Subject category: Integrated genomics and post-genomics approaches in microbial ecology

Heterotrophic resourcefulness and unusual sulfur biogeochemistry in a hypersaline Antarctic lake

Running title: Heterotrophic resourcefulness and unusual sulfur cycling

Sheree Yau[1], Timothy J. Williams[1], Federico M. Lauro[1], Matthew Z. DeMaere[1], Mark V. Brown[1,2], John Rich[3], John A.E. Gibson[4] and Ricardo Cavicchioli[1*]

[1] School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, New South Wales, Australia.
[2] Evolution and Ecology Research Centre, The University of New South Wales, Sydney, New South Wales, Australia.
[3] P. O. Box 5824, Albany 6332, Western Australia, Australia.
[4] Marine Research Laboratories, Tasmanian Aquaculture and Fisheries Institute, University of Tasmania, Hobart, Tasmania, Australia.

* To whom correspondence should be addressed: Ricardo Cavicchioli, School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, NSW, 2052, Tel. (+61 2) 9385 3516, Fax. (+61 2) 9385 2742, Email. r.cavicchioli@unsw.edu.au

Keywords: Metagenomics, Organic Lake, Antarctic microbial ecology, nutrient cycles, dimethylsulfide

Formatted: Numbering: Continuous

Moved (insertion) [1]

Formatted: Font: Not Bold

Deleted: Federico M. Lauro[1],

Deleted: **********

Formatted: Font: Times New Roman, 12 pt

Moved up [1]: Running title: Heterotrophic resourcefulness and unusual sulfur cycling¶
¶

Deleted: ¶

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Deleted: cycle

Abstract

Organic Lake is a shallow marine-derived hypersaline lake in the Vestfold Hills, Antarctica that has the highest reported concentration of dimethylsulfide (DMS) in a natural body of water (Franzmann *et al.*, 1987b). To determine the composition and functional potential of the microbial community and learn about the unusual sulfur chemistry in Organic Lake, shotgun metagenomics (2.4 Gbp titanium 454) was performed on size fractionated samples (3.0, 0.8 and 0.1 µm) collected along a depth profile. Eucaryal phytoflagellates were the main photosynthetic organisms. Bacteria were dominated by the globally distributed heterotrophic taxa *Marinobacter*, *Roseovarius* and *Psychroflexus.* Candidate division RF3 was overrepresented at the oxycline and OD1 in the lake bottom. The dominance of heterotrophic degradation coupled with low fixation potential indicates possible net carbon loss. However, abundant marker genes for aerobic anoxygenic phototrophy, sulfur oxidation, rhodopsins and CO oxidation were also linked to the dominant heterotrophic bacteria and may indicate use of photo- and lithoheterotrophy as mechanisms for conserving organic carbon. Similarly, a high genetic potential for the recycling of nitrogen compounds likely functions to retain fixed nitrogen in the lake. Dimethylsulfoniopropionate (DMSP) lyase genes (*dddD*, *dddL* and *dddP*) were abundant indicating DMSP is a significant carbon and energy source. Unlike marine environments, DMSP demethylases (*dmdA*) were less abundant than DMSP lyases indicating that DMSP cleavage is the likely source of the high DMS concentration. Strategies of nutrient resourcefulness such as DMSP cleavage, carbon mixotrophy (photoheterotrophy and lithoheterotrophy) and nitrogen remineralization in dominant Organic Lake bacteria are potentially important adaptations to nutrient constraints. In particular, carbon mixotrophy reduces the extent of carbon oxidation for energy production allowing more carbon to be used for biosynthetic processes. The study sheds light on how microbial communities and the functional processes they perform evolve in response to unusual environmental conditions.

Deleted: ¶

Formatted: Normal1

Deleted:

Deleted: , lithoheterotrophyand

Deleted: This

Introduction

Life in the Antarctic is constrained by low temperature, and water, nutrient and light availability. In the Antarctic frozen desert, ice-free regions containing liquid water in lakes and ponds are rare oases for life (Wilkins *et al.*, 2012). The Vestfold Hills, on the eastern shore of Prydz Bay, East Antarctica (Supplementary Figure S1), is a unique region where hundreds of lakes are present. The lakes were formed from seawater, trapped less than 10 000 BP when the continental ice-sheet receded and the land rose above sea-level (Zwartz *et al.*, 1998; Gibson, 1999). Differing local conditions has led each lake to develop unique physical and chemical properties, and life in the lakes tends to be entirely microbial with low levels of diversity (Bowman *et al.*, 2000). The Vestfold Hills contains the highest density of meromictic (permanently stratified) water bodies in Antarctica (Gibson, 1999). The strong physico-chemical stratification within a single, largely closed system, provides the opportunity to investigate how microbial communities and ecosystem processes have evolved in the cold and in response to gradients of nutrients, oxygen, salinity and solar irradiance.

Molecular biology approaches have proven useful for describing the diversity and gene content of microorganisms in Antarctic lakes and for inferring the functional roles of the taxa present (Laybourn-Parry and Pearce, 2007; Wilkins *et al.*, 2012). However to date, only a few large scale shotgun metagenome studies have been performed on the Antarctic continent and in the surrounding Southern Ocean (reviewed in Wilkins *et al.*, 2012). In the Vestfold Hills, metagenomics and metaproteomics have been used to study Ace Lake (64°28′23.2″S, 78°11′20.8″E)) and Organic Lake (68°27′23.4″S, 78° 11′ 22.6″E) (Ng *et al.*, 2010; Lauro *et al.*, 2011; Yau *et al.*, 2011). For Ace Lake, a comprehensive assessment of the community structure, biogeochemical fluxes and responses to resource limitation have been described (Lauro *et al.*, 2011). The metabolism of abundant green sulfur bacteria (Ng *et al.*, 2010) was found to play a central role in nutrient cycling and a mathematical model was developed that showed its dominance was dependent on synchronicity with the polar light cycle leading to absence of phage predation (Lauro *et al.*, 2011). For Organic Lake, a member of the virophage virus family was discovered that potentially regulates microbial loop dynamics (Yau *et al.*, 2011). The Organic Lake virophage likely depends on phycodnaviruses (algal viruses) and it was predicted that the virophage would reduce infective phycodnaviruses leading to an increased frequency of algal blooms and thus carbon flux (Yau *et al.*, 2011). Virophage sequences were also identified in a range of aquatic metagenomes revealing that they are likely to play ecologically important roles in many aquatic systems (Yau *et al.*, 2011). These studies on Ace and Organic lakes both used shotgun metagenomics, and the unanticipated nature of the discoveries serve to illustrate the value of adopting a metagenomics approach for learning about microbial ecology in Antarctic environments.

Due to polar light cycle and low overall levels of photosynthetically active radiation, phototrophic growth and biomass production are restricted, being relatively high in summer and negligible in winter (Laybourn-Parry *et al.*, 2005). In Ace Lake, phototrophic algae, particularly phytoflagellates, engage in mixotrophy, thereby supplementing their carbon and nutrient requirements to enable them to remain active during winter and poised to photosynthesize in summer (Laybourn-Parry *et al.*, 2005). Resourcefulness has also been demonstrated by heterotrophic marine bacteria that gain from light energy, either directly through photoheterotrophic processes involving aerobic anoxygenic photosynthesis (AAnP), or indirectly through lithoheterotrophic processes utilizing inorganic compounds (*e.g.* CO) formed from

Deleted: ¶
Page Break
¶
¶

Deleted: 2000b

Formatted: Font: Times New Roman, 12 pt

Deleted: &

Deleted: (OLV)

Deleted: OLV

Deleted: OLV-like

Deleted: coastal marine, hypersaline and freshwater

Deleted: virophages

Deleted: (e.g. OLV)

Deleted: "look and see"

Formatted: Font: Times New Roman, 12 pt

dissolved organic carbon reacting with light (Moran and Miller, 2007). Gaining energy by not consuming organic carbon or liberating CO_2 can result in the more efficient utilization of carbon by microbial populations and the conservation of carbon within marine systems (Moran and Miller, 2007). In Ace Lake, nutrient cycles have been interpreted from the relative abundance of marker genes (Lauro *et al.*, 2011). For the carbon cycle, fermentation, sulfate-reduction and methanogenesis is inferred to lead to particulate organic carbon breakdown through to CO_2 and CH_4, with CO oxidation of incompletely oxidized organic compounds being used for energy generation (Lauro *et al.*, 2011). In this system, high levels of CO oxidation may explain why the inorganic carbon content in the lake remains high (Rankin et al., 1999; Lauro *et al.*, 2011). Aside from this study (Lauro *et al.*, 2011), shotgun metagenomics has not been employed to examine the versatility of metabolic and energy generation pathways of microbial communities in Antarctic lake systems.

Organic Lake is shallow (6.8 m) and has variable surface water temperatures (−14 to +15 °C) while remaining sub-zero throughout most of its depth (Franzmann *et al.*, 1987b; Gibson *et al.*, 1991; Roberts *et al.*, 1993; Gibson, 1999). The lake has a high organic load generated from autochthonous production and input from penguins and terrestrial algae, and nutrient turnover is expected to be slow due to the constraints imposed on microbial activity by the lake's hypersalinity (≈230 g L^{-1} maximum salinity) and low temperature (Franzmann *et al.*, 1987b; Gibson *et al.*, 1991; Roberts *et al.*, 1993; Gibson, 1999). The salt and marine biota in the lake originate from seawater that was trapped in a basin ~ 3 000 y B.P. (Zwartz *et al.*, 1988; Bird *et al.*, 1991). The bottom waters of Organic Lake are unusual due to the absence of hydrogen sulfide and the high concentration of the volatile gas dimethylsulfide (DMS) (Deprez *et al.*, 1986; Franzmann *et al.*, 1987; Gibson *et al.*, 1991; Roberts and Burton 1993a; Roberts *et al.*, 1993b). Concentrations of DMS as high as 5 000 nM have been recorded in Organic Lake (Gibson *et al.*, 1991), 100 times the maximum concentration recorded from seawater in the adjacent Prydz Bay and at least 1000 times that of the open Southern Ocean (Curran and Jones, 1998).

More than forty years ago atmospheric DMS was proposed to have a regulatory effect on global cloud cover as it is a precursor of cloud condensation nuclei (Lovelock and Maggs, 1972; Charlson *et al.*, 1987). However, the first enzymes involved in DMS production were only identified in the last six years (Todd *et al.*, 2007). Rapid progress has been made in this short period and the pathways and organisms involved in DMS transformations have been extensively reviewed (Johnston *et al.*, 2008; Schäfer *et al.*, 2010; Curson *et al.*, 2011b; Reich *et al.*, 2011b; Moran *et al.*, 2012). The main source of DMS in the marine environment is from the breakdown of DMSP. Eucaryal phytoplankton, in particular diatoms, dinoflagellates and haptophytes, produce large quantities of DMSP, which is thought to function principally as an osmolyte. DMSP is released due to cell lysis, grazing or leakage and follows two known fates: DMSP cleavage by DMSP lyases (DddD, -L, -P, -Q, -W and -Y) or demethylation by DMSP demethylase (DmdA). Both pathways are associated with diverse microorganisms that can utilize DMSP as a sole carbon and energy source. However, it is only the cleavage pathway that releases volatile DMS that can lead to sulfur loss through ventilation to the atmosphere.

The very high levels of DMS in Organic Lake make it an ideal system for identifying the microorganisms and the processes involved in DMS accumulation. The previous Organic Lake metagenome study examined viruses from the 0.1 µm fraction of surface water that was collected from Organic Lake in December 2006, and November and December 2008 (Yau *et al.*, 2011). In the present study we focused on the cellular population rather than viruses. Metagenomic

Deleted: ca.

Comment [U1]: Sheree please check all Franzmann refs, e.g. 1987a or b?

Deleted: &

Deleted: &

Deleted: Over

Deleted: ,

Deleted: &

analyses were performed on biomass captured by sequential filtration through a 20 µm pre-filter onto 3.0, 0.8, 0.1 µm filters, from a depth profile (1.7, 4.2, 5.7, 6.5 and 6.7 m) taken in November 2008. This filtration and shotgun sequencing approach was originally adopted by the Global Ocean Sampling (GOS) expedition (Rusch *et al.*, 2007) and has proven to be a powerful approach for studying Antarctic aquatic microbial communities (Ng *et al.*, 2010; Lauro *et al.*, 2011; Yau *et al.*, 2011; Brown *et al.*, 2012; Williams *et al.*, 2012b; Wilkins *et al.*, 2012b). Our study determined the composition and functional potential of Organic Lake microbiota and, in conjunction with historic and contemporary physico-chemical data, generated an integrative understanding of the whole lake ecosystem.

