Understanding globally important biogeochemical processes from a study of a hypersaline Antarctic lake

Yau S, Lauro FM, DeMaere MZ, Brown MV, Williams T???, Cavicchioli R

# Potential Aims and Scope

Microbial population and community ecology, Integrated genomics and post-genomics approaches in microbial ecology, **Microbial ecology and functional diversity of natural habitats**. ORIGINAL ARTICLE 5 000 word limit. TITLE: 50 characters.

# Abstract (250 word limit)

Organic Lake is a shallow (6.75 m deep) marine-derived hypersaline lake in the Vestfold Hills, East Antarctica with a high concentration of the cloud-forming gas dimethylsulfide (DMS). During sampling, the lake was stratified into an aerobic mixed zone and a suboxic deep zone, with a peak of C, S and ammonia below the oxycline (6.5 m). Environmental DNA from size fractionated samples (3.0, 0.8 and 0.1 µm) along the depth profile was sequenced and the taxonomic composition and functional diversity determined. The microbial composition resembles other hypersaline and cold environments indicating a strong selection for species. Eucaryotic phytoflagellates related to *Dunaliella* and *Pseudopedinella* are the dominant photosynthetic organisms.Bacterioplankton throughout the water column was dominated by heterotrophic *Marinobacter*, *Roseovarius* and *Psychroflexus.* Candidate division RF3, *Halomonas* and *Psychromonas* were overrepresented at 6.5 m and associated with high potential for fermentation of particulate matter. The bottom samples were abundant in candidate divisions OD1 and TM7. Genetic potential for nitrogen cycling showed large denitrification potential, limited fixation and no nitrification. Nitrogen limitation was evident in the predominance of ammonia uptake and remineralization pathways. A diverse set of functional genes were assigned to the *Marinobacter* and *Roseovarius* clades including rhodopsin, DMSP lyase (*dddD*, *dddL* and *dddP*),Calvin cycle, anaerobic respiration and CO oxidation genes indicating they are metabolically versatile generalists that may contribute to primary production and the unusual sulfur cycle. This study has allowed a rigorous description of the microbial community within a natural environment and sheds light on globally relevant biogeogemical processes such as DMS generation, lithoheterotrophy, chemolithoautotrophy and photoheterotrophy. It shows potential for diverse strategies in may be the rule, not the exception for marine-derived surface water *Gammaproteobacteria* and *Roseobacter* with implications untangling microbial roles in production.

# Introduction

Life in the Antarctic is constrained by extremes of temperature and salinity under a polar light cycle. In the frozen desert, ice-free regions containing liquid water in lakes and ponds are rare oases for life. The Vestfold Hills, on the eastern shore of the Prydz Bay, East Antarctica (Figure S1), is one such region where hundreds of lakes are found. The lakes were formed from seawater, trapped approximately 10 000 BP when the continental ice-sheet receded and the land rose above sea-level (Zwartz *et al*., 1998; Gibson, 1999). Life in these lakes can be entirely microbial and of reduced diversity (Bowman *et al.*, 2000b). Differing local conditions has lead each lake to develop unique physical and chemical properties, making them fitting sites to study microbial ecology, biogeochemistry and evolution. The Vestfold Hills contains the highest density of meromictic (permanently stratified) water bodies in Antarctica (Gibson, 1999). They are advantageous study sites as environmental gradients exist within a single, largely closed system allowing species to be more easily related to abiotic factors. By using molecular techniques, a large proportion of the species diversity and gene content can be covered allowing better inference of the functional roles for the taxa present (Laybourn-Parry & Pearce, 2007).

A metagenomic approach, complemented with metaproteomics, has been successfully applied to two lakes in the Vestfold Hills (Ng *et al.*, 2010; Lauro *et al.*,2011; Yau *et al.*, 2011). The first of these was Ace Lake, where a comprehensive description of the community structure, biogeochemical fluxes and responses to resource limitation was achieved (Lauro *et al.*, 2011). The metabolism of the 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). In the surface water of the second lake, Organic Lake, a member of the virophage virus family was discovered that potentially regulates microbial loop dynamics (Yau *et al*., 2011). Virophage require a helper virus to replicate but are detrimental to their helper (La Scola *et al.*, 2008). The Organic Lake virophage (OLV) likely depends on phycodnaviruses (algal viruses). The presence of OLV would reduce infective phycodnaviruses leading to increased algal blooms and thus carbon flux (Yau *et al.*, 2011). These studies have achieved exceptional insight into Antarctic lakes but are also relevant to other aquatic systems. For example, OLV-like sequences were found in coastal marine, hypersaline and freshwater metagenomes indicating virophages have a wider ecological role (Yau *et al.*, 2011).

This study extends the previous metagenomic analysis of the surface water viriome of Organic Lake to examine the entire microbial community along a depth profile from a whole ecosystem perspective. The bottom waters of Organic Lake are unusual due to the high concentration of the volatile gas dimethylsulfide (DMS) (Deprez *et al*., 1986;Franzmann *et al.*, 1987; Gibson *et al.*, 1991; Roberts & Burton 1993a; Roberts *et al.*, 1993b). Aerosols derived from atmospheric DMS act as cloud condensation nuclei and are hypothesized to affect climate (Charlson *et al.*, 1987). Although the importance of DMS was proposed forty years ago (Lovelock & Maggs, 1972) the first enzymes involved in DMS production were only identified in the last five years (Todd *et al.*, 2007). Concentrations of DMS as high as 5000 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 & Jones, 1998). This makes it an ideal location to identify the microbes involved in and potentially the basis for DMS accumulation. This study sought to determine the composition and the functional potential of Organic Lake microbiota and, in conjunction with historic and contemporary physico-chemical data, generate and an integrative understanding of the whole lake ecosystem.

# Materials and Methods

## Sample collection and preparation

Water was collected from Organic Lake on 10 November 2008 (68º27'22.15"S, 78º11'23.95"E) through a 30 cm hole in the 0.8 m thick ice cover above the deepest point in the lake. Samples were collected at 1.7, 4.2, 5.7, 6.5 and 6.7 m depths. Lake water was passed through a 20 µm pore size pre-filter, and then microbial biomass was captured by sequential filtration onto 3.0 µm, 0.8 µm and 0.1 µm pore size membrane filters. Between 1–2 L of lake water was sufficient to clog the filters. DNA was extracted from the filters as previously described (Ng *et al.*, 2010; Lauro *et al.*,2011). DNA from all samples was sequenced using the Roche GS-FLX titanium sequencer. Reads were processed to remove low quality bases, assembled and annotated as previously described (Lauro *et al.*, 2011). Water was also collected for microscopic and chemical analysis at the same sample depths and frozen −80ºC.

