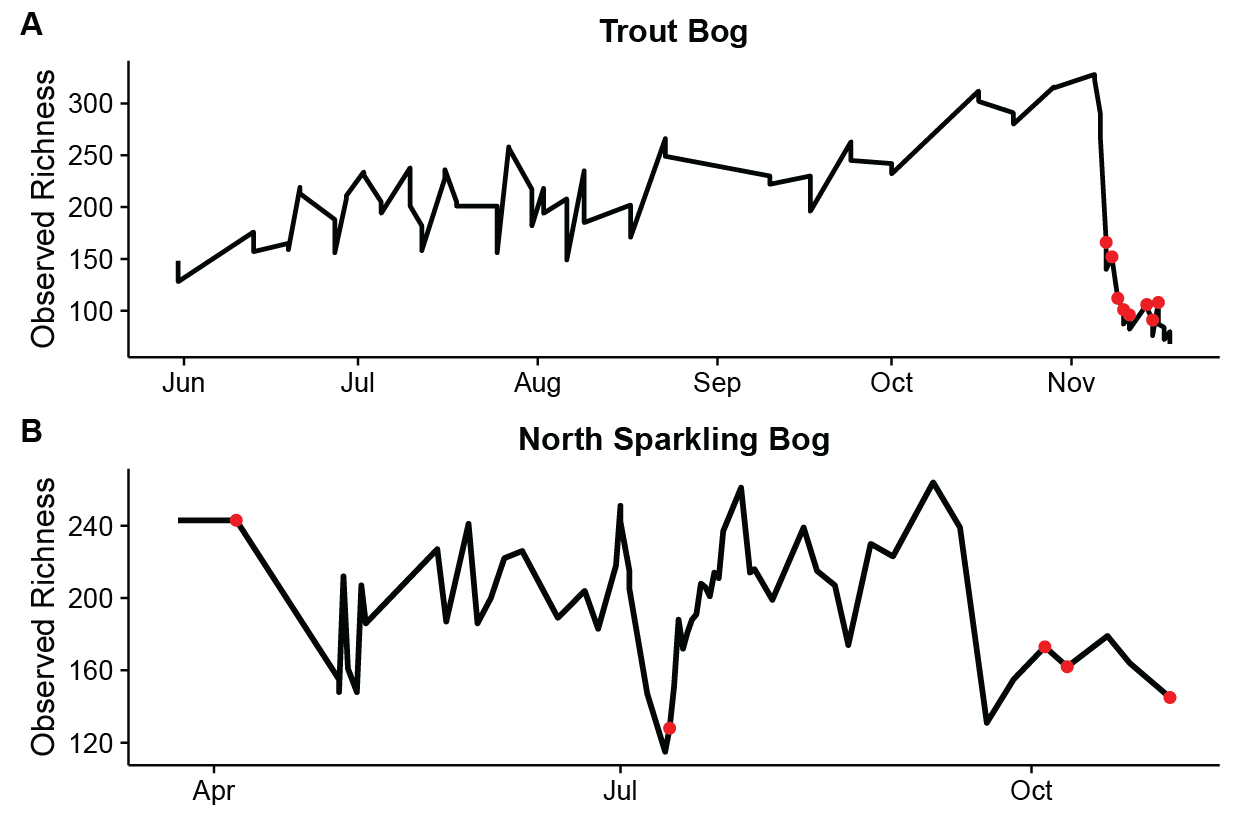
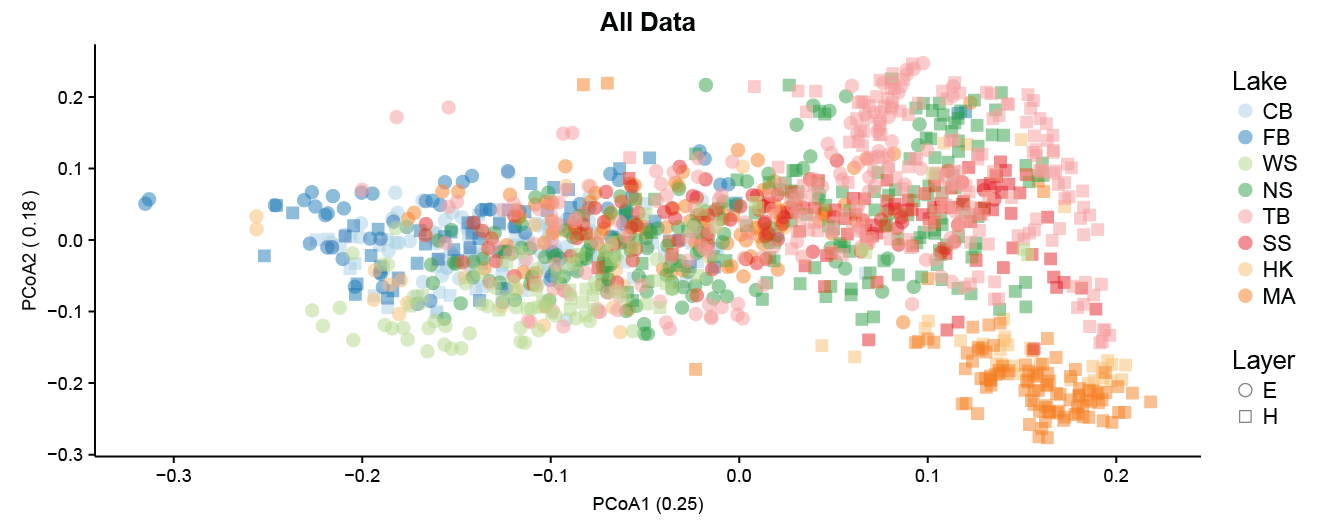
# Supplemental Figures