Materials and methods

Characteristics of the lake and samples taken

The water level of Organic Lake was measured by surveying as +1.886 m relative to the survey mark (NMV / S / 53) located at S 68°27′28.3″ E078°11′20.9″. Water was collected from Organic Lake on 10 November 2008 through a 30 cm hole in the 0.8 m thick ice cover above the deepest point in the lake. The sampling hole (S 68°27′22.2″ E 78°11′23.9″) was established following bathymetry measurements constructed on a metric grid. Samples were collected for metagenomics, microscopy and chemical analyses at 1.7, 4.2, 5.7, 6.5 and 6.7 m depths (maximum lake depth 6.8 m). For metagenomics, lake water was passed through a 20 µm pore size pre-filter, and microbial biomass captured by sequential filtration onto 3.0 µm, 0.8 µm and 0.1 µm pore size 293 mm polyethersulfone membrane filters, and samples immediately preserved in buffer and cryogenically frozen in liquid nitrogen, as described previously (Ng *et al.*, 2010; Lauro *et al.*, 2011). Between 1–2 L of lake water was sufficient to saturate the holding capacity of the filters. DNA was extracted from the filters, samples sequenced using the Roche GS-FLX titanium sequencer, and reads processed to remove low quality bases, assembled and annotated, as previously described (Ng *et al.*, 2010; Lauro *et al.*, 2011).

Physical and chemical analyses

An *in situ* profile of pH, conductivity, turbidity, dissolved oxygen (DO) and pressure was measured using a submersible probe (YSI sonde model V6600). A temperature profile was measured using a maximum-minimum mercury thermometer as the YSI probe did not have a capacity to record temperature below –10ºC. The 5.7 m sample corresponded to the turbidity maximum and the 6.5 m sample to the turbidity minimum. Conductivity at *in situ* temperature was converted to conductivity at 15ºC as described previously (Gibson, 1999). The adjusted conductivity brings the temperature to within a range suitable for estimating practical salinity using the formula of Fofonoff and Millard (1983). Salinity was likely to have been underestimated as it is higher than the range (2–42) for which the conductivity–salinity relation holds. However, the relative difference in salinity between the samples would be accurate. Density was calculated from the *in situ* conductivity and temperature using the equations described by Gibson *et al.* (1990) and expressed at temperature T as:

$$\sigma_T = (1000 - \text{density})\ \text{kg/m}^3$$

Ammonia, nitrate, nitrite, total nitrogen (TN), total dissolved nitrogen (TDN), dissolved reactive phosphorus (DRP), total phosphorus (TP), total dissolved phosphorus (TDP), total organic carbon (TOC), total dissolved carbon (DOC), total sulfur (TS) and total dissolved sulfur (TDS) were determined by American Public Health Associations Standard Methods at the Analytical

Deleted: Methods
Deleted: 27’28
Deleted: ”
Deleted: 11’20
Deleted: ”.
Deleted: 27’22
Deleted: ”
Deleted: 11’23
Deleted: ”)
Deleted: (Figure ****).
Deleted: (
Deleted:).

Services, Tasmania. Values for dissolved nutrients were measured after filtration through a 0.1 µm pore size membrane filter. All other nutrients were measured from water collected after filtration through the on-site 20 µm pore size pre-filter. Ammonia, nitrate, nitrite, DRP, TN, TDN, TP and TDP were measured in a Flow Injection Analyser (Lachat Instruments, Colorado, USA). TOC and DOC were determined in the San++ Segmented Flow Analyser (Skalar, Breda, Netherlands). TS and TDS were analyzed in the 730ES Inductively Coupled Plasma–Atomic Emission Spectrometer (Agilent Technologies, California, USA). Principal Component Analysis (PCA) was performed using the PRIMER Version 6 statistical package (Clarke and Gorley, 2006) on the normalized physical and chemical parameters.

Epifluorescence microscopy

Water samples collected for microscopy were preserved in formaldehyde (1% v/v). Cells and virus-like particles (VLPs) were vacuum filtered onto 25 mm polycarbonate 0.015 µm pore-size membrane filters (Nuclepore Track-etched, Whatman, GE Healthcare, USA) with a 0.45 µm pore-size backing filter. The 0.015 µm filter was mounted onto a glass slide with ProLong® Gold anti fade reagent (Invitrogen, Life Technologies, NY, USA) and 2 µl (25 × dilution in sterile filtered milliQ water <0.015 µm) SYBR® Gold nucleic acid stain (Invitrogen, Life Technologies, NY, USA). Prepared slides were visualized in an epifluorescence microscope (Olympus BX61, Hamburg, Germany) under excitation with blue light (460–495 nm, emission 510–550 nm). Cell and VLP counts were performed on the same filter over 30 random fields of view.

Cellular diversity analyses

Diversity of *Bacteria*, *Archaea* and *Eucarya* was assessed using ribosomal small subunit (SSU) gene sequences. Metagenomic reads that matched the 16S and 18S rRNA genes were retrieved using Metaxa (Bengtsson *et al.*, 2011). Only sequences longer than 200 bp were accepted for downstream analysis. The Quantitative Insights Into Microbial Ecology (QIIME) pipeline (version 1.4.0) (Caporaso *et al*., 2010) implementing UCLUST, was used to group SSU sequences into operational taxonomic units (OTUs) at 97% percent identity against the SILVA SSU reference database (release 108) (www.arb-silva.de). SSU sequences that did not cluster with sequences from SILVA were allowed to form new OTUs (no suppression). A representative sequence from each OTU was chosen and classified to the genus level using QIIME implementing the RDP classifier (Wang *et al*., 2007) trained against SILVA. Assignments were accepted to the lowest taxonomic rank with bootstrap value ≥85%. To allow comparison of the relative abundance of taxa, the number of SSU matches per sample filter was normalized to the average number of reads (403 577). Statistical analysis on the relative SSU abundances was performed using the PRIMER Version 6 package (Clarke and Gorley, 2006). The SSU counts of each sample filter were aggregated to the genus level and square root transformed to reduce the contribution of highly abundant taxa. A resemblance matrix was computed using Bray-Curtis similarity. The upper mixed zone (1.7, 4.2 and 5.7 m) and deep zone (6.5 and 6.7 m) samples were designated as separate groups and an analysis of similarity (ANOSIM) performed to test for difference between the two groups. BEST analysis was performed with the abiotic variables: conductivity, temperature, turbidity, DO, pH, TOC, TN, TP, TS, total C:N, total C:P, total N:P, cell counts and VLP counts. The Bio-Env procedure in BEST looks at all the abiotic variables in combination and finds a subset sufficient to best explain the biotic structure. A heat map with bi-

Deleted: &

Deleted: &

clustering dendogram was generated using R and the package 'seriation' (Hahsler *et al*., 2008) on the normalized square-root transformed SSU counts.

Analysis of functional potential

The relative abundance and taxonomic origin of functional marker genes was used to determine the potential for carbon, nitrogen and sulfur conversions. Open reading frames (ORFs) were predicted from trimmed metagenomic reads using MetaGene (Noguchi *et al*., 2006) accepting those >90 bp in length. ORFs were translated using the standard bacterial/plastid translation table and compared to protein sequences from the Kyoto Encyclopedia of Genes and Genomes (KEGG) GENES database (release 58) using the Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990). The BLAST output was processed using KEGG Orthology Based Annotation System (KOBAS) version 2.0 (Xie *et al.*, 2011) accepting assignments to KEGG Orthology (KO) groups with e-value <1e−05 and rank >5. KO groups used as functional markers are listed in Supplementary Table S1. Marker enzymes were assigned to taxonomic groups based on the species of origin of the best KEGG GENES BLASTp match. Marker genes not represented by a KO group were assessed by BLASTp queries of marker gene sequences with experimentally confirmed function (Supplementary Table S2) against a database of translated ORFs predicted from metagenomic reads. Matches were accepted if the e-value was <1e−10 and sequence identity was within the range shared by homologs of the query sequence(s) (Supplementary Table S2). Matches to marker genes were normalized to 100 Mbp per sample and counted. Normalized frequencies of markers from the same pathway were averaged and those from different pathways were summed.

The same marker genes and BLAST procedure was used to compare the DMSP catabolism and photoheterotrophy potential of Organic Lake with nearby Ace Lake (Lauro *et al.*, 2011), Southern Ocean (Wilkins *et al.*, 2012b) and GOS metagenomes (Rusch *et al.*, 2007). Counts of single copy gene *recA* were also determined to estimate the percentage of genomes containing each marker gene (percentage of marker genes relative to *recA*). Matches to *recA* were accepted with e-value <1e-20 (Howard *et al.*, 2008). For GOS samples, the BLAST database was generated from peptide sequences retrieved from CAMERA (camera.calit2.net) while the other BLAST databases were produced as for Organic Lake. The total number of trimmed base pairs for GOS samples was estimated by multiplying the number of reads from each sample by the average read length (822 bp) (Rusch *et al.*, 2007).

Marker gene sequences for phylogenetic analysis were clustered using the CD-HIT web server (Huang *et al*., 2010) at 90% global amino acid identity. A representative sequence from the clusters that resided within a desired conserved region and homologs from cultured strains were used in phylogenetic analyses performed in MEGA 5.05 (Tamura *et al.* 2011). Sequences were aligned with MUSCLE (Robert, 2004) using default parameters (gap opening penalty: −2.9, gap extension penalty: 0). Neighbor-joining was used to compute the phylogenies with a Poisson substitution model, uniform rates of change and complete deletion of alignment gaps. Node support was tested with bootstrap analysis (500 replicates).

Results and discussion

Abiotic properties and water column structure

In situ physico-chemical profiles (Supplementary Figure S2) measured over the deepest point in the lake (Supplementary Figure S3) determined the existence of two zones: an upper mixed zone

Comment [SY2]: I've normalized to number of base pairs per sample rather than reads. This allows comparison with Sanger sequenced GOS metagenomes that have reads 2x the length of the 454 sequenced metagenomes.

Deleted:

Formatted: Indent: First line: 0.75 cm

Deleted: the

Deleted:

Deleted: as established by

Deleted: . (

Deleted: strainswere

above 5.7 m and a suboxic deep zone below 5.7 m (Figure 1A). The separation of the two zones was indicated by a pycnocline and oxycline starting at 5.7 m. The pH also decreased with DO, likely due to fermentation products such as acetic, formic and lactic acids that have been reported in the bottom waters (Franzmann *et al*., 1987b; Gibson *et al*., 1994). The deep zone was not completely anoxic (Supplementary Figure S2). Oxygen may be episodically introduced to bottom waters as a result of currents of cold dense water sinking during surface ice-formation (Ferris *et al*., 1999). In comparison to meromictic lakes such as Ace Lake that have strong pycnoclines and a steep salt gradient in the anoxic zone, Organic Lake is shallow and has relatively weak stratification (Gibson, 1999). Samples were collected from the upper mixed (1.7, 4.2 and 5.7 m) and deep (6.5 m and 6.7 m) zones.

All nutrients, except for nitrate and nitrite reached maximum concentrations at 6.5 m (Table 1) suggestive of a layer of high biological activity above the lake bottom. Consistent with this, cell and VLP counts were highest at 6.5 m. However, turbidity was lowest at this depth demonstrating turbidity was not principally determined by cell density (Figure 1B). Microscopy images did not show a shift in cell morphology that could account for the large drop in turbidity (Supplementary Figure S4), which suggests particulate matter primarily contributed to turbidity readings. The low turbidity and peak in cell counts and nutrients at the oxycline at 6.5 m may be caused by an active microbial community degrading particulate matter. This inference is supported by the report of high concentrations of dissolved organic acids and free amino acids in the deep zone (Gibson *et al*., 1994) as these nutrients are indicative of the breakdown of high molecular weight carbohydrates, lipids and proteins. Furthermore, the C:N and C:P ratios throughout the lake were high compared to the Redfield ratio (Redfield *et al*., 1963) except at 6.5 m indicating this was the only depth where dissolved nitrogen and phosphorous were not relatively limited (Table 1). PCA analysis of physico-chemical parameters showed all samples, except the 6.5 m sample, separated with depth along the PC1 axis (Supplementary Figure S5). Accordingly, turbidity, TS and cell density were the strongest explanatory variables for the separation of the 6.5 m sample from the other deep sample, indicating that increased activity at 6.5 m was related to breakdown of particulate matter and sulfur chemistry.