## 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 minimum-maximum mercury thermometer. 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 according to the relation described by Gibson (1999). The adjusted conductivity brings the temperature within the acceptable range to estimate practical salinity by the formula of Fofonoff and Millard (1983). However, salinity was likely underestimated as it is higher than the range (2–42) for which the conductivity–salinity relation holds. 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:

σT = (1000–density) kg/m3

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 Services Tasmania. Values for dissolved nutrients were measured after filtration through 0.1 µm pore size membrane. All other nutrients were measured from water collected after filtration through 20 µm pore size membrane. 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 & Gorley, 2006) on the normalized physical and chemical parameters to visualize how abiotic factors varied with depth.

## 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.

## Biological diversity analyses

### Cellular diversity

Diversity of cellular life was assessed using ribosomal small subunit (SSU) gene sequences. Metagenomic reads that matched the 16S and 18S rRNA gene 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). 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 the RDP classifier (Wang *et al*., 2007) trained against SILVA (release 108) sequences (www.arb-silva.de). 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 & 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 mixed zone (1.7, 4.2 and 5.7 m) and from the 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 considering following 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 biclustering 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

Open reading frames (ORFs) were predicted from trimmed metagenomic reads using MetaGene (Noguchi *et al*., 2006) accepting those > 90 bp. 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). KEGG GENES is a collection of genes from all complete genomes from public resources, primarily NCBI RefSeq. The BLAST output was processed using KEGG Orthology Based Annotation System (KOBAS) version 2.0 (Xie *et al.*, 2011) accepting assignments to KEGG orthologs with e-value < 1e−05 and rank > 5. Matches to KEGG orthologs that are functional markers for carbon, nitrogen and sulfur cycling (\*Table CNS genes) were counted. Normalized frequencies of markers from the same pathway were averaged and those from different pathways were summed. Marker enzymes were assigned to species based on the origin of the best KEGG GENES BLASTp match.

Marker genes not represented by KEGG orthologs were retrieved by alternative strategies. Organic Lake rhodopsin homologs were retrieved if they had a top BLAST match to any in a list of 139 entries in the KEGG GENES database affiliated with bacteriorhodopsin, xanthorhodopsin, halorhodopsin or proteorhodopsin. The DMSP lyases with experimentally confirmed function were retrieved from National Center for Biotechnology Information (NCBI) ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) sequence databases (Table S2). These sequences were used to query a BLAST database of translated ORFs predicted from Organic Lake metagenomic reads. Matches were accepted if e-value was <1e−10 and sequence identity was within the range shared by the query enzymes of the same family.

## Phylogenetic analyses

Marker gene sequences for phylogenetic analysis were clustered using the CD-HIT web server (Huang *et al*., 2010) at 90% global amino acid identity. The longest sequence of the cluster was used in a BLASTp query against the NCBI non-redundant (NR) database to retrieve full-length sequenced homologs. Organic Lake representative sequences that resided within a desired conserved region and homologs from NR 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 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 of Organic Lake

During sampling on November 2008, Organic Lake had a maximum depth of 6.75 m and the surface measured 3.87 m above mean sea level. *In situ* physico-chemical profiles (Figure S2) were measured over the deepest point in the lake (Figure S3) to evaluate the water column properties and structure. Two distinct zones were apparent: a mixed zone above 5.7 m and a suboxic deep zone below (Figure 1A). The separation of these two zones was indicated by a pycnocline starting at 5.7 m. The presence of an oxycline at the same depth indicates depletion of DO due to respiration in the stagnant waters. The pH also decreased with DO, likely due to accumulated fermentation products such as acetic, formic and lactic acids that have been previously recorded in the bottom waters (Franzmann *et al*., 1987b; Gibson *et al*., 1994). The deep zone was not completely anoxic as has been recorded in the past (Franzmann *et al.*, 1987b; Gibson *et al.*, 1991) indicating DO had entered the bottom waters in the last 13 years. Oxygen may be episodically introduced with currents of cold dense water that are generated during ice-formation in the lake shallows (Ferris *et al*., 1999). Samples were collected from the mixed (1.7, 4.2 and 5.7 m) and deep (6.5 m and 6.7 m) zones to examine their nutrient content and microbiology.

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 (Figure 1B). 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. Microscopy images do not show a shift in cell morphology that could account for the large drop in turbidity (\*Figure S3), which suggests particulate matter primarily contributed to turbidity readings. (\*Trophic status). The low turbidity and peak in cell counts and nutrients at 6.5 m suggest increased degradation of particulate matter by the microbial community via processes enriched in the microaerophilic environment. This is supported by the high concentrations of dissolved organic and free amino acids found in the deep zone (Gibson *et al.*, 1994) indicative of breakdown of high molecular weight compounds such as carbohydrates, lipids and proteins.

The C:N and C:P ratios were high compared to the Redfield ratio (Redfield *et al.*, 1963; \*others?) except at 6.5 m indicating this was the only depth where dissolved N and P 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 (Figure S4). 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 increased activity at this depth was related to breakdown of particulate matter and sulfur chemistry.

