthesis and growth of Phaeocystis antarctica and Fragilariopsis cylindrus under different irradiance levels. *Integr. Comp. Biol.* **50**, 950–966 (2010).
42. Friedman, J. & Alm, E. J. Inferring correlation networks from genomic survey data. *PLoS Comput Biol* **8**, e1002687 (2012).
43. Dong, J. & Horvath, S. Understanding network concepts in modules. *BMC Syst. Biol.* **1**, 24 (2007).
44. Ma, B. *et al*. Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. *ISME J.* **10**, 1891 (2016).
45. Faust, K. & Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* **10**, 538 (2012).
46. Tortell, P. D. *et al*. Metabolic balance of coastal Antarctic waters revealed by autonomous pCO2 and ΔO2/Ar measurements. *Geophys. Res. Lett.* **41**, 6803–6810 (2014).
47. Zhu, F., Massana, R., Not, F., Marie, D. & Vaulot, D. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* **52**, 79–92 (2005).
48. Godhe, A. *et al*. Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. *Appl. Environ. Microbiol.* **74**, 7174–7182 (2008).
49. Reuer, M. K., Barnett, B. A., Bender, M. L., Falkowski, P. G. & Hendricks, M. B. New estimates of Southern Ocean biological production rates from O 2/Ar ratios and the triple isotope composition of O2. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **54**, 951–974 (2007).
50. Goldman, J. A. L. *et al*. Gross and net production during the spring bloom along the Western Antarctic Peninsula. *New Phytol.* **205**, 182–191 (2015).
51. Ho, D. T. *et al*. Measurements of air-sea gas exchange at high wind speeds in the Southern Ocean: Implications for global parameterizations. *Geophys. Res. Lett.* **33** (2006).
52. de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A. & Iudicone, D. Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res. Ocean.* **109** (2004).
53. Laws, E. A. Photosynthetic quotients, new production and net community production in the open ocean. *Deep Sea Res. Part A. Oceanogr. Res. Pap.* **38**, 143–167 (1991).
54. Eveleth, R. *et al*. Ice melt influence on summertime net community production along the Western Antarctic Peninsula. *Deep Sea Res. Part II Top. Stud. Oceanogr* (2016).
55. Hansell, D. A. & Carlson, C. A. Net community production of dissolved organic carbon. *Global Biogeochem. Cycles* **12**, 443–453 (1998).
56. Hansell, D. A. & Carlson, C. A. *Biogeochemistry of marine dissolved organic matter*. (Academic Press, 2014).
57. Estapa, M. L. *et al*. Decoupling of net community and export production on submesoscales in the Sargasso Sea. *Global Biogeochem. Cycles* (2015).
58. Li, Z. & Cassar, N. A mechanistic model of an upper bound on oceanic carbon export as a function of mixed layer depth and temperature. *Biogeosciences Discuss.* **2017**, 1–27 (2017).
59. Stoeck, T. *et al*. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* **19**, 21–31 (2010).
60. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).

61. Fadrosh, D. W. *et al*. An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* **2**, 1–7 (2014).
62. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* **13**, 31 (2012).
63. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–998 (2013).
64. Pruesse, E. *et al*. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **35**, 7188–7196 (2007).
65. Caporaso, J. G. *et al*. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
66. George, E. I. & Mcculloch, R. E. Variable Selection Via Gibbs Sampling. *J. Am. Stat. Assoc.* **88**, 881–889 (1993).
67. Medlin, L., Elwood, H. J., Stickel, S. & Sogin, M. L. The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* **71**, 491–499 (1988).
68. Petrou, K., Trimborn, S., Rost, B., Ralph, P. J. & Hassler, C. S. The impact of iron limitation on the physiology of the Antarctic diatom Chaetoceros simplex. *Mar. Biol.* **161**, 925–937 (2014).
69. Hoppe, C. J. M., Holtz, L., Trimborn, S. & Rost, B. Ocean acidification decreases the light-use efficiency in an Antarctic diatom under dynamic but not constant light. *New Phytol.* **207**, 159–171 (2015).
70. Strzepek, R. F., Maldonado, M. T., Hunter, K. A., Frew, R. D. & Boyd, P. W. Adaptive strategies by Southern Ocean phytoplankton to lessen iron limitation: Uptake of organically complexed iron and reduced cellular iron requirements. *Limnol. Oceanogr.* **56**, 1983–2002 (2011).
71. Hillebrand, H., Dürselen, C., Kirschtel, D., Pollingher, U. & Zohary, T. Biovolume calculation for pelagic and benthic microalgae. *J. Phycol.* **35**, 403–424 (1999).
72. Shannon, P. *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
73. Legendre, P. & Legendre, L. F. J. *Numerical ecology*. **24** (Elsevier, 2012).

Acknowledgements

This work was supported by NSF OPP-1043339 to Cassar, NSF PLR-1440435 to Ducklow (Palmer LTER) and NSF OPP-1341479 to Marchetti. This work was also in part (NC and YL) supported by the "Laboratoire d'Excellence" LabexMER (ANR-10-LABX-19) and co-funded by a grant from the French government under the program "Investissements d'Avenir". The authors thank Oscar Schofield for providing chlorophyll data, Duke IGSP for sequencing the DNA samples, and Kuan Huang for discussion. The authors are grateful to all the scientists, scientific support personnel and the officers and crew on ARSV L.M. Gould for collecting samples and shipboard assistance.

Author Contributions

N.C. conceived the study. Y.L., N.C. and A.M. designed the study. Y.L. and H.D. collected field samples and underway data. Y.L. processed the DNA samples. C.M. and A.M. conducted the isolation and culture experiments. Y.L. and Z.L. analyzed the data. Y.L. and N.C. wrote the manuscript with contributions from all other authors.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-14109-1.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017