# Title

Alexandra M. Linz, Shaomei He, Samuel Schmitz, Sarah Stevens, Rex R. Malmstrom, Robin R. Rohwer, Joshua J. Hamilton, Karthik Anantharaman, Tijana Glavina del Rio, Stefan Bertilsson, Katherine D. McMahon

# Abstract

# Introduction

Why nutrient cycling in lakes is important globally (lakes as integrators of the landscape, lakes as hotspots of carbon cycling, small lakes disproportionately active in carbon cycle, lakes as sentinels of change)

Description of study sites

How can metagenomic/genome analysis shed light on ecosystems?

Review other papers that have used this dataset (and why this paper is novel)

# Results/Discussion

**Overview of Dataset**

Analyzing the genomes of uncultured microbes can provide insight into the potential metabolic functions of those organisms. 205 bacterial metagenome assembled genomes (MAGs) were recovered from a metagenomic time series in Trout Bog and Lake Mendota as described in Bendall, et al (Bendall et al. 2016). These MAGs range in completeness from 50 to 99% complete, and passed quality checks for contamination using CheckM (Parks et al. 2015) (supp table). Of the 205 MAGs, 102 were recovered from Lake Mendota, 31 were recovered from the epilimnion of Trout Bog, and 69 were recovered from the hypolimnion of Trout Bog. Bacteroidetes and Actinobacteria comprised the highest numbers of MAGs in Lake Mendota. In the Trout Bog epilimnion, Betaproteobacteria and Actinobacteria were the mostly frequently observed MAGs, while Betaproteobacteria, Verrucomicrobia, Actinobacteria, and Bacteroidetes comprised most MAGs in the hypolimnion. Several MAGs in the epilimnion and hypolimnion of Trout Bog appeared to be from the same population based on high average nucleotide identities (supp table). Other groups of interest found in this dataset were Planctomycetes, Cyanobacteria, and Tenericutes in Lake Mendota and Elusimicrobia, Saccharibacteria, and Ignavibacteria in Trout Bog. The phylogenetic distribution of MAGs was consistent with the classifications of 16S ribosomal rRNA gene amplicon sequencing results (Figure 1). These results are consistent with other 16S-based studies in these sites (Hall et al. 2017; Linz et al. 2017).

**A close up of a map

Description generated with high confidence**

**[Figure 1B: barchart of community composition by 16S]**

**Figure 1. How representative are the MAGs of the microbial community?** The taxonomic classifications of MAGs (A) reflect the community composition observed via 16S rRNA ribosomal amplicon sequencing (B).