## Organic Lake microbial community composition and vertical distribution

### Overall microbial diversity

Metagenomic sequences were obtained from size fractionated (3.0 µm, 0.8 µm and 0.1 µm) microbial biomass from each sample depth (Table S1). To determine the microbial composition, a total of 3 959 reads matching to the SSU gene were retrieved from the metagenomic sequences. These grouped into 983 OTUs. Bacteria were numerically dominant comprising 76.2% of all SSU sequences. 16.3% of sequences were assigned as Eucarya and 7.5% of SSU sequences could not be classified. Only 2 reads, assigned to a deep sea hydrothermal clade of *Halobacteriales*, were classified as Archaea (Table S3) indicating Archaea were rare in Organic Lake. Overall microbial diversity was fairly low, with 15 bacterial phyla and 6 eucaryotic superkingdom divisions in total. Of these, only 7 bacterial phyla and 4 eucaryotic phyla were predominant. (\*diversity indices in primer)

Three bacterial classes, *Gammaproteobacteria*, *Alphaproteobacteria* and *Flavobacteria*, were the most abundant and were found on all filter sizes at all depths (Figure 2A). Each of these three classes consisted of one dominant genus, which was *Marinobacter*, *Roseovarius* and *Psychroflexus* respectively (Figure 2C). Cyanobacterial sequences were all classified as chloroplasts (Figure 2A), except for three reads that could not be assigned to any lower rank (Table S3) indicating free-living *Cyanobacteria* were absent or extremely rare. Moderately abundant bacterial classes were *Actinobacteria*, *Deltaproteobacteria*, *Epsilonproteobacteria*, and candidate divisions OD1 and RF3. Lower abundance divisions included the *Bacilli*, *Clostridia*, *Spirochaetes*, *Lentisphaera*, TM7, *Opitutae*, *Verrucomicrobia*, Bhi80-139, Bd1-5, SR1 and *Chlamydiae* (Figure 2A).

The dominant Eucarya were photosynthetic *Chlorophyta* (green algae) and *Dictyochophyceae* (silicoflagellate algae) (Figure 2B) principally assigned to the genus *Dunaliella* and the order *Pedinellales* respectively (Table S3)*.* Lower abundance Eucarya included *Bacillariophyta* (diatoms), *Dinophyceae* (dinoflagellates), *Fungi* and heterotrophic *Choanoflagellida* (choanoflagellates) and *Ciliophora* (ciliates). *Bacillariophyta* and *Dinophyceae* were related to *Chaetoceros* and *Gymnodinium* respectively (Table S3). Both of these classes would contribute to primary production, although dinoflagellates are potentially bacteriovorous.

### Selection for psychrophilic and halophilic *Eucarya* and *Bacteria*

Cultured relatives of taxa detected in Organic Lake are known to be halophilic and/or psychrophilic (Gauthier *et al.*, 1992; Dobson *et al.*, 1991; Labrenz *et al*., 1999;\*\*\*\* Aquiluna\*Kang *et al*., 2012). Similarly, uncultured taxa have highest identity to sequences from saline and/or cold environments (\*OD1 Mosier *et al*., 2007; RF3 Demergasso *et al.*, 2010;\*\*) (Table S3). The consistent association with phylotypes from similar environments indicates a strong selection for species by common environmental factors. This is further supported by the persistence of the same taxa in Organic Lake over time, such as *Dunaliella*, *Chaetoceros, Psychroflexus*,(Franzmann *et al*., 1987b)  *Marinobacter*, *Halomonas* and *Roseovarius* (Bowman *et al*., 2000b), which indicates they are particularly adapted to the lake conditions.

### Variation of microbial composition according to size and depth

Community composition varied with size fraction and depth (Figure 2). This was supported by seriation analysis that showed samples clearly clustered according to size fraction, and those clusters further separated into mixed and deep zone groups (Figure 3). Furthermore, it identified genera that were differentially distributed with size and depth and those that co-varied (Figure 3). A significant difference in genus-level composition between mixed and deep zone samples was supported by ANOSIM analysis (Rho: 0.53, significance: 0.1%). Differential vertical distribution of taxa is consistent with partitioning of ecological functions in the lake and in conjunction with the physical and chemical data, provided insight into the functional roles of those taxa.

#### 20–3.0 µm fraction

The mixed zone samples had a relatively high abundance of *Dunaliella* chloroplasts and chlorophyte algae consistent with large phototrophic cells concentrating near surface light. *Dunaliella* have been previously isolated from Organic Lake and were reported to be the dominant eucaryotic alga (Franzman *et al.*, 1987b). Signatures of algae found at the bottom of the lake are likely due to sedimentation of dead cells or resting cysts.

*P*. *gondwanensis* (previously *Flavobacterium*) has been isolated from Organic Lake (Franzmann *et al*., 1987b) and ranges in length from 1.5–11.5 µm (Dobson *et al*., 1991) consistent with enrichment on the 3.0 µm filter. *Psychroflexus* was overrepresented in the surface and the 6.7 m sample. *Flavobacteria* have been associated with phytoplankton blooms in the Southern Ocean (Abell & Bowman 2005a; Abell & Bowman 2005b), which is hypothesized to be related to their ability to degrade high molecular weight carbon from algal exudates and detritus (reviewed in Kirchman *et al.*, 2009). Likely, Organic Lake *Psychroflexus* fills a similar ecological role. In support of this, *Psychroflexus* clusters with *Dunaliella* chloroplasts in the seriation analysis (Figure 3) and *P. gondwanese* abundance in Organic Lake has been correlated with average hours of sunshine per day (James *et al.*, 1994). Its presence in the deep zone was most likely due to sedimentation as *P. gondwanense* is non-motile and strictly aerobic (Dobson *et al.*, 1991).

*Roseovarius* was enriched at 4.2 m and 6.5 m suggesting adaptation to aerobic and microaerophilic conditions. *R. tolerans*, an isolate from the meromictic Ekho Lake, also in the Vestfold Hills and a close relative of Organic Lake *Roseovarius*,has a large cell size (1.1–2.2 μm long) (Labrenz *et al*., 1999) accounting for accumulation on this size fraction. The abundance of *Roseovarius* in the deep zone is somewhat unexpected as *Roseovarius* isolates are strictly aerobic. However one strain from Ekho Lake is capable of microaerophilic growth (Labrenz *et al.*, 1999). This is consistent with their overrepresentation at 6.5 m rather than at the lake bottom. *Roseovarius* is a member of the *Roseobacter* clade whose diverse metabolic capabilities. These include DMSP degradation and aerobic anoxygenic phototrophy (AAnP) (reviewed in Wagner-Döbler & Biebl, 2006) which are relevant to the Organic Lake system (\*see below).

