Title: Microbial community structure corresponds to nutrient gradients and human impact within coastal wetlands of the Great Lakes.

Dean J. Horton[1], Kevin R. Theis[2], Donald G. Uzarski[1], and Deric R. Learman[1]*

[1]Institute for Great Lakes Research and Department of Biology, Central Michigan University, Mt. Pleasant, MI, 48859, U.S.A.

[2]Department of Microbiology, Immunology, and Biochemistry, Wayne State University, Detroit, MI 48201, U.S.A

*Corresponding author: Department of Biology, 3111 Biosciences Building, Central Michigan University, Mt. Pleasant, MI, 48859. Email: deric.learman@cmich.edu; Telephone: 989-774-4409

14 **Originality-Significance Statement**

15 This research is original in providing an initial, geographically wide-ranging
16 characterization of microbial communities of the Great Lakes coastal wetlands, an understudied
17 system of wetlands directly bordering the North American Laurentian Great Lakes. Relationships
18 between microbial communities within coastal wetland soils and geochemistry highlighted that
19 anthropogenic impacts may be significantly influencing microbial community structure, as well
20 as unique subnetworks of microbial taxa. Negative anthropogenic impacts to these coastal
21 wetland communities could influence natural biogeochemical cycles which occur within coastal
22 wetland soils, and by extension would directly influence the Great Lakes themselves.

23

24 **Summary**

25 Microbial communities within the soil of Great Lakes coastal wetlands drive
26 biogeochemical cycles and provide several other ecosystems services. However, there exists a
27 lack of understanding of how microbial communities respond to nutrient gradients and human
28 activity in these systems. This research sought to address this lack of understanding through
29 exploration of microbial community diversity and networks among coastal wetlands throughout
30 the Great Lakes. Significant differences in microbial community structure were illuminated
31 between Lake Erie and all other wetlands, and chemical and biological structure did not vary
32 within Lake Erie with increasing soil depth. Beyond this, alpha diversity levels were highest
33 within Lake Erie coastal wetlands. These diversity differences coincided with higher nutrient
34 levels within the Lake Erie region. Site-to-site variability existed within Lake Erie, East and
35 North Saginaw Bay regions, suggesting site-scale heterogeneity may impact microbial
36 community structure. Several subnetworks of microbial communities and individual OTUs were

related to chemical gradients among wetland regions, revealing several candidate indicator communities and taxa which may be useful for Great Lakes coastal wetland management. This research provides an initial characterization of microbial communities among Great Lakes coastal wetlands, and demonstrates that microbial communities could be negatively impacted by anthropogenic activities.

Introduction

The Laurentian Great Lakes of North America are one of the largest freshwater systems on Earth, and are critical in supporting biogeochemical cycles, freshwater resources, biodiversity, and economic viability of the surrounding region. Coastal wetlands bordering the Great Lakes are integral in sustaining the proper functioning of the Great Lakes themselves, making up nearly 200,000 ha of habitat between the United States and Canada despite reduction of habitat by approximately 50% since European colonization (Dahl, 1990; Hecnar, 2004; Sierszen *et al.*, 2012). In addition, Great Lakes coastal wetlands are vital in the retention of chemical pollutants (e.g., heavy metals), sediments, and excess nutrients (e.g., N and P) caused by anthropogenic activity, and provide critical habitat for countless biological species (Wang & Mitsch, 1998; Sierszen *et al.*, 2012). The economy of the Great Lakes is contingent on the existence and proper functioning of coastal wetlands. In providing ecosystem services and promoting biodiversity, these wetlands have an estimated annual worth of $69 billion USD; the value of recreational fishing alone is valued at $7.4 billion USD per year (Krantzberg & de Boer, 2008; Campbell *et al.*, 2015). Notably, the Great Lakes region has been impacted by negative anthropogenic pressure, with cumulative stress having a particular impact on the western basin of Lake Erie (Danz *et al.*, 2007; Uzarski *et al.*, 2017). These negative impacts extend to the coastal

wetlands, as pressure from agricultural runoff, atmospheric deposition, and urbanization influence water chemistry, thereby reducing water quality and impairing wetlands (Trebitz *et al.*, 2007; Morrice *et al.*, 2008). This has stoked a surge in research assessing biodiversity and anthropogenic pressure on coastal wetlands of the Great Lakes since the Great Lakes Water Quality Agreement (GLWQA) was established in 1972 (Hackett *et al.*, 2017). While much research on coastal wetlands has flourished in the wake of this international agreement, microbial communities within Great Lakes coastal wetlands remain almost entirely uncharacterized (Hackett *et al.*, 2017).

As ecological transition zones between upland and aquatic environments (Uzarski, 2009), wetlands host a unique suite of chemical cycles while providing ecosystem services to bordering environments. Most notably, carbon mineralization occurs within wetland soils via redox processes mediated by microbial communities, and these processes contribute to pollution mitigation and atmospheric greenhouse gas flux (Conrad, 1996; Reddy & DeLaune, 2008). Wetland soil often becomes chemically structured with increasing depth through sequential reduction of electron acceptors that decrease in metabolic favorability to microbes due to thermodynamic constraints (Conrad, 1996; Reddy & DeLaune, 2008; Kögel-Knabner *et al.*, 2010). As microbial community metabolism will change in concert with soil chemical profiles, microbial community structural shifts commonly result (Lüdemann *et al.*, 2000; Edlund *et al.*, 2008; Lipson *et al.*, 2015). However, concentration of carbon electron donors can influence the vertical stratification of redox processes (Achtnich *et al.*, 1995; Alewell *et al.*, 2008), and by extension, vertical microbial community structure (defined as relative proportions of microbial taxa within a community). As an example of how this may apply to natural environments, increased nutrient influx from anthropogenic activities (such as agricultural pressure) may

impact microbial community structure within coastal wetlands, which may extend to redox processes. While sequential reduction chains and vertical microbial community structuring can be predictable in lab experiments, complex heterogeneity within wetland soil can often influence microbial community processes and structure to a large degree (Alewell *et al.*, 2008; Lamers *et al.*, 2012). As such, it is relevant to understand relationships between microbial community structure and vertical soil depth *in situ* along a gradient of nutrient input, as microbial community beta diversity could be indirectly linked to carbon mineralization.

As the microbial communities within Great Lakes coastal wetlands have yet to be fundamentally described, it is of importance to gather baseline data on what microbes exist within these systems, elucidate how these microbes could be interacting, and determine to what extent microbial diversity may already be impacted by anthropogenic chemical disturbance. Additionally, if specific groups of microbial taxa are found to be predictive of environmental gradients, this could aid in understanding specific impacts of anthropogenic activity in generating these gradients and selecting for specific microbial networks or taxa (Sims *et al.*, 2013; Urakawa & Bernhard, 2017). This study sought to provide an initial characterization of microbial communities within soils of Great Lakes coastal wetlands bordering the western basin of Lake Erie, Saginaw Bay of Lake Huron, and northern Lake Michigan. Additionally, this study explored how environmental conditions of these coastal wetlands could act as potential drivers of microbial community structure among and within wetlands. It was predicted that microbial community structure would be related to environmental gradients among and within coastal wetland regions of the Great Lakes. Further, it was anticipated that high nutrient levels within wetlands would decouple the relationship between microbial community structure and soil depth, as has been suggested in previous studies (Achtnich *et al.*, 1995; Alewell *et al.*, 2008). Through

high-throughput sequencing of the 16S rRNA gene and network analyses, variations in key microbial taxa and subcommunities related to environmental gradients established by wetlands were identified.

Results

Chemical analyses

Significant correlations ($r > 0.7$, $p \leq 0.001$) were found among NH_4^+, OM, OC, and TN. Thus, downstream analysis combined these values into one parameter, "NUTR", represented by OM values as this variable was the most strongly correlated with each of the other variables. PC1 and PC2 explained 56.2% and 20.6% of the variation among samples, respectively (Fig. 1). MANOVA found significant differences in physicochemical profiles based on region ($F = 2.71$, $p \leq 0.001$) and depth ($F = 6.85$, $p \leq 0.01$). Ninety-five percent confidence intervals suggested that significant variation among regions was driven primarily by physicochemical distinctness of Lake Erie coastal wetlands. This separation was related to increased NUTR, NO_3^-, and S.

Increasing depth within cores showed a consistent shift in environmental variables, except those sites located in the western basin of Lake Erie (Supplemental Fig. 1). Specifically, OM, OC, and TN consistently decreased with increasing depths within all wetlands except Lake Erie sites. Similarly, C:N increased with depth in all wetlands except Lake Erie sites, where C:N ratio remained relatively low (~ 12) and stable with increasing soil depth. Within the Lake Erie wetland region, pH was more acidic in the overlying water with respect to all other wetland regions (Supplemental Table 1). However, pH was still relatively neutral within Lake Erie (average pH = 7.26, standard deviation = 0.24), where other wetland regions experienced slightly more basic pH, with averages ranging between 7.72 – 8.39.

129

130 ***Alpha Diversity***

131 Sufficient depth of sampling was also reinforced by rarefaction curve analysis
132 (Supplemental Fig. 2). Good's coverage reported averages ranging between 89.3 – 93.5% for
133 each region at the subsampled value of 48,226 sequences. Chao 1 richness estimates varied
134 significantly among wetland regions ($F = 8.38$, $p \leq 0.05$), as well as wetland sites ($F = 16.78$, $p \leq$
135 0.001). Pairwise comparisons found that the LE region had significantly higher ($p \leq 0.01$) Chao1
136 estimates than NSB and WSB regions (Fig. 2; Supplemental Table 2). Additionally, pairwise
137 comparisons found a high degree of significant variability ($p \leq 0.01$) in Chao1 estimates among
138 wetland sites (Supplemental Table 2). Further, Shannon diversity levels also significantly varied
139 among wetland sites ($F = 4.57$, $p \leq 0.001$), with site LE_D having significantly higher ($p \leq 0.01$)
140 Shannon diversity levels than sites ESBT_A and WSB_B (Supplemental Table 2). Soil depth
141 was not found to be significant in influencing alpha diversity levels.

142 Shannon diversity and Chao1 both significantly correlated with measured environmental
143 variables (Table 1). Specifically, Chao1 estimates increased with NO_3^-, P, and S concentrations
144 ($p \leq 0.01$), and were weakly positively correlated ($p \leq 0.05$) with NUTR. Additionally, Shannon
145 diversity levels increased alongside NUTR and S ($p \leq 0.001$), and were weakly positively
146 correlated with NO_3^- ($p \leq 0.05$). There were no significant relationships between alpha diversity
147 and C:N, and there were no negative relationships between alpha diversity and any of the
148 measured environmental variables.

149

150 ***Beta Diversity***

151 *Beta diversity among regions*

NMDS demonstrated separation of microbial communities based on wetland site, region, and soil depth (Fig. 3). perMANOVA demonstrated that differences in microbial community structure were highly significantly related to wetland site (pseudo-F = 2.57, $p \leq 0.001$), region (pseudo-F = 5.91, $p \leq 0.001$), and soil depth (pseudo-F = 6.35; $p \leq 0.001$). Post-hoc pairwise perMANOVA found that community structure within the LE region was significantly distinct ($p \leq 0.01$) from all other wetland regions (Table 2). No significant differences in community structure were found between any other wetland regions compared. Additionally, microbial community beta diversity was distinct ($p \leq 0.003$) between the top soil depth and the middle and bottom soil depths. However, no significant differences in microbial community structure were found between the middle and bottom soil depths (Table 2). Variation in microbial community structure was strongly significantly correlated ($p \leq 0.001$) to depth ($r = 0.41$), NO_3^- ($r = 0.20$), NUTR ($r = 0.60$), and S ($r = 0.41$), and weakly correlated ($p \leq 0.016$) with C:N ($r = 0.11$) and P ($r = 0.14$) (Supplemental Table 3).

Beta dispersion tests suggested weakly significant variation in structural variance among regions ($p \leq 0.05$), however, Tukey's HSD test using adjusted p-values for multiple comparisons did not find any significance ($p > 0.05$) between pairwise comparisons of regional groups. There were no differences in community structural dispersion among soil depths.

