PNAS
www.pnas.org

Supplementary Information for

Oceanographic boundaries constrain microbial diversity gradients in the South Pacific Ocean

Authors: Eric J. Raes[1,2,3*], Levente Bodrossy[3], Jodie van de Kamp[3], Andrew Bissett[3], Martin Ostrowski[4], Mark Brown[5], Swan L.S. Sow[3,6], Bernadette Sloyan[3], and Anya M. Waite[1,7]

1. Alfred Wegener Institute for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany
2. The Oceans Institute M047, University of Western Australia, Crawley, 6009 WA, Australia
3. CSIRO Oceans and Atmosphere, GPO Box 1538, Hobart, Tasmania, 7001, Australia
4. Macquarie University, Faculty of Science and Engineering, NSW 2109, Australia
5. School of Biotechnology and Biomolecular Sciences, University of NSW Kensington, Sydney, NSW 2052 Australia
6. Institute for Marine and Antarctic Studies, University of Tasmania, Hobart 7001, Tasmania, Australia
7. Universität Bremen, Department Biology-Chemistry, Bibliothekstraße, 1 D-28334, Bremen, Germany

* Corresponding author: eric.raes@csiro.au current address[3]

www.pnas.org/cgi/doi/10.1073/pnas.1719335115

This PDF file includes:

Supplementary text covering materials and methods:

1) Biophysical parameters
2) Nutrient analysis
3) DNA extraction protocol
4) Pigment analysis and Table S1
   4.1. Nanoplankton
5) Carbon assimilation rates
6) Richness and statistical analyses Table S2 and S3

Figures S1 to S9

Table S2-S3

References for SI

Supplementary Information Text

1. Biophysical parameters

A 36-Niskin bottle rosette was fitted with a Seabird SBE32 conductivity-temperature-depth (CTD) profiler, a SBE43 O_2 sensor, a Chelsea Aqua tracker Fluorometer and a Wetlabs C-Star™ transmissometer. Transmission was used as a proxy for particle concentration; areas with high particle concentrations denoted by a low transmission signal (1) (See figure S1b). Physical, biogeochemical data and metadata can be accessed through the GO-SHIP data portal (https://cchdo.ucsd.edu/search?q=GO-SHIP; voyage 096U20160426) and through the CSIRO data portal (https://www.cmar.csiro.au/data/trawler/index.cfm; voyage IN2016_V03). Mixed layer depths (MLD) were calculated according to Talley, Pickard, Emery and Swift (2), as $\Delta T = 0.2$ °C compared to a surface reference depth of 7.3 ± 4.5 m.).

2. Nutrient analysis

Silicate analysis (detection limits to 0.2 µmol L^{-1}) was based on Armstrong et al. (1967) with minor changes to the coil sequence as recommended in Practical Guidelines for the Analysis of Seawater see Aminot, Kérouel and Coverly (3). PO_4^{3-} was analysed via the AA3 method G-297-03 rev3 according to Murphy and Riley (4) with detection limits to 0.01 µmol L^{-1}, using 10mm flow cell and a LED lamp. NO_3^-/ NO_2^- were analysed per AA3 method G-172-96 rev15 according to Armstrong, Stearns and Strickland (5) and

Grasshoff, Kremling and Ehrhardt (6) with detection limits to 0.015 µmol L^{-1} using 10mm flow cells and a LED lamp. Unless elsewhere stated NO_3^- and NO_x are the sum of $NO_3^- + NO_2^-$. NH_4^+ was analysed per AA3 method G-327-05 rev4 according to Kérouel and Aminot (7) with detection limits to 0.015 µmol L^{-1}, using a JASCO fluorometer FP-2020. Certified reference materials (CRM) for SiO_4, PO_4^{3-}, NO_3^-, and NO_2^- in seawater, produced by KANSO Japan, were used in each nutrient analysis to ensure the accuracy of results. The Reference CRMs were run 4 times after the calibration standards. Accuracy was determined by comparing the new standard batch with the old and tracking to ensure the concentration was within 1% accuracy between batches. The GO-SHIP criteria (Hydes et al., 2010), reference section 5.3, specifies using 1-3 % of full scale (depending on the nutrient) as acceptable limits of accuracy.

3. Extraction Protocol for Sterivex Samples as modified in Appleyard, Abell and Watson (8)

Materials

Lysis buffer

200mM $NaH_2PO_42H_2O$ (monobasic)

200mM Na_2HPO_4 (dibasic) MW 142 142g/1L =1M 5.68g/200mL = 200mM

To make up 200mL lysis buffer

39mL 200mM NaH_2PO_4

61mL 200mM Na_2HPO_4

17.54g NaCl

2g CTAB

4g PVP K30 + ddH20 to make up to 200ml

Adjust to pH 7.0 (using NaOH – try couple of mL of 10M NaOH)

Lysozyme

Proteinase K – 20mg/ml

From FastDNA™ Spin Kit for Soil (MP Biomedicals)

MT buffer From PowerWater® Sterivex™ DNA Isolation Kit (Qiagen)

Columns and sample recovery tubes, 3ml and 20ml syringes, 5ml tubes

Buffer ST4 (warmed to 65 °C before use)

Buffer ST5 and ST6

Inlet and outlet caps for Sterivex filters

Phenol:Chloroform:Isoamyl (25:24:1) (PCI)

Chloroform:Isoamyl (24:1) (CI)
TE buffer

Protocol

1. weigh 125 mg lysozyme into 50 mL falcon tube and add 25 mL Lysis Buffer to dissolve (lysozyme final conc. 5 mg/ml).
2. remove filters from -80 °C, remove inlet cap and using a pipette add 1.875 ml Lysis buffer (containing 5mg/ml final concentration of lysozyme) and 0.125 ml MT buffer.
3. recap the Sterivex filter and attach filter (with inlet end facing out) to the horizontal vortexer, Speed 6 for 60 min (turning the filter a couple of times during the hr)
4. using 3 ml syringe, draw back plunger and attach to inlet end of filter – release plunger and buffer in filter should flow into syringe. Divide approx 2 ml of buffer evenly into 2 × 2.0 ml tubes
5. in fume hood, add 900 µl PCI to each tube, invert several times, spin down 13000/10 min/RT
6. combine the aqueous phases from both tubes into one 2.0 ml tube (~ 1.5 – 1.8 ml), add 20 µl Prot K, and incubate on heat block for 2 hr at 60 °C
7. in fume hood, add 500 µl CI, spin down 13000/10 min/RT – remove aqueous phase to new tube
8. in fume hood, add 500 µl CI, spin down 13000/5 min/RT – remove aqueous phase to new tube
9. after 2nd spin, remove 1 ml of aqueous phase, add to 5 ml tube
10. add 3 ml of warmed ST4 buffer (65 °C), mix by inversion
11. attach column to barrel of 20 ml syringe and attach to vacuum manifold
12. pour contents of 5 ml tube into barrel while still warm
13. using vacuum, pull contents through the column
14. while keeping column attached to the manifold, remove barrel and add 800 µl ST5 to column
15. using vacuum, pull contents through the column
16. add 800 µl ST6 to column
17. using vacuum, pull contents through the column, then keep on vacuum for 2 mins
18. turn vacuum off, put column into new 2.0 ml elution tube and let air dry on bench for 10 mins
19. add 80 µl 0.1 x TE to column, incubate at 37 °C for 45 min
20. spin down column and tube at 13000/2 min/RT to elute DNA
21. Quantify 2 µl on the Nanodrop and record concentration and 260:280 ratio in spreadsheet

Storage and Downstream Applications

1. Transfer DNA (total volume approx. 78 µl) to 96 well plate and record well location of each sample in spreadsheet.
2. Using multi-channel pipette, transfer aliquots to further 96 well plates (1 x amplicon sequencing, 1 x shotgun metagenomics, then 5 x 10 µl, 2 x 5 µl, and remainder for long term storage at -80 °C).

3. Plates for long term storage are dried down in vacuum centrifuge at 30 °C. Seal with AlumaSeal CS™ foil film. Place plates in for storage at -80 °C (Environmental Genomics ultra freezer, CSIRO Hobart).
4. Plates for amplicon sequencing and shotgun metagenomics are sent to Ramaciotti Centre for Genomics (on dry ice).