**Figure S1. Phylum rank abundance in entire dataset.** When OTUs are grouped by phylum and read abundances summed over the entire dataset, Proteobacteria, Actinobacteria, and Bacteroidetes are the most abundant phyla. Unclassified Bacteria are the fifth largest group. Members of the candidate phyla radiation such as OD1 (Parcubacteria) and OP3 (Omnitrophica) are also well-represented in this dataset.

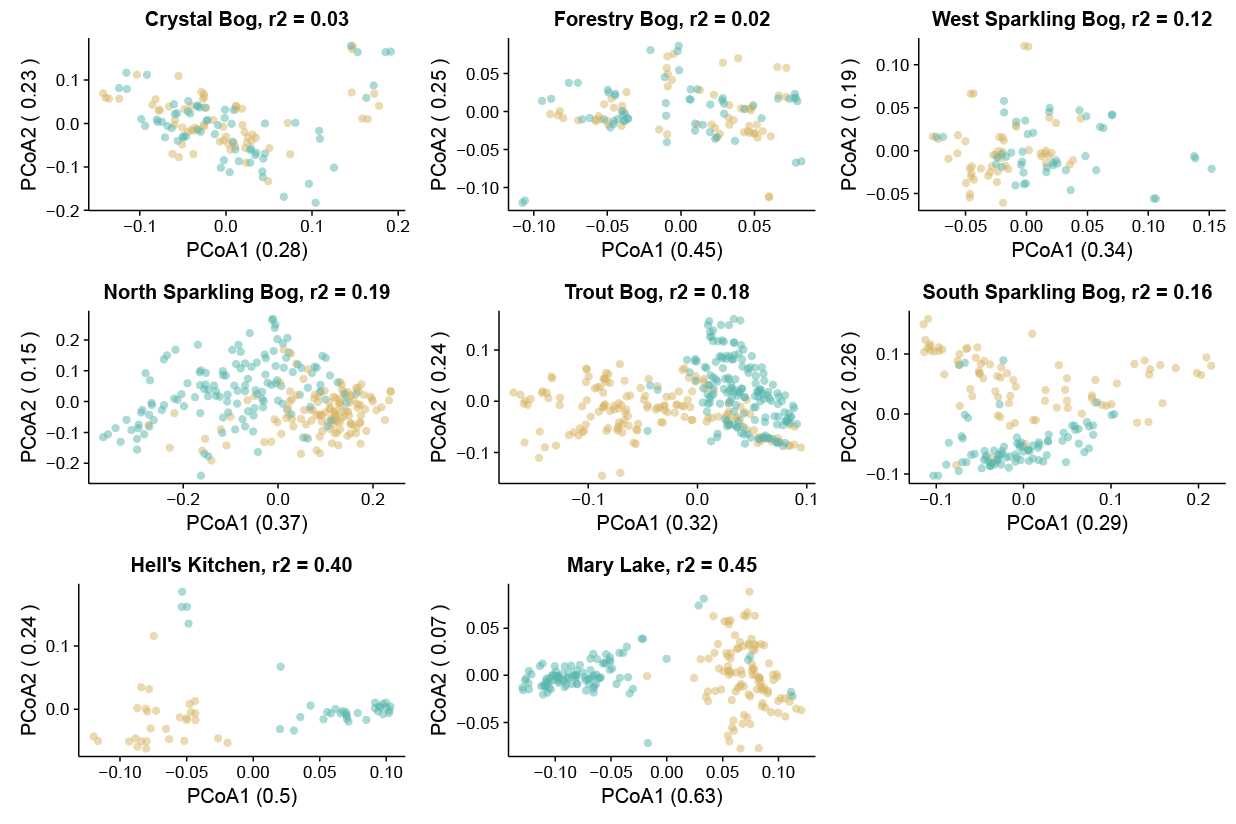
**Table S1. Means and standard deviations in the number of taxa by lake and layer.** In order to better interpret the results of Figure 1, statistics about richness are presented here.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Epilimnion  Mean | Standard Deviation | Hypolimnion  Mean | Standard deviation |
| Crystal Bog (CB) | 129 | 28 | 148 | 31 |
| Forestry Bog (FB) | 109 | 32 | 145 | 57 |
| West Sparkling Bog (WS) | 150 | 45 | 182 | 56 |
| North Sparkling Bog (NS | 143 | 33 | 178 | 40 |
| Trout Bog (TB) | 148 | 38 | 186 | 38 |
| South Sparkling Bog (SS) | 191 | 57 | 191 | 54 |
| Hell’s Kitchen (HK) | 199 | 67 | 397 | 124 |
| Mary Lake (MA) | 259 | 67 | 477 | 110 |

