Introduction

Methods

## Results

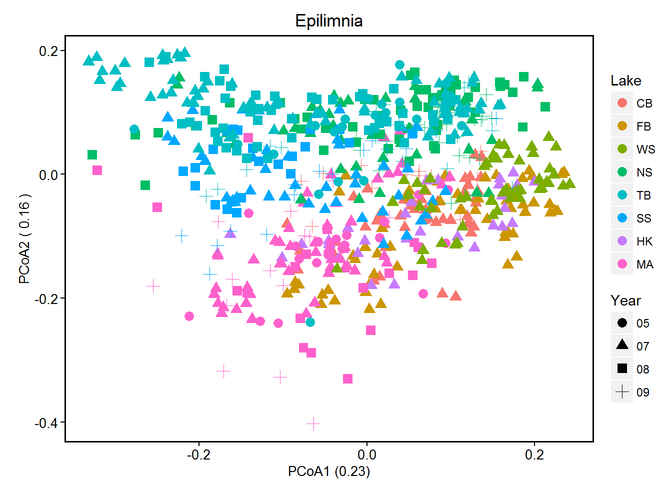
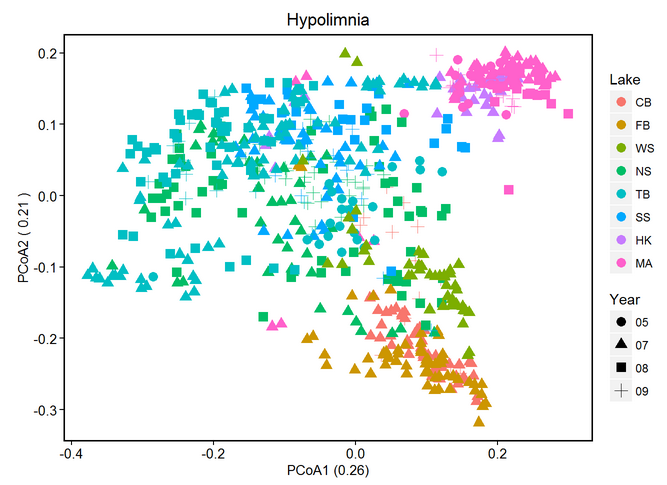
Overview of community composition

A multi-year time series of 16S data from multiple lakes was used to investigate bacterial community dynamics over long time scales. 9,000 OTUs were detected in 1,300 samples. In this time series, *Proteobacteria, Actinobacteria, Bacteroidetes,* and *Verrucomicrobia* were the most abundant phyla. Within these phyla, OTU abundance is highly uneven. For example, much of the abundance of *Proteobacteria* can be attributed to OTUs belonging to the well-known freshwater groups *Polynucleobacter* and *Limnohabitans,* and the freshwater generalist clade acI contributes disproportionately to the observed abundance of *Actinobacteria*. Unevenness is present throughout the dataset, which has a long rare tail of OTUs and trends driven largely by the most abundant OTUs.

Clusters of community composition

When differences in community composition are quantified using weighted UniFrac distance, several trends emerge. Communities from the epilimnion and hypolimnion layers are distinct in all lakes. Epilimnion communities have a higher proportion of OTUs classified using the established freshwater taxonomy, while hypolimnion communities have more OTUs that cannot be classified past the family level. OTUs belonging to candidate phyla are more frequently observed in hypolimnia. These differences likely reflect the bias of freshwater microbial ecology to the surface waters of lakes, and these differences are less pronounced in lakes that mix frequently. Within layers, mixing regime is the next category of community composition. This effect is stronger in hypolimnia, which experience major changes in oxygen content during mixing events. The polymictic lakes Crystal Bog, West Sparkling Bog, and Forestry Bog are more similar to each other than to the other lakes in the dataset, as are the meromictic lakes Mary Lake and Hell’s Kitchen. The dimictic lakes North Sparkling Bog, South Sparkling Bog, and Trout Bog have more variation in community composition, but still have a unique signature of ad dimictic mixing regime. In 2009, Crystal Bog remained stratified all summer (check buoy data), and in this year its hypolimnion community composition is more similar to that found in dimictic lakes than to that found in samples from Crystal Bog in 2007.

