

## RESEARCH ARTICLE

# Flooding and ecological restoration promote wetland microbial communities and soil functions on former cranberry farmland

Rachel L. Rubin<sup>1</sup>, Kate A. Ballantine<sup>1</sup>, Arden Hegberg<sup>2</sup>, Jason P. Andras<sup>2\*</sup>

**1** Department of Environmental Studies, Mount Holyoke College, South Hadley, Massachusetts United States of America, **2** Department of Biology, Mount Holyoke College, South Hadley, Massachusetts, United States of America

\* [jandras@mtholyoke.edu](mailto:jandras@mtholyoke.edu)

## Abstract

Microbial communities are early responders to wetland degradation, and instrumental players in the reversal of this degradation. However, our understanding of soil microbial community structure and function throughout wetland development remains incomplete. We conducted a survey across cranberry farms, young retired farms, old retired farms, flooded former farms, ecologically restored former farms, and natural reference wetlands with no history of cranberry farming. We investigated the relationship between the microbial community and soil characteristics that restoration intends to maximize, such as soil organic matter, cation exchange capacity and denitrification potential. Among the five treatments considered, flooded and restored sites had the highest prokaryote and microeukaryote community similarity to natural wetlands. In contrast, young retired sites had similar communities to farms, and old retired sites failed to develop wetland microbial communities or functions. Canonical analysis of principal coordinates revealed that soil variables, in particular potassium base saturation, sodium, and denitrification potential, explained 45% of the variation in prokaryote communities and 44% of the variation in microeukaryote communities, segregating soil samples into two clouds in ordination space: farm, old retired and young retired sites on one side and restored, flooded, and natural sites on the other. Heat trees revealed possible prokaryotic (*Gemmatimonadetes*) and microeukaryotic (*Rhizaria*) indicators of wetland development, along with a drop in the dominance of *Nucleomycea* in restored sites, a class that includes suspected mycorrhizal symbionts of the cranberry crop. Flooded sites showed the strongest evidence of wetland development, with triple the soil organic matter accumulation, double the cation exchange capacity, and seventy times the denitrification potential compared to farms. However, given that flooding does not promote any of the watershed or habitat benefits as ecological restoration, we suggest that flooding can be used to stimulate beneficial microbial communities and soil functions during the restoration waiting period, or when restoration is not an option.

## OPEN ACCESS

**Citation:** Rubin RL, Ballantine KA, Hegberg A, Andras JP (2021) Flooding and ecological restoration promote wetland microbial communities and soil functions on former cranberry farmland. PLoS ONE 16(12): e0260933. <https://doi.org/10.1371/journal.pone.0260933>

**Editor:** Luigimaria Borruso, Free University of Bozen-Bolzano, ITALY

**Received:** October 27, 2020

**Accepted:** November 19, 2021

**Published:** December 17, 2021

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0260933>

**Copyright:** © 2021 Rubin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Metadata is attached in the [supplementary material](#). Sequence files are available under the NCBI sequence read archive (accession #PRJNA693513), and under JGI's

Genome Portal (projects 1191228 & 1191229) (<https://genome.jgi.doe.gov/portal/Theeffcowetlands/Theeffcowetlands.info.html>).

**Funding:** Funding was provided by Joint Genome Institute Project 1188787. The funder had no role in study design, data collection, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

A paradigm shift has occurred in society's perspective on wetlands. The "drain the swamp" mentality of the early 1900's contrasts with today's rallying calls to preserve, restore, and rewild. In recognition of the role of wetlands for critical ecosystem services, such as carbon storage and nitrogen removal, the United Nations declared 2021–2030 the decade on ecological restoration. While progress is evident, science is still unclear on how to best restore wetlands because the return of plant and animal communities does not guarantee the return of desirable ecosystem functions [1, 2]. Soil microbial community surveys can be used to gauge whether treated sites are on track to meet restoration goals, particularly those that accrue slowly and are difficult to measure, such as soil organic matter accumulation and nitrogen removal.

The physical soil environment influences microbial community composition, which in turn influences the physical soil environment. Despite this established feedback loop, it is still unclear whether discrete community types align closely with soil biogeochemistry (indicator species; [3–5]), or whether community members are redundant to one another, their identity uncoupled from their biogeochemistry [6, 7]. Evidence to support these theories remains incomplete, with much of this research effort focusing on prokaryotes (bacteria and archaea). Microeukaryotes—the microscopic fungi, soil animals, and protists that occur in high numbers in soil and water, have received less attention. Characterizing prokaryotes and microeukaryotes together should yield a more complete picture of the relationships between land management, soil physicochemical gradients, and microbial community structure. Within this field of inquiry, restored peat wetlands are a particularly underexplored ecosystem, and evaluating how microbial communities align with soil biogeochemistry should also provide insights for managing these unique ecosystems.

Cranberries (*Vaccinium macrocarpon*) are native to New England and occur naturally in kettle hole bogs and marshes. Indigenous groups including the Wampanoag have been cultivating and harvesting cranberries for over 12,000 years before commercial cranberry farming spread rapidly across the region in the 1800's. At least half of Massachusetts cranberry farms are built on existing peat wetlands [8], exploiting natural springs for irrigation and harvesting purposes. On modern cranberry farms, water is supplied through a system of dams, ditches, and culverts. Pesticides are applied regularly, and thin layers of sand, often accumulating up to a meter in depth, are applied to promote drainage and prevent root dieback [9]. Until the 1990's, Massachusetts was the worldwide leader in total harvestable acres, but due to climate and economic factors, at least 40% of existing farmland will be retired in the next decade [8].

The cranberry retirement wave presents a widespread opportunity, both for wetland restoration and for research on ecosystem trajectories. Active ecological restoration on former cranberry farms facilitates ecosystem recovery through earth moving, ditch filling, surface grading, and channel or floodplain reconstruction. When no action is taken, retired farms are typically drained of existing water. Over time, it appears that retired sites will often transition to an upland maple or pine forest, never regaining their historical wetland status [10].

Prior to restoration, cranberry farms are typically placed into conservation easements, and it is not uncommon for farms to be retired for 15 years or longer before they can be restored. Rather than allowing retired farms to develop into upland forests, the USDA National Resources Conservation Service (NRCS) experimented with leaving retired farms permanently flooded, in the harvest condition. While flooding clearly maintains ponded water and wet soils, the soil characteristics and microbial communities of flooded sites have not yet been characterized. If flooded sites have comparable soil-based benefits as restored sites, then flooding could be used as a mitigation strategy for sites awaiting restoration, or as an alternative when restoration is not an option.

We conducted a survey of soil microbial communities and ecosystem functions related to soil development and nutrient cycling across 23 sites and five treatments (farm, young retired, old retired, flooded, and restored) in southeastern Massachusetts. Our research questions were: 1) Which of the five treatments had the most similar microbial communities to natural sites? 2) How much variation in community composition can be explained by soil variables, and which of these variables align most closely to microbial community composition? 3) Which taxa are indicators of wetland development? Given the results of our previous work [11], we hypothesized that restored sites would harbor similar microbial communities to natural sites. While we had not examined flooded sites before, we hypothesized that flooded sites would also have similar characteristics to natural wetlands, but to a lesser extent than restored sites, due to their permanently flooded (unnatural) hydrologic regime. We expected that denitrification potential and soil organic matter would be important for explaining variation in microbial communities across the six treatment categories, and that these variables would be highest at the wet sites (natural, restored, and flooded) and lowest at the drier sites (farm, young retired, and old retired).