## Photosynthesis and Carbon Fixation

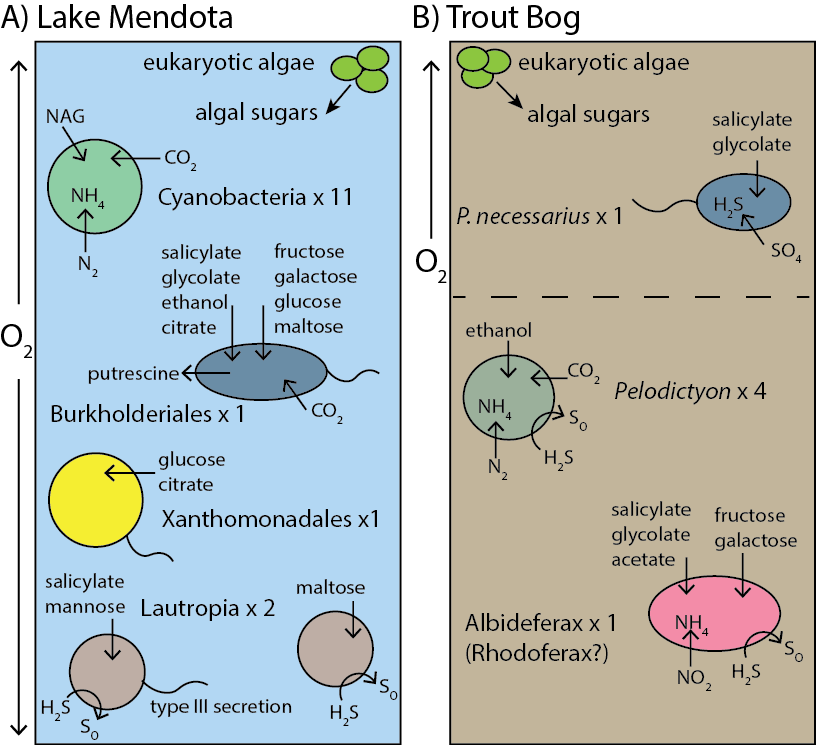
Primary production is a critical component of the carbon cycle in lakes. Therefore, we looked at potential routes of primary production within the microbial community, expecting to find differences between our two ecosystems. In Lake Mendota, MAGs classified as Cyanobacteria comprised the majority of photoautotrophs in the dataset. These populations contained genes encoding enzymes in the Calvin-Benson-Bassham (CBB) pathway, as well as nitrogenase-containing operons and genes encoding a pathway for N-acetyl-glucosamine (NAG) degradation. In Trout Bog, genomes appearing to be from photoautotrophic organisms were classified as *Chlorobium clathratiforme*, a species of Chlorobiales widespread in humic lakes (Karhunen et al. 2013). These green sulfur bacteria also contained the putative genes encoding nitrogen and carbon fixation. However, the Chlorobiales MAGs in Trout Bog contained genes encoding citrate lyase and other key enzymes in the reductive TCA cycle, an alternative carbon fixation method to the CBB pathway commonly found in green sulfur bacteria (Kanao et al. 2002; Tang and Blankenship 2010). Genes encoding this pathway have been identified in other humic lakes (Peura et al. 2015). These MAGs also appear to possess pathways for the degradation of ethanol, citrate, and a variety of sugars. The primary producers Cyanobacteria and Chlorobiales seem to perform similar ecosystem functions, such as nitrogen and carbon fixation, in their respective lakes; however, oxygen availability drives both the type of microbe acting in this role and the pathways that it uses for primary production.

A second type of putative primary producer found in Lake Mendota is a member of Burkholderiales. This population possess genes encoding the CBB pathway and genes annotated as bacteriochlorophyllidae reductase subunit Y (bchlY) and photosynthetic reaction center M (pufM), markers of anoxygenic photosynthesis previously used in analyses of freshwater genomes (Martinez-Garcia et al. 2012). Interestingly, the marker genes for anoxygenic photosynthesis were identified in several other MAGs of Burkholderiales from both lakes (classified as groups such as *Polynucleobacter necessarius*, *Lautropia*, and *Albidoferax*); however, no putative carbon fixation genes were found. A member of the Gammaproteobacteria Xanthomonadales was also predicted to perform anoxygenic photosynthesis without carbon fixation.

While Xanthomonadales is not a known phototroph, Gammaproteobacteria contains the purple sulfur bacteria, anaerobic anoxygenic phototrophs found in aquatic habitats (Tank, Thiel, and Imhoff 2009). Members of Burkholderiales, including *P.necessarius* and *Rhodoferax* (*Albidoferax* is characterized by a lack of phototrophy in its single species (Venkata Ramana and Sasikala 2009); our MAG is therefore more likely a member of the closely related phototrophic genus *Rhodoferax*), are members of the polyphyletic purple nonsulfur bacteria and have been detected in freshwater lakes previously (Mašín et al. 2011). These aerobic anoxygenic photosynthesizers are photoheterotrophs, meaning that they produce energy from photosynthesis rather than fixing carbon like the primary producers. While many questions remain about these unique bacteria, their high levels of abundance and activity in aquatic systems suggest that they are important to carbon cycling (Koblizek 2015). The prevalence of MAGs predicted to be aerobic anoxygenic photosynthesizers in our dataset support this. Furthermore, these MAGs have the heterotrophic potential to degrade a variety of carbon substrates such as salicylate, glycolate, and carbohydrates.

