Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI) Vol. 28:1-13

Microbial Diversity and Population Density is Positively Correlated in New York State Freshwater Wetlands

Tanya Balaji, Azraa Banka, Daven Kuder, Raveen Sidhu

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY Increasing anthropogenic activities, particularly those associated with urbanization, have increased ecological strain on near-urban freshwater wetlands. Despite wetlands being ecologically significant environments with key functions like habitat preservation, water quality enhancement, nutrient cycling properties, and regulatory roles in maintaining greenhouse gas budgets, research characterizing the microbial diversity of wetland soils and the impact of urbanization on these communities is not well established. To address this gap, this paper aims to investigate the impact of proximity to population centers on the microbial community and environmental conditions of wetlands. To explore the relationships between human population density and freshwater wetland microbial environments, alpha and beta diversity metrics between urban and rural sites in the state of New York were assessed. Moreover, differential abundance of microbes between sites were quantified, and a correlation analysis between pH, nitrogen content and soil saturation was conducted. We show that richness and evenness are significantly higher in the urban sites compared to the rural sites. In addition, differential analysis revealed an enrichment of taxa such as Spirochaetales and Methanosarcinales at urban and rural sites, respectively. Lastly a positive correlation was observed with soil saturation and pH across genera *Aidingimonas*, *Flavobacterium*, and *Sphingomonas* in Owego, a rural site. These findings are inconsistent with our hypothesis and demonstrate that microbial diversity in urban sites increases compared to rural sites which permits the selection of various taxa. Ultimately, this paper highlights the intricate nature of microbial wetland systems, emphasizing the need for further research into the impact of increased urbanization and population density on native wetland microbial communities and their functionality at an ecological scale.

INTRODUCTION

Wetlands are defined as soil areas covered by water all-year or for varying periods of time during the year (https://www.epa.gov/wetlands/what-wetland). Wetlands are ecologically significant environments and play important roles in protecting and improving water quality, providing habitat to aquatic and terrestrial wildlife, and act as a buffer against storms and other natural catastrophes that could disrupt coastlines (1). Wetlands play an influential ecological role because of their high productivity, their nutrient recycling properties, and their role in the regulation of global greenhouse gas budgets (2).

Many of the qualities of wetlands stem from their dependence on soil microbial communities, which play a notable role in regulating the cycling, retention, and release of major nutrients and carbon in freshwater wetlands, while also influencing water quality and global carbon cycling. Wetland plant roots often create oxic-anoxic conditions, which allows for the simultaneous activity of both aerobic and anaerobic microbial communities (2). There are various environmental factors that contribute to microbial community structure in wetlands, such as soil salinity, pH, temperature, organic carbon (CH_4) content, inorganic carbon (CO_2) content, ion concentrations, and nutrient concentration (3). In this study, we focused on the following environmental variables: soil nitrate and ammonia concentrations, soil saturation, nitrogen in microbial biomass, and pH and investigated how these factors change with population density.

Previous studies have shown that nitrate and nitrite concentrations are important as they serve as major substrates for nitrifying and denitrifying bacteria, which play important biogeochemical cycling roles in wetland ecosystems, especially when experiencing periodic

Published Online: September 2023

Citation: Balaji, Banka, Kuder, Sidhu. 2023. Microbial Diversity and Population Density is Positively Correlated in New York State Freshwater Wetlands. UJEMI 28:1-13

Editor: Lauren Pugsley and Shruti Sandilya, University of British Columbia

Address correspondence to: https://jemi.microbiology.ubc.ca/

flooding and drying (4). In general, the addition of soil nitrogen indirectly affects soil microbial functions by altering carbon turnover rates and microbial growth rates more generally (5). Nitrogen availability enhances other soil carbon processes that determine what bacterial genera dominate the soil environment. However, not enough research has been conducted to find explicit correlations between microbial diversity and nitrogen in microbial biomass (5). Regarding soil saturation, this cannot be easily correlated to increases or decreases in microbial diversity (6). This is true because moisture content has both indirect and direct impacts on other variables, such as ion concentrations and nutrient concentrations. This does not necessarily impact diversity, however, it may produce changes in which bacterial groups dominate a specific environment (6). Soil pH has also been found to be a major variable controlling microbial community structure. Previous studies have shown that different pHs can select for different groups of bacteria to dominate, which may determine differences in microbial community structure between rural and urban sites (7).

In wetlands, dominant bacterial taxa typically include Chloroflexi, Proteobacteria (Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, and Deltaproteobacteria), Actinobacteria, Acidobacteria, Firmicutes, and Cyanobacteria (8). Generally, Acidobacteria have been found to be the most abundant phylum in freshwater wetland soils, followed by Alphaproteobacteria, Actinobacteria, Betaproteobacteria, and Chloroflexi (Green Non-Sulfur Bacteria) (8). Some research has suggested that there is decreased diversity in less disturbed wetlands (rural wetlands), as the abundance of some phyla (eg. Acidobacteria*)* increases relative to the other phyla (8). However, the diversity and functionality of microbial communities in wetland systems is highly unexplored and the factors that contribute to variance in microbial composition, diversity, and functionality still require more research to draw more conclusive relationships between variables (2).

Soils near urban systems tend to have higher nutrient and pollutant content which microbial communities are extremely sensitive to (9). The scale of difference will depend on the mixing and diffusion rates of the wetlands being sampled; higher mixing and diffusion would result in a more homogenous environment further from a chemical source, and a subsequently more homogenous microbial community (9, 10). Previous studies have had contrasting results when assessing the relationship between proximity to population dense areas and microbial community structure. Some studies have suggested that environments with higher heterogeneity (urban areas) can promote the dispersal of bacterial species and the diversification of bacterial community functionality, which would lead to higher diversity (11). However, other studies have shown that urbanization increases habitat heterogeneity, and have shown to result in lower levels of diversity due to extensive selective pressures (11). Nevertheless, the association between population density and microbial community structure still remains unclear.

With this knowledge gap in mind, we wanted to determine the impact of population density on microbial community composition and diversity in samples collected from both rural and urban wetland samples. Prior to analyzing our dataset, we hypothesized that (i) microbial community composition and diversity would be higher in rural wetland samples, compared to urban samples, and (ii) microbial functional diversity would be affected by specific environmental variables, with some being more influential than others in altering community structure (eg. pH, nitrogen content). We found that microbial community composition and diversity between both rural sites and between both urban sites that we chose were similar. However, surprisingly, we found that there were greater differences in microbial community composition and diversity in urban wetland samples closer to population dense areas.

METHODS AND MATERIALS

Dataset and metadata. The original dataset was compiled by Ballantine *et al*, but no research article was published with this dataset. The study's metadata file contained information on several environmental variables such as latitude, longitude, elevation, exchange capacity, conductivity, nitrate and ammonium levels, carbon content, CH_4 flux, CO_2 flux, microbial biomass, pH, etc. We focused particularly on the metadata categories of pH, nitrogen content (nitrate and ammonium), nitrogen in microbial biomass, and soil saturation (moisture).