4. Pigment analysis

Chlorophyll *a* extractions were carried out according to Parsons, Takahashi and Hargrave (9) on 0.525 L of sample water, using 25-mm GF/F filters through gentle vacuum filtration (pressure drop <10 kPa) at 5 sampling depths within the mixed layer. Samples were measured on a Turner Trilogy laboratory fluorometer (See figure S1a).

Four liters of sample water from the surface at each station were filtered in dim-lit light on 25 mm Whatman GF/F filters for analysis of taxonomically significant chlorophylls and carotenoids. Samples were snap frozen and stored in liquid N. Photosynthetic pigments were analyzed using high-performance liquid chromatography (HPLC) according to Hooker, *et al.* (10). HPLC data were analysed using diagnostic pigments of dominant phytoplankton functional guilds as well as size classes according to Hirata, *et al.* (11). Phytoplankton size classes (microplankton, nanoplankton and picoplankton) were calculated according to Hirata, *et al.* (11); See table S1. Processed HPLC data were analysed using diagnostic pigments of dominant phytoplankton functional guilds as well as size classes according to Hirata, *et al.* (11). Our data showed that: (1) the total chl a (TChl a) made up at least 70% of the total pigment concentration and (2) we had a regression between TChl a and the accessory pigments with a slope of 1.1 and $r^2 > 0.9$. ∑DP is the sum of all diagnostic pigments and is equal to 1.41 Fuco (Fucoxanthin) + 1.41 Perid (Perdidin) + 1.27 Hex-fuco (19'-hexanoyloxyfucoxanthin) + 0.6 Allo (Alloxanthin) + 0.35 But-fuco (19′-but-fucoxanthin) + 1.01 Chl-b (Chlorophyll b) + 0.86 Zea (Zeaxanthin) = Chl-a (Chlorophyll a) (12). The calculated size classes were microplankton (20 – 200 µm), nanoplankton (2 – 20 µm) and picoplankton (0.2 – 2 µm) (13). We note that diagnostic pigment analysis has its ambiguities (e.g. fucoxanthin is a precursor for 19'-butanoyloxyfucoxanthin and 19'-hexanoyloxyfucoxanthin); for more

detail, see Aiken, *et al.* (14) and Uitz, Claustre, Morel and Hooker (12). Also the size of diatoms can range between the micro and nanoplankton size fraction. We note that some diagnostic pigments are shared by several phytoplankton groups and some groups may cover a broad size (see also Kheireddine, *et al.* (15)). Nevertheless, this approach has shown to give valuable insights in the dominant trends of the phytoplankton community and size structure at the regional and seasonal scales (16).

Table S1: Formulas used for calculation of phytoplankton functional guild and size class fractions based on diagnostic pigments after Hirata et al. 2011.

Reference	Phytoplankton Size Class / Functional Type	Diagnostic Pigments	Estimation Formula
Hirata et al. 2011	Diatoms	Fucoxanthin (Fuco)	1.41*Fuco / ∑DP
	Dinoflagellates	Peridinin (Perid)	1.41*Perid / ∑DP
	Microplankton	Fuco, Perid	1.41*(Fuco + Perid) / ∑DP
	Nanoplankton	Hex, Chl-*b*, But-fuco, Alloxanthin (Allo)	(1.27*Hex + 1.01*Chl-*b* + 0.35*But-fuco + 0.6*Allo) / ∑DP
	Picoplankton	Zea, Hex, Chl-*b*	(0.86*Zea + 1.27*Hex) / ∑DP

The HPLC data set are available at https://doi.pangaea.de/10.1594/PANGAEA.884052 (17).

4.1 Nanoplankton

We found that nanoplanktonic guilds (prymnesiophytes in particular) were present in all oceanic provinces. The proportion of nanoplankton within the community varied between 15-20 % at 60 °S and in the equatorial region, but up to 65 % at 55 °S (near the STF). The sharp increase in prymnesiophytes coincided with the abrupt disappearance of diatoms, while the minima corresponded to maximal microplankton and picoplankton abundances, in the Southern Ocean and equatorial regions respectively. Both the prymnesiophytes and the pico-eukaryotes contributed up to 83% to the phytoplankton guild fraction within the STF. The overall prevalence of prymnesiophytes reaffirms that

this group could be an important player in CO_2 fixation, as mentioned by Jardillier, Zubkov, Pearman and Scanlan (18) in subtropical and tropical ecosystem. Our data highlights that the prymnesiophytes are even present south of the STF.

5. Carbon assimilation rates

At each CTD station, water samples to measure C-assimilation rates were taken from the clean underway flow through system (intake at 6m). Triplicate C-assimilation experiments were initiated by adding 20 μmol L^{-1} of $NaH^{13}CO_3$ to polycarbonate one-liter incubation bottles. All incubation bottles were acid rinsed three times, rinsed two times with deionized water, and rinsed three times with seawater directly from the sample point prior to incubation. Polycarbonate bottles were placed in on-deck incubators where temperature regulation was maintained by a continuous surface seawater flow. After 24 h incubation the experiments were terminated by filtering each bottle (pressure drop <10 kPa) through a 25 mm precombusted GF/F filter. The natural abundance of the particulate organic carbon (POC), used as t-zero values to calculate assimilation rates, were obtained by filtering four liters of water onto pre-combusted GF/F filters for each station. All filters, enriched and natural abundance, were snap frozen in liquid N and stored at -80°C. Filters were later acidified and dried overnight at 60 °C. Determination of total C and $\delta^{13}C$ was carried out using a continuous flow system consisting of a SERCON 20-22 mass spectrometer connected with an Automated C analyser at the Stable Isotope Facility of the University of California Davis. The external error of analyses (1 std dev) was 0.15 permille (‰) for $\delta^{13}C$. C assimilation rates (nmol C L^{-1} h^{-1}) were calculated following the equations described in Knap, Michaels, Close, Ducklow and Dickson (19). Primary productivity are available at https://doi.pangaea.de/10.1594/PANGAEA.884052 (17).

6. Richness and statistical analyses

All analyses utilised sub-sampled OTU tables as described above. Pro- and eukaryotic richness were calculated as the number of 97% OTUs observed per sample (20). Absolute latitudinal ranges of OTUs were calculated as the difference between the maximum and

minimum latitudes where an OTU was present (eg., an OTU occurring at 50° S, 40° S and 30° S would have an absolute latitudinal range of 20° (21). Correlations were calculated using the Spearman coefficient r_s as a measure of rank correlation using the Corrplot and Vegan packages in R studio (22), see Figs. S2 and Fig. S3. In contrast to the Pearson coefficient, it assesses monotonic relationships, not purely linear relationships. Multiple error rates were corrected using Holm' method (23). Productivity and richness relationships were assessed with linear, cubic and quadratic regressions.

Boosted regression trees

Boosted regression trees Predictor variable influence on pro and eukaryotic richness was assessed using boosted regression trees (see Elith, Leathwick and Hastie (24)). Through computerised boosting, BRT optimises model predictive performance through forward step-wise fitting of many simple models, combining their predictive capability to give more weight to residual variation that remains unexplained (Elith et al., 2008). Like regression trees, BRT is robust to missing and unbalanced data, outliers and non-linear relationships (24). Cross-validation (cv correlation in the manuscript Tables 1 and 2) is a technique to evaluate how predictive the model is by partitioning the original data into a training set to train the model, and a test set to evaluate it. In k-fold cross-validation, the original sample is randomly partitioned into k equal size subsamples. The tenfold cross validation correlation value is comparable to an r^2 value. See also 'A working guide to boosted regression trees' (24). We did not consider day lengths as a predictor in our BRT models as these increased linerarly with latitude (see Fig. S9). BRT was implemented in the R software environment using the gbm.step function, a guassian error structure, 10 fold cross-validation and the following settings which resulted in models with > 2000 trees in all cases; learning.rate = 0.001, tree.complexity = 10, bag.fraction = 0.5 (24).

Supplementary Table S2.