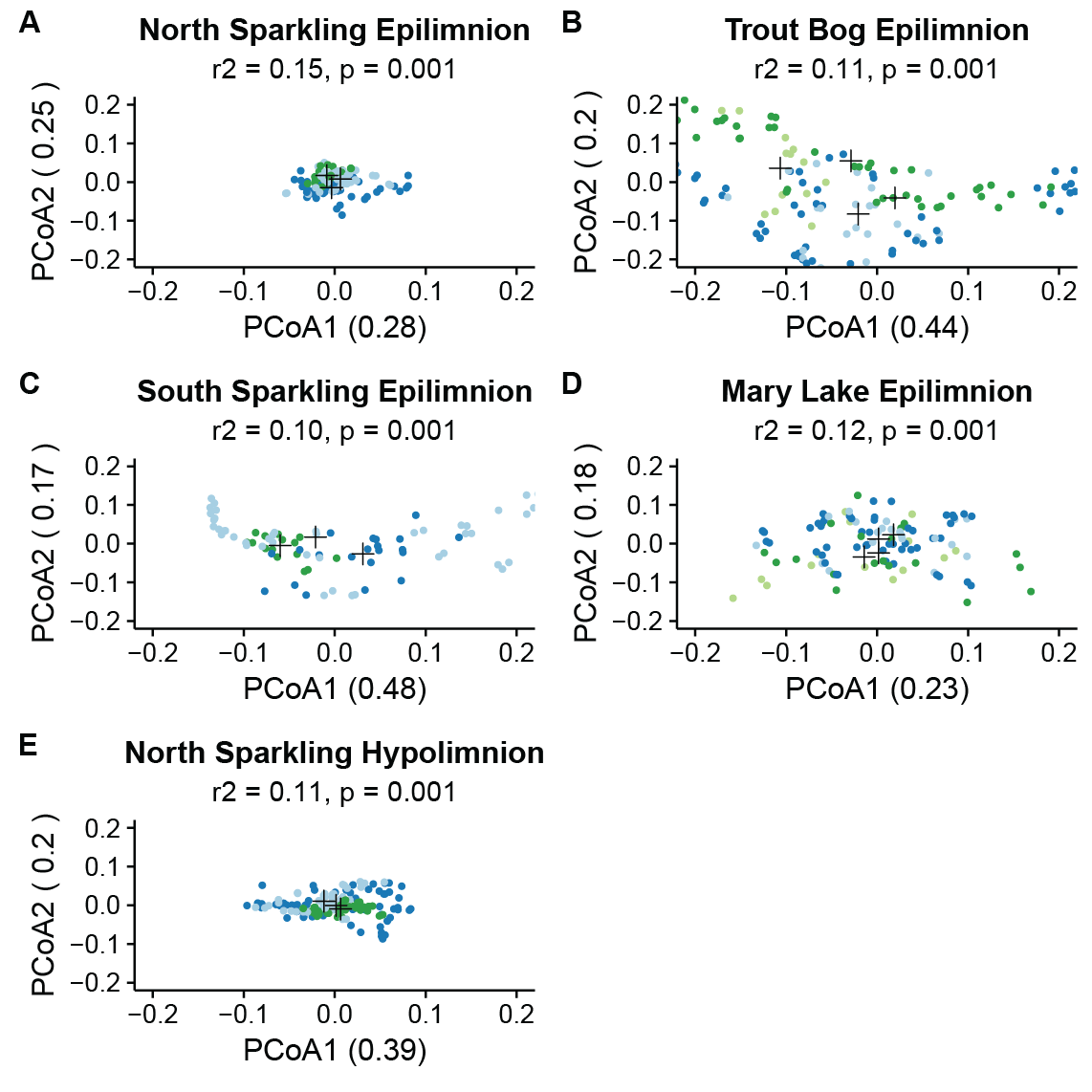
**Figure S2. Richness over time during mixing events.** In both panels, the black line traces the number of OTUs observed at each time point in the hypolimnion, and the red dots indicate dates on which the water column was uniform in temperature (mixed). Sharp decreases in richness are observed during 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.



