# Methods

**Table 1. Location and characteristics of study sites.** The lakes included in this time series are small, humic bog lakes in the boreal region near Minocqua, Wisconsin, USA. They range in depth from 2 to 21.5 meters and encompass a range of water column mixing frequencies (termed regimes). Dimictic lakes mix twice per year, typically in fall and spring, while polymictic lakes can mix more than twice throughout the spring, summer, and fall. Meromictic lakes have no recorded mixing events. pH was measured in 2007, while nutrient data was measured in 2008. When two values are present in a single box, the first represents the epilimnion value and the second represents the hypolimnion value.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Forestry Bog | Crystal Bog | North Sparkling Bog | West Sparkling Bog | Trout Bog | South Sparkling Bog | Hell’s Kitchen | Mary Lake |
| ID | FB | CB | NS | WS | TB | SS | HK | MA |
| Depth *(m)* | 2.0 | 2.5 | 4.5 | 4.6 | 7.0 | 8.0 | 19.3 | 21.5 |
| Surface area *(m2)* | 1300 | 5600 | 4700 | 11900 | 10100 | 4400 | 30000 | 12000 |
| Mixing regime | Polymictic | Polymictic | Dimictic | Polymictic | Dimictic | Dimictic | Meromictic | Meromictic |
| GPS coordinates | 46.047776, -89.651248 | 46.007639, -89.606341 | 46.004819, -89.705214 | 46.004633, -89.709082 | 46.041140, -89.686352 | 46.041140, -89.709082 | 46.186674, -89.702510 | 46.250764, -89.900419 |
| Years sampled | 2007 | 2007, 2009 | 2007, 2008, 2009 | 2007 | 2005, 2007, 2008, 2009 | 2007, 2008, 2009 | 2007 | 2005, 2007, 2008, 2009 |
| pH | 4.97, 4.85 | 4.49, 4.41 | 4.69, 4.80 | 5.22, 5.14 | 4.60, 4.78 | 4.46, 4.94 |  | 5.81, 5.72 |
| Dissolved inorganic carbon (ppm) |  | 0.69, 1.72 | 1.12, 2.31 |  | 1.73, 4.47 | 1.97, 6.42 |  | 5.54, 12.38 |
| Dissolved organic carbon (ppm) |  | 15.47, 13.6 | 10.05, 10.40 |  | 19.87, 20.58 | 12.40, 21.92 |  | 20.63, 67.10 |
| Total nitrogen (ppb) |  | 620.57, 846.00 | 629.09, 809.45 |  | 737.71, 1121.00 | 813.88, 1498 |  | 1332.57, 3652.38 |
| Total phosphorus (ppb) |  | 30.00, 38.86 | 78.00, 135.45 |  | 50.57, 53.25 | 48.63, 69.14 |  | 78.00, 303.50 |
| Total dissolved nitrogen (ppb) |  | 1290.19, 490.13 | 442.39, 586.56 |  | 582.5, 820.21 | 451.63, 1179.21 |  | 1024.5, 3220.14 |
| Total dissolved phosphorus (ppb) |  | 84.25, 14.88 | 70.22, 22.67 |  | 34.5, 31.57 | 16.25, 18.29 |  | 71.13, 228 |

