Organic Lake

# 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.

# Introduction

Antarctic biota live at the extremes of temperature and salinity under a polar light regime. The lakes are a rare source of liquid water and an oasis for life. They are ideal location to study questions of biogeography, evolution and a potential source of novel life and genes. Organic Lake is known to have high levels of reduced sulfur compounds and a member of the newly discovered virophage family that may be mediate ecosystem stability. The ability to encapsulate a large proportion of the species diversity allows us to infer which species may be mediating particular biological processes. Meromictic lakes are ideal systems to link species to microbial processes as abiotic variables exist along a spatial gradient.

In this study we aimed to:

1. Determine the microbial composition population structure of Organic Lake. Determine if microbial composition varies along the depth/chemical gradient and if so, which taxa differ.
2. Describe the functional capacity of the microorganisms in the lake.
3. Link microbial processes to phylogenetic groups or to physical-chemical parameters.
4. Examine possible microbe-microbe interactions (link microbial groups with each other).

# Materials and Methods

## Background on sample site

Organic Lake is located on Long Peninsula in the Vestfold Hills, an ice-free oasis on the eastern shore of the Prydz Bay, Princess Elizabeth Land, East Antarctica (figure: vestfold map). It consists of remnant seawater that was trapped approximately 10,000 BP when the continental ice-shelf retreated and isostatic rebound caused the land to rise above sea-level (Zwartz *et al*., 1998; Gibson, 1999). Complete separation from the ocean occurred approximately 3,000 years before present (Bird *et al*., 1991) and the water has since concentrated to approximately 6 × the salinity of seawater. Organic Lake was recorded to have an extremely high concentration of dimethyl sulfide (DMS) (Franzmann *et al.*, 1987) and other polysulfides (Roberts *et al*., 1993) in the bottom waters.

## Sample collection and preparation

Water was collected on 10 November 2008 along a depth profile when the lake was covered by a 0.8 m layer of ice through a 30cm hole above the deepest point. Lake water was passed through a 20 µm pore size pre-filter 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. The samples were collected at 1.7, 4.2, 5.7, 6.5 and 6.7 m depths. Filters were placed in a storage buffer (\*composition) and kept frozen at -80ºC. Water was also collected at the same sample depths for microscopic and chemical analysis and frozen -80ºC. Samples for microscopy were preserved in formaldehyde (1% v/v).

An *in situ* vertical profile of pH, conductivity, turbidity, dissolved oxygen (DO), pressure and depth 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 depth sample corresponded to the turbidity maximum and 6.5 m 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 using the formula described by Fofonoff and Millard (1983). However, salinity was likely underestimated as the conductivity to salinity relation holds for practical salinity between 2 and 42 and Organic Lake salinity was much higher. Nitrate, nitrite, ammonia, total nitrogen, dissolved reactive phosphorus, total phosphate, total carbon, total dissolved carbon and total sulfur were determined by the Analytical Centre\* (Tasmania). Principal Component Analysis (PCA) was performed using the Primer V.6 statistical package (\*ref) on the (\*ranked?) normalized physical and chemical parameters to visualize how abiotic factors varied with depth.

DNA was extracted from the membrane 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 processed to remove low quality bases as previously describes (Lauro *et al.*, 2011). Total protein was extracted from the membrane filters as previously described (Ng *et al.*, 2010; Lauro *et al.*,2011).

## Epifluorescence microscopy

Fixed water samples were vacuum filtered onto a 0.015 µm pore-size filter (Whatman\*) with a 0.45 µm pore-size backing filter. The 0.015 µm filter was mounted onto a glass slide with ProLong Gold (\*brand) and cells and virus-like particles (VLPs) stained with SYBR gold (\*brand) by adding 2 µl m (25 × dilution) into the mountant. All water used in filtration was autoclaved and filtered to <0.015 µm. Prepared slides were visualized in an epifluorescence microscope (Olympus BX61) under excitation with blue light (\* nm, filter set 2). Cell and VLP counts were performed on the same filter with over 30\* random fields of view. Cell density was calculated by \*.

## Diversity analyses

### Cellular diversity analysis

Cellular diversity 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). This software implements hidden markov model based searches to retrieve 12S/16S/18S sequences from Bacteria, Eucarya, Archaea and plastids and trims off regions outside of the SSU gene. 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). QIIME was then used to choose a representative sequence from each set of OTUs and classify the representative set to the \*family level by using the RDP classifier (Wang *et al*., 2007) trained against the SILVA (release 108) sequences (www.arb-silva.de). Assignments were accepted to the highest taxonomic rank with bootstrap value ≥ 85 %. This prevented low confidence matches contributing to counts of high-confidence phylogenetic groups while avoiding grouping all the unclassified taxa together. QIIME was used to calculate alpha diversity indices: Chao1, Simpson, Shannon and observed species.