Overall microbial diversity

SSU genes (3 959 reads) that were retrieved from the metagenome data grouped into 983 OTUs. OTUs for *Bacteria* comprised 76.2%, *Eucarya* 16.3% and 7.5% of SSU sequences could not be classified. Only 2 reads, assigned to a deep sea hydrothermal clade of *Halobacteriales* (Supplementary Table S4) indicating *Archaea* were rare in Organic Lake. Microbial diversity was low, consisting of 15 bacterial phyla and 6 eucaryal superkingdom divisions (Figure 2). Of these, only 7 bacterial phyla and 4 eucaryal phyla were predominant.

The most abundant bacterial classes, *Gammaproteobacteria*, *Alphaproteobacteria* and *Flavobacteria*, were represented by OTUs on all filter sizes at all depths (Figure 2A) and each consisted of one dominant genus, *Marinobacter*, *Roseovarius* and *Psychroflexus*, respectively (Figure 2C). Essentially all OTUs for *Cyanobacteria*/chloroplasts were classified as chloroplasts (Figure 2A), except for three reads that could not be assigned to any lower rank (Supplementary Table S4) indicating free-living *Cyanobacteria* were rare or absent. OTUs for moderately abundant bacterial classes were *Actinobacteria*, *Deltaproteobacteria*, *Epsilonproteobacteria*, and candidate divisions OD1 and RF3. Lower abundance divisions included OTUs for *Bacilli*, *Clostridia*, *Spirochaetes*, *Lentisphaeria*, TM7, *Opitutae*, *Verrucomicrobia*, Bhi80-139, Bd1-5, SR1 and *Chlamydiae* (Figure 2A). The dominant eucaryal OTUs were for photosynthetic

Deleted: previously
Deleted: recorded
Deleted: .
Deleted: in
Deleted: N
Deleted: P
Deleted: .
Deleted: (*diversity indices)

Chlorophyta (green algae) and *Dictyochophyceae* (silicoflagellate algae) (Figure 2B) principally assigned to the genus *Dunaliella* and the order *Pedinellales*, respectively (Supplementary Table S4). Lower abundance eucaryal OTUs included *Bacillariophyta* (diatoms), *Dinophyceae*, *Fungi* and heterotrophic *Choanoflagellida* and *Ciliophora* (see Supplementary Table S4 for lower taxonomic rank assignments).

Variation of microbial composition according to size and depth

Community composition varied with size fraction and depth. This was supported by seriation analysis that showed samples clustered according to size fraction, and those clusters further separated into upper mixed and deep zone groups (Figure 3). A significant difference in genus-level composition between the upper mixed and deep zones was supported by ANOSIM test (Rho: 0.53, significance: 0.1%). Differential vertical distribution of taxa is consistent with partitioning of ecological functions in the lake and in association with the physical and chemical data, described functional roles of those taxa.

20–3.0 µm fraction community composition

The upper mixed zone samples had a relatively high OTU abundance of *Dunaliella* chloroplasts and chlorophyte algae consistent with large active photosynthetic organisms concentrating near surface light. They are likely the main source of primary production in Organic Lake and have previously been reported to be the dominant algae (Franzman *et al.*, 1987b). The SSU sequences for these algae at the bottom of the lake are likely to be due to sedimentation of dead cells or resting cysts.

Psychroflexus OTUs were overrepresented in the surface and 6.7 m samples. Consistent with enrichment on the 3.0 µm filters, *Psychroflexus* (formerly *Flavobacterium*) *gondwanensis* (Bowman *et al.*, 1998) isolated from Organic Lake (Franzmann *et al.*, 1987b) had cells 1.5–11.5 µm in length (Dobson *et al.*, 1991). *Flavobacteria* associate with phytoplankton blooms in the Southern Ocean (Abell and Bowman 2005a; Abell and Bowman 2005b; Williams *et al.*, 2012b), and have specialized abilities to degrade polymeric substances from algal exudates and detritus (reviewed in Kirchman, 2002; Williams *et al.*, 2012b). It is likely that Organic Lake *Psychroflexus* fills a similar ecological role. In support of this, *Psychroflexus* OTUs cluster with *Dunaliella* chloroplasts in the seriation analysis (Figure 3) and *P. gondwanensis* abundance in Organic Lake has been correlated with average hours of sunshine per day indicating population dynamics that is related to summer algal blooms (James *et al.*, 1994). The *Psychroflexus* OTUs in the deep zone are most likely due to sedimentation as *P. gondwanensis* is non-motile and strictly aerobic (Dobson *et al.*, 1991).

Roseovarius OTUs were enriched at 4.2 m and 6.5 m suggesting different ecotypes may be present in the upper mixed zone compared to the deep zone. *Roseovarius tolerans*, an isolate from Ekho Lake in the Vestfold Hills, Antarctica has a cell size (1.1–2.2 µm; Labrenz *et al.*, 1999) that would be expected to be captured on the 0.8 µm filter. The *Roseovarius* captured on the 3 µm filter may therefore be a different species, or a strain similar to *R. tolerans* from Ekho Lake that exhibits different growth characteristics (*i.e.* larger cell size or forms aggregates). A strain of this species from Ekho Lake is capable of microaerophilic growth (Labrenz *et al.*, 1999). Overrepresentation at 6.5 m may therefore be indicative of growth at that depth rather than sedimentation because sinking cells would be more abundant close to the lake bottom at 6.7 m. *Roseovarius* OTUs cluster with *Dunaliella* chloroplast and *Psychroflexus* OTUs in the seriation analysis (Figure 3), suggesting that Organic Lake *Roseovarius* may be utilizing compounds released from algal-derived particulate matter, or made available by processing of

Deleted: &

Deleted: &

Deleted: consistent with *Roseovarius* capture on the 3 µm filter. A strain

complex organic matter by *Psychroflexus*. *Roseovarius* is a member of the *Roseobacter* clade, which is inferred to have an opportunistic ecology frequently associated with nutrient-replete plankton aggregates, including by-products of flavobacterial exoenzymatic attack (Moran *et al*., 2007; Teeling *et al*., 2012). Additionally, the diverse metabolic capabilities of the *Roseobacter* clade include DMSP degradation, AAnP and CO oxidation (reviewed in Wagner-Döbler and Biebl, 2006). All of these capabilities should facilitate growth in both the upper mixed and deep zones of Organic Lake (see *Carbon resourcefulness in dominant heterotrophic bacteria* below).

3–0.8 µm size fraction community composition

On the 0.8 µm filter, OTUs for *Marinobacter* dominated at all depths except 6.5 m. Their capture on this size fraction is consistent with the cell size of isolates (1.2–3 µm) (Gauthier *et al.*, 1992). The genus is metabolically versatile, which likely permits it to occupy the entire water column. *Marinobacter* is heterotrophic and the genus includes hydrocarbon-degrading strains (e.g., Gauthier *et al*., 1992; Huu *et al*., 1999), although deep-sea metal-oxidizing autotrophs have also been reported (Edwards *et al.*, 2003). Some isolates are capable of interacting with diatoms (Gärdes *et al*., 2010) and dinoflagellates (Green *et al.*, 2006). *Marinobacter* isolates from Antarctic lakes are capable of anaerobic respiration using dimethyl sulfoxide (DMSO) (Matsuzaki *et al*., 2006) or nitrate (Ward and Priscu, 1997). Analysis of functional potential linked to *Marinobacter* revealed additional metabolic capabilities potentially related to its dominance in Organic Lake (see *Carbon resourcefulness in dominant heterotrophic bacteria* and *Molecular basis for unusual sulfur chemistry* below).

OTUs for RF3 and *Halomonas* were overrepresented at 6.5 m, and RF3 sequences were more abundant (Figure 2 and 3). Their relative abundance in the deep zone indicates a role in microaerophilic processes. The majority of RF3 sequences to date are from anaerobic environments including mammalian gut (Tajima *et al.*, 1999; Ley *et al*., 2006; Samsudin *et al*., 2011), sediment (Yanagibayashi *et al.*, 1999; Röske *et al.*, 2012), municipal waste leachate (Huang *et al.*, 2005), anaerobic sludge (Chouari *et al.*, 2005; Goberna *et al*., 2009; Rivière *et al.*, 2009; Tang *et al*., 2011), a subsurface oil well head (Yamane *et al.*, 2011), and the anaerobic zone of saline lakes (Humayoun *et al*., 2003; Schmidtova *et al*., 2009; Bowman *et al.*, 2000). However, some members have been found in surface waters (Demergasso *et al*., 2008; Xing *et al.*, 2009; Yilmaz *et al*., 2012) suggesting not all members are strict anaerobes. Several *Halomonas* isolates have been sourced from Organic Lake including two described species *Halomonas subglaciescola* and *H. meridiana*, both of which grow as rods with dimensions consistent with capture on this size fraction (Franzmann *et al.*, 1987a; James *et al.*, 1990). Despite these isolates being aerobic, *Halomonas* has been reported to be enriched at the oxycline in Organic Lake (James *et al.*, 1994) indicating *Halomonas* in the lake plays an ecological role in the suboxic zone. This capacity may be linked to the ability of free amino acids and organic acids (which are abundant in the deep zone) to stimulate the growth of isolates (Franzmann *et al.*, 1987a).

0.8–0.1 µm size fraction community composition

A large number of eucaryal sequences were evident in the 0.1 µm size fraction. The upper zone was overrepresented by OTUs for *Pedinellales* (silicoflagellate algae) that co-varied with chloroplasts (Figure 2 and 3). *Pedinellales* have only been detected in Antarctic lakes from molecular studies (Unrein *et al.*, 2005; Lauro *et al.*, 2011) including Organic Lake (Yau *et al.*, 2011), and light microscopy studies of Antarctic Peninsular freshwater lakes reported 5–8 µm diameter cells resembling *Pseudopedinella* (Unrein *et al.*, 2005). It is possible that in Organic

Deleted: aerobic anoxygenic phototrophy (

Deleted:)

Deleted: &

Deleted: *Marinobacter*is

Deleted: affiliated with this genus

Deleted: &

Deleted: 2000b

Deleted:). Subsequent

Deleted: analyses in freshwater lakes

Deleted: the

Deleted: that were 5–8 µm in diameter

Lake small (0.8–0.1 µm) free-living members or chloroplast containing cyst forms (Thomsen, 1988) exist. However, without evidence to support this (*e.g.* by microscopy) it seems more likely that the lake sustains a relatively small number of active photosynthetic cells and the sequences detected arise from cysts or degraded cellular material.

OTUs for *Candidatus* "Aquiluna", in the Luna-1 cluster of *Actinobacteria* (Hahn *et al.*, 2004; Hahn *et al.*, 2009) were most abundant at 1.7 m. The genus has small cells (<1.2 µm; Hahn *et al.*, 2009), accounting for their concentration on this size fraction. Although originally described in freshwater lakes, the same clade was detected in abundance in Ace Lake (Lauro *et al.*, 2011) and surface Artic seawater (Kang *et al.*, 2012) demonstrating that they play ecological roles in polar saline systems. In Ace Lake surface waters they were associated with utilization of labile carbon and nitrogen substrates (Lauro *et al.*, 2011), and in Organic Lake surface waters they probably perform similar functions. The presence of this clade in the deep zone implies a facultative anaerobic lifestyle or sedimented cells.

The bottom of the water column was distinguished by the presence of OTUs for candidate divisions OD1 and TM7. OD1 was more abundant, and its prevalence on this size fraction is consistent with similar findings for size fractionation of ground water (Miyoshi *et al.*, 2005). OD1 is consistently associated with reduced, sulfur-rich, anoxic environments (Harris *et al.*, 2004; Elshahed *et al.*, 2005). OD1 from Zodletone Spring, Oklahoma, was reported to possess enzymes related to those from anaerobic microorganisms (Elshahed *et al.*, 2005). Genomic analyses identified OTUs for OD1 in the anoxic zone of Ace Lake (Lauro *et al.*, 2011). The distribution of OD1 in Organic Lake is consistent with an anaerobic metabolism and potential involvement in sulfur chemistry.