# Results

## Overview of community composition

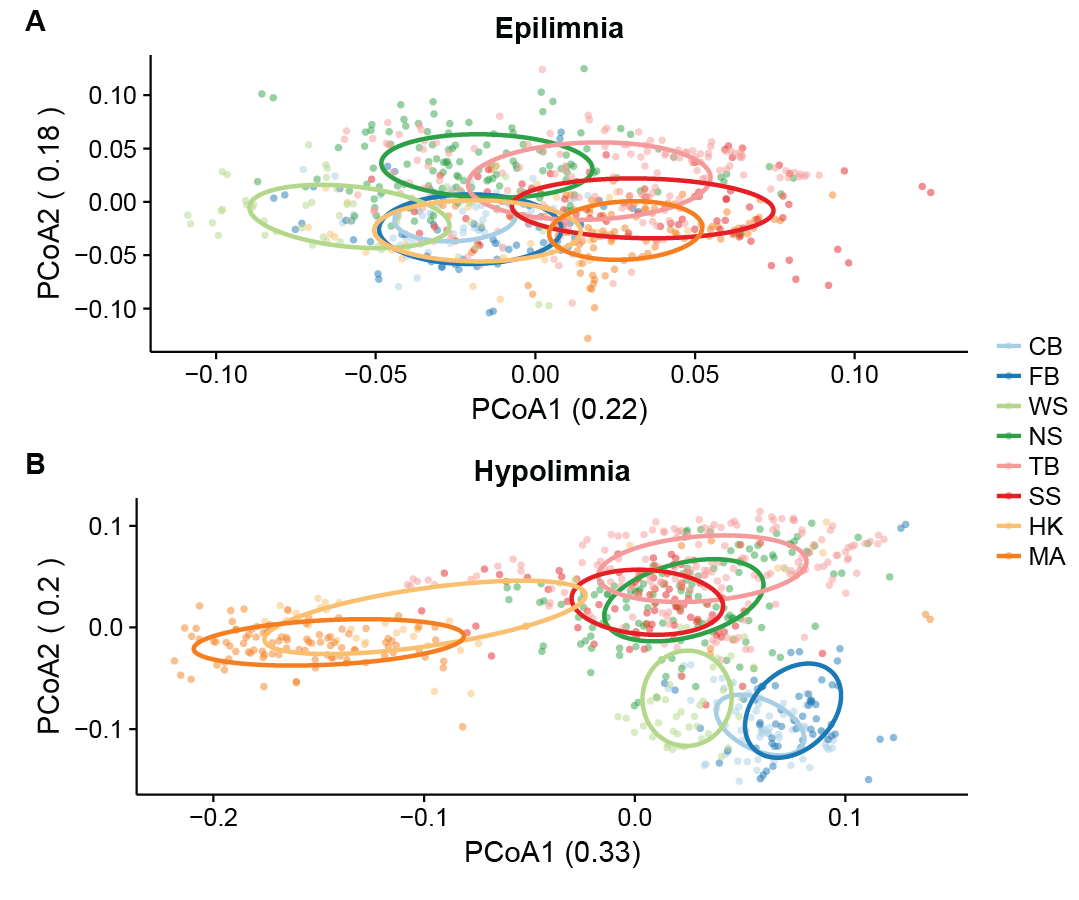
A multi-year, multi-site time series of 16S amplicon data was used to investigate bacterial community composition over long time scales. 8,795 OTUs were detected in 1,387 samples. *Proteobacteria, Actinobacteria, Bacteroidetes,* and *Verrucomicrobia* were the most abundant phyla (Figure S1). Within these 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 clade acI contributes disproportionately to the observed abundance of *Actinobacteria*. Unevenness is a recurring theme in this dataset, which has a long rare tail of OTUs and trends driven largely by the most abundant OTUs.

## Community richness

We hypothesized that disturbance frequency, indicated by mixing regime, determines biodiversity levels. Observed richness was calculated for every sample at the OTU level, and samples were aggregated by lake and layer. Hypolimnia typically contained more unique OTUs than epilimnia. Significant differences in richness between lakes were detected. For both layers, polymictic lakes had the fewest taxa, meromictic lakes had the most taxa, and dimictic lakes had intermediate numbers of taxa (Figure 1). This does not support our initial hypothesis that intermediate disturbance leads to the greatest amount of biodiversity. Instead, richness appears to increase with depth. As many variables are co-dependent with lake depth (such as mixing regime, volume of integrated water column, dissolved carbon concentrations and total nitrogen concentration), it is not clear what is driving this trend.

## C:\Users\Alex\AppData\Local\Microsoft\Windows\INetCacheContent.Word\fig1.png

**Figure 1. Richness by layer and lake.** Lakes on the x axis are arranged by depth (see Table 1 for lake abbreviations and depth measurements). Lakes CB, FB, and WS are polymictic, lakes NS, TB, and SS are dimictic, and lakes HK and MA are meromictic. Colored bars above each plot represent significant differences in richness between lakes, with each colored bar matching the color of a lakes boxplot. For example, in Panel A, the boxplot for CB has the colored bars matching FB, NS, TB, SS, HK, and MA above it. This indicates that it is significantly different from these lakes, but not significantly different from the missing colored bar, WS.



**Figure 2.** Weighted UniFrac distance was used to perform principal coordinates analysis on epilimnion (A) and hypolimnion (B) samples. In both layers, samples cluster significantly by lake and mixing regime as tested using ANOSIM. (See Table 1 for lake abbreviations; CB, FB, and WS are polymictic, NS, TB, and SS are dimictic, HK and MA are meromictic). Ellipses indicating the clustering of each lake were calculate based on standard error using a 95% confidence interval. Differences in bacterial community composition between lakes and mixing regimes are more pronounced in hypolimnia than epilimnia.