Metadata filtering and Binning. The first step was to remove any sample locations that had "Missing" or "Not Provided" data values for any of our variables of interest (soil nitrogen concentrations, soil saturation, nitrogen in microbial biomass, and pH). The resulting filtered data locations were then mapped and compared to New York State population data (2021) from the US Census Bureau to determine binning criteria (12). Four binning locations were chosen based on population density and data clumping (Fig 1). Two urban bins were created by binning data within a 25km radius of Auburn and Ithaca respectively; these sites were chosen as both Auburn and Ithaca have a population density more than 3000 people/km^2. Two rural bins were created by grouping data within a 30km radius of Owego and Cuyler. These locations had a population density less than 700 people/km^2. We used a larger radius for rural bins to maintain an even sampling depth between all binning locations. After filtering, 49 sites remained for downstream analysis compared to the original 192 sites sampled.

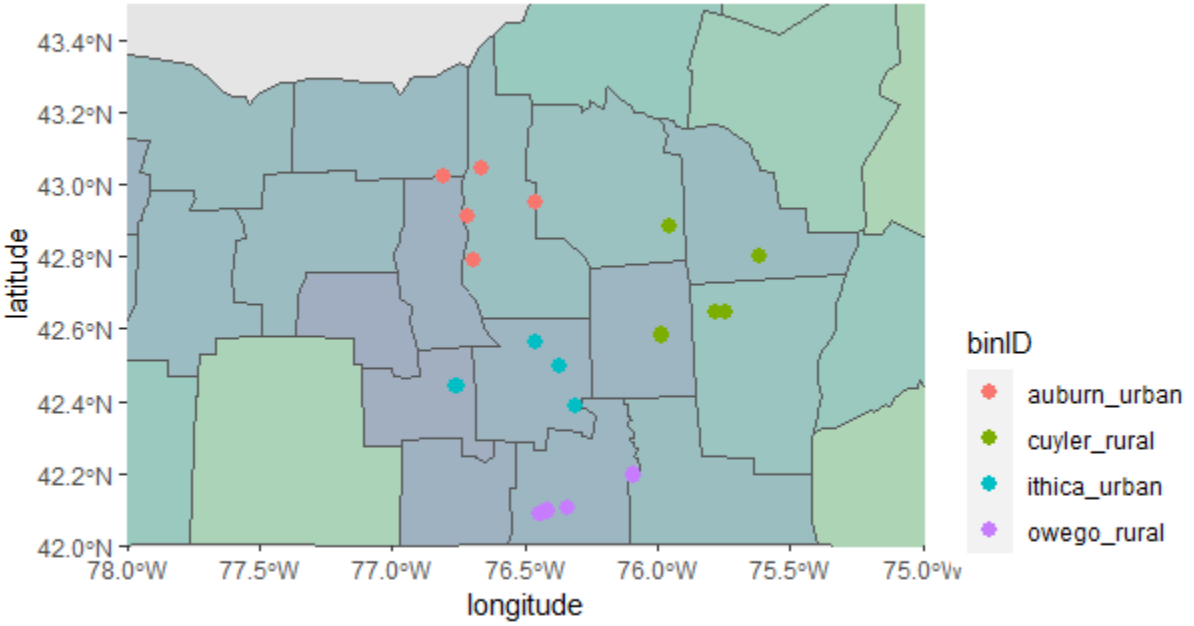

FIG. 1 Binning sites in New York Wetlands. Binning locations are grouped by color and circle. After determining population centers, sample sites located within a 25km radius of population centers were binned together (Auburn and Ithaca). Sample sites within a 30km radius of designated urban areas were grouped together (Owego and Cuyler). Red represents sample sites which were placed in the auburn_urban bin. Green represents sample sites placed in the cuyler_rural bin. Blue represents sample sites placed in the ithaca_urban bin. Purple represents sample sites placed in the owego_rural bin.

Data processing using the QIIME2 pipeline & R Studio. We utilized the QIIME2 software to import and demultiplex the 16S rRNA single-ended sequences from the wetlands dataset (13). We then conducted sequence quality control using Divisive Amplicon Denoising Algorithm 2 (DADA2), and chose a truncation length of 125 nucleotides (14). This length was chosen because it was the maximum sequence length that could be applied while retaining sufficient sequence quality. We then generated a feature table of sequences that further went taxonomy-based filtering to remove mitochondrial and chloroplast sequences from our dataset. The final step we conducted in QIIME2 was the generation of phylogenetic trees, which we then imported into R, along with our feature sequence table, our taxonomy file, and our metadata file (13). In R, we generated an alpha rarefaction plot using a sampling depth of 250,000 which was chosen to sufficiently capture our sample richness (15). This value was chosen to be the minimum sequencing depth at which all samples showed optimal richness, while retaining enough samples for each rural and urban site that we analyzed. These steps are outlined in Script #0 (see Supplemental 3).

Alpha Diversity analysis. RStudio was used to conduct alpha diversity analysis on our filtered datasets generated in QIIME2 (13, 15). Using our filtered feature table, Shannon, Chao1, ACE, Simpson, InvSimpson, and Fisher diversity metrics to assess whether the results were significant between Auburn and Cuyler, and Ithaca and Owego.

Beta Diversity analysis. RStudio was used to produce beta diversity metrics based on the binID column for the filtered dataset (15). Bray-Curtis, Jaccard, and Unifrac diversity statistics were used to compare samples based on binID. Adonis2 was then used to measure the statistical significance of the results comparing Auburn and Cuyler, and Ithaca and Owego (16). When using Adonis2, we also set the total number of permutations to 10,000 (16). RStudio steps for this step are outlined in Script #1R (see Supplemental 4).

Classify differential relative abundance through taxonomic analysis. Bioconductor packages 'Phyloseq' and 'DESeq2' were used in RStudio to analyze the differential abundance of microbial communities between urban and rural site pairings (17, 18). A phyloseq object was created with the generated taxonomic information, OTU table, phylogenetic tree, and provided metadata of samples of interest; OTU values were then transformed using a pseudocount of one (18). This object was then filtered to provide isolated comparisons between rural and urban site pairings (urban site Auburn and rural site Cuyler; urban site Ithaca and rural site Owego). Filtered sites were subjected to a Wald test in DESeq2 for an analysis of how microbial abundance changes between samples (17). The results were refined by filtering for significant microbial abundance changes ($p < 0.001$) which exhibited a log2fold change cutoff of 2 between sample pairings. Comparisons between the Auburn and Cuyler, and Ithaca and Owego, were then conducted at the Order level and plotted in R (17). These steps are outlined in Script #1R.

