Organic Lake

# Introduction

The presence of exclusively microbial biota of reduced richness makes Organic Lake an ideal site to study microbial diversity and dynamics and its relationship with ecosystem function.

In this study we aimed to:

1. Describe the microbiota of Organic Lake
2. Determine if microbial composition varies along the depth gradient and if so, which taxa differ.
3. Determine if there microbial composition at the surface changes over time and if it changes reproducibly over time.
4. Describe the functional capacity of the microorganisms in the lake.
5. Link microbial processes to phylogentic groups or to physical-chemical parameters.
6. 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 (Gibson 1999, Zwartz). Complete separation from the ocean occurred approximately 3,000 years BP (Bird 1991) and the water has since concentrated to approximately 6 × the salinity of seawater.

## Sample collection and preparation

The following water samples were collected from Organic Lake (see Table 1 for sampling details):

1) Surface water when the lake was completely ice-free, December 24, 2006.

2) A depth profile collected through a 30cm hole above the deepest point when the lake was frozen, November 10, 2008

3) Surface water when the lake was partially ice-covered, December 12, 2008.

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 as described previously (Ng 2010, Lauro 2010, Yau 2011). Between 1 – 2 L of lake water was sufficient to clog the filters. Lake water for subsequent microscopic analysis was preserved in formaldehyde (1% v/v) and frozen at -80 ºC.

During the November 2008 sampling, water was collected after each stage of filtration and similarly preserved for microscopic analysis. A vertical profile of pH, conductivity, turbidity, dissolved oxygen (DO), pressure and depth was measured (YSI sonde model v6600). Temperature was measured using a minimum-maximum mercury thermometer. The samples were collected at 1.7, 4.2, 5.7, 6.5 and 6.7 m: of which 5.7 m corresponded to the turbidity maximum. Conductivity at *in situ* temperature was converted to conductivity at 15 ºC according to the relation described by (Gibson 1999). The adjusted conductivity brings the temperature within the acceptable range to estimate salinity using the formula described by Fofonoff and Millard (1983). However, salinity is likely underestimated as the conductivity to salinity relation holds for practical salinity between 2 and 42 while the practical salinity of Organic Lake is much higher. Principal Component Analysis (PCA) was performed using the Primer v.6 statistical package on the (\*ranked?) normalized physical and chemical parameters to visualize how samples variation with depth.

DNA was extracted from the membrane filters as previously described (Ng 2010, Lauro 2010, Yau 2011). DNA from all samples was sequenced using the Roche GS-FLX titanium sequencer. Additional Sanger sequence (\*ABI sequencer) was obtained for the 2006 sample 0.1 µm filter from \*4 Kb the fragments cloned into \*vector as described previously (Rusch). Reads processed to remove low quality bases as previously describes (Lauro 2010).

Total protein was extracted from the membrane filters as previously described (Ng 2010, Lauro 2010, Yau 2011).

## Diversity analyses

### Bacterial diversity analysis

Bacterial diversity was assessed using the 16S rRNA gene. Metagenomic reads that matched the 16S rRNA gene were retrieved by BLASTn comparison of all reads against the Ribosomal Database Project (RDP) (release 10) sequences and accepting matches with an expectation value (E-value) less than 1e-03 and alignment length greater than 200 bp. 16S rRNA sequences were assigned to a taxonomic rank using the RDP classifier (Wang 2007) accepting assignments with bootstrap value ≥ 75 %. Below this cut-off, reads were labeled with their closest RDP match pre-fixed with ‘unclassified’. This prevented low confidence matches contributing to counts of high-confidence phylogenetic groups while avoiding grouping all the unclassified taxa together. Reads that did not align to the 16S rRNA gene using RDP classifier or aligned with low confidence (<75 % bootstrap value on the kingdom level) and were removed from the analysis. To allow comparison of the relative abundance of taxa between sample filters, the number of 16S matches per sample filter was normalized to the average number of reads (403,577) obtained for each sample filter.

Statistical analysis on the relative bacterial abundances was performed using the PRIMER v6 package (Clarke 1993, Clarke 2006). The 16S rRNA gene counts of each sample filter were aggregated to a genus level and square root transformed to reduce the contribution of highly abundant taxa in subsequent analyses. 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 **M**ulti**d**imensional **S**caling (MDS) routines in the PRIMER v6 statistical package. 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. 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. Since many abiotic values varied linearly, for example, depth and pressure, only the following variables were considered: salinity, temperature, turbidity, dissolved oxygen saturation and pH. 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).