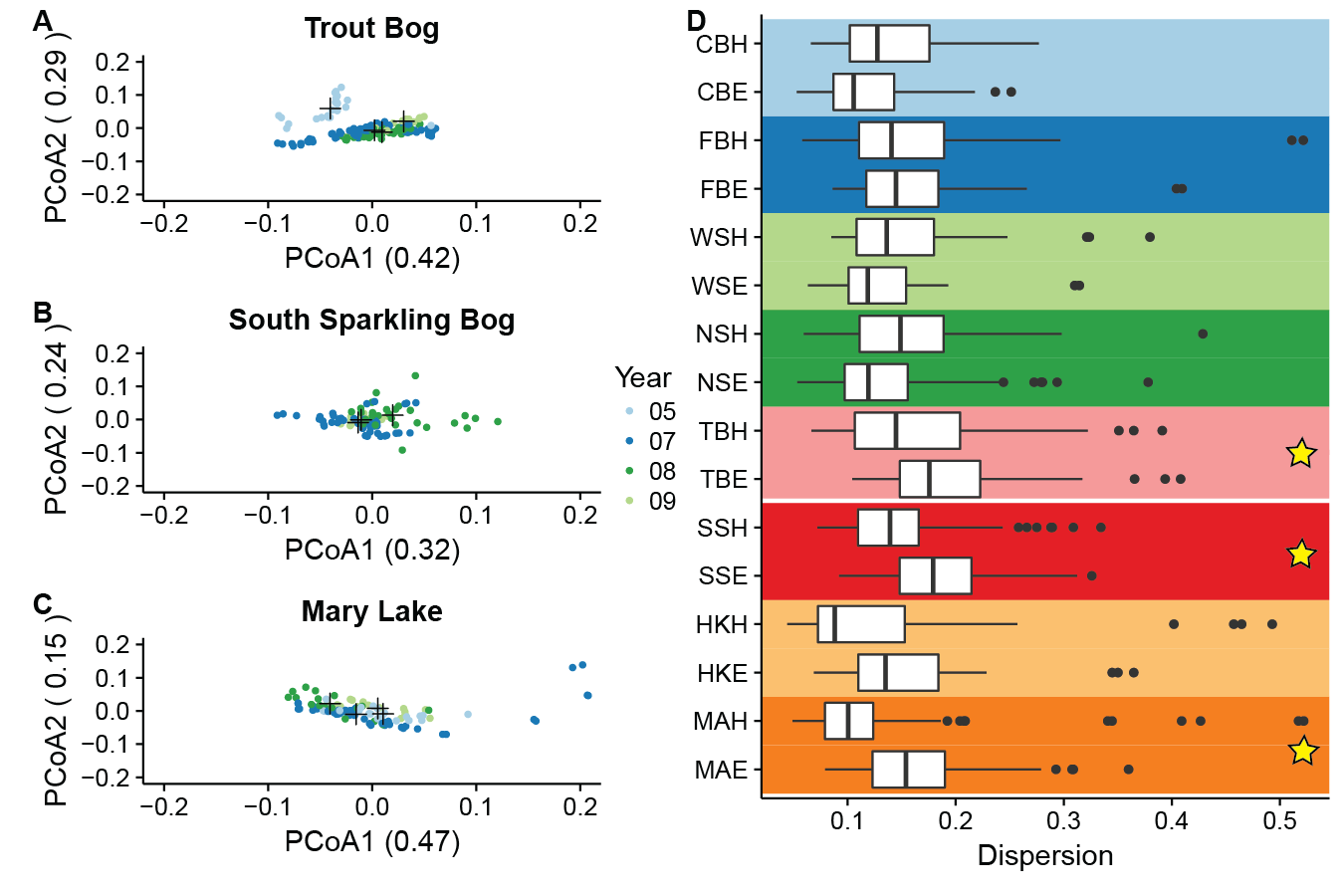
## Clusters of community composition

When differences in community composition are quantified using weighted UniFrac distance and visualized using principal coordinates analysis, several trends emerge (Figure 2). Communities from the epilimnion and hypolimnion layers are significantly distinct from each other at p < 0.05 in all lakes except Forestry Bog (FB) (p = 0.10). The strength of this clustering by layer within lakes increase with depth, as quantified by the r2 statistic produced by ANOSIM.

