Diversity and distribution of sediment bacteria across an ecological and trophic gradient

Hailey M. Sauer^{1,2}, Trinity L. Hamilton^{*1,3}, Rika E. Anderson⁴, Charles E. Umbanhowar Jr.⁵, Adam J. Heathcote²

1. Department of Plant and Microbial Biology, University of Minnesota, St. Paul, MN, 55108 USA
2. St. Croix Watershed Research Station, Science Museum of Minnesota, Marine on St. Croix, MN, 55047, USA
3. The Biotechnology Institute, University of Minnesota, St. Paul, MN, 55108 USA
4. Biology Department, Carleton College. Northfield, MN, 55057 USA
5. Departments of Biology and Environmental Studies, St. Olaf College, Northfield, MN, USA, 55057 USA

* Corresponding Author

Email: trinityh@umn.edu (TLH)

Abstract

Can we cluster bacterial sediment communities based on lake size, depth, and trophic status? Or, are bacterial microbial communities an emergent property of their geography, integrating regional physical and climatic conditions? Lakes in Minnesota are uniquely situated to address these questions because of their wide geographic range and variability in size and basin land-use. In this study, we selected twenty lakes with varying morphological and chemical properties across four ecological regions of Minnesota. Our objectives were to (i) evaluate the diversity and spatial variation of the bacterial community at the sediment-water interface and (ii) determine how lake location and watershed land-use impact aqueous chemistry and influence community structure. Our data indicate that sediment communities from similar depth intervals are more likely to cluster by ecological region rather than any individual lake properties (*e.g.*, trophic status, TP concentration, lake depth). However, composition is tied to a given lake, wherein samples from the same core were more alike than samples at similar depths across lakes. Our results illustrate the diversity within lake sediment microbial communities and provide insight into relationships between taxonomy, physicochemical, and geographic properties of north temperate lakes.

Introduction

A community of microorganisms living together in a particular environment or habitat are referred to as a microbiome. In the past decade, studies into the microbiomes of human organs, plants, soils, waters, and even space station astronauts have enhanced our understanding of how microbiota protect us from pathogens, increase agricultural production,

and ultimately cycle nutrients in and throughout the natural environment (1–4). Microbiomes are connected to the biogeochemical cycling of nutrients at both local and global levels, and as a result, disturbances or variations in their community composition can result in gains or losses of functional attributes, changes in nutrient availability, and shifts in ecosystem adaptability (5–7). Understanding the selective pressures on microbiome community composition can provide insight into an environment’s ability to support higher trophic levels and respond to anthropogenic change. Freshwater lakes are an ideal system to explore these insights as they provide a variety of regulating and cultural services, which hinge on the composition of their microbiome (8).

Despite their relatively small surface area, lakes contribute disproportionately to biogeochemical cycles, including the essential macronutrients carbon, nitrogen, and phosphorus (9–13). Nutrients in lakes are recycled through several biotic and abiotic processes, but ultimately a significant proportion end up in sediments, where bacterial abundance and diversity typically exceeds that of the water column (14–16). As organic matter reaches the sediments it is degraded by select bacteria and archaea. Oxygen is consumed, and anaerobic respiration proceeds. These respiratory processes occur along a redox gradient, eventually leading to the transformation of nitrogen, iron, and sulfur compounds. The complementary metabolisms of sediment microbiomes make sediments a global biogeochemical hotspot, one in which there has been a concerted effort to understand the environmental factors that regulate composition and function (17–20).

Chemical and physical characteristics of the lake such as salinity, pH, temperature, and nutrient concentrations select for specific bacteria, a process commonly referred to as species

sorting (21). The physicochemical characteristics of the system are partly based on the external inputs of both organic matter and nutrients, such as phosphorus and nitrogen, from the surrounding watershed (22). Different land uses (e.g., agricultural, urban, forested, etc.) in the watershed strongly influence the amount and types of terrestrial organic matter and nutrients that enter the water, and therefore land use may subsequently affect community composition (23). While there have been several studies that address the effects that local environmental factors have on microbiome species selection (e.g., eutrophic reservoirs, alpine lakes), few have examined the effects of the land-use of the watershed on community assembly (22–26).

In this study we selected twenty lakes with varying morphological and chemical properties across four ecological regions with varying land use in Minnesota (U.S.A.) (Fig 1). We sought to (i) evaluate the diversity and spatial variation of the bacterial community at the sediment-water interface and (ii) determine how lake location and watershed land use impact aqueous chemistry and influence community structure. We hypothesize that community composition of lake sediments will appear homogeneous across ecological regions and land use at higher taxonomic levels; however, when analyzed at lower resolutions we hypothesize increased structure by eco-region at lower taxonomic levels. To test this, we compare the alpha and beta diversity of bacteria across taxonomic scales (Phylum to Order) and ecological regions, and we highlight important regional and local factors that influence community composition.

Fig 1. Map of Sampling Locations

Location of the twenty study lakes that were cored across the state of Minnesota (U.S.A.) between the summer of 2018 and 2019, shaded by ecoregion. Map was created using QGIS and data were made available by the MN Geospatial Commons.

Materials & Methods

Site description

For this study, we selected twenty lakes within Minnesota's Sentinel Lakes in a Changing Environment (SLICE) program. SLICE is a collaborative research initiative providing long-term data on a representative sub-sampling of Minnesota's lakes that span the diverse geographic, land-use, and climatic gradients present in Minnesota (Fig 1). The lakes span four of the seven Environmental Protection Agency/Commission for Environmental Cooperation's (level III) ecological regions. These regions can be characterized by their underlying geology, soils, vegetation, and land use. This is the first comprehensive sediment bacterial survey of these lakes.

Water Sample Collection & Analysis

From each site we collected an integrated (0-2m) epilimnetic water sample, and a hypolimnetic (maximum lake depth - 1m) water sample when thermal stratification was present. We analyzed samples using standard methods for total nitrogen (TN) (4500-N), total phosphorus (TP) (4500-P), soluble reactive phosphorus (SRP) (4500-P), NH₃ (4500-NH₃), NO₃ (4500-NO₃) on a SmartChem 170 (Unity Scientific, Inc.) discrete analyzer (APHA 2012). We measured carbon as dissolved organic carbon (DOC), and dissolved inorganic carbon (DIC) using a Torch Combustion TOC Analyzer (Teledyne Tekmar, Inc.) following Standard Method 5310-C (27). Additionally, we measured chlorophyll-a concentrations via fluorometry following the EPA method 445.0 (28). We provided a complete summary of aqueous chemistry results in the Supplemental Information (S1 Table).

