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. σT (1000−density) was calculated from temperature and conductivity. (**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). Abbreviations: cond, conductivity; 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 (1.7–6.7m) of Organic Lake. The x-axis shows normalized counts of SSU sequences. Taxa that belong to the same higher rank are shown grouped with a square bracket in the legend. Abundant taxa are numbered in the plot with corresponding numbered boxes in the legend. (**C**) Composition of abundant bacterial classes. SSU sequences were classified to the genus level or to the lowest rank with bootstap confidence >85% (see materials and methods).

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 level of genus. SSU counts were normalized to the average number of reads from each sample and square root transformed. Taxa that comprised <2% of the sample were not included.

Figure 4. Genetic potential for carbon, nitrogen and sulfur cycling in Organic Lake. The number of matches from each sample to KEGG orthologs (Ks) of marker genes involved in C, N and S cycles was normalized across samples (100 000 reads) and plotted. The 3.0, 0.8 and 0.1 µm size fractions are shown as green, red and blue respectively. The sum of matches from the whole lake to each process is shown proportional to the arrow size (log10 scale).

(**A**) Carbon cycle. Aerobic carbon fixation = (K00855 + K01602)/2, aerobic respiration = (K02256 + K02262 + K02274 + K02276)/4, CO oxidation = (K03518 + K03519 + K03520)/3, methane oxidation = K08684, methanogenesis = (K00401 + K00400)/2, fermentation = K00016 + (K00169 + K00170)/2 and anaerobic carbon fixation = (K01648 + K00174 + K00175 + K00244)/4 + (K00194 + K00197)/2.

Carbon cycle marker genes: phosphoribulokinase (K00855), RuBisCO small chain (K01602), cytochrome C oxidase subunit I (*coxI*) (K02256), cytochrome C oxidase subunit III (*coxIII*) (K02262), cytochrome C oxidase subunit I (*coxA*) (K02274), cytochrome C oxidase subunit III (*coxC*) (K02276), carbon monoxide dehydrogenase small subunit (*coxS*)(K03518), carbon monoxide dehydrogenase medium subunit (*coxM*) (K03519), carbon monoxide dehydrogenase large subunit (*coxL*) (K03520), soluble methane monooxygenase (K08684), coenzyme M methyl reductase beta subunit (*mcrB*) (K00401), methyl coenzyme M reductase system, component A2 (K00400), L-lactate dehydrogenase (K00016), pyruvate:ferredoxin oxidoreductase alpha subunit (K00169), pyruvate:ferredoxin oxidoreductase beta subunit (K00170), ATP citrate lyase (K01648), 2-oxoglutarate:ferredoxin oxidoreductase subunit alpha (K00174), 2-oxoglutarate:ferredoxin oxidoreductase subunit beta (K00175), fumarate reductase flavoprotein subunit (K00244), CO dehydrogenase subunit delta (K00194) and CO dehydrogenase subunit gamma (K00197). Note that methane monooxygenase (K08684) orthologs may also hydrolyze other hydrocarbons.

(**B**) Nitrogen cycle. Nitrogen assimilation (K00360 + K00367)/2 + (K01915 + K00265 + K00284)/3, mineralization = K00260 + K00261 + K00262, nitrogen fixation = (K00531 + K02586 + K02588 + K02591)/4, denitrification = (K02305 +K04561+ K00376)/3, ammonification = K05904 + K03385, nitrification (K10944 + K10945 + K10946)/3 and Anammox = K10535.