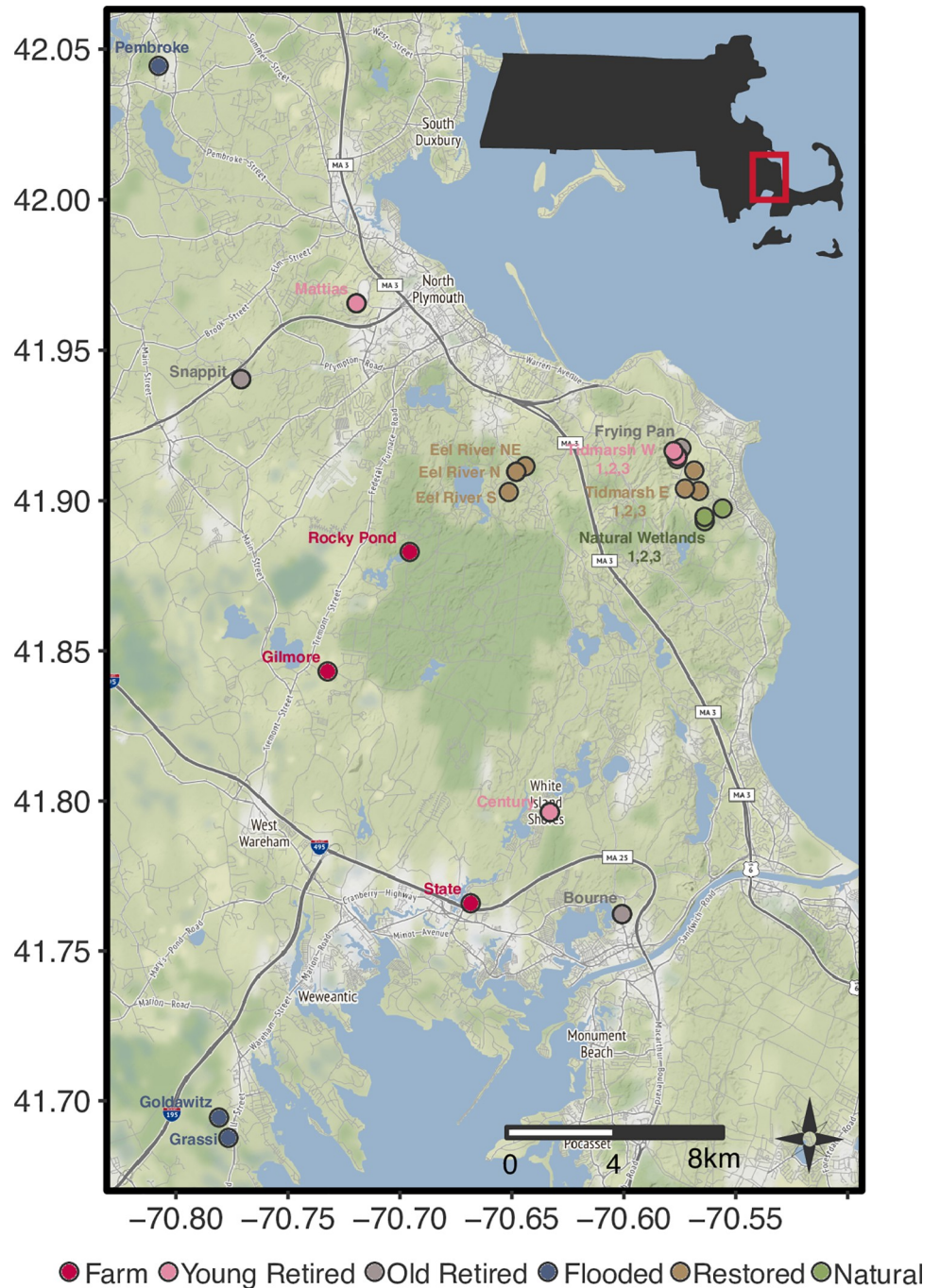
## Methods

### Field sampling and site histories

We sampled twenty-three field sites across southeastern Massachusetts during July 2017, spanning a perimeter of 100 km and total area of 500 km<sup>2</sup>. Treatments included: cranberry farm (3 sites), young retired (5 sites), old retired (3 sites), flooded (3 sites), restored (6 sites), and natural sites (3 sites) (Fig 1). Natural sites have no history of cranberry farming. Restored sites, flooded sites and young retired sites were roughly comparable in age; restored sites had been restored either 1 year (Tidmarsh) or 6 years (Eel River) prior to sampling, flooded sites were flooded 4 years prior to sampling, and young retired sites were retired 1–4 years prior to sampling. Old retired sites were retired 17–20 years prior to sampling, and farms had been continuously farmed for at least 50 years. In our study, “treatments” include multiple sites with varying site ages, and conclusions are drawn with this in mind.

The flooded sites known as “Grassi”, “Goldawitz” and “Pembroke” are NRCS conservation easement sites. Early in the Wetlands Reserve Program, the agency made use of existing water control structures (flumes and wooden boards) to flood the farms and create artificial wetlands (S1 Fig in S1 File). Ecologically restored sites were manipulated according to a methodology developed by the Massachusetts Department of Environmental Restoration (DER). Procedures include removal of dams and water control structures, plugging irrigation ditches with sand, surface roughening and microtopography grading, and, when needed, channel and floodplain reconstruction. No seeding was performed at either the restored or the flooded sites, and previous work has shown that native plants recolonize these sites through activation of the below-ground seedbank [12, 13].

Some sampling sites were clustered together whereas others were more spread out (Fig 1). This sampling scheme was deployed to capture variation within particularly large restoration projects (Tidmarsh West and Eel River), or when sites co-occurred with other outside long-term monitoring objectives (Tidmarsh East). To validate that clustered sampling still captured more variation across these areas, we conducted a sensitivity test by systematically dropping all possible combinations of two sites from Tidmarsh West, Eel River and Tidmarsh East, thereby equalizing site replication across treatment categories (S2 Fig in S1 File). For the majority of cases (8 cases out of 12), dropping two sites at a time from Tidmarsh East (1, 2, and 3), Tidmarsh West (1, 2, and 3) and Eel River (North, Northeast, and South) decreased variation rather than increased variation. Therefore, we retained all 23 sites throughout subsequent analyses.



**Fig 1. Map of study area.** 23 sites were surveyed in southeast Massachusetts, spanning five treatments. Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap (WGS-84), under ODbL.

<https://doi.org/10.1371/journal.pone.0260933.g001>

Eight soil samples (dysic, mesic Typic Haplosaprists of the Freetown series), were randomly selected at each site, spaced 30 m apart. Plants and mosses were removed from the soil surface, and soils were collected using a 1.5-cm diameter soil corer to a depth of 10 cm, and flame sterilized between each sample. Soil samples were homogenized in their own bags and subsampled in the field: half of each sample was flash frozen in liquid nitrogen and stored at -80 °C until DNA extraction, and the other half was stored at 5 °C for the remaining 35 soil analyses.

## DNA sequencing and bioinformatics

DNA was extracted using a DNeasy Powersoil Kit (QIAGEN Inc, Hilden DE) using 0.2 g of soil. For prokaryotic amplicons, libraries were prepared by amplifying the V4-V5 hypervariable region of the 16S rRNA gene using primers 515F (5'-GTGYCAGCMGCCGCGGTAA-3'), and 926R (5'-CCGYCAATYMTTTRAGTTT-3) [14] and sequenced at the Joint Genome Institute (Walnut Creek, CA, USA) on an Illumina MiSeq 96 cycle cartridge, across four runs. For eukaryotic amplicons, libraries were prepared by amplifying the V9 hypervariable region of the 18S rRNA gene using primers 1391F (5'-GTACACACCGCCCGTC-3) and reverse primer EukBR (5'-TGATCCTTCTGCAGGTTTCACCTAC-3) [15, 16] and sequenced at Argonne National Lab on an Illumina MiSeq 300 cycle V2 cartridge, on a single run.

Raw sequences were processed using Quantitative Insights into Microbial Ecology 2 (QIIME2) [17]. Amplicon sequence variants (ASVs) were identified using the Divisive Amplicon Denoising Algorithm (DADA2), which merges sequences, removes chimeric sequences, removes reads with ambiguous bases, and corrects sequence errors so that amplicon sequence variants can be resolved to the level of single-nucleotide differences [18].

For 16S amplicons, reverse reads were low quality (this is common for this gene region), so only forward reads were used. Sequences were truncated to 277 base pairs, and denoising was performed separately on each run, resulting in 22,176,127 total sequences (an average of 129,857 sequences per sample), and 13,978 unique ASVs across 183 samples (an average of 1,579 ASV's per sample). One sample, TE3.2, was removed at the denoising step due to low sequence count. Feature tables and representative sequence files from each of the four runs were merged, and a denovo phylogenetic tree was made using the align-to-tree-mafft-fastree pipeline. Reads from the V4-V5 hypervariable region were extracted from the SILVA ver. 132 database [19], and a feature classifier was trained on that gene region. Taxonomy was assigned at 99% similarity using the SILVA ver. 132 QIIME2 release. Following taxonomy assignment, sequences and feature tables were filtered to remove mitochondrial and chloroplast sequences, as well as sequences that were unassigned at the domain level.

For 18S amplicons, forward and reverse reads were trimmed to 144 bp and denoised, resulting in 11,430,013 total sequences (average of 36,859 sequences per sample) amounting to 70,836 unique ASVs across 183 samples (average of 407 ASVs per sample). One sample (noted as M.3 in the [S1 Data](#)) was removed during denoising due to low sequence count. Phylogenetic trees and feature classification were performed in a similar manner as the 16S analyses, this time using a feature classifier trained on the V9 hypervariable region of the 18S gene. Taxonomy was assigned at 99% similarity, and sequences and feature tables were filtered to remove plants and unassigned sequences at the domain level.

## Soil variables

We measured 35 soil variables to characterize the soil physicochemical environment ([S1 File](#)). Soil analyses were performed by the University of Massachusetts Plant and Nutrient Testing Laboratory, using standard protocols. Macro and micronutrient concentrations were determined using the Modified Morgan extraction procedure [20]. Cation exchange capacity was determined using hydrochloric acid cation displacement method [21], and base saturation of calcium, magnesium, and potassium was determined using atomic absorption spectrometry. Soil organic matter was determined through loss on ignition at 360° C.

Denitrification potential was measured at the Cary Institute of Ecosystem Studies using the denitrification enzyme assay method [22]. This method adds an abundance of carbon and nitrate to each sample in an oxygen free container, and uses an acetylene inhibitor to prevent reduction of N<sub>2</sub>O to N<sub>2</sub>. The amount of N<sub>2</sub>O produced is proportional to the amount of

denitrification enzyme present. Potential methane emissions were measured following ten-day soil incubations, and were quantified using a Shimadzu GC-8 gas chromatograph with a flame ionization detector. Detailed methods for all soil variables are available in [S1 File](#).

## Statistical analyses

QIIME2 feature tables, representative sequences, taxonomy tables, and phylogenetic trees were imported into R (v. 1.2.1335, R Core Team 2020) using the `qiime2R` package [23]. Following conversion to `phyloseq` objects, we removed samples with less than 1000 sequences, which dropped 14 additional samples from the prokaryote dataset and no additional samples from the microeukaryote dataset.