Within layers, mixing regime is the next driver of community composition. This effect is stronger in hypolimnia, which experience major changes in oxygen content during mixing events. As previously mentioned, depth and mixing regime are co-dependent with several other environmental variables such as carbon and nitrogen concentrations. These differences in environmental conditions likely lead to unique bacterial community compositions by mixing regime, with taxa endemic to certain types of lakes.

## Variability and dispersion

While community composition was distinct by layer, lake, and mixing regime, there was still variability in community composition over time. Each year in each lake had a significantly different community composition, indicating interannual variability in the community composition (Figure 3a-c, Figures S2-3). We found no evidence of repeating seasonal trends during the stratified summer months in these lakes. The abundance trends of individual OTUs were not observed to repeat each year, even when abundances in each year were normalized using z-scores. The lack of evidence for repeating annual trends during stratification is not surprising given prior published literature; however, this result emphasizes the need for multiple years of sampling to determine the true community composition of a body of freshwater.



**Figure 3. Internannual variability and dispersion by lake.** Principal coordinates analysis using weighted UniFrac as the distance metric was used to measure the amount of interannual variation in lake hypolimnia (A-C). Black crosses indicated the centroid for each year. ANOSIM was used to measure significance and to quantify the variation explained by year for each lake (r2). All hypolimnia showed significant clustering by year, but less of the variation in Mary Lake was explained by year, indicating that it has less interannual variation than the other hypolimnia. Six outliers in Mary Lake from 2007 are not shown, as their coordinates lie outside the range of PCoA1; these points were included in the ANOSIM significance test. Panel D shows dispersion of layers from PCoAs run on samples from each lake individually (Lake abbreviations found in Table 1; E indicates epilimnion and H indicates hypolimnion). Stars indicate significant differences between layers at p < 0.05 by Wilcoxon signed rank test. Layers were signficantly different in TB, SS, and MA (dimictic, dimictic, and meromictic, respectively). No significant differences in dispersion between layers in the polymictic lakes (CB, FB, and WS), or in NS (dimictic) and HK (meromictic).

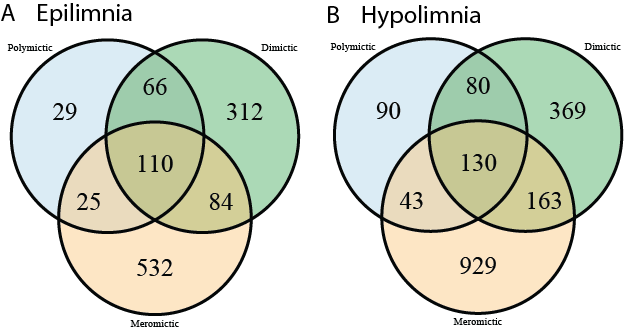
Varibility can also be assessed by measuring the dispersion of groups in PCoA. Dispersion is the distance of each point from the centroid of a group on an ordination plot. This analysis showed that layers had significantly different dispersion in two of the dimictic lakes (Trout Bog and South Sparkling Bog) and a meromictic lake (Mary Lake). It was not significantly different in the polymictic lakes, dimictic North Sparkling Bog, and meromictic Hell’s Kitchen. North Sparkling Bog likely does not significant differences between layers because it was artificially mixed during stratification in 2008, essentially making in polymictic for one year. While Hell’s Kitchen was meromictic in the year measured, it does not have as many data points as Mary Lake. When dispersion between layers is significant, the hypolimnion is on average more dispersed than the epilimnion, indicating higher variability. This is consistent with previously published results, and confirms that hypolimnia have more variability than epilimnia.

## The core community of bog lakes

One of the goals of this study was to determine the core bacterial community of bog lakes in general, and of specific types of lakes. Our previous analyses in Figure 2 showed that that community composition is distinct in each layer and lake, while the data presented in Figure 3 shows substantial variability in community composition within the same lake and layer. Still, rarefaction curves generated for the entire dataset and for each layer begin to level off, suggesting that we have sampled the majority of taxa found in our study sites (Figure S4). To identify the taxa that comprise the core community of a bog lake, we defined “core” as being present in 90% of a group of samples, regardless of abundance. Four OTUs met this criteria for all samples in the dataset: OTU0076 (bacI-A1), OTU0097 (PnecC), OTU0813 (acI-B2), and OTU0678 (LD28). These taxa were therefore also core in both epilimnia and hypolimnia. Additional core taxa in epilimnia also included OTU0004 (betI), OTU0184 (acI-B3), OTU0472 (Lhab-A4), and OTU0522 (alfI-A1), while additional core taxa in hypolimnia included OTU0042 (Rhodo), OTU0053 (unclassified Verrucomicrobia), and OTU0189 (acI-B2).