Archaeal richness - MLD		Bacterial richness - MLD		Eukaryote richness - MLD	
Predictor variable	Relative influence (%)	Predictor variable	Relative influence (%)	Predictor variable	Relative influence (%)
Temperature	21.4	Temperature	24.3	Si	34.7
PP	14.1	PP	13.4	NO_3^-	12.1
NO_2^-	13.6	NO_2^-	12.5	Temperature	11.1
Salinity	9.9	MLD	10.1	TChl_a	7.4
Si	8.5	Salinity	8.2	PO_4^{3-}	7.3
NO_3^-	7.8	NO_3^-	7.6	NH_4^+	5.9
MLD	7.2	Si	6.9	MLD	5.5
TChl_a	6.7	NH_4^+	6.5	NO_2^-	5.1
NH_4^+	6.1	TChl_a	6.0	PP	4.6
Oxygen	3.0	Oxygen	2.7	Salinity	4.4
PO_4^{3-}	1.8	PO_4^{3-}	1.8	Oxygen	1.9
training data correlation	0.959	training data correlation	0.934	training data correlation	0.896
cv correlation ± se	0.844±0.022	cv correlation ± se	0.826±0.038	cv correlation ± se	0.793±0.04

Table S2. Percentage relative influence of 11 predictor variables in explaining variation in pro- and eukaryotic richness from samples just below the mixed layer depth (n=89) along the p15S GO-SHIP transect determined by BRT (see also Elith, Leathwick and Hastie (24)). Predictor variables: Primary productivity of the surface layer (PP); Mixed Layer Depth (MLD); NOx (NO_3^- + NO_2^-); Note: phytoplankton size classes and taxonomic groups derived from HPLC analysis are not included as these were obtain from surface samples only.

Non multi-dimensional scaling (nMDS) and Analysis of Similarities (ANOSIM)

Table S3: Tests for differences between unordered Oceanographic Biomses.

ANOSIM – OCEANOGRAPHIC BIOMES **16S - BACTERIA**

PAIRWISE TESTS

GROUPS	**R Statistic**	**Significance Level %**
SO, STF	0.844	0.1
SO, TROP. OLIGO	0.997	0.1
SO, EQ. UPWELLING	1	0.1
STF, TROP. OLIGO	0.827	0.1
STF, EQ. UPWELLING	0.981	0.1
TROP. OLIGO, EQ. UPWELLING	0.344	0.1

ANOSIM – OCEANOGRAPHIC BIOMES **16S - ARCHAEA**

PAIRWISE TESTS

GROUPS	**R Statistic**	**Significance Level %**
SO, STF	0.828	0.1
SO, TROP. OLIGO	0.993	0.1
STF, TROP. OLIGO	0.821	0.1
SO, EQ. UPWELLING	1	0.1
STF, EQ. UPWELLING	1	0.1
TROP. OLIGO, EQ. UPWELLING	0.216	0.2

ANOSIM – OCEANOGRAPHIC BIOMES **18S - EUKARYOTES**

PAIRWISE TESTS

GROUPS	**R Statistic**	**Significance Level %**
SO, STF	0.824	0.1
SO, TROP. OLIGO	0.984	0.1
SO, EQ. UPWELLING	0.991	0.1
STF, TROP. OLIGO	0.731	0.1
STF, EQ. UPWELLING	0.882	0.1
TROP. OLIGO, EQ. UPWELLING	0.598	0.1

Groups: SO = Southern Ocean; STF = Subtropical Front; Trop. Oligo = Tropical Oligotrophic; Eq. upwelling = Equatorial upwelling.

ANOSIM provides a way to test statistically whether there is a significant difference between two or more groups of sampling units. If two groups of sampling units are really different in their species composition, then compositional dissimilarities between the groups ought to be greater than those within the groups. The ANOSIM statistic R is based on the difference of mean ranks between groups (r_B) and within groups (r_W).

$$R = (r\_B - r\_W)/(N\ (N-1)\ /\ 4)$$

The divisor is chosen so that R will be in the interval -1 ... +1, value 0 indicating completely random grouping. Note the significance level is express as %, to obtain the actual P value you need to dive by 100. For more information see Clarke and Warwick (25).

Supplementary Figures

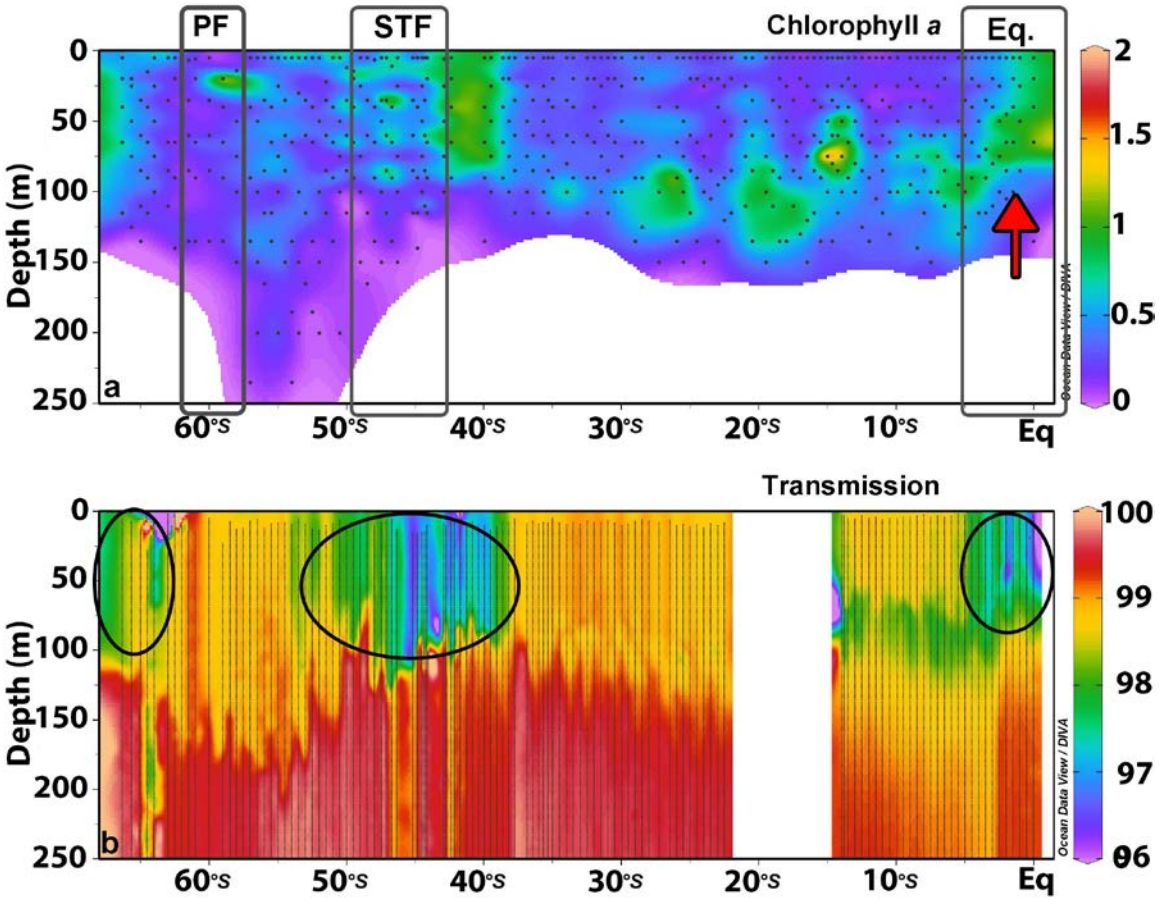

Fig. S1: a) Chlorophyll *a* concentrations (in µg L^{-1}) along the P15S GO-SHIP transect, **b)** Transmission (in %), used as a proxy for particle concentration (1) along the P15S GO-SHIP transect. Black dots on **a)** denote sampling depths for chlorophyll *a* measurements. The polar front, subtropical front and the equatorial upwelling provinces are highlighted by grey rectangles. The red arrow highlights upwelling of colder, relatively nutrient richer waters. Thin grey lines on **b)** denote sampling stations, and black circles highlight low transmission signals (= areas with high particle concentrations).