Beta diversity within regions

Individual NMDS plots of each region identified a high degree of variability of significant correlations ($p \leq 0.05$) between microbial community structure and several environmental variables, as well as depth, dependent upon the wetland region explored (Fig. 4; Supplemental Table 3). Depth was significantly correlated with microbial community structure

in all wetland regions except NSB and LE. However, microbial community structure was more strongly related to depth in NSB ($r = 0.35$, $p = 0.071$) than LE ($r = 0.19$, $p = 0.40$). NUTR was significantly correlated ($p \leq 0.01$) with community structure within regions BA ($r = 0.82$), ESBT ($r = 0.51$), and LE ($r = 0.66$). C:N was significantly correlated ($p \leq 0.01$) with community structure within regions within Saginaw Bay (i.e., ESBT [$r = 0.65$], NSB [$r = 0.58$], and WSB [$r = 0.58$]). Beta diversity was not significantly associated with concentrations of NO_3^- in any region.

To test for significant differences in microbial beta diversity within regions, perMANOVA was used to evaluate differences in microbial community structure among soil depths and sites within wetland regions (Supplemental Table 3). Depth did not significantly explain microbial community structure within the region LE (pseudo-F = 0.87; p = 0.65), however, it did explain differences in microbial community structure within the other wetland regions. Specifically, BA (pseudo-F = 2.12; p = 0.006), ESBT (pseudo-F = 2.49; p = 0.001), NSB (pseudo-F = 1.24; p = 0.093), and WSB (pseudo-F = 1.53; p = 0.014). Significant differences in microbial community structure were found among different wetland sites within regions ESBT (pseudo-F = 3.11; p = 0.001), LE (pseudo-F = 2.89; p = 0.004), and NSB (pseudo-F = 2.16; p = 0.003). As only one site was sampled within the BA region, testing for differences among wetland sites within the BA region could not be accomplished.

Network Analyses

Weighted Correlation Network Analysis (WGCNA) was used to explore strong relationships between subcommunities and individual OTUs with environmental parameters within Great Lakes coastal wetlands. After removal of OTUs which did not have at least two

representative sequences in at least 10% of samples, a total of 7,562 OTUs remained for WGCNA. In determining scale-free topology of the OTU network, a soft power threshold of 4 was reached, and an R^2 of 0.87 was established as linear fit from the regression of the frequency distribution of node connectivity against node connectivity (Supplemental Fig. 3). Of the 33 constructed subnetworks, the same one (subnetwork "orange") was found to be most strongly correlated to both NUTR ($r = 0.94$) and NO_3^- ($r = 0.55$) (Supplemental Fig. 4). A separate subnetwork ("pink") was strongly correlated ($r = 0.74$) to C:N. All correlations of subnetworks to environmental variables were significant ($p \leq 0.001$). OTU VIP values ≤ 1 were removed due to the large amount of OTUs within subnetworks correlated with C:N for visualization purposes.

For subnetwork relationships to NUTR (including OM, OC, NH_4^+, and TN), partial least square analysis (PLS) found that 69 OTUs were 93.8% predictive of variance in NUTR (Supplemental Fig. 5). OTU co-correlation networks were constructed using an OTU co-correlation threshold of 0.25, with strong correlations ($r > 0.59$) between all OTUs and NUTR (Fig. 5). Of the top 15 OTUs contributing to PLS regression by VIP score, seven were related to *Betaproteobacteria*, five were related to *Anaerolineaceae* (within *Chloroflexi*), and one representative OTU was related to each of *Bellilinea* (*Chloroflexi*), *Desulfobacterales* (*Deltaproteobacteria*), and *Rhizobiales* (*Alphaproteobacteria*).

For subnetwork relationships to C:N, PLS found that 144 OTUs were 59.0% predictive of variance in C:N (Supplemental Fig. 6). Networks were constructed using an OTU co-correlation threshold of 0.1, within positive or negative correlations ($r > +/- 0.2$) between OTUs (VIP > 1) and C:N (Fig. 6). Of the top 15 OTUs by VIP score within the network, two OTUs related to *Bacteroidetes* were negatively correlated with C:N. Other top OTUs were positively related to

220 C:N, including seven OTUs related to *Anaerolineaceae*, four OTUs which were unclassified
221 *Bacteria*, and one representative OTU related to each of *Bacillus* (*Firmicutes*) and *Chloroflexi*.
222 While a correlation strength of r = 0.55 was found between a subnetwork and NO_3^-
223 concentrations, PLS did not find a strong relationship between predicted values and actual values
224 of NO_3^- (Supplemental Fig. 7). As such, individual OTUs were correlated to NO_3^- (Fig. 7) and
225 were further explored through this method rather than WGCNA. Correlation strength to NO_3^-
226 ranged from +/-.37 to +/-.55 for positive and negative correlations, respectively. OTUs related to
227 *Anaerolineaceae* simultaneously were most positively and most negatively related to NO_3^-
228 concentration. OTUs related to *Acidobacteria*, *Chlorobi*, *Bacteroidetes*, *Proteobacteria*, OP11,
229 and *Verrucomicrobia* were predominantly positively related to NO_3^-. OTUs related to
230 *Actinomycetales* (*Actinobacteria*) and *Euryarchaeota* were negatively related to NO_3^-.
231
232 **Discussion**
233 Great Lakes coastal wetlands are biological diversity hotspots that mediate fluxes of
234 biogeochemical cycles, as well as buffer the Great Lakes themselves from terrestrial pollutants.
235 Despite the importance of coastal wetland microbial communities in mediating many of these
236 ecosystem services, a lack of fundamental information on these communities has hindered full
237 understanding of these critical environments (Hackett *et al.*, 2017). Previous research on
238 microbial communities in Great Lakes coastal wetlands has focused on the use of microbial
239 enzymatic assays as a tool to explore decomposition rates and nutrient limitation in wetlands
240 (Jackson *et al.*, 1995; Hill *et al.*, 2006). This study is the first to deeply characterize microbial
241 communities among multiple regions across the Laurentian Great Lakes basin, while
242 simultaneously delineating the importance of environmental conditions in structuring these

243 microbial communities. Additionally, this study highlights important keystone subcommunities
244 of potentially interacting OTUs, which may serve as indicators of ecosystem health.
245
246 *Microbial diversity driven by chemistry within Great Lakes coastal wetlands*
247 Taxonomic groups such as plants, birds, fish, and invertebrates within the Great Lakes
248 coastal wetlands have been impacted by human practices within the Great Lakes watershed
249 (Howe *et al.*, 2007; Tulbure *et al.*, 2007; Uzarski *et al.*, 2009; Cooper *et al.*, 2012; Uzarski *et al.*,
250 2017). This study is the first to establish that these patterns appear consistent with microbial
251 communities in these ecosystems as well. Microbial community structure was significantly
252 dissimilar between LE and all other wetland regions. It is therefore likely that anthropogenic
253 stressors related to nutrient loading (and potentially other pollutants) are driving these trends, as
254 nutrient levels were substantially elevated within LE sites. Likewise, Lake Erie costal wetland
255 sites had distinct microbial communities. Previous research has found that nutrient levels (e.g.,
256 C, N, P, etc.), to varying degrees, can influence microbial community composition and structure
257 (Hartman *et al.*, 2008; Peralta *et al.*, 2013; Ligi *et al.*, 2014; Arroyo *et al.*, 2015). Indeed, PCA
258 and MANOVA found that Lake Erie coastal wetlands were chemically distinct from all other
259 wetland regions, primarily driven by elevated nutrient (C, N, P), NO_3^-, and S concentrations
260 within the soil. Additionally, Lake Erie coastal wetlands (and the watershed which drains into
261 them) have been historically impacted by anthropogenic pollution and agricultural practices,
262 particularly in comparison to other coastal wetlands within the Laurentian Great Lakes region.
263 This has been demonstrated by multiple ecological indices (e.g., Cvetkovic & Chow-Fraser,
264 2011; Uzarski *et al.*, 2017) and physicochemical uniqueness (increased levels of nutrients and
265 particulate matter) within the western basin of Lake Erie (Danz *et al.*, 2007; Trebitz *et al.*, 2007;

Cvetkovic & Chow-Fraser, 2011; Uzarski *et al.*, 2017). Data presented in this study corroborate this historical evidence of human impact and nutrient loading in the western basin of Lake Erie.

High nutrient influx could also be influencing chemical and microbial vertical structure within coastal wetland soils. Vertical microbial community structure was not evident within the first 6 cm of soil of Lake Erie coastal wetlands, unlike other coastal wetlands regions. Also unique to Lake Erie, vertical chemical structure was not evident for nutrient levels (C, N, P) or C:N. These data provide evidence that microbial community structural shifts corresponding to wetland vertical profiles are related to concentrations of organic carbon and other nutrients (e.g., NO_3^-, NUTR, P, S). This is meaningful, as it has been previously postulated that concentrations of carbon electron donors may influence redox gradients within wetland soils (Achtnich *et al.*, 1995). Redox gradients are uniquely tied to carbon cycling rates, and as microbial community structure corresponds with redox gradients (Lüdemann *et al.*, 2000; Edlund *et al.*, 2008; Lipson *et al.*, 2015), distinct patterns of vertical community structure within the soil may be indicative of biogeochemical processes being affected within Great Lakes coastal wetlands. Connections between microbial community metabolic shifts with soil depth and levels of dissolved organic carbon *in situ* were incompletely resolved (Alewell *et al.*, 2008). The results presented here suggest that a connection between microbial community metabolism and organic carbon concentration may exist within Great Lakes coastal wetlands, however, it is necessary to better link microbial community diversity, microbial activity, and chemical cycles to establish this connection. As a caveat, it is possible that chemical and microbial structuring still existed within Lake Erie wetlands, yet they were not evident within the first 6 cm of soil or at the spatial scale measured. Nevertheless, microbial communities within Lake Erie coastal wetlands did not follow the same pattern of vertical structure, either chemically or biologically, evident in other regions,

suggesting that the biological integrity of coastal wetland systems is susceptible to negative anthropogenic pressure. Furthermore, carbon and nutrient levels are stable with increasing soil depth in Lake Erie coastal wetlands, which could be indicative of low carbon cycling rates or elevated sedimentation rates.

Similar to beta diversity, patterns in alpha diversity of microbial communities may also be related to soil chemical characteristics within the sampled freshwater coastal wetlands. Microbial community alpha diversity was highest within Lake Erie soils, which experienced the highest levels of nutrient concentrations of the sampled wetland regions. Significant correlations between alpha diversity and physicochemical properties (Table 1) suggest that the higher diversity in Lake Erie coastal wetlands could be driven by its distinct soil characteristics. It has been established in some cases that productivity and alpha diversity increase in a linear relationship on regional scales, but not at local scales (Chase & Leibold, 2002). This pattern of diversity is commonly the result of relatively few species existing within productive habitats on the local scale, while dissimilarity in beta diversity among productive habitat patches increases alpha diversity at larger spatial resolution. It is possible that this relationship exists within our data, with the Lake Erie region being the most OTU-diverse, coinciding with high nutrient concentrations. Habitat patches within Lake Erie wetlands may be smaller in spatial scale than was sampled within this study (where we might expect lower alpha diversity), and high dissimilarity in OTU composition among these hypothetical habitat patches may be responsible for higher alpha diversity in Lake Erie coastal wetlands. Substantiating this, taxa-area relationships have been established for microbial communities which suggest that microbial community turnover can occur at distances as short as millimeters to centimeters within soils and sediments (Grundmann & Debouzie, 2000; Ettema & Wardle, 2002; Horner-Devine *et al.*, 2004;