To allow comparison of the relative abundance of taxa between samples, the number of SSU matches per sample filter was normalized to the average number of reads (403,577) obtained for each sample filter. Statistical analysis on the relative SSU gene abundances was performed using the PRIMER Version.6 package (\*Clarke 1993, Clarke 2006). The SSU gene counts of each sample filter were square root transformed to reduce the contribution of highly abundant taxa. The Bray-Curtis similarity of the 16S assemblage between each sample filter was computed. Patterns in the resulting similarity matrix were visualized using hierarchical clustering (CLUSTER) and non-parametric Multidimensional Scaling (MDS) routines. The CLUSTER analysis groups samples at successively smaller number of clusters at decreasing thresholds of similarity. Statistical significance of the clusters was determined by the ‘similarity profile’ (SIMPROF) permutation test. To determine if physical and chemical parameters from the November 2008 depth profile and the patterns in bacterial and chloroplast assemblages were correlated, BEST and LINKTREE analyses were performed considering following abiotic variables: salinity, temperature, turbidity, percent dissolved oxygen saturation, pH, nitrate, nitrite, ammonia, dissolved reactive phosphate, total carbon, total dissolved carbon, total nitrogen, total phosphorus and total sulfur. 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. LINKTREE takes the subset of variables that explains the biotic structure and finds the thresholds that maximizes the separation of the assemblage into groups by binary separation. Significance of the LINKTREE separations tested with SIMPROF (see above).

### Viral diversity

## Functional potential of Organic Lake

### Kegg

# Results

## Physical and chemical properties of Organic Lake

During the November 2008 sampling Organic Lake had a maximum depth of 6.75 m (figure: bathymetry) and measured 3.874 m above mean sea level. Water temperature was below zero (average of -13 ºC) and the estimated maximum practical salinity was 188 (figure:profiles\*cut off graphs to 1.5 m). From 1.5 m to approximately 5 m the physical and chemical parameters remained fairly constant indicating mixing in this zone. Below 5.7 m dissolved oxygen, salinity, temperature and pH varied continuously with depth. However, cell density, VLP density, ammonia, dissolved reactive phosphate, total C, total N, total P and total S were at a maximum at 6.5 m while turbidity, nitrate and nitrite were at a minimum (figure: profiles) indicating some stratification below 5.7 m. We designated 0–5.7 m to be the mixolimnion, as that appears to be the deepest extent of mixing, and below to be the monimolimnion (\*density converter stability of stratification calculation). The separation between the mixolimnion and monimolimnion is evident through PCA analysis, which shows the samples in the mixolimnion grouped together while the other samples separated along the PC1 axis with depth. PC1 described 63.1% of the variation with a roughly equal contribution from DO, conductivity, temperature, nitrite and TOC. PC2 described 26.1% of the separation, mainly between the 6.5 m and 6.7 m samples and was driven primarily by nitrate, turbidity and total S (\*include cell counts?).

The C:N:P ratio in the mixolimnion is approximately 60:5:1 and peaks in the monimolimnion at 22:3: 1 (table: sample enviroparams). This shows that nitrogen is limiting throughout the water column (compared to the ocean) but is least limited at 6.5 m indicating relatively high microbial production. The pattern of low nitrate and high ammonia at the oxycline is consistent with nitrate reduction under suboxic conditions, which occurs in other Antarctic lakes such as the west lobe of Lake Bonney (Voytek *et al*., 1999). Several taxa identified in Organic Lake (\*see section below) are related to Antarctic bacteria capable of nitrate reduction. However, Bowman *et al.* (2000b) hypothesized that redox potential was too high in Organic Lake for anaerobic respiration to occur (\*relict activity?). Organic Lake is enriched in sulfur compared to similar Antarctic Lakes (\*check). Salinity is purportedly too high for sulfate reducing bacteria (Franzmann *et al*., 1987a)or phototrophic sulfur bacteria to occur (Burke & Burton, 1988)(\*check other lakes such as Pendant, Burton? and Bonney, Vida for the presence of sulfate reducers and GSB). This is consistent with the lack of these species in the taxonomic analysis (see below) and may have lead to accumulation of sulfur compounds in Organic Lake.

## Cellular diversity and distribution

A total of 3,959 reads matching to SSU were retrieved from the November 2008 profile. Of the 16S rDNA sequences, X reads were classified as bacterial or chloroplast compared to a 2 archaeal sequences indicating Bacteria were numerically dominant in Organic Lake. Direct microscopic counts of cells in the size range typical of bacteria or archaea (figure 3:microscopy images) were taken to be bacteria.