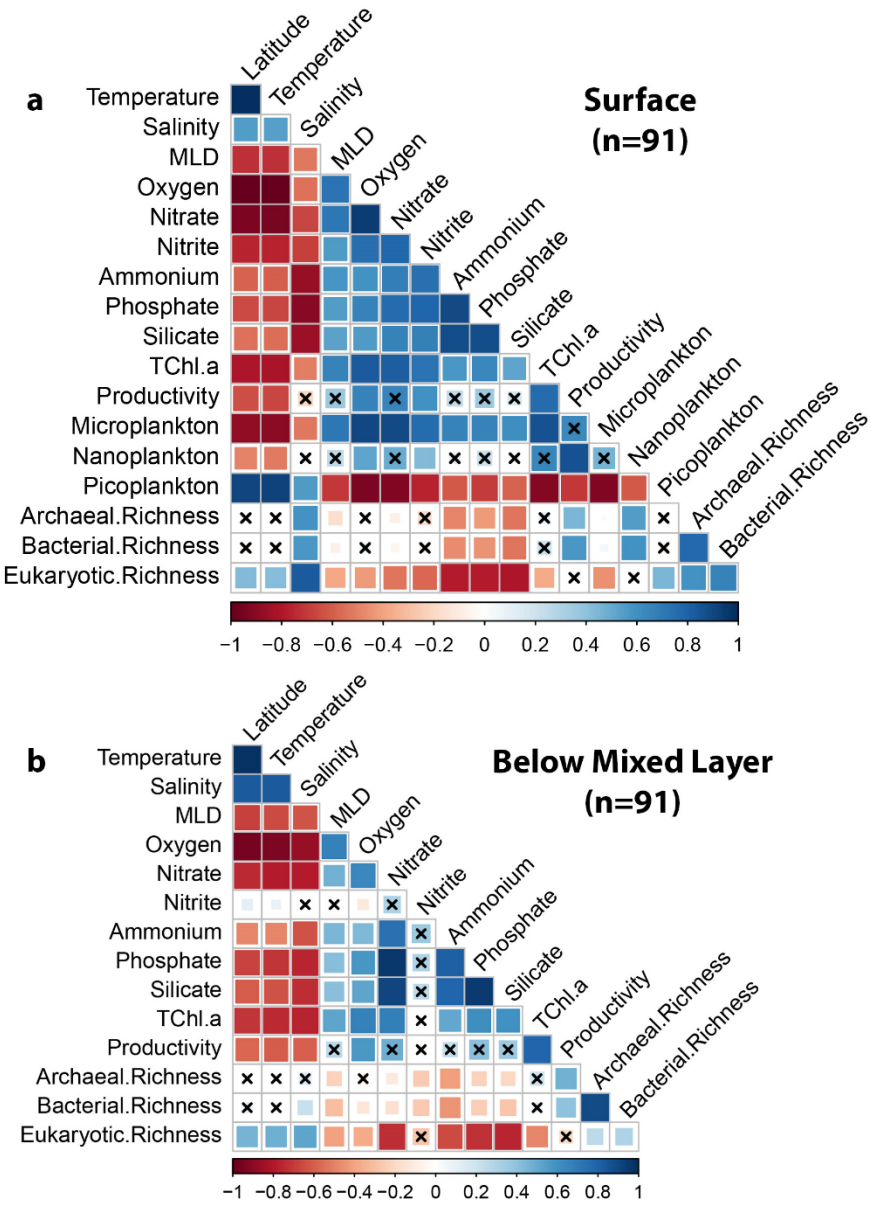

Fig. S2: Spearman correlation coefficients (rs) plots for pro-and eukaryotic richness and biogeochemical parameters. Top figure **a)** shows the samples from the surface waters and bottom figure **b)** shows the samples just below the mixed layer depth. Phytoplankton size classes derived from HPLC analysis are denoted: Microphytoplankton (20 – 200 µm), Nanophytoplankton (2 – 20 µm) and Picophytoplankton (0.2 – 2 µm). Positive correlations are displayed in blue and negative correlations in red. Colour intensity and the size of the squares are proportional to the correlation coefficients. P values > 0.05 are indicated by black crosses. Note primary productivity values used in the correlation plot are averages from three replicates.

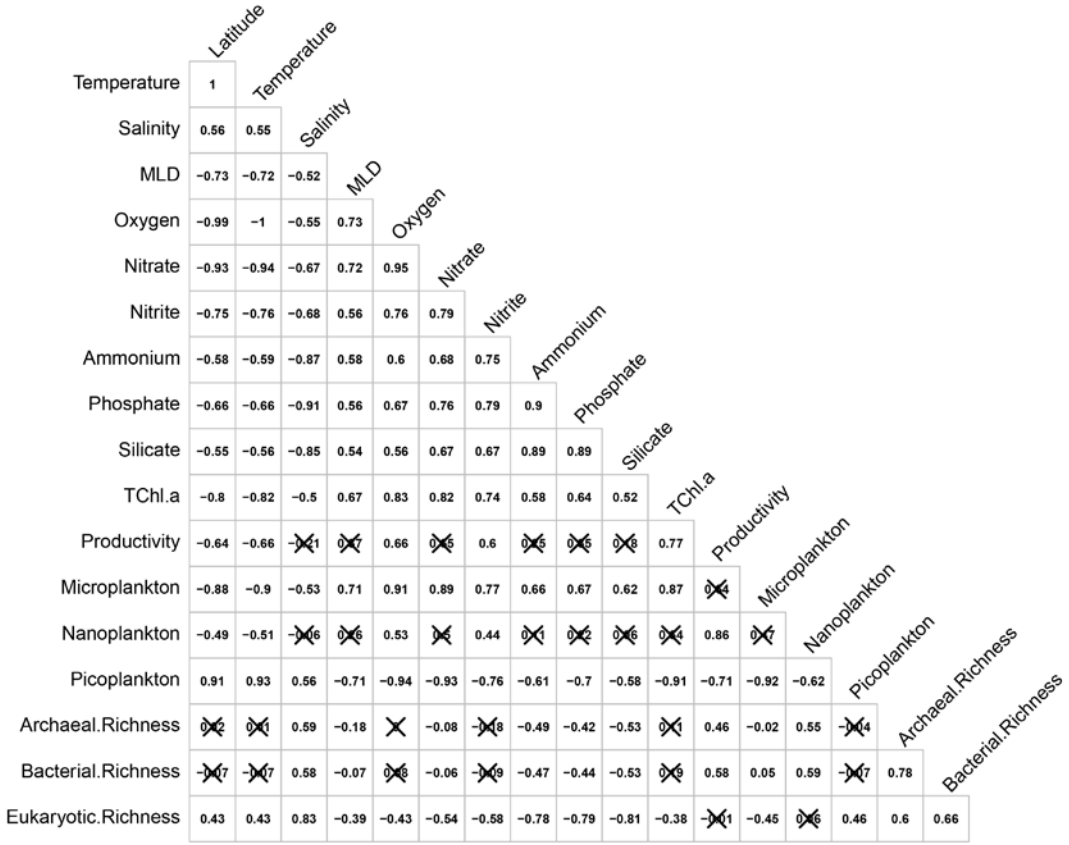

Fig. S2c: Spearman correlation coefficients (rs) plot for pro-and eukaryotic richness and biogeochemical parameters from the surface waters P values > 0.05 are indicated by black crosses. Note primary productivity values used in the correlation plot are averages from three replicates.