Correlation analysis between variables of interest (NO_3 and NH_4 soil content, N in biomass, pH, and soil moisture) and microbial abundance regression. A correlation analysis between various metadata categories (NO_3, NH_4, N in microbial biomass, pH, and soil moisture) and significant orders was conducted using RStudio (15). Changes in metadata were graphed against microbial abundance in each of the four binnings sites. Orders were chosen based on differential expression from the DESeq2 analysis. A p-value of 0.05 was used to determine significant correlations, and anything below this threshold was designated as significant. Graphs were produced using ggplot2 in RStudio (15).

RESULTS

Urban sites show significantly higher alpha diversity than rural sites in wetland samples. To measure significance between the sample locations, a Kruskal-Wallis test using the Chao1, ACE, and indices was conducted for each binning location (Auburn, Ithaca, Cuyler, and Owego). When conducting our alpha diversity analysis (Fig 2), Ithaca showed significantly higher p-values for Chao1, Shannon, and ACE metrics compared to Owego (with 0.010, 0.014, and 0.011 values respectively). When comparisons were made between Auburn and Cuyler, no significant results were found for Chao1, Shannon, and ACE (p-values of 0.17, 0.064, and 0.017). This was contrary to our hypothesis and we were surprised to find that this significance was only observed in a single comparison between Ithaca and Owego, and not between Auburn and Cuyler.

Beta diversity is insignificant between urban and rural wetland sites. With respect to beta diversity, no significant clustering patterns were observed for either of the jaccard, bray-curtis, and unifrac metrics. We generated a PCoA plot using the bray-curtis metric, but found no distinct clustering patterns to be observed (see Supplemental 5).

Consistency of Nitrogen content across binning sites depends on N Speciation. The distribution of nitrogen species was much more site-specific, and there were general trends between both rural and both urban sites. Nitrate was found to be elevated in rural binning locations, with $[NO_3]_{ave}$=0.71 and =1.09ugN/gDry soil for Cuyler and Owego, respectively (Fig 3A). The rural locations had $[NO_3]_{ave}$=0.46 and =0.30 ugN/gDry soil for Auburn and Ithaca, respectively (Fig 3A). $[NO_3]$ were found to significantly correlate with binning location (p-value = 0.005). The ammonium concentrations for both rural sites were fairly

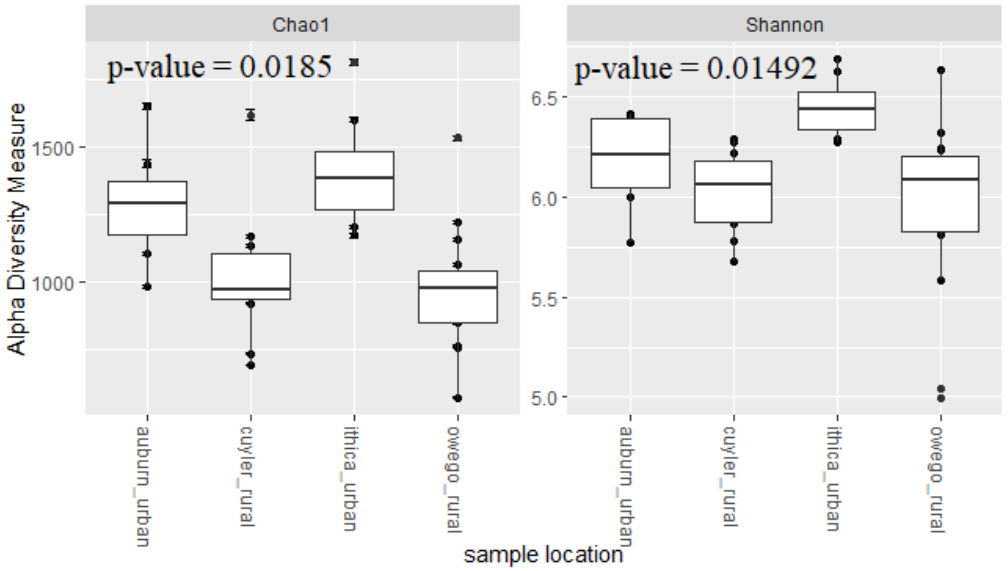

FIG. 2 Alpha Diversity is observed to be significant between rural and wetland samples, as indicated by the Chao1 and Shannon diversity metrics. Chao1 and Shannon alpha diversity metrics for all four binning sites (auburn_urban, cuyler_rural, ithaca_urban, and owego_rural). Both Chao1 and Shannon metrics were significant between binning locations. There is a distinct difference in alpha diversity between rural and urban sites. A p-value threshold of <0.05 was used to determine significance. Chao1 had a p-value = 0.011. Shannon diversity had a p-value = 0.015.

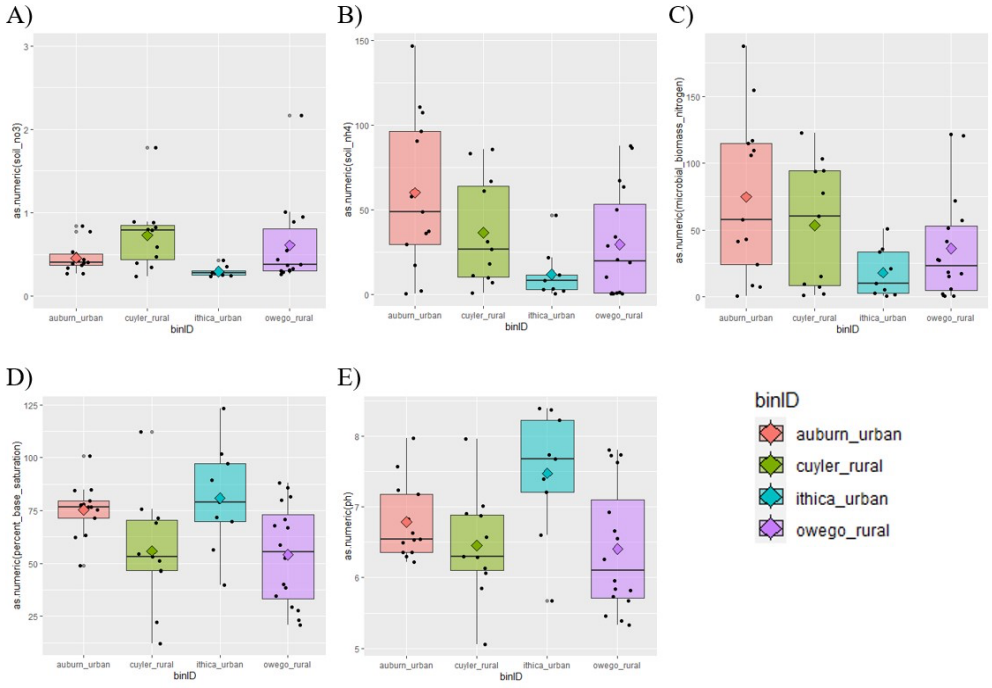

FIG. 3 Environmental variables variables measured for each location show differences in NO3, pH, and base saturation. The distribution of environmental variables within the four bins (auburn_urban, cuyler_rural, ithaca_urban, and owego_rural). Only pH and [NO_3] were found to be significantly correlated with binning location. Red box plots correspond with Auburn data points, green box plots correspond with Cuyler data points, blue box plots correspond with Ithaca data points, and purple corresponds to Owego data points. The diamonds of corresponding color represent the average measurement for each variable at each binning site. A) shows the concentration of NO_3 in soil samples in (ugN/g dry soil) across binning locations. B) shows the soil concentration of NH4 in (ugN/g dry soil) across binning sites. C) shows the amount of nitrogen in microbial biomass (ugN/g dry soil). D) shows the percent base saturation across sample bins. E) Shows the soil pH across sample bins.

