Supplementary Material

**Metabolic pathways inferred from the 16S ribosomal RNA gene illuminate ecological changes across South Pacific frontal boundaries**

**Eric J. Raes1\*, Kristen Karsh1, Swan L. S. Sow1,2,3, Martin Ostrowski4, Mark Brown5, Jodie van de Kamp1, Rita M. Franco-Santos2, Levente Bodrossy1, Anya M. Waite6**

1)CSIRO Oceans and Atmosphere, GPO Box 1538, Hobart, TAS, 7001 Australia

2) Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia

3) Royal Netherlands Institute for Sea Research, Texel, The Netherlands

4) Climate Change Cluster, University of Technology Sydney, Australia
5) School of Biotechnology and Biomolecular Sciences, University of NSW Kensington, Sydney, NSW 2052 Australia
6) Ocean Frontier Institute and Dept. of Oceanography, Dalhousie University, Halifax, B3H 4R2 Nova Scotia, Canada

**\* Corresponding author:**Eric Jorden Raes
eric.raes@csiro.au

## Materials and Methods

## DNA sequencing and bioinformatics

Amplicon sequence variant (ASV) tables were prepared after Bissett, Fitzgerald 1 and as outlined in <https://www.australianmicrobiome.com/protocols/16sanalysisworkflow/>. Briefly, FLASH v1.2.11 2 was used to merge paired-end reads and unique sequences were denoised into 79073 ASVs with the UNOISE algorithm 3 using USEARCH 64 bit v8.0.1517 4. ASV abundance profiles per sample were constructed by mapping all the reads to the unique ASVs using the USEARCH ‘otutab’ command. Genomic data are available at <https://www.ncbi.nlm.nih.gov/bioproject/385736> under the accession no. PRJNA385736

## Functional microbial community composition

The main idea of PICRUSt2 is to infer the genomic content for a given ASV based on averaging over the reference genomes that are most similar in 16S rRNA sequences 5, 6, 7, 8, 9. The PICRUSt2 genome database (as of November 8, 2017) is based on 41,926 bacterial and archaeal genomes from the Integrated Microbial Genomes (IMG) database 10. The nearest sequenced taxon index (NSTI) score in PICRUSt2 is a metric to evaluate the degree to which microorganisms in a sample are related to available sequenced genomes. PICRUSt2 outputs MetaCyc pathway predictions (Caspi et al., 2018) which are comparable with common shotgun metagenomics output abundances. MetaCyc pathways were linked to their respective secondary superclass levels in the MetaCyc pathway hierarchy system using the ‘smart tables‘ and selecting the ontology ‘parent class‘ at <https://metacyc.org/>.

## Nutrient analyses of seawater

Seawater samples were collected for nutrient analyses from the Niskin bottles at all 36 depth horizons. Dissolved inorganic nutrient (DIN) analyses were conducted at sea on a Bran + Luebbe AA3 HR segmented flow analyser by the CSIRO Hydrochemistry group, following standard spectrophotometric methods 11, 12, 13, 14, 15. Detection limits for dissolved inorganic concentrations of silicate (Si) were 0.2 µmol L-1, for phosphate (PO43-) 0.01 µmol L-1), for nitrate (NO3-) 0.01 µmol L-1, for nitrite (NO2-) 0.015 µmol L-1, and for ammonium (NH4+) 0.015 µmol L-1.

## Rate measurements

Stable isotope tracers (15N and 13C) were used to measure N2 fixation rates, NO3- and NH4+ uptake rates and C assimilation rates (primary productivity). The rates have been published in Raes, van de Kamp 16 and are used in this study as a setting for the inferred functional metabolic pathways. Primary productivity and nitrogen assimilation data from this study are available at <https://doi.pangaea.de/10.1594/PANGAEA.884052> (Raes et al., 2017) and <https://doi.pangaea.de/10.1594/PANGAEA.885169> (Raes et al., 2018b).

## Pigment analyses

Pigment analyses have been published in Raes, van de Kamp 16 and data are available at <https://doi.pangaea.de/10.1594/PANGAEA.884052>.

**Supplementary Tables:**

**Supplementary Table 1:** ANOSIM results for OTU and MetaCyc pathways abundances.

|  |
| --- |
| **Bacterial 16S rRNA data** |
| **Ocean Provinces** | **R-value**  | **Significance** | **Permutations** |
| Southern Ocean and STF | 0.823  | 0.001 | 999  |
| Southern Ocean and SPSG | 0.997  | 0.001 | 999  |
| Southern Ocean and PED | 1  | 0.001 | 999  |
| STF and SPSG | 0.859  | 0.001 | 999  |
| STF and PED | 0.999  | 0.001 | 999  |
| SPSG and PED | 0.448  | 0.001 | 999  |
|  |  |  |  |
|  |  |  |  |
| **MetaCyc data** |  |  |  |
| **Ocean Provinces** | **R-value**  | **Significance** | **Permutations** |
| Southern Ocean and STF | 0.594  | 0.001 | 999  |
| Southern Ocean and SPSG | 0.988  | 0.001 | 999  |
| Southern Ocean and PED | 0.955  | 0.001 | 999  |
| STF and SPSG | 0.91  | 0.001 | 999  |
| STF and PED | 0.927  | 0.001 | 999  |
| SPSG and PED | 0.495  | 0.001 | 999  |

R values indicate the level of similarity between all sampled ocean provinces, with R=0 indicating strong similarity and R=1 indicative of a strong dissimilarity. Significant differences at the p<0.001 level.