Sediment Sample Collection & DNA Isolation

We collected sediment cores from July 2018 through June 2019 using a rod-driven piston corer with a 7cm diameter polycarbonate tube (29). We determined coring locations (i.e., flat areas near the deepest basin) using publicly available bathymetric maps (<https://www.dnr.state.mn.us/lakefind/index.html>), avoiding steep-sided “holes” where sediment-focusing may be high. After sediment core retrieval, we stabilized core tops in the field using a gelling agent (e.g., Zorbitrol) and returned intact cores to the laboratory where we stored them vertically at 4°C until processing. In cases where the upper sediments were extremely flocculent, we immediately sectioned the upper most sections (~0-30 cm) in the field to prevent mixing during transport.

We vertically extruded the cores in the lab in 1 to 2 cm intervals, depending on lake productivity, and took subsamples from two intervals for DNA analysis. The subsamples collected were from the 0-2cm (hereafter referred to as shallow) and either the 3-4cm or 4-6cm interval (hereafter referred to as deep). We extracted DNA from 0.25g of wet sediment from each subsample using a PowerSoil DNA Isolation Kit (Qiagen, Inc.) following the manufacturer’s recommendations. We determined final bulk DNA concentrations using a Qubit dsDNA HS Assay kit (Molecular Probes, Eugene, OR, USA) and Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA).

Nucleic Acid Preparation, Amplification, and Sequencing

We submitted the DNA samples to the University of Minnesota’s Genomic Center (UMGC) where they performed library preparation for Illumina high-throughput sequencing using a Nextera XT workflow and a 2x300 bp chemistry. The workflow utilizes transposome-based shearing which fragments the DNA and adds adapter sequences in one step. The DNA

was amplified and dual-indexed with adapter sequences through PCR, using primers 515Ff and 806rB to target the V4 hypervariable region of bacterial 16S SSU rRNA gene sequences. The indexed samples were then sequenced using an Illumina MiSeq at the UMGC. A total of 3.29 million (3,290,170) raw reads were obtained from 40 samples.

Data Processing

We conducted post-sequence processing in mothur following the MiSeq SOP (30,31). Briefly, we merged forward and reverse reads, and screened, trimmed and removed ambiguous bases. We then aligned the reads to references in the SILVA database (v.132), and identified and removed chimeras using UCHIME (32,33). Finally, given the nature of the study (i.e., broad scale patterns of diversity), we classified the sequences as operational taxonomic units (OTUs) using a 97% similarity threshold and assigned taxonomy using the SILVA database (34,35).

Community Analysis & Statistics

Unless otherwise stated, we conducted all statistical analyses in R (36,37). We loaded both the environmental and community data into R using Phyloseq (38) and removed any reads classified as mitochondrial or chloroplast. Our final dataset after all post-processing contained 2,181,132 reads assigned to 53,854 taxa across 40 samples.

Alpha Diversity

We removed all singletons (OTUs observed only once across all 40 samples) from the data before calculating alpha diversity statistics. Given the observed correlation of richness based on sample read depth across sequencing batches (S1 Fig.), we chose to rarefy the data to 90% the read depth of the lowest samples (15,771 reads). Our final dataset for alpha diversity

included 630,840 read counts of 25,563 taxa across 40 samples. We calculated all available alpha diversity measures using the Phyloseq package in R (S2 Fig. and S2 Table) (38). We compared the richness (observed number of OTUs) and evenness (Shannon) of the samples based on sample depth using a Wilcox test, and trophic status and ecological status using a Kruskal-Wallis test with a Dunn Post Hoc test and Bonferroni correction. In all tests, one outlying sample (Trout, Deep) was removed due to uncharacteristically low diversity. Finally, we assessed the predictive capabilities of environmental parameters using multiple regression and determined the significance and variance partitioned by each regressor using the relaimpo and vegan packages in R (39,40). We selected the final models based on AIC scores for both richness (observed) and evenness (Shannon).

Beta Diversity

Prior to beta diversity analysis we denoised the samples by removing any OTU that did not have 2 or more counts and occur in at least 10% of the samples. Post filtering, the average number of reads per sample was reduced to 47,605, the minimum read depth was 15,150, and the maximum read depth was 99,561. Since OTU data have a strong positive skew, we attempted to diminish the effects using a variance stabilizing transformation (VST) (41). Log-like transformations, like VST, have been shown to transform count data to near-normal distributions and produce larger eigengap values, ultimately leading to more consistent correlation estimates which influence downstream analyses (42). After denoising and transformation, the final dataset for beta diversity analysis included 5,512 taxa across 40 samples.

We visualized the sample dissimilarity using principal component analysis (PCA) and the ordinate function in Phyloseq (38). After ordination, we further analyzed the distribution of taxa based on the ecological regions using permutational analysis of variance (PERMANOVA) and the “adonis” function in vegan (40). We used a Bray-Curtis dissimilarity to test for group differences and assessed dispersion within groups using permutations and vegan’s “betadisp” and “permutest” functions. Prior to creating the dissimilarity matrix, we converted negative VST values to zero because negative values after transformation likely represent zero counts or very few counts and for the distances and hypothesis in future tests these values would be negligible. We performed a cluster analysis using Ward’s (D2) method and the same dissimilarity matrix generated for the PERMANOVA analysis.

Results and Discussion

Alpha Diversity

We used alpha diversity metrics to summarize the structure of the bacterial communities in terms of the number of OTUs (richness) and the distribution of their abundances (evenness) for all samples. We then compared the observed diversity (a measure of richness) and Shannon diversity (a measure of evenness) across sampling locations and ecological regions (Fig 2). Sample richness varied from ~2000-4000 OTUs and sample evenness varied from 5.5 to 7.5. Sediments, both shallow and deep, from Carrie Lake were the most diverse in terms of richness (4116 OTUs and 4100 OTUs; shallow and deep respectively) and evenness (7.41, 7.39; shallow and deep respectively). The least diverse shallow sample in terms of richness was Pearl Lake (2371 OTUs) and evenness was Greenwood Lake (6.44). Trout Lake

was the least diverse deep sample in both the total number of OTUs (552 observed) and Shannon diversity (5.49). There was no significant difference within lake diversity between shallow and deep lake sediments across all samples (Wilcoxon test $p > 0.05$); however, the deeper interval sample was more diverse in both richness and evenness in a majority of lakes. The exception to that pattern were the samples from lakes in the Canadian Shield (CS) where all of the shallow interval samples were more diverse.