#### 3–0.8 µm size fraction

*Marinobacter* dominated the 0.8 µm size fraction at all depths except 6.5 m. Their abundance on this size fraction is consistent with the cell size of isolates (Gauthier *et al.*, 1992). The genus is extraordinarily metabolically versatile and described as a “generalist” or “opportunitroph” (\*Singer *et al.*, 2011). Some isolates are capable of interacting with diatoms (Gärdes *et al*., 2010) and dinoflagellates (Green *et al.*, 2006). Others are iron/manganese-oxidizing facultative chemoautotrophs (Wang *et al.*, 2011) (\*perhaps sulfur oxidizing Swan 2011) or capable of unusual redox cycling (Handley *et al*., 2009). *Marinobacter* isolates from Antarctic Lakes are similarly capable of anaerobic respiration using dimethyl sulfoxide (DMSO) (Matsuzaki *et al*., 2006) or nitrate (Ward & Priscu, 1997). Either faculty would allow for the presence of *Marinobacter* throughout the water column in Organic Lake and the possibility of occupying multiple functional roles.

*Saccharospirillum*, like the name suggests degrades polysaccharides (\*ref) and is likely associated with algal blooms (\*ref).

RF3, *Halomonas*, *Psychromonas* were concentrated on the 6.5 m sample and are the most likely involved in degradation of particulate matter. RF3 very likely has an anaerobic lifestyle as most 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 (Bowman *et al.*, 2000b; Humayoun *et al*., 2003; Schmidtova *et al*., 2009). However, some members have been found in aerobic environments such as surface waters of hypersaline systems (Demergasso *et al*., 2008; Yilmaz *et al.*, 2012), surface of a freshwater lake (Xing *et al.*, 2009) and compost (Partanen *et al.*,2010).

*\*Halomonas*, *Psychromonas*

*Bacilli* and *Clostridia* were also exclusively present

#### 0.8–0.1 µm size fraction

There was had a large number of eucaryotic sequences on the 0.1 µm size fraction from fungi, *Dictyochophyceae*, *Dinophyceae* and choanoflagellates. These taxa were all found on larger size fractions, except fungi which were uniquely present in the 1.7 m 0.1 µm. The presence of these *Eucarya* on the smallest filter may due to sampling particular stages in their life history; such as a cysts or spores, or degraded cellular material. The mixed zone was dominated by *Pedinellales* that co-varied with unclassified chloroplasts consistent with active photosynthetic cells. *Pedinellales* have only been previously detected in Antarctic lakes from molecular studies (Unrein *et al.*, 2005; Lauro *et al.*, 2011; Yau *et al.*, 2011). They are known to be small, but again their concentration on the 0.1 µm was unusual (\*check).

“*Candidatus* Aquiluna”, in the Luna-1 cluster of *Actinobacteria* (Hahn *et al.*, 2004; Hahn *et al.*, 2009) was most abundant at 1.7 m; however, it was also present in the deep zone. The genus has small cells, <1.2 µm in length (Hahn *et al.*, 2009), consistent with their concentration on the smallest size fraction. Although originally described in freshwater lakes, the same cluster was detected in abundance in Ace Lake (Lauro *et al.*, 2010) and surface Artic seawater (Kang *et al*., 2012) demonstrating they are relevant to polar saline systems. Isolates are aerobic chemoheterotrophs (Hahn *et al.*, 2009) and in Ace Lake surface water they were associated with utilization of labile C and N substrates (Lauro *et al.*, 2010) indicating their ecological function in Organic Lake surface waters. Presence of this clade in both mixed and deep zones implies it may be a facultative anaerobic or sedimenting to the bottom.

The deep samples were distinguished by the presence of candidate divisions OD1 and TM7 which were concentrated on the lake bottom. Their prevalence on the smallest size due to small cell size is supported by other studies where OD1 was found to predominate in <0.2 µm fraction of ground water (Miyoshi *et al.*, 2005). OD1 has been consistently associated with anoxic environments (Harris *et al*., 2004) and genomic fragments from Zodletone Spring, Oklahoma showed oxygen sensitive enzymes related to anaerobic or facultative anaerobic bacteria (Elshahed *et al.*,2005). In the marine environment, it has also been associated with reduced conditions with high sulfur such as sulfate and sulfides (Harris *et al*., 2004; Elshahed *et al.*,2005;\*other ref). The distribution of OD1 is consistent with an anaerobic metabolism and potential involvement in sulfur chemistry.

(\*TM7 sizes).

## Organic Lake functional gene complement and links to taxonomic groups

To determine the functional capacity in Organic Lake, functional molecular markers for C, N and S conversions (Figure 4), as well markers for other processes of interest were retrieved from metagenomic reads. There were differences in the distribution of functional genes according to size fraction and depth indicating a strong link between taxa and function (\*ANOSIM test of mixed vs deep functional complement). (\*RELATED of taxa and functional tables). Variation in the population structure was significantly correlated (Rho: 0.519, significance: 0.3%) by BEST analysis with the abiotic parameters: DO, temperature, TS and TN. The DO gradient has an obvious effect of separating aerobic taxa from anaerobes as described above. It also generates diversity of N and S conversions in the deep zone as oxygen sensitive processes such as N fixation and dissimilatory nitrate and sulfate reduction (discussed below) can occur and is consistent with a correlation between these parameters and community stratification.

## Reduced potential for nitrogen cycling related to N limitation

Potential for nitrogen conversions was greatly reduced in Organic Lake. Low abundance of nitrogenase genes and of diazotrophs indicates a low capacity for N fixation in Organic Lake, which would contribute to N limitation. N fixation was confined to the deep zone (Figure 2B) and principally linked to anerobic *Epsilonproteobacteria*, *Deltaproteobacteria* and *Clostridia* (\*functional table) consistent with their restriction to the lake bottom. Potential for aerobic ammonia oxidation was not detected, nor were ammonia-oxidizing bacteria or archaea supporting a lack of nitrification potential in Organic Lake. This was also the case in nearby Ace Lake (Lauro *et al*., 2011) indicating some limiting factor in the lakes in the Vestfold Hills. Similarly, anaerobic ammonia oxidation (anammox) potential, indicated by the hydroxylamine oxidase (HAO/HZO), was extremely low. Known anammox organisms were not present (\*check) and HAO was linked instead to sulfate reducing bacteria (\*functional table) and may in fact be involved in other N pathways (\*check). This indicates an inability for nitrification to occur in the mixed zone and a very limited potential for ammonia loss in the deep zone.

