The North Temperate Lakes - Microbial Observatory dataset is a comprehensive 16S amplicon survey spanning four years, eight lakes, and two depths. We found that multiple years of sampling were necessary to describe the community of bog lake ecosystems. Richness and membership in these communities were structured by layer, mixing regime, and lake. We identified specific bacterial taxa present throughout the dataset, as well as taxa endemic to certain depths or mixing regimes. Mixing events were associated with reduced richness and an increase in the proportion of certain taxa. High levels of variability were detected in this dataset; each year in each lake harbored a unique bacterial community. Our results emphasize the importance of multiple sampling events to assess full bacterial community membership and variability.

The bog lakes in this study have been model systems for freshwater microbial ecology for many years. Early studies used Automated Ribosomal Intergenic Spacer Analysis (ARISA), a fingerprinting technique for identifying unique bacterial taxa in environmental samples (1). Our research built upon these studies and added information about the taxonomic identities of bacterial groups. For example, persistent and unique bacterial groups were detected in the bog lakes using ARISA (2); using 16S amplicon sequencing, we determined that these groups are the ubiquitous freshwater bacteria LD28, acI-B2, PnecC, and bacI-A1. Differences in richness and community membership were previously detected in Crystal Bog, Trout Bog, and Mary Lake, three sites representative of the three mixing regime categories of polymictic, dimictic, and meromictic (2). Our data supported these results and suggest that these trends are indeed linked with mixing regime, as we included multiple lakes of each type in this study.

We also supported previous research on the characteristics of bacterial communities in the epilimnion and hypolimnion, and the impacts of lake mixing on these communities. We confirmed that epilimnia tended to be more variable than hypolimnia, potentially due to increased exposure to climatic events (2). Mixing was disruptive to both epilimnion and hypolimnion communities, selecting for only a few taxa that can thrive during this disturbance, but quickly recovering diversity (3, 4). Comparing richness between lakes of different mixing regimes did not support the intermediate disturbance hypothesis, our initial inspiration for the collection of this dataset; rather, the least mixed lakes had the most diverse communities. As many variables are co-dependent with mixing regime (such as depth, volume of integrated water column, dissolved carbon concentrations and total nitrogen concentration), it is not clear which variables are driving this trend.

We were not able to detect seasonal trends in bog lakes in our multiple years of sampling. While seasonality in marine and river systems has been well-established by our colleagues, previous research on seasonality in freshwater lakes has produced mixed results (5–8). Nelson identified distinct, repeatable community types in his seminal paper on seasonality in alpine lakes, but noted that stratified summer communities were distinct each year (9). Seasonal trends were detected in a time series from Lake Mendota similar to ours, but summer samples in Lake Mendota were more variable then those collected in other seasons (10). In the previous ARISA-based research on the bog lakes included in our dataset, community properties such as richness and rate of change were consistent each year, and the phytoplankton community hypothesized to drive seasonal trends in the bacterial community based on correlation studies (11–13). Synchrony in seasonal trends was observed by Kent, et al. (12); however, in a second year of sampling for seasonal trends in Crystal Bog and Trout Bog, these findings were not reproduced (14). Successional trends were studied in Crystal Bog and Lake Mendota in 2000-2001 and “dramatic changes” in community composition associated with drops in biodiversity are described during the summer months, while spring, winter, and fall had more stable community composition (11). Because the majority of our dataset was collected during the summer stratified period, increased summer variability may explain why we see a different community each year and a lack of seasonal trends. However, we cannot disprove the presence of seasonality in temperate freshwater lakes.

One of the biggest benefits of 16S amplicon sequecing over ARISA is the ability to classify sequences. In addition to a core of persistent taxa found in nearly every sample collected, we also identified taxa endemic to either the epilimnion or hypolimnion and to specific mixing regimes. These endemic taxa likely reflect the biogeochemical differences driven by mixing regime. Dimictic and meromictic hypolimnia, which are consistently anaerobic, harbor putative sulfur reducers not present in polymictic hypolimnia, which are more frequently oxygenated. Members of clade acI partition by mixing regime in epilimnia, though the functional traits driving this filtering effect are the subject of active study (15). Interestingly, the hypolimnion of meromictic Mary Lake contains several taxa classified into the candidate phyla radiation and a larger proportion of completely unclassified reads than other hypolimnia (16). This is consistent with the findings of other 16S and metagenomics studies of meromictic lakes, and suggests that the highly reduced and consistently anaerobic conditions in meromictic hypolimnia support taxa would make excellent study systems for research on members of the candidate phyla radiation and “microbial dark matter” (17, 18).

Perhaps the biggest implication of this study is the importance of repeated sampling of the same locations. A similar dataset spanning only a single year would not have captured the full extent of variability observed, and therefore would not have detected as many of the taxa belonging to the bog lake community. While we found no evidence for seasonal trends or repeated annual trends, it is possible that there are cycles or variables acting on scales greater than the five years covered in this dataset, or that annual differences are driven by environmental factors that do not occur every year. Biotic interactions between bacterial taxa may also contribute to the observed variability. Understanding the factors that contribute to variability in bog lake communities will lead to improved predictive modelling in freshwater systems, allowing forecasting of bloom events and guiding better management strategies. Additionally, these systems may be ideal for addressing some of the core questions in microbial ecology, such as how community assembly occurs, how interactions between taxa shape community composition, and how resource partitioning drives the lifestyles of bacterial taxa.