constant with $[NH_4]_{ave}$=36.4 and =29.4 ugN/gDry soil for Cuyler and Owego respectively (Fig 3B). The urban locations had widely different ammonium concentrations, likely due to runoff accumulation in Auburn, which has a much lower elevation than the other binning locations (Fig 3B). The average ammonium concentration in Auburn was found to be 60.1 ugN/gDry soil, whereas the average ammonium concentration in Ithaca was only 11.8 ugN/gDry soil. Ammonium concentrations did not significantly correlate with binning locations (p-value = 0.30). This trend continued with Auburn having the highest amount of nitrogen found in biomass (at 74.8 ugN/gDry) and Ithaca having the lowest N content held in biomass (at 18.0ug/gDry). Both rural binning sites had relatively similar N content in biomass, with 53.4 ugN/gDry found in Cuyler biomass and 36.2 ugN/gDry soil found in Owego biomass (Fig 3C). Nitrogen in biomass did not significantly correlate with binning locations (p-value = 0.36). Significance values were calculated using a Kriskal Wallis test and a significance p-value <0.05.

Soil pH and Saturation tightly linked to Rural and Urban binning. Sampling locations had soil pH ranging from 5-8.5 and, similar to observed N-species distribution, rural sites had conserved pH levels while urban pH was site-specific. Rural binning sites had pH_{ave}=6.45 and =6.40 for Cuyler and Owego respectively (Fig 3E). The urban locations had less variation in pH measurements between samples than rural locations, but there was no consistency in average pH for urban bins as a whole. pH_{ave} for Auburn was 6.79, which was relatively consistent with rural pH measurements. Ithaca, however, had pH_{ave}=7.74. Soil pH was found to significantly correlate with binning location (p-value = 0.028). Soil saturation was conserved between rural and urban binning locations (Fig 3D). Surprisingly, Urban bins had a higher soil saturation (on average) than rural locations. The average percent soil saturation for rural locations was 55.9 and 54.1 for Cuyler and Owego respectively (Fig 3D). The average percent soil saturation for urban sites was 75.3 and 80.9 for Auburn and Ithaca respectively. Owego had the widest range of soil saturation with 87.8-20.9%. Owego also had the biggest bin size, being composed of 15 samples. Auburn had the smallest range in soil saturation, with samples ranging from 48.9% to 100.8%. Due to the methods used to calculate percent base saturation there are values which exceed 100% saturation (Fig 3D). Base saturation values did not significantly correlate with all binning locations (p-value = 0.12).

Consistent enrichment of bacterial orders at urban and rural sites. Comparing the differential abundance analysis reveals orders which are enriched across both urban sites and urban sites (Fig 4, 5). Urban sites Auburn and Ithaca shared a significant enrichment of Rokubacteriales, Spirochaetales, Vicinamibacterales, Rhizobiales, and NB1-j (Deltaproteobacteria). Rural sites Cuyler and Owego shared a significant enrichment of Defluviicoccales, Methanosarciniales, Bacteroidetes_VC2.1_Bac22, Flavobacteriales, Cytophagales, and Sphingobacteriales.

Defluviicoccale, Spirochaetales, Flavobacteriales, and Rokubacteriales all show significant correlations between cell abundance and [NO3]. After performing DESeq2 to identify differentially expressed orders between urban and rural binning sites, a correlation analysis was run for the most differentially expressed orders and select metadata categories. We choose to run the abundance of differentially expressed organisms against soil pH, percent base saturation, soil $[NO_3]$, and soil $[NH_4]$. Using a significance metric of $p < 0.05$, we found that the Order of Defluviicoccale had significant correlations between cell abundance and pH (positive correlation in Auburn, Ithaca, and Owego bins), NO_3 (negative correlation in Ithaca and positive correlation in Owego bins), NH_4 (negative correlations in Auburn and Owego bins), and soil saturation (positive correlation in Owego bin). The Order Spirochaetales also had significant correlations between cell abundance and pH (negative correlation in Cuyler bin), NO_3 (positive correlation in Cuyler bins), and base saturation (positive and negative correlations in Auburn and Cuyler bins respectively). Abundances of the order Flavobacteriales were found to have significant positive correlations between cell abundance and pH and base saturation in the Owego bin. Abundances of Rhizobiales had significant positive correlations with NO_3 in Ithaca and base saturation in both rural bins (Auburn and Ithaca). Abundances of Rokubacteriales had a significant positive correlation with base saturation in the Auburn bin (Supplemental 2).

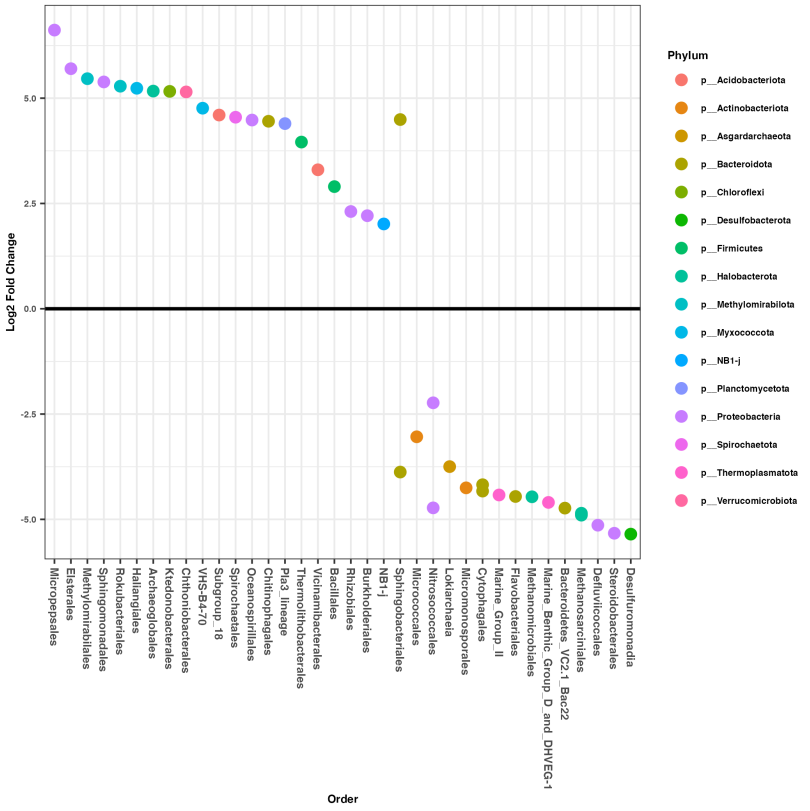

FIG. 4 The relative abundance of significant genus between urban site Auburn and rural site Cuyler. Orders that exhibited a significant log2fold change of 2 between the Auburn and Cuyler site were plotted above using DESeq2. Colours depict distinct phyla.