Although eukaryotic genomes were not included in this analysis, eukaryotic algae are known photoautotrophs in both lakes (Descy et al. 2000; Hurley and Armstrong 1990) Algae produce amino acids, carbohydrates, and carboxylic acids that fuel growth of the heterotrophic community (Salcher, Posch, and Pernthaler 2013). Results of incubation assays found that algae in boreal wetlands can release up 38% of their net productivity as DOC, stimulating growth of the heterotrophic bacterial community (Wyatt and Turetsky 2015). Although we do not have data from algal genomes, we observed many MAGs in both Lake Mendota and Trout Bog containing putative pathways for the degradation of carbohydrates such as glucose, galactose, maltose, rhamnose, mannose, and xylose, as well as other low molecular weight compounds such as glycolate and citrate. These compounds are all documented algal exudates in freshwater (Giroldo, Augusto, and Vieira 2005; Juttner and Matuschek 1977).

Another form of harvesting sunlight for energy is the use of light-activated proteins such as rhodopsins. One of the most abundant and ubiquitous freshwater bacteria, the Actinobacteria lineage acI, possess genes encoding putative actinorhodopsins that were found to be highly expressed in Lake Mendota (Hamilton et al. 2017). Additionally, rhodopsins were observed in many phylogenetically diverse MAGs in both Trout Bog and Lake Mendota, and will be the subject of further study (cite Shaomei’s preprint when it comes out).

.

**Figure 2. Photosynthesis in Lake Mendota and Trout Bog.** Both systems contain photoautotrophs and photoheterotrophs, but taxa and pathways vary between lakes. Lake Mendota phototrophs are aerobic and fix carbon via the RuBisCo enzyme, while Trout Bog phototrophs include anaerobes and fix carbon via the reductive TCA cycle. Both lakes host anoxygenic photosynthesis and a variety of photoheterotrophs.

## Degradation of Aquatic and Terrestrial Carbon

One important contrast in freshwater carbon cycling is the degradation of carbon produced in a lake (autocthonous) versus terrestrially-derived carbon (allocthonous). The relative importance of terrestrial vs. aquatic carbon at the ecosystem level in freshwater is a matter of debate in the literature; however, from a microbial standpoint, degradation of allochthonous carbon often requires more specialized enzymes than autochthonous carbon (Brett et al. 2017). Much of the allocthonous carbon is derived from plant biopolymers such as cellulose and lignin or from humic acids, all containing aromatic structures (Hutalle-Schmelzer et al. 2010). However, this high-molecular weight DOC can be degraded into low-molecular weight DOC through photochemical processes, making that carbon available without specialized enzymes (Bertilsson and Tranvik 1998), and autochthonous DOC contains its own difficult biopolymer in the form of chitin (Beier and Bertilsson 2011). We sought to use our MAGs to investigate preferences for autochthonous or allochthonous carbon in populations of bacteria from two lakes with different carbon inputs.

MAGs from both lakes contained genes encoding pathways for the degradation of cellulose and aromatic compounds such as salicylate. These genes are found in MAGs classified as Verrumicrobia, Bacteroidetes, and Burkholderiales in Lake Mendota and in Trout Bog. Additionally, Tenericutes and Cyanobacteria degrade this type of carbon in Lake Mendota, while it is degraded by Gallionella, Solirubrobacterales, and Actinobacteria in Trout Bog.

A type of autochthonous carbon degradation in freshwater is methylotrophy, the degradation of one carbon compounds such as methane, methanol, formaldehyde, or methylamines. One carbon compounds are transformed to formaldehyde or formate and then acetate using either the serine cycle or the RuMP pathway. The resulting acetate can then be used either for energy generation in the TCA cycle, or for biosynthesis via gluconeogenesis; in fact, methylotrophy by definition requires that the methylotroph can grow solely on one carbon compounds. Multiple MAGs classified as well-studied methylotrophs Methylococcales (an aerobic methanotroph) and Methylophilaceae (which likely degrades methanol and methylamines) containing genes for methylotrophic pathways were found in both lakes (Kalyuzhnaya et al. 2012; Salcher et al. 2015) (Figure 3). Consistent with studies of related cultured isolates, MAGs belonging to Methylococcales often contained operons encoding nitrogen fixation (Dedysh, Ricke, and Liesack 2004). Lake Mendota additionally had MAGs containing potential methylotrophs belonging to Planctomyces and Rhodocyclaceae, while additional potential methylotrophs in Trout Bog included Burkholderiales, Rhizobiales, Nitrosomonadales, Geobacteraceae, and Solirubrobacterales.