Spatial dependency is a familiar challenge in microbial studies, and occurs in two ways: 1) environmental gradients are spatially structured, causing spatial structuring of taxa, or 2) biological processes such as speciation, extinction, dispersal or species interactions are distance-related. To test for the first type of spatial dependency, we compared the Euclidean distance matrix of GPS coordinates against the Euclidean distance of soil variables (Wisconsin double-transformed to control for different measurement scales). To test for the second type of spatial dependency, we compared the Euclidean distance matrix of GPS coordinates against the weighted UniFrac microbial community distance matrix. We performed both analyses across the whole dataset, and within each treatment category. There was no relationship between geography and soil variables in either case. In comparing geography and community composition, there was no autocorrelation for prokaryotes, but there was significant autocorrelation for microeukaryotes. We controlled for spatial autocorrelation when possible in further analyses, but acknowledge that spatial relationships cannot be fully decoupled from the treatments, which were also spatially structured ([Fig 1](#)).

To assess microbial community similarity to natural sites, we used a phylogenetically meaningful distance measure, weighted UniFrac, because microbial phylogeny is highly correlated with functional traits [24]. Weighted UniFrac incorporates the proportion of shared branch lengths amongst the total branch length, as well as the relative abundance of each taxon [25]. Prior to distance calculations, we resolved multichotomies to dichotomies [26] and used a proportional transformation to turn ASV counts into relative abundance, in lieu of rarefaction [27]. We compared unweighted and weighted UniFrac measures for the 16S and 18S datasets, and chose weighted UniFrac for the final presentation because it produced the best model fit ( $R^2$ ) for prokaryotes and microeukaryotes in the canonical analyses of principal coordinates (CAP; methods to follow).

We conducted a similar analysis for the 35 soil variables. For this analysis, we used a Canberra (weighted Manhattan) distance matrix, calculated from Wisconsin double transformed data to account for different measurement scales for each soil variable. We chose Canberra distance rather than the more commonly used Euclidean distance because our dataset had several zeros (values that were below instrument detection limits) and Canberra distance is less sensitive to shared zeros than Euclidean distance. In this particular case, all distance values happened to fall between 0 and 1, so we converted these numbers to similarity (1-dissimilarity) to ease interpretation. We caution that this can only be done when similarity scores fall between 0 and 1, and similarity scores from this study and should not be compared to similarity scores from other studies. For each of the three analyses described above ([Fig 2A–2C](#)), we also conducted linear mixed effects models on similarity scores using the `lme4` package [28], with treatment coded as fixed effect and site included as random effect to control for non-independence of technical replicates taken from the same site (a square-root transformation was applied when necessary to achieve homogeneity of residuals). When omnibus tests were significant ( $\alpha$

= 0.05), pairwise comparisons between treatments were estimated using the `glht` function in the `multcomp` package [29].

To visualize community differences and to identify the soil variables that relate most strongly to microbial community composition, we used a constrained ordination approach using all 183 technical replicates (canonical analysis of principal coordinates; `capscale` function in `phyloseq`) [30, 31]. Like canonical correlation analysis (CCA), CAP can be used with any distance measure and can also remove variance from nuisance variables (we removed latitude and longitude). Unlike CCA, which assumes a unimodal distribution of species along gradients, CAP assumes a linear distribution of species along gradients, which is well suited to environmental gradients at the regional scale (i.e. southeastern Massachusetts). Using weighted UniFrac distance measures, we used a model selection approach to maximize model fit ( $R^2$ ). Redundant covariates were removed, as were covariates that contributed a variance inflation factor of greater than 10. Forward selection was run on soil variables using the `ordiR2step` function with 200 permutations [32]. Through this process, we achieved a final model with the highest explanatory power and the least number of variables for each of the 16S and 18S datasets. To visualize how environmental gradients align with communities, the top six continuous predictors were plotted as directional vectors.

To visualize differences in within-sample dominance of prokaryotic and microeukaryotic taxa, we constructed differential heat trees using the `metacoder` package [33]. This method expands on the heat map concept by incorporating both the differential abundance and the hierarchical organization of each taxon. We used the  $\log_2$  median ratio to compare the differential abundance of each taxon, using natural wetlands as the universal reference group to compare against the other treatments. For instance, the median abundance for a given bacterial class across all the natural wetland samples was divided by the median abundance of that same bacterial class across all the farm samples, and a  $\log_2$  transformation was applied to normalize the data. Finally, a Wilcoxon Rank-Sum test was used, such that significant differences were shown in color, and non-significant comparisons were shown in grey.

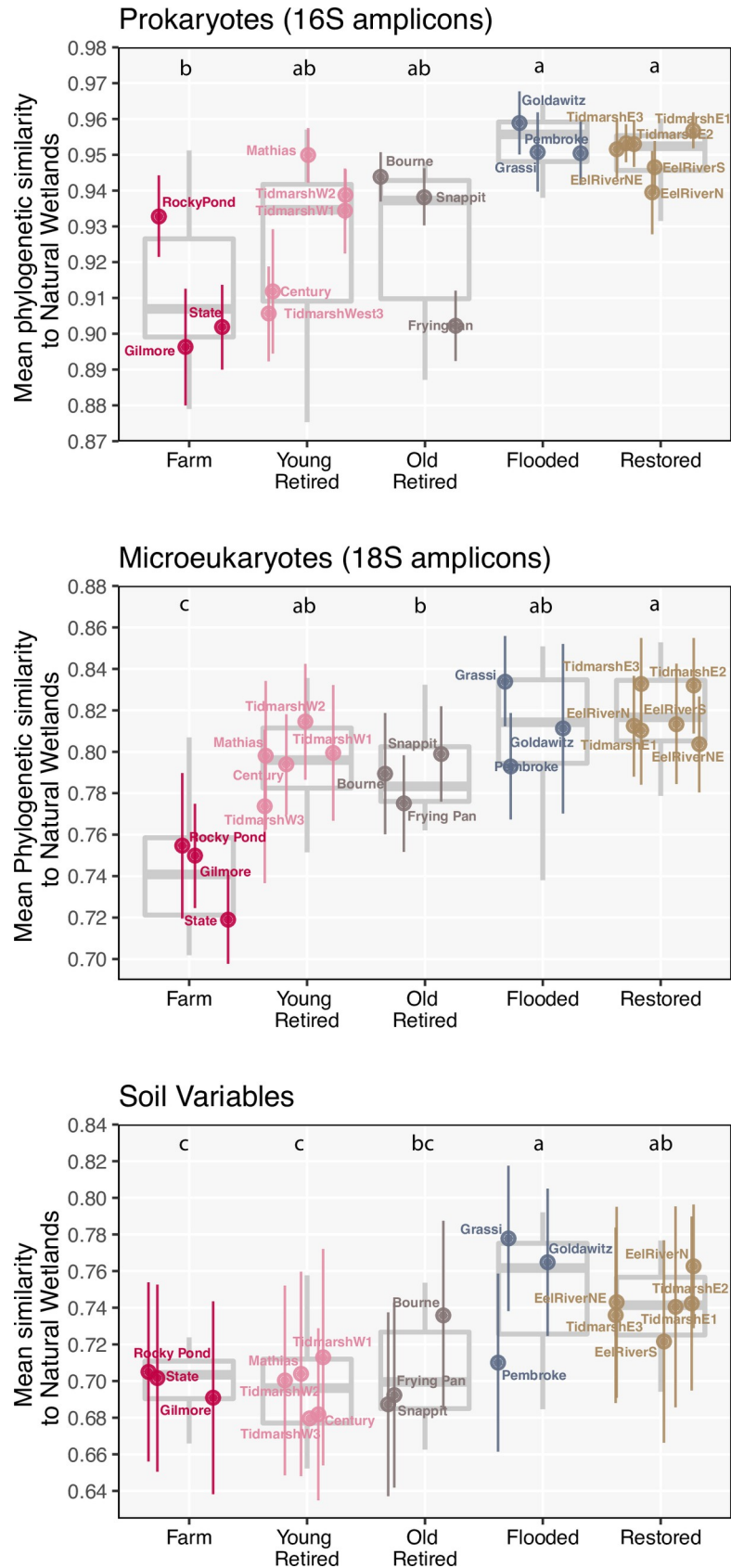
## Results

### Question 1: Which treatments had the most similar microbial communities to natural wetlands?

For prokaryote communities, flooded sites and restored sites had the highest median phylogenetic similarity to natural sites, followed by old retired sites, young retired sites, and farms, which had the lowest similarity (Fig 2A). The same pattern held for microeukaryote communities; restored sites and flooded sites had the highest median phylogenetic similarity to natural wetlands, followed by young retired sites, old retired sites, and farm sites (Fig 2B). In a similar fashion, the multivariate analysis of all 35 soil variables revealed that flooded sites were the most similar to natural sites, followed by restored sites (Fig 2C). Young retired sites, old retired sites and farms had equally lower similarity to natural sites (Fig 2C).

### Question 2: How much variation in community composition can be explained by soil variables, and which of these variables have the strongest association with microbial community structure?

Canonical analysis of principal coordinates revealed microbial community differences and the soil variables that explained the greatest variation. The six strongest soil predictors of prokaryote community composition were: potassium base saturation, sodium, zinc, cation exchange capacity, denitrification potential, and iron. Potassium base saturation and sodium structured





**Fig 2. Similarity between treatments and natural wetland sites.** Plots depict similarity scores for (A) prokaryote communities; (B) microeukaryote communities; and (C) Soil variables. Points indicate the mean similarity of each site to all samples drawn from natural sites, and vertical error bars are standard deviations of similarity scores (eight technical replicates are represented). Grey box plots indicate the median and interquartile range of all soil samples for each treatment category.