Fig 2. Bacterial Alpha Diversity Across Samples

Observed diversity, a measure of richness, and Shannon diversity, measure of evenness, for all samples where shape indicates the sediment depth, color indicates the ecological region, and sites are ordered based on ecological region then latitude. One sample (Trout, Deep - CS), with lower diversity, was removed for visualization.

While all samples were highly diverse (Richness > 2250 OTUs), there were differences in the levels of richness and evenness when comparing samples across the ecological region (Fig 3). Shannon diversity (evenness) levels were statistically different across the ecological regions ($p = 0.008$). Across groups the samples from lakes in the Western Cornbelt Plains (CB) were more diverse than both the Northern Lakes and Forests (NLF) ($p=0.0206$) and CS ($p=0.0099$) samples. Observed diversity (richness) was also statistically significant across the ecological regions ($p=0.003$). Again, there were statistical differences in the diversity levels of the CB and NLF ($p=0.007$) and CS ($p=0.006$) wherein the CB samples had greater richness.

Fig 3. Bacterial Alpha Diversity by Ecological Region

Box plots show mean alpha level diversity of the observed Operational Taxonomic Units (OTUs) and Shannon indices of the four distinct ecological regions present within the study area: Western Cornbelt Plains (CB), North Central Hardwood Forests (NCHF), Northern Lakes and Forests (NLF), and Canadian Shield (CS). Open (shallow) and closed (deep) circles indicate unique samples and color indicates the ecological region. One sample was removed from both plots for due to uncharacteristically low diversity. Significance between regions was calculated nonparametrically using a Kruskal Wallis H test followed by a Dunn post hoc test with a Bonferroni correction. Reported p values indicate significant differences in Observed and Shannon diversity (respectively) across ecological regions, specifically the diversity of CB lake

sediments when compared to NLF ($p=0.007$ & $p=0.0206$) and CS ($p=0.006$ & $p=0.0099$)
sediments.

The lakes in the CB are highly impacted by agricultural activity and have cultivated or
pastureland comprising approximately 50-90% of their total watershed. This land use contrasts
with other ecological regions like the NLF (2-24% agriculture) and the CS (<3% agriculture)
where species richness was statistically less rich. Agricultural runoff and drainage can carry a
variety of contaminants (e.g., herbicides, pesticides) which can stimulate the growth of certain
bacteria (e.g., *Planktothrix*) (7,43). The selecting pressure of land use on microbial taxa and food
webs varies within and across ecosystems from highly selective to uninformative, and often
land use is described as having indirect effects on microbial composition and diversity (i.e.,
differing land covers lead to differing nutrient loads in runoff) (22,44,45). Ultimately, our data
confirm patterns observed by others: that eco-regional or eco-zone concepts can affect alpha
diversity and species richness (45). We explored the eco-regional-diversity relationship further
by examining within phyla richness across the four ecoregions (S3 Table). From this we found
15 phyla which were statistically different in terms of richness in one ecoregion. Several of
these phyla (e.g., Bacteroidetes, Spirochaetes and Lentisphaerae) exhibited patterns of richness
across ecological regions like that seen in the entire microbiome (i.e., a decreasing richness
from CB to CS) while only one phylum, Armatimonadetes, showed an opposite richness pattern.
Other phyla like Epsilonbacteraeota and Modulibacteria were more rich in the CB, while
Kiritimatiellaeota were richer in CS samples. The identification of specific phyla that display
varying richness depending on the broader ecological region of their environment may provide
insight into the more nuanced indirect effects that geography plays in microbial assembly. For

example, a shift in Bacteroidetes richness across these ecological regions may be indicative of the changing land use and subsequent nutrient regimes which could lead to increased algal biomass within the lake (as discussed below). However, given the low resolution of 16S rRNA gene sequencing (particularly partial gene amplicons) and an inability to confidently determine unique species and potential functional differences we could not specifically address the mechanisms that lead to increased richness across ecological regions.

Because lakes in these ecological regions also tend to vary based on their trophic status, we compare the differences in alpha diversity based on proxies for lake productivity. Using previously reported (yearly average values) of chlorophyll-a and total phosphorous (Chl-a, TP), and Secchi depth provided by the MN Department of Natural Resources, we classified the lakes as hypereutrophic, eutrophic, mesotrophic, and oligotrophic (S1 Table). We found that both Shannon diversity and observed OTUs were greater in hypereutrophic systems compared to oligotrophic systems, and eutrophic systems were also statistically richer than oligotrophic systems (S3 Fig.). Since most of the lakes were classified as eutrophic or mesotrophic and the values of TP and Chl-a vary seasonally depending on sampling times, we further assessed the effects of local water chemistry, measured at time of sampling, on alpha diversity. Using an exhaustive search with AIC selection criteria we modeled Shannon and Observed diversity using all aqueous chemistry measures, lake latitude, depth, trophic status, ecoregion, and land cover use in the watershed. The lake's latitude as well as the concentrations of total phosphorus (TP), total nitrogen (TN), and dissolved organic carbon (DOC) best predicted Shannon diversity ($\text{adj } R^2 = 0.5713$). Whereas the lake's latitude as well as the concentrations of TP, TN, DOC, and Chl-a best predicted observed diversity ($\text{adj } R^2 = 0.6365$).

The results of our alpha diversity analysis indicate that sediments in more eutrophic systems, like those of the CB, are more diverse. This is in contrast to unimodal diversity-productivity relationships seen among other freshwater communities such as phytoplankton, zooplankton, and fish but similar to other studies on freshwater bacterial communities along a trophic gradient (46–50). Previous work addressing the diversity-productivity relationship of bacterial communities highlights the importance of rare or dormant taxa, in that as trophic status increases the diversity of rare/dormant taxa increases (51). In our study we deemed rare taxa at the phylum level as those not comprising more than 1% relative abundance of the sample. We then compared the richness of these phyla individually across ecological region and trophic status (S3 Table). From this, we found that no rare taxa (at the phylum level) exhibited a statistically significant ($p < 0.01$) increase in richness due to increased trophic status. Nevertheless, there were three common phyla that did increase as a function of trophic status: Chloroflexi, Spirochaetes, and TA06. Among these three phyla only Chloroflexi, which play an important role in the degradation of labile carbon and secretion of organic acids in subsurface sediments, have previously been shown to increase in abundance and diversity with eutrophication in aquatic environments (52,53). By examining the trends in richness across trophic status, like ecological region, we may begin to uncover indicator taxa for over enrichment and eutrophication; however, coarser species and function relationships need to be considered.