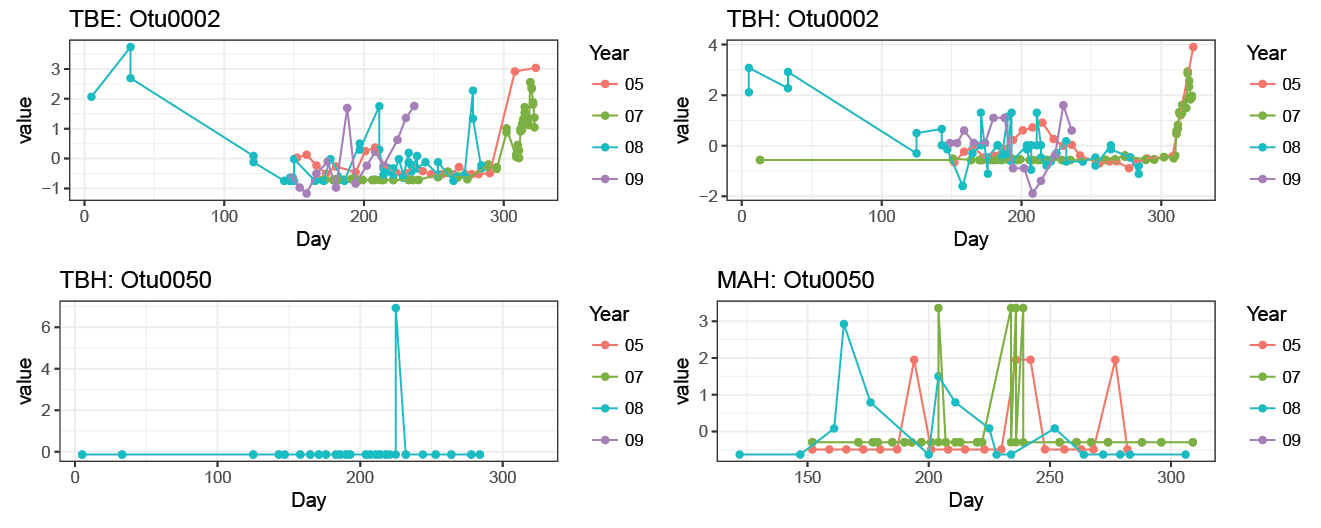
**Figure S3. PCoA of all data points with a layer designation.** As an overview of dataset, an ordination was performed on all datapoints with a layer designation using UniFrac distance and principle coordinates analysis. Samples cluster by lake, layer, and mixing regime. Polymictic epilimnia (E) and hypolimnia (H) samples are found together. Dimictic hypolimnia samples are distinct from dimictic epiliminia samples, but still overlap in the ordination. Meromictic hypolimnia cluster separately from the rest of the dataset.



**Figure S4. Layers cluster within lakes**. All clustering by layer is significant at p < 0.005 except Forestry Bog, where p = 0.101. Given that it is polymictic, shallow, and only includes one year of sampling, this is not surprising. Clustering is especially prominent in the meromictic lakes.



**Figure S5. PCoA of extra epilimnia and hypolimnia by lake by year**. Each year in each lake has a unique community composition, regardless of layer. Plots for lakes and layers not shown in Figure 3 are presented here, using the same analysis as in the main text. Only sites with at least three years of sampling were analyzed.