312 Martiny *et al.*, 2011). Microbial diversity has also been found to have a strong relationship with
313 pH within forest and wetland soils (Fierer & Jackson, 2005; Hartman *et al.*, 2008; Lauber *et al.*,
314 2009). The water directly above the soil in the LE region had a pH that was relatively more
315 acidic (although circumneutral) with respect to other sampled wetland regions. This may suggest
316 that more OC was being oxidized as CO_2 reacts with water to form carbonic acid. This idea
317 would then suggest that LE sites had a higher deposition rate replacing OC as oxidation
318 occurred. While previous research has recognized links between pH and microbial community
319 diversity within wetland soils, these relationships have been highly variable, ranging from a lack
320 of relationship to a strong predictive relationship (Hartman *et al.*, 2008; Peralta *et al.*, 2013; Ligi
321 *et al.*, 2014; Arroyo *et al.*, 2015). To date, research has been variable with respect to
322 environments studied and methods by which pH was measured; this may partly explain the wide
323 range of estimates of relationships between wetland soil pH and microbial community structure.
324 As such, more research is needed to appreciate the influence of pH on microbial communities
325 within wetland soils. Standardization of methodology for assessing environmental characteristics
326 such as pH may also be required.
327
328 *Relationships between microbial subnetworks and environmental gradients*
329 Development of biological indices and establishment of indicator taxa have been
330 suggested as necessary for microbial communities within wetlands, particularly through the use
331 of high-throughput sequencing technologies which now allow for deep assessment of microbial
332 community composition and structure within environmental samples (Sims *et al.*, 2013; Urakawa
333 & Bernhard, 2017). In addition to their importance as biological signals for environmental
334 health, indicator taxa may play prominent roles in bioremediation of excess nutrients and

pollutants found within anthropogenically impacted coastal wetlands. Through network analyses, we have delineated multiple subcommunities which were significantly related to environmental gradients (such as nutrients C, N, and P) measured by Uzarski *et al.* (2017) among coastal wetlands sampled in this study. Specifically, a subnetwork of 69 microbial taxa was 93.8% predictive of nutrient level variation among coastal wetland soils. Several microbial taxa within this subcommunity were individually predictive of nutrient levels to a high degree, including several OTUs related to *Anaerolineaceae*, one OTU related to genus *Anaerolinea*, and another related to genus *Bellilinea*. From the genus *Anaerolinea*, two thermophilic chemoorganotrophs (*Anaerolinea thermophila* and *Anaerolinea thermolimosa*) have been isolated (Sekiguchi *et al.*, 2003; Yamada *et al.*, 2006). Only one isolated member has been established within the genus *Bellilinea* (*Bellilinea caldifistulae*); it has been described as a thermophilic, fermentative, obligate anaerobe which thrives in co-culture with methanogens (Yamada *et al.*, 2007). It is unlikely that the OTUs found in our study are the same species as the isolated *Anaerolinea* and *Bellilinea* species, as coastal wetland soils are not high-temperature environments necessary for thermophilic species. Additionally, no OTUs related to methanogenic archaea were found within this subnetwork, suggesting that *Anaerolineacea* OTUs within coastal wetland soils may fluctuate independently of any specific methanogenic OTUs. It is possible that the *Bellilinea* OTU found within the subnetwork is related to nutrient level concentrations. This would support fermentative metabolism as noted within *Bellilinea caldifistulae*. It is important to note that several other studies have discovered OTUs related to *Anaerolineaceae* within wetland soils, with upwards of 90% relative abundance among *Chloroflexi* OTUs within these systems (Ansola *et al.*, 2014; Deng *et al.*, 2014; Hu *et al.*, 2016). This suggests that there are probable mesophilic

357 species yet to be isolated within this ubiquitous family of bacteria, which may be of high
358 importance within wetland soils.

359 *Betaproteobacteria* were also found to significantly predict nutrient levels among coastal
360 wetlands. Hu *et al.* (2016) found that both *Betaproteobacteria* and *Anaerolineae* were positively
361 related to TN levels, which is consistent with the data presented here, and these two taxa were
362 suggested to contribute to higher levels of heterotrophic activity. Further, *Anaerolineaceae*
363 OTUs were consistently related to increasing C:N, suggesting that many taxa within this family
364 have preference for recalcitrant carbon sources. As C:N also tends to increase with soil depth, it
365 is also probable that the putatively obligate anaerobic *Anaerolineaceae* are coinciding with
366 decreasing oxygen levels and/or changing metabolism requirements with increasing soil depth.

367 While there was no subnetwork of microbial taxa predictive of NO_3^- in the studied coastal
368 wetland systems, there were several taxonomic groups that were either positively or negatively
369 related to NO_3^- independently. Taxa spanning ten phyla were found to either positively or
370 negatively correlate to NO_3^- gradients (Fig. 8). As nitrate reduction is a process which is
371 undertaken by a wide breadth of phylogenetic groups (Reddy & DeLaune, 2008), it is possible
372 that several taxa are capable of nitrate respiration within Great Lakes coastal wetlands. Notably,
373 some taxa which positively related to NO_3^- are known nitrate reducers in other systems, such as
374 members of *Deltaproteobacteria*, including *Myxococcales* and *Anaeromyxobacter* (Sanford *et*
375 *al.*, 2002). Also noteworthy are the high number of OTUs within *Anaerolineaceae* which
376 positively correlated with NO_3^- . However, several *Anaerolineaceae* OTUs were negatively
377 related to NO_3^- gradients, suggesting that functional idiosyncrasy may exist among members of
378 this family.

379

Conclusions

This study marks the first characterization of microbial communities within Great Lakes coastal wetlands. Coastal wetlands are integral in the proper functioning of biogeochemical cycles and environmental sustainability of the Great Lakes. While it has long been known that anthropogenic pressure can impact animal and plant communities within these coastal wetlands, this is the first evidence that these pressures may also be influencing microbial communities, and may be influencing biogeochemical cycles by extension. Alpha and beta diversity were both related to nutrient gradients among and within regions, suggesting that variability in microbial community structure is highly coupled to geochemistry within wetland soil. Further, this study provides insight on microbial community subnetworks and individual OTUs, which were predictive of chemical concentrations, and may be useful for future management of Great Lakes coastal wetland systems.

Beyond taxonomic assessments of microbial communities with relation to wetland monitoring, we propose that wetland microbial community structure can also potentially be used to assess a wetland for management purposes. As illustrated within this study, wetland microbial community structure and depth are decoupled within the wetlands experiencing the highest nutrient levels, likely originating from terrestrial inputs due to human activity. As such, multivariate statistics (as used in the methods of this study) may prove useful in examining relationships between wetland soil depth and microbial community structure alongside taxonomic analyses, which could provide indicators of nutrient loading stress on coastal wetland habitats.

Experimental Procedures

In the summer of 2014, wetland soil cores were collected within Laurentian Great Lakes coastal wetland ecosystems across several sites within several regions. Specifically, soil cores were collected from ten sites across five regions, including two sites in the western basin of Lake Erie (LE), three sites in eastern Saginaw Bay (ESBT), two sites in northern Saginaw Bay (NSB), two sites in western Saginaw Bay (WSB) in Lake Huron, and one site in the Beaver Island archipelago (BA) in Lake Michigan (Fig. 8). These sites were selected as they corresponded to environmental gradients, as well as a human impact gradients based upon SumRank scores as described in Uzarski *et al.* (2017) (Supplemental Fig. 8). Soil cores were collected by hand-driving plastic core tubes (~ 5 cm diameter) vertically into the soil. Among wetlands, samples were collected within the same vegetation zone (either dominated by cattails, genus *Typhus*, or bulrush, genus *Shoenoplectus*) as an attempt to control for collection bias across sites. Cores were sampled to a depth of at least 6 cm (except for one core which was sampled to a depth of 4 cm) and were immediately flash frozen in a dry ice ethanol bath. Samples were transported on dry ice to Central Michigan University wherein they were stored at -80 ˚C.

Triplicate cores were taken at five wetland sites while duplicate cores were taken at five other wetland sites. For sample extraction and sectioning, cores were extruded while still frozen via a custom-built core extruder. The edge of the core was warmed with a heat gun to allow the soil core to pass efficiently through the plastic container, however, the inner-core did not thaw during extrusion. Ice was applied to the plastic core liner to prevent accelerated thawing. Beginning from the top surface of soil, 2 cm sections were cut via an ethanol and flame-sterilized hacksaw blade and the sectioned core samples were placed into whirl pak bags and stored at -80 ˚C. The extruder face plate was sterilized between cuts of the same core with ethanol. The

425 extruder device was fully cleaned and sterilized between cores with physical scrubbing and
426 ethanol sterilization.
427
428 ***Microbial community analysis***
429 Each soil sample was analyzed independently for microbial community analyses. DNA
430 was extracted from ~ 0.25 g of soil using a MoBio PowerSoil DNA Isolation Kit (Mo Bio,
431 Carlsbad, CA) following the standard manufacturer's protocol. Concentrations of extracted DNA
432 were assessed using a Qubit® 2.0 fluorometer (Life Technologies, Carlsbad, CA) to ensure
433 successful DNA extraction and quantification for sequence library preparation. DNA samples
434 were sent to Michigan State University (East Lansing, MI) for library preparation and sequence
435 analysis at the Research Technology Support Facility. The V4 region of the 16S rRNA gene was
436 amplified for downstream sequencing with commonly used primers 16Sf-V4 (515f) and 16Sr-V4
437 (806r) and a previously developed protocol (Caporaso *et al.*, 2012; Kozich *et al.*, 2013). Paired-
438 end 250 bp sequencing was accomplished via a MiSeq high-throughput sequencer (Illumina, San
439 Diego, CA). Acquired DNA sequences were filtered for quality and analyzed using MOTHUR v
440 1.35.1 (Schloss *et al.*, 2009) following the MiSeq SOP (available at https://www.mothur.org/)
441 with modifications. Scripts used for sequence processing can be found at the GitHub repository
442 associated with this study (https://github.com/horto2dj/GLCW/). Briefly, paired end sequences
443 were combined into single contigs. Sequences that contained homopolymers > 8 bases, and those
444 less than 251 or greater than 254 bp were removed. Sequences were aligned against the Silva (v.
445 119) rRNA gene reference database (Quast *et al.*, 2012). Sequences which did not align with the
446 V4 region were also subsequently removed from analysis. Chimeric DNA was searched for and
447 removed via UCHIME (Edgar *et al.*, 2011). Sequences were classified via the Ribosomal

448 Database Project (training set v. 9; Cole *et al.*, 2013) using the 'wang' method with a confidence
449 threshold of 80. Sequences classified as chloroplast, mitochondria, eukaryotic, or unknown were
450 removed. Remaining sequences were clustered into Operational Taxonomic Units (OTUs) at
451 0.03 sequence dissimilarity using the opticlust clustering algorithm. Sequence data associated
452 with this research have been submitted to the GenBank database under accession numbers
453 SRR6261304 – SRR6261377.
454
455 ***Chemical analysis***
456 Each soil layer (top, middle, and bottom) was analyzed separately for local chemistry at
457 each site. Within each site, soil samples of the same depth (i.e., top, middle, and bottom soil
458 samples) among duplicate/triplicate cores were combined and homogenized to obtain enough
459 soil for chemical analyses. Soil samples were analyzed for percent total N ("TN"), total P ("TP",
460 ppm), total S ("TS", ppm), NO_3^- (ppm), NH_4^+ (ppm), percent organic matter ("OM"), percent
461 organic carbon ("OC"), and C:N at the Michigan State University Soil & Plant Nutrient Lab
462 (East Lansing, MI). A YSI multiprobe (YSI Inc., Yellow Springs, OH) was used to measure pH
463 of the water residing directly above each collected soil core.
464
465 ***Statistical analyses***
466 Statistical analyses were completed using R statistical software version 3.2.2 (R Core
467 Team, 2015) unless otherwise stated. Code used for statistical analyses (and bioinformatic
468 workflow) in this study can be found in the associated GitHub repository
469 (https://github.com/horto2dj/GLCW/).
470

Physicochemical Analysis

Differences in chemical profiles between samples within and among wetland regions were visualized using Principal Component Analysis (PCA). Prior to PCA, percentages were arcsin square root transformed and ratios were log transformed. Additionally, Pearson correlation analyses were performed to search for significant correlations between chemical variables. Collinearity in the dataset was addressed by combining highly correlated environmental variables ($r > 0.7$, $p \leq 0.001$). Only one of the correlated variables was included in PCA to remove exaggeration of correlated variables in PCA structure. MANOVA was used to determine the influences of region and soil depth on physicochemical composition of samples, and 95% confidence intervals were established to compare differences among groups. Chemical depth profiles were also visualized for each wetland site to understand shifts in measured environmental variables with increasing soil depth.