Beta Diversity

To further distinguish trends in the data based on ecoregion and other environmental measures, we examined beta diversity (diversity between samples). Using our variance

stabilized data, we conducted a principal component analysis (PCA) to explore the differences in community composition of samples across sites and depths. The first two components explained ~30% of the variation in the samples (Fig 4). In addition to PCA we performed a cluster analysis to determine which samples were most similar (Fig 5). Using both approaches, we concluded that there was a clear distinction between community composition based on ecological region and lake depth.

Fig 4. Beta-Diversity: Principal Component Analysis of Samples

Principal Component Analysis (PCA) of surface sediment microbiome samples where color represents ecological region. Components one and two explained 31.4% of the total variance. Environmental variables were fit using linear regressions where each component was plotted as a function of an environmental vector and those with $p < 0.01$ were plotted. Solid line ellipses are the outer sample bounds for each region and the shaded ellipses are the standard error of the weighted centroids for the data. Abbreviations: Dissolved Inorganic Carbon (DIC), Total Nitrogen (TN), Total Phosphorus (TP), Chlorophyll-a (Chl-a).

Fig 5. Beta-Diversity: Hierarchical Clustering of Samples

Hierarchical clustering analysis of sediment bacterial communities using Ward's D linkages. Clusters reflect the dissimilarities (Bray-Curtis) between variance stabilized 16S rRNA OTUs within each sediment sample where shape indicates depth of samples and color ecological region. Bars along the bottom highlight the first four clusters. These clusters highlight differences in ecological region and depth, where bars 1 & 4 respectively separate Canadian Shield (CS) and Western Cornbelt Plains (CB) samples and bar 2 clusters samples from lakes ~20m or deeper. Bar 3 represents the remaining samples from the Northern Lakes and Forests (NLF) and Northcentral Hardwood Forest (NCHF) regions.

In terms of ecological region, samples from the CS were the first to cluster out. The second cluster consists of six samples from three lakes that are best characterized as deep (max depth > ~20m) meso-eutrophic systems. After depth, the CB samples cluster and the remaining cluster consists of samples from both the NLF and NCHF. Because the most predominant clusters were based on ecological region, we used a non-parametric multivariate analysis of variance (PERMANOVA) with the four regions as our independent variable and tested for

differences in community dissimilarities at the OTU level using the same Bray-Curtis dissimilarity matrix as in the cluster analysis. The results of the PERMANOVA indicate there was an ecoregional difference in composition. To ensure these results reflected distinct groups and not over dispersion within groups, we used vegan's "betadisp" function to test for homogeneity of variance. The resulting insignificant values led us to conclude that community composition is a function of ecoregion and sample sites within these regions are not over dispersed.

Because ecoregion was only a consistent explanatory variable for the more geographically distant ecological regions (CB, CS), we wanted to determine if any of the local physicochemical or watershed land-use factors were also driving community composition. We passively fit these factors to the PCA, treating them as dependent variables explained by the scores from the ordination. Each variable was analyzed separately and added to the plot where the direction of the arrow indicates the gradient direction, and the length indicates the strength of the correlation. After correcting the p-values for multiple comparisons (i.e., Bonferroni), the concentrations for DIC, TP, TN, and Chl-a and lake depth and latitude were the most significant local physicochemical factors differentiating the samples. The four land uses include pastured, cultivated, developed, and forested of which developed, forested, and cultivated lands were the most significant. Forested land use was correlated in the direction of the CS whereas cultivated and urban land-uses were correlated in the opposite direction. Beyond land use, the concentration of DIC was negatively correlated with samples from lakes in the CS. Finally, the cluster of deep lakes was best explained by a combination of lake depth and TP. These are likely related, as the greater depths of these lakes can lead to stronger and prolonged periods of

stratification wherein redox conditions of the sediment change to release phosphorus bound to reducible forms of iron.

The results from our study indicate that beta diversity among lake sediment microbiomes is determined by a combination of land use and productivity. These results are consistent with previous studies examining inter-lake microbiome variability across a variety of spatial scales (22,26,45,50,54). More specifically, our work parallels findings that nitrogen is a selective variable for community composition and that it covaries with urban and agricultural land use (44). Beyond nutrients, lake depth is commonly identified as a partitioning factor for microbial communities, as it was in our system (45). DOC concentrations have also been strongly associated with beta diversity of microbial communities; however, our data do not reflect this trend, potentially highlighting the importance of other micronutrients and abiotic factors for explaining community variation from lake to lake (55). Importantly, while our findings suggest a combination of productivity measures and land use are drivers of lake sediment microbiome structure, the abiotic measures used to assess these relationships were taken from water column measurements. While surface sediments are suitable for estimating site diversity, the specific vertical abiotic properties of sediments may provide deeper context to the microbiome as a whole (20,26,56).

Community Composition

From the 40 samples we recovered 55 unique phyla—22 of which were dominant (~90% of the total sample abundance) (S4 and S5 Figs). In all samples Proteobacteria were the most abundant, comprising approximately 5-20% of the total population of a given sample.

Proteobacteria are often the most abundant phylum in sediment and soil ecosystems given their diverse metabolisms and role in the degradation of organic matter (57–59). At the class level, samples were predominated largely by Deltaproteobacteria and Gammaproteobacteria and there was no clear pattern in their distribution across ecological or trophic gradients. One exception was the presence and/or high abundance of the order MBNT15 (class Deltaproteobacteria) in CS sediments (Fig 6). While there is little known about the ecological significance of MBNT15, these organisms are obligate anaerobes commonly found in stable sediments and known to reduce nitrate; as such, their presence correlates negatively with rates of nitrogen cycling (60,61). Additionally, they have been found to be minor constituents in sediments overburdened with organic matter (62). Given the physicochemical properties of lakes in the CS region (low DOC, low DOM, and TN) our observations of greater MBNT15 abundance in the CS are not unexpected. Additionally, the presence of orders Syntrophobacterales and Desulfobacterales in several of our shallow (0-2cm) samples indicate active sulfur cycling and the depletion of oxygen in the surface sediment, as the genera within these orders are strict anaerobes (63).

Fig 6. Deltaproteobacteria Abundance Across Samples

Abundance comparison of sediment deltaproteobacterial communities where shape indicates depth of samples, color ecological region, and size the relative abundance in percent. Bars along the left group the OTUs by order. OTUs were selected if they comprised >0.05% of the total relative abundance of the sample. The OTUs bolded are mentioned in the text.