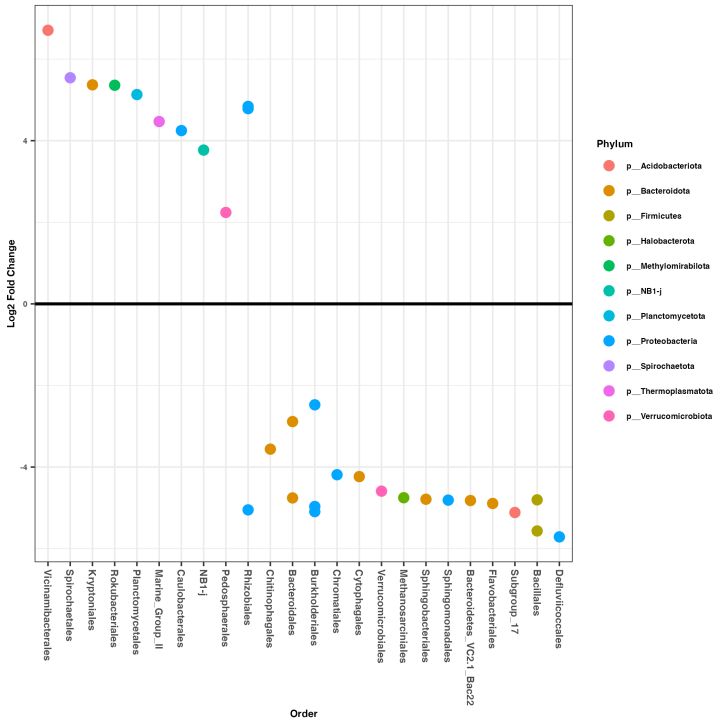

FIG. 5 The relative abundance of significant genus between urban site Ithaca and rural site Owego. Orders that exhibited a significant log2fold change of 2 between the Ithaca and Owego site were plotted above using DESeq2. Colours depict distinct phyla.

DISCUSSION

In this study, we aimed to explore potential differences in microbial community composition and environmental conditions of wetlands in relation to the proximity of two major population centers. We conducted alpha and beta diversity analysis as well as differential abundance with DESeq2. We also conducted a correlation analysis to compare the effects of five environmental variables (soil nitrogen content, soil saturation, N in microbial biomass, and pH) on microbial abundance for each rural and urban binning site.

Significant alpha diversity between urban and rural sites. We sought to determine whether the microbial community composition varied at the sampling sites depending on whether the sites were in proximity to population centers. To investigate this, we evaluated alpha diversity between Auburn (urban), Cuyler (rural), Ithaca (urban), and Owego (rural). As shown in Fig 2, we found significant differences between urban sites and rural sites, with urban sites exhibiting greater measures of alpha diversity (using a p-value threshold <0.05). This suggests that wetlands in proximity to population centers have greater microbial diversity. These results contradict our original hypothesis based on literature stating that urbanization results in a decrease in microbial diversity (11). Existing literature suggests that higher environmental heterogeneity can impact the establishment of dispersed bacterial species, which can decrease the similarities in community composition, and increase diversity (11). Environmental heterogeneity is a term used to describe the non-uniformity present within environments with regards to vegetation and physical factors, which may include soil variables, climate, topography (19). In the context of our study, environmental heterogeneity could result from anthropogenic activity altering natural wetland heavy metal, fertilizer, and CO_2 and CH_4 concentrations. Conversely, too much environmental heterogeneity can result in decreased microbial diversity as the altered environment would then select for specific taxa, over others (11, 20). This may favor select bacterial taxa that exhibit specific metabolic functions that allow them to adapt to conditions resulting from urbanization (eg. increased pollutant content, increased runoff, etc).

Urban and rural sites were enriched with consistent bacterial orders.

Several enriched taxa may be explained when considered alongside our correlation analysis.

Urban community trends: Rokubacteriales is notably enriched in both urban sites, Auburn and Ithaca (Fig 4, 5). Previous studies have characterized Rokubacteriales as a neutrophilic non-methanotrophic found in nitrogen-rich peatlands (21). Peatlands are a classification of wetlands which are highly saturated with water and produce strong anaerobic conditions (22). Our linear regression found that the urban wetlands exhibited greater soil saturation than rural sites (Supplemental 2), which may explain why Rokubacteriales were more abundant in urban sites versus rural sites. This increased abundance might be because the high soil saturation of urban sample sites replicates optimal conditions of peatlands for Rokubacteriales. The same study which characterized the presence of Rokubacteriales at peatlands also found that the relative abundance of Rokubacteriales was positively correlated with pH (22). This is consistent with our finding that Rokubacteriales is differentially abundant at urban sites (Supplemental 2), which feature higher pH levels than the rural sites.

Spirochaetales is another order that is enriched across urban sites (Fig 4, 5). Rieke et al. found that the relative abundance of Spirochaetales in soil increases following the agricultural application of manure to crop soils (23). Both urban sites Ithaca and Auburn are located in counties associated with productive farming with 169, 969 and 91, 277 acres of cropland, respectively (24, 25). Both the counties that Auburn and Ithaca reside in have documented plans on mitigation of manure runoff, with local county rulings suggesting a need for reducing soil leaching and runoff of fertilizer products like manure (26, 27). As we find that Spirochaetales is enriched at sites with documented manure management and runoff concerns, our findings may align with Rieke et al., where Spirochaetales is enriched at sites of manure runoff (23, 28). This grows more plausible when considering the greater elevation of the agriculture land surrounding Auburn site, which may result in a greater accumulation of greater manure runoff (29), and an enrichment of manure-associated microbes like Spirochaetales.

Vicinamibacterales and NB1-j are enriched at all urban sites (Fig 4, 5). Both the orders Vicinamibacterales and NB1-j have been described as taxa present in heavily contaminated

environments, ranging from marine sediments to the Canadian prairies (30, 31). Moreover, independent studies have concluded that Vicinamibacterales and NB1-j are positively correlated with copper contamination (30, 31). A link between agricultural activity and copper contamination is well-documented around global farmlands (32, 33, 34). Moreover, agricultural contamination of copper has been shown to contaminate water systems (34). It may be possible that the proximity of the urban sites to agricultural centers is resulting in greater heavy-metal contamination, which may be influencing the microbiome of wetlands to enrich for species like Vicinamibacterales and NB1-j. Similarly, the order Rhizobiales is enriched across both Auburn and Ithaca urban sites and is also associated with metal and metalloid contaminated soils (35).