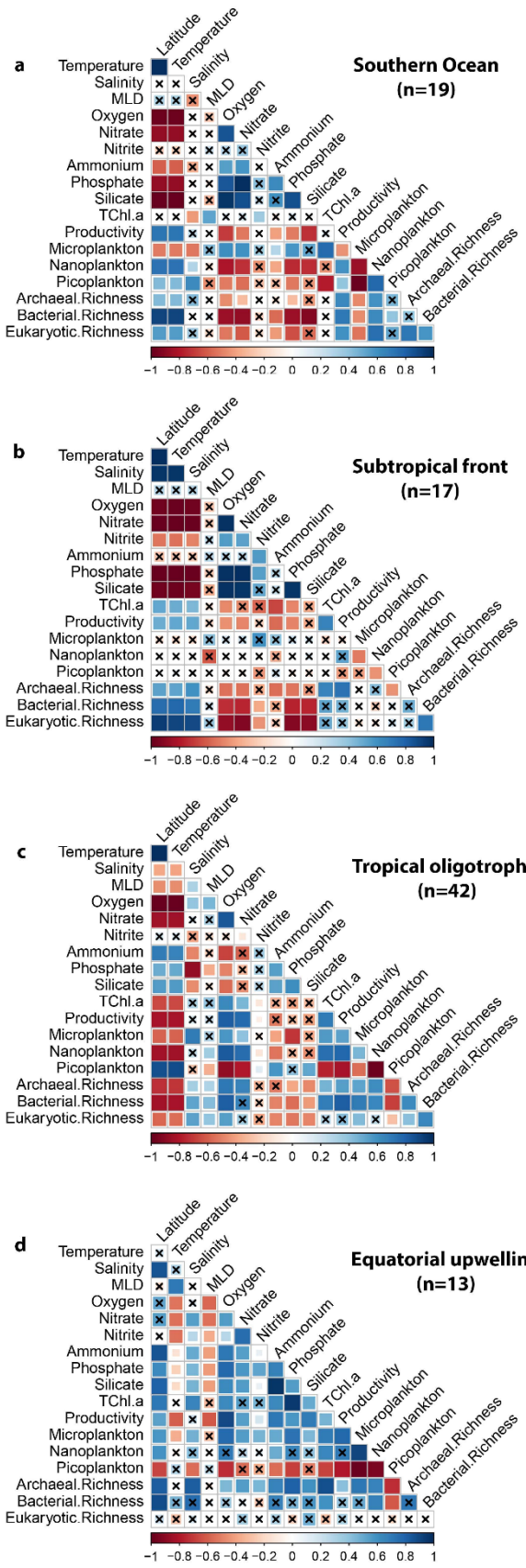
a
Southern Ocean
(n=19)
b
Subtropical front
(n=17)
c
Tropical oligotrophic
(n=42)
d
Equatorial upwelling
(n=13)
Latitude
Temperature
Salinity
MLD
Oxygen
Nitrate
Nitrite
Ammonium
Phosphate
Silicate
TChl.a
Productivity
Microplankton
Nanoplankton
Picoplankton
Archaeal.Richness
Bacterial.Richness
Eukaryotic.Richness
−1 −0.8 −0.6 −0.4 −0.2 0 0.2 0.4 0.6 0.8 1

Fig. S3 Spearman correlation coefficients (r_s) plots for pro-and eukaryotic richness and biogeochemical parameters for four different oceanic provinces in the surface waters of the South Pacific Ocean. **a)** Southern Ocean, **b)** Subtropical front, **c)** Tropical oligotrophic and **d)** Equatorial upwelling. Phytoplankton size classes derived from HPLC analysis: Microplankton (20 – 200 µm), Nanoplankton (2 – 20 µm) and Picoplankton (0.2 – 2 µm). Positive correlations are displayed in blue and negative correlations in red. Colour intensity and the size of the squares are proportional to the correlation coefficients. P values > 0.05 are indicated by black crosses. Note primary productivity values used to make the correlation plots are averages from three replicates.

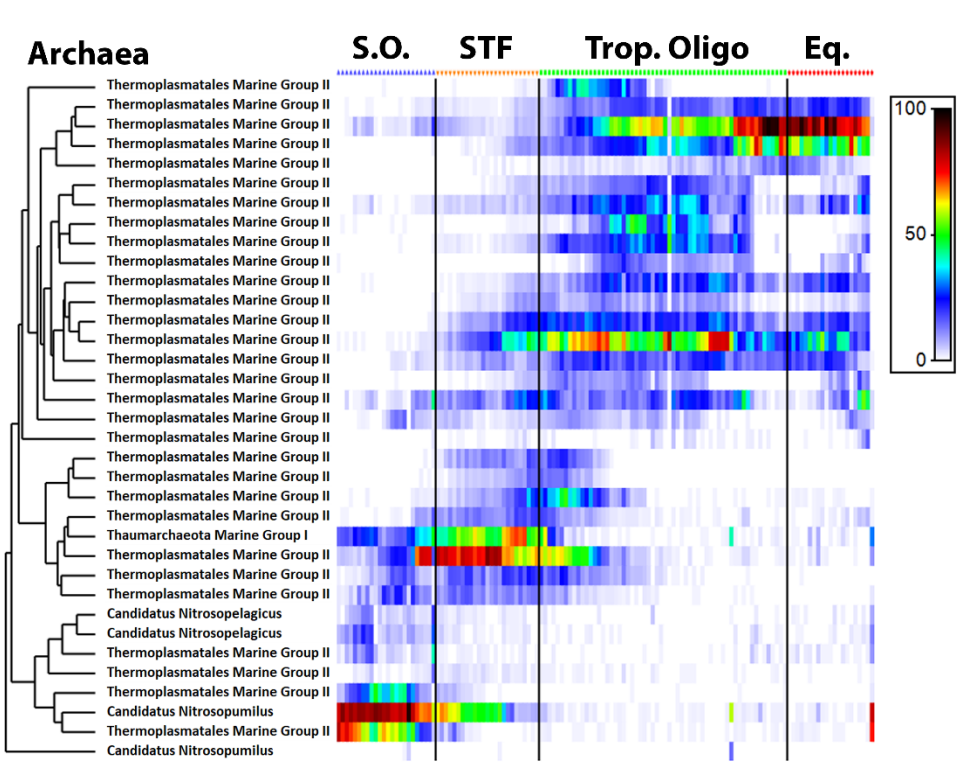

Fig. S4 a) Heat map of the archaeal Domain. Most abundant sequences of the OTU matrix are displayed and sequence reads for both archaea and bacteria are square root transformed. Colour bar denotes sequence abundance. Dendogram is not a phylogentic tree, but illustrates the clustering of strongly correlated OTUs. Blue, organge, green and red symbols above the heatmaps denote CTD stations in their respecitive oceanogrpahic provinces.

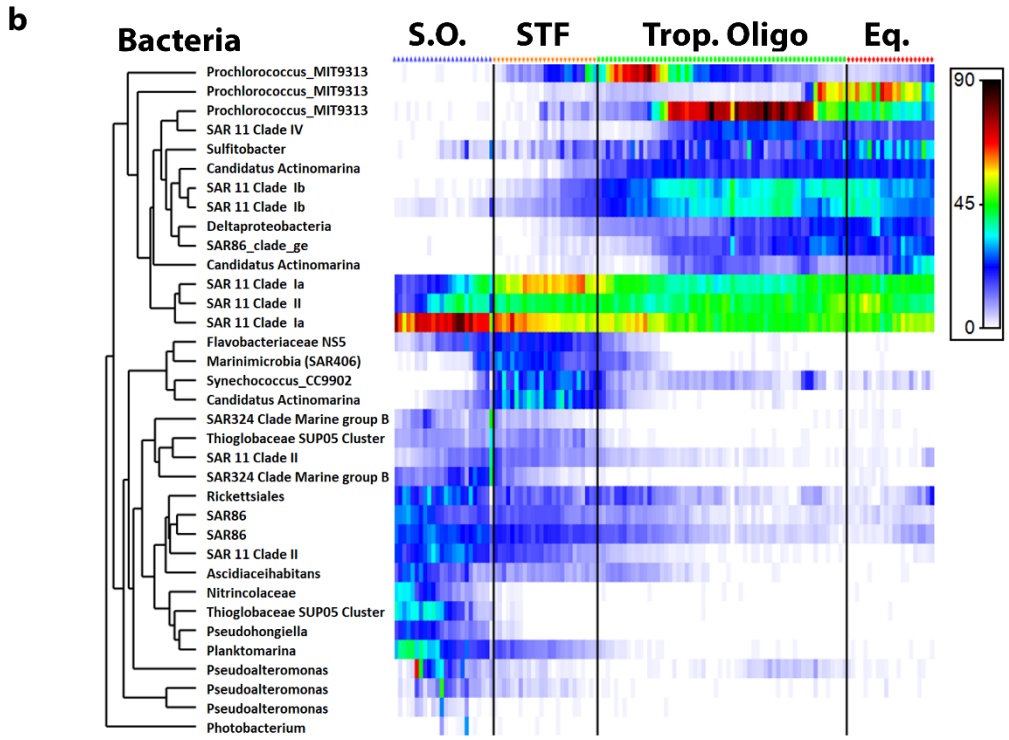

Fig. S4 b) Heat map of the bacterial Domain. Most abundant sequences of the OTU matrix are displayed and sequence reads for both archaea and bacteria are square root transformed. Colour bar denotes sequence abundance. Dendogram is not a phylogentic tree, but illustrates the clustering of strongly correlated OTUs. Blue, organge, green and red symbols above the heatmaps denote CTD stations in their respecitive oceanogrpahic provinces.