Organic Lake functional potential

To determine the potential for functional processes in Organic Lake, gene markers for carbon, nitrogen and sulfur conversions (Figure 4) were retrieved from metagenomic reads. BEST analysis showed that variation in the population structure was significantly correlated (Rho: 0.519, significance: 0.3%) with the abiotic parameters, DO, temperature, TS and TN. The DO gradient has an obvious effect of separating aerobic from anaerobic taxa, and allows oxygen sensitive nitrogen and sulfur processes to occur in the deep zone. Functional potential, taxonomic composition and the physico-chemical data were integrated to infer the carbon, nitrogen and sulfur cycles.

Carbon resourcefulness in dominant heterotrophic bacteria

In both the upper mixed and deep zones, potential for carbon fixation was much lower than for degradative processes, indicating potential for net carbon loss (Figure 4A). Potential for carbon fixation via the oxygen-tolerant Calvin cycle (Figure 4A) was assessed by presence of the marker genes ribulose-bisphosphate carboxylase (RuBisCO) and phosphoribulokinase (*prkB*) (Hügler and Sievert, 2011). The majority of RuBisCO homologs were related to *Viridiplantae* (Table 2) supporting the ecological role of green algae as the principle photosynthetic organisms. RuBisCO was only associated with a small proportion of *Gammaproteobacteria* (Table 2), principally from sulfur-oxidizing *Thiomicrospira*, indicating some *Gammaproteobacteria* are autotrophs. However, the majority of *prkB* matched to *Gammaproteobacteria* (Table 2), predominantly *Marinobacter*. Although deep-sea, iron-oxidizing autotrophic members of *Marinobacter* have been isolated (Edwards *et al.*, 2003), all genomes reported for *Marinobacter* have *prkB* but lack RuBisCO genes. Across *Marinobacter* genomes the *prkB* homolog is

Deleted: C
Deleted: N
Deleted: C, N
Deleted: S
Deleted: N
Deleted: S
Deleted: C, N
Deleted: S
Deleted: in Organic Lake
Deleted: &

consistently adjacent to a gene for a putative phosphodiesterase, suggesting that the enzymes expressed by these genes may be involved in a pathway involved in pentose phosphate metabolism unrelated to carbon fixation. Albeit exceptional, this decoupling of *prkB* from RuBisCO involved in carbon fixation (forms I and II), also observed in *Ammonifex* (Hügler and Sievert, 2011), undermines the utility of *prkB* as a marker gene for the Calvin cycle within certain groups. Thus, there is no evidence for autotrophy in Organic Lake mediated by *Marinobacter*.

Evidence for carbon fixation via the reverse tricarboxylic acid (rTCA) cycle was also indicated, with genes for ATP citrate lyase (*aclAB*) linked to sulfur-oxidizing *Epsilonproteobacteria* (Supplementary Figure S6A). In general, the rTCA cycle is restricted to anaerobic and microaerophilic bacteria (Hügler and Sievert, 2011), which is consistent with the detection of *Epsilonproteobacteria* in the lake bottom where oxygen is lowest, and the microaerophilic/anaerobic metabolisms characteristic of the group (Campbell *et al.*, 2006). Anaerobic carbon fixation was represented by potential for the Wood-Ljungdahl (WL; or reductive acetyl-CoA) pathway (Figure 4A). WL-mediated carbon fixation, for which CO dehydrogenase/acetyl-CoA synthase is the key enzyme, was linked to *Firmicutes* and *Deltaproteobacteria* that are known to grow autotrophically using this pathway (Hügler and Sievert, 2011).

Potential for carbon loss by via respiration was indicated by an abundance of cytochrome C oxidase genes (*coxAC*) throughout the water column. In the deep zone, potential for fermentation was greatest at 6.5 m (Figure 4A) and likely the main biological activity that was occurring at that depth. Fermentation was indicated by the marker gene lactate dehydrogenase (*ldh*) that competes for pyruvate at the acetate-lactate branch point. These genes were linked to *Firmicutes* (Table 2), which was only present at 6.5 m and represented by the classes *Clostridia* and *Bacilli* (Figure 2A). As the related candidate division RF3 (Tajima *et al.*, 1999) also has relatively high abundance in this zone (Figure 2A) (see *0.8–3.0 µm size fraction community composition* above), there is circumstantial evidence that RF3 possesses fermentative metabolism and may therefore play an important ecological role in Organic Lake by degrading high molecular weight compounds to organic acids that other organisms can utilize. Assimilation of fermentation products appears to play a greater role in Organic Lake rather than complete anaerobic oxidation involving methanogens or sulfate-reducing bacteria; the former were absent and the latter were present in low abundance (Figure 2A, 2C).

Alphaproteobacteria, predominantly *Roseovarius* (Figure 2C), were implicated in CO oxidation (Table 2; Supplementary Figure S6A), which is used to generate energy for lithoheterotrophic growth (Moran and Miller, 2007), although CO oxidation may also be involved in anaplerotic C fixation (Moran *et al.*, 2007). The CO oxidation capacity was at a maximum at 6.5 m (Figure 4A), and therefore associated with the deep-zone *Roseovarius* ecotype of Organic Lake. CO oxidation can function as a strategy to limit oxidation of organic carbon for energy so that a greater proportion can be directed towards biosynthesis (Moran and Miller, 2007).

Photosynthesis reaction center genes *pufLM*, involved in photoheterotrophy via AAnP, were abundant in Organic Lake (Figure 4A; Table 2). These were linked to the *Roseobacter* clade of *Alphaproteobacteria* (Table 2), major contributors to AAnP in ocean surface waters (Béjà *et al.*, 2002; Moran et al., 2007). This is consistent with the known metabolic potential of bacteriochlorophyll A (BchlA) producing *Roseovarius tolerans* from Ekho Lake (Labrenz *et al.*, 1999). Photoheterotrophy can also be rhodopsin-dependent, with proteorhodopsins (PRs) of

Deleted: &

Deleted:

Comment [TW3]: Wrong figure.

Deleted: &

Deleted: C

Deleted: &

Deleted: as

Deleted:)(

Deleted: could

Deleted: &

Deleted: &

Formatted: Line spacing: single

Deleted:

marine *Flavobacteria* and *Vibrio* previously linked to light-dependent energy generation to supplement heterotrophic growth, particularly during carbon limitation (Gómez-Consarnau *et al*., 2007; Gómez-Consarnau *et al*., 2010). However, the function(s) of rhodopsins are diverse, and PRs are also hypothesized to be involved in light or depth sensing (Fuhrman *et al.*, 2008).

Rhodopsin genes were abundant in Organic Lake (Figure 4A), and were associated with all the dominant Organic Lake aerobic heterotrophic lineages (Supplementary Figure S6). Phylogenetic analysis revealed six well-supported Organic Lake rhodopsin groups (Supplementary Figure S6). All groups had an L or M residue at position 105 (*vs* the SAR86 PR), denoting tuning to surface green light (Man *et al.*, 2003; Gomez-Consarnau *et al.*, 2007), and is characteristic of oceanic coastal samples (Rusch *et al.*, 2007). Four of the groups clustered with homologs of genera detected in the lake, namely *Marinobacter*, *Psychroflexus*, *Octadecabacter* and "*Ca.* Aquiluna" (Supplementary Figure S6 and Table S4). Another group (SAL-R group) originates from the sphingobacterium *Salinibacter ruber*, which produces xanthorhodopsin (Balashov *et al*., 2005); it is therefore likely that Organic Lake *Sphingobacteria* (Supplementary Table S4) were the origin of this rhodopsin group. The most abundant group (OL-R1; Supplementary Figure S6) had no close homologs from GENBANK, but it was abundant on the 3.0 µm fraction and has a distribution suggesting it originates from Organic Lake members of the *Roseobacter* clade (Figure 4A). All ORFs adjacent to OL-R1 rhodopsin containing scaffolds were related to *Octadecabacter* further supporting their *Roseobacter* clade provenance (Supplementary Figure S7). Genes downstream of OL-R1 were involved in carotenoid synthesis, indicating OL-R1 is a xanthorhodopsin, occuring as a retinal protein or in a carotenoid complex (Balashov *et al.*, 2005).

Photoheterotrophic potential of Organic Lake was compared with other aquatic environments including nearby Ace Lake, Southern Ocean (SO) and GOS expedition samples. The Organic Lake 0.1 µm fraction had the lowest rhodopsin counts and percentage of rhodopsin containing cells of all size-matched samples surveyed (Table 3). Non-marine GOS samples from the 0.1 µm fraction have been noted to have lower rhodopsin abundance (Sharma *et al.*, 2008), which was similarly evident from our analysis (Table 3). In contrast, the 3.0 µm Organic Lake size fractions had higher rhodopsin counts than Ace Lake and comparable counts to the SO samples, although the percentage of rhodopsin containing cells was still lower than that of the SO. The paucity of rhodopsins in the Organic Lake 0.1 µm fraction is likely due to the lack of SAR11 clade, which is expected to be the main source of rhodopsin genes in Ace Lake and marine samples. This indicates that although Organic Lake has an overall lower frequency of rhodopsin genes compared to sites for which size fraction-matched metagenomes are available, the rhodopsins associated with larger or particle-associated cells are as abundant as in the marine environment.

Counts of *pufLM* genes in the Organic Lake 0.1 µm size fraction were similar to GOS samples, except for Punta Cormorant hypersaline lagoon which had the highest *pufLM* counts and percentage of AAnP cells (Table 3). However, the highest overall counts of *pufLM* were from the 3.0 µm size fraction of Organic Lake, likely due to the high proportion of members of the *Roseobacter* clade. Notably, *pufLM* genes were not detected in high abundance in Ace Lake or the Southern Ocean samples, indicating AAnP is a unique adaptation in Organic Lake among these polar environments. The similarly high abundance of *pufLM* genes in Punta Cormorant hypersaline lagoon indicates AAnP may be advantageous in environments with salinity above marine levels.

Deleted: ,

Deleted: was

Deleted: which occur

Deleted: /

Deleted: the

Deleted: in

Deleted: Interestingly

Deleted: likely

Deleted: .

Deleted: overall

Deleted: -

Deleted: at all

Deleted: of

Deleted: salinity.

The contribution of light-driven energy generation processes to the carbon budget is difficult to infer from genetic potential alone. For example, the relative abundance of AAnP and PR genes in Arctic bacteria has been reported to be the same in winter and summer (Cottrell and Kirchman, 2002). Furthermore, regulation of pigment synthesis is complex; for example, BchlA expression in *R. tolerans* occurs in the dark but is inhibited by continuous dim light (Labrenz *et al.*, 1999). However, it is possible that the apparent negative balance in carbon conversion potential could be augmented by photoheterotrophy performed by bacterial groups that are abundant in Organic Lake. In particular, the Organic Lake *Psychroflexus* could play a particular role as it has a PR related to *Dokdonia*, which was shown to function under carbon-limitation (Gómez-Consarnau *et al.*, 2007). Furthermore, detection of higher AAnP potential in Organic Lake than other aquatic environments linked with taxa known to be capable of AAnP, suggests it may have a greater influence in the carbon budget of Organic Lake.