The stark contrast between community composition between years in Crystal Bog is mirrored to a lesser extent in the other lakes in this study. Each year of sampling from the same lake is unique and distinct, particularly in hypolimnia. While the change in Crystal Bog can be attributed to mixing frequency, the reasons for this trend in other lakes is less clear. Some possible drivers could include climatic factors, landscape level events, differences in geography, or stochastic community assembly.



**Figure 1. Principal components analysis based on weighted UniFrac distance of samples split by layer.** Clustering by mixing regime and year are observed.

Indicator taxa

To further investigate potential drivers of the differences in community composition between mixing regime, we performed indicator analysis on this dataset (Table 2). This technique identifies taxa that are found significantly more often in one group of samples than another. There is substantial overlap between the indicator taxa of polymictic epilimnia and hypolimnia, which makes sense as this group is defined by high mixing frequency. These taxa include Actinobacteria clade acI-B3, *Limnohabitans* tribe Lhab-A4, *Methylophilales*, and tribe PnecC (most closely related to *Polynucleobacter necessarius*). In dimictic epilimnia, top indicators included *Methylophilales*, acV, acI-B2, Muci (a clade of *Mucilaginibacter*), Rhodo (a member of *Rhodoferax*), and larger taxonomic groups Acidobacteria, and Chlorobia. Dimictic hypolimnia indicator taxa also included Methylophilales, Acidobacteria, and Rhodo, but are additionally indicated by *Holophagaceae*, PnecC, Verrucomicrobial group Ellin515, and members of *Desulfobacterales*. The indicator taxa for meromictic epilimnia include many well-characterized freshwater groups such as Lhab-A, alfI-B2, betIII-A, acI\_A, and acTH2. However, meromictic hypolimnia share none of these taxa as indicators, and instead have candidate phyla OP3 and OP8 as indicators, along with Syntrophobacterales, Verruco-5, Desulfobacteraceae, and archaeal group Parvarcheota. Planctomycetes are indicators of both layers in meromictic lakes. The presence of taxa endemic to each mixing regime likely reflects the environmental filtering taking place in these ecosystems.