<https://doi.org/10.1371/journal.pone.0260933.g002>

communities along CAP axis 1 (Fig 3A), in which farm sites, young retired sites and old retired sites had a higher percent potassium base saturation, and restored and flooded sites had higher sodium. Cation exchange capacity was also a large predictor, structuring communities along CAP axis 1 and CAP axis 2. Together, soil variables explained 45% of the variation of the weighted UniFrac distance, whereas 52% of the variation was unconstrained. Conditional variables (latitude and longitude) explained the remaining 3% of the variation in this dataset.

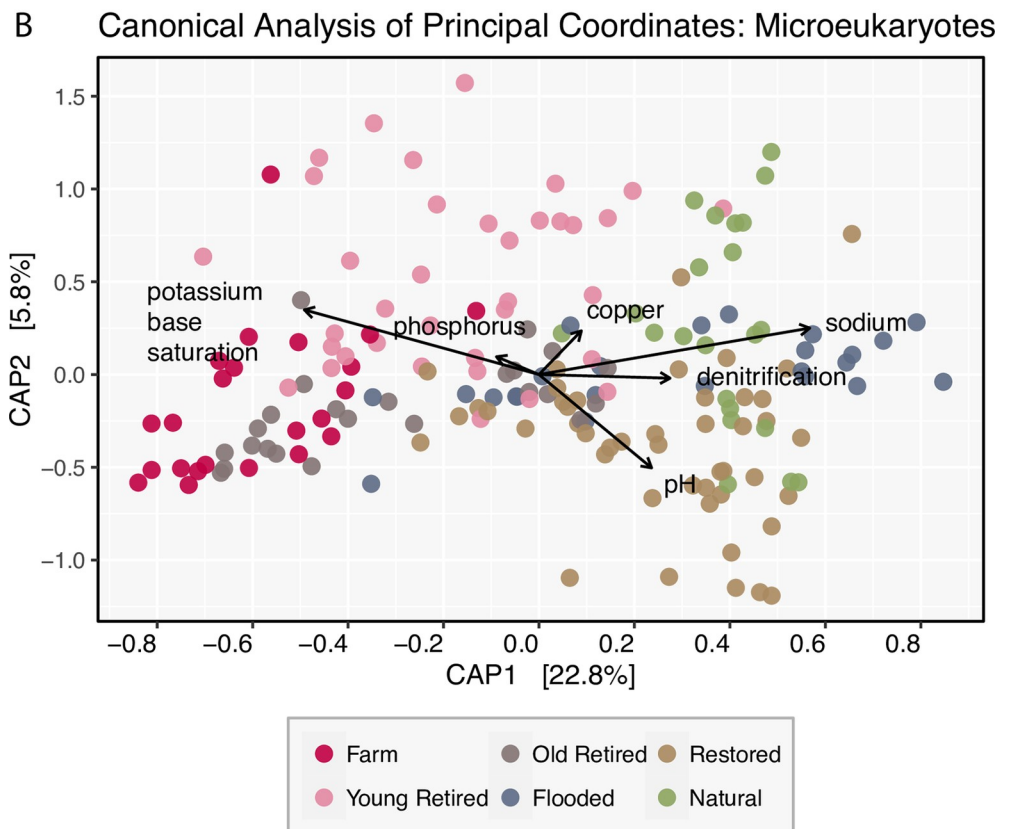
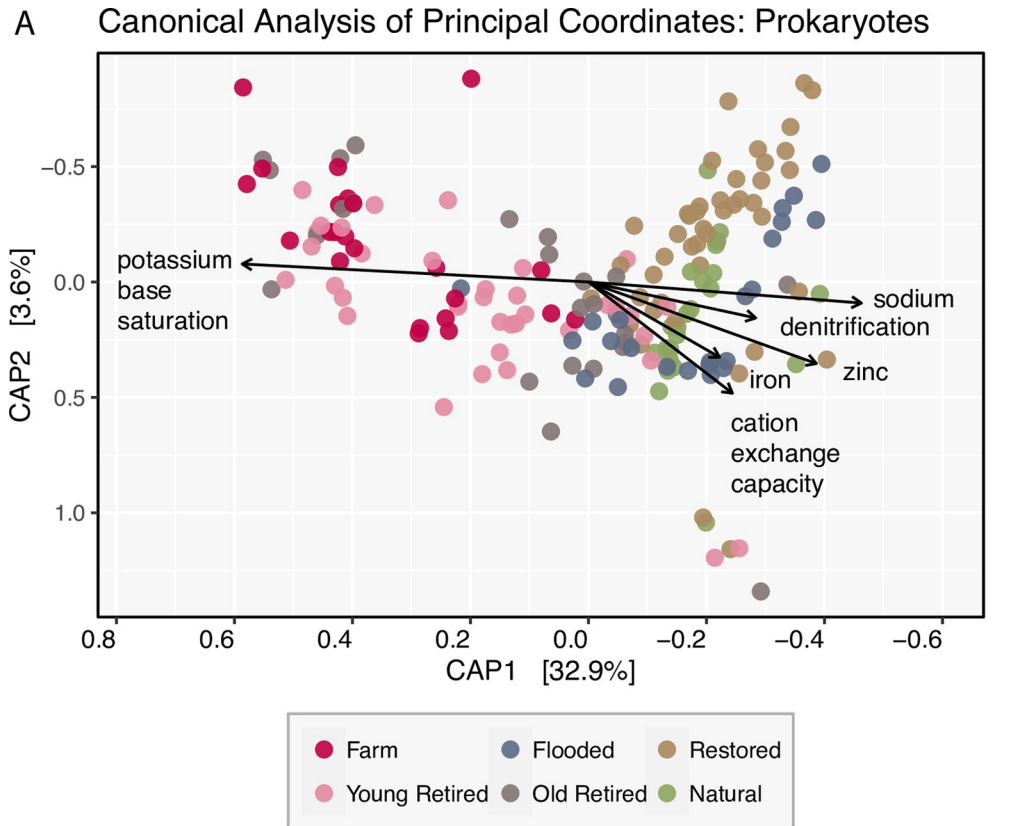
For microeukaryotes, the six largest predictors of community composition were: potassium base saturation, sodium, pH, denitrification potential, copper and phosphorus. Similar to the prokaryotes, microeukaryote communities in farmed, young retired, and old retired sites were correlated with potassium base saturation and sodium along CAP axis 1 (Fig 3B), whereas the higher pH at restored and passively restored sites was correlated with communities at restored sites. Together, soil variables explained 44% of the variation in the weighted UniFrac distance, whereas 49% of the variation was unconstrained. Conditional variables (latitude and longitude) explained the remaining 7% of the variation in this dataset.

The twelve soil variables that were identified as important in the CAP analysis also varied significantly across treatments (Table 1). Three variables important to wetland development—denitrification potential, cation exchange capacity, and soil organic matter—were highest in flooded sites compared to farm sites, young retired sites, and restored sites. However, we note that the difference in denitrification potential between flooded sites and restored sites was not statistically significant, and all treatments had less than 12% of the soil organic matter found in natural sites.

### Question 3: Which taxa are indicators of wetland development?

We used heat trees to visualize how the within-sample dominance of taxa differed between treatments and natural wetlands. For prokaryotes, *Parcubacteria* and *Methanobacteria* were more dominant in natural wetlands, shown in teal (Fig 4), whereas *Actinobacteria* and *Bacteroidia* were typically less dominant in natural wetlands, shown in orange. We identified a possible indicator of wetland development in the phylum *Gemmatimonadetes*, which was more dominant in natural sites when compared to the farm and retired sites (shown in teal), but equally dominant when compared to flooded sites and restored sites (shown in grey). Another possible indicator was a decline in FCPU426, which was more dominant in farm sites and retired sites when compared to natural sites (shown in orange), but equally dominant when compared to flooded sites and restored sites (shown in grey). Generally speaking, heat trees for farm sites and retired sites were similar (top row), and heat trees for flooded sites and restored sites were similar (bottom row), corroborating our earlier findings from the CAP analysis.

For microeukaryotes, natural sites consistently lacked *Nucleotmycea*, a superclass that includes the classes *Fungi* and *Cristidicoidea* (Fig 5). While *Nucleotmycea* was a clear indicator of cranberry farming and its legacy, *Rigifilida* and *Metamonada* were consistently more dominant in natural wetlands compared to all other treatments. *Rhizaria* was a possible indicator of wetland development; it was more abundant in natural sites, but differences in abundance were much smaller when compared to flooded sites and restored sites (indicated by muted teal and orange shading). Generally speaking, heat trees for farm sites and retired sites were similar



**Fig 3. Canonical analysis of principal coordinates.** (A) Prokaryotic communities were distinct (capscale,  $p < 0.01$ ), and the six largest predictors of community composition were: potassium base saturation, sodium, zinc, cation exchange capacity, denitrification potential, and iron. Together, soil variables explained 45% of the variation of the weighted UniFrac distance, whereas 52% of the variation was unconstrained. Conditional variables (latitude and longitude) explained the remaining 3% of the variation in this dataset. (B) Microeukaryote communities were distinct (capscale,  $p < 0.01$ ), and the six largest predictors of community composition were: potassium base saturation, sodium, pH, denitrification potential, copper and phosphorus. Together, soil variables explained 44% of the variation in the weighted UniFrac distance, whereas 49% of the variation was unconstrained. Conditional variables (latitude and longitude) explained the remaining 7% of the variation in this dataset.