Regenerated nitrogen is predominant in the nitrogen cycle

Nitrogen cycling potential throughout the lake was dominated by assimilation and mineralization/assimilation pathways (Figure 4B). Glutamate dehydrogenase (GDH) genes (*gdhA*) were abundant (Figure 4B), and linked predominantly to *Alpha-* and *Gammaproteobacteria* and to a lesser extent *Bacteroidetes* (Table 2). However, the significance of the readily reversible GDH depends on its origin; *Bacteroidetes* are likely to use GDH in the oxidative direction for glutamate catabolism (Takahashi *et al.*, 2000; Williams *et al.*, 2012b), whereas the use of GDH in the oxidative or reductive directions by *Proteobacteria* is likely to depend upon the source of reduced nitrogen (ammonia *vs* amino acids). Glutamine synthetase (*glnB*) and glutamate synthase genes (*gltBS*), were predominantly linked to *Alpha-* and *Gammaproteobacteria* (Table 2), indicating the potential for high-affinity ammonia assimilation by these groups in Organic Lake. The high ammonia concentration in the deep zone (Figure 1B; Table 1) would result from a higher rate of mineralization (ammonification) than assimilation. This is consistent with abundant OTUs for *Psychroflexus* (*Bacteroidetes*) in this zone, and due to either turnover of organic matter or lysis of *Bacteroidetes* cells after sedimentation in anoxic water. In addition, the gene for ammonia-generating nitrite reductase (*nrfA*) was linked to *Bacteroidetes* and *Planctomycetes* (Table 2), indicating ammonia may also be produced by these putative aerobic heterotrophs. Overall, the data suggest that ammonia is actively assimilated in the aerobic upper mixed zone, but is permitted to accumulate in the anaerobic deep zone.

Potential for nitrogen conversions typically found in other aquatic environments was greatly reduced in Organic Lake. There was a very low potential for nitrogen fixation that was confined to the deep zone (Figure 2B) and principally linked to anaerobic *Epsilonproteobacteria* (Table 2). This diazotrophic potential may not be realized by nitrogen-fixing *Epsilonproteobacteria*, given the high ammonia concentration present in the deep zone. No ammonia monooxygenase genes (*amoA*) were detected. The potential for ammonia oxidation was only represented by hydroxylamine/hydrazine oxidase-like (*hao*) genes, which were in low abundance and linked to *Deltaproteobacteria* (Table 2). *hao* genes are present in non-ammonia-oxidizing bacteria (Bergmann *et al.*, 2005), and those from Organic Lake belong to a family of multiheme cytochrome c genes present in sulfate-reducing *Deltaproteobacteria* that have no proven role in ammonia oxidation. In the genomes of sulfate-reducing *Deltaproteobacteria* the *hao* gene is invariably situated adjacent to a gene for a NapC/NirT protein, which suggests a role in dissimilatory reduction. Collectively these data indicate an inability for nitrification to occur in the upper mixed zone and no potential for ammonia loss in the deep zone.

Deleted: &

Deleted: exampleBchlA

Deleted: C

Deleted: Also

Deleted:

Deleted: which

Deleted: indicate

Deleted: of the bacteroidetan

Deleted: bacteroidetan

Deleted:); this offers another potential avenue for

Deleted: production

Deleted: have been documented

Deleted:)

Deleted: inspection of the genes in

Deleted: reveal they

Deleted: cytochromes

Deleted: this

Denitrification genes (*norCB* and *nozB*) and genes for nitrate assimilation (*nasA*) were present throughout the water column (Figure 4B) and were linked primarily to *Gammaproteobacteria* (Table 2). Low nitrate and nitrite in the deep zone (Figure 1B, Table 1) indicates oxidized nitrogen has been depleted by dissimilatory or assimilitory reduction by heterotrophic *Gammaproteobacteria*. Denitrification genes are phylogenetically widespread and usually induced by low oxygen or oxidized nitrogen species (Kraft *et al.*, 2011) and thus expected to be active in the deep zone or oxycline. However, denitrification may be inhibited even if conditions appear appropriate. For example, in Lake Bonney, Antarctica, denitrification occurs in the west lobe, but not in the east lobe of the lake despite the presence of anoxia, nitrate and denitrifying *Marinobacter* species (Ward and Priscu, 1997; Ward *et al.*, 2005). Moreover, in the absence of nitrification, denitrification and nitrate assimilation would be limited by the lack of potential to re-form oxidized nitrogen. The preponderance of assimilation/mineralization pathways geared towards reduced nitrogen appears to reflect a "short circuit" of the typical nitrogen cycle that would conserve nitrogen in a largely closed system. Hence, the predominant nitrogen source is regenerated fixed nitrogen. Similar findings were also made for Ace Lake, although in this system the presence of a dense layer of green sulfur bacteria with the potential to fix nitrogen augments the nitrogen cycle (Lauro *et al.*, 2011).

Molecular basis for unusual sulfur chemistry

Several meromictic hypersaline lakes in the Vestfold Hills, including Organic Lake, with practical salinity greater than ~150 are characterized by an absence of hydrogen sulfide and photoautotrophic sulfur bacteria (Burke and Burton, 1988). Although sulfate is present (Franzmann *et al.*, 1987b), geochemical conditions of these lakes are not conducive to dissimilatory sulfur cycling between sulfur oxidizing and sulfate reducing bacteria typical of other stratified systems such as Ace Lake (Ng *et al.*, 2010; Lauro *et al.*, 2011). Consistent with this, potential for dissimilatory sulfate reduction represented by dissimilatory sulfite reductase (*dsrAB*) and adenylylsulfate reductase (*aprAB*) linked to sulfate-reducing *Deltaproteobacteria* (Table 2) was low in Organic Lake. Sulfate-reduction potential was confined to the 6.7 m sample (Figure 4C) where oxygen concentration was lowest and *Deltaproteobacteria* were present (Figure 2A).

Capacity for oxidation of reduced sulfur compounds, represented by the sulfur oxidation multienzyme genes (*soxAB*), was present throughout the water column (Figure 4C) and linked primarily to *Alpha-* and *Gammaproteobacteria* (Table 2). Sulfur-oxidizing *Gammaproteobacteria* and *Alphaproteobacteria* are known to oxidize sulfur compounds, such as thiosulfate, aerobically. Although a small proportion of *Gammaproteobacteria* had the capacity for autotrophy (see *Carbon resourcefulness in dominant heterotrophic bacteria*), the majority of sulfur-oxidizers were likely chemolithoheterotrophs as they were related to heterotrophic *Marinobacter* and *Roseobacter*. The sulfur dehydrogenase genes *soxCD* linked to *Alpha-* and *Gammaproteobacteria* were similarly present throughout the water column. *soxCD* are accessory components of the Sox enzyme system without which complete oxidation of thiosulfate cannot occur (Friedrich *et al.*, 2005). Thus the presence of *soxCD* indicates complete oxidization likely occurs, although the different distribution of *soxAB* and *soxCD* in the water column (Figure 4C) suggests a proportion of the community may lack *soxCD* and deposit sulfur. Sulfur-oxidizing *Epsilonproteobacteria* possessing *soxAB* genes (Table 2) were present only in the deep zone of Organic Lake (Figure 2A,C) and were related to autotrophic deep sea sulfur-oxidizers, some members of which are capable of anaerobic sulfur oxidation using nitrate (Yamamoto and Takai,

Deleted: &

Comment [TW4]: I have to admit, I don't fully understand this section… perhaps FL can check.

Deleted: &

Deleted: in the

Deleted: &

2011). It is unlikely that appreciable sulfur oxidation occurs in the deep zone as the known terminal electron acceptors, oxygen and nitrate, are deplete and the abundance of sulfur oxidizing *Epsilonproteobacteria* is low (Figure 2A).

It is likely that the limited anaerobic dissimilatory sulfur cycle contributes to the accumulation of DMS in Organic Lake in the deep zone. In the upper mixed zone, DMS could potentially be oxidized as a carbon and energy source or utilized as an electron donor by sulfur-oxidizing bacteria (Schäfer *et al.*, 2010). In anoxic zones, methanogenic *Archaea* or sulfate-reducing bacteria are the main organisms known to breakdown DMS (Schäfer *et al.*, 2008). Methanogens and genes involved in methanogenesis were not detected, nor has methane been detected (Gibson *et al.*, 1994) leaving sulfate-reduction the most likely route of DMS catabolism. The low dissimilatory sulfate reduction potential in the deep zone coupled with the relatively stagnant waters would likely minimize DMS oxidation and loss by ventilation. DMS would therefore be expected to accumulate in the deep zone if production rates were higher than breakdown.

To determine the source of high DMS in the bottom waters of Organic Lake, the genes involved in DMS formation were surveyed. Genes for DMSP lyases *dddD*, *dddL* and *dddP*, were detected in Organic Lake at levels comparable to other dominant processes such as respiration and fermentation (Figure 4C) indicating DMSP is an important carbon and energy source in Organic Lake. *dddD* was the most abundant of the Organic Lake DMSP lyases (Table 3) and comprised two main types: MAR-dddD and OL-dddD (Supplementary Figure S8). Neither of these types clustered with the non-functional *Dinoroseobacter shibae* DFL 12 and *Ruegeria pomeroyi* DSS-3 *dddD* homologs (Todd *et al.*, 2011) or carnitine coenzyme A transferase outgroups, thereby providing support for their proposed role as functional DMSP lyases. The MAR-dddD type includes the *Marinobacter* sp. ELB17 *dddD* homolog, and MAR-dddD sequences were most abundant on the 0.8 µm fraction where *Marinobacter* OTUs were also abundant, indicating MAR-dddD derives from Organic Lake *Marinobacter* (Supplementary Figure S8). OL-dddD did not have a close relative from cultured bacteria making its precise taxonomic origins uncertain. The abundance of OL-dddD on the 3.0 µm fraction suggests it originates from *Alphaproteobacteria*. OL-dddD containing contigs carried genes of mixed *Alpha*- and *Gammaproteobacterial* origin supporting its provenance from one of these classes and consistent with the "pick 'n mix" arrangement of genes found beside sequenced *dddD* regions (Johnston *et al.*, 2008). Upstream of OL-dddD was *dddT* (Supplementary Figure S**), a betaine, choline, carnitine transporter (BCCT) family protein that likely functions in substrate import, demonstrating OL-dddD forms an operon-like structure similar to *Halomonas* sp. HTNK1 (Todd *et al.* 2010).

Two *dddL* groups were detected in Organic Lake: SUL-dddL and MAR-dddL (Supplementary Figure S9). The former includes the *Sulfitobacter* sp. EE-36 *dddL* and the latter the *Marinobacter manganoxydans* MnI7-9 homolog indicating they originate from *Roseobacter*-clade and *Gammaproteobacteria*, respectively. *Sulfitobacter* sp. EE-36 has demonstrated DMSP lyase activity and the *dddL* gene alone is sufficient for DMS generation (Curson *et al.*, 2008). These data indicate that the Organic Lake members of the SUL-dddL group perform the same functional role. The MAR-dddL clade appears to be an uncharacterized branch of the *dddL* family. *dddP* was detected as the least abundant of the DMSP lyases (Table 3). Phylogenetic analyses showed Organic Lake *dddP* likely originate from *Roseovarius* (Supplementary Figure S10). The Organic Lake sequences formed a clade with the functionally verified *Roseovarius nibinhibens* ISM *dddP* (Todd *et al.*, 2009).

Deleted: S

Formatted: Indent: First line: 0.76 cm

Formatted: Font: Times New Roman, 12 pt

Deleted: ¶ ¶

Formatted: Line spacing: single

Deleted: ,.

Deleted:

Comment [U5]: Which figure or delete?

Formatted: Font: Not Italic

Formatted: Pattern: Clear

Deleted: ***),

Deleted: like that of

A single type of DMSP demethylase, *dmdA* was identified. It clustered with *Roseobacter*-clade *dmdA* (Supplementary Figure S11), corresponding to the marine clade A (Howard *et al.*, 2006), and includes the functionally verified *R. pomeroyi* DSS-3 homolog. These data indicate that the Organic Lake sequences correspond to true DMSP demethylases and not related glycine cleavage T proteins or aminomethyltransferases (Howard *et al.*, 2006).

DMSP cleavage appears to be a significant source of DMS in Organic Lake. DMSP likely originates from *Bacillariophyta* or *Dinoflagellida* as Organic Lake *Dunaliella* have been reported not to produce DMSP in culture (Franzmann *et al.*, 1987b). Based on the abundance of marker genes, DMSP cleavage is predicted to occur at highest levels in the deep zone (Figure 4C) where the DMS concentration has been measured to be highest (Deprez *et al.*, 1986; Franzmann *et al.*, 1987; Gibson *et al.*, 1991; Roberts and Burton 1993a; Roberts *et al.*, 1993b). DMS can also be produced in anoxic environments from the reduction of DMSO, degradation of sulfur containing amino acids, and sulfide methylation (Schäfer *et al.*, 2010). Our data indicate that some DMSO reduction linked to *Firmicutes* could occur, but is not likely a major pathway (Figure 4C), and the potential for the other DMS yielding processes could not be determined because the pathways have not been established. When cultivated, *Halomonas* isolates from Organic Lake produced DMS from cysteine (Franzmann *et al.*, 1987b) providing some evidence that DMS production from anaerobic degradation of amino acids can occur.