N cycling potential in Organic Lake was concentrated between assimilation and mineralization/uptake pathways. Assimilation potential was indicated by the GS-GOGAT enzymes and was linked to dominant microbes throughout the lake profile. Assimilatory nitrite reductase was not abundant indicating a predominance of reduced N uptake in the community. Potential for mineralization to ammonia, indicated by glutamate dehydrogenase (GDH), was similarly present throughout the whole lake and associated abundant taxonomic groups. GDH may function in reverse as an ammonium uptake mechanism, particularly in high ammonium concentrations (\*ref). The genetic potential for N metabolism indicates a shift away from oxidized N forms and increased cycling between organic N and ammonia. The higher ammonia concentration in the deep zone, particularly at 6.5 m would result from a higher rate of mineralization than assimilation. Ammonia would also be produced from dissimilatory nitrate reduction to ammonia (DNRA), associated mainly to *Sphingobacteria* and other anaerobic bacteria, (\*Table) in addition to Stickland fermentation (\*figure) by *Clostridia*. Higher rates of organic N catabolism in the deep zone is consistent with the increased potential for degradative processes such as fermentation (\*see C cycle) and anaerobic organisms such as *Clostridia* and the related division RF3 (Figure 2A) (Figure 1).

Denitrification genes and were present throughout the water column (Figure 4B) and was linked to numerous denitrifiying organisms (\*table). Low nitrate and nitrite in the deep zone (Figure 1B, Table1) indicates of depletion by dissimilatory nitrate reduction. The potential for net N loss may have contributed to the establishment of N limitation in the lake. Denitrification enzymes are usually induced by low oxygen or oxidized N (\*ref) and thus expected to only be active in the deep zone. However, denitrification may be inhibited even if conditions appear appropriate; this is the case with Lake Bonney, Antarctica where denitrification occurs in the west lobe, but not in the east lobe of the same lake despite the anoxia, available nitrate and presence of denitrifying *Marinobacter* species (Ward & Priscu, 1997; Ward *et al*., 2005). Furthermore, in the absence of nitrification, denitrification would be limited in Organic Lake by the lack of potential to re-formed oxidized N. The preponderance of assimilation/mineralization pathways geared towards reduced N reflects a “short circuit” of the typical N cycle that would conserve N in a largely closed system, similar to what was proposed to occur in Ace Lake (Lauro *et al*., 2011). Denitrification appears to be the only major pathway for N loss, thus *in situ* rate measurements would be necessary to determine the Organic Lake N budget.

## Carbon cycling

**Primary production by photo and chemoautotrophy**

Oxygenic photosynthesis was presumably mediated out by phytoflagellates and diatoms (Figure 2). Rubisco, a marker gene for the Calvin carbon fixation cycle, was primarily assigned to *Chlorophyta*, consistent with this assertion. However, phosphoribulose kinase (*prk*A), another diagnostic gene of the Calvin cycle, was assigned to *Marinobacter*. This implies the *Gammaproteobacteria*, even the surface heterotrophic lineages, have the capacity for autotrophic carbon fixation. Most likely they have chemolithoautotrophic capacity such as metal oxidation. Aerobic carbon fixation which appears to be linked primarily to *Marinobacter* or Alteromondales in general and so is most abundant in the mixed zone and the very bottom sample. This is consistent with *Marinobacter* having a generalist metabolic strategy and thus many members possessing carbon fixation genes, likely involved in chemolithoautotrophic iron or manganese oxidation. Genes for respiration were abundant throughout the water column and assigned to the major heterotrophic bacterial lineages.

**High biological activity at 6.5 m**Anaerobic carbon fixation and fermentation were processes associated with the increased biological activity at 6.5 m. (\*why not at 6.7 m too?)

**Absence of C cycle genes**: Genes for methanogenesis were absent (Figure 4A). This was expected as methanogenesis usually only occurs when alternate electron acceptors are depleted, but Organic Lake sulfate concentrations in the deep zone are high (Franzman *et al.*, 1987). (\*ref paper SRB vs methanogens) Moreover, methane oxidation enzymes that were detected are not indicative of active methane production. They related to alkane hydroxylases and therefore most likely involved in hydrolysis of compounds such as phenol, which has been previously detected in the bottom waters of Organic Lake (Roberts & Burton 1993a; Roberts *et al.*, 1993b).

**Net loss:** A net loss in essential elements implies that a there may be an influx of exogenous nutrients occurs to sustain the lake system. However, external input, such as from glacial melt-water, could only occur in the summer months when the lake is ice-free. Furthermore, the water column structure is characteristic of a negative water balance (\*Gibson) indicating the Organic Lake system has been largely closed in the recent past. Thus, if external inputs occur, they are episodic and would necessitate interim strategies for C, N and S conservation.

### Potential for lithoheterotrophy

CO is an indirect green house gas as it contributes to methane and nitrous oxide concentrations (\*refMoran). It is formed in aquatic environments during photochemical degradation of organic molecules (\*ref). Purely carboxydotrophic bacteria oxidize CO to CO2 using the enzyme CO dehydrogenase and fix a proportion of the CO2 with ribulose 1,5-bisphosphate carboxylase (RuBisCo). Alternatively, it has been proposed that roseobacters may fix CO2 by anaplerotic mechanisms (\*ref Moran 2007). Many roseobacters such as *R. pomeroyi* only possess CO dehydrogenase and are able to oxidize CO at low concentrations as an energy source and assimilate organic carbon for growth thus limiting organic carbon oxidation. (\*link CO oxidation to taxon).

### Diverse proteorhodopsin homologs are linked to most bacterial lineages present

The first rhodopsin found in bacteria, termed proteorhodopsin (PR) acts as a light-driven proton pump and was hypothesized to be used for energy generation (Béjà *et al*., 2000). Metagenomic studies have since shown PR are diverse, widely distributed in the surface ocean (Rusch *et al*., 2007) and associated with diverse bacterial clades including *Alphaproteobacteria* (de la Torre *et al.*, 2003) and *Bacteroidetes* (Venter *et al.*, 2004) as well as *Euryarchaeota* (Frigaard *et al.*, 2006). Actinorhodopsins (Sharma *et al.*, 2008; Sharma *et al.*, 2009) and xanthorhodopsin (\*ref) form a clade related to PR that we will refer to as actino-xanthorhodopsins.