<https://doi.org/10.1371/journal.pone.0260933.g003>

(top row), and heat trees for flooded sites and restored sites were similar (bottom row), corroborating our earlier findings from the CAP analysis.

We further examined which taxa contributed to the class level patterns detected above. While no prokaryotic genera showed any particular trends during data exploration, a strong pattern appeared within the eukaryote dataset: the fungal genera *Archaeorhizomyces* and *Cairneyella* were the two most abundant eukaryotic genera across the dataset; together, these genera occupied an average of 36% of the total microeukaryote gene abundance per sample, and up to 78% of the total microeukaryote community in one sample. These genera were dominant in cranberry farms, and decreased markedly in retired, restored and natural wetlands (Fig 6).

## Discussion

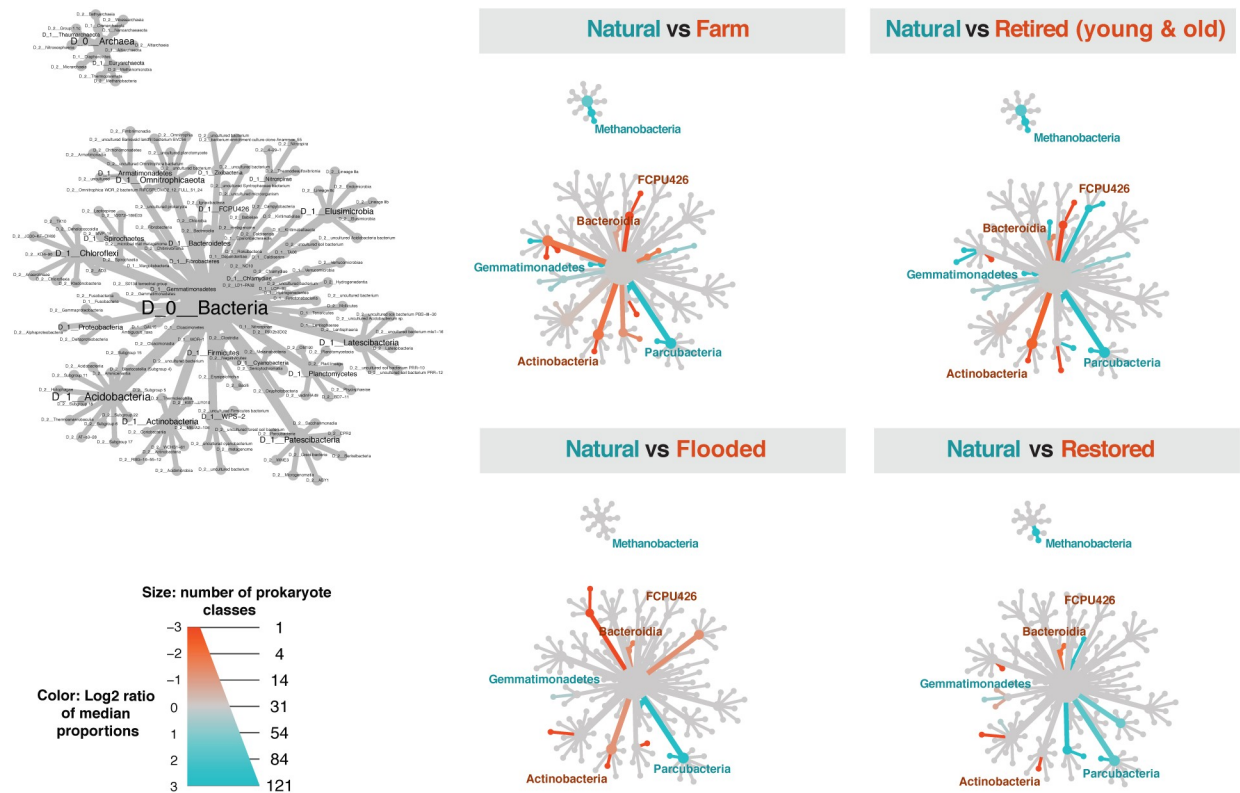
Restoration success is often defined by the convergence of biological and functional variables towards natural reference conditions. Ecological restoration aims to achieve this by filling in drainage ditches of retired cranberry farms and breaking up the sand layer through surface roughening, thereby retaining water, slowing decomposition, and facilitating soil organic matter accumulation on top of the sandy agricultural substrate. We compared ecological restoration to the retired condition and the permanently flooded condition, which is achieved by retaining the agricultural dams that were formerly used for the cranberry harvest. We found that flooding was remarkably effective at jumpstarting wetland biological and functional

**Table 1. Top twelve soil variables that were identified as important predictors of communities of either prokaryotes or microeukaryotes, as determined by canonical analysis of principal coordinates, along with soil moisture.**

Variable	Farm	New Retired	Old Retired	Flooded	Restored	Natural	p-value <sup>2</sup>
Cation exchange capacity (meq 100 g <sup>-1</sup> soil)	6.44 (1.75) <sup>b</sup>	10.57 (3.09) <sup>b</sup>	8.75 (2.53) <sup>b</sup>	12.48 (1.65) <sup>b</sup>	9.02 (3.46) <sup>b</sup>	24.84 (12.60) <sup>a</sup>	<0.001
Copper content (ppm)	0.11 (0.02)	0.77 (1.16)	0.70 (0.95)	0.78 (0.44)	0.19 (0.15)	1.00 (0.63)	ns
Denitrification potential (ng N g dry soil <sup>-1</sup> hour <sup>-1</sup> )	6.2 (12.3)	51.1 (131.3)	124.9 (253.6)	456.5 (727.3)	256.4 (423.8)	883.2 (1,519.8)	ns
Iron content (ppm)	42.75 (22.75)	67.45 (41.87)	68.43 (56.34)	138.01 (49.20)	107.5 (68.32)	561.55 (855.79)	ns
Magnesium base saturation (%)	3.61 (0.65)	3.36 (1.29)	2.32 (1.02)	3.63 (1.14)	4.34 (1.76)	9.61 (8.18)	ns
Methane (ug CH <sub>4</sub> g dry soil <sup>-1</sup> hour <sup>-1</sup> )	0.01 (0.00) <sup>b</sup>	0.02 (0.01) <sup>b</sup>	0.01 (0.00) <sup>b</sup>	0.06 (0.14) <sup>ab</sup>	0.03 (0.04) <sup>b</sup>	0.29 (0.84) <sup>a</sup>	<0.001
pH	4.57 (0.19) <sup>b</sup>	4.32 (0.25) <sup>b</sup>	4.53 (0.38) <sup>b</sup>	4.67 (0.52) <sup>ab</sup>	4.95 (0.27) <sup>a</sup>	4.17 (0.48) <sup>bc</sup>	<0.001
Phosphorus content (ppm)	3.84 (1.62)	5.59 (2.03)	3.68 (1.55)	4.56 (2.10)	3.43 (2.72)	4.03 (2.00)	ns
Potassium base saturation (%)	1.94 (0.39) <sup>a</sup>	1.47 (0.59) <sup>a</sup>	0.86 (0.31) <sup>b</sup>	0.85 (0.33) <sup>b</sup>	0.76 (0.32) <sup>b</sup>	1.16 (0.46) <sup>ab</sup>	<0.001
Sodium content (ppm)	10.80 (3.78) <sup>b</sup>	11.58 (5.19) <sup>b</sup>	8.35 (5.89) <sup>b</sup>	51.33 (62.77) <sup>b</sup>	19.69 (11.27) <sup>b</sup>	90.47 (41.16) <sup>a</sup>	<0.001
Soil moisture (%)	19.1 (6.6) <sup>b</sup>	24.4 (9.7) <sup>b</sup>	14.0 (10.2) <sup>b</sup>	34.7 (16.9) <sup>b</sup>	36.6 (12.7) <sup>b</sup>	74.4 (21.4) <sup>a</sup>	<0.001
Soil organic matter (%)	2.22 (0.89) <sup>b</sup>	3.87 (2.20) <sup>b</sup>	2.71 (1.42) <sup>b</sup>	6.02 (2.80) <sup>b</sup>	3.96 (2.81) <sup>b</sup>	50.76 (33.45) <sup>a</sup>	<0.001
Zinc content (ppm)	0.87 (0.38) <sup>b</sup>	2.01 (1.23) <sup>b</sup>	0.90 (0.68) <sup>b</sup>	5.24 (3.02) <sup>b</sup>	2.18 (1.86) <sup>b</sup>	9.64 (5.26) <sup>a</sup>	<0.001

Variables are listed in alphabetical order, values are means and standard deviations, and N is the total number of technical replicates within each treatment category. P-values are derived from linear mixed effects models conducted on each response variable, which were approximated using Type II Wald chisquare tests as implemented in the lme4 and car packages [28]. When omnibus tests were significant, pairwise comparisons were estimated using Tukey contrasts in the multcomp package [29].