**Supplementary Table 2:** The 399 inferred MetaCyc pathways from the PICRUSt2 analysis. MetaCyc pathways are collapsed into 41 secondary superclasses and ontology parents of MetaCyc pathways are also shown at two levels.

**Supplementary Table 3:** BRT model outputs from 75 observations and 22 predictors. Attached as a separate .csv file. See table for the relative influence (%) for the 22 biotic and abiotic predictors.

**Supplementary Table 4:** is attached as a separate .csv file and details the relative abundance of each sequence at the 97% similarity threshold with an NSTI >2.

**Supplementary Figures**



**Supplementary Figure 1:** Nitrateconcentrations (µmol L-1) along the P15S GO-SHIP transect from 66 oS to the equator. Top panel highlights data from the surface to 350m, and the bottom panel show nitrateconcentrations from 350m to the bottom. The thick white line represents the MLD. The black circles on the top panel show where the DNA samples were collected and the little black dots on both panels shows denote the three sampling depths of 140 CTD stations. The depth range for Depth 1 was 6.6±4.1 m, for Depth 2 it was 35.3±13.1 m and for Depth 3, 70.7±18.0 m. The boxes on top of the figures mark the oceanographic provinces; the Southern Ocean, the Subtropical Frontal zone (STF), the South Pacific Subtropical Gyre (SPSG), and the Pacific Equatorial Divergence Province (PED).



**Supplementary Figure 2:** Fermentation pathways



**Supplementary Figure 3:** Microbial metabolism pathways presented as the secondary superclass in MetaCyc classification system. Figures **A**, **E**, **H** and **L** present our hypotheses. The cumulative relative abundances of the pathways in figures **B, C, D, F, G, I, J, K, M** and **N** present trends in the core ecosystem functions from 66 oS to the equator. The expected trends are shown above each panel. (**1**) Eppley 17, (**2**) Boyd, Watson 18, **(3)** Browning, Achterberg 19, (**4**) Rivkin and Legendre 20, (**5**) Van Mooy, Rocap 21. Non-significant trends are denoted by ‘ns’ in the trend plots.

****

**Supplementary Figure 4:**  ‘indicator’ pathways in the Southern Ocean



**Supplementary Figure 5:**  ‘indicator’ pathways in the STF



**Supplementary Figure 6:**  ‘indicator’ pathways in the PED

**References:**

1. Bissett A*, et al.* Introducing BASE: the Biomes of Australian Soil Environments soil microbial diversity database. *GigaScience* **5**, 21 (2016).

2. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957-2963 (2011).

3. Edgar RC, Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* **31**, 3476-3482 (2015).

4. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460-2461 (2010).

5. Barbera P*, et al.* EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. *Systematic Biology* **68**, 365-369 (2018).

6. Czech L, Barbera P, Stamatakis A. Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics*, (2020).

7. Douglas GM*, et al.* PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* **38**, 669-688 (2020).

8. Louca S, Doebeli M. Efficient comparative phylogenetics on large trees. *Bioinformatics* **34**, 1053-1055 (2017).

9. Ye Y, Doak TG. A Parsimony Approach to Biological Pathway Reconstruction/Inference for Genomes and Metagenomes. *PLOS Computational Biology* **5**, e1000465 (2009).

10. Markowitz VM*, et al.* IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic acids research* **40**, D115-D122 (2012).

11. Aminot A, Kérouel R, Coverly SC. Nutrients in seawater using segmented flow analysis. *Practical guidelines for the analysis of seawater*, 143-178 (2009).

12. Armstrong F, Stearns C, Strickland J. The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer® and associated equipment. In: *Deep Sea Research and Oceanographic Abstracts*). Elsevier (1967).

13. Hydes D*, et al.* Determination of Dissolved Nutrients (N, P, SI) in Seawater With High Precision and Inter-Comparability Using Gas-Segmented Continuous Flow Analysers., (2010).

14. Kérouel R, Aminot A. Fluorometric determination of ammonia in sea and estuarine waters by direct segmented flow analysis. *Marine Chemistry* **57**, 265-275 (1997).

15. Grasshoff K, Kremling K, Ehrhardt M. *Methods of seawater analysis*. John Wiley & Sons (2009).

16. Raes EJ*, et al.* N2 Fixation and New Insights Into Nitrification From the Ice-Edge to the Equator in the South Pacific Ocean. *Frontiers in Marine Science* **7**, (2020).

17. Eppley RW. Temperature and phytoplankton growth in the sea. *Fish bull* **70**, 1063-1085 (1972).

18. Boyd PW*, et al.* A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilization. *Nature* **407**, 695-702 (2000).

19. Browning TJ*, et al.* Nutrient co-limitation at the boundary of an oceanic gyre. *Nature* **551**, 242-246 (2017).

20. Rivkin RB, Legendre L. Biogenic carbon cycling in the upper ocean: effects of microbial respiration. *Science* **291**, 2398-2400 (2001).

21. Van Mooy BA, Rocap G, Fredricks HF, Evans CT, Devol AH. Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments. *Proceedings of the National Academy of Sciences* **103**, 8607-8612 (2006).