Methylotrophs such as these in freshwater have been identified as members in syntrophic relationships. For example, Methylococcales was discovered at high abundances in the anoxic regions of a stratified lake by consuming oxygen produced by algae (Milucka et al. 2015), and cooperative behavior between members of Methylophilaceae and Methylococcaceae from Lake Washington, potentially involving denitrification and/or the exchange of methanol, has been observed (Beck et al. 2013). Methylotrophy appears to be an important step in the freshwater carbon cycle in Lake Mendota and Trout Bog, where it may be a link between carbon cycling and other nutrients.

As previously mentioned, algal exudates are a major source of autochthonous carbon in freshwater. Genes encoding the pathways and transporters for degradation of amino acids and carbohydrates were widespread in MAGs from Lake Mendota and Trout Bog, as were genes encoding pathways for the degradation of chitin and its breakdown products, chitobiose and NAG (check chitobiose). These rich sources of carbon and nitrogen may be used for biosynthesis, or they may be used for energy generation, likely depending on the current needs of a bacterium. MAGs classified as Verrucomicrobia and Bacteroidetes possessed genes potentially encoding the degradation of chitin and its derivatives in both lakes. MAGs from Lake Mendota with similar genes included Cyanobacteria, Planctomyces, and Actinobacteria, while Trout Bog MAGs with these potential pathways included Holophagales, Ignavibacteria, and Helicobacterales. Interestingly, a MAG from bacterial predator Bdellovibrionales was recovered from Trout Bog (cite info on bdello).

The balance of autochthonous to allochthonous carbon degradation was similar in both of our study sites, and the relative importance of allochthonous to autochthonous carbon on the ecosystem level is unclear from our dataset. However, other studies show that autochthonous carbon is more likely to be found in higher trophic levels, suggesting a preference for prey that consume carbon produced in situ and a higher importance for autochthonous carbon (Brett et al. 2017; Guillemette, McCallister, and del Giorgio 2015).

A close up of a map

Description generated with high confidence

**Figure 3. Methylotrophy in freshwater.** *Methylococcaceae* and *Methylotenera* are taxa capable of methylotrophy in both Lake Mendota and Trout Bog. Methanol degradation by *Methylotenera* has been extensively studied in freshwater sediments; our genomes show similar pathways. *Methylococcaceae* also shows pathways consistent with cultured relatives.

**Nitrogen Cycling**

While carbon cycling was relatively similar between lakes, the drastically different concentrations of nitrogen in Trout Bog versus Mendota lead us to hypothesize that steps in the water column nitrogen cycle may be altered between these two systems. One key difference was that in Mendota, very few MAGs had genes encoding nitrogen fixation, and they belong mainly to Cyanobacteria. Conversely, many MAGs in Trout Bog contained these genes, and they were in phylogenetically diverse populations. Genes annotated as nitrate and nitrite reductases, key enzymes in denitrification, were found in MAGs from both lakes. However, nitrate reductases were far less common than nitrite reductases in Lake Mendota (19 vs 53, respectively), and found primarily in Cyanobacteria. Genes annotated as ammonia monooxygenase were not found in MAGs from either lake, aside from the ammonia/methane monooxygenases found in MAGs classified as the methanotroph Methylococcales, which are likely not involved in nitrogen cycling. No genes potentially encoding the anammox pathway were identified any of the MAGs. Genes encoding steps in the urea cycle or ammonia assimilation were found in nearly every MAG.

The pattern of which pathways were encoded in which MAGs likely reflects water column nitrogen cycling in Trout Bog and Lake Mendota. The increased numbers of genes and operons encoding nitrogen fixation in Trout Bog is easily explained by the severe nitrogen limitation in this system, where harvesting atmospheric nitrogen would convey a significant competitive advantage. Conversely, Lake Mendota receives high levels of nitrate and ammonia, making nitrogen fixation more energetically expensive than it is worth.

Genes annotated as nitrate and nitrite reductases, both assimilatory and otherwise, were found in MAGs from both lakes, suggesting that these nitrogen compounds could be utilized as terminal electron acceptors or as sources of ammonia in both ecosystems.