<https://doi.org/10.1371/journal.pone.0260933.t001>



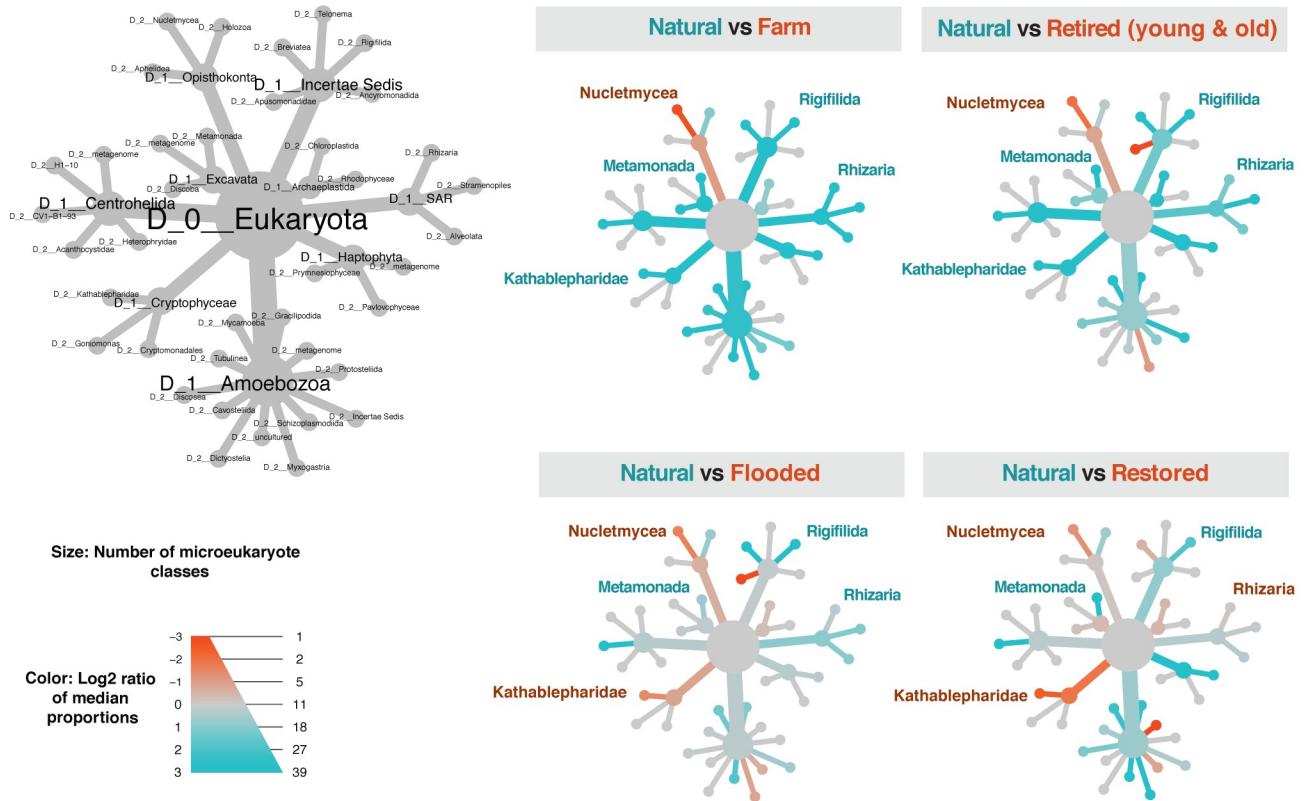
**Fig 4. Differences in prokaryotic taxa between treatments and natural wetlands.** Heat trees were constructed using the metacoder package [33], at the class taxonomic level. The size of the node in each cladogram is proportional to the number of unique classes within each phylum. Color intensity is proportional to the difference in abundance between natural sites and the other treatments, as calculated from the log<sub>2</sub> ratio of median abundance. Young retired and old retired samples were pooled together for simplicity, since they had similar communities. Taxa that have a higher within-sample dominance in natural sites are shown in teal whereas taxa that have a higher within-sample dominance in farm, retired, flooded, and restored wetlands are shown in orange. Nonsignificant comparisons are shown in grey. Colored labels showcase example taxa that were consistently more (teal) or less (orange) abundant in natural sites or displayed variation across the four heat trees.

<https://doi.org/10.1371/journal.pone.0260933.g004>

characteristics, even exceeding restoration in terms of soil organic matter accumulation and cation exchange capacity, after only four years. In contrast, retired sites failed to achieve natural wetland conditions, even after several decades following retirement. These results suggest that farms that are slated for retirement should be left in the flooded condition rather than the drained condition.

Spatial autocorrelation is often inherent in observational studies, and should be considered in the interpretation of results. The Tidmarsh East (restored sites) were close in proximity to the natural sites so there is some uncertainty as to whether the patterns observed here are geographical artifacts or representative of restoration activities as a whole. Yet, while we found spatial autocorrelation (for the microeukaryote dataset only), it was unlikely to have affected our major conclusions: there was no relationship between geography and soil variables, but treatments clearly had a consistent effect on soil variables, suggesting that treatment effects overwhelmed intrinsic environmental gradients (Fig 1, Table 1). Likewise, the flooded sites were the furthest geographically from the natural sites, yet they often had the highest similarity to the natural sites, also indicating that the effect of treatments were much stronger than microbial dispersal limitation.

While some soil-based functions, such as respiration and soil organic matter formation are widespread among microbial taxa, other functions such as methanogenesis and denitrification

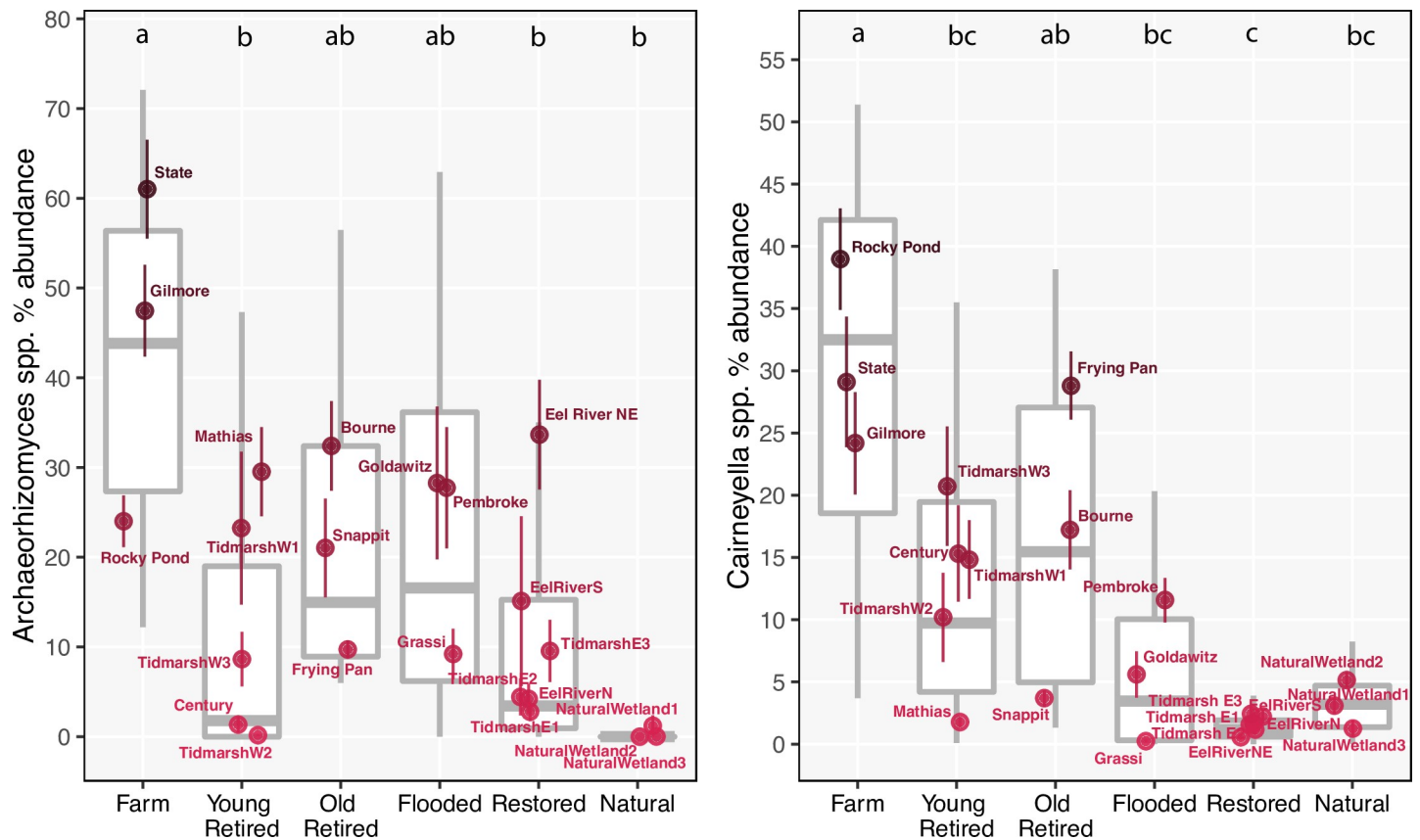


**Fig 5. Differences in eukaryotic taxa between management histories and natural wetlands.** Heat trees were constructed using the metacoder package, at the class level. The size of the nodes in each cladogram is proportional to the number of unique classes within each phylum, and color intensity is proportionate to the difference in abundance between natural sites and each of the other management categories, calculated from the log<sub>2</sub> ratio of median abundance. Young retired and old retired samples were pooled together for simplicity, since they had similar communities. Taxa that have a higher within-sample dominance in natural sites are shown in teal whereas taxa that have a higher within-sample dominance in farm, retired, flooded, and restored wetlands are shown in orange. Nonsignificant comparisons are shown in grey. Colored labels showcase example taxa that were consistently more (teal) or less (orange) abundant in natural sites or displayed variation across the four heat trees.