A total of 399 reads matching to rhodopsins were detected in Organic Lake, which formed 124 clusters at 90% amino acid identity. Phylogenetic analysis revealed six well-supported rhodopsin groups named for their taxonomic affiliation: MAR-R, OL-R1, OCT-R, SAL-R, AQU-R and PSY-R (Figure S8). Only the PSY-R clustered with the PRs showing most Organic Lake rhodopsin diversity was within the actino-xanthorhodopsin clade. All groups had an L or M residue corresponding to position 105 in the SAR86 PR denoting tuning to surface green light (Man *et al.*, 2003; Gomez-Consarnau *et al.*, 2007), which is consistent with the shallow water in Organic Lake and is characteristic of coastal samples (Rusch *et al.*, 2007).

MAR-R, PSY-R, OCT-R and AQU-R groups (Figure S8) all clustered with homologs of genera detected in the lake, namely *Marinobacter*, *Psychroflexus*, *Octadecabacter* and “*Candidatus* Aquiluna” (Figure 2C, Table S3). Xanthorhodopsin was described from the sphingomonad *Salinibacter* *ruber* (\*ref), thus SAL-R likely originates from other *Sphingobacteria*, such as *Lewinella* or the environmental clade E6ac02 (Table S3). The distribution of the abundant MAR-R and PSY-R rhodopsins agrees with the distribution of *Marinobacter* and *Psychroflexus* (\*Figure) further supporting to their phylogenetic origins. However, the most abundant group, OL-R1, had no close homologs from GENBANK. From its high abundance and concentration on the 3.0 µm fraction, OL-R1 group most likely originated from the *Rhodobacterales* including *Roseovarius* (\*Figure). Although it is possible that OL-R1 was encoded by *Flavobacteria* as they are similarly abundant in the 3.0 µm fraction, all known *Flavobacteria* possess PRs, not actino-xanthorhodopsins.

If there are approximately 3 000 bacterial SSU sequences, assuming PR to be single copy and SSU copy number to range from one to ten, 13–100% of Organic Lake bacteria have a PR. This is comparable to the Mediterranean Sea estimates of 13% (Sabehi *et al.*, 2005).(\*use recA or radA, as Sharma *et al.*, 2008 which saw Punta Cormorant has 36% vs open ocean 63%).

Recently, proteorhodopsins of marine Flavobacteria and *Vibrio* have been associated with light-dependent energy generation (Gomez-Consarnau *et al*., 2007), especially under low carbon conditions (\*ref). This is a potential mechanism for conserving carbon for growth or inhabiting low oxygen environments. This is most likely to be the case for Organic Lake *Psychroflexus* as it is both taxonomically related to *Dokdonia* and has a PR of the same type. These data suggest there is selection for rhodopsins in the dominant Organic Lake bacterial lineages. All sequenced homologs similar to Organic Lake rhodopsins originated from polar and/or lake species (Gosink *et al.*, 1997;Ward & Priscu, 1997; Bowman *et al.*,1998; Antón *et al.*, 2002;Hahn, 2009; Kang *et al.*, 2012). If these proteorhodopsin homologs in Organic Lake add to energy generation, this would indicate mixotrophy is a common and may be an important adaptive strategy for surviving polar lake environments.

### Aerobic Anoxygenic Photosynthesis

Aerobic and anaerobic anoxygenic photosynthesis may be occurring and would be mediated by the Roseobacters present such as *Roseovarius*. *R. tolerans* is the type species of the genus and was isolated from Ekho Lake, a meromictic hypersaline lake in the Vestfold Hills (\*Labrenz *et al.*, 1999). It was found to produce bacteriochlorophyll A when grown in the dark, but continuous dim light inhibited production (\*Labrenz *et al.*, 1999). (\*Look for the pufLM genes).

## Molecular basis for unusual sulfur chemistry

### Absence of typical S cycle

In Organic Lake, sulfur oxidation genes were undetectable and dissimilatory sulfate reduction extremely limited (Figure 4C). Consistent with this, sulfur-oxidizing *Epsilonproteobacteria* and sulfate-reducing *Deltaproteobacteria* were present at very low abundance (Figure 2A, Table S3). Despite the presence of sulfate, sulfate-reducers appear to be limited and potential for sulfur cycling typical in other stratified water bodies is greatly reduced. Several reasons have been suggested such as high salinity (check other saline lakes\*), oxidizing environment (\*check), cold? In the absence of sulfide, sulfur-oxidizers would be limited to utilizing other substrates (\*double check the genes for sulfur cycles\*check what else they can use). As described for the N cycle, Organic Lake appears to have a reduction in pathways for cycling inorganic S typical in other meromictic lake systems and a prevalence of assimilation/mineralization pathways. Absence of these pathways has perhaps given rise to the development of unusually high levels of alternative reduced sulfur compounds.

### DMSP and DMS metabolism

To determine the potential source of high DMS in the bottom waters of Organic Lake, the presence of DMSP lyases were investigated. These enzymes catalyse the breakdown of DMSP forming DMS as a by-product and this process is the major source of DMS in the marine environment (reviewed in Curson *et al.*, 2011b). Homologs of DMSP lyases DddD,DddL andDddP, were detected in Organic Lake at levels comparable to other dominant processes such as respiration and fermentation (\*figure:DMS\_cycle). DddD, was the most abundant of the Organic Lake DMSP lyases (70%).

There were two main DddDtypes: MAR-dddD and OL-dddD (Figure S6). MAR-dddD grouped with a *Marinobacter* sp. ELB17 homolog and had a distribution consistent with the distribution of *Gammaproteobacteria* including *Marinobacter* (Figure S6). OL-dddD was the major DddD type. It did not cluster with good support with sequenced DddD, but had highest identity (~80%) to *ddd*Dfrom *Halomonas* sp. HTNK. The distribution of OL-dddD did not reflect that of *Halomonas* (\*figure), rather its abundance on 3.0 µm fraction suggests the *Rhodobacteraceae* as the most likely origin (Figure S6).