Rural community trends: The rural sites Cuyler and Owego depict an enrichment of several orders (Fig 4, 5). Of these enriched rural taxa, Bacteroidetes VC2.1 Bac22 (also known as Bacteroidetes_VC2.1_Bac2 and *Candidatus* Sulfidibacteriales) and Defluviicoccales have yet to be described in soils (36, 37). Sphingobacteriales is also minimally characterized in soil environments, but it is associated in highly diversified and complex soil microaggregates containing organic matter (38). Given the minimal information available on Bacteroidetes VC2.1 Bac22, Defluviicoccales, and Sphingobacteriales, understanding their enrichment at rural sites requires further research.

The enriched orders Flavobacteriales and Cytophagales are associated with the decomposition of organic compounds (39, 40, 41). It is unclear why these decomposing microbes are found enriched at the rural wetlands rather than the urban wetlands. It may be that different microbes fill the decomposer niche at urban sites (Fig 4, 5; 42). Another reason why Cytophagales may be found enriched at the rural sites is that they are typically found in soils with pH levels that range from neutral to slightly acidic (43). As the urban sites had a higher pH, it might be that the near-neutral pH of the rural sites selects for Cytophagales enrichment at rural sites. It is unclear why rural and urban sites vary in their pH levels. A potential list of reasons includes agricultural runoff, gravel road input which increases with proximity to human population centers, and pesticide usage (28, 30). The most common pesticide used in agriculture across all four sites is SuperChlor, which uses the active ingredient sodium hypochlorite (44, 45); sodium hypochlorite is a weak base that has a pH of 12 before it is diluted for agriculture applications (45). In 2009, the counties in which the urban sites Auburn and Ithaca reside in used 9337 and 1636 GL of this chemical, respectively (44). Conversely, the rural sites Cuyler and Owego respectively used 1163 and 148 GL (44). It is possible that this variation in pesticide use may result in a more basic runoff at the urban sites, meaning the near-neutral pH of the rural sites could select for groups like Cytophagales.

Methanosarciniales is enriched across both rural sites (Fig 4, 5). Methanosarciniales are methane-producing archaea which selectively inhabit environments devoid of oxygen (46). Previous literature has described them as a rare taxonomic order present in wetlands in low abundances (47, 48, 49). Wang et al. demonstrated that higher levels of anthropogenic pollution can result in a less resilient archaeal community of wetland methanogens like Methanosarciniales (50). Thus, the relative abundance of Methanosarciniales at rural sites compared to urban sites may be due to a decreased amount of exposure to anthropogenic pollution as a function of population density. Wetland remediation efforts for the rural site Cuyler have been documented (51, 52). A conscious effort to restore the rural wetland sites may enable the growth of more sensitive taxa like Methanosarciniales, which require undisturbed environments due to their sensitivity to oxygen and anthropogenic pollutants (46, 50).

Limitations Our study was limited by the ambiguous sampling methods that Ballantine *et al*. used to retrieve data. There is no associated publication with the dataset used, making it difficult to determine how data was collected and the heuristics used to filter the collected data. We also are unclear about the sequencing procedure used, which could have artificially impacted the measured abundances of certain taxa. In addition to the ambiguities in dataset processing, the data collected by Ballantine et al. was not intended to be used to test the impacts of anthropogenic activity on wetland ecology. If the study was more focused on intentionally testing anthropogenic activity such as farming activity, industrialization, and pollution, the sampling and data processing procedures would likely be different, and would

likely be more reflective of the differences between urban and rural wetland soils. Our adaptation of the dataset produced by Ballantine et al. required us to remove a lot of the sequencing data due to lack of metadata. This decreased the amount of sequencing data we had to work with substantially. If the original sampling procedure was conducted with the intention of measuring the impact of anthropogenic activity on microbial diversity, we would have more data to work with, resulting in a more holistic conclusion.

Conclusions Our results refute our original hypothesis and suggest that microbial alpha diversity may be higher in urban binning locations, closer to population dense areas. We found orders, which have been previously understudied in soil communities, to be significantly different between urban and rural sites. Of particular interest were the orders Rokubaceriales, Defluviicocales, Spirochaetales, and Methanosarciniales. We also found pH and soil saturation to differ between urban and rural locations. Further research is required to determine the full impact of population centers on wetland soil ecology and microbial diversity. Our findings provide a basis for future studies to conduct a more in-depth analysis of microbial diversity and population density.

Future Directions There is conflicting evidence regarding whether microbial diversity increases or decreases with proximity to population dense areas. This draws attention to the need for more research efforts to (i) characterize and study wetland soils and (ii) intentionally study the impacts of anthropogenic activity on wetland microbial soil communities. This can be accomplished by measuring anthropogenic impacts on wetland soil ecology by tracking wastewater runoff, agricultural runoff, stormwater runoff, and measuring pollution content within particular sites. Conducting a more holistic omics approach (proteomics, metabolomics, and transcriptomics) to measure functional community metrics would also allow us to understand the intricacies of microbial communities and anthropogenic impacts. It would be particularly rewarding to investigate the ecological metabolism of known recalcitrant pollutants (such as aromatic compounds, alkanes, and various chlorinated hydrocarbon compounds). These chemicals are known to persist in the natural environment and accumulate to toxic levels (11). Understanding the *in situ* pathways for breakdown of these recalcitrant compounds would allow for better cleanup efforts in the future. In addition, there is minimal information regarding specific taxa in soil environments such as Bacteroidetes VC2,1 Bac22, Defluviicoccales, and Sphingobacteriales; further research can help to understand their enrichment and characterize their roles.

ACKNOWLEDGEMENTS

We would like to thank the Department of Microbiology and Immunology at the University of British Columbia for funding our research. We would also like to express our gratitude to the MICB 475 teaching team at the University of British Columbia, in particular, Dr. Evelyn Sun, Dr. Melissa Chen, and Christopher Lee, for their continuous support throughout this project. Finally, we would like to thank Ballantine *et al.* for their dataset which we used in this study.

CONTRIBUTIONS

TB focused on the QIIME pipeline along with the alpha/beta diversity metrics (with help from DK) in RStudio. DK primarily contributed to the filtering of the metadata and was responsible for generating Figures 1, 2, and 3A-E included in our manuscript. DK also contributed to the correlation analysis in RStudio. RS and AB contributed to the differential abundance analysis in RStudio and helped generate Figures 4, 5. All authors contributed equally to the synthesis of the manuscript, which included the abstract, introduction, methods, results, discussion, conclusion, and future directions sections. Furthermore, TB, DK, and RS contributed to the editing of the draft manuscript to produce the final version.