<https://doi.org/10.1371/journal.pone.0260933.g005>

may be narrowly distributed. In our study we found a strong correlation between microbial communities and soil functions; half of the variation in prokaryotes and microeukaryote communities could be explained by soil variables, including cation exchange capacity and denitrification. It has been previously found that soils with high denitrification potential also have a discrete microbial community [11, 34, 35]. In our study, denitrification was among the top six predictors for both prokaryotes and microeukaryotes, suggesting that the microeukaryotes surveyed here are contributing to denitrification themselves [36], or that microeukaryote communities sort with specific denitrifying bacterial communities. Furthermore, high methane fluxes in natural sites were associated with the dominance of *Methanobacteria*, a class of archaea containing several methanogens. These results lend further support to the idea that microbial taxa (or taxonomic groups) play deterministic roles in soil functions that are important to restoration practitioners, rather than being redundant to one another.

Microbial communities sorted with denitrification and cation exchange capacity, variables that restoration seeks to maximize, as well as underlying chemical gradients; prokaryote communities aligned with potassium base saturation, sodium, zinc and iron, and microeukaryote communities aligned with potassium base saturation, sodium, pH and copper. The high potassium base saturation in farm sites and retired sites may reflect a legacy of fertilizer application,



**Fig 6. Relative dominance of two most common eukaryotic genera.** The two most common eukaryotic genera were *Archaeorhizomyces* and *Cairneyella*, suspected mycorrhizal symbionts of the cranberry crop. Both genera had lower median within-sample dominance in young retired farms, old retired farms, flooded, restored, and natural sites than in cranberry farms. Points indicate mean abundance for each site and vertical lines are standard errors calculated from eight technical replicates. Grey box plots indicate the median and interquartile range of all soil samples within each treatment category.

<https://doi.org/10.1371/journal.pone.0260933.g006>

whereas higher amounts of iron and zinc in flooded, restored and natural sites may be a direct result of the greater amount of soil organic matter found in these soils [37]. Natural sites were also saltier and more acidic than the other treatments, both of which are known environmental filters of microbial communities [38, 39]. Taken together, these results suggest that the effects of restoration treatments on microbial communities can be mediated by indirect effects on soil chemistry.

Differences in taxa dominance across treatments provided context for interpreting differences in biogeochemical variables. For instance, the high within-sample dominance of the methanogen *Methanobacteria* in natural and flooded sites makes sense, given that these sites both had saturated soils and some standing water. It has been previously noted that *Euryarchaeota*, a phylum containing methanogens, dominates wet ecosystems whereas *Thaumarchaeota*, a phylum containing ammonia oxidizers, dominates in drier systems [40]. Our finding that *Gemmatimonadetes* is a possible indicator of wetland development was supported by [41], which found *Gemmatimonadetes* in vegetated, but not bare wetlands. *Actinobacteria* and *Bacteroidia*, which were consistently less dominant in natural sites compared with the other treatments, may be indicators of disturbance or early succession, as noted by [42]. There are fewer ecological data sets on wetland microeukaryotes to compare our results to, but we note that *Rhizaria*, a large group that includes predators of both bacteria and autotrophic protists [43], was also identified as possible indicator of wetland development in our study. Longitudinal

studies over time could reveal early, mid, and late successional indicator communities, and complementary isotope labeling studies could provide insights into trophic interactions [44].

Cranberries are a unique agricultural crop because they are a native wetland perennial, grown in monocultures. We detected a belowground legacy of cranberry cultivation on the soil microbiome, findings which are echoed in investigations of *Vaccinium* patches in natural bogs [45, 46]. The two most abundant genera across the microeukaryote dataset, *Archaeorhizomyces* spp. and *Cairneyella* spp., are suspected ectomycorrhizal and ericoid mycorrhizal symbionts of the cranberry crop [47, 48]. The decline of these genera in restored sites in particular was likely due to soil disturbance, uprooting of cranberry vines, and loss of host plants. Previous vegetation surveys in ecologically restored, former farms indicate that native vegetation establishes quickly following restoration, replacing most cranberry plants [10, 12, 49]. While it appears that these fungal symbionts track with aboveground cranberry abundance, future work evaluating rhizosphere community assembly around controlled plantings of cranberry and native wetland plants could reveal whether soil microbes control aboveground plant succession, or vice versa.

In previous studies, we have described the desirable impacts that active ecological restoration of retired cranberry farms has on both biogeochemical characteristics [50, 51] and microbial community structure [11]. Here, we found that flooding also promotes beneficial wetland functions, even exceeding wetland restoration for soil organic matter formation and cation exchange capacity. While flooding clearly promotes wetland biogeochemistry, flooding does not meet other restoration goals, such as repairing stream connectivity for fish and other aquatic life. Flooding is achieved by leaving artificial dams and water control structures in place, which necessarily impedes the flow of any streams or rivers on site. Many coastal cranberry farms are located on historically important migration and spawning grounds for anadromous fish such as alewife, blueback herring, rainbow smelt, American shad, and white perch. Restoration of these migratory habitats is one of the explicit primary goals of ecological restoration, and these goals are definitively not met when impoundments are left in place.

In conclusion, flooded and restored sites approached soil attributes of natural wetland sites, including prokaryote communities, microeukaryote communities, and 35 key soil variables. Effects were surprisingly strong for flooded sites, likely due to wetter conditions that support soil organic matter accumulation. While these results underscore the foundational role that water retention plays on soil development, we do not suggest that flooding should be a substitute for active ecological restoration. Rather, retired farms should be flooded during the restoration waiting period, or when ecological restoration is not an option.

## Supporting information

**S1 Data. This file is an excel spreadsheet containing site metadata, the 35 soil variables, and sequence counts for each sample.**

(XLSX)

**S1 File. This file contains S1 and S2 Figs as well as detailed methods for the soil variable analyses.**

(DOCX)

## Acknowledgments

We thank Glorianna Davenport, Evan Schulman, Alex Hackman, Helen Castles, Martha Sylvia, Mass Audubon, and several landowners for field access and support. The Cary Institute of Ecosystem Studies provided technical assistance for soil analyses. This work was conducted as

part of the Living Observatory and the Mount Holyoke Restoration Ecology Program, in collaboration with the MA DER Cranberry Bog Program. We thank the hosts of the 2020 QIIME2 Workshop in Bethesda, MD, and Rachael Lappan for the great 16S and 18S QIIME2 tutorial. Finally, we acknowledge the Wampanoag Nation, and their ancestral homeland on which this survey was conducted.

## Author Contributions

**Conceptualization:** Rachel L. Rubin, Kate A. Ballantine, Jason P. Andras.

**Data curation:** Rachel L. Rubin, Kate A. Ballantine, Arden Hegberg, Jason P. Andras.

**Formal analysis:** Rachel L. Rubin.

**Funding acquisition:** Kate A. Ballantine, Jason P. Andras.

**Investigation:** Rachel L. Rubin, Kate A. Ballantine, Arden Hegberg, Jason P. Andras.

**Methodology:** Rachel L. Rubin, Kate A. Ballantine, Jason P. Andras.

**Project administration:** Kate A. Ballantine, Jason P. Andras.

**Resources:** Rachel L. Rubin, Kate A. Ballantine, Jason P. Andras.

**Software:** Rachel L. Rubin.

**Supervision:** Rachel L. Rubin, Kate A. Ballantine, Jason P. Andras.

**Validation:** Rachel L. Rubin, Kate A. Ballantine, Jason P. Andras.

**Visualization:** Rachel L. Rubin.

**Writing – original draft:** Rachel L. Rubin, Kate A. Ballantine, Arden Hegberg, Jason P. Andras.

**Writing – review & editing:** Rachel L. Rubin, Kate A. Ballantine, Arden Hegberg, Jason P. Andras.

## References

1. Ballantine KA, Schneider R. Fifty-five years of soil development in restored freshwater depressional wetlands. *Ecological Applications*. 2009; 19:1467–1480. <https://doi.org/10.1890/07-0588.1> PMID: 19769095
2. Strickland MS, Callahan MA, Gardiner ES, Stanturf JA, Leff JW, et al. Response of soil microbial community composition and function to a bottomland forest restoration intensity gradient. *Appl Soil Ecol*. 2017; 119:317–326.
3. Sims A, Zhang Y, Gajaraj S, Brown PB, Ju Z. Toward the development of microbial indicators for wetland assessment. *Water Res*. 2013;1711–1725. <https://doi.org/10.1016/j.watres.2013.01.023> PMID: 23384515
4. Astudillo-García C, Hermans SM, Stevenson B, Buckley HL, Lear G. Microbial assemblages and bioindicators as proxies for ecosystem health status: potential and limitations. *Appl Microb Biotechnol*. 2019; 103:6407–6421. <https://doi.org/10.1007/s00253-019-09963-0> PMID: 31243501
5. Hermans SM, Buckley HL, Case BS, Curran-Cournane F, Taylor M, Lear. Using soil bacterial communities to predict physico-chemical variables and soil quality. *Microbiome* 2020; 8:79. <https://doi.org/10.1186/s40168-020-00858-1> PMID: 32487269
6. Frossard A, Gerull L, Mutz M, Gessner MO. Disconnect of microbial structure and function: enzyme activities and bacterial communities in nascent stream corridors. *ISME J* 2011; 6:680–691. <https://doi.org/10.1038/ismej.2011.134> PMID: 22030674
7. Louca S, Polz MF, Mazel F, Albright MN, Huber J, et al. Function and functional redundancy in microbial systems. *Nat Ecol Evo* 2018; <https://doi.org/10.1038/s41559-018-0519-1> PMID: 29662222
8. Hoekstra BR, Neill C, Kennedy CD. Trends in the Massachusetts cranberry industry create opportunities for the restoration of cultivated riparian wetlands. *Rest Ecol*. 2020; 28:185–195.