Non-proteinogenic amino acids and polyamines have been suggested as a major pool of dissolved organic nitrogen in freshwater. Pathways for biosynthesis of putrescine, spermidine, and canavanine, as well as their corresponding transporters, were widespread in both lakes, supporting the importance of non-proteinogenic amino acids and amino acid derivatives in these systems. Overall….

**Sulfur Cycling**

Sulfur cycling is again relatively similar between Trout Bog and Lake Mendota in broad functions, if not the taxa responsible. Assimilatory sulfates reduction (where sulfates are incorporated into cell components) was more common than dissimilatory sulfate reduction (where sulfate is used as a terminal electron acceptor and sulfide is expelled outside the cell) in both systems. More pathways for sulfide oxidation were found in Trout Bog than in Lake Mendota.

**Table 2. Sulfur cycling in Lake Mendota and Trout Bog.** Sulfur oxidation and reduction are found in both lakes. However, Lake Mendota contains more pathways for reduction than oxidation, while Trout Bog leans towards oxidation of reduced sulfur.

|  |  |  |
| --- | --- | --- |
|  | ME | TB |
| Sulfate reduction (assimilatory) | Planctomyces, Mycobacteraceae, Cyanobacteria, Verrucomicrobia, Sphingobacterales, unclassified, Acidimicrobiales, Rhodocyclaceae, Methylophilales, Methylococcales, Chloroflexi, Sphingomonadales, Chlamydiales, Burkholderiales, Xanthomonadales | Methylococcales, Verrucomicrobia, Burkholderiaceae, Chlorobiales, Methylophilaceae, Acidimicrobiaceae, Actinobacteria, Solibacterales, Helicobacteracaea, Sphingobacterales, Holophagales, Geobacteraceae, Ignavibacteriaceae, Bacteroidales |
| Sulfate reduction (dissimilatory) | None | Gallionellaceae, Desulfobacterales |
| Sulfide oxidation | Burkholderiales, Planctomyces, Rhodocyclaceae, Methylophilales, Acidimicrobia, Actinomycetales | Methylococcales, Burkholderiales, Chlorobi, Methylophilaceae, Actinobacteria, Holophagales, Helicobacteraceae, Sphingobacterales, Gallionellaceae, Solirubrobacterales, Ignavibacteraceae, Bacteroidales, Nitrosomonadales, Gammaproteobacteria |

**Unusual microbes**

Although our primary goal was to use genome content to investigate differences in nutrient cycling between lakes, we recovered the genomes of unusual micro-organisms in this process and report their genome content here. One MAG from Elusimicrobiales was recovered from Trout Bog. While this genome is only 44% complete, we can propose that uses sugars such as maltooligosaccharides, maltose, and arabinogalactan as a carbon source. This population of Elusimicrobiales reduces sulfate via the assimilatory pathway. It also contains one nitrogenase subunit, suggesting that it may be capable of fixing nitrogen.

Thirteen MAGs classified as Planctomycetes were recovered from Lake Mendota. A significant number of metagenomics reads in Lake Mendota were also classified as Planctomycetes, suggesting that this is an abundant group. While some Planctomycetes isolates have been known to perform anammox or degrade one-carbon compounds, no genes encoding the enzymes necessary for this pathway were found in the Lake Mendota MAGs. The gene content of a Planctomyces MAG from Trout Bog, however, suggested methylotrophy. The Plantcomycetes MAGs contained genes encoding glucoside hydrolases and sulfatase enzymes. These pathways could be used to break down complex polysaccharides, such as those produced by Cyanobacteria or eukaryotic algae. This suggests a new potential role for Planctomyces in freshwater as a specialist in polysaccharide degradation.