REFERENCES

1. Cherry, JA. 2011. Ecology of Wetland Ecosystems: Water, Substrate, and Life. Nature Education Knowledge 3(10):16. https://www.nature.com/scitable/knowledge/library/ecology-of-wetland-ecosystems-water-substrate-and-17059765/
2. Bodelier LEP, Dedysh NS. 2013. Microbiology of wetlands. Front Microbiol 4(79). doi: 10.3389/fmicb.2013.00079.
3. Li W, Lv X, Ruan J, Yu M, Song YB, Yu J, Dong M. 2019. Variations in soil bacterial composition and diversity in newly formed coastal wetlands. Front Microbiol 9. https://doi.org/10.3389/fmicb.2018.03256
4. Liu Y, Tong T, Li B, Xie S. 2019. Dynamics of bacterial communities in a river water treatment wetland. Ann Microbiol 69: 637–645. https://doi.org/10.1007/s13213-019-01454-x
5. Wu R, Cheng X, Zhou W, Han H. 2019. Microbial regulation of soil carbon properties under nitrogen addition and plant inputs removal. Peer J (7). doi: 10.7717/peerj.7343
6. Van Horn DJ, Okie JG, Buelow HN, Gooseff MN, Barrett JE, Takacs-Vesbach CD. 2014. Soil microbial responses to increased moisture and organic resources along a salinity gradient in a polar desert. Appl Environ Microbiol 80(10):3034-43. doi: 10.1128/AEM.03414-13
7. Asabere SB, Zeppenfeld T, Nketia KA, Sauer D. 2018. Urbanization leads to increases in pH, carbonate, and soil organic matter stocks of arable soils of Kumasi, Ghana (West Africa). Front Environ Sci 6. https://doi.org/10.3389/fenvs.2018.00119.
8. Hartman WH, Richardson CJ, Vilgalys R, Bruland GL. 2008. Environmental and anthropogenic controls over bacterial communities in wetland soils. PNAS 105(46): 17842-17847. https://doi.org/10.1073/pnas.0808254105
9. Naylo A, Almeida Pereira SI, Benidire L, El Khalil H, Castro PM, Ouvrard S, Schwartz C, Boularbah A. 2019. Trace and major element contents, microbial communities, and enzymatic activities of urban soils of Marrakech City along an anthropization gradient. Journal of Soils and Sediments 19:2153–2165. https://doi.org/10.1007/s11368-018-2221-y
10. Clunes J, Valle S, Dörner J, Campos M, Medina J, Zuern S, Lagos L. 2022. Changes in Soil Quality of an Urban Wetland as a Result of Anthropogenic Disturbance. Land 11(3):394. https://doi.org/10.3390/land11030394
11. Yi X, Ning C, Feng S, Gao H, Zhao J, Liao J, Peng Y, Zhao S, Liu S. 2022. Urbanization-induced environmental changes strongly affect wetland soil bacterial community composition and diversity. Environ Res Lett 17. doi 10.1088/1748-9326/ac444f.
12. U.S. Census Bureau quickfacts: New York.https://www.census.gov/quickfacts/fact/table/NY/POP060220?fbclid=IwAR1X86v40gxCb8rQ8kxdKmPE85_sBxUXu07xVJNgvSfL-T0qGAkPHQW9NyI
13. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, and Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology 37: 852–857.
14. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Met. 13: 581-583. https://doi.org/10.1038/nmeth.3869
15. RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.
16. Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, TerBraak C, Weedon J. 2022. vegan: Community Ecology Package. R package version 2.6-4. https://CRAN.R-project.org/package=vegan.
17. DESEqLove MI, Huber W, Anders S. 2014. "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biology, 15(12):550.
18. McMurdie PJ, Holmes S. 2013. "phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data." *PLoS ONE*, 8(4), e61217. http://dx.plos.org/10.1371/journal.pone.0061217.

19. Dronova I. 2017. Environmental heterogeneity as a bridge between ecosystem service and visual quality objectives in management, planning and design. Landsc Urban Plan, 163:90-106. https://doi.org/10.1016/j.landurbplan.2017.03.005
20. Zhang Y, Ji G D, Wu T, Qiu J. 2020. Urbanization significantly impacts the connectivity of soil microbes involved in nitrogen dynamics at a watershed scale Environ. Pollut 258:113708. https://doi.org/10.1016/j.envpol.2019.113708
21. Ivanova, A, Oshkin, I, Danilova, O., Philippov, D., Ravin, N, Dedysh, S. 2021. Rokubacteria in Northern Peatlands: Habitat Preferences and Diversity Patterns. Microorganisms. 10: 35056460. https://doi.org/10.3390/microorganisms10010011
22. Lindsay, R. 2015. Peatland Classification. In The Wetland Book (pp. 1515-1528). Springer. https://doi.org/10.1007/978-90-481-9659-3_341
23. Rieke EL, Soupir ML, Moorman TB, Yang F, Howe AC. 2018. Temporal dynamics of bacterial communities in soil and leachate water after swine manure application. Front Microbiol 9:3197. https://doi.org/10.3389/fmicb.2018.03197
24. United States Department of Agriculture, National Agricultural Statistics Service. 2017. Census of Agriculture County Profile: Tompkins County. https://www.nass.usda.gov/Publications/AgCensus/2017/Online_Resources/County_Profiles/New_York/cp36109.pdf
25. United States Department of Agriculture, National Agricultural Statistics Service. 2017. Census of Agriculture County Profile: Cayuga County. https://www.nass.usda.gov/Publications/AgCensus/2017/Online_Resources/County_Profiles/New_York/cp36011.pdf
26. Wightman, J, Woodbury, P. 2022. Dairy Manure Storage & Greenhouse Gas Mitigation Opportunities (Project Proposal). Retrieved from https://ccetompkins.org/resources/dairy-manure-storage-greenhouse-gas-mitigation-opportunities
27. Ketterings, Q., Czymmek, K, & Bittman, S. 2009. Economic and environmentally sound manure management option for reduced till systems (Project Report). Northern NY Agricultural Development Program. Retrieved from https://www.nnyagdev.org/wp-content/uploads/2012/03/NNYADP-2009-Manure-Application-Report.pdf
28. Floc'h, J, Hamel, C, Laterrière, M, Tidemann, B, St-Arnaud, M, & Hijri, M. 2022. Long-Term Persistence of Arbuscular Mycorrhizal Fungi in the Rhizosphere and Bulk Soils of Non-host Brassica napus and Their Networks of Co-occurring Microbes. Frot. in Plant. Sci. 13:828145. https://doi.org/10.3389/fpls.2022.828145
29. Government of Ontario. 2020. Handling runoff from solid agricultural source material storages and outside livestock areas. Retrieved from https://www.ontario.ca/page/handling-runoff-solid-agricultural-source-material-storages-and-outside-livestock-areas
30. Dell'Anno, F, Rastelli, E, Tangherlini, M, Corinaldesi, C, Sansone, C, Brunet, C, Balzano, S, Ianora, A, Musco, L, Montereali, M, Dell'Anno, A. 2021. Highly contaminated marine sediments can host rare bacterial taxa potentially useful for bioremediation. Front. In Micb. 12:584850. https://doi.org/10.3389/fmicb.2021.584850
31. Genova, G, Della Chiesa, S, Mimmo, T, Borruso, L, Cesco, S, Tasser, E, Matteazzi, A, Niedrist, G. 2022. Copper and zinc as a window to past agricultural land-use. Journ. of Hazard. Mat. 3: 126631. https://doi.org/10.1016/j.jhazmat.2021.126631
32. Liao Z, Chen Y, Ma J, Islam MS, Weng L, Li Y. 2019. Cd, Cu, and Zn accumulations caused by long-term fertilization in greenhouse soils and their potential risk assessment. Int. J. Environ. Res. 15:2805. https://doi.org/10.3390/ijerph16152805
33. Okereafor, U, Makhatha, M, Mekuto, L, Uche-Okereafor, N, Sebola, T, Mavumengwana, V. 2020. Toxic Metal Implications on Agricultural Soils, Plants, Animals, Aquatic life and Human Health. Int. J. of Env. Res. 7: 2204. https://doi.org/10.3390/ijerph17072204
34. Shiyab, S. 2018. Phytoaccumulation of Copper from Irrigation Water and Its Effect on the Internal Structure of Lettuce. Agric. 2:29. https://doi.org/10.3390/agriculture8020029
35. Teng Y, Wang X, Li L, Li Z, Luo Y. 2015. Rhizobia and their bio-partners as novel drivers for functional remediation in contaminated soils. Front. Plant. Sci. 6:32. https://doi.org/10.3389/fpls.2015.00032
36. Leng H, Zhao W, Xiao X. 2022. Cultivation and metabolic insights of an uncultured clade, Bacteroidetes VC2.1 Bac22 (Candidatus Sulfidibacteriales ord. nov.), from deep-sea hydrothermal vents. Environ Microbiol Res. https://doi.org/10.1111/1462-2920.15931
37. Rubin-Blum, M, Sisma-Ventura, G, Yudkovski, Y, Belkin, N, Kanari, M, Herut, B, Rahav, E. 2021. Diversity, activity and abundance of benthic microbes in the southeastern Mediterranean Sea: A baseline for monitoring. Israel Ocean. Lim. Res., 3108000. https://doi.org/10.1101/2021.01.27.428413 Preprint.
38. Bach, E, Williams, R, Hargreaves, S, Yang, F. 2018. Greatest soil microbial diversity found in micro-habitats. Soil Biol. & Bioc. 118:217-226. https://doi.org/10.1016/j.soilbio.2017.12.018
39. Tondello, A, Fasolo, A, Marcato, S, Treu, L, Bonato, T, Zanardi, W, Concheri, G, Squartini, A, Baldan, B. 2022. Characterization of bacterial communities isolated from municipal waste compost and screening of their plant-interactive phenotypes. Sci. Total Environ. 806:150592. https://doi.org/10.1016/j.scitotenv.2021.150592