Beyond Proteobacteria, other phyla like Acidobacteria, Actinobacteria, Bacteroidetes, Latescibacteria, Nitrospirae, and Spirochaetes exhibit shifts in abundance based on ecological region (S5 Fig). Latescibacteria, for example, show a subtle increase in abundance moving from

the CB to the CS across the NLF and NCHF. The presence of Latescibacteria in all systems is likely due to their ability to degrade several different polymers (proteins, lipids, polysaccharides, fatty acids) and their presence in freshwater sediments supports their proposed role in algal detritus turnover (64). The increase in abundance as a function of ecoregion may be related to the number of stratified lakes within the CS region, especially given the increase in abundance with sediment depth for most lakes.

Other phyla like Acidobacteria and Bacteroidetes display greater changes in abundance across ecological regions (S5 Fig). For example, Acidobacteria comprise <1% of the total population in CB sediments; however, in CS sediments they can make up as much as 6% of the total population. Acidobacteria, like Proteobacteria, are well distributed across environments and metabolically diverse (65). In soil systems these bacteria have been shown to partition based on regional land use. Notably, the order subgroup 6 (phylum Acidobacteria) has previously been observed more frequently in pastured regions compared to forested region; however, in our sediment samples these trends do not exist (66). Subgroup 6 predominates the Acidobacteria population, is found throughout all geographic regions, and is more abundant in White Iron lake – which has one of the most heavily forested watersheds (83%). Bacteroidetes populations exhibit an opposite trend to Acidobacteria – decreased abundance in the CS sediments compared to the CB. Bacteroidetes species specialize in the degradation of organic polymers (e.g., *Bacteroides* in the gut microbiome) (67). Given their proclivity for polymers over monomers, members of Bacteroidetes are often more abundant in aquatic systems during and shortly after algal blooms (68,69). Our data confirm the relationship between Bacteroidetes and

productivity, as lakes in the CB region receive excess nutrient loads and often experience algal blooms during the open water season.

In addition to shifts in composition across ecological regions, there were certain phyla that were not observed at our filtering level (OTUs comprising >0.5% of the total abundance within a sample). Most notably were the absence of Cyanobacteria in CS lake sediments. Land use models have previously been used to partially estimate total cyanobacterial biomass and examine cyanobacterial community structure and genes related to the production of the cyanotoxin microcystin (70,71). Our data show similar findings to these previous works in which the populations of toxin-forming members like those found in the order Nostocales (S6 Fig) were more abundant in heavily agriculture and urban settings where they are well adapted to handle higher nutrient levels in the water (72).

Caveats and Conclusion

The data obtained during this study provide unique insight into the structure, diversity, and distribution of sediment bacteria across both a trophic and ecological region gradient. A primary caveat is the limited explanation of variance (~30%). While we were able to explain a similar amount of the variability to that observed in other studies, there is still a large amount unexplained (21,44,73). This could be in part due to measurements that were not collected. For example, specific parameters like pH and redox potential at the sediment water interface within each lake would provide greater context for the sediment microbiome structure. Our analysis was also limited to abiotic factors as explanatory variables of alpha and beta diversity for bacterial communities. However, biotic interactions exert a selective force on community

structure through a variety of control methods (e.g., grazing, phage infection) (74). Moreover, the horizontal structuring of sediment communities as well as the overall food web dynamics especially given the differing productivity-diversity relationships could be considered in future studies.

In conclusion, we examined the lake sediment bacterial communities of 20 lakes to determine the influence of land-use and large-scale land classifications on community structure and diversity. We observed that ecological region with more agricultural land use and greater eutrophication exhibited higher diversity. Likewise, we found that toxin-forming community members were more abundant in heavily agriculture and urban settings.

Many land managers have access to land use maps, and remote sensing is improving our means of evaluating land use in poorly accessible parts of the world. Our results, along with future studies, offer opportunities to connect land use with sediment microbial structure and ultimately to understand lakes' abilities to adapt to anthropogenic changes.

Acknowledgements

Funding for this project was provided by a grant to AJH from the Minnesota Environment and Natural Resources Trust fund as recommended by Legislative-Citizen Commission on Minnesota Resources (LCCMR). HMS and TLH were also supported by NSF grant #1948058. The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported within this paper. Michele

455 Natarajan, Erin Mortenson, and Alaina Fedie of the St. Croix Watershed Research Station
 456 coordinated laboratory analyses.

457

458

459

460

461

462

463

464

465

466

467 **References**

- 468 1. Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM, Relman DA. The application of ecological
469 theory toward an understanding of the human microbiome. *Science*. 2012 Jun 8;336(6086):1255–
470 62.
- 471 2. Voorhies AA, Mark Ott C, Mehta S, Pierson DL, Crucian BE, Feiveson A, et al. Study of the impact of
472 long-duration space missions at the International Space Station on the astronaut microbiome.
473 *Scientific Reports*. 2019 Jul 9;9(1):9911.
- 474 3. Bakker MG, Manter DK, Sheflin AM, Weir TL, Vivanco JM. Harnessing the rhizosphere microbiome
475 through plant breeding and agricultural management. *Plant Soil*. 2012 Nov 1;360(1):1–13.
- 476 4. Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature*
477 *Reviews Microbiology*. 2017 Oct;15(10):579–90.
- 478 5. Ossowicki A, Raaijmakers JM, Garbeva P. Disentangling soil microbiome functions by perturbation.
479 *Environmental Microbiology Reports* [Internet]. [cited 2021 Aug 20];n/a(n/a). Available from:
480 <https://sfamjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1758-2229.12989>
- 481 6. Comte J, Giorgio PAD. Links between resources, C metabolism and the major components of
482 bacterioplankton community structure across a range of freshwater ecosystems. *Environmental*
483 *Microbiology*. 2009;11(7):1704–16.
- 484 7. Zeglin LH. Stream microbial diversity in response to environmental changes: review and synthesis of
485 existing research. *Frontiers in Microbiology*. 2015;6:454.
- 486 8. Janssen ABG, Hilt S, Kosten S, Klein JJM de, Paerl HW, Waal DBV de. Shifting states, shifting services:
487 Linking regime shifts to changes in ecosystem services of shallow lakes. *Freshwater Biology*.
488 2021;66(1):1–12.
- 489 9. Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S. A Guide to the Natural History of
490 Freshwater Lake Bacteria. *Microbiology and Molecular Biology Reviews*. 2011 Mar 1;75(1):14–49.
- 491 10. Downing JA, Prairie YT, Cole JJ, Duarte CM, Tranvik LJ, Striegl RG, et al. The global abundance and
492 size distribution of lakes, ponds, and impoundments. *Limnology and Oceanography*.
493 2006;51(5):2388–97.
- 494 11. Tranvik LJ, Downing JA, Cotner JB, Loiselle SA, Striegl RG, Ballatore TJ, et al. Lakes and reservoirs as
495 regulators of carbon cycling and climate. *Limnology and Oceanography*. 2009;54(6part2):2298–314.
- 496 12. Harrison JA, Maranger RJ, Alexander RB, Giblin AE, Jacinthe P-A, Mayorga E, et al. The regional and
497 global significance of nitrogen removal in lakes and reservoirs. *Biogeochemistry*. 2009 Mar
498 1;93(1):143–57.
- 499 13. Knoll LB, Vanni MJ, Renwick WH, Kollie S. Burial rates and stoichiometry of sedimentary carbon,
500 nitrogen and phosphorus in Midwestern US reservoirs. *Freshwater Biology*. 2014;59(11):2342–53.
- 501 14. Heathcote AJ, Anderson NJ, Prairie YT, Engstrom DR, del Giorgio PA. Large increases in carbon burial
502 in northern lakes during the Anthropocene. *Nature Communications*. 2015 Nov 26;6(1):10016.