Alpha-diversity Analysis

Alpha diversity analyses were performed to explore variation in OTU richness and evenness among wetland sites, regions, and soil depths, as well as to determine whether observed trends were driven by environmental variables. Prior to alpha diversity analyses, sequence abundance for each sample was subsampled to the lowest sequence abundance for any one sample (n = 48,226 sequences). Singletons were maintained within the sequence dataset for alpha diversity analyses, as alpha diversity indices can be reliant on the presence of singletons for proper estimation. Alpha diversity was calculated for each site using MOTHUR, including Chao 1 richness, non-parametric Shannon diversity, and Good's coverage indices. Linear mixed-effect models and ANOVAs were used to test influences of wetland site, region, and soil depth

494 on alpha diversity, controlling for wetland site as a random effect. Linear models and ANOVAs
495 were used to test for variation in alpha diversity among wetland sites. If significant variation was
496 found within an ANOVA result, post-hoc comparisons were implemented between sample
497 groups using Tukey's Honest Significant Differences (HSD) tests with Bonferroni adjustments
498 (p-values obtained by number of comparisons) for pairwise comparisons.
499
500 *Beta diversity Analysis*
501 Beta diversity analyses were used to evaluate variation in microbial community structure
502 among wetland sites, regions, and soil depths, and to assess the extent to which observed
503 variation was explained by environmental conditions. Singletons and doubletons were removed
504 from the dataset for beta diversity analyses. All sequence data were maintained for beta diversity
505 analyses and transformed using the *DeSeq2* (Love *et al.*, 2014) package, which normalized OTU
506 abundances among samples using a variance stabilizing transformation (VST) (McMurdie &
507 Holmes, 2014). The *phyloseq* (McMurdie & Holmes, 2013) and *Vegan* (Oksanen *et al.*, 2007)
508 packages were used to compare beta diversity among samples. Dissimilarity in microbial
509 community structure among samples within and among sites was visualized using Non-metric
510 Multidimensional Scaling (NMDS) plots based on pairwise Bray-Curtis dissimilarity estimates.
511 The function *envfit* of the Vegan package was used to evaluate correlation between chemical
512 parameters and microbial community structure among samples according to NMDS. "Depth"
513 was also implemented as a dummy variable to test correlation between depth and microbial
514 community structure.
515 To test for significant differences in beta diversity among wetland sites, regions, and soil
516 depth, Permutational Multivariate Analysis of Variance (perMANOVA) (Anderson, 2001) were

implemented. Specifically, these tests evaluate significant variation among within group and between group means (Clarke, 1993; Anderson, 2001; Anderson & Walsh, 2013). If perMANOVA found significant differences among groups at the global level, pairwise perMANOVA tests between groups were implemented with Bonferroni significance adjustments to control for multiple pairwise comparisons. Anderson's permutation of dispersions test (PERMDISP; Anderson, 2006; Anderson *et al.*, 2006) was used to test for differences in variance of community structure among sample groups (i.e. sites, regions, soil depths). Tukey's Honest Significant Difference (HSD) tests were implemented with adjusted p-values for multiple pairwise comparisons if significant differences in dispersion were found among groups.

To explore relationships between regional microbial community structure and environmental variables, NMDS plots were generated for each individual region. Applying NMDS to each region also allowed for the assessment of the correlational relationship between community structure and soil depth (as a dummy variable) and other environmental variables (using the *envfit* function) within individual regions. To test for differences in microbial community structure between/among sites within a region, as well as among depths within a region, perMANOVA was implemented individually for each region.

Network and OTU Correlation Analyses

To explore relationships between microbial sub-communities and OTUs to environmental variables, Weighted Correlation Network Analysis (WGCNA) was implemented on OTU relative abundances using the *WGCNA* package (Langfelder & Horvath, 2008; Langfelder & Horvath, 2012), executed as previously described (Guidi *et al.*, 2016; Henson *et al.*, 2016) with modifications. OTUs which did not possess at least 2 sequences across 10% of samples were

removed from network analyses. These OTUs were removed to eliminate OTUs with potentially spurious correlations to environmental variables or other OTUs, as well as to reduce computational stress of analyses. Remaining OTU abundances across samples were normalized using variance stabilizing transformation (VST) performed as described previously for beta diversity analyses. To ensure scale-free topology of the network, the dissimilarity matrix generated through VST was transformed to an adjacency matrix by raising this dissimilarity matrix to a soft threshold power. A threshold power of $p = 4$ was chosen to meet scale-free topology assumptions based upon criterion established by Zhang & Horvath (2005). Scale-free topology of network relationships was further ensured through regression of the frequency distribution of node connectivity against node connectivity; a network is scale-free if an approximate linear fit of this regression is evident (see Zhang & Horvath, [2005] for more in-depth explanation). A topological overlap matrix (TOM) was generated using the adjacency matrix, and subnetworks of highly connected and correlated OTUs were delineated with the TOM and hierarchical clustering. Representative eigenvalues of each subnetwork (i.e., the first principal component) were correlated (Pearson) with values of measured environmental variables to identify the subnetworks most related to said environmental variables. The subnetworks with the highest positive correlations to environmental variables of interest (e.g., NO_3^-, C:N, etc.) were selected for further analyses of relationships among subnetwork structure, individual OTUs, and environmental variables. Partial Least Square regression (PLS) was used to test predictive ability of subnetworks in estimating variability of environmental parameters, which allowed for delineation of potential indicator subnetworks and OTUs. Pearson correlations were calculated between response variables and leave-one-out cross-validation (LOOCV) predicted values. If PLS found that regression between actual and predicted values was below minimum

563 threshold of R^2 = 0.3, WGCNA analysis was halted for that network, as the network was deemed
564 to lack predictive ability of that environmental variable. Variable Importance in Projection (VIP)
565 (Chong & Jun, 2005) analysis was used to determine the influence of individual OTUs in PLS. A
566 high VIP value for an OTU indicates high importance in prediction of the environmental
567 response variable for that OTU. For network construction and visualization purposes, the
568 minimum correlation value required between two OTUs to constitute an "edge" between them
569 was delineated at different *r* values for each network related to an individual environmental
570 variable (ranging between 0.1 – 0.25), as co-correlations between OTUs within some networks
571 were stronger than others. The number of co-correlations an OTU has with other OTUs within a
572 network is defined as "node centrality" (as described by Henson *et al.*, 2016).

573 If an environmental variable of interest was not found to have a network strongly
574 associated with it, Spearman's correlations were implemented between individual OTUs and that
575 environmental variable. Significance was only considered for OTUs with strong relationships ($r > 0.3$, $p \leq 0.001$) to environmental variables to limit the impact of potentially spurious
577 correlations. Correlations to OTUs with unknown taxonomic identification were excluded from
578 correlational analysis.

579

580 **Acknowledgments**

581 Special thanks to Alexandra Bozimowski and Thomas Langer for their assistance in collecting
582 the wetland samples and data acquisition, as well as Dr. Matthew Cooper for providing valuable
583 insight and conversation on wetland ecology and the custom-built core extruder built by Mr.
584 Gary Cooper. Special thanks also to Mike Henson for providing assistance on network statistical
585 analyses. Funding was provided, in part, by the CMU College of Science and Engineering.

586 Additional funding for this work was provided by the Great Lakes National Program Office
587 under the United States Environmental Protection Agency, grant number GL-00E00612-0 as part
588 of the US federal government's Great Lakes Restoration Initiative. Although the research
589 described in this work has been partly funded by the United States Environmental Protection
590 Agency, it has not been subjected to the agency's required peer and policy review, and therefore,
591 does not necessarily reflect the views of the agency and no official endorsement should be
592 inferred. This paper is contribution number XX of the Central Michigan University Institute for
593 Great Lakes Research.

594

595

596 **References**

597 Achtnich, C., Bak, F., and Conrad, R. (1995). Competition for electron donors among nitrate
598 reducers, ferric iron reducers, sulfate reducers, and methanogens in anoxic paddy soil. Biol Fert
599 Soils. 19(1): 65-72.

600

601 Alewell, C., Paul, S., Lischeid, G., and Storck, F.R. (2008). Co-regulation of redox processes in
602 freshwater wetlands as a function of organic matter availability?. Sci Total Environ 404(2): 335-
603 342.

604

605 Anderson, M.J. (2001). A new method for non□parametric multivariate analysis of variance.
606 Austral Ecol 26(1): 32-46.

607

608 Anderson, M.J. (2006). Distance□based tests for homogeneity of multivariate dispersions.
609 Biometrics 62(1): 245-253.

Anderson, M.J., Ellingsen, K.E., and McArdle, B.H. (2006). Multivariate dispersion as a measure of beta diversity. Ecol Lett 9(6): 683-693.

Anderson, M.J., and Walsh, D.C. (2013). PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: what null hypothesis are you testing?. Ecol Monogr 83(4): 557-574.

Ansola, G., Arroyo, P., and de Miera, L.E.S. (2014). Characterisation of the soil bacterial community structure and composition of natural and constructed wetlands. Sci Total Environ 473: 63-71.

Arroyo, P., de Miera, L.E.S., and Ansola, G. (2015). Influence of environmental variables on the structure and composition of soil bacterial communities in natural and constructed wetlands. Sci Total Environ 506: 380-390.

Campbell, M., Cooper, M.J., Friedman, K., and Anderson, W.P. (2015). The economy as a driver of change in the Great Lakes–St. Lawrence River basin. J Great Lakes Res 41: 69-83.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., *et al.* (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6(8): 1621-1624.

633 Chase, J.M., and Leibold, M.A. (2002). Spatial scale dictates the productivity–biodiversity
634 relationship. Nature 416(6879): 427-430.
635
636 Chong, I.G., and Jun, C.H. (2005). Performance of some variable selection methods when
637 multicollinearity is present. Chemometr Intell Lab 78(1): 103-112.
638
639 Clarke, K.R. (1993). Non-parametric multivariate analyses of changes in community structure.
640 Aust J Ecol 18: 117-143.
641
642 Cole J.R., Wang Q., Fish J.A., Chai B., Mcgarrell D.M., Sun Y., *et al*. (2013) Ribosomal
643 Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res 42(D1):
644 doi:10.1093/nar/gkt1244
645
646 Conrad, R. (1996). Soil microorganisms as controllers of atmospheric trace gases (H2, CO, CH4,
647 OCS, N2O, and NO). Microbiol Rev 60(4): 609-640.
648
649 Cooper, M.J., Gyekis, K.F., and Uzarski, D.G. (2012). Edge effects on abiotic conditions,
650 zooplankton, macroinvertebrates, and larval fishes in Great Lakes fringing marshes. J Great
651 Lakes Res 38(1): 142-151.
652
653 Cvetkovic, M., and Chow-Fraser, P. (2011). Use of ecological indicators to assess the quality of
654 Great Lakes coastal wetlands. Ecological Indicators 11(6): 1609-1622.
655

656 Dahl, T.E. (1990). Wetlands losses in the United States 1780's to 1980's. U.S. Department of the
657 Interior, Fish and Wildlife Service, Washington. D.C. 13pp.
658
659 Danz, N.P., Niemi, G.J., Regal, R.R., Hollenhorst, T., Johnson, L.B., Hanowski, J.M., *et al*.
660 (2007). Integrated measures of anthropogenic stress in the US Great Lakes basin. Environ
661 Manage 39(5): 631-647.
662
663 Deng, Y., Cui, X., Hernández, M., and Dumont, M.G. (2014). Microbial diversity in hummock
664 and hollow soils of three wetlands on the Qinghai-Tibetan Plateau revealed by 16S rRNA
665 pyrosequencing. PLoS One 9(7): e103115.
666
667 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves
668 sensitivity and speed of chimera detection. Bioinformatics 27(16): 2194-2200.
669
670 Edlund, A., Hårdeman, F., Jansson, J.K., and Sjöling, S. (2008). Active bacterial community
671 structure along vertical redox gradients in Baltic Sea sediment. Environ Microbiol 10(8): 2051-
672 2063.
673
674 Ettema, C.H., and Wardle, D.A. (2002). Spatial soil ecology. Trends in ecology and evolution
675 17(4): 177-183.
676
677 Fierer, N., and Jackson, R.B. (2006). The diversity and biogeography of soil bacterial
678 communities. P Natl Acad Sci USA 103(3): 626-631.

679

680 Grundmann, G.L., and Debouzie, D. (2000). Geostatistical analysis of the distribution of NH4+
681 and NO2−-oxidizing bacteria and serotypes at the millimeter scale along a soil transect. FEMS
682 Microbiol Ecol 34(1): 57-62.

683

684 Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., *et al*. (2016). Plankton
685 networks driving carbon export in the oligotrophic ocean. Nature 532(7600): 465.

686

687 Hackett, R.A., Babos, H.B., Collins, E.E., Horton, D., Schock, N., and Schoen, L.S. (2017).
688 Researcher disciplines and the assessment techniques used to evaluate Laurentian Great Lakes
689 coastal ecosystems. J Great Lakes Res 43(1): 9-16.

690

691 Hartman, W.H., Richardson, C.J., Vilgalys, R., and Bruland, G.L. (2008). Environmental and
692 anthropogenic controls over bacterial communities in wetland soils. P Natl Acad Sci 105(46):
693 17842-17847.

694

695 Hecnar, S.J. (2004). Great Lakes wetlands as amphibian habitats: a review. Aquat Ecosyst Health
696 7(2): 289-303.

697

698 Henson, M.W., Hanssen, J., Spooner, G., Flemming, P., Pukonen, M., Stahr, F., and Thrash, J.C.
699 (2016). Nutrient dynamics and stream order influence microbial community patterns along a
700 2914 km transect of the Mississippi River. bioRxiv, 091512.