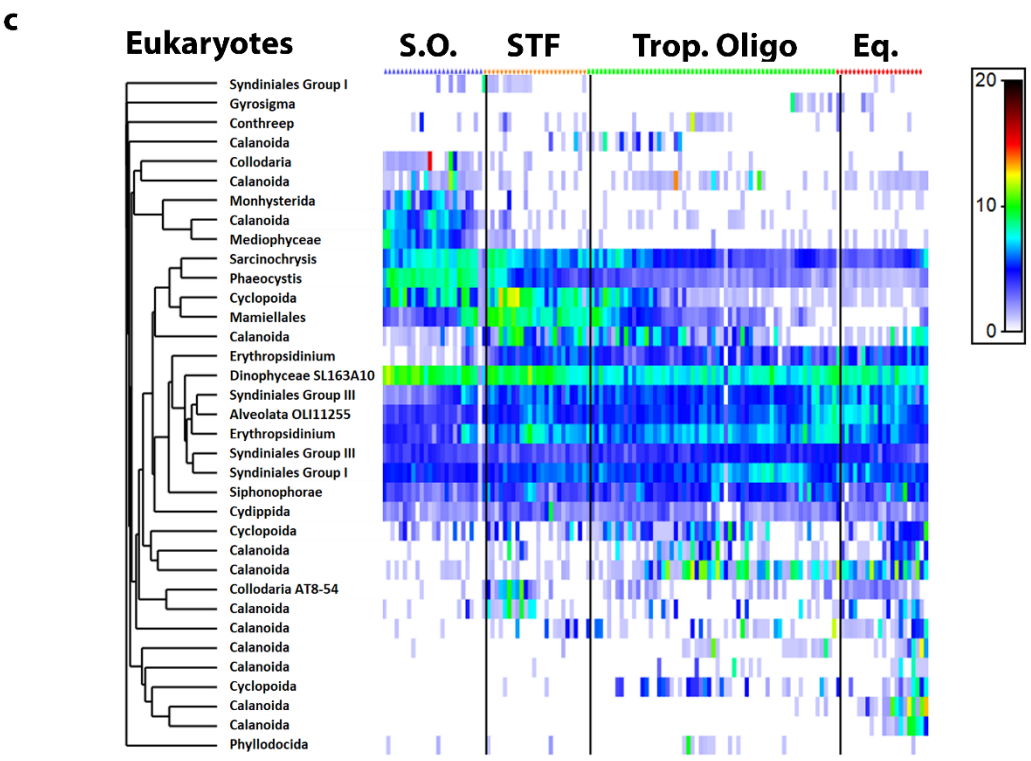

Fig. S4 c) Heat map of the eukaryotic Domain. Most abundant sequences of the OTU matrix are displayed and sequence reads for both archaea and bacteria are square root transformed. Colour bar denotes sequence abundance. Dendogram is not a phylogentic tree, but illustrates the clustering of strongly correlated OTUs. Blue, organge, green and red symbols above the heatmaps denote CTD stations in their respecitive oceanogrpahic provinces.

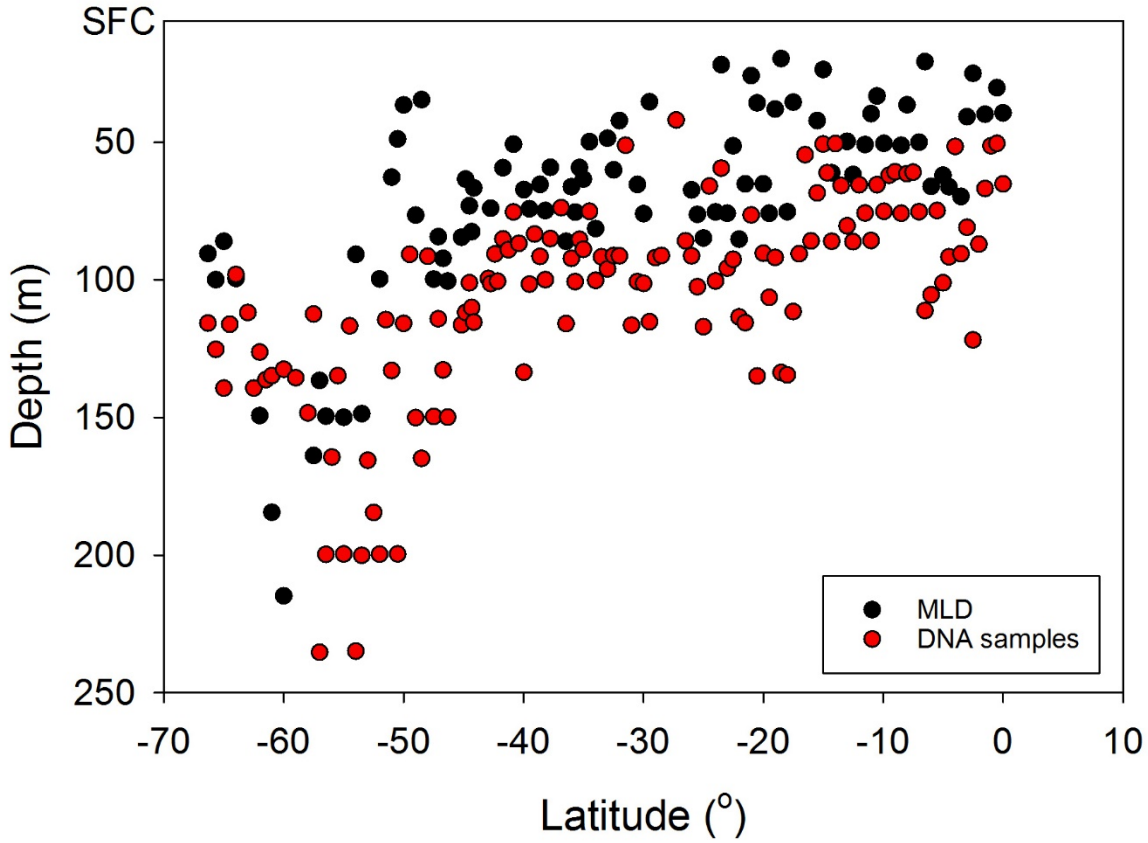

Fig. S5. Mixed layer depths were calculated according to Talley, Pickard, Emery and Swift (2), as a $\Delta T = 0.2$ °C compared to a surface reference depth. Mixed layer depths are denoted in black circles on the graph. DNA sampling depths are highlighted in red circles.