15. Anderson NJ, Heathcote AJ, Engstrom DR, Globocarb data contributors. Anthropogenic alteration of nutrient supply increases the global freshwater carbon sink. *Sci Adv.* 2020 Apr;6(16):eaaw2145.
16. Haglund A-L, Lantz P, Törnblom E, Tranvik L. Depth distribution of active bacteria and bacterial activity in lake sediment. *FEMS Microbiology Ecology.* 2003 Oct 1;46(1):31–8.
17. Nealson KH. SEDIMENT BACTERIA: Who's There, What Are They Doing, and What's New? Annual Review of Earth and Planetary Sciences. 1997;25(1):403–34.
18. Frindte K, Allgaier M, Grossart H-P, Eckert W. Microbial Response to Experimentally Controlled Redox Transitions at the Sediment Water Interface. Vopel KC, editor. *PLoS ONE.* 2015 Nov 24;10(11):e0143428.
19. Xiong W, Xie P, Wang S, Niu Y, Yang X, Chen W. Sources of organic matter affect depth-related microbial community composition in sediments of Lake Erhai, Southwest China. *J Limnol* [Internet]. 2014 Nov 24 [cited 2021 Aug 24];73(AoP). Available from: <http://www.jlimnol.it/index.php/jlimnol/article/view/jlimnol.2014.1106>
20. Wurzbacher C, Fuchs A, Attermeyer K, Frindte K, Grossart H-P, Hupfer M, et al. Shifts among Eukaryota, Bacteria, and Archaea define the vertical organization of a lake sediment. *Microbiome.* 2017 Dec;5(1):41.
21. Niño-García JP, Ruiz-González C, del Giorgio PA. Interactions between hydrology and water chemistry shape bacterioplankton biogeography across boreal freshwater networks. *ISME J.* 2016 Jul;10(7):1755–66.
22. Chen W, Wilkes G, Khan IUH, Pintar KDM, Thomas JL, Lévesque CA, et al. Aquatic Bacterial Communities Associated With Land Use and Environmental Factors in Agricultural Landscapes Using a Metabarcoding Approach. *Front Microbiol* [Internet]. 2018 [cited 2021 Aug 10];0. Available from: <https://www.frontiersin.org/articles/10.3389/fmicb.2018.02301/full>
23. Bertilsson S, Burgin A, Carey CC, Fey SB, Grossart H-P, Grubisic LM, et al. The under-ice microbiome of seasonally frozen lakes. *Limnology and Oceanography.* 2013;58(6):1998–2012.
24. Bienhold C, Zinger L, Boetius A, Ramette A. Diversity and Biogeography of Bathyal and Abyssal Seafloor Bacteria. *PLOS ONE.* 2016 Jan 27;11(1):e0148016.
25. King AJ, Freeman KR, McCormick KF, Lynch RC, Lozupone C, Knight R, et al. Biogeography and habitat modelling of high-alpine bacteria. *Nat Commun.* 2010 Aug 10;1(1):53.
26. Wobus A, Bleul C, Maassen S, Scheerer C, Schuppler M, Jacobs E, et al. Microbial diversity and functional characterization of sediments from reservoirs of different trophic state. *FEMS Microbiology Ecology.* 2003 Dec 1;46(3):331–47.
27. American public health association, American water works association, Water environment federation, editors. Standard methods for the examination of water and wastewater. 22nd ed. Washington (D.C.): American public health association; 2012.

538 28. Arar EJ, Collins GB. Method 445.0 In Vitro Determination of Chlorophyll a and Pheophytin a in
539 Marine and Freshwater Algae by Fluorescence. U.S. Environmental Protection Agency; 1997.

540 29. Wright HE. Coring tips. *J Paleolimnol*. 1991 Jan;6(1):37–49.

541 30. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur:
542 Open-Source, Platform-Independent, Community-Supported Software for Describing and
543 Comparing Microbial Communities. *Applied and Environmental Microbiology*. 2009 Dec
544 1;75(23):7537–41.

545 31. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index
546 sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq
547 Illumina sequencing platform. *Appl Environ Microbiol*. 2013 Sep;79(17):5112–20.

548 32. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene
549 database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013
550 Jan;41(Database issue):D590–6.

551 33. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of
552 chimera detection. *Bioinformatics*. 2011 Aug 15;27(16):2194–200.

553 34. Stackebrandt E, Goebel BM. Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA
554 Sequence Analysis in the Present Species Definition in Bacteriology. *International Journal of*
555 *Systematic and Evolutionary Microbiology*. 1994 Oct 1;44(4):846–9.

556 35. Glassman SI, Martiny JBH. BROADSCALE Ecological Patterns Are Robust to Use of Exact Sequence
557 Variants versus Operational Taxonomic Units. Tringe SG, editor. *mSphere* [Internet]. 2018 Aug 29
558 [cited 2021 Sep 13];3(4). Available from: <https://journals.asm.org/doi/10.1128/mSphere.00148-18>

559 36. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R
560 Foundation for Statistical Computing; 2018. Available from: <https://www.R-project.org/>

561 37. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse.
562 *JOSS*. 2019 Nov 21;4(43):1686.

563 38. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics
564 of Microbiome Census Data. *PLOS ONE*. 2013 Apr 22;8(4):e61217.