Seasonal trends

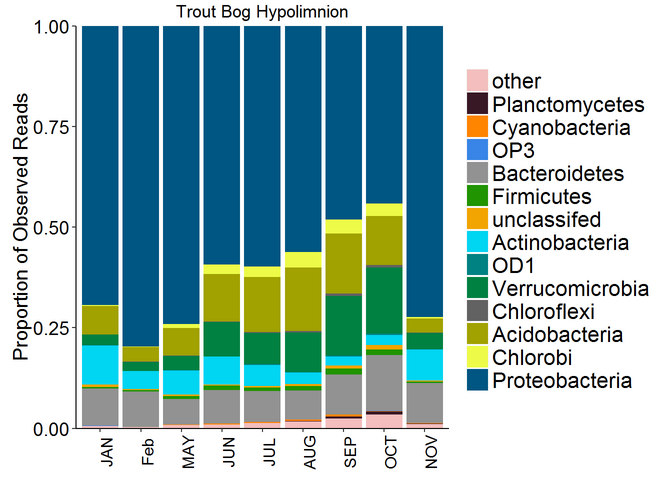
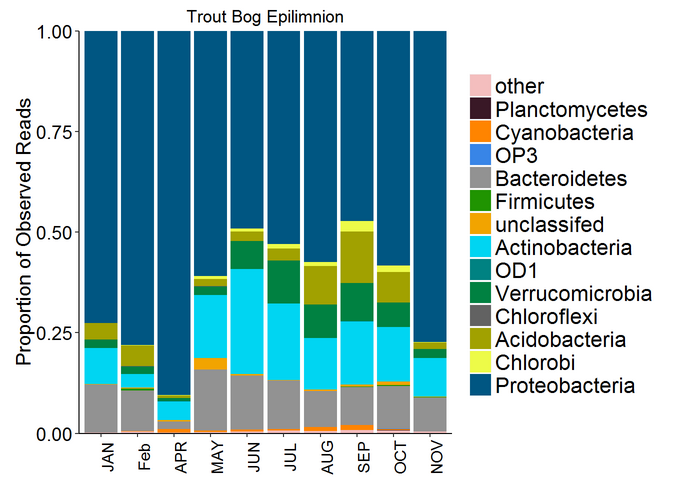
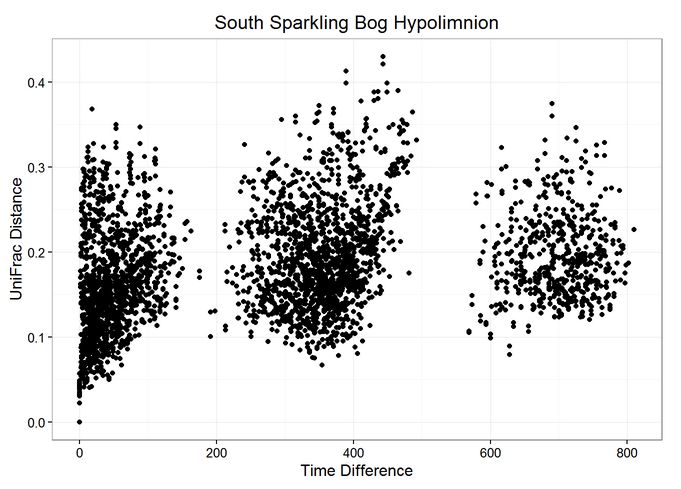
Despite distinct community compositions each year, there are high level seasonal trends. Both biodiversity and evenness increase over time while stratification is in place. Sharp decreases in both of these community properties are observed during the well-sampled Trout Bog 2007 fall mixing event and the North Sparkling Bog artificial mixing event. At the phylum level, the abundance of *Proteobacteria* is particularly high in the few winter samples in this dataset and during mixing events. After spring stratification, the abundance of *Proteobacteria* decreases as the abundances of other phyla increase. Because this is relative abundance data, this trend may reflect the observed increase in biodiversity over time, and the survivability of dominant Proteobacterial groups *Polynucleobacter* and *Limnohabitans* during the ice-on period of the year.

Figure 2. Phylum-level seasonal trends. Polynucleobacter decreases at the expense of everything else.

However, seasonal trends breakdown at lower taxonomic levels. In other aquatic ecosystems, OTUs have similar trends each year, resulting in high similarity between samples taken at the same time each year. This is frequently depicted by plotting a beta diversity metric such as UniFrac distance between samples versus the time between collection of those samples, which results in a sine wave-like pattern with decreasing amplitude. When we perform the same analysis on this dataset, we do not see that pattern. Instead, we see similarity decreasing over time in a logarithmic fashion, so that samples close together in time decrease in similarity quickly and that samples taken a year apart are as dissimilar as samples taken four years apart. The only instance of a sine wave-like pattern is observed in Trout Bog, which has increases in similarity at approximately 1 year and 2 years. This pattern is driven by the higher number of fall, winter, and spring samples collected from Trout Bog than from other lakes. Samples from these times of year are more similar from year to year, while samples collected during the summer stratification period do not show annual trends at the OTU level.