**Figure S6. Annual trends in OTUs.** We could not identify repeating seasonal trends in OTU abundances. While OTUs tended to show a consistent response to mixing events, their abundance during summer stratification was variable. Example plots showing abundance trends in OTUs over multiple years in the same site are presented here, and readers curious about other OTUs and sites can run the code below for any combination of OTU and location.

library(OTUtable) # You will need these three packages

library(ggplot2)

library(reshape2)

data(otu\_table) # Load the OTU table

# Write function to plot multiple years at once.

annual\_trends <- function(lake, otu){

bog <- bog\_subset(lake, otu\_table)

year1 <- year\_subset("05", bog)

year2 <- year\_subset("07", bog)

year3 <- year\_subset("08", bog)

year4 <- year\_subset("09", bog)

# Since sites have different years sampled, these if statements identify which years are present

if(dim(year1)[2] > 0){

# Once years present are identified, normalize and combine into a single table

year1 <- zscore(year1)

year2 <- zscore(year2)

year3 <- zscore(year3)

year4 <- zscore(year4)

ztable <- cbind(year1, year2, year3, year4)

}else if(dim(year1)[2] == 0 & dim(year3)[2] > 0){

year2 <- zscore(year2)

year3 <- zscore(year3)

year4 <- zscore(year4)

ztable <- cbind(year2, year3, year4)

}else if(dim(year1)[2] == 0 & dim(year3)[2] == 0 & dim(year4)[2] > 0){

year2 <- zscore(year2)

year4 <- zscore(year4)

ztable <- cbind(year2, year4)

}else{

ztable <- zscore(year2)

}

# Format the final table

ztable <- melt(ztable)

ztable$Year <- substr(ztable$Var2, start = 9, stop = 10)

ztable$Day <- format(extract\_date(ztable$Var2), format = "%j")

# Save the results for plotting

plot <- ggplot(data = ztable[which(ztable$Var1 == otu), ], aes(x = Day, y = value, group = Year, color = Year)) + geom\_point() + geom\_line() + theme\_bw() + labs(title = paste(lake, otu, sep = ": "))

return(plot)

}

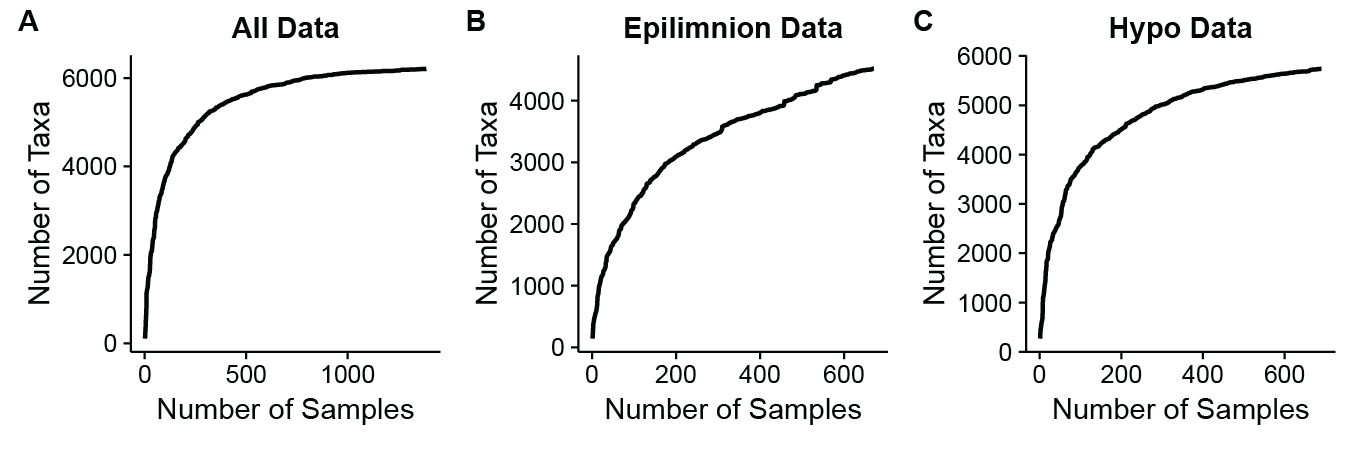
# Example Usage – 3 letter site code includes 1st 2 for site (see Table 1) and letter 3 for layer (E = epilimnion, H = hypolimnion. OTU designation is case sensitive, and number must contain 4 digits.

plot\_this <- annual\_trends(“TBE”, “Otu0012”)