40. Zhang, Y, Li, Y, Wang, L, Tang, Y, Chen, J, Hu, Y, Fu, X, Le, Y. 2012. Soil microbiological variability under different successional stages of the Chongming Dongtan wetland and its effect on soil organic carbon storage. Ecol. Eng. 49: 247-253. https://doi.org/10.1016/j.ecoleng.2012.10.002
41. Galitskaya, P, Biktasheva, L, Saveliev, A, Grigoryeva, T, Boulygina, E, Selivanovskaya, S. 2017. Fungal and bacterial successions in the process of co-composting of organic wastes as revealed by 454 pyrosequencing. PLoS One. 12:10. https://doi.org/10.1371/journal.pone.0186051
42. Zhan, P, Liu, Y, Wang, H, Wang, C, Xia, M, Wang, N, Cui, W, Xiao, D, Wang, H. 2021. Plant litter decomposition in wetlands is closely associated with phyllospheric fungi as revealed by microbial community dynamics and co-occurrence network. Sci. Total Environ. 753:142194. https://doi.org/10.1016/j.scitotenv.2020.142194
43. Reichenbach, H. 2006. The Order Cytophagales. In The Prokaryotes. Springer. https://doi.org/10.1007/0-387-30747-8_20
44. New York State Department of Environmental Conservation. 2009. Pesticide Sales and Use Reporting Program. Cornell University. Retrieved from https://psur.cce.cornell.edu/
45. Giuseppe, F. 2009. Hypochlorite based biocide combination for use in agricultural applications. Google Patents. Retrieved from https://patents.google.com/patent/WO2009087188A2/en.
46. DasSarma, S, Coker, J, DasSarma, P. 2009. Archaea. In Encyclopedia of Microbiology. Academic Press. 1-23. https://doi.org/10.1016/B978-012373944-5.00108-5
47. Zhang G, Tian J, Jiang N, Guo X, Wang Y, Dong X. 2008. Methanogen community in Zoige wetland of Tibetan plateau and phenotypic characterization of a dominant uncultured methanogen cluster ZC-I. Environ Microbiol 10(7):1850-1860. https://doi.org/10.1111/j.1462-2920.2008.01606.x
48. Yang, S, Liebner, S, Winkel, M, Alawi, M, Horn, F, Dörfer, C, Ollivier, J, He, J, Jin, H, Kühn, P, Schloter, M, Scholten, T, Wagner, D. 2017. In-depth analysis of core methanogenic communities from high elevation permafrost-affected wetlands. Soil Biol. & Bioc. 111:66-77. https://doi.org/10.1016/j.soilbio.2017.03.007
49. Narrowe, A., Borton, M., Hoyt, D., Smith, G, Daly, R, Angle, J. Eder, C. 2019. Uncovering the diversity and activity of methylotrophic methanogens in freshwater wetland soils. mSystems. 4:6. https://doi.org/10.1128/mSystems.00320-19
50. Wang Y, Hong Y, Ma M, Wu S, Op den Camp HJM, Zhu G, Zhang W, Ye F. 2019. Anthropogenic pollution intervenes the recovery processes of soil archaeal community composition and diversity from flooding. Front Microbiol 10:2285. https://doi.org/10.3389/fmicb.2019.02285
51. Native Orchid Conservation Network. 2022. Conservation. Native Orchid Conservation Network. Retrieved from http://orchidsandconservation.org/conservation.
52. The Wetland Trust. 2023. Our Partners. The Wetland Trust. Retrieved from https://www.thewetlandtrust.org/partners/.
53. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. "Welcome to the tidyverse." Journal of Open Source Software, 4(43): 1686. https://doi.org/10.21105/joss.01686.
54. Paradis E, Schliep K. 2019. "ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R." Bioinformatics, 35: 526-528. https://doi.org/10.1093/bioinformatics/bty633
55. Wickham H. 2016. "ggplot2: Elegant Graphics for Data Analysis". Springer-Verlag, New York.
56. Appelhans T, Detsch F, Reudenbach C, Woellauer S. 2022. "mapview: Interactive Viewing of Spatial Data in R." R package version 2.11.0, https://CRAN.R-project.org/package=mapview