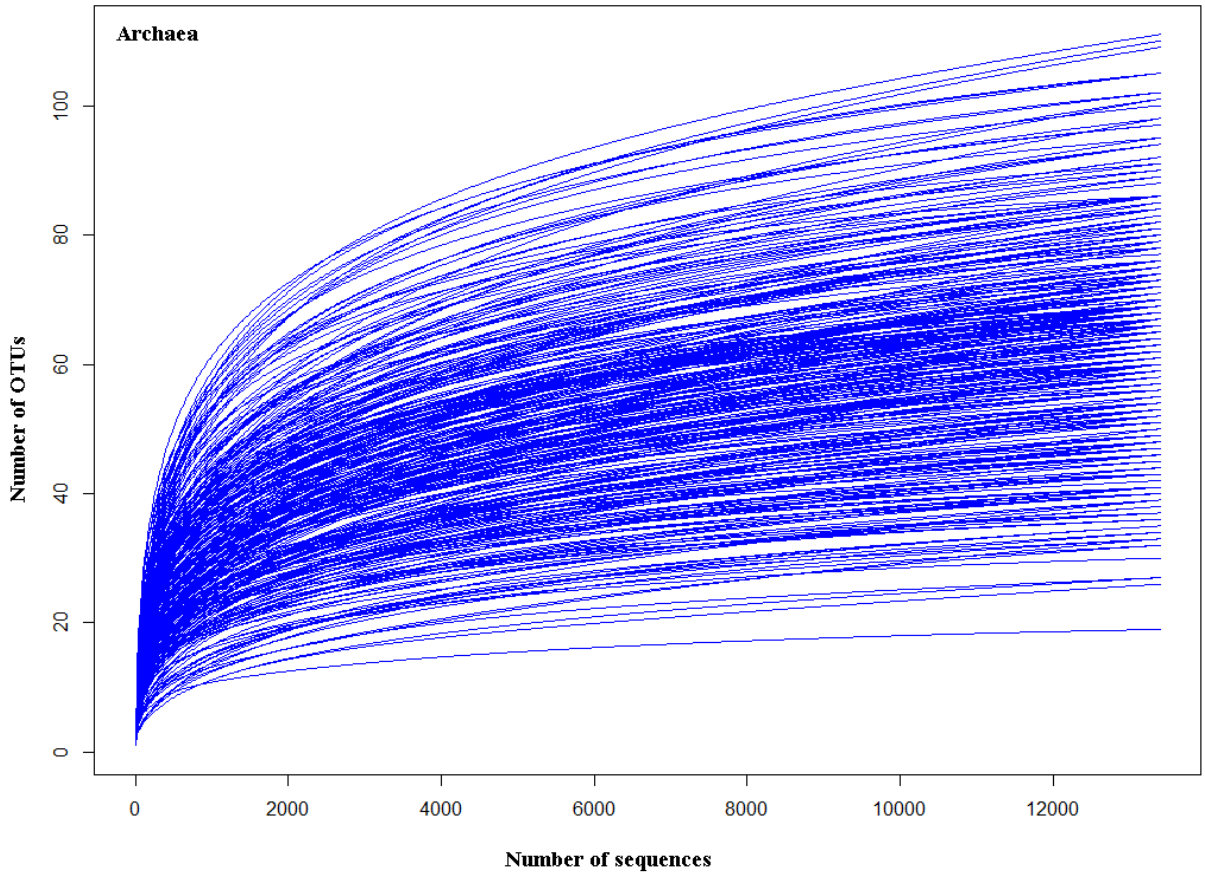

Fig. S6: Rarefaction curves for Archaea, subsampled to a depth of 13400 sequences.

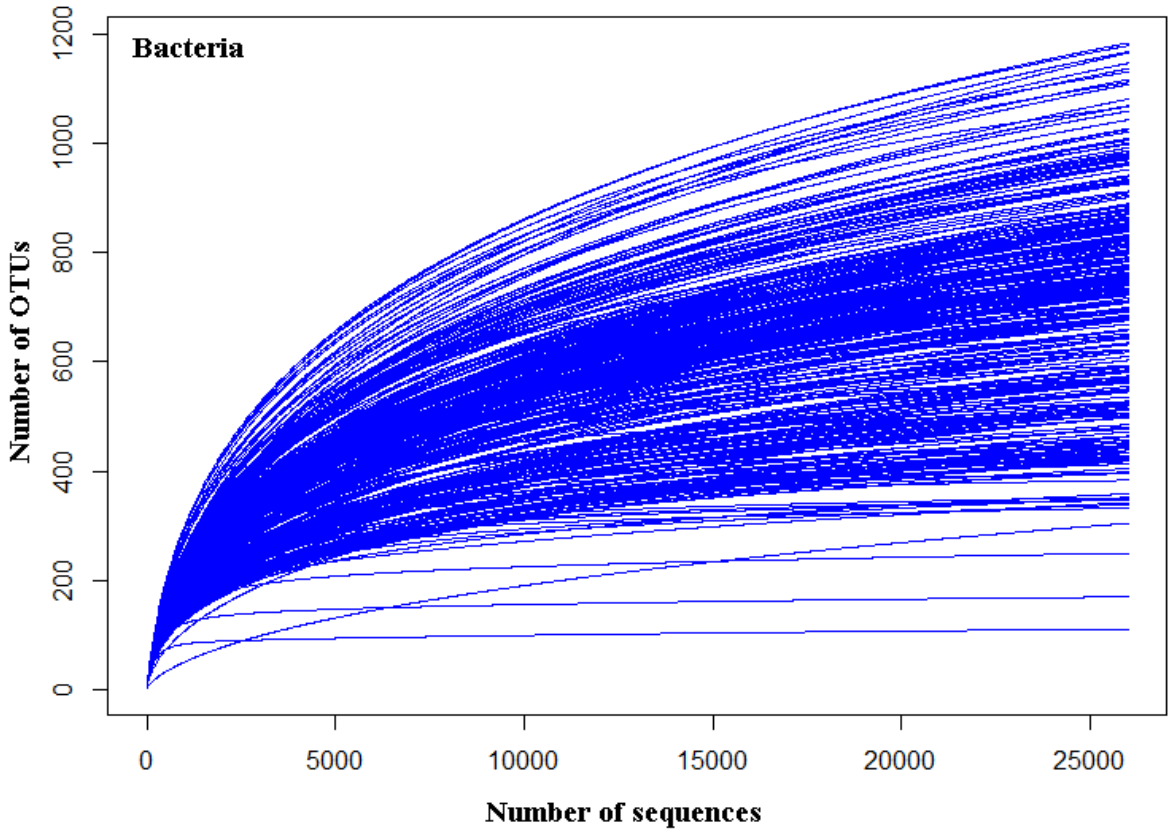

Fig. S7: Rarefaction curves for bacteria, subsampled to a depth of 13400 sequences.