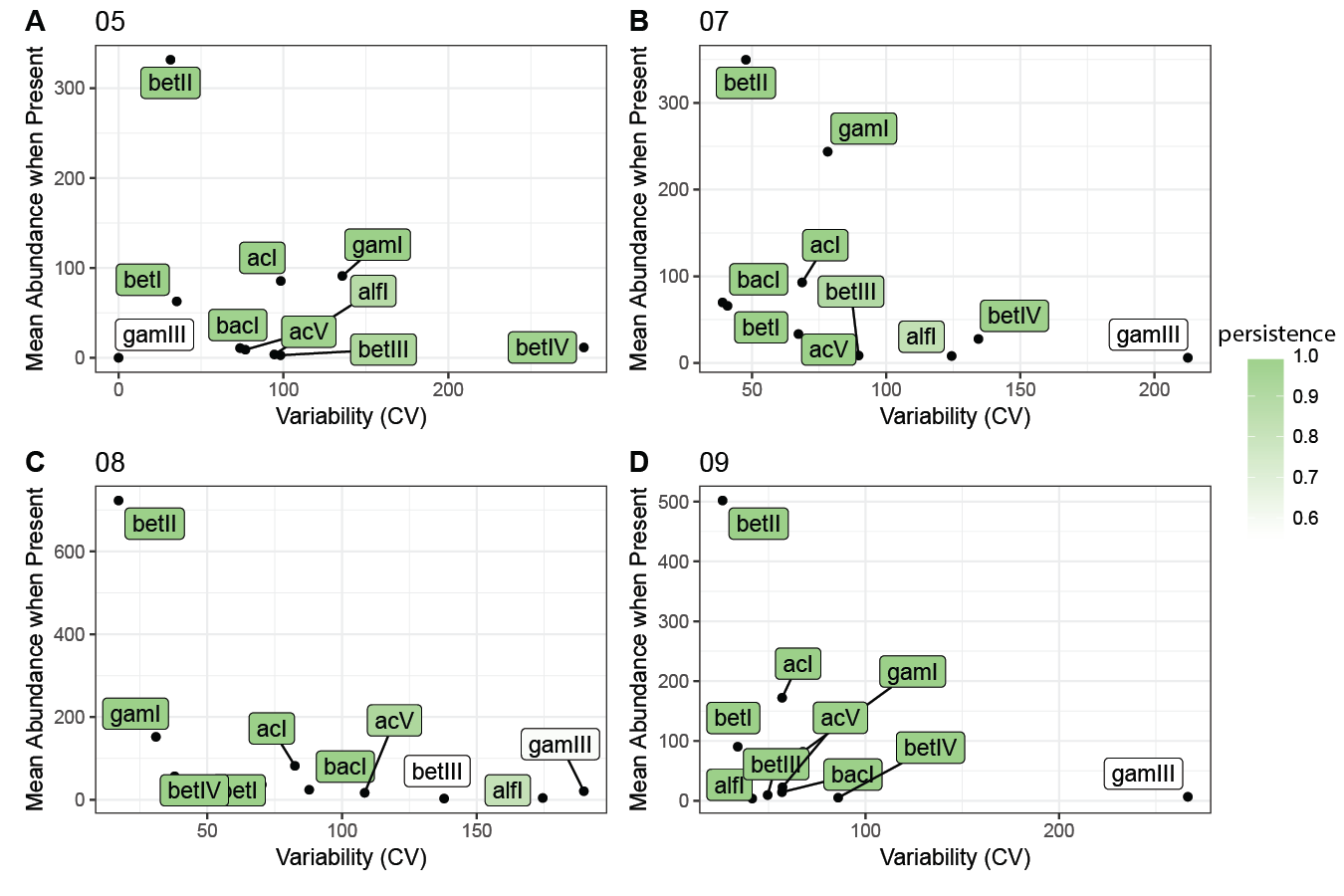
plot\_this

# You may get warning messages about points being removed. That means the OTU was not present in those points

# If all points were removed and no plot is produced, it was not present in that site



**Figure S7. Rarefaction curves.** Rarefaction curves generated by randomly ordering samples and counting the cumulative number of taxa show logarithmic rather than linear trends as the number of samples included increases. This suggests that we are beginning to exhaustively sample the community of bog lakes.

****

**Figure S8. Lineage traits by year.** Figure 5 demonstrates that lineages show consistent traits in different lakes; this plot shows that those traits are relatively consistent between years as well, as in the Trout Bog hypolimnion.