701

Hill, B.H., Elonen, C.M., Jicha, T.M., Cotter, A.M., Trebitz, A.S., and Danz, N.P. (2006). Sediment microbial enzyme activity as an indicator of nutrient limitation in Great Lakes coastal wetlands. Freshwater Biol 51(9): 1670-1683.

Horner-Devine, M.C., Lage, M., Hughes, J.B., and Bohannan, B.J. (2004). A taxa-area relationship for bacteria. Nature 432(7018): 750.

Howe, R.W., Regal, R.R., Hanowski, J., Niemi, G.J., Danz, N.P., and Smith, C.R. (2007). An index of ecological condition based on bird assemblages in Great Lakes coastal wetlands. J Great Lakes Res 33(sp3): 93-105.

Hu, Y., Wang, L., Fu, X., Yan, J., Wu, J., Tsang, Y., *et al*. (2016). Salinity and nutrient contents of tidal water affects soil respiration and carbon sequestration of high and low tidal flats of Jiuduansha wetlands in different ways. Sci Total Environ 565: 637-648.

Jackson, C.R., Foreman, C.M., and Sinsabaugh, R.L. (1995). Microbial enzyme activities as indicators of organic matter processing rates in a Lake Erie coastal wetland. Freshwater biol 34(2): 329-342.

Kögel-Knabner, I., Amelung, W., Cao, Z., Fiedler, S., Frenzel, P., Jahn, R., *et al*. (2010). Biogeochemistry of paddy soils. Geoderma 157(1): 1-14.

724 Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., Schloss, P.D. (2013) Development
725 of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence
726 Data on the MiSeq Illumina Sequencing Platform. Appl Environ Microb 79(17): 5112-5120.
727
728 Krantzberg, G., and De Boer, C. (2008). A valuation of ecological services in the Laurentian
729 Great Lakes Basin with an emphasis on Canada. Am Wat Works Assoc J 100(6): 100.
730
731 Lamers, L.P., Van Diggelen, J.M., den Camp, H.J.O., Visser, E.J., Lucassen, E.C., Vile, M.A., *et*
732 *al*. (2012). Microbial transformations of nitrogen, sulfur, and iron dictate vegetation composition
733 in wetlands: a review. Front Microbiol, 3:156.
734
735 Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation
736 network analysis. BMC Bioinformatics 9(1): 559.
737
738 Langfelder, P., and Horvath, S. (2012). Fast R functions for robust correlations and hierarchical
739 clustering. J Stat Softw 46(11).
740
741 Lauber, C.L., Hamady, M., Knight, R., and Fierer, N. (2009). Pyrosequencing-based assessment
742 of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl
743 Environ Microb 75(15): 5111-5120.
744

745 Ligi, T., Oopkaup, K., Truu, M., Preem, J.K., Nõlvak, H., Mitsch, W.J., *et al*. (2014).
746 Characterization of bacterial communities in soil and sediment of a created riverine wetland
747 complex using high-throughput 16S rRNA amplicon sequencing. Ecol Eng 72: 56-66.
748
749 Lipson, D.A., Raab, T.K., Parker, M., Kelley, S.T., Brislawn, C.J., and Jansson, J. (2015).
750 Changes in microbial communities along redox gradients in polygonized Arctic wet tundra soils.
751 Env Microbiol Rep 7(4): 649-657.
752
753 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
754 dispersion for RNA-seq data with DESeq2. Genome Biol 15(12): 550.
755
756 Lüdemann, H., Arth, I., and Liesack, W. (2000). Spatial changes in the bacterial community
757 structure along a vertical oxygen gradient in flooded paddy soil cores. Appl Environ Microb
758 66(2): 754-762.
759
760 Martiny, J.B., Eisen, J.A., Penn, K., Allison, S.D., and Horner-Devine, M.C. (2011). Drivers of
761 bacterial β-diversity depend on spatial scale. P Natl Acad Sci 108(19): 7850-7854.
762
763 McMurdie, P.J., and Holmes, S. (2013) Phyloseq: an R package for reproducible interactive
764 analysis and graphics of microbiome census data. PLoS One 8(4).
765
766 McMurdie, P.J., and Holmes, S. (2014). Waste not, want not: why rarefying microbiome data is
767 inadmissible. PLoS Comput Biol, 10(4): e1003531.

Morrice, J.A., Danz, N.P., Regal, R.R., Kelly, J.R., Niemi, G.J., Reavie, E.D., *et al*. (2008). Human influences on water quality in Great Lakes coastal wetlands. Environ Manage 41(3): 347-357.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M.H.H., Oksanen, M.J., and Suggests, M.A.S.S. (2007). The vegan package. Community ecology package, 10, 631-637.

Peralta, R.M., Ahn, C., and Gillevet, P.M. (2013). Characterization of soil bacterial community structure and physicochemical properties in created and natural wetlands. Sci Total Environ 443: 725-732.

Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., *et al*. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41(D1): D590-D596.

R Core Team (2015) R: A language and environment for statistical computing. Vienna, Austria. https://www.R-project.org/.

Reddy, K.R., and DeLaune, R.D. (2008). Biogeochemistry of wetlands: science and applications. Boca Raton, FL: CRC Press.

790 Sanford, R.A., Cole, J.R., and Tiedje, J.M. (2002). Characterization and description of
791 Anaeromyxobacter dehalogenans gen. nov., sp. nov., an aryl-halorespiring facultative anaerobic
792 myxobacterium. Appl Environ Microb 68(2): 893-900.
793
794 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., *et al.* (2009).
795 Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for
796 Describing and Comparing Microbial Communities. Appl Environ Microbiol 75(23): 7537-7541.
797
798 Sekiguchi, Y., Yamada, T., Hanada, S., Ohashi, A., Harada, H., and Kamagata, Y. (2003).
799 Anaerolinea thermophila gen. nov., sp. nov. and Caldilinea aerophila gen. nov., sp. nov., novel
800 filamentous thermophiles that represent a previously uncultured lineage of the domain Bacteria
801 at the subphylum level. Int J Syst Evol Micr 53(6): 1843-1851.
802
803 Sierszen, M.E., Morrice, J.A., Trebitz, A.S., and Hoffman, J.C. (2012). A review of selected
804 ecosystem services provided by coastal wetlands of the Laurentian Great Lakes. Aquat Ecosyst
805 Health 15(1): 92-106.
806
807 Sims, A., Zhang, Y., Gajaraj, S., Brown, P.B., and Hu, Z. (2013). Toward the development of
808 microbial indicators for wetland assessment. Water Res 47(5): 1711-1725.
809
810 Trebitz, A.S., Brazner, J.C., Cotter, A.M., Knuth, M.L., Morrice, J.A., Peterson, G.S., *et al.*
811 (2007). Water quality in Great Lakes coastal wetlands: basin-wide patterns and responses to an
812 anthropogenic disturbance gradient. J Great Lakes Res 33(sp3): 67-85.

Tulbure, M.G., Johnston, C.A., and Auger, D.L. (2007). Rapid invasion of a Great Lakes coastal wetland by non-native Phragmites australis and Typha. J Great Lakes Res 33(sp3): 269-279.

Urakawa, H., and Bernhard, A.E. (2017). Wetland management using microbial indicators. Ecol Eng 108(B): 456-476.

Uzarski, D.G. (2009). Wetlands of Large Lakes. Encyclopedia of Inland Waters. Oxford: Elsevier. p. 599-606.

Uzarski, D.G., Brady, V.J., Cooper, M.J., Wilcox, D.A., Albert, D.A., Axler, R.P, *et al*. (2017). Standardized measures of coastal wetland condition: Implementation at a Laurentian Great Lakes basin-wide scale. Wetlands 37(1): 15-32.

Uzarski, D.G., Burton, T.M., Kolar, R.E., and Cooper, M.J. (2009). The ecological impacts of fragmentation and vegetation removal in Lake Huron's coastal wetlands. Aquat Ecosyst Health 12(1): 45-62.

Wang, N., and Mitsch, W.J. (1998). Estimating phosphorus retention of existing and restored coastal wetlands in a tributary watershed of the Laurentian Great Lakes in Michigan, USA. Wetl Ecol Manag 6(1): 69-82.

Yamada, T., Sekiguchi, Y., Hanada, S., Imachi, H., Ohashi, A., Harada, H., and Kamagata, Y. (2006). Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the bacterial phylum Chloroflexi. Int J Syst Evol Micr 56(6): 1331-1340.

Yamada, T., Imachi, H., Ohashi, A., Harada, H., Hanada, S., Kamagata, Y., and Sekiguchi, Y. (2007). Bellilinea caldifistulae gen. nov., sp. nov. and Longilinea arvoryzae gen. nov., sp. nov., strictly anaerobic, filamentous bacteria of the phylum Chloroflexi isolated from methanogenic propionate-degrading consortia. Int J Syst Evol Micr 57(10): 2299-2306.

Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol 4(1).

858

859

860

861

862

863

864

865

866 **Tables**

867 Table 1. Correlations between alpha diversity metrics and measured environmental variables.

868 Asterisks represent significance values where $p \leq 0.001$ (***), $p \leq 0.01$ (**), and $p \leq 0.05$ (*).

	Chao 1	Shannon
P	0.31**	-0.09
S	0.42***	0.45***
NO3	0.42***	0.24*
C:N	-0.03	-0.2
NUTR	0.24*	0.41***

869

870 Table 2. Pairwise perMANOVA results comparing pairwise differences between wetland regions

871 and differences between wetland soil depths. Asterisks represent significance values where $p \leq$

872 0.001 (***), $p \leq 0.01$ (**), $p \leq 0.05$ (*), and *n.s.* representing 'not significant' below p-value

873 thresholds.

Region	BA	ESBT	LE	NSB	WSB
BA	-				
ESBT	*n.s.*	-			
LE	**	**	-		
NSB	*n.s.*	*n.s.*	**	-	
WSB	*n.s.*	*n.s.*	**	*n.s.*	-

Depth	Top	Middle	Bottom
Top	-		
Middle	**	-	
Bottom	**	*n.s.*	-

874

875

876

877

878

879

880 **Figures**

881 Figure 1. Principal Component Analysis (PCA) illustrating separation of samples based upon soil
882 geochemistry. Shapes and colors correspond to different wetland depths and regions,
883 respectively, as listed in the legend. Percentages on axes represent explained variance of that
884 principal component. Vectors represent impact of specific environmental variables on sample
885 distribution. NUTR represents OM values, which correlated significantly ($p \leq .01$, $r > 0.56$) to
886 NO_3^-, OC, OM, S, and TN. Ellipses represent 95% confidence intervals of region groupings.

887

888 Figure 2. Boxplot diagram comparing Chao1 diversity among wetland regions. Boxes with the
889 same letter are not significantly different, while those with no common letters are significantly
890 different ($p \leq 0.01$). Lines within boxes represent the median, hinges represent +/- 25% quartiles,
891 whiskers represent up to 1.5x the interquartile range. Colors represent wetland region.

892

893 Figure 3. Nonmetric Multidimensional Scaling (NMDS) plot illustrating separation of samples
894 based upon differences in microbial community structure. Shapes and colors correspond to

different depths and wetland regions, respectively, as listed in the legend. Vectors represent correlations of environmental variables to the distribution of the microbial communities represented in the plot.

Figure 4. NMDS plots of each wetland region demonstrating separation of samples based upon differences in microbial community structure, including a. BA, b. ESBT, c. LE, d. NSB, and e. WSB. Shapes and colors correspond to different depths and wetland sites, respectively, as listed in the legends. Vectors represent correlations of environmental variables to the distribution of microbial communities represented in the plots.

Figure 5. Network visualization and results of partial least squares analysis on the subnetwork most correlated with NUTR. The y-axis represents correlation of OTU to OC values, whereas the x-axis represents the node centrality. Points represent OTUs, and the color of points corresponds to the phylum to which an OTU belongs. Point size corresponds to VIP score of that OTU. The top 15 OTUs are labeled within the graph with corresponding lowest taxonomic identification possible, and the level of that classification. D = Domain; P = Phylum, C = Class, O = Order, F = Family, G = Genus.