Because some OTUs were endemic to specific lakes, we performed the same core analysis on groups created by combining OTUs with the same tribe level classification. OTUs not classified at the tribe level were not included. Results were similar to those at the OTU level. Core tribes of all samples included bacI-A1, PnecC, acI-B2, and LD28, but also betIII-A1 and acI-B4. In epilimnia, the core tribes were bacI-A1, PnecC, betIII-A1, acI-B3, acI-B2, Lhab-A4, alfI-A1, LD28, and acI-B4, while in hypolimnia, they were Rhodo, bacI-A1, PnecC, betIII-A1, acI-B2, and acI-B4. These results show that despite lake differences and interannual variability, there are bacterial taxa that are consistently present in bog lakes.

Principal coordinates analysis in Figure 2 suggested that samples clustered by mixing regime as well as by lake and layer. To better understand how the bacterial community changes with mixing regime, we began by making Venn diagrams of OTUs shared and unique to each mixing regime (Figure 4). In this analysis, an OTU needs to appear in only one sample at any abundance to be considered present. In both epilimnia and hypolimnia, meromictic lakes have the greatest numbers unique OTUs while polymictic lakes have the least, consistent with the differences in richness between lakes found in Figure 1. Meromictic and dimictic lakes share the most OTUs in both layers, followed by dimictic and polymictic lakes, and then by meromictic and polymictic lakes. This suggests that there is a gradient of community composition across mixing frequencies.



**Figure 4. Numbers of unique and shared OTUs by mixing regime.** To better understand how community composition differs by mixing regime, we quantified the number of shared and unique OTUs in each category. An OTU needs only to appear in one sample at any abundance to be considered present in a category. We found that in both layers, meromictic lakes have the greatest numbers of unique OTUs and polymictic lakes have the least. Meromictic and dimictic lakes share the most OTUs, while meromictic and polymictic lakes share the least. Dimictic lakes share more OTUs with meromictic lakes than with polymictic lakes.

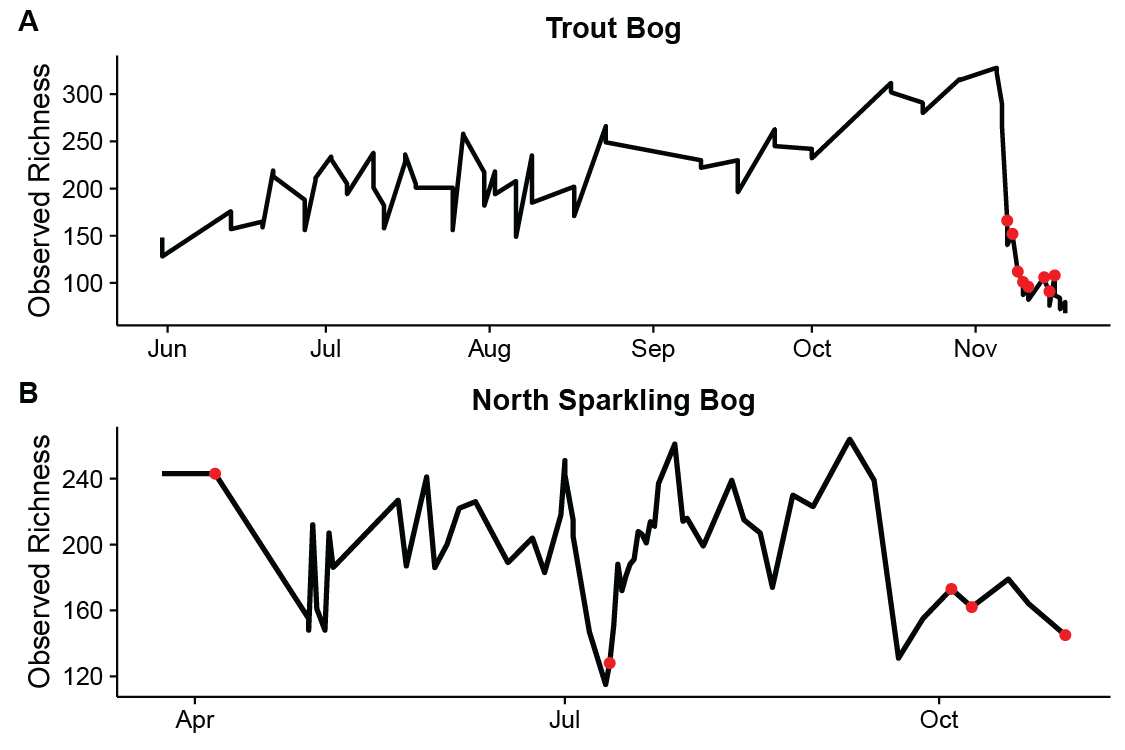
We next used indicator analysis to identify the taxa unique to each mixing regime. Indicator analysis is a statistical method used to determine if taxa are found significantly more often in certain pre-determined groups of samples than in others. In this case, the groups were defined by mixing regime, and normalization was applied to account for different numbers of samples in each group. OTUs were combined at every taxonomic level, and all taxonomic levels were run in the indictor analysis at once to account for differences in the ability of these levels as indicators (for example, the order Actinomycetales is a better indicator of polymictic hypolimnia than the phylum Actinobacteria). An abundance threshold of 500 reads was imposed on each taxonomic group. The full table of results from the indicator analysis are available in the supplemental material, while a few indicator taxa of interest are highlighted here.