9. Caruso FL, Wilcox WF. *Phytophthora cinnamomi* as a cause of root rot and dieback of cranberry in Massachusetts. *Plant Dis.* 1990; 74:664–667.
10. Klee RJ, Zimmerman KI, Daneshgar PP. Community Succession after Cranberry Bog Abandonment in the New Jersey Pinelands. *Wetlands.* 2019; 39:777–788.
11. Andras JP, Rodriguez-Reillo WG, Truchon A, Blanchard JL, Pierce EA, Ballantine KA. Rewilding the small stuff: the effect of ecological restoration on prokaryotic communities of peatland soils. *FEMS Microb Ecol.* 2020; 96:1–37. <https://doi.org/10.1093/femsec/fiaa144> PMID: 32804239
12. Noyce G, Neill C. Biogeochemistry and vegetation in a cranberry bog chronosequence. *Semantic Scholar.* 2017;519123.
13. Ballantine KA, Davenport G, Deegan L, Gladfelter E, Hatch C, Kennedy C, et al. Learning from the restoration of wetlands on cranberry farmland: preliminary benefits assessment. *Living Observatory,* 2020.
14. Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Env Microb.* 2015; 18:1403–1404. <https://doi.org/10.1111/1462-2920.13023> PMID: 26271760
15. Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *Plos One.* 2009; 4:e6372. <https://doi.org/10.1371/journal.pone.0006372> PMID: 19633714
16. Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, et al. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol.* 2010; 19:21–31. <https://doi.org/10.1111/j.1365-294X.2009.04480.x> PMID: 20331767
17. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, et al. Reproducible, interactive scalable and extensive microbiome data science using QIIME2. 2019; 37:852–857.
18. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene analysis. *ISME J.* 2017; 12:2639–2643.
19. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nuc Acid Res.* 2013; 41:590–596. <https://doi.org/10.1093/nar/gks1219> PMID: 23193283
20. Morgan MF. Chemical soil diagnosis by the universal soil testing system. Connecticut Agricultural Experimental Station. Bulletin 450, 1941. Connecticut, USA.
21. Hendershot WH, Lalonde H, Duquette M. 2008. Ion exchange and exchangeable cations. In Carter MR, and Gregorich EG (eds). *Soil Sampling and Methods of Analysis.* 2nd ed. Canadian Society of Soil Science, CRC Press and Taylor & Francis Group. Oxford, UK.
22. Groffman PM, Holland E, Myrold DD, Robertson GP, Zou X. Denitrification. Pages 272–288 in Robertson GP, Bledsoe CS, Coleman DC, and Sollins P (eds). *Standard soil methods for long term ecological research.* 1999. Oxford University Press, New York, New York, USA.
23. Bisanz JE. qiime2R: Importing QIIME2 artifacts and associated data into R sessions. 2018; <https://github.com/bisanz/qiime2R>.
24. Shankar V, Paliy O. Advantages of phylogenetic distance based constrained ordination analyses for the examination of microbial communities. *Sci Rep.* 2017; 7:6481. <https://doi.org/10.1038/s41598-017-06693-z> PMID: 28743891
25. Lozupone C, Knight R. Unifrac: a new phylogenetic method for comparing microbial communities. *Appl Env Microb.* 2005; 71:8228–8235. <https://doi.org/10.1128/AEM.71.12.8228-8235.2005> PMID: 16332807
26. Paradis E. & Schliep K. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526–528.
27. McMurdie PJ, Homes S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLOS Comp Biol* 2014; 4:e1003531.
28. Bates D, Machler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *Statistical Software* 2015; 67:1–48.
29. Hothorn T, Bretz F, Westfall P. Simultaneous Inference in General Parametric Models. *Biometrical Journal* 2008; 50:346–363. <https://doi.org/10.1002/bimj.200810425> PMID: 18481363
30. Anderson MJ, Willis TJ. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology.* 2003; 85:511–525.
31. McMurdie PJ, Holmes S. phyloseq: An R package for Reproducible interactive analysis and graphics of microbiome census data. *Plos One* 2013; 8:e6121.
32. Blanchet FG, Legendre P, Borcard D. Forward selection of explanatory variables. *Ecology.* 2008; 89:2623–2632. <https://doi.org/10.1890/07-0986.1> PMID: 18831183

33. Foster ZSL, Sharpton TJ, Grünwald NJ. Metacoder: an R package for visualization and manipulation of community taxonomic diversity data. *PLoS Comput Biol*. 2017; 13(2):e1005404. <https://doi.org/10.1371/journal.pcbi.1005404> PMID: 28222096
34. Peralta AL, Matthews JW, Kent AD. Microbial community structure and denitrification in a wetland mitigation bank. *Appl Env Microb*. 2010; 76:4207–4215.
35. Ligi T, Oopkaup K, Truu M, Preem JK, Nölvak H, et al. Characterization of bacterial communities in soil and sediment of a created riverine wetland complex using high-throughput 16S rRNA amplicon sequencing. *Ecol Eng*. 2014; 72:56–66.
36. Maeda K, Spor A, Edel-Hermann V, Heraud C, Breuil M, Bizouard F, et al. N<sub>2</sub>O production, a wide-spread trait in fungi. *Sci Rep*. 2015; 5:9697. <https://doi.org/10.1038/srep09697> PMID: 25894103
37. Yarwood SA. The role of wetland microorganisms in plant-litter decomposition and soil organic matter formation: a critical review. *FEMS Microb Eco* 98:1–17. <https://doi.org/10.1093/femsec/fiy175> PMID: 30169564
38. Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, et al. Soil bacterial and fungal communities across a pH gradient in arable soil. *ISME J*. 2010; 4:1340–1351. <https://doi.org/10.1038/ismej.2010.58> PMID: 20445636
39. Lynam CA, Bulesco AN, Dunphy CM, Osborne CM, Vineis JH, Bowen JL. Microbial community response to a passive salt marsh restoration. *Est. Coasts* 2020; 43:1439–1455.
40. Jiao S, Xu Y, Zhang J, Lu Y. Environmental filtering drives distinct continental atlases of soil archaea between dryland and wetland agricultural ecosystems. *Microbiome* 2019; 7:15. <https://doi.org/10.1186/s40168-019-0630-9> PMID: 30709414
41. Wang J, Wang J, Zhang Z, Li Z, Zhang Z, et al. Shifts in the bacterial population and ecosystem functions in response to vegetation in the yellow river delta wetlands. *mSystems*. 2020; 5:e00412. <https://doi.org/10.1128/mSystems.00412-20> PMID: 32518198
42. Yan D, Bissett A, Gellie N, Mills JG, Lowe AJ, Breed MF. Soil bacterial community differences along a coastal restoration chronosequence. *Plant Ecol*. 2019; 221:795–811.
43. Sisson C, Gulla-Devaney B, Katz L, Grattepanche JD. Seed bank and seasonal patterns of the eukaryotic SAR (Stramenopila, Alveolata and Rhizaria) clade in a New England vernal pool. *J Plankton Res*. 2018; 40:376–390.
44. Seyler LM, Tuorto S, McGuinness LR, Gong D, Kerkhof LJ. Bacterial and archaeal specific-predation in the North Atlantic Basin. *Front Mar Sci*. 2019; 6:1–10.
45. Timonen S, Sinkko H, Sietö O, Rinta-Kanto JM, Kiheri H, et al. Ericoid roots and mycospheres govern plant-specific bacterial communities in boreal forest humus. *Microb Ecol*. 2017; 73:939–953. <https://doi.org/10.1007/s00248-016-0922-6> PMID: 28025668
46. Chronaková A, Bárta J, Kastovská E, Urbanová Z, Pícek T. Spatial heterogeneity of belowground microbial communities linked to peatland microhabitats with different plant dominants. *FEMS Microb Ecol*. 2019; 95:fiz130. <https://doi.org/10.1093/femsec/fiz130> PMID: 31425589
47. Rosling A, Cox F, Cruz-Martinez K, Ihrmark K, Grelet G, et al. Archaeorhizomycetes: unearthing an ancient class of ubiquitous soil fungi. *Science*. 2011; 333:876–879. <https://doi.org/10.1126/science.1206958> PMID: 21836015
48. Midgley D, Rosewarne CP, Greenfield P, Li D, Vockler CJ, et al. Genomic insights into the carbohydrate catabolism of *Cairneyella variabilis* gen. nov. sp. nov., the first reports from a genome of an ericoid mycorrhizal fungus from the southern hemisphere. *Mycorrhiza* 2016; 26:345–352. <https://doi.org/10.1007/s00572-016-0683-6> PMID: 26861481
49. Sandler HA, Alpert P, Shumaker D. Invasion of natural and agricultural cranberry bogs by introduced and native plants. *Plant Ecol*. 2007; 190:219–231.
50. Ballantine KA, Anderson TR, Pierce EA, Goffman PM. Restoration of denitrification in agricultural wetlands. *Ecol Eng*. 2017; 106:570–577.
51. Bartolucci NN, Anderson TR, Ballantine KA. Restoration of retired agricultural land to wetland mitigated greenhouse gas emissions. *Restoration Ecology*. 2020; 29:e13314.