Figure 6. Network visualization and results of partial least squares analysis on the subnetwork most correlated with C:N. The y-axis represents correlation of OTU to C:N, whereas the x-axis represents the node centrality. Points represent OTUs, and the color of points corresponds to the phylum to which an OTU belongs. Point size corresponds to VIP score of that OTU. Only OTUs with a VIP score > 1 were displayed for visualization purposes. The top 15 OTUs are labeled

within the graph with corresponding lowest taxonomic identification possible, and the level of that classification. D = Domain; P = Phylum, C = Class, O = Order, F = Family, G = Genus.

Figure 7. Significant ($p \leq 0.001$) positive and negative correlations between microbial taxonomic groups and NO_3^-. Abundance of OTUs corresponding to taxonomic groups is plotted on the y-axis. Positive correlations are plotted above the "0" line, while negative correlations are plotted below the "0" line. Colors of bars correspond to associated phylum of the taxonomic group on the x-axis.

Figure 8. Geographic map displaying location of sites sampled within this study. Colors of points correspond to region sampled.

Supplemental Material

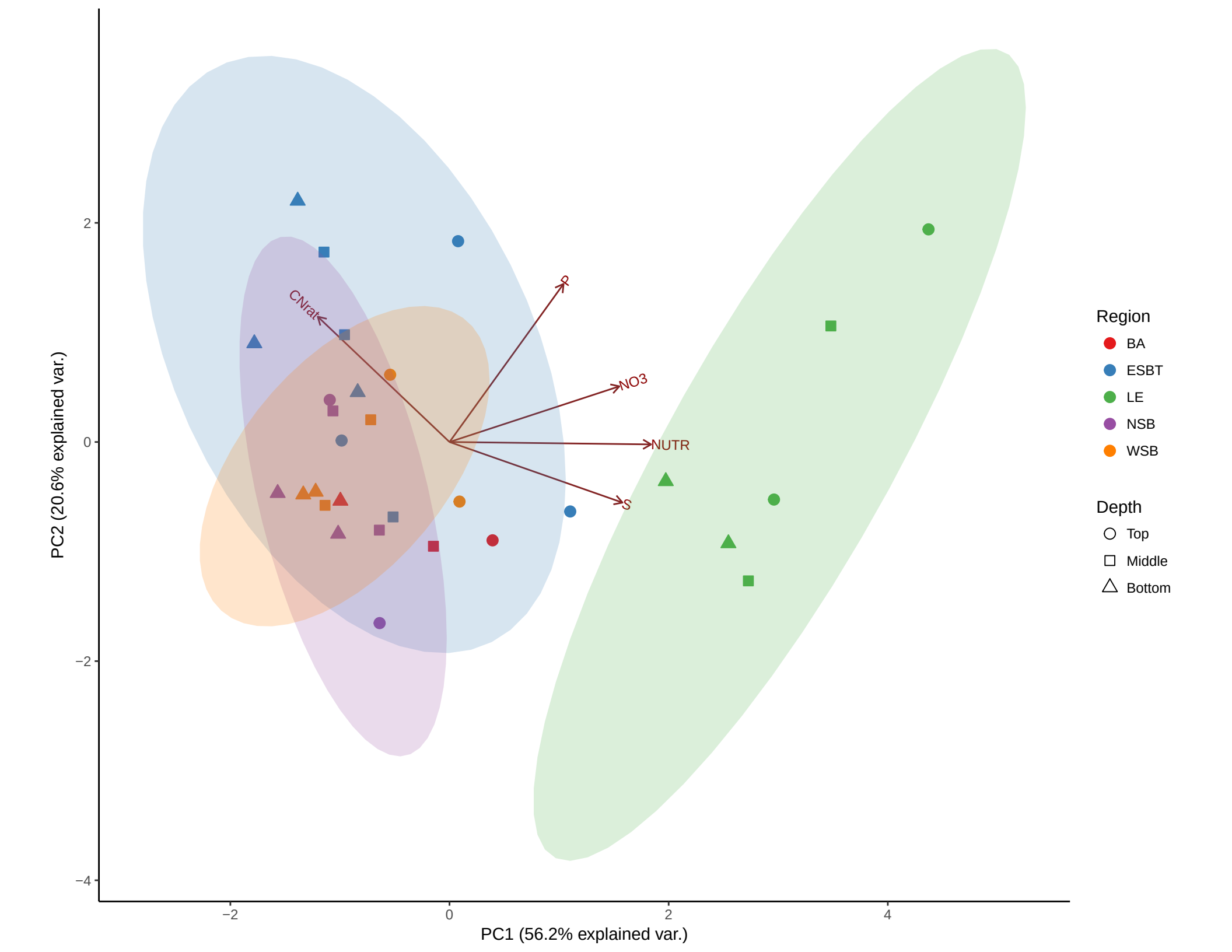
Supplemental Figure 1. Depth profiles demonstrating trends in measured environmental variables with increasing depth among wetland regions and sites. Colors represent wetland regions, whereas point shapes represent distinct wetland sites.

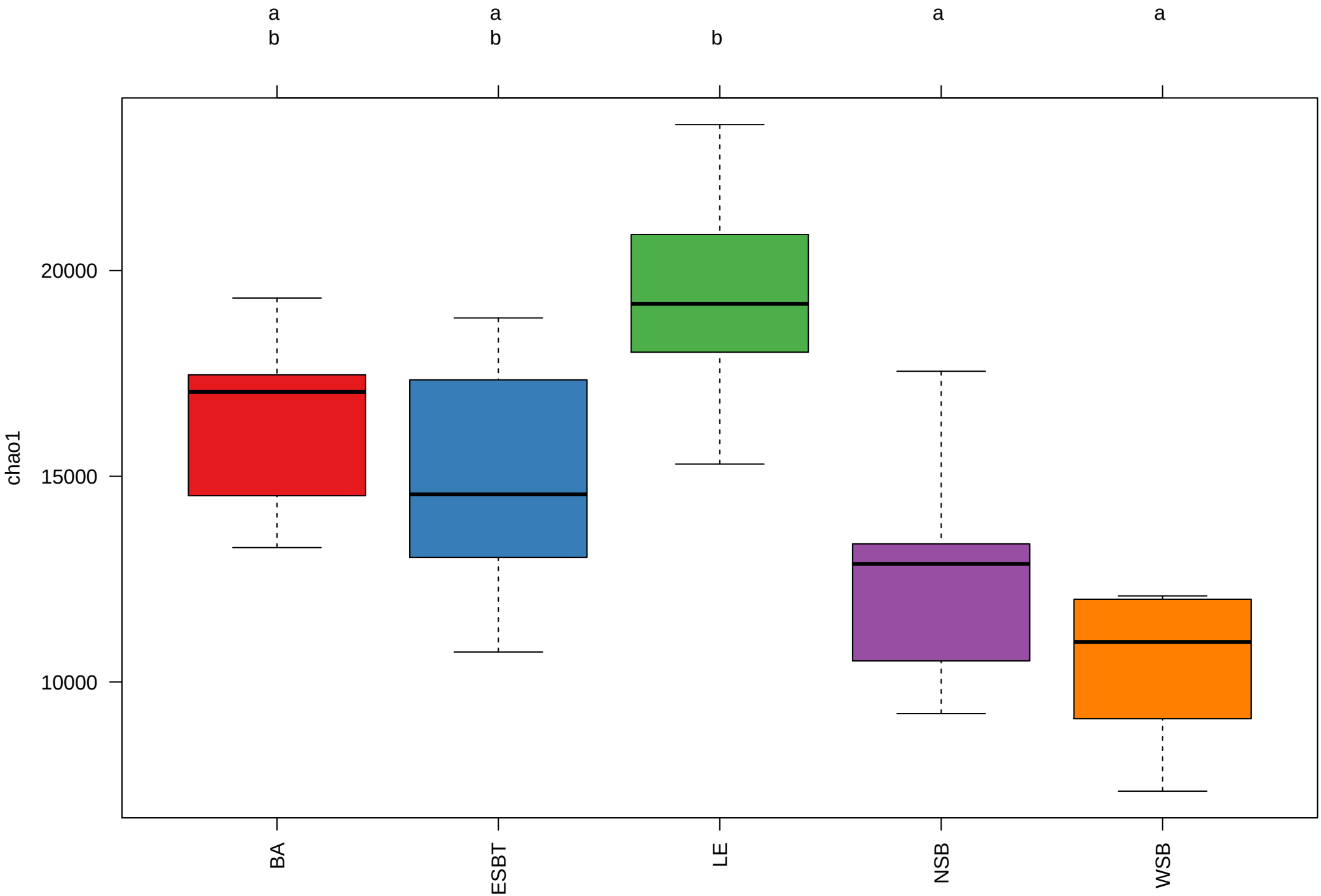
Supplemental Figure 2. Rarefaction curve analysis demonstrating sequencing depth for alpha diversity analyses. Different colored lines represent different samples. The vertical black line represents the sequencing depth used to standardize all samples for alpha diversity analysis. The dashed line represents the 1:1 slope.

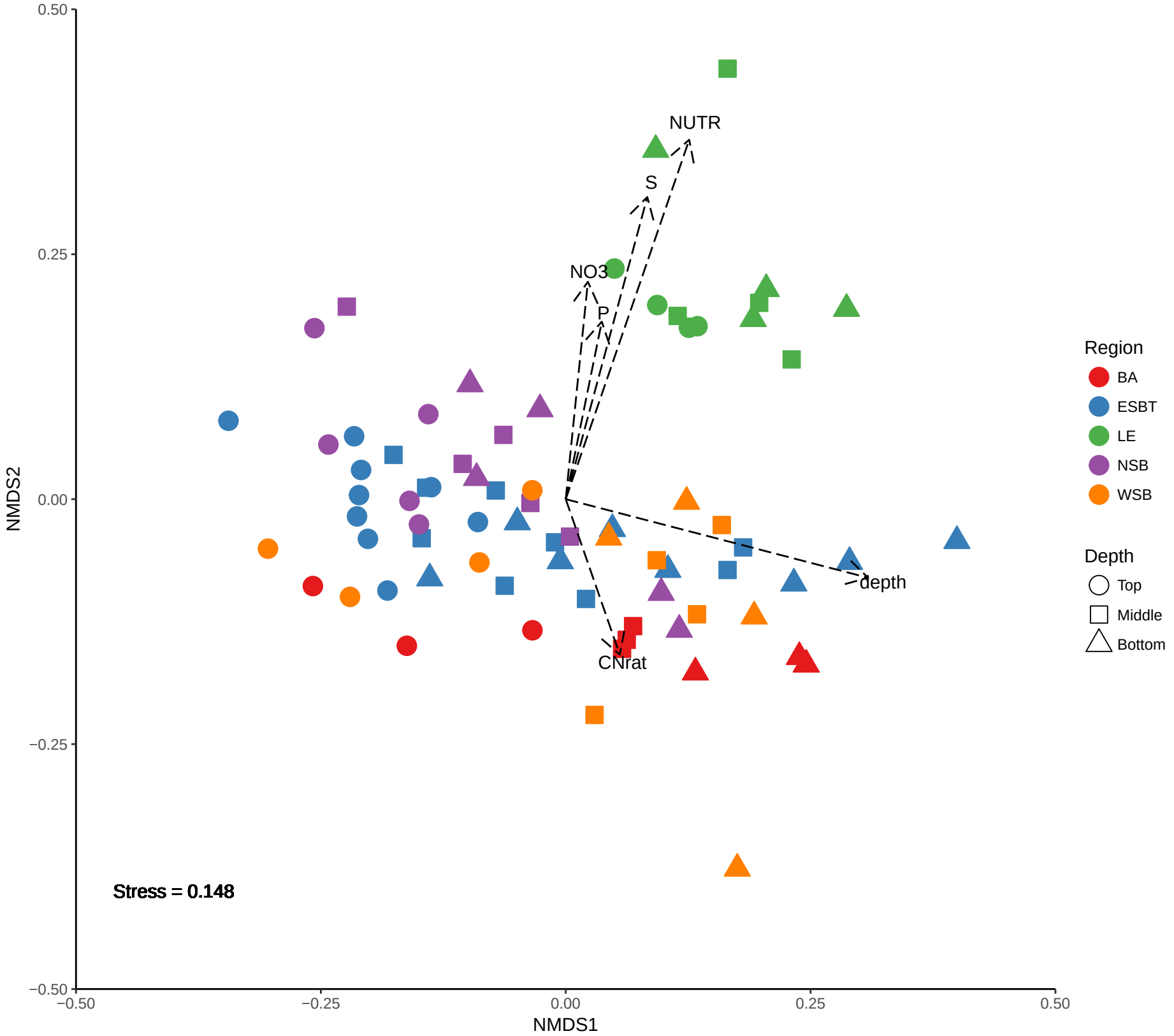
941 Supplemental Figure 3. Plots demonstrating regression of the frequency distribution of node
942 connectivity against node connectivity.
943
944 Supplemental Figure 4. Individual correlations of established subnetworks with environmental
945 parameters.
946
947 Supplemental Figure 5. Partial least squares analysis results predicting OC (NUTR) values using
948 relative abundance values of OTUs within the module most connected to OC (NUTR).
949
950 Supplemental Figure 6. Partial least squares analysis results predicting C:N values using relative
951 abundance values of OTUs within the module most connected to C:N.
952
953 Supplemental Figure 7. Partial least squares analysis results predicting NO_3^- values using relative
954 abundance values of OTUs within the module most connected to NO_3^-.
955
956 Supplemental Figure 8. Principal Components Analysis demonstrating separation of coastal
957 wetland sampling locations by site water quality data. Points are color-coded by region.
958 Percentages on axes represent explained variance of that principal component. Vectors represent
959 impact of specific environmental variables on sample distribution. Choloro.a = chlorophyll A,
960 DO = dissolved oxygen, Temp = water temperature, Turb = water turbidity. SumRank represents
961 SumRank values calculated for each wetland site as outlined by Uzarski et al. (2017). DO also
962 represents redox potential values, which correlated significantly ($r > 0.7$, $p \leq .001$). Temp also