The potential for DMSP cleavage was more than twice that of DMSP demethylation (Figure 4C). This is unusual compared to the marine environment or Ace Lake where DMSP demethylation potential is much higher than cleavage (Table 3). Previous estimates have similarly shown marine environments to have demethylation potential up to two orders of magnitude higher than cleavage (Howard *et al.*, 2008; Todd *et al.*, 2009; Todd *et al.*, 2011b; Reisch *et al.*, 2011). The frequency of DMSP lyase genes *dddD* and *dddL* in Organic Lake exceeded those of all other environments, except Punta Cormorant hypersaline lagoon, where *dddL* abundance was comparable (Table 3). This suggests selection in Organic Lake for DMSP cleavage due to functional advantage and/or selection for taxa that carry DMSP lyase genes. There is evidence that high DMSP cleavage potential is adaptive in hypersaline systems, as a high proportion of *ddd* genes were similarly detected in Punta Cormorant hypersaline lagoon and saltern ponds (Raina *et al*., 2010). The accumulated DMS in Organic Lake suggests conditions in Organic Lake favor the relatively wasteful lysis pathway, where both sulfur and carbon is lost to the organism performing the DMSP lysis, over the more 'thrifty' demethylation pathway. This is particularly pertinent to the *Roseobacter* lineages that can also perform either process. One possibility that has been proposed is that when sulfur is in excess and the organism can easily assimilate alternative sulfur sources, the lysis pathway may be competitive (Johnston *et al.*, 2008). This may be particularly the case in hypersaline systems if higher concentrations of DMSP are being produced as an osmolyte.

Conclusion

Through the use of shotgun metagenomics and size partitioning of samples, we discovered that the Organic Lake system is dominated by heterotrophic bacteria related to *Psychroflexus*, *Marinobacter* and *Roseovarius* with primary production provided largely by chlorophyte algae related to *Dunaliella*. Genetic potential for oxidation of fixed carbon by heterotrophic bacteria occurs greatly in excess of carbon fixation, suggesting possible net carbon loss. However, by linking key metabolic processes to the dominant heterotrophic lineages we uncovered processes that were unusually abundant in Organic Lake that may serve to maximize exploitation of limited

Deleted: &

Deleted: as well as

Deleted:

Deleted: .(

Deleted:)

Deleted: is

Deleted: partioning

Deleted: the

resources and minimize loss. Recalcitrant polymeric algal material and particulate matter is likely remineralized by *Psychroflexus* in the upper mixed zone and by *Firmicutes* in the deep zone to provide labile substrates for use by other heterotrophic bacteria. The generalist *Marinobacter* and *Roseovarius* lineages were associated with abundant genes involved in rhodopsin-mediated and AAnP photoheterotrophy; the latter of which was more abundant in Organic Lake than any other system surveyed. Potential for chemolithoheterotrophy, sulfur oxidation and CO oxidation was also high, and along with photoheterotrophy, may provide a supplementary energy source if organic carbon becomes limiting.

In addition to being able to describe the functional capacities and potential importance of poorly understood microbial processes occurring in the lake (*e.g.* photoheterotrophy by *Alphaproteobacteria*), we were able to answer targeted questions about the biology of the unusual lake sulfur chemistry. The low potential for dissimilatory sulfur cycling (both S oxidation and DSR) and relatively stable waters of the deep zone, combined with the generation of DMS from DMSP, facilitate the accumulation of a high level of DMS in the lake. It appears *Marinobacter* and *Roseovarius* play a key role in DMS formation by cleaving DMSP generated by upper mixed zone eucaryal algae. The remarkable abundance of DMSP lyase genes suggests DMSP is a significant carbon source in Organic Lake and the cleavage pathway provides a selective advantage under the unique constraints of the Organic Lake environment.

In view of the the minimal capacity for biological fixation of carbon and nitrogen, and yet organic richness, including high levels of DMS, in Organic Lake, we evaluated what input the lake may have received throughout its relatively brief ~3 000 year history. The volume of the lake is small (~6 x 10^4 m^3), and exogenous input may occur from guano deposited in a small penguin rookery nearby the lake, through Giant Petrel or Skua grazing and defecation, and/or by decaying animal carcasses such as elephant seals which can weigh on the order of 1 ton and are present near the lake. It is also possible that during isolation from the ocean, the base of the water column in the marine basin that formed the lake may have acted as a sump for organic material. Phytoplankton blooms and benthic mats tend to make coastal marine basins very productive, and organic matter that sediments out of the surface waters will become trapped in the denser, more saline bottom layers (Bird *et al.*, 1991). Retention of captured organic matter in the lake may also have been facilitated by Organic Lake having become highly saline quickly (Bird *et al.*, 1991). Studies in the future that experimentally determine exogenous input and historical lake dynamics (*e.g.* stable isotope and biomarker analyses of lake sediment), and metaproteogenomic analyses of interannual community composition and function, will provide improved knowledge of the unusual biogeochemistry of Organic Lake and better enable predictions to be made about how the lake may be affected by ecosystem changes.

Acknowledgements

This work was supported by the Australian Research Council and the Australian Antarctic Science program. The authors acknowledge assistance from the J. Craig Venter Institute and the Gordon and Betty Moore Foundation. We thank John Bowman for providing unpublished rhodopsin sequence data.

References

Abell GCJ, Bowman JP (2005a) Ecological and biogeographic relationships of class *Flavobacteria* in the Southern Ocean. *FEMS Microbiol Ecol* **51**: 265–277.

Deleted: providing

Deleted: (

Deleted: ¶

Deleted: .

Deleted: by eucaryotic

Deleted: in

Deleted: the

Deleted: did not anticipate the extent to which the lake microbial community is orientated towards a net negative C and N balance. In contemplating this we examined

Deleted: relatively

Deleted: (*****). It is possible that the C

Deleted: N balance is sporadically readdressed by

Deleted: therefore contribute substantial organic material.

Deleted: REF

Deleted: REF

Deleted: determining

Formatted: Space After: 0 pt, Line spacing: single

Formatted: Font: Times New Roman, 12 pt, Not Bold, Font color: Black

Formatted: Indent: Hanging: 0.75

Deleted: and

Deleted: .(2005b

Formatted: Font: Italic

Formatted: Font: Times New Roman, 12 pt

Abell GCJ, Bowman JP. (2005b) Colonization and community dynamics of class *Flavobacteria* on diatom detritus in experimental mesocosms based on Southern Ocean seawater. *FEMS Microbiol Ecol* **53**: 379–391.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990) Basic Local Alignment Search Tool. *J Mol Biol* **215**: 403–410.

Balashov SP, Imasheva ES, Boichenko VA, Antón J, Wang JM, Lanyi JK. (2005) Xanthorhodopsin: a proton pump with a light-harvesting carotenoid antenna. *Science* **309**: 2061–2064.

Béjà O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP *et al.* (2000) Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**: 1902–1906.

Béjà O, Suzuki MT, Heidelberg JF, Nelson WC, Preston CM, Hamada T *et al.* (2002) Unsuspected diversity among marine aerobic anoxygenic phototrophs. *Nature* **6872**: 630–633.

Bengtsson K, Eriksson KM, Hartmann M, Wang Z, Shenoy BD, Grelet G-A *et al.* (2011) Metaxa: a software tool for automated detection and discrimination among ribosomal small subunit (12S/16S/18S) sequences of archaea, bacteria, eukaryotes, mitochondria, and chloroplasts in metagenomes and environmental sequencing datasets. *Antonie Van Leeuwenhoek* **100**: 471–475.

Bergmann DJ, Hooper AB, Klotz MG. (2005) Structure and sequence conservation of *hao* cluster genes of autotrophic ammonia-oxidizing bacteria: evident for their evolutionary history. *Appl Environ Microbiol* **71**: 5371–5382.

Bird MI, Chivas AR, Radnell CJ, Burton HR. (1991) Sedimentological and stable-isotope evolution of lakes in the Vestfold Hills, Antarctica. *Palaeogeogr Palaeoclimatol Palaeoecol* **84**: 109–130.

Bowman JP, McCammon SA, Lewis T, Skerratt JH, Brown JL, Nichols DS *et al.* (1998) *Psychroflexus torquis* gen. nov., sp. nov., a psychrophilic species from Antarctic sea ice, and reclassification of *Flavobacterium gondwanense* (Dobson et al. 1993) as *Psychroflexus gondwanense* gen. nov., comb. nov. *Microbiology* **144**: 1601–1609.

Bowman JP, McCammon SA, Rea SM, McMeekin TA. (2000) The microbial composition of three limnologically disparate hypersaline Antarctic lakes. *FEMS Microbiol Lett* **183**: 81–88.

Burke CM and Burton HR. (1988) Photosynthetic bacteria in meromictic lakes a stratified fjords of the Vestfold Hills, Antarctica. *Hydrobiologia* **165**: 13–23.

Campbell BJ, Engel AS, Porter ML, Takai K. (2006) The versatile ε-proteobacteria: key players in sulfidic habitats. *Nat Rev Microbiol* **4**: 458–468.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al.* (2010) QIIME allows analysis of high-throughput community sequence data. *Nat Methods* **7**: 335–336.

Charlson RJ, Lovelock JE, Andreae MO, Warren SG. (1987) Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature* **326**: 655–661.

Chouari R, Le Paslier D, Daegelen P, Ginestet P, Weissenbach J, Sghir A. (2005) Novel predominant archaeal and bacterial groups revealed by molecular analysis of an anaerobic sludge digester. *Environ Microbiol* **7**: 1104–1115.

Clarke KR and Gorley RN. (2006) PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth.

Cottrell MT and Kirchman DL. (2009) Photoheterotrophic microbes in the Arctic Ocean in summer and winter. *Appl Environ Microbiol* **75**: 4958–4966.

Formatted: Indent: Hanging: 0.75 cm

Deleted: , Jovanovich SB

Deleted: , Eisen JA

Deleted: , McMeekin TA.

Deleted: 2000b)

Curran MAJ and Jones GB. (1998) Spatial distribution of dimethylsulfide and dimethylsulfoniopropionate in the Australasian sector of the Southern Ocean. *J Geophys Res* **103**: 16 677–16 689.

Curson ARJ, Rogers R, Todd JD, Bearley CA, Johnston AWB (2008) Molecular genetic analysis of a dimethysulfoniopropionate lyase that liberates the climate-changing gas dimethylsulfide in several marine α-proteobacteria and *Rhodobacter sphaeroides*. *Environ Microbiol* **10**: 757–767.

Curson ARJ, Sullivan MJ, Todd JD, Johnston AWB. (2010) Identification of genes for dimethyl sulfide production in bacteria in the gut of Atlantic Herring (*Clupea harengus*). *ISME J* **4**: 144–146.

Curson ARJ, Sullivan MJ, Todd JD, Johnston AWB. (2011a) DddY, a periplasmic dimethylsulfoniopropionate lyase found in taxonomically diverse species of Proteobacteria. *ISME J* **5**: 1191–1200.

Curson ARJ, Todd JD, Sullivan MJ, Johnston AWB. (2011b) Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes. *Nat Rev Microbiol* **9**: 849–859.

Demergasso C, Escudero L, Casamayor EO, Chong G, Balagué V, Pedrós-Alió. (2008) Novelty and spatio-temporal heterogeneity in the bacterial diversity of hypersaline Lake Tebenquiche (Salar de Atacama). *Extremophiles* **12**: 491–504.

Deprez PP, Franzmann PD, Burton HR. (1986) Determination of reduced sulfur gases in Antarctic lakes and seawater by gas chromatography after solid adsorbent preconcentration.*J Chromatogr* **362**: 9–21.

Dobson SJ, James SR, Franzmann PD, McMeekin TA. (1991) A numerical taxonomic study of some pigmented bacteria isolated from Organic Lake, an antarctic hypersaline lake. *Arch Microbiol* **156**: 56–61.