Nitrogen cycle marker genes: assimilatory nitrate reductase (K00360), assimilatory nitrate reductase (K00367), glutamine synthetase (*glnA*) (K01915), glutamate synthase (NADPH/NADH) large chain (*gltB*) (K00265), glutamate synthase (ferredoxin-dependent) (*gltS*) (K00284), glutamate dehydrogenase (K00260, K00261, K00262), nitrogenase (K00531), nitrogenase molybdenum-iron protein alpha chain (*nifD*) (K02586), nitrogenase iron protein (*nifH*) (K02588), nitrogenase molybdenum-iron protein beta chain (*nifK*) (K02591), nitric-oxide reductase (*norC*) (K02305), nitric-oxide reductase (*norB*) (K04561), nitrous oxide reductase (*nosZ*) (K00376), Cytochrome c nitrite reductase (*nrfA*) (K05904), formate-dependent nitrite reductase periplasmic cytochrome c552 subunit (*nrfA*) (K03385), ammonia monooxygenase subunit A (*amoA*) (K10944), ammonia monooxygenase subunit B (*amoB*) (K10945), ammonia monooxygenase subunit C (*amoC*) (K10946) and hydroxilamine oxidase (*hao*) (K10535).

(**C**) Sulfur cycle. Assimilatory sulfate reduction (K00860 + K00956 + K00957)/3, mineralization = K00456 + K01011 and dissimilatory sulfate reduction/sulfur oxidiation = (K00394 + K00395 + K00396)/3. Dissimilatory sulfate reduction and sulfur oxidation utilize orthologous genes and were distinguished based on the phylogenetic affiliation of the gene match.

Sulfur cycle marker genes: adenylyl sulfate kinase (*cysC*) (K00860), sulfate adenylyltransferase subunit 1 (*cysN*) (K00956), sulfate adenylyltransferase subunit 1 (*cysD*) (K00957), cysteine dioxygenase (K00456), 3-mercaptopyruvate sulfurtransferase (K01011), adenylylsulfate reductase subunit A (*aprA*) (K00394), adenylylsulfate reductase subunit B (*aprB*) (K00395) and sulfite reductase (*dsrA*) (K00396).

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. Phylogenetic tree of dddD DMSP lyase homologs. *E. coli* carnitine coenzyme A transferase was used as an out-group. 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. Bootstrap values are shown at the nodes. Organic Lake sequences from this study are 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: EBA01716.1, AEV37420.1, ACY01992.1, ADZ91595.1, EAQ63474.1, ABR72937.1, ACV84065.1, ACY02894.1, ABI89851.1, YP\_002822700.1, EEE36156.1, ABV95365.1, AAV94987.1 and EGB36199.1.

Figure S7. 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. Bootstrap values are shown at the nodes. Organic Lake sequences from this study are 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: EEB86351.1, ADK55772.1, EAQ07081.1, EEE47811.1, EAV43167.1, EAU41122.1, EAQ10619.1, ABV95046.1, EAQ04071.1, ABA77574.1 and EHJ04839.1.

Figure S8. Phylogenetic tree of rhodopsin homologs including proteorhodopsin, bacteriorhodopsin, actinorhodopsin and xanthorhodopsin. The *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 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 light-dependent growth are shown in bold. Accession numbers from top to bottom are: EAZ99241.1, EDP63929.1, EGF32634.1, ZP\_09955974.1, AEG32267.1, EDY76405.1, EDY88259.1, YP\_445623.1, ACN42850.1, EIC91904.1, AAZ21446.1, AAT38609.1, AEE49633.1, EAS71907.1, John Bowman personal correspondence, EAQ40507.1, EAQ40925.1, EAR12394.1, EHQ04368.1, EAZ94876.1, EIA08356.1, AEE20201.1, EEG43331.1, ZP\_09501337.1 and YP\_001689404.1

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 and 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.

Table S1.Summary of metagenomic data for Organic Lake samples.

Table S2. Full list of KEGG Orthologs (KO) involved in carbon, nitrogen and sulfur conversions searched for in the Organic Lake metagenome. rTCA: tricarboxylic acid cycle, WL: Wood-Ljungdahl pathway, AAnP: aerobic anoxygenic phototrophy, DNRA: dissimilatory nitrate reduction to ammonia, Anammox: anaerobic ammonia oxidation, ASR: assimilatory sulfate reduction, DSR: dissimilatory sulfate reduction, SRB: sulfate-reducing bacteria.

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

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).