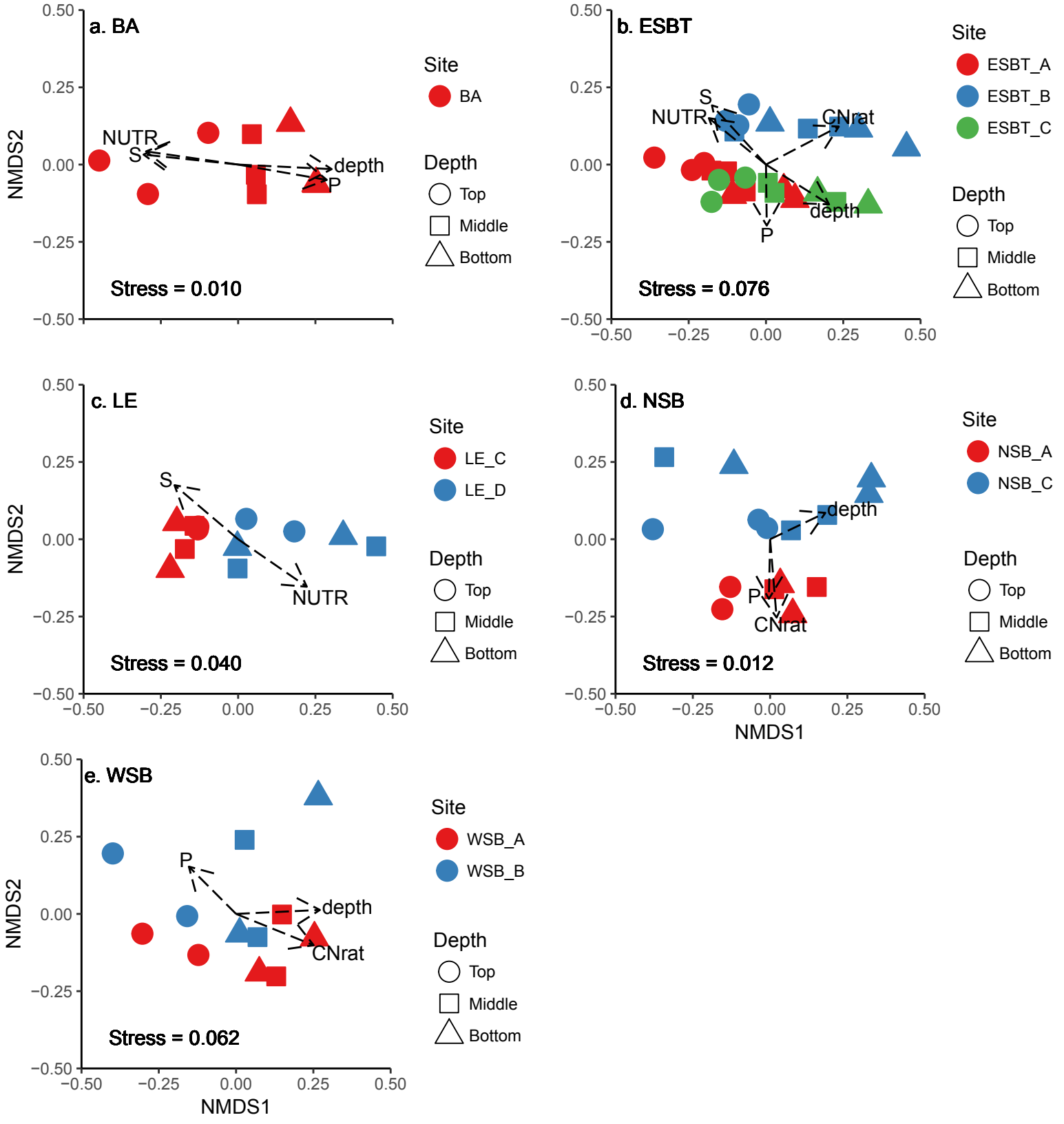
963 represents conductivity and total dissolved solids values, which correlated significantly ($r > 0.7$,

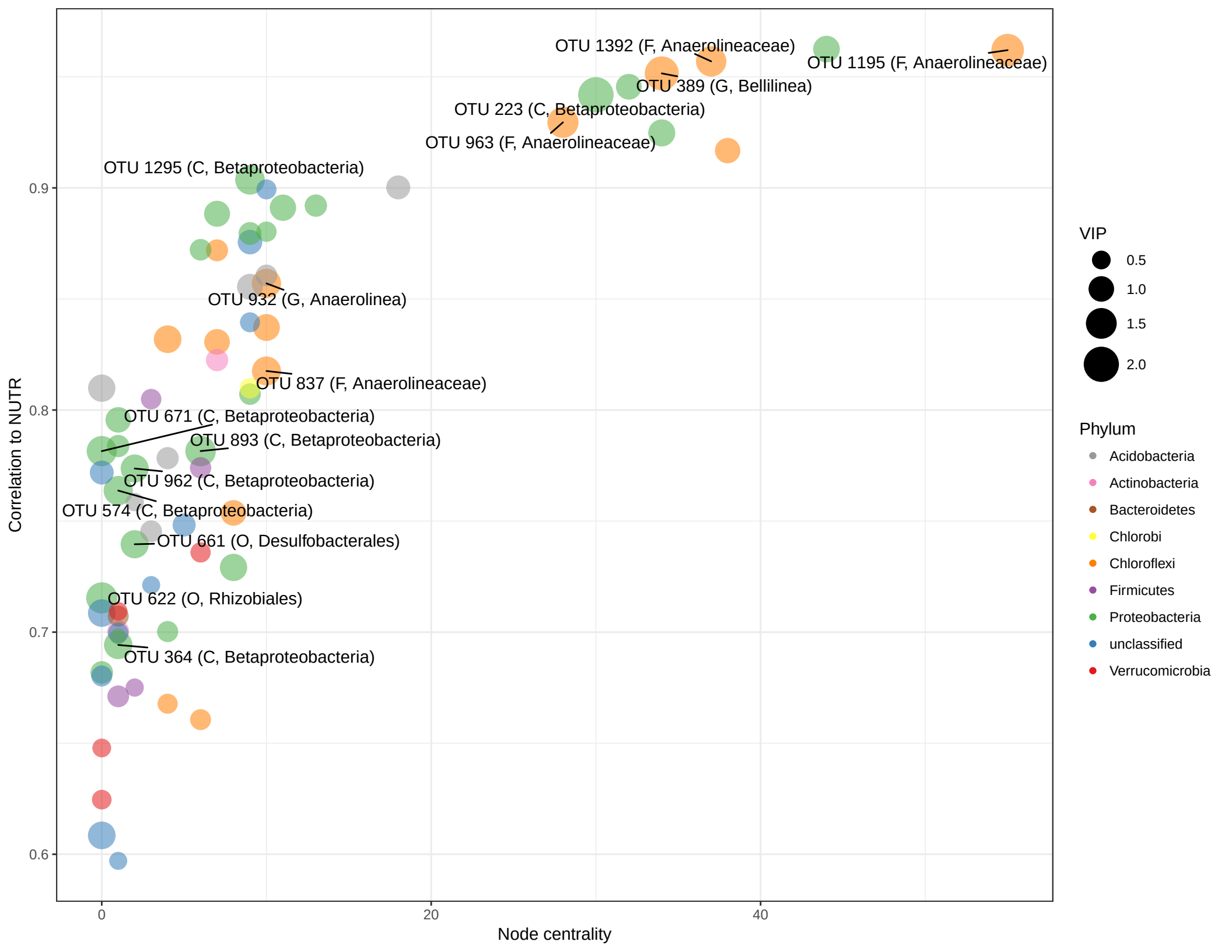
964 $p \leq .001$).

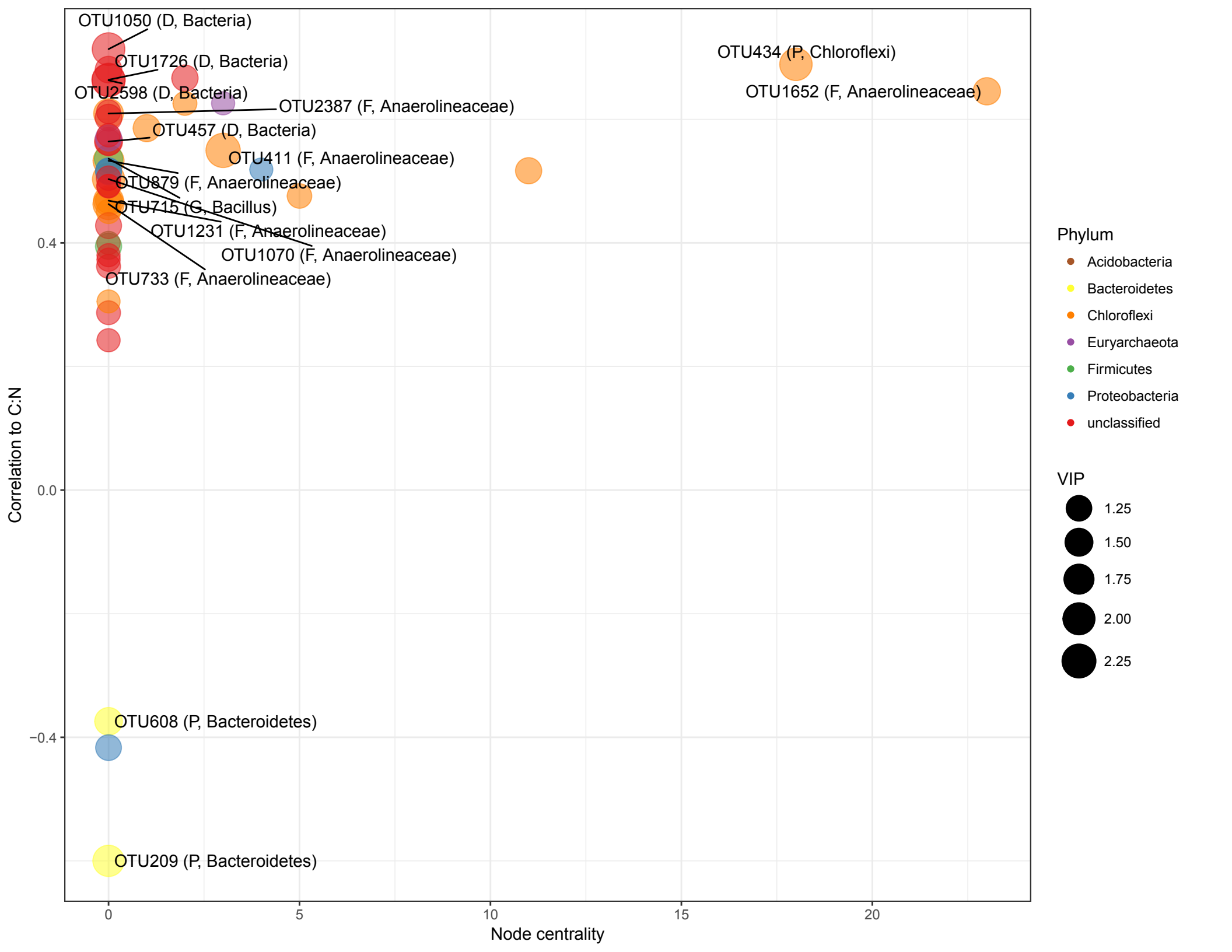

PC2 (20.6% explained var.)
PC1 (56.2% explained var.)
2
0
−2
−4
−2
0
2
4
CNrat
P
NO3
NUTR
S
Region
BA
ESBT
LE
NSB
WSB
Depth
Top
Middle
Bottom