The clade acI is a ubiquitous freshwater group, with specific lineages and tribes showing a preference for bog lakes in previous studies. Our dataset shows a further distinction of acI by mixing regime in epilimnia; acI-B3 is an indicator of polymictic epilimnia and acI-B2 is an indicator in dimictic epilimnia, while acI-A6 is found preferentially in meromictic epilimnia. Other indicators of polymictic epilimnia include other ubiquitous groups such as Limnohabitans and Polynucleobacter. Methylophilales, a putative methylotroph, is an indicator of dimictic epilimnia, while the phylum Planctomyces is found primarily in meromictic epilimnia.

The indicators for polymictic hypolimnia include several of the same indicators for epilimnia, such as PnecC, betI-A, and verI-A. Dimictic hypolimnia preferentially contain putative sulfate reducers Desulfobulbaceae, as well as the epilimnion indicator Methylophilales. Meromictic hypolimnia also contain putative sulfur reducers such as Syntrophobacteriales and Desulfobacteraceae, but also contain more unusual taxa such as Omnitrophica (formerly candidate phylum OP3), OP8, Verrucomicrobia, and Planctomyces. These indicator taxa likely reflect the environmental conditions unique to each mixing regime.

## Community composition during mixing events

One of the reasons bog lakes were chosen for this study was their fast transitions from mixed to stratified and back to mixed. Despite the transient nature of mixing events, two were captured in this dataset. One is a natural fall mixing event in dimictic Trout Bog in 2007, and the other is the artificial mixing of dimictic North Sparkling Bog in the summer of 2008. These two events give us a unique perspective on how the bacterial community changes in response to mixing events.

**Figure 5. Richness over time during mixing events.** In both panels, the black line traces the number of OTUs observed at each time point, and the red dots indicate dates on which the water column was uniform in temperature (mixed). Sharp decreases in richness are observed during both the fall mixing in Trout Bog, 2007 (A) and the artificial mixing in July in North Sparkling Bog, 2008 (B). Transient mixing dates in the fall of 2008 in North Sparkling Bog also show lower richness.

Observed richness was traced over time in the years of both mixing events. This analysis revealed sharp decreases in both lakes. In the case of the Trout Bog fall mixing, richness remained low until the end of the sampling season. In North Sparkling Bog, richness quickly returned to the levels observed before the artificial mixing. This suggests that mixing acts as a bottleneck that only certain taxa can tolerate, but that a source of diversity exists that allows fast re-establishment of the community.

If a reduced number of taxa can survive or even thrive during mixing events, which taxa are they? In Trout Bog hypolimnion in 2007, the top ten most abundant OTUs belonged to groups such as PnecC, Methylococcales, bacI-A1, acI-B2, Methylophilales, Acetobacteraceae, and acV-A2. The North Sparkling Bog hypolimnion mixed dates shared PnecC, Methylococcales, bacI-A1, acI-B2, and acV-A2, suggesting that these are the taxa selected for by the conditions of a mixing event. However, North Sparkling Bog also had high abundances of acI-B3, acI-B4, and Lhab-A4, but did not contain Acetobacteraceae and Methylophilales in its top ten most abundant OTUs during mixed dates.

Supplemental

Figure S1. Phylum rank abundance in entire dataset

Figure S2. PCoA of extra epilimnia by lake by year (extension of Fig 3)

Figure S3. PCoA of extra hypolimnia by lake by year (extension of Fig 3)

Figure S4. Rarefaction curves

Supplemental document – indicator analysis results