Two DddLgroups were detected in Organic Lake: SUL-dddL and a MAR-dddL (Figure S7). SUL-dddL clusters with *Sulfitobacter* sp. EE-36 and other *Rhodobacteraceae* (\*check if they are from different Roseobacter clades). MAR-dddL forms a separate clade from the known DddLand includes a hypothetical protein from *Marinobacter manganoxydans* MnI7-9. This finding suggests MAR-dddL clade is an unrecognized branch of this enzyme family and is the first report of DddLlinked to *Gammaproteobacteria*. Whether it confers the Ddd phenotype requires further confirmation although in *Sulfitobacter* sp. EE-36 the *dddL* gene alone is sufficient for DMS generation (Curson *et al.*, 2008). The SUL-dddL is most abundant at 6.5 m depth while MAR-dddL is most abundant at 5.7 m. Both are predominantly located on the 0.8 µm fraction, which suggests the origin of both Organic Lake *ddd*Ltypes are *Gammaproteobacteria*. The MAR-dddL fits the distribution of *Marinobacter* while the abundance of SUL-dddL at 6.5 m indicates this homolog originated from *Psychromonas* or *Halomonas* which are predominant at 6.5 m. (\*none of the Roseobacters on 0.8 um???)

\*dddP \*Check other DMSP pathways.

DMSP degradation appears to be the main source of the high concentration of DMS in Organic Lake. The most abundant DMSP lyase, OL-dddD, was concentrated in the deep zone which is where the DMS concentration is highest. As in the marine environment, this function is likely performed by *Roseobacter-*cladeand *Gammaproteobacteria* such as *Marinobacter*, *Psychromonas* and *Halomonas*. In addition, reduction of DMSO to DMS (\*figure) or a yet undefined pathway of anaerobic production from cysteine (\*ref) may be further sources of DMS in the deep zone. Usually methanogenic Archaea or sulfate-reducing bacteria break down DMS in anoxic conditions (\*ref). Since only a very low abundance of sulfate-reducing bacteria and *dsr*A genes were detected and methanogens are absent, DMS can accumulate in Organic Lake as a metabolic end-product. DMSP in Organic Lake likely originates from degraded algal material.

## Discussion points

What are the sulfate reducing Deltaproteobacteria doing if not sulfate reduction?? Could they make DMS anaerobically? Can sulfur go to DMS? Amino acids to DMS?

{\*intro: The bottom waters of Organic Lake were found to be anoxic, but not sulfidic or methanogenic (Franzmann *et al.*, 1987b; Gibson *et al.*, 1991). Although sulfates and organic acids have been recorded (Franzmann *et al.*, 1987b; Gibson *et al.*, 1994), the cold and salinity six times that of seawater, appears to preclude the establishment of sulfate-reducing bacteria and phototrophic sulfur bacteria (Burke & Burton, 1988) indicating other microbes are involved in the unusual sulfur chemistry. }

{\*selection for cold species: The Organic Lake 16S composition was most like that of meromictic hypersaline Antarctic lakes Ekho and Bonney. They are characterized by an abundance of *Gammaproteobacteria*, *Alphaproteobacteria* and *Bacteroidetes* and scarce or absent photolithoautotrophic bacteria as well as haloarchaea (Bowman *et al.*, 2000b; Glatz *et al*., 2006). The salinity in these lakes (150–180) appears to be too high for planktonic photosynthetic bacteria such as *Synechococcus* relatives (Powell *et al.*, 2005) and anoxygenic sulfur bacteria (Burke & Burton, 1988). In contrast, it appears too low for haloarchaea found to dominate Deep Lake (Bowman *et al.*, 2000b). These findings correspond well to a study of solar salterns along a salinity gradient that showed Cyanobacteria were confined to salinity <65 and haloarchaea at salinity >190 (Ghai *et al.*, 2011). Salinity is therefore a crucial constraining factor for Organic Lake species composition; a factor that is inter-related to other variables such as freezing point and thus ice-cover and available light. This has lead to species diversity that is reduced to the point that entire divisions present in other Antarctic lakes are excluded in Organic Lake. }

{\**functional:*The majority of the genetic potential was restricted to the 0.8 and 3.0 µm size fractions evident in the higher percentage of ORFs with matches to KEGG genes ( average 55%) compared to 0.1 µm fraction (average 28%) (Table S1). The lack of ascribed functional genes in the 0.1 µm filter reflects the paucity of cellular life in that fraction and possibly the high representation of candidate divisions, with poorer representation of functional sequences.}

## Acknowledgements

## References

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

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

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

Antón J, Oren A, Benlloch S, Rodríguez-Valera F, Amann R, Roselló-Mora R. (2002) *Salinibacter ruber* gen. nov., sp. nov., a novel extremely halophilic member of the *Bacteria* from saltern crystallizer ponds. *Int J Syst Evol Microbiol* **52**: 485–491.

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

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.

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, McMeekin TA. (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. (2000b) 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.

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.

Chen YG, Cui XL, Wang YX, Tang SK, Zhang YQ, Li WJ, Liu JH *et al*. (2009) *Psychroflexus sediminis* sp. nov., a mesophilic bacterium isolated from salt lake sediment in China. *Int J Syst Evol Microbiol* **59**: 569–573.

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. (1993) Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* **18**: 117–143.

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

Curran MAJ and Jones GB. (1998) Spatial distribution of dimethylsulfide and dimethylsulfonioproprionate 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 dimethysulfonioproprionate 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 dimethylsulfonioproprionate lyase found in taxonomically diverse species of Proteobacteria. *ISME J* **5**: 1191–1200.

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

de la Torre JR, Christianson LM, Béjà O, Suzuki MT, Karl DM, Heidelberg J *et al.* (2003) Proteorhodopsin genes are distributed among divergent bacterial taxa. *PNAS* **100**: 12830–12835.

DeSantis Jr. TZ, Hugenholtz P, Keller K, Brodie EL, Larsen N, Piceno YM *et al*. (2006) NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Res* **34**:W394–399.

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.

Demergasso C, Dorador C, Meneses D, Blamey J, Cabrol N, Escudero L, Chong G. (2010) Prokaryotic diversity pattern in high-altitude ecosystems of the Chilean Altiplano. *J Geophys Res* **115**: G00D09

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.

Donachie SP, Bowman JP, Alam M. (2005) *Psychroflexus tropicus* sp. nov., an obligately halophilic *Cytophaga-Flavobacterium-Bacteroides* group bacterium from an Hawaiian hypersaline lake. *Int J Syst Evol Microbiol* **54**: 935–940.

Edgar RC. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nuc Acids Res* **32**: 1792–1797.

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.

Frigaard NU, Martinez A, Mincer TJ, DeLong EF. (2006) Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. *Nature* **439**: 847–850.

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.