To answer these questions and more, we are continuing to collect and sequence samples for the North Temperate Lakes – Microbial Observatory, and we are expanding our sequencing repertoire beyond 16S sequencing. All of the 16S data we have currently generated can be found in the R package “OTUtable” which is available on CRAN for installation via the R command line, or on our GitHub page. We hope that this dataset and its future expansion will be used as a resource for researchers investigating their own questions about how bacterial communities behave on long time scales.

1. **Fisher MM**, **Triplett EW**. 1999. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. Appl Environ Microbiol **65**:4630–4636.

2. **Shade A**, **Jones SE**, **McMahon KD**. 2008. The influence of habitat heterogeneity on freshwater bacterial community composition and dynamics. Environ Microbiol **10**:1057–1067.

3. **Shade A**, **Read JS**, **Welkie DG**, **Kratz TK**, **Wu CH**, **McMahon KD**. 2011. Resistance, resilience and recovery: Aquatic bacterial dynamics after water column disturbance. Environ Microbiol **13**:2752–2767.

4. **Shade A**, **Read JS**, **Youngblut ND**, **Fierer N**, **Knight R**, **Kratz TK**, **Lottig NR**, **Roden EE**, **Stanley EH**, **Stombaugh J**, **Whitaker RJ**, **Wu CH**, **McMahon KD**. 2012. Lake microbial communities are resilient after a whole-ecosystem disturbance. ISME J **6**:2153–2167.

5. **Crump BC**, **Hobbie JE**. 2005. Synchrony and seasonality in bacterioplankton communities of two temperate rivers. Limnol Oceanogr **50**:1718–1729.

6. **Gilbert JA**, **Steele JA**, **Caporaso JG**, **Steinbruck L**, **Reeder J**, **Temperton B**, **Huse S**, **McHardy AC**, **Knight R**, **Joint I**, **Somerfield P**, **Fuhrman JA**, **Field D**. 2012. Defining seasonal marine microbial community dynamics. ISME J **6**:298–308.

7. **Fuhrman JA**, **Hewson I**, **Schwalbach MS**, **Steele JA**, **Brown M V**, **Naeem S**. 2006. Annually reoccurring bacterial communities are predictable from ocean conditions. Proc Natl Acad Sci USA **103**:13104–13109.

8. **Cram J**, **Chow C**, **Sachdeva R**, **Needham D**, **Parada A**, **Steele J**, **Fuhrman J**. 2015. Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. ISME J **9**:563–580.

9. **Nelson CE**. 2009. Phenology of high-elevation pelagic bacteria: the roles of meteorologic variability, catchment inputs and thermal stratification in structuring communities. ISME J **3**:13–30.

10. **Kara EL**, **Hanson PC**, **Hu YH**, **Winslow L**, **McMahon KD**. 2013. A decade of seasonal dynamics and co-occurrences within freshwater bacterioplankton communities from eutrophic Lake Mendota, WI, USA. ISME J **7**:680–4.

11. **Yannarell AC**, **Kent AD**, **H. LG**, **Kratz TK**, **Triplett EW**. 2003. Temporal Patterns in Bacterial Communities in Three Temperate Lakes of Different Trophic Status. Microb Ecol **46**:391–405.

12. **Kent AD**, **Yannarell AC**, **Rusak JA**, **Triplett EW**, **McMahon KD**. 2007. Synchrony in aquatic microbial community dynamics. ISME J **1**:38–47.

13. **Kent AD**, **Jones SE**, **Lauster GH**, **Graham JM**, **Newton RJ**, **McMahon KD**. 2006. Experimental manipulations of microbial food web interactions in a humic lake: Shifting biological drivers of bacterial community structure. Environ Microbiol **8**:1448–1459.

14. **Rusak JA**, **Jones SE**, **Kent AD**, **Shade AL**, **Mcmahon TM**. 2009. Spatial synchrony in microbial community dynamics : testing among-year and lake patterns **30**:936–940.

15. **Garcia SL**, **Buck M**, **McMahon KD**, **Grossart H-P**, **Eiler A**, **Warnecke F**. 2015. Auxotrophy and intra-population complementary in the “interactome” of a cultivated freshwater model community. Mol Ecol n/a–n/a.

16. **Hug LA**, **Baker BJ**, **Anantharaman K**, **Brown CT**, **Probst AJ**, **Castelle CJ**, **Butter CN**, **Hernsdorf AW**, **Amano Y**, **Ise K**, **Suzuki Y**, **Dudek N**, **Relman DA**, **Finstad KM**, **Amundson R**, **Thomas BC**, **Ban JF**. 2016. OPEN A new view of the tree of life **1**:1–6.

17. **Gies EA**, **Konwar KM**, **Beatty JT**, **Hallam SJ**. 2014. Illuminating microbial dark matter in meromictic Sakinaw Lake. Appl Environ Microbiol **80**:6807–6818.

18. **Borrel G**, **Lehours AC**, **Bardot C**, **Bailly X**, **Fonty G**. 2010. Members of candidate divisions OP11, OD1 and SR1 are widespread along the water column of the meromictic Lake Pavin (France). Arch Microbiol **192**:559–567.