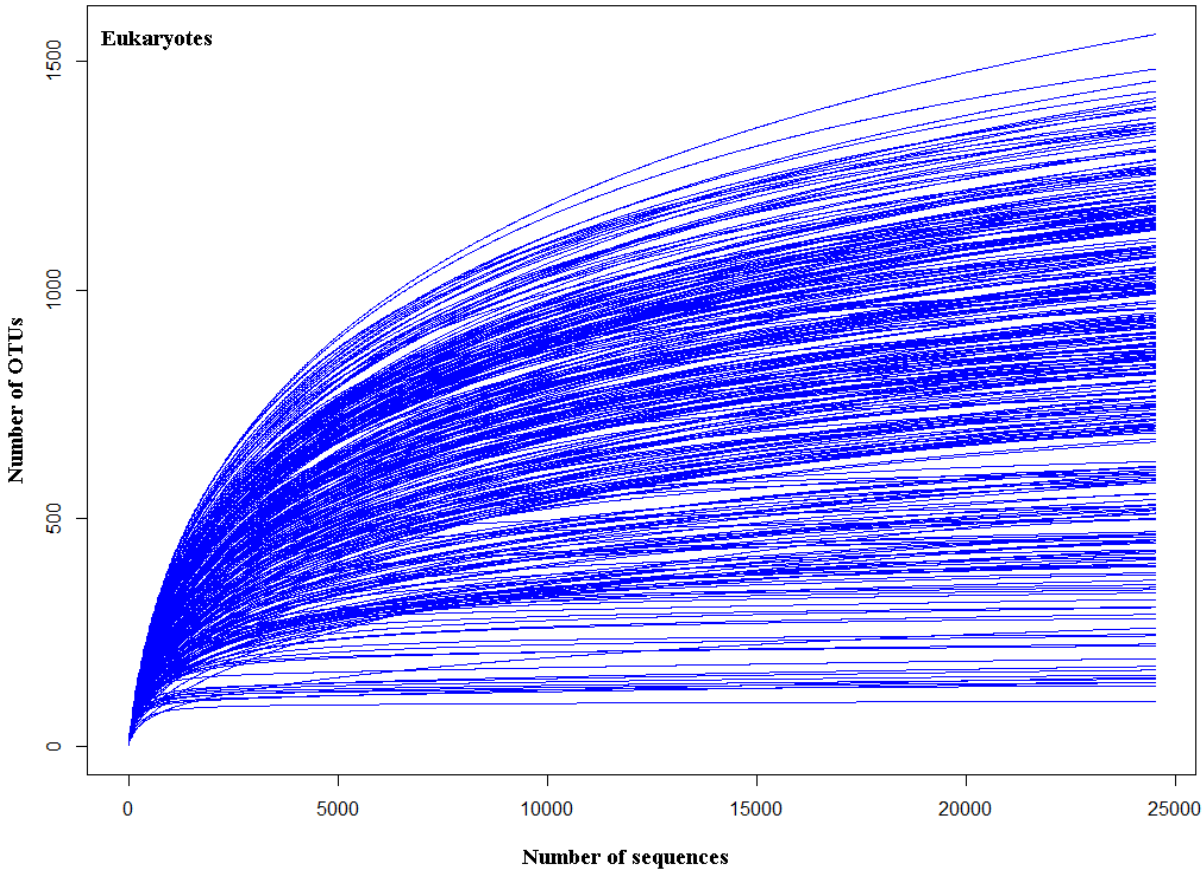

Fig. S8: Rarefaction curves for eukaryotes, subsampled to a depth of 24500 sequences.

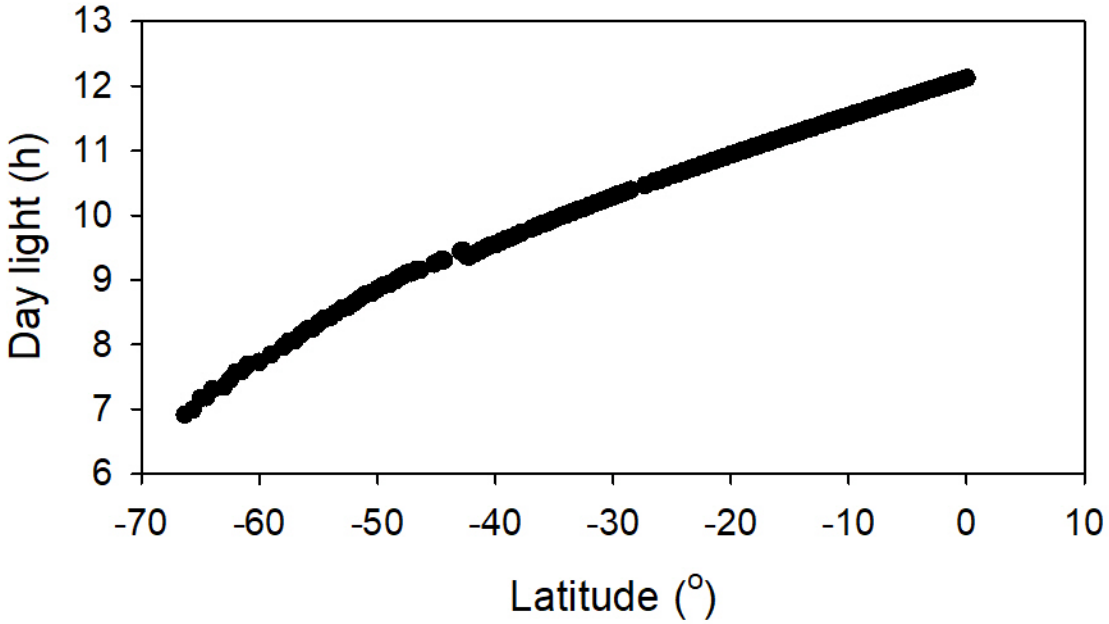

Fig. S9: Hours of day light along the latitudinal P15S GO-SHIP transect in early Austral winter in the South Pacific Ocean.

References

1. Karageorgis AP, *et al.* (2008) Particle dynamics in the Eastern Mediterranean Sea: A synthesis based on light transmission, PMC, and POC archives (1991–2001). *Deep Sea Research Part I: Oceanographic Research Papers* 55(2):177-202.
2. Talley L, Pickard G, Emery W, & Swift J (2011) Descriptive Physical Oceanography - An Introduction.
3. Aminot A, Kérouel R, & Coverly SC (2009) Nutrients in seawater using segmented flow analysis. *Practical guidelines for the analysis of seawater*:143-178.
4. Murphy J & Riley J (1962) A modified single solution method for the determination of phosphate in natural waters. *Anal. Chim. Acta* 27:31 - 36.
5. Armstrong FAJ, Stearns CR, & Strickland JDH (1967) The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer® and associated equipment. *Deep Sea Res. Oceanogr. Abstr.* 14:381 - 389.
6. Grasshoff K, Kremling K, & Ehrhardt M eds (2009) *Methods of seawater analysis* (Wiley-VCH, Weinheim, Germany), 3rd edition Ed.
7. Kérouel R & Aminot A (1997) Fluorometric determination of ammonia in sea and estuarine waters by direct segmented flow analysis. *Mar. Chem.* 57:265 - 275.
8. Appleyard S, Abell G, & Watson R (2013) Tackling microbial related issues in cultured shellfish via integrated molecular and water chemistry approaches. in *Seafood CRC Final Report (2011/729) April 2013. ISBN ISBN 9781922173980 (ebook), 89pp.*
9. Parsons TR, Takahashi M, & Hargrave B (2013) *Biological oceanographic processes* (Elsevier).
10. Hooker SB, *et al.* (2005) The second SeaWiFS HPLC analysis round-robin experiment (SeaHARRE-2). *NASA Tech. Memo* 212785:124.
11. Hirata T, *et al.* (2011) Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* 8(2):311.
12. Uitz J, Claustre H, Morel A, & Hooker SB (2006) Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll. *Journal of Geophysical Research: Oceans* 111(C8).
13. Sieburth JM, Smetacek V, & Lenz J (1978) Pelagic ecosystem structure: heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnology and oceanography* 23(6):1256-1263.
14. Aiken J, *et al.* (2009) Phytoplankton pigments and functional types in the Atlantic Ocean: a decadal assessment, 1995–2005. *Deep Sea Research Part II: Topical Studies in Oceanography* 56(15):899-917.
15. Kheireddine M, *et al.* (2017) Assessing pigment-based phytoplankton community distributions in the Red Sea. *Frontiers in Marine Science* 4:132.
16. Ras J, Claustre H, & Uitz J (2008) Spatial variability of phytoplankton pigment distributions in the Subtropical South Pacific Ocean: comparison between in situ and predicted data. *Biogeosciences* 5(2):353-369.

17. Raes EJ, Clementson L, Bodrossy L, Strutton P, & Waite AM (2017) Pigment and primary productivity on the P15S GO-SHIP transect: From the ice edge to the equator along 170°W. (PANGAEA).
18. Jardillier L, Zubkov MV, Pearman J, & Scanlan DJ (2010) Significant CO2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. *The ISME journal* 4(9):1180-1192.
19. Knap A, Michaels A, Close A, Ducklow H, & Dickson A (1996) Protocols for the joint global ocean flux study (JGOFS) core measurements.
20. Fuhrman JA, *et al.* (2008) A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences* 105(22):7774-7778.
21. Amend AS, *et al.* (2013) Macroecological patterns of marine bacteria on a global scale. *Journal of Biogeography* 40(4):800-811.
22. Team R (2015) RStudio: integrated development for R. *RStudio, Inc., Boston, MA URL http://www. rstudio. com*.
23. Holm S (1979) A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*:65-70.
24. Elith J, Leathwick JR, & Hastie T (2008) A working guide to boosted regression trees. *Journal of Animal Ecology* 77(4):802-813.
25. Clarke KR & Warwick R (2001) Change in marine communities. *An approach to statistical analysis and interpretation*.