Diversity indices (table or figure: diversity indices) between sample filters and sample depths were not significantly different from one another indicating diversity is similar throughout the water column. The estimate of total species richness (Chao1) was much higher than previously calculated from a 16S clone library of the sediment (Bowman *et al*., 2000b). (\*Calculate percentage of diversity sampled?) This is due to the use of metagenomic reads when forming OTUs inflating the apparent number of OTUs and occurs for several reasons. Non-overlapping reads that cover different sections of the SSU gene will not be grouped as the same OTU if that gene is not present in the SILVA release 108 reference set. A read may match group with a partial sequence in the SILVA reference database, but if a large proportion of the read is outside the reference sequence, it will form its own OTU. Furthermore, even if two reads originate from the same SSU gene, some regions are more unique so a read that matches to a less unique region may not cluster with the correct OTU.

**Bacteria**

The majority of bacteria were from oxidative heterotrophic groups except for a very small number of cyanobacteria detected. The dominant bacterial classes were Gammaproteobacteria and Flavobacteria.Flavobacteria consisted primarily of *Psychroflexus*, (\*more?) Previously, several strains of *Psychroflexus* have been isolated from Organic Lake. Gammaproteobacteria: *Marinobacter*. Alphaproteobacteria were also fairly abundant, especially in the 3.0 µm fraction, and were primarily of the family Rhodobacteraceae. (\*complete table of taxa and what they do.)

Other bacterial classes present in all samples include Actinobacteria, Cytophagia, Sphingobacteria, Opitutae, unclassified cyanobacteria, OD1 and Rf3.

(\*look up what these genera do, what metabolic class? Linked to sulphur? Linked to ice and algae?).

**Eucarya**

The dominant eucarya were the photosynthetic flagellates from the families Chlorophyceae (green algae) andDictyochophyceae (silicoflagelates) predominantly of the genera *Dunaliella* and *Pseudopedinella* respectively. Bacillariophyceae (diatoms), Dinophyceae (dinoflagellates) (\*photosynthetic dinos?) and heterotrophic choanoflagellates (class Codonosigidae) (\*Check the choanoflagellate in Organic Lake taxonomy) were present at low abundances throughout the water column. Two groups showed highly localized distributions: fungi and ciliates. Fungi were restricted to the 0.1 µm fraction of the 1.7 m sample while ciliate signatures (Intramaculata; Spirotrichea) were detected only the 0.1 µm fraction on the 6.8 m. Eucaryal abundance, more than composition, varied down the water column with the highest abundance restricted to the mixolimnion. This is consistent with the majority of eucarya being phototrophs or aerobic heterotrophs and thus localizing to regions of high light and oxygen respectively. Signatures in the monimolimnion may represent sedimenting cells.

(\*discuss: Franzmann paper that they are Dunaliella are dominant and choanoflagellates are present but no Chaetoceros. Chaetoceros may be transient members of the population from an ice community because (Wright and Burton, 1981) made no mention of them either. Composition is quite different from 2006 samples in the way that there are no prasinophytes like pyramimonas detected and Dunaliella is at much higher numbers. Mention that silicoflagellates were not previously detected in Organic, first report of dictyochophyceae in Antarctic lakes was from Unrein 2005, so may be important in Antarctic but were often missed. Fungi and ciliates being in small size fractions is perplexing. Fungi found in Bielewicz 2010 and Unrein 2005. Discuss the possible succession of eucarya in the lake. Perhaps link to Fedes models of strain cycling due to viral pressures. Also potential link to the polar night transition (Bielewics 2010) that some taxa a more light tolerant. Eucarya occupying a main role as primary producers. How do the heterotrophs survive the low oxygen??)

**Cellular composition is vertically stratified. The smaller size fractions separate into mixolimnion and monimolimnion, 3.0 um bottom sample clusters with surface.** MDS (figure:MDS) and CLUSTER (figure:16scluster) (\*try hierarchical clustering) analysis of the bacterial and eucaryal community composition down the water column of Organic Lake showed samples separated according to both size fraction and depth. The 0.1 µm samples displayed the largest degree of separation with the mixolimnion samples (1.7-5.7m ) forming one cluster and the two monimolimnion samples (6.5 m and 6.7 m) another. Relative abundance of some groups varied vertically (figure:simper). Verrucomicrobia, Actinobacteria and Candidate divisions OD1, TM7 and SR1 were most abundant in the 6.7 m sample. Alphaproteobacteria and Rf3 peaked at 6.5 m depth. Chloroplasts and eucarya were largely restricted to 1.6 to 5.7 m ( figure: SIMPER).

**Relationship between community assemblage patterns and environmental parameters measured:** Bacterial assemblages were significantly correlated (0.516 correlation, 0.3% significance) with DO%, temperature and total sulfur. This indicates the bacterial community structure, is to a small extent, determined by these factors. (\*RELATE to the species composition?).

## Viral diversity and distribution

**What viruses are present in Organic Lake profile? Does the viral composition vary with depth?Is the viral composition correlated with any physical or chemical parameters?**

## Whole system functions

**What can we piece together about nutrient cycles in Organic Lake from the taxa we know to be present?Do any taxa co-vary over depth? Does the viral composition vary with and cellular taxa?**