Gärdes A, Kaeppel E, Shehzad A, Seebah S, Teeling H, Yarza P, Glöckner FO *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, Betrand JC. (1992) *Marinobacter hydrocarbonoclasticus* gen. nov., sp. nov., a new, extremely halotolerant, hydrocarbon-degrading marine bacterium. *Int J Syst Bacteriol* **42**: 568–576.

Ghai R, Pašić L, Fernández AB, Martin-Cuadrado A-B, Mizuno CM, McMahon KD, Papke RT *et al.* (2011) New abundant microbial groups in aquatic hypersaline environments. *Sci Rep* **1**: srep00135.

\*Gibson JAE *et al.* (1989) Temperature profiles of saline lakes of the Vestfold Hills. *ANARE Research Notes,* No.67, 75pp

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, Burton HR, Gallagher JB. (1995) Meromictic Antarctic lakes as indicators of local water balance: structural changes in Organic Lake, Vestfold Hills 1978–1994.  *ANARE Research Notes*, No.94, 16pp.

\*Gibson JAE *et al.* (1996) Meromictic Antarctic lakes as recorders of climate change: the structures of Ace and Organic Lakes, Vestfold Hills, Antarctica. *Papers and Proceedings of the Royal Society of Tasmania* **130**:73–78.

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.

Gosink JJ, Herwig RP, Staley JT. (1997) *Octadecabacter articus* gen. nov., sp. nov., and *O. antarcticus*, sp. nov., nonpigmented, psychrophilic gas vacuolate bacteria from polar sea ice and water. *System Appl Microbiol* **20**: 356–365.

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.

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.

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.

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.

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.

Kang I, Lee K, Yang S-J, Choi A, Kang D, Lee YK, Cho J-C. (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 Ecol* **39**: 91–100.

La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, Merchat C *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 Bacter* **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.

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.

\*Ludwig W., *et al.* (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.

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.

Lee ZM, Bussema C 3rd, Schmidt TM. (2009) rrnDB: documenting the number of rRNA and tRNA genes in bacteria and archaea. *Nucleic Acids Res* **37** (Database issue): D489–D493.

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.

McCammon SA and Bowman JP. (2000) Taxonomy of Antarctic *Flavobacterium* species: description of *Flavobacterium gillisiae* sp. nov., *Flavobacterium tegetincola* sp. nov.and *Flavobacterium xanthum* sp.nov., nom. rev. and reclassification of [*Flavobacterium*] *salegens* as *Salegentibacter salegens* gen. nov., comb. nov. *Int J Syst Evol Microbiol* **50**: 1055–1063.

\*Millero FJ, Chen CT, Bradshaw A, Schleicher K. (1980) A new high pressure equation of state for seawater. *Deep Sea Res A* **27**: 255–264.

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.

Naganuma T, Hua PN, Okamoto T, Ban S, Imura S, Kanda H. (2005) Depth distribution of euryhaline halophilic bacteria in Suribati Ike, a meromictic lake in East Antarctica. *Polar Biosci* **28**: 964–970.

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.

Pagaling E, Wang H, Venables M, Wallace A, Grant WD, Cowan DA, Jones BE *et al.* (2009) Microbial biogeography of six salt lakes in Inner Mongolia, China and a Salt Lake in Argentina. *Appl Environ Microbiol* **75**: 5750–5760.

Partanen P, Hultman J, Paulin L, Auvinen P, Romantschuk M. (2010) Bacterial diversity at different stages of the composting process. *BMC Microbiol* **10**: 94.

Powell LM, Bowman JP, Skerratt JH, Franzmann PD, Burton HR. (2005) Ecology of a novel *Synechococcus* clade occurring in dense populations in saline Antarctic lakes. *Mar Ecol Prog Ser* **291**: 65–80.

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.

Rivière D, Desvignes V, Pelletier E, Chaussonnerie S, Guermazi S, Weissenbach, Li T *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, Heidelbergg 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.

Sabehi G, Loy A, Jung K-H, Partha R, Spudich JL, Isaacson T, Hirschberg J *et al*. (2005) New insights into metabolic properties of marine bacteria encoding proteorhodopsins. *PLoS Biol* **3**: e273.

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.

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.

Sharma AK, Sommerfeld K, Bullerjahn GS, Matteson AR, Wilhelm SW, Jezbera J, Brandt U *et al*. (2009) Actinorhodopsin genes discovered in diverse freshwater habitats and among cultivated freshwater *Actinobacteria*. *ISME J* **3**: 726–737.

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.

Tian F, Yu Y, Chen B, Li H, Yao Y-F, Guo X-K. (2009) Bacterial, archaeal and eukaryotic diversity in Artic sediment as revealed by 16S rRNA and 18S rRNA gene clone libraries analysis. *Polar Biol* **32**: 93–103.

Todd JD, Rogers R, Li YG, Wexler M, Bond PL, Sun L, Curson ARJ *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 dimethylsulfonioproprionate 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, Page PCB *et al.* (2010) Molecular dissection of bacterial acrylate catabolism – unexpected links with dimethylsulfonioproprionate 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, dimethylsulfonioproprionate 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.

Van Trappen S, Mergaert J, Van Eygen S, Dawyndt P, Cnockaert MC, Swing J. (2002) Diversity of 746 heterotrophic bacteria isolated from microbial mats from ten Antarctic lakes. *System Appl Microbiol* **25**: 603–610.

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA *et al*. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.

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.

Wang H, Li H, Shao Z, Liao S, Johnstone L, Rensing C, Wang G. (2011) Genome sequence of deep-sea Manganese-oxidizing bacterium *Marinobacter manganoxydans*. *J Bacteriol* **194**: 899–900.

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.

\*Wu J, Mao X, Cai T, Luo J, Wei L. (2006) KOBAS server: a web-based platform for automated annotation and pathway identification. *Nucleic Acids Res* **34**: W720–W724.

Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li CY, Wei L. (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.

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.

Yoon JH, Kang SJ, Jun YT, Oh TK. (2009) *Psychroflexus salinarum* sp. nov., isolated from a marine solar saltern. *Int J Syst Evol Microbiol* **59**: 2404–2407.

Zhang H, Hosoi-Tanabe S, Nagata S, Ban S, Imura S. (2010) *Psychroflexus* lacisalsi sp. nov., a moderate halophilic bacterium isolated from a hypersaline lake (Hunazoko-Ike) in Antarctica. *J Microbiol* **48**: 160­–164.

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.