565 39. Groemping U. Relative Importance for Linear Regression in R: The Package relaimpo. *Journal of*
566 *Statistical Software*. 2006 Sep 1;17(1):1–27.

567 40. Oksanen J, Kindt R, Legendre P, Hara B, Simpson G, Solymos P, et al. The vegan Package. 2009 Jan
568 19;

569 41. Love M, Ahlmann-Eltze C, Forbes K, Anders S, Huber W. DESeq2: Differential gene expression
570 analysis based on the negative binomial distribution [Internet]. Bioconductor version: Release
571 (3.12); 2020 [cited 2020 Nov 25]. Available from: <https://bioconductor.org/packages/DESeq2/>

572 42. Badri M, Kurtz ZD, Müller CL, Bonneau R. Normalization methods for microbial abundance data
573 strongly affect correlation estimates. :21.

574 43. Saxton MA, Morrow EA, Bourbonniere RA, Wilhelm SW. Glyphosate influence on phytoplankton
575 community structure in Lake Erie. *Journal of Great Lakes Research*. 2011 Dec;37(4):683–90.

576 44. Marmen S, Blank L, Al-Ashhab A, Malik A, Ganzert L, Lalzar M, et al. The Role of Land Use Types and
577 Water Chemical Properties in Structuring the Microbiomes of a Connected Lake System. *Frontiers in*
578 *Microbiology*. 2020;11:89.

579 45. Kraemer SA, Barbosa da Costa N, Shapiro BJ, Fradette M, Huot Y, Walsh DA. A large-scale
580 assessment of lakes reveals a pervasive signal of land use on bacterial communities. *ISME J*. 2020
581 Dec;14(12):3011–23.

582 46. Vallina SM, Follows MJ, Dutkiewicz S, Montoya JM, Cermeno P, Loreau M. Global relationship
583 between phytoplankton diversity and productivity in the ocean. *Nat Commun*. 2014 Jul 1;5(1):4299.

584 47. Barnett A, Beisner BE. Zooplankton Biodiversity and Lake Trophic State: Explanations Invoking
585 Resource Abundance and Distribution. *Ecology*. 2007;88(7):1675–86.

586 48. Vonlanthen P, Bittner D, Hudson AG, Young KA, Müller R, Lundsgaard-Hansen B, et al.
587 Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature*. 2012
588 Feb;482(7385):357–62.

589 49. Bell T, Newman JA, Silverman BW, Turner SL, Lilley AK. The contribution of species richness and
590 composition to bacterial services. *Nature*. 2005 Aug;436(7054):1157–60.

591 50. Kiersztyn B, Chróst R, Kaliński T, Siuda W, Bukowska A, Kowalczyk G, et al. Structural and functional
592 microbial diversity along a eutrophication gradient of interconnected lakes undergoing
593 anthropopressure. *Sci Rep*. 2019 Jul 31;9(1):11144.

594 51. Jankowski K, Schindler DE, Horner-Devine MC. Resource Availability and Spatial Heterogeneity
595 Control Bacterial Community Response to Nutrient Enrichment in Lakes. *PLOS ONE*. 2014 Jan
596 28;9(1):e86991.

597 52. Zeng J, Jiao C, Zhao D, Xu H, Huang R, Cao X, et al. Patterns and assembly processes of planktonic
598 and sedimentary bacterial community differ along a trophic gradient in freshwater lakes. *Ecological*
599 *Indicators*. 2019 Nov 1;106:105491.

600 53. Huang W, Chen X, Jiang X, Zheng B. Characterization of sediment bacterial communities in plain
601 lakes with different trophic statuses. *MicrobiologyOpen*. 2017;6(5):e00503.

602 54. Xing P, Hahn MW, Wu QL. Low Taxon Richness of Bacterioplankton in High-Altitude Lakes of the
603 Eastern Tibetan Plateau, with a Predominance of Bacteroidetes and Synechococcus spp. *Applied and*
604 *Environmental Microbiology*. 2009 Nov 15;75(22):7017–25.

605 55. Muscarella ME, Boot CM, Broeckling CD, Lennon JT. Resource heterogeneity structures aquatic
606 bacterial communities. *ISME J*. 2019 Sep;13(9):2183–95.

- 607 56. Pla-Rabes S, Flower RJ, Shilland EM, Kreiser AM. Assessing microbial diversity using recent lake
608 sediments and estimations of spatio-temporal diversity. *Journal of Biogeography*.
609 2011;38(10):2033–40.
- 610 57. Spain AM, Krumholz LR, Elshahed MS. Abundance, composition, diversity and novelty of soil
611 Proteobacteria. *ISME J*. 2009 Aug;3(8):992–1000.
- 612 58. Zhang J, Yang Y, Zhao L, Li Y, Xie S, Liu Y. Distribution of sediment bacterial and archaeal
613 communities in plateau freshwater lakes. *Appl Microbiol Biotechnol*. 2015 Apr 1;99(7):3291–302.
- 614 59. Kersters K, De Vos P, Gillis M, Swings J, Vandamme P, Stackebrandt E. Introduction to the
615 Proteobacteria. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, editors. *The*
616 *Prokaryotes: Volume 5: Proteobacteria: Alpha and Beta Subclasses* [Internet]. New York, NY:
617 Springer; 2006 [cited 2021 Aug 26]. p. 3–37. Available from: [https://doi.org/10.1007/0-387-30745-](https://doi.org/10.1007/0-387-30745-1_1)
618 [1_1](https://doi.org/10.1007/0-387-30745-1_1)
- 619 60. Repert DA, Underwood JC, Smith RL, Song B. Nitrogen cycling processes and microbial community
620 composition in bed sediments in the Yukon River at Pilot Station. *Journal of Geophysical Research:*
621 *Biogeosciences*. 2014;119(12):2328–44.
- 622 61. Chen Y-J, Leung PM, Wood JL, Bay SK, Hugenholtz P, Kessler AJ, et al. Metabolic flexibility allows
623 bacterial habitat generalists to become dominant in a frequently disturbed ecosystem. *ISME J*. 2021
624 May 3;1–19.
- 625 62. Hori T, Kimura M, Aoyagi T, R. Navarro R, Ogata A, Sakoda A, et al. Biodegradation Potential of
626 Organically Enriched Sediments under Sulfate- and Iron-Reducing Conditions as Revealed by the 16S
627 rRNA Deep Sequencing. *J of Wat ^|^ Envir Tech*. 2014;12(4):357–66.
- 628 63. Whitman WB, Rainey F, Kämpfer P, Trujillo M, Chun J, DeVos P, et al., editors. *Bergey’s Manual of*
629 *Systematics of Archaea and Bacteria* [Internet]. 1st ed. Wiley; 2015 [cited 2021 Aug 26]. Available
630 from: <https://onlinelibrary.wiley.com/doi/book/10.1002/9781118960608>
- 631 64. Farag IF, Youssef NH, Elshahed MS. Global Distribution Patterns and Pangenomic Diversity of the
632 Candidate Phylum “Latescibacteria” (WS3). *Appl Environ Microbiol* [Internet]. 2017 May 15 [cited
633 2021 Mar 11];83(10). Available from: <https://aem.asm.org/content/83/10/e00521-17>
- 634 65. Kielak AM, Barreto CC, Kowalchuk GA, van Veen JA, Kuramae EE. The Ecology of Acidobacteria:
635 Moving beyond Genes and Genomes. *Front Microbiol* [Internet]. 2016 May 31 [cited 2021 Aug
636 27];7. Available from: <http://journal.frontiersin.org/Article/10.3389/fmicb.2016.00744/abstract>
- 637 66. Navarrete AA, Venturini AM, Meyer KM, Klein AM, Tiedje JM, Bohannon BJM, et al. Differential
638 Response of Acidobacteria Subgroups to Forest-to-Pasture Conversion and Their Biogeographic
639 Patterns in the Western Brazilian Amazon. *Front Microbiol*. 2015 Dec 22;6:1443.
- 640 67. Karlsson FH, Ussery DW, Nielsen J, Nookaew I. A Closer Look at Bacteroides: Phylogenetic
641 Relationship and Genomic Implications of a Life in the Human Gut. *Microb Ecol*. 2011 Apr
642 1;61(3):473–85.