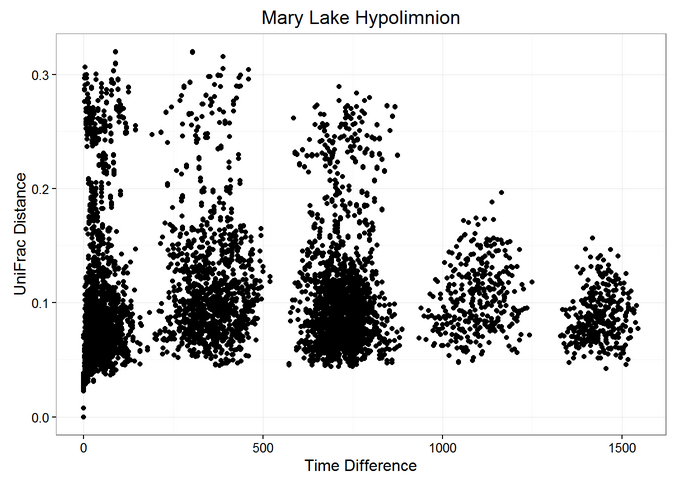
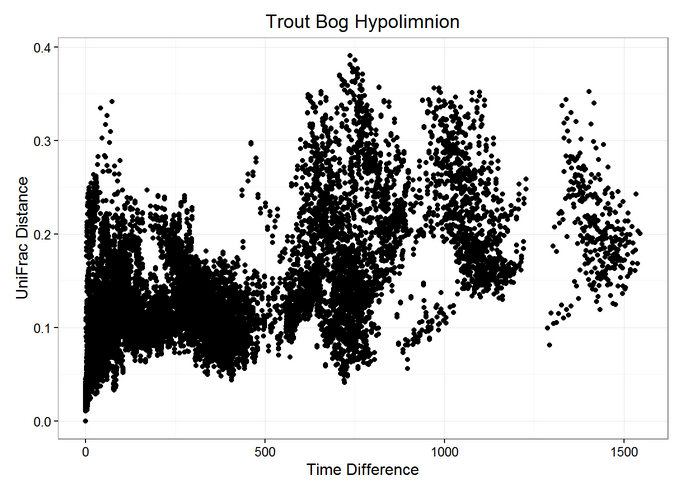
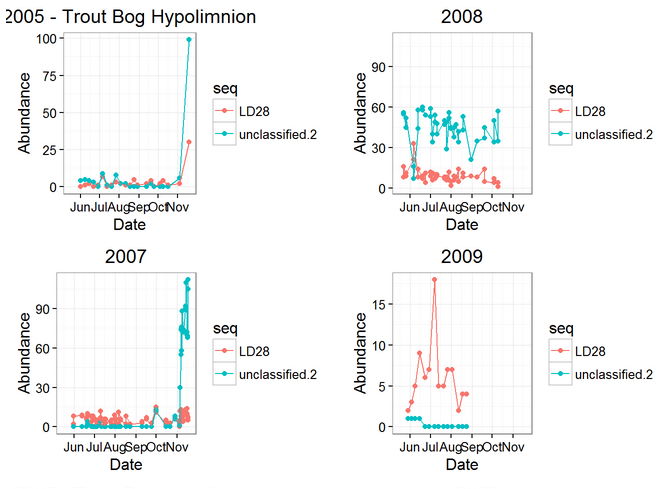


Figure 3. Seasonal trends disappear below the phylum level.

Trends in specific OTUs do not repeat annually

 As suggested by the lack of similarity in community composition between samples taken one year apart, trends in OTU abundance over time do not repeat annually. They do not peak in abundance at consistent times of year or show the same relationship to time each year. OTUs would likely show a consistent response to mixing events because this is such a large disturbance; however, this dataset does not capture a sufficient number of mixing events to demonstrate this. To further complicate OTU abundances, closely related OTUs have different interactions in different years. We have observed the same OTUs to be strongly correlated in some years while strongly anti-correlated in other years. The reasons for these observations are unclear; with only four years of data, differences between years cannot be linked to climatic factors, events in the surrounding landscape, or changes in nutrient levels. It is also possible that biotic drivers and stochastic community assembly are at play in these ecosystems.

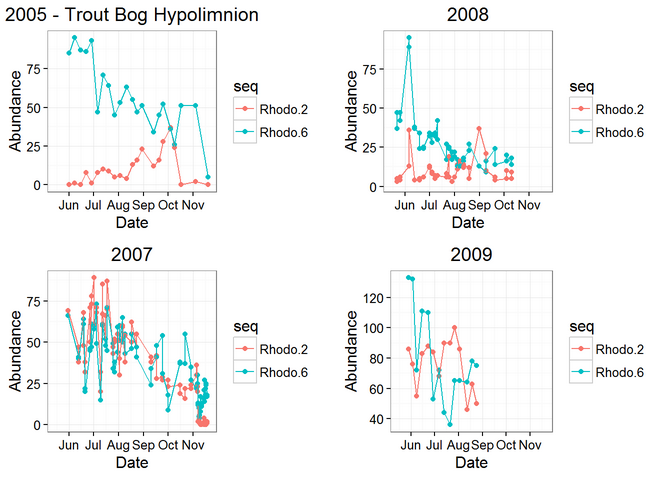


Figure 4. Examples of different trends in abundance with closely related OTUs.