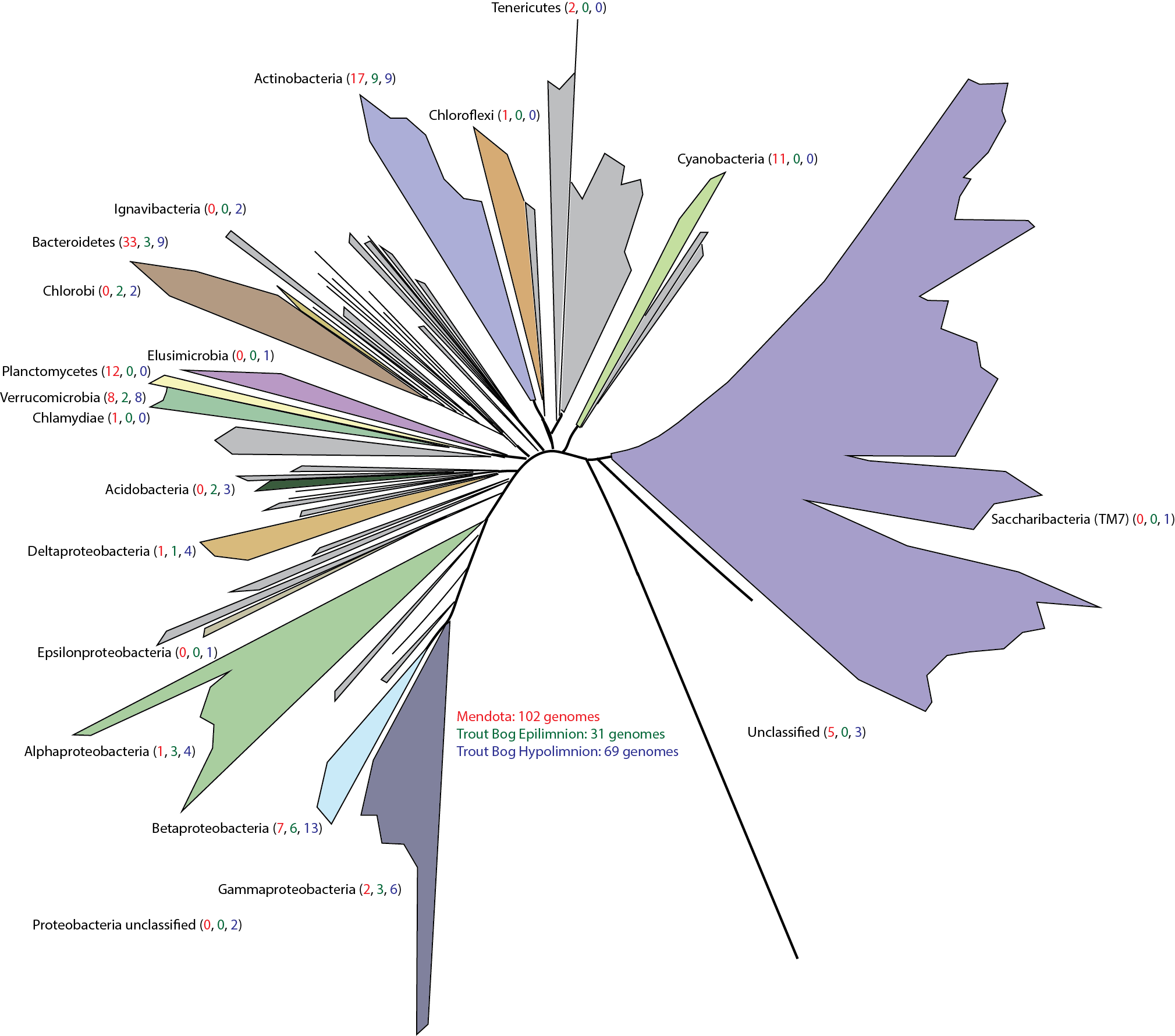
A picture containing text, map

Description generated with very high confidence

**Figure 6B. Proposed functions of Planctomycetes in freshwater.** Observation of genes annotated as glycoside hydrolases and sulfatases suggest a role for Planctomycetes as polysaccharide degraders.

Methods (last)

# Supplemental Figures

**Figure S1.** MAGs were recovered from our freshwater metagenomic time series. The taxonomic classification of each MAG is overlaid on the tree of life adapted from Hug, et al. Colored branches indicate the presence of a MAG from that lineage, and colored numbers after the classification report how many MAGs were recovered from each site in the following order: Lake Mendota (red), Trout Bog epilimnion (green), and Trout Bog hypolimnion (blue).