68. Cottrell MT, Kirchman DL. Natural Assemblages of Marine Proteobacteria and Members of the *Cytophaga-Flavobacter* Cluster Consuming Low- and High-Molecular-Weight Dissolved Organic Matter. *Appl Environ Microbiol*. 2000 Apr;66(4):1692–7.
69. Pinhassi J, Sala MM, Havskum H, Peters F, Guadayol Ò, Malits A, et al. Changes in Bacterioplankton Composition under Different Phytoplankton Regimens. *Applied and Environmental Microbiology*. 2004 Nov 1;70(11):6753–66.
70. Catherine A, Troussellier M, Bernard C. Design and application of a stratified sampling strategy to study the regional distribution of cyanobacteria (Ile-de-France, France). *Water Research*. 2008 Dec;42(20):4989–5001.
71. Marmen S, Aharonovich D, Grossowicz M, Blank L, Yacobi YZ, Sher DJ. Distribution and Habitat Specificity of Potentially-Toxic Microcystis across Climate, Land, and Water Use Gradients. *Front Microbiol* [Internet]. 2016 Mar 15 [cited 2021 Aug 27];7. Available from: <http://journal.frontiersin.org/Article/10.3389/fmicb.2016.00271/abstract>
72. Cirés S, Ballot A. A review of the phylogeny, ecology and toxin production of bloom-forming *Aphanizomenon* spp. and related species within the Nostocales (cyanobacteria). *Harmful Algae*. 2016 Apr;54:21–43.
73. Lear G, Washington V, Neale M, Case B, Buckley H, Lewis G. The biogeography of stream bacteria. *Global Ecology and Biogeography*. 2013;22(5):544–54.
74. Louca S, Parfrey LW, Doebeli M. Decoupling function and taxonomy in the global ocean microbiome. *Science*. 2016 Sep 16;353(6305):1272–7.

Supporting Information

Sequencing Data Available

S1 Table. Aqueous Chemistry Data.

S2 Table. Alpha Diversity Measures by Sample. Values for all alpha diversity measure calculated by sample.

S3 Table. Kruskal Wallis P-values for Phyla Richness Across Ecological Regions and Trophic Status. Table of Kruskal Wallis p-values for individual phyla within the alpha diversity dataset (rarefied). P-values for both the significance based on ecological region and trophic status are reported and bolded values are significant at $p=0.001$. Dunn post hoc test with Bonferroni correction p-values are reported for those KW $p<0.001$. Phyla below the bolded line are phyla that comprise of $<1\%$ of the total relative abundance of a given sample (see S4 Fig).

S1 Fig. Batch Effect on Richness. Observed richness (total number of OTUs) based on the total number of reads recovered per sample where color indicates the sequencing batch. There was a statistically significant Pearson's correlation between the number of total reads and the observed richness; $p < 0.001$ and $R^2=0.63$.

S2 Fig. Alpha Diversity Measures by Sample. Alpha diversity measures across samples, where shape indicates depth of sample and color indicates ecological regions. All measures were calculated using Phyloseq and exhibit similar patterns in diversity; decreasing diversity across a northeasterly transect.

S3 Fig. Bacterial Alpha Diversity by Trophic Status. Box plots show mean alpha level diversity of the observed Operational Taxonomic Units (OTUs) and Shannon indices of the four trophic status classifications within the study area: Oligotrophic, Mesotrophic, Eutrophic, and Hypereutrophic. Each point represents a given sample where shape indicates depth of sample and color indicates the trophic status. One sample with extremely low richness and diversity was removed from both plots for visualization. Significance between regions was calculated nonparametrically using a Kruskal Wallis H test followed by a Dunn post hoc test with a Bonferroni correction. Reported p-values indicate significant differences in Observed and Shannon diversity, respectively, across trophic status, specifically the diversity of Oligotrophic lake sediments when compared to Eutrophic ($p=0.0224$) and Hypereutrophic ($p=0.0038$ & $p=0.013$) sediments

S4 Fig. Bacterial Relative Abundance. Bar plots of phyla that comprise $>1\%$ of the total relative abundance of a given sample. Samples are sorted along the X axis by ecological region.

S5 Fig. Bacterial Abundance at the Phylum Level Across Ecological Regions. Box plots show mean relative abundance for the phyla across ecological region. Each point is a sample. Abbreviations: Western Cornbelt Plains (CB), North Central Hardwood Forests (NCHF), Northern Lakes and Forests (NLF), and Canadian Shield (CS).

S6. Cyanobacterial Abundance Across Samples

Abundance comparison of sediment cyanobacterial communities where shape indicates depth of samples, color ecological region, and size the relative abundance in percent. Bars along the left group the OTUs by order. OTUs were selected if they comprised $>0.01\%$ of the total relative abundance of the sample.

Figure

Figure

Figure

Figure

Figure

Figure