Dobson SJ, Colwell RR, McMeekin TA, Franzmann PD. (1993) Direct sequencing of the polymerase chain reaction-amplified 16S rRNA gene of *Flavobacterium gondwanense* sp. nov. and*Flavobacterium salegens* sp. nov., two new species from a hypersaline Antarctic lake. *Int J Syst Bacteriol* **43**: 77–83.

Edwards KJ, Rogers DR, Wirsen CO, McCollom TM. (2003) Isolation and characterization of novel psychrophilic, neutrophilic, Fe-oxidizing, chemolithoautotrophic α- and γ-*Proteobacteria* from the deep sea. *Appl Environ Microbiol* **69**: 2906–2913.

Ferris JM, Gibson JAE, Burton HR. (1991) Evidence of density currents with the potential to promote meromixis in the ice-covered saline lakes. *Palaeogeogr Palaeoclimatol Palaeoecol* **84**: 99–107.

Franzmann PD, Burton HR, McMeekin TA. (1987a) *Halomonas subglaciescola*, a new species of halotolerant bacteria isolated from Antarctica. *Int J Syst Bacteriol* **37**: 27–34.

Franzmann PD, Deprez PP, Burton HR, van den Hoff J. (1987b) Limnology of Organic Lake, Antarctica, a meromictic lake that contains high concentrations of dimethyl sulfide. *Aust J Mar Freshw Res* **38**:409–417.

Friedrich CG, Bardischewsky F, Rother D, Quentmeier A, Fischer J. (2005) Prokaryotic sulfur oxidation. *Curr Opin Microbiol* **8**: 253–259.

Fofonoff NP and Millard RC Jr. (1983) Algorithms for computation of fundamental properties of seawater. *UNESCO Technical Papers in Marine Science*, no.**44**.

Fuhrman JA, Schwalbach MS, Stingl U. (2008) Proteorhodopsins: an array of physiological roles? *Nat Rev Microbiol* **6**: 488–494.

Comment [U6]: Sheree please check all Curson refs as I did not see all cited – e.g 2011 and 2011a

Deleted: .(

Deleted: ¶

Deleted: .(

Gärdes A, Kaeppel E, Shehzad A, Seebah S, Teeling H, Yarza P *et al*. (2010) Complete genome sequence of *Marinobacter adhaerens* type strain (HP15), a diatom-interacting marine microorganism. *Stand Genomic Sci* **3**: 97–107.

Gauthier MJ, Lafay B, Christen R, Fernandez L, Acquaviva M, Bonin P *et al*. (1992) *Marinobacter hydrocarbonoclasticus* gen. nov., sp. nov., a new, extremely halotolerant, hydrocarbon-degrading marine bacterium. *Int J Syst Bacteriol* **42**: 568–576.

Gibson JAE, Ferris JM, Burton HR. (1990) Temperature density, temperature conductivity and conductivity-density relationships for marine-derived saline lake waters. *ANARE Research Notes*, No. 78.

Gibson JAE, Garrick RC, Franzmann PD, Deprez PP, Burton H. (1991) Reduced sulfur gases in saline lakes of the Vestfold Hills, Antarctica. *Palaeogeo Palaeoclimatol Palaeoecol* **84**:131–140.

Gibson JAE, Qiang XL, Franzmann PD, Garrick RC, Burton HR. (1994) Volatile fatty and dissolved free amino acids in Organic Lake, Vestfold Hills, East Antarctica. *Polar Biol* **14**: 545–550.

Gibson JAE. (1999) The meromictic lakes and stratified marine basins of the Vestfold Hills, East Antarctica. *Antarct Sci* **11**:175–192.

Glatz RE, Lepp PW, Ward BB, Francis CA. (2006) Planktonic microbial community composition across steep physical/chemical gradients in permanently ice-covered Lake Bonney, Antarctica. *Geobiology* **4**: 53–67.

Goberna M, Insam H, Franke-Whittle IH. (2009) Effect of biowaste sludge maturation on the diversity of thermophilic bacteria and archaea in an anaerobic reactor. *Appl Environ Microbiol* **75**: 2566–2572.

Gómez-Consarnau L, González JM, Coll-Lladó M, Gourdon P, Pascher T, Neutze R *et al.* (2007) Light stimulates growth of proteorhodopsin-containing marine *Flavobacteria*. *Nature* **445**: 210–213.

Gómez-Consarnau L, Akram N, Lindell K, Pedersen A, Neutze R, Milton DL *et al.* (2010) Proteorhodopsin phototrophy promotes survival of marine bacteria during starvation. *PLoS Biol*. **8**: e1000358.

Green DH, Bowman JP, Smith EA, Gutierrez T, Bolch CJ. (2006) *Marinobacter algicola* sp. nov., isolated from laboratory cultures of paralytic shellfish toxin-producing dinoflagellates. *Int J Syst Evol Microbiol* **56**: 523–527.

Hahn MW, Stadler P, Wu QL, Pöckl. (2004) The filtration–acclimatization method for isolation of an important fraction of the not readily cultivable bacteria. *J Microbiol Methods* **57**: 379–390.

Hahn MW. (2009) Description of seven candidate species affiliated with the phylum *Actinobacteria*, representing planktonic freshwater bacteria. *Int J Syst Evol Microbiol* **59**: 112–117.

Hahsler M, Hornik K, Buchta C. (2008) Getting things in order: an introduction to R package seriation. *J Stat Softw* **25**:1–34.

Harris JK, Kelley ST, Pace NR. (2004) New Perspective on uncultured bacterial phylogenetic division OP11. *Appl Environ Microbiol* **70**: 845–849.

Huang L, Zhu S, Zhou H, Qu L. (2005) Molecular phylogenetic diversity of bacteria associated with the leachate of a closed municipal solid waste landfill. *FEMS Microbiol Lett* **242**: 297–303.

Deleted: , Glöckner FO

Deleted: , Betrand JC

Deleted: , Pedrós-Alió C, Pinhassi J.

Formatted: Font: Italic

Deleted: , González JM

Formatted: Font: Times New Roman, 12 pt

Deleted: ¶

Formatted: Indent: Hanging: 0.75 cm, Line spacing: single

Formatted: Indent: Hanging: 0.75 cm

Huang Y, Niu B, Gao Y, Fu L, Li W. (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**: 680–682.

Hügler M and Sievert SM. (2011) Beyond the Calvin cycle: autotrophic carbon fixation in the ocean. *Annu Rev Mar Sci* **3**: 261–289.

Humayoun SB, Bano N, Hollibaugh JT. (2003) Depth distribution of microbial diversity in Mono Lake, a meromictic soda lake in California. *Appl Environ Microbiol* **69**: 1030–1042.

Huu NB, Denner EB, Ha DT, Wanner G, Stan-Lotter H. (1999) *Marinobacter aquaeolei* sp. nov., a halophilic bacterium isolated from a Vietnamese oil-producing well. *Int J Syst Bacteriol* **49**: 367–375.

James SR, Dobson SJ, Franzmann PD, McMeekin TA. (1990) *Halomonas meridiana*, a new species of extremely halotolerant bacteria from Antarctic saline lakes. *System Appl Microbiol* **13**: 270–278.

James SR, Burton HR, McMeekin TA, Mancuso CA. (1994) Seasonal abundance of *Halomonas meridiana*, *Halomonas subglaciescola*, *Flavobacterium gondwanense* and *Flavobacterium salegens* in four Antarctic Lakes. *Antarctic Sci* **6**: 325–332.

Johnston AWB, Todd JD, Sun L, Nikolaidou-Katsaridou MN, Curson ARJ, Rogers R. (2008) Molecular diversity of bacterial production of the climate changing gas, dimethyl sulphide, a molecule that impinges on local and global symbioses. *J Exp Bot* **59**: 1059–1067.

Kang I, Lee K, Yang S-J, Choi A, Kang D, Lee YK *et al.* (2012) Genome sequence of "*Candidatus* Aquiluna" sp. strain IMCC13023, a marine member of the *Actinobacteria* isolated from an Artic Fjord. *J Bacteriol* **194**: 3550–3551.

Kirchman DL. (2002) The ecology of *Cytophaga-Flavobacteria* in aquatic environments. *FEMS Microbiol Eco l***39**: 91–100.

Kraft B, Stous M, Tegetmeyer HE. (2011) Microbial nitrate respiration – genes, enzymes and environmental distribution. *J Biotechnol* **155**: 104–117.

La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G *et al.* (2008) The virophage as a unique parasite of the giant mimivirus. *Nature* **455**: 100–105.

Labrenz M, Collins MD, Lawson PA, Tindall BJ, Schumann P, Hirsch P. (1999) *Roseovarius tolerans* gen. nov., sp. nov., a budding bacterium with variable bacteriochlorophyll *a* production from hypersaline Ekho Lake. *Int J Syst Bacteriol* **49**: 137–147.

Lauro FM, DeMaere MZ, Yau S, Brown MV, Ng C, Wilkins D *et al.* (2011) An integrative study of a meromictic lake ecosystem in Antarctica. *ISME J* **5**:879–895.

Laybourn-Parry J, Marshall WA, Marchant HJ. (2005). Flagellate nutritional versatility as a key to survival in two contrasting Antarctic saline lakes. *Freshw Biol* **50**: 830–838.

Ley RE, Turnbaugh PJ, Klein S, Gordon JI. (2006) Human gut microbes associated with obesity. *Nature* **444**: 1022–1023.

Lovelock JE and Maggs RJ. (1972) Atmospheric dimethyl sulfide and the natural sulphur cycle. *Nature* **237**: 452–453.

Laybourn-Parry J and Pearce D. (2007) The biodiversity and ecology of Antarctic lakes: models for evolution. *Phil Trans R Soc B* **364**: 2273–2289.

Man D, Wang W, Sabehi G, Aravind L, Post AF, Massana R *et al.* (2003) Diversification and spectral tuning in marine proteorhodopsins. *EMBO J* **22**:1725–1731.

Matsuzaki M, Kubota K, Satoh T, Kunugi M, Ban S, Imura S. (2006) Dimethyl sulfoxide-respiring bacteria in Suribati Ike, a hypersaline lake, in Antarctica and the marine environment. *Polar Biosci* **20**: 73–87.

Deleted: .(

Deleted: , Cho J-C

Deleted: , Merchat C

Deleted: D*et*

Formatted: Font: Times New Roman, 12 pt

Formatted: Indent: Hanging: 0.75 cm

Deleted: .(

Miyoshi T, Iwatuski T, Naguma T. (2005) Phylogenetic characterization of 16S rRNA gene clones from deep-groundwater microorganisms that pass through 0.2 µm-pore-size filters. *Appl Environ Microbiol* **71**: 1084–1088.

Moran MA, Belas R, Schell MA, González JM, Sun F, Binder BJ *et al.* (2007) Ecological genomics of marine Roseobacters. *Appl Environ Microbiol* **73**: 4559–4569.

Moran MA and Miller WL. (2007) Resourceful heterotrophs make the most of light in the coast ocean. *Nat Rev Microbiol* **5**: 792–799.

Moran MA, Reisch CR, Kiene RP, Whitman WB. (2012) Genomic insights into bacterial DMSP transformations.*Ann Rev Marine Sci* **4**: 523–542.

Ng C, DeMaere MZ, Williams TJ, Lauro FM, Raftery M, Gibson JAE *et al.* (2010) Metaproteogenomic analysis of a dominant green sulfur bacterium from Ace Lake, Antarctica. *ISME J* **4**: 1002–1019.

Noguchi H, Park J, Takagi T. (2006) MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. *Nucleic Acids Res* **34**: 5623–5630.

Raina J-P, Dinsdale EA, Willis BL, Bourne DG. (2010) Do the organic sulfur compounds DMSP and DMS drive coral microbial associations? *Trends Microbiol* **3**: 101–108.

Redfield AC, Ketchum BH, Richards FA. (1963) The influence of organisms on the composition of seawater, In: Hill MN (ed). The sea. John Wiley and Sons: New York, pp 26–77.

Reisch CR, Moran MA, Whitman WB. (2011) Bacterial catabolism of dimethylsulfoniopropionate (DMSP). *Front Microbiol* **2**: 1–12.

Rivière D, Desvignes V, Pelletier E, Chaussonnerie S, Guermazi S, Weissenbach *et al.* (2009) Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *ISME J* **3**: 700–714.

