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

Figure 2. Diversity of (**A**) Bacteria and (**B**) Eucarya from each size fraction (0.1, 0.8 and 3.0 µm) at each sample depth (1.7, 4.2, 5.7, 6.5 and 6.7 m) of Organic Lake aggregated at a class level. The x-axis shows normalized counts of SSU sequences. Taxa that belong to the same rank are shown grouped with a square bracket in the legend. Abundant taxa are labeled in the plot with a number that corresponds to the numbered boxes in the legend. (**C**) Composition of abundant bacterial classes.

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

Figure 4. Genetic potential for (**A**) carbon (**B**) nitrogen (**C**) and sulfur cycling in Organic Lake. The diagrams on the left hand side are simplified biogeochemical pathways. Black arrows represent pathways for which marker genes were detected where arrow width is proportional to the normalized counts for those genes in all samples (log10 scale). Black arrows that start with an open square are pathways that may increase flux in the indicated direction. Dashed blue arrows are inferred pathways. Dashed black arrows marked with a red cross are pathways for which the marker genes and taxa known to mediate the process were not detected. The plots on the right show genetic potential for biogeochemical pathways along the depth profile for each size fraction. The y-axis shows sample depths (m) and the x-axis shows normalized counts of marker genes where counts from the 0.1, 0.8 and 3.0 µm size fractions are shown as blue, red and green respectively. SCFA, short chain fatty acids; DNRA, dissimilatory nitrate reduction to ammonia; anammox, anaerobic ammonia oxidation; MT, methanethiol; DMSP, dimethylsulfoniopropionate; DMS, dimethylsulfide; DMSO dimethylsulfoxide; ASR, assimilatory sulfate reduction; DSR, dissimilatory sulfate redution.

Figure S1. Map of the Vestfold Hills showing Organic Lake (circled in red), fjords, bays and meromictic lakes (numbered). Inset is the position of the Vestfold Hills relative to Australia and other Antarctic coastal oases. Meromictic lakes, black fill; seasonally isolated marine basins and other lakes, lined fill; the Southern Ocean, gray fill; continental ice sheet; stippled. The lakes are: 1. unnamed lake 2; 2.Organic Lake; 3. Pendant Lake; 4. Glider lake; 5. Ace Lake; 6. unnamed lake 1; 7. Williams Lake; 8. Abraxas Lake; 9. Johnstone Lake; 10. Ekho Lake; 11. Lake Farrell; 12. Sheild Lake; 13. Oval Lake; 14. Ephyra Lake; 15. Scale Lake; 16. Lake Anderson; 17. Oblong Lake; 18. Lake McCallum; 19. Clear Lake; 20. Laternula Lake; 21. South Angle Lake. Map is adapted from Gibson *et al.*, 1999.

Figure S2. Bathymetry of Organic Lake 9 November 2008.

Figure S3. Vertical profiles of physical and chemical parameters of Organic Lake taken *in situ* at the deepest point in the lake on 9 November 2008. σT (1000−density) was calculated from temperature and conductivity.

Figure S4. Epifluorescence microscopy images of Organic Lake microbiota filtered onto 0.01 µm polycarbonate membrane and stained with SYBR Gold. From top to bottom, 1.7, 4.2, 5.7, 6.5 and 6.7 m sample depths.

Figure S5. PCA analysis of physico-chemical parameters and cell/VLP counts of Organic Lake profile. Data points are the sampling depths 1.7, 4.2, 5.7, 6.5 and 6.7 m. The overlaid vector diagram shows the relative contribution of the variables to explaining the difference between samples. PC1 explained 74.3% and PC2 14.7% of the variation between samples. Abbreviations: cond, conductivity; temp, temperature; turb, turbidity.

Figure S6. Barplots showing the frequencies of taxonomic assignments to the KEGG orthologs (KO) that were markers for (**A**) carbon, (**B**) nitrogen and (**C**) sulfur conversions. The x-axis shows normalized counts of the marker genes across all lake samples. Enzyme name and corresponding KO identification number is shown on the left of each plot. See Table S2 for full KOs search list and gene marker gene descriptions.

Figure S7. Phylogenetic tree of rhodopsin homologs including proteorhodopsin, bacteriorhodopsin, actinorhodopsin and xanthorhodopsin. *Halobacterim salinarum* R1 halorhodopsin was used as an out-group. The tree was computed from a 78 amino acid region spanning the motif involved in ‘spectral tuning’ using the neighbor-joining algorithm. Organic Lake sequences from this study are shown in red and marked with an asterisk (\*). Numbers in parentheses are counts of sequences which clustered with the Organic Lake homolog shown in the tree with 90% amino acid identity. Sequences with confirmed activity are shown in bold. Accession numbers from top to bottom are: EAZ99241, EDP63929, EGF32634, ZP\_09955974, AEG32267, EDY76405, EDY88259, YP\_445623, ACN42850, EIC91904, ZP\_02194911, AAZ21446, AAT38609, AEE49633, EAS71907, sequence from John Bowman (personal correspondence), EAQ40507, EAQ40925, EAR12394, EHQ04368, EAZ94876, EIA08356, AEE20201, EEG43331, ZP\_09501337 and YP\_001689404.