Consistent traits of OTUs

Even though OTUs do not show the same trends each year, they do have traits that are consistent between years and lakes. We quantified mean abundance when present, persistence, and the coefficient of variance for clades classified using the freshwater taxonomy. Using only these well-defined freshwater groups allowed better taxonomic resolution. This analysis showed that low persistence is associated with high variability, and that low variability is associated with high abundance. We rarely observe “bloomers,” situations where a clade has both high abundance and low persistence. Most freshwater clades are highly persistent at low abundances with low variability. Clade gamIII of the Gammaproteobacteria is an exception, with low persistence, low abundance, and high variability. Clades gamI and verI-A occasionally also exhibit this profile. Clades betII and acI are highly abundant and persistent with low variability, consistent with their profile as ubiquitous freshwater generalists. Knowledge of the general manner in which these clades behave can begin to shed light on their lifestyles and the reasons for their observed abundance trends.

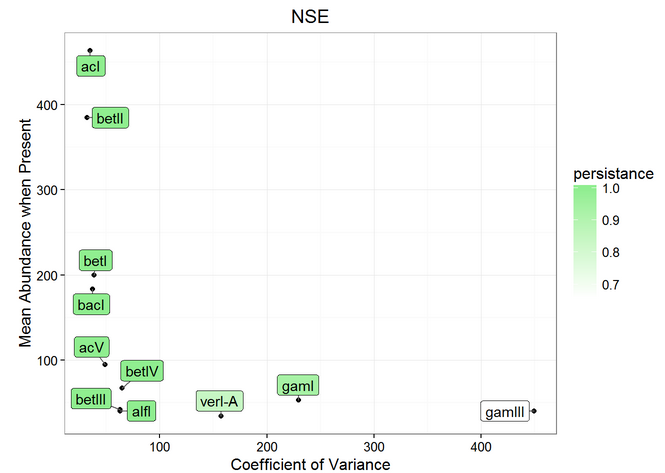
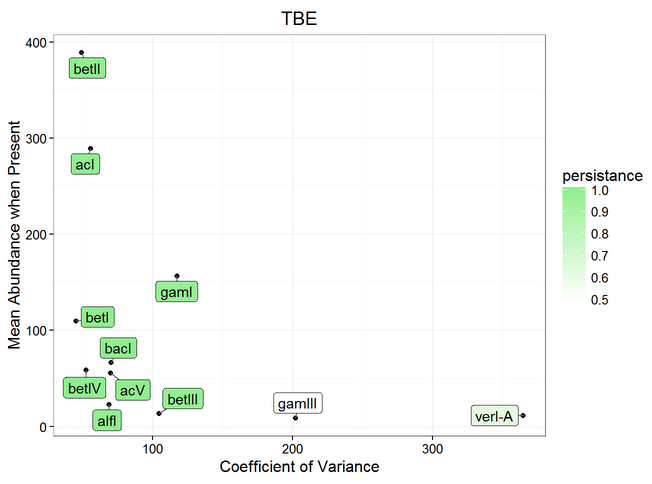


Figure 5. Traits of freshwater clades

Discussion

One of the major findings of this research is that bacterial community composition clusters by lake, layer, mixing regime, and year (Figure 1). This effect is stronger in hypolimnia than in epilimnia, likely due to the higher abundance of freshwater generalists in epilimnia and the increased impact of mixing events on hypolimnia. The separation of samples by site suggests a strong environmental filtering effect. The unique community composition observed in each year within the same lake