Roberts NJ and Burton HR. (1993a) Sampling volatile organics from a meromictic Antarctic lake. *Polar Biol* **13**: 359–361.

Roberts NJ, Burton HR, Pitson GA. (1993b) Volatile organic compounds from Organic Lake, an Antarctic hypersaline, meromictic lake. *Polar Biol* **13**: 361–366.

Röske K, Sachse R, Scheerer C, Röske I. (2012) Microbial diversity and composition of the sediment in the drinking water reservoir Saidenbach (Saxonia, Germany). *Syst Appl Microbiol* **35**: 35–44.

Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S *et al.* (2007) The *Sorcerer II* Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* **5**: 398–431.

Samsudin AA, Evans PN, Wright AG, Al Jassim R. (2011) Molecular diversity of the foregut bacteria community in the dromedary camel (*Camelus dromedariusi*). *Environ Microbiol* **13**: 3024–3035.

Schmidtova J, Hallam SJ, Baldwin SA. (2009) Phylogenetic diversity of transition and anoxic zone bacterial communities within a near-shore anoxic basin: Nitinat Lake. *Environ Microbiol* **11**: 3233–3251.

Scott KM, Sievert SM, Abril FN, Ball LA, Barrett CJ, Blake RA *et al.* (2006) The genome of deep-sea vent chemolithoautotroph *Thiomicrospira crunogena* XCL-2. *PLoS Biol* **4**: e383.

Sharma AK, Zhaxybayeva O, Papke RT, Doolittle WF. (2008) Actinorhodopsins: proteorhodopsin-like gene sequences found predominantly in non-marine environments. *Environ Microbiol* **10**: 1039–1056.

Deleted: , Edmonds J

Deleted: .(

Formatted: Font: 12 pt

Deleted: , Li T

Deleted: , Boller AJ

Formatted: Font: Times New Roman, 12 pt

Sievert SM, Scott KM, Klotz MG, Chain PSG, Hauser LJ, Hemp J *et al.* (2008) Genome of the Epsilonproteobacterial chemoautotroph *Sulfurimonas denitrificans. Appl Environ Microbiol* **74**: 1145–1156.

Tajima K, Aminov RI, Nagamine T, Ogata K, Nakamura M, Matsui H *et al*. (1999) Rumen bacterial diversity as determined by sequence analysis of 16S rDNA. *FEMS Microbiol Ecol* **29**: 159–169.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.

Tang Y, Ji P, Hayashi J, Koike Y, Wu X, Kida K. (2011) Characteristic microbial community of a dry thermophilic methanogenic digester: its long-term stability and change with feeding. *Appl Microbiol Biotechnol* **91**: 1477–1461.

Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM *et al.* (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**: 608–611.

Thomsen HA. (1988) Ultrastructural studies of the flagellate and cyst stages of Pseudopedinella tricostata (Pedinellales, Chrysophyceae). *Brit Phycol J* **23**: 1–16.

Todd JD, Rogers R, Li YG, Wexler M, Bond PL, Sun L *et al.* (2007) Structural and regulatory genes required to make the gas dimethyl sulfide in bacteria. *Science* **315**: 666–669.

Todd JD, Curson ARJ, Dupont CL, Nicholson P, Johnston AWB. (2009) The *dddP* gene, encoding a novel enzyme that converts dimethylsulfoniopropionate into dimethyl sulfide, is widespread in ocean metagenomes and marine bacteria and also occurs in some Ascomycete fungi. *Environ Microbiol* **11** :1376–1385.

Todd JD, Curson ARJ, Nikolaidou-Kataraidou N, Brearley CA, Watmough NJ, Chan Y *et al.* (2010) Molecular dissection of bacterial acrylate catabolism – unexpected links with dimethylsulfoniopropionate catabolism and dimethyl sulfide production. *Environ Microbiol* **12**: 327–343.

Todd JD, Curson ARJ, Kirkwood M, Sullivan MJ, Green RT, Johnston AWB. (2011) DddQ, a novel, cupin-containing, dimethylsulfoniopropionate lyase in marine roseobacters and in uncultured marine bacteria. *Environ Microbiol* **13** :427–438.

Todd JD, Kirkwood M, Newton-Payne S, Johnston AWB. (2012) DddW, a third DMSP lyase in model Roseobacter marine bacterium, *Ruegeria pomeroyi* DSS-3. *ISME J* **6** :223–226.

Unrein F, Izaguirre I, Massana R, Balagué V, Gasol JM. (2005) Nanoplankton assemblages in maritime Antarctic lakes: characterisation and molecular fingerprinting comparison. *Aquat Microb Ecol* **40**: 269–282.

Wagner-Döbler I and Biebl H. (2006) Environmental biology of the marine *Roseobacter* lineage. *Ann Rev Microbiol* **60**: 255–280.

Wang Q, Garrity GM, Tiedje JM, Cole JR. (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.

Ward BB and Priscu JC. (1997) Detection and characterization of denitrifying bacteria from a permanently ice-covered Antarctic lake. *Hydrobiologia* **347**: 57–68.

Ward BB, Granger J, Maldonado MT, Casciotti KL, Harris S, Wells ML. (2005) Denitrification in the hypolimnion of permanently ice-covered Lake Bonney, Antarctica. *Aquat Microb Ecol* **38**: 295–307.

Deleted: ¶
Deleted: , Hügler M
Deleted: , Kassabgy M
Deleted: *British Phycological Journal*
Deleted: , Curson ARJ
Deleted: , Page PCB
Deleted: .(

Wilkins D, Yau S, Williams TJ, Allen MA, Brown MV, DeMaere MZ, *et al.* (2012a) Key microbial drivers in Antarctic aquatic environments. *FEMS Microbiol Rev* doi:10.1111/1574-6976.12007.

Williams TJ, Wilkins D, Long E, Evans F, DeMaere MZ, Raftery MJ, *et al.* (2012b) The role of planktonic *Flavobacteria* is processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environ Microbiol* doi:10.1111/1462-2920.12017.

Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, *et al.* (2011) KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* **39**: W316–W322.

Xing P, Hahn MW, Wu QL. (2009) Low taxon richness of bacterioplankton in high-altitude lakes of the eastern Tibetan Plateau, with a predominance of *Bacteroidetes* and *Synechoccocus* spp. *Appl Environ Microbiol* **75**: 7017–7025.

Yamamoto M and Takai K. (2011) Sulfur metabolisms in epsilon- and gamma-*Proteobacteria* in deep-sea hydrothermal fields. *Front Microbiol* **2**: 1–8.

Yamane K, Hattori Y, Ohtagaki H, Fujiwara K. (2011) Microbial diversity with dominance of 16S rRNA genes sequences with high GC contents at 74 and 98°C subsurface crude oil deposits in Japan. *FEMS Microbiol Ecol* **76**: 220–235.

Yanagibayashi M, Nogi Y, Li L, Kato C. (1999) Changes in the microbial community in Japan Trench sediment from a depth of 6292 m during cultivation without decompression. *FEMS Microbiol Lett* **170**: 271–279.

Yau S, Lauro FM, DeMaere MZ, Brown MV, Thomas T, Raftery MJ *et al.* (2011) Virophage control of antarctic algal host-virus dynamics. *Proc Natl Acad Sci USA* **108**: 6163–6168.

Yilmaz P, Iversen MH, Hankeln W, Kottman R, Quast C, Glöckner FO, (2012) Ecological structuring of bacterial and archaeal taxa in surface ocean waters. *FEMS Microbiol Ecol* **81**: 373–385.

Zwartz D, Bird M, Stone J, Lambeck K. (1998) Holocene sea-level change and ice-sheet history in the Vestfold Hills, East Antarctica. *Earth Planet Sci Lett* **155**: 131–145.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Figure legends and table titles

Figure 1 Vertical structure of Organic Lake. (**A**) Parameters that varied unimodally with depth showed two zones: an aerobic mixed zone above 5.7 m and a denser suboxic deep zone below. (**B**) Additional factors that revealed stratification within the deep zone. The peak in concentration at 6.5 m for ammonia was also observed for all other nutrients assayed except nitrate and nitrite (see Table 1). σ_T = (1000−density); cond, conductivity; DO, dissolved oxygen; turb, turbidity.

Figure 2 Diversity of *Bacteria* and *Eucarya* in Organic Lake. (**A**) *Bacteria* and (**B**) *Eucarya* from each size fraction (0.1, 0.8 and 3.0 μm) at each sample depth (1.7, 4.2, 5.7, 6.5 and 6.7 m) aggregated according to class. The x-axis shows counts of SSU sequences normalized to average reads acquired per sample filter. Taxa that belong to the same higher rank are shown grouped

Deleted: , Lauro FM, Cavicchioli R.

Deleted: , Cavicchioli R.

Deleted: , Kong L, Gao G, Li CY, Wei L.

Deleted: .(

with a square bracket in the legend. Abundant taxa are labeled in the plot with a number that corresponds to the numbered boxes in the legend. (**C**) Composition of abundant bacterial classes.

Figure 3 Heatmap and biclustering plot of the SSU gene composition in Organic Lake. Samples are shown according to size fractionate (3.0, 0.8 and 0.1 µm) and depth (1.7, 4.2, 5.7, 6.5 and 6.7 m). SSU genes were classified to the lowest taxonomic rank that gave bootstrap confidence > 85% until the rank of genus. SSU gene counts were normalized and square root transformed. Taxa that comprised <2% of the sample were not included.

Figure 4 Carbon, nitrogen and sulfur cycles in Organic Lake. Vertical profiles of genetic potential for (**A**) carbon (**B**) nitrogen (**C**) and sulfur conversions for each size fraction. The y-axis shows sample depths (m) and the x-axis shows counts of marker genes normalized to 100 Mbp of DNA sequence. The 0.1, 0.8 and 3.0 µm size fractions are shown as blue, red and green, respectively. Counts for marker genes for the same pathway or enzyme complex were averaged and those from different pathways were summed. For marker gene descriptions see Table S1 and Table S3. AAnP, aerobic anoxygenic photosynthesis; rTCA, reverse TCA; WL, Wood-Ljungdahl; DNRA, dissimilatory nitrate reduction to ammonia; DSR, dissimilatory sulfate reduction, DMSO dimethylsulfoxide; ASR, assimilatory sulfate reduction; DMSP, dimethylsulfoniopropionate.

Table 1 Physico-chemical properties, cell counts and VLP counts of Organic Lake samples.

Table 2 Contribution of different taxonomic groups to counts of marker genes involved in carbon, nitrogen and sulfur conversions.

Table 3 Counts of genes involved in DMSP catabolism and photoheterotrophy.

Supplementary Figure legends and supplementary table titles

Figure S1 Map of the Vestfold Hills showing Organic Lake.

Figure S2 Bathymetry of Organic Lake 9 November 2008.

Figure S3 Vertical profiles of physical and chemical parameters of Organic Lake taken *in situ* at the deepest point in the lake on 9 November 2008.

Figure S4 Epifluorescence microscopy images of Organic Lake microbiota (<20 µm) filtered onto 0.01 µm polycarbonate membrane and stained with SYBR Gold.

Figure S5 PCA analysis of physico-chemical parameters and cell/VLP counts of Organic Lake profile.

Figure S6 Phylogenetic tree of rhodopsin homologs including proteorhodopsin, bacteriorhodopsin, actinorhodopsin and xanthorhodopsin.

Figure S7 Genomic maps of Organic Lake scaffolds containing the OL-R1 rhodopsin homolog.

Figure S8 Phylogenetic tree of DddD DMSP lyase homologs.

Figure S9 Phylogenetic tree of DddL DMSP lyase homologs from Organic Lake and public databases.

Figure S10 Phylogenetic tree of DddP DMSP lyase homologs from Organic Lake and public databases.

Figure S11 Phylogenetic tree of DmdA DMSP demethylase homologs from Organic Lake and public databases.

Table S1 Full list of KEGG Orthologs (KO) involved in carbon, nitrogen and sulfur conversions that were searched for in the Organic Lake metagenome.

Table S2 Sequences used in this study as BLAST queries for retrieving homologs in the Organic Lake metagenomes.

Table S3 Summary of metagenomic data for Organic Lake samples.

Table S4 Microbial taxa detected in the Organic Lake water column profile.

Formatted: Normal, Indent: Left: 0 cm

Formatted: Font: Times New Roman, 12 pt