Figure S8. Phylogenetic tree of DddD DMSP lyase homologs. *E. coli* carnitine coenzyme A transferase was used as an out-group. *Dinoroseobacteria shibae* DFL 12 and *Ruegeria pomeroyi* DSS-3 homologs are non-functional outgroup (Todd *et al.*, 2011). The tree was computed from a 75 amino acid region within the conserved amino-terminal class III coenzyme A domain (CaiB) using the neighbor-joining algorithm. Organic Lake sequences from this study are shown in red and marked with an asterisk (\*).Numbers in parentheses are counts of sequences which clustered with the Organic Lake homolog shown in the tree with 90% amino acid identity. Sequences with confirmed DMSP lyase activity are shown in bold. Accession numbers from top to bottom are: EBA01716, AEV37420, ACY01992, ADZ91595, EAQ63474, ABR72937, ACV84065, ACY02894, ABI89851, YP\_002822700, EEE36156, ABV95365, AAV94987 and EGB36199.

Figure S9. Phylogenetic tree of DddL DMSP lyase homologs from Organic Lake and public databases. The tree was computed from an 84 amino acid N-terminal region using the neighbor-joining algorithm. Organic Lake sequences from this study are shown in red and marked with an asterisk (\*). Numbers in parentheses are counts of sequences which clustered with the Organic Lake homolog shown in the tree with 90% amino acid identity. Sequences with confirmed DMSP lyase activity are shown in bold. Accession numbers from top to bottom are: EEB86351, ADK55772, EAQ07081, EEE47811, EAV43167, EAU41122, EAQ10619, ABV95046, EAQ04071, ABA77574 and EHJ04839.

Figure S10. Phylogenetic tree of DddP DMSP lyase homologs from Organic Lake and public databases. The tree was computed from a 129 amino acid C-terminal region including the predicted catalytic sites using the neighbor-joining algorithm. Organic Lake sequences from this study are shown in red and marked with an asterisk (\*). Numbers in parentheses are counts of sequences which clustered with the Organic Lake homolog shown in the tree with 90% amino acid identity. Sequences with confirmed DMSP lyase activity are shown in bold. Accession numbers from top to bottom are: ZP\_01755203, YP\_167522, YP\_613011, YP\_682809, EAP77700, ZP\_01741265, ZP\_01036399, ZP\_01881042, ZP\_05063825, AFO91571, YP\_509721, ZP\_01448542, AEQ39103, AEQ39091, XP\_001823911, XP\_389272 and ACF19795.

Figure S11. Phylogenetic tree of DmdA DMSP demethylase homologs from Organic Lake and public databases. The tree was computed from a 128 amino acid region using the neighbor-joining algorithm. Organic Lake sequences from this study are shown in red and marked with an asterisk (\*). Numbers in parentheses are counts of sequences which clustered with the Organic Lake homolog shown in the tree with 90% amino acid identity. Sequences with confirmed DMSP lyase activity are shown in bold. Accession numbers from top to bottom are: EDZ60447, YP\_265671, EDZ61098, EAU51039, YP\_003550401, EDP61332, EAQ26389, ABV94056, AAV94935, AAV95190, EDY79173, EDY89914, EAW42451, AAV94935 and AAV97197.

Table 1. Physico-chemical properties, cell counts and VLP counts of Organic Lake samples. ND, data not determined; SRP, soluble reactive phosphate; TOC, total organic carbon; DOC, dissolved organic carbon; TN, total nitrogen; TDN, total dissolved nitrogen; TP, total phosphorus; TDP, total dissolved phosphorus; TS, total sulfur; TDS, total dissolved sulfur; VLP, virus like particles. One standard deviation shown for cell and VLP counts.

Table 2. Marker genes involved in carbon, nitrogen and sulfur cycling detected in Organic Lake metagenomes and frequently associated taxonomic groups. See Figure S6 for frequencies of all taxonomic groups that matched to KEGG orthologs. *Mollicutes* are likely to actually be the related candidate division RF3.

Table 3. Counts of of homologs for genes per 100 Mbp involved in DMSP catabolism and photoheterotrophy in Organic Lake, Ace Lake and GOS metagenomes. Percentages shown in parentheses are calculated from ratio of each homolog to the single-copy gene *recA*. (\*I’ve shown the percentages for GOS sites as calculated by Howard *et al*., 2008 for the moment).

1

Table S2.Summary of metagenomic data for Organic Lake samples.

Table S3. DMSP lyase and demethylase sequences used in this study as BLAST queries for retrieving homologs in the Organic Lake metagenomes. (%ID) is the minimum amino acid sequence identity for matches to be considered a homolog.

Table S4. Microbial taxa detected in the Organic Lake water column profile by analysis of SSU sequences. SSU sequences were classified to the genus level or to the lowest rank with bootstap confidence >85% (see materials and methods).