Reads with an RDP classification with bootstrap value below 75 % at the phylum level were flagged for manual phylogenetic classification as they likely belong to novel lineages. Those reads were aligned to the 16S gene using the NAST aligner (\*ref), their 7,682 bp alignments were imported into the Greengenes ARB database (release date\*). The phylogenetic affiliation determined by its position within the 16S phylogenetic tree created in ARB by neighbor-joining (Ludwig 2004).

### Eucaryal diversity

Reads that encode the 18S rRNA gene were retrieved by performing a BLASTn search of all reads against the SILVA (release 98) database and accepting those reads below and expectation value of 1e-05 and with an alignment length greater than 300 bp. Longer fragments were desired as taxonomic resolution for eucarya is limited from using the 18S ribosomal subunit alone (\*reference). Since SILVA contains bacteria, archaeal and eucaryal ribosomal subunit sequences, reads that matched non-eucaryal sequences were removed. This was achieved by submitting the list of the accession numbers of the top BLAST matches to Batch Entrez (<http://www.ncbi.nlm.nih.gov/sites/batchentrez>) and using the Taxonomic Groups filter to identify non-eucaryal sequences and subsequently removing those reads. The accession descriptions were used to further remove reads which matched plastid and non-18S ribosomal subunit sequences from the eucaryal sequences. The taxonomic origin of reads that matched the 18S rRNA gene was determined by BLASTn comparison against a custom ARB database containing cultured eucaryal representatives, environmental clone sequences (primarily of marine) and Antarctic sequences. Reads were assigned to their best BLAST match if they had an e-value < 1e-05 and ≥97 % identity. Those below 97 % identity were manually classified by fitting those sequences to the custom ARB tree.

# Results

## Physical and chemical properties of Organic Lake

**Is the lake physically stratified?** 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 (maximum: -8 ºC, minimum: -13.5 ºC) and the estimated maximum practical salinity was 188 (figure:profiles). Between the surface and 1.5 m there was a rapid change in temperature, dissolved oxygen, turbidity and salinity. From 1.5 m to approximately 5 m the physical and chemical parameters remained fairly constant indicating mixing in this zone. Below 5 m salinity and temperature increased rapidly while pH and dissolved oxygen decreased (\*check DO error of probe) (figure: profiles). Dissolved oxygen is so low that the lake could be considered hypoxic\*. These data are consistent with layer of fresher melt water directly below the 0.83 m thick ice penetrating to 1.5m, mixed waters at 1.5 to 5 m and an unmixed bottom waters below 5 m. We designated 0 - 5 m to be the mixolimnion and below 5 m to be the monimolimnion as 5 m appears to be the deepest extent of mixing (\*density converter).

In the monimolimnion all parameters measured except for turbidity vary continuously with depth indicating limited mixing within this zone. This is evident through PCA analysis (\*redo PCA with nutrient data. figure: PCA physical) which shows the samples in the mixolimnion grouped together while the other samples separated along the PC1 axis with depth. PC1 described 75.1% of the variation with a roughly equal contribution from DO, salinity, conductivity, and pH. A sharp spike in turbidity occurred at 5.7 m, after which it dropped 3 fold before increasing again near the lake bottom (figure: profiles). The turbidity maxima at 5.7 m and 6.75 m are likely due to particles or cells settling at the lower limits of the mixolimnion and monilimnion respectively, or zones high production or a combination of the two. (\*turbidity could also vary with some unmeasured environmental factor or be due to some sort of Goldilocks effect). (\*see if cell counts match turbidity).

**Nutrient analyses, Are the nutrient profiles normal? Are any nutrients limiting?** All nutrient concentrations increased with depth (figure: profiles). However, ammonia, dissolved reactive phosphorus (DRP), total carbon, total nitrogen, total phosphorus and total sulfur all reached maximum concentrations at 6.5 m, corresponding to the turbidity minimum. Only nitrate and nitrite were at minimal concentrations at this depth. This implies either nutrient utilization decreases or production increases with depth. If turbidity is assumed to correspond to cell density and cell density to productivity, the inverse relationship of turbidity and nutrient concentration at 6.5 m indicates a lack of utilization is leading to nutrient accumulation at depth.

Salinity is purportedly too high for sulfate reducing bacteria to occur (\*reference? Also too saline for GSB according to Burton) so sulfate is not reduced to hydrogen sulfide (\*check other lakes such as Pendant, Burton? and Bonney that are high salinity and dysoxic for the presence of sulfate reducers. There may be an excess of sulfate compared to similar systems. Bear in mind, it is possible that sulfate may be reduced to sulfite and not completely on to sulfide.) If complete reduction of sulfate to sulfide did occur, any sulfide would be oxidized to sulfate due to the presence of oxygen throughout the lake. (\*Check out if the absolute amount is in excess).

C:N ratio in the profile is on average 10:1 including cells and 11:1 dissolved. This is much higher than the ratio defined in the ocean (6.6:1) indicating total carbon is in excess. The N:P ratio is average 4.2:1 (in cells) and 7.75:1 (dissolved) which is much lower than the Redfield ratio in the ocean (16:1) showing nitrogen is limiting throughout the water column of Organic Lake. (\*Check out the ratios for other Antarctic Lakes) The C:N:P ratio in the mixolimnion is approximately 60:5:1 and this decreases down the watercolumn to 22:3: 1 at 6.5 m (table: sample enviroparams). This shows that while carbon decreases at depth, it is never limiting while nitrogen is limiting throughout the water column and moreso at depth. This could be due to the lower amount of oxygen available which leads to nitrate reduction by *Marinobacter*. (Ward 1997) (\*Look for nitrate reductase, assimilatory NAS, respiratory NAR and periplasmic Nap? What about fixation?)

## Bacterial diversity and distribution

All 16S rRNA gene sequences, except for a single Euryarchaeal read, were classified as bacterial or chloroplast indicating Archaea are near-absent and Bacteria are numerically dominant in Organic Lake.

A total of X 16S rRNA gene sequences were retrieved. What percentage of them could be assigned?

Direct microscopic counts of cells in the size range typical of bacteria or archaea (Figure 3) were taken to be bacteria. (\*Diversity indices, species richness etc.)

How many 16S reads were there? What percentage of reads could be classified to a phylum level?

**The following classes were present in all samples, bacteria present in all samples might perform essential functions** (Figures 4 and 5): Flavobacteria, Gammaproteobacteria, Alphaproteobacteria, Deltaproteobacteria, Actinobacteria, Sphingobacteria, candidate division TM7, candidate division OD1, Opitutae, and chloroplasts.

Flavobacteria and Gammaproteobacteria were found in all samples and together were the most abundant bacterial classes. Alphaproteobacteria were also fairly abundant, especially in the 3.0 µm fraction. Flavobacteria consisted of *Psychroflexus*, \*more? Gammaproteobacteria: *Marinobacter*. Alpha-proteobacteria\*, Deltaproteobacteria.

Three other classes, Actinobacteria, Sphingobacteria and candidate division TM7, OD1, Opitutae and chloroplasts were present in all samples in low abundance but not in all size fractions. Actinobacteria mainly consisted of unclassified genera belonging to the family Microbacteriaceae. Verrucomicrobia principally consisted of *Coraliomargarita* (14) and other lower abundance genera.

These classes of bacteria likely are part of a stable summer consortia as they were present from 2006 until 2008. \*They are mainly heterotrophic?

Unclassified bacteria: unclassified bacteria Bacteroidetes, unclassified bacteria Chlamydiae, unclassified bacteria Firmicutes, unclassified bacteria Fusobacteria, unclassified bacteria Lentisphaerae, unclassified bacteria OD1, unclassified bacteria Planctomycetes, unclassified bacteria Proteobacteria and unclassified bacteria Spirochaetes.

(\*look up what these genera do, what metabolic class? Linked to sulphur? Linked to ice and algae?. Think of a better way to show these data).

**Does the bacterial composition vary with depth?** MDS (Figure 5) and CLUSTER (Figure S2) (\*Find better representation eg hierarchical clustering?) analysis of the bacterial community composition down the water column of Organic Lake showed samples separated according to both size fraction and depth. There was a broad division between the larger size fractions and the 0.1 µm samples. There was more separation between the 0.1 µm samples which were divided into one cluster comprising the surface and interface samples (GS374-GS376) clearly separated from the two bottom samples (GS377 and GS378). The 0.8 µm samples were similarly separated according to depth but with less difference between the two groups. All the 3.0 µm samples clustered together except the 6.5 m sample (GS377-3.0). This indicates the 0.1 and the 0.8 µm bacterial populations are strongly stratified vertically while the 3.0 µm community is much less so.

**Is there a relationship between community assemblage patterns and environmental parameters measured?:** Bacterial assemblages were weakly, but significantly correlated (0.374 correlation, 1.3% significance) with turbidity, temperature, pH and dissolved oxygen. This reflects a correlation between based on depth related parameters (temperature, pH and dissolved oxygen) and turbidity. This indicates the bacterial community structure, is to a small extent, determined by these factors.

**Does the bacterial community vary with time?**

The dominance of Bacteroidetes compared to Proteobacteria varied with the time of sampling. Bacteroidetes were clearly dominant in the December 2006 samples, particularly in the 0.8 µm size fraction. Proteobacteria were consistently the most abundant phyla in the November 2008 profile. In the December 2008 sample both Bacteroidetes and Proteobacteria were dominant with the relative abundance of one over the other varying with size fraction. The majority of Proteobacteria and Bacteriodetes consisted of the genera *Marinobacter* and *Psychroflexus* (\*previously Halomonas?) respectively (\*figure). Although, other Flavobacteriaceae (\*such as x, y and z) were also present. While within the proteobacteria, other gammaproteobacteria including *Psychromonas* and alphaproteobacteria such as *Roseovarius* were also fairly abundant.

*Psychroflexus* appears negatively correlated with ice and *Marinobacter* positively correlated. (\*test? Flavobacteria are epiphytic to algal species in sea-ice, so can those be correlated? Marinobacter may be a Gymnodinium symbiont, Melt water populations?)

Firmicutes, Planctomycetes and Spirochaetes were present only in 2008 samples at low abundances. Firmicutes were mainly unclassified genera, the most abundant of which is related to *Erysipelotrichi*. Planctomycetes consisted of only one genus related to *Schlesneria*. Similarly, the only genus of Spirochaetes was *Spirochaeta*.

## Eucaryal diversity and distribution

A total of 655 reads that matched the 18S rRNA gene were retrieved from all the Organic Lake metagenomes. 583 (89%) of those reads matched with ≥97 % identity to isolates, environmental clones or to 18S rRNA gene sequences from Ace and Organic Lake previously phylogenetically analysed (Yau 2011, Lauro 2010). The remaining reads matched to known sequences with ≥90 % identity except for three reads (gs375-0.1 83% NaeItali, gs379-0.1:88,89% MsgBrev3). Those reads were manually classified.

Eucarya were of the superkingdoms: Chromalveolata (stramenopiles, alveolata), Archaeplastida (Chloroplastida, Opistokonta fungi, choanomonada).

**What sort of euks are in Organic Lake?**

**Does the eukaryotic composition vary with depth?**

**Is the eukaryotic composition correlated with any physical or chemical parameters?**

**Does the eukaryotic composition vary with time?**

## Viral diversity and distribution

**How does the Phycodnaviral diversity change over time? Can it be linked to virophage diversity? Can it be linked to eucaryotes?**

**What sort of viruses are in Organic Lake? Determine by GAAS or by viral marker?**

**Does the viral composition vary with depth?**

**Is the viral composition correlated with any physical or chemical parameters?**

**Does the viral composition vary with time?**

## Read-based analysis of diversity vs molecular markers

**How does read-based analysis compare to 16S/18S analysis in what is there? In terms of the relative abundances? and in terms of statistical differences?**

## Whole system functions

**Do any taxa co-vary over time or depth? Eg. Virophage and phycodnavirus, phycodnavirus and eukaryotes**

**What can we piece together about nutrient cycles in Organic Lake from the taxa we know to be present?**

**What can we piece together about nutrient cycles in Organic Lake from the functional genes?**

**What can we piece together about nutrient cycles from the Proteome?**

## 

# Discussion

Physicochemistry of Organic Lake:

The mixolimnion is thermally and chemically stratified showing two zones: one just beneath the ice to 1.5 m and the second from 1.5 m to 5 m. This is typical of meromictic Antarctic lakes in the early summer when mixing due to ice-formation and convection ceases in the mixolimnion (Gibson 1996).

The physiochemistry of Organic Lake has varied greatly over last 30 years during which it has been monitored. These changes, particularly in temperature, salinity and water column structure, have been related to variations in water level (Gibson 1995, Gibson 1996). When the limnology of Organic Lake first described 1978 and 1984 (Franzmann 1987), it was considered meromictic due to the stable temperatures (-5 to -6 ºC), increased density and anoxia of water below 5 m. A thermocline and oxycline occurred at approximately 3.5 m. During this time, the water level rose from 6.5 m to 7.5 m and density at 3 m decreased, likely due to dilution of salts (Franzmann 1987)(Gibson 1995). Between 1989 and 1994 the water level dropped 0.81 m, temperatures fell (minimum annual temperature of -16.5 ºC in the mixolimnion and -8 ºC in the monimolimnion) mixolimnion penetrated slightly deeper than 4 m (Figure X)(Gibson 1996). \*Other bathymetry files that you can’t open. \*density? Temperature and salinity profiles from this study are similar to those in the 1990’s so it is possible Organic Lake has remained somewhat stable since that time. However, there is no data available between 1997 and this study to definitively confirm this. Franzmann *et al.* (1987) were not able to successfully isolate any anaerobic bacteria which may be due to periodic oxygenation or microaerophilic nature of the bottom waters preventing anaerobes from establishing permanently.

16S diversity: The bacterial population in the water column of Organic Lake differs according to seasonal changes more than by vertical stratification. Variation in the bacterial population down the depth profile mainly occurs in the 0.1um size fraction while the composition of the larger size fractions remained fairly homogeneous.

18S diversity:

## References

Bird M.I., Chiva A.R., Radnell C.J., Burton H.R. (1991) Sedimentaological and stable-isotope evolution of lakes in the Vestfold Hills, Antarctica. *Palaeogeography, Palaeoclimatology, Palaeoecology* **84**: 109-130

Clarke K.R. (1993) Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* **18**, pp. 117-43.

Clarke K.R., Gorley R.N. (2006) PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth

DeSantis T.Z., *et al.* (2006) NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Research* **34**:W394-9

Franzmann P. D., *et al.* (1987) Limnology of Organic Lake, Antarctica, a meromictic lake that contains high concentrations of dimethyl sulfide. *Australian Journal of Marine and Freshwater Research* **38**:409-17

Fofonoff N.P. and Millard R.C. Jr. (1983) Algorithms for computation of fundamental properties of seawater. *UNESCO Technical Papers in Marine Science* **44**: 53

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

Gibson J.A.E., Burton H.R., Gallagher J.B. (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 J.A.E. *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**:73-8

Gibson J.A.E. (1999) The meromictic lakes and stratified marine basins of the Vestfold Hills, East Antarctica. *Antarctic Science* **11**(2):175-92

Lauro F.M., *et al.* (2010) An integrative study of a meromictic lake ecosystem in Antarctica. *Journal of the International Society for Microbial Ecology* **5**(5):879-95

Ludwig W., *et al.* (2004) ARB: a software environment for sequence data. *Nucleic Acids Research* **32**(4):1363-71

Ng C., *et al.* (2010) Metaproteogenomic analysis of a dominant green sulfur bacterium from Ace Lake, Antarctica. *Journal of the International Society for Microbial Ecology* **4**(8):1002-19

Wang Q., *et al.* (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into new bacterial taxonomy. *Applied and Environmental Microbiology* **73**(16):5261-7

Yau S., *et al.* (2011) Virophage control of antarctic algal host-virus dynamics. *Proceedings of the National Academy of Sciences* **108**(15): 6163-8

## Extra 16S stuff

**All sample MDS:** Patterns in the bacterial composition of all Organic Lake samples filters were visualized using MDS and CLUSTER analyses. A broad division was evident between 0.8 and 3.0 µm bacterial compositions and the 0.1 µm population. The 0.1 µm samples were the most divergent from one another with the December 2006 surface water sample (GS233-0.1 µm) clearly the most divergent. The sample taken in December 2008, although taken at approximately the same time of year as the 2006 sample, tended to cluster more closely with the profile samples (GS374-GS379). Differences in composition could be seen according to depth with the deeper profile samples (GS377-0.1 and GS378-0.8) grouping together but these spatial differences were much smaller. Not much variation within the 0.8 and 3.0 µm cluster was evident except that the 2006 sample (GS233-0.8) was again the most divergent and the GS377-0.8 sample also separated.

**GS233:** Within GS233, community composition on the three size fractions differed taxonomically and in relative abundance. 0.1µm had the lowest hit rate, followed by the 3.0 µm fraction and the 0.8 µm fraction although the 3.0 µm filter was more species rich\*. The 0.1 µm fraction was dominated by candidate division OD1 (\*what?), 0.8 µm by Flavobacteriaceae and the 3.0 µm by Flavobacteriaceae and Rhodobacteraceae. In 0.8 and 3.0 µm most of the Flavobacteriaceae was classified to the genus level as *Psychroflexus* but other (\*amount) of cryptic genera were present indication heterogeneity in the population*.* The abundance of *Psychroflexus* in 0.8 µm was the highest of all the samples. The 0.1 µm fraction also had a relatively high abundance of Flavobacteriaceae, but mainly of the genus *Persicivirga*. The presence of *Persicivirga* was unique this sample. Rhodobacteraceae were of diverse genera, those that could be classified included *Roseovarius*, *Loktanella* and *Sulfitobacter* but again, there were several cryptic genera. *Glaciecola* is characteristic of this membrane size.

**GS374:** In contrast, the surface sample taken during ice cover GS374 was dominated by *Marinobacter, Psychroflexus* (and relatives) and *Roseovarius*. The dominance of these clades varied with filter size with GS374-0.1 and and GS374-0.8 clearly dominated by *Marinobacter* while the 3.0 um filter had a larger proportion of *Psychroflexus* and *Roseovarius*, although *Marinobacter* was still quite abundant.

**GS375:** The dominant genera on GS375 followed a similar pattern to GS374 with the exception of *Roseovarius* being slightly more abundant in the 3.0 um fraction than *Psychroflexus*.

At the bottom of the profile GS378, the dominant 16S reads on the 0.1um fraction are now candidate division OD1. The larger size fractions are again dominated by *Marinobacter* and *Psychroflexus.*

*Psychroflexus* appears to dominate in the samples that are partially ice-free, or totally ice-free. It is by far the dominant bacterium in GS233 and GS379. It is mainly found on the larger size fractions. When it is dominant, it is primarily found on the 0.8 um fraction. *Marinobacter* appears to vary inversely with *Psychroflexus*. It is found to dominate the samples from the profile and are always more numerous on the 0.8 um filter. Chloroplasts? *Psychromonas* occurs in the deeper samples and reach larger populations in the partially ice free resample. Not surprisingly *Bacillariophyta* chloroplasts are mainly associated with surface samples \*up to 3.5 m and on the bottom, probably due to sedimentation. In the surface, *Bacillariophyta* could be inversely proportional to the *Chlorophyta*.

*Roseovarius* is consistently more abundant on the 3.0 um fraction of a sample. Its abundance between samples doesn’t appear to vary much except that it is most dominant in the profile. Chlorophyta seems to vary with depth and doesn’t appear to be the most dominant in the bottom, or when the lake is completely thawed.

**Is total bacterial abundance related to any physical parameter in the profile?** The relative abundance of bacteria estimated from 16S frequencies in the metagenome showed counts from 0.1 µm size fractions were consistently lower than the 0.8 and the 3.0 µm size fractions (Figure 4, Figure S3). 16S counts from the larger size fractions were similar (\*test) to one another except in the December 2006 samples where there the highest counts were obtained from the 0.8 µm size fraction. Microscopic examination revealed the majority of bacteria were filamentous or chain forming (Figure 3), with an average cell size (\*) and were indeed be captured on the larger size fractions. This is consistent with morphologies of bacteria isolated previously from the lake (\*ref). The 16S counts also varied with turbidity in a size dependent fashion (Figure S3). Bacterial abundances on the 0.1 µm fraction varied inversely with turbidity. In the mixolimnion the bacterial abundances on the 0.8 µm and 3.0 µm size fractions varied directly with turbidity, although the 0.8 µm counts peaked at 4 m rather than at 5.6 m while in the monimolimnion only abundances in the 0.8 µm fraction varied directly (Figure S2). This suggests bacteria on the 0.8 µm and 3.0 µm size fractions contribute to turbidity in the mixolimnion but in the monimolimnion the 0.8 µm sized bacteria are the main contributors. Assuming that turbidity is directly related to bacterial density, this suggests bacteria on the 0.1 µm filters do not contribute significantly to cell density while the larger size fractions do. (\*test if cell counts varies with any of the envirodata. This may be related to Organic Lake being relatively copiotrophic? cells are motile and actively search nutrient particles and they have a cell size that varies with food availability, feast famine strategy. Chemical data to show it is copiotrophic?).