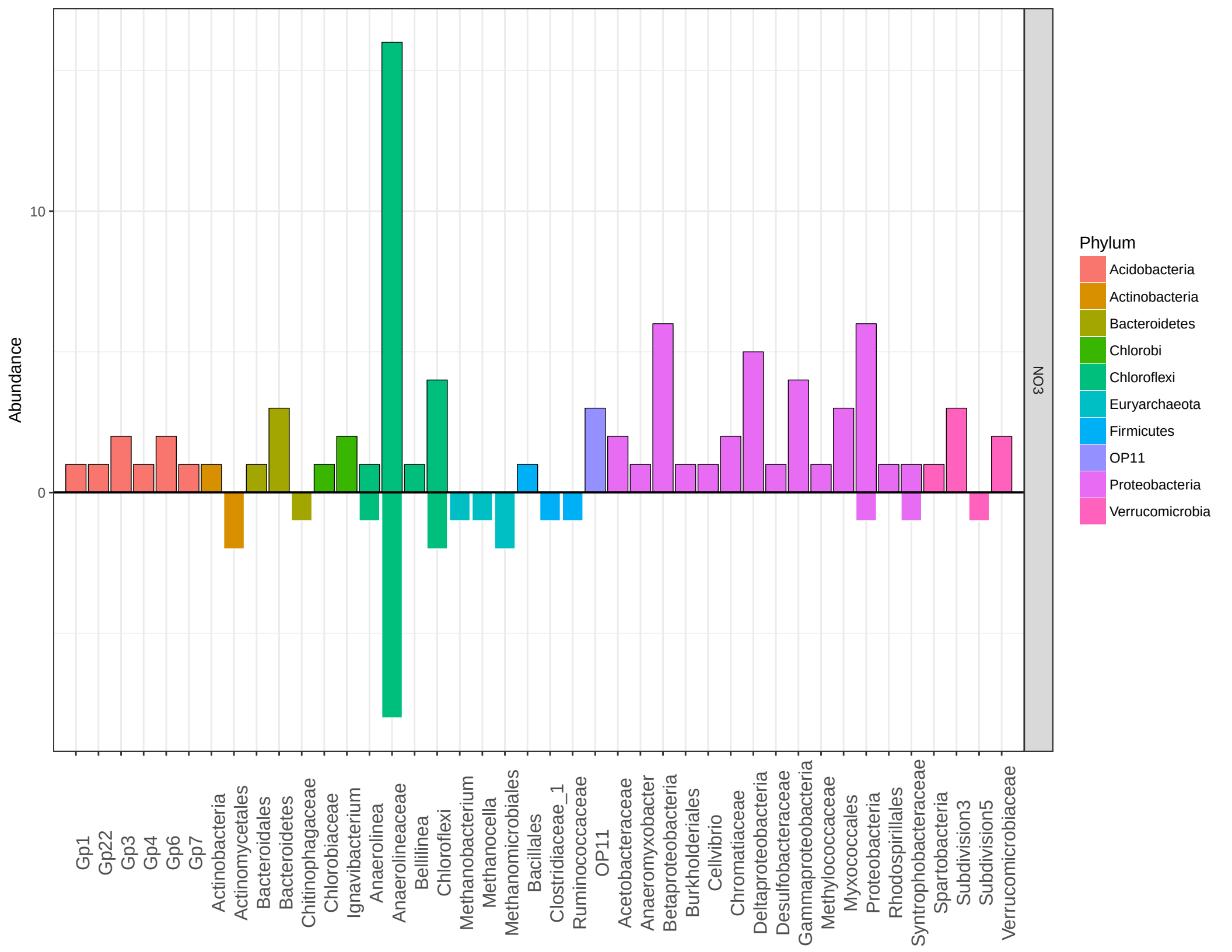

a
b
a
b
b
a
a
chao1
20000
15000
10000
BA
ESBT
LE
NSB
WSB

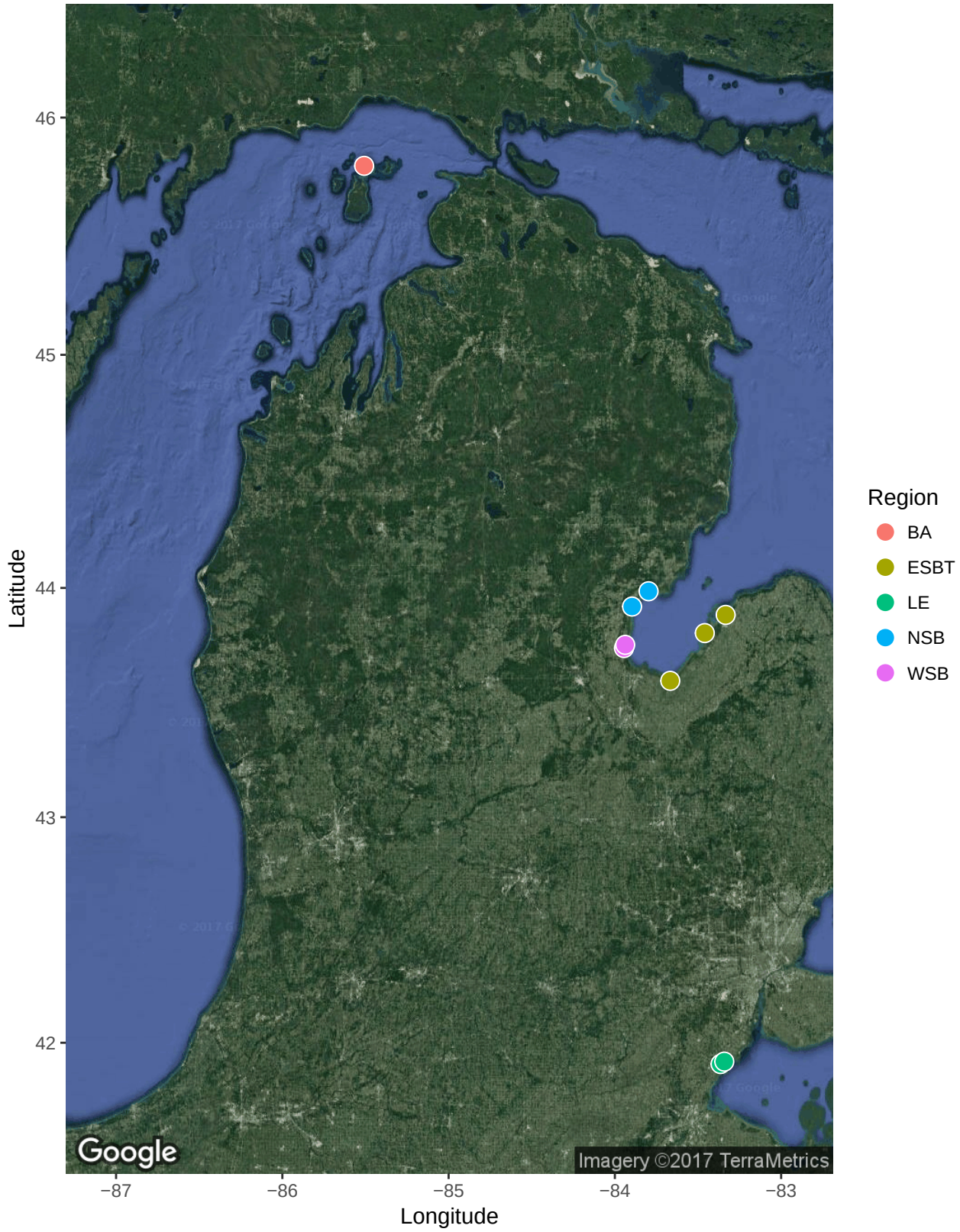

NUTR
S
NO3
P
depth
CNrat
Region
BA
ESBT
LE
NSB
WSB
Depth
Top
Middle
Bottom
Stress = 0.148
NMDS1
NMDS2
0.50
0.25
0.00
−0.25
−0.50
−0.50
−0.25
0.00
0.25
0.50


a. BA
NUTR
S
depth
P
Stress = 0.010
Site
BA
Depth
Top
Middle
Bottom
b. ESBT
S
NUTR
CNrat
depth
P
Stress = 0.076
Site
ESBT_A
ESBT_B
ESBT_C
Depth
Top
Middle
Bottom
c. LE
S
NUTR
Stress = 0.040
Site
LE_C
LE_D
Depth
Top
Middle
Bottom
d. NSB
depth
P
CNrat
Stress = 0.012
Site
NSB_A
NSB_C
Depth
Top
Middle
Bottom
e. WSB
P
depth
CNrat
Stress = 0.062
Site
WSB_A
WSB_B
Depth
Top
Middle
Bottom
NMDS1
NMDS2
0.50
0.25
0.00
−0.25
−0.50


OTU 1392 (F, Anaerolineaceae)
OTU 1195 (F, Anaerolineaceae)
OTU 389 (G, Bellilinea)
OTU 223 (C, Betaproteobacteria)
OTU 963 (F, Anaerolineaceae)
OTU 1295 (C, Betaproteobacteria)
OTU 932 (G, Anaerolinea)
OTU 837 (F, Anaerolineaceae)
OTU 671 (C, Betaproteobacteria)
OTU 893 (C, Betaproteobacteria)
OTU 962 (C, Betaproteobacteria)
OTU 574 (C, Betaproteobacteria)
OTU 661 (O, Desulfobacterales)
OTU 622 (O, Rhizobiales)
OTU 364 (C, Betaproteobacteria)
Correlation to NUTR
0.9
0.8
0.7
0.6
0
20
40
Node centrality
VIP
0.5
1.0
1.5
2.0
Phylum
Acidobacteria
Actinobacteria
Bacteroidetes
Chlorobi
Chloroflexi
Firmicutes
Proteobacteria
unclassified
Verrucomicrobia


OTU1050 (D, Bacteria)
OTU1726 (D, Bacteria)
OTU2598 (D, Bacteria)
OTU2387 (F, Anaerolineaceae)
OTU457 (D, Bacteria)
OTU411 (F, Anaerolineaceae)
OTU879 (F, Anaerolineaceae)
OTU715 (G, Bacillus)
OTU1231 (F, Anaerolineaceae)
OTU1070 (F, Anaerolineaceae)
OTU733 (F, Anaerolineaceae)
OTU434 (P, Chloroflexi)
OTU1652 (F, Anaerolineaceae)
OTU608 (P, Bacteroidetes)
OTU209 (P, Bacteroidetes)
Correlation to C:N
0.4
0.0
−0.4
0
5
10
15
20
Node centrality
Phylum
Acidobacteria
Bacteroidetes
Chloroflexi
Euryarchaeota
Firmicutes
Proteobacteria
unclassified
VIP
1.25
1.50
1.75
2.00
2.25


Abundance
10
0
NO3
Gp1
Gp22
Gp3
Gp4
Gp6
Gp7
Actinobacteria
Actinomycetales
Bacteroidales
Bacteroidetes
Chitinophagaceae
Chlorobiaceae
Ignavibacterium
Anaerolinea
Anaerolineaceae
Bellilinea
Chloroflexi
Methanobacterium
Methanocella
Methanomicrobiales
Bacillales
Clostridiaceae_1
Ruminococcaceae
OP11
Acetobacteraceae
Anaeromyxobacter
Betaproteobacteria
Burkholderiales
Cellvibrio
Chromatiaceae
Deltaproteobacteria
Desulfobacteraceae
Gammaproteobacteria
Methylococcaceae
Myxococcales
Proteobacteria
Rhodospirillales
Syntrophobacteraceae
Spartobacteria
Subdivision3
Subdivision5
Verrucomicrobiaceae
Phylum
Acidobacteria
Actinobacteria
Bacteroidetes
Chlorobi
Chloroflexi
Euryarchaeota
Firmicutes
OP11
Proteobacteria
Verrucomicrobia


Latitude
46
45
44
43
42
−87
−86
−85
−84
−83
Longitude
Google
Imagery ©2017 TerraMetrics
Region
BA